Generalizing Action Justification and Causal Links to Policies

Sarath Sreedharan[1], **Christian Muise**[2], **Subbarao Kambhampati** [3]

[1] Department of Computer Science, CSU
[2] School of Computing, Queen's University
[3] School of Computing & AI, ASU
sarath.sreedharan@colostate.edu, christian.muise@queensu.ca, rao@asu.edu

Abstract

We revisit two concepts popularly used within the context of classical planning, namely action justification and causal links. While these concepts have come to underpin some of the most popular notions of explanations in classical planning, these notions are still restricted to sequential plans. To address this shortcoming, we propose a generalization of these concepts that is applicable to state-action policies. We introduce algorithms that can identify justified actions and causal links contributed by such actions for policies generated for Fully Observable Non-Deterministic (FOND) planning problems. We also present an empirical evaluation that demonstrates the computational characteristics of these algorithms on standard FOND benchmarks.

1 Introduction

The question of whether an action is 'required' in a plan, is one that has received considerable attention within classical planning (Fink and Yang 1993; Kambhampati 1995). Roughly speaking, an action is generally understood to be required (or more commonly referred to as *justified*) if it contributes in some way to the achievement of the goal. While the notion of justified actions was introduced as a way to formalize the concept of minimal plans, the need to understand whether an action is truly required in a plan has become one of the most widely studied explanatory queries in planning (Seegebarth et al. 2012; Bercher et al. 2014). A related but distinct concept is that of a causal link (McAllester and Rosenblitt 1991), where a causal link captures the effects contributed by an action to a plan (in the form of a future precondition satisfied by that action). While a non-justified action could have a causal link associated with it, for planning problems without conditional-effects a well-justified action must contribute a causal link. As such, causal links have become the primary way of explaining why an action is required (Seegebarth et al. 2012).

Even with their wide use, the definitions of these concepts and related algorithms are closely tied to the fact that a solution to a classical planning problem takes the form of a sequence of actions. However, as we move to more general planning formalisms we will need to adopt more complex solution concepts. To the best of our knowledge, there has not been any formal efforts on mapping these concepts to more general planning solutions such as policies or controllers. This is unfortunate as the underlying phenomena captured by these concepts are present in these more general settings as well. After all, an action that was unnecessary in a plan can't automatically become required when one represents it using a policy.

In this paper, we will revisit the concept of justified actions, in particular well-justified actions, and propose a generalization that applies to stationary pure policies that map factored states to actions. In the same vein, we will provide a generalization of causal links which can be used as the basis to explain the contributions made by a well-justified action. Our choice to focus on state-action policies was primarily motivated by their generality. Many of the other popular solution representation schemes, including partial state policies, non-stationary policies, and controllers of various forms like history-based controllers and full programmatic controllers (which can support conditional statements and iterations), can be mapped back into a state-action policy. Thus one could directly use our methods for those solution forms as well. In keeping with the spirit of generality, we decided to center our formulation around Fully Observable Non-Deterministic (FOND) planning problems, as it makes the fewest assumptions about the underlying non-determinism of the problem. As such, one could directly apply all the results and algorithms provided in this paper to reinforcement learning and stochastic planning problems (by mapping it into a FOND problem where every action effect with non-zero probability is mapped into a possible non-deterministic outcome of the action). Our work represents novel contributions to both the theoretical understanding of non-deterministic action justifiability and explainability in this richer execution setting. In fact, our work represents an essential first step required for both formalizing and answering the explanatory query, "Why is an action a required in state s per the policy?".

To summarize, the main contributions of the paper are as follows:

1. We propose a generalization of well-justified actions that apply to policies. We additionally propose an algorithm to detect whether a given state and action pair is justified for a given FOND policy.
2. We present a generalization of causal links that applies to

policies and also develop an algorithm to extract causal chains contributed by a well-justified action. To the best of our knowledge, this represents the first formalization of the term causal links within the context of policies in general.

3. Finally, we present an evaluation of both algorithms on policies generated for some standard FOND benchmarks.

2 Background

We will be focusing on cases where the planning problem may be represented as a fully observable non-deterministic planning problem (FOND). Such models may be represented in declarative form using PDDL variants that use 'oneof' effects (Bryce and Buffet 2008). Mathematically, we expect a FOND model to be represented by a tuple of the form $\mathcal{M} = \langle F, A, I, G \rangle$, where F is a set of propositional fluents that is used to define the state space for the planning problem ($S = 2^F$); A is the set of actions available to the agent; $I \subseteq F$ is the initial state from which the agent needs to try achieving the goal; and $G \subseteq F$ is the goal specification and any state that satisfies the goal specification (i.e., $G \subseteq s$) is considered to be a valid goal state. To simplify the discussion, without loss of generality, we will assume G is a singleton set consisting of a single goal atom. Overloading the notation a bit, we will also use the symbol G to denote the goal atom. Each action $a \in A$ is further defied by a tuple $a = \langle pre_a, \mathbb{E}(a) \rangle$. In this action definition, pre_a stands for the preconditions for executing the action and $\mathbb{E}(a)$ the set of possible effects. In this paper, we will exclusively focus on positive conjunctive preconditions, as such we will represent each preconditions as a subset of fluents. $\mathbb{E}(a) = \{\langle add_a^1, del_a^1 \rangle, ..., \langle add_a^k, del_a^k \rangle\}$ represents the set of mutually exclusive effects that could occur as the result of executing the action a and $add_a^i \subseteq F$ and $del_a^i \subseteq F$ correspond to the add and delete effect corresponding to the i^{th} effect. With the action definitions in place we can also define the set of transitions possible under this action definition (denoted as $\mathbb{T}$). In particular, we will define a transition $\langle s_1, a, s_2 \rangle$ to be possible (denoted as $\langle s_1, a, s_2 \rangle \in \mathbb{T}$) if

$$pre_a \subseteq s_1 \text{ and } \exists j, \text{ such that}$$
$$\langle add_a^j, del_a^j \rangle \in \mathbb{E}(a),\ s_2 = (s_1 \setminus del_a^j) \cup add_a^j$$

Throughout this paper, we will focus on cases where non-determinism is considered to be fair, i.e., for every non-deterministic action every possible effect is guaranteed to occur infinitely often if the action is executed infinitely often. We can also use the same formalism to capture deterministic domains by restricting ourselves to cases where the effect set for each action is a singleton set. This paper will only focus on cases where action effects are free from conditional-effects.

A solution to a FOND problem takes the form of a policy that maps a state to an action, usually denoted by a function $\pi : S \rightarrow A \cup \{a^\emptyset\}$, where $a^\emptyset$ is an artificial empty action assigned to states that are either not supported by the policy or are goal states (with empty preconditions and actions). In this paper, we will focus on deterministic and stationary policies, where the deterministic term refers to the fact that the policies map a state to a single action and the stationarity refers to the fact that the mapping from state to action does not change with time. Additionally, for notational simplicity we will sometimes use set notations to capture the policy. Specifically, we will say $\langle s_i, a_i \rangle \in \pi$, if $\pi(s_i) = a_i$. Additionally, we will refer to any state action pair $\langle s_i, a_i \rangle \in \pi$ as a policy step.

A concept that will be central to the main of the techniques are traces supported by a given policy. We will refer to a state action state sequence of the form $\tau = \langle s_1, a_1, ..., s_k \rangle$ as a trace supported by a policy π if for every s_i, where $i \neq k$, we have $\pi(s_i) = a_i$, $\langle s_i, a_i, s_{i+1} \rangle \in \mathbb{T}$. A trace is said to be a goal achieving trace if $G \subseteq s_k$ and a state state s_j is said to be reachable from s_i if there exists a trace of the form $\tau = \langle s_i, a_i, ..., a_j, s_j \rangle$. We will also refer to the sequence of action that appears in the trace, as the action sequence corresponding to the given trace.

In terms of a valid policy for a FOND problem, the literature generally differentiates between weak solutions, strong and strong-cyclic solutions (Cimatti et al. 2003). Weak solutions are policies such that there exists at least one goal-achieving trace from the initial state. A policy is said to be strong-cyclic if the goal is reachable from all states reachable from the initial state. Finally, a policy is said to be a strong solution if we can again guarantee that goal is reachable from all states reachable from the initial state, but additionally, now we require that a state can never be repeated in any given goal-reaching trace. However, in this paper we will not differentiate between these specific classes of solutions and all methods studied here are equally valid for all policy classes.

While many flavors of action justification has been studied in the literature (Fink and Yang 1993), this paper will focus exclusively on ***Well-Justified*** actions. *An action is said to be well-justified if the removal of that action will cause the resultant plan to be invalid*. A plan is said to be well-justified if every action in the plan is well-justified. Our choice of well-justified actions was motivated by the fact that it corresponded to one of the most widely used within the explanation literature. Section 7 provides a discussion of the various instances of the use of these explanation types within classical planning.

Next let us take a quick look at causal links, particularly as applied to totally ordered plans. Specifically a totally ordered plan $\mathcal{P}$ can be represented as a sequence of the form $\mathcal{P} = \langle a{:}1, ..., a'{:}j \rangle$, where each step of the plan ($a{:}i$) is captured using the action label (a) and the step count (i). Now a causal link is said to exist between two steps $a{:}i$ and $a'{:}k$ if there exists a precondition for the action a' that is provided by the add effect of a. The causal link between the action is denoted as $a{:}i \rightarrow_p a'{:}j$, where p is the fact being 'produced' by action $a{:}i$ and being 'consumed' by $a'{:}j$. Each causal link is assumed to be not threatened by any other action between the two steps. However, attaching a causal link to an action doesn't mean it is well-justified. In fact one could associate a causal link to a redundant action if its effect corresponds to a future precondition, even if a future action could also satisfy that effect or if that fact is already true in the current state. So we will focus on a stronger notion, that we will term ***required causal links***, which will additionally require that the causal link contributed by the current action cannot have been contributed by any action before the consumer action (including actions prior to the producer action). This is equivalent to the notion of exhaustive causal links as discussed by Kambhampati (1994). Since in this paper, one of our focus will be on identifying a single causal link contributed by the action, we will restrict our attention to ***minimal-length required causal links***, i.e., ones where the consumer is the closest to the producer action per their position in the plan.

Causal Chain Explanations As per Seegebarth et al. (2012), causal chain explanations are designed to address explanatory queries of the type

"Why is the plan step 'a:o' "necessary" for π to constitute a solution?"

Here the explanation takes the form of a sequence of causal links that originates at the plan step in question and terminates at the goal. We will denote a causal link chain explanation as

$$a{:}o \rightarrow f_i \rightarrow ... \rightarrow f_k,$$

where each link of the form $f_i \rightarrow f_{i+1}$ correspond to a causal link and f_k is part of the goal. Specifically, for each link of the form $f_i \rightarrow f_{i+1}$ there would exist a causal link of the form $a{:}i \rightarrow_{f_{i+1}} a'{:}i+1$, where f_i is part of the precondition for the action a. In this paper, we will see how our generalized notions of causal links can be leveraged to also generate a generalized causal link explanation.

3 Motivating Example

As a running example throughout the paper consider the policy generated by a futuristic daily planner that takes into account all the possible contingencies of the day and comes up with a policy that will get you to the office in time. The policy starts with you at home and as the first action, the daily planner recommends you to start the day by placing a call to your local baker for a dozen of the day's special donuts. At the end of this action, you will find yourself at home having ordered a dozen of maple-glazed donuts or a dozen strawberry sprinkle donuts with a coupon for a free milkshake. Now based on the outcome of this action the policy now requires you to take different routes to the office, with different potential branching points owing to the various non-determinisms in the world. Figure 1 presents a high-level overview of this policy with its various contingencies. Regardless of your personal feelings toward fried pastry, you may be confused as to why the planner might be asking you to take the time to buy donuts when you should be trying to get as early as possible to the office parking lot to get a free parking spot. Looking at the immediate actions that follow, one may be forgiven to think that the action is just a random action thrown into a seemingly bloated plan created by a faulty planner. But as we will see throughout the rest of the paper, this action is in fact required in this policy in so far as it is required for the goal achievement. Additionally, you may not have the patience to go through each possible trace corresponding to the multitude of ways the world may evolve and how they may feed into your goal of getting to your office. Ideally, you would want to be able to leverage mechanisms like causal link explanations that demonstrate how the action contributes to the eventual goal.

4 Generalizing Well-Justified Actions

As discussed in the motivational example, our problem starts with a user of a planning system being presented with a policy. Once the policy is provided, the user may identify some policy step whose role they are unsure of. Thus the user may turn to the system to understand why that step is required in the policy. Once such a step is presented, the first order of business is to identify whether the step is in fact required in the policy. We will say that a step is required in the policy if it is well-justified. Repeating the definition in the context of sequential plans, one could informally say that an action 'a' is well-justified at state 's' if without the execution of action 'a' at state 's' the goal could not have been achieved by the rest of the policy. However, this is not an operationalizable description of the property as by the very nature of policy as a solution concept, the execution of an action is necessary as the change of the state is needed to enable the execution of the rest of the policy. While in the context of sequential plans, one could meaningfully talk about removing an action from the sequence and then testing whether the remaining plan is valid or not, it is unclear how one could perform such transformations over a policy. At the same time, it is worth remembering that the concept of whether or not an action is well-justified at a particular policy step is still a relevant question to ask. After all, it would make no sense to claim that one could make a non-well-justified sequential plan well-justified by just mapping it to a policy. In this paper, we will try to propose a formal definition of this concept that leverages the fact that from any given state, one could characterize how the policy contributes to the goal by considering all the goal-achieving traces.

Definition 1. *An action 'a' is said to be* **required** *(or equivalently* **well-justified**) *at a state 's' for a policy 'π' to achieve a goal G under a given model $\mathcal{M}$, if for every every action sequence corresponding to a goal-reaching trace originating from state 's', removal of actions corresponding to the policy step in question will result in an invalid sequence, in so far that there exists no valid trace possible under $\mathcal{M}$ that corresponds to that sequence.*

In our example discussed above, the action '`order_donuts`' will be required if each of it outcome contributes at least one useful fact that may be needed by some future actions. It is easy to see why the above definition is a possible generalization of the existing notions of well-justified actions in classical planning. In fact we can see that when one maps a plan into a policy, the well-justified actions of the original plan aligns with well-justified actions for the policy. Specifically, we can state that

Proposition 1. *Let $\mathcal{P} = \langle a_1{:}1, ..., a_k{:}n\rangle$ be a valid plan for a deterministic model $\mathcal{M}$ and π a policy such that, we will have $\langle s_i, a_j{:}i\rangle \in \pi$ when $a_j{:}i$ is part of $\mathcal{P}$ (say corresponding to step i) and s_i can be obtained by executing plan*

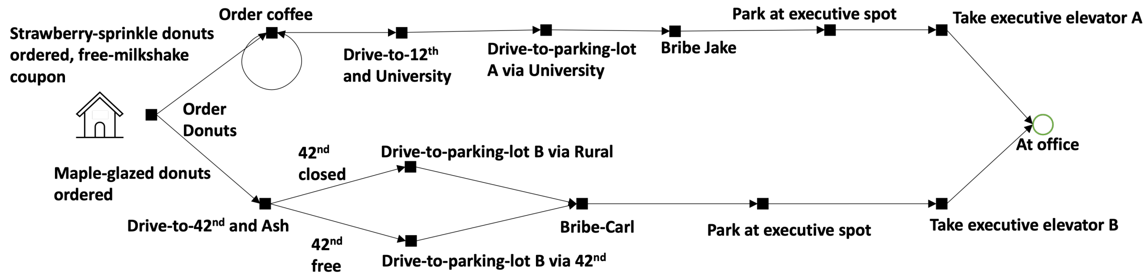

Figure 1: A simple overview of the daily planner policy, which highlights the actions that are determined by the policy along with some of the non-deterministic effects caused by the action.

prefix of length $i-1$ at the initial state I. Any state action pair $\langle s_i, a_j{:}i\rangle \in \pi$, is well-justified per Definition 1 if and only if the corresponding step in plan $\mathcal{P}$ is well-justified per the definition provided by Fink and Yang (1993) (in that the removal of that plan step results in an invalid solution).

The proof for the proposition is trivial given the fact that for a deterministic domain, the policy π can only generate a single goal reaching trace. Additionally, the action sequence corresponding to that trace is the same as the original plan $\mathcal{P}$. Thus an action being well-justified in the policy means that its removal from the trace action sequence and by extension $\mathcal{P}$, renders it invalid. Thus establishing the 'if' part of the statement and we can employ a similar line of reasoning to establish the 'only if' part of the statement.

However, it is worth noting that the notion of an action being required is an extremely strong condition, and there could very well be goal-reaching policies where none of the actions are required (a fact that is true for "well-justified" actions in classical planning as well). One could also look at weaker notions of how an action contributes to a goal (for example if the action 'a' is well-justified in at least one of the traces or 'a' may be well-justified for some subset of traces), however, that also means that one could in principle build a valid weak solution with the rest of the policy while ignoring the current action. We will leave the investigation of such weaker forms of justifications and their correspondence to existing notions of justifications for future work.

4.1 Identifying Well-Justified Actions

To identify whether an action is required at a state, we will leverage a compilation based method that will generate a modified deterministic planning problem, whose unsolvability will help us detect whether an action is required at a given policy step.

The compilation will form an updated planning model that will only allow actions allowed under a given policy. Additionally the compilation will maintain two copies of each fluent, one of which we will use to detect whether an action contributes to a future precondition.

For a given model $\mathcal{M} = \langle F, A, I, G\rangle$ and a query regarding the use of action a in state s for policy π, we will be creating a new model

$$\mathcal{M}^J_{(\pi,\langle s,a\rangle)} = \langle F^\phi, A^J_{(\pi,\langle s,a\rangle)}, I^J_{(\pi,\langle s,a\rangle)}, G\rangle.$$

In regards to the duplicate fluents, the new fluent set F^ϕ, now contains two copies for each fluent $f \in F$ that was part of the original problem definition. We will use the label f to capture one of the copy and denote the other copy with the notation f^ϕ. We will maintain the mapping between the two copies using the function $\phi : f \mapsto f^\phi$ and also overload the function to also apply to sets, i.e, $F^\phi = F \cup \phi(F)$

In regards to the actions, $A^J_{(\pi,\langle s,a\rangle)}$ represents the new set of actions. In particular, the new model will have an action for each state action pair that is part of the policy in question, i.e.,

$$A^J_{(\pi,\langle s,a\rangle)} = \{a_i^{s_j} \mid \langle s_j, a_i\rangle \in \pi \setminus \{\langle s,a\rangle\}\} \cup \{a^J_{\langle s,a\rangle}\}$$

. Where an action $a_i^{s_j} \in A^\pi_{\langle s,a\rangle}$ will encode the fact that this copy of the action a_i is meant to be executed only in state s_j and may also have preconditions that need only be a subset of the fluents that are true in s_j. The former is captured in terms of the fluents belonging to the set $\phi(F)$ and the latter by using fluents from the original set F. Similarly, action effects will now include copies of the original effects of the action in terms of both fluent sets, thereby allowing us to capture both the action's capability of allowing the continued execution of the policy while allowing us to maintain a separate accounting of how the action contributes to the preconditions of future actions. On the other hand the definition of action copy $a^J_{\langle s,a\rangle}$, corresponding to the query state action pair, is similar except that the action effect only includes a copy corresponding to the set $\phi(F)$. This means, that the action can only contribute to the policy state component of the precondition of the future actions.

More formally, the action will be defined as

$$a_i^{s_j} = \langle pre_{a_i^{s_j}}, \mathbb{E}(a_i^{s_j})\rangle,$$

such that

$$pre_{a_i^{s_j}} = \phi(s_j) \cup pre_{a_i}, \text{ and}$$

$$\mathbb{E}(a_i^{s_j}) = \{\langle \phi(add^m_{a_i}) \cup add^m_{a_i}, \phi(del^m_{a_i}) \cup del^m_{a_i}\rangle \mid \langle add^m_{a_i}, del^m_{a_i}\rangle \in \mathbb{E}(a_i)\}.$$

Similarly $a^J_{\langle s,a\rangle} = \langle pre_{a^J_{\langle s,a\rangle}}, \mathbb{E}(a^J_{\langle s,a\rangle})\rangle$, such that

$$pre_{a^J_{\langle s,a\rangle}} = \phi(s) \cup pre_a$$

$$\mathbb{E}(a^J_{\langle s,a\rangle}) = \{\langle \phi(add^m_a), \phi(del^m_a)\rangle \mid \langle add^m_a, del^m_a\rangle \in \mathbb{E}(a)\}$$

One point to note here is that to effectively constrain application of actions to specific states in the policy, we have to not only consider facts that are true in the state but also the ones that are false. We can still use our positive precondition formulation to support this by using the standard compilation technique to compile away negative preconditions. Since this is a standard technique, we will not include this as part of our formalization, but the reader is advised to keep in mind that when we say $\phi(s)$ is part of the precondition it also includes new positive fluents that correspond to the fluents that may be false in state s (with the necessary changes made to the effects as well).

Finally, the new initial state corresponds to the state s part of the query and contains fluents from both the original fluent set and the new copy fluent set and the goal is the same as the original problem.

$$I^J_{(\pi,\langle s,a\rangle)} = s \cup \phi(s).$$

Now the resultant model $\mathcal{M}^J_{(\pi,\langle s,a\rangle)}$ is still a non-deterministic planning domain. However, for the purposes of identifying whether an action is well-justified we only need to consider a determinization of this model. In particular, we will consider an all outcome determinization (Yoon, Fern, and Givan 2007) of the model $\mathcal{D}(\mathcal{M}^J_{(\pi,\langle s,a\rangle)})$. Effectively the determinization will create a separate deterministic action for each outcome of the action. So in the case of our running example, now there will be two different copies for the order_donuts action. One whose effect only makes has_maple_glazed_donuts true and the other one which makes has_strawberry_sprinkle_donuts and has_milkshake_coupon. Given the nature of the determinization, the set of goal-reaching traces for the model $\mathcal{M}^J_{(\pi,\langle s,a\rangle)}$ would exactly correspond to the set of valid plans for the deterministic model $\mathcal{D}(\mathcal{M}^J_{(\pi,\langle s,a\rangle)})$. This brings us to the proposition

Proposition 2. *An action 'a' is required at a state 's' for a policy 'π' (where requirement is defined as per Definition 1), if and only if the modified planning problem $\mathcal{D}(\mathcal{M}^J_{(\pi,\langle s,a\rangle)})$ is unsolvable.*

Proof Sketch. To show the validity the statement, it is worth remembering that by construction every valid plan in $\mathcal{D}(\mathcal{M}^J_{(\pi,\langle s,a\rangle)})$ corresponds to a possible trace for the policy by π. Additionally through the way we specify $a^J_{\langle s,a\rangle}$, we have essentially removed the ability to the step in question to support future preconditions of future actions specified with fluents F. If in fact the action was well-justified then the removal of the step from every trace would render them invalid and by extension our construction of $\mathcal{D}(\mathcal{M}^J_{(\pi,\langle s,a\rangle)})$ for a well-justified policy state action pair would render every possible valid plan invalid. This establishes the 'if' part of the statement. Similarly unsolvability of the planning problem points to the fact that the removal of the step renders every valid plan invalid and by extension each possible action sequence corresponding to the possible goal reaching traces. Thus establishing the fact the state action pair was well-justified per Definition 1 and hence proving the only if part of the statement. □

5 Generalizing Required Causal Links

With the question of how one could detect when an action may be contributing to the goal solved, the next obvious question to ask would be to see how one could capture the contributions made by the action. As such, the next concept we are interested in generalizing is that of causal links contributed by the action. One obvious strategy could be to just enumerate all possible goal-reaching traces and present a causal link contributed by the action in each trace. However, in this paper our primary interest would be in developing a method that summarizes the actions' contribution to the policy as a whole. The conciseness of our proposed generalization also makes it better suited for applications like explanations, since any techniques that might try to iterate over possible traces could overwhelm users even in the simplest domains. Also it is worth remembering that even a non-well-justified action can contribute a causal link (depending on how one assigns the causal link), so our focus will be in identifying and generalizing required causal links as discussed earlier.

Our abstract representation will leverage necessary subgoals made feasible by the execution of the action. In particular, for grounding the concept of a necessary subgoal we will build on the notion of a policy landmark introduced in (Sreedharan, Srivastava, and Kambhampati 2020), which defined policy landmarks as being facts and their corresponding ordering that needs to be satisfied by every valid goal reaching trace that can be sampled from the initial state. In our case, we will use a more restricted version of policy landmarks, one that additionally requires that the landmarks we focus on are required as preconditions for different actions or in the goal. We will refer to such policy landmarks as *causal policy landmarks*. Additionally we will only focus on traces originating from the policy state s in question. We will capture such a causal policy landmark set with the notation $\mathbb{L} = \langle \mathcal{L}, \prec\rangle$. Here $\mathcal{L} \subseteq F$ is the set of possible landmark facts and $\prec$ defines a partial ordering over these facts, such that $f_1 \prec f_2$ captures the fact that f_1 should appear before f_2.

By focusing on causal policy landmarks, we effectively filter out any facts that just appear as side-effects of some actions and only focus on the facts used by actions in the policy. In the case of our daily planner domain, a subgoal made possible by the action would be `security-guard-bribed` and our causal link would be represented as $\langle I,$`Order-donuts`$\rangle$ $\rightarrow$`security-guard-bribed`. This is because Carl can be bribed only with map glazed donuts, while Jake is happy to take a strawberry sprinkle donut provided you also give him a cup of coffee. More formally

Definition 2. *For an action 'a' required at state 's' for a* *policy 'π' under a given model $\mathcal{M}$ and a causal policy landmark set $\mathbb{L}$, a fact $f \in \mathcal{L}$ is said to be the **proximal required subgoal** if f cannot be achieved in any goal reaching trace once action a is removed and there exist no fluent f' such that $f' \prec f$ that satisfies the same requirement. Additionally, we will refer to f as constituting a **minimal-length generalized required causal link** contributed by the action, represented as $\langle s, a\rangle \rightarrow f$.*

This brings us to the first property regarding generalized required causal link, namely there always exist one for a well-justified action.

Proposition 3. *For an action 'a' required at state 's' for a policy 'π' under a given model $\mathcal{M}$ and a causal policy landmark set $\mathbb{L}$, there must exist a fluent $f \in \mathcal{L}$ such that $\langle s, a\rangle \rightarrow f$.*

This result trivially follows from the fact that the goal fluent is part of the causal landmarks and we know that the removal of the action leads to goal not being achievable anymore. So if all the landmarks that are preceding the goal in the landmark set fail, the goals will satisfy the requirements outlined in Definition 2.

5.1 Identifying Proximal Required Subgoal

We will again be leveraging a compilation based method as in the previous section. In particular for a given model $\mathcal{M} = \langle F, A, I, G\rangle$ and a query regarding the use of action a in state s for policy π, we will be creating a new model

$$\mathcal{M}^C_{(\pi,\langle s,a\rangle)} = \langle F^\phi, A^C_{(\pi,\langle s,a\rangle)}, I^C_{(\pi,\langle s,a\rangle)}, G\rangle$$

.

Where F^ϕ and $I^C_{(\pi,\langle s,a\rangle)}$ stays the same as the previous compilation for $\mathcal{M}^J_{(\pi,\langle s,a\rangle)}$ (with $I^C_{(\pi,\langle s,a\rangle)} = I^J_{(\pi,\langle s,a\rangle)}$). Unlike $\mathcal{M}^J_{(\pi,\langle s,a\rangle)}$, for the actions in this setting we will not remove the effects defined over F from the action corresponding to the policy state action pair being queried about, i.e.,

$$A^C_{(\pi,\langle s,a\rangle)} = \{a_i^{s_j} \mid \langle s_j, a_i\rangle \in \pi\},$$

where each $a_i^{s_j}$ is defined as before.

As before, we will be focusing an all outcome determinization of the model $\mathcal{D}(\mathcal{M}^C_{(\pi,\langle s,a\rangle)})$ and use it to identify the causal landmark set (Keyder, Richter, and Helmert 2010). Causal landmarks are landmarks that always appear in the precondition of an action. However, for our purposes, we can't directly use the causal landmark set generated from $\mathcal{D}(\mathcal{M}^\pi_{\langle s,a\rangle})$ as the preconditions of the actions in the model also contain state descriptions. As such the landmarks directly calculated from $\mathcal{D}(\mathcal{M}^\pi_{\langle s,a\rangle})$ may contain facts that are not part of any action preconditions. Our use of a distinct set of fluents to capture the state and preconditions allows us to filter out such landmarks. Specifically, let $\mathbb{L} = \langle \mathcal{L}, \prec\rangle$ be the landmark set, where $\mathcal{L} \subseteq F^\phi$ is the set of landmark fluents and $\prec$ is the ordering between the fluents (we will specifically focus on sound orderings derived from delete relaxations of the problem (Richter, Helmert, and Westphal 2008)), then we will use the set $\mathbb{L}' = \langle \mathcal{L}', \prec\rangle$, where $\mathcal{L}' = \mathcal{L} \setminus \phi(F)$.

Proposition 4. *The landmark set $\mathbb{L}'$ for the model $\mathcal{D}(\mathcal{M}^\pi_{\langle s,a\rangle})$ corresponds to the causal policy landmark set for policy π.*

We can establish this proof by following a slightly modified version of the proof described in (Sreedharan, Srivastava, and Kambhampati 2020). It's also worth noting that we are guaranteed that $G \subseteq \mathcal{L}'$

To generate our required subgoal, we need to identify the landmarks whose achievement actually requires the execution of the action in question (a) at state s. To identify whether a landmark $f \in \mathcal{L}$ requires the action, we will be using a formulation similar to the one we used to test whether the policy step was well-justified. Specifically we will use the model $\mathcal{M}^J_{(\pi,\langle s,a\rangle)\rightarrow f}$ defined as

$$\mathcal{M}^J_{(\pi,\langle s,a\rangle)\rightarrow f}\langle F^\phi, A^J_{(\pi,\langle s,a\rangle)}, I^J_{(\pi,\langle s,a\rangle)}, G_f\rangle.$$

The difference from the original formulation $\mathcal{M}^J_{(\pi,\langle s,a\rangle)}$, being the goal which is now set as $G_f = \{f\}$. Similar to the earlier formulation we are trying to see if the removal of F copy of the action effects results in the landmark being unreachable. More formally we can state,

Proposition 5. *A causal policy landmark $f \in \mathcal{L}$ is said to be the proximal required subgoal (per Definition 2), if and only if the modified planning problem $\mathcal{D}(\mathcal{M}^\pi_{\langle s,a\rangle,\rightarrow f})$ is unsolvable and there exists no $f' \in \mathcal{L}$, $f' \prec f$ and $\mathcal{D}(\mathcal{M}^\pi_{\langle s,a\rangle,\rightarrow f'})$ is unsolvable.*

The proof for this proposition is symmetric to the one used for Proposition 2.

5.2 Relationship to Required Causal Links

Now to see how these extracted relationship compare against the causal links, we will constrain ourselves to deterministic settings, where from any state, there can at most be one goal-achieving trace. We will assume the same policy structure. Now we will show that every valid minimal-length generalized required causal link (per Definition 2) corresponds to a minimal-length required causal link for the corresponding action in the original plan and vice-versa.

Proposition 6. *Let $\mathcal{P} = \langle a_1{:}1, ..., a_k{:}n\rangle$ be a valid plan for a deterministic model $\mathcal{M}$ and π the corresponding policy representation, let $a{:}i$ be a well justified step in $\mathcal{P}$ and let s_i, a be the corresponding well-justified policy step in π, then we have*

1. *if $a{:}i \rightarrow_f a'{:}j$ is a minimal-length required causal link for in $\mathcal{P}$, then there must exist a a minimal-length generalized required causal link of the form $\langle s_i, a\rangle \rightarrow f$ for π.*
2. *Similarly, if $\langle s_i, a\rangle \rightarrow f$ is a minimal-length generalized required causal link, then there must exist a minimal-length required causal link for $\mathcal{P}$ of the form $a{:}i \rightarrow_f a'{:}j$*

Proof Sketch. We can prove this statement by leveraging the fact that the causal policy landmarks of π consists of just the preconditions (along with the goal) of the actions in $\mathcal{P}$ and the ordering corresponds to the ordering of the actions in the totally-ordered plan. Now with that fact in mind, we can see that if there is a required subgoal is found, it is a precondition of a future action ($a' : j$) and the effects of the current action is required for satisfying that precondition and no action between $a : i$ and $a' : j$ can contribute this fact (this follows from the Definition 2). Thus this must be an effect of the action a, thus establishing the fact that $a : i \rightarrow_f a : j$ must be a required causal link. The fact that it is of minimal-length required causal link comes from the fact that there were no required subgoals that preceded f (again a requirement from Definition 2). This proves the first part of the proposition, the second part of the proposition can be established by inverting the arguments and thus proving the statement as a whole □

5.3 Generating Causal Link Chain Explanations

While the above approach identifies the first causal link contributed by the action, the most common explanation for a required action takes the form of a chain of causal links that extends to the goal. To provide such an explanation we have to not only identify a single landmark that is required, but ideally, we would like to present a chain of facts each requiring the last fact to be achieved. Note that here we can't just rely on the landmark ordering as it may also encode the relationship being enforced by the state part of the preconditions. So we will build a variation of $\mathcal{M}^{\pi}_{\langle s,a\rangle}$ denoted as $\hat{\mathcal{M}}^{\pi}_{\langle s,a\rangle,f_1\rightarrow f_2}$ that will try to identify such requirement relationship between landmarks. Specifically, we will have $\hat{\mathcal{M}}^{\pi}_{\langle s,a\rangle,f_1\rightarrow f_2} = \langle F^{\pi}_{\langle s,a\rangle}, \hat{A}^{\pi}_{\langle s,a\rangle,f_1\rightarrow f_2}, I^{\pi}_{\langle s,a\rangle}, \{f_2\}\rangle$. Now the goal is to achieve f_2, and we will form $\hat{A}^{\pi}_{\langle s,a\rangle,f_1\rightarrow f_2}$ from $A^{\pi}_{\langle s,a\rangle}$ by removing f_1 from all add effects while preserving $\phi(f_1)$. More formally, let $a^{s'}_j \in A^{\pi}_{\langle s,a\rangle}$, then we have a correspond action $\hat{a}^{s'}_{j,f_1\rightarrow f_2} \in \hat{A}^{\pi}_{\langle s,a\rangle,f_1\rightarrow f_2}$, such that $\mathbb{E}(\hat{a}^{s'}_{j,f_1\rightarrow f_2}) = \{\langle add^{s',m}_{a_j} \setminus f_1, del^{s',m}_{a_j} \setminus f_1\rangle \mid \langle add^{s',m}_{a_j}, del^{s',m}_{a_j}\rangle \in \mathbb{E}(a^{s'}_j)\}$

Since the requirement ordering will be a subset of the landmark ordering, we will only need to run this test between landmarks when there already exists an ordering. We will denote this requirement ordering with the notation $\prec_R$.

Finally, to generate the explanation chain itself, we will iterate over a topological sort over $\mathcal{L}'$ and find the first landmark f_1 that requires action a and build a chain consisting of a set of totally ordered landmarks over the requirement ordering that terminates with the goal G. More formally

Definition 3. *A chain of facts $\mathcal{E} = \langle f_1, ..., f_j, ..., f_n\rangle$, such that all $f_i \in F$ is considered a **valid explanation** for the query $\langle s, a, \pi, \mathcal{M}\rangle$, if*

1. *Every fact f_i in $\mathcal{E}$ is a causal policy landmark for the policy π and model $\mathcal{M}$*
2. *f_1 requires the action 'a' to be executed in state s*
3. *For all pairs of landmarks, f_i and f_{i+1}, we have $f_i \prec_R f_{i+1}$*
4. *Finally, we have $f_n = G$.*

The above definition presents a general description for a valid explanation. Note that the set of valid explanation covered by the above definition may not be equivalent in how effective the user may find them to be. As such one may need to use additional criteria to choose an explanation from this set of valid explanations. Choosing a landmark with the least number of preceding facts as the first element in the chain being one such possible criteria.

In our running example, the full causal link chain would be

$$\langle I, \texttt{order-donuts}\rangle \rightarrow \texttt{security-guard-bribed} \rightarrow \texttt{parked-at-executive-parking-spot} \rightarrow \texttt{in-executive-elevator} \rightarrow \texttt{at-office}.$$

Where each relation captured by $\rightarrow$ in the above sequence now correspond to a generalized causal link. This causal link chain can then be used to generate the exact explanatory message that can be provided to the user. An example explanatory message that can be generated from the above causal link chain may be

The step is needed to achieve the fact 'security guard bribed', which is needed to achieve the fact parked 'at executive parking spot', which is needed to achieve the fact 'in executive elevator', which is required to achieve the goal 'at office'.

Relationship to Causal Link Explanations Now to see how these explanations compare against the causal chains, we will constrain ourselves to deterministic settings, where every action has a single outcome. Thus from any state, there can at most be one goal-achieving trace. We will assume the same policy structure. Now we will show that every valid explanation (per Supplementary Definition 3) corresponds to the fact that are part of a causal chain explanation and every causal chain explanation correspond to an explanation of the form described in Supplementary Definition 3.

Proposition 7. *For a given causal chain explanation $\langle s_1 : a_1 \rightarrow_{f_1} s_2 : a_2, ..., s_m : a_m \rightarrow_g s_g : a^{\emptyset}\rangle$, the chain $\mathcal{E} = \langle f_1, ..., g\rangle$ is a valid explanation for the requirement query $\langle s_1, a_1, \pi, \mathcal{M}\rangle$, when $\mathcal{M}$ is completely deterministic.*

Proof Sketch. To see why this is true, we can see that all three requirements of a valid explanation provided in Definition 3 are met here. (1.) directly holds as all the facts are causal policy landmarks (they all appear in the precondition and there is only one path). (2.) holds automatically as this is a fact that is contributed by the action and per our definition of causal link explanation no action between the producer and consumer would generate the fact f_1. Thus f_1 would cause $\mathcal{M}^{\pi}_{\langle s,a\rangle,\rightarrow f_1}$ to be unsolvable as model will disallow any use of actions after s_2 to be used. (3.) holds because the causal links are preconditions and as such removal of them causes the subsequent fact to be unachievable at the subsequent step. □

Similarly, we can also show that

Proposition 8. *For any valid explanation chain $\mathcal{E} = \langle f_1, ..., g\rangle$ for the query $\langle s_1, a_1, \pi, \mathcal{M}\rangle$ (where $\mathcal{M}$ is completely deterministic), there exist a causal chain explanation of the form $\langle s_1 : a_1 \rightarrow_{f_1} s_2 : a_2, ..., s_m : a_m \rightarrow_g s_g : a^{\emptyset}\rangle$, for some action set $\{a_2, ..., a_m\}$.*

The proof for this proposition follows a similar line of argument to the one described in Proposition 6.

Remark: While a well-justified action always contribute a required causal link, it need not be part of a causal link chain that terminates at goal. Guaranteeing this requires us to look at stronger notions of action justification, which we hope to study as future work. It is also worth noting that most existing landmark generation methods identify landmarks from some relaxation of the original planning model (usually delete-relaxed models). Under these conditions, the extracted causal link would still remain a required subgoal, in that the policy step was needed to compute it. However, the subgoal need not be a proximal required subgoal anymore.

6 Empirical Evaluation

As a way to provide a computational evaluation of the explanation generation methods discussed in this paper, we ran our method on several standard FOND benchmarks (Muise 2018). Our evaluation has two primary goals: to identify the frequency with which well-justified actions occur in policies generated for these planners and to evaluate the computational requirements. The first would, in fact, also provide insights into the kind of policies that are generated by the state-of-the-art planner. The second allows us to identify the time taken to both determine whether an action is required and to generate a required causal link. The results reported here are from experiments run on a 12 core Intel(R) Xeon(R) CPU with an E5-2643 v3@3.40GHz processor and a 64 GB RAM.

For generating the policies, we used the PRP planner (Muise, McIlraith, and Beck 2012) which by default produces a policy defined over partial states. We generate the full state policy by executing this policy defined over partial states from the initial state (favoring actions with lower distance when multiple partial states match).

For generating the landmarks, we made use of the implementation of (Keyder, Richter, and Helmert 2010) provided by the Fast Downward planning system (Helmert 2006). Additionally, we used the FastDownward planner to test the unsolvability of the various subgoals. In particular, used the LAMA 2011 IPC configuration (–alias seq-sat-lama-2011). Table 1 provides an overview of the various statistics we calculated from the experiments. The experiments were run on five domains. For each domain, we selected the first fifteen problems from PRP FOND benchmark list and then selected the ones that were solvable by PRP planner within a time limit of 30 minutes (the exact problems are provided in the supplementary package).

One of the first things to note is that, for all domains except triangle tireworld not all non-trivial state-action pairs were well-justified. By non-trivial state-action pairs, we refer to any reachable state action pair where the state didn't already satisfy the goal. The presence of such non-justified state-action pairs refers to the fact that there is possible space for PRP planners to improve the quality of the policies generated by taking into account notions of justified actions and by extension policy minimality. For most domains, the generation time for the causal links was quite within the limits to be applicable for systems that require quick response time. In fact for all domains except Triangle-tireworld the average time taken for explanations generated were less than five seconds. Even in the case of Triangle-tireworld, the time taken was mostly due to the last four problems. In fact, the average time taken to establish whether an action is required for the first five problems was 6.7 seconds. On further analyses, the main limiting factor seems to the number of fluents. For the largest problem instance we tested with in triangle-tireworld there were close to four hundred thousand fluents. While Zenotravel domain had much larger policies, their fluent numbers were not close to the count in triangle-tireworld (largest count was still less than three thousand). In future, we hope to leverage approximations (including approximate unsolvability tests), which would be able to handle such large fluent counts. Please note that in the particular compilation considered in the evaluation, we made no distinction between static predicates and fluents. This means that our compilation considers more fluents than what was required. So, we expect to improve upon the existing results by considering more efficient implementations.

7 Related Work

The concept of justified actions and plans has a long history within automated planning literature, with some of the earliest works being introduced in the context of abstraction (cf. (Yang and Tenenberg 1990)). Since then, many flavors of action justification have been studied in the literature including *Backward Justification*, *Greedy Justification* and *Perfect Justification*, with perfectly justified plans presenting the strongest notion. In fact, establishing whether a plan is perfectly justified is NP-complete even for totally ordered plan, as in the worst case you have to check if any subset of actions are redundant together. However, it is worth noting that even within the same justification class the complexity of establishing justification could differ with solution representation. Policy cases notwithstanding, the complexity of establishing whether an action is required within a partially ordered plan is NP-complete (Olz and Bercher 2019), while it's polynomial for totally-ordered plans.

Causal links were first introduced into the mainstream planning vocabulary by McAllester and Rosenblitt (1991) (however similar concepts existed before as in the use of the term GOST by Tate (1977)). The primary use of causal links remains in the context of partial-order or plan-space planning. However, the intuitive nature of these data-structures have lent themselves to their use in numerous applications within the context of plan explanations. A common one we saw was the use of causal link chains to explain the contributions made by an action to the plan. However, the history of causal chain explanation starts much earlier than their latest incarnation in (Seegebarth et al. 2012). One of the earliest works to look at a similar form of information was the PRIAR system (Kambhampati 1991) that introduced the notion of validation structures that encodes such information in the form of plan annotations. Validation structures were proposed as a correctness explanation that could then be used to guide various tasks including plan retrieval, refitting, and

Domains	Problem Count	Average % of Non-Trivial Well-Justified State-action Pairs	Average Time Taken to Establish Action is Well-Justified (S)	Average Time Taken to Identify Causal Link (S)
Elevator	15	99.8%	1.34 ± 0.17	2.76 ± 0.36
Exploding Blocksworld	8	75.25%	2.08 ± 0.84	3.15 ± 0.17
Tireworld	15	94.4%	1.31 ± 0.15	2.64 ± 0.31
Zenotravel	14	98.19%	2.34 ± 0.67	4.85 ± 0.31
Triangle Tire-World	9	100%	81.47 ± 78.93	180.53 ± 177.144

Table 1: The evaluation of the proposed method on standard FOND benchmarks. Here non-trivial state action pair refers to state action pairs where the policy didn't assign the $a^{\emptyset}$ goal action.

modification (Kambhampati 1990). Another early work that looked at the introduction of similar information was that of (Veloso 1992), which looked at performing regression-based analyses to determine initial state conditions relevant to the goal. In more recent work, such information was also used by (Chakraborti et al. 2019) to provide an overview of the plan as a whole. (Bryce et al. 2017) also looks at similar information to visualize plans by visualizing causal link chains in the style of metro rail maps. (Bercher et al. 2014) presents human subject studies to verify the effectiveness of such explanations by grounding these explanations in the context of the application of an assistive system for putting together a home theater system.

Another closely related work, we briefly mentioned earlier is the method discussed in Sreedharan, Srivastava, and Kambhampati (2020). However it is worth noting that the primary objective of this previous work is to provide a summary of the whole policy and does not provide any insights into specific roles played by individual steps, which is our main objective. In addition to focusing on individual steps, we also address the challenge of filtering side-effects from fluents that are actually needed by future actions. This is a step that is central to our ability to generalize causal links from classical planning literature.

In addition to justified actions, extant literature have also considered other notions of required actions particularly at the level of specific planning problems. One particular example is the notion of action landmarks (Karpas and Domshlak 2009), where action landmarks are actions that must occur in every plan to a goal. Note that a justified action need not be a landmark action at all either for an all-outcome determinization of the original problem or for the version of the problem that restricts plans that corresponds to traces possible under a given policy (we will refer to this version of the planning problem as the policy restricted planning problem). After all, even for a given policy there might be paths from initial states to goal that circumvents the step in question completely. On the other hand, if one were to root the problem in the state that is part of the step in question and regenerate the action landmarks, then even non-justified actions would turn into an action landmark for the policy restricted formulation, since the policy requires the action to be executed (while justified actions still may not be a landmark in the case of the original problem). A closely related, but distinct notion is that of strong and weak stubborn sets (Alkhazraji et al. 2012) , which can again be seen to include non-justified actions. While each of these notions are in some sense capturing cases where an action may be required, it is interesting to note that they are in fact capturing complementary concepts.

8 Conclusion

The paper presents a generalization of causal chains and action justification, which could act as basis for explaining why an action is required in a plan. While this paper focuses on formalizing and establishing the properties of the generalized versions of the notions, the next step would be to run user studies to verify the effectiveness of explanations built out of these components. While results from Sreedharan, Srivastava, and Kambhampati (2020), already provides some positive evidence to the utility of such information, we hope to run a large scale user study to verify the effectiveness of such explanations for FOND and stochastic policies, similar to the one carried out by Bercher et al. (2014) for line plans. It is worth noting that these generalized causal links now present temporally ordered (but not necessarily adjacent) facts that must hold in a set of traces. As such, one could equivalently capture these causal links using LTL formulas over these traces. One interesting thread of future work might be to investigate whether there are particular LTL templates that people prefer and whether we can map our explanations (for well-justified actions or for other more general forms of justification) to such templates. Some of the other next steps, already mentioned in the paper, include investigating weaker notions of justifications. We would also be interested in seeing if we could extract such causal structures for other forms of planning including numeric and temporal planning. Recent works (c.f. (Lindner and Olz 2022)) have also highlighted the usefulness of distinguishing between different roles played by actions within a single causal link explanations. We also hope to investigate how to incorporate such considerations into our formulations.

Acknowledgements

Kambhampati's research is supported in part by ONR grants N00014-16-1-2892, N00014-18-1- 2442, N00014-18-1-2840, N00014-9-1-2119, AFOSR grant FA9550-18-1-0067, DARPA SAIL-ON grant W911NF-19- 2-0006, and a JP Morgan AI Faculty Research grant. Muise's research is supported in part by funding from Natural Sciences and Engineering Research Council of Canada (NSERC).

References

Alkhazraji, Y.; Wehrle, M.; Mattmüller, R.; and Helmert, M. 2012. A stubborn set algorithm for optimal planning. In *ECAI 2012*, 891–892. IOS Press.

Bercher, P.; Biundo, S.; Geier, T.; Hoernle, T.; Nothdurft, F.; Richter, F.; and Schattenberg, B. 2014. Plan, Repair, Execute, Explain – How Planning Helps to Assemble Your Home Theater. In *ICAPS*.

Bryce, D.; Bonasso, P.; Adil, K.; Bell, S.; and Kortenkamp, D. 2017. In-situ domain modeling with fact routes. In *Proceedings of the Workshop on User Interfaces and Scheduling and Planning, UISP*, 15–22.

Bryce, D.; and Buffet, O. 2008. 6th international planning competition: Uncertainty part. *Proceedings of the 6th International Planning Competition (IPC'08)*.

Chakraborti, T.; Fadnis, K. P.; Talamadupula, K.; Dholakia, M.; Srivastava, B.; Kephart, J. O.; and Bellamy, R. K. E. 2019. Planning and Visualization for a Smart Meeting Room Assistant – A Case Study in the Cognitive Environments Laboratory at IBM T.J. Watson Research Center, Yorktown. *AI Communication*.

Cimatti, A.; Pistore, M.; Roveri, M.; and Traverso, P. 2003. Weak, strong, and strong cyclic planning via symbolic model checking. *Artif. Intell.*, 147(1-2): 35–84.

Fink, E.; and Yang, Q. 1993. A spectrum of plan justifications. In *Proceedings of the AAAI 1993 Spring Symposium*, 23–33.

Helmert, M. 2006. The fast downward planning system. *Journal of Artificial Intelligence Research*, 26: 191–246.

Kambhampati, S. 1990. Mapping and Retrieval During Plan Reuse: A Validation Structure Based Approach. In *AAAI*.

Kambhampati, S. 1991. *Flexible reuse and modification in hierarchical planning: A validation structure-based approach.* Ph.D. thesis, University of Maryland College Park.

Kambhampati, S. 1994. Multi-Contributor Causal Structures for Planning: A Formalization and Evaluation. *Artif. Intell.*, 69(1-2): 235–278.

Kambhampati, S. 1995. Admissible pruning strategies based on plan minimality for plan-space planning. In *IJCAI*, 1627–1635. Citeseer.

Karpas, E.; and Domshlak, C. 2009. Cost-Optimal Planning with Landmarks. In *IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA, July 11-17, 2009*, 1728–1733.

Keyder, E.; Richter, S.; and Helmert, M. 2010. Sound and Complete Landmarks for And/Or Graphs. In *ECAI 2010 - 19th European Conference on Artificial Intelligence, Lisbon, Portugal, August 16-20, 2010, Proceedings*, volume 215 of *Frontiers in Artificial Intelligence and Applications*, 335–340. IOS Press.

Lindner, F.; and Olz, C. 2022. Step-by-Step Task Plan Explanations Beyond Causal Links. In *31st IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2022, Napoli, Italy, August 29 - Sept. 2, 2022*, 45–51. IEEE.

McAllester, D. A.; and Rosenblitt, D. 1991. Systematic Nonlinear Planning. In *Proceedings of the 9th National Conference on Artificial Intelligence, Anaheim, CA, USA, July 14-19, 1991, Volume 2*, 634–639. AAAI Press / The MIT Press.

Muise, C. 2018. planner-for-relevant-policies. https://github.com/QuMuLab/planner-for-relevant-policies/tree/master/prp-scripts. Accessed: 2023-04-4.

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)*.

Olz, C.; and Bercher, P. 2019. Eliminating redundant actions in partially ordered plans—a complexity analysis. In *Proceedings of the International Conference on Automated Planning and Scheduling*, volume 29, 310–319.

Richter, S.; Helmert, M.; and Westphal, M. 2008. Landmarks Revisited. In *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, AAAI 2008, Chicago, Illinois, USA, July 13-17, 2008*, 975–982. AAAI Press.

Seegebarth, B.; Müller, F.; Schattenberg, B.; and Biundo, S. 2012. Making Hybrid Plans More Clear to Human Users - A Formal Approach for Generating Sound Explanations. In McCluskey, L.; Williams, B. C.; Silva, J. R.; and Bonet, B., eds., *Proceedings of the Twenty-Second International Conference on Automated Planning and Scheduling, ICAPS 2012, Atibaia, São Paulo, Brazil, June 25-19, 2012*. AAAI.

Sreedharan, S.; Srivastava, S.; and Kambhampati, S. 2020. TLdR: Policy summarization for factored SSP problems using temporal abstractions. In *Proceedings of the International Conference on Automated Planning and Scheduling*, volume 30, 272–280.

Tate, A. 1977. Generating Project Networks. In *Proceedings of the 5th International Joint Conference on Artificial Intelligence. Cambridge, MA, USA, August 22-25, 1977*, 888–893. William Kaufmann.

Veloso, M. M. 1992. *Learning by analogical reasoning in general problem-solving*. Ph.D. thesis, CMU.

Yang, Q.; and Tenenberg, J. D. 1990. ABTWEAK: Abstracting a Nonlinear, Least Commitment Planner. In *Proceedings of the 8th National Conference on Artificial Intelligence. Boston, Massachusetts, USA, July 29 - August 3, 1990, 2 Volumes*, 204–209. AAAI Press / The MIT Press.

Yoon, S. W.; Fern, A.; and Givan, R. 2007. FF-Replan: A Baseline for Probabilistic Planning. In *Proceedings of the Seventeenth International Conference on Automated Planning and Scheduling, ICAPS 2007, Providence, Rhode Island, USA, September 22-26, 2007*, 352. AAAI.