Abstraction of Nondeterministic Situation Calculus Action Theories

Bita Banihashemi\textsuperscript{1}, Giuseppe De Giacomo\textsuperscript{2}, Yves Lespérance\textsuperscript{3}

\textsuperscript{1}Ronin Institute
\textsuperscript{2}University of Oxford
\textsuperscript{3}York University
bita@ronininstitute.org, giuseppe.degiacomo@cs.ox.ac.uk, lesperan@eecs.yorku.ca

Abstract

We develop a general framework for abstracting the behavior of an agent that operates in a nondeterministic domain, i.e., where the agent does not control the outcome of the nondeterministic actions, based on the nondeterministic situation calculus and the ConGolog programming language. We assume that we have both an abstract and a concrete nondeterministic basic action theory, and a refinement mapping which specifies how abstract actions, decomposed into agent actions and environment reactions, are implemented by concrete ConGolog programs. This new setting supports strategic reasoning and strategy synthesis, by allowing us to quantify separately on agent actions and environment reactions. We show that if the agent has a (strong FOND) plan/strategy to achieve a goal/complete a task at the abstract level, and it can always execute the nondeterministic abstract actions to completion at the concrete level, then there exists a refinement of it that is a (strong FOND) plan/strategy to achieve the refinement of the goal/task at the concrete level.

1 Introduction

When working in realistic dynamic domains, the use of abstraction has proven essential in many areas of artificial intelligence, for example, in improving the efficiency of planning (e.g., [Chen and Bercher, 2021]), explaining agents’ behavior (e.g., [Seegebarth et al., 2012]), and in reinforcement learning (e.g., [Sutton et al., 1999]).

Recently, [Banihashemi et al., 2017] (BDL17) proposed a formal account of agent abstraction based on the situation calculus [McCarthy and Hayes, 1969; Reiter, 2001] and the ConGolog agent programming language [De Giacomo et al., 2000]. They assume that one has a high-level/abstract action theory, a low-level/concrete action theory, both representing the agent’s behavior at different levels of detail, and a refinement mapping between the two. The refinement mapping specifies how each high-level action is implemented by a low-level ConGolog program and how each high-level fluent can be translated into a low-level state formula. This work defines notions of abstractions between such action theories in terms of the existence of a suitable bisimulation relation [Milner, 1971] between their respective models. Abstractions have many useful properties that ensure that one can reason about the agent’s actions (e.g., executability, projection, and planning) at the abstract level, and refine and concretely execute them at the low level. The framework can also be used to generate high-level explanations of low-level behavior.

This framework was formulated assuming a deterministic environment, as usual in the situation calculus. Hence, the only nondeterminism was coming from ConGolog programs and was angelic in nature, i.e., under the control of the agent [Levesque et al., 1997; De Giacomo et al., 2000].

However, many agents operate in nondeterministic environments where the agent does not fully control the outcome of her actions (e.g., flipping a coin where the outcome may be heads or tails). Recently, [De Giacomo and Lespérance, 2021] (DL21) proposed a simple and elegant situation calculus account of nondeterministic environments where they clearly distinguish between the nondeterminism associated with agent choices and that associated with environment choices, the first being angelic for the agent, and the second being devilish, i.e., not under the agent’s control.\textsuperscript{1} The presence of environment nondeterminism deeply influences reasoning about action: if the agent wants to complete a task, she cannot simply do standard logical reasoning (i.e., satisfiability/logical implication), but needs to do realizability/synthesis, i.e., devise a strategy that, in spite of the uncontrollable reactions of the environment, guarantees successful completion of the task [Church, 1963; Pnueli and Rosner, 1989; Abadi et al., 1989]. We see this, for example in planning, where in the classical deterministic setting the problem can be solved by heuristic search, while in fully observable nondeterministic (FOND) domains forms of adversarial (AND/OR) search are required [Cimatti et al., 1998; Haslum et al., 2019].

In this paper, we develop an account of agent abstraction in presence of nondeterministic environments based on (DL21). The refinement mapping between high-level fluents and low-level state formulas remains as in (BDL17). Instead we consider each high-level action, now nondeterministic, as being

\textsuperscript{1}Earlier accounts dealing with the topic in the situation calculus were more complex and did not clearly distinguish between these two forms of nondeterminism, e.g., [Pinto et al., 2000] which deals with stochastic actions, and [Bacchus et al., 1999] which deals with uncertainty, noisy acting, and noisy sensing.
composed of an agent action and an (implicit) environment reaction. As a consequence, we map the agent action (without the environment reaction) into a low-level agent program that appropriately reflects the nondeterminism of the environment, and the complete high-level action, including both the agent action and the environment reaction, into a low-level system program that relates the high-level environment reaction to the low-level ones.

We show that the (BDL17) notion of m-bisimulation extends naturally to this new setting. This allows for exploiting abstraction when reasoning about executions as in (BDL17), in spite of the nondeterministic environment. But, notably, this new setting now supports strategic reasoning and strategy synthesis, by allowing us to quantify separately on agent actions and environment reactions. As a result, we can exploit abstraction by finding abstract strategies at the high-level and then refine them into concrete low-level ones. In particular, we show that if the agent has a (strong FOND) plan/strategy to achieve a goal/complete a task at the high level (i.e., no matter what the environment does), and she can always execute the nondeterministic high-level actions to completion at the low level (even if not controlling their outcome), then there exists a refinement of that it is a (strong FOND) plan/strategy to achieve the refinement of the goal/task at the low level.

2 Preliminaries

Situation calculus. The situation calculus is a well known predicate logic language for representing and reasoning about dynamically changing worlds [McCarthy and Hayes, 1969; Reiter, 2001]. All changes to the world are the result of actions, which are terms in the language. A possible world history is represented by a term called a situation. The constant \(S_0\) is used to denote the initial situation. Sequences of actions are built using the function symbol do, such that do\((a, s)\) denotes the successor situation resulting from performing action \(a\) in situation \(s\). Predicates and functions whose value varies from situation to situation are called fluents, and are denoted by symbols taking a situation term as their last argument (e.g., \(\text{Open}(\text{Door}1, s)\)). The abbreviation do\((\langle a_1, \ldots, a_n \rangle, s)\) stands for do\((a_1, \text{do}(a_2, \ldots, \text{do}(a_n, s))\)) for an action sequence \(\vec{a}\), we often write do\((\vec{a}, s)\) for do\((\vec{a}, s)\). In this language, a dynamic domain can be represented by a basic action theory (BAT), where successor state axioms represent the causal laws of the domain and provide a solution to the frame problem [Reiter, 2001]. A special predicate Poss\((a, s)\) is used to state that action \(a\) is executable in situation \(s\); the precondition axioms characterize this predicate. Abbreviation Executable\((s)\) means that every action performed in reaching situation \(s\) was possible in the situation in which it occurred.

Nondeterministic situation calculus. A major limitation of the standard situation calculus and BATs is that atomic actions are deterministic. [De Giacomo and Lespérance, 2021] (DL21) propose a simple extension of the framework to handle nondeterministic actions while preserving the solution to the frame problem. For any primitive action by the agent in a nondeterministic domain, there can be a number of different outcomes. (DL21) takes the outcome as being determined by the agent’s action and the environment’s reaction to this action. This is modeled by having every action type/function \(A(\vec{x}, e)\) take an additional environment reaction parameter \(e\), ranging over a new sort Reaction of environment reactions. The agent cannot control the environment reaction, so it performs the reaction-suppressed version of the action \(A(\vec{x})\) and the environment then selects a reaction \(e\) to produce the complete action \(A(\vec{x}, e)\). We call the reaction-suppressed version of the action \(A(\vec{x})\) an agent action and the full version of the action \(A(\vec{x}, e)\) a system action.

Nondeterministic basic action theories (NDBATs). These can be seen as a special kind of BAT, where every action function takes an environment reaction parameter, and moreover, for each agent action \(A(\vec{x})\), we

- have its agent action precondition denoted by: Poss\(_{ag}(A(\vec{x}), s) \equiv A(\vec{x}, \text{Poss}(\vec{x}, e), s)\);
- have a reaction independence requirement, stating that the precondition for the agent action is independent of any environment reaction: \(\forall e.\text{Poss}(A(\vec{x}, e), s) \supset \text{Poss}_{ag}(A(\vec{x}), s)\);
- have a reaction existence requirement, stating that if the precondition of the agent action holds then there exists a reaction to it which makes the complete system action executable: Poss\(_{ag}(A(\vec{x}), s) \supset \exists e.\text{Poss}(A(\vec{x}, e), s)\).

The above requirements must be entailed by the action theory for it to be an NDBAT.

A NDBAT \(D\) is the union of the following disjoint sets: foundational, domain independent, axioms of the situation calculus (\(\Sigma\)) as in standard BATs, axioms describing the initial situation \(D_{\text{init}}\) as in standard BATs, unique name axioms for actions (\(D_{\text{unam}}\)) as in standard BATs, successor state axioms (SSAs) describing how fluents change after system actions are performed (\(D_{\text{ssa}}\)), and system action precondition axioms, one for each action type, stating when the complete system action can occur (\(D_{\text{pos}}\)); these are of the form: Poss\((A(\vec{x}, e), s) \equiv A(\vec{x}, e, s)\).

Example. Our running example is based on a triangle tire-world domain (see Fig. 1). The agent’s goal is to drive from location 11 to location 13. When driving from one location to the next, the possibility of a tire going flat exists. If there is a spare tire in the location where the car is at (indicated by circles in Fig. 1), the agent can use it to fix a flat.

NDBAT \(D^{\text{f}}\) models the domain at the low level. The system action drive\((o, d, r)\) can be performed by the agent to drive from origin location \(o\) to destination location \(d\), and \(r\) indicates environment’s reaction. This action is executable

![Figure 1: A 3-Level Triangle Tireworld](image-url)
when the agent action \textit{drive}(o, d) is executable (i.e., when the agent is at o, a road connects o to d, and agent does not have a flat tire), and \( r \) can take on two values: \textit{FlatTire} if the tire goes flat and \textit{NoFlatTire} otherwise. The fluent \textit{At}(l, s) indicates agent’s location. \textit{Road}(o, d, s) specifies the road connections in Fig. 1. \textit{Flat}(l, s) describes whether there is a flat tire, and \textit{Visited}(l, s) indicates if the location \( l \) has already been visited by the agent.

The system action \textit{fixFlatTire}(l, r) fixes a flat tire, and the environment reaction \( r \) is \textit{Success}_LF (assumed for simplicity). It is executable if the fluent \textit{Spare}(l, s) holds, i.e., there is a spare tire. The action \textit{wait} can be performed by the agent to remain idle at a location.

\( D_f \) includes the following action precondition axioms (throughout the paper, we assume that free variables are universally quantified from the outside):

\[
\text{Poss}_{\text{dr}}(\text{drive}(o, d), s) \equiv \\
\neg d \land \text{At}(o, s) \land \text{Road}(o, d, s) \land \neg \text{Flat}(l, s)
\]

\[
\text{Poss}_{\text{fr}}(\text{fixFlatTire}(l), s) \equiv \\
\text{At}(l, s) \land \text{Spare}(l, s) \land \text{Flat}(l, s)
\]

\[
\text{Poss}_{\text{wa}}(\text{wait}(l, s), s) \equiv \text{true}
\]

\[
\text{Poss}_{\text{fr}}(\text{fixFlatTire}(l), s) \land r = \text{FlatTire} \lor r = \text{NoFlatTire}
\]

\[
\text{Poss}_{\text{wa}}(\text{wait}(l, s), s) \land r = \text{Success}_LF
\]

For other fluent, we have SSAs specifying that they are unaffected by any action.

\( D_f \) also contains the following initial state axioms:

\[
\text{At}(l, S_0) \equiv \exists o, r, a = \text{drive}(o, l, r) \lor \text{At}(l, S_0) \land \forall d, r, a \neq \text{drive}(d, r)
\]

\[
\text{Flat}(l, S_0) \equiv \\
\exists o, d, a = \text{drive}(o, d, \text{FlatTire}) \lor \text{Flat}(l, S_0) \land r, l, a \neq \text{fixFlatTire}(l, r)
\]

\[
\text{Visited}(l, S_0) \equiv \\
\exists o, r, a = \text{drive}(o, l, r) \lor \text{Visited}(l, S_0)
\]

For the other fluents, we have SSAs specifying that they are unaffected by any action.

\( D_f \) also contains the following initial state axioms:

\[
\text{Road}(o, d, S_0) \equiv \{ (11, 12), (11, 21), (12, 21), (12, 22), (12, 13), (13, 22), (22, 31), (21, 31), (21, 12), (21, 11), (21, 22), (12, 22), (13, 12), (12, 23), (31, 22), (31, 21) \}
\]

\[
\text{Spare}(l, S_0) \equiv \{ (21, 22, 31), \text{Visited}(l, S_0) \} \land 1 = l
\]

\[
\text{At}(l, S_0) \equiv (l = 11) \land \text{Flat}(l, S_0) \land l = 13 \land \neg \text{Flat}(l, S_0)
\]

\textbf{FOND planning and synthesis in NDBATs.} We start with some definitions. A weak plan is one that achieves the goal when the environment “cooperates” and selects environment reactions to nondeterministic actions that make this happen. Formally, we say that a sequence of agent actions \( \bar{a} \) is a weak plan to achieve \textit{Goal} if \textit{Goal}(s’) holds, i.e., there exists an execution of \( \bar{a} \) that takes us from the initial situation \( S_0 \) to situation \( s’ \) where the goal holds. \( \text{Do}_{\text{dr}}(\bar{a}, s, s’) \) means that the system may reach situation \( s’ \) when the agent executes the sequence of agent actions \( \bar{a} \) depending on environmental reactions:

\[
\text{Do}_{\text{dr}}(\epsilon, s, s’) \equiv s’ = s (\text{where } \epsilon \text{ is the empty sequence of actions})
\]

\[
\text{Do}_{\text{fr}}([\text{fixFlatTire}(l)], \sigma, s, s’) \equiv \\
\exists e. \text{Poss}(\text{fixFlatTire}(l), e, s) \land \text{Do}_{\text{dr}}(\sigma, \text{do}(\text{fixFlatTire}(l), e), s, s’)
\]

A strong plan for the agent guarantees the achievement of a goal no matter how the environment reacts; it is a strategy for the agent to follow to ensure that the goal is achieved.

\textbf{Example cont.} Strategies use the \textit{stop} action, hence we define a \textit{stop} action at the low level; this action always terminates with the \textit{Success}_LF reaction. The strategy that guarantees reaching location 13 is defined as follows:

\[
\text{D}_f \models \text{AgtCanForceBy}(\text{At}(13, S_0, f))
\]

where

\[
\begin{align*}
\text{stop}_{\text{LF}}(l, s) & \equiv \text{stop}_{\text{LF}}(l, s) \land l \equiv 13 \land \text{Flat}_{\text{LF}}(s) \\
\text{fixFlatTire}(l, s) & \equiv \text{fixFlatTire}(l, s) \land l \equiv 13 \land \text{Flat}_{\text{LF}}(s) \\
\text{drive}(a, d) & \equiv \text{drive}(a, d) \land l \equiv 13 \land \neg \text{Flat}_{\text{LF}}(s) \\
\text{Spare}(l, s) & \equiv \text{Spare}(l, s) \land l \equiv 13 \land \neg \text{Flat}_{\text{LF}}(s) \\
\text{Visited}(l, s) & \equiv \text{Visited}(l, s) \land l \equiv 13 \land \neg \text{Flat}_{\text{LF}}(s)
\end{align*}
\]

That is, the agent should stop if she is already at location 13, otherwise she drives to a location she has not visited before and that has a spare tire; in case of a flat tire, action \textit{fixFlatTire} is executed; in all other cases the agent waits.

\textbf{High-level programs and ConGolog.} To represent and reason about complex actions or processes obtained by suitably executing atomic actions, various so-called high-level programming languages have been defined. Here we concentrate on (a variant of) ConGolog [De Giacomo et al., 2000] that includes the following constructs:

\[
\delta ::= \alpha \mid \varphi \mid \delta_1 \mid \delta_2 \mid \pi.x.\delta \mid \delta^* \mid \delta_1 \parallel \delta_2
\]

In the above, \( \alpha \) is an action term, possibly with parameters, and \( \varphi \) is a situation-suppressed formula, i.e., a formula with all situation arguments in fluents suppressed. As usual, we denote by \( \varphi[s] \) the formula obtained from \( \varphi \) by restoring the situation argument \( s \) into all fluents in \( \varphi \). The sequence of program \( \delta_1 \) followed by program \( \delta_2 \) is denoted by \( \delta_1; \delta_2 \). Program \( \delta_1 \parallel \delta_2 \) allows for the nondeterministic choice between programs \( \delta_1 \) and \( \delta_2 \), while \( \pi.x.\delta \) executes program \( \delta \) for some nondeterministic choice of a binding for object variable \( x \) (observe that such a choice is, in general, unbounded). \( \delta^* \) performs \( \delta \) zero or more times. Program \( \delta_1 \parallel \delta_2 \) expresses
the concurrent execution (interpreted as interleaving) of programs $\delta_1$ and $\delta_2$. The construct if $\phi$ then $\delta_1$ else $\delta_2$ endif is defined as $[\phi\?;\delta_1]\mid[\neg\phi\?;\delta_2]$. We also use $\text{nil}$, an abbreviation for $\text{True}$, to represent the empty program, i.e., when nothing remains to be performed.

Formally, the semantics of ConGolog is specified in terms of single-step transitions, using the following two predicates: (i) $\text{Trans}(\delta, s, \delta', s')$, which holds if one step of program $\delta$ in situation $s$ may lead to situation $s'$ with $\delta'$ remaining to be executed; and (ii) $\text{Final}(\delta, s)$, which holds if program $\delta$ may legally terminate in situation $s$. The definitions of $\text{Trans}$ and $\text{Final}$ we use are as in [De Giacomo et al., 2010], differently from [De Giacomo et al., 2000], the test construct $\varphi\,?$ does not yield any transition, but is final when satisfied.

Predicate $\text{Do}(\delta, s, s')$ means that program $\delta$, when executed starting in situation $s$, has as a legal terminating situation $s'$, and is defined as $\text{Do}(\delta, s, s') \triangleq \exists \delta'.\text{Trans}(\delta, s, \delta', s') \land \text{Final}(\delta', s')$ where $\text{Trans}$ denotes the reflexive transitive closure of $\text{Trans}$. We use $\mathcal{C}$ to denote the axioms defining the ConGolog programming language.

For simplicity in this paper, we use a restricted class of ConGolog programs which are situation-determined (SD) [De Giacomo et al., 2012], i.e., for every sequence of actions, the remaining program is uniquely determined by the resulting situation:

$$\text{SituationDetermined}(\delta, s) \triangleq \forall s', \delta', \delta''.\ \text{Trans}^*(\delta, s, \delta', s') \land \text{Trans}^*(\delta, s, \delta'', s') \supset \delta' = \delta''$$

**ConGolog program execution in ND domains.** To specify agent behaviors in nondeterministic domains, (DL21) use ConGolog agent programs, which are composed as usual, but only involve agent (reaction-suppressed) atomic actions. The semantics for such agent programs is a variant of the one above where a transition for an atomic agent action may occur whenever there exists a reaction that can produce it.

**Strategic reasoning in executing programs in ND domains.** To represent the ability of the agent to execute an agent program in a ND domain (DL21) introduce $\text{AgtCanForceBy}(\delta, s, f)$ as an adversarial version of $\text{Do}$ in presence of environment reactions. This predicate states that strategy $f$, a function from situations to agent actions (including the special action $\text{stop}$), executes SD ConGolog agent program $\delta$ in situation $s$ considering its nondeterminism angelic, as in the standard $\text{Do}$, but also considering the nondeterminism of environment reactions devilish-adversarial:

$$\text{AgtCanForceBy}(\delta, s, f) \triangleq \forall P,[\ldots \supset P(\delta, s)]$$

where $\ldots$ stands for:

- $[f(\delta) = \text{stop} \land \text{Final}(\delta, s)] \supset P(\delta, s)$
- $[\exists A, \exists f(\delta) = A(t) \neq \text{stop} \land \exists e. \exists \delta'. \text{Trans}(\delta, s, \delta', \text{do}(A(t, e), s)) \supset \forall e. (\exists \delta'. \text{Trans}(\delta, s, \delta', \text{do}(A(t, e), s)) \supset P(\delta, s))$.

We say that predicate $\text{AgtCanForceBy}(\delta, s, f)$ holds iff there exists a strategy $f$ s.t. $\text{AgtCanForceBy}(\delta, s, f)$ holds.

**Example cont.** Suppose we have a program $\delta_{\text{go}}^{\text{LL}}(l)$ which lets the agent drive to any location, or fix a flat any number of times until a destination $l$ is reached: $\delta_{\text{go}}^{\text{LL}}(l) = (\pi_0.d\text{rive}(a, d) \mid \pi_0.d\text{FlatTire}(d))^{*}.\text{At}_{\text{LL}}(l)$.

Now suppose the agent has been assigned the task of going to location 13. We can show that she has a strategy for executing the task $\delta_{\text{go}}^{\text{LL}}(13)$:

$$D_{13}^{\text{f}} \models \text{AgtCanForceBy}(\delta_{\text{go}}^{\text{LL}}(13), S_0, f)$$

Here, the strategy $f$ to execute the task is same as $f_1$ above.

### 3 Abstraction in Nondeterministic Domains

In this section, we show how we can extend the agent abstraction framework of (BDL17) to handle nondeterministic domains. As in (BDL17), we assume that there is a high-level/abstract (HL) action theory $D_h$ and a low-level/concrete (LL) action theory $D_l$ representing the agent’s possible behaviors at different levels of detail. In (BDL17), these are standard BATs; here, we assume that they are both NDBATs. $D_h$ (resp. $D_l$) involves a finite set of primitive action types $A_h$ (resp. $A_l$) and a finite set of primitive fluent predicates $F_h$ (resp. $F_l$). The terms of sort $\text{Object}$ are assumed be a countably infinite set $N$ of standard names for which we have the unique name assumption and domain closure. We also assume that $\text{Reaction}$ is a sub-sort of $\text{Object}$. Also, $D_h$ and $D_l$ are assumed to share no domain specific symbols except for the set of standard names for objects $N$. For simplicity and w.l.o.g., it is assumed that there are no functions other than constants and no non-fluent predicates.

**Refinement mapping.** As in (BDL17), we assume that one relates the HL and LL theories by defining a refinement mapping that specifies how HL atomic actions are implemented at the low level and how HL fluents can be translated into LL state formulas. In deterministic domains, one can simply map HL atomic actions to a LL program that the agent uses to implement the action. In nondeterministic domains however, we additionally need to specify what HL reaction the environment performs in each LL refinement of the HL action.

**Definition 1 (NDBAT Refinement Mapping).** A NDBAT refinement mapping $m$ is a triple $\langle m_a, m_s, m_f \rangle$ where $m_a$ associates each HL primitive action type $A$ in $A_h$ to a SD ConGolog agent program $\delta_h^A$ defined over the LL theory that implements the agent (i.e., reaction-suppressed) action, i.e., $m_a(A(\bar{x})) = \delta_h^A(\bar{x})$, $m_s$ associates each $A$ to a SD ConGolog system program $\delta_s^A$ defined over the LL theory that implements the system action, i.e., $m_s(A(\bar{x}, \bar{e})) = \delta_s^A(\bar{x}, \bar{e})$, thus specifying when the HL reaction occurs (system programs are interpreted using the standard transition semantics), and (as in (BDL17)) $m_f$ maps each situation-suppressed HL fluent $F(\bar{x})$ in $F_h$ to a situation-suppressed formula $\phi_F(\bar{x})$ defined over the LL theory that characterizes the concrete conditions under which $F(\bar{x})$ holds in a situation, i.e., $m_f(\bar{x}) = \phi_F(\bar{x})$.

We can extend a mapping to a sequence of agent actions in the obvious way, i.e., $m_a(\alpha_1, \ldots, \alpha_n) = m_a(\alpha_1) \ldots m_a(\alpha_n)$ for $n \geq 1$ and $m_a(\bar{e}) = \text{nil}$, and similarly for sequences of system actions. Note that $m_a(\bar{\alpha})$ is well defined even if the action parameters are not ground; but
the action functions must be given. We also extend the notation so that \( m_f(\phi) \) stands for the result of substituting every fluent \( F(\bar{x}) \) in situation-suppressed formula \( \phi \) by \( m_f(F(\bar{x})) \).

Agent actions and system actions must be mapped in a consistent way. To ensure this, we require the following:

**Constraint 2. (Proper Refinement Mapping)**

For every high-level system action sequence \( \alpha \) and every high-level action \( A \), we have that:

\[
\mathcal{D}_h \cup \mathcal{C} \models \forall s.(Do(m_s(\alpha), S_o, s) \supset \\
\forall \bar{x}, \bar{e}.(Do(m_s(A(\bar{x}), e), s, s') \equiv \exists e. Do(m_s(A(\bar{x}, e), s, s'))) 
\]

This ensures that (1) for every situation \( s' \) that can be reached from \( s \) by executing a refinement of the HL agent action \( A(\bar{x}) \), there is a HL reaction \( e \) that generates it, and (2) that for every situation \( s' \) that can be reached from \( s \) by executing a refinement of the HL system action \( A(\bar{x}, e) \), there is a refinement of the HL agent action that generates it (in the above, \( s \) is any situation reached by executing a sequence of HL system actions). Note that (1) essentially states that the reaction existence requirement holds at the low level for refinements of HL actions and (2) amounts to having the reaction independence requirement hold. We say that an NDBAT refinement mapping \( m \) is proper wrt low-level NDBAT \( \mathcal{D}_l \) if this constraint holds.

**Example cont.** NDBAT \( \mathcal{D}^l_h \) abstracts over driving and fixing potential flat tires. Action \( driveAndTryFix(o, d, r) \) can be performed to drive from origin \( o \) to destination \( d \), and fix a flat if one occurs and a spare is available. This action is executable when \( driveAndTryFix(o, d, r) \) is executable, i.e., when the agent is at \( o \), a road connects \( o \) to \( d \), and agent does not have a flat tire; the environment reaction \( r \) may take on the following values: \( DruNoFlat \), if the tire does not go flat, \( DruFlat \) if the tire goes flat and it is not possible to fix it, or \( DruFixFlat \) if the tire went flat and the flat was fixed. Action \( wait\_HL \) lets the agent remain idle at a location. \( \mathcal{D}^l_h \) includes the following system and agent action precondition axioms:

\[
\text{Poss}_{s}(drive\_And\_Try\_Fix(o, d), s) \equiv \\
\text{Poss}_{s}(wait\_HL, s) \equiv \text{TRUE} \\
\text{Poss}(drive\_And\_Try\_Fix(o, d, r), s) \equiv \\
\text{Poss}_{s}(drive\_And\_Try\_Fix(o, d), s) \land (r = DruNoFlat) \lor \\
\text{Poss}(wait\_HL, r), s) \equiv \\
\text{Poss}_{s}(wait\_HL, s) \land (r = Success\_HL) 
\]

\( \mathcal{D}^l_h \) also includes the following SSAs:

\[
\text{At\_HL}(l, do(a, s)) \equiv \exists a, r.a = drive\_And\_Try\_Fix(o, l, r) \lor \\
\text{At\_HL}(l, s) \land \forall d, r.a \neq drive\_And\_Try\_Fix(l, d, r) \\
\text{Flat\_HL}(o, s) \equiv \\
\exists a, do = drive\_And\_Try\_Fix(o, d, DruFlat) \lor Flat\_HL(s) \\
\text{Visited\_HL}(l, do(a, s)) \equiv \\
\exists a, r.a = drive\_And\_Try\_Fix(o, l, r) \lor Visited\_HL(l, s) 
\]

For the other fluents, we have SSAs specifying that they are unaffected by any action.

The initial state axioms of \( \mathcal{D}^l_h \) are same as \( \mathcal{D}^l_l \), with high-level fluents having a distinct name (with HL suffix); e.g., \( Road\_HL \) is axiomatized exactly as \( Road\_LL \).

We specify the relationship between the HL and LL ND-BATs through the following refinement mapping \( m^tt \):

\[
m^tt(drive\_And\_Try\_Fix(o, d)) = \\
drive\_HL(o, d) ; \\
\text{if } \neg Flat\_HL \text{ then nil else } \\
\text{if } \neg Spar\_HL(d) \text{ then nil else } fix\_Flat\_Tire(d) \text{ endIf} \\
\text{endif} \\
m^tt(wait\_HL) = wait\_LL \\
m^tt(drive\_And\_Try\_Fix(o, d, r)) = \\
\pi_t\_drive\_HL(o, d, r) ; \\
\text{if } \neg Flat\_HL \text{ then } r_h = Dru\_No\_Flat? \text{ else } \\
\text{if } \neg Spar\_HL(d) \text{ then } r_h = Dru\_Flat? \\
\text{else } fix\_Flat\_Tire(d, Success\_HL); r_h = Dru\_Flat\_Fix? \text{ endIf} \\
\text{ endif} \\
m^tt(wait\_HL(r_h)) = \pi_t\_wait\_HL(r_h); r_h = Success\_HL; \text{ if and only if } \\
m^tt(Flat\_HL) = Flat\_LL
\]

Thus, the HL agent action \( drive\_And\_Try\_Fix(o, d) \) is implemented by an LL program where the agent first performs \( drive\_HL(o, d) \), and depending on whether the tire has gone flat and a spare exists at location \( d \), fixes the tire or does nothing. Fluents \( At\_HL(l) \), \( Spar\_HL(l) \), \( Road\_HL(o, d) \), \( Dest\_HL(l) \), and \( Visited\_HL(l) \) are mapped to their low-level counterparts similar to \( Flat\_HL \). We can show that:

**Proposition 3.** NDBAT refinement mapping \( m^tt \) is proper wrt \( \mathcal{D}^l_l \).

\( m \)-Bisimulation. To relate the HL and LL models/theories, (BDL17) define a variant of bisimulation [Milner, 1971; Milner, 1989]. The base condition for the bisimulation is:

**Definition 4.** (\( m \)-isomorphic situations). Let \( M_h \) be model of the HL theory \( \mathcal{D}_h \), and \( M_l \) a model of the LL theory \( \mathcal{D}_l \cup \mathcal{C} \).

We say that situation \( s_h \in M_h \) is \( m \)-isomorphic to situation \( s_l \in M_l \), written \( s_h \sim^m M_h, M_l s_l \), if and only if

\[
M_h, v[s/s_h] \models F(\bar{x}, s) \text{ iff } M_l, v[s/s_l] \models m(F(\bar{x}))[s] 
\]

for every high-level primitive fluent \( F(\bar{x}) \) in \( \mathcal{F}_h \) and every variable assignment \( v \).

If \( s_h \sim^m M_h, M_l s_l \), then \( s_h \) and \( s_l \) evaluate all HL fluents the same.

**Definition 5.** (\( m \)-Bisimulation). Given \( M_h \) a model of \( \mathcal{D}_h \), and \( M_l \) a model of \( \mathcal{D}_l \cup \mathcal{C} \), a relation \( B \subseteq \Delta^m_\phi \times \Delta^m_\phi \) (where \( \Delta^m_\phi \) stands for the situation domain of \( M \)) is an \( m \)-bisimulation relation between \( M_h \) and \( M_l \) if \( (s_h, s_l) \in B \) implies that: (i) \( s_h \sim^m M_h, M_l s_l \); (ii) for every HL primitive action type \( A \) in \( \mathcal{A}_h \), if there exists \( s_h' \) such that \( M_h, v[s/s_h, s'/s'_h] \models \text{Poss}(A(\bar{x}), s) \land s' = do(A(\bar{x}), s) \), then there exists \( s_l' \) such that \( M_l, v[s/s_l, s'/s'_l] \models \text{Do}(m(A(\bar{x})), s, s') \) and \( (s'_h, s'_l) \in B \); and (iii) for every HL primitive action type \( A \) in \( \mathcal{A}_l \), if there exists \( s_l' \) such that \( M_l, v[s/s_l, s'/s'_l] \models \text{Do}(m(A(\bar{x})), s, s') \), then there exists

\[3\]For proofs, see [Banihashemi et al., 2023]

\[2\]As usual, \( M, v \models \phi \) means that model \( M \) and assignment \( v \) satisfy formula \( \phi \) (where \( \phi \) may contain free variables that are interpreted by \( v \)); also \( v[x/e] \) represents the variable assignment that is just like \( v \) but assigns variable \( x \) to entry \( e \).
\( s'_h \) such that \( M_h, v[s/s_h, s'/s'_h] \models \text{Poss}(A(x), s) \land s' = \text{do}(A(x), s) \) and \( (s'_h, s'_h') \in B \). \( M_h \) is m-bisimilar to \( M_l \), written \( M_h \sim_m M_l \), if and only if there exists an m-bisimulation relation \( B \) between \( M_h \) and \( M_l \) such that \( (s'_h, s'_h') \in B \).

The definition of m-bisimulation can remain as in (BDL17) where conditions (ii) and (iii) are applied to high-level system primitive actions and their mapping \( m_s \).

The definition of m-bisimulation ensures that performing a HL system action results in m-bisimilar situations. We can show that with the restriction to proper mappings, this automatically carries over to HL agent actions (so there is no need to change the definition of m-bisimulation to get this):

**Theorem 6.** Suppose that \( M_h \sim_m M_l \), where \( M_h \models \mathcal{D}_h \), \( M_l \models \mathcal{D}_l \cup \mathcal{C} \), \( D_h \) and \( D_l \) are both NDBATs, and \( m \) is proper wrt \( D_l \). Then for any HL system action sequence \( \alpha \) and any HL primitive action \( A \), we have that:

1. If there exist \( s_h \) and \( s'_h \) such that \( M_h, v[s/s_h, s'/s'_h] \models \text{Do}(\bar{\alpha}, S_0, s) \land \text{Do}_{x_0}(A(x), s, s') \), then there exist \( s_l \) and \( s'_l \) such that \( M_l, v[s/s_l, s'/s'_l] \models \text{Do}(\bar{\alpha}, S_0, s) \land \text{Do}_{x_0}(m_{\bar{\alpha}}(A(x)), s, s') \), \( s_h \sim_{m, M_h} s_l \) and \( s'_h \sim_{m, M_h} s'_l \);
2. If there exist \( s_l \) and \( s'_l \) such that \( M_l, v[s/s_l, s'/s'_l] \models \text{Do}(\bar{\alpha}, S_0, s) \land \text{Do}_{x_0}(m_{\bar{\alpha}}(A(x)), s, s') \), then there exist \( s_h \) and \( s'_h \) such that \( M_h, v[s/s_h, s'/s'_h] \models \text{Do}(\bar{\alpha}, S_0, s) \land \text{Do}_{x_0}(A(x), s, s') \), \( s_h \sim_{m, M_h} s_l \) and \( s'_h \sim_{m, M_h} s'_l \).

(BDL17) use m-bisimilarity to define notions of sound/complete abstraction between a high-level action theory and a low-level one: \( \mathcal{D}_h \) is a sound abstraction of \( \mathcal{D}_l \) relative to refinement mapping \( m \) if and only if, for all models \( M_l \) of \( \mathcal{D}_l \cup \mathcal{C} \), there exists a model \( M_h \) of \( \mathcal{D}_h \) such that \( M_h \sim_m M_l \). With a sound abstraction, whenever the high-level theory entails that a sequence of actions is executable and achieves a certain condition, then the low level must also entail that there exists an executable refinement of the sequence such that the "translated" condition holds afterwards. Moreover, whenever the low level considers the executability of a high-level action is satisfiable, then the high level does also. A dual notion is also defined: \( \mathcal{D}_h \) is a complete abstraction of \( \mathcal{D}_l \) relative to refinement mapping \( m \) if and only if, for all models \( M_h \) of \( \mathcal{D}_h \), there exists a model \( M_l \) of \( \mathcal{D}_l \cup \mathcal{C} \) such that \( M_l \sim_m M_h \).

The notion of m-bisimulation provides the semantic underpinning of these notions of abstraction. (BDL17) also prove the following results that identify a set of properties that are necessary and sufficient to have a sound/complete abstraction (here we adjust the notation to use system actions) and can be used to verify that one has a sound/complete abstraction, e.g., by model checking or theorem proving techniques:

**Theorem 7 (BDL17).** \( \mathcal{D}_h \) is a sound abstraction of \( \mathcal{D}_l \) relative to mapping \( m \) if and only if for any sequence of high-level system actions \( \bar{\alpha} \):

(a) \( D^s_{\mathcal{D}_h} \cup D^s_{\mathcal{D}_l} \cup D^s_{\mathcal{C}_l} = m(\phi) \), for all \( \phi \in D^s_{\mathcal{D}_h} \);
(b) \( \mathcal{D}_l \cup \mathcal{C} = \forall s. \text{Do}(m_{\bar{\alpha}}(\bar{\alpha}), S_0, s) \supset \text{Do}(m_{\bar{\alpha}}(\bar{\alpha}), S_0, s) \supset \text{Do}(m_{\bar{\alpha}}(\bar{\alpha}), S_0, s) \models m(\phi) \).

(c) \( D^s \cup \mathcal{C} \models \forall s. \text{Do}(m_{\bar{\alpha}}(\bar{\alpha}), S_0, s) \supset \text{Do}(m_{\bar{\alpha}}(\bar{\alpha}), S_0, s) \supset \text{Do}(m_{\bar{\alpha}}(\bar{\alpha}), S_0, s) \models m(\phi) \).

Here and below, sequences of high-level action systems \( \bar{\alpha} \) may contain free variables in the action parameters, but the action functions must be given, and similarly for sequences of high-level agent actions (this generalizes (BDL17) which only considers ground action sequences).

We can extend this result to sequences of agent actions by exploiting the fact that \( m \) is a proper mapping:

**Theorem 12.** If \( M_h \sim_m M_l \), then for any sequence of high-level agent actions \( \bar{\alpha} \) and any high-level situation-suppressed formula \( \phi \), we have that:

\( M_h, v = \exists s'. \text{Do}(m_{\bar{\alpha}}(\bar{\alpha}), S_0, s') \land m_{\phi}(\phi)[s'] \) if and only if \( M_h, v = \exists s'. \text{Do}(\bar{\alpha}, S_0, s', \phi)[s'] \).
This means that if the agent has a weak plan to achieve a goal at the high level, then there exists a refinement of it that is a weak plan to achieve the mapped goal at the low level.

**Example cont.** At the high level, a weak plan to achieve the goal of getting to location 13 is to first drive to 12 and then to 13:

$$D_h^t = \exists s'.D_{ao}(\tilde{a}, S_0, s') \land At_{HL}(13, s')$$

where

$$\tilde{a} = driveAndTryFix(11, 12); driveAndTryFix(12, 13)$$

This plan worked provided no flat occurs when driving to 12, as there is no spare there. By Th. 12, there exists a weak plan at the low level that refines this high-level plan:

$$D_l^t = \exists s'.D_{ao}(m_a(\tilde{a}), S_0, s') \land At_{LL}(13, s') \land D_{ao}(\{\text{drive}(11, 12); \text{drive}(12, 13)\}, S_0, s')$$

### 5 Results about Strategic Reasoning

Let us now discuss how NDBAT abstractions can be used in FOND domains to synthesize strategies to fulfill reachability/achievement goals as well as temporally extended goals/tasks. First of all, we need to consider how much strategic reasoning the agent needs to do to execute a high-level atomic action at the low level. A given HL agent action \(A(\tilde{x})\) is mapped to a LL agent program \(m_a(A(\tilde{x}))\) that implements it. As we have seen, in \(m\)-bisimilar models, Constraint 2 ensures that if \(A(\tilde{x})\) is executable at the HL, then there exists a terminating execution of \(m_a(A(\tilde{x}))\) at the LL (this holds for HL system actions as well). But this does not mean that all executions of \(m_a(A(\tilde{x}))\) terminate, as some may block or diverge, due to either agent or environment choices. In general, the agent must do strategic reasoning to ensure that the execution of \(m_a(A(\tilde{x}))\) terminates (and the environment may need to cooperate as well). But we can impose further constraints on the mapping of HL actions to avoid this or ensure that a strategy exists. Note that ensuring that the execution of \(m_a(A(\tilde{x}))\) terminates does not mean that the agent controls the action’s outcome (e.g., not having a flat); the outcome is still determined by the environment reactions.

**Constraints on HL action implementation.** First, we may want to require that the mapping of HL actions is such that the implementation program always terminates and no LL strategic reasoning is required to ensure termination. To do this, we first define:

$$InevTerminates(\delta, s) \equiv \forall \delta\{\ldots \supset P(\delta, s)\}$$

where \(\ldots\) stands for

$$[\text{Final}(\delta, s) \supset P(\delta, s)] \land \exists s'. \exists t'. \text{Trans}(\delta, s, d', s') \land \forall s'. \forall t'. \lfloor(\text{Trans}(\delta, s, d', s') \supset P(\delta', s')) \supset P(\delta, s)\rfloor$$

**InevTerminates(\delta, s)** means that program \(\delta\) executed starting in situation \(s\) inevitably terminates, i.e., all its executions eventually terminate.\(^4\)

Then we can ensure that for any HL agent action that is possibly executable at the low level, all executions terminate (i.e., they never block or diverge) by requiring the following:

**Constraint 13. (HL actions Inevitably Terminate)** For every high-level system action sequence \(\tilde{a}\) and every high-level action \(A\), we have that:

$$D_l \cup C \models \forall s. (\text{Do}(m_a(\tilde{a}), S_0, s) \supset \exists \tilde{x}. (\exists s'. D_{ao}(\{m_a(A(\tilde{x}), s, s') \supset InevTerminates(m_a(A(\tilde{x})), s))\))$$

**Proposition 14. NDBAT \(D_l^t\) and mapping \(m^t\) satisfy Constraint 13.**

If we impose Constraint 13, then the agent that executes the program that implements the HL action can be a dumb executor. But this may seem too restrictive. An alternative is to impose the weaker requirement that for any HL agent action that is possibly executable at the LL, the agent has a strategy to execute it to termination no matter how the environment reacts (even if not controlling its outcome). Formally:

**Constraint 15. (Agt Can Always Execute HL actions)** For every high-level system action sequence \(\tilde{a}\) and every high-level action \(A\), we have that:

$$D_l \cup C \models \forall s. (\text{Do}(m_a(\tilde{a}), S_0, s) \supset \exists \tilde{x}. (\exists s'. D_{ao}(\{m_a(A(\tilde{x}), s, s') \supset \text{AgtCanForce}(m_a(A(\tilde{x})), s))\))$$

Note that if we only impose Constraint 15, then there is no guarantee that when a HL system action \(A(\tilde{x}, e)\) is possibly executable, the environment can actually ensure that reaction \(e\) occurs when the agent executes \(m_a(A(\tilde{x}))\). We can define an additional constraint that ensures this; see [Banihashemi et al., 2023] for details.

**Planning for achievement goals.** Returning to planning, we can now show that if Constraint 15 holds (i.e., the agent always knows how to execute high-level primitive actions at the low level, even if not controlling their outcome) and the agent has a strong plan to achieve a goal at the high level, then there exists a refinement of the high-level plan that is a strong plan to achieve the refinement of the goal at the low level:

**Theorem 16. If \(M_h \sim_m M_l\) and Constraint 15 holds, then for any high-level system action sequence \(\tilde{a}\) and any high-level situation-suppressed formula \(\phi\), we have that:**

$$\text{if } M_h, v \models \text{Executable(do(\tilde{a}, S_0))} \land \text{AgtCanForce(\phi, do(\tilde{a}, S_0))}$$

then $$M_l, v \models \exists s. \text{Do}(m_a(\tilde{a}), S_0, s) \land \forall s. \text{Do}(m_a(\tilde{a}), S_0, s) \supset \text{AgtCanForce}(m_f(\phi), s)$$

**Proof sketch.** The proof is by induction on the length of the HL strategy \(f_h\) at \(do(\tilde{a}, S_0)\), where the length of a strategy is the length of its longest branch, i.e.,

$$\text{length}(f, s) = 0$$

if \(f(s) = \text{stop}\) and \(\text{length}(f, s) = \max_{D_{ao}(f(s), s, s')} (\text{length}(f, s'))\) otherwise.

Assume the antecedent. It follows that there is a strategy \(f_h\) such that \(\text{AgtCanForceBy}(f_h, s_h, f_h)\), where \(s_h = do(\tilde{a}, S_0)\). By proof of Theorem 11 we have that at the LL \(\exists s_l. \text{Do}(m_a(\tilde{a}), S_0, s_l)\) and for all such \(s_l\) we have that \(s_h \sim_m \text{AgtCanForce}(m_f(\phi), s_l, s_l)\). Take an arbitrary such \(s_l\). We show that there exists a LLI strategy \(f_l\) such that \(\text{AgtCanForceBy}(m_f(\phi), s_l, f_l)\).

Base case, when \(\text{length}(f_h, s_h) = 0\), and \(f_h(s_h) = \ldots\)
At the HL, by the definition of $AgtCanForceBy$.

Inductive case: Assume the result holds for any HL strategy $f_h$ of length at most $k$. We show that it must also hold for a HL strategy $f_h$ of length $k+1$.

At the HL, by the definition of $AgtCanForceBy$, we have that there exists an high-level action $A(\bar{t})$ such that $f_h(s_h) = A(\bar{t}) \land \exists s'_h. DoAgg(A(\bar{t}), s_h, s'_h) \land \forall s'_h. DoAgg(A(\bar{t}), s_h, s'_h) \supset \text{AgtCanForceBy}(\phi, s'_h, f_h)$.

By Constraint 15, there exists a LL strategy $g_l$ such that $AgtCanForceBy(m_a(A(\bar{t}), s_l, g_l))$. For any $s'_l$ such that $DoAgg(g_l(s), s'_l)$, it follows by Constraint 2 that there exists $s'_h$ such that $s'_h \sim m_s(m_l(s'), s'_l)$. The length of $f_h$ at $s'_h$ is at most $k$.

Thus by the induction hypothesis, there exists a LL strategy $f'_l$ such that $AgtCanForceBy(m_f(\phi), s'_l, f'_l)$. It follows that $AgtCanForceBy(m_f(\phi), s_l, f_l)$ for $f_l = f'_l \circ g_l$.

Example cont. As strategies use the stop action, we first define a stopHL action at the high level which is mapped to the stopLL action at the concrete level; stopHL always terminates with the SuccessHS reaction. By Proposition 14, Constraint 13 (and thus also the weaker Constraint 15) is satisfied for our example domain.

We can show that the agent has a strong plan to achieve the goal of being at location 13 at the high level (dfh abbreviates driveAndTryFix):

$D_{hl}^t \models \text{AgtCanForceBy}(At_{HL}(13), S_0, f_h)$

for

$$f_h(s) :=\begin{cases} \text{stopHL} & \text{if } At_{HL}(13, s) \\ df(\sigma, \delta) & \text{if } At_{HL}(\sigma, \delta) \land \sigma \neq 13 \land \text{RoadHL}(\delta, \sigma, s) \\ \text{waitHL} & \text{otherwise} \end{cases}$$

The agent’s high-level strategy $f_h$ is to stop if she is already at location 13, otherwise to drive to a location which has not been visited previously and has a spare tire and fix a flat if one occurs; in all other cases she waits.

By Th. 16 there exists a strategy at the low level that is refinement of the high-level strategy; this strategy is the same as strategy $f_l$ in Section 2. It is easy to show that $f_l$ is a refinement of the high-level strategy $f_h$.

Planning for temporally extended goals/tasks. We can also show a similar result to the previous theorem for high-level programs/tasks/temporally extended goals: if Constraint 15 holds and the agent has a strategy to successfully execute an agent program (without concurrency) at the high level, then the agent also has a strategy to successfully execute some refinement of it at the low level:

Theorem 17. If $M_h \sim_m M_l$ and Constraint 15 holds, then for any high-level system action sequence $\alpha$ and any SD ConGolog high-level agent program $\delta$ without the concurrent composition construct, we have that:

if $M_h, v \models \text{Executable}(do(\bar{a}, S_0)) \land \text{AgtCanForceBy}(\delta, do(\bar{a}, S_0))$

then $M_l, v \models \exists s. Do(m_a(\bar{a}), S_0, s) \land \forall s. Do(m_a(\bar{a}), S_0, s) \supset \text{AgtCanForceBy}(m_a(\delta), s)$

Here, we apply the mapping to a high-level agent program $\delta$ without concurrency to produce a low level agent program $m_a(\delta)$; this can be defined in the obvious way, using $m_a$ to map atomic actions and $m_f$ for tests as usual, and by mapping the components and composing the result for other constructs, e.g., $m_a(\alpha_1 \alpha_2) = m_a(\alpha_1)m_a(\alpha_2)$ (we discuss concurrency in [Banihashemi et al., 2023]).

Proof sketch. The proof is by induction on the length of the high-level strategy $f_h$ for executing $\delta$ at $do(\bar{a}, S_0)$ and is similar to that of the above theorem. Note that since $\delta$ is SD, does not involve concurrency, and the high-level atomic actions are all mapped to SD programs, $m_a(\delta)$ must also be SD (we execute a high-level action using the appropriate strategy essentially as a procedure call and the high-level program continues after the call has completed).

Example cont. Consider the high-level program $\delta_{HL}^H(l) = (\pi_0, d. \text{driveAndTryFix}(\sigma, \delta))$: at $HL(l)$? which goes to location $l$ by repeatedly picking adjacent locations and doing $\text{driveAndTryFix}(\sigma, \delta)$ until the agent is at $l$. We can show that the agent has a strategy to successfully execute this program to get to location 13:

$D_{hl}^t \models \text{AgtCanForceBy}(\delta_{HL}^H(13), S_0, f_h)$

Here, the high-level strategy $f_h$ to execute the program is same as the one we saw above. It can also be shown that the low-level strategy $f_l$ seen earlier is a refinement of $f_h$ and can be used to to execute $m_h(\delta_{HL}^H(13))$. To further illustrate the use of high-level programs, notice that we could give the agent the program $\delta_{HL}^H(31); \delta_{HL}^H(13)$ to execute. This is a more complex task as she must first go to location 31 and then to 13; but it is easier to to find a strategy to do it.

In [Banihashemi et al., 2023], we show how we can handle more complex domains where reactions represent (sets of) exogenous actions.

6 Monitoring and Explanation

Being able to monitor what a low-level agent is doing and describe it in abstract terms (e.g., to a client or manager) is particularly important in the context of Explainable AI (BDL17) show how a sound abstraction wrt a mapping $m$ can be used to do this when certain conditions hold. They define an inverse mapping $m^{-1}$ that takes a sequence of low-level actions $\bar{\alpha}$ and returns the sequence of high-level actions $\alpha$ that it refines in a given model. In [Banihashemi et al., 2023], we show that we can easily adapt this approach for NDBAT abstractions.

Example cont. Consider a variant of our running example where, if the agent has a flat and there is no spare tire in a location (i.e., at locations 11, 12 and 13), she can either buy one, or she can use a roadside assistance service to get one provided the location is within the covered area. We can have HL actions for these with the mapping:

$m_{hl}^t(\text{buyAndFix}(l, r_0)) = \pi_r.\text{order}(l, r); \pi_r.\text{pay}(l, r); \pi_r.\text{fixFlatTire}(l, r); r_h = \text{SuccessBuy}?

m_{hl}^t(\text{serviceAndFix}(l, r_0)) = \pi_r.\text{callServiceLL}(l, r); \pi_r.\text{fixFlatTire}(l, r); r_h = \text{SuccessServ}?
Suppose that at the LL, the agent knows that both locations 12 and 13 fall under the roadside service coverage, while at the HL, the agent only knows this for 13 (we have imperfect information). We can define NDBATs for this and show $D_{h}^{t+}$ is a sound abstraction of $D_{h}^{t+}$ [Banihashemi et al., 2023].

Now suppose that at the LL, the sequence of system actions $\alpha = drive(11, 21, NoFlatTire); drive(21, 12, FlatTire)$ has occurred. In location 12, no spare tire is available, and if the agent wants to move on with her journey, she needs to either call for service or order a new tire. The inverse mapping tells us that HL system action sequence $\alpha = driveAndTryFix(11, 21, DvrNoFlat); driveAndTryFix(21, 12, DvrFlat)$ has occurred. Since $D_{h}^{t+} \models \text{At}_{HL}(12, do(\alpha, S_{0}))$, we can conclude at the HL that the agent is now at location 12. Since $D_{h}^{t+} \cup \{\text{Poss}(\text{serviceAndFix}(12, \text{Succ}\_\text{Serv}), do(\alpha, S_{0}))\}$ is satisfiable, we can also conclude that $\text{serviceAndFix}(12, \text{Succ}\_\text{Serv})$ might occur next, and similarly for $\text{buyAndFix}(12, \text{Succ}\_\text{Buy})$; in fact we know at the HL that the latter is executable: $D_{h}^{t+} \models \text{Poss}(\text{buyAndFix}(12, \text{Succ}\_\text{Buy}), do(\alpha, S_{0}))$.

7 Discussion

There has been much previous work on fully/partially observable nondeterministic planning, e.g., [Ghallab et al., 2004; Cimatti et al., 2003; Muise et al., 2014], strategy logics such as ATL [Alur et al., 2002], and strategic reasoning [Xiong and Liu, 2016b; Xiong and Liu, 2016a]. But such work assumes that the domain is represented at a single level of abstraction.

In planning, approaches such as hierarchical task networks (HTNs) [Erol et al., 1994] support abstract tasks that facilitate search. [Kuter et al., 2009] propose an algorithm which combines HTN-based search-control strategies with Binary Decision Diagram-based state representations for planning in nondeterministic domains. [Chen and Bercher, 2021] extend HTN planning with nondeterministic primitive tasks. [Bonet et al., 2017] investigate nondeterministic abstractions for generalized planning, where one looks for a (typically iterative) plan that solves a whole class of related planning problems. [Cui et al., 2021] apply the abstraction framework of (BDL17) to generalized planning. They use Golog programs to represent nondeterministic actions, and our results should make it possible to simplify their framework.

[Banihashemi et al., 2018] generalized the approach of (BDL17) to deal with agents that may acquire new knowledge during execution, and formalized (piecewise) hierarchical refinements of an agent’s ability (i.e., strategy) to achieve a goal. However, this work does not consider nondeterministic actions and is based on a more complex framework involving online executions.

Abstraction is important for efficient reasoning and explainability. This paper presents foundational results, i.e., a generic framework for abstraction in nondeterministic domains, which can be used for many reasoning tasks, such as planning, execution monitoring, and explanation. One potential practical application area is smart manufacturing. For instance, [De Giacomo et al., 2022] present a “manufacturing as a service” framework where facilities (made up of concrete resources) bid to produce products, given an abstract product recipe, which is based on an abstract information model in the cloud. Our work shows how one can develop an abstract domain model where planning for arbitrary high-level goals is easier and where we have guarantees that a high-level plan can be refined into a low-level one. It also allows one to monitor the low-level system by generating high-level descriptions of low-level system executions from which one can reason at the high level about what might happen next. Note that often manufacturing processes depend on the data and objects (parts) they produce and consume: to formalize this aspect, the situation calculus which provides a first-order representation of the state of the processes can indeed be used. Another potential practical application area is business process management, to support execution monitoring, explanation, and failure handling, see e.g., [Marrella et al., 2017].

As discussed in Section 3, (BDL17) identifies a set of properties that are necessary and sufficient to have a sound complete abstraction wrt a mapping, which can be used to verify that one has such an abstraction. This means that in the finite domain/propositional case, verifying that one has a sound abstraction is decidable and one can use theorem proving techniques to do it. Moreover, when one has complete information and a single model, one can use model-checking techniques. Bounded basic action theories [De Giacomo et al., 2016] constitute an important class of infinite domain theories where verification is decidable. In the general infinite-domain case, there are sound but incomplete reasoning methods that can be used [De Giacomo et al., 2010].

Automatically synthesizing sound complete abstractions is another interesting problem. [Luo et al., 2020] show that one can use the well-explored notion of forgetting (of low-level fluent and action symbols) to automatically obtain a sound and complete high-level abstraction of a low-level BAT for a given mapping under certain conditions. They also show that such an abstraction is computable in the propositional case. However in general, there may be many different abstractions of a low-level theory, each of which may be useful for a different purpose. So defining an abstract language and mapping for a domain is not a trivial problem. Some human intervention is likely to be required, e.g., the modeler might specify the goals of the abstraction, or which details can be considered unimportant.

In this paper, we have not considered fairness constraints and strong cyclic plans [D’Ippolito et al., 2018; Aminof et al., 2020]; this is a topic for future work. Another interesting direction for future research is to consider a notion of multi-tier planning [Ciolek et al., 2020], where planning may be done using a simple model without unlikely contingencies, but where one can fall back to a more complex model when such contingencies do occur or more robustness is required. Related to this is the notion of “best effort strategies” (in presence of multiple [contradictory] assumptions about the environment) as proposed by [Aminof et al., 2021]. These are agent plans which, for each of the environment specifications individually, realize the agent’s goal against a maximal set of environments satisfying that specification. We are also interested in extending the current framework to deal with partial observability and sensing.
Acknowledgments

Work supported by the ERC Advanced Grant WhiteMech (No. 834228), by the EU ICT-48 2020 project TAILOR (No. 952215), as well as by the National Science and Engineering Research Council of Canada.

References


[De Giacomo et al., 2010] Giuseppe De Giacomo, Yves Lespérance, and Adrian R. Pearce. Situation calculus based programs for representing and reasoning about game structures. In Principles of Knowledge Representation and...


