# State Defaults and Ramifications in the Unifying Action Calculus


**Ringo Baumann***, **Gerhard Brewka***, **Hannes Strass**†, **Michael Thielscher**‡, **Vadim Zaslawski***

*Department of Intelligent Systems, University of Leipzig
{baumann,brewka}@informatik.uni-leipzig.de, vz@uni-leipzig.de
†Department of Computer Science, Dresden University of Technology
hannes.strass@inf.tu-dresden.de
‡School of Computer Science and Engineering, The University of New South Wales
mit@cse.unsw.edu.au



## Abstract

We present a framework for reasoning about actions that not only solves the frame and ramification problems, but also the *state default problem*—the problem to determine what *normally* holds at a given time point. Yet, the framework is general enough not to be tied to a specific time structure. This is achieved as follows: We use effect axioms that draw ideas both from Reiter's successor state axioms and the non-monotonic causal theories by Giunchiglia et al. These axioms are formulated in a recently proposed unifying action calculus to guarantee independence of a specific underlying notion of time. Reiter's default logic is then wrapped around the resulting calculus and plays a key role in solving the ramification as well as the state default problem.


## Introduction

Reasoning about actions is one of the fundamental abilities cognitive agents must possess in order to be successful in their environments. Although KR research in this area has gone on for at least four decades now and has produced a plethora of competing approaches, there are still challenges to be met. In the light of the large variety of existing approaches unifying treatments are called for. Moreover, although several of the standard problems like the frame problem or the ramification problem have been solved individually in restricted settings, a combination of the solutions in more general settings is still lacking.

The Unifying Action Calculus (UAC) (Thielscher 2010) is an important step towards the much needed unifying treatment of action theories. The UAC is a first order calculus which abstracts away from a particular solution of the frame problem. Moreover, and maybe even more importantly, it abstracts away from a particular time structure. As demonstrated in (Thielscher 2010) various specific approaches like situation calculus (McCarthy 1963), event calculus (Kowalski and Sergot 1986; Shanahan 1997), or fluent calculus (Thielscher 1999) can be reconstructed as special instantiations of the UAC. Results obtained for UAC then immediately apply to all instantiations, which is a tremendous advantage.

The main goal of this paper is to develop an integrated solution to three different problems: the frame problem, the ramification problem, and the state default problem. The frame problem (how to represent what does not change) and the ramification problem (how to represent indirect effects of actions) have received considerably more attention than the state default problem (how to represent what normally holds in a state under certain conditions). This is at least partly due to the fact that, based on the work of Reiter (Reiter 1991), we now have robust monotonic solutions to the frame problem. The basic underlying idea is that, rather than relying on a non-monotonic logic, the frame problem is solved by adequate (first order) effect axioms which exactly specify the effects of actions on fluents.

Also our approach will rely on monotonic—yet carefully reformulated—effect axioms. However, non-monotonic reasoning still is highly relevant. In most realistic settings agents have partial knowledge about their environment only. Extending incomplete knowledge about a particular state with plausible conclusions, based on adequate default rules, is often the best an agent can do—and action theories should be able to capture this aspect, in addition to providing a solution to the frame and ramification problems.

To handle these non-monotonic aspects adequately our approach will be based on Reiter's default logic (Reiter 1980), one of the most prominent and most expressive non-monotonic formalisms. It will turn out that default logic not only allows us to represent state defaults, but also to handle ramifications adequately. A notorious problem arising with the treatment of indirect effects are self-justifying cycles. These can elegantly be avoided using default logic.

Since we aim at a broadly applicable solution, UAC is the obvious starting point. However, we need to give up some of UAC's generality. Since we want an integrated solution of the three mentioned problems, we cannot abstract away from solving the frame problem. For this reason our approach will be based on a particular solution of the frame problem. However, it will still be completely independent of the underlying time structure. Hence, it will be general enough to be applicable to different specific formalisms like situation calculus or event calculus.

The technical challenge then is to come up with adequately modified effect axioms which are able to accommodate default conclusions and indirect effects. This may look simpler than it actually is. Simply representing a state default like "if $f$ holds then normally also $f'$ holds" as the

Reiter default $Holds(f, s) : Holds(f', s)/Holds(f', s)$ does not work since this default is at odds with the effect axiom. Our solution will be based on the use of certain special predicates which, intuitively, allow us to decouple information about potential default conclusions and what actually is accepted in the successor state.

A fluent $f$ can hold, respectively not hold, in a state $s$ for different reasons: (a) because it is a direct or indirect effect of an action leading to $s$, (b) because an applicable state default says so, or (c) because of persistence. Of course, these reasons can be in conflict, for instance if an effect contradicts a default conclusion, or if a default conclusion contradicts persistence. We thus need to introduce priorities among the potential reasons to solve these conflicts. The most natural priorities which are also implicit in our approach are (from most to least preferred)

direct/indirect effects < default conclusions < persistence.

To capture these priorities we not only need an adequate effect axiom, but also a rather sophisticated representation of state defaults and indirect effects in default logic. For this reason we introduce a domain description language with simple constructs for defaults and ramifications. Specifications in this language are then translated to default theories.

Since the primary purpose of the paper is to illustrate how problematic interactions of action effects, default conclusions, and the persistence assumption can be handled, we restrict the presentation to actions with unconditional, local effects. Our approach can, with minor adjustments, be extended to conditional and non-local action effects as well; we however fear that this would complicate the reader's understanding of the principles underlying our integrated solution to the three aforementioned problems.

The paper is organized as follows. We first give the necessary background about UAC and default logic. We then discuss an elegant solution of the frame problem within UAC based on a particular effect axiom. In the subsequent section we present our solution to the state default problem based on a modification of this effect axiom and a particular representation of state defaults in default logic. We then introduce a default logic-based solution to the ramification problem, followed by a discussion of the combination of these approaches. Related work and open research topics are discussed in the concluding section.

## Background

### The Unifying Action Calculus

The unifying action calculus was proposed in (Thielscher 2010) to allow for a treatment of problems in reasoning about actions that is independent of a particular calculus. It is based on a finite, sorted logic language with equality which includes the sorts FLUENT, ACTION, and TIME along with the predicates $< : \text{TIME} \times \text{TIME}$, that denotes a (possibly partial) ordering on time points; $Holds : \text{FLUENT} \times \text{TIME}$, that is used to state that a fluent is true at a given time point; and $Poss : \text{ACTION} \times \text{TIME} \times \text{TIME}$, that means "action $a$ is possible starting at time $s$ and ending at time $t$."

The following definition introduces the fundamental types of formulas of the UAC: they allow to express properties of action domains at given time points and applicability conditions and effects of actions.

**Definition 1.** Let $\vec{s}$ be a sequence of variables of sort TIME.

- A *state formula* $\Phi[\vec{s}]$ *in* $\vec{s}$ is a first-order formula with free variables $\vec{s}$ where
  - for each occurrence of $Holds(f, s)$ in $\Phi[\vec{s}]$ we have $s \in \vec{s}$ and
  - predicate $Poss$ does not occur.

Let $s, t$ be variables of sort TIME and $A$ be a function into sort ACTION.

- A *precondition axiom* for $A(\vec{x})$ is of the form

$$Poss(A(\vec{x}), s, t) \equiv \pi_A[s] \quad (1)$$

  where $\pi_A[s]$ is a state formula in $s$ with free variables among $s, t, \vec{x}$.
- An *effect axiom* for $A(\vec{x})$ is of the form

$$Poss(A(\vec{x}), s, t) \supset (\exists \vec{y})((\forall f)(\Upsilon^+[s, t] \supset Holds(f, t)) \wedge (\forall f)(\Upsilon^-[s, t] \supset \neg Holds(f, t))) \quad (2)$$

  in which both $\Upsilon^+[s, t]$ and $\Upsilon^-[s, t]$ are state formulas in $s, t$ with free variables among $f, s, t, \vec{x}, \vec{y}$.[1]

It is clear that effect axiom (2) can only be used to conclude about the involved time points $s$ and $t$. Alas, this is no inherent restriction of the UAC: should the user desire to make conclusions about time points $t'$ with $s < t' < t$, they have to add appropriate axioms.[2]

We next formalize how action domains are axiomatized in the unifying action calculus.

**Definition 2.** A *(UAC) domain axiomatization* consists of a finite set of foundational axioms $\Omega$ (by which the UAC is instantiated by a concrete time structure, e.g. the branching situations along with the usual ordering from Situation Calculus), a set $\Pi$ of precondition axioms (1), and a set $\Upsilon$ of effect axioms (2); the latter two for all functions into sort ACTION.

A domain axiomatization is *progressing*, if

- $\Omega \models (\exists s : \text{TIME})(\forall t : \text{TIME}) s \leq t$ and
- $\Omega \cup \Pi \models Poss(a, s, t) \supset s < t$ .

In this paper, we are only concerned with progressing domain axiomatizations. To be able to reference the unique initial time point, we use the macro $Init(t) \stackrel{\text{def}}{=} \neg(\exists s) s < t$. We will then equip our domain axiomatizations with a set $\Sigma_0$ of *initial state axioms* describing the state of the world at the initial time point. These initial state axioms are state formulas of the form $Init(t) \supset \Phi[t]$ where $\Phi[t]$ is a state formula in $t$.

[1] The original definition of UAC effect axioms is more general; in this paper, we restrict their syntax for the sake of clarity. Variables $\vec{x}$ and $\vec{y}$ can be of any sort.

[2] For example, by instantiating UAC to the event calculus.

## Default Logic

In default logic (Reiter 1980) a default $d$ is an expression

$$\frac{A : B_1, \ldots, B_n}{C}$$

where $A, B_i$, and $C$ are first-order formulas. $A$ is the *prerequisite*, $B_1, \ldots, B_n$ are *consistency conditions* or *justifications*, and $C$ is the *consequent*. Note that we explicitly include justification-free defaults, that is the case where $n = 0$. Such defaults behave like classical inference rules and will be needed for handling ramifications later on. To save space we often write the default as $A{:}B_1, \ldots, B_n/C$.

A *default theory* is a pair $(\mathcal{D}, W)$, where $W$ is a set of sentences in first-order logic and $\mathcal{D}$ is a set of defaults.

A default is *closed* if its prerequisite, justifications, and consequent are sentences, that is, have no free variables. Otherwise, it is *open*. A default theory is *closed* if all its defaults are closed; otherwise, it is *open*.

Extensions are deductively closed sets of formulas which contain all elements of $W$, are closed under "applicable" defaults, and which are grounded in the sense that each formula has a non-cyclic derivation. For closed default theories this is captured by the following definition:

**Definition 3.** Let $(\mathcal{D}, W)$ be a closed default theory. The operator $\Gamma$ assigns to every set $S$ of formulas the smallest deductively closed set $U$ of formulas such that:

1. $W \subseteq U$,
2. if $A{:}B_1, \ldots, B_n/C \in D$, $U \models A$, $S \not\models \neg B_i$, $1 \leq i \leq n$, then $C \in U$.

A set $E$ of formulas is an *extension* of $(\mathcal{D}, W)$ if and only if $E = \Gamma(E)$, that is, $E$ is a fixed point of $\Gamma$.

We will interpret open defaults as schemata representing all of their ground instances. Therefore, open default theories can be viewed as shorthand notation for their closed counterparts.

We write $W \mathrel{|\!\!\approx}^{skept}_{\mathcal{D}} \Psi$ to express that the formula $\Psi$ is contained in each extension of the default theory $(\mathcal{D}, W)$.

## The Effect Axiom

This section presents the effect axiom that will be employed and elaborated throughout the paper. An axiom of this form was first presented in (Thielscher 2010) and is, as mentioned there, inspired by the work of (Giunchiglia et al. 2004). In the most simple form of the effect axiom, we allow two causes to determine a fluent's truth value: persistence and direct effects. Before introducing the axiom itself, we define two macros that formalize the individual causes. The first pair of macros expresses that a fluent $f$ persists from $s$ to $t$.

$$\mathit{FrameT}(f, s, t) \stackrel{\text{def}}{=} \mathit{Holds}(f, s) \land \mathit{Holds}(f, t) \tag{3}$$

$$\mathit{FrameF}(f, s, t) \stackrel{\text{def}}{=} \neg\mathit{Holds}(f, s) \land \neg\mathit{Holds}(f, t) \tag{4}$$

Suppose the direct (positive and negative) effects of an action are given as a set of fluent literals. These can easily be translated into "causes" for the purpose of designing an effect axiom for that action.

**Definition 4.** Let $A$ be a function into sort ACTION and $\Gamma_A$ be a set of fluent literals with free variables in $\vec{x}$ that denote the positive and negative direct, unconditional effects of $A(\vec{x})$, respectively.

$$\mathit{DirT}_{A(\vec{x})}(f, s, t) \stackrel{\text{def}}{=} \bigvee_{F(\vec{x}') \in \Gamma_A, \vec{x}' \subseteq \vec{x}} f = F(\vec{x}') \tag{5}$$

$$\mathit{DirF}_{A(\vec{x})}(f, s, t) \stackrel{\text{def}}{=} \bigvee_{\neg F(\vec{x}') \in \Gamma_A, \vec{x}' \subseteq \vec{x}} f = F(\vec{x}') \tag{6}$$

Recall that the restriction to unconditional, local effects is solely for presentation purposes; there is no inherent property of the approach that forbids more general action effects.

Now putting the causes "persistence" and "direct effect" together yields the basic version of our effect axiom.

**Definition 5.** Let $A$ be a function into sort ACTION. An *effect axiom with unconditional effects and the frame assumption* is of the form

$$\begin{aligned} &\mathit{Poss}(A(\vec{x}), s, t) \supset \\ &\quad (\forall f)(\mathit{Holds}(f, t) \equiv \mathit{CausedT}(f, A(\vec{x}), s, t)) \land \\ &\quad (\forall f)(\neg\mathit{Holds}(f, t) \equiv \mathit{CausedF}(f, A(\vec{x}), s, t)) \end{aligned} \tag{7}$$

$$\begin{aligned} &\mathit{CausedT}(f, A(\vec{x}), s, t) \stackrel{\text{def}}{=} \\ &\qquad \mathit{FrameT}(f, s, t) \lor \mathit{DirT}_{A(\vec{x})}(f, s, t) \end{aligned} \tag{8}$$

$$\begin{aligned} &\mathit{CausedF}(f, A(\vec{x}), s, t) \stackrel{\text{def}}{=} \\ &\qquad \mathit{FrameF}(f, s, t) \lor \mathit{DirF}_{A(\vec{x})}(f, s, t) \end{aligned} \tag{9}$$

The macros $\mathit{CausedT}$, $\mathit{CausedF}$ will be re-defined several times throughout the paper. (We will explicitly only define $\mathit{CausedT}$, the analogous macro definition of $\mathit{CausedF}$ can readily be obtained by replacing all causes by their negative versions.) When speaking about effect axiom (7), we will understand it retrofitted with their "latest version."

The design principle underlying our axiomatization technique is that of causation: a fluent holds at a time point that is the end point of an action if and only if there is a cause for that; similarly, a fluent does not hold if and only if there is a cause for that, too.[3]

From now on, when speaking about domain axiomatizations, we will understand all effect axioms to be of the form (7) and have the domain axiomatizations include uniqueness-of-names axioms for all finitely many function symbols into sorts FLUENT and ACTION.

For presentation purposes, we will make use of the concept of *fluent formulas*, where terms of sort FLUENT play the role of atomic formulas, and complex formulas can be built using the usual first-order connectives. We will denote by $\Phi[s]$ the state formula that is obtained by replacing all fluent literals in a fluent formula $\Phi$ by $\mathit{Holds}$ literals in $s$. The operator $|\cdot|$ will be used to extract the affirmative component

[3] The attentive reader will have noticed that the syntax of axiom (7) does not quite correspond to Definition 1. Simple syntactical manipulations can however be conducted to transform the effect axiom into a form that matches the structure of (2).

of a fluent literal, that is, $|\neg f| = |f| = f$. For reasons of expressiveness, we will allow fluent formulas to also contain equality atoms, that will remain unchanged by the translation into state formulas.

**Example 1.** Consider the very simple action domain that uses the fluents $\mathsf{SOP}(x)$ (object $x$ is a sheet of paper) and $\mathsf{PA}(x)$ (object $x$ is a paper airplane) along with the action $\mathsf{Fold}(x)$ that turns a sheet of paper $x$ into the paper airplane $x$. This is formulated as the (progressing) domain axiomatization $\Sigma = \Omega \cup \Pi \cup \Upsilon \cup \Sigma_0$, where $\Pi$ contains the precondition axiom

$$Poss(\mathsf{Fold}(x), s, t) \equiv Holds(\mathsf{SOP}(x), s) \wedge s < t,$$

$\Upsilon$ contains effect axiom (7) characterized by $\Gamma_{\mathsf{Fold}(x)} = \{\mathsf{PA}(x), \neg\mathsf{SOP}(x)\}$, and the initial state is given by $\Sigma_0 = \{Init(s) \supset Holds(\mathsf{SOP}(\mathsf{P}), s)\}$. We can now employ logical entailment to infer that after folding it in the initial time point, the object $\mathsf{P}$ is no longer a sheet of paper but a paper airplane:

$$\Sigma \models (Init(t_0) \wedge Poss(\mathsf{Fold}(\mathsf{P}), t_0, t_1)) \supset Holds(\mathsf{PA}(\mathsf{P}), t_1) \wedge \neg Holds(\mathsf{SOP}(\mathsf{P}), t_1)$$

The first formal result about our effect axiom shows that it correctly establishes action effects while still providing a solution to the frame problem. For a ground action $\alpha = A(\vec{a})$, we use the abbreviation $\Gamma_\alpha \stackrel{\text{def}}{=} \Gamma_A\{\vec{x} \mapsto \vec{a}\}$.

**Proposition 1.** *Let $\Sigma$ be a domain axiomatization such that $\Sigma \models Poss(\alpha, \sigma, \tau)$ for some ground action $\alpha$ and time points $\sigma, \tau$, and let $\varphi$ be a ground fluent literal.*

1. *Direct effects override persistence:*
   *Let $\varphi \in \Gamma_\alpha$. Then $\Sigma \models \varphi[\tau]$.*
2. *The frame assumption is correctly implemented:*
   *Let $|\varphi|, \neg|\varphi| \notin \Gamma_\alpha$. Then $\Sigma \models \varphi[\sigma] \equiv \varphi[\tau]$.*

*Proof.*

1. We make a case distinction on $\varphi$'s sign.
   (a) $\varphi = |\varphi|$:
   By Definition 4, $DirT_\alpha(\varphi, \sigma, \tau) \equiv \varphi = \varphi \vee \ldots$, hence $CausedT(\varphi, \alpha, \sigma, \tau) \equiv \varphi = \varphi \vee \ldots$. By axiom (7) and the assumption $\Sigma \models Poss(\alpha, \sigma, \tau)$, we have $\Sigma \models Holds(\varphi, \tau)$.
   (b) $\varphi = \neg|\varphi|$: Symmetric.
2. $DirT_\alpha(\varphi, \sigma, \tau) \equiv \bot$ and $DirF_\alpha(\varphi, \sigma, \tau) \equiv \bot$ due to the assumption. Hence, by expanding macros (8) and (9), we get $CausedT(\varphi, \alpha, \sigma, \tau) \equiv FrameT(\varphi, \sigma, \tau)$ and $CausedF(\varphi, \alpha, \sigma, \tau) \equiv FrameF(\varphi, \sigma, \tau)$. Together with assumption $\Sigma \models Poss(\alpha, \sigma, \tau)$, this yields $\Sigma \models (\varphi[\tau] \equiv (\varphi[\sigma] \wedge \varphi[\tau])) \wedge (\neg\varphi[\tau] \equiv (\neg\varphi[\sigma] \wedge \neg\varphi[\tau]))$ and, in consequence, $\Sigma \models (\varphi[\tau] \supset \varphi[\sigma]) \wedge (\neg\varphi[\tau] \supset \neg\varphi[\sigma])$. ☐

The following theorem shows an interesting relationship between our effect axioms with their particular solution to the frame problem and Reiter's successor state axioms (Reiter 1991) for the situation calculus.

**Theorem 2.** *If unconditional action effects are represented as in Definition 4, effect axioms (7) and successor state axioms are logically equivalent up to consistency of the effect specifications.*

*Proof.* See appendix. ☐

## State Defaults

An intelligent agent endowed with an internal world model will virtually never have complete knowledge about the world it is situated in. Therefore, an agent should have the technical ability to resolve its uncertainty towards the state of the world by making sensible default assumptions. In our approach, this is done by easy-to-write condition/conclusion pairs, that we call *state defaults*.

State defaults are a way of specifying conditions under which a fluent *normally* holds (or does not hold, respectively). These conditions are formulated by the user in a logical language and then translated into rules of default logic.

**Definition 6.** A *state default* is of the form $\Phi/\psi$, where $\Phi$, the *prerequisite*, is a fluent formula and $\psi$, the *consequent*, is a fluent literal.

Concluding a single fluent literal will not allow us to model disjunctive default knowledge; however, concluding a conjunction of literals can be emulated by distributing them over several rules with the same prerequisite.

Assuming a user has given their impression of how the world normally behaves by specifying a set of state defaults, we can now turn to translating these into the logical language we employ in this work. The special predicate symbol $DefT(f, s, t)$ will be used to express that a fluent $f$ is normally true at time point $t$. Likewise, $DefF(f, s, t)$ means that $f$ is normally false at $t$. Note that this is not the same as $\neg DefT(f, s, t)$, which only means that $f$ is not normally true at time point $t$. The additional TIME argument $s$ is used to keep track of the starting time point of the action that led to $t$. With this in mind, we can now define how to (automatically) create Reiter defaults from user-specified state defaults. (Observe that all default rules thus created are normal.)

**Definition 7.** Let $\delta = \Phi/\psi$ be a state default.

$$\delta_{Init} \stackrel{\text{def}}{=} \frac{Init(t) \wedge \Phi[t] : \psi[t]}{\psi[t]} \quad (10)$$

$$\delta_{Poss} \stackrel{\text{def}}{=} \frac{\Phi[t] \wedge \neg Viol_\delta(s) : Def(\psi, s, t)}{Def(\psi, s, t)} \quad (11)$$

$$Viol_\delta(s) \stackrel{\text{def}}{=} \Phi[s] \wedge \neg\psi[s]$$

$$Def(\psi, s, t) \stackrel{\text{def}}{=} \begin{cases} DefT(\psi, s, t) & \text{if } \psi = |\psi| \\ DefF(|\psi|, s, t) & \text{otherwise} \end{cases}$$

For a set $\mathcal{D}$ of state defaults, the corresponding set of default rules is defined as

$$\mathcal{D}_{Init,Poss} \stackrel{\text{def}}{=} \{\delta_{Init}, \delta_{Poss} \mid \delta \in \mathcal{D}\}. \quad (12)$$

The intuition behind the *Init* default rules for the initial time point should be clear: whenever, initially, the prerequisite is fulfilled and there is no reason to believe otherwise, we can safely assume the consequent. Note that we bypass the $Def$ predicate in this case—this is not a problem since there is no action leading to the initial time point and thus no effect axiom interfering with the default conclusion. For the $Poss$ defaults concerning two time points $s, t$ connected via action application, we require that (1) the state default's prerequisite hold at the resulting time point $t$, and (2) the state default not be violated at the starting time point $s$. (A default violation occurs when the prerequisite of a state default is known to be met, yet the opposite of the consequent prevails.) We then conclude that the consequent holds unless there is information to the contrary. The reason we watch out whether the default was violated at the starting time point is to prevent application of initially definitely violated state defaults through irrelevant actions. Without this precaution, any default conclusion, however unrealistic, could be enforced by applying a dummy action (like Wait)—a fine axiomatization of the ostrich algorithm.

Up to here, default conclusions and "hard facts" live in completely different, disconnected worlds ($Init$ defaults aside). It is the first modification to our effect axiom that brings them together: if a fluent is *normally* true (false) after applying an action we accept this as a cause for its being *actually* true (false).

**Definition 8.** Let $A$ be a function into sort ACTION. An *effect axiom with unconditional effects, the frame assumption, and simple normal state defaults* is of the form (7), where

$$CausedT(f, A(\vec{x}), s, t) \stackrel{\text{def}}{=} FrameT(f, s, t) \vee DirT_{A(\vec{x})}(f, s, t) \vee DefT(f, s, t) \quad (13)$$

Note that, whenever it is definitely known that $Holds(f, t)$ after $Poss(a, s, t)$, it follows from the effect axiom that $\neg DefF(f, s, t)$; a symmetrical argument applies if $\neg Holds(f, t)$. This means that definite knowledge about a fluent inhibits the opposite default conclusion.

But now imagine the following scenario: we know that a fluent $f$ holds at a time point $s$: $Holds(f, s)$. Nothing further is known about $f$—in particular no default information. Then an action $a$ occurs and leads to time point $t$, i.e. $Poss(a, s, t)$. The effects of action $a$ do not involve $f$, that is, $\Gamma_a \cap \{f, \neg f\} = \emptyset$. Intuitively, by persistence, we should be able to conclude that $f$ still holds at $t$: $Holds(f, t)$. But we only get the weaker conclusion $Holds(f, t) \vee (\neg Holds(f, t) \wedge DefF(f, s, t))$, which means that $f$ either stays true or becomes false due to a default conclusion. Since we know the latter is impossible, we would like to incorporate this information somewhere, rather than let the solution of one problem (the state default problem) disrupt the solution of another (the frame problem). The following addition ensures that knowledge about impossibility of default conclusions is adequately represented in our automatic translations.

**Definition 9.** Let $\mathcal{D}$ be a set of state defaults, $\psi$ be a fluent literal, and $s, t$ be variables of sort TIME. The *default closure axiom for $\psi$ with respect to $\mathcal{D}$* is

$$\left[ \bigwedge_{\Phi/\psi \in \mathcal{D}} \neg\Phi[t] \vee Viol_{\Phi/\psi}(s) \right] \supset \neg Def(\psi, s, t) \quad (14)$$

Note that for a fluent literal $\psi$ not mentioned as a consequent in $\mathcal{D}$ the default closure axiom is just $\top \supset \neg Def(\psi, s, t)$. For a domain axiomatization $\Sigma$ and a set $\mathcal{D}$ of state defaults, we denote by $\Sigma_{\mathcal{D}}$ the default closure axioms with respect to $\mathcal{D}$ and the fluent signature of $\Sigma$.

We are now ready to define the fundamental notion of our solution to the state default problem: a default theory where the incompletely specified world consists of a UAC domain axiomatization augmented by suitable default closure axioms, and the default rules are the automatic translations of user-specified, domain-dependent state defaults.

**Definition 10.** Let $\Sigma$ be a domain axiomatization and $\mathcal{D}$ be a set of state defaults. The corresponding *domain axiomatization with state defaults* is the pair $(\Sigma \cup \Sigma_{\mathcal{D}}, \mathcal{D}_{Init,Poss})$.

The workings of all of the preceding definitions are best understood with the help of an example domain.

**Example 2.** Recall the domain axiomatization $\Sigma$ from Example 1. We augment this domain by a state default saying that paper airplanes normally fly.

$$\delta = \mathsf{PA}(z)/\mathsf{Flies}(z) \quad (15)$$

The corresponding default rules for the initial time point and action executions are, respectively:

$$\frac{Init(t) \wedge Holds(\mathsf{PA}(z), t) : Holds(\mathsf{Flies}(z), t)}{Holds(\mathsf{Flies}(z), t)}$$

$$\frac{Holds(\mathsf{PA}(z), t) \wedge \neg Viol_\delta(s) : DefT(\mathsf{Flies}(z), s, t)}{DefT(\mathsf{Flies}(z), s, t)}$$

where $Viol_\delta(s) = Holds(\mathsf{PA}(z), s) \wedge \neg Holds(\mathsf{Flies}(z), s)$. The default closure axiom for (15) is

$$(\neg Holds(\mathsf{PA}(z), t) \vee Viol_\delta(s)) \supset \neg DefT(\mathsf{Flies}(z), s, t)$$

for the fluent literal $\mathsf{Flies}(z)$ and $\top \supset \neg Def(\psi, s, t)$ for all other literals $\psi$. Putting all of this together yields the domain axiomatization with state defaults $(\Sigma \cup \Sigma_{\mathcal{D}}, \mathcal{D}_{Init,Poss})$. Using default knowledge and skeptical reasoning now gives us the desired conclusion that a sheet of paper initially folded into a paper airplane indeed flies.

$$\Sigma \cup \Sigma_{\mathcal{D}} \approx^{skept}_{\mathcal{D}_{Init,Poss}} (Init(t_0) \wedge Poss(\mathsf{Fold}(\mathsf{P}), t_0, t_1)) \supset Holds(\mathsf{Flies}(\mathsf{P}), t_1)$$

Much like it was the case for the simple form of our effect axiom, there is also a formal result which shows that axiom (7), apart from solving the frame problem, implements a particular preference ordering among potential reasons for a fluent to hold or not to hold. Taking state defaults into account, the priorities become

direct effects < default conclusions < persistence.

**Theorem 3.** *Let $\Sigma$ be a domain axiomatization, $\mathcal{D}$ be a set of state defaults, $\delta = \Phi/\psi \in \mathcal{D}$ be a state default, $E$ be an extension for the domain axiomatization with state defaults $(\Sigma \cup \Sigma_{\mathcal{D}}, \mathcal{D}_{Init,Poss})$, $\varphi$ be a ground fluent, and $E \models Poss(\alpha, \sigma, \tau)$ for some ground action $\alpha$ and time points $\sigma, \tau$.*

*1. Effects override everything:*

$$\varphi \in \Gamma_\alpha \textit{ implies } E \models \varphi[\tau]$$

*2. Defaults override persistence:*

*(A) Let $\psi, \neg\psi \notin \Gamma_\alpha$;*

*(B) for each $\delta' = \Phi'/\neg\psi \in \mathcal{D}$, let $E \not\models \Phi'[\tau]$; and*

*(C) $E \models \Phi[\tau] \wedge \neg\, Viol_\delta(\sigma)$.*

*Then $E \models \psi[\tau]$.*

*3. The frame assumption is correctly implemented: Let $\psi, \neg\psi \notin \Gamma_\alpha$ and for all defaults $\delta_1 = \Phi_1/\psi, \delta_2 = \Phi_2/\neg\psi \in \mathcal{D}$, let $E \not\models \Phi_i[\tau]$ or $E \models Viol_{\delta_i}(\sigma)$. Then*

$$E \models \psi[\sigma] \equiv \psi[\tau]$$

*Proof.* See appendix. □

As important and nice as these properties are, a default theory would be useless if it did not admit any extension at all. But the existence of extensions for our default theories follows immediately from a result by Reiter (1980), since the default rules automatically created by Definition 7 are all normal. If the involved domain axiomatization is consistent, we can even guarantee all its extensions are consistent, too.

**Theorem 4.** *Let $\Sigma$ be a domain axiomatization and $\mathcal{D}$ be a set of state defaults. Then the corresponding domain axiomatization with state defaults $(\Sigma \cup \Sigma_{\mathcal{D}}, \mathcal{D}_{Init,Poss})$ has an extension. If furthermore $\Sigma$ is consistent, then so are all extensions for $(\Sigma \cup \Sigma_{\mathcal{D}}, \mathcal{D}_{Init,Poss})$.*

*Proof.* Existence of an extension is a corollary of (Reiter 1980, Theorem 3.1) since all defaults in $\mathcal{D}_{Init,Poss}$ are normal. If $\Sigma$ is consistent, then so is $\Sigma \cup \Sigma_{\mathcal{D}}$: for any model $I$ for $\Sigma$ with $I \models (\exists f, s, t) DefT(f, s, t)$, we can build a model $I'$ for $\Sigma$ with $|DefT^{I'}| < |DefT^{I}|$ by removing an element from $DefT^{I}$ and adjusting $Holds^{I'}$ and $Poss^{I'}$ accordingly. The same can be done with respect to $DefF$, hence for any $\mathcal{D}$ there exists a model for $\Sigma$ which is by the syntactic structure of (14) also a model for $\Sigma \cup \Sigma_{\mathcal{D}}$. Consistency of all extensions now follows from (Reiter 1980, Corollary 2.2). □

## Ramifications

To tackle the ramification problem, that is, the problem of determining the indirect effects of actions (Ginsberg and Smith 1987), we make use of *causal relationships*, a concept employed in the solution to the ramification problem in the fluent calculus (Thielscher 1997). These relationships express under which conditions a change of the truth value of a particular fluent induces a change of the truth value of another fluent.

**Definition 11.** A *causal relationship* is of the form

$$\chi \texttt{ causes } \psi \texttt{ if } \Phi \tag{16}$$

where $\Phi$, the *context*, is a fluent formula; $\chi$, the *trigger*, and $\psi$, the *effect*, are fluent literals.

Causal relationships are straightforwardly translated into "causes" in the sense of our effect axioms: whenever, for an action execution from $s$ to $t$, the context holds at $t$ and the truth value of the trigger has changed, then the effect $\psi$ of the causal relationship is an indirect effect of the action, written $Ind(\psi, s, t)$.

**Definition 12.** Let $r = \chi \texttt{ causes } \psi \texttt{ if } \Phi$ be a causal relationship. Its associated *ramification default* is

$$\delta_r \stackrel{\text{def}}{=} \frac{\Phi[t] \wedge \neg\chi[s] \wedge \chi[t] :}{Ind(\psi, s, t)} \tag{17}$$

$$Ind(\psi, s, t) \stackrel{\text{def}}{=} \begin{cases} IndT(\psi, s, t) & \text{if } \psi = |\psi| \\ IndF(|\psi|, s, t) & \text{otherwise} \end{cases}$$

For a set $\mathcal{R}$ of causal relationships, the corresponding set of ramification defaults is defined as $\mathcal{D}_{\mathcal{R}} \stackrel{\text{def}}{=} \{\delta_r \mid r \in \mathcal{R}\}$.

The additional causes of being an indirect positive or negative effect are easily integrated into effect axiom (7).

**Definition 13.** An *effect axiom with unconditional effects, the frame assumption, and indirect effects* is of the form (7), where

$$\begin{aligned} CausedT(f, A(\vec{x}), s, t) \stackrel{\text{def}}{=}\ & FrameT(f, s, t) \vee \\ & DirT_{A(\vec{x})}(f, s, t) \vee IndT(f, s, t) \end{aligned} \tag{18}$$

As before, the incorporation of ramifications into our effect axiom interferes with our solution of the frame problem, but we can deal with this in a manner similar to the one for state defaults.

**Definition 14.** Let $\mathcal{R}$ be a set of causal relationships, $\psi$ be a fluent literal, and $s, t$ be variables of sort TIME. The *ramification default closure axiom for $\psi$ with respect to $\mathcal{R}$* is

$$\left[ \bigwedge_{\chi \texttt{ causes } \psi \texttt{ if } \Phi \in \mathcal{R}} \neg\Phi[t] \vee \chi[s] \vee \neg\chi[t] \right] \supset \neg Ind(\psi, s, t) \tag{19}$$

Again, for a fluent literal $\psi$ not mentioned as an effect in $\mathcal{R}$ the ramification default closure axiom is just $\top \supset \neg Ind(\psi, s, t)$. The fluent signature of a domain axiomatization $\Sigma$ defines its associated set of ramification default closure axioms, which we will write as $\Sigma_{\mathcal{R}}$ for a particular set $\mathcal{R}$ of causal relationships.

The fundamental notion of our solution to the ramification problem is now a default theory comprised of a domain axiomatization extended by suitable ramification default closure axioms and default rules automatically created from causal relationships.

**Definition 15.** Let $\Sigma$ be a domain axiomatization and $\mathcal{R}$ be a set of causal relationships. The corresponding *domain axiomatization with ramification defaults* is the pair $(\Sigma \cup \Sigma_{\mathcal{R}}, \mathcal{D}_{\mathcal{R}})$.

Our example domain will illustrate the way indirect effects are treated in our framework.

**Example 3.** The fact that a paper airplane that is wet does not fly is expressed by the following causal relationships:

$$\mathsf{Wet}(y)\ \texttt{causes}\ \neg\mathsf{Flies}(y)\ \texttt{if}\ \mathsf{PA}(y)$$
$$\mathsf{PA}(y)\ \texttt{causes}\ \neg\mathsf{Flies}(y)\ \texttt{if}\ \mathsf{Wet}(y)$$

They are to be read as "whenever a paper airplane becomes wet, it does not fly" and "whenever something that is wet becomes a paper airplane, it does not fly," respectively. The corresponding ramification defaults are

$$\frac{Holds(\mathsf{PA}(y),t) \wedge \neg Holds(\mathsf{Wet}(y),s) \wedge Holds(\mathsf{Wet}(y),t) :}{IndF(\mathsf{Flies}(y),s,t)}$$
$$\frac{Holds(\mathsf{Wet}(y),t) \wedge \neg Holds(\mathsf{PA}(y),s) \wedge Holds(\mathsf{PA}(y),t) :}{IndF(\mathsf{Flies}(y),s,t)}$$

We again use the domain axiomatization $\Sigma$ from Example 1 and slightly modify the description of the initial time point $\Sigma_0$ to contain $Init(t_0) \supset (Holds(\mathsf{PA}(\mathsf{P}),t_0) \wedge Holds(\mathsf{Flies}(\mathsf{P}),t_0))$; that is, there initially exists a paper airplane that flies. We add an action $\mathsf{Dip}(x)$ with the intended meaning that object $x$ is dipped into water. It is characterized by precondition axiom $Poss(\mathsf{Dip}(x),s,t) \equiv s < t$ and set of effects $\Gamma_{\mathsf{Dip}(x)} = \{\mathsf{Wet}(x)\}$. Our formalism enables us to infer that when this action occurs to an initially flying paper airplane, then the plane cannot fly anymore:

$$\Sigma \cup \Sigma_{\mathcal{R}} \mathrel{|\!\!\approx}^{skept}_{\mathcal{D}_{\mathcal{R}}} (Init(t_0) \wedge Poss(\mathsf{Dip}(\mathsf{P}),t_0,t_1)) \supset \neg Holds(\mathsf{Flies}(\mathsf{P}),t_1)$$

The existence of extensions for default theories of this kind can again be guaranteed. Yet, for justification-free default rules, consistency of the domain axiomatization does not imply consistency of the extension: a causal relationship might be inconsistent with the action domain in that an indirect effect openly contradicts a direct one. Having a single inconsistent extension is therefore a clear indication of an axiomatization error.

**Theorem 5.** *Let $\Sigma$ be a domain axiomatization and $\mathcal{R}$ be a set of causal relationships. Then the corresponding domain axiomatization with ramification defaults $(\Sigma \cup \Sigma_{\mathcal{R}}, \mathcal{D}_{\mathcal{R}})$ has an extension.*

Due to the groundedness of extensions in default logic, this approach to dealing with ramifications is quite expressive: it can deal with instantaneous change and circular dependencies in between indirect effects.

**Example 4** (Gear Wheel Domain)**.** There are two interlocked gear wheels, that can be separately turned and stopped. Let the fluents $\mathsf{W}_1, \mathsf{W}_2$ express that the first (resp. second) gear wheel is turning. The actions to initiate/end this are $\mathsf{Turn}_i, \mathsf{Stop}_i$ with effects $\Gamma_{\mathsf{Turn}_i} = \{\mathsf{W}_i\}$, $\Gamma_{\mathsf{Stop}_i} = \{\neg\mathsf{W}_i\}$, $i = 1, 2$; there also exists a trivial action $\mathsf{Wait}$ without any direct effects, $\Gamma_{\mathsf{Wait}} = \emptyset$. The causality relating the interlocked gear wheels is described as follows: whenever the first wheel is turned, it causes the second one to turn, and vice versa; whenever the first wheel is stopped, it causes the second one to stop as well, and vice versa. The respective causal relationships $\mathcal{R}$

$$\mathsf{W}_1\ \texttt{causes}\ \mathsf{W}_2\ \texttt{if}\ \top,\ \neg\mathsf{W}_1\ \texttt{causes}\ \neg\mathsf{W}_2\ \texttt{if}\ \top,$$
$$\mathsf{W}_2\ \texttt{causes}\ \mathsf{W}_1\ \texttt{if}\ \top,\ \neg\mathsf{W}_2\ \texttt{causes}\ \neg\mathsf{W}_1\ \texttt{if}\ \top$$

are straightforwardly translated into ramification default rules $\mathcal{D}_{\mathcal{R}}$. We take the domain axiomatization $\Sigma$ to be comprised of precondition axioms $Poss(A,s,t) \equiv s < t$ for all above-mentioned actions $A$, effect axioms (7) according to Definitions 4 and 13, and the initial state characterized by $Init(t) \supset \neg Holds(\mathsf{W}_1,t) \wedge \neg Holds(\mathsf{W}_2,t)$, that is, both wheels stand still. According to the domain axiomatization with ramification defaults $(\Sigma \cup \Sigma_{\mathcal{R}}, \mathcal{D}_{\mathcal{R}})$ constructed by our approach, turning one wheel causes both of them to turn:

$$\Sigma \cup \Sigma_{\mathcal{R}} \mathrel{|\!\!\approx}^{skept}_{\mathcal{D}_{\mathcal{R}}} (Init(t_0) \wedge Poss(\mathsf{Turn}_1,t_0,t_1)) \supset Holds(\mathsf{W}_1,t_1) \wedge Holds(\mathsf{W}_2,t_1)$$

On the other hand, during waiting, the wheels keep their current state.

$$\Sigma \cup \Sigma_{\mathcal{R}} \mathrel{|\!\!\approx}^{skept}_{\mathcal{D}_{\mathcal{R}}} (Init(t_0) \wedge Poss(\mathsf{Wait},t_0,t_1)) \supset \neg Holds(\mathsf{W}_1,t_1) \wedge \neg Holds(\mathsf{W}_2,t_1)$$

We remark that this default logic-based, calculus-independent solution of the ramification problem subsumes the solution of (Lin 1995): instantiating UAC to situation calculus and translating Lin's "causal rules" into Reiter defaults allows us to draw all conclusions that can be drawn by the circumscription policy proposed there; in case of cyclic dependencies, we can even draw more conclusions.

## Combining State Defaults and Ramifications

The individual solutions of the state default and the ramification problem can easily be integrated into a joint solution. Both state defaults and causal relationships are translated to Reiter defaults of a particular form, and these defaults just need to be combined to a single default theory. However, as before the effect axioms need to be modified adequately:

**Definition 16.** An *effect axiom with unconditional effects, the frame assumption, simple normal state defaults, and indirect effects* is of the form (7), where

$$CausedT(f,A(\vec{x}),s,t) \stackrel{\text{def}}{=} FrameT(f,s,t) \vee DirT_{A(\vec{x})}(f,s,t) \vee DefT(f,s,t) \vee IndT(f,s,t) \quad (20)$$

**Definition 17.** Let $\Sigma$ be a domain axiomatization, $\mathcal{D}$ be a set of state defaults, and $\mathcal{R}$ be a set of causal relationships. The corresponding *domain axiomatization with state defaults and ramification defaults* is the pair $(\Sigma \cup \Sigma_{\mathcal{D}} \cup \Sigma_{\mathcal{R}}, \mathcal{D}_{Init,Poss} \cup \mathcal{D}_{\mathcal{R}})$.

There is an issue that needs to be mentioned, however. The default theories obtained for solving the state default and the ramification problems individually are such that existence of extensions is guaranteed. In the former case this is so because all defaults are normal, in the latter case because all defaults are justification free. Combining normal and justification-free defaults can, however, lead to nonexistence of extensions.

**Example 5.** Consider the default theory with $W = \{A\}$ and $\mathcal{D} = \{A{:}B/B, B{:}/\neg A\}$. All defaults are either normal or justification-free, yet no extension exists.

Intuitively, the application of a default may indirectly create a conflict with the prerequisite of the very same default. Two different views seem possible:

- A default creating a conflict with its own prerequisite should not be applicable when constructing an extension.
- Situations like this can be viewed as indicating a modeling error. In terms of skeptical consequence, non-existence of extensions is the same as inconsistency, and like the latter it shows that the axiomatization needs to be debugged.

We believe both views are reasonable and refrain from taking sides. Nevertheless, we offer a technical solution representing the first view:

**Definition 18.** Let $AX = (\Sigma \cup \Sigma_{\mathcal{D}} \cup \Sigma_{\mathcal{R}}, \mathcal{D}_{Init,Poss} \cup \mathcal{D}_{\mathcal{R}})$ be a domain axiomatization with state defaults and ramification defaults. $E$ is a *weak extension of* $AX$ iff it is an extension of $(\Sigma \cup \Sigma_{\mathcal{D}} \cup \Sigma_{\mathcal{R}}, \mathcal{D}^*_{Init,Poss} \cup \mathcal{D}_{\mathcal{R}})$, where $\mathcal{D}^*_{Init,Poss}$ is a maximal subset of $\mathcal{D}_{Init,Poss}$ such that an extension exists.

Note that the minimization of defaults to be disregarded implicit in this definition is to be read in terms of open defaults. Thus, if the number of state defaults is finite we can guarantee existence of weak extensions.

## Discussion

In this paper we proposed an integrated solution to the frame, state default, and ramification problems based on a unifying action calculus. The solution abstracts from the underlying time structure and relies on a monotonic solution of the frame problem. This is achieved through adequate effect axioms, together with a particular representation of state defaults and causal connections in terms of Reiter defaults.

Our present work builds on our own preliminary results on the use of mere atomic, normal defaults (and without considering ramifications) in the UAC (Strass and Thielscher 2009a; 2009b). Apart from that, the representation of state defaults in general action theories has received surprisingly little attention so far. A notable exception is (Lakemeyer and Levesque 2009) where an approach is introduced based on a variant of the situation calculus using the logic of "only knowing." Several essential differences between this work and our approach exist:

- our approach is not restricted to the time structure of situation calculus;
- we handle defaults as well as ramifications, while the latter are not covered in (Lakemeyer and Levesque 2009);
- the treatment of defaults implicit in "only knowing" is the one of AEL (Moore 1985). AEL has well-known problems with self-justifying cycles which do not arise in default logic. This qualifies the latter as much more adequate for our purposes.

The non-monotonic causal logic of (Giunchiglia et al. 2004) (originating itself in (McCain and Turner 1995)) provides a constructor `default` $\psi$ `if` $\Phi$ that could be interpreted as implementing the state default $\Phi/\psi$. However, upon closer inspection, it turns out that mapping a default like our $\mathsf{PA}(z)/\mathsf{Flies}(z)$ into `default` $\mathsf{Flies}(z)$ `if` $\mathsf{PA}(z)$ won't do, for the following reason. In Causal Logic, the default `default` $\psi$ `if` $\Phi$ ultimately translates to the causal rule $i : \psi \Leftarrow i : \psi \wedge i : \Phi$ (for all time points $i$). Since similar causal rules are also used to address the Frame and Ramification Problems, these defaults can easily lead to a conflict with other causal rules if the consequent of a default is a "regular" fluent. For instance, suppose our example default were understood as

$$\texttt{caused}\ \mathsf{Flies}(z)\ \texttt{if}\ \mathsf{Flies}(z) \wedge \mathsf{PA}(z)$$

with Flies being a regular fluent, which is subject to persistence. Persistence is expressed by the two causal rules

$$i+1 : \mathsf{Flies}(z) \Leftarrow i : \mathsf{Flies}(z) \wedge i+1 : \mathsf{Flies}(z)$$
$$i+1 : \neg\mathsf{Flies}(z) \Leftarrow i : \neg\mathsf{Flies}(z) \wedge i+1 : \neg\mathsf{Flies}(z)$$

Now, consider a time point $t$ at which $\neg\mathsf{Flies}(\mathsf{P}) \wedge \neg\mathsf{PA}(\mathsf{P})$ holds, and take a successor time point $t+1$ at which $\mathsf{PA}(\mathsf{P})$ becomes true. The causal theory would then support a model in which $\mathsf{Flies}(\mathsf{P})$ becomes true (since there would be a "cause" for this, namely the default rule), but also a model in which $\mathsf{Flies}(\mathsf{P})$ stays false (there would be a "cause" for this, too, namely the persistence rule). Hence, it would not follow that $\mathsf{Flies}(\mathsf{P})$ holds (by default) at $t+1$. To handle this example right, one would have to leave out the causal rule stating persistence of the fluent Flies. But this would lead to problems in other situations where there is no action or default that says the truth value of $\mathsf{Flies}(\mathsf{P})$ should change. In such situations persistence of Flies, and thus the causal persistence rule leading to problems in the example above, is obviously needed. (For the case of fluents that are not subject to persistence, "statically determined" or "defined" fluents, formulating a default for them has no effect since their truth value cannot vary given fixed truth values of the defining fluents.) Now even if we ignore interactions of default rules with persistence rules and just look at a single time point, the way a default causal law translates into a Reiter default in the logic of (Giunchiglia et al. 2004) models the correct intuition only for complete extensions (extensions that, for every formula, entail either the formula or its negation): for example, the default causal law

$$\texttt{default}\ \mathsf{Flies}(\mathsf{P})\ \texttt{if}\ \mathsf{PA}(\mathsf{P})$$

becomes the causal rule

$$\mathsf{Flies}(\mathsf{P}) \Leftarrow \mathsf{Flies}(\mathsf{P}) \wedge \mathsf{PA}(\mathsf{P}) \tag{21}$$

that is identified with the Reiter default

$$\frac{: \mathsf{Flies}(\mathsf{P}) \wedge \mathsf{PA}(\mathsf{P})}{\mathsf{Flies}(\mathsf{P})} \tag{22}$$

We firstly observe that the causal theory $\{(21)\}$ has no model; put in another way, the corresponding default theory $(\{(22)\}, \emptyset)$ has no complete extension. Secondly, the only (incomplete) extension of this default theory is the set $E = Th(\{\mathsf{Flies}(\mathsf{P})\})$, from which we can conclude the object P flies although we do not even know whether it is a paper airplane (since $E$ does not say anything about $\mathsf{PA}(\mathsf{P})$). Hence, one cannot straightforwardly use (Giunchiglia et al. 2004) to default reason even about a single time point without somehow sorting out the incomplete extensions first.

Two further approaches exist that are related to our work. In (Thielscher 2001) the fluent calculus is extended by default rules that enable an agent to assume away abnormal qualifications of actions by default. The main differences to our present work are: we are not restricted to a specific calculus or time structure, and we handle general state defaults rather than dealing with the specific case of action qualifications. In (Kakas, Michael, and Miller 2008) reasoning about actions is combined with default reasoning in the framework of an Action Description Language. The main difference to our work is the use of a special-purpose syntax for default and effect laws along with a tailor-made semantics with its own definition of a linear time structure, models and entailment. In contrast, our approach is formulated entirely within classical first-order logic and default logic. The advantages are, first, that the standard semantics for these logics apply; second, that we can make use of known results, e.g. on the existence of extensions; and, third, that standard inference methods and theorem proving techniques can be employed for automated reasoning with our domain axiomatizations.

Regarding future work, we already mentioned that there is nothing inherent in the approach to state defaults presented here that hinders us to incorporate more general action descriptions. Still, this is not immediate—the effect axioms have to be modified once again—, and we have not yet found a solution that is conceptually as satisfactory to us as the material presented here. What has not yet been considered, and could in fact constitute more of a change, are disjunctive default conclusions and disjunctive action effects.

As a second (and more important) future research goal, we will use the insights gained through this theoretical framework to develop an actual, practical implementation. This might seem a bit startling at first: after all, extension existence for closed normal first-order default theories is not even semi-decidable. Thus, we will have to make some restricting assumptions towards the expressivity of the implemented fragment of our theory. This is not a drawback but only common practice in action theory-inspired implementations: both situation calculus and fluent calculus are second-order logical formalisms, yet the programming languages based on them, Golog (Levesque et al. 1997) and Flux (Thielscher 2005), are successfully used in practice. Still, the reader might argue that default logic retains its high computational complexity even through restriction to propositional logic.[4] But in spite of such discouraging worst-case complexity results, the close relationship of default reasoning with logic programming (Bidoit and Froidevaux 1987) and the remarkable maturity of current implementations of the answer set programming (ASP) paradigm (Gelfond and Lifschitz 1991) make ASP the perfect candidate for putting our framework into practice.

**Acknowledgements.** We thank the anonymous reviewers for helpful comments and suggestions on the paper. This work was partially supported by *Deutsche Forschungsgemeinschaft* under grants Br 1817/3 and Th 541/15 and by the Australian Research Council under grant FT0991348.

[4]Skeptical default reasoning is $\Pi_2^P$-complete even for propositional normal default theories (Gottlob 1992).

## References

Bidoit, N., and Froidevaux, C. 1987. Minimalism Subsumes Default Logic and Circumscription in Stratified Logic Programming. In *Symposium on Logic in Computer Science: Proceedings*. IEEE Press, New York, NY, USA.

Gelfond, M., and Lifschitz, V. 1991. Classical Negation in Logic Programs and Disjunctive Databases. *New Generation Computing* 9:365–385.

Ginsberg, M. L., and Smith, D. E. 1987. Reasoning about Action I: A Possible Worlds Approach. *Artificial Intelligence* 35:233–258.

Giunchiglia, E.; Lee, J.; Lifschitz, V.; McCain, N.; and Turner, H. 2004. Nonmonotonic Causal Theories. *Artificial Intelligence* 153(1-2):49–104.

Gottlob, G. 1992. Complexity Results for Nonmonotonic Logics. *Journal of Logic and Computation* 2(3):397–425.

Kakas, A.; Michael, L.; and Miller, R. 2008. Fred meets Tweety. In *Proceedings of the 18th European Conference on Artificial Intelligence (ECAI 2008)*, 747–748. Amsterdam, The Netherlands: IOS Press.

Kowalski, R. A., and Sergot, M. J. 1986. A Logic-based Calculus of Events. *New Generation Computing* 4(1):67–95.

Lakemeyer, G., and Levesque, H. 2009. A Semantical Account of Progression in the Presence of Defaults. In *Proceedings of the Twenty-first International Joint Conference on Artificial Intelligence (IJCAI'09)*, 842–847.

Levesque, H. J.; Reiter, R.; Lin, F.; and Scherl, R. B. 1997. Golog: A Logic Programming Language for Dynamic Domains. *Journal of Logic Programming* 31.

Lin, F. 1995. Embracing Causality in Specifying the Indirect Effects of Actions. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95)*, 1985–1993. Morgan Kaufmann.

McCain, N., and Turner, H. 1995. A Causal Theory of Ramifications and Qualifications. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95)*, 1978–1984.

McCarthy, J. 1963. Situations and Actions and Causal Laws. Stanford Artificial Intelligence Project: Memo 2.

Moore, R. 1985. Semantical Considerations of Nonmonotonic Logic. *Artificial Intelligence* 25(1):75–94.

Reiter, R. 1980. A Logic for Default Reasoning. *Artificial Intelligence* 13:81–132.

Reiter, R. 1991. The Frame Problem in the Situation Calculus: A Simple Solution (Sometimes) and a Completeness Result for Goal Regression. In *Artificial Intelligence and Mathematical Theory of Computation – Papers in Honor of John McCarthy*, 359–380. Academic Press.

Shanahan, M. 1997. *Solving the Frame Problem: A Mathematical Investigation of the Common Sense Law of Inertia*. The MIT Press.

Strass, H., and Thielscher, M. 2009a. On Defaults in Action Theories. In Mertsching, B.; Hund, M.; and Aziz, Z., eds., *Proceedings of the 32nd German Annual Conference on Artificial Intelligence (KI'09)*, 298–305. Paderborn, Germany: Springer-Verlag Berlin Heidelberg.

Strass, H., and Thielscher, M. 2009b. Simple Default Reasoning in Theories of Action. In Nicholson, A., and Li, X., eds., *22nd Australasian Joint Conference on Artificial Intelligence (AI'09)*, 31–40. Melbourne, Australia: Springer-Verlag Berlin Heidelberg.

Thielscher, M. 1997. Ramification and Causality. *Artificial Intelligence* 89(1–2):317–364.

Thielscher, M. 1999. From Situation Calculus to Fluent Calculus: State Update Axioms as a Solution to the Inferential Frame Problem. *Artificial Intelligence* 111(1–2):277–299.

Thielscher, M. 2001. The Qualification Problem: A Solution to the Problem of Anomalous Models. *Artificial Intelligence* 131(1-2):1–37.

Thielscher, M. 2005. FLUX: A Logic Programming Method for Reasoning Agents. *Theory and Practice of Logic Programming* 5(4-5):533–565.

Thielscher, M. 2010. A Unifying Action Calculus. To appear in *Artificial Intelligence*.

## Appendix

***Proof of Theorem 2***. Successor state axioms with direct effect specifications (5, 6) can be expressed as UAC effect axioms by replacing $Do(A(\vec{x}), s)$ with $t$:

$$\begin{aligned} &Poss(A(\vec{x}), s, t) \supset \\ &\quad (\forall f)[(Holds(f,t) \equiv DirT_{A(\vec{x})}(f,s,t) \vee \\ &\qquad Holds(f,s) \wedge \neg DirF_{A(\vec{x})}(f,s,t)] \end{aligned} \tag{23}$$

Since (7) and (23) share the same predicates, it is quite useful to introduce them through the following abbreviations:

- $S$ for $Holds(f,s)$,
- $D$ for $Holds(f,t)$,
- $P$ for $DirT_{A(\vec{x})}(f,s,t)$, and
- $N$ for $DirF_{A(\vec{x})}(f,s,t)$.

Now, (23) can be represented as

$$Poss(A(\vec{x}), s, t) \supset F(S, D, P, N)$$

where $F(S,D,P,N) \stackrel{\text{def}}{=} D \equiv P \vee S \wedge \neg N$.
Analogously, (7) is represented as

$$Poss(A(\vec{x}), s, t) \supset G(S, D, P, N)$$

where $G(S,D,P,N) \stackrel{\text{def}}{=}$

$$(D \equiv S \wedge D \vee P) \wedge (\neg D \equiv \neg S \wedge \neg D \vee N).$$

The formulae $F$ and $G$ are logically equivalent to the unique formulae $\widehat{F}$ and $\widehat{G}$, respectively, that are in canonical conjunctive normal form (CCNF), where uniqueness is understood to be up to the order of the factors. The trivial conversion into CCNF is omitted here. Uniqueness of CCNF guarantees that two formulae are identical if they share the same CCNF. Comparison of $\widehat{F}$ and $\widehat{G}$ reveals that a set of ten disjuncts of $\widehat{G}$ strictly includes a set of eight disjuncts of $\widehat{F}$. Therefore, supplementing $\widehat{F}$ with the missing disjuncts from $\widehat{G}$ one can establish an identity

$$\begin{aligned} &\widehat{G}(S,D,P,N) \equiv \\ &\quad \left.\begin{array}{l} \left.\begin{array}{l} \left.\begin{array}{l} (\neg S \vee \neg D \vee \neg P \vee \neg N) \\ \wedge(\ \ S \vee \neg D \vee \neg P \vee \neg N) \end{array}\right\} \neg D \vee \neg P \vee \neg N \\ \left.\begin{array}{l} \wedge(\neg S \vee \ \ D \vee \neg P \vee \neg N) \\ \wedge(\ \ S \vee \ \ D \vee \neg P \vee \neg N) \end{array}\right\} \ D \vee \neg P \vee \neg N \end{array}\right\} \neg P \vee \neg N \end{array}\right. \\ &\quad \left.\begin{array}{l} \wedge(\neg S \vee \ \ D \vee \neg P \vee \neg N) \\ \wedge(\ \ S \vee \ \ D \vee \neg P \vee \neg N) \\ \wedge(\neg S \vee \neg D \vee \ \ P \vee \neg N) \\ \wedge(\ \ S \vee \neg D \vee \ \ P \vee \neg N) \\ \wedge(\neg S \vee \ \ D \vee \neg P \vee \ \ N) \\ \wedge(\ \ S \vee \ \ D \vee \neg P \vee \ \ N) \\ \wedge(\neg S \vee \ \ D \vee \ \ P \vee \ \ N) \\ \wedge(\ \ S \vee \neg D \vee \ \ P \vee \ \ N) \end{array}\right\} \widehat{F}(S,D,P,N) \ . \end{aligned}$$

Taking into account the idempotency law, the two leftmost disjuncts of $\widehat{F}$ can be once more conjoined with the right part of the identity. After that, combining together the four leftmost disjuncts of $\widehat{G}$ and replacing formulae in CCNF with original ones, the identity reduces to

$$G(S,D,P,N) \equiv F(S,D,P,N) \wedge (\neg P \vee \neg N)$$

Thus, the effect axiom (7) is logically equivalent to:

$$\begin{aligned} &Poss(A(\vec{x}), s, t) \supset \\ &\quad (\forall f)[(Holds(f,t) \equiv DirT_{A(\vec{x})}(f,s,t) \vee \\ &\qquad Holds(f,s) \wedge \neg DirF_{A(\vec{x})}(f,s,t)) \wedge \\ &\qquad \neg(DirT_{A(\vec{x})}(f,s,t) \wedge DirF_{A(\vec{x})}(f,s,t))] \end{aligned}$$

□

***Proof of Theorem 3***. If $E$ is inconsistent, the claims are immediate, so in what follows assume that $E$ is consistent.

1. We make a case distinction on the sign of $\psi$.
   (a) $\psi = |\psi|$:
   By Definition 8, $DirT_\alpha(\varphi, \sigma, \tau) \equiv \varphi = \varphi \vee \ldots$ and consequently $CausedT(\varphi, \alpha, \sigma, \tau) \equiv \varphi = \varphi \vee \ldots$, thus by effect axiom (7) and assumption $E \models Poss(\alpha, \sigma, \tau)$, we get $E \models Holds(\varphi, \tau)$.
   (b) $\psi = \neg\,|\psi|$: Analogous.
2. Assume $E \not\models \neg Def(\psi, \sigma, \tau)$. Together with assumption (C) this means that default $\delta_{Poss}$ is applicable to $E$. Since $E$ is an extension, we have $E \models Def(\psi, \sigma, \tau)$. Invoking effect axiom (7) yields the claim. It remains to establish $E \not\models \neg Def(\psi, \sigma, \tau)$. By (Reiter 1980, Theorem 2.1), there exist $E_i$, $i \geq 0$, with $E_0 = \Sigma \cup \Sigma_{\mathcal{D}}$ and for $i \geq 0$, $E_{i+1} = Th(E_i) \cup \left\{ \beta \mid \frac{\alpha:\beta}{\beta} \in \mathcal{D}_{Init,Poss}, \alpha \in E_i, \neg\beta \notin E \right\}$ such that $E = \bigcup_{i=0}^{\infty} E_i$. We first show $E_0 \not\models \neg Def(\psi, \sigma, \tau)$. Due to (A) and effect axiom (7), we know that $E_0 \not\models \psi[\tau]$ and $E_0 \not\models \neg\psi[\tau]$. By assumption (C), consistency of $E$, and $E_0 \subseteq E$, we get $E_0 \not\models \neg\Phi[\tau] \vee Viol_\delta(\sigma)$. In combination with $E_0 \not\models \neg\psi[\tau]$ this yields $E_0 \not\models \neg Def(\psi, \sigma, \tau)$,

since default closure axioms and $\alpha$'s effect axiom are the only ways to conclude $\neg Def(\psi, \sigma, \tau)$. Now assume to the contrary that $E \models \neg Def(\psi, \sigma, \tau)$. Then there is a minimal integer $i \geq 0$ such that $E_i \not\models \neg Def(\psi, \sigma, \tau)$ and $E_{i+1} \models \neg Def(\psi, \sigma, \tau)$. This must be due to a default $\delta' = \Phi' / \neg\psi \in \mathcal{D}$ with $E_i \models \Phi'[\tau] \wedge \neg Viol_{\delta'}(\sigma)$. But then $E_i \subseteq E$ implies $E \models \Phi'[\tau]$, which is a contradiction to assumption (B).

3. By the assumption that $\psi$ and $\neg\psi$ are not amongst $\alpha$'s direct effects, $\Sigma \models \neg Dir(\psi, \sigma, \tau) \wedge \neg Dir(\neg\psi, \sigma, \tau)$. The second assumption ensures that the left-hand sides of all relevant default closure axioms become true, hence $E \models \neg Def(\psi, \sigma, \tau) \wedge \neg Def(\neg\psi, \sigma, \tau)$. In consequence, $CausedT(|\psi|, \alpha, \sigma, \tau)$ reduces to $FrameT(|\psi|, \sigma, \tau)$ (similarly for $CausedF$), which proves the claim. □