Deriving properties of belief update from theories of action

Alvaro del Val
Robotics Lab
Stanford University
Stanford, CA 94305
delval@scottie.stanford.edu

Yoav Shoham
Computer Science Department
Stanford University
Stanford, CA 94305
shoham@cs.stanford.edu

Abstract

Two areas that have attracted much interest in recent years, belief update and reasoning about action, have so far been largely disjoint. Indeed, at first glance there appears to be little connection between them. In this paper we argue that this first impression is wrong; specifically, we show that the postulates for belief update recently proposed in [Katsuno and Mendelzon, 1991], can in fact be analytically derived, using the formal theory of action proposed in [Lin and Shoham, 1991].

Introduction

In this paper we tie together theories of belief change and theories of action, two areas that have attracted much interest in recent years. 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? Initial work 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, it has been 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] proposed 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. We will describe this work in more detail later in the paper.

In a largely independent line of enquiry, researchers have been interested in formal theories of time and action. 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. For an overview of the literature on nonmonotonic temporal reasoning, cf. [Shoham and Baker, 1992].

As we have said, these two research areas have been 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 be derived analytically. 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 empoverished 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 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 that proposed in [Lin and Shoham, 1991], which is described later. Finally, anything inferred about the resulting situation can be backtranslated to the timeless framework of belief update. In this way the KM-postulates can be proved.

This approach relies crucially on the meaning of update, and is not applicable in any straightforward way to belief revision. We note, however, that [Grahne *et al.*, 1992] have recently proposed a connection between revision and update, which suggests that the revision postulates can nonetheless be derived; we return to this topic in the summary section.

The structure of this article is as follows. Section 2 briefly reviews the results of [Katsuno and Mendelzon, 1991] on belief update. Section 3 does the same for the theory of action proposed in [Lin and Shoham, 1991]. The main contribution of the paper lies in section 4, where the belief update problem is encoded as a theory of action, and the KM-postulates are derived as theorems. In section 5 we consider the case of update in the presence of "integrity constraints", and in section 6 we provide a characterization of the propositional update operators determined by our construction. We discuss related work and open problems in the concluding section.

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 with a finite number of propositional variables. 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 in terms of the following representation theorem. 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 interpretation J, if $I \neq J$ then $I \leq_I J$ and $J \not\leq_I I$. In what follows, we use $Min(S, \leq)$, for any set S and preorder $\leq$ over S, to denote the set of elements of S that are minimal under $\leq$.

Theorem 1 *(Katsuno and Mendelzon, 1990) An update operator $\diamond$ satisfies conditions (U1)-(U8) iff there exists a faithful assignment that maps each interpretation to a partial preorder $\leq_I$ such that:*

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

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, which has been proposed in [Lin and Shoham, 1991]. Beside demonstrating that their formulation yields the desired results in particular examples that had been discussed in the literature previously, 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 (called "epistemological completeness"), and showed their formulation adequate relative to this criterion. We do not have the space to explain this criterion further here, but mention it by way of justifying our selection of a theory of action. In the remainder of the section we review the theory.

We use the situation calculus formalism. To be precise, our language $\mathcal{S}$ 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 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 $\mathbf{P}$, and, following [Lifschitz, 1990], use a predicate *frame*, whose extension is exactly the set of fluents denoted by some fluent term in **P**. We also use the predicate $ab(p;s;a)$ as an abbreviation for:

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

We assume $frame$ to be explicitly defined by means of some axiom (F). Since we are going to use circumscription, we need unique names axioms for the set of fluents **P** and for situations. We will denote the set of unique names axioms for fluents by (N1). For situations, we use the unique names axiom (N2):

$$\begin{aligned}&\forall a, s.\ earlier(s, result(a,s)) \wedge \\ &\forall s, s', s''.\ earlier(s,s') \wedge earlier(s', s'') \supset earlier(s,s'') \wedge \\ &\forall s, s'.\ earlier(s,s') \supset s \neq s'.\end{aligned}$$

In order to apply the circumscription policy, we consider the language $\mathcal{S}'$, which is the extension of $\mathcal{S}$ 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:

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

where $Circum(W(s); ab(p;s;a); holds)$ stands for the circumscription of ab in W with $holds$ allowed to vary. Intuitively, what this circumscription 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:

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

Then $Comp(T_0)$ entails:

$$\begin{aligned}\forall s, p.\ \ &frame(p) \wedge p \neq P_1 \supset \\ &holds(p,s) \equiv holds(p, result(toggle, s)),\end{aligned}$$

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

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 $holds(or(P, and(Q,R)), S)$, 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 $holds(not(R), result(A^S_{\neg r}, S))$.

For simplicity, we consider only the finitary case, that is, our initial propositional language contains only a finite number of variables. In addition, we will assume that the set of frame fluents used below is finite.

We will be using the situation calculus, defined in the previous section. The situation calculus allows arbitrary terms. We will first fix the terms that we will be using in order to express the initial database and the update formula in situation calculus. Then we will specify the tranlation process. Finally, we encode the update problem as a causal theory.

We first introduce the sets of situation and action terms. For any situation term S and any *satisfiable* formula $\mu \in \mathcal{L}$, we will introduce an action constant A^S_μ, with the intuitive meaning described before. Situation terms consist of the constant S_0, intuitively denoting the initial situation, and of terms of the form $result(A^S_\mu, S')$ for any action term A^S_μ and situation term S'.

The set of fluent terms is obtained quite directly from the propositional language. Consider a propositional language $\mathcal{L}$ with a set of primitive propositional symbols $\mathcal{P}_\mathcal{L}$ and closed under negation ($\neg$) and disjunction ($\vee$). The set $\mathcal{P}$ of fluent terms of $\mathcal{S}$ is defined as follows:

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

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

$$\begin{aligned}&\forall p, s.\ holds(not(p), s) \equiv \neg holds(p,s); \\ &\forall p, q, s.\ holds(or(p,q), s) \equiv holds(p,s) \vee holds(q,s).\end{aligned}$$

Now that the terms of the language are fixed, we can 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 $\mathcal{S}$. Thus, rather than defining "the" translation of a formula ψ into the language $\mathcal{S}$, we define the translation of ψ at a situation S, denoted by ψ^S. The easiest way to do it is to first map ψ into a fluent term ψ^t as follows:

$$\begin{aligned}&p^t = P \text{ if } p \in \mathcal{P}_\mathcal{L} \\ &(\neg\psi)^t = not(\psi^t) \\ &(\psi \vee \phi)^t = or(\psi^t, \phi^t).\end{aligned}$$

We can now define ψ^S, for any formula $\psi \in \mathcal{L}$, and any situation term S in $\mathcal{S}$, simply as:

$$\psi^S = holds(\psi^t, S).$$

The causal theory for the actions we have introduced is given by 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_μ^S and S are closed terms of the appropriate sorts[1].

We will now apply the policy of the previous section. Fix first the fluent terms P for which $frame(P)$ holds. The following results assume the set of frame fluents is kept fixed but, unless explicitly noted, are independent of what choice do we make in this regard; it also assumes that this set is finite[2]. Second, let (N1′) be the unique names axiom for the frame fluents. For situations, the matter is slightly more complex than before. (N2′) stands for the conjunction of (N2) with:

$$\forall s, s', A_\mu^S, A_\phi^{S'}.(result(A_\mu^S, s) = result(A_\phi^{S'}, s')) \equiv (s = s' \wedge \forall s''.\ holds(\mu^t, s'') \equiv holds(\phi^t, s'')).$$

Intuitively, this makes any two situations distinct except in the case in which they are the result of performing "equivalent" actions in identical situations. (The recursion to determine whether two situations are identical bottoms out in S_0.)

Let W be as described earlier[3], replacing (N1) and (N2) by (N1′) and (N2′), respectively, and let $Comp(T)$ be as before.

The circumscription of ab for each situation and action results in a set of strict partial orderings $<_{ab,S',A_\mu^S}$ over the interpretations of $\mathcal{S}$, such that $I <_{ab,S',A_\mu^S} J$ iff I and J agree on everything except $holds$ and ab, and the extension of $ab(p)(s/S'; a/A_\mu^S)$ in I is a proper subset of its extension in J.

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. Formally, for any situation term S we say that M_S is a *state* of situation S iff there is some $\mathcal{P}' \subseteq \mathcal{P}$ such that

$$M_S = \{holds(P, S) \mid P \in \mathcal{P}'\} \cup \{\neg holds(P, S) \mid P \in \mathcal{P} - \mathcal{P}'\},$$

and M_S is consistent. Intuitively, a state of a given situation is a complete specification of the values of all fluents at that situation.

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 states and the action corresponding to the update formula.

Definition 1 *Let $I_{r(A_\mu,S)}$ and $J_{r(A_\mu,S)}$ be two states of the situation $result(A_\mu^{S'}, S)$, and let M_S be a state of the situation S. $I_{r(A_\mu,S)} <_{M_S} J_{r(A_\mu,S)}$ iff for any interpretation J, if $J \in Mods(W(S))$, $J(result(A_\mu^{S'}, S)) = J_{r(A_\mu,S)}$ and $J(S) = M(S)$ then there exists an interpretation $I \in Mods(W(S))$ such that $I <_{ab,S',A_\mu^S} J$ and $I(result(A_\mu^{S'}, S)) = I_{r(A_\mu,S)}$.*

In what follows, for any set of formulas Γ (in either language) we will use the notation $Mods(\Gamma)$ for the set of models of Γ. Similarly, let $\Gamma^S \subseteq \mathcal{S}$ be a set of situation calculus formulas, with S as the only situation term occurring in Γ^S: we use $States(\Gamma^S)$ for the set of states M_S of S such that $M_S \models \varphi^S$ for every $\varphi^S \in \Gamma^S$. Intuitively, 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 following lemma tells us the sense in which these orderings capture the result of the circumscription.

Lemma 2
For every $M \in Mods(\forall s.W(s))$, $M \in Mods(Comp(T))$ iff for every S' and A_μ^S, $M(result(A_\mu^S, S')) \in Min(States(holds(\mu_t, result(A_\mu^S, S'))), <_{M(S')})$.

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

$$R_{\psi^{S_0},A_\mu^{S_0}} = \{\varphi \mid Comp(T) \cup \psi^{S_0} \models \varphi \text{ and } \varphi \text{ contains } result(A_\mu^{S_0}, S_0) \text{ as only situation term}\}.$$

The next lemma draws us very close to the representation theorem of Katsuno and Mendelzon.

Lemma 3 $States(R_{\psi^{S_0},A_\mu^{S_0}}) =$

$$\bigcup_{M_S \in States(\psi^{S_0})} Min(States(holds(\mu^t, result(A_\mu^{S_0}, S_0))), <_{M_S}).$$

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.

Given this result and our translation, it is easy to see how we can derive the KM-postulates. For some

[1]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_μ^S in situation S', resulting in an "update" which leaves the database unchanged.

[2]This restriction, as well as the restriction to a finitary propositional language, can be lifted by requiring that the circumscriptive ordering resulting from choosing an infinite set of frame fluents be *smooth* in the sense of [Lehmann and Magidor, 1990]. Analogous results can then be proved for an infinitary version of the postulates. See [del Val, 1992].

[3]Taken as a set rather than as a conjunction, since T now consists of an infinite number of axioms.

fixed choice for the predicate *frame*, define the update operator $\diamond$ as follows, for any formula μ and database ψ:

Definition 2
$\psi \diamond \mu \models \phi$ *iff* $holds(\phi^t, result(A_\mu^{S_0}, S_0)) \in R_{\psi s_0, A_\mu^{s_0}}$.

We are now ready for the main results of this paper.

Theorem 4 *The update operator $\diamond$ satisfies postulates (U1) and (U3)-(U8).*

In order to satisfy (U2) we need a further condition.

Definition 3 (Frame completeness condition) *A choice of frame fluents is* complete *iff for any situation s and states R and T of s consistent with $W(s)$, if R and T agree on all frame fluents then $R = T$.*

Intuitively, the frame completeness condition ensures that the values of the frame fluents are sufficient to completely characterize a state, and plays the same role as the faithfulness condition of section 2.

Theorem 5 *The update operator $\diamond$ satisfies (U2) if and only if it satisfies the frame completeness condition.*

Integrity constraints

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. There is often a set of formulas which play the role of "integrity constraints", to use a term common in the database literature; these are formulas which the database should always satisfy. [Katsuno and Mendelzon, 1989] postulate, in the context of AGM revision rather than KM-update, that revision operator $\circ_\gamma$ under constraints γ 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 are easily handled in our framework. Suppose we are given, in addition to the initial database ψ, a constraint γ (we assume $\psi \models \gamma$). 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 $Comp(T')$. Let $<_{M_S}^\gamma$ be the ordering obtained by replacing $W(S)$ by $W'(S)$ in definition 1. Finally, let $R_{\psi s_0, \gamma, A_\mu^{s_0}}$ be the result of replacing $Comp(T)$ by $Comp(T')$ in the definition of $R_{\psi s_0, A_\mu^{s_0}}$, and define the update operator $\diamond_\gamma$ under constraints γ analogously. Then all the lemmas of section 4 still hold, after making the appropriate substitutions. We will therefore not repeat them here. Rather, we remark that satisfaction of the constraints is already built into the definition of the orderings $<_{M_S}^\gamma$.

Lemma 6
$Min(States(holds(\mu^t, result(A_\mu^S, S))), <_{M_S}^\gamma) =$
$\quad Min(States(holds(\mu^t \wedge \gamma^t, result(A_\mu^S, S))), <_{M_S}^\gamma)$.

As a result, lemma 3 can also be written as:

Lemma 7 $States(R_{\psi s_0, \gamma, A_\mu^{s_0}}) =$

$$\bigcup_{M \in States(\psi^{S_0})} Min(States(holds(\mu^t \wedge \gamma^t, result(A_\mu^{S_0}, S_0))), <_M^\gamma)$$

Corollary 8 $\psi \diamond_\gamma \mu \equiv \psi \diamond (\mu \wedge \gamma)$

Notice however that constraints add an additional degree of freedom in the design of update operators satisfying (U2), by making it easier for the frame completeness condition to be satisfied.

Propositional update operators

The constraints imposed on update operators by our construction seem to be tighter than the KM-postulates, and thus it appears that the converse of theorem 4 does not hold. However, we can characterize the set of propositional update operators determined by our construction as follows.

Definition 4 *A propositional update operator $\diamond_\gamma$ (under constraints γ) is "action based" iff there exists a set $\Gamma \subseteq \mathcal{L}$ and an ordering $<_M$ over interpretations for each interpretation M such that:*

1. *$\psi \diamond_\gamma \mu \models \sigma$ iff $\bigcup_{I \in Mods(\psi)} Min(Mods(\mu \wedge \gamma), \leq_I) \subseteq Mods(\sigma)$*
2. *Γ is a set of "frame" formulas satisfying: if $I \models \gamma$, $J \models \gamma$ and for every $\theta \in \Gamma$, $I \models \theta$ iff $J \models \theta$, then $I = J$.*
3. *$I <_M J$ iff $Diff_\Gamma(I, M) \subset Diff_\Gamma(J, M)$, where $Diff_\Gamma(I, M) = \{\theta \in \Gamma \mid I \models \theta \text{ iff } M \not\models \theta\}$.*

Theorem 9 *The class of action-based update operators and the class of operators definable with our construction and satisfying the frame completeness condition are identical.*

In the special case in which $frame(p^t)$ iff p is a primitive propositional symbol, we have Winslett's (non-prioritized) "PMA" update operator. We do not consider in this paper the introduction of "priorities" in the definition of update operators, which would correspond in our framework to using prioritized circumscription. For epistemologically complete theories of action, prioritization has no role to play. For non-epistemologically complete theories, however, it might be desirable to use it, to capture the information that some changes are more likely than others. We expect our results to extend easily to this case, including the "generalized prioritized circumscription" introduced in [Grosof, 1991]. (See [del Val, 1992] for details).

Conclusions and further work

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. 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 proposal.

Independently, [Reiter, 1992] has proposed an account of database update in terms of his recent proposal for solving the frame problem in [Reiter, 1991]. Though his proposal appears to be somewhat more limited in the type of updates that it allows and might run into limitations in dealing with the ramification problem, the connections between his work and ours are still unclear.

It would be desirable to obtain similar results for AGM revision, and to establish the connections between AGM revision and KM update. The work reported in [Grahne *et al.*, 1992] is an important step in the direction of a solution to the second problem, but the first one remains open.

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. How should we deal with the case in which these expectations turn out to be wrong? Should we use AGM revision, or is there a more promising approach based on research in reasoning about action?

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. Of special interest in this context is the question of the persistence of derived information. The definition of parallel updates on the basis of a treatment of parallel actions is another open problem[4]. Finally, 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.

Acknowledgements

We thank John McCarthy and Ray Reiter for useful comments on previous versions of this paper.

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

References

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, Alvaro 1992. *Belief Revision and Update.* Ph.D. Dissertation, Stanford University. In preparation.

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

Grahne, Gösta; Mendelzon, Alberto; and Reiter, Raymond 1992. On the semantics of belief revision systems. In *Proceedings of the Conference on Theoretical Aspects of Reasoning about Knowledge.*

Grosof, Benjamin 1991. Generalizing prioritization. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning.*

Katsuno, Hirofumi and Mendelzon, Alberto O. 1989. A unified view of propositional knowledge base updates. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence.*

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.*

Lehmann, Daniel and Magidor, Menachem 1990. What does a conditional knowledge base entail? Technical Report TR-90-10, CSD, Hebrew University.

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

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

Lin, Fanghzen and Shoham, Yoav 1992. Concurrent actions in situation calculus. In *Proceedings of the AAAI-92.*

Reiter, Raymond R. 1991. The frame problem in the situation calculus. In *Working Notes of the AAAI Spring Symposium.*

Reiter, Raymond R. 1992. On formalizing database updates: Preliminary report. In *Proceedings of the 3rd International Conference on Extending Database Technology.*

Shoham, Yoav and Baker, Andrew 1992. Nonmonotonic temporal reasoning. In Gabbay, D., editor 1992, *Handbook of Artificial Intelligence and Logic Programming.* Forthcoming.

Winslett, Marianne 1988. Reasoning about action using a possible models approach. In *Proceedings of the 7th Conference of the AAAI.*

Winslett, Marianne 1989. Sometimes updates are circumscription. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence.*