Projection in the Epistemic Situation Calculus with Belief Conditionals

Christoph Schwering and **Gerhard Lakemeyer**

Knowledge-Based Systems Group
RWTH Aachen University
(schwering,gerhard)@kbsg.rwth-aachen.de

Abstract

A fundamental task in reasoning about action and change is projection, which refers to determining what holds after a number of actions have occurred. A powerful method for solving the projection problem is regression, which reduces reasoning about the future to reasoning about the initial state. In particular, regression has played an important role in the situation calculus and its epistemic extensions. Recently, a modal variant of the situation calculus was proposed, which allows an agent to revise its beliefs based on so-called belief conditionals as part of its knowledge base. In this paper, we show how regression can be extended to reduce beliefs about the future to initial beliefs in the presence of belief conditionals. Moreover, we show how any remaining belief operators can be eliminated as well, thus reducing the belief projection problem to ordinary first-order entailments.

1 Introduction

A fundamental task in reasoning about action and change is *projection,* which refers to determining what holds after a number of actions have occurred. A powerful method for solving the projection problem is *regression,* which reduces reasoning about the future to reasoning about the initial state. In particular, regression has played an important role in the situation calculus (McCarthy 1963; Reiter 2001) and its epistemic extensions (Scherl and Levesque 2003; Lakemeyer and Levesque 2011).

One drawback of these epistemic extensions is that beliefs cannot be revised in the presence of conflicting information. For example, imagine a person in a restaurant who (mistakenly) believes that the restaurant's specialty is burger. She also believes that the specialties of Italian restaurants are pasta or pizza. If she then figures out that the restaurant is indeed Italian, she should revise her beliefs about the specialties of this particular restaurant. In (Shapiro et al. 2011) and, more recently, (Schwering and Lakemeyer 2014), it is shown how such belief revision scenarios can be modelled in the framework of the situation calculus using so-called *belief conditionals*. However, it remains open how to address the projection problem in the presence of belief conditionals.

In this paper, we provide a solution using regression. The work is carried out within the logic $\mathcal{ESB}$ of (Schwering and

Lakemeyer 2014), which is a modal variant of the situation calculus with belief revision proposed earlier by (Shapiro et al. 2011). We further draw on a result from (Levesque and Lakemeyer 2001) to reduce reasoning about beliefs to (non-modal) first-order entailments.

While the situation calculus is perhaps the most thoroughly studied action formalism, there are others such as the event calculus (Kowalski and Sergot 1989), the fluent calculus (Thielscher 1999), and the family of action languages $\mathcal{A}$ (Gelfond and Lifschitz 1993). An alternative to projection by regression is *progression* (Lin and Reiter 1997), where a query about the future is answered by first updating the initial knowledge base to account for the effects of actions and then checking whether the query follows from the resulting knowledge base. While progression may be preferable over regression for long sequences of action, it has the disadvantage that the resulting knowledge base is not always first-order representable. Besides (Shapiro et al. 2011; Schwering and Lakemeyer 2014), a number of other proposals such as (Demolombe and Pozos Parra 2006; Delgrande and Levesque 2012; Fang and Liu 2013) have dealt with extending action formalisms with belief revision in the spirit of AGM (Alchourron, Gärdenfors, and Makinson 1985). In (Pagnucco et al. 2013) a limited solution to the projection problem in action theories derived from (Shapiro et al. 2011) is proposed by a translation to Default Logic extended with preferences (Baumann et al. 2010), together with an implementation using Answer Set Programming (Gelfond 2008). The limitations include a finite domain and a restriction to non-nested beliefs. To the best of our knowledge, there currently are no general solutions to the projection problem in the context of belief revision.

The rest of the paper is organized as follows. The next section presents the logic $\mathcal{ESB}$. Next, we define Reiter's basic action theories within $\mathcal{ESB}$. In Section 4, we devise a regression theorem for $\mathcal{ESB}$. Afterwards we show how beliefs can be reduced to first-order entailments. Then we conclude.

2 The Logic $\mathcal{ESB}$

$\mathcal{ESB}$ is a logic for reasoning about actions and beliefs. It is based on the modal variant of the situation calculus $\mathcal{ES}$ (Lakemeyer and Levesque 2011) and the belief revision framework by (Shapiro et al. 2011). *Only-believing* allows to express all the agent believes through belief conditionals

in the spirit of Lewis' counterfactuals (Lewis 1973) in a way closely related to Pearl's System Z (Pearl 1990). These beliefs may be updated or revised by actions. Semantically, beliefs are represented by possible worlds ranked by their plausibility, where only the most plausible worlds determine the current beliefs. Given new information, some possible worlds may be dropped, thus leading to new beliefs.

The Language

The language $\mathcal{ESB}$ is a first-order modal language with equality and sorts of type action and object. It comes with a countably infinite set of *standard names* for both sorts, which can be thought of as special constants that satisfy the unique name assumptions and an infinitary version of domain closure.

The set of *terms* of sort action or object is the least set that contains all variables and standard names of the corresponding sort. Function symbols are left out to ease the presentation.

The set of *formulas* is the least set that contains $P(t_1, \ldots, t_k)$, $(t_1 = t_2)$, $(\alpha \wedge \alpha')$, $\neg\alpha$, $\forall x.\alpha$, $[r]\alpha$, $\square\alpha$, $\mathbf{P}\alpha$, $\mathbf{K}\alpha$, $\mathbf{B}\alpha$, $\mathbf{B}(\phi_1 \Rightarrow \psi_1)$, $\mathbf{O}(\alpha, \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\})$, where P is a predicate symbol, $t_1, \ldots, t_k$ are terms, r is an action term, x is variable, and α, α', $\phi_1, \ldots, \phi_m$, $\psi_1, \ldots, \psi_m$ are formulas. We read $[r]\alpha$ as "α holds after action r," $\square\alpha$ as "α holds after any sequence of actions," and $\mathbf{P}\alpha$ as "α was true before the last action." We read $\mathbf{K}\alpha$ as "α is known" and $\mathbf{B}\alpha$ as "α is believed." The belief conditional $\mathbf{B}(\phi \Rightarrow \psi)$ is read like Lewis' counterfactuals (Lewis 1973): "it is believed that, if ϕ was true, then ψ would be true." The only-believing operator $\mathbf{O}(\alpha, \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\})$ means that α is known, the counterfactuals $\phi_i \Rightarrow \psi_i$ are believed, but nothing else is known or believed.

We will use $\vee$, $\exists$, $\supset$, $\equiv$, TRUE, and FALSE as the usual abbreviations. Given natural numbers $p, p_1, \ldots, p_m$ and formulas $\alpha_1, \ldots, \alpha_m$, we let $\bigwedge_{i:p_i \geq p} \alpha_i$ stand for the conjunction of all α_i for which $p_i \geq p$. By α_n^x we denote the formula α with all free occurrences of x replaced with n. We sometimes write $\vec{t}$ for $t_1, \ldots, t_k$.

When we omit brackets, the operator precedence is in decreasing order $[r], \neg, \wedge, \vee, \exists, \forall, \supset, \equiv, \square$. Free variables are implicitly universally quantified with maximal scope unless said otherwise. Thus $\square[a]D(x) \equiv a = \text{odr} \wedge S(x) \vee D(x)$ means $\forall a.\forall x.\square(([a]D(x)) \equiv (a = \text{odr} \wedge S(x)) \vee D(x))$. We use sans-serif font for standard names like odr.

There are two distinguished predicates, $Poss(a)$ to capture an action's precondition and $SF(a)$ to represent an action's binary sensing result.

A formula with no $[r]$, $\square$, or $\mathbf{P}$ is called *static*. A formula with no $\mathbf{K}$, $\mathbf{B}$, or $\mathbf{O}$ is called *objective*. An objective, static formula with no $Poss$ or SF is called a *fluent* formula. A formula with no free variable is called a *sentence*.

The Semantics

Truth of a formula α is defined wrt a sequence actions z, the actual world w, and the agent's epistemic state f. We write $f, w, z \models \alpha$. More precisely, z is a sequence of action standard names that represent the actions executed so far and

thus denotes to the current situation. A world w is a function that determines a truth value $w[P(\vec{n}), z] \in \{0, 1\}$ for each ground atom $P(\vec{n})$ and sequence of actions z. An epistemic state f is a function that maps each plausibility $p \in \mathbb{N}$ to a set of worlds $f(p)$ considered possible at plausibility level p. The most plausible possible worlds are only in $f(0)$, less plausible worlds in $f(1)$, and so on. When z is the empty sequence $\langle\rangle$, we often omit it. When f or w are irrelevant to the truth of α we may leave them out as well.

We now present the objective part of the semantics:

1. $f, w, z \models P(n_1, \ldots, n_k)$ iff $w[P(n_1, \ldots, n_k), z] = 1$;

2. $f, w, z \models (n_1 = n_2)$ iff n_1 and n_2 are identical;

3. $f, w, z \models (\alpha_1 \wedge \alpha_2)$ iff $f, w, z \models \alpha_1$ and $f, w, z \models \alpha_2$;

4. $f, w, z \models \neg\alpha$ iff $f, w, z \not\models \alpha$;

5. $f, w, z \models \forall x.\alpha$ iff $f, w, z \models \alpha_n^x$
   for all standard names n of the corresponding sort;

6. $f, w, z \models [n]\alpha$ iff $f, w, z \cdot n \models \alpha$;

7. $f, w, z \models \square\alpha$ iff $f, w, z \cdot z' \models \alpha$ for all z';

8. $f, w, z \models \mathbf{P}\alpha$ iff $f, w, z' \models \alpha$ and $z = z' \cdot n$
   for some z' and n.

To characterize what is believed after an action sequence z, we define the relation $w' \simeq_z w$ for any given w (read: w' agrees with w on the sensing throughout z) as follows:

- $w' \simeq_{\langle\rangle} w$ for all worlds w';

- $w' \simeq_{z \cdot n} w$ iff $w' \simeq_z w$ and $w'[SF(n), z] = w[SF(n), z]$.

To ease the presentation of the following semantic rules, it is convenient to write $f, w, z \models \mathbf{K}^p\alpha$ as shorthand for "for all $w' \simeq_z w$, if $w' \in f(p)$, then $f, w', z \models \alpha$" for any $p \in \mathbb{N}$. In other words, the macro expresses knowledge at plausibility level p. Notice that $\mathbf{K}^p$FALSE holds if no world is considered possible at plausibility level p, and $\neg\mathbf{K}^p\neg\alpha$ means that there is at least one world which satisfies α at plausibility level p.

Similarly we write $f, w, z \models \mathbf{O}^p\alpha$ to abbreviate "for all $w' \simeq_z w$, $w' \in f(p)$ iff $f, w', z \models \alpha$" for any $p \in \mathbb{N}$. This macro thus expresses that α is all that is known at plausibility level p, which corresponds to *only-knowing* in (Levesque and Lakemeyer 2001).

The semantics of the epistemic operators follows:

9. $f, w, z \models \mathbf{K}\alpha$ iff for all $p \in \mathbb{N}$, $f, w, z \models \mathbf{K}^p\alpha$;

10. $f, w, z \models \mathbf{B}\alpha$ iff for all $p \in \mathbb{N}$,
    if $f, w, z \models \mathbf{K}^q$FALSE for all $q < p$,
    then $f, w, z \models \mathbf{K}^p\alpha$;

11. $f, w, z \models \mathbf{B}(\phi \Rightarrow \psi)$ iff for all $p \in \mathbb{N}$,
    if $f, w, z \models \mathbf{K}^q\neg\phi$ for all $q < p$,
    then $f, w, z \models \mathbf{K}^p(\phi \supset \psi)$;

12. $f, w, z \models \mathbf{O}(\alpha, \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\})$ iff
    for some $p_1, \ldots, p_m \in \mathbb{N} \cup \{\infty\}$,

    (a) $f, w, z \models \mathbf{O}^p(\alpha \wedge \bigwedge_{i:p_i \geq p}(\phi_i \supset \psi_i))$ for all $p \in \mathbb{N}$,

    (b) $f, w, z \models \mathbf{K}^p\neg\phi_i$ for all i and for all $p < p_i$,

    (c) $f, w, z \not\models \mathbf{K}^{p_i}\neg\phi_i$ for all i with $p_i \neq \infty$.

Intuitively, $\mathbf{B}\alpha$ holds if all worlds of the first non-empty plausibility level satisfy α. $\mathbf{B}(\phi \Rightarrow \psi)$ holds if all worlds of the first plausibility level that is consistent with ϕ satisfies $\phi \supset \psi$. The purpose of only-believing is to determine the agent's epistemic state:

Theorem 1 ((Schwering and Lakemeyer 2014)) *Let α, $\phi_1, \ldots, \phi_m$, and $\psi_1, \ldots, \psi_m$ be objective. Then there is a unique f such that $f \models \mathbf{O}(\alpha, \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\})$.*

A set of sentences Σ entails α (written as $\Sigma \models \alpha$) iff for every f, for every w, if $f, w \models \alpha'$ for every $\alpha' \in \Sigma$, then $f, w \models \alpha$. A sentence is valid (written as $\models \alpha$) iff $\{\} \models \alpha$.

3 Basic Action Theories

To axiomatize a dynamic domain we use the modal variant of Reiter's *basic action theories* (Lakemeyer and Levesque 2011; Reiter 2001). A basic action theory over a finite set of *fluent predicates* $\mathcal{F}$ consists of a static and a dynamic part. The dynamic axioms express when an action is executable (Σ_{pre}), how actions change the truth values of fluents (Σ_{post}), and how actions produce knowledge (Σ_{sense}):

- Σ_{pre} contains a single sentence $\Box Poss(a) \equiv \pi$ where π is a fluent formula;

- Σ_{post} contains a sentence $\Box[a]F(\vec{x}) \equiv \gamma_F$ for all $F \in \mathcal{F}$ where γ_F is a fluent formula;

- Σ_{sense} contains a single sentence $\Box SF(a) \equiv \varphi$ where φ is a fluent formula.

The sentences in Σ_{post} are called *successor state axioms* because they relate the state after an action a to the one before a. They incorporate Reiter's solution to the *frame problem* (Reiter 2001). We refer to the dynamic axioms as Ω.

The static part of a basic action theory expresses what is true initially in the real world (Σ_0) or what the agent believes to be true (Σ_{bel}), respectively:

- Σ_0 contains finitely many fluent sentences;

- Σ_{bel} contains finitely many belief conditionals $\phi \Rightarrow \psi$ where ϕ and ψ are fluent sentences.

In the rest of the paper, we use two basic action theories about the same set of fluents $\mathcal{F}$ to represent the *actual* world (Σ_0, Ω) and the agent's *beliefs* (Σ'_{bel}, Ω'), respectively. Then the projection problem is to decide whether this theory entails a sentence α involving actions and beliefs: $\Omega \wedge \Sigma_0 \wedge \mathbf{O}(\Omega', \Sigma'_{bel}) \models \alpha$.[1]

Example Imagine a person at a restaurant who wants to order the specialty. Her model of the domain's dynamics shall be correct, so we have one set of dynamic axioms Ω for both the actual and the believed basic action theory.

The specialties are represented by $S(x)$ and the ordered dishes by $D(x)$. The agent can order the specialty through action odr and she can sense whether or not she is at an Italian restaurant, indicated by I, through action loc. Hence

---
[1]We identify a finite set of sentences with their conjunction.

the dynamic axioms Ω are the union of:

$$\Sigma_{pre} = \{\Box Poss(a) \equiv \text{TRUE}\}$$
$$\Sigma_{post} = \{\Box[a]D(x) \equiv a = \text{odr} \wedge S(x) \vee D(x),$$
$$\Box[a]S(x) \equiv S(x), \quad \Box[a]I \equiv I\}$$
$$\Sigma_{sense} = \{\Box SF(a) \equiv a = \text{loc} \supset I\}$$

In fact, the restaurant is Italian, but the agent does not know that and believes the specialty to be burgers. But she also believes that if the restaurant was Italian, the specialty would be pasta or pizza. Thus we have:

$$\Sigma_0 = \{I\}$$
$$\Sigma'_{bel} = \{\text{TRUE} \Rightarrow S(x) \equiv x = \text{burger},$$
$$I \Rightarrow S(\text{pasta}) \vee S(\text{pizza})\}$$

Therefore the agent believes to get a burger after ordering the specialty. After sensing that the restaurant is Italian, she changes that belief to pasta or pizza. That is, she believes to be served a specialty x, but she is unsure what x is. Hence we have the following entailments:

$$\Omega \wedge \Sigma_0 \wedge \mathbf{O}(\Omega, \Sigma'_{bel}) \models [\text{odr}]\mathbf{B}D(\text{burger})$$
$$\Omega \wedge \Sigma_0 \wedge \mathbf{O}(\Omega, \Sigma'_{bel}) \models [\text{odr}][\text{loc}]\mathbf{B}(\exists x.(D(x) \wedge \neg \mathbf{B}D(x)))$$

We will use the second entailment to illustrate the techniques developed in the paper.

4 Projection by Regression

Regression rewrites a formula about future situations to a formula about the current situation (Reiter 2001). The idea is to successively replace subformulas $[r]F(\vec{t})$ with the right-hand side of F's successor state axiom γ_F. As we shall see in this section, we can similarly regress action occurrences in front of belief modalities. More precisely, our regression operator can handle $\mathcal{ESB}$ formulas with no $\Box$ or $\mathbf{O}$, provided predicates are taken from $\mathcal{F} \cup \{Poss, SF\}$. We call such a formula *regressable*.

As there is a faithful translation of $\mathcal{ES}$ basic action theories to Reiter's classical situation calculus (Lakemeyer and Levesque 2011), the regression result from this section carries over to the framework by (Shapiro et al. 2011).

Regression of Objective Formulas

For objective regressable formulas, our regression operator closely follows the one presented by (Lakemeyer and Levesque 2011) for the logic $\mathcal{ES}$, except that we add a rule to handle $\mathbf{P}\alpha$. We assume from now on that all basic action theories and formulas are *rectified*, that is, all quantifiers have distinct variables. Then the regression of an objective formula α after actions z wrt a basic action theory with dynamic axioms Ω is defined as follows:

1. $\mathcal{R}[z, (t_1 = t_2)] = (t_1 = t_2)$;

2. $\mathcal{R}[z, (\alpha_1 \wedge \alpha_2)] = (\mathcal{R}[z, \alpha_1] \wedge \mathcal{R}[z, \alpha_2])$;

3. $\mathcal{R}[z, \neg\alpha] = \neg\mathcal{R}[z, \alpha]$;

4. $\mathcal{R}[z, \forall x.\alpha] = \forall x.\mathcal{R}[z, \alpha]$;

5. $\mathcal{R}[z, [r]\alpha] = \mathcal{R}[z \cdot r, \alpha]$;

6. $\mathcal{R}[\langle\rangle, \mathbf{P}\alpha] = \text{FALSE}$;

   $\mathcal{R}[z \cdot r, \mathbf{P}\alpha] = \mathcal{R}[z, \alpha]$;

7. $\mathcal{R}[z, Poss(r)] = \mathcal{R}[z, \pi_r^a]$;

8. $\mathcal{R}[z, SF(r)] = \mathcal{R}[z, \varphi_r^a]$;

9. $\mathcal{R}[\langle\rangle, F(\vec{t})] = F(\vec{t})$;

   $\mathcal{R}[z \cdot r, F(\vec{t})] = \mathcal{R}[z, \gamma_F{}_{\vec{t}}^{\vec{x}}{}_r^a]$.

The most interesting rule is 9, which substitutes a fluent $F(\vec{t})$ after an action r with the appropriately instantiated right-hand side of F's successor state axiom $\gamma_F{}_{\vec{t}}^{\vec{x}}{}_r^a$. Notice that $Poss$ and SF atoms are replaced with their corresponding right-hand sides. It is easy to see that the regression of an objective regressable formula is independent of Ω:

Theorem 2 *Let α be an objective regressable sentence. Then: $\Omega \wedge \Sigma_0 \models \alpha$ iff $\Sigma_0 \models \mathcal{R}[\langle\rangle, \alpha]$.*

Proof. The proof is by induction on z and sub-induction on the length of α. We only do the proof for formulas of the form $\mathbf{P}\alpha$, for the other cases we refer to (Lakemeyer and Levesque 2011). Suppose $w_1 \models \Omega \wedge \Sigma_0$ and $w_2 \models \Sigma_0$.

For the base case let $z = \langle\rangle$. Then we have: $w_1, \langle\rangle \models \mathbf{P}\alpha$ iff (by semantic rule 8) $w_1, \langle\rangle \models \text{FALSE}$ iff (by definition of $\mathcal{R}$) $w_2 \models \mathcal{R}[\langle\rangle, \text{FALSE}]$ iff (by definition of $\mathcal{R}$) $w_2 \models \mathcal{R}[\langle\rangle, \mathbf{P}\alpha]$.

Now consider $z \cdot n$ for the induction step. Then we have: $w_1, z \cdot n \models \mathbf{P}\alpha$ iff (by semantic rule 8) $w_1, z \models \alpha$ iff (by induction) $w_2 \models \mathcal{R}[z, \alpha]$ iff (by definition of $\mathcal{R}$) $w_2 \models \mathcal{R}[z \cdot n, \mathbf{P}\alpha]$. $\qquad\square$

Example (cont.) We regress $[odr][loc]\exists x. D(x)$, which says that the agent gets a dish after ordering the specialty and sensing in what restaurant she is. Via $\exists x. \mathcal{R}[odr \cdot loc, D(x)]$ we get $\exists x. \mathcal{R}[odr, loc = odr \wedge S(x) \vee D(x)]$, which is equivalent to $\exists x. \mathcal{R}[odr, D(x)]$. Then we get $\exists x. \mathcal{R}[\langle\rangle, odr = odr \wedge S(x) \vee D(x)]$, which is equivalent to $\exists x.(S(x) \vee D(x))$.

Regression of Subjective Formulas

Now we extend the regression operator to handle epistemic operators for knowledge and belief. To begin with, the following two lemmas show that $\mathbf{K}\alpha$ and $\mathbf{B}\alpha$ can both be reduced to belief conditionals $\mathbf{B}(\phi \Rightarrow \psi)$:

Lemma 3 $\models \Box \mathbf{K}\alpha \equiv \mathbf{B}(\neg\alpha \Rightarrow \text{FALSE})$

Proof. For the only-if direction suppose $f, w, z \models \mathbf{K}\alpha$. Thus $f, w, z \models \mathbf{K}^p\alpha$ for all $p \in \mathbb{N}$, and therefore $f, w, z \models \mathbf{K}^p(\neg\alpha \supset \text{FALSE})$. Thus $f, w, z \models \mathbf{B}(\neg\alpha \Rightarrow \text{FALSE})$.

For the if direction suppose $f, w, z \models \mathbf{B}(\neg\alpha \Rightarrow \text{FALSE})$. We show that $f, w, z \models \mathbf{K}^p\alpha$ for all $p \in \mathbb{N}$ by induction. $f, w, z \models \mathbf{B}(\neg\alpha \Rightarrow \text{FALSE})$ implies $f, w, z \models \mathbf{K}^0(\neg\alpha \supset \text{FALSE})$, which simplifies to $f, w, z \models \mathbf{K}^0\alpha$. For the induction step let $p \in \mathbb{N}$ and $f, w, z \models \mathbf{K}^q\alpha$ for all $q < p$. Thus $f, w, z \models \mathbf{K}^q\neg\neg\alpha$, and since $f, w, z \models \mathbf{B}(\neg\alpha \Rightarrow \text{FALSE})$, $f, w, z \models \mathbf{K}^p(\neg\alpha \supset \text{FALSE})$. Thus $f, w, z \models \mathbf{K}^p\alpha$. $\qquad\square$

Lemma 4 $\models \Box \mathbf{B}\alpha \equiv \mathbf{B}(\text{TRUE} \Rightarrow \alpha)$

Proof. Follows from semantic rules 10 and 11. $\qquad\square$

Hence the really interesting epistemic operator is the belief conditional $\mathbf{B}(\phi \Rightarrow \psi)$. The following result can be considered a successor state axiom for belief conditionals, as it relates beliefs *after* some action to the beliefs *before* that action. It thus lies the ground of regression of beliefs:

Theorem 5

$$\models \Box[a]\mathbf{B}(\phi \Rightarrow \psi) \equiv$$
$$SF(a) \wedge \mathbf{B}(\ SF(a) \wedge [a]\phi \Rightarrow [a]\psi) \vee$$
$$\neg SF(a) \wedge \mathbf{B}(\neg SF(a) \wedge [a]\phi \Rightarrow [a]\psi)$$

Proof. By contradiction. Let $f, w, z \models SF(n)$; the case for $\neg SF(a)$ is completely analogous.

For the only-if direction let $f, w, z \models [n]\mathbf{B}(\phi_n^a \Rightarrow \psi_n^a)$ but $f, w, z \not\models \mathbf{B}(SF(n) \wedge [n]\phi_n^a \Rightarrow [n]\psi_n^a)$. Then there is some $p \in \mathbb{N}$ such that for all $q < p$, $f, w, z \models \mathbf{K}^q\neg(SF(n) \wedge [n]\phi_n^a)$ and $f, w, z \not\models \mathbf{K}^p(SF(n) \wedge [n]\phi_n^a \supset [n]\psi_n^a)$. Therefore firstly for all $q < p$, for all $w' \in f(q)$ with $w' \simeq_z w$, $f, w', z \models \neg(SF(n) \wedge [n]\phi_n^a)$ and secondly for some $w'' \in f(p)$ with $w'' \simeq_z w$, $f, w'', z \not\models SF(n) \wedge [n]\phi_n^a \supset [n]\psi_n^a$, which means that $f, w'', z \models SF(n) \wedge [n]\phi_n^a \wedge \neg[n]\psi_n^a$. Thus for all $q < p$, for all $w' \in f(q)$ with $w' \simeq_z w$, we have $w' \not\simeq_{z \cdot n} w$ or $f, w', z \cdot n \models \neg\phi_n^a$, but for some $w'' \in f(p)$ with $w'' \simeq_{z \cdot n} w$, we have $f, w'', z \cdot n \models \phi_n^a \wedge \neg\psi_n^a$. Therefore $f, w, z \cdot n \models \mathbf{K}^q\neg\phi_n^a$ for all $q < p$ and $f, w, z \cdot n \not\models \mathbf{K}^p(\phi_n^a \supset \psi_n^a)$. Hence $f, w, z \cdot n \not\models \mathbf{B}(\phi_n^a \Rightarrow \psi_n^a)$, and this is $f, w, z \not\models [n]\mathbf{B}(\phi_n^a \Rightarrow \psi_n^a)$, which is a contradiction.

For the if direction suppose $f, w, z \models \mathbf{B}(SF(n) \wedge [n]\phi_n^a \Rightarrow [n]\psi_n^a)$ but $f, w, z \not\models [n]\mathbf{B}(\phi_n^a \Rightarrow \psi_n^a)$. Then there is some $p \in \mathbb{N}$ such that for all $q < p$, $f, w, z \cdot n \models \mathbf{K}^q\neg\phi_n^a$ and $f, w, z \cdot n \not\models \mathbf{K}^p(\phi_n^a \supset \psi_n^a)$. Therefore for all $q < p$, there is no $w' \in f(q)$ with $w' \simeq_{z \cdot n}$ and $f, w', z \cdot n \models \phi_n^a$ and there is some $w'' \in f(p)$ with $w'' \simeq_{z \cdot n}$ such that $f, w'', z \cdot n \not\models \phi_n^a \supset \psi_n^a$. Thus for all $q < p$, for all $w' \in f(q)$ with $w' \simeq_z w$, $f, w', z \models \neg(SF(n) \wedge [n]\phi_n^a)$ and for some $w'' \in f(p)$ with $w'' \simeq_z w$, $f, w'', z \models SF(n) \wedge [n]\phi_n^a \wedge \neg[n]\psi_n^a$. Hence for all $q < p$ we have $f, w, z \models \mathbf{K}^q\neg(SF(n) \wedge [n]\phi_n^a)$ and $f, w, z \not\models \mathbf{K}^p(SF(n) \wedge [n]\phi_n^a \supset [n]\psi_n^a)$. Therefore $f, w, z \not\models \mathbf{B}(SF(n) \wedge [n]\phi_n^a \Rightarrow [n]\psi_n^a)$, which is a contradiction. $\qquad\square$

In our example, the agent believes to get pasta or pizza after she finds out the restaurant is Italian. By Theorem 5 this is because prior to that, she had believed that if she *was* at an Italian restaurant, she would get pasta or pizza.

We now extend the regression operator to deal with beliefs. To this end, we add to the regression operator $\mathcal{R}[\Omega', \Omega, z, \alpha]$ two new arguments with the understanding that the previous rules are retrofitted with these arguments as well and regression is performed wrt Ω. Intuitively, Ω' is what the agent believes to be the world's laws of dynamics, while Ω represents the actual laws of dynamics. Then epistemic modalities are regressed as follows:

10. $\mathcal{R}[\Omega', \Omega, z, \mathbf{K}\alpha] = \mathcal{R}[\Omega', \Omega, z, \mathbf{B}(\neg\alpha \Rightarrow \text{FALSE})]$;

11. $\mathcal{R}[\Omega', \Omega, z, \mathbf{B}\alpha] = \mathcal{R}[\Omega', \Omega, z, \mathbf{B}(\text{TRUE} \Rightarrow \alpha)]$;

12. $\mathcal{R}[\Omega', \Omega, \langle\rangle, \mathbf{B}(\phi \Rightarrow \psi)] = \mathbf{B}(\phi' \Rightarrow \psi')$
    where $\phi' = \mathcal{R}[\Omega', \Omega', \langle\rangle, \phi]$ and $\psi' = \mathcal{R}[\Omega', \Omega', \langle\rangle, \psi]$;

    $\mathcal{R}[\Omega', \Omega, z \cdot r, \mathbf{B}(\phi \Rightarrow \psi)] = \mathcal{R}[\Omega', \Omega, z, \beta_r^a]$
    where β is the right-hand side of Theorem 5.

We need some preparatory work before we can extend Theorem 2 to the subjective case in Theorem 12:

Definition 6 For a world w, the world w_Ω satisfies the following conditions:

- $w_\Omega[F(\vec{n}), \langle\rangle] = w[F(\vec{n}), \langle\rangle]$ and
  $w_\Omega[F(\vec{n}), z \cdot n'] = 1$ iff $w_\Omega, z \models \gamma_F{}^{\vec{x}}_{\vec{n}}{}^{a}_{n'}$ for all $F \in \mathcal{F}$;
- $w_\Omega[P(\vec{n}), z] = w[P(\vec{n}), z]$ for all $P \notin \mathcal{F} \cup \{Poss, SF\}$;
- $w_\Omega[Poss(n), z] = 1$ iff $w_\Omega, z \models \pi_n^a$;
- $w_\Omega[SF(n), z] = 1$ iff $w_\Omega, z \models \varphi_n^a$.

Definition 7 For an epistemic state f, f_Ω is such that $f_\Omega(p) = \{w_\Omega \mid w \in f(p)\}$ for all $p \in \mathbb{N}$.

In other words, w_Ω and f_Ω comply with w and f for the initial truth values, which in the face of actions change according to Ω. They are uniquely determined by the basic action theory and w or f, respectively, because successor state axioms prevent nondeterminism. We thus have the following lemmas, which generalize results from (Lakemeyer and Levesque 2011) to the case of beliefs:

Lemma 8 If $w \models \Omega \wedge \Sigma_0$, then $w_\Omega = w$.

Lemma 9 If $w \models \Sigma_0$, then $w_\Omega \models \Omega \wedge \Sigma_0$.

Lemma 10 If $f \models \mathbf{O}(\text{TRUE}, \Sigma_{bel})$, then $f_\Omega \models \mathbf{O}(\Omega, \Sigma_{bel})$.

Lemma 11 Let α be a regressable sentence.
Then: $f, w \models \mathcal{R}[\Omega', \Omega, z, \alpha]$ iff $f_{\Omega'}, w_\Omega, z \models \alpha$.

Proof. The proof is by induction on z and sub-induction on the length α where the length $\mathbf{B}(\phi \Rightarrow \psi)$ is the length of $\phi \supset \psi$ plus 1, the length of $\mathbf{K}\alpha$ is the length of $\mathbf{B}(\neg\alpha \Rightarrow \text{FALSE})$ plus 1, and the length $\mathbf{B}\alpha$ is the length of $\mathbf{B}(\text{TRUE} \Rightarrow \alpha)$ plus 1.

For the base case let $z = \langle\rangle$. For objective formulas we refer to (Lakemeyer and Levesque 2011).

For formulas $\mathbf{B}(\phi \Rightarrow \psi)$ we have: $f_{\Omega'}, w_\Omega, \langle\rangle \models \mathbf{B}(\phi \Rightarrow \psi)$ iff (by semantic rule 11) for all $p \in \mathbb{N}$, if for all $q < p$, for all $w' \in f_{\Omega'}(q)$, $f_{\Omega'}, w', \langle\rangle \models \neg\phi$, then for all $w' \in f_{\Omega'}(p)$, $f_{\Omega'}, w', \langle\rangle \models \phi \supset \psi$ iff (by definition of $f_{\Omega'}$) for all $p \in \mathbb{N}$, if for all $q < p$, for all $w' \in f(q)$, $f_{\Omega'}, w'_{\Omega'}, \langle\rangle \models \neg\phi$, then for all $w' \in f(p)$, $f_{\Omega'}, w'_{\Omega'}, \langle\rangle \models \phi \supset \psi$ iff (by sub-induction) for all $p \in \mathbb{N}$, if for all $q < p$, for all $w' \in f(q)$, $f, w' \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \neg\phi]$, then for all $w' \in f(p)$, $f, w' \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \phi \supset \psi]$ iff (by definition of $\mathcal{R}$) for all $p \in \mathbb{N}$, if for all $q < p$, for all $w' \in f(q)$, $f, w' \models \neg\mathcal{R}[\Omega', \Omega, \langle\rangle, \phi]$, then for all $w' \in f(p)$, $f, w' \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \phi] \supset \mathcal{R}[\Omega', \Omega, \langle\rangle, \psi]$ iff (by semantic rule 11) $f, w \models \mathbf{B}(\mathcal{R}[\Omega', \Omega, \langle\rangle, \phi] \Rightarrow \mathcal{R}[\Omega', \Omega, \langle\rangle, \psi])$ iff (by definition of $\mathcal{R}$) $f, w \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \mathbf{B}(\phi \Rightarrow \psi)]$.

For formulas $\mathbf{K}\alpha$ we have: $f_{\Omega'}, w_\Omega, \langle\rangle \models \mathbf{K}\alpha$ iff (by Lemma 3) $f_{\Omega'}, w_\Omega, \langle\rangle \models \mathbf{B}(\neg\alpha \Rightarrow \text{FALSE})$ iff (by sub-induction) $f, w \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \mathbf{B}(\neg\alpha \Rightarrow \text{FALSE})]$ iff (by definition of $\mathcal{R}$) $f, w \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \mathbf{K}\alpha]$.

For formulas $\mathbf{B}\alpha$ we have: $f_{\Omega'}, w_\Omega, \langle\rangle \models \mathbf{B}\alpha$ iff (by Lemma 4) $f_{\Omega'}, w_\Omega, \langle\rangle \models \mathbf{B}(\text{TRUE} \Rightarrow \alpha)$ iff (by sub-induction) $f, w \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \mathbf{B}(\text{TRUE} \Rightarrow \alpha)]$ iff (by definition of $\mathcal{R}$) $f, w \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \mathbf{B}\alpha]$.

Now consider $z \cdot n$ for the main induction step. For all except for $\mathbf{B}(\phi \Rightarrow \psi)$ the sub-induction is the same as in

the base case. $f_{\Omega'}, w_\Omega, z \cdot n \models \mathbf{B}(\phi \Rightarrow \psi)$ iff (by Theorem 5) $f_{\Omega'}, w_\Omega, z \models \beta_n^a$ where β is the right-hand side of from Theorem 5 iff (by induction) $f, w \models \mathcal{R}[\Omega', \Omega, z, \beta_n^a]$ iff (by definition of $\mathcal{R}$) $f, w \models \mathcal{R}[\Omega', \Omega, z, \mathbf{B}(\phi \Rightarrow \psi)]$. $\square$

We are now ready for the regression theorem for the situation calculus with beliefs, which reduces reasoning about *future* situations to reasoning about the *initial* situation:

Theorem 12 *Let α be a regressable sentence. Then:*

$$\Omega \wedge \Sigma_0 \wedge \mathbf{O}(\Omega', \Sigma'_{bel}) \models \alpha \qquad \text{iff}$$
$$\Sigma_0 \wedge \mathbf{O}(\text{TRUE}, \Sigma'_{bel}) \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \alpha]$$

Proof. For the only-if direction suppose the left-hand side holds and $f, w \models \Sigma_0 \wedge \mathbf{O}(\text{TRUE}, \Sigma'_{bel})$. By Lemmas 9 and 10, $f_{\Omega'}, w_\Omega \models \Omega \wedge \Sigma_0 \wedge \mathbf{O}(\Omega', \Sigma'_{bel})$. Since $f_{\Omega'}, w_\Omega \models \alpha$ and by Lemma 11, $f, w \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \alpha]$.

For the if direction suppose the right-hand side holds and $f, w \models \Omega \wedge \Sigma_0 \wedge \mathbf{O}(\Omega', \Sigma'_{bel})$. Then $w \models \Sigma_0$. Let $f' \models \mathbf{O}(\text{TRUE}, \Sigma'_{bel})$. By assumption, $f', w \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \alpha]$. By Lemma 11, $f'_{\Omega'}, w_\Omega \models \alpha$. By Lemmas 9 and 10, $f'_{\Omega'}, w_\Omega \models \Omega \wedge \Sigma_0 \wedge \mathbf{O}(\Omega', \Sigma'_{bel})$. By Lemma 8, $w_\Omega = w$, and by Theorem 1, $f'_{\Omega'} = f$. $\square$

Example (cont.) Consider $[\text{odr}][\text{loc}]\mathbf{B}(\exists x.(D(x) \wedge \neg\mathbf{B}D(x)))$, which says that after ordering a specialty and sensing she is in an Italian restaurant, the agent believes that she will get a dish, but is unsure which one. Regressing this sentence leads to a disjunction for the sensing outcomes for odr and loc. To ease the presentation, we only consider the positive outcomes here. Since $\mathcal{R}[\langle\rangle, SF(\text{odr})]$ is vacuously TRUE, we can simplify the intermediate regression result to

$$\mathcal{R}[\langle\rangle, [\text{odr}]SF(\text{loc})] \wedge$$
$$\mathcal{R}[\langle\rangle, \mathbf{B}([\text{odr}]SF(\text{loc}) \Rightarrow [\text{odr}][\text{loc}]\exists x.(D(x) \wedge \neg\mathbf{B}D(x)))].$$

Regression then proceeds with the antecedent and consequent within $\mathbf{B}$, where we can reuse the objective regression example from the previous subsection, and eventually obtain a static sentence equivalent to

$$I \wedge \mathbf{B}(I \Rightarrow \exists x.((S(x) \vee D(x)) \wedge \neg\mathbf{B}(I \Rightarrow S(x) \vee D(x)))).$$

In the next section we will see that this sentence is indeed entailed by our example $\Sigma_0 \wedge \mathbf{O}(\text{TRUE}, \Sigma'_{bel})$.

5 Reducing Beliefs to Non-Modal Reasoning

In this section we reduce reasoning about beliefs to ordinary first-order reasoning. Combined with regression, this means that all modal operators can be eliminated from the query.

By Theorem 12, the projection problem can be reduced to the entailment $\Sigma_0 \wedge \mathbf{O}(\text{TRUE}, \Sigma'_{bel}) \models \alpha$ where α is the result of regressing some regressable formula. It is easy to see that α is static and contains no epistemic operators other than $\mathbf{B}(\phi \Rightarrow \psi)$. We call such a formula *regressed*. The goal is now to reduce entailment of $\mathbf{B}(\phi \Rightarrow \psi)$ by $\mathbf{O}(\text{TRUE}, \Sigma'_{bel})$ to objective entailments. To this end, we first need to represent $\mathbf{O}(\text{TRUE}, \Sigma'_{bel})$ by means of objective sentences:

Definition 13 Let $\delta = \mathbf{O}(\alpha, \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\})$ and $f \models \delta$. Then $\vec{\gamma} = \gamma_0, \ldots, \gamma_k$ is an *objective representation* of δ iff γ_p is objective, $f(p) = \{w \mid w \models \gamma_p\}$ for all $0 \leq p \leq k$, and $f(p) = \{w \mid w \models \gamma_k\}$ for all $p > k$.

Hence, γ_p represents what is believed at the pth plausibility level, plus what is believed at all following levels in case γ_p is the last sentence of $\vec{\gamma}$. Fortunately, an objective representation of $\mathbf{O}(\text{TRUE}, \Sigma'_{bel})$ always exists and is straightforward to generate:

Theorem 14 *Let* $\delta = \mathbf{O}(\alpha, \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\})$ *for objective* α, ϕ_i, ψ_i. *A representation* $\vec{\gamma}$ *of* δ *exists, is unique (up to logical equivalence), and can be generated using first-order entailment.*

Proof. To generate $\vec{\gamma}$, let $p_1 = 0, \ldots, p_m = 0$. Then repeat the following two steps from $p = 0$ to $p = m$:

- Let $\gamma_p = \alpha \wedge \bigwedge_{i:p_i \geq p}(\phi_i \supset \psi_i)$.

- For all i, if $\gamma_p \models \neg\phi_i$, let $p_i = p + 1$.

Now let $f(p) = \{w \mid w \models \gamma_p\}$ for all $0 \leq p \leq m$ and $f(p) = \{w \mid w \models \gamma_m\}$ for all $p > m$. Then f satisfies semantic rule 12. Hence $\vec{\gamma}$ is an objective representation.

For uniqueness of $\vec{\gamma}$, let $\vec{\gamma}'$ be another representation. Since $f \models \delta$ is unique by Theorem 1, $\{w \mid w \models \gamma_p\} = \{w \mid w \models \gamma'_p\}$ for all $0 \leq p \leq m$ and therefore $\models \gamma_p \equiv \gamma'_p$. $\square$

We will now reduce testing if $\mathbf{O}(\text{TRUE}, \Sigma'_{bel})$ entails $\mathbf{B}(\phi \Rightarrow \psi)$ to non-modal entailments involving only γ_p, ϕ, and ψ for an objective representation $\vec{\gamma}$ of $\mathbf{O}(\text{TRUE}, \Sigma'_{bel})$. First we need to address variables that occur freely in ϕ or ψ, that is, variables quantified outside of $\mathbf{B}$. The idea is to produce an objective formula $\text{RES}[\![\alpha, \beta]\!]$ which is valid for those instances of the free variables for which $\beta \supset \alpha$ is valid (Levesque and Lakemeyer 2001):

Definition 15 Let α be a regressed formula and β be an objective sentence.

- If α has no free variables:

$$\text{RES}[\![\alpha, \beta]\!] = \begin{cases} \text{TRUE} & \text{if } \beta \models \alpha \\ \text{FALSE} & \text{otherwise} \end{cases}$$

- If x is a free variable in α, $n_1, \ldots, n_k$ are the standard names of the same sort as x occurring in α and β, and n' is a standard name of the same sort as x different from $n_1, \ldots, n_k$:

$$\text{RES}[\![\alpha, \beta]\!] =$$
$$((x = n_1) \wedge \text{RES}[\![\alpha^x_{n_1}, \beta]\!]) \vee \ldots \vee$$
$$((x = n_k) \wedge \text{RES}[\![\alpha^x_{n_k}, \beta]\!]) \vee$$
$$((x \neq n_1) \wedge \ldots \wedge (x \neq n_k) \wedge \text{RES}[\![\alpha^x_{n'}, \beta]\!]^{n'}_x)$$

Recall that $\mathbf{B}(\phi \Rightarrow \psi)$ is true iff for all $p \in \mathbb{N}$, if $\neg\phi$ holds at all plausibility levels $q < p$, then $\phi \supset \psi$ holds at plausibility level p. Provided that ϕ and ψ are objective, we can reformulate whether $\mathbf{B}(\phi \Rightarrow \psi)$ holds wrt an objective representation $\vec{\gamma}$ of the epistemic state: for all $0 \leq p \leq m$, if $\text{RES}[\![\neg\phi, \gamma_q]\!]$ is valid for all $q < p$, then $\text{RES}[\![\phi \supset \psi, \gamma_p]\!]$ is valid. We can thus define a procedure $\|\alpha\|_{\vec{\gamma}}$ to eliminate all $\mathbf{B}(\phi \Rightarrow \psi)$ from α. To cope with non-objective ϕ or ψ, we simply apply $\| \cdot \|_{\vec{\gamma}}$ recursively:

Definition 16 Let α be a regressed formula and $\vec{\gamma} = \gamma_0, \ldots, \gamma_m$ be objective sentences. Then $\|\alpha\|_{\vec{\gamma}}$ is defined inductively:

1. $\|\alpha\|_{\vec{\gamma}} = \alpha$ if α is an objective formula

2. $\|\neg\alpha\|_{\vec{\gamma}} = \neg\|\alpha\|_{\vec{\gamma}}$

3. $\|(\alpha_1 \wedge \alpha_2)\|_{\vec{\gamma}} = (\|\alpha_1\|_{\vec{\gamma}} \wedge \|\alpha_2\|_{\vec{\gamma}})$

4. $\|\forall x.\alpha\|_{\vec{\gamma}} = \forall x.\|\alpha\|_{\vec{\gamma}}$

5. $\|\mathbf{B}(\phi \Rightarrow \psi)\|_{\vec{\gamma}} = \bigwedge_{p=0}^{m} \big(\big(\bigwedge_{q=0}^{p-1} \text{RES}[\![\|\neg\phi\|_{\vec{\gamma}}, \gamma_q]\!] \big) \supset \text{RES}[\![\|\phi \supset \psi\|_{\vec{\gamma}}, \gamma_p]\!] \big)$

Theorem 17 *Let* α *be a regressable sentence and* $\vec{\gamma}$ *be an objective representation of* $\mathbf{O}(\text{TRUE}, \Sigma'_{bel})$. *Then:*

$$\Omega \wedge \Sigma_0 \wedge \mathbf{O}(\Omega', \Sigma'_{bel}) \models \alpha \qquad \textit{iff}$$
$$\Sigma_0 \models \|\mathcal{R}[\Omega', \Omega, \langle\rangle, \alpha]\|_{\vec{\gamma}}$$

Proof sketch. By Theorem 12, the left-hand side holds iff $\Sigma_0 \wedge \mathbf{O}(\text{TRUE}, \Sigma'_{bel}) \models \mathcal{R}[\Omega', \Omega, \langle\rangle, \alpha]$. The remaining equivalence can be established through results similar to Lemmas 7.2.2 and 7.3.2 from (Levesque and Lakemeyer 2001). It needs to be shown that rule 5 of $\| \cdot \|_{\vec{\gamma}}$ mimics the semantics of $\mathbf{B}(\phi \Rightarrow \psi)$. Roughly, this is true because $\vec{\gamma}$ is sufficient to represent all plausibility levels induced by $\mathbf{O}(\text{TRUE}, \Sigma'_{bel})$. $\square$

Example (cont.) In the previous section, we regressed the query $[\text{odr}][\text{loc}]\mathbf{B}(\exists x.(D(x) \wedge \neg\mathbf{B}D(x)))$ and thus obtained $I \wedge \mathbf{B}(I \Rightarrow \exists x.((S(x) \vee D(x)) \wedge \neg\mathbf{B}(I \Rightarrow S(x) \vee D(x))))$. We will now eliminate the remaining belief modalities using the results from this section.

An objective representation of $\mathbf{O}(\text{TRUE}, \Sigma'_{bel})$ is

$$\gamma_0 = \neg I \wedge (S(x) \equiv x = \text{burger})$$
$$\gamma_1 = I \supset (S(\text{pasta}) \vee S(\text{pizza}))$$
$$\gamma_2 = \text{TRUE}.$$

Note that $\neg I$ occurs in γ_0 because $S(x) \equiv x = \text{burger}$ is inconsistent with the consequent of $I \Rightarrow S(\text{pasta}) \vee S(\text{pizza})$.

Recall that $\| \cdot \|_{\vec{\gamma}}$ works its way from the inside to the outside. $\|\mathbf{B}(I \Rightarrow S(x) \vee D(x))\|_{\vec{\gamma}}$ expands to the sentence

$$(\text{RES}[\![I \supset S(x) \vee D(x), \gamma_0]\!]) \wedge$$
$$(\text{RES}[\![\neg I, \gamma_0]\!] \supset \text{RES}[\![I \supset S(x) \vee D(x), \gamma_1]\!]) \wedge$$
$$(\text{RES}[\![\neg I, \gamma_0]\!] \wedge \text{RES}[\![\neg I, \gamma_1]\!] \supset$$
$$\text{RES}[\![I \supset S(x) \vee D(x), \gamma_2]\!]).$$

The first and third conjuncts are equivalent to TRUE because $\gamma_0 \models \neg I$ and $\gamma_1 \not\models \neg I$, respectively. However, the second conjunct is equivalent to FALSE because the antecedent $\text{RES}[\![\neg I, \gamma_0]\!]$ is TRUE, but $\text{RES}[\![I \supset S(x) \vee D(x), \gamma_1]\!]$ is equivalent to FALSE because $\gamma_1 \not\models I \supset (S(x) \vee D(x))^x_n$ for all $n \in \{\text{pasta}, \text{pizza}, \text{tantuni}\}$, which are the standard names from γ_1 and the query plus one additional (tantuni).

$\|\mathbf{B}(I \Rightarrow \exists x.((S(x) \vee D(x)) \wedge \neg\text{FALSE}))\|_{\vec{\gamma}}$ expands to a similar formula except that x is bound within the $\text{RES}[\![\cdot, \cdot]\!]$ operator, and hence is equivalent to TRUE.

Therefore to answer the original query it only remains to check $\Sigma_0 \models I \wedge \text{TRUE}$, which is true.

6 Conclusion

This paper presents reduction theorems to solve the projection problem for beliefs. The work is carried out within a modal variant of the belief revision framework for the situation calculus by (Shapiro et al. 2011). Firstly, we presented a regression mechanism that uses belief conditionals to relate future beliefs to the initial beliefs. Secondly, we showed how belief modalities can be eliminated from the query as well. That way, the belief projection problem can be reduced to a finite number of non-modal first-order entailments.

The next step is to employ these techniques within a decidable fragment of the situation calculus such as (Lakemeyer and Levesque 2014). In future, we plan to investigate whether or not similar results can be obtained for formalisms that use Spohn-style (Spohn 1988) plausibility rankings such as (Delgrande and Levesque 2012). Another open problem is to address belief projection through progression (Lin and Reiter 1997).

Acknowledgments

References

Alchourron, C. E.; Gärdenfors, P.; and Makinson, D. 1985. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic* 50(2):510–530.

Baumann, R.; Brewka, G.; Strass, H.; Thielscher, M.; and Zaslawski, V. 2010. State defaults and ramifications in the unifying action calculus. In *Proc. KR*.

Delgrande, J. P., and Levesque, H. J. 2012. Belief revision with sensing and fallible actions. In *Proc. KR*.

Demolombe, R., and Pozos Parra, M. d. P. 2006. Belief revision in the situation calculus without plausibility levels. In *Foundations of Intelligent Systems*, LNCS. 504–513.

Fang, L., and Liu, Y. 2013. Multiagent knowledge and belief change in the situation calculus. In *Proc. AAAI*.

Gelfond, M., and Lifschitz, V. 1993. Representing action and change by logic programs. *The Journal of Logic Programming* 17(2):301–321.

Gelfond, M. 2008. Answer sets. *Handbook of knowledge representation* 1:285–316.

Kowalski, R., and Sergot, M. 1989. A logic-based calculus of events. In *Foundations of knowledge base management*. 23–55.

Lakemeyer, G., and Levesque, H. J. 2011. A semantic characterization of a useful fragment of the situation calculus with knowledge. *Artificial Intelligence* 175(1):142–164.

Lakemeyer, G., and Levesque, H. J. 2014. Decidable reasoning in a fragment of the epistemic situation calculus. In *Proc. KR*.

Levesque, H. J., and Lakemeyer, G. 2001. *The Logic of Knowledge Bases*. MIT Press.

Lewis, D. 1973. *Counterfactuals*. John Wiley & Sons.

Lin, F., and Reiter, R. 1997. How to progress a database. *Artificial Intelligence* 92(1):131–167.

McCarthy, J. 1963. Situations, Actions, and Causal Laws. Technical Report AI Memo 2, AI Lab, Stanford University.

Pagnucco, M.; Rajaratnam, D.; Strass, H.; and Thielscher, M. 2013. Implementing belief change in the situation calculus and an application. In *Logic Programming and Nonmonotonic Reasoning*. 439–451.

Pearl, J. 1990. System Z: A natural ordering of defaults with tractable applications to nonmonotonic reasoning. In *Proc. TARK*, 121–135.

Reiter, R. 2001. *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems*. The MIT Press.

Scherl, R., and Levesque, H. J. 2003. Knowledge, action, and the frame problem. *Artificial Intelligence* 144(1–2):1–39.

Schwering, C., and Lakemeyer, G. 2014. A semantic account of iterated belief revision in the situation calculus. In *Proc. ECAI*, 801–806.

Shapiro, S.; Pagnucco, M.; Lespérance, Y.; and Levesque, H. J. 2011. Iterated belief change in the situation calculus. *Artificial Intelligence* 175(1):165–192.

Spohn, W. 1988. Ordinal conditional functions: A dynamic theory of epistemic states. In Harper, W. L., and Skyrms, B., eds., *Causation in Decision, Belief Change, and Statistics*. 105–134.

Thielscher, M. 1999. From situation calculus to fluent calculus: State update axioms as a solution to the inferential frame problem. *Artificial Intelligence* 111(1):277–299.