Features, Projections, and Representation Change for Generalized Planning

Blai Bonet\(^1\) and Hector Geffner\(^2\)

\(^1\) Universidad Simón Bolívar, Caracas, Venezuela
\(^2\) ICREA & Universitat Pompeu Fabra, Barcelona, Spain
bonet@usb.ve, hector.geffner@upf.edu

Abstract

Generalized planning is concerned with the characterization and computation of plans that solve many instances at once. In the standard formulation, a generalized plan is a mapping from features or observation histories into actions, assuming that the instances share a common pool of features and actions. This assumption, however, excludes the standard relational planning domains where actions and objects change across instances. In this work, we extend the standard formulation of generalized planning to such domains. This is achieved by projecting the actions over the features, resulting in a common set of abstract actions which can be tested for soundness and completeness, and which can be used for generating general policies such as “if the gripper is empty, pick the clear block above \(x\) and place it on the table” that achieve the goal \(\text{clear}(x)\) in any Blocksworld instance. In this policy, “pick the clear block above \(x\)” is an abstract action that may represent the action \(\text{Unstack}(a, b)\) in one situation and the action \(\text{Unstack}(b, c)\) in another. Transformations are also introduced for computing such policies by means of fully observable non-deterministic (FOND) planners. The value of generalized representations for learning general policies is also discussed.

1 Introduction

Generalized planning is concerned with the characterization and computation of plans that solve many instances at once [Srivastava et al., 2008; Bonet et al., 2009; Srivastava et al., 2011a; Hu and De Giacomo, 2011; Belle and Levesque, 2016; Segovia et al., 2016]. For example, the policy “if left of the target, move right” and “if right of the target, move left”, solves the problem of getting to the target in a \(1 \times n\) environment, regardless of the agent and target positions or the value of \(n\).

The standard, semantic, formulation of generalized planning due to Hu and De Giacomo [2011] assumes that all the problems in the class share a common set of features and actions. Often, however, this assumption is false. For example, in the Blocksworld, the policy “if the gripper is empty, pick the clear block above \(x\) and place it on the table” eventually achieves the goal \(\text{clear}(x)\) in any instance, yet the set of such instances do not have (ground) actions in common. Indeed, the expression “pick up the clear block above \(x\)” may mean “pick block \(a\) from \(b\)” in one case, and “pick block \(c\) from \(d\)” in another. A similar situation arises in most of the standard relational planning domains. Instances share the same action schemas but policies cannot map features into action schemas; they must select concrete, ground, actions.

In this work, we show how to extend the formulation of generalized planning to relational domains where the set of actions and objects depend on the instance. This is done by projecting the actions over a common set of features, resulting in a common set of general actions that can be used in generalized plans. We also address the computation of the resulting general policies by means of transformations and fully observable non-deterministic (FOND) planners, and discuss the relevance to work on learning general policies.

Generalized planning has also been formulated as a problem in first-order logic with solutions associated with programs with loops, where actions schemas do not need to be instantiated [Srivastava et al., 2011a]. The resulting formulation, however, is complex and cannot benefit from existing propositional planners. First-order decision theoretic planning can also be used to generate general plans but these are only effective over finite horizons [Boutilier et al., 2001; Wang et al., 2008; van Otterlo, 2012].

The paper is organized as follows. We review the definition of generalized planning and introduce abstract actions and projections. We then consider the computation of generalized policies, look at examples, and discuss related work and challenges.

2 Generalized Planning

The planning instances \(P\) that we consider are classical planning problems expressed in some compact language as a tuple \(P = (V, I, A, G)\) where \(V\) is a set of state variables that can take a finite set of values (boolean or not), \(I\) is a set of atoms over \(V\) defining an initial state \(s_0\), \(G\) is a set of atoms or literals over \(V\) describing the goal states, and \(A\) is a set of actions \(a\) with their preconditions and effects that define the set \(A(s)\) of actions applicable in any state \(s\), and the successor state function \(f(a, s)\), \(a \in A(s)\). A state is a valuation over \(V\) and a solution to \(P\) is an applicable action sequence
\[ \pi = a_0, \ldots, a_n \] that generates a state sequence \( s_0, s_1, \ldots, s_n \) where \( s_n \) is a goal state (makes \( G \) true). In this sequence, \( a_i \in A(s_i) \) and \( s_{i+1} = f(a_i, s_i) \) for \( i = 0, \ldots, n - 1 \). A state \( s \) is reachable in \( P \) if \( s = s_n \) for one such sequence.

A general planning problem \( Q \) is a collection of instances \( P \) assumed to share a common set of actions and a common set of features or observations [Hu and De Giacomo, 2011]. A boolean feature \( p \) for a class of problems represents a function \( \phi_p \) that takes an instance \( P \) from \( Q \) and a reachable state \( s \) in \( P \), and results in a boolean value denoted as \( \phi_p(s) \), with the reference to the problem \( P \) omitted. Features are sometimes associated with observations, and in such cases \( \phi_p \) is a sensing function. For example, if \( Q \) is the class of all block instances with goal \( on(x, y) \) where \( x \) and \( y \) are any blocks. Two instances \( P \) and \( P' \) with goals \( on(a, b) \) and \( on(c, d) \) respectively, are part of \( Q \) with values for \( x \) and \( y \) being \( a \) and \( b \) in \( P \), and \( c \) and \( d \) in \( P' \). General features using such parameters can be used, with the value of the parameters in an instance \( P \) being determined by matching the goal of \( P \) with the generic goal. A parametric feature like \( n(x) \) that tracks the number of blocks above \( x \) thus represents the feature \( n(a) \) in \( P \) and the feature \( n(c) \) in \( P' \).

2.1 Numerical Features

For extending the formulation to domains where actions take arguments that vary from instance to instance, we need numerical features. A numerical feature \( n \) for a generalized problem \( Q \) represents a function \( \phi_n \) that takes an instance \( P \) and a state \( s \), and results in a non-negative integer value denoted as \( \phi_n(s) \). For the problem \( Q \) representing the Blocksworld instances with goal \( on(a, b) \), a numerical feature \( n(a) \) can represent the number of blocks above \( a \). The set of features \( F \) becomes thus a pair \( F = (B, N) \) of boolean and numerical features \( B \) and \( N \). A valuation over \( F \) is an assignment of truth values to the boolean features \( p \in B \) and non-negative integer values to the numerical features \( n \in N \). A boolean valuation over \( F \), on the other hand, is an assignment of truth values to the boolean features \( p \in B \) and to the atoms \( n = 0 \) for \( n \in N \). The feature valuation for a state \( s \) is denoted as \( \phi_F(s) \), while the boolean valuation as \( \phi_B(s) \). The number of feature valuations is infinite but the number of boolean valuations is \( 2^{|P|} \). Policies for a generalized problem \( Q \) over a set of features \( F = (B, N) \) are defined as:

Definition 1 (Generalized Planning). A policy for a generalized problem \( Q \) over the features \( F = (B, N) \) is a partial mapping \( \pi \) from boolean valuations over \( F \) into the common actions in \( Q \). The policy \( \pi \) solves \( Q \) if \( \pi \) solves each instance \( P \) in \( Q \); i.e., if the trajectory \( s_0, \ldots, s_n \) induced by the policy \( \pi \) in \( P \) is goal reaching.

2.2 Parametric Goals and Features

We are often interested in finding a policy for all instances of a domain whose goal has a special form. For example, we may be interested in a policy for the class \( Q \) of all Blocksworld instances with goal of the form \( on(x, y) \) where \( x \) and \( y \) are any blocks. Two instances \( P \) and \( P' \) with goals \( on(a, b) \) and \( on(c, d) \) respectively, are part of \( Q \) with values for \( x \) and \( y \) being \( a \) and \( b \) in \( P \), and \( c \) and \( d \) in \( P' \). General features using such parameters can be used, with the value of the parameters in an instance \( P \) being determined by matching the goal of \( P \) with the generic goal. A parametric feature like \( n(x) \) that tracks the number of blocks above \( x \) thus represents the feature \( n(a) \) in \( P \) and the feature \( n(c) \) in \( P' \).

3 Abstract Actions

A generalized planning problem like “all Blocksworld instances with goal of the form \( on(x, y) \)” admits simple and general plans yet such plans cannot be accounted for in the current framework. The reason is that such instances share no common pool of actions. This is indeed the situation in relational domains where objects and actions change across instances. For dealing with such domains, we define the notion of abstract actions that capture the effect of actions on the common set of features.

The abstract actions for a generalized problem \( Q \) are defined as pairs \( \bar{a} = (Pre; Eff) \) where \( Pre \) and \( Eff \) are the action precondition and effects expressed in terms of the set of features \( F = (B, N) \). The syntax is simple: preconditions may include atoms \( p \) and \( n = 0 \) or their negations, for \( p \in B \) and \( n \in N \), and effects may be boolean, \( p \) or \( \neg p \) for \( p \in B \), or numerical increments and decrements expressed as \( n + 1 \) and \( n - 1 \) respectively for \( n \in N \). In the language of abstract actions, features represent state variables (fluents), not functions, with \( n \) representing a non-negative integer, and \( n + 1 \) and \( n - 1 \) representing updates \( n := n + \Delta \) and \( n := n - \Delta \) for random positive integers \( \Delta \) and \( \Delta' \) that are not allowed to make \( n \) negative. The numerical updates are thus non-deterministic. The negation of atom \( n = 0 \) is expressed as \( n > 0 \) and it must be a precondition of any abstract action that decrease \( n \) for ensuring that \( n \) remains non-negative.

Definition 2. An abstract action \( \bar{a} \) over the features \( F = (B, N) \) is a pair \( (Pre; Eff) \) such that 1) each precondition in \( Pre \) is a literal \( p \) or \( \neg p \) for \( p \in B \), or a literal \( n = 0 \) or \( n > 0 \) for \( n \in N \), 2) each effect in \( Eff \) is a boolean literal over \( B \), or a numerical update \( n + 1 \) or \( n - 1 \) for \( n \in N \), and 3) \( n > 0 \) is in \( Pre \) if \( n = 1 \) is in \( Eff \).

We want abstract actions that represent the diverse set of concrete actions in the different instances \( P \) of the generalized problem \( Q \). The number of concrete actions across all instances in \( Q \) is often infinite but the number of abstract actions is bounded by \( 3^{|P|} \).

Abstract actions operate over abstract states \( \bar{s} \) that are valuations over the feature variables. An abstract action \( \bar{a} = \bar{a} \). 

4668
represents a concrete action \( a \) over a state \( s \) in instance \( P \) if the two actions are applicable in \( s \), and they affect the values of the features in a similar way. Let us say that \( \text{Pre} \) is true in \( s \) if \( \text{Pre} \) is true in the valuation \( \phi_B(s) \). Then:

**Definition 3.** The abstract action \( \tilde{a} = \langle \text{Pre}; \text{Eff} \rangle \) over a set of features \( F \) represents the action \( a \) on the state \( s \) of instance \( P \) if the two actions are applicable in \( s \), and they affect the values of the features in a similar way.

Let: for each boolean feature \( p \) in \( B \), if \( p \) changes from true to false (resp. from false to true) in transition \( s \sim f(a, s) \) then \( \neg p \) in \( \text{Eff} \) (resp. \( p \) in \( \text{Eff} \)).

b) for each boolean feature \( p \) in \( B \), if \( p \) (resp. \( \neg p \)) is in \( \text{Eff} \), then \( p \) is true (resp. false) in \( f(a, s) \), and
c) for each numerical feature \( n \) in \( N \), \( n \leq \) (resp. \( n \geq \)) in \( \text{Eff} \) iff \( \phi_n(f(a, s)) < \phi_n(s) \) (resp. \( \phi_n(f(a, s)) > \phi_n(s) \)).

**Example.** Let \( Q_{\text{clear}} \) be the set of Blocksworld instances \( P \) with goal of the form \( \text{clear}(x) \) and initial situation where the arm is empty, and let \( F = \{ H, n(x) \} \) be the set of features where \( H \) holds iff the arm is holding some block, and \( n(x) \) counts the number of blocks above \( x \). The abstract action \( \tilde{a} = \langle \neg H; n(x) > 0; H, n(x) \leq ) \) (1)

represents any action that picks up a block from above \( x \), as it makes \( H \) true and decreases the number of blocks above \( x \). If \( P \) is an instance with goal \( \text{clear}(a) \) and \( s \) is a state where \( \text{on}(b, a) \), \( \text{on}(c, b) \), and \( \text{clear}(c) \) are all true, \( \tilde{a} \) represents the action \( \text{Unstack}(c, b) \) as both actions, the abstract and the concrete, are applicable in \( s \), make \( H \) true, and decrease \( n(x) \). Likewise,

\[ \tilde{a}' = \langle H; \neg H \rangle \]

represents any action that places the block being held anywhere but above \( x \) (as it does not affect \( n(x) \)). In the state \( s' \) that results from the state \( s \) \( \text{Putdown}(c) \), \( \tilde{a}' \) represents the action \( \text{Putdown}(c) \), and also \( \text{Stack}(c, d) \) if \( d \) is a block in \( P \) that is clear in both \( s \) and \( s' \).

We say that an abstract action \( \tilde{a} \) is sound when it represents some concrete action in each (reachable) state \( s \) of each instance \( P \) of \( Q \) where \( \tilde{a} \) is applicable:

**Definition 4.** An abstract action \( \tilde{a} = \langle \text{Pre}; \text{Eff} \rangle \) is sound in the problem \( Q \) over the features \( F \) iff for each instance \( P \) in \( Q \) and each reachable state \( s \) in \( P \) where \( \text{Pre} \) holds, \( \tilde{a} \) represents one or more actions \( a \) on \( P \) on the state \( s \).

The abstract action (1) is sound in \( Q_{\text{clear}} \) as in any reachable state of any instance where the arm is empty and there are blocks above \( x \), there is an unstack action that makes the feature \( H \) true and decreases the number of blocks above \( x \). The abstract action (2) is sound too. The two actions, however, do not provide a complete representation:

**Definition 5.** A set \( A \) of abstract actions over the set of features \( F \) is complete for \( Q \) if for any instance \( P \) in \( Q \), any reachable state \( s \) of \( P \), and any action \( a \) that is applicable at \( s \), there is an abstract action \( \tilde{a} \) in \( A \) that represents \( a \) on \( s \).

---

Throughout, Blocksworld refers to the encoding with action schemas \( \text{Stack}(x, y) \), \( \text{Unstack}(x, y) \), \( \text{Pickup}(x) \), and \( \text{Putdown}(x) \). The set made of the two actions above is not complete as they cannot represent concrete actions that, for example, pick up the target block \( x \) or a block that is not above \( x \).

**Example.** A sound and complete set of abstract actions \( A_F \) for \( Q_{\text{clear}} \) can be obtained with the features \( F' = \{ X, H, Z, n(x), m(x) \} \) where \( X \), \( H \), and \( Z \) represent that block \( x \) is being held, that some other block is being held, and that there is a block below \( x \) respectively. The new counter \( m(x) \) tracks the number of blocks that are not in the same tower as \( x \) or being held. The abstract actions in \( A_F \), with names for making their meaning explicit, are:

- Pick-x-some-below = \langle \neg H, \neg X, n(x) = 0; Z; X, \neg Z, m(x) \rangle,
- Pick-x-none-below = \langle \neg H, \neg X, n(x) = 0; \neg Z; X \rangle,
- Pick-above-x = \langle \neg H, \neg X, n(x) > 0; H, n(x) \rangle,
- Pick-other = \langle \neg H, \neg X, m(x) > 0; H, m(x) \rangle,
- Put-x-on-table = \langle X; \neg X \rangle,
- Put-x-above-some = \langle X; \neg X, Z, m(x) \rangle,
- Put-aside = \langle H; \neg H, m(x) \rangle,
- Put-above-x = \langle H; \neg H, n(x) \rangle.

**4 Generalized Planning Revisited**

The notion of generalized planning can be extended to relational domains, where instances share no common pool of actions, by moving to the abstract representation provided by abstract actions:

**Definition 6.** Let \( Q \) be a generalized planning problem, \( F \) be a set of features, and \( A_F \) be a set of sound abstract actions. A policy for \( Q \) over \( F \) is a partial mapping \( \pi \) from the boolean valuations over \( F \) into \( A_F \). The (abstract) policy \( \pi \) solves \( Q \) if \( \pi \) solves each instance \( P \) in \( Q \); i.e., if all the trajectories \( s_0, \ldots, s_n \) induced by \( \pi \) in \( P \) are goal reaching.

Abstract policies \( \pi \) map boolean valuations over \( F \) into sound abstract actions. We write \( a_i \in \pi(\phi_B(s_i)) \) to express that \( a_i \) is one of the concrete actions represented by the abstract action \( \tilde{a}_i = \pi(\phi_B(s_i)) \) in the state \( s_i \) of instance \( P \). Since the abstract actions in \( A_F \) are assumed to be sound, there must be one such concrete action \( a_i \) over any reachable state \( s_i \) of any problem \( P \) where \( \tilde{a}_i \) is applicable. Such concrete action, however, is not necessarily unique. The trajectories \( s_0, \ldots, s_n \) induced by the policy \( \pi \) in \( P \) such that \( a_i \in \pi(\phi_B(s_i)) \), \( s_{i+1} = f(a_i, s_i) \) for \( i < n \), and \( s_n \) is the first state in the sequence where the goal of \( P \) is true, \( \pi(\phi_B(s_n)) = \top \), or \( a_n \notin A(s_n) \).

**Example.** For the generalized problem \( Q_{\text{clear}} \) with features \( F = \{ H, n(x) \} \), and actions (1) and (2), the following policy, expressed in compact form in terms of two rules, is a solution:

\[ \neg H, n(x) > 0 \Rightarrow \tilde{a}, \quad H, n(x) > 0 \Rightarrow \tilde{a}'. \]

The policy picks blocks above \( x \) and puts them aside (not above \( x \)) until \( n(x) \) becomes zero.
5 Computation

We focus next on the computation of general policies. We proceed in two steps. First, we map the generalized problem $Q$ given features $F = (B, N)$ and a set $A_F$ of sound actions into a numerical non-deterministic problem $Q_{\text{clear}}$. While numerical planning problems can be undecidable [Helmert, 2002], these numerical problems are simpler and correspond to the so-called qualitative numerical problems (QNPs) whose solutions can be obtained using standard, boolean FOND planners [Srivastava et al., 2011b; Bonet et al., 2017].

5.1 Feature Projection $Q_F$

For defining the first reduction, we assume a set $A_F$ of sound abstract actions and formulas $I_F$ and $G_F$, defined over the atoms $p$ for $p \in B$ and $n = 0$ for $n \in N$, that provide a sound approximation of the initial and goal states of the instances in $Q$. For this, it must be the case that for any instance $P$ in $Q$: a) the truth valuation $\phi_B(s)$ for the initial state $s$ of $P$ satisfies $I_F$, and b) the reachable states $s$ in $P$ with a truth valuation $\phi_B(s)$ that satisfies $G_F$ are goal states of $P$.

Definition 7. For a given set of sound abstract actions $A_F$, and sound initial and goal formulas $I_F$ and $G_F$ for $Q$, the projection $Q_F = (V_F, I_F, G_F, A_F)$ of $Q$ is a numerical, non-deterministic planning problem with actions $A_F$, initial and goal situations $I_F$ and $G_F$, and state variables $V_F = F$.

The states in a projection $Q_F$ are valuations $\bar{s}$ over the features $F$ that in $Q_F$, like in abstract actions, represent state variables and not functions. The possible initial states are the valuations that satisfy $I_F$, the goal states are the ones that satisfy $G_F$, and the actions are the abstract actions in $A_F$.

Solutions to a projection $Q_F$ are partial policies $\pi$ that map boolean valuations over $F$ into actions in $A_F$ such that all the state trajectories induced by $\pi$ in $Q_F$ are goal reaching. A state trajectory $\bar{s}_0, \ldots, \bar{s}_n$ is induced by $\pi$ in $Q_F$ if it is a state that satisfies $I_F$, $\bar{a}_i = \pi(\bar{s}_i)$, $\bar{s}_{i+1} \in F(\bar{a}_i, \bar{s}_i)$ for $i < n$, and $\bar{s}_n$ is the first state in the sequence where $G_F$ is true, $\pi(\bar{s}_n)$ is not defined, or $\bar{a}_n$ is not applicable in $\bar{s}_n$, $F(\cdot, \cdot)$ is a non-deterministic transition function defined by the actions in $A_F$ in the usual way. The soundness of $I_F$, $G_F$, and $A_F$ imply the soundness of the projection $Q_F$.

Theorem 8. If $A_F$ is a set of sound abstract actions for the features $F$, and the formulas $I_F$ and $G_F$ are sound for $Q$, then a solution $\pi$ for $Q_F$ is also a solution for $Q$.

To see why this results holds, notice that if $\bar{a}_i$ is an abstract action applicable in $\bar{s}_i$ and $s$, where $s$ is a state for some instance $\bar{P}$ in $Q$, then there is at least one concrete action $a_i$ in $P$ that is represented by $\bar{a}_i$ in $s$. Also, for every state trajectory $\bar{s}_0, \ldots, \bar{s}_n$ induced by the policy $\pi$ on $P$, there is one abstract state trajectory $\bar{s}_0, \ldots, \bar{s}_n$ induced by $\pi$ in $Q_F$ that tracks the values of the features, i.e. where $\bar{s}_i = \phi_F(\bar{s}_i)$ for $\bar{i} = 0, \ldots, n$, as a result of the soundness of $A_F$ and the definition of $I_F$. Since the latter trajectories are goal reaching, the former must be too since $G_F$ is sound.

The projections $Q_F$ are sound but not complete. The incompleteness is the result of the abstraction (non-deterministic feature increments and decrements), and the choice of $A_F$, $I_F$, and $G_F$ that are only assumed to be sound.

Example. Let us consider the features $F = \{H, n(x)\}$, the set of abstract actions $A_F = \{a, \bar{a}\}$ given by (1) and (2), $I_F = \{-H, n(x) > 0\}$, and $G_F = \{n(x) = 0\}$. The policy given by (3) solves the projection $Q_F = (V_F, I_F, G_F, A_F)$ of $Q_{\text{clear}}$, and hence, by Theorem 8, also $Q_{\text{clear}}$. □

5.2 Boolean Projection and FOND Problem $Q_F$

The second piece of the computational model is the reduction of the projection $Q_F$ into a standard (boolean) fully observable non-deterministic (FOND) problem. For this we exploit a reduction from qualitative numerical planning problems (QNPs) into FOND [Srivastava et al., 2011b; Bonet et al., 2017]. This reduction replaces the numerical variables $n$ by propositional symbols named “$n = 0$” that are meant to be true when the numerical variable $n$ has value zero. The negation of the symbol “$n = 0$” is denoted as “$n > 0$”.

Definition 9. Let $Q$ be a generalized problem and $Q_F$ be a projection of $Q$ for $F = (B, N)$. The boolean projection $Q_F'$, associated with $Q_F$ is the FOND problem obtained from $Q_F$, by replacing 1) the numerical variables $n \in N$ by the symbols $n = 0$, 2) first-order literals $n = 0$ by propositional literals $n = 0$, 3) effects $n \rightarrow$ by deterministic effects $n > 0$, and 4) effects $n \leftarrow$ by non-deterministic effects $n > 0 | n = 0$.

The boolean projection $Q_F'$ is a FOND problem but neither the strong or strong cyclic solutions of $Q_F'$ [Cimatti et al., 2003] capture the solutions of the numerical problem $Q_F$.

The reason is that the non-deterministic effects $n > 0 | n = 0$ in $Q_F'$ are neither fair, as assumed in strong cyclic solutions, nor adversarial, as assumed in strong solutions. They are conditionally fair, meaning that from any time point on, infinite occurrences of effects $n > 0 | n = 0$ (decrements) imply the eventual outcome $n = 0$, on the condition that from that point on, no action with effect $n > 0$ (increment) occurs.

The policies of the boolean FOND problem $Q_F'$ that capture the policies of the numerical problem $Q_F$ are the strong cyclic solutions that terminate [Srivastava et al., 2011b]. We call them the qualitative solutions of $Q_F'$ and define them equivalently as:

Definition 10. A qualitative solution of the FOND $Q_F' = (V_{F}', I_{F}', G_{F}', A_{F}')$ is a partial mapping $\pi$ from boolean feature valuations into actions in $A_{F}'$ such that the state-action trajectories induced by $\pi$ over $Q_F'$ that are conditionally fair are all goal reaching.

A state-action trajectory $s_{t_0}^{t_1}, a_{0}^{t_1}, s_{1}^{t_1}, \ldots$ over $Q_F'$ is not conditionally fair iff a) it is infinite, b) after a certain time step $i$, it contains infinite actions with effects $n > 0 | n = 0$ and no action with effect $n > 0$, and c) there is no time step after $i$ where $n = 0$. The qualitative solutions of $Q_F'$ capture the solutions of the numerical projection $Q_F$ exactly:

Theorem 11. $\pi$ is a qualitative solution of the boolean FOND $Q_F'$ iff $\pi$ is a solution of the numerical projection $Q_F$.

This is because for every trajectory $s_0, s_1, \ldots$ induced by policy $\pi$ over the numerical problem $Q_F$, there is a trajectory $s_0, s_1, \ldots$ induced by $\pi$ over the boolean problem $Q_{F}'$, and vice versa, where $s_0$ is the boolean projection of the state $s_i$; namely, $p$ is true in $s_i'$ iff $p$ is true in $s_i$, and $n = 0$ is true (resp. false) in $s_i'$ iff $n$ has value (resp. greater than) 0 in $s_i$. 4670
We say that $Q_F^+$ is a qualitative FOND associated with the boolean projection $Q_F^0$ if the strong cyclic solutions $\pi$ of $Q_F^+$ represent qualitative solutions of $Q_F^0$; i.e., strong cyclic solutions of $Q_F^+$ that terminate [Srivastava et al., 2011b]. Under some conditions, roughly, that variables that are decreased by some action are not increased by other actions, $Q_F^+$ can be set to $Q_F^0$ itself. In other cases, a suitable translation is needed [Bonet et al., 2017]. Provided a sound translation, solutions to a generalized problem $Q$ can be computed from $Q_F^+$ using off-the-shelf FOND planners.\footnote{The translation by Bonet et al. [2017] is actually not sound in general as claimed. We’ll report the fix elsewhere.}

**Theorem 12.** Let $Q_F$ be a feature projection of a generalized problem $Q$, and let $Q_F^+$ be a qualitative FOND problem associated with the boolean projection $Q_F^0$. The strong cyclic solutions of $Q_F^+$ are solutions of the generalized problem $Q$.

This is a soundness result. Completeness is lost already in the reduction from $Q$ to $Q_F$ as discussed above.

**Example.** For $Q_{\text{clear}}$ and the projection $Q_F$ above with $F = \{H, n(x)\}$, the boolean projection $Q_F^0 = \langle V_F, I_F, G_F, A_F^0 \rangle$ has boolean variables $V_F = \{H, n(x) = 0\}$, initial and goal formulas $I_F = \{\neg H, n(x) > 0, H, n(x) = 0\}$, and actions $A_F^0 = \{\Delta_1, \Delta_2\}$ where $\Delta_1 = (\neg H, n(x) > 0, H, n(x) = 0)$ and $\Delta_2 = (H, \neg H)$. Since there are no actions in $Q_F^0$ that increment the numerical variable $n(x)$, $Q_F^+$ is $Q_F^0$ and hence, by Theorem 12, the strong cyclic solutions to $Q_F^0$ are solutions to $Q_{\text{clear}}$. The policy shown in (3) was computed from $Q_F^0$ by the FOND planner MyND [Mattnueier et al., 2010] in 58 milliseconds. □

### 6 Examples and Experiments

We illustrate the representation changes and the resulting methods for computing policies in four problems. Experiments were done on an Intel i5-4670 CPU with 8Gb of RAM.

#### 6.1 Moving in Rectangular Grids

The generalized problem $Q_{\text{move}}$ involves an agent that moves in an arbitrary $n \times m$ grid. The instances $P$ are represented with atoms $a(t, x, y)$ and actions $\text{Move}(x, y, x', y')$, and have goals of the form $a(x', y')$. A general policy for $Q_{\text{move}}$ can be obtained by introducing the features $\Delta_X = |x' - x|$ and $\Delta_Y = |y' - y|$ where $a(x, y)$ is true in the state $s$. A projection $Q_F^0$ of $Q_{\text{move}}$ is obtained from this set of features $F$, the goal formula $G_F = \{\Delta_X = 0, \Delta_Y = 0\}$, and the initial DNF formula $I_F$ with four terms corresponding to the four truth valuations of the atoms $\Delta_i = 0$ and $\Delta_i = 1$, and their negations, interpreted as propositional literals, and the two actions transformed as:

- $\text{Move-blank} = (\Delta_b > 0; \Delta_b \downarrow)$,
- $\text{Move-tile} = (\Delta_b = 0, \Delta_i = 0; \Delta_i \downarrow, \Delta_b \uparrow)$.

The boolean projection $Q_F^+$ is $Q_F^0$ with the atoms $\Delta_i = 0$ and $\Delta_i = 1$, and their negations, interpreted as propositional literals, and the two actions transformed as:

- $\text{Move-blank} = (\Delta_b > 0; \Delta_b \downarrow)$,
- $\text{Move-tile} = (\Delta_b = 0, \Delta_i = 0; \Delta_i \downarrow, \Delta_b \uparrow)$.

As before, $Q_F^+$ is equal to $Q_F^0$, because the only action that increments a variable, $\Delta_b$, cannot be used until $\Delta_b = 0$. MyND solves $Q_F^+$ in 65 milliseconds, producing the policy below which by Theorem 12 solves the generalized problem $Q_{\text{move}}$:

- $\Delta_b > 0, \Delta_i > 0 \Rightarrow \text{Move-blank}$,
- $\Delta_b = 0, \Delta_i > 0 \Rightarrow \text{Move-tile}$.

#### 6.2 Sliding Puzzles

We consider next the generalized problem $Q_{\text{slide}}$ where a designated tile $t^*$ must be moved to a target location $(x_0^*, y_0^*)$ in a sliding puzzle. The STRIPS encoding of an instance $P$ contains atoms $a(t, x, y)$ and $a(B(x, y)$ for the location of tiles and the “blank”, and actions $\text{Move}(t, x, y, x', y')$ for exchanging the location of the tile $t$ and the blank if in adjacent cells. For solving $Q_{\text{slide}}$, we consider two numerical features: the Manhattan distance $\Delta$ from the current location of the target tile to its target location, and the minimal total distance $\Delta_b$ that the blank must traverse without going through the current target location so that that target tile can be moved and decrement the value of $\Delta$.

The feature projection $Q_F^0$ has goal formula $G_F = \{\Delta = 0\}$, initial formula $I_F$ given by the DNF with four terms corresponding to the four truth valuations of the atoms $\Delta_i = 0$ and $\Delta_i = 1$, and the two abstract actions

- $\text{Move-blank} = (\Delta_b > 0; \Delta_b \downarrow)$,
- $\text{Move-tile} = (\Delta_b = 0, \Delta_i > 0; \Delta_i \downarrow, \Delta_b \uparrow)$.

The boolean projection $Q_F^+$ is $Q_F^0$ with the atoms $\Delta_i = 0$ and $\Delta_i = 1$, and their negations, interpreted as propositional literals, and the two actions transformed as:

- $\text{Move-blank} = (\Delta_b > 0; \Delta_b \downarrow)$,
- $\text{Move-tile} = (\Delta_b = 0, \Delta_i > 0; \Delta_i \downarrow, \Delta_b \uparrow)$.

As before, $Q_F^+$ is equal to $Q_F^0$, because the only action that increments a variable, $\Delta_b$, cannot be used until $\Delta_b = 0$. MyND solves $Q_F^+$ in 65 milliseconds, producing the policy below which by Theorem 12 solves the generalized problem $Q_{\text{move}}$:

- $\Delta_b > 0, \Delta_i > 0 \Rightarrow \text{Move-blank}$,
- $\Delta_b = 0, \Delta_i > 0 \Rightarrow \text{Move-tile}$.

#### 6.3 Blocksworld: Achieving on($x, y$)

The problem $Q_{\text{on}}$ is about achieving goals of the form $\text{on}(x, y)$ in Blocksworld instances where for simplicity, the gripper is initially empty, and the blocks $x$ and $y$ are in different towers with blocks above them. We use a set of features $F$ given by $n(x)$ and $n(y)$ for the number of blocks above $x$ and $y$, booleans $X$ and $H$ that are true when the gripper is holding $x$ or another block, and $\text{on}(x, y)$ that is true when $x$ is on $y$. As before, we include a sound but incomplete set of actions $A_F$ needed to solve $Q_{\text{on}}$ where $E$ abbreviates the conjunction $\neg X$ and $\neg H$:
– Pick-\(x = (E, n(x) = 0; X)\),
– Pick-above-\(x = (E, n(x) > 0; H, n(x)\not\in i)\),
– Pick-above-\(y = (E, n(y) > 0; H, n(y)\not\in i)\),
– Put-x-on-\(y = (X, n(y) = 0; \neg X, on(x, y), n(y)\not\in i)\),
– Put-other-aside = (\(H; \neg H)\).

The projected problem \(Q_F\) has this set of actions \(A_F, I_F = \{n(x) > 0, n(y) > 0, E, \neg on(x, y)\}\), and \(G_F = \{on(x, y)\}\).

The projection \(Q_F\) is \(Q_F\) but with the propositional reading of the atoms \(n(\cdot) = 0\) and their negations, and actions \(A_F:\)
– Pick-\(x' = (E, n(x) = 0; X)\),
– Pick-above-\(x' = (E, n(x) = 0; H, n(x) > 0 | n(x) = 0)\),
– Pick-above-\(y' = (E, n(y) = 0; H, n(y) > 0 | n(y) = 0)\),
– Put-x-on-\(y' = (X, n(y) > 0; \neg X, on(x, y), n(y) > 0)\)
– Put-other-aside' = (\(H; \neg H)\).

Since the effects that increment a variable also achieve the goal, the qualitative problem \(Q_F^+\) is equal to \(Q_F^+\). The planner MyND over \(Q_F^+\) yields the policy \(\pi\) below in 70 milliseconds, that solves \(Q_F^+\) and hence \(Q_m\). The negated goal condition \(-on(x, y)\) is part of the following rules but it is for reasons of clarity:

– \(E, n(x) > 0, n(y) > 0 \Rightarrow \text{Pick-above-}x'\),
– \(H, \neg X, n(x) > 0, n(y) > 0 \Rightarrow \text{Put-other-aside}'\),
– \(H, \neg X, n(x) = 0, n(y) > 0 \Rightarrow \text{Put-other-aside}'\),
– \(E, n(x) = 0, n(y) > 0 \Rightarrow \text{Pick-above-y}'\),
– \(H, \neg X, n(x) = 0, n(y) = 0 \Rightarrow \text{Put-other-aside}'\),
– \(E, n(x) = 0, n(y) = 0 \Rightarrow \text{Pick-above-x}'\),
– \(X, \neg H, n(x) = 0, n(y) = 0 \Rightarrow \text{Put-x-on-y}'\).

6.4 Blocksworld: Building a Tower

We consider a final generalized blocks problem, \(Q_{\text{tower}}\), where the task is building a tower with all the blocks. For this, we consider the feature set \(F'\) and the set of abstract actions \(A_{F'}\) above (end of Sect. 3), with \(I_{F'} = \{\neg X, \neg H, Z, n(x) > 0, m(x) > 0\}\) and \(G_{F'} = \{\neg X, \neg H, m(x) = 0\}\). For space reasons we do not show the projections \(Q_{F'}^+\) and \(Q_{F'}^-\), or the problem \(Q_{F'}^\ast\) fed to the planner. The resulting policy

– \(\neg X, \neg H, m(x) > 0 \Rightarrow \text{Pick-other}\),
– \(\neg X, H, m(x) > 0 \Rightarrow \text{Put-above-x}\),
– \(\neg X, H, m(x) = 0 \Rightarrow \text{Put-above-x}\).

is obtained with the FOND-SAT-based planner [Geffner and Geffner, 2018] in 284 milliseconds (the FOND planner MyND produces a buggy plan in this case). Interestingly, the addition of the atom \(\neg Z\) to \(G_{F'}\) yields a very different policy that builds a single tower but with \(x\) at the bottom.

7 Discussion

Arbitrary Goals, Concepts, Indexicals, and Memory

Most of the examples above deal with instances involving atomic goals (except \(Q_{\text{tower}}\)). The definition of a generalized problem that would yield a policy for solving any Blocksworld instance is more challenging. Inductive, as opposed to model-based approaches, for obtaining such policies have been reported [Martín and Geffner, 2004; Fern et al., 2006]. These approaches learn generalized policies from sampled instances and their plans. They do not learn boolean and numerical features but unary predicates or concepts like “the clear block that is above \(x\)”, yet features, concepts, and indexical or deictic representations [Chapman, 1989; Ballard et al., 1997] are closely related to each other. The execution of general policies requires tracking a fixed number of features, and hence constant memory, independent of the instance size, that is given by the number of features. This is relevant from a cognitive point view where short term memory is bounded and small [Ballard et al., 1995].

General Policy Size, Polynomial Features, and Width

The numerical feature \(h^*\) that measures the optimal distance to the goal can be used in the policy \(\pi_{\text{Gen}}\) given by the rule \(h^* > 0 \Rightarrow MoveToGoal\) for solving any problem. Here MoveToGoal is the abstract action \(a = (h^* > 0; h^*)\) that is sound and represents any concrete optimal action. The problem with the feature \(h^*\) is that its computation is intractable in general, thus it is reasonable to impose the requirement that features should be computable in (low) polynomial time. Interestingly, however, instances over many of the standard classical domains featuring atomic goals have a bounded and small width, which implies that they can be solved optimally in low polynomial time [Lipovetzky and Geffner, 2012]. This means that the general policy \(\pi_{\text{Gen}}\) is not useless after all, as it solves all such instances in polynomial time. The general policy \(\pi\) computed by FOND planners above, however, involves a fixed set of boolean variables that can be tracked in time that is constant and does not depend on the instance size.

Deep Learning of Generalized Policies, and Challenge

Deep learning methods have been recently used for learning generalized policies in domains like Sokoban [Groshev et al., 2017] and 3D navigation [Mirowski et al., 2016]. Deep learning methods however learn functions from inputs of a fixed size. Extensions for dealing with images or strings of arbitrary size have been developed but images and strings have a simple 2D or linear structure. The structure of relational representations is not so uniform and this is probably one of the main reasons that we have not yet seen deep (reinforcement) learning methods being used to solve arbitrary instances of the blocks world. Deep learning methods are good for learning features but in order to be applicable in such domains, they need the inputs expressed in terms of a fixed number of features as well, such as those considered in this work. The open challenge is to learn such features from data. The relevance of this work to learning is that it makes precise what needs to be learned. A crisp requirement is that the set of
features $F$ to be learned for a generalized problem $Q$ should support a sound set of actions $A_F$ sufficient for solving the problem.

8 Summary

We have extended the standard semantic formulation of generalized planning to domains with instances that do not have actions in common by introducing abstract actions: actions that operate on the common pool of features, and whose soundness and completeness can be determined. General plans map features into abstract actions, which if sound, can always be instantiated with a concrete action. By a series of reductions, we have also shown how to obtain such policies using off-the-shelf FOND planners. The work relates to a number of concepts and threads in AI, and raises a number of crisp challenges, including the automatic discovery of the boolean and integer features that support general plans in relational domains.

Acknowledgements

We thank the anonymous reviewers for useful comments. H. Geffner is partially funded by grant TIN-2015-67959-P, MINECO, Spain.

References


