Deriving properties of belief update from theories of action

Alvaro del Val and Yoav Shoham
Computer Science Department
Stanford University
Stanford, CA 94305
{delval,shoham}@cs.stanford.edu

Abstract

We present an approach to database update as a form of non monotonic temporal reasoning, the main idea of which is the (circumscriptive) minimization of changes with respect to a set of facts declared "persistent by default." The focus of the paper is on the relation between this approach and the update semantics recently proposed by Katsuno and Mendelzon. Our contribution in this regard is twofold:

- We prove a representation theorem for KM semantics in terms of a restricted subfamily of the operators defined by our construction.
- We show how the KM semantics can be generalized by relaxing our construction in a number of ways, each justified in certain intuitive circumstances and each corresponding to one specific postulate. It follows that there are reasonable update operators outside the KM family.

Our approach is not dependent for its plausibility on this connection with KM semantics. Rather, it provides a relatively rich and flexible framework in which the frame and ramification problems can be solved in a systematic way by reasoning about default persistence of facts.

1 Introduction

Theories of belief change address the following general question: Given an initial database Γ and a new piece of information μ to be incorporated into it, what should the new database be? Much previous work has concentrated on normative theories of belief revision, postulating a number of conditions that a 'rational' belief-revision operator should satisfy (cf. [Gärdenfors, 1988; Alchourrón *et al.*, 1985]). These postulates aim to capture stability properties, eliminating unnecessary perturbations to the original database. For example, one postulate states that if μ is consistent with Γ then the new database is simply the addition of μ to Γ.

It has recently been proposed that the operation of incorporating a new piece of information into an existing database might take different meanings. In particular, Katsuno and Mendelzon [1991] have suggested to distinguish between belief *revision* and belief *update*; loosely speaking, the former says that the beliefs may have been wrong and in need of revision, whereas the latter says that the beliefs were correct, but the world has in the meanwhile evolved and the beliefs must be updated. Katsuno and Mendelzon [1991] propose a set of belief-update postulates, which are similar to, but distinct from, the belief-revision postulates; they also provide the model theory for these postulates, in the form of a representation theorem.

According to the KM proposal, therefore, the problem of update is fundamentally one of *reasoning about change*. This is a problem which has received substantial attention over the years from researchers interested in formal theories of time and action, in what was until now a largely independent line of enquiry. Of particular interest have been theories of *nonmonotonic temporal reasoning*, and associated problems such as the *frame*, *qualification* and *ramification* problems. The essential issue in nonmonotonic temporal reasoning is that fully specifying the conditions needed to make predictions (or other temporal inferences) might be impossible to do explicitly. For example, one would not want to have to explicitly state that after starting the engine of a yellow car, the car remains yellow; that should follow 'by default.' Research in this area consists primarily of formal methods for achieving these default conclusions.

As said, these two research areas, belief change and non-monotonic temporal reasoning, have been until now largely disjoint. In this paper we tie them together, and in particular show that the KM-postulates need not be postulated at all, but can instead, under certain conditions, be derived analytically. After defining our approach to update, we present a representation theorem for KM update operators in terms of restrictions of the class of operators allowed in our construction. By relaxing these restrictions (which, unlike many of the KM postulates, have a fairly clear intuitive interpretation in terms of default persistence of facts), we can point to circumstances in which we need to move beyond KM semantics. Thus, our results provide an answer to the question: *Under what circumstances should the KM postulates be accepted?*

The basic idea is simple, we believe, and is as follows. Although update is supposed to reflect changes that have taken place in the world over time, the update problem (like that of belief revision) is formulated using a language incorporating no model of time or change. The sentences describe a single state of the world, a snapshot of it at a given situation. There is a clear computational advantage to this, since one does not need to store the whole history of the domain under consideration. Time is only implicit in the succession of theories resulting from a series of updates, but old theories are simply discarded. However, the price of this conciseness is impoverished semantical content, in which the rules of update must be postulated from outside the theory. In order to recover the lost information, in this paper we translate the update problem into a richer language, which explicates the temporal information: The initial database is taken to describe a particular situation, and the update formula is taken to describe the effects of a particular action. A formal theory of action is then used to infer facts about the result of taking the particular action in the particular situation; the formal theory of action we will employ is a generalization of the one proposed in [Lin and Shoham, 1991], which is described later. Finally, anything inferred about the resulting situation can be translated back to the timeless framework of belief update.

To illustrate the connection between non-monotonic temporal reasoning and update, suppose we have a database of facts about university life, and that we want to update the database with the fact that Smith is enrolled in CS205 this quarter. Clearly, this new fact should not affect many other facts in the KB, such as the composition of the faculty, the course offerings, or the color of the university buildings. On the other hand, it *should* have an effect on, say, her schedule and the total number of units she is taking; and these changes can in turn have other indirect effects, e.g. she might become unenrollable in CS210 due to a scheduling conflict, and there might be changes in her status as a part- or full-time student and in the amount of tuition she has to pay. These indirect effects might depend on a variety of circumstances: scheduling conflicts depend on the course offerings for the quarter, the university might define part- and full-time status differently for undergraduate and graduate students, the tuition owed might depend on whether she is a resident of the state, etc. These are clear instances of the frame and ramification problems: it would be unreasonable to require the user to explicitly state all what does not change as a result of the update (the frame problem), or to provide individual axioms for each of the myriad ways in which the effects of a particular update might depend on the circumstances in which the update takes place (the ramification problem). We believe, therefore, that any proposed update semantics must be defended in terms of its ability to provide solutions to these problems. A similar argument, restricted to the frame problem, has been independently advanced by Reiter [1992a].

The structure of this paper is as follows. Sections 2 and 3 briefly review, respectively, KM semantics and the theory of action proposed in [Lin and Shoham, 1991]. In section 4 we generalize Lin and Shoham's class of theories of action to allow for explicit reasoning about default persistence, and encode the update problem as a theory of action. The main results of the paper are in sections 5 and 6. In section 5, a representation theorem for KM semantics is proved by placing some restrictions on the theory of action. The consequences of lifting these restrictions are discussed in section 6. The main body of the paper concludes with section 7, where we turn our attention to the ramification problem by considering the case of update in the presence of 'constraints'. Some formal extensions of our results, related work, and topics for further research are discussed in the concluding section.

There are two appendixes, which are included so as to avoid too many technical detours in the presentation of the main results. The first one presents an alternative (non circumscriptive) characterization of the update operators determined by our definition of update in section 4. This characterization helps understand how one could go about to compute the consequences of updates with the theory of action presented here, and it is used in the proof of the main results of the paper. The second appendix contains the proofs of all lemmas and theorems in the paper, from the main text and the first appendix.

2 Update in propositional languages: Review

Katsuno and Mendelzon proposed eight postulates that should be satisfied by update operators. Let $\diamond$ be an update operator for a propositional language $\mathcal{L}$ with a finite number of propositional variables $\mathcal{P}_{\mathcal{L}}$. The KM postulates are the following:

(U1) $\psi \diamond \mu$ implies μ.

(U2) If ψ implies μ then $\psi \diamond \mu$ is equivalent to ψ.

(U3) If ψ and μ are satisfiable then $\psi \diamond \mu$ is also satisfiable.

(U4) If $\models \psi_1 \equiv \psi_2$ and $\models \mu_1 \equiv \mu_2$ then $\psi_1 \diamond \mu_1$ is equivalent to $\psi_2 \diamond \mu_2$.

(U5) $(\psi \diamond \mu) \wedge \phi$ implies $\psi \diamond (\mu \wedge \phi)$.

(U6) If $\psi \diamond \mu_1$ implies μ_2 and $\psi \diamond \mu_2$ implies μ_1 then $\psi \diamond \mu_1$ is equivalent to $\psi \diamond \mu_2$.

(U7) If ψ is complete then $(\psi \diamond \mu_1) \wedge (\psi \diamond \mu_2)$ implies $\psi \diamond (\mu_1 \vee \mu_2)$.

(U8) $(\psi_1 \vee \psi_2) \diamond \mu$ is equivalent to $(\psi_1 \diamond \mu) \vee (\psi_2 \diamond \mu)$

Update operators satisfying these postulates can be characterized as follows. An *update assignment* is a function which assigns to each interpretation I a relation $\leq_I$ over the set of interpretations of the language. We say that this assignment is *faithful* iff for any J, if $I \neq J$ then $I \leq_I J$ and $J \not\leq_I I$. Let $Mod(\phi)$ stand for the models of the formula ϕ, and $Min(S, \leq)$, for any set S and relation $\leq$ over S, denote the set of elements of S that are minimal under $\leq$.

Theorem 1 *[Katsuno and Mendelzon, 1991] An update operator $\diamond$ satisfies conditions (U1)-(U8) iff there exists a faithful update assignment that maps each interpretation I to a partial order $\leq_I$ such that:*

$$Mod(\psi \diamond \mu) = \bigcup_{I \in Mod(\psi)} Min(Mod(\mu), \leq_I).$$

3 Provably correct theories of action: Review

As mentioned in the introduction, in tying together the theory of belief update with theories of action, we will use a particular theory of action, a generalization of the one proposed in [Lin and Shoham, 1991]. Beside demonstrating that their formulation yields the desired results in particular examples that had been previously discussed in the literature, Lin and Shoham were the first to offer a formal justification for a theory of action. Specifically, they defined a formal criterion for the adequacy of theories of action and showed their formulation adequate relative to this criterion. The criterion, 'epistemological completeness', intuitively guarantees that a causal theory about the effects of (deterministic) actions will preserve completeness in the following sense: given a complete description of a situation, the axioms in the theory will result in a complete description of subsequent situations. Lin and Shoham provide a sufficient condition for a theory to be epistemologically complete, and prove how a wide class of theories can be made epistemologically complete by providing either frame axioms or a non-monotonic completion. This can be seen as providing a functionality for theories of action which is analogous to that of Clark's completion and various forms of closed world assumption for logic programming.

Though their monotonic completion is intended and adequate only for deterministic actions, the non-monotonic completion can also be applied to non-deterministic actions. In the absence of constraints, deterministic actions roughly correspond to conjunctive updates (those in which the update formula is a conjunction of literals);[1] in choosing to allow for arbitrary updates and constraints (and thus for actions with arbitrary effects) we will leave the realm of epistemologically complete theories. Since the theory proposed by Lin and Shoham is a special case of the more general construction presented here, we can still preserve its virtues for applications in which the underlying completeness assumptions are reasonable. Our generalization of this theory adds an additional degree of flexibility, needed in order to handle the frame and ramification problem in the more complicated circumstances in which we cannot appeal to epistemological completeness. We review Lin and Shoham's approach next.

We use the situation calculus formalism. To be precise, our language $\mathcal{L}_{\mathcal{S}}$ is a three-sorted predicate calculus language. The three sorts partition the terms of the language into situation, action and (propositional) fluent terms. In addition, there is a binary function *result*, whose first argument is of action sort and whose second argument and value are of situation sort; and a binary predicate *holds*, such that its first argument is of fluent sort and its second argument is of situation sort. The semantics of the language is the standard one for sorted predicate calculus.

Lin and Shoham consider a class of *causal theories* for deterministic actions, defined in the standard situation calculus language. Formally, a causal theory T for action A with the domain constraint C and the direct effects $P_1, \ldots, P_n$ under preconditions $R_1, \ldots, R_n$ is the

[1] Whether an action is deterministic depends not only on its direct effects but on the indirect ones as well, as determined by the constraints.

set of causal axioms:

$$\forall s.\ R_i(s) \supset holds(P_i, result(A, s))$$

and the constraint involving no situation terms other than s:

$$\forall s.\ C(s)$$

The causal theory T for an action A tells us what changes as a result of action A. It does not tell us what does not change; for that we need either a set of frame axioms, or some way of non-monotonically specifying them. We describe the latter next.

We fix a set of fluent terms $\mathbf{F}$, and, following [Lifschitz, 1990], use a predicate $frame$, whose extension is exactly the set of fluents denoted by some fluent term in $\mathbf{F}$. We assume $p \in \mathbf{F}$ can be formalized with a first-order formula F. We also use $ab(p, a, s)$ as an abbreviation for the formula:

$$frame(p) \wedge (holds(p, s) \equiv \neg holds(p, result(a, s))).$$

In order to apply the circumscription policy, we add unique names axioms[2] (N1) for the set of fluents $\mathbf{F}$ and (N2) for situations, and extend the language with a new predicate symbol $holds'$, with same sorts as $holds$ for its arguments. Given a causal theory T, let $W(s)$ be an abbreviation for the formula:

$$(\forall p.\ holds(p, s) \equiv holds'(p, s)) \wedge (\bigwedge T) \wedge N1 \wedge N2 \wedge F.$$

Finally, $Comp(T)$ is an abbreviation for:

$$\forall s, a.\ Circum(W(s); ab(p, a, s); holds),$$

where $Circum(W(s); ab(p, a, s); holds)$ stands for the circumscription of ab in W with $holds$ allowed to vary. Intuitively, what the policy does is to minimize changes one situation at a time. For any situation S, the minimization will allow $holds$ to vary at any other point except at S, since $holds'$ is kept fixed. As a very simple example, suppose we have an action $toggle$, whose effect is to change the value of a fluent P_1, formulated in a theory T_0 with no constraints and the single causal axiom:

$$\forall s.\ holds(P_1, s) \equiv \neg holds(P_1, result(toggle, s)).$$

[2] Unique names axioms force any two different terms to denote different objects.

Then $Comp(T_0)$ entails:

$$\forall s, p.\ frame(p) \wedge p \neq P_1 \supset holds(p, s) \equiv holds(p, result(toggle, s)),$$

i.e. $toggle$ causes no change in the value of any (frame) fluent other than P_1.

4 The update problem in situation calculus

The update problem can be formulated in situation calculus as follows. The initial database is taken to describe some particular situation S. The update formula is taken to describe the occurrence of a special action, denoted by A^S_μ, whose intuitive reading is 'that action which, when taken in S, causes μ.' The updated database is taken to describe the situation $result(A^S_\mu, S)$.

As an illustration of our approach, suppose we are given an initial database $(p \vee (q \wedge r))$, which we want to update with the formula $\neg r$. Using for example Winslett's 'PMA' update operator (defined later), the updated database is then $((p \vee q) \wedge \neg r)$. Our approach to obtain this result is to translate the database into the situation calculus formula $holds(p, S) \vee (holds(q, S) \wedge holds(r, S))$, for some situation S, and to compute the set of consequences about the situation $result(A^S_{\neg r}, S)$ entailed by the circumscription of a theory similar to the one described in the previous section, containing the causal axiom $holds(not(r), result(A^S_{\neg r}, S))$.

We will be using the language $\mathcal{L}_\mathcal{S}$ of situation calculus, defined in the previous section. Because our focus is in establishing a mapping with propositional update operators, we will define the set of fluent terms of $\mathcal{L}_\mathcal{S}$ in terms of an initial propositional language $\mathcal{L}$; for simplicity, we consider only the finitary case, that is, $\mathcal{L}$ contains only a finite set $\mathcal{P}_\mathcal{L}$ of variables.[3] In addition, we need to introduce the action terms corresponding to the update formulas. Once the terms of the language are fixed, we describe the translation process. Finally, we encode the update problem as a causal theory.

Situation and action terms.

- *S_0 is a (closed) situation term, intuitively denoting the initial situation.*
- *If S is a closed situation term, and μ is a satisfiable formula of the propositional language $\mathcal{L}$, then A^S_μ is a closed action term, with the intuitive meaning described.*

[3] But see the concluding section.

- *If S is a situation term and A is an action term then $result(A, S)$ is a situation term.*

The set of fluent terms is obtained quite directly from the propositional language $\mathcal{L}$, which we assume is obtained by closing its set $\mathcal{P}_{\mathcal{L}}$ of propositional symbols under negation ($\neg$) and disjunction ($\vee$).

Fluent terms. *The set $\mathcal{P}$ of (closed) fluent terms of $\mathcal{L}_{\mathcal{S}}$ is defined as follows:*

- *p is a (primitive) fluent term if $p \in \mathcal{P}_{\mathcal{L}}$*
- *If ψ is a fluent term, then so is $not(\psi)$.*
- *If ψ and ϕ are fluent terms, then so is $or(\psi, \phi)$.*

Non-primitive fluent terms are required to satisfy the following axioms:

$$\forall p, s.\ holds(not(p), s) \equiv \neg holds(p, s);$$
$$\forall p, q, s.\ holds(or(p, q), s) \equiv holds(p, s) \vee holds(q, s).$$

After fixing the terms of the language, we translate the database into situation calculus. To translate a propositional formula ψ, we have to think of it as holding at a particular point of time or situation. This is quite natural, since the database is subject to change through updates, but forces us to make a choice about what is the situation in which ψ holds, since this needs to be expressed in $\mathcal{L}_{\mathcal{S}}$. Thus, rather than defining 'the' translation of a formula ψ into the language $\mathcal{L}_{\mathcal{S}}$, we define the translation of ψ at a situation S, denoted by ψ^S. The easiest way to do it is to first map every propositional formula $\psi \in \mathcal{L}$ into a fluent term ψ^t as follows:

- $p^t = p$ *if* $p \in \mathcal{P}_{\mathcal{L}}$
- $(\neg\psi)^t = not(\psi^t)$
- $(\psi \vee \phi)^t = or(\psi^t, \phi^t)$

and then define ψ^S simply by $\psi^S = holds(\psi^t, S)$.

After the translation is defined, we can present the theory of action that we use to encode update. The causal theory T for the actions we have introduced is the set of instantiations of the axiom schema:

$$holds(\mu^t, result(A_\mu^S, S)),$$

where μ^t is the fluent term corresponding to a satisfiable propositional formula μ, and A^S_μ and S are closed terms of the appropriate sorts.[4]

We will now apply the circumscription policy of the previous section, after an important generalization. As it is clear from the definition of the abnormality predicate ab, the role of the $frame$ predicate is to select a set of fluents for which changes must be minimized, or in other words, a set of fluents whose value 'persists by default'. But the set of facts that we think likely to persist in any given situation might well depend on the state of the world at the time. To allow for this dependence, we need to add a situation argument to the predicate $frame$. For greater clarity, we will in fact replace $frame$ by a new binary predicate $persistent$ with first argument of sort fluent and second argument of sort situation, and redefine the abnormality predicate $ab(p, a, s)$ to be an abbreviation for:

$$persistent(p, s) \wedge (holds(p, s) \equiv \neg holds(p, result(a, s))).$$

We assume that knowledge about which facts should be treated as persistent in each state and situation is encoded in our theory of action by means of some *persistence axiom* (P) $\in \mathcal{L}_\mathcal{S}$, which we leave as a parameter in the theory of action. This axiom might, as a very simple example, declare every primitive fluent (and no other fluent) to be persistent at every situation. More generally, the persistence axiom might be thought of as associating to each possible state of the world at a given time a (possibly partial) specification of which fluents should be treated as persistent in that state. We want to constrain the choice of (P) as little as possible, yet some requirements must be imposed to ensure that it plays its intended role. These requirements are informally presented at the end of this section; their formal statement is deferred to appendix A.

Finally, in order to apply the circumscription policy, we add suitable unique names axioms for fluents (N1$'$) and situations (N2$'$), and consider the extended language $\mathcal{L}'_\mathcal{S}$ obtained from $\mathcal{L}_\mathcal{S}$ by adding two new binary predicates $holds'$ and $persistent'$, with same sorts as their respective namesakes.

Let then $W(s)$ be the set of formulas with free variable s:

$$T \cup \{N1', N2', P, \forall p.\ holds(p, s) \equiv holds'(p, s),$$
$$\forall p.\ persistent(p, s) \equiv persistent'(p, s)\}.$$

[4] Nothing in our results depends on actions being parametrized by situations. But as presented here, if S and S' denote different situations, there is no axiom characterizing the effect of A^S_μ in situation S', resulting in an 'update' which leaves the database unchanged. The intuition behind this parametrization is simply that the same fact might be brought about by different actions in different situations.

Our final theory $Comp(T)$ is the union over all closed terms A and S of the appropriate sorts of $Circum(W(S); ab(p, a/A, s/S); holds, persistent)$, *i.e.* the circumscription of $ab(p, a/A, s/S)$ in $W(S)$ with *holds* and *persistent* allowed to vary.[5]

Suppose now that we are given an initial propositional database ψ. Let ψ^{S_0} be the translation of ψ into situation calculus as holding at S_0. We can then take result of updating ψ with μ as the set of consequences about the situation $result(A_\mu^{S_0}, S_0)$ entailed by $Comp(T) \cup \psi^{S_0}$. To capture this, let

$$\begin{aligned} U(\psi, <\mu>) \quad = \quad & \{\varphi \mid Comp(T) \models \psi^{S_0} \supset \varphi \text{ and } \varphi \text{ contains } result(A_\mu^{S_0}, S_0) \\ & \text{as only situation term}\} \end{aligned}$$

Define now the update operator $\diamond$, for any satisfiable formula μ and database ψ, as follows:

Definition 1 $\psi \diamond \mu \models \phi$ *iff* $holds(\phi^t, result(A_\mu^{S_0}, S_0)) \in U(\psi, <\mu>)$.

Iterated updates can be similarly defined in terms of the result of executing the corresponding sequence of actions. Formally, for any formula ψ and sequence of satisfiable formulas $\mu_1, \ldots, \mu_n$, let

$$\begin{aligned} U(\psi, <\mu_1, \ldots, \mu_n>) \quad = \quad & \{\varphi \mid Comp(T) \models \psi^{S_0} \supset \varphi, \text{ and } \varphi \text{ contains } S_n \text{ as only situation} \\ & \text{term, where for } n \geq j > 0,\ S_j = result(A_{\mu_j}^{S_{j-1}}, S_{j-1})\}. \end{aligned}$$

(As a special case, define $U(\psi, <>) = \psi^{S_0}$.) The following definition generalizes definition 1 to sequences of updates. For any non empty sequence of satisfiable formulas $\mu_1, \ldots, \mu_n$, and any database ψ, define:

Definition 2 $(\ldots((\psi \diamond \mu_1) \diamond \mu_2) \ldots) \diamond \mu_n \models \phi$ *iff* $holds(\phi^t, S_n)) \in U(\psi, <\mu_1, \ldots, \mu_n>)$.

As said, these definitions assume the persistence axiom (P) to satisfy certain assumptions. Informally, we can describe them as follows. For any closed situation term S, a *state* of S is a complete specification of *holds* at S; and an S-history is a complete specification of the extension of both *holds* and *persistent* from S_0 up to S consistent with the theory $T \cup \{N1', N2', P\}$. The assumptions are:

[5] As the reader will have noticed, the role of $persistent'$ is analogous to that of $holds'$. It allows us to keep the extension of *persistent* fixed at a situation S while allowing it to vary at any other point. This is needed because the result of the update should be independent of what fluents become persistent *as a result* of the update.

1. Only named fluents are in the extension of *persistent*; this can be achieved, for example, by adding a domain closure axiom.

2. The theory $T \cup \{N1', N2', P\}$ is a conservative extension of $T \cup \{N1', N2'\}$ with respect to the language obtained from $\mathcal{L}_{\mathcal{S}}$ by removing the predicate *persistent*.

3. Any state of $result(A, S)$ consistent with the causal axiom for A (if any) can be used to extend any S-history into a $result(A, S)$-history.

4. A set of S-histories such that its elements agree with each other on the situations in which they overlap is consistent with $T \cup \{N1', N2', P\}$.

These assumptions we view as *mandatory*, well-behavedness requirements on (P), and leave ample room for different choices of persistence axiom, as might best suit a given application. Conditions 1 and 2 need no explanation. Conditions 3 and 4 are not needed in the presence of the TI condition (introduced in the next section), and are thus inessential for the derivation of a correspondence between theories of action and KM semantics. They are only needed when the set of persistent fluents can be different in two states which are identical except for their situation argument. Roughly speaking, condition 3 ensures that the history of *persistent* does not make a possible future state impossible; condition 4, that there is no interaction among different time branches.

The formal statement of these mandatory conditions on (P) is notationally cumbersome, and is thus left to section A.1 in the first appendix. All that needs to be emphasized at this point is that an update operator is not completely defined in our construction unless some *specific* persistence axiom is provided, and that this axiom should be provided in the simplest form compatible with the requirements of the domain of application. For example, the update operator determined by the previous definitions when we use the example axiom suggested earlier (namely, persistent fluents = primitive fluents, for every situation) exactly captures the PMA update operator [Winslett, 1988], as can easily verified with the results in appendix A.

5 From theories of action to KM semantics, and back

As said, then, the definition of update we have provided depends on the persistence axiom (P), which specifies the set of persistent fluents at each particular state. Definition 2 does not therefore characterize any specific operator, but a family of them. Interesting subfamilies can be obtained by placing additional restrictions on the persistence axiom, as we do in this section. These new restrictions are *optional*, not required by our construction. However, they turn out to be crucial to our mapping between KM semantics and theories of action.

We will consider three such restrictions. The first one (SDP, for "state determines persistent fluents") requires that complete knowledge of the state of the world at any given time be sufficient to uniquely determine the set of persistent facts at that time. The second one (PDS, for "persistent fluents determine state") ensures that the values of the persistent fluents (at a situation S, in a model of $W(S)$) are sufficient to completely characterize a state. Finally, the "time independence" condition (TI) requires the set of persistent fluents to depend only on the current state of the world, so that identical states at different times determine the same set of persistent fluents. Formally:

Definition 3 (SDP condition) *A theory of action satisfies the SDP condition iff the persistence axiom is such that for any situation term S, any state R of S consistent with $W(S)$, and any fluent term θ, either $W(S), R \models persistent(\theta, S)$ or $W(S), R \models \neg persistent(\theta, S)$.*

Definition 4 (PDS condition) [6] *A theory of action satisfies the PDS condition iff for any situation term S, any two states R_1 and R_2 of S, and any set of fluent terms P such that $\{persistent(\theta, S) \mid \theta \in P\} \cup \{\neg persistent(\theta, S) \mid \theta \in \mathcal{P} \setminus P\}$ is consistent with $W(S) \cup R_1$: if R_1 and R_2 agree on the value of every $\theta \in P$ then $R_1 = R_2$.*

Definition 5 (TI condition) *A theory of action satisfies the TI condition if the persistence axiom has the form $\forall s.\ P(s)$, where $P(s)$ contains no situation term other than s.*

Before we consider the effects of these restrictions, it is natural to ask about the properties of our update operators in the general case, without any of these optional conditions on the persistence axiom. The answer is provided by the next theorem. In it, like in all the results that follow, we say that a postulate is satisfied iff it is satisfied *throughout the evolution of the database.* That is, individual updates within sequences of updates also satisfy the given postulates with respect to the previous state of the database.

Theorem 2 *The update operator $\diamond$ satisfies (U1), (U3), (U5), (U6) and (U8), but does not in general satisfy (U2), (U4) or (U7).*

[6] This condition generalizes the "frame completeness condition" of [del Val and Shoham, 1992].

Proof. (Sketch) The key step in the proof are lemmas A.3 and A.4 in appendix A, that show that the results of update can be computed in terms of the states that are minimal under any one of a set of partial orders over states induced by the circumscription policy. The proof is then similar to the proof of the KM representation theorem (theorem 1 in this paper). $\Box$

The next theorem presupposes some of the results of the next section. We have chosen to present this result first, in order to keep separate in our presentation the establishment of a tight correspondence with KM semantics from the identification of circumstances in which we need to move beyond the KM framework.

Theorem 3 *Suppose the SDP, PDS and TI conditions are satisfied. Then the update operator $\diamond$ satisfies postulates (U1)–(U8).*

Proof. From theorems 2, 5, 6, and 7. SDP singles out a single partial order in lemma A.4, PDS ensures than an analogous to the faifthfulness condition holds, and TI ensures satisfaction of (U4). $\Box$

The converse of theorem 3 also holds: Every KM operator can be captured in our framework. Taken together, these results provide a representation theorem for KM semantics in terms of theories of action.

Theorem 4 *For any update operator $\diamond'$ satisfying (U1)-(U8) there exists an operator $\diamond$ based on definition 2 and satisfying the PDS, SDP and TI conditions such that for any ψ and sequence of* satisfiable *formulas $\mu_1, \ldots, \mu_n$:*

$$Mod((\ldots((\psi \diamond \mu_1) \diamond \mu_2) \ldots) \diamond \mu_n) = Mod((\ldots((\psi \diamond' \mu_1) \diamond' \mu_2) \ldots) \diamond' \mu_n).$$

Proof. (Construction.) Since $\diamond$ satisfies (U1)-(U8), there exists a faithful update assignment of a partial order $\leq_M$ to every interpretation M such that $\diamond$ can be defined in terms of the representation theorem 1. The crucial part on the proof is providing an adequate persistence axiom. Let $\mathcal{W}$ be the set of all propositional interpretations, and for any $W \subseteq \mathcal{W}$, choose one formula θ_W such that $Mod(\theta_W) = W$. For any propositional model M, let $\Sigma_M = \{\theta_W \mid W = \{I \mid I \leq_M J\}$ for some $J \in \mathcal{W}\}$, and let $M(s)$ be a finite axiomatization of the state of situation s such that for any $\theta \in \mathcal{L}$, $M \models \theta$ iff $M(s) \models holds(\theta^t, s)$. Let $\mathcal{M}_s$ be the (finite) set of all such finitely axiomatized states. For any $M(s) \in \mathcal{M}_s$, let $P_M(s)$ be a finite first order axiomatization of '$persistent(\theta^t, s)$ iff $\theta \in \Sigma_M$'. (The finite vocabulary of $\mathcal{L}$ guarantees finitariness in all these cases.) Then the persistence axiom $\forall s.\ P(s)$ is:

$$\forall s. \bigwedge_{M(s) \in \mathcal{M}_s} M(s) \supset P_M(s).$$

The TI condition is clearly satisfied by this persistence axiom, and it is easy to see that the SDP and PDS conditions are satisfied as well, the latter because of the faithfulness of $\leq_M$. The details are provided in the appendix. $\Box$

Thus, in the presence of the SDP, PDS and TI conditions, our 'update semantics' is equivalent to Katsuno and Mendelzon's. In so far as our approach is based on a solution to the frame and ramification problem and makes explicit the temporal evolution of the database, we believe our approach provides a solid foundation to the KM semantics for applications in which the conditions on the persistence axiom are reasonable. However, as we have already anticipated and will argue in detail in the next section, we do *not* believe these conditions are always reasonable.

6 Beyond KM semantics

So far we have achieved half the goal of this paper, with theorems 3 and 4 providing a very tight correspondence between theories of action and KM semantics. As said, though, since our intuition is that theories of action should form the basis for update, we believe that further insight can be achieved by relaxing some of the assumptions embodied by the theories of action we have considered. The KM postulates identify an interesting class of operators, but there are important operators outside this class, as we argue next.

6.1. (U2), PDS, and indeterministic actions. How reasonable is (U2)? Consider the following example from [Goldszmidt and Pearl, 1992]. Suppose we order a robot to paint a wall in blue or white. If the wall is initially white, then (U2) entails that it will remain white after the action. If the robot has no way of knowing the original color, however, there is no reason why the color of the wall should persist after the action. For another example, suppose we know that Fred decided today whether to leave his current job to accept another offer, but we do not know his specific decision; according to (U2), the result of updating the database *current-job* with *current-job* $\vee$ *new-job* should be that Fred rejects the offer!

In both cases, the problem arises because (U2) requires the database to remain unchanged when a disjunctive update arrives and the disjunction is already satisfied by the database.

Both Goldszmidt and Pearl, and Katsuno and Mendelzon, have a possible solution to this type of problem (the latter through the operation of "erasure" [Katsuno and Mendelzon, 1991]). In our view, the key issue is whether certain facts should persist or not when disjunctive updates arrive; our construction allows for an explicit axiomatization of default persistence, and thus we suggest that it provides the greatest flexibility in handling this problem.

The connection between (U2) and assumptions about persistence is stated in the next theorem. Here, the *weak PDS condition* means the restriction of PDS to R_1, R_2, and P such that: $P \in CP_\psi(R_1)$, and for some ψ, satisfiable $\mu_1, \ldots, \mu_n$, $R_1 \in States(U(\psi, <\mu_1, \ldots, \mu_n>))$, but $R_2 \notin States(U(\psi, <\mu_1, \ldots, \mu_n>))$.

Theorem 5

1. *If the PDS condition is satisfied then $\diamond$ satisfies (U2).*
2. *If $\diamond$ satisfies (U2), then the weak PDS condition is satisfied.*
3. *If the TI condition is satisfied and $\diamond$ satisfies (U2) then the PDS condition is satisfied.*

In particular, therefore, in the presence of TI the PDS condition provides a necessary and sufficient condition for (U2) to be satisfied.

We remark that the problem with (U2) can be seen as an instance of a more general and fundamental problem, namely: it will not always be reasonable to assume that the world changes minimally as a result of an update. If we learn that Fred broke his left or right arm, it is reasonable to assume that he broke *just* one arm, not both. On the other hand, if we learn that he rented movie A or movie B at the video store, why should we assume that he rented exactly one of them, rather than both? Similarly, if a robot is ordered to make some room in a workspace, it might, or it might not, be reasonable to assume that it will remove as few objects as possible from the work area. Thus, the assumption of minimal change is not always adequate.

Our approach allows to explicitly state which fluents should play a role in the minimization of changes, so we might expect it to be able to handle both minimal and non minimal change. (In particular, there is the degenerate special case in which the persistence axiom declares every fluent to be *non*-persistent at every situation; clearly, this determines an update operator embodying no principle of minimal change whatsoever.) In general, given an update $p \vee q$, treating p and q as persistent will enforce satisfaction of (U2) iff either p or q already hold, and will force exactly one of p or q to become true otherwise; while treating both p and q as non persistent will result in an updated database which entails $p \vee q$, but nothing stronger than that. Note that a difficulty arises because, for example, treating *current-job* and *new-job* as non-persistent at a situation S entails also that they are non-persistent with respect to *any* update, for example, the update $true$, so that facts become false "without reason". It is straightforward to solve this problem by making default persistence dependent on the update action, and not just on the situation, which in turn requires adding an action argument to $persistent$. The details are outside the scope of this paper.

6.2. (U4), TI, and time varying persistence. The TI condition can also result overrestrictive in forcing the set of persistent fluents at any state to be independent of past states. This assumption is connected to postulate (U4), according to which the update of equivalent databases with equivalent formulas should produce equivalent results. In the timeless framework in which the update problem is formulated in the KM proposal, the postulate appears to encode a principle of syntax independence. If we consider the evolution of the database, however, it is clear that it implies more than that. Suppose the database is represented by a formula ψ_1 at some time 1, is updated with μ to yield a database ψ_2, and after a series of updates the database ends up at time t in a state in which it can be represented by some formula ψ_t which is equivalent to ψ_1; according to (U4), updating ψ_t with μ should result in a database equivalent to ψ_2.

Clearly, this is not always reasonable. Myers and Smith [1988] present a number of examples in which the reasonableness of treating a fact as persistent by default depends on the way in which we came to know that fact. Similarly, it is easy to find examples in which some fact should be treated as persistent or not in virtue of what causes it. For example, the practice of mountain climbing results in a high risk of injury, but the risk will disappear as soon as the practice is given up; whereas the professional practice of tennis might result in a high risk of (elbow) injury long after the practice is quit. Thus, whether "high injury risk" needs to be treated as persistent depends on the circumstances that cause the risk.

Since the results of our circumscriptive policy at any given situation depend only on the set of persistent fluents at that situation, it is easy to drop this assumption as well, while preserving all postulates except (U4). This does *not* mean that updates depend on the syntactic form of the database, only that they depend on the set of persistent facts, which might differ for identical states at different times.

As the next theorem states, there is a very close connection between (U4) and the TI condition. In order to establish this connection, we need a way to compare theories of action with different persistence axioms, so as to filter out "inconsequential" violations of TI. Let us first explicitly include the persistence axiom in our notation for the update function, writing $U_P(\psi, <\mu_1, \ldots, \mu_n>)$ to indicate that the underlying theory of action contains the persistence axiom (P). We say that two theories of action with persistence axioms (P) and (P^*), respectively, are *update-equivalent* iff they are equivalent with respect to update, *i.e.* iff for any propositional ψ, satisfiable propositional $\mu_1, \ldots, \mu_n$, we have:

$$U_P(\psi, <\mu_1, \ldots, \mu_n>) = U_{P^*}(\psi, <\mu_1, \ldots, \mu_n>).$$

The connection between TI and (U4) is then the following:

Theorem 6 *The update operator $\diamond$ satisfies (U4) iff it is based on a theory of action which is update equivalent to a theory of action satisfying the TI condition.*

The basic gist of Myers and Smith's approach to this problem can be easily incorporated in our framework. Myers and Smith use default logic, and propose the use of default frame axioms of the form:

$$\frac{Q(p,s) \wedge holds(p,s) : holds(p, result(a,s))}{holds(p, result(a,s))}$$

Here, $Q(p,s)$ is some complicated precondition (the details of which need not concern us) which depends on the current and past states of the database, and p is a primitive fluent. This default can be read as saying that if $Q(p,s)$ is true, and it is consistent to assume that p remains true after executing the action a, then p remains true. In our terms, this means that whenever $Q(p,s)$ is true, p should be declared *persistent* at s. Assuming, as they appear to do, that $Q(p,s)$ is also necessary for p to be persistent, the spirit of their proposal can be captured by having the persistence axiom in our theory of action entail:

$$\forall s.\ Q(p,s) \equiv persistent(p,s).$$

Notice that because $Q(p,s)$ can be determined by considering only the current and past states of the database, this kind of axiom satisfies the mandatory requirements on the persistence axiom, even though it does not satisfy the TI condition. This is just, therefore, a special case of our construction.

6.3. (U7) and SDP. Finally, (U7). It is easy to find theories of action violating SDP but which define an operator $\diamond$ satisfying most of the other postulates. Intuitively, a (non-trivial) violation of SDP corresponds to situations in which we do not regard our knowledge of the domain to be sufficient to uniquely determine which facts are likely to persist, or in which we perhaps want to consider the effects of alternative assumptions about persistence. This is a perhaps rare but certainly conceivable circumstance, which is not covered by KM. We have:

Theorem 7

1. *If the SDP condition is satisfied then $\diamond$ satisfies (U7).*
2. *If the PDS condition is satisfied and $\diamond$ satisfies (U7) then $\diamond$ is update equivalent to a theory of action satisfying the SDP condition.*

There appears to be some subtle interaction between (U2) and (U7), which is why the PDS condition appears to be needed in order to derive SDP from (U7) in the second part of the theorem.

7 Constraints and the ramification problem

In Lin and Shoham's proposal for reasoning about action, the ramification problem (roughly, the problem of specifying both direct and indirect effects of actions) is solved to a great extent by means of the constraint $\forall s.\ C(s)$. Indirect effects of actions are simply those that follow from the direct effects by using this constraint and the frame axioms (or its non-monotonic equivalent). Similarly, constraints can play a crucial role in the update problem, since there will be often a set of formulas which the database should always satisfy. For this reason, in the context of belief change these constraints are sometimes called "protected formulas," a term that we will later use. [Katsuno and Mendelzon, 1989] postulate, in the context of AGM revision rather than KM-update, that a revision operator under constraints γ, written $\circ_\gamma$, should be defined in terms of a standard revision operator $\circ$ as:

$$\psi \circ_\gamma \mu \equiv \psi \circ (\mu \wedge \gamma).$$

Our framework allows us, once again, to *prove* that the analogous approach for update under constraints is correct; there is no need to 'postulate' it.

Constraints and ramifications are easily handled in our framework, as they are automatically handled by the same overall minimization principle. Suppose we are given, in addition to the initial database ψ, a constraint γ. We assume that ψ satisfies the constraint, *i.e.* that $\psi \models \gamma$. Remove from the language all terms A_μ^S such $\mu \wedge \gamma$ is unsatisfiable, and let T' be the theory obtained by restricting T to the new language and adding the constraint $\forall s.\ \gamma^s$. Let W' be the formula obtained by replacing T by T' in W, and similarly for $Comp(T')$. Denote by $U_\gamma(\psi, < \mu_1, \ldots, \mu_n >)$ the result of replacing $Comp(T)$ by $Comp(T')$ in the definition of $U(\psi, < \mu_1, \ldots, \mu_n >)$, and define the update operator $\diamond_\gamma$ under constraints γ analogously.[7] Update under constraints can then be characterized in a very similar way as in the case without constraints (see in particular lemma A.5), and we can prove:

Theorem 8 *If* $\psi \models \gamma$ *then* $\psi \diamond_\gamma \mu \equiv \psi \diamond (\mu \wedge \gamma)$

Notice however that constraints add an additional degree of freedom in the design of update operators satisfying (U2), by making it easier for the PDS condition to be satisfied. Consider the following example, slightly adapted to the context of update from [Lifschitz, 1990], and which also illustrates the interaction between the choice of persistent fluents and constraints. Suppose a light is on exactly when two switches are in the same position, either both on or both off. We can write this as the constraint:

$$\forall s.\ holds(On, s) \equiv (holds(Sw_1, s) \equiv holds(Sw_2, s)).$$

Which fluents should be treated as persistent? Suppose On, Sw_1 and Sw_2 are persistent at S_0 (*i.e.* $P \models \forall \sigma.\ persistent(\sigma, S_0) \equiv (\sigma = On \vee \sigma = Sw_1 \vee \sigma = Sw_2)$). Then turning a switch from an initial situation in which both are in the same position does not yield the expected results:

$$Comp(T) \not\models holds(Sw_1, S_0) \wedge holds(Sw_2, S_0) \supset \neg holds(On, result(A_{\neg Sw_1}^{S_0}, S_0)).$$

In some models of $Comp(T)$ the light does go off, as expected, but in others it is the second switch Sw_2 which changes its position while the light remains on. The desired result can be

[7]These substitutions should also be made in the definition of S-histories, admissible S-histories, CP, CP_ψ, and in the requirements on the persistence axiom in appendix A. In these definitions, whenever the expression "R satisfies the causal axiom, if any, for S" is used, replace it by "R satisfies γ^S and the causal axiom, if any, for S".

obtained, however, by simply excluding On from the set of persistent fluents at S_0. Note that the resulting theory satisfies PDS.

8 Discussion

By analytically deriving the KM-postulates for update from a rigorous theory of action, we have linked two previously unrelated fields of research and provided a foundation for the KM update proposal, which at the same time allows us to delineate the limits of its applicability. Not only our approach to belief update subsumes KM semantics; we have also argued that by focusing on the default persistence of facts and on the frame and ramification problems, our approach can provide insights on database update that cannot be easily gathered from theorem 1 and KM semantics.

We have restricted our attention in this paper to update in a finitary propositional language, in order to analyze the connections with the KM proposal. Though the postulates can be used for any logic, it is only for a finitary propositional logic that we have a representation theorem characterizing KM semantics. Obviously, our proposal is not restricted to propositional languages; nothing prevents us from introducing a distinct action for each predicate calculus sentence, just as we have done with propositional formulas. The question is whether the technical results presented so far would still hold. There are two separate questions: whether the extension of our construction to these cases will preserve the same postulates, and whether arbitrary operators satisfying the postulates in these more general cases can be captured in our construction.

As for the first question, the infinitary case can be handled by requiring that the circumscriptive orderings are smooth.[8] It can then be shown that the theory of action satisfies the same postulates (in their infinitary version) as in the finitary case. Similarly, it is easily shown that a FOL update operator defined as in theorem 1 satisfies the same postulates as in the propositional case. Though we have not worked out the details, there is no reason to doubt that our results still hold after extension to FOL, since satisfaction of the postulates hinges mostly on the properties of partial orders.

The second question, whether arbitrary (infinitary or FOL) operators satisfying the

[8] See lemma B.2 in the appendix for a definition of smoothness. The smoothness of the circumscriptive orderings, in particular, is a sufficient condition for the satisfiability of the circumscription [Etherington *et al.*, 1985; Lifschitz, 1986], and is easily shown to be sufficient for the key lemma B.2 to hold in the infinitary case.

postulates can be captured in the theory of action, is trickier. The construction used in theorem 4 can be applied to any update operator defined as in theorem 1, provided that for every two interpretations M and N, the set of models $K = \{I \mid I \leq_M N\}$ is finitely axiomatizable. However, it is an open question whether infinitary or FOL operators satisfying the postulates can always be defined as in theorem 1 or analogously.

In the category of related work, a formal connection between non-monotonic reasoning and update was first studied in [Winslett, 1989], for Winslett's update operator, but with no relation to theories of action. She has also suggested in [Winslett, 1988] to use update for reasoning about action; our results can be seen as providing formal support to this idea.

Independently, [Reiter, 1992b; Reiter, 1992a] has proposed an account of database update in terms of recent proposals for solving the frame problem [Schubert, 1990]. Reiter's proposal requires the database designer to specify a set of predefined "update transactions" (actions), which restrict the set of updates allowable to the user. For deterministic actions and in the absence of constraints, the main case he considers, his approach appears to be equivalent to that of [Lin and Shoham, 1991].[9] He provides some suggestive examples of how to handle non deterministic actions in the absence of constraints, but leaves the problem somewhat unexplored. But it is the introduction of constraints that appears to pose the most serious obstacle for Reiter's approach, though it is too early to tell whether it is a fundamental one. The approach requires providing, for each fluent, an axiom specifying all conceivable circumstances in which the value of the fluent can change as a result of an action. In the presence of constraints, the effects of actions might depend in a potentially infinite number of ways on the particular details of the situation in which the action occurs; attempting to explicitly axiomatize them away appears to be in direct conflict, in our opinion, with the goal of attaining a *concise* representation of ramifications. Finally, we remark that in our framework arbitrary updates and action terms denoting "real" events can be straightforwardly combined, which in our opinion invalidates some of the criticisms directed by Reiter to the KM approach in [Reiter, 1992a], at least with respect to our construction.

In [del Val and Shoham, 1992], we presented a less general version of this approach, in which the set of persistent fluents was not allowed to vary. As a result, we were only able to prove that our construction satisfied the KM postulates, yet we still claimed that the "postulates need not be postulated at all," but can instead be derived analytically. This claim could be subject to a number of criticisms, some of which were articulated by

[9] Reiter, personal communication.

[Goldszmidt and Pearl, 1992] and anonymous referees. First, our construction satisfied the postulates, but so do others, including the one presented by Goldszmidt and Pearl. Second, it could be argued that the generality of the representation theorem for KM update provided a sufficient justification for this approach, one, moreover, that appeared to be clearly more general than ours and subsume it. These concerns were all reasonable, but are in our view addressed by the results presented here. We can now capture every KM update operator,[10] and also show that reasonable update operators fall outside this family. This provides further justification for the reasonableness of our encoding in this particular theory of action, and shows the limitations of the KM representation theorem in providing a justification for the whole approach.

There are other issues that suggest themselves for further work. For example, updates can be seen as providing the *expected* changes in the domain as a result of a change. In this sense, computing an update is an exercise in temporal prediction/projection. Will the CS building change its color after Smith enrolls in CS205? Well, of course, the color "should persist" after this update in (*almost*) every conceivable circumstance. Will Smith be able to afford her education after this enrollment? This is not so clear; it might depend *e.g.* on whether the enrollment changes her tuition category. Assumptions about default persistence will often help us to come up with correct predictions, but there is clearly no guarantee that this will always be the case. The question is then, how should we deal with the case in which these predictions turn out to be wrong? Should we use AGM revision, or is there a more promising approach based on research in reasoning about action? Notice also that the adequacy of this exercise in temporal projection will also depend on the accuracy and completeness of the information received. If recent rain made the grass and the street wet, but only the former fact is observed and reported to the database, then the updated database will contain incorrect or incomplete information. Similarly, in a rapidly changing world, *e.g.* if a ball is rolling down a hill, some mechanisms must be devised so that the database keeps a reasonably accurate model of the state of the world. One could, for example, keep updating the database with the constantly changing position of the ball; more practically, one might provide some way to predict changes in some fundamental magnitudes, such as whether the ball's velocity is increasing or decreasing, and request an update each time a transition is predicted. Update facilities might therefore have to be integrated within a more general

[10]This includes, of course, Goldszmit and Pearl's operator, at least in so far as disjunctive updates are handled in their framework as required for KM compliance.

problem solver.

Theories of action provide an excellent framework in which to deal in a principled way with the persistence of facts, a topic which lies at the heart of the update problem. An interesting open problem is the definition of parallel updates on the basis of a treatment of parallel actions.[11] In addition, the framework of propositional update loses some of the information encoded by the non-monotonic approach for reasoning about action, specially information about the past. An interesting issue is whether hybrid representations could be defined to benefit from this information without incurring in the full representational cost of keeping the whole history of the database encoded in situation calculus.

References

[Alchourrón *et al.*, 1985] Alchourrón, Carlos E.; Gärdenfors, Peter; and Makinson, David 1985. On the logic of theory change: Partial meet functions for contraction and revision. *Journal of Symbolic Logic* 50:510–530.

[del Val and Shoham, 1992] del Val, Alvaro and Shoham, Yoav 1992. Deriving properties of belief update from theories of action. In *Proceedings of the Tenth Conference of the AAAI.*

[Etherington *et al.*, 1985] Etherington, D.W.; Mercer, R.E.; and Reiter, Raymond 1985. On the adequacy of predicate circumscription for closed world reasoning. *Computational Intelligence* 1:11–15.

[Gärdenfors, 1988] Gärdenfors, Peter 1988. *Knowledge in Flux.* The MIT Press.

[Goldszmidt and Pearl, 1992] Goldszmidt, Moisés and Pearl, Judea 1992. Rank-based systems: A simple approach to belief revision, belief update and reasoning about evidence and actions. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning (KR'92).*

[Katsuno and Mendelzon, 1989] Katsuno, Hirofumi and Mendelzon, Alberto O. 1989. A unified view of propositional knowledge base updates. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence.*

[11] Cf. [Lin and Shoham, 1992].

[Katsuno and Mendelzon, 1991] Katsuno, Hirofumi and Mendelzon, Alberto O. 1991. On the difference between updating a knowledge database and revising it. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning (KR'91)*.

[Lifschitz, 1986] Lifschitz, Vladimir 1986. On the satisfiability of circumscription. *Artificial Intelligence* 28:17–27.

[Lifschitz, 1990] Lifschitz, Vladimir 1990. Frames in the space of situations. *Artificial Intelligence* 46:365–376.

[Lin and Shoham, 1991] Lin, Fanghzen and Shoham, Yoav 1991. Provably correct theories of action (preliminary report). In *Proceedings of the Ninth Conference of the AAAI*.

[Lin and Shoham, 1992] Lin, Fanghzen and Shoham, Yoav 1992. Concurrent actions in situation calculus. In *Proceedings of the Tenth Conference of the AAAI*.

[Myers and Smith, 1988] Myers, Karen L. and Smith, David E. 1988. The persistence of derived information. In *Proceedings of the Seventh Conference of the AAAI*.

[Reiter, 1992a] Reiter, Raymond 1992a. On specifying database updates. Technical Report KRR-TR-92-5, Computer Science Department, University of Toronto.

[Reiter, 1992b] Reiter, Raymond R. 1992b. On formalizing database updates: Preliminary report. In *Proceedings of the Third International Conference on Extending Database Technology*.

[Schubert, 1990] Schubert, L. K. 1990. Monotonic solution of the frame problem in situation calculus: An efficient method for worlds with fully specified actions. In Kyberg, H. E.; Loui, R. P.; and Carlson, G. N., editors 1990, *Knowledge Representation and Defeasible Reasoning*. Kluwer Academic Press.

[Winslett, 1988] Winslett, Marianne 1988. Reasoning about action using a possible models approach. In *Proceedings of the Seventh Conference of the AAAI*.

[Winslett, 1989] Winslett, Marianne 1989. Sometimes updates are circumscription. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*.

A Appendix: Characterizing updates

A.1 Mandatory requirements on the persistence axiom and notational conventions.

In this subsection we describe the mandatory requirements on the persistence axiom, after introducing some necessary concepts and notation. The next subsection will then lay the formal groundwork for the study of the properties of our construction by providing an alternate characterization of the family of update operators covered by definition 2.

State. *For any situation term S we say that R_S is a* state *of situation S iff there is some set of closed fluent terms F_t such that*

$$R_S = \{holds(\theta, S) \mid \theta \in F_t\} \cup \{\neg holds(\theta, S) \mid \theta \in \mathcal{P} \setminus F_t\},^{12}$$

and R_S is consistent with the axioms for non primitive fluents.

Let $\Gamma(S)$ be a set of formulas with holds as only predicate and S as only situation term. $\mathrm{States}(\Gamma(S))$ *denotes the set of states R of S such that $R \models \varphi$ for every $\varphi \in \Gamma(S)$.*

Finally, we will use the notation $Min(K, \leq)$, for some set of states K (e.g. $K = States(holds(\mu^t, S))$) and any binary relation $\leq$ defined over K, as introduced before theorem 1.

Intuitively, a state of a given situation is a complete specification of the values of all fluents at that situation. We will consider only models of these axioms, and for any such model M and any situation term S, we denote by $M(S)$ the state of S determined by M. Similarly, the set of *States* of a set of situation calculus formulas containing a single situation term corresponds to the set of models of the translation of these formulas into $\mathcal{L}$ (the notation *States*() is introduced by analogy with the notation *Mod*(Γ) for the set of models of a set of formulas $\Gamma \subseteq \mathcal{L}$).

We also want shorthands for specifying facts about *persistent*. Given any set of closed fluent terms P_t and any closed situation term S, we define $Pers(P_t, S) = \{persistent(\theta, S) \mid \theta \in P_t\} \cup \{\neg persistent(\theta, S) \mid \theta \in \mathcal{P} \setminus P_t\}$. Given an interpretation M and a closed situation term S, we define $P_t(M(S)) = \{\theta \in \mathcal{P} \mid M \models persistent(\theta, S)\}$, and $P(M(S)) = Pers(P_t(M(S)), S)$. Thus $P(M(S))$ gives the complete specification of *persistent* determined by M at S.

[12] Recall that $\mathcal{P}$ is the set of closed fluent terms of the language

Next, we need a way to identify the specifications of persistent that are acceptable for any given state. We will allow them to depend only on the previous "history," but not on alternate "branches of time" or on future states. In order to make this precise we need the concept of an S-history, introduced next, which intuitively is a complete specification of *holds* and *persistent* from S_0 up to S which is consistent with the causal axioms and the persistence axiom.

S-history. *For any closed situation term S, an S-history is a set of formulas $H_S = R_0 \cup \ldots \cup R_k \cup P_0 \cup \ldots \cup P_k$ such that:*

- *Each R_i is a state of situation S_i satisfying the causal axiom, if any, "for" S_i,*[13] *where $S_0, \ldots, S_k$ is the set of subterms of S.*
- *For each P_i there exists a set of fluent terms P_i^t such that $P_i = Pers(P_i^t, S_i)$.*
- *$H_S \cup T \cup \{N1', N2', P\}$ is consistent.*

We use similar notation as before to single out parts of an S-history. Given an S-history H, and $S' \in subterms(S)$,[14] $H(S')$ is the state of S' determined by H; $P(H(S'))$ is the complete specification of *persistent* determined by H at S'; and $P_t(H(S'))$ is such that $Pers(P_t(H(S')), S') = P(H(S'))$.

Requirements on the persistence axiom.

1. Only named fluents are in the extension of *persistent*; this can be achieved, for example, by adding a domain closure axiom.
2. The theory $T \cup \{N1', N2', P\}$ is a conservative extension of $T \cup \{N1', N2'\}$ with respect to the language obtained from $\mathcal{L}_\mathcal{S}$ by removing the predicate *persistent*.
3. For every situation term S', action term A_μ^S, state R of the situation $result(A_\mu^S, S')$, and interpretation M of $\mathcal{L}_\mathcal{S}$: if R satisfies the causal axiom, if any, for $result(A_\mu^S, S')$, and $M \in Mod(T \cup \{P, N1', N2'\})$, then there exists an interpretation N of $\mathcal{L}_\mathcal{S}$ with same domains as M and agreeing with M on everything except *ab*, *persistent*,

[13] For any S, if S has the form $result(A_\mu^{S'}, S')$, then the causal axiom "for" S is $holds(\mu^t, S)$; otherwise there is no causal axiom for S.

[14] $subterms(S')$ denotes the set of subterms of the situation term S'.

and *holds*, such that $N \in Mod(T \cup \{P, N1', N2'\} \cup R \cup \bigcup_{S'' \in subterms(S')}(M(S'') \cup P(M(S''))))$.

4. For any set of closed situation terms $\{S_1, \ldots, S_n\}$ and any set $\mathcal{H} = \{H_1, \ldots, H_n\}$ such that each H_i is an S_i-history, if $\bigcup \mathcal{H}$ is consistent then $\bigcup \mathcal{H} \cup T \cup \{N1', N2', P\}$ is consistent.

The reader is referred to section 4 for an intuitive explanation of these requirements.

A.2 Characterizing updates

It will later be convenient, in order to establish the properties of our update operators, to have an alternative characterization of the operators determined by our construction, in a format more suitable for comparison with KM semantics. This is the goal of this subsection. The characterization is obtained by mapping the results of the circumscription policy into orderings over *states*, orderings used in very much the same way as in the KM representation theorem (theorem 1 in this paper). Details follow.

The circumscription of *ab* for each situation and action results in a set of partial orders $\leq_{ab, A^S_\mu, S'}$ over the interpretations of $\mathcal{L}'_{\mathcal{S}}$, such that $M \leq_{ab, A^S_\mu, S'} N$ iff M and N have the same domains and agree on everything except *holds*, *persistent* and *ab*, and the extension of $ab(p, a/A^S_\mu, s/S')$ in M is a subset of its extension in N. In order to keep the correspondence with the propositional case as close as possible, however, we choose to characterize the models of $Comp(T)$ in terms of a different set of orderings. Specifically, we will use orderings over states rather than over interpretations. The intuition here is that the results of propositional update should only depend on the theory and the update formula, which in our framework means that it should only depend on the immediately preceding state and the action corresponding to the update formula (and, of course, on the set of persistent fluents at that state).

Definition 6 *Let R and T be states of the situation $result(A^S_\mu, S')$, let M_S be a state of the situation S', and let F be a set of fluent terms. We say that $R \leq^F_{M_S} T$ iff $Diff_F(R, M_S) \subseteq Diff_F(T, M_S)$, where for any state U of $result(A^S_\mu, S')$*

$$Diff_F(U, M_S) = \{\theta \in F \mid U \models \neg holds(\theta, result(A^S_\mu, S')) \text{ iff } M_S \models holds(\theta, S')\}.$$

The following lemma tells us the sense in which these orderings capture the result of the circumscription (using the notation introduced in the definition of *state* in the previous subsection).

Lemma A.1 *For every $M \in Mod(\forall s.W(s))$, $M \in Mod(Comp(T))$ iff for every situation term S' and every action term A_μ^S, either $M \models S \neq S'$ and $M(result(A_\mu^S, S'))$ and $M(S')$ assign the same value to each $\theta \in P_t(M(S'))$, or $M \models S = S'$ and $M(result(A_\mu^S, S')) \in Min(States(holds(\mu^t, result(A_\mu^S, S'))), \leq_{M(S')}^{P_t(M(S'))})$.*

The set of persistent fluents at each state is captured in this lemma using $P_t(M(S'))$, *i.e.* by direct appeal to models M of the whole theory of action, models which somewhat irrelevantly contain information about future states of the database. More conveniently, the function $CP(R, H)$ defined below identifies the set of specifications of *persistent* that can be consistently associated with the state R in view of the previous history H. Define

$$CP(R, H) = \{P_t \mid (H \cup R \cup Pers(P_t, S^*) \cup T \cup \{N1', N2', P\}) \textit{ is consistent}\},$$

where: either R is a state of S_0 and $H = \emptyset$, in which case $S^* = S_0$; or H is an S-history for some S, and R is a state of the situation $S^* = result(A_\mu^{S'}, S)$ for some action term $A_\mu^{S'}$, which satisfies the causal axiom, if any, for S^*.

As a reassuring lemma, we have:

Lemma A.2 *$CP(R, H) \neq \emptyset$, whenever it is defined.*

More importantly, the next lemma draws us pretty close to the representation theorem of Katsuno and Mendelzon for non-iterated updates.

Lemma A.3

$$States(U(\psi, <\mu>)) = \bigcup_{\substack{F \in CP(M_S, \emptyset) \\ M_S \in States(\psi^{S_0})}} Min(States(holds(\mu^t, result(A_\mu^{S_0}, S_0))), \leq_{M_S}^F)$$

As in the representation theorem for propositional update, this can be seen as selecting for each state of the original theory (for each model, in the propositional case) the set of closest states (models) satisfying the update formula. There are important differences, though. Specifically, lemma A.3 allows a degree of indeterminacy in the measure of closeness; has no analogue to the "faithfulness condition" of theorem 1; and though this can only be seen when we consider iterated updates, in our construction the measure of closeness can be different for identical states at different times. As we will see, each of these differences is reflected in one specific postulate, and each is closely related to the optional requirements on the persistence axiom stated at the beginning of next section.

Iterated updates can be characterized in a similar way as non-iterated updates. Specifically, if a sequence of updates has transformed the initial database ψ into some other database ψ' at situation S, the only states on which the result of a subsequent update depends are those states of S consistent with ψ'. We still have the specification of $persistent$ at those states as an additional parameter, and as remarked at the end of section 4, we allow this specification to depend on previous states, not just on states of S. The reasons behind this choice, in addition to generality, will become apparent in section 6. Since we allow the set of persistent fluents to vary for identical states at different times (in the absence of the TI condition, defined below), it seems reasonable to allow for this variation to depend in some way on what preceded the state at each time.

The set of specifications of persistent fluents which are acceptable at a given state can be defined using a set valued function CP_ψ on states, such that $P_t \in CP_\psi(R)$, for a state R, intuitively means that there exists some history of the world consistent with the evolution of the database in which it is consistent to take the fluent terms in P_t as persistent at R. The evolution of the database is captured in terms of the notion of admissible S-histories. Intuitively, this a history of the world up to S consistent with the evolution of the database, as determined by the circumscription. Formally, an S-history H is *admissible* iff for every term $result(A_\mu^{S'}, S') \in subterms(S)$, $H(result(A_\mu^{S'}, S')) \in Min(States(holds(\mu^t, result(A_\mu^{S'}, S'))), \leq_{H(S')}^{P_t(H(S'))})$.

CP$_\psi$. *The function CP_ψ is defined as follows:*

- *If R is a state of S_0 then $CP_\psi(R) = CP(R, \emptyset)$.*
- *If R is a state of $result(A_\mu^S, S)$ then:*
  $CP_\psi(R) = \{P_t \mid P_t \in CP(R, H)$ for some admissible S-history H such that $H \models \psi^{S_0}$ and $R \in Min(States(holds(\mu^t, result(A_\mu^S, S))), \leq_{H(S)}^{P_t(H(S))})\}$.

The appeal to the orderings $\leq_M^F$ in the definition of both CP_ψ and of admissible histories derives from the fact that the evolution of the database is characterized in terms of these orderings.

The next lemma generalizes lemma A.3 in a natural way, which will make easy to establish the properties satisfied by our construction throughout the evolution of the database.

Lemma A.4 *For* $0 < j \leq n$, *let* $S_j = result(A_{\mu_j}^{S_{j-1}}, S_{j-1})$.

$$States(U(\psi, <\mu_1, \ldots, \mu_n>)) = \bigcup_{\substack{F \in CP_\psi(M_S) \\ M_S \in States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>))}} Min(States(holds(\mu_n^t, S_n))), \leq_{M_S}^F)$$

Notice that this lemma also makes formal the claim that the set of persistent fluents at any state depends only on the past, but not on alternate branches of time or future states, since $CP_\psi(M_S)$ can be determined by considering only the past history of the database.

We end this appendix with the case of constraints. After making all the replacements indicated in the text in section 7, it is easy to show that all the lemmas in this appendix still hold after replacing $States(holds(\mu^t, S))$ by $States(holds((\mu \wedge \gamma)^t, S))$ whenever this or a similar expression occurs. In particular, we can prove in exactly the same way as before:

Lemma A.5 *Suppose* $\psi \models \gamma$, *and for* $0 < j \leq n$, *let* $S_j = result(A_{\mu_j}^{S_{j-1}}, S_{j-1})$.

$$States(U_\gamma(\psi, <\mu_1, \ldots, \mu_n>)) = \bigcup_{\substack{F \in CP_\psi(M_S) \\ M_S \in States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>))}} Min(States(holds((\mu_n \wedge \gamma)^t, S_n))), \leq_{M_S}^F)$$

B Appendix: Proofs

We will need a number of auxiliary lemmas, that we will introduce as needed. Their numbers are prefixed with the appendix letter.

Lemma B.1 *For every situation term* S', *action term* A_μ^S, *state* R *of* $result(A_\mu^S, S')$, *and interpretation* M *of* $\mathcal{L}'_\mathcal{S}$: *if* R *satisfies the causal axiom, if any, for* $result(A_\mu^S, S')$, *and* $M \in Mod(W(S'))$, *then there exists* $N \in Mod(W(S') \cup R \cup \bigcup_{S'' \in subterms(S')}(M(S'') \cup P(M(S''))))$ *such that* N *has the same domains as* M *and agrees with* M *on everything except* ab, $persistent$, *and* $holds$.

Proof. Let M^* be the restriction of M to the language $\mathcal{L}_\mathcal{S}$. By the third assumption on the persistence axiom, there exists an interpretation N^* of $\mathcal{L}_\mathcal{S}$ with same domains as M^* and agreeing with M^* on everything except ab, $persistent$, and $holds$, such that $N^* \in Mod(T \cup \{P, N1', N2'\} \cup R \cup \bigcup_{S'' \in subterms(S')}(M(S'') \cup P(M(S''))))$. So let N be the interpretation of the extended language $\mathcal{L}'_S$ which is exactly like N^* in the common portion of the language and such that N interprets *holds'* and *persistent'* identically with M. By construction of N, it suffices to show that $N \in \mathit{Mod}(W(S'))$. Since $S' \in \mathit{subterms}(S')$, $N(S') = M(S')$ and $P(N(S')) = P(M(S'))$, and since N and M interpret *holds'* and *persistent'* identically, N satisfies both $\forall p.\ \mathit{holds}(p, S') \equiv \mathit{holds}'(p, S')$ and $\forall p.\ \mathit{persistent}(p, S') \equiv \mathit{persistent}'(p, S')$. Since $N \in \mathit{Mod}(T \cup \{P, N1', N2'\})$, it follows that $N \in \mathit{Mod}(W(S'))$, as desired. $\Box$

Lemma A.1 *For every $M \in \mathit{Mod}(\forall s.W(s))$, $M \in \mathit{Mod}(\mathit{Comp}(T))$ iff for every situation term S' and action term A^S_μ, either $M \models S \neq S'$ and $M(\mathit{result}(A^S_\mu, S'))$ and $M(S')$ assign the same value to each $p \in P_t(M(S'))$, or $M \models S = S'$ and $M(\mathit{result}(A^S_\mu, S')) \in \mathit{Min}(\mathit{States}(\mathit{holds}(\mu^t, \mathit{result}(A^S_\mu, S'))), \leq^{P_t(M(S'))}_{M(S')})$.*

Proof. ($\Rightarrow$) Suppose $M \in \mathit{Mod}(\mathit{Comp}(T))$ holds but the right hand side does not for some situation S' and action A^S_μ. If $M \models S \neq S'$ then by the hypothesis for indirect proof the extension of $\mathit{ab}(p, a/A^S_\mu, s/S')$ in M is not empty. It follows from lemma B.1 that there exists $N \in \mathit{Mod}(W(S'))$ with same domains as M and agreeing with M on everything except *ab*, *holds* and *persistent* but which assigns an empty extension to $\mathit{ab}(p, a/A^S_\mu, S')$, in contradiction with $M \in \mathit{Mod}(\mathit{Comp}(T))$. Thus, assume $M \models S = S'$. Since $M \in \mathit{Mod}(\forall s.W(s))$, we have that $M(\mathit{result}(A^S_\mu, S')) \in \mathit{States}(\mathit{holds}(\mu^t, \mathit{result}(A^S_\mu, S')))$, so there exists some $R \in \mathit{States}(\mathit{holds}(\mu^t, \mathit{result}(A^S_\mu, S')))$ such that $R <^{P_t(M(S'))}_{M(S')} M(\mathit{result}(A^S_\mu, S'))$. Letting $N_{ab} = \mathit{Diff}_{P_t(M(S'))}(R, M(S'))$ and $M_{ab} = \mathit{Diff}_{P_t(M(S'))}(M(\mathit{result}(A^S_\mu, S')), M(S'))$, this means that $N_{ab} \subset M_{ab}$. Notice that M_{ab} can be rewritten as

$$M_{ab} = \{\theta \in P_t(M(S')) \mid M \models \neg \mathit{holds}(\theta, \mathit{result}(A^S_\mu, S')) \equiv \mathit{holds}(\theta, S')\}.$$

By lemma B.1, there exists $N \in \mathit{Mod}(W(S') \cup R)$ with same domains as M and agreeing with M on everything except *ab*, *holds* and *persistent*. It follows that $R = N(\mathit{result}(A^S_\mu, S'))$, and since $N, M \in \mathit{Mod}(W(S'))$ and they agree on *holds'* and *persistent'*, $N(S') = M(S')$ and $P_t(N(S')) = P_t(M(S'))$. Using the last three equalities for substitution in

$$N_{ab} = \{\theta \in P_t(M(S')) \mid R \models \neg \mathit{holds}(\theta, \mathit{result}(A^S_\mu, S')) \text{ iff } M(S') \models \mathit{holds}(\theta, S')\}$$

and rewriting as we did with M_{ab}, we obtain:

$$N_{ab} = \{\theta \in P_t(N(S')) \mid N \models \neg \mathit{holds}(\theta, \mathit{result}(A^S_\mu, S')) \equiv \mathit{holds}(\theta, S')\}.$$

Because $M \in Mod(W(S'))$, only named fluents are in the extension of $persistent$ in M, so by definition of ab, the extension of $ab(p, a/A_\mu^S, S')$ in M is $\{\theta^* \mid \theta^*$ is in M's fluent domain and there exits $\theta \in \mathcal{P} \cap M_{ab}$ such that M interprets θ as $\theta^*\}$. By the same reasoning, the extension of $ab(p, a/A_\mu^S, S')$ in N is $\{\theta^* \mid \theta^*$ is in N's fluent domain and there exits $\theta \in \mathcal{P} \cap N_{ab}$ such that N interprets θ as $\theta^*\}$. Using these characterizations of the extension of ab in N and M, the unique names axiom for fluents, and the fact that N and M have the same domains and agree on all terms, it then follows from $N_{ab} \subset M_{ab}$ that the extension of $ab(p, a/A_\mu^S, s/S')$ in N is a strict subset of its extension in M. Since $N \in Mod(W(S'))$, this contradicts the hypothesis that $M \in Mod(Comp(T))$.

($\Leftarrow$) Suppose that the right hand side holds but $M \notin Mod(Comp(T))$. Since $M \in Mod(\forall s.\ W(s))$, there exist S', A_μ^S and $N \in Mod(W(S'))$ such that $N <_{ab, A_\mu^S, S'} M$. If $M \models S \neq S'$, this is impossible, for in this case by hypothesis the extension of $ab(p, a/A_\mu^S, s/S')$ in M is empty. So assume $M \models S = S'$. It suffices to show that $N(result(A_\mu^S, S')) <_{M(S')}^{P(M(S'))} M(result(A_\mu^S, S'))$, since this contradicts the hypothesis. It follows from $N <_{ab, A_\mu^S, S'} M$ that N and M agree on everything except ab, $holds$ and $persistent$. Since only named fluents can be in the extension of $persistent$ and N and M agree on all terms, it follows from $N <_{ab, A_\mu^S, S'} M$, the unique names axiom for fluents and the definition of ab that $\{\theta \in P_t(N(S')) \mid N(result(A_\mu^S, S')) \models \neg holds(\theta, result(A_\mu^S, S'))$ iff $N(S') \models holds(\theta, S')\} \subset \{\theta \in P_t(M(S')) \mid M(result(A_\mu^S, S') \models \neg holds(\theta, result(A_\mu^S, S))$ iff $M \models holds(\theta, S)\}$. It also follows from $M, N \in Mod(W(S'))$ and the agreement of M and N on $holds'$ and $persistent'$ that $N(S') = M(S')$ and $P_t(N(S')) = P_t(M(S'))$. Substituting $M(S')$ for $N(S')$ and $P_t(M(S'))$ for $P_t(N(S'))$ in the above inclusion statement immediately gives us $N(result(A_\mu^S, S')) <_{M(S')}^{P(M(S'))} M(result(A_\mu^S, S'))$, as desired. □

Lemma A.2 $CP(R, H) \neq \emptyset$.

Proof. Recall first that $CP(R, H)$ is only defined if $H = \emptyset$ and R is a state of S_0, or H is an S-history for a given S, and R is a state of $result(A_\mu^{S'}, S)$ satisfying the causal axiom (if any) for $result(A_\mu^{S'}, S)$.

In the first case, since $R \cup T \cup \{N1', N2'\}$ is consistent, it follows from the second assumption on the persistence axiom that $R \cup T \cup \{N1', N2', P\}$ is consistent, so it has a model M. Then $P_t(M(S)) \in CP(R, \emptyset)$.

In the second case, by definition of S-histories, $H \cup T \cup \{N1', N2', P\}$ is consistent. Let M be a model of this set of formulas. Then $H = \bigcup_{S' \in subterms(S)} (M(S') \cup P(M(S')))$, and since R satisfies the corresponding causal axiom (if any), it follows from the third assumption on the persistence axiom that there exists $N \in Mod(R \cup H \cup T \cup \{N1', N2', P\})$. It follows that $P_t(N(result(A_\mu^{S'}, S))) \in CP(R, H)$. $\Box$

We say that a partial order $\leq$ is *smooth* over one of its subdomains D iff for every $s \in D$, either s is minimal under $\leq$ in D or there exists $t \in D$ such that $t \leq s$ and t is minimal under $\leq$ in D.

Lemma B.2 *If the vocabulary of the underlying propositional language $\mathcal{L}$ is finite, then the ordering $\leq_M^F$ is smooth over any set D of states of $result(A_\mu^S, S')$, any state M of S', and any set of fluent terms F. In particular, $Min(D, \leq_M^F) \neq \emptyset$.*

Proof. Partition F into the set of equivalence classes $[\theta] \subseteq F$ determined by propositional equivalence of the translation of the fluent terms of F into the initial propositional language $\mathcal{L}$, and let F^* be obtained from F by picking exactly one representative element from each equivalence class. (In particular, $F^* \subseteq F$.) For any $\theta \in F$, let θ^* be the representative of $[\theta]$. We claim that $\leq_M^F = \leq_M^{F^*}$. From left to right, suppose $R \leq_M^F T$, *i.e.* $Diff_F(R, M) \subseteq Diff_F(T, M)$, and suppose $\theta^* \in Diff_{F^*}(R, M)$. Then $\theta^* \in F^* \subseteq F$, so $\theta^* \in Diff_F(R, M) \subseteq Diff_F(T, M)$, so $T \models \neg holds(\theta^*, result(A_\mu^S, S'))$ iff $M \models holds(\theta^*, S')$. Since $\theta^* \in F^*$, this means that $\theta^* \in Diff_{F^*}(T, M)$. From right to left, suppose $R \leq_M^{F^*} T$, *i.e.* $Diff_{F^*}(R, M) \subseteq Diff_{F^*}(T, M)$, and suppose $\theta \in Diff_F(R, M)$. The axioms for non primitive fluents, together with some straightforward propositional reasoning, entail $\forall s. holds(\theta, s) \equiv holds(\theta^*, s)$. Hence for every $\varphi \in [\theta]$, $M \models holds(\varphi, S')$ iff $M \models holds(\theta^*, S')$; similarly $T \models \neg holds(\varphi, result(A_\mu^S, S'))$ iff $T \models \neg holds(\theta^*, result(A_\mu^S, S'))$, and similarly for R. It then follows from $\theta \in Diff_F(R, M)$ that $\theta^* \in Diff_{F^*}(R, M) \subseteq Diff_{F^*}(T, M)$, and from $\theta^* \in Diff_{F^*}(T, M)$ that $\theta \in Diff_F(T, M)$, as desired. We conclude that $\leq_M^F = \leq_M^{F^*}$, and thus for any set D of states of $result(A_\mu^S, S')$ we have $Min(D, \leq_M^F) = Min(D, \leq_M^{F^*})$.

Because of the assumption that the initial propositional language contains only a finite number of variables, the set F^* is finite, from which the conclusion follows. $\Box$

Lemma A.3

$$States(U(\psi, <\mu>)) = \bigcup_{\substack{F \in CP(M_S, \emptyset) \\ M_S \in States(\psi^{S_0})}} Min(States(holds(\mu^t, result(A_\mu^{S_0}, S_0))), \leq_{M_S}^F)$$

Proof. If ψ^{S_0} is unsatisfiable, then the lemma is obvious, so we assume it is satisfiable. The proof uses the fact that

$$States(U(\psi, <\mu>)) = \{M(result(A_\mu^{S_0}, S_0)) \mid M \in Mod(Comp(T) \cup \{\psi^{S_0}\})\}, \qquad (1)$$

which we prove first. The right to left inclusion is obvious. For the other inclusion, suppose $R \in States(U(\psi, <\mu>))$ but for every $M_i \in Mod(Comp(T) \cup \{\psi^{S_0}\})$ we have that $R \neq M_i(result(A_\mu^{S_0}, S_0))$. Then for any such M there exists θ_i such that $M \models holds(\theta_i, result(A_\mu^{S_0}, S_0))$ and $R \not\models holds(\theta_i, result(A_\mu^{S_0}, S_0))$. Since there is only a finite number n of states (because of the finitariness of the underlying propositional language), there are only at most a finite number n of non-equivalent θ_i with this property. Thus $Comp(T) \models \psi^{S_0} \supset holds(\theta_1, result(A_\mu^{S_0}, S_0)) \vee \ldots \vee holds(\theta_n, result(A_\mu^{S_0}, S_0))$ but $R \not\models holds(\theta_1, result(A_\mu^{S_0}, S_0)) \vee \ldots \vee holds(\theta_n, result(A_\mu^{S_0}, S_0))$. But this contradicts $R \in States(U(\psi, <\mu>))$.

($\Rightarrow$) Suppose $R \in States(U(\psi, <\mu>))$. Expression 1 entails that there exists $M \in Mod(Comp(T) \cup \{\psi^{S_0}\})$ such that $M(result(A_\mu^{S_0}, S_0)) = R$. Using $M \in Mod(Comp(T))$ and lemma A.1, it follows that $R \in Min(States(holds(\mu^t, result(A_\mu^{S_0}, S_0)), \leq_{M(S_0)}^{P_t(M(S_0))})$. Trivially, $M(S_0) \in States(\psi^{S_0})$, and by definition of CP, $P_t(M(S_0)) \in CP(M(S_0), \emptyset)$.

($\Leftarrow$) Suppose $R \in Min(States(holds(\mu^t, result(A_\mu^{S_0}, S_0))), \leq_{M_S}^F)$ for some state $M_S \in States(\psi^{S_0})$ and some set of fluent terms $F \in CP(M_S, \emptyset)$. Using again expression 1, it suffices to show that for some $M \in Mod(Comp(T) \cup \{\psi^{S_0}\})$, $M(result(A_\mu^{S_0}, S_0)) = R$. We construct M as follows. First, we fix the set of states and corresponding specifications of persistent fluent terms that M should determine for each situation term S. Set $M(S_0) = M_S$, and $P_t(M(S_0)) = F$. Inductively, for any situation term $result(A_\phi^{S'}, S)$,

(a) If $S = S'$, choose $M(result(A_\phi^S, S)) \in Min(States(holds(\phi^t, result(A_\phi^S, S))), \leq_{M(S)}^{P_t(M(S))})$, with the special case $M(result(A_\mu^{S_0}, S_0)) = R$; otherwise $M(result(A_\phi^{S'}, S))$ is obtained by substituting $result(A_\phi^{S'}, S)$ for S in every formula in $M(S)$.

(b) choose $P_t(M(result(A_\phi^S, S))) \in CP(M(result(A_\phi^S, S)), H_S)$, where
$H_S = \bigcup_{S' \in subterms(S)} (M(S') \cup P(M(S')))$
$P(M(S')) = Pers(P_t(M(S')), S')$.

We claim that this specification is well-defined. The base case is that $M(result(A_\phi^{S'}, S_0))$ and $P_t(M(result(A_\phi^{S'}, S_0)))$ are defined. It follows from lemma B.2 that for any action term $A_\phi^{S_0}$, $Min(States(holds(\phi^t, result(A_\phi^{S_0}, S_0))), \leq_{M(S)}^{P_t(M(S))}) \neq \emptyset$. Thus $M(result(A_\phi^{S_0}, S_0))$ is well-defined for any $A_\phi^{S_0}$; for any $A_\phi^{S'}$ with $S' \neq S_0$, it is trivially well-defined as well. Second, to show that $P_t(M(result(A_\phi^{S'}, S_0)))$ is defined it suffices to show, by construction, that $CP(M(result(A_\phi^{S'}, S_0)), H_{S_0})$ is not empty; by lemma A.2, it suffices to show that it is defined, *i.e.* that (a) $M(result(A_\phi^{S'}, S_0))$ satisfies the causal axiom, if any, for $result(A_\phi^{S'}, S_0)$, and that (b) H_{S_0} is an S_0-history. (a) is satisfied by construction. As for (b), by hypothesis, $F \in CP(M_S, \emptyset)$; thus, by construction, $P_t(M(S_0)) \in CP(M(S_0), \emptyset)$. By definition of CP, then, $M(S_0) \cup P(M(S_0)) \cup T \cup \{N1', N2', P\}$ is consistent or, in other words, H_{S_0} is an S_0-history.

Inductively, let $S^* = result(A_\phi^{S'}, S)$ for some ϕ, and $S^{**} = result(A_\theta^{S''}, S^*)$, and suppose $M(S^*)$ and $P_t(M(S^*))$ are defined. Then $M(S^{**})$ is well-defined either trivially or by lemma B.2. As in the base case, to show that $P_t(M(S^{**}))$ is defined, it suffices to show that $CP(M(S^{**}), H_{S^*})$ is defined. (a) By construction $M(S^{**})$ satisfies the causal axiom, if any, for S^{**}. (b) By construction, $P_t(M(S^*)) \in CP(M(S^*), H_S)$, so by definition of CP, $H_S \cup M(S^*) \cup P(M(S^*)) \cup T \cup \{N1', N2', P\}$ is consistent. This means that $H_{S^*} = H_S \cup M(S^*) \cup P(M(S^*))$ is an S^*-history. We conclude that $CP(M(S^{**}), H_{S^*})$ is defined, therefore not empty, so $P_t(M(S^{**}))$ is defined.

Thus, this partial specification of M is well defined. We give it a name: Let $G = \bigcup(M(S) \cup P(M(S))$, with the union taken over the set of all closed situation terms. We claim that $G \cup T \cup \{N1', N2', P\}$ is consistent. Suppose not. By compactness of FOL, there exists some finite subset $K \subseteq (G \cup T \cup \{N1', N2', P\})$ such that $K \models false$. Let $G_F = K \setminus (T \cup \{N1', N2', P\})$, let $Sit(G_F)$ be the set of situation terms occurring in G_F at the top level (*i.e.* not just as subterms of some other term), and let $\mathcal{H} = \{H_{S_i} \mid S_i \in Sit(G_F)\}$. Notice that $G_F \subseteq \bigcup \mathcal{H}$, and thus by monotonicity of FOL consequence, $\bigcup \mathcal{H} \cup T \cup \{N1', N2', P\} \models false$.

We claim that this is impossible. It is easy to see that $\bigcup \mathcal{H}$ is consistent, since for every $S_i \in Sit(G_F)$, and every $S \in subterms(S_i)$, $M(S)$ and $P(M(S))$ are individually consistent, and since for every $S_i, S_j \in Sit(G_F)$, the common portions of H_{S_i} and H_{S_j} are identical (if S_{ij} is the greatest common subterm of S_i and S_j, the common portion is the restriction of both to formulas whose situation argument is a subterm of S_{ij}). Since we showed earlier that for every S, H_S is an S-history, it follows from the fourth assumption on the persistence axiom that $\bigcup \mathcal{H} \cup T \cup \{N1', N2', P\}$ is consistent. This contradicts our previous conclusion, thus the hypothesis for indirect proof is false, hence $G \cup T \cup \{N1', N2', P\}$ is consistent.

It follows that there exists some interpretation M of $\mathcal{L}_{\mathcal{S}}$ such that $M \in Mod(G \cup T \cup \{N1', N2', P\})$. Let M^* be the interpretation of $\mathcal{L}'_{\mathcal{S}}$ identical to M in the common part of the language and which interprets $holds'$ and $persistent'$ identically with $holds$ and $persistent$, respectively. Then $M^* \in Mod(\forall s.W(s))$. By construction of G and lemma A.1, $M^* \in Mod(Comp(T) \cup \{\psi^{S_0}\})$, and $M^*(result(A_\mu^{S_0}, S_0)) = R$, as desired. $\square$

Lemma B.3 *Suppose* $R \in Min(States(holds(\mu^t, result(A_\mu^S, S))), \leq_{H(S)}^{P_t(H(S))})$ *for some admissible* S*-history* H*, and that every subterm of* S *other than* S_0 *has the form* $result(A_\mu^{S'}, S')$*. Then there exists* $M \in Mod(Comp(T) \cup H \cup R)$*.*

Proof. The proof is very similar to that of the right to left direction of lemma A.3, so we only sketch it. Again, we specify a model M by giving first $M(S')$ and $P(M(S'))$ for every closed situation term S', and then showing that it is consistent with $T \cup \{N1', N2', P\}$. Set $M(S_0) = H(S_0)$, $P(M(S_0)) = P(H(S_0))$, and more generally for every $S' \in subterms(S)$, set $M(S') = H(S')$ and $P(M(S')) = P(H(S'))$. For every $result(A_\theta^{S'}, S'')$ such that $S \in subterms(S'')$, choose $M(result(A_\theta^{S'}, S''))$ and $P(M(result(A_\theta^{S'}, S'')))$ as in lemma A.3, with the special case $M(result(A_\mu^S, S)) = R$. It only remains to specify $M(S')$ and $P(M(S'))$ for situation terms S' such that $S \notin subterms(S')$ and $S' \notin subterms(S)$. This can be done again as in lemma A.3. The proof that this specification is well defined is almost identical. For the part shared by M and H, well definition is given. For the rest, we need to repeat identical proofs as in lemma A.3 for the "future" of H (base case: $M(result(A_\mu^{S'}, S)))$ and for alternative branches of time (base case: $M(result(A_\theta^{S'}, S_0))$, for every $A_\theta^{S'}$ such that $H(result(A_\theta^{S'}, S_0))$ is not defined). After well-definition, the proof that this specification is consistent with $T \cup \{N1', N2', P\}$ is again identical. Thus we can show that there exists a model M of $\forall s.\ W(s)$ such that $M \in Mod(H \cup R)$. By definition of admissible S-histories, the fact that any subterm of S has the form indicated in the hypothesis of the lemma, and by construction, M satisfies the conditions of lemma A.1, and thus $M \in Mod(Comp(T))$ as well. $\square$

Lemma A.4 *For* $0 < j \leq n$*, let* $S_j = result(A_{\mu_j}^{S_{j-1}}, S_{j-1})$*.*

$$States(U(\psi, <\mu_1, \ldots, \mu_n>)) = \bigcup_{\substack{F \in CP_\psi(M_S) \\ M_S \in States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>))}} Min(States(holds(\mu_n^t, S_n)), \leq_{M_S}^F)$$

Proof. As in the proof of lemma A.3, we can prove that for every $0 < j \leq n$:

$$States(U(\psi, <\mu_1, \ldots, \mu_j>)) = \{M(S_j) \mid M \in Mod(Comp(T) \cup \{\psi^{S_0}\})\}. \qquad (2)$$

($\Rightarrow$) Suppose $R \in States(U(\psi, <\mu_1, \ldots, \mu_n>))$ for some $n > 0$. Expression 2 entails that there exists $M \in Mod(Comp(T) \cup \{\psi^{S_0}\})$ such that $M(S_n) = R$. Using $M \in Mod(Comp(T))$ and lemma A.1, it follows that $R \in Min(States(holds(\mu_n^t, S_n)), \leq_{M(S_{n-1})}^{P_t(M(S_{n-1}))})$. Using expression 2 on S_{n-1}, we obtain that $M(S_{n-1}) \in States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>))$ (for the case $n = 1$, recall that by convention $U(\psi, <>) = \psi^{S_0}$). We have to verify that $P_t(M(S_{n-1})) \in CP_\psi(M(S_{n-1}))$. If $n = 1$ this follows from the definition of CP_ψ for states of S_0, and the fact that $M \in Mod(Comp(T))$. For $n > 1$, we have to show that there exists an admissible S_{n-2}-history H such that $M(S_{n-1}) \in Min(States(holds(\mu_{n-1}, S_{n-1})), \leq_{H(S_{n-2})}^{P_t(H(S_{n-2}))})$, $H \models \psi^{S_0}$, and $P_t(M(S_{n-1})) \in CP(M(S_{n-1}), H)$. Since $M \in Mod(Comp(T))$, it follows from lemma A.1 that the S_{n-2}-history H determined by M is admissible, and clearly $H \models \psi^{S_0}$. By lemma A.1, furthermore, $M(S_{n-1}) \in Min(States(holds(\mu_{n-1}, S_{n-1})), \leq_{H(S_{n-2})}^{P_t(H(S_{n-2}))})$. By definition of CP and the fact that $M \in Mod(Comp(T))$, we also have that $P_t(M(S_{n-1})) \in CP(M(S_{n-1}), H)$. Therefore $P_t(M(S_{n-1})) \in CP_\psi(M(S_{n-1}))$.

It follows then that $R \in Min(States(holds(\mu_n^t, S_n)), \leq_{M(S_{n-1})}^{P_t(M(S_{n-1}))})$ for some $M(S_{n-1}) \in States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>))$ and $P_t(M(S_{n-1})) \in CP_\psi(M(S_{n-1}))$, as desired.

($\Leftarrow$) Suppose $R \in Min(States(holds(\mu_n^t, S_n)), \leq_{M_S}^F)$, where M_S is some state of S_{n-1} such that $M_S \in States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>))$ and F is some set of fluent terms such that $F \in CP_\psi(M_S)$. Using again expression 2, it suffices to show that for some $M \in Mod(Comp(T) \cup \{\psi^{S_0}\})$, $M(S_n) = R$. If $n = 1$, this is just lemma A.3. For $n > 1$, since $F \in CP_\psi(M_S)$ for $M_S \in States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>))$, $F \in CP(M_S, H)$ for some admissible S_{n-2}-history H such that (*) $M_S \in Min(States(holds(\mu_{n-1}^t, S_{n-1})), \leq_{H(S_{n-2})}^{P_t(H(S_{n-2}))})$, and $H \models \psi^{S_0}$. Since $F \in CP(M_S, H)$, $H \cup M_S \cup Pers(F, S_{n-1}) \cup T \cup \{N1', N2', P\}$ is consistent, and therefore $H' = H \cup M_S \cup Pers(F, S_{n-1})$ is an S_{n-1}-history. Using lemma B.3 and fact (*) in the previous sentence, H' is also admissible. It follows then from lemma B.3 and the assumption on R that there exists $M \in Mod(Comp(T) \cup H' \cup R)$. Since $H' \models \psi^{S_0}$ and $M(S_n) = R$, the conclusion follows from expression 2. $\square$

Theorem 2. *The update operator $\diamond$ satisfies (U1), (U3), (U5), (U6) and (U8), but does not in general satisfy (U2), (U4) or (U7).*

Proof. Using the same notation for situation terms as in the previous lemma, let ψ_{n-1} be a propositional formula such that $States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>)) = States(\psi_{n-1}^{S_{n-1}}, S_{n-1}))$, with $\psi_0 = \psi$. That is $\psi_{n-1} \in \mathcal{L}$ represents the state of the database in the propositional language after this series of updates. (Such a formula must exist by the finitariness of the propositional language.) We show that the update of this database satisfies the postulates for any n.

We first "translate" lemma A.4 into the propositional language $\mathcal{L}$. Define a propositional analogue of the function Diff_F of definition 6. For any set of formulas Σ of the propositional language $\mathcal{L}$, and any two propositional interpretations I and J, let $\mathit{Diff}_\Sigma(I,J) = \{\theta \in \Sigma \mid I \models \theta \text{ iff } J \not\models \theta\}$, and define $I \leq_M^\Sigma J$ iff $\mathrm{Diff}_\Sigma(I,M) \subseteq \mathrm{Diff}_\Sigma(J,M)$. Finally, for any propositional interpretation M and situation term S, let $CP_S(M) = \{\Sigma_F \mid F \in CP_\psi(M_S), \Sigma_F = \{\varphi \mid \varphi^t \in F\}\}$, where for any propositional interpretation M, M_S is the state of situation S such that $M_S \models \varphi^S$ iff $M \models \varphi$.

It follows from definition 2 that for any formula $\theta \in \mathcal{L}$, $\mathit{Mod}(\psi_{n-1} \diamond \mu_n) \subseteq \mathit{Mod}(\theta)$ iff $\mathit{States}(U(\psi, < \mu_1, \ldots, \mu_n >)) \subseteq \mathit{States}(\mathit{holds}(\theta^t, S_n))$. It is straightforward to show, on the other hand, that, for $\Sigma \in CP_S(M)$, $I \in \mathit{Min}(\mathit{Mod}(\mu_n), \leq_M^\Sigma)$ if and only if $I_{S_n} \in \mathit{Min}(\mathit{States}(\mathit{holds}(\mu_n^t, S_n)), \leq_{M_{S_{n-1}}}^F)$, and thus, using lemma A.4, $\mathit{Mod}(\psi_{n-1} \diamond \mu_n) \subseteq \mathit{Mod}(\theta)$ iff $\bigcup_{\substack{\Sigma \in CP_{S_{n-1}}(M) \\ M \in \mathit{Mod}(\psi_{n-1})}} \mathit{Min}(\mathit{Mod}(\mu_n), \leq_M^\Sigma) \subseteq \mathit{Mod}(\theta)$. In the finitary case, this entails:

$$\mathit{Mod}(\psi_{n-1} \diamond \mu_n) = \bigcup_{\substack{\Sigma \in CP_{S_{n-1}}(M) \\ M \in \mathit{Mod}(\psi_{n-1})}} \mathit{Min}(\mathit{Mod}(\mu_n), \leq_M^\Sigma). \tag{3}$$

Notice that we can assume without loss of generality that for any $\Sigma \in CP_S(M)$, the set Σ is finite, using the same reasoning as in lemma B.2, and thus that the ordering $\leq_M^\Sigma$ is smooth.

We prove the postulates. (U1) and (U8) are obvious from expression 3. (Recall that updates with unsatisfiable formulas are undefined.)

We prove (U3) by induction on n. For $n = 1$, assume ψ is satisfiable, and let $M \in \mathit{Mod}(\psi)$. Then by lemma A.2 and definition of $CP_\psi(R)$ for R a state of S_0, $CP_\psi(M_{S_0}) \neq \emptyset$, so $CP_{S_0}(M) \neq \emptyset$. There exists, therefore, $\Sigma \in CP_{S_0}(M)$, and because the ordering $\leq_M^\Sigma$ is smooth, the conclusion follows. For $n > 1$, assume inductively that (U3) holds for every $m < n$. Because ψ is satisfiable and updates are only defined for μ satisfiable, it follows from the inductive hypothesis that there exists some $N \in \mathit{Mod}(\psi_{n-1})$. Therefore $N_{S_{n-1}} \in \mathit{States}(U(\psi, < \mu_1, \ldots, \mu_{n-1} >))$. Using expression 2, there exists $J \in \mathit{Mod}(\mathit{Comp}(T) \cup \{\psi^{S_0}\})$ such that $J(S_{n-1}) = N_{S_{n-1}}$. It is then straightforward to verify, using lemma A.1, that $P_t(J(S_{n-1})) \in CP_\psi(J(S_{n-1}))$, and thus that $CP_{S_{n-1}}(N_{S_{n-1}}) \neq \emptyset$. The conclusion is then an immediate consequence of expression 3 and the smoothness of the orderings.

We prove (U5). If $(\psi_{n-1} \diamond \mu_n) \wedge \theta$ is inconsistent then (U5) holds trivially. Otherwise, let $J \in \mathit{Mod}((\psi_{n-1} \diamond \mu_n) \wedge \theta)$. Then there exists $I \in \mathit{Mod}(\psi_{n-1})$ and $\Sigma \in CP_{S_{n-1}}(I)$ such that $J \in Min(Mod(\mu_n), \leq_M^\Sigma)$. Since $Mod(\mu_n \wedge \theta) \subseteq Mod(\mu_n)$, $J \in Min(Mod(\mu_n \wedge \theta), \leq_M^\Sigma)$, as required by (U5).

We prove (U6). Suppose $Mod(\psi_{n-1} \diamond \mu_n^1) \subseteq Mod(\mu_n^2)$ and $Mod(\psi_{n-1} \diamond \mu_n^2) \subseteq Mod(\mu_n^1)$, and suppose that, contrary to (U6), $J \in Mod(\psi_{n-1} \diamond \mu_n^1)$ but $J \notin Mod(\psi_{n-1} \diamond \mu_n^2)$. Then, $J \in Mod(\mu_n^2)$. Using the smoothness of $\leq_I^\Sigma$ and expression 3, for every $I \in Mod(\psi_{n-1})$ and every $\Sigma \in CP_{S_{n-1}}(I)$ there exists some $J_{I,\Sigma} \in Min(Mod(\mu_n^2), \leq_I^\Sigma)$ such that $J_{I,\Sigma} <_I^\Sigma J$. Using expression 3 again, for each $J_{I,\Sigma}$ we have that $J_{I,\Sigma} \in Mod(\psi_{n-1} \diamond \mu_n^2)$, so by assumption for every $J_{I,\Sigma}$, $J_{I,\Sigma} \in Mod(\mu_n^1)$. Thus $J \notin \bigcup_{\substack{\Sigma \in CP_{S_{n-1}}(M) \\ M \in Mod(\psi_{n-1})}} Min(Mod(\mu_n^1), \leq_M^\Sigma)$, which by expression 3 contradicts the hypothesis that $J \in Mod(\psi_{n-1} \diamond \mu_n^1)$.

The proof that (U2) and (U4) are not satisfied is provided in subsequent theorems. As for (U7), suppose the initial propositional vocabulary is $\{p, q\}$, and let the persistence axiom be:

$$\forall s.\ (\forall \theta.\ persistent(\theta, s) \equiv \theta = p) \vee (\forall \theta.\ persistent(\theta, s) \equiv \theta = q).$$

Using expression 3, verify that $Mod((p \wedge q) \diamond \neg p) \cap Mod((p \wedge q) \diamond \neg q) = Mod(\neg p \wedge \neg q)$, yet $Mod((p \wedge q) \diamond (\neg p \vee \neg q)) = Mod((p \wedge \neg q) \vee (\neg p \wedge q))$, in contradiction with (U7). $\square$

Theorem 3 *Suppose the SDP, PDS and TI conditions are satisfied. Then the update operator $\diamond$ satisfies postulates (U1)-(U8).*

Proof. From theorems 2, 5, 6, and 7. $\square$

Theorem 4 *For any update operator $\diamond'$ satisfying (U1)-(U8) there exists an operator $\diamond$ based on definition 2 and satisfying the PDS, SDP and TI conditions such that for any ψ and sequence of* satisfiable *formulas $\mu_1, \ldots, \mu_n$:*

$$Mod((\ldots((\psi \diamond \mu_1) \diamond \mu_2) \ldots) \diamond \mu_n) = Mod((\ldots((\psi \diamond' \mu_1) \diamond' \mu_2) \ldots) \diamond' \mu_n).$$

Proof. Since $\diamond$ satisfies (U1)-(U8), there exists a faithful update assignment of a partial order $\leq_M$ to every interpretation M such that $\diamond$ can be defined in terms of the representation theorem 1 as:

$$Mod(\psi \diamond' \mu) = \bigcup_{M \in Mod(\psi)} Min(Mod(\mu), \leq_M) \tag{4}$$

Let $\mathcal{W}$ be the set of all propositional interpretations, and for any $W \subseteq \mathcal{W}$, choose one formula θ_W such that $Mod(\theta_W) = W$. For any propositional model M, let $\Sigma_M = \{\theta_W \mid W = \{I \mid I \leq_M J\}$ for some $J \in \mathcal{W}\}$. Let $\leq_M^\Sigma$, for any set of formulas Σ and propositional interpretation M, be defined as in the proof of theorem 2. We first establish the following lemma:

Lemma B.4 *$I \leq_M J$ iff $I \leq_M^{\Sigma_M} J$.*

Proof. Note first that by theorem 1, the partial orders $\leq_M$ are faithful, so $M \leq_M N$ for any N and M, and thus $M \in \{K \mid K \leq_M N\}$ for any N. It follows that $\theta \in \Sigma_M$ implies $M \in \mathit{Mod}(\theta)$.

($\Rightarrow$) If $I \leq_M J$ then by transitivity of $\leq_M$: for any L, if $J \in \{K \mid K \leq_M L\}$, then $I \in \{K \mid K \leq_M L\}$. It follows that for any $\theta \in \Sigma_M$, if $J \in \mathit{Mod}(\theta)$ then $I \in \mathit{Mod}(\theta)$, or contrapositively, if $I \notin \mathit{Mod}(\theta)$ then $J \notin \mathit{Mod}(\theta)$. Since $\theta \in \Sigma_M$ implies $M \in \mathit{Mod}(\theta)$, we obtain that $\mathit{Diff}_{\Sigma_M}(I, M) \subseteq \mathit{Diff}_{\Sigma_M}(J, M)$. Thus $I \leq_M^{\Sigma_M} J$.

($\Leftarrow$) Suppose $I \leq_M^{\Sigma_M} J$, *i.e.* $\mathit{Diff}_{\Sigma_M}(I, M) \subseteq \mathit{Diff}_{\Sigma_M}(J, M)$. Using again the fact that $M \in \mathit{Mod}(\Sigma_M)$, we obtain that for every $\theta \in \Sigma_M$, if $I \notin \mathit{Mod}(\theta)$ then $J \notin \mathit{Mod}(\theta)$. Taking the contrapositive, it follows that whenever $J \in \{K \mid K \leq_M L\}$ for some L then $I \in \{K \mid K \leq_M L\}$ as well. In particular, $I \in \{K \mid K \leq_M J\}$ and thus $I \leq_M J$. $\square$

We can now complete the proof of theorem 4. The persistence axiom given in the text is such that for any interpretation M and any state M_S of some situation S, if M_S is the translation of M at S (*i.e.* if $M_S \models \varphi^S$ iff $M \models \varphi$ for any $\varphi \in \mathcal{L}$) and $CP_\psi(M_S)$ is defined then $CP_\psi(M_S)$ is the singleton set $\{\{\varphi^t \mid \varphi \in \Sigma_M\}\}$. Letting $\diamond$ be the update operator determined by the theory of action with this persistence axiom, expression 3 can be simplified to become:

$$\mathit{Mod}(\psi_{n-1} \diamond \mu_n) = \bigcup_{M \in \mathit{Mod}(\psi_{n-1})} \mathit{Min}(\mathit{Mod}(\mu_n), \leq_M^{\Sigma_M}), \tag{5}$$

where ψ_{n-1} is, as in expression 3, the result of updating ψ with the sequence $\mu_1, \ldots, \mu_n$. Thus, it follows from lemma B.4 that the right hand sides of expressions 4 and 5 are identical when the same formulas are considered. Hence

$$\mathit{Mod}(\psi_{n-1} \diamond \mu_n) = \mathit{Mod}(\psi_{n-1} \diamond' \mu_n)$$

for arbitrary n. It only remains to be shown that $\diamond$ satisfies the various conditions. TI is satisfied by construction, SDP because $CP_\psi(M_S)$ is always a singleton. As for PDS, by the faithfulness of $\leq_M$, $M <_M N$ for any $N \neq M$, and by reflexivity $M \leq_M M$. Thus

$\theta_{\{M\}} \in \Sigma_M$ for any M. Using the same notation as in the definition of the persistence axiom (P), by construction we have that for any state $M(s) \in \mathcal{M}_s$:

$$(P) \models \forall s.\ M(s) \supset persistent(\theta_{\{M\}}, s).$$

Clearly, no state $N(s)$ can agree with $M(s)$ on $\theta_{\{M\}}$ yet be distinct from $M(s)$. Thus the PDS condition is satisfied. $\Box$

Theorem 5

1. *If the PDS condition is satisfied then $\diamond$ satisfies (U2).*
2. *If $\diamond$ satisfies (U2), then the weak PDS condition is satisfied.*
3. *If the TI condition is satisfied and $\diamond$ satisfies (U2) then the PDS condition is satisfied.*

Proof. As usual, we want to prove this for arbitrary updates in a sequence of updates. Let, as usual, $S_n = result(A_{\mu_n}^{S_{n-1}}, S_{n-1})$ for any $n > 0$. Satisfaction of (U2) is equivalent to the following: for any ψ, $\mu_1, \ldots, \mu_n$, if $holds(\mu_n^t, S_{n-1}) \in U(\psi, <\mu_1, \ldots, \mu_{n-1}>)$ then $States(U(\psi, <\mu_1, \ldots, \mu_n>)) = \{R(S_{n-1}/S_n) \mid R \in States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>))\}$, where the notation $R(S/S')$ indicates the substitution of S' for S in state R.

1. Assume $holds(\mu_n^t, S_{n-1}) \in U(\psi, <\mu_1, \ldots, \mu_{n-1}>)$, and that the PDS condition is satisfied. Using expression 2 from the proof of lemma A.4 above, it suffices to show that for any $M \in Mod(Comp(T) \cup \{\psi^{S_0}\})$, $M(S_{n-1}/S_n) = M(S_n)$. So let $M \in Mod(Comp(T) \cup \{\psi^{S_0}\})$. Using lemma A.1, it suffices to show in turn that $Min(States(holds(\mu_n^t, S_n)), \leq_{M(S_{n-1})}^{P_t(M(S_{n-1}))}) = \{M(S_{n-1}/S_n)\}$. For the left to right inclusion, we have by assumption that $M(S_{n-1}/S_n) \in States(holds(\mu_n^t, S_n))$; therefore $Diff_{P_t(M(S_{n-1}))}(M(S_{n-1}/S_n), M(S_{n-1})) = \emptyset$, and the conclusion follows. For the other inclusion, let $R \in Min(States(holds(\mu_n^t, S_n)), \leq_{M(S_{n-1})}^{P_t(M(S_{n-1}))})$. Then $Diff_{P_t(M(S_{n-1}))}(M(S_{n-1}/S_n), M(S_{n-1})) = \emptyset$, or else R would not be minimal. Thus $R(S_n/S_{n-1})$ agrees with $M(S_{n-1})$ on the value of every $\theta \in P_t(M(S_{n-1}))$. Using the PDS condition, $R(S_n/S_{n-1}) = M(S_{n-1})$, from which the conclusion follows.

2. Suppose the weak PDS condition is violated for some R_1, R_2 and $P \in CP_\psi(R_1)$. Then for any tautology μ_n, we have that $Diff_P(R_2(S_{n-1}/S_n), R_1) = \emptyset$. Since $R_2(S_{n-1}/S_n)$ satisfies $holds(\mu_n^t, S_n)$, $R_2(S_{n-1}/S_n) \in Min(States(holds(\mu_n^t, S_n)), \leq_R^P)$. By assumption, $R_1 \in States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>))$ and $P \in CP_\psi(R_1)$, so by lemma A.4, $R_2(S_{n-1}/S_n) \in States(U(\psi, <\mu_1, \ldots, \mu_n>))$. But since $R_2 \notin States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>))$, this implies that $\diamond$ does not satisfy (U2).

3. Suppose (U2) is satisfied but the PDS condition is not for some states R_1 and R_2 of some S, and some set of fluent terms P satisfying the assumptions of the PDS condition. Because the TI condition is satisfied, the set of allowable specifications of persistent fluents at any state is identical for states which are equal modulo substitution of the situation argument. This implies that we can assume without loss of generality that $S = S_0$, since R_1 and $R_1(S/S_0)$ will have the same set of allowable specifications of the persistent fluents. Let $\mu_1 \in \mathcal{L}$ be a formula such that $Mod(\mu_1) = \{M_1\}$ where M_1 is such that for any φ, $M_1 \models \varphi$ iff $R_1 \models \varphi^{S_0}$, and let μ_2 be a propositional tautology. Let $S_1 = result(A_{\mu_1}^{S_0}, S_0)))$, $S_2 = result(A_{\mu_2}^{S_1}, S_1)))$. The choice of μ_1 implies that $States(holds(\mu_1^t, S_1)) = \{R_1(S_0/S_1)\}$, hence $States(U(\psi, < \mu_1 >)) = \{R_1(S_0/S_1)\}$. Since R_1 and R_2 agree on all fluents in P, and $R_2(S_0/S_2)$ trivially satisfies the causal axiom for μ_2, $R_2(S_0/S_2) \in Min(States(holds(\mu_2^t, S_2)), \leq_{R_1(S_0/S_1)}^P)$. It follows from the TI condition that $P \in CP_\psi(R_1(S_0/S_1))$, and thus by lemma A.4, $R_2(S_0/S_2) \in States(U(\psi, < \mu_1, \mu_2 >))$, even though $R_2(S/S_1) \neq R_1(S/S_1)$. Since $holds(\mu_2^t, S_1) \in U(\psi, < \mu_1 >)$ (trivially), this contradicts (U2). $\Box$

Theorem 6 *The update operator $\diamond$ satisfies (U4) iff it is based on a theory of action which is update equivalent to a theory of action satisfying the TI condition.*

Proof. ($\Leftarrow$) Without loss of generality, assume $\diamond$ is based on a theory of action satisfying the TI condition (rather than on a theory of action which is update equivalent to one satisfying the TI condition). Then the persistence axiom has the form $\forall s.\ P(s)$, where $P(s)$ contains no situation term other than s. It follows that the set $CP_\psi(R)$, whenever it is defined, is unique modulo substitutions of the situation argument. The right to left direction of the theorem is then an immediate consequence of this and lemma A.4.

($\Rightarrow$) Suppose $\diamond$ satisfies (U4), and let (P) be the persistence axiom of the underlying theory of action. Let us superscript the functions CP^P and CP_ψ^P with the persistence axiom as well. We construct an alternative persistence axiom (P*) with the same specification of persistent fluents at S_0. Specifically, for any propositional interpretation M, let $M(s)$ be a finite axiomatization of the state of s such that for any $\varphi \in \mathcal{L}$, $M \models \varphi$ iff $M(s) \models \varphi^s$. Let $\mathcal{M}_s$ be the set of all such finitely axiomatized states. Without loss of generality (using a reasoning similar to that of lemma B.2), we can assume that for any $M(s) \in \mathcal{M}_s$, the set $CP^P(M(S_0), \emptyset)$ is finite, say $\{P_M^1, \ldots, P_M^m\}$, where each P_M^i is finite as well. Let then $P_M(s)$ be a finite axiomatization of "[$persistent(\theta, s)$ iff $\theta \in P_M^1$] or ... or [$persistent(\theta, s)$

iff $\theta \in P^m_M$]". The persistence axiom (P*) is:

$$\forall s. \bigwedge_{\mathcal{M}_s} M(s) \supset P_M(s).$$

We have to show that these two theories are update equivalent, that is, that for any $\psi, \mu_1, \ldots, \mu_n$, $U_P(\psi, <\mu_1, \ldots, \mu_n>) = U_{P^*}(\psi, <\mu_1, \ldots, \mu_n>)$. For $n = 1$, this trivially holds, since for any state R of S_0 and any ψ we have $CP^P_\psi(R) = CP^{P^*}_\psi(R)$. Assume inductively that $U_P(\psi, <\mu_1, \ldots, \mu_{n-1}>) = U_{P^*}(\psi, <\mu_1, \ldots, \mu_{n-1}>)$. Let φ represent the current state of the database, *i.e.* $States(\varphi^{S_{n-1}}) = States(U_P(\psi, <\mu_1, \ldots, \mu_{n-1}>))$ (using the by now customary notation for the situation terms of update sequences). Let $S^* = result(A^{S_0}_{\mu_n}, S_0)$. Because the theory of action containing (P) satisfies (U4), we have that $States(U_P(\psi, <\mu_1, \ldots, \mu_n>)) = \{R(S^*/S_n) \mid R \in States(U_P(\varphi, <\mu_n>))\}$. By inductive hypothesis, $States(\varphi^{S_{n-1}}) = States(U_{P^*}(\psi, <\mu_1, \ldots, \mu_{n-1}>))$ as well, and because (P*) satisfies the TI condition, we also have $States(U_{P^*}(\psi, <\mu_1, \ldots, \mu_n>)) = \{R(S^*/S_n) \mid R \in States(U_{P^*}(\varphi, <\mu_n>))\}$. But $States(U_{P^*}(\varphi, <\mu_n>)) = States(U_P(\varphi, <\mu_n>))$, from which the conclusion follows. □

Theorem 7

1. *If the SDP condition is satisfied then $\diamond$ satisfies (U7).*

2. *If the PDS condition is satisfied and $\diamond$ satisfies (U7) then $\diamond$ is update equivalent to a theory of action satisfying the SDP condition.*

Proof. 1. Suppose the SDP condition is satisfied. Then $CP_\psi(R)$ is always a singleton. Suppose $States(U(\psi, <\mu_1, \ldots, \mu_n>)) = \{M_S\} = States(\varphi^{S_n})$ for some φ (so that φ is complete). Say $CP_\psi(M_S) = \{F\}$. Using expression 3:

$$Mod(\varphi \diamond \mu) = Min(Mod(\mu), \leq^{\Sigma_F}_M), \tag{6}$$

where: $\Sigma_F = \{\varphi \mid \varphi^t \in F\}$ is the set of formulas corresponding to the fluents in F; $M \models \theta$ iff $M_S \models holds(\theta^t, S_n)$; and $\leq^{\Sigma_F}_M$ is the partial order defined in the text preceding expression 3. Suppose $I \in Mod(\varphi \diamond \mu_1) \cap Mod(\varphi \diamond \mu_2)$, but $I \notin Mod(\varphi \diamond (\mu_1 \vee \mu_2))$, in contradiction with (U7). By expression 6, there exists $J \in Mod(\mu_1 \vee \mu_2)$ such that $J <^{\Sigma_F}_M I$. If $J \in Mod(\mu_1)$, this contradicts the minimality of I in $Mod(\mu_1)$ according to $\leq^{\Sigma_F}_M$; if $J \in Mod(\mu_2)$, this contradicts the minimality of I in $Mod(\mu_2)$ according to $\leq^{\Sigma_F}_M$.

2. Suppose $\diamond$ satisfies (U7) and the PDS condition. By theorems 2 and 5, $\diamond$ satisfies (U1)–(U3) and (U5)–(U8). Furthermore, it is easily seen that (U4) is satisfied when we do not consider iterated updates, so *for each situation* there exists a faithful update assignment of a partial preorder $\leq_M$ to each interpretation M such that, for any given situation:

$$Mod(\psi \diamond \mu) = \bigcup_{M \in Mod(\psi)} Min(Mod(\mu), \leq_M).$$

We can then apply the same construction as in theorem 4, with possibly one persistence axiom for each situation (so "the" persistence axiom would in this case be a possibly infinite set of axioms), to show that this operator can be captured by a theory of action satisfying the PDS and SDP conditions. $\square$

Lemma A.5 *Suppose $\psi \models \gamma$, and for $0 < j \leq n$, let $S_j = result(A_{\mu_j}^{S_{j-1}}, S_{j-1})$.*

$$States(U_\gamma(\psi, <\mu_1, \ldots, \mu_n>)) = \bigcup_{\substack{F \in CP_\psi(M_S) \\ M_S \in States(U(\psi, <\mu_1, \ldots, \mu_{n-1}>))}} Min(States(holds((\mu_n \wedge \gamma)^t, S_n))), \leq_{M_S}^F)$$

Proof. The proof is identical to the proof of the analogous lemma without constraints (lemma A.4), after making all the substitutions described in the text (section 7. In particular, the replacement of any expression of the form $States(holds(\mu^t, S))$ by $States(holds((\mu \wedge \gamma)^t, S))$ should be done throughout the proofs as well. $\square$

Theorem 8 *If $\psi \models \gamma$ then $\psi \diamond_\gamma \mu \equiv \psi \diamond (\mu \wedge \gamma)$*

Proof. Immediate from lemmas A.4 and A.5. $\square$