

# Weak and Strong Reversibility of Non-deterministic Actions: Universality and Uniformity

Jakub Med<sup>1,2</sup>, Lukáš Chrpa<sup>1</sup>, Michael Morak<sup>3\*</sup>, Wolfgang Faber<sup>3\*</sup>

<sup>1</sup>Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague

<sup>2</sup>Faculty of Electrical Engineering, Czech Technical University in Prague

<sup>3</sup>University of Klagenfurt

jakub.med@cvut.cz, chrpaluk@cvut.cz, Michael.Morak@aau.at, Wolfgang.Faber@aau.at

## Abstract

Classical planning looks for a sequence of actions that transform the initial state of the environment into a goal state. Studying whether the effects of an action can be undone by a sequence of other actions, that is, *action reversibility*, is beneficial, for example, in determining whether an action is safe to apply. This paper deals with action reversibility of non-deterministic actions, i.e., actions whose application might result in different outcomes. Inspired by the established notions of weak and strong plans in non-deterministic (or FOND) planning, we define the notions of *weak and strong reversibility* for non-deterministic actions. We then focus on the universality and uniformity of action reversibility, that is, whether we can always undo all possible effects of the action by the same means (i.e., policy), or whether some of the effects can never be undone. We show how these classes of problems can be solved via classical or FOND planning and evaluate our approaches on FOND benchmark domains.

## Introduction

*Automated planning* is a subfield of Artificial Intelligence that deals with the problem of whether there exists a sequence of actions that leads from an initial state to some goal state of the environment (Ghallab, Nau, and Traverso 2004, 2016). In Fully Observable Non-deterministic (FOND) planning, actions have non-deterministic effects, i.e., if an action is applied, one of the effects will (randomly) occur, and hence the formalism can model the uncertainty associated with applying such actions (Cimatti et al. 2003).

In recent years, a problem concerning whether the effects of an action can be undone by means of other actions has drawn attention. In the literature, this problem is referred to as *action reversibility* (Morak et al. 2020). One of the main motivations for its study is that actions identified as reversible are safe to apply. In other words, if a reversible action is applied, its application does not lead to a dead-end state (unless the action was already applied in a dead-end state). This property is especially useful in online planning, where an agent plans and acts w.r.t. a short time horizon, as it might provide information about safe states (Cserna et al. 2018). Another possible benefit of action reversibility

can be in post-planning plan optimization, in which it might contribute to more efficient computations of redundant “action cycles” (Med and Chrpa 2022). In FOND planning, the potential benefits of action reversibility also include determining whether recovery from undesirable effects is possible, which might lead to more efficient generation of strong cyclic plans (Muise, McIlraith, and Beck 2012; Camacho, Muise, and McIlraith 2016), or whether we can safely perform deterministic replanning in unknown states (cf. FF-replan (Yoon, Fern, and Givan 2007)).

A specific form of action reversibility involving the search for an *inverse action* has been investigated (Koehler and Hoffmann 2000; Chrpa, McCluskey, and Osborne 2012). The existence of a pair of inverse actions is a special case of action reversibility, where the reverse plan contains only one action. Eiter, Erdem, and Faber (2008) introduced the concept of *reverse plans* that corresponds to the notion of *uniform reversibility* established by Morak et al. (2020) referring to reversibility of action effects by the same means (i.e., by the same action sequence). A more general notion of (non-uniform) reversibility has been tackled by compiling the problem into contingent planning (Daum et al. 2016) or logic programs (Faber, Morak, and Chrpa 2021). Recent work of Chrpa, Faber, and Morak (2021) investigates under which condition we can find universal reverse plans.

This paper studies the reversibility of non-deterministic actions, which is, to the best of our knowledge, the first work in this area. Inspired by the established notions of *weak and strong plans* (Cimatti et al. 2003) we propose notions of *weak and strong reversibility* that, in a nutshell, refer to “all effects might be undone” and “all effects can always be undone”, respectively. We then focus on a subclass of the problem concerning *universal uniform reversibility*, that is, whether we can always undo all effects of a (non-deterministic) action by the same (weak or strong) policy, or whether some of the effects can never be undone, i.e., they are *universally irreversible*. We show how these problems can be compiled into classical planning (for weak reversibility and irreversibility) or FOND planning (for strong reversibility). Our approaches are evaluated on a range of FOND benchmark domains.

\*These authors contributed equally.

Copyright © 2024, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

## Preliminaries

We adopt terminology from the Simplified Action Structures (SAS<sup>+</sup>) formalism (Bäckström and Nebel 1995) and the FOND planning (Cimatti et al. 2003).

Let  $v$  be a **variable** (or a **state variable**) and  $dom(v)$  be its domain. A **fact**  $(v, x)$  is a pair that contains a variable  $v$  and its value  $x \in dom(v)$ . The **set of all facts** over the set of variables  $\mathcal{V}$  is denoted as  $\mathcal{F}(\mathcal{V})$ . A set  $\Sigma \subseteq \mathcal{F}(\mathcal{V})$  is called a **variable assignment** over  $\mathcal{V}$  if and only if for all  $(v, x) \in \Sigma$  does not exist  $y \in dom(v)$  such that  $y \neq x$  and  $(v, y) \in \Sigma$ . The **value** of the variable  $v$  in  $\Sigma$ , denoted as  $\Sigma[v]$ , is equal to  $x$  if and only if  $(v, x) \in \Sigma$ . The **set of all variables** in  $\Sigma$ , denoted as  $vars(\Sigma)$ , is defined as  $vars(\Sigma) = \{v \mid (v, x) \in \Sigma\}$ . A variable assignment  $\Sigma$  over the set of variables  $\mathcal{V}$  is called a **complete variable assignment** if and only if  $vars(\Sigma) = \mathcal{V}$ . Otherwise, it is called a **partial variable assignment**. The **set of all complete variable assignments** over  $\mathcal{V}$  is denoted as  $\mathcal{S}(\mathcal{V})$ .

An **action**  $a$  over a set of variables  $\mathcal{V}$  is a pair  $(pre(a), eff(a))$ , where  $pre(a)$  is a variable assignment over  $\mathcal{V}$  representing the **precondition** of the action  $a$ , and  $eff(a)$  is a non-empty set of variable assignments over  $\mathcal{V}$  representing the **set of possible effects** of the action  $a$ . The action  $a$  is called **deterministic** if and only if  $|eff(a)| = 1$ . A **determinization** of an action  $a$  with respect to the effect  $e \in eff(a)$ , denoted as  $a_e^d$ , is an action  $a_e^d = (pre(a), \{e\})$ . The set of variables related to an action  $a$ , denoted  $vars(a)$ , is the set of variables  $vars(pre(a)) \cup \bigcup_{e \in eff(a)} vars(e)$ . We say that an action  $a$  is **applicable** in a variable assignment  $\Sigma$  if and only if  $pre(a) \subseteq \Sigma$ . The **set of applicable actions** in the variable assignment  $\Sigma$  from the set of actions  $\mathcal{A}$  is denoted as  $\alpha(\Sigma, \mathcal{A})$ . The **application of a deterministic action**  $a = (pre(a), \{e\})$  in a variable assignment  $\Sigma$  such that  $pre(a) \subseteq \Sigma$ , denoted as  $\gamma(\Sigma, a)$ , is the variable assignment  $\gamma(\Sigma, a) = \{(v, x) \in \Sigma \mid v \notin vars(e)\} \cup e$  (the application is undefined if  $a$  is not applicable in  $\Sigma$ ). Given a deterministic action  $a = (pre(a), \{e\})$ , we define  $ha(a)$  as the variable assignment  $\gamma(pre(a), a)$ . It is easy to prove  $ha(a)$  is the largest set of facts that satisfies  $ha(a) \subseteq \gamma(\Sigma, a)$  for all variable assignment  $\Sigma$  that  $a$  is applicable in; for this reason, we say that  $ha(a)$  **necessarily holds** after applying action  $a$ .

The **application of a (non-deterministic) action**  $a$  in a variable assignment  $\Sigma$ , denoted as  $\delta(\Sigma, a)$ , is the set of variable assignments  $\delta(\Sigma, a) = \{\gamma(\Sigma, a_e^d) \mid e \in eff(a)\}$ .

Let  $\mathcal{V}$  be a set of variables. A **FOND planning domain**  $\mathcal{D}$  is a pair  $\langle \mathcal{V}, \mathcal{A} \rangle$ , where  $\mathcal{A}$  is a set of actions over  $\mathcal{V}$ . An **SAS<sup>+</sup> planning domain**  $\mathcal{D}^d$  is a pair  $\langle \mathcal{V}, \mathcal{A}^d \rangle$ , where  $\mathcal{A}^d$  is a set of deterministic actions over  $\mathcal{V}$ . A **state**  $s$  of the domain  $\mathcal{D}$  is a complete variable assignment over the set  $\mathcal{V}$ .

A **FOND (resp. SAS<sup>+</sup>) planning task**  $\mathcal{T}$  (resp.  $\mathcal{T}^d$ ) is a triple  $\langle \mathcal{D}, s_I, G \rangle$ , where  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  is a FOND (resp. SAS<sup>+</sup>) planning domain,  $s_I \in \mathcal{S}(\mathcal{V})$  is an **initial state** and  $G \subseteq \mathcal{F}(\mathcal{V})$  is a variable assignment representing a **goal**. When we refer to a planning task or a domain, we mean a non-deterministic task or domain, unless stated otherwise.

A sequence of deterministic actions  $\pi = \langle a_1, \dots, a_n \rangle$ ,  $\forall i \in \mathbb{N}, 1 \leq i \leq n : a_i \in \mathcal{A}^d$ , is called a **plan** for the SAS<sup>+</sup> planning domain  $\mathcal{D}^d$ . An **application of the plan**  $\pi = \langle a_1,$

$\dots, a_n \rangle$  in a variable assignment  $\Sigma$ , denoted as  $\gamma(\Sigma, \pi)$ , is a variable assignment  $\gamma(\Sigma, \pi) = \gamma(\gamma(\Sigma, a_1), \langle a_2, \dots, a_n \rangle)$ . If  $\pi = \langle \rangle$ , then  $\gamma(\Sigma, \langle \rangle) = \Sigma$ .

A state  $s$  is **reachable** in the SAS<sup>+</sup> planning task  $\mathcal{T}^d$  if and only if there exists a plan  $\pi$  for the domain  $\mathcal{D}^d$  such that  $\gamma(s_I, \pi) = s$ . Otherwise,  $s$  is **unreachable**. We say that  $\mathcal{T}^d$  is **solvable** if and only if some goal state  $s_G \supseteq G$  is reachable in the planning task  $\mathcal{T}^d$ . Otherwise,  $\mathcal{T}^d$  is **unsolvable**. A plan  $\pi$  is called a **goal plan** of the SAS<sup>+</sup> planning task  $\mathcal{T}^d$  if and only if  $\gamma(s_I, \pi) \supseteq G$ .

A **policy**  $\Pi$  for the domain  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  is a binary relation over the set of states  $\mathcal{S}(\mathcal{V})$  and the set of applicable actions of  $\mathcal{A}$ , i.e.,  $\Pi \subseteq \{(s, a) \mid s \in \mathcal{S}(\mathcal{V}), a \in \alpha(s, \mathcal{A})\}$ . The **set of all states related in**  $\Pi$  is the set  $\sigma(\Pi) = \{s \mid (s, a) \in \Pi\}$ . The  **$n$ -step application** of  $\Pi$  in a state  $s$ , denoted as  $\delta^n(s, \Pi)$ , is the set  $\delta^n(s, \Pi) = \bigcup_{s' \in \delta^{n-1}(s, \Pi)} \bigcup_{(s', a) \in \Pi} \delta(s', a)$  for  $n \geq 1$ , and  $\delta^0(s, \Pi) = \{s\}$ . We say that  $s'$  is **reachable** from  $s$  with a policy  $\Pi$  if and only if  $s' \in \delta^i(s, \Pi)$  for some  $i \geq 0$ . We say that  $s'$  is a **terminal state** for  $\Pi$  with respect to a state  $s$  if and only if  $s'$  is reachable from  $s$  with  $\Pi$  and  $s' \notin \sigma(\Pi)$ . The **set of all such terminal states** is denoted as  $\tau(\Pi, s)$ .

A policy  $\Pi$  is called a **weak goal policy** for  $\mathcal{T}$  if and only if  $\tau(\Pi, s_I) \cap \{s_G \in \mathcal{S}(\mathcal{V}) \mid G \subseteq s_G\} \neq \emptyset$ . A policy  $\Pi$  is called a **strong goal policy** for  $\mathcal{T}$  if and only if  $\tau(\Pi, s_I) \subseteq \{s_G \in \mathcal{S}(\mathcal{V}) \mid G \subseteq s_G\}$  and for each  $s \in \sigma(\Pi)$  at least one state that satisfies the goal  $G$  is reachable by  $\Pi$ .<sup>1</sup> A task  $\mathcal{T}$  is called **solvable** if and only if there exists a weak goal policy for the task  $\mathcal{T}$ . Otherwise, it is called **unsolvable**.

Let  $\Psi$  be a set of pairs  $(\Sigma, a)$ , where  $\Sigma$  is a variable assignment, and  $a \in \mathcal{A}$  is an action that is applicable in  $\Sigma$ . The set  $\Psi$  is called an **implicitly-defined policy** for the domain  $\mathcal{D}$ . We say that  $\Psi$  **implicitly defines** the policy  $\Pi$  if and only if  $\Pi = \bigcup_{(\Sigma, a) \in \Psi} \{(s, a) \mid s \in \mathcal{S}(\mathcal{V}), \Sigma \subseteq s\}$ .

## Non-deterministic Action Reversibility

Action reversibility, concerning the problem of whether the effects of an action can be undone by means of other actions, has been studied in deterministic settings (see Eiter, Erdem, and Faber (2008); Daum et al. (2016); Morak et al. (2020)).

In our work, we introduce the concept of action reversibility for non-deterministic actions in the FOND planning formalism. Non-deterministic action reversibility, in contrast to the deterministic variant, has to deal with non-deterministic actions—not only those whose effects we try to undo, but also those that might be caused by actions we apply to reverse the effects of the former action. In FOND planning, we usually consider two types of solutions, *weak* and *strong* (Cimatti et al. 2003). For a weak solution, we have a chance to achieve the goal if everything “goes well”, while for a strong solution, there is a guarantee that the goal will eventually be achieved. We adopt these notions to establish *weak and strong action reversibility* such that for weak reversibility we require that there is a chance to undo all action effects, while for strong reversibility this has to be guaran-

<sup>1</sup>Note that our definition coincides with the definition of strong cyclic solution (Def. 2.10) of Cimatti et al. (2003).

teed. Note that in this paper, we consider strong cyclic solutions that guarantee success if each effect of each non-deterministic action has a nonzero chance to occur if the action is applied (in other words, we consider the *fairness assumption*) (Aminof, De Giacomo, and Rubin 2020).

Inspired by the work of Morak et al. (2020), we classify the notion of (non-deterministic) action reversibility into three categories, *S-reversibility*, *uniform reversibility* and *universal reversibility*, which will be now introduced.

## S-reversibility

Naturally, it might be useful to study whether an action  $a$  is reversible in some subset of states  $S \subseteq \mathcal{S}(\mathcal{V})$ . That is, for each state from  $S$  in which  $a$  is applicable, we have to find a sequence of other actions that undoes the effects of  $a$ . This property is called *S-reversibility* (Morak et al. 2020).

To address non-deterministic actions, we introduce the notions of *weak S-reversibility* and *strong S-reversibility*, in the context of work of Cimatti et al. (2003) as described above. Informally speaking, we call an action *weakly S-reversible* if, in each state from  $S$  in which the action is applicable, the action’s effects can be undone by a sequence of (non-deterministic) actions while assuming the occurrence of the specific effect for each of the actions. We call an action *strongly S-reversible* if, in each state  $s \in S$  where  $a$  is applicable, the action’s effects are eventually undone by some policy  $\Pi_s$  having precisely  $s$  as its terminal state.

We also establish the notion of *action S-irreversibility*. It refers to a situation in which there is no way to undo any of the action’s effects (even by means of a weak solution). The following definition formalizes the above notions.

**Definition 1.** Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain,  $a \in \mathcal{A}$  be an action, and  $S \subseteq \mathcal{S}(\mathcal{V})$  be a set of states of the domain  $\mathcal{D}$ . The action  $a$  is called **weakly (resp. strongly) S-reversible** in the domain  $\mathcal{D}$  if and only if for each state  $s \in S$ ,  $pre(a) \subseteq s$ , there exists a policy  $\Pi$  for the domain  $\mathcal{D}$  such that for each  $s' \in \delta(s, a)$  it holds that  $s \in \tau(\Pi, s')$  (resp.  $\{s\} = \tau(\Pi, s')$ ) and for each state  $s''$  that is reachable by  $\Pi$  from the state  $s'$  it holds that  $s$  is reachable by  $\Pi$  from  $s''$ ). The action  $a$  is called **S-irreversible** in the domain  $\mathcal{D}$  if and only if for each state  $s \in S$ ,  $pre(a) \subseteq s$ , there does not exist a policy  $\Pi$  for the domain  $\mathcal{D}$  such that for each  $s' \in \delta(s, a)$  it holds that  $s \in \tau(\Pi, s')$ .

**Example 1.** Let  $\mathcal{D}$  be a planning domain with two variables door and window. Let door have domain  $\{\text{open}, \text{closed}\}$  and window have domain  $\{\text{open}, \text{closed}, \text{broken}\}$ .

Let VENT be a deterministic action with  $\{(\text{window}, \text{closed})\}$  as its precondition and  $\{(\text{door}, \text{open}), (\text{window}, \text{open})\}$  as effects. Furthermore, there are additional non-deterministic actions: CLOSE-WINDOW and CLOSE-DOOR with preconditions of a relevant entity being open and effects  $eff(\text{CLOSE-WINDOW}) = \{(\text{window}, \text{closed}), (\text{window}, \text{broken})\}$ ,  $eff(\text{CLOSE-DOOR}) = \{(\text{door}, \text{closed}), \emptyset\}$ .

The policy that applies CLOSE-DOOR in a state where both the door and the window are open and close-window in a state where the window is open and the door is closed is a weak  $\{(\text{window}, \text{closed}), (\text{door}, \text{closed})\}$ -reverse policy for the action VENT. Similarly, there is a weak  $\{(\text{window},$

closed), (door, open) $\}$ -reverse policy for the action VENT that applies close-window in a state where both the door and the window are open. Therefore, the action VENT is weakly  $\{(\text{window}, \text{closed}), (\text{door}, \text{closed})\}, \{(\text{window}, \text{closed}), (\text{door}, \text{open})\}$ -reversible.

As the application of both policies may eventually end in a broken window, neither of them is a strong reverse policy for the above set of states. Intuitively, close-window action is irreversible in any state in which the action is applicable since no action can “fix” the broken window.

Now, consider the same domain, except that there are two variants of the former CLOSE-WINDOW. First, CLOSE-WINDOW-DRAUGHT-FREE, which cannot break the window, and on top of the former action requires the door to be closed; and second, close-window, on top of the former action requires the door to be open, with the same effects as the former. In such a domain, the policy that applies CLOSE-DOOR if the door is not closed and CLOSE-WINDOW-DRAUGHT-FREE if the window is not closed (and the door is) is a strong  $\{(\text{window}, \text{closed}), (\text{door}, \text{closed})\}$ -reverse policy for the action VENT.

## Uniform Reversibility

A special case of *S-reversibility* is *uniform S-reversibility*, which refers to the existence of a common solution concept that can undo the effects of the action in question, for each state of  $S$  in which the action can be applied (Morak et al. 2020), i.e., the same solution concept reverses the effects of the action in question applied in any state of the set  $S$  where the action is applicable. Formally, the difference lies in the order of universal and existential quantifiers. The following definition formalizes the notions of weak and strong uniform *S-reversibility* for non-deterministic actions.

**Definition 2.** Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain,  $a \in \mathcal{A}$  be an action,  $S \subseteq \mathcal{S}(\mathcal{V})$  be a set of states of the domain  $\mathcal{D}$ . The action  $a$  is called **weakly (resp. strongly) uniformly S-reversible** in the domain  $\mathcal{D}$  if and only if there exists a policy  $\Pi$  for  $\mathcal{D}$  such that for each state  $s \in S$ ,  $pre(a) \subseteq s$ , it holds that for each  $s' \in \delta(s, a)$  we have  $s \in \tau(\Pi, s')$  (resp.  $\{s\} = \tau(\Pi, s')$ ) and for each state  $s''$  that is reachable by  $\Pi$  from the state  $s'$  it holds that  $s$  is reachable by  $\Pi$  from  $s''$ .

Having a single policy that can undo the effects of an action in all relevant situations (i.e., states from  $S$  in which  $a$  is applicable) is practical, as we do not need to consider multiple state-specific policies. Analogously to the notion of reverse plan in classical planning (Morak et al. 2020), we define the notion of *reverse policy*.

**Definition 3.** Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain and  $a \in \mathcal{A}$  be an action. A policy  $\Pi$  for the domain  $\mathcal{D}$  is called a **weak (resp. strong) S-reverse policy** for the action  $a$  if and only if the action  $a$  is weakly (resp. strongly) uniformly *S-reversible* by the policy  $\Pi$ .

**Example 2.** Consider the domain of Example 1. The policy which applies CLOSE-WINDOW if the window is open and CLOSE-DOOR if both are open is a weak uniform  $\{(\text{window}, \text{closed}), (\text{door}, \text{closed})\}, \{(\text{window}, \text{closed}), (\text{door}, \text{open})\}$ -reverse policy for VENT action (however, the policy relates some states with multiple actions).

## Universal Reversibility

Universal reversibility is a specific case of  $S$ -reversibility covering all states, i.e.,  $S = \mathcal{S}(\mathcal{V})$ . Informally speaking, if an action is universally reversible (or irreversible), its effects can always (or never) be undone.

**Definition 4.** Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain and  $a \in \mathcal{A}$  be an action. The action  $a$  is called **weakly (resp. strongly) universally (uniformly) reversible** in the domain  $\mathcal{D}$  if and only if the action  $a$  is weakly (resp. strongly) (uniformly)  $\mathcal{S}(\mathcal{V})$ -reversible in the domain  $\mathcal{D}$ . The action  $a$  is called **universally irreversible** in the domain  $\mathcal{D}$  if and only if the action  $a$  is  $\mathcal{S}(\mathcal{V})$ -irreversible in the domain  $\mathcal{D}$ .

**Example 3.** Consider the domain of Example 1. The states with a closed window and the door being open or closed are the only states where the action VENT is applicable. Recall that there is a weak reverse policy for any such state. Therefore, the action VENT is weakly universally reversible. The common policy of Example 2 proves the weak universal uniform reversibility of the action VENT.

**Example 4.** Now, as an example of strong universal uniform reversibility, let us consider the well-known Transport domain (Helmert, Do, and Refanidis 2010). In the domain, packages have to be transported from one location to another through various vehicles. Packages can be loaded and unloaded into vehicles without restrictions and the loading or unloading process can fail, leaving the package in its former position (at the location or inside the truck, respectively). The actions of the LOAD and UNLOAD “families” are strongly universally uniformly reversible, as if the package is successfully loaded or unloaded, respectively, we can always undo it by unloading, or loading the package back.

## On Combining Policies

Having defined the relevant notions, we will first establish some theoretical properties that follow from our definitions.

It can be easily derived that weak or strong (uniform)  $S$ -(ir)reversibility implies weak or strong (uniform)  $S'$ -(ir)reversibility if  $S' \subseteq S$ . Weak or strong uniform  $S$ -reversibility implies weak or strong  $S$ -reversibility, and strong (uniform)  $S$ -reversibility implies weak (uniform)  $S$ -reversibility. Furthermore, we can observe that if an action is weakly (resp. strongly)  $S_1$ - and  $S_2$ -reversible, then it is weakly (resp. strongly)  $(S_1 \cup S_2)$ -reversible. However, for uniform reversibility such an implication does not hold in general.

A more interesting property of any  $S$ -reverse policy is that it cannot assign any action to any state in  $S$  in which the “to be reversed” action is applicable.

**Example 5.** Consider a domain that describes the result of coin tossing. It has a variable whose domain is {heads, tails}, and TOSS-COIN (without precondition). The action has a strong universal uniform {(coin, head)}-reverse policy that applies TOSS-COIN as long as the state is {(coin, tails)}. The same holds symmetrically for tails. Hence, the action is weakly and strongly universally reversible. However, there is no weak or strong universal reverse policy, since, e.g., the union of the mentioned policies has no terminal state (it relates each state with some action).

**Proposition 6.** Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain,  $\Pi$  be a policy for the domain  $\mathcal{D}$ , and  $a \in \mathcal{A}$  be an action. If  $\Pi$  is a weak (resp. strong)  $S$ -reverse policy for the action  $a$ , then  $\{s \mid s \in S, pre(a) \subseteq s\} \cap \sigma(\Pi) = \emptyset$ .

*Proof.* Proof by contradiction. Assume that  $\{s' \mid s' \in S, pre(a) \subseteq s'\} \cap \sigma(\Pi) \neq \emptyset$ . Then, there is a state  $s \in \{s' \mid s' \in S, pre(a) \subseteq s'\} \cap \sigma(\Pi)$ . According to the assumption and since  $\{s\} \subseteq S$ , the action  $a$  is weakly (resp. strongly) uniformly  $\{s\}$ -reversible by the policy  $\Pi$ . Since  $s$  is also in  $\sigma(\Pi)$ ,  $s$  cannot be a terminal state of the policy  $\Pi$ . Therefore, for each  $s' \in \delta(s, a)$  we have  $s \notin \tau(\Pi, s')$ . This is in contradiction with the assumption of  $\Pi$  being a weak (resp. strong)  $S$ -reverse policy for the action  $a$ .  $\square$

An important consequence of Proposition 6 concerns universal uniform reversibility. It simply follows that when the set  $S$  is equal to the set  $\mathcal{S}(\mathcal{V})$ , then the situation of Proposition 6 can be simplified to the fact that the universal reverse policy can contain only the states in which the action we want to reverse is not applicable.

In contrast to deterministic planning which deals with action sequences, policies have a different structure as they assign appropriate actions to apply in relevant states. This property gives grounds to raise the question of how combining two policies affects  $S$ -reversibility of non-deterministic actions. In the case of weak reversibility, in which reverse policies are constructed along a sequence of actions, we can safely combine policies as long as the other policy does not compromise the assumption of Proposition 6.

**Theorem 7.** Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain,  $a \in \mathcal{A}$  be an action,  $\Pi_1$  be a weak  $S$ -reverse policy for the action  $a$  and  $\Pi_2$  be a policy for the domain  $\mathcal{D}$ . The policy  $\Pi_1 \cup \Pi_2$  is a weak  $S$ -reverse policy for the action  $a$  if and only if  $\{s \mid s \in S, pre(a) \subseteq s\} \cap \sigma(\Pi_2) = \emptyset$ .

*Proof.* The “if” part is proven by a contraposition. Let us assume that there is some state  $s' \in \{s \mid s \in S, pre(a) \subseteq s\} \cap \sigma(\Pi_2)$ . Then, as  $\sigma(\Pi_2) \subseteq \sigma(\Pi_1 \cup \Pi_2)$ ,  $s' \in \{s \mid s \in S, pre(a) \subseteq s\} \cap \sigma(\Pi_1 \cup \Pi_2)$ . That violates the assumption of Proposition 6 and hence the policy  $\Pi_1 \cup \Pi_2$  cannot be a weak  $S$ -reversible policy for the action  $a$ .

For the “only if” part, we can observe that extending a policy (by adding a pair of some state and some action) does not affect the reachability of any state that was reachable before (including the terminal states). From the assumption  $\{s \mid s \in S, pre(a) \subseteq s\} \cap \sigma(\Pi_2) = \emptyset$  and the fact that  $\{s \mid s \in S, pre(a) \subseteq s\} \cap \sigma(\Pi_1) = \emptyset$  (see Proposition 6), we get  $\{s \mid s \in S, pre(a) \subseteq s\} \cap \sigma(\Pi_1 \cup \Pi_2) = \emptyset$ . Therefore, all states  $\{s \mid s \in S, pre(a) \subseteq s\}$  are terminal with respect to a given starting state for the policy  $\Pi_1 \cup \Pi_2$  (see the definition of terminal states). Hence, the policy  $\Pi_1 \cup \Pi_2$  is a weak  $S$ -reverse policy for the action  $a$ .  $\square$

Regarding the strong  $S$ -reverse policy, since every strong  $S$ -reverse policy is also a weak  $S$ -reverse policy, the same conditions apply in this case as well. However, the conditions of Theorem 7 are not sufficient since, if the second policy assigns an action to a state that is considered by the first strong  $S$ -reverse policy, the union of such policies could

introduce a different terminal state; and this would violate the definition of strong  $S$ -reverse policy. Hence, to combine a strong  $S$ -reverse policy with a second one, the latter should not interfere with the former, as summarized in the following theorem.

**Theorem 8.** *Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain,  $a \in \mathcal{A}$  be an action,  $\Pi_1$  be a strong  $S$ -reverse policy for the action  $a$  and  $\Pi_2$  be a policy for the domain  $\mathcal{D}$ . If  $\sigma(\Pi_2) \cap \{s \mid s \in S, \text{pre}(a) \subseteq s\} = \emptyset$  and  $\sigma(\Pi_1) \cap \sigma(\Pi_2) = \emptyset$ , then  $\Pi_1 \cup \Pi_2$  is a strong  $S$ -reverse policy for the action  $a$ .*

*Proof.* Let  $s$  be a state  $s \in \{s'' \mid s'' \in S, \text{pre}(a) \subseteq s''\}$ . From the assumption of  $\Pi_1$  being a strong  $S$ -reverse policy for action  $a$ , we get  $\{s\} = \tau(\Pi_1, s')$  for all  $s' \in \delta(s, a)$ . Let  $\Delta^{s'}$  be a set of reachable states from the state  $s'$  by policy  $\Pi_1$ . The set can be divided into two disjoint sets  $\Delta^{s'} \cap \sigma(\Pi_1)$  and  $\Delta^{s'} \setminus \sigma(\Pi_1) = \tau(\Pi_1, s') = \{s\}$ .

Now, we show that if we add an arbitrary state-action pair  $(s'', a'') \in \Pi_2$  to the policy  $\Pi_1$ , the set of reachable states  $\Delta^{s'}$  from the state  $s'$  by the extended policy remains unchanged. Therefore, if the policy  $\Pi_1$  is extended by any state-action pair  $(s'', a'') \in \Pi_2$ , i.e.,  $\Pi'_1 = \Pi_1 \cup \{(s'', a'')\}$ ,  $\Delta^{s'}$  remains unchanged, because  $s'' \notin \Delta^{s'} \cap \sigma(\Pi'_1)$  (due to the fact that  $\sigma(\Pi_1) \cap \sigma(\Pi_2) = \emptyset$ ) and neither is in  $\Delta^{s'} \setminus \sigma(\Pi'_1) = \tau(\Pi_1, s') = \{s\}$  (due to the fact that  $\sigma(\Pi_2) \cap \{s''' \mid s''' \in S, \text{pre}(a) \subseteq s'''\} = \emptyset$ ).

Therefore, we can derive that  $\Delta^{s'}$  remains the same for the combined policy  $\Pi_1 \cup \Pi_2$ . Furthermore, we can see that no other state could become terminal since no other state became reachable, i.e.,  $\tau(\Pi_1 \cup \Pi_2, s') \subseteq \{s\}$ . From Theorem 7 we know  $s \in \tau(\Pi_1 \cup \Pi_2, s')$ . Hence,  $\Pi_1 \cup \Pi_2$  is a strong  $S$ -reverse policy for the action  $a$ .  $\square$

A practical consequence of the above theorems is that they provide conditions under which we can merge two reverse policies into one. If  $\Pi_1$  is a weak (resp. strong)  $S_1$ -reverse policy for some action  $a$  and  $\Pi_2$  is a weak (resp. strong)  $S_2$ -reverse policy for the action  $a$ ,  $\Pi_1 \cup \Pi_2$  is a weak (resp. strong)  $(S_1 \cup S_2)$ -reverse policy for  $a$  if the conditions of Theorem 7 (resp. Theorem 8) are satisfied.

## Determining Universal Uniform Reversibility

To determine universal uniform reversibility as well as universal irreversibility, we took inspiration from the work of Chrupa, Faber, and Morak (2021) who studied such cases of reversibility in the deterministic context. In summary, to determine universal uniform reversibility (for deterministic actions), it is sufficient to consider only actions that contain only variables present in the precondition of a “to be reversed” action. Such a property is practically very useful as it usually (considerably) simplifies the problem.

However, in the non-deterministic context, the theoretical findings of Chrupa, Faber, and Morak (2021) can only be partially adopted. Due to the different nature of policies, we show that allowing only actions operating over the set of variables present in the precondition of a “to be reversed” action while looking for weak or strong universal reverse policies is a sufficient but not necessary condition. In other

words, if a strong or weak reverse policy containing only “restricted” actions is found, then such a policy is universal, but the opposite implication generally does not hold (as we argue later in this section).

## Conditions for Universal Uniform Reversibilities

At first, we show that for an action to be strongly universally uniformly reversible, its effects may modify only variables that are also present in its preconditions. Note that an analogous claim also holds for the deterministic case (Chrupa, Faber, and Morak 2021). The intuition behind the claim is that if a variable is present in the action’s effects but not present in its precondition, then multiple states (each refers to a different value of the variable while assuming that the variable can have at least two different values) “collapse” into a single state referring to the value of the variable present in the action’s effects.

**Lemma 9.** *Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain such that  $\forall v \in \mathcal{V} : |\text{dom}(v)| \geq 2$  and  $a \in \mathcal{A}$  be an action. If  $\bigcup_{e \in \text{eff}(a)} \text{vars}(e) \not\subseteq \text{vars}(\text{pre}(a))$ , then the action  $a$  is not strongly universally uniformly reversible.*

*Proof.* Assume that there exists a strong universal reverse policy  $\Pi$  for the action  $a$ .

W.l.o.g., let  $e \in \text{eff}(a)$  be an action effect for which there exists  $v \in \text{vars}(e) \setminus \text{vars}(\text{pre}(a))$ . Since  $|\text{dom}(v)| \geq 2$ , there exist two distinct states  $s_1, s_2$  in which  $a$  is applicable and which differ in the value of the variable  $v$ . As states  $s_1$  and  $s_2$  differ only in the value of the variable  $v$  and since the effect  $e$  modifies  $v$ , we have  $s' = \gamma(s_1, a_e^d) = \gamma(s_2, a_e^d)$ .

As we initially assumed,  $\Pi$  is a strong universal reverse policy for the action  $a$  and this implies  $\tau(\Pi, \gamma(s_1, a_e^d)) = \{s_1\}$  and  $\tau(\Pi, \gamma(s_2, a_e^d)) = \{s_2\}$  (see Definition 2). However, since  $s' = \gamma(s_1, a_e^d) = \gamma(s_2, a_e^d)$ , we get  $\tau(\Pi, s') = \{s_1, s_2\}$ . This is in contradiction with  $s_1 \neq s_2$ .  $\square$

In the case of weak universal uniform reversibility, an analogous claim cannot be made. For a weak universal reverse policy  $\Pi'$ , we can derive an analogous claim as we did in the proof of Lemma 9, that is,  $\tau(\Pi', s') = \{s_1, s_2\}$ . This claim does not contradict the definition of weak uniform reversibility, as both cases, that is,  $s_1 \in \tau(\Pi', s')$  and  $s_2 \in \tau(\Pi', s')$ , might be satisfied.

As an example, consider a bowling domain, with a variable *bowled* with the domain  $\{\text{true}, \text{false}\}$  and an action that resets bowling pins that restores any combination of pins that are down or standing with all pins standing, and that sets *bowled* to *false*. The action is applicable if and only *bowled* is *true*. The action’s effects contain variables of pins that are not in its preconditions. There is another action that rolls the bowling ball, without any preconditions, and which can non-deterministically push over none, some, or all standing pins, and always sets *bowled* to *true*. The rolling action can weakly universally uniformly undo the reset action, as each state before the reset can be obtained, but it cannot strongly universally uniformly undo the reset action, as no state before the reset can be guaranteed to be obtained.

The following theorem shows that if we generate a weak or strong universal reverse policy in a restricted state and

action space that takes into account only variables that are present in the precondition of “to be reversed” action, then such a reverse policy is universal.

**Theorem 10.** *Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain,  $a \in \mathcal{A}$  be an action such that  $\bigcup_{e \in \text{eff}(a)} \text{vars}(e) \subseteq \text{vars}(\text{pre}(a))$ . If there exists an implicitly defined policy  $\Psi$  for the domain  $\mathcal{D}$  such that  $\forall (\Sigma', a') \in \Psi : \text{vars}(\Sigma') \subseteq \text{vars}(\text{pre}(a))$  and that the policy  $\Pi$  implicitly defined by  $\Psi$  is weak (resp. strong)  $S$ -reverse policy for the action  $a$  such that  $\{s \mid s \in S, \text{pre}(a) \subseteq s\} \neq \emptyset$  and  $\forall (s', a') \in \Pi : \text{vars}(a') \subseteq \text{vars}(\text{pre}(a))$ , then  $\Pi$  is a weak (resp. strong) universal reverse policy for the action  $a$ .*

*Proof.* Let  $s \in S$  be a state such that  $\text{pre}(a) \subseteq s$  and  $\Pi$  be a weak (resp. strong)  $\{s\}$ -reverse policy for the action  $a$  implicitly defined by  $\Psi$ .

To show that  $\Pi$  is a weak (resp. strong) universal reverse policy for the action  $a$ , we need to prove that for each state  $s' \in \mathcal{S}(\mathcal{V})$  such that  $\text{pre}(a) \subseteq s'$  the policy  $\Pi$  is also a weak (resp. strong)  $\{s'\}$ -reverse policy for the action  $a$ . If there does not exist any other state  $s' \in \mathcal{S}(\mathcal{V})$  such that  $\text{pre}(a) \subseteq s'$  and  $s \neq s'$ , then the policy  $\Pi$  is a weak (resp. strong) universal reverse policy for the action  $a$ . If there exists such a state  $s'$ , then we can observe that  $s'$  differs from  $s$  only in values of variables that are not part of  $\text{vars}(\text{pre}(a))$  (otherwise it would compromise the applicability of  $a$ ).

Assumption  $\bigcup_{e \in \text{eff}(a)} \text{vars}(e) \subseteq \text{vars}(\text{pre}(a))$  states that any effect  $e$  of the action  $a$  can modify only the variables present in  $\text{pre}(a)$ . Therefore, the values of the variables of  $\mathcal{V} \setminus \text{vars}(\text{pre}(a))$  remain unchanged after the application of the action  $a$  regardless of what effect takes place.

As for any state-action pair  $(s_x, a_x) \in \Pi$  it is the case that  $\text{vars}(a_x) \subseteq \text{vars}(\text{pre}(a))$ , none of the actions of the policy  $\Pi$  can change any variable from  $\mathcal{V} \setminus \text{vars}(\text{pre}(a))$  or require a specific value of any of such variables.

From the assumption stating  $\forall (\Sigma', a') \in \Psi : \text{vars}(\Sigma') \subseteq \text{vars}(\text{pre}(a))$  it holds that each variable assignment  $\Sigma'$  does not contain variables from  $\mathcal{V} \setminus \text{vars}(\text{pre}(a))$ . Therefore, the policy  $\Pi$  behaves independently on values of variables  $\mathcal{V} \setminus \text{vars}(\text{pre}(a))$  which means that  $\Pi$  yields the same outcome (apart from the values of the variables of  $\mathcal{V} \setminus \text{vars}(\text{pre}(a))$  that remain constant) from both states  $s$  and  $s'$ .

According to the assumption, if we apply  $\Pi$  in any state in  $\delta(s, a)$ , we may (resp. have to) return to the state  $s$ , which is terminal. As we have shown that  $\Pi$  and  $a$  can modify only the variables present in  $\text{pre}(a)$ , we can derive that for each state  $s' \in \mathcal{S}(\mathcal{V})$ ,  $\text{pre}(a) \subseteq s'$ ,  $\Pi$  is weak (resp. strong)  $\{s'\}$ -reverse policy for  $a$ . Hence, the policy  $\Pi$  is a weak (resp. strong) universal reverse policy for the action  $a$ .  $\square$

In contrast to deterministic universal uniform reversibility (Chrapa, Faber, and Morak 2021), the opposite implication, i.e., there is a weak/strong universal reverse policy only if Theorem 10’s assumption is true, does not hold.

Regarding weak universal uniform reversibility, Lemma 9 indicates that a “to be reversed” action might modify variables that are not present in its precondition and still be (potentially) weakly universally uniformly reversible. This observation can also be applied to other actions that are part of

the weak universal reverse policy. As an example, we might have an action whose non-deterministic effects contain all states in which the “to be reversed” action is applicable.

In the case of strong universal uniform reversibility, the issue with the opposite implication of Theorem 10 is in the possibility of merging policies. If the conditions of Theorem 8 are satisfied, we can merge two strong reverse policies to obtain another strong reverse policy applicable to the union of states of the individual policies. Hence, it is possible to obtain a strong universal reverse policy by combining more specific strong reverse policies.

As an example, let us have a domain where an agent can move between three locations –  $A, B, C$ . However, there is another variable that determines whether the agent can move from  $B$  to  $A$ , or from  $B$  to  $C$ . Moving from  $A$  to  $B$  and from  $C$  to  $A$  is not restricted. Depending on the value of the other variable, we undo the effects of moving from  $A$  to  $B$  by either moving directly back to  $A$ , or moving through  $C$ . These two reverse policies can be combined, as the states they are operating over are disjoint (because of different values of the other variable). The combined policy is a strong universal reverse policy for the action that moves us from  $A$  to  $B$ . As the other variable is involved in the policy and there is no other strong universal reverse policy, there is no  $\Psi$  of Theorem 10.

## Compilations

Theorem 10 provides us with a blueprint on how classical (respectively, FOND planning) can be leveraged in finding weak (resp. strong) universal reverse policies.

The following theorem shows how a weak universal reverse policy can be generated by means of classical planning by specifying multiple classical planning tasks (one for each determinization of the “to be reversed” action). Since, for weak reversibility, we only need to decide reachability of the former state, we replace each non-deterministic action with its respective determinizations. Moreover, we may consider only determinizations that operate only on the variables present in the precondition of the “to be reversed” action.

**Theorem 11.** *Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain and  $a \in \mathcal{A}$  be an action such that  $\bigcup_{e \in \text{eff}(a)} \text{vars}(e) \subseteq \text{vars}(\text{pre}(a))$ . If for each determinization  $a_e^d$  of the action  $a$  a plan  $\pi_e$  is a goal plan for the SAS<sup>+</sup> planning task  $\mathcal{T}_e^d = \langle \langle \text{vars}(\text{pre}(a)), \{(a')_e^d \mid a' \in \mathcal{A}, e \in \text{eff}(a'), \text{vars}((a')_e^d) \subseteq \text{vars}(\text{pre}(a)) \} \rangle, \text{ha}(a_e^d), \text{pre}(a) \rangle$ , then the action  $a$  is weakly universally uniformly reversible.*

*Proof (Sketch).* Each plan  $\pi_e = \langle a_{e_1}, \dots, a_{e_n} \rangle$  can be transformed into a policy  $\Pi_e = \{(\gamma(\text{ha}(a_e^d), \langle a_{e_1}, \dots, a_{e_{i-1}} \rangle), a_{e_i}) \mid i \in \mathbb{N}, 1 \leq i \leq n\}$ . Since for each effect  $e_i$  determined by the corresponding determinization  $a_{e_i}$  it is the case that  $\text{vars}(e_i) \subseteq \text{vars}(\text{pre}(a))$ , we can observe that none of the variables other than  $\text{vars}(\text{pre}(a))$  is modified in the process. Also, each determinization  $a_e^d$  of the action  $a$  modifies only the variables of  $\text{vars}(\text{pre}(a))$ . All policies resulting from transformed plans that are solutions of the SAS<sup>+</sup> tasks for each determinization  $a_e^d$  contain  $\text{pre}(a)$  as one of the terminal states. By combining all the policies into one (meet-

ing the conditions of Theorem 7) we obtain an implicitly-defined policy for the domain  $\langle \mathcal{V}, \{(a'_e)^d \mid a' \in \mathcal{A}, e \in \text{eff}(a'), \text{vars}((a'_e)^d) \subseteq \text{vars}(\text{pre}(a))\} \cup \{a\} \rangle$ . If we replace the determinizations by their corresponding stochastic actions of  $\mathcal{A}$ , we obtain an equivalent (from the perspective of weak reversibility) implicitly-defined policy  $\Psi$  for the domain  $\langle \mathcal{V}, \mathcal{A} \rangle$ . Then, a policy implicitly defined by  $\Psi$  is a weak universal reverse policy for the action  $a$ .  $\square$

A similar claim can be made about strong universal uniform reversibility, albeit in the context of FOND planning, since strong universal uniform reversibility needs to guarantee that the only terminal state is the state in which a “to be reversed” action is applied. The compilation of the problem of finding a strong universal reverse policy to a FOND planning task is analogous to the weak universal uniform reversibility case. In addition, to be able to merge the found policies, their sets of related states have to be pairwise disjoint, as Theorem 8 suggests.

**Theorem 12.** *Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain and  $a \in \mathcal{A}$  be an action such that  $\bigcup_{e \in \text{eff}(a)} \text{vars}(e) \subseteq \text{vars}(\text{pre}(a))$ . If for each determinization  $a_e^d$  exists a strong goal policy  $\Pi_e$  for the FOND planning task  $\langle \langle \text{vars}(\text{pre}(a)), \{a' \mid a' \in \mathcal{A}, \text{vars}(a') \subseteq \text{vars}(\text{pre}(a)) \} \rangle, \text{ha}(a_e^d), \text{pre}(a) \rangle$  such that the sets  $\sigma(\Pi_e)$  are pairwise disjoint, then action  $a$  is strongly universally uniformly reversible.*

*Proof (Sketch).* It can be observed that each policy  $\Pi_e$  is an implicitly defined policy that strongly universally uniformly reverses the determinization  $a_e^d$ . These policies can be combined (the sets of related states are pairwise disjoint) according to Theorem 8. Hence, the policy implicitly defined by  $\bigcup_{e \in \text{eff}(a)} \Pi_e$  is a strong universal reverse policy for  $a$ .  $\square$

We also consider the possibility of identifying universal irreversibility. The idea is derived from the work of Chrapa, Faber, and Morak (2021) that concerns the universal irreversibility of deterministic actions. Informally speaking, if for some effect of an action, we cannot reach the precondition of that action, then the action is universally irreversible. We can “project” the problem onto the variables related to a “to be reversed” action  $a$ , i.e., omit the others from the problem. In contrast to the above theorems, we do not discard actions that also operate on the other variables, but we project their preconditions and effects to  $\text{vars}(a)$ .

Let  $\Sigma$  be a variable assignment over the set of variables  $\mathcal{V}$  and let  $\mathcal{V}' \subseteq \mathcal{V}$  be a set of variables. The projection of  $\Sigma$  on the set of variables  $\mathcal{V}'$  is a variable assignment  $\Sigma|_{\mathcal{V}'} = \{(v, x) \in \Sigma \mid v \in \mathcal{V}'\}$ . Furthermore, the projection of an action  $a$  on the set of variables  $\mathcal{V}'$  is an action  $a|_{\mathcal{V}'} = (\text{pre}(a)|_{\mathcal{V}'}, \{e|_{\mathcal{V}'} \mid e \in \text{eff}(a)\})$ .

**Theorem 13.** *Let  $\mathcal{D} = \langle \mathcal{V}, \mathcal{A} \rangle$  be a planning domain and  $a \in \mathcal{A}$  be an action. If for any determinization  $a_e^d$  of the action  $a$  the SAS<sup>+</sup> planning task  $\langle \langle \text{vars}(a), \{((a'_e)^d)|_{\text{vars}(a)} \mid a' \in \mathcal{A}, e' \in \text{eff}(a')\} \rangle, \text{ha}(a_e^d), \text{pre}(a_e^d) \rangle$  is unsolvable, then the action  $a$  is universally irreversible.*

*Proof.* If the specified SAS<sup>+</sup> planning task is unsolvable, it means that the precondition of  $a$  is unreachable from some

of its effects. Since we focus only on a subset of variables, the applicability of actions is more optimistic than in the general case. So, if an abstract task created by projecting into a subset of variables is unsolvable, then the original task is unsolvable as well (Helmert, Haslum, and Hoffmann 2007). Hence, we can derive that if for any determinization of  $a$ , the specified SAS<sup>+</sup> planning task is unsolvable, then  $a$  is universally irreversible.  $\square$

## Experiments

The section presents empirical evidence on the existence of investigated phenomena in many benchmark domains. Based on the claims of previous sections, we have designed and performed the experiments for the resolution of the investigated classes of non-deterministic action reversibility.

We have evaluated our approaches on 20 FOND domains of two sets of benchmarks: one from the repository of the PRP planner (Muise, McIlraith, and Beck 2012; Muise, Belle, and McIlraith 2014; Muise, McIlraith, and Belle 2014), and second set proposed by Geffner and Geffner (2018); namely, Acrobatics (A.), Beam Walk (BE. W.), Blocks World (BL.), Bus Fare (B. F.), Climber (C.), Doors (D.), Earth Observation (E. O.), Elevators (EL.), Exploding Blocks World (E. BL.), Faults (FA.), First Responders (F. R.), Forest (FO.), Islands (IS.), Miner (M.), River (R.), Spiky Tire World (S. T.), Tire World (T.), Tire World Truck (T. TRU.), Triangle Tire World (TRI. T.), and Zeno Travel (Z.). To translate lifted benchmark domains into (grounded) formalism considered in the paper, we have selected one problem instance for each considered (lifted) domain and instantiated it with (modified) translator of PRP planner.

For each action, we initially check whether they satisfy the condition  $\bigcup_{e \in \text{eff}(a)} \text{vars}(e) \subseteq \text{vars}(\text{pre}(a))$ . If the action does not satisfy the condition, we use Theorem 13 to check its universal irreversibility. Based on the result, we either conclude “universally irreversible” or “we have not identified anything, besides the action is not strongly universally uniformly reversible” (since the action does not satisfy the conditions, from Theorem 10 we know that the action is not strongly universally uniformly reversible; Theorem 11 is inapplicable and Theorem 13 is an implication only). If the action satisfies the initial condition, we check if it is weakly universally uniformly reversible (which is a necessary condition for strong reversibility) by leveraging Theorem 11. Based on the result, strong universal uniform reversibility or universal irreversibility is checked using Theorems 12 and 13, respectively. If all of them “fail”, we conclude that “we have not identified anything”, denoted in Table 1 as “?”. To check weak reversibility and irreversibility, we use the LAMA planner (Richter and Westphal 2010) and to check strong reversibility, we use the PRP planner (both are built on top of the Fast Downward planner (Helmert 2006)).

All experiments<sup>2</sup> ran on a machine with an Intel® Core™ i7-7700HQ processor and 32 GB of DDR4 RAM operating at a frequency of 2400 MHz. The operating system was Ubuntu 22.04.3 in WSL 2 (version 2.0.9.0) of Windows 10.

<sup>2</sup>publicly available at: <https://gitlab.com/automated-planning/reversibility/papers/icaps-24-med-et-al>

Dom.	A	$\not\subseteq$ ( $\neg S$ )		$\subseteq$				
		I	?	W	FM	S	?	I
A.	5	0	0	4	1	3	1	0
BE. W.	7	0	0	6	2	4	0	1
BL.	190	0	185	5	0	5	0	0
B. F.	5	0	0	2	0	2	0	3
C.	3	0	0	0	0	0	0	3
D.	5	4	1	0	0	0	0	0
E. O.	27	0	0	21	0	21	0	6
EL.	41	3	30	8	0	8	0	0
E. BL.	85	35	45	0	0	0	5	0
FA.	51	25	26	0	0	0	0	0
F. R.	46	6	4	22	0	22	8	6
FO.	150	7	119	4	0	4	10	10
IS.	24	0	0	20	0	20	0	4
M.	179	0	3	158	0	158	0	18
R.	3	2	0	0	0	0	0	1
S. T.	211	3	78	124	0	124	6	0
T.	52	7	1	0	0	0	44	0
T. TRU.	24	0	10	10	0	10	0	4
TRI. T.	11	3	0	0	0	0	0	8
Z.	740	0	96	504	0	504	140	0

Table 1: Results of identification of action reversibility. The table headers depict proven properties: number of actions in the domain ( $|A|$ ), weak (resp. strong) universal uniform reversibility (W) (resp. S), universal irreversibility (I), whether  $vars(a) \subseteq vars(pre(a))$  ( $\subseteq$  or  $\not\subseteq$ ), and “?” indicates “no class of action (ir)reversibility was identified”. The column “FM” represents a proven weak universal uniform reversibility, where strong universal reverse policies were found for all determinizations, but due to the conditions of Theorem 8, we were unable to merge them. “ $\neg S$ ” stands for actions which are proven not strongly universally uniformly reversible.

## Results

Table 1 provides an overview of what actions have been identified as (weakly or strongly) universally uniformly reversible, universally irreversible, or unidentified.

The results indicate that the weakly universally uniformly reversible actions identified by Theorem 11 are likely also identified as strongly universally uniformly reversible by Theorem 12, since this happened in the vast majority of situations in our experiments. The only exceptions are three actions for which we have found strong universal reverse plans for each determinization (see Theorem 12), but we were unable to merge them into a general policy since the reverse plans had conflicting states.

On average, we were able to identify some class of (ir)reversibility in approximately  $68.30 \pm 33.64\%$  of actions, with the lowest relative amount of approx. 2.63% in the BL. domain, while in the BE. W., B. F., C., E. O., IS., R. and TRI. T. domains all the actions were identified in some class. In domains such as FA., FO. or E. BL., most of the actions did not satisfy the “ $\subseteq$ ” condition and hence could not be identified as weakly universally uniformly reversible through our theoretical study.

Overall, on average, approx.  $60.91 \pm 40.13\%$  of actions satisfied the “ $\subseteq$ ” condition on action’s preconditions. In an

“average” case, approx. 33.49% of actions were identified as universally irreversible, 34.81% as weakly universally uniformly reversible, 32.39% as strongly universally uniformly reversible, and 31.70% of actions remained undetermined (columns labeled with the question mark in Table 1) out of which 23.88% is proven not to be strongly universally uniformly reversible (with respect to the total number of actions). The relatively high number of undetermined actions is caused by the requirement for universality and uniformity. Although these requirements are practically desirable, actions might not always conform to them. Also, as we have shown, the methods are not theoretically complete, and hence some cases might not have been identified (as we observed in the case of universal irreversibility in T. domain).

On average, the measured mean compilation time per action in a domain is  $7.037 \pm 5.748$  milliseconds for weak,  $7.028 \pm 5.772$  milliseconds for strong, and  $9.113 \pm 5.715$  milliseconds for irreversibility reformulation with minimal, resp. maximum values, 0.994, 0.846, 3.031 milliseconds, resp. 17.337, 17.743, 22.441 milliseconds. As for runtimes for solving the reformulations, we have measured the time required to find a plan or policy, or to decide the (un)solvability of an abstract SAS<sup>+</sup> planning task for all determinizations. The mean runtimes are  $156.111 \pm 60.408$ ,  $47.712 \pm 18.495$  and  $186.015 \pm 81.118$  milliseconds to determine weak universal uniform reversibility, strong universal uniform reversibility and universal irreversibility according to Theorems 11, 12 and 13, respectively. The minimum values are 118.091, 22.970, 120.439, and the maximum values are 462.132, 87.549, and 711.832 milliseconds, respectively. The results show the practical viability of our methods.

## Conclusion

In this paper, we have conceptualized the notions of *weak and strong reversibility* of non-deterministic actions in FOND planning. These notions are inspired by weak and strong plans in FOND planning (Cimatti et al. 2003) and share the same meaning, i.e., weak reversibility refers to the possibility of undoing all effects of an action, while strong reversibility refers to the certainty of undoing all effects. We specifically focused on *universal uniform* cases that refer to the fact that effects of a non-deterministic action can always be undone by the same reverse policy. We proposed methods based on compiling the weak and strong universal uniform reversibility problem into classical or FOND planning, respectively, and we proposed a method for determining universal irreversibility via classical planning.

An experimental evaluation that we conducted on existing FOND benchmarks has shown that we were able to identify a type of reversibility for about 56.32% actions and the running time of any of the methods was in the lower hundreds of milliseconds on average. These results demonstrated practical usefulness of our methods in spite of their narrow focus.

In the future, we plan to investigate computational complexity of proposed classes of non-deterministic action reversibility. We also plan to focus on more general subclasses of non-deterministic action reversibility (e.g., strong *S*-reversibility) and we would also like to generalize the results for lifted representation of FOND planning tasks.



## Acknowledgements

This research is supported by the Czech Science Foundation (project no. 23-05575S), by the European Union under the project ROBOPROX (reg. no. CZ.02.01.01/00/22.008/0004590), by the Czech Ministry of Education, Youth and Sports under the Czech-Austrian Mobility programme (project no. 8J22AT003) and by the Grant Agency of Czech Technical University (project no. SGS24/140/OHK3/3T/13).

## References

- Aminof, B.; De Giacomo, G.; and Rubin, S. 2020. Stochastic Fairness and Language-Theoretic Fairness in Planning in Non-deterministic Domains. In Beck, J. C.; Buffet, O.; Hoffmann, J.; Karpas, E.; and Sohrabi, S., eds., *Proceedings of the Thirtieth International Conference on Automated Planning and Scheduling, Nancy, France, October 26-30, 2020*, 20–28. AAAI Press.
- Bäckström, C.; and Nebel, B. 1995. Complexity Results for SAS+ Planning. *Comput. Intell.*, 11: 625–656.
- Camacho, A.; Muise, C. J.; and McIlraith, S. A. 2016. From FOND to Robust Probabilistic Planning: Computing Compact Policies that Bypass Avoidable Deadends. In Coles, A. J.; Coles, A.; Edelkamp, S.; Magazzeni, D.; and Sanner, S., eds., *Proceedings of the Twenty-Sixth International Conference on Automated Planning and Scheduling, ICAPS 2016, London, UK, June 12-17, 2016*, 65–69. AAAI Press.
- Chrupa, L.; Faber, W.; and Morak, M. 2021. Universal and Uniform Action Reversibility. In Bienvenu, M.; Lakemeyer, G.; and Erdem, E., eds., *Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning, KR 2021, Online event, November 3-12, 2021*, 651–654.
- Chrupa, L.; McCluskey, T. L.; and Osborne, H. 2012. Determining Redundant Actions in Sequential Plans. In *IEEE 24th International Conference on Tools with Artificial Intelligence, ICTAI 2012, Athens, Greece, November 7-9, 2012*, 484–491. IEEE Computer Society.
- Cimatti, A.; Pistore, M.; Roveri, M.; and Traverso, P. 2003. Weak, strong, and strong cyclic planning via symbolic model checking. *Artificial Intelligence*, 147(1-2): 35–84.
- Cserna, B.; Doyle, W. J.; Ramsdell, J. S.; and Ruml, W. 2018. Avoiding Dead Ends in Real-Time Heuristic Search. In McIlraith, S. A.; and Weinberger, K. Q., eds., *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, AAAI-18*, 1306–1313. AAAI Press.
- Daum, J.; Torralba, Á.; Hoffmann, J.; Haslum, P.; and Weber, I. 2016. Practical Undoability Checking via Contingent Planning. In Coles, A. J.; Coles, A.; Edelkamp, S.; Magazzeni, D.; and Sanner, S., eds., *Proceedings of the Twenty-Sixth International Conference on Automated Planning and Scheduling, ICAPS 2016, London, UK, June 12-17, 2016*, 106–114. AAAI Press.
- Eiter, T.; Erdem, E.; and Faber, W. 2008. Undoing the effects of action sequences. *Journal of Applied Logic*, 6(3): 380–415.
- Faber, W.; Morak, M.; and Chrupa, L. 2021. Determining Action Reversibility in STRIPS Using Answer Set and Epistemic Logic Programming. *Theory and Practice of Logic Programming*, 21(5): 646–662.
- Geffner, T.; and Geffner, H. 2018. Compact Policies for Fully Observable Non-Deterministic Planning as SAT. In de Weerd, M.; Koenig, S.; Röger, G.; and Spaan, M. T. J., eds., *Proceedings of the Twenty-Eighth International Conference on Automated Planning and Scheduling, ICAPS 2018, Delft, The Netherlands, June 24-29, 2018*, 88–96. AAAI Press.
- Ghallab, M.; Nau, D. S.; and Traverso, P. 2004. *Automated planning - theory and practice*. Elsevier. ISBN 978-1-55860-856-6.
- Ghallab, M.; Nau, D. S.; and Traverso, P. 2016. *Automated Planning and Acting*. Cambridge University Press. ISBN 978-1-107-03727-4.
- Helmert, M. 2006. The Fast Downward Planning System. *Journal of Artificial Intelligence Research*, 26: 191–246.
- Helmert, M.; Do, M.; and Refanidis, I. 2010. International Planning Competition 2008 Deterministic Part. <https://ipc08.icaps-conference.org/deterministic/index.html>. Accessed: 2023-11-27.
- Helmert, M.; Haslum, P.; and Hoffmann, J. 2007. Flexible Abstraction Heuristics for Optimal Sequential Planning. In *Proceedings of the Seventeenth International Conference on Automated Planning and Scheduling, ICAPS 2007*, 176–183.
- Koehler, J.; and Hoffmann, J. 2000. On Reasonable and Forced Goal Orderings and their Use in an Agenda-Driven Planning Algorithm. *Journal of Artificial Intelligence Research*, 12: 338–386.
- Med, J.; and Chrupa, L. 2022. On Speeding Up Methods for Identifying Redundant Actions in Plans. In Kumar, A.; Thiébaux, S.; Varakantham, P.; and Yeoh, W., eds., *Proceedings of the Thirty-Second International Conference on Automated Planning and Scheduling, ICAPS 2022, Singapore (virtual), June 13-24, 2022*, 252–260. AAAI Press.
- Morak, M.; Chrupa, L.; Faber, W.; and Fiser, D. 2020. On the Reversibility of Actions in Planning. In Calvanese, D.; Erdem, E.; and Thiébaux, M., eds., *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning, KR 2020, Rhodes, Greece, September 12-18, 2020*, 652–661.
- Muise, C.; Belle, V.; and McIlraith, S. A. 2014. Computing Contingent Plans via Fully Observable Non-Deterministic Planning. In *The 28th AAAI Conference on Artificial Intelligence*.
- Muise, C.; McIlraith, S. A.; and Beck, J. C. 2012. Improved Non-deterministic Planning by Exploiting State Relevance. In *The 22nd International Conference on Automated Planning and Scheduling (ICAPS)*.
- Muise, C.; McIlraith, S. A.; and Belle, V. 2014. Non-Deterministic Planning With Conditional Effects. In *The 24th International Conference on Automated Planning and Scheduling*.
- Richter, S.; and Westphal, M. 2010. The LAMA Planner: Guiding Cost-Based Anytime Planning with Landmarks. *Journal of Artificial Intelligence Research*, 39: 127–177.
- Yoon, S. W.; Fern, A.; and Givan, R. 2007. FF-Replan: A Baseline for Probabilistic Planning. In Boddy, M. S.; Fox, M.; and Thiébaux, S., eds., *Proceedings of the Seventeenth International Conference on Automated Planning and Scheduling, ICAPS 2007, Providence, Rhode Island, USA, September 22-26, 2007*, 352. AAAI.