Causality Revisited: Reifying Effects

Mike Stilman
Department of Computer Science
Stanford University
Stanford, CA 94305-9020
mstilman@cs.stanford.edu
Formal Reasoning Group Tech. Report Aug. 2001

Abstract

There is substantial research regarding the frame problem in reasoning about change. Recent work has illustrated the application of Causal Logic to solving this problem in a variety of domains. By means of reifying effects, we provide a framework for reasoning about the causes of change directly in a first-order logic formulation of the situation calculus. We demonstrate that this language augmented with circumscription is sufficient for an intuitive yet highly expressive solution to the frame problem. Avoiding a number of pitfalls of earlier attempts at such a framework, our formalism is structured to handle a wider set of domains. These include nondeterminism and context-dependent effects. In order to accomplish this solution, we discuss properties of effects and how they can be expressed in first order logic. We show a relationship between these properties and causal explanations for change. We study the effectiveness of this style of representation on benchmark problems and note the possible integration of other work regarding the situation calculus.

1 Introduction

The frame problem in formal reasoning about change was first described by McCarthy and Hayes (McCarthy and Hayes, 1969). The introduction of formal default reasoning methods such as circumscription gave rise to a number of potential solution (McCarthy, 1986). However, these initial solutions proved to be unsuccessful in that they would yield anomalous results in scenarios such as the Yale Shooting Problem (Hanks and McDermott, 1986).

In order to resolve YSP a number of alternative applications of circumscription were introduced in the late 80s. The most prominent of these were Causal Minimization (Lifschitz, 1987), Chronological Minimization (Shoham, 1988) and State-Based Minimization (Baker, 1989). However, each of these have been shown to have severe restrictions on their expressiveness. Furthermore, researchers began to seek solutions to the frame problem which could handle more complicated domains involving nondeterminism, ramifications and concurrency.

More recent solutions to the frame problem have often resorted to the use of alternative logics and formalisms other than situation calculus. Representative of these is the event calculus (Shanahan, 1995). While this approach has resolved a number of issues connected to the frame problem, it has also added a considerable amount of ontology and syntax. Among other intricacies, the event calculus demands that times be explicitly attached to situations. Though there is no consensus on a single standard for formal reasoning about change, the simplicity and versatility of the situation calculus is appealing. From (McCarthy, 1995) and various other works we see that the situation calculus can succinctly be expanded to handle concurrency, narratives, internal events among other complicated domains. However the frame problem in the language of the situation calculus remains an open question.

In this paper we explore an alternative avenue in approaching this problem. In particular, we hope to show that first order logic augmented with circumscription is sufficient for a considerably more intuitive, expressive solution to the frame problem in the situation calculus than earlier results.

Our exploration is highly related to recent work by (McCain and Turner, 1997) and (Lifschitz, 1997) which uses Causal Logic in achieving a solution. Though (McCain and Turner, 1997) relies on a "time-based" representation, more developments (Lifschitz, 1998) show that this approach can be adapted to the situation calculus. The cornerstone of this formalism is the requirement that the truth value of a fluent expression in a situation have a cause. Actions with the effect of giving such a truth value constitute a cause for the fluent expression to have this value. (Lifschitz, 1998) provides an axiomatization that uses *causes* to resolve the frame problem in domains with ramifications and nondeterminism.

(McCain and Turner, 1997) and (Lifschitz, 1998) additionally state that the process of "literal completion" could be used to translate the solution into first-order logic. In this paper we examine an alternative. We look into causality in an intuitive fashion which is directly through first-order logic. Early attempts to do so (Lifschitz, 1987) yielded a number of side-effects. Causal Minimization did not allow for actions to have context-dependent effects. In problems of post-diction, actions could not be described to have abnormal effects for a single situation. At that time questions of nondeterminism and ramifications were also left largely unexplored.

This has led us to the following questions. Is it possible to intuitively represent causes for the values of fluent expressions in first-order logic? If so, then how would this be useful in dealing with the frame problem? We have found that reifying effects is an extremely effective tool in answering these questions. In this paper we will demonstrate how our approach can be used to address the frame problem in its various forms.

2 Overview

As mentioned earlier, we will restrict ourselves to a standard version of the situation calculus in the language of first order logic augmented by circumscription. The language includes sorts for situations s, fluent expressions β and actions a. The predicate $Holds(\beta, s)$ represents the value of fluent β in situation s. The function $Result(a, s)$ maps the action a in s to the resulting situation.

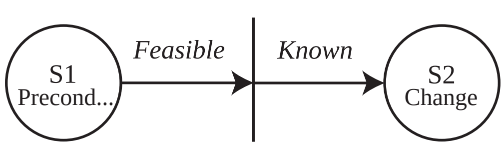
Of particular interest to us is any scenario in which a fluents value changes from situation s to $Result(a, s)$. As in previous first-order logic approaches we will consider the lack of change normal and not require any cause. In cases where change does occur we will require an explanation.

(Lifschitz, 1987) proposed that this explanation be given by the $Causes$ predicate associated with an action. Our approach regards the following observation. Explanations should not come directly from the action, but rather from the effect of the action. The rest of this section will focus on describing the intuition behind reifying effects. To do so, we first explain the relationship between effects and actions.

We propose that actions be viewed simply as occurrences in the world. However, a given action may be associated with a class of effects. Each effect is a cause for the value of a fluent expression to change. For example the single action of drinking may have the effects of quenching thirst and emptying a glass; hence causing the values of $Thirsty$ and $Empty$ to change respectively. Furthermore, we make the assumption that the state of the fluents in situation s where the action takes place determines which effects the action can and will have. In order to express this, we will add a new sort δ for constant symbols that designate the preconditions for a particular effect.

This leads us to the following definition of an effect. An *effect* is a 4-tuple (a, β, ν, δ). Given the occurrence of action a and the conditions designated by δ the effect is a cause for β to have the truth value ν in the resulting situation.

In the following sections we will describe a relatively simple set of formulas based on our notion of effects. The foundation for these formulas lies in two predicates that are used to express the properties of effects. Primarily we are interested in whether the effects are *known* to the reasoning system and whether they are *feasible* in a given situation. These properties of an effect are sufficient and necessary conditions for an effect to be a cause for explaining change.



The diagram above illustrates the conditions for explaining a change. Suppose a fluent expression in $S1$ changes value at $S2 = Result(a, S1)$ for some action a. Though we will look into other causes later in the paper, here we are solely concerned with the direct effects of actions. Suppose an effect of action a is known to produce the required change. If the preconditions in $S1$ make it feasible, then this effect is a cause for the change. When the effect is not known or it is not feasible in $S1$ it cannot be a cause for the change.

The following sections describe how we can formalize these concepts in first-order logic. Given such a formalization, we can use the standard form of circumscription (McCarthy, 1980) to minimize change that lacks a cause. In problems of projection this has an entirely intuitive effect. For actions without indirect effects the circumscribed theory explains changes using only known and feasible effects. This is given more attention in Section 4. The same occurs for problems of post-diction, with the following exception. If the problem contains a change which is not explainable through the known/feasible effects the minimization will contain a change without a known cause.

In our solution we choose to fix the timeline and state that the fluent must have changed during one of the actions between the situation of its initial value and the one in which its value changes. Hence we conclude that it was an abnormal effect of one of these actions, and that its value is uncertain during the situations in-between. Whether or not this is reasonable is a separate subject of debate. We could follow (Baker, 1989) or (Reiter, 1991) and add situations where the change took place. Since the notion of correctness here is rather vague we have chosen to leave this to a more detailed analysis, perhaps also involving the possibility of change due to concurrency. Furthermore, the formula indicating abnormal effects serves as a convenient marker for analyzing the causal structure of the formalism. Unlike Chronological Minimization it shows that the change may have occurred in any situation prior to the final one. Also, in contrast to (Lifschitz, 1987) it is situation dependent and does not carry over to other instances of actions.

In the following sections we will introduce the formalism that allows us to have the stated representation. We will discuss its features and provide some examples of its applications. We will conclude with a look at nondeterminism, ramifications and future projections for this approach.

3 Formalizing Effects

The purpose of this formalization is to explicitly represent known and defined effects. This allows us to define them as causes for change and therefore minimize change that is not caused.

3.1 Preliminaries

The domain is largely that of the situation calculus. It differs in the following respects. We add three predicates: $KnownEff, FeasEff$ and $AbEff$. Additionally we include a sort for objects that designate the preconditions for effects.

1. Variables: We will use the usual variables s for situations, a for actions, β for fluents and ν for truth values. Additionally the variable δ will range over the sort for effect designations.

2. Function: The $Result(a, s)$ function will be used in the standard way of mapping to the resulting situation of action a in situation s.

3. Predicates: Our language contains the standard $Holds(\beta, s)$ predicate of the situation calculus. We also add the following:

   $KnownEff(a, \beta, \nu, \delta)$ is a 4-tuple that is affirmed precisely when the effect of action a characterized by preconditions δ is known to give fluent β truth value ν.

   $FeasEff(a, \beta, \nu, \delta, s)$ is a 5-tuple that intuitively corresponds to the feasibility of the effect in situation s. It is logically equivalent to the existence of preconditions for the effect δ of action a on fluent β in situation s.

   $AbEff(a, \beta, \nu, s)$ is an abnormality predicate. It is affirmed precisely when the change in value of β to ν resulting from action a is not caused by a known and feasible effect. Hence we state that action a had an abnormal effect.

4. Circumscription: We will require circumscription as described in (McCarthy, 1980). Circumscription is a second order method which minimizes the extension of a particular predicate P in a theory Λ. Minimizing the extension of a predicate means that the extension only contains the objects which are entailed by Λ. Circumscription also allows us to vary the extensions of certain other predicates Q to find the minimal extension of P. We will write this CIRC[Λ;P;Q].

3.2 Formalization

The new formulas describing the properties of effects will take the place of effect axioms. Each effect is described with a pair of axioms. The first will simply affirm $KnownEff$ stating that the effect is known.

The second is a biconditional between $FeasEff$ and the preconditions for the effect. Here Π_+ and Π_- indicate the preconditions for an effect which affirms or negates the value of a fluent expression respectively. [1] Usually Π is a conjunction of Holds and ¬Holds formulas. If an effect is always feasible we affirm $FeasEff$ for all situations.

Each of the following pairs of formulas expresses a single effect in the theory. The δ in the formulas are constant objects designating the preconditions for the effect. Within the set of effects with equivalent a,β and ν we require that each effect have a unique δ. For simplicity we distinguish these constants by subscript, however in use they should simply be named from the syntax of the preconditions.

$KnownEff(a, \beta, T, \delta_i)$	$\Pi_+ \Leftrightarrow FeasEff(a, \beta, T, \delta_i, s)$
$KnownEff(a, \beta, F, \delta_j)$	$\Pi_- \Leftrightarrow FeasEff(a, \beta, F, \delta_j, s)$
$KnownEff(a, \beta, T, \delta_k)$	$FeasEff(a, \beta, T, \delta_k, s)$
$KnownEff(a, \beta, F, \delta_l)$	$FeasEff(a, \beta, F, \delta_l, s)$

The axioms described above are necessary to define effects as causes for change In the next section we will use circumscription to get closure for the *known* property of effects. Lacking, however are axioms which state that change does occur as a result of these effects. We therefore introduce two universal effect axioms.

These axioms apply to the simplest deterministic case. In Section 6 we discuss a minor modification that contributes to expressiveness.

$$KnownEff(a, \beta, T, \delta) \wedge FeasEff(a, \beta, T, \delta, s) \Rightarrow Holds(\beta, Result(a, s)) \quad (1)$$

$$KnownEff(a, \beta, T, \delta) \wedge FeasEff(a, \beta, F, \delta, s) \Rightarrow \neg Holds(\beta, Result(a, s)) \quad (2)$$

The intuition behind these axioms is that if a known effect of a is feasible in s then the change this effect causes takes place in $Result(a, s)$.

We also include the following *explanation-style* axioms:

$$[Holds(\beta, s) \wedge \neg Holds(\beta, Result(a, s))] \Rightarrow \exists \delta [FeasEff(a, \beta, F, \delta, s) \wedge KnownEff(a, \beta, F, \delta)] \vee AbEff(a, \beta, F, \delta, s) \quad (3)$$

$$[\neg Holds(\beta, s) \wedge Holds(\beta, Result(a, s))] \Rightarrow \exists \delta [FeasEff(a, \beta, T, \delta, s) \wedge KnownEff(a, \beta, T, \delta)] \vee AbEff(a, \beta, T, \delta, s) \quad (4)$$

These axioms require that every change in the value of a fluent be explained by some cause. In this formalization we allow the cause to only be an effect of some action. If the change occurs but it cannot be explained by a known and feasible effect of a we introduce an abnormal effect of a for this situation.

3.3 Applying Circumscription

We will want to use circumscription in order to complete the $KnownEff$ predicate and consequently minimize change which is not caused. In formalizing this, let Γ be a theory consisting of Axioms 1-4 as well as the pairs of definition axioms for the effects and unique names for actions and fluents. We let Σ be any further descriptions of the problem or query. Other than the definitions of $FeasEff$, Σ will include all axioms that are indexed on situations.

The theory we will actually query is then:

$$CIRC[\Sigma \wedge CIRC[\Gamma; KnownEff]; AbEff; Holds]$$

The inner circumscription of $KnownEff$ gives us completeness for this predicate. The minimal extension of $KnownEff$ in Γ only includes the $KnownEff$ objects stated as axioms. This is the case because no axioms in Γ require the existence of situations, and hence Axioms 3, 4 need not be used to affirm $KnownEff$.

The outer circumscription is equally simple. It requires that $AbEff$ have a minimal extension in the full theory. Since affirming $AbEff$ occurs when a change cannot be explained by any known/feasible effect this results in the minimization of change without cause.

[1]The preconditions for an effect should not be confused with preconditions for actions. The occurrence of the action does not depend on the effects it may have.

3.4 Unique Names for δ

We have previously described δ as a sort for constant symbols that designate effects. In particular the need for δ is to distinguish between effects which are equivalent aside from their preconditions. As in most situation calculus formalisms, we will demand unique names for fluents and actions. However, we can show that this is not necessary for the δ designations given our axioms.

The reason we would like the designations to be unique has to do with minimizing $AbEff$. Explaining change without adding tuples to satisfy $AbEff(a, \beta, \nu, \delta, s)$ is more minimal. We need to prove that this minimization will never introduce an anomalous equality between δ objects in order to explain change without adding $AbEff$.

Consider the situation where none of the known effects that explain a change are feasible in s. We would like this to always entail $AbEff$. To yield $AbEff(a, \beta, \nu, \delta, s)$ after the minimization, Axioms 3, 4 require the following:

$$\neg\exists\delta[FeasEff(a, \beta, \nu, \delta, s) \wedge KnownEff(a, \beta, \nu, \delta)]$$

where a,β,ν and s are determined by the particular change and a,β have unique names. Suppose there is some number of known effects in the *class* defined by the (a, β, ν) above. Let these effects be designated by $\delta_1 \ldots \delta_n$ respectively. Let $\Pi_1 \ldots \Pi_n$ be the preconditions for these effects such that no Π_i holds in s.

Now suppose some δ' still satisfies the formula above. Then it must satisfy $KnownEff(a, \beta, \nu, \delta')$ and $FeasEff(a, \beta, \nu, \delta', s)$. By the completeness of $KnownEff$ this means $\delta' = \delta_i$ for some δ_i above.

However since all δ_i are designations of known effect preconditions, the theory contains the following axiom for each.

$$FeasEff(a, \beta, \nu, \delta_i, s) \Leftrightarrow \Pi_i$$

Since we know $\neg\Pi_i$ we conclude $\neg FeasEff(a, \beta, \nu, \delta_i, s)$. However we also have $FeasEff(a, \beta, \nu, \delta', s)$. This establishes that for all δ_i: $\delta' \neq \delta_i$ contradicting our previous conclusion. Hence if none of the known effects are feasible in s, there is no δ' that satisfies the explanation formula.

This weak form of unique names for δ is sufficient for restricting explanations to known and feasible effects.

3.5 Framework Observations

The description of the axioms yields itself to an intuitive explanation of this solution. When discussing change in terms of effects, it is useful to consider the various properties of effects. By making $FeasEff$ logically equivalent to the preconditions for an effect, we give this predicate the precise meaning that its arguments describe an effect which is *feasible* in the situation. The completeness of $KnownEff$ given by the circumscription establishes that the only *known* effects are those that are axiomatized. This corresponds directly to the concept that the effect is *known* to the reasoning system. These associations give the *existence of effect* element of Axioms 3, 4 the same meaning as the requirements for an effect to be a cause.

In addition to being able to analyze the properties of effects, we are also capable of expressing various types of effects. This formalism places no restrictions on context-dependence in the form of contradicting preconditions for effects. Early Causal Minimization did not allow this (Lifschitz, 1987). In fact a single action can be expressed to have numerous various effects in different circumstances. When defining effects, however, one should be careful that the changes they cause do not contradict. In other words it would not be reasonable to define two effects for a single action which are feasible in the same situation that assign contradicting values to a fluent. Some such problems can be resolved by nondeterministic effects, which we discuss in Section 6.

We note that the formalism remains elaboration tolerant in terms of actions and effects. When adding a new effect exactly two axioms and a single δ' constant will need to be added to the theory. This applies to adding effects of existing actions or creating additional actions.

4 Projection: Example

In order to demonstrate the applications of this formalism, we will first look at the benchmark Yale Shooting Problem (Hanks and McDermott, 1986). We reformulate the axioms for this problem as follows:

$$KnownEff(Load, Loaded, T, \delta_1) \quad (5a)$$
$$FeasEff(Load, Loaded, T, \delta_1, s) \quad (5b)$$
$$KnownEff(Shoot, Alive, F, \delta_2) \quad (5c)$$
$$Holds(Loaded, s) \Leftrightarrow FeasEff(Shoot, Alive, F, \delta_2, s) \quad (5d)$$
$$UNA[Load, Wait, Shoot] \quad (5e)$$
$$UNA[Alive, Loaded] \quad (5f)$$
$$(5g)$$

In addition to Axioms 1-4 these axioms comprise Γ. The sentences describing the problem are unchanged from the original formulation.

$$Holds(Alive, S0) \quad (6a)$$
$$\neg Holds(Loaded, S0) \quad (6b)$$
$$UNA[S0, Result] \quad (6c)$$

Let Σ be these description axioms. We are interested in the following consequence:

$$CIRC[\Sigma \wedge CIRC[\Gamma; KnownEff]; AbEff; Holds] \models \neg Holds(Alive, Result(Shoot, Result(Wait, Result(Load, S0))))$$

The result of the inner circumscription is obvious. Since nothing in Γ affirms $KnownEff$ other than Axioms 5a, 5c, the minimal theory must satisfy:

$$KnownEff(a, \beta, \nu, \delta) \Leftrightarrow (a = Load, \beta = Loaded, \nu = T, \delta = \delta_1) \; (a = Shoot, \beta = Alive, \nu = F, \delta = \delta_2)$$

The outer circumscription requires more explanation. Suppose the circumscribed theory contained a change other than the two described above. By the loose unique names property of δ, there is no δ that would satisfy:

$$FeasEff(a, \beta, \nu, \delta, s) \wedge KnownEff(a, \beta, \nu, \delta)$$

Hence the theory would contain $AbEff(a, \beta, \nu, s)$. Then the minimal extension of $AbEff$ in $\Sigma \wedge CIRC[\Gamma; KnownEff]$ is non-empty.

Let us assign constants to the $Result$ of each action chronologically $(S0, S1 \ldots)$. Consider the intuitive sequence of events:

$$\begin{array}{ll} \neg Holds(Loaded, S0) & Holds(Alive, S0) \\ Holds(Loaded, S1) & Holds(Alive, S1) \\ Holds(Loaded, S2) & Holds(Alive, S2) \\ Holds(Loaded, S3) & \neg Holds(Alive, S3) \end{array}$$

In this interpretation of $Holds$, the only changes that occur are given by the two axiomatized $KnownEff$ objects. Furthermore $Holds(Loaded, S2)$ provides $FeasEff(Shoot, Alive, F, \delta_2, S2)$ by Axiom 5d. Hence we can explain both changes in Axioms 3, 4 by the effects with designations δ_1 and δ_2. It follows that the theory does not entail $AbEff$ for any change. This contradicts the premise that the minimal extension of $AbEff$ is non-empty. Therefore there can be no changes other than those given by the two axioms.

In fact, Axioms 1,2 ensure that these two changes must take place. Hence the sequence described above must be the extension of $Holds$ in the circumscribed theory. This gives us the correct result for this scenario.

In this study we are most interested in exploring the various properties of reifying effects. Briefly, we mention that this example generalizes to other projection problems. Without indirect effects of actions or domain constraints, it is always possible to vary $Holds$ so that $AbEff$ is not introduced. In doing so, the deterministic nature of these actions will yield a single extension of $Holds$ in the circumscribed theory.

5 Explanations: Grand Theft Auto

Let us explore how this solution handles post-diction. The scenario will involve leaving a car which is not stolen and then returning to find it stolen. We will represent the passage of time with two $Wait$ actions. The action $Wait$ will have no known effects. We would like our solution to reflect that the car must have been stolen during either of the two actions.

There are no effects so we need no axioms to define them. Thus Γ will be empty. Here is the statement of the problem (Σ), similar to (Baker, 1989).

$$\begin{array}{l} \neg Holds(Stolen, S0) \\ S1 = Result(Wait, S0) \\ S2 = Result(Wait, S1) \\ Holds(Stolen, S2) \end{array}$$

We know the car is stolen in $S2$ but we have no information regarding $S1$. Hence it would be reasonable that the circumscription allow:

$$Holds(Stolen, S1) \vee \neg Holds(Stolen, S1)$$

This simply says that the car was stolen during either of the Wait actions. The inner circumscription will negate all $KnownEff$ objects. In particular it will produce:

$$\forall \delta, s.[\neg KnownEff(Wait, Stolen, T, \delta, s)]$$

For $s = S0$ or $S1$, we clearly cannot satisfy the following with any δ.

$$\begin{array}{l} FeasEff(Wait, Stolen, T, \delta, s) \wedge \\ KnownEff(Wait, Stolen, T, \delta) \end{array}$$

Since a change must occur from either $S0 - S1$ or $S1 - S2$, Axiom 3 yields:

$$\begin{array}{l} AbEff(Wait, Stolen, T, S0) \vee \\ AbEff(Wait, Stolen, T, S1) \end{array}$$

The outer circumscription of the theory including this sentence with $AbEff$ will entail the following:

$$\begin{array}{l} \forall s.[AbEff(Wait, Stolen, T, s) \Rightarrow (s = S0)] \vee \\ \forall s.[AbEff(Wait, Stolen, T, s) \Rightarrow (s = S1)] \end{array}$$

Hence, the car may or may not have been stolen before S1. This problem generalizes to a variety of scenarios in which the known information about causes for change is insufficient to explain the changes. The result of the circumscription will always be an exclusive disjunction of abnormal effects. Since none of the actions are known to have the desired effect, every action that could potentially have had the effect is included in the disjunction.

Though the intuition of our conclusion is debatable, we will argue that it is justified. Given an axiomatized timeline, this problem asks us to explain the scenario. As a form of Causal Minimization we would like all changes to have a cause. If we do not have a known and feasible effect of an action that provides this cause then we assume that the cause is an abnormal effect of some action. This allows us to have indefinite information regarding the situations where the change may have occurred. We also reveal all possible causes for the change within a set timeline.

As stated previously, we could have followed (Baker, 1989) or other solutions which vary the timeline and introduce actions and/or situations. Possibly that would be a more useful solution. Our purpose in minimizing $AbEff$ is to simply minimize the unknown explanations. We could therefore consider $AbEff$ to be a place holder for the minimization of these other techniques. A more interesting possibility, however, is to use the current method as part of a higher-level structure. If the circumscription reveals that certain abnormal effects are possible explanations, the reasoning system could itself introduce further information regarding the situation. Aside from capturing some intuition, the exclusive disjunction would serve to point out where the introduction should take place.

6 Nondeterministic Effects

Leaving the discussion of future reasoning systems, let us return to a more concrete analysis of reified effects. In the overview of this formalism we discussed the value of allowing effects to have various properties. The properties could be used to define causes and heighten expressiveness. A rather simple but useful application of this comes up with nondeterministic actions.

Suppose action a' can have the effect of changing the value of fluent β' given certain preconditions. However we do not want to state that the action always has this effect given the preconditions. (McCain and Turner, 1997) and (Lifschitz, 1998) refer to this as a nondeterministic action. In our case this is a nondeterministic effect. Intuitively, this is an effect which is sufficient to act as a cause for explanation. Yet when it is feasible and its action is performed, it does not actually assert that the change takes place.

Expressing such effects in our formalism is simple and intuitive. Consider determinism to be a new property of effects. We will represent it by the predicate $DetEff(a', \beta', \nu', \delta')$. The arguments as with the *known* and *feasible* properties just represent the effect. For deterministic effects, we will affirm this formula in the axiomatization.

We will also need to adjust Axioms 1,2 as follows.

$$\begin{array}{c} KnownEff(a, \beta, T, \delta) \wedge FeasEff(a, \beta, T, \delta, s) \\ \wedge DetEff(a, \beta, T, \delta) \Rightarrow Holds(\beta, Result(a, s)) \end{array} \quad (7)$$

$$\begin{array}{c} KnownEff(a, \beta, T, \delta) \wedge FeasEff(a, \beta, F, \delta, s) \\ \wedge DetEff(a, \beta, F, \delta) \Rightarrow \neg Holds(\beta, Result(a, s)) \end{array} \quad (8)$$

The new formulation of the Effect Axioms simply formalizes the earlier intuition. An effect must only change the value of a fluent if it is deterministic. Clearly both deterministic and nondeterministic effects can still be used to explain change through Axioms 3,4. For the previous examples and other deterministic problems, declaring all effects to be deterministic will yield identical results. In cases where nondeterminism is useful, such as (Lifschitz, 1998) $Damaged$ block, this provides a very intuitive method for adding nondeterministic effects.

In addition to being intuitive, the way in which we axiomatize nondeterminism is also somewhat elaboration tolerant. If we wish to make a particular nondeterministic effect into a deterministic one, we can simply add a $DetEff$ formula for the effect to the theory.

7 Ramifications

We will not venture to formalize indirect effects of actions in this paper. However, ramifications are another extremely important factor in reasoning about change. They have also been the subject of substantial discussion in recent work. Consequently we introduce some thoughts regarding these sorts of effects and how they relate to this formalism.

Currently our formalism contains Axioms 3,4 which are used to indicate that change must have a cause. At this stage of development the cause can solely consist of the direct effects of actions. One possibility for handling ramifications is to expand this axiom to allow both direct and indirect effects.

In order to describe indirect effects we could reify them analogously to direct effects. This could be done by adding a causal structure to state constraints. Suppose an indirect effect is a triple consisting of (β, ν, δ), where the arguments are the same as a direct effect, without the action term. Then we can also ascribe it the properties of being *known* and *feasible*. For example, consider the following relationship.

$$\neg Holds(Alive, s) \rightarrow Holds(Dead, s)$$

This can additionally be described as:

$$\begin{array}{c} KnownEff(Dead, T, \delta_1) \\ KnownEff(Alive, T, \delta_2) \end{array}$$

Furthermore we can define the feasibility of each of these effects within the situation as follows.

$$FeasEff(Dead, T, \delta_1, s) \Leftrightarrow \neg Holds(Alive, s)$$

This would give us the necessary expansion of Axioms 3,4 where we would add the existence of δ for indirect effects to the disjunction of explanations.

It is not clear whether the quantity of axioms that would need to be added is justifiable for handling a single state constraint. It is possible that this will depend on the domain and the types/quantities of domain constraints present.

In order to avoid these extra semantics it may be more instrumental to consider alternatives. For example, we could implement ideas regarding state-based minimization (Baker, 1989). Since this sort of minimization inherently handles certain types of ramifications, we could connect it with our formalism.

Alternatively, the problem could be addressed using the concept of internal events (McCarthy, 2002). Our form of Causal Minimization only minimizes those effects which are unknown or undefined. Potentially this allows us to add internal events to handle indirect effects of actions. The effects of these internal events would be known/defined and would therefore have no impact on the minimization.

Given various options, the lack of expressiveness for ramifications is a current shortcoming but not a limitation of this approach. Incorporating work such as (McCarthy, 2002) would greatly contribute to the expressiveness of our formalism. Handling ramifications would only be one of the gains.

8 Concluding Remarks

This paper is largely an exploration of an alternative representation style for the situation calculus. We introduced the reification of effects as a potential method for attaining greater expressiveness in a solution to the frame problem. We demonstrated how some properties of effects could be expressed in first-order logic augmented with circumscription. This provided us with an effective tool for expressing the causes for change. Incorporating some ideas from Causal Logic (McCain and Turner, 1997),(Lifschitz, 1998) we were able to overcome a number of obstacles previously encountered by first order formalisms.

There are numerous directions for further research on this topic. First of all we find it particularly interesting to investigate other properties of effects. Possibly the interactions between these properties could lead to more robust and expressive results.

Additionally, although we discussed adding tools to handle indirect effects of actions, we only opened the door to further research. A practical approach would be to attempt the integration of existing representation styles with the one presented in this paper. From this we could analyze the potential of various methods in the context of reified effects.

Finally, perhaps by combining methods or further innovation it would be useful to construct a general solution to explanations in the situation calculus. The approach of our method is fairly intuitive. However, different styles of reasoning offer different sorts of explanations. We hope that a closer look at effects is a step towards the ability to express a variety of such styles.

9 Acknowledgments

I am very grateful to John McCarthy and Sheila McIlrath for their insightful discussions. I want to extend a particularly warm thank you to Aarati Parmar for her helpful suggestions, editing and encouragement.

References

[Baker, 1989] Baker, A. B. (1989). A simple solution to the yale shooting problem. In *KR 89*.

[Hanks and McDermott, 1986] Hanks, S. and McDermott, D. (1986). Default reasoning, nonmonotonic logics and frame problem. In *Proceedings of AAAI-86*, pages 328–333. Morgan Kaufmann.

[Lifschitz, 1987] Lifschitz, V. (1987). Formal theories of action. In Brown, F., editor, *The Frame Problem in Artificial Intelligence: Proceedings of the 1987 Workshop*, pages 35–58, Los Altos, CA. Morgan Kaufmann Publishers.

[Lifschitz, 1997] Lifschitz, V. (1997). On the logic of causal explanation. *Artificial Intelligence*, 96(2):451–465.

[Lifschitz, 1998] Lifschitz, V. (1998). Situation calculus and causal logic. In Cohn, A. G., Schubert, L., and Shapiro, S. C., editors, *KR'98: Principles of Knowledge Representation and Reasoning*, pages 536–546. Morgan Kaufmann, San Francisco, California.

[McCain and Turner, 1997] McCain, N. and Turner, H. (1997). Causal theories of action and change. In Shrobe, H. and Senator, T., editors, *Proceedings of the Thirteenth National Conference on Artificial Intelligence and the Eighth Innovative Applications of Artificial Intelligence Conference*, pages 460–465, Menlo Park, California. AAAI Press.

[McCarthy, 1980] McCarthy, J. (1980). Circumscription – A form of non-monotonic reasoning. *Artificial Intelligence*, 13(1,2):27–39.

[McCarthy, 1986] McCarthy, J. (1986). Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence*, 28:89–118.

[McCarthy, 1995] McCarthy, J. (1995). Situation calculus with concurrentevents and narrative. Contents subject to change. URL will remain.

[McCarthy, 2002] McCarthy, J. (2002). Actions and other events in situation calculus. In *Proceedings of KR 2002*. To be published.

[McCarthy and Hayes, 1969] McCarthy, J. and Hayes, P. J. (1969). Some philosophical problems from the standpoint of artificial intelligence. In Meltzer, B. and Michie, D., editors, *Machine Intelligence 4*, pages 463–502. Edinburgh University Press.

[Reiter, 1991] Reiter, R. (1991). The frame problem in the situation calculus: A simple solution (sometimes) and a completeness result for goal regression. In Lifschitz, V., editor, *Artificial Intelligence and Mathematical Theory of Computation*, pages 359–380. Academic Press.

[Shanahan, 1995] Shanahan, M. (1995). A circumscriptive calculus of events. *Artificial Intelligence*, 77(2):249–284.

[Shoham, 1988] Shoham, Y. (1988). Chronological ignorance: Experiments in nonmonotonic temporal reasoning. *Artificial Intelligence*, 36(3):279–331.