



HAL
open science

Sensing and revision in a modal logic of belief and action

Andreas Herzig, Dominique Longin

► **To cite this version:**

Andreas Herzig, Dominique Longin. Sensing and revision in a modal logic of belief and action. [Rapport de recherche] IRIT/2002-01-R, IRIT : Institut de Recherche en Informatique de Toulouse, France. 2002. hal-03534104

HAL Id: hal-03534104

<https://hal.science/hal-03534104>

Submitted on 19 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sensing and revision in a modal logic of belief and action (preliminary report)

Andreas Herzig Dominique Longin

Institut de Recherche en Informatique de Toulouse
Équipe Logic, Interaction, LAnguage and Computation
118 Route de Narbonne, F-31062 Toulouse cedex 4
`{herzig,longin}@irit.fr`
<http://www.irit.fr/ACTIVITES/LILaC/>

2 janvier 2002

Technical Report – IRIT/2002-01-R

Abstract

We propose a modal logic of belief and nondeterministic actions, where sensing is in terms of test actions. It is not supposed that the action laws are known by the agent.

We propose successor state axioms, which in the case of deterministic actions allow for regression.

For the case where sensing is inconsistent with the agent's beliefs we propose a new solution that does not resort to orderings of plausibility.

Keywords: Reasoning about Action, Belief Revision, Modal Logic.

1 Introduction

Intelligent agents have two fundamental abilities: perception and action. Since the beginning of the 90ies, research in reasoning about actions has analyzed perception in terms of actions. Such *perception actions* have been studied as part of a larger class of actions, called *sensing actions* [16, 11], which can be defined as actions involving perception. Actions not involving perception are *non-sensing actions*. Perception actions are thus ‘pure’ sensing actions: they do not modify the environment.

It is noted in several places (e.g. [20, footnote 10], [16]) that actions can be decomposed into a sequence of non-sensing actions and perception actions. For example, the action of tossing a coin is the non-sensing action of tossing without observing the result – eyes shut –, followed by the perception action of checking the result.

The simplest perception action we can think of is testing whether some proposition is true: checking whether the light in some room is on, whether tossing a coin resulted in heads or tails, etc. We call such actions *test actions*.

In this paper we study a logic where atomic actions are restricted to be either non-sensing actions or test actions, and shall thus study the simplest logic of actions allowing for sensing. It will turn out that both classes of actions obey the same principles. These principles have been called *successor state axioms* in reasoning about actions, and it has been shown that they enable simple reasoning mechanisms [13]. The extension to non-sensing actions requires introducing mental attitudes into the language such has knowledge [16, 22] and goals [19]. In almost all approaches successor state axioms are formulated in terms of a modal operator of knowledge. Knowledge being viewed as true belief, it follows that if perception is correct then *surprises* are impossible: if an agent knows that A then he can never perceive that $\neg A$. (Indeed, as A was known before, A must be true, and given that perception doesn’t change the environment, A still holds after the perception action; as perception is correct $\neg A$ cannot be perceived.) It follows that to reason about the dynamics of knowledge we only need two operations: KM-updates [9] to take into account non-sensing actions, and AGM-expansions [1] to take into account perception actions.

This is an idealisation: perception often contradicts what is held to be true. I might be convinced that I have a coin in my pocket, but on checking I find out I don’t, I think my watch is waterproof but when

trying it out it isn't, etc. It is therefore rather the notion of *belief* that is appropriate here:

It is non trivial to extend the above solutions in order to handle belief. Expansion operations do not suffice to take into account the effects of perception on belief: we need belief revision operations \tilde{A} *la* AGM. The only proposal up to now is that of Shapiro *et al.* [20], which is based on orderings of plausibility. Our proposal does without such a device, and differs in several ways. Most importantly, in [20] it is supposed that while beliefs about contingent facts might be wrong, beliefs about action laws are correct. We go beyond that and suppose that action laws as believed by agents might differ from the 'real' action laws. Aiming at a logic that is as simple as possible, we propose a logic that is based on a Dynamic Logic solution to the frame problem [2].

In Sect. 2 we introduce a logic for belief, action, and time. Then we focus on non-sensing and testing actions and the corresponding action laws (Sect. 3), and show how the frame problem can be solved (Sect. 4). Finally we discuss related work, in particular the recent Situation Calculus based approaches of Shapiro *et col.* (Sect. 5).

2 Dynamic Doxastic Logic

Here we describe a basic logic, which is similar to that of [18, 17], augmented with a temporal operator.

We suppose that our language contains a modal operator of belief *Bel*. The formula $BelA$ is read "the agent believes A ". $BelIf A$ reads "the agent knows¹ whether A is true or not", and can be defined as $BelIf A \stackrel{\text{def.}}{=} BelA \vee Bel\neg A$.

We adopt the modal logic KD45 as the logic of belief, i.e. we suppose agents do not entertain inconsistent beliefs, and are aware of their beliefs and disbeliefs.

We use a simple version of PDL [6] to speak about actions. To each action a there is associated a modal operator $After_a$. An example of a formula is $Bel\neg After_a \perp \wedge After_a \perp$, expressing that the agent believes that a can be executed, while this is not the case. The operator $Feasible_a$ is introduced as an abbreviation: $Feasible_a A \stackrel{\text{def.}}{=} \neg After_a \neg A$.

We adopt the standard axiomatics of PDL, which for our fragment is

¹We use the term "knows" here because "the agent believes whether A " sounds odd.

nothing but the multimodal logic K. ($After_a$ corresponds to the Dynamic Logic operator $[a]$, and $Feasible_a$ to $\langle a \rangle$.)

We finally suppose our language contains a temporal operator $Always$. $AlwaysA$ is read “ A is always true.” We adopt the principles of S4 for $Always$, plus

$$AlwaysA \rightarrow After_a A$$

It follows that $AlwaysA \rightarrow After_{a_1} \dots After_{a_n} A$, and also $AlwaysA \rightarrow After_a AlwaysA$.

Note that that we do not suppose that $AlwaysA \rightarrow BelA$. Hence $Always$ is not as powerful as the universal modal operator [12].

We adopt the standard possible worlds semantics, with models having a set of states or possible worlds W , and accessibility relations R_{Bel} , R_{Always} , and R_a respectively associated to the modal operators Bel , $Always$, and $After_a$.

R_{Bel} is reflexive, transitive and euclidean. In a state $w \in W$, the set of states $\{u \in W : wR_{Bel}u\}$ is called the *belief state* of the agent in w .

R_{Always} is reflexive and transitive, and contains every accessibility relation R_a associated to the action a . Hence R_{Always} contains the reflexive and transitive closure of the accessibility relations associated to the actions. (We might have supposed it is identical, but this is not necessary here.) Hence $Always$ allows to quantify over the states that are reachable by sequences of actions.

The R_a are not necessarily functions, hence actions might be non-deterministic.

3 Action laws

Action laws describe action types, as opposed to action instances. We suppose that to each action type a there is associated a set of *effect laws* and a set of *executability laws*. Basically, the former are of the form $A \rightarrow After_a C$ and the latter are of the form $A \rightarrow Feasible_a \top$ where A and C do not contain action operators.

3.1 Laws, not axioms

In most of the related approaches [16, 20, 22] it is supposed that actions laws are known by all agents. We do not make that hypothesis here, and admit that agents can entertain wrong beliefs not only about facts,

but also about action laws. I might e.g. consider that action a is always executable, while a is in fact executable only in context A ; I believe that a coin is biased, and that **toss** deterministically results in **Heads**, while in fact **toss** is non-deterministic.

Then what is the logical status of action laws? On the one hand, they are ‘eternal’ in the sense that they remain true after every action execution. On the other hand, we want to allow for false beliefs about laws. We therefore cannot expect action laws to be global axioms in the standard sense [4], because we do not want necessitation by Bel to be applied to them.

It is the modal operator *Always* which helps us out. $Always A \wedge \neg Bel A$ is consistent, and thus action laws might be ignored by agents. Nevertheless we have $Always A \rightarrow After_a Always A$, i.e. laws keep on being true after actions have been executed.

Given an action law A , we can now distinguish *objective action laws* of the form $Always A$ from *subjective action laws* of the form $Bel Always A$. For tossing with a biased coin we have the objective $Always After_{\text{toss}}(\text{Heads} \vee \text{Tails})$, and the subjective $Bel Always After_{\text{toss}} \text{Heads}$.

3.2 Non-sensing actions

A non-sensing action does not involve perception, and has no effect on beliefs others than those stipulated by the law itself. Its objective and subjective effect laws respectively take the form

$$\begin{aligned} & Always(A \rightarrow After_a C) \\ & Bel Always(A \rightarrow After_a C) \end{aligned}$$

where C must be *factual*, i.e. without any modal operator.

Take the action of tossing a coin eyes shut. As the agent cannot observe the effects of **toss**, he predicts them in an a priori way, according to his mental state and the action laws he assumes. The agent can thus be said to ‘mentally execute’ **toss**. Suppose the agent believes that **toss** is not biased. Then one of the subjective effects of **toss** is that both heads and tails are possible outcomes. Hence $Bel Always After_{\text{toss}} \neg Bel If \text{Heads}$.

Note that actions might have doxastic preconditions, such as knowing my phone number to be able to call me.

3.3 Test actions

Perception actions leave the environment unchanged. They can therefore be characterized as having the property $B \rightarrow After_a B$ for every objective formula B , and objective formula being a formula without occurrences of the doxastic modal operator Bel .

According to our definitions the distinction between non-sensing and perceptive actions is not exclusive, and at the intersection there is a class of actions that do not change the world: their effect laws only specify when they are inexecutable. Such actions can be seen as *test actions*: the action of testing that A fails if A is false, and succeeds else. In the latter case, the agent learns that A . Such actions are the simplest perception actions.

We note $test(A)$ the action of testing that A .² Testing-that actions can be characterised by the following *logical axioms* (cf. [7, 8]).

$$\begin{aligned}
 A &\rightarrow Feasible_{test(A)} \top && (\text{TestAct}_1) \\
 \neg A &\rightarrow After_{test(A)} \perp && (\text{TestAct}_2) \\
 &After_{test(A)} A && (\text{TestAct}_3) \\
 C &\rightarrow After_{test(A)} C \quad \text{if } C \text{ is objective} && (\text{TestAct}_4) \\
 &Feasible_{test(A)} C \rightarrow After_{test(A)} C && (\text{TestAct}_5)
 \end{aligned}$$

The first two axioms together say that $test(A)$ is executable iff A is true. Therefore learning that $test(A)$ has been executed amounts to learning that A . (TestAct₃) says that A holds after testing that A . Together with the more general principles of section 4 it will guarantee that testing that A leads to believing that A . (TestAct₄) expresses that test actions are perception actions, and the fifth says they are deterministic.

Note that we have formulated the principles for tests as logical axioms. They are thus known by the agent. (This is obtained by the rule of necessitation of KD45.)

Testing-that actions behave as expansions in the AGM-theory: they make shrink the belief state by ‘throwing out’ possible states.

$test(A)$ is indeed a non-sensing action: the reason is that when notified that $test(A)$ has been executed the agent already can immediately

² $test(A)$ is similar to the PDL test “ $A?$ ”. The difference is that for the latter $After_{A?} C$ is defined as $A \rightarrow C$. Hence such tests validate $B \rightarrow After_{A?} B$ for every formula. However, consider $B = \neg BelIf A$: the formula $\neg BelIf A \rightarrow After_{A?} \neg BelIf A$ should not be valid. Therefore such a principle must be restricted to factual B ’s, which is what we did here.

conclude that A is true in the environment, without further sensing. In some sense perception of A is ‘encoded’ in the public announcement of actions.

Nondeterministic composition of $test(A)$ and $test(\neg A)$ can ‘simulate’ the perhaps more intuitive action $testIf(A)$ of testing-if A : testing if the coin is heads amounts to nondeterministically choose between $test(A)$ and $test(\neg A)$ and execute the chosen action. Therefore testing-if can be viewed as an abbreviation of testing-that:

$$After_{testIf(A)}B \stackrel{\text{def.}}{=} After_{test(A)}B \wedge After_{test(\neg A)}B$$

It can then be proved that $Feasible_{testIf(A)}\top$ holds, as well as $After_{testIf(A)}BelIf A$ and $A \rightarrow After_{testIf(A)}BelA$. Note that while $test(A)$ is a non-sensing action, $testIf(A)$ is not.

We have provided a detailed analysis of the relationship between these two classes of test actions in [7]. We just note that an action $testIfCond(A, C)$ of testing whether A that is only executable under condition C can be simulated as well: $After_{testIfCond(A, C)}B$ can be viewed as an abbreviation of $C \rightarrow After_{testIf(A)}B$. Testing-that is thus expressive enough to capture a large class of perception actions.

For example, checking whether the outcome of tossing is heads or not is done by executing $testIf(\text{Heads})$. Hence we should have $After_{testIf(\text{Heads})}BelIf\text{Heads}$, or equivalently $After_{test(\text{Heads})}BelIf\text{Heads} \wedge After_{test(\neg\text{Heads})}BelIf\neg\text{Heads}$. This will be derived from the above axioms for testing-that by means of the axioms of the next section.

4 Solving the frame problem

From now on we make three hypotheses:

- All atomic actions are non-sensing actions (allowing thus for testing-that actions).
- Actions are *public*: their occurrence is perceived both correctly and completely. This means that (1) if an agent believes that some action a occurred then a indeed occurred (correctness), and (2) if a occurred then the agent believes a occurred (completeness).

This is supposed in many approaches, e.g. [16, 20, 22]. Intuitively, we can think of such actions as being publicly announced to all agents. Note that this makes that perception of testing-that actions is accurate.

- Non-sensing actions do not affect the agent’s cognition. Hence we exclude actions such as modifying the agent’s memory.

This hypothesis is related to our definition of effect laws for non-sensing actions, where the Bel cannot occur in postconditions.

Under these hypotheses simple and uniform axioms can be defined. In particular, revision can be reduced to the problem of accommodating executability of actions that were believed to be inexecutable: the agent believed that a is inexecutable, and learns that a has been executed.

4.1 The axiom for updates

The basic axiom accounting for updates has been given in [8]. It is the generalisation of the successor state axiom for knowledge of [16] to non-deterministic actions.

$$\begin{aligned} (\neg After_a \perp \wedge \neg Bel After_a \perp) \rightarrow \\ (Feasible_a Bel A \leftrightarrow Bel After_a A) \end{aligned} \quad (SSA_1)$$

Semantically, this means that the possible states after a are obtained by ‘mentally executing a ’, i.e. applying a to the states that were possible before a , and collecting the resulting states.

Consider e.g. $a = \text{toss}$. If we read the equivalence from the left to the right, (SSA_1) expresses that the agent’s uncertainty about the nondeterministic result of toss is preserved through its execution: before executing toss the agent ignores whether heads or tails will result, and this disbelief $\neg Bel If \text{Heads}$ is preserved through the execution of toss . (Remember toss is non-sensing.) From the right to the left, (SSA_1) expresses that for non-sensing actions there are no a posteriori beliefs the agent didn’t already hold a priori: if the agent believes that the coin is heads after toss then – as he had no means to check whether Heads is true – he must have believed before that the coin is biased, i.e. $Bel After_{\text{toss}} \text{heads}$.

As said in section 3, (SSA_1) is intuitive for testing-that actions, too: if after testing that Heads I believe that $\text{Heads} \wedge A'$ for some A' , and $test(\text{Heads})$ is executable, then I believe before $test(\text{Heads})$ that $\text{Heads} \wedge A'$ will be true afterwards.

As announced in the end of section 3, from (SSA_1) we can prove $After_{test(\text{Heads})} Bel \text{Heads}$ and $After_{test(\neg \text{Heads})} Bel \neg \text{Heads}$ using $(TestAct_3)$ and standard modal principles.

Remark 1 *The only case where the \rightarrow direction of the equivalence in (SSA_1) cannot be accepted is when a erases all or part of the memory of the agent (e.g. taking off the batteries of a robot). The \leftarrow direction is counter-intuitive only if the agent knows that a adds unjustified information to his memory. This is the case e.g. when he is hypnotized or takes drugs. We have excluded such extreme cases by hypothesis.*

Remark 2 *The plausibility of (SSA_1) heavily relies on our hypothesis that actions are public. For example, suppose $Bel(\neg\text{Heads} \wedge \neg\text{Tails})$. Suppose that a **toss**-action takes place but the agent isn't notified. Hence we have $After_{\text{toss}}Bel(\neg\text{Heads} \wedge \neg\text{Tails})$. But from axiom (SSA_1) and the action laws for **toss** it follows that $After_{\text{toss}}Bel(\text{Heads} \vee \text{Tails})$, which is contradictory.*

Remark 3 *(SSA_1) does not hold for testing-if actions. Indeed, suppose $A \wedge \neg Bel A$. From our axioms for testing we get the theorem $After_{\text{testIf}(A)}Bel A$, as well as $Feasible_{\text{testIf}(A)}\top$ (cf. section 3.3) and $BelFeasible_{\text{testIf}(A)}\top$ (obtained from the preceding by necessitation). Then with (SSA_1) we would get the counter-intuitive $BelAfter_{\text{testIf}(A)}A$.*

4.2 Enabling actions and the axiom for revision

Suppose the agent believes a is inexecutable, and learns that a has nevertheless been executed. Axiom (SSA_1) says nothing about that case. Such surprising occurrences of actions are indeed problematic, because the agent cannot just mentally execute a , but must first change his beliefs about a 's preconditions.

We propose to formalize the operation of changing beliefs about preconditions by the mental execution of a particular action whose effect is to make the executability preconditions of a true. Formally, we associate to every atomic action a an action $enable_a$, and we say that $enable_a$ makes a executable. We suppose that the action laws for every $enable_a$ are known by the agent.

We postulate that $enable_a$ can occur in every state. Hence we have the executability laws

$$\begin{array}{ll} \text{AlwaysFeasible}_{enable_a}\top & (\text{Def}_{\text{Always}X}(enable_a)) \\ \text{BelAlwaysFeasible}_{enable_a}\top & (\text{Def}_{\text{BelAlways}X}(enable_a)) \end{array}$$

This means that for every action there is at least one state where it is executable, which amounts to excluding from our actions the action $test(\perp)$ which is always inexecutable.

As $enable_a$ makes a executable, the set of executability laws for a determines the effect laws for $enable_a$. Let the set of objective and subjective executability laws for a respectively be

$$\begin{aligned} & \{Always(A_1 \rightarrow Feasible_a \top), \dots, Always(A_n \rightarrow Feasible_a \top)\} \\ & \{BelAlways(A'_1 \rightarrow Feasible_a \top), \dots, BelAlways(A'_m \rightarrow Feasible_a \top)\} \end{aligned}$$

Then we have:

$$\begin{aligned} & AlwaysAfter_{enable_a}(A_1 \vee \dots \vee A_n) && (\text{Def}_{E(enable_a)}1) \\ & BelAlwaysAfter_{enable_a}(A'_1 \vee \dots \vee A'_m) && (\text{Def}_{E(enable_a)}2) \end{aligned}$$

For example, for the above version of the toss action we get $After_{enable_{\text{toss}}}HasCoin$.

The precondition of executability of $test(A)$ being A , to make A executable amounts to updating the actual state by A . Hence we have the axiom

$$After_{enable_{test(A)}}A \quad (\text{TestAct}_6)$$

Then we postulate the following axiom for revision:

$$\begin{aligned} & (BelAfter_a \perp \wedge \neg After_a \perp) \rightarrow \\ & (Feasible_a BelA \leftrightarrow BelAfter_{enable_a} After_a A) \end{aligned} \quad (\text{SSA}_2)$$

Semantically, this means that if the agent believes a is inexecutable then the possible states after a are obtained by first enabling a in the possible states before a , then applying a to these states, and finally collecting the resulting states.

Let us illustrate (SSA₂) by our running example. Suppose the agent believes that some coin is biased: $BelAlwaysAfter_{\text{toss}}Heads$, and suppose he executes toss . toss being non-sensing, we obtain $After_{\text{toss}}BelHeads$ from (SSA₁) and the executability laws for toss .

Now suppose the agent subsequently perceives that $\neg Heads$ via learning the occurrence of $test(\neg Heads)$ (figure 1). As he believes $test(\neg Heads)$ to be inexecutable, we have $BelAfter_{test(\neg Heads)}\perp$, and (SSA₂) applies. Therefore

$$After_{test(\neg Heads)}Bel\neg BelHeads \leftrightarrow BelAfter_{enable_{test(\neg Heads)}}After_{test(\neg Heads)}\neg BelHeads$$

This illustrates that the agent has a consistent belief state after learning that $Heads$.

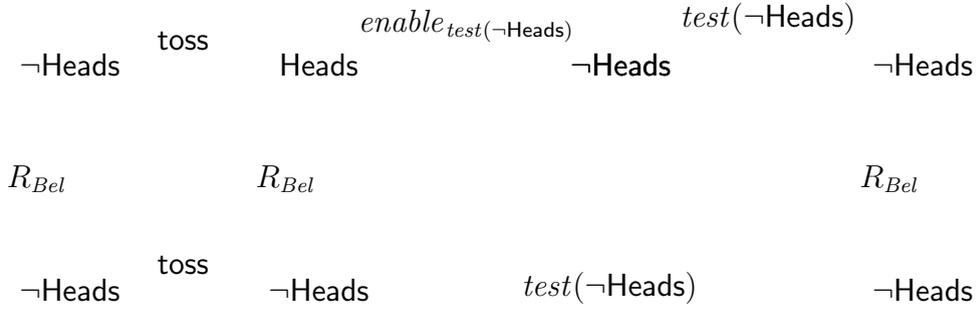


Figure 1: believing the coin is biased

From (SSA_1) and (SSA_2) it can be proved

$$Feasible_a Bel A \rightarrow After_a Bel A$$

expressing what may be called *doxastic determinism* of non-sensing actions.

4.3 Solving the non-doxastic frame problem

Given our successor state axioms we can reuse non-epistemic solutions to the frame problem. Just as Scherl and Levesque have applied Reiter’s solution [16] we use the solution of [2] in order to stay within propositional logic.

Which truths can be preserved after the performance of a non-sensing action? Our key concept is that of the *influence of an action*. If there exists a relation of influence between the action and an atom p , then p cannot be preserved. The relation $a \rightsquigarrow p$ is read “the action a influences the truth value of p ”. In our example, $\rightsquigarrow = \{\text{toss} \rightsquigarrow \text{Heads}, \text{toss} \rightsquigarrow \text{Tails}, \text{toss} \rightsquigarrow \text{HasCoin}\}$. Note that \rightsquigarrow is in the metalanguage. We extend \rightsquigarrow to formulas by stipulating that $a \rightsquigarrow A$ if there is an atom p occurring in A such that $a \rightsquigarrow p$.

The concept of influence (or dependence) is close to notions that have recently been studied in the field of reasoning about actions in order to solve the frame problem, e.g. Sandewall’s [15] occlusion, Thielscher’s [21] influence relation, or the ‘possibly changes’ operators of Giunchiglia *et al.* [5].

The preservation of formulas that are not influenced by an action is formalized by the influence-based *logical axiom*

$$A \rightarrow After_a A \quad \text{if } a \not\rightsquigarrow A \text{ and } A \text{ is factual} \quad (\text{Preserv})$$

This expresses that if a does not influence A then A is preserved. The restriction that A be factual avoids e.g. $Feasible_{a'}\top \rightarrow After_a Feasible_{a'}\top$, which is not necessarily the case because a might modify the executability preconditions of a' .

Remark 4 *As we have formulated (Preserv) as an axiom, it is not only true, but also believed by the agent. It might nevertheless be the case that the alleged influence of an action can differ from its real influence. We therefore should have an objective relation \rightsquigarrow_{obj} and a subjective relation \rightsquigarrow_{subj} .*

We can code these relations into our framework by putting $\rightsquigarrow = \rightsquigarrow_{obj} \cap \rightsquigarrow_s$, and adding frame axioms for those atoms where the relations differ: e.g. if $a \rightsquigarrow_{obj} p$ and $a \not\rightsquigarrow_{subj} p$, then we add to the set of effect axioms of a that p and $\neg p$ are believed to persist, i.e. $BelAlways(p \rightarrow After_a p)$ and $BelAlways(\neg p \rightarrow After_a \neg p)$.

It can also be argued that influence should be conditional. This can be encoded in a similar way by means of conditional frame axioms.

5 Discussion and related work

We have defined a modal logic of belief and nondeterministic actions where the agent's beliefs about the action laws might be inaccurate. Our central axioms (SSA₁) and (SSA₂) have the form of successor state axioms. When actions are deterministic, (SSA₁) is exactly the syntactic counterpart of the successor state axiom of [16].

In our framework belief-contravening information can be restricted to learning that some action a has been executed. Inconsistency with the agent's beliefs means that the agent believes a to be inexecutable, and learns that a has occurred. We have shown that such a revision operation can be implemented by an updating operation enabling the execution of a . Our second axiom (SSA₂) is a new solution that does not resort to orderings of plausibility.

When restricted to deterministic actions our axioms allow for regression. In the case of nondeterministic actions it is not clear how this could be done. An alternative is to use the famous modular completeness result due to Sahlqvist [14], which applies here almost immediately (because our axioms are of the required form). We thus get for free soundness and completeness results, as well as a tableau algorithm. If the tableau algorithm terminates then we get a decision procedure for our logic. We

are currently working on that, aiming at applying recent results on modal axioms of confluence and permutation (of which our SSA_1 and SSA_2 are instances).

In [20], Shapiro *et col.* add to the Scherl and Levesque framework a revision-like operation based on plausibility orderings. They define $BelA$ as truth of A in the most plausible among the possible states. If a sensing action eliminates the most plausible of the possible states, then previously less plausible states become the most plausible ones. The plausibility ordering should be kept fixed.

While being intuitively appealing, such a solution has several drawbacks. (1) As the authors note, it is restricted to deterministic actions. (2) “The specification of [the plausibility ordering] over the initial situation is the responsibility of the axiomatizer of the domain.” [20] This is particularly demanding because (3) in order to guarantee that after a the set of possible states is nonempty, the authors require the set of possible states to contain enough situations initially, restricting thus the agent’s ‘doxastic freedom’. (4) As pointed out in [3], such a solution to the problem of revision endangers the solution to the frame problem: suppose you firmly believe $p \wedge q$, say that the printer is on (p), and that the proceedings of KR’00 are on your shelf (q). Hence the set of most plausible states is the singleton $\{p, q\}$. Conforming to the recommendations of [20] you nevertheless consider the states $\{p, \neg q\}$, $\{\neg p, q\}$, and $\{\neg p, \neg q\}$ to be plausible, too. As you firmly believe $p \wedge q$, it seems natural not to distinguish these three states and to associate to them the same very low plausibility. Now suppose you learn that $\neg p$: according to [20] you then not only abandon your belief that p , but also your belief that q . It might be argued that the state $\{\neg p, q\}$ should be more plausible than $\{\neg p, \neg q\}$. It nevertheless seems to be fair to say that specifying a satisfactory plausibility ordering is a delicate task, involving a lot of imponderabilities in what concerns the relative plausibility of independent propositions. (5) The approach is unsatisfactory when applied to communication. Consider the following example: agent k is competent at p , and j is not. Agent i is completely ignorant initially: Hence all possible states are equally plausible for i . Then (under adequate hypotheses of cooperation) we can expect that when j asserts p , then i adopts p , i.e. $After_{asserts(j,p)} Bel p$. Moreover, as all states were equally plausible, p holds in every state possible for i . Therefore when subsequently k asserts $\neg p$, i will unavoidably move to an empty set of possible states.

It seems to us that another problem will show up if one tries to extend

the approach of [20] to account for imperfect knowledge about laws, which is what we have supposed here.

We have supposed that actions are public. Almost all the approaches make that assumption. It has been relaxed in [3], where drawbacks of the earlier solution in [10] are pointed out. We can integrate their solution via the introduction of particular atoms $Observable(a)$, for every action a . $Observable(a)$ is true in a state when the agent is able to observe a . Now $Observable(a)$ conditions axioms (SSA₁) and (SSA₂). In turn, if $Observable(a)$ is false then the agent's belief state doesn't change:

$$\neg Observable(a) \rightarrow (After_a Bel A \leftrightarrow Bel A)$$

which has the form of a successor state axiom.

The normative framework for belief revision being the AGM theory [1], which of their postulates do we satisfy? It can be shown that just as Shapiro *et col.* we satisfy the basic postulates (K*1) – (K*4), and (K*6). (The names of the postulates are as in [20]). If we define update actions as in [20] we satisfy the update postulates (K◊1), (K◊2), (K◊4), and (K◊5) just as there. If we define updating by A as $enable_{test(A)}$ then we moreover satisfy (K◊3).

References

- [1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50(2), June 1985.
- [2] Marcos A. Castilho, Olivier Gasquet, and Andreas Herzig. Formalizing action and change in modal logic I: the frame problem. *Journal of Logic and Computation*, 9(5):701–735, 1999.
- [3] Robert Demolombe and Maria del Pilar Pozos Parra. Formalisation de l'évolution de croyances dans le Calcul des Situations. In B. Chaib-draa and P. Enjalbert, editors, *Proc. Journées Francophones Modèles Formels de l'Interaction (MFI'01)*, volume 2, pages 205–217, 2001.
- [4] Melvin Fitting. *Proof methods for modal and intuitionistic logics*. D. Reidel, Dordrecht, 1983.
- [5] E. Giunchiglia, G. N. Kartha, and V. Lifschitz. Representing action: indeterminacy and ramifications. *Artificial Intelligence Journal*, 95, 1997.
- [6] David Harel. Dynamic logic. In D. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic*, volume II. D. Reidel Publishing Company, 1984.
- [7] Andreas Herzig, Jérôme Lang, and Thomas Polacsek. A modal logic for epistemic tests. In *Proc. Eur. Conf. on Artificial Intelligence (ECAI'2000)*, Berlin, August 2000.
- [8] Andreas Herzig, Jérôme Lang, Dominique Longin, and Thomas Polacsek. A logic for planning under partial observability. In *Proc. Seventeenth National Conf. on Artificial Intelligence (AAAI-2000)*, 2000.
- [9] Hirofumi Katsuno and Alberto O. Mendelzon. Propositional knowledge base revision and minimal change. Technical Report KRR-TR-90-3, Department of Computer Science, University of Toronto, March 1990.
- [10] Yves Lespérance, Hector J. Levesque, and Raymond Reiter. A Situation Calculus approach to modeling and programming agents. In A. Rao and M. Wooldridge, editors, *Foundations and theories of Rational Agents*. Kluwer Academic Publishers, 1999.
- [11] Hector J. Levesque. What is planning in the presence of sensing? In *Proc. of the 13th Nat. Conf. on Artificial Intelligence (AAAI-96)*, Portland, Oregon, August 1996.
- [12] Solomon Passy and Tinko Tinchev. An essay in combinatory dynamic logic. *Information and Computation*, 93:263–332, 1991.
- [13] Ray Reiter. The frame problem in the Situation Calculus: A simple solution (sometimes) and a completeness result for goal regression. In Vladimir Lifschitz, editor, *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy*, pages 359–380. Academic Press, San Diego, CA, 1991.

- [14] H. Sahlqvist. Completeness and correspondence in the first and second order semantics for modal logics. In S. Kanger, editor, *Proc. 3rd Scandinavian Logic Symposium*, volume 82 of *Studies in Logic*, 1975.
- [15] Erik Sandewall. *Features and Fluents*. Oxford University Press, 1994.
- [16] Richard Scherl and Hector J. Levesque. The frame problem and knowledge producing actions. In *Proc. Nat. Conf. on AI (AAAI'93)*, pages 689–695. AAAI Press, 1993.
- [17] Krister Segerberg. Two traditions in the logic of belief: bringing them together. to appear.
- [18] Krister Segerberg. Belief revision from the point of view of doxastic logic. *Bulletin of the IGPL*, 3:534–553, 1995.
- [19] S. Shapiro, Yves Lespérance, and Hector J. Levesque. Specifying communicative multi-agent systems. In W. Wobcke, M. Pagnucco, and C. Zhang, editors, *Agents and Multi-Agent Systems - Formalisms, Methodologies, and Applications*, pages 1–14. Springer-Verlag, LNAI 1441, 1998.
- [20] S. Shapiro, M. Pagnucco, Y. Lespérance, and H. J. Levesque. Iterated belief change in the situation calculus. In *Proc. Seventh Int. Conf. on Principles of Knowledge Representation and Reasoning (KR 2000)*, pages 527–538, 2000.
- [21] Michael Thielscher. Computing ramifications by postprocessing. In *Proc. 14th Int. Joint Conf. on Artificial Intelligence (IJCAI'95)*, pages 1994–2000, Montreal, Canada, 1995.
- [22] Michael Thielscher. Representing the knowledge of a robot. In A. Cohn, F. Giunchiglia, and B. Selman, editors, *Proc. KR'00*, pages 109–120. Morgan Kaufmann, 2000.