Embracing Causality in Specifying the Indeterminate Effects of Actions

Fangzhen Lin
Department of Computer Science
University of Toronto
Toronto, Canada M5S 3H5
email: fl@ai.toronto.edu

Abstract

This paper makes the following two contributions to formal theories of actions: Showing that a causal minimization framework can be used effectively to specify the effects of indeterminate actions; and showing that for certain classes of such actions, regression, an effective computational mechanism, can be used to reason about them.

Logical Preliminaries

We shall investigate the problem in the framework of the situation calculus [8]. Our version of it employs a many sorted second-order language. We assume the following sorts: situation for situations, action for actions, fluent for propositional fluents, truth-value for truth values true and false, and object for everything else.

We use the following domain independent predicates and functions:

- The binary function do - for any action a and any situation s, \( do(a, s) \) is the situation resulting from performing a in s.
- The binary predicate H - for any propositional fluent p and any situation s, \( H(p, s) \) is true if p holds in s.
- The binary predicate Poss - for any action a and any situation s, \( Poss(a, s) \) is true if a is possible (executable) in s.
- The ternary predicate Caused - for any fluent p, any truth value v, and any situation s, \( Caused(p, v, s) \) is true if the fluent p is caused to have the truth value v in the situation s. For instance, \( Caused(loaded, true, do(\text{load}, s)) \) means that the action load causes loaded to be true in the resulting situation.

We shall make use some additional special predicates and functions, and will introduce them when they are needed.

Our contributions in this paper are two folds. We first show that the causal minimization framework of (Lin [5]) can be used effectively to specify the effects of indeterminate actions. We then show that for certain classes of such actions, regression, an effective computational mechanism, can be used to reason about them.

Introduction

Much recent work on theories of actions has concentrated on primitive, determinate actions. In this paper, we pose ourselves the problem of specifying directly the effects of indeterminate actions, like we do for the primitive, determinate ones.

There are several reasons why we think this is an important problem. First of all, there are actions whose effects, when described at a natural level, are indeterminate. Secondly, one can argue that there is no absolute defining line between determine and indeterminate actions. The differences have a lot to do with the levels of descriptions. The effects of an action may be determinate at one level of description, but indeterminate at another. So a theory that treats determine and indeterminate actions in fundamentally different ways will have difficulties coping with language changes. Finally, even if all the primitive actions have determinate effects, there are still needs for specifying directly the effects of complex actions that are often indeterminate. For instance, these specifications may be part of the inputs to a program synthesizer.

Our contributions in this paper are two folds. We first show that the causal minimization framework of (Lin [5]) can be used effectively to specify the effects of indeterminate actions. We then show that for certain classes of such actions, regression, an effective computational mechanism, can be used to reason about them.

Our contributions in this paper are two folds. We first show that the causal minimization framework of (Lin [5]) can be used effectively to specify the effects of indeterminate actions. We then show that for certain classes of such actions, regression, an effective computational mechanism, can be used to reason about them.

1 For the purpose of this paper, the phrases “the effects of indeterminate actions” and “the indeterminate effects of actions” are considered to be synonyms.

2 We use the convention that in displayed formulas, free variables are implicitly universally quantified from the outside.
The unique names assumptions for fluent and action names (we assume there are only finitely many of them). Specifically, if \( \{F_1, \ldots, F_n\} \) is the set of the fluent names, then we have:

\[
F_i(\vec{x}) \neq F_j(\vec{y}), \text{ i and j are different},
\]

\[
F_i(\vec{x}) = F_i(\vec{y}) \supset \vec{x} = \vec{y}.
\]

Similarly for action names. In the following, we shall denote this set of unique names axioms by \( D_{una} \).

- The set \( \Sigma \) of foundational axioms in [6] for the discrete situation calculus. These axioms characterize the structure of the space of situations. For the purpose of this paper, it is enough to mention that they include the following unique names axioms for situations:

\[
s \neq do(a, s),
\]

\[
do(a, s) = do(a', s') \supset (a = a' \land s = s').
\]

In the rest of this paper, we shall frequently make use of the following shorthand notation: If \( F \) is a fluent name of arity \( n \rightarrow \text{fluent} \), then we define the expression \( F(t_1, \ldots, t_n, t_s) \) to be a shorthand for the formula \( H(F(t_1, \ldots, t_n), t_s) \), where \( t_1, \ldots, t_n \) are terms of sort \( \text{object} \), and \( t_s \) a term of sort \( \text{situation} \). So if \( \text{white} \) is a fluent, then \( \text{white}(s) \) is a shorthand for \( H(\text{white}, s) \).

**Minimizing Causation**

The approach of (Lin [5]) to specifying the effects of actions can be summarized as follows:

1. Formalize the causal laws and constraints of the domain by a set \( T \) of axioms.
2. Circumscribe (minimize) \( \text{Caused} \) in \( T \cup \Sigma \cup D_{una} \cup \{1, 2, 3\} \) with all other predicates fixed.
3. The resulting theory, \( T' \), together with the following generic frame axiom: Unless caused otherwise, a fluent's truth value will persist:

\[
\exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exists v \exist...
a = pickup \land [(p = holding \land v = \text{true}) \lor \\
(p = \text{white} \land v = \text{false}) \lor \\
(p = \text{black} \land v = \text{false})] \\
\lor \\
a = \text{drop} \land [(p = \text{holding} \land v = \text{false}) \lor \\
p = \text{white} \lor p = \text{black}].

From this and the generic frame axiom (4), we can deduce the following successor state axiom (Reiter [lo]) for holding:

\[ \text{Poss}(a, s) \supset \text{holding}(do(a, s)) \equiv \\
a = \text{pickup} \lor (\text{holding}(s) \land a \neq \text{drop}). \]

We don’t get successor state axioms for white and black. But we have the following explanation closure axioms:

\[ \text{Poss}(a, s) \supset -[\text{white}(s) \equiv \text{white}(do(a, s))] \supset \\
(a = \text{pickup} \lor a = \text{drop}), \]
\[ \text{Poss}(a, s) \land -[\text{black}(s) \equiv \text{black}(do(a, s))] \supset \\
(a = \text{pickup} \lor a = \text{drop}). \]

These axioms, together with the effect axioms, yield the following disjunction of successor state axioms:

\[ \text{Poss}(a, s) \supset \\
\{[\text{white}(do(a, s)) \equiv \\
(a = \text{drop} \lor (\text{white}(s) \land a \neq \text{pickup}))] \land \\
[\text{black}(do(a, s)) \equiv \\
(\text{black}(s) \land a \neq \text{pickup} \land a \neq \text{drop})]) \lor \\
[\text{white}(do(a, s)) \equiv \\
(\text{white}(s) \land a \neq \text{pickup} \land a \neq \text{drop})] \land \\
[\text{black}(do(a, s)) \equiv \\
(a = \text{drop} \lor (\text{black}(s) \land a \neq \text{pickup}))]) \lor \\
[\text{white}(do(a, s)) \equiv \\
(a = \text{drop} \lor (\text{white}(s) \land a \neq \text{pickup}))] \land \\
[\text{black}(do(a, s)) \equiv \\
(a = \text{drop} \lor (\text{black}(s) \land a \neq \text{pickup}))).\]

Notice the correspondences between the three cases and those in drop’s effect axiom for white and black.

**Classes of Indeterminate Actions**

The indeterminate effects of drop are inclusive in that the pin may land on a white square, a black square, or both. To see how such inclusive indeterminate effects can be represented succinctly, notice first that under the two general axioms (1) and (2) about \textit{Caused}, the effect axiom (9) is equivalent to the following three axioms:

\[ \text{Poss}(\text{drop}, s) \supset \mathbf{\text{Caused}}(\text{white}, \text{true}, do(\text{drop}, s)) \lor \\
\mathbf{\text{Caused}}(\text{black}, \text{true}, do(\text{drop}, s))), \]
\[ \text{Poss}(\text{drop}, s) \supset \{\mathbf{\text{Caused}}(\text{white}, \text{true}, do(\text{drop}, s)) \lor \\
\mathbf{\text{Caused}}(\text{white}, \text{false}, do(\text{drop}, s))) \lor \\
\text{Poss}(\text{drop}, s) \supset \{\mathbf{\text{Caused}}(\text{black}, \text{true}, do(\text{drop}, s)) \lor \\
\mathbf{\text{Caused}}(\text{black}, \text{false}, do(\text{drop}, s))). \]

Notice that under the domain closure and unique names axiom (3) for truth values, the last axiom is equivalent to

\[ \text{Poss}(\text{drop}, s) \supset (\exists n)\mathbf{\text{Caused}}(\text{black}, n, do(\text{drop}, s)). \]

This axiom is like the \textit{releases} propositions in the action description language of [3]. Notice here the necessity of something like the predicate \textit{Caused}. The corresponding sentence in terms of \textit{H}:

\[ \text{Poss}(\text{drop}, s) \supset (\text{H}(\text{black}, do(\text{drop}, s)) \lor \\
\neg \text{H}(\text{black}, do(\text{drop}, s))) \]

is just a tautology.

In general, if the action \(a\) has inclusive indeterminate effects on the fluent terms \(P_1, \ldots, P_n\), i.e. causes at least one of them to be true and the rest of them to be false, under the context \(\gamma\), then we have the following causal laws:

\[ \text{Poss}(a, s) \land \gamma \supset \{\mathbf{\text{Caused}}(P_i, \text{true}, do(a, s)) \lor \cdots \lor \\
\mathbf{\text{Caused}}(P_n, \text{true}, do(a, s))\}, \]
\[ \text{Poss}(a, s) \land \gamma \supset \{\mathbf{\text{Caused}}(P_i, \text{false}, do(a, s)) \lor \\
\text{Caused}(P_i, \text{false}, do(a, s)))\}, \]

where \(1 \leq i \leq n\).

The number of indeterminate effects need not be finite. If, under the context \(\gamma\), the action \(a\) has the inclusive indeterminate effects on \(F(x)\) for those \(x\) that satisfies \(\varphi\), then we have the following causal laws:

\[ \text{Poss}(a, s) \land \gamma \land (\exists x)\varphi(x) \supset \\
(\exists x)[\varphi(x) \land \text{caused}(F(x), \text{true}, do(a, s))], \]
\[ \text{Poss}(a, s) \land \gamma \land (\forall x)[\varphi(x) \lor \\
\text{Caused}(F(x), \text{false}, do(\alpha, s))]\].

For instance, playing loud rock and roll music will make some of the nearby people (including the person who plays it) happy, and the rest of them unhappy: let \(\gamma\) be true, \(\varphi(x)\) be \text{nearby}(x, s), and \(F(x)\) be \text{happy}(x).

Contrast to the inclusive indeterminate effects, we have the exclusive ones. For instance, flipping a coin causes exactly one of \{head, tail\} to be true. Generally, if the action \(a\) has exclusive indeterminate effects on the fluent terms \(P_1, \ldots, P_n\), i.e. causes exactly one of them to be true and the rest of them to be false, under the context \(\gamma\), then we have the following causal laws:

\[ \text{Poss}(a, s) \land \gamma \supset \{\mathbf{\text{Caused}}(P_i, \text{true}, do(a, s)) \lor \cdots \lor \\
\mathbf{\text{Caused}}(P_n, \text{true}, do(a, s))\}, \]
\[ \text{Poss}(a, s) \land \gamma \supset \{\mathbf{\text{Caused}}(P_i, \text{false}, do(a, s)) \lor \\
\text{Caused}(P_i, \text{false}, do(a, s)))\}, \]

where \(1 \leq i \leq n\), and \(\lor\) is the exclusive or operator:

\[ \varphi_1 \lor \cdots \lor \varphi_k \equiv (\varphi_1 \lor \cdots \lor \varphi_k) \land \bigwedge_{1 \leq i \leq k} (\varphi_i \land \neg \varphi_i). \]
Again, the number of indeterminate effects need not be finite.

There are, of course, actions with indeterminate effects that are neither inclusive or exclusive. In general, if the number of the indeterminate effects of an action $A(x)$ is finite, then its effect axioms can be written of the following forms:

$$
\text{Poss}(A(x), s) \supset (\forall p, s)[\varphi(x, p, v, s) \supset \text{Caused}(p, v, do(A(x), s))],
$$

(10)

$$
\text{Poss}(A(x), s) \supset \{ (\forall p, v)[\varphi_1(x, p, v, s) \supset \text{Caused}(p, v, do(A(x), s))] \} \lor \cdots \lor (\forall p, v)[\varphi_n(x, p, v, s) \supset \text{Caused}(p, v, do(A(x), s))],
$$

(11)

where $\varphi$ and $\varphi_i$ are formulas that do not quantify over situation variables, and do not mention any other situation dependent atomic formulas except those of the form $H(t, s)$.

For instance, the two effect axioms about $\text{drop}$ can be rewritten as:

$$
\text{Poss}(\text{drop}, s) \supset (\forall p, s)[p = \text{holding} \land v = \text{false} \supset \text{Caused}(p, v, do(d\text{rop}, s))],
$$

(12)

$$
(\forall p, v)[p = \text{white} \land v = \text{true} \lor \text{black} \land v = \text{false} \supset \text{Caused}(p, v, do(d\text{rop}, s))] \lor \cdots \lor (\forall p, v)[p = \text{white} \land v = \text{true} \lor \text{black} \land v = \text{false} \supset \text{Caused}(p, v, do(d\text{rop}, s))].
$$

(13)

Notice that (10) and (11) can be combined into a single axiom of the latter form. But as we shall see later, it is beneficial to have a separate axiom for indeterminate effects.

### Computing Successor State Axioms

We now consider how to reason with the theories of the actions whose effects are specified by axioms of the forms (10) and (11).

Let $T_{ea}$ be a given set of the effect axioms of the forms (10) and (11). Then the conjunction of the sentences in $T_{ea}$ is separable (Lifschitz [4]) w.r.t. $\text{Caused}$. Therefore, according to a result in [4], $\text{Circum}(T_{ea}, \text{Caused})$, the circumscription of $\text{Caused}$ in $T_{ea}$, is computable by a first-order sentence. In general, this sentence, together with $D_{una}$, will yield a disjunction of successor state axioms, which is often large and cumbersome to reason with. In particular, it is not clear how to compute regression, a computationally effective mechanism for tasks such as planning and temporal projection [11, 9, 10], w.r.t. such disjunctions.

### A Transformation

To overcome this, we introduce a new ternary predicate $\text{Case}$ of the arity $\text{object} \times \text{action} \times \text{situation}$, and a distinguished constant 0 and a unary function $\text{succ}$ over sort $\text{object}$. We use the convention that if a natural number $n$ occurs as an object term in a formula, then it is considered to be a shorthand for the term obtained from 0 by applying $n$ times the function $\text{succ}$. For instance, in $\text{Case}(2, a, s)$, the number 2 is a shorthand for $\text{succ}(\text{succ}(0))$.

For now we shall consider $\text{Case}$ to be an auxiliary predicate introduced for computational purposes. Later, we shall consider some possible interpretations of this predicate.

Using $\text{Case}$, we transform the indeterminate effect axiom (11) into the following sentences that have the form of a determinate effect axiom:

$$
\text{Poss}(A(x), s) \land \text{Case}(1, A(x), s) \supset (\forall p, v)[\varphi_1(x, p, v, s) \supset \text{Caused}(p, v, do(A(x), s))],
$$

(14)

$$
\vdots
$$

$$
\text{Poss}(A(x), s) \land \text{Case}(n, A(x), s) \supset (\forall p, v)[\varphi_n(x, p, v, s) \supset \text{Caused}(p, v, do(A(x), s))],
$$

(15)

together with the following constraints on $\text{Case}$:

$$
\text{Case}(1, A(x), s) \lor \cdots \lor \text{Case}(n, A(x), s),
$$

(16)

$$
\{ (\forall p, v)[\varphi_1(x, p, v, s) \lor \varphi_2(x, p, v, s) \lor \cdots \lor \varphi_n(x, p, v, s)] \} \land
$$

$$
\neg(\forall p, v)[\varphi_1(x, p, v, s) \lor \cdots \lor \varphi_n(x, p, v, s)]
$$

(17)

for any $1 \leq i \neq j \leq n$.

Notice the exclusive or in (16). This is because when $\text{Case}$ is circumscribed, the logical or in (11) will become exclusive. The intuitive meaning of (17) is that if the extension of $(\lambda p, v)\varphi_i$ $\lor \varphi$ strictly contains that of $(\lambda p, v)\varphi_i$, then the conjunct corresponding to $\text{Case}(i, A(x), s)$ cannot be minimal, so $\text{Case}(i, A(x), s)$ must not hold. These constraints are best understood in light of the following Theorem 1 which will establish the correctness of the above transformation.

Notice also that this transformation applies only to the indeterminate effect axiom (11). This is why it is beneficial to put as much information as possible into (10).

In the following, we shall denote by $T_{ea}$ the set of axioms obtained from $T_{ea}$ by replacing every indeterminate effect axiom in it of the form (11) by the axioms (14) - (15). We shall denote by $D_{case}$ the set of constraints (16) and (17). Notice that this set is also dependent on $T_{ea}$.

Given two theories $T_1$ and $T_2$ such that $T_1$'s language is $T_2$'s augmented by a new predicate $P$, we say
that these two theories are equivalent with respect to $T_2$'s language if $T_1$ is a conservative extension of $T_2$: a structure is a model of $T_2$ if it can be extended into a model of $T_1$. As it turns out, this is the same as saying that $T_2$ is the result of forgetting $P$ in $T_1$ according to (Lin and Reiter [7]), and according to a result there, when $T_1$ is finite, this is the same as saying that $T_2$ is logically equivalent to the sentence $(\exists P) \land T_1$, where $\land T_1$ is the conjunction of the sentences in $T_1$.

We have:

**Theorem 1** Under the unique names axioms $D_{una}$, the result of circumscribing Caused in $T_{ea}' \cup D_{case}$ is a conservative extension of the result of circumscribing Caused in $T_{ea}$:

$$D_{una} \models \text{Circum}(T_{ea}, \text{Caused}) \equiv (\exists \text{Case})\text{Circum}(T_{ea}' \cup D_{case}, \text{Caused}).$$

**Corollary 1.1** Under the unique names assumptions, for any formula $\varphi$ that does not mention Case, $D_{una} \cup \text{Circum}(T_{ea}, \text{Caused}) \models \varphi$ iff $D_{una} \cup \text{Circum}(T_{ea}' \cup D_{case}, \text{Caused}) \models \varphi$.

### Computing Successor State Axioms

Having established the correctness of the above transformation, we now proceed to show how to generate successor state axioms from the resulting axioms.

Notice first that the sentence (10) can be rewritten into an axiom of the following form:

$$\text{Poss}(A(Z), s) \land \varphi \land \text{Poss}(?, p, s) \supset \text{Caused}(p, v, do(A(?), s)).$$

Similarly, we can do the same for axioms of the form (14) (15). Now from these axioms in $T_{ea}'$, we can generate, for each fluent $F$, two axioms of the following forms:

$$\text{Poss}(a, s) \land \gamma_F^+(\overline{x}, a, s) \supset \text{Caused}(F(\overline{x}), \text{true}, do(a, s)),$$

(18)

$$\text{Poss}(a, s) \land \gamma_F^-(\overline{x}, a, s) \supset \text{Caused}(F(\overline{x}), \text{false}, do(a, s)),$$

(19)

where $\gamma_F^+$ and $\gamma_F^-$ do not quantify over situation variables, and the only situation dependent atomic formulas in them are either of the form $H(t, s)$ or of the form $\text{Case}(t_1, t_2, s)$.

Given these two effect axioms, we generate the following successor state axiom for $F$:

$$\text{Poss}(a, s) \supset F(\overline{x}, do(a, s)) \equiv \gamma_F^+(\overline{x}, a, s) \lor (F(\overline{x}, s) \land \neg \gamma_F^-(\overline{x}, a, s)).$$

(20)

Now let $D_{ss}$ be the set of successor state axioms, one for each fluent, so generated. Our claim is that, under some reasonable conditions, $D_{ss}$ captures all the information about the truth values of the fluents in $\text{Circum}(T_{ea}', \text{Caused}) \cup \{1, 2, 4\}$. More precisely, we have:

**Theorem 2** Under the assumption that the following consistency condition [10] is satisfied for each fluent $F$:

$$D_{una} \cup D_{ap} \cup D_{case} \models (\forall \overline{x}, a, s).\text{Poss}(a, s) \supset \neg (\gamma_F^+(\overline{x}, a, s) \land \gamma_F^-(\overline{x}, a, s)),$$

the theory

$$\Sigma \cup D_{una} \cup D_{ap} \cup \{\text{Circum}(T_{ea}', \text{Caused})\} \cup D_{case} \cup \{1, 2, 3, 4\}$$

is a conservative extension of the theory

$$\Sigma \cup D_{una} \cup D_{ap} \cup D_{ss} \cup D_{case} \cup \{3\}.$$

**Corollary 2.1** Under the assumptions in the theorem, for any formula $\varphi$ that does not mention Caused, $\Sigma \cup D_{una} \cup D_{ap} \cup \{\text{Circum}(T_{ea}', \text{Caused})\} \cup D_{case} \cup \{1, 2, 3, 4\} \models \varphi$ iff $\Sigma \cup D_{una} \cup D_{ap} \cup D_{ss} \cup D_{case} \cup \{3\} \models \varphi$.

**Corollary 2.2** Under the assumptions in the theorem, for any formula $\varphi$ that does not mention Caused and Case, $\Sigma \cup D_{una} \cup D_{ap} \cup \{\text{Circum}(T_{ea}', \text{Caused})\} \cup \{1, 2, 3, 4\} \models \varphi$ iff $\Sigma \cup D_{una} \cup D_{ap} \cup D_{ss} \cup D_{case} \cup \{3\} \models \varphi$.

**Proof:** Apply Theorem 1 and Theorem 2.

Theorem 2 informs us that if we are only concerned with the truth values of fluents, then the original effect axioms as well as the basic axioms about Caused can all be discarded. In particular, this is the case with the projection problem.

Technically, the consistency conditions are needed because without these conditions, the successor state axiom (20) may not entail the formula

$$\text{Poss}(a, s) \supset \gamma_F^+(\overline{x}, a, s) \supset \neg F(\overline{x}, do(a, s)),$$

which is a consequence of the effect axiom (19) and the two basic axioms (1) and (2) about causality.

**Example 1** Consider again our checkerboard example. We shall consider only the successor state axioms for white and black. The indeterminate effect axiom (13) of drop is translated into:

$$\text{Poss}(\text{drop}, s) \land \text{Case}(1, \text{drop}, s) \supset$$

$$\text{Caused(white, true, do(drop, s)) \land Caused(black, false, do(drop, s))},$$

$$\text{Poss}(\text{drop}, s) \land \text{Case}(2, \text{drop}, s) \supset$$

$$\text{Caused(white, false, do(drop, s)) \land Caused(black, true, do(drop, s))},$$

$$\text{Poss}(\text{drop}, s) \land \text{Case}(3, \text{drop}, s) \supset$$

$$\text{Caused(white, true, do(drop, s)) \land Caused(black, true, do(drop, s))}.$$
Together with the original determinate effect axioms, we have:

\[
\text{Poss}(a,s) \supset \{ \text{white}(do(a,s)) \equiv [a = \text{drop} \land (\text{Case}(1, \text{drop}, s) \lor \text{Case}(3, \text{drop}, s))] \lor \text{Caused}(\text{white}, \text{true}, \text{do}(a,s)),
\]

Thus we have the following successor state axiom for white:

\[
\text{Poss}(a,s) \supset \{ \text{white}(do(a,s)) \equiv [a = \text{drop} \land (\text{Case}(1, \text{drop}, s) \lor \text{Case}(3, \text{drop}, s))] \lor \text{white}(s) \land \\
\neg[a = \text{pickup} \lor (a = \text{drop} \land \text{Case}(2, \text{drop}, s))].
\]

A similar successor axiom can be obtained for black. It can be seen that the consistency conditions are satisfied for both white and black.

We shall not get into details regarding the accompanied constraints about Case, but note that for this example, all constraints of the form (17) are logical consequence of the unique names assumptions. So the following is the only nontrivial constraint about Case:

\[
\text{Case}(1, \text{drop}, s) \lor \text{Case}(2, \text{drop}, 2) \lor \text{Case}(3, \text{drop}, 3).
\]

### Regression and Some of Its Properties

Once we have a successor state axiom for each fluent, regression becomes syntactic substitutions [10]: for any formula \( \varphi(s) \) that does not quantify over situation, and action \( \alpha \), the regression of a formula \( \varphi(s) \) over \( \alpha \), written \( R(\varphi(s), \alpha) \), is the result of replacing in \( \varphi(s) \) every atomic formula of the form \( H(F(i), s) \) by \( \Phi_F(i, \alpha, s) \), where

\[
\text{Poss}(a,s) \supset [F(x, do(a,s)) \equiv \Phi_F(x, a, s)]
\]

is the successor state axiom for \( F \).

The following result is immediate:

**Lemma 2** Let \( D_{ss} \) be a set of successor state axioms, one for each fluent. We have:

\[
D_{ss} \models (\forall s).\text{Poss}(\alpha,s) \supset [\varphi(s) \equiv R(\varphi,\alpha)].
\]

In the rest of this section, we assume that we’re given an action theory of the form:

\[
D = \Sigma \cup D_{una} \cup D_{ap} \cup D_{ss} \cup D_{case} \cup D_{S0},
\]

where \( D_{case} \) is a set of Case constraints of the form (16) or of the form (17), and \( D_{S0} \) is a set of sentences that do not mention any other situation term except \( S_0 \), and do not mention Poss, Caused, and Case. The other components of \( D \) have the usual meaning.

Our main concern is the soundness and completeness of regression for doing temporal projection with respect to the initial database. Our first positive result is about Case independent temporal projections:

**Theorem 3** Let \( \varphi(s) \) be a formula that does not quantify over situation variables, does not mention any other situation term except \( s \), and does not mention Poss, Caused, and Case. Let \( \alpha \) be an action term. If, under \( D_{una}, R(\varphi,\alpha) \) is equivalent to a formula that does not mention Case, then

\[
D \models \varphi(do(\alpha,S_0))
\]

iff

\[
D_{ap} \cup D_{una} \cup D_{case} \models \Psi(S_0) \land R(\varphi,\alpha)(S_0),
\]

where \( D_{ap} \models \text{Poss}(\alpha,S_0) \equiv \Psi(S_0), R(\varphi,\alpha)(S_0) \) is obtained from \( R(\varphi,\alpha) \) by replacing \( s \) by \( S_0 \), and \( \varphi(do(\alpha,S_0)) \) is obtained from \( \varphi(s) \) by replacing \( s \) by \( do(\alpha,S_0) \).

Notice that this theorem depends on the particular form the constraints in \( D_{case} \) have: they are about Case only, that is, the result of forgetting it will yield a tautology: \( (\exists \text{Case})D_{case} = \text{true}. \)

One of the conditions in Theorem 3 is that under the unique names assumptions, \( R(\varphi,\alpha) \) be equivalent to a formula that does not mention Case. This condition holds if the action \( \alpha \)'s effects on the fluents in \( \varphi \) are definite. Thus Theorem 3 informs us for reasoning about the determinate effects of actions, the auxiliary predicate Case can be rightly ignored.

When either \( \varphi(s) \) or its regression mentions Case, we need to include constraints on Case:

**Theorem 4** Let \( \varphi(s) \) be a formula that does not quantify over situation variables, and does not mention Poss and Caused. Let \( \alpha \) be an action term. If \( D_{case} \) does not mention \( H \), then

\[
D \models \varphi(do(\alpha,S_0))
\]

iff

\[
D_{ap} \cup D_{una} \cup D_{case} \models \Psi(S_0) \land R(\varphi,\alpha)(S_0).
\]

Given the forms (16) and (17) the constraints in \( D_{case} \) must take, \( D_{case} \) does not mention \( H \) if all the indeterminate effects of actions are context free. This condition is needed because although \( D_{case} \) itself contains no information about \( H \), it can when used together with some assumptions about Case that can be easily incorporated into the query \( \varphi(s) \).

Finally, notice that Theorem 3 and Theorem 4 can be generalized to temporal projections with sequences of actions.

### The Ramification Problem

This section discusses how to represent indirect effects of actions in our framework. Example: whenever white and black are both true, happy will be true as well:

\[
\text{white}(s) \land \text{black}(s) \supset \text{Caused}(\text{happy}, \text{true}, s). \quad (21)
\]

Adding this axiom will make happy a possible indirect effect of the action drop. Due to the space limitation, we omit the details which can be found in the online version of this paper at:

Related Work and Discussions

Epistemologically, we have shown how the causal minimization framework of [5] can be used to specify the indeterminate effects of actions. Computationally, we have shown how goal regression can be used to reason about them.

There have been other proposals in the literature (e.g. [1, 2, 3, 12]) for specifying the effects of indeterminate actions. To the best of our knowledge, the computational contribution of this work is novel.

Among the extant approaches, the ones in [3] and [1] seem closest to ours. As we mentioned in Section 4, the releases propositions of [3]: A releases F corresponds to the following axiom in our language:

\[ \text{Poss}(A, s) \supset \text{Caused}(F, \text{true}, \text{do}(A, s)) \lor \text{Caused}(F, \text{false}, \text{do}(A, s)). \]

Regarding the work of [1], the In(F) and Out(F) predicates there correspond to Caused(F, true, do(A, s)) and Caused(F, false, do(A, s)), respectively, in our language. However, the formalism of [3] is limited because no complex releases propositions are allowed. For instance, one cannot write expressions like

\[ (\forall a). a \text{ releases } F \leftrightarrow a \text{ releases } F'. \]

The formalism of [1] is also limited because the action parameters of its In and Out predicates are not made explicit, thus cannot be quantified over.

Finally, we want to remark on the auxiliary predicate Case. In this paper, we have used it entirely for computational purposes. However, there are some interesting possible interpretations of this predicate.

There is a sense that Case can be interpreted in probabilistic terms. For instance, if

\[ \text{Poss}(\text{drop}, s) \land \text{Case}(1, \text{drop}, s) \supset \text{Caused}(\text{white}, \text{true}, \text{do}(	ext{drop}, s)) \land \text{Caused}(\text{black}, \text{false}, \text{do}(	ext{drop}, s)), \]

then Case(1, drop, s) may stand for the probability of the pin lying entirely within a white square after it has been dropped. Under this interpretation, the first constraint (16) on Case, in this example the following one:

\[ \text{Case}(1, \text{drop}, s) \lor \text{Case}(2, \text{drop}, s) \lor \text{Case}(3, \text{drop}, s), \]

says that the explicitly enumerated possible outcomes are both exclusive and exhaustive, and the constraints (17) simply eliminate redundant outcomes. In this regard, it would be interesting to formally connect our approach to probabilistic ones. This is a future research that we’re pursuing.

Another possible interpretation of Case is based on the view that in principle, it is always possible to reduce indeterminate actions to determinate ones, and one way of doing this is to introduce new fluents to name those low level contexts under which the effects of actions will be determinate. According to this view, Case can be seen as playing the role of such new fluents. For instance, Case(1, drop, s) may name the context under which drop has the effect of causing the pin lying entirely within a white square. We are currently exploring the possible impact of this interpretation as well.

Acknowledgements

Thanks to the other members of the University of Toronto Cognitive Robotics group (Yves Lesperance, Hector Levesque, Daniel Marcu, and Ray Reiter), to Vladimir Lifschitz, and to Yan Zhang for helpful discussions and comments. This research was supported by grants from the Government of Canada Institute for Robotics and Intelligent Systems, and from the National Science and Engineering Research Council of Canada.

References