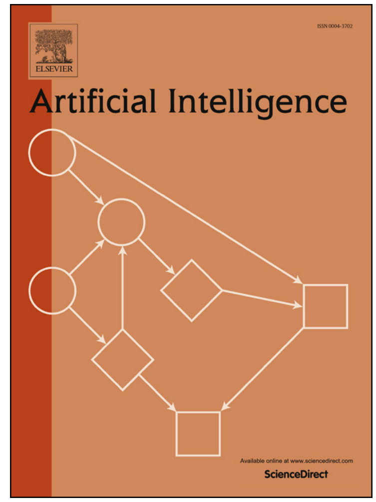
Accepted Manuscript



Belief revision and projection in the epistemic situation calculus

Christoph Schwering, Gerhard Lakemeyer, Maurice Pagnucco

PII: S0004-3702(17)30085-1
DOI: http://dx.doi.org/10.1016/j.artint.2017.07.004
Reference: ARTINT 3024

To appear in: *Artificial Intelligence*

Received date: 7 March 2016
Revised date: 7 July 2017
Accepted date: 19 July 2017

Please cite this article in press as: C. Schwering et al., Belief revision and projection in the epistemic situation calculus, *Artif. Intell.* (2017), http://dx.doi.org/10.1016/j.artint.2017.07.004

Belief revision and projection in the epistemic situation calculus

Christoph Schwering[a,*], Gerhard Lakemeyer[b], Maurice Pagnucco[a]

[a] *School of Computer Science and Engineering, The University of New South Wales, NSW, 2052, Australia*
[b] *Department of Computer Science, RWTH Aachen University, 52056 Aachen, Germany*

Abstract

This article considers defeasible beliefs in dynamic settings. In particular, we examine the belief projection problem: what is believed after performing an action and/or receiving new information? The approach is based on an epistemic variant of Reiter's situation calculus, where actions not only have physical effects but may also provide new information to the agent. The preferential belief structure is initially determined using conditional statements. New information is then incorporated using two popular belief revision schemes, namely natural and lexicographic revision. The projection problem is solved twofold in this formalism: by goal regression and by knowledge base progression.

Keywords: Knowledge representation, Reasoning about actions, Belief revision

1. Introduction

Knowledge and actions have long been identified as two key aspects of an intelligent system: McCarthy's pioneering 1959 paper [1] envisions a computer program that chooses its actions based on knowledge about its current situation. Ensuing from McCarthy's original *situation calculus* [2], numerous logical languages for modeling such dynamic systems have been developed. One of the most successful approaches is Reiter's variant of the situation calculus [3, 4], whose popularity is due to its simple yet powerful solutions for the *frame problem* and the *projection problem*. Projection refers to determining whether a certain formula is true after a sequence of actions; it is the fundamental operation in reasoning about actions and plays an essential role in planning. The versatility of Reiter's framework has been proven by a wide range of extensions that accommodate concepts such as time, concurrency, complex actions, decision theory, and, of particular relevance to this paper, knowledge and sensing. An explicit notion of knowledge allows for modeling both knowledge and lack thereof within the object language. For instance, we could express that a gift box is known to contain an unknown gift.

Unequivocal knowledge however is rare in everyday situations. More often than not, intelligent agents merely have *beliefs* which may or may not hold true in actuality. In fact, agents often consider both cases possible, but regard one of the options to be more plausible than the other. For example, an agent might believe that the gift box presumably is empty, but that if it is not empty, then most likely it contains a gift (whatever that gift may be). The second belief here is called *conditional* because it is constrained by a hypothesis (namely the box not being empty). Such conditionals are an intuitive way of expressing beliefs about different contingencies and, implicitly, their plausibility.

In a dynamic setting, beliefs are subject to *change*. Change comes in two types: *physical* change reflects what actually happens in the environment; *epistemic* change occurs when the agent receives new information about its environment. In our scenario, dropping the box could have the physical effect of breaking the objects inside the box. But it might be only after the following clinking noise that agent realizes the box was not empty and something actually broke inside it. Possible inconsistencies among beliefs and such new information can be resolved by *belief revision*, which aims to give up just enough old beliefs in order to accommodate the new information [5].

*Corresponding author.
Email addresses: c.schwering@unsw.edu.au (Christoph Schwering), gerhard@kbsg.rwth-aachen.de (Gerhard Lakemeyer), m.pagnucco@unsw.edu.au (Maurice Pagnucco)

For an intelligent agent—a robot, for example—it is important to behave reasonably in such dynamic and uncertain conditions. The key problem that arises when reasoning about actions and beliefs is the *belief projection problem*: what does the agent believe after a sequence of actions brings about physical and/or epistemic change? More formally, the question can be cast as one of logical entailment: given a knowledge base comprising initial beliefs and knowledge about actions, a query formula about the agent's beliefs and a sequence of actions, does the knowledge base entail the query after the action sequence is performed? The goal is to reduce this problem to an equivalent static entailment, that is, to another entailment problem where no actions are involved, which can therefore be tackled without having to take the dynamics into consideration.

The projection problem can be approached from two opposite directions:

- *Regression* rewrites the query to roll back the action sequence and checks this new, action-less query against the original knowledge base.
- *Progression*, on the other hand, applies the effects of the actions to the knowledge base in order to roll the knowledge base forward, so that the original query can be tested against that updated knowledge base.

Regression and progression both have their own strengths and weaknesses. While the procedure to regress a query is often simple and elegant, in a long-lived system with a history of thousands or even millions of actions, regression is not feasible as it would need to revert this huge sequence of actions for every single query. By contrast, a knowledge base only needs to be progressed once for every action; however, this may be very expensive computationally and is not even first-order definable in general [6, 7]. Given this duality, a reasoning system might use regression for short-term planning, and periodically progress the knowledge base during its "mental idle time" to keep up with the actual change of its environment [8].

In this paper we study the belief projection problem and present results for both regression and progression.[1] While the question of modeling belief change in the situation calculus has also been addressed by others in various ways [12, 13, 14, 15], these approaches focus more on representational aspects and do not provide the sort of expressivity that we found is necessary for belief projection. For this reason, instead of adopting one of the referenced frameworks, we propose a novel formalism that can be characterized in the following way:

- Beliefs about different contingencies are expressed through if-then statements such as "if the box is not empty, then it presumably contains a gift"; so-called *conditional beliefs*. Furthermore, by way of a concept called *only-believing* it is possible to capture that a conditional knowledge base is *all* the agent believes. Only-believing hence implicitly specifies the agent's non-beliefs, which are essential to infer, for instance, that the box is believed to contain an unknown gift.
- Physical and epistemic change is caused by *actions*. Physical change means that an action may directly affect and change the value of predicates, like dropping a box has the effect of breaking the fragile items inside the box. Epistemic change, on the other hand, means that an action may convey new information that leads to a reassessment of how plausible the agent considers its different beliefs to be, like the clinking noise tells the agent that something seems to be broken in the box; this information is taken into account by *belief revision*.

As we shall see, belief projection in such a framework comprising physical and revision effects fundamentally relies on conditional belief: our regression operator will exploit relationships between the conditional beliefs before and after an action, and the progression results relate the original and the updated conditional knowledge base by means of only-believing. This comes as no surprise considering the relationship between belief revision and conditional belief manifested in Ramsey's test [16, 17], according to which a conditional is accepted iff the consequent is true after revising by its antecedent.

Revision operators can be differently radical in terms of how firmly they hold on to the new information after subsequent revisions. We consider two particular operators in this paper: natural [18, 19] as well as lexicographic revision [20, 21], which we refer to as *weak* and *strong* revision here because lexicographic revision leads to firmer belief in the new information than natural revision. While these are only two among many revision schemes proposed in the literature and especially natural revision has drawn criticism, they are also perhaps the most well-known operators and stand out due to their straightforward definitions.

[1] This paper unifies and extends results by the authors that first appeared in [9, 10, 11].

The language we propose builds on first-order logic, which makes reasoning in our framework undecidable in general. Nevertheless the paper's contribution goes beyond the conceptual level, as fragments of the language are decidable. In particular, the ideas presented here are compatible with the idea of limited belief, where decidability is achieved through semantic restrictions [22, 23].

On the technical side, we abandon Reiter's axiomatic formulation of the situation calculus in favor of Lakemeyer and Levesque's semantical characterization [24], which amalgamates Levesque's logic of only-knowing [25, 26] with actions in the spirit of Reiter's situation calculus. We see three benefits of the semantical over the traditional axiomatic approach. Firstly, proofs in the possible-worlds semantics are often much more straightforward than in the comparable epistemic formalisms that use Tarskian semantics and reify possible worlds. Secondly, the language and the semantics naturally support arbitrary nesting of actions, quantification, and beliefs. Finally, the semantical approach allows for a much clearer definition of only-believing, which in classical first-order logic, if expressible at all, requires a complex meta-logical belief closure operation [4].

The rest of this paper is organized as follows. In the next section we review related work in the fields of reasoning about actions, belief revision, and their intersection. Section 3 introduces our formalism and presents some relevant properties. In the next section we introduce the Reiter-style action theories and the belief projection problem in our variant of the situation calculus. The following two sections are devoted to solving the belief projection problem, first by goal regression and then by theory progression. Finally we compare our work with the standard belief revision frameworks in Section 7. Then we conclude.

2. Related work

We begin our survey of related work with the situation calculus and its epistemic extensions, proceed with particular focus on other variants that accommodate belief change, and finally review literature on classical belief revision.

2.1. Reasoning about actions and knowledge

The situation calculus was originally proposed by McCarthy [2] to allow programs like the fictitious Advice Taker program [1] to reason about which action to take next. The most thoroughly investigated variant of the situation calculus however is due to Reiter [4]; his approach views situations as sequences of actions which represent the "world history." Other significant action formalisms include the fluent calculus [27] and more distant relatives like the event calculus [28] and the family of action languages $\mathcal{A}$ [29].

Reiter's situation calculus is formulated in classical predicate logic with Tarskian semantics. His axiomatic approach requires standardization of models through unique-name axioms and other foundational axioms, including a second-order one, and makes extensive use of reification to represent situations. This can become quite cumbersome in particular when combined with epistemic reasoning as introduced by Scherl and Levesque [30], where not just one but potentially many possible worlds need to be considered.

Lakemeyer and Levesque address this issue in their epistemic situation calculus $\mathcal{ES}$ [24], a first-order modal logic with possible-worlds semantics that amalgamates Levesque's logic of only-knowing [25, 26] with actions in the spirit of Reiter's situation calculus and Scherl and Levesque's epistemic extension [30]. Actions occur in this language as modal operators, whereas possible worlds and situations are purely semantic concepts. While $\mathcal{ES}$ has been shown to faithfully reconstruct Reiter's situation calculus [24], the semantical formulation arguably yields a cleaner language and usually leads to more straightforward proofs.

Besides the classical notion of knowledge, $\mathcal{ES}$ also provides a way of expressing that a knowledge base is *all* the agent knows, implicitly meaning that anything that does not follow from the knowledge base is not known. In Reiter's axiomatization of the situation calculus, such a closed-world assumption on knowledge can be achieved, too, albeit through a meta-logical operation that augments the knowledge base with (infinitely many) sentences about what the agent does not know [4]. In this paper we generalize this notion of only-knowing to handle conditional knowledge bases. As we shall see in Section 3.4, the natural extension of Reiter's meta-logical operation to the case of conditional knowledge bases is not well-behaved.

Reiter proposed his situation calculus along with basic action theories in order to solve the frame problem and the projection problem [3, 4]. His regression mechanism was later generalized for knowledge [30, 24]; in Section 5 we take it another step further to conditional beliefs. Lin and Reiter introduced progression of basic action theories as an alternative solution of the projection problem [6]. Lakemeyer and Levesque extended Lin and Reiter's progression to $\mathcal{ES}$ by using a dynamic variant of only-knowing [31]. Our progression of conditional knowledge bases from Section 6 generalizes their results further to conditional beliefs. While progression has been shown to be first-order definable for certain classes of basic action theories [6, 32], it remains an open question if and how these results generalize to knowledge bases with conditional beliefs.

Scherl and Levesque [30] and most subsequent work on epistemic reasoning in the situation calculus, including $\mathcal{ES}$, only consider unrevisable knowledge. Sensing in these formalisms means that all worlds that disagree with the actual world on the value of a specific formula are deemed impossible. In its binary form, sensing answers yes–no questions such as "is the gift broken?"; these answers are definitive and cannot be revised. A sensing result that contradicts the agent's knowledge therefore irrevocably leads to a state of logical inconsistency, where no meaningful inferences can be made. This sort of sensing corresponds to belief expansion in the belief revision literature.

Our approach avoids this limitation by accommodating new information through a different mechanism where the information carried by an action need not be grounded by the actual world and may hence be incorrect. Such new information is incorporated using classical belief revision schemes, which give up a minimal set of previous beliefs when necessary. These differences notwithstanding, our mechanism of informing can simulate Scherl and Levesque's sensing for the most part as we shall briefly argue in Section 4.1.

2.2. Belief change in the situation calculus

A number of extensions of the situation calculus that address belief change have been proposed [12, 13, 14, 15, 33]. These papers however focus more on representational aspects than on belief projection. The main difference between the present and the referenced papers is consequently our investigation of the belief projection problem. There are also major differences among the formalisms; particularly the prominent role of conditional belief in our approach stands out compared to the others.

The formalism by Shapiro et al. [12] is arguably the closest to ours. Shapiro et al. follow Scherl and Levesque [30] in their use of sensing, which eliminates possible worlds that disagree with sensed information, which they combine with a plausibility ranking of possible worlds. As usual, belief is defined as what is true in the most-plausible worlds. Belief revision hence occurs when a sensing result eliminates all most-plausible worlds, so that other worlds become the most-plausible ones. Unfortunately, this scheme implies that sensed information itself cannot be revised: after sensing information that contradicts an earlier sensing, all possible worlds are eliminated, and the agent is caught in a state of inconsistency. This is due to the use of classical sensing, which irrevocably eliminates worlds, as opposed to updating the plausibility ranking as done by classical revision operators.

Shapiro et al. also briefly discuss a conditional belief operator for the purpose of specifying the initial beliefs. Capturing a knowledge base with that operator is however quite cumbersome as typically also a (potentially huge) number of negative conditionals is required to also capture the intended non-beliefs. As it stands, these conditionals need to be determined by hand; a closure operation analogous to Reiter's knowledge closure [4] seems inappropriate as it may be inconsistent as we shall illustrate in Section 3.4. In our approach the need for such a meta-logical belief closure is obviated by only-believing.

Shapiro et al. opt for classical sensing instead of belief revision because any plausibility updating scheme would conflict with belief introspection in their approach. The anomalies they mention do not occur in our logic because here a world's plausibility is a property of that world alone (and of the epistemic state), independent of the currently considered actual world. Theorem 17 shows that introspection as well as quantifying-in are well-behaved in our framework.

The next-closest relative is by Delgrande and Levesque [14]. In this formalism, actions can inform the agent that some information is true like in our approach; new information is then incorporated by a revision scheme based on Spohn's ranking functions [34]. Among the mentioned approaches, this is the only one that follows a traditional revision scheme like ours. However, it differs in that actions may be fallible, which means that an agent may do a different action than it actually intended to do. This is modeled using a binary fluent relation between actions that may be mistaken for another. Delgrande and Levesque focus on these representational aspects and do not consider projection. To what extent the techniques developed in this paper could be transferred to their formalism is an open question.

Fang and Liu [15] use plausibility rankings on worlds and actions; the plausibility ranking then changes according to the executed actions. Both rankings are explicitly specified in the beginning. Fang et al. also consider progression [33], however only in the propositional case.

Demolombe and Pozos-Parra's proposal [13] avoids any plausibility ranking by compiling physical and epistemic effects on predetermined beliefs of interest to special successor-state axioms. Our approach however shows that ranking the worlds by plausibility can be considered a purely semantic concept; plausibilities are not needed in the object language.

Another framework to deal with faulty sensors in the situation calculus is the Bayesian approach by Bacchus et al. [35]. They also use extra action parameters to indicate the real-world outcome, similar to our mimicking of classical sensing described in Section 4.1.

Belief revision has also been addressed in dynamic epistemic logic [36] by several authors [37, 38, 39, 40]. Closest to our work is van Benthem's [40], which gives semantics to conditional belief through a plausibility ranking of possible worlds and considers natural and lexicographic revision, like we do. Most importantly, van Benthem reduces beliefs after revision to beliefs before that revision in a regression-like fashion; these reduction theorems resemble our Theorems 27 and 28.

2.3. Belief revision and conditional belief

The seminal work on belief revision is the system of eight postulates proposed by Alchourrón, Gärdenfors, and Makinson [41, 42] and named AGM after their initials. These postulates aim to govern how a belief revision operator should incorporate new information into an existing set of beliefs in a reasonable way.

Building on AGM's work on one-time revision, Darwiche and Pearl later proposed postulates to address iterated revision as well [43]. What is being revised in their framework is no longer merely a set of beliefs, but an epistemic state where some beliefs are more plausible than others; revision updates the associated plausibility ranking and produces a new epistemic state. The question of iterated revision however remains disputed and has produced numerous alternative postulate systems, including [21, 44, 45, 46].

It is no surprise given this controversy about the postulates that there is no commonly agreed-on revision operator. Among the proposed revision operators are natural revision [18, 19], lexicographic revision [20, 21], restrained revision [45], irrevocable revision [47], and Darwiche and Pearl's operator [43]. While these operators all give highest priority to the new information, they vary in how firmly they hold on to these beliefs after new information comes in. For instance, natural revision very readily gives up information from an earlier revision after a subsequent revision, whereas lexicographic revision is much more reluctant to do so. For that reason, we refer to natural as weak revision and to lexicographic as strong revision. Similarly, Rott refers to them as conservative and moderate revision, respectively, as natural revision minimally modifies the plausibility ranking.

Natural revision in particular has drawn criticism for giving up information from previous revisions too readily [43, 45, 48, 49]. Booth and Meyer's restrained revision attempts to behave more intuitively here. They propose a more restricted class of iterated revision operators called admissible operators. In this class, restrained revision replaces natural revision as the most conservative operator.

Despite these valid criticisms we opt for natural and lexicographic revision in this paper mainly for pragmatic reasons. On the one hand, they are perhaps the best-known revision operators and, on the other, they feature very simple definitions which make them quite easy to work with.

Which of the many known revision operators is most suitable is a hard question; it appears to depend at least on the context and the sort of information that was obtained [45]. A meta-theory about when to choose which revision operator as called for by Rott [50] is however beyond the scope of this paper. Instead, it is up to the modeler in our framework, which revision operator (among natural and lexicographic) to choose for which kind of information.

Another popular framework for iterated revision is due to Spohn [34]. His ordinal conditional functions map worlds to plausibilities, which are not only taken to order the worlds but also interpreted quantitatively. Revision then not only takes a new formula, but also a parameter that indicates how firmly this new information is believed.

AGM-style belief revision is not intended and in fact inappropriate for dealing with physical change. While we follow Reiter's situation calculus here to account for physical change, Katsuno and Mendelzon have proposed a postulate system that addresses physical change in the spirit of AGM's framework [51]. The belief update approach differs fundamentally from the situation calculus. For instance, it is possible to update beliefs by any disjunctive formula in Katsuno and Mendelzon's framework, whereas Reiter's basic action theories only allow for deterministic change. On the other hand, belief update is a meta-logical operation and as such has no concept of actions at the object-language level. Boutilier has folded belief revision and update into a single framework [52]: when an event happens, the beliefs are revised by the event's precondition and then updated by its effect. This view is also applicable to actions in our framework.

Conditional belief and belief revision are closely related through what is known as the Ramsey test [17]: the conditional "if α, then plausibly β" shall be believed iff β is believed after revision by α. This relationship also underlies our Theorems 27 and 28 that relate beliefs before and after an action. The most influential work on conditional logic is Lewis' book on counterfactuals [53]. The system-of-spheres model we use for epistemic states to capture conditional belief is chiefly due to him and to Grove [54]. While Lewis argues against a finiteness assumption for the number of spheres on philosophical grounds, we make this assumption here as it simplifies the technical treatment and since finitely many spheres suffice for our purposes.

Our semantics of (nested) conditionals is somewhat unconventional; an antecedent has no repercussions on any nested beliefs in the consequent (Property 15 in Theorem 17). Semantically, the reason is that the system of spheres is independent of the actual world, a condition Lewis calls uniformity [53]. This is different to both Boutilier's original account of natural revision [18] as well as Levi's philosophical considerations on nested beliefs [55]. Our semantics has some advantages. Firstly, making nested beliefs in the consequent independent of the outer antecedent greatly simplifies Theorems 27 and 28, which are fundamental for belief regression. Without that independence, action operators would have to be pushed inside of all possibly nested beliefs in order to avoid redundant revisions. Although possible, this is quite cumbersome. Secondly, the simple semantics makes it possible to reduce conditional beliefs to a series of non-modal reasoning tasks as shown in [10, 56], which we do not discuss in this paper. Considering these issues and that there is no commonly-agreed-on semantics for nested beliefs [53, 55], we believe our simpler (as it does not modify the epistemic state) semantics of conditional beliefs is preferable for the purposes of this paper.

Only-believing has relations to both Levesque's only-knowing [26] and Pearl's meta-logical System Z [57], as we shall discuss in Section 3.4. Like only-knowing, only-believing works by including as many worlds as possible into the epistemic state. Additionally, it imposes a plausibility ranking on these worlds, which can be shown to be equivalent to Pearl's Z-ranking [56, 57]. A concept similar to only-believing was used by Boutilier [58] to semantically capture System Z, too. Only-knowing can be considered a combination of (ordinary) knowing *at least* and knowing *at most* [26]. Boutilier introduces a conditional belief variant of the latter, by which he forces worlds to be as normal as possible and ultimately replicates System Z.

3. The logic $\mathcal{ESB}$

In this section we introduce a first-order[2] logic with modal operators for both actions and beliefs. The language features predicates, equality, functions, and countably many standard names, which serve as rigid designators of all elements in the domain. Predicates are either *fluent*, meaning their value varies as the result of actions, or *rigid*. Functions are assumed to be rigid for simplicity. The terms of the language come in two sorts, *action* and *object*. Actions are further divided into the subsorts *weak-revision* and *strong-revision actions*: an action may, besides its (physical) effect on fluents, also carry new information, and the action's subsort determines *how* this new information is incorporated into the current beliefs. The language supports conditional belief, including a variant called *only*-belief which implicitly also specifies what is *not* believed.

We define the language in Section 3.1 and give the semantics in Section 3.2. In Sections 3.3 and 3.4 we further investigate conditional beliefs and only-believing, respectively. In particular, we prove a unique-model property, which is foundational for many subsequent results of this paper.

3.1. The syntax

Definition 1. The symbols of our language are taken from the following vocabulary:

- infinitely many object standard names $^{\#}1, {}^{\#}2, \ldots$;

[2] We will add a second-order extension in Section 6.1.

- infinitely many first-order variables, written schematically as x:
  - of sort object, written schematically as y;
  - of sort action, written schematically as a;
- infinitely many function symbols:
  - of sort object, written schematically as g;
  - of sort action, written schematically as A and divided in weak- and strong-revision actions;
- infinitely many predicate symbols:
  - of type rigid, written schematically as R;
  - of type fluent, written schematically as F;
- connectives and other symbols: $=$, $\vee$, $\neg$, $\exists$, $\square$, **B**, **O**, $\Rightarrow$, round/square/curly brackets, comma.

Each function or predicate symbol has an *arity*, which indicates how many arguments it takes. There shall be a special fluent *IF* which takes a single action as an argument. Variables as well as function and predicate symbols may be decorated with subscripts or superscripts.

Definition 2. The terms of the language are of sorts either *object* or *action*, where actions are divided in the subsorts *weak*- and *strong*-revision actions. The terms are formed from first-order variables, standard names, and function symbols. *Standard names* serve as unique identifiers of all individuals; they are defined as follows:

- the object standard names are $^{\#}1, {}^{\#}2, {}^{\#}3, \ldots$;
- the action standard names are of the form $A(n_1, \ldots, n_k)$ where A is an action function symbol and the n_i are standard names.

The set of *terms* of sort object or action is the least set such that:

- every variable and standard name is a term of the corresponding sort;
- $g(t_1, \ldots, t_k)$ is an object term if g is an object function symbol and the t_i are terms;
- $A(t_1, \ldots, t_k)$ is an action term if A is an action function symbol and the t_i are terms.

An object term $g(n_1, \ldots, n_k)$ is *primitive* when all of the n_i are standard names.

For example, if *unbox* is a unary action symbol, *unbox*($^{\#}5$) is an action standard name. If *gift* is an object constant, *unbox*(*gift*) is an action term, but not a standard name. Throughout the paper, we use n for standard names and t for terms.

Definition 3. The set of well-formed *formulas* of the language is the least set such that:

- $R(t_1, \ldots, t_k)$ and $F(t_1, \ldots, t_k)$ is a formula where R and F are fluent and rigid predicate symbols and t_i are terms;
- $(t_1 = t_2)$ is a formula where t_1 and t_2 are terms;
- $\neg\alpha$, $(\alpha \vee \beta)$, and $\exists x\,\alpha$ are formulas where α and β are formulas and x is a variable;
- $[t]\alpha$ and $\square\alpha$ are formulas where α is a formula and t is an action term;
- $\mathbf{B}(\alpha \Rightarrow \beta)$ is a formula where α and β are formulas;
- $\mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ is a formula where ϕ_i and ψ_i are formulas without **B** or **O**.

We read $[t]\alpha$ as "α holds after action t" and $\square\alpha$ as "α holds after any sequence of actions." Formulas without $[t]$ or $\square$ operators are called *static*. A conditional belief $\mathbf{B}(\alpha \Rightarrow \beta)$ can be read as "it is believed that if α is true, then presumably β is also true." The only-believing operator $\mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ means "the conditionals $\phi_i \Rightarrow \psi_i$ are *all* that is believed." Formulas without **B** or **O** operators are called *objective*; as a notational convention we will use $\phi, \psi, \upsilon, \tau$ to denote objective formulas. For the purposes of this paper, the conditionals within **O** are restricted to objective formulas. A set $\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ is called objective when all ϕ_i and ψ_i are objective. Finally, $R(t_1, \ldots, t_k)$ and $F(t_1, \ldots, t_k)$ are called rigid and fluent *atoms*; an atom is *primitive* when all arguments t_i are standard names.

By α^x_t we denote the formula α with t substituted for all free occurrences of the variable x. A *sentence* is a formula with no free variable. We use the usual abbreviations $(\alpha \wedge \beta)$, $(\alpha \supset \beta)$, $(\alpha \equiv \beta)$, $\forall x\alpha$, TRUE, FALSE, and $(t_1 \neq t_2)$. Additionally,

- $\mathbf{B}\alpha$ stands for $\mathbf{B}(\text{TRUE} \Rightarrow \alpha)$;
- $\mathbf{K}\alpha$ stands for $\mathbf{B}(\neg\alpha \Rightarrow \text{FALSE})$.

$\mathbf{B}\alpha$ and $\mathbf{K}\alpha$ are read as "α is believed" and "α is known," respectively.

The language features one special fluent predicate, $IF(t)$. The general idea is that when action t is executed, this *informs* the agent that most plausibly $IF(t)$ holds. In a sense, $IF(t)$ represents the most-plausible precondition of t, so when t is executed, this tells the agent that it is plausible that $IF(t)$ was true. For example, a knowledge base might stipulate that $IF(clink)$ is true when something breaks to pieces, and that $IF(unbox(\#5))$ is true only if #5 actually is in the box (we formalize this example in Section 4). It is noteworthy that this information does not necessarily depend on what is true in the real world; in particular, the information may actually be mistaken. Unlike the classical approach to sensing [30, 24], we can therefore model that incoming information contradict each other. Such inconsistencies are then resolved by means of belief revision. We elaborate on the intuition behind this model in Section 4.1.

To ease the readability of formulas we often omit brackets. Then the convention is that $\square$ shall be the weakest binding operator of all; the other unary operators bind stronger than the logical connectives; and the connective $\wedge$ binds stronger than $\vee$. Free variables are implicitly universally quantified with maximal scope unless noted otherwise. For example, $\square[a]Broken(y) \equiv Broken(y) \vee InBox(y) \wedge Fragile(y) \wedge a = dropbox$ abbreviates the formula $\forall a \forall y \square(([a]Broken(y)) \equiv (Broken(y) \vee ((InBox(y) \wedge Fragile(y)) \wedge (a = dropbox))))$. Intuitively this formula means that after any sequence of actions ($\square$), an object y is broken after another action a iff it either was broken already or it is fragile and in the box which is dropped.

3.2. The semantics

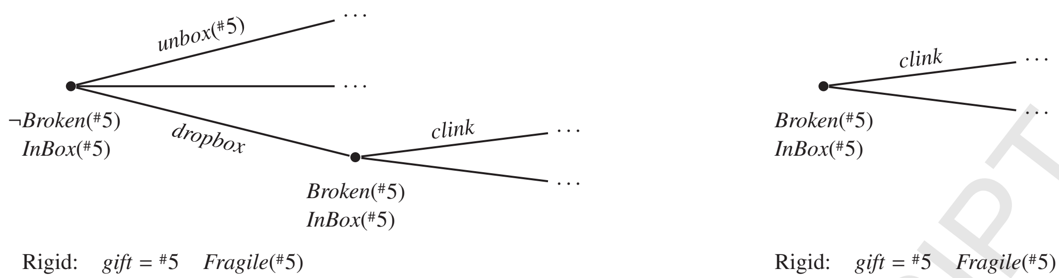
We now give a semantics for this language using possible worlds. We will begin with the objective fragment of the logic, that is, the part of the language that is not concerned with beliefs. The key ingredient of our semantics is worlds. A world assigns a standard name to each primitive term and a truth value to each primitive atom, which in case of fluent atoms also depends on the sequence of actions executed thus far. A world is therefore not just a snapshot but also determines the future. In the following we formalize these notions.

Definition 4. An *action sequence* is the empty sequence $\langle\rangle$ or the concatenation $z \cdot n$ of an action sequence z and an action standard name n. A *world* w is a function:

- from the primitive object terms $g(n_1, \ldots, n_k)$ to object standard names;
- from the primitive rigid atoms $R(n_1, \ldots, n_k)$ to truth values $\{0, 1\}$;
- from the primitive fluent atoms $F(n_1, \ldots, n_k)$ and action sequences to truth values $\{0, 1\}$.

The *progression* of a world w by an action standard name n is a world $w \gg n$ such that:

- $(w \gg n)[g(n_1, \ldots, n_k)] = w[g(n_1, \ldots, n_k)]$ for all object function symbols g;
- $(w \gg n)[R(n_1, \ldots, n_k)] = w[R(n_1, \ldots, n_k)]$ for all rigid predicate symbols R;
- $(w \gg n)[F(n_1, \ldots, n_k), z] = w[F(n_1, \ldots, n_k), n \cdot z]$ for all fluent predicate symbols F and action sequences z.



(a) A world w where $^{\#}5$ is the gift, fragile, initially in the box and not broken, but is broken by dropping the box.

(b) The progression $w \gg \mathit{dropbox}$ of the world w from Figure 1a by the action *dropbox*.

Figure 1: A world can be thought of as a tree: the root node is the initial valuation of fluent predicates; edges represent actions that lead to a new valuation. Additionally a world is a valuation of rigid predicates and functions, which are not subject to change. The progression of a world by an action is the subtree rooted in the child node reached by the corresponding edge.

We abbreviate $w \gg n_1 \gg \ldots \gg n_k$ by $w \gg \langle n_1, \ldots, n_k \rangle$.

For example, to say that $^{\#}5$ is the gift, is fragile, and breaks after *dropbox*, a world w might stipulate $w[\mathit{gift}] = {}^{\#}5$, $w[\mathit{Fragile}({}^{\#}5)] = 1$, $w[\mathit{Broken}({}^{\#}5), \langle\rangle] = 0$, and $w[\mathit{Broken}({}^{\#}5), \langle \mathit{dropbox} \rangle] = 1$; such a world is depicted in Figure 1a. In that world's progression by *dropbox*, the object is broken: $(w \gg \mathit{dropbox})[\mathit{Broken}({}^{\#}5), \langle\rangle] = 1$, as depicted in Figure 1b. Note that *dropbox* is not interpreted by the world, as it is an action standard name.

Definition 5. The *denotation* $w(t)$ of a term t in a world w is defined as follows:

- if n is a standard name, $w(n) = n$;
- if g is an object function symbol, $w(g(t_1, \ldots, t_k)) = w[g(n_1, \ldots, n_k)]$ where $n_i = w(t_i)$;
- if A is an action function symbol, $w(A(t_1, \ldots, t_k)) = A(n_1, \ldots, n_k)$ where $n_i = w(t_i)$.

For example, if $w[\mathit{gift}] = {}^{\#}5$, then the denotation of $\mathit{unbox}(\mathit{gift})$ is $w(\mathit{unbox}(\mathit{gift})) = \mathit{unbox}({}^{\#}5)$.

Definition 6. Truth of an objective sentence is defined with respect to a world w:

S'1. $w \models R(t_1, \ldots, t_k)$ iff $w[R(n_1, \ldots, n_k)] = 1$ where $n_i = w(t_i)$ for every rigid predicate symbol R;

S'2. $w \models F(t_1, \ldots, t_k)$ iff $w[F(n_1, \ldots, n_k), \langle\rangle] = 1$ where $n_i = w(t_i)$ for every fluent predicate symbol F;

S'3. $w \models (t_1 = t_2)$ iff n_1 and n_2 are identical standard names where $n_i = w(t_i)$;

S'4. $w \models \neg\phi$ iff $w \not\models \phi$;

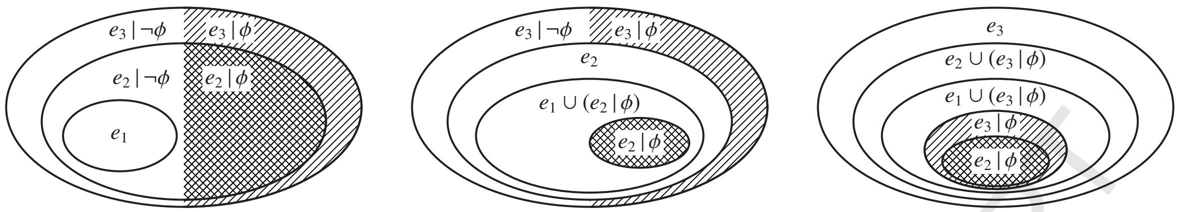
S'5. $w \models (\phi \vee \psi)$ iff $w \models \phi$ or $w \models \psi$;

S'6. $w \models \exists x \phi$ iff $w \models \phi^x_n$ for some standard name n of the same sort as x;

S'7. $w \models [t]\phi$ iff $w \gg n \models \phi$ where $n = w(t)$;

S'8. $w \models \Box\phi$ iff $w \gg z \models \phi$ for every action sequence z.

The above rules capture the objective fragment of our logic, that is, the sentences without **B** or **O**. It is largely identical to the objective semantics of $\mathcal{ES}$ [24], yet it departs from its ancestor in two respects to simplify the formalism: firstly, action functions are not interpreted but form the action standard names instead; secondly, the actions executed thus far are accounted for by progressing the world instead of making the history of actions an explicit part of the model. Unlike classical first-order logic, Rule S'6 handles quantification substitutionally, made possible by the use of standard names as both terms of the language and the universe of discourse. While there are philosophical arguments against substitutional quantification [59], it greatly simplifies the formal machinery. The computational aspects remain however unaltered by substitutional quantification: a static and objective sentence without standard names is satisfiable in $\mathcal{ESB}$ iff it is satisfiable in classical first-order logic [26].



(a) An epistemic state e. (b) Weak revision $e *_{\mathrm{w}} \phi$. (c) Strong revision $e *_{\mathrm{s}} \phi$.

Figure 2: The original epistemic state e (left) has three different spheres, denoted by e_1, e_2, e_3. Hatched area indicates ϕ-worlds; the most-plausible ones occur in e_2, and e_3 contains additional ones. Weak (natural) revision by ϕ (center) promotes these most-plausible ϕ-worlds, namely $e_2 \mid \phi$, to the first level, but leaves the ranking otherwise unchanged. Strong (lexicographic) revision (right) promotes all ϕ-worlds, namely $e_2 \mid \phi$ and $e_3 \mid \phi$, over all $\neg\phi$-worlds, but preserves the relative ordering of the ϕ-worlds and $\neg\phi$-worlds, respectively.

Before we proceed with the semantics of beliefs, let us briefly familiarize ourselves with the objective semantics by showing that

$$\forall a \forall x \forall y (\Box[a]Broken(y) \equiv Broken(y) \lor InBox(y) \land Fragile(y) \land a = dropbox)$$

logically entails

$$\forall y' (InBox(y') \land Fragile(y') \supset [dropbox]Broken(y')).$$

The entailment is proved by showing that every world that satisfies the first sentence also satisfies the second. Suppose w satisfies the former sentence. Recall that $\forall$, $\land$, $\supset$, and $\equiv$ are the usual abbreviations for expressions involving $\exists$, $\lor$, and $\neg$. Following the rules of the semantics for the quantifiers and for $\Box$, we obtain that $w \gg z \models [dropbox]Broken(n) \equiv Broken(n) \lor InBox(n) \land Fragile(n) \land dropbox = dropbox$ for every object standard name n and every action sequence z. In particular, since $w = w \gg z$ for $z = \langle\rangle$, we obtain $w \models [dropbox]Broken(n) \equiv Broken(n) \lor InBox(n) \land Fragile(n)$ for every object standard name n. It is easy to see that this implies $w \models InBox(n) \land Fragile(n) \supset [dropbox]Broken(n)$ for every object standard name n. Introducing a universal quantifier for n yields $w \models \forall y' (InBox(y') \land Fragile(y') \supset [dropbox]Broken(y'))$, which is what we wanted to prove.

We now extend Definition 6 to also give a semantics to beliefs. Following the *system of concentric spheres* model popular in belief revision [53, 54], the possible worlds are stratified into spheres. The first, or innermost, sphere represents the actual beliefs; the outer spheres add less-plausible beliefs. Semantically, each sphere is modeled as a set of possible worlds, and a system of spheres is hence represented as a sequence of sets of possible worlds, also referred to as *epistemic state*. An example is depicted in Figure 2a. To keep matters simple, we assume only finitely many different spheres. The spheres are numbered consecutively with natural numbers, also called *plausibilities*. A smaller value indicates higher plausibility. The concentricity means that every sphere, or *(plausibility) level*, subsumes the less-plausible levels. The plausibility of a world is the plausibility of the first level that contains that world. The plausibility of a sentence corresponds to the most-plausible world that satisfies that sentence; we denote the plausibility of α in an epistemic state e by $\lfloor e \mid \alpha \rfloor$. In case there is no such world in e, $\lfloor e \mid \alpha \rfloor$ cannot be a natural number. For that purpose, we use $\infty \notin \mathbb{N}$ to denote an "undefined" plausibility, with the understanding that $p < \infty$ and $\infty + p = \infty$ for all $p \in \mathbb{N}$ and $\infty + \infty = \infty$. Thus, $\lfloor e \mid \alpha \rfloor = \infty$ indicates that all worlds in e satisfy $\neg\alpha$. To avoid confusion, we always make explicit when an expression may take the value ∞.

When new information comes in, a system of spheres needs to be revised to take into account that information. In that process, less-plausible worlds may be promoted to become the new most-plausible worlds, and the beliefs may thus change. There are several ways to revise a plausibility ranking [50]. The ones we consider here are *natural revision* [18, 19] and *lexicographic revision* [20, 21]. After a single revision, both natural and lexicographic revision lead to the same beliefs, as they both promote to the innermost sphere the most-plausible worlds that satisfy the new information. However, the outer spheres are determined differently, and therefore natural and lexicographic revision differ in how strongly the agent holds on to the new information. Natural revision by ϕ only makes the most-plausible ϕ-worlds the new most-plausible worlds (Figure 2b). Lexicographic revision, on the other hand, promotes *all* ϕ-worlds over all $\neg\phi$-worlds (Figure 2c). Consequently, after lexicographic revision logically stronger new evidence is needed to give up belief in ϕ compared to natural revision. In other words, belief in ϕ after natural revision is weaker than after lexicographic revision. In the following we therefore refer to natural revision as *weak revision* and to lexicographic revision as *strong revision*. The following definition makes these intuitions precise.

Definition 7. An *epistemic state* e is an infinite sequence of sets of worlds e_p, $p \in \mathbb{N} = \{1, 2, \ldots\}$, that

- is concentric, that is, $e_p \subseteq e_{p+1}$ for all $p \in \mathbb{N}$;
- converges, that is, $e_q = e_p$ for some $q \in \mathbb{N}$ and all $p \geq q$.

We use $\langle e_1, \ldots, e_q \rangle$ as a short notation for e when it converges at level q or earlier. We furthermore define

- $\lfloor e \rfloor = \min\{p \mid p = \infty \text{ or } e_p \neq \{\}\}$;
- $\lceil e \rceil = \max\{p \mid p = 1 \text{ or } e_{p-1} \neq e_p\}$;
- $e \,|\, \phi = \langle e_1 \cap W, \ldots, e_{\lceil e \rceil} \cap W \rangle$ where $W = \{w \mid w \models \phi\}$ for objective ϕ.

Intuitively, $\lfloor e \rfloor \in \mathbb{N} \cup \{\infty\}$ denotes the first non-empty level of e, and $\lceil e \rceil \in \mathbb{N}$ refers to the last distinct level. Note that $\lfloor e \rfloor \leq \lceil e \rceil$ holds when there is a non-empty level; otherwise, when all levels are empty, we have $\lfloor e \rfloor = \infty$ and $\lceil e \rceil = 1$. Together, $\lfloor e \rfloor$ and $e \,|\, \phi$ express the plausibility $\lfloor e \,|\, \phi \rfloor$ of ϕ in e. Epistemic states are defined to be infinite sequences of sets of worlds in order to simplify the technical treatment (it saves us from normalizing the length of these sequences). The convergence condition however guarantees that every epistemic state is determined by a finite prefix. For convenience, we shall use the notation $\langle e_1, \ldots, e_q \rangle$ for some $q \in \mathbb{N}$ as a shorthand for the infinite sequence of sets of worlds e with levels $e_1, \ldots, e_q$ and $e_p = e_q$ for all $p \geq q$. Note that q need not be minimal: for example, $\langle e_1, e_2 \rangle = \langle e_1, e_2, e_2, e_2 \rangle$; and generally, $e = \langle e_1, \ldots, e_{\lceil e \rceil} \rangle$.

With these definitions, we can formalize revision of e.

Definition 8. The *weak revision* $e *_{\mathrm{w}} \phi$ of an epistemic state e by an objective sentence ϕ is defined as follows:

- if $\lfloor e \,|\, \phi \rfloor = \infty$: $\quad e *_{\mathrm{w}} \phi = \langle \{\} \rangle$;
- if $\lfloor e \,|\, \phi \rfloor \neq \infty$: $\quad e *_{\mathrm{w}} \phi = \langle W, e_1 \cup W, \ldots, e_{\lceil e \rceil} \cup W \rangle$ where $W = (e \,|\, \phi)_{\lfloor e \,|\, \phi \rfloor}$.

The *strong revision* $e *_{\mathrm{s}} \phi$ of an epistemic state e by an objective sentence ϕ is defined as follows:

- if $\lfloor e \,|\, \phi \rfloor = \infty$: $\quad e *_{\mathrm{s}} \phi = \langle \{\} \rangle$;
- if $\lfloor e \,|\, \phi \rfloor \neq \infty$: $\quad e *_{\mathrm{s}} \phi = \langle (e \,|\, \phi)_{\lfloor e \,|\, \phi \rfloor}, \ldots, (e \,|\, \phi)_{\lceil e \rceil}, (e \,|\, \neg\phi)_{\lfloor e \,|\, \neg\phi \rfloor} \cup W, \ldots, (e \,|\, \neg\phi)_{\lceil e \rceil} \cup W \rangle$ where $W = (e \,|\, \phi)_{\lceil e \rceil}$.

When the revision mechanism is clear from context or irrelevant, we just write $e * \phi$. In particular, we mean by $e * IF(n)$ the revision according to the sort of n, that is, $e *_{\mathrm{w}} IF(n)$ if n is a weak-revision action and $e *_{\mathrm{s}} IF(n)$ if n is a strong-revision action.

An epistemic state is progressed by an action n by first revising it by $IF(n)$ and then progressing the individual worlds by n.

Definition 9. The *progression* of a set of worlds W and of an epistemic state e are defined as follows:

- $W \gg n = \{w \gg n \mid w \in W\}$;
- $e \gg n = \langle e'_1 \gg n, \ldots, e'_q \gg n \rangle$ where $\langle e'_1, \ldots, e'_q \rangle = e * IF(n)$.

We abbreviate $e \gg n_1 \gg \ldots \gg n_k$ by $e \gg \langle n_1, \ldots, n_k \rangle$.

These definitions introduce the toolbox we need for the semantics of beliefs: epistemic states, their revision, and their progression. The following lemma says that the revision and the progression of an epistemic state are well-behaved.

Lemma 10. *$e *_{\mathrm{w}} \phi$, $e *_{\mathrm{s}} \phi$, and $e \gg n$ are epistemic states.*

Proof. Let $e = \langle e_1, \ldots, e_q \rangle$ be an epistemic state. Clearly, $e *_\mathrm{w} \phi$ is a finite sequence and $(e *_\mathrm{w} \phi)_p \subseteq (e *_\mathrm{w} \phi)_{p+1}$ for all $p \in \mathbb{N}$, so it is an epistemic state. Likewise, $e *_\mathrm{s} \phi$ is a finite sequence and $(e *_\mathrm{s} \phi)_p \subseteq (e *_\mathrm{s} \phi)_{p+1}$ for all $p \in \mathbb{N}$, so it is an epistemic state, too.

Now consider $e \gg n$, which simply progresses the individual worlds in $e * IF(n)$. It is immediate from Definition 4 that the progression $w \gg n$ of a world w again is a world. Thus $W \gg n$ is a set of worlds if W is one, and if $W \subseteq W'$, then $W \gg n \subseteq W' \gg n$. Thus and since $e * IF(n)$ is an epistemic state, $e \gg n$ is one, too. □

As mentioned above, weak and strong revision of the same epistemic state lead to the same most-plausible beliefs (but typically differ in the less-plausible beliefs). The next lemma records this.

Lemma 11. $(e *_\mathrm{w} \phi)_1 = (e *_\mathrm{s} \phi)_1$ *and either* $\lfloor e * \phi \,|\, \phi \rfloor = 1$ *or* $(e *_\mathrm{w} \phi) = (e *_\mathrm{s} \phi)$.

Proof. First suppose $\lfloor e \,|\, \phi \rfloor = \infty$. Then $e *_\mathrm{w} \phi = \langle\{\}\rangle = e *_\mathrm{s} \phi$, so the lemma follows. Now suppose $\lfloor e \,|\, \phi \rfloor = \infty$. Then $(e \,|\, \phi)_{\lfloor e \,|\, \phi \rfloor} \neq \{\}$. Thus $(e *_\mathrm{w} \phi)_1 = (e *_\mathrm{s} \phi)_1$ and $\lfloor e * \phi \,|\, \phi \rfloor = 1$. □

We can now give a semantics to the full language. Rules S2–S8 below are simply the ones from Definition 6 retrofitted with an additional epistemic state e, which in case of $[t]$ and $\Box$ needs to be progressed on the right-hand side. The relevant addition are the ones for beliefs, Rules S9 and S10.

Definition 12. Truth of a sentence is defined with respect to an epistemic state e and a world w:

S1. $e, w \models R(t_1, \ldots, t_k)$ iff $w[R(n_1, \ldots, n_k)] = 1$ where $n_i = w(t_i)$ for every rigid predicate symbol R;

S2. $e, w \models F(t_1, \ldots, t_k)$ iff $w[F(n_1, \ldots, n_k), \langle\rangle] = 1$ where $n_i = w(t_i)$ for every fluent predicate symbol F;

S3. $e, w \models (t_1 = t_2)$ iff n_1 and n_2 are identical standard names where $n_i = w(t_i)$;

S4. $e, w \models \neg\alpha$ iff $e, w \not\models \alpha$;

S5. $e, w \models (\alpha \vee \beta)$ iff $e, w \models \alpha$ or $e, w \models \beta$;

S6. $e, w \models \exists x \alpha$ iff $e, w \models \alpha^x_n$ for some standard name n of the same sort as x;

S7. $e, w \models [t]\alpha$ iff $e \gg n, w \gg n \models \alpha$ where $n = w(t)$;

S8. $e, w \models \Box\alpha$ iff $e \gg z, w \gg z \models \alpha$ for every action sequence z;

S9. $e, w \models \mathbf{B}(\alpha \Rightarrow \beta)$ iff for all $p \in \mathbb{N}$ and every world w', if $p \leq \lfloor e \,|\, \alpha \rfloor$ and $w' \in e_p$, then $e, w' \models (\alpha \supset \beta)$;

S10. $e, w \models \mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ iff for all $p \in \mathbb{N}$ and every world w', $w' \in e_p$ iff $e, w' \models \bigwedge_{i:\lfloor e \,|\, \phi_i \rfloor \geq p} (\phi_i \supset \psi_i)$;

where $e \,|\, \alpha = \langle e_1 \cap W, \ldots, e_{\lceil e \rceil} \cap W \rangle$ for $W = \{w \mid e, w \models \alpha\}$ generalizes $e \,|\, \phi$ to arbitrary formulas.

It is immediate that Definitions 6 and 12 agree on the truth of objective sentences, because the only difference between Rules S'1–S'8 and Rules S1–S8 is the extra parameter for the epistemic state carried around.

By Rule S9 the conditional belief $\mathbf{B}(\alpha \Rightarrow \beta)$ holds when the most-plausible α-worlds satisfy β. It thus captures our reading "if α is true, then presumably β is also true." Notice that the belief vacuously holds if there is no α-world. We elaborate on conditional beliefs in Section 3.3.

Rule S10 defines the semantics of only-believing $\mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$. This means that, for one thing, $\mathbf{B}(\phi_i \Rightarrow \psi_i)$ shall be believed, that is, $(\phi_i \supset \psi_i)$ must be satisfied by all worlds up to the first sphere that contains a ϕ_i-world and, for another, the spheres shall be maximal with that property. This maximization intuitively means that any world shall be considered as plausible as possible provided that it is compatible with the conditional beliefs; it thus captures the "only" in only-believing. Only-believing generalizes Levesque's only-knowing [26] to conditional belief, and we will also see a close relation to Pearl's System Z [57] in Section 3.4.

In the rest of the paper, we sometimes omit e or w in $e, w \models \alpha$ when it is irrelevant to the truth of α. For example, when α is objective, we may leave out e. A set of sentences Σ *entails* a sentence α, written $\Sigma \models \alpha$, iff for all e and w, if $e, w \models \beta$ for all $\beta \in \Sigma$, then $e, w \models \alpha$. We write $\models \alpha$ when $\Sigma = \{\}$.

3.3. Properties of conditional beliefs

The conditional belief operator $\mathbf{B}(\alpha \Rightarrow \beta)$ is used to form queries about the agent's beliefs. It expresses the agent's belief that if α was true, then β would be true as well. Or in terms of possible worlds, the most-plausible α-worlds must satisfy β as well. This operator turns out to be a quite versatile tool. In particular, it can capture the following intuitions:

- $\mathbf{B}(\textsc{true} \Rightarrow \alpha)$ represents actual belief in α: it holds when all most-plausible worlds satisfy α. It is therefore abbreviated by $\mathbf{B}\alpha$.
- $\mathbf{B}(\neg\alpha \Rightarrow \textsc{false})$ captures the usual semantics of indefeasible knowledge of α: all worlds at all plausibility levels satisfy α. It is therefore abbreviated by $\mathbf{K}\alpha$.
- $\mathbf{B}(\alpha \vee \beta \Rightarrow \neg\beta)$ asserts that α is strictly more plausible than β: the first $(\alpha \vee \beta)$-worlds must be $\neg\beta$-worlds.
- $\neg\mathbf{B}(\alpha \Rightarrow \neg\beta)$ says that β would be considered possible if α were true: among the most-plausible α-worlds at least one is a β-world. In particular, $\neg\mathbf{B}(\alpha \Rightarrow \textsc{false})$ and $\neg\mathbf{K}\neg\alpha$ express that there is at least one α-world.

In this section we examine a few properties of conditional belief. En route, we shall familiarize ourselves with the formalism. To begin with, we observe the following alternative formulation of its semantics, which sometimes is more convenient to work with than Rule S9.

Theorem 13. *$e \models \mathbf{B}(\alpha \Rightarrow \beta)$ iff $\lfloor e \mid \alpha \rfloor = \infty$ or $e, w \models (\alpha \supset \beta)$ for all $w \in e_{\lfloor e \mid \alpha \rfloor}$.*

Proof. For the *only-if* direction let $e \models \mathbf{B}(\alpha \Rightarrow \beta)$. Then by Rule S9, for all $p \in \mathbb{N}$, if $p \leq \lfloor e \mid \alpha \rfloor$ and $w \in e_p$, then $e, w \models (\alpha \supset \beta)$. If $\lfloor e \mid \alpha \rfloor = \infty$, the right-hand side of the theorem trivially holds. Otherwise $e, w \models (\alpha \supset \beta)$ for all $w \in e_{\lfloor e \mid \alpha \rfloor}$, and the right-hand side holds again.

For the *if* direction first let $\lfloor e \mid \alpha \rfloor = \infty$. Then $e, w \not\models \alpha$ for all $w \in e_p$ and $p \in \mathbb{N}$. Hence $e, w \models (\alpha \supset \beta)$ for all $w \in e_p$ and $p \in \mathbb{N}$, and so $e \models \mathbf{B}(\alpha \Rightarrow \beta)$ by Rule S9. Now let $\lfloor e \mid \alpha \rfloor \neq \infty$ and $e, w \models (\alpha \supset \beta)$ for all $w \in e_{\lfloor e \mid \alpha \rfloor}$. By Definition 7, $e_1 \subseteq \ldots \subseteq e_{\lfloor e \mid \alpha \rfloor}$. Thus for all $p \in \mathbb{N}$, if $p \leq \lfloor e \mid \alpha \rfloor$ and $w \in e_p$, then $e, w \models (\alpha \supset \beta)$, which by Rule S9 gives $e \models \mathbf{B}(\alpha \Rightarrow \beta)$. □

Another easy exercise is to confirm that $\mathbf{K}\alpha$ indeed expresses knowledge of α as claimed above.

Theorem 14. *$e \models \mathbf{K}\alpha$ iff $e, w \models \alpha$ for all $w \in e_p$ and $p \in \mathbb{N}$.*

Proof. For the *only-if* direction, let $e \models \mathbf{K}\alpha$. By Rule S9, for all $p \in \mathbb{N}$, if $p \leq \lfloor e \mid \neg\alpha \rfloor$ and $w \in e_p$, then $e, w \models (\neg\alpha \supset \textsc{false})$, which simplifies to $e, w \models \alpha$ (*). We show by induction on p that $p \leq \lfloor e \mid \neg\alpha \rfloor$ for all $p \in \mathbb{N}$, which immediately gives us the right-hand side of the theorem. The base case holds trivially. For the induction step, suppose $p \leq \lfloor e \mid \neg\alpha \rfloor$. Then $e, w \models \alpha$ for all $w \in e_p$ by (*), and thus $\lfloor e \mid \neg\alpha \rfloor > p$, that is, $p + 1 \leq \lfloor e \mid \neg\alpha \rfloor$.

Conversely, let $e, w \models \alpha$ for all $w \in e_p$ and $p \in \mathbb{N}$. Then $\lfloor e \mid \neg\alpha \rfloor = \infty$, and by Rule S9, $e \models \mathbf{K}\alpha$. □

Next, we prove that $\mathbf{B}(\alpha \vee \beta \Rightarrow \neg\beta)$ says that α is more plausible than β.

Theorem 15. *$e \models \mathbf{B}(\alpha \vee \beta \Rightarrow \neg\beta)$ iff $\lfloor e \mid \alpha \rfloor < \lfloor e \mid \beta \rfloor$ or $\lfloor e \mid \alpha \rfloor = \lfloor e \mid \beta \rfloor = \infty$.*

Proof. $e \models \mathbf{B}(\alpha \vee \beta \Rightarrow \neg\beta)$ iff (by Theorem 13) $\lfloor e \mid \alpha \vee \beta \rfloor = \infty$ or $e, w \models (\alpha \vee \beta \supset \neg\beta)$ for all $w \in e_{\lfloor e \mid \alpha \vee \beta \rfloor}$. The former is equivalent to $\lfloor e \mid \alpha \rfloor = \lfloor e \mid \beta \rfloor = \infty$. The latter holds iff $e, w \not\models \beta$ for all $w \in e_{\lfloor e \mid \alpha \vee \beta \rfloor}$ iff $\lfloor e \mid \beta \rfloor > \lfloor e \mid \alpha \vee \beta \rfloor = \lfloor e \mid \alpha \rfloor$. □

Before we turn to more general properties of the conditional belief operator, we need the following simple lemma.

Lemma 16. *$\models \Box\alpha$ iff $\models \alpha$.*

Proof. For the *only-if* direction suppose $\models \Box\alpha$. Then by Rule S8, $e \gg z, w \gg z \models \alpha$ for all e, w, and z. In particular, this holds for $z = \langle\rangle$, and since $e \gg \langle\rangle = e$ and $w \gg \langle\rangle = w$, we have $e, w \models \alpha$ for all e, w, so $\models \alpha$. Conversely, suppose $\models \alpha$. Therefore and by Lemma 10, for all e, w, z, we have $e \gg z, w \gg z \models \alpha$, and by Rule S8 $e, w \models \Box\alpha$. Thus $\models \Box\alpha$. □

The following theorem establishes several general properties of the conditional belief operator. Notice that, by the above lemma, all the validities also hold after an arbitrary sequence of actions.

Theorem 17.

1. $\not\models \mathbf{K}\alpha \supset \alpha$;
2. $\not\models \mathbf{B}(\alpha \Rightarrow \beta) \land \mathbf{B}(\beta \Rightarrow \gamma) \supset \mathbf{B}(\alpha \Rightarrow \gamma)$;
3. $\not\models \mathbf{B}(\alpha \Rightarrow \gamma) \supset \mathbf{B}(\alpha \land \beta \Rightarrow \gamma)$;
4. $\not\models \mathbf{B}(\alpha \Rightarrow \beta) \equiv \mathbf{B}(\neg\beta \Rightarrow \neg\alpha)$;
5. $\models \mathbf{B}(\alpha \Rightarrow \beta) \supset \mathbf{B}(\alpha \supset \beta)$;
6. $\models \mathbf{B}\alpha \land \mathbf{B}\beta \supset \mathbf{B}(\alpha \Rightarrow \beta)$;
7. $\models \mathbf{B}\alpha \land \mathbf{B}(\alpha \supset \beta) \supset \mathbf{B}\beta$;
8. $\models \mathbf{K}\alpha \land \mathbf{K}(\alpha \supset \beta) \supset \mathbf{K}\beta$;
9. $\models \mathbf{B}\alpha \land \mathbf{B}(\alpha \Rightarrow \beta) \supset \mathbf{B}\beta$;
10. $\models \mathbf{K}\alpha \supset \mathbf{B}\alpha$;
11. $\models \mathbf{B}(\alpha \Rightarrow \beta) \supset \mathbf{K}\mathbf{B}(\alpha \Rightarrow \beta)$;
12. $\models \neg\mathbf{B}(\alpha \Rightarrow \beta) \supset \mathbf{K}\neg\mathbf{B}(\alpha \Rightarrow \beta)$;
13. $\models \forall x \mathbf{B}(\alpha \Rightarrow \beta) \supset \mathbf{B}(\alpha \Rightarrow \forall x \beta)$ *where x does not occur freely in α;*
14. $\models \mathbf{K}\alpha$ *if* $\models \alpha$;
15. $\models \mathbf{B}(\alpha \Rightarrow \mathbf{B}(\beta \Rightarrow \gamma)) \land \neg\mathbf{K}\neg\alpha \supset \mathbf{B}(\beta \Rightarrow \gamma)$.

Let us explain the properties in English before we prove them. To begin with, our logic does not require the agent's knowledge to be actually true, that is, the epistemic state does not have to contain the actual world (Property 1). As usual, neither transitivity nor monotonicity nor contraposition hold for conditional beliefs (Properties 2, 3, 4). A conditional belief does however imply believing the material implication, which corresponds to Lewis' [53] weak centering, and strong centering holds as well (Properties 5 and 6). Knowledge and belief are closed under modus ponens from material implications and from belief conditionals (Properties 7, 8, 9), and what is believed is a subset of what is known (Property 10). The abbreviations $\mathbf{B}\alpha$ and $\mathbf{K}\alpha$ both are negatively and positively introspective (Properties 10, 11, 12). Hence, they are weak-S5 operators [60]. The Barcan formula is satisfied as well (Property 13) and the agent is moreover omniscient (Property 14). Somewhat surprising is perhaps Property 15: when a conditional is nested in another conditional's consequent, then the outer conditional's antecedent is irrelevant to the inner conditional. Alternatively one could condition the nested belief on the outer conditional's antecedent as well. Our simple semantics has some advantages as we discuss in Section 2.

Proof.

1. Let R be a rigid atom and $e = \langle e_1 \rangle$ with $e_1 = \{w \mid w \models R\}$, and let w' a world such that $w' \not\models R$. We show that $e, w' \not\models \mathbf{K}R \supset R$, that is, $e, w' \models \mathbf{K}R \land \neg R$. Firstly, $e \models \mathbf{K}R$ because (by Theorem 14) $w \models R$ for all $w \in e_p$ and $p \in \mathbb{N}$. Secondly, by construction, $w' \models \neg R$.
2. Let R be a rigid atom and $e = \langle e_1, e_2 \rangle$ with $e_1 = \{w \mid w \models R\}$ and $e_2 = \{w \mid w \models \textsc{true}\}$. We show that $e \not\models \mathbf{B}(\neg R \Rightarrow \textsc{true}) \land \mathbf{B}(\textsc{true} \Rightarrow R) \supset \mathbf{B}(\neg R \Rightarrow R)$. Firstly, $e \models \mathbf{B}(\neg R \Rightarrow \textsc{true})$ iff (by Theorem 13) $\lfloor e \mid \neg R \rfloor = \infty$ or $w \models \neg R \supset \textsc{true}$ for all $w \in e_{\lfloor e \mid \neg R \rfloor}$, which trivially holds. Secondly, $e \models \mathbf{B}(\textsc{true} \Rightarrow R)$ iff (by Theorem 13) $\lfloor e \mid \textsc{true} \rfloor = \infty$ or $w \models \textsc{true} \supset R$ for all $w \in e_{\lfloor e \mid \textsc{true} \rfloor}$, which holds by definition of e_1. However, $e \models \mathbf{B}(\neg R \Rightarrow R)$ iff $\lfloor e \mid \neg R \rfloor = \infty$ or $w \models \neg R \supset R$ for all $w \in e_{\lfloor e \mid \neg R \rfloor}$, which is false since $\lfloor e \mid \neg R \rfloor = 2$ and $w \not\models R$ for some $w \in e_2$.
3. Let R and e be as in the previous case. We showed that $e \models \mathbf{B}(\textsc{true} \Rightarrow R)$, but $e \not\models \mathbf{B}(\neg R \Rightarrow R)$, so clearly strengthening the premise in $\mathbf{B}(\textsc{true} \Rightarrow R)$ by $\neg R$ renders it false: $e \not\models \mathbf{B}(\textsc{true} \land \neg R \Rightarrow R)$.
4. Again let R and e be as in the first case. We show $e \not\models \mathbf{B}(\textsc{true} \Rightarrow R) \equiv \mathbf{B}(\neg R \Rightarrow \neg\textsc{true})$, which is just what the abbreviation $\mathbf{B}R \equiv \mathbf{K}R$ stands for. In the first case we already showed that $e \models \mathbf{B}(\textsc{true} \Rightarrow R)$. However, $w \not\models R$ for some $w \in e_2$, so $e \not\models \mathbf{K}R$ by Theorem 14.
5. We show that $e \models \mathbf{B}(\alpha \Rightarrow \beta) \supset \mathbf{B}(\alpha \supset \beta)$ for all e. According to Rule S9 we have
    - $e \models \mathbf{B}(\alpha \Rightarrow \beta)$ iff for all $p \in \mathbb{N}$, if $p \leq \lfloor e \mid \alpha \rfloor$ and $w' \in e_p$, then $e, w' \models (\alpha \supset \beta)$;

- $e \models \mathbf{B}(\alpha \supset \beta)$ iff for all $p \in \mathbb{N}$, if $p \leq \lfloor e \,|\, \textsc{true} \rfloor$ and $w' \in e_p$, then $e, w' \models \textsc{true} \supset (\alpha \supset \beta)$.

We show that the right-hand side of the first line subsumes the right-hand side of the second. Clearly, $e, w' \models \textsc{true} \supset (\alpha \supset \beta)$ iff $e, w' \models (\alpha \supset \beta)$. So it only remains to be shown that the second line's if-condition is at least as strong as the first line's, that is, that $\lfloor e \,|\, \alpha \rfloor \geq \lfloor e \,|\, \textsc{true} \rfloor$, which is clearly true.

6. We show that $e \models \mathbf{B}\alpha \wedge \mathbf{B}\beta \supset \mathbf{B}(\alpha \Rightarrow \beta)$ for all e. Let $e \models \mathbf{B}\alpha \wedge \mathbf{B}\beta$. By Theorem 13, either $\lfloor e \,|\, \textsc{true} \rfloor = \infty$, or $e, w \models (\textsc{true} \supset \alpha)$ and $e, w \models (\textsc{true} \supset \beta)$ for all $w \in e_{\lfloor e \,|\, \textsc{true} \rfloor}$. If $\lfloor e \,|\, \textsc{true} \rfloor = \infty$, then also $\lfloor e \,|\, \alpha \rfloor = \infty$ and hence $e \models \mathbf{B}(\alpha \Rightarrow \beta)$ by Theorem 13. Otherwise, $e_{\lfloor e \,|\, \textsc{true} \rfloor} \neq \{\}$ and $e, w \models (\alpha \wedge \beta)$ for all $w \in e_{\lfloor e \,|\, \textsc{true} \rfloor}$. Then clearly $\lfloor e \,|\, \alpha \rfloor = \lfloor e \,|\, \textsc{true} \rfloor$ and $e, w \models (\alpha \supset \beta)$ for all $w \in e_{\lfloor e \,|\, \alpha \rfloor}$, and therefore $e \models \mathbf{B}(\alpha \Rightarrow \beta)$ by Theorem 13.
7. We show that $e \models \mathbf{B}\alpha \wedge \mathbf{B}(\alpha \supset \beta) \supset \mathbf{B}\beta$ for all e. Let $e \models \mathbf{B}\alpha \wedge \mathbf{B}(\alpha \supset \beta)$. We need to show that $e \models \mathbf{B}\beta$, which by Theorem 13 holds iff $\lfloor e \,|\, \textsc{true} \rfloor = \infty$ or $e, w \models \beta$ for all $w \in e_{\lfloor e \,|\, \textsc{true} \rfloor}$. Suppose $\lfloor e \,|\, \textsc{true} \rfloor \neq \infty$, for otherwise $e \models \mathbf{B}\beta$ follows trivially. From the antecedent $e \models \mathbf{B}\alpha \wedge \mathbf{B}(\alpha \supset \beta)$ we obtain by Theorem 13 that $e, w \models \alpha \wedge (\alpha \supset \beta)$ for all $w \in e_{\lfloor e \,|\, \textsc{true} \rfloor}$. Thus $e, w \models \beta$ for all $w \in e_{\lfloor e \,|\, \textsc{true} \rfloor}$, and so $e \models \mathbf{B}\beta$.
8. We show that $e \models \mathbf{K}\alpha \wedge \mathbf{K}(\alpha \supset \beta) \supset \mathbf{K}\beta$ for all e. Let $e \models \mathbf{K}\alpha \wedge \mathbf{K}(\alpha \supset \beta)$. By Theorem 14, $e, w \models \alpha$ and $e, w \models (\alpha \supset \beta)$ for all $w \in e_p$ and $p \in \mathbb{N}$. Hence, $e, w \models \beta$ for all $w \in e_p$ and $p \in \mathbb{N}$, and so $e, w \models \mathbf{K}\beta$ by Theorem 14.
9. We show that $e \models \mathbf{B}\alpha \wedge \mathbf{B}(\alpha \Rightarrow \beta) \supset \mathbf{B}\beta$ for all e. Let $e \models \mathbf{B}\alpha \wedge \mathbf{B}(\alpha \Rightarrow \beta)$. By Property 5, $e \models \mathbf{B}(\alpha \supset \beta)$, and thus by Property 7, $e \models \mathbf{B}\beta$.
10. Let $e \models \mathbf{K}\alpha$. Then by Theorem 14, $e, w \models \alpha$ for all $w \in e_p$ and $p \in \mathbb{N}$, particularly when $p \leq \lfloor e \,|\, \textsc{true} \rfloor$. Thus by Rule S9, $e \models \mathbf{B}\alpha$.
11. Let $e \models \mathbf{B}(\alpha \Rightarrow \beta)$. Then $e, w \models \mathbf{B}(\alpha \Rightarrow \beta)$ for arbitrary w, and particularly for all $w \in e_p$ and $p \in \mathbb{N}$. Thus by Theorem 14, $e \models \mathbf{KB}(\alpha \Rightarrow \beta)$.
12. Let $e \not\models \mathbf{B}(\alpha \Rightarrow \beta)$. Then similar to the above, $e, w \not\models \mathbf{B}(\alpha \Rightarrow \beta)$ for arbitrary w, and particularly for all $w \in e_p$ and $p \in \mathbb{N}$. Thus by Theorem 14, $e \models \mathbf{K}\neg\mathbf{B}(\alpha \Rightarrow \beta)$.
13. Let $e \models \forall x \mathbf{B}(\alpha \Rightarrow \beta)$. By Rules S4, S6, and S9, for all standard names n, for all $p \in \mathbb{N}$, if $p \leq \lfloor e \,|\, \alpha \rfloor$ and $w \in e_p$, then $e, w \models \alpha \supset \beta^x_n$. Reintroducing the quantifier by Rules S4 and S6 in front of β yields that for all $p \in \mathbb{N}$, if $p \leq \lfloor e \,|\, \alpha \rfloor$ and $w \in e_p$, then $e, w \models \alpha \supset \forall x \beta$. Thus by Rule S9, $e \models \mathbf{B}(\alpha \Rightarrow \forall x \beta)$.
14. Let $e, w \models \alpha$ for all e, w. Then $e, w \models \neg\alpha \supset \textsc{false}$ for all $w \in e_p$ and $p \in \mathbb{N}$ for all e, so by Theorem 14, $e \models \mathbf{K}\alpha$ follows.
15. Let $e \models \mathbf{B}(\alpha \Rightarrow \mathbf{B}(\beta \Rightarrow \gamma)) \wedge \neg\mathbf{K}\neg\alpha$. The first assumption implies by Theorem 13 that $\lfloor e \,|\, \alpha \rfloor \neq \infty$ or $e, w \models \alpha \supset \mathbf{B}(\beta \Rightarrow \gamma)$ for all $w \in e_{\lfloor e \,|\, \alpha \rfloor}$. The second assumption implies by Theorem 14 that $\lfloor e \,|\, \alpha \rfloor \neq \infty$, and thus $e, w \models \alpha$ for some $w \in e_{\lfloor e \,|\, \alpha \rfloor}$. Hence $e, w \models \alpha \wedge (\alpha \supset \mathbf{B}(\beta \Rightarrow \gamma))$ for that w, so $e \models \mathbf{B}(\beta \Rightarrow \gamma)$. □

3.4. Properties of only-believing

The idea behind only-believing is to specify that a given set of conditional beliefs is *all* the agent believes. Only-believing hence not only captures what the agent believes, but also what she does not believe. Such an account not only of belief but also non-belief is vital to obtain meaningful results for queries that involve meta-belief such as $\mathbf{B}(\mathit{InBox}(\mathit{gift}) \wedge \neg\exists y \mathbf{B}\,\mathit{gift} = y)$, which says that the agent believes an unidentified gift is in the box.

In this section we show that only-believing exhibits a *unique* model. As it turns out, this model can be determined in a greedy fashion, namely by maximizing the spheres one after another, starting with the inner-most. While this is not to say that this is the only imaginable way to express only-belief, we will observe a close relationship between only-believing and Levesque's only-knowing as well as Pearl's System Z. We will further see that an ordinary closed-world assumption for conditional belief is not well-behaved and hence not a viable alternative to only-believing.

To begin with, we show the unique-model property for only-believing in the next two theorems.

Theorem 18. *Let* $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ *be objective. If* $e \models \mathbf{O}\Gamma$ *and* $e' \models \mathbf{O}\Gamma$*, then* $e = e'$.

Proof. Let $e \models \mathbf{O}\Gamma$ and $e' \models \mathbf{O}\Gamma$. We show by induction on $p \in \mathbb{N}$ that $e_p = e'_p$ and that $\lfloor e \mid \phi_i \rfloor > p$ iff $\lfloor e' \mid \phi_i \rfloor > p$ for all i. For the base case consider $p = 1$. By Rule S10, $w \in e_1$ iff $w \models \bigwedge_{1 \le i \le m}(\phi_i \supset \psi_i)$ iff $w \in e'_1$. Thus $e_1 = e'_1$, and $\lfloor e \mid \phi_i \rfloor > 1$ iff $w \not\models \phi_i$ for all $w \in e_1 = e'_1$ iff $\lfloor e' \mid \phi_i \rfloor > 1$. For the induction step suppose the statement holds for $p-1$. By induction, $\lfloor e \mid \phi_i \rfloor \ge p$ iff $\lfloor e' \mid \phi_i \rfloor \ge p$ for all i (*). By Rule S10, $w \in e_p$ iff $w \models \bigwedge_{i:\lfloor e \mid \phi_i \rfloor \ge p}(\phi_i \supset \psi_i)$ iff (by (*)) $w \models \bigwedge_{i:\lfloor e' \mid \phi_i \rfloor \ge p}(\phi_i \supset \psi_i)$ iff $w \in e'_p$. Thus $e_p = e'_p$, and $\lfloor e \mid \phi_i \rfloor > p$ iff $w \not\models \phi_i$ for all $w \in e_p = e'_p$ iff $\lfloor e' \mid \phi_i \rfloor > p$. □

Theorem 19. *Let $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ be objective. Then there is an e such that $e \models \mathbf{O}\Gamma$.*

Proof. Let $e = \langle e_1, \ldots, e_{m+1} \rangle$, where $e_1 = \{w \mid w \models \bigwedge_i(\phi_i \supset \psi_i)\}$ and $e_p = \{w \mid w \models \bigwedge_{i:\lfloor \langle e_1, \ldots, e_{p-1} \rangle \mid \phi_i \rfloor \ge p}(\phi_i \supset \psi_i)\}$ for $p > 1$. This is well-defined as the right-hand side for e_p only refers to $e_1, \ldots, e_{p-1}$. Note that $\lfloor \langle e_1, \ldots, e_{p-1} \rangle \mid \phi \rfloor \ge p$ iff $\lfloor e \mid \phi \rfloor \ge p$ for any objective ϕ (*). To see that for all i either $\lfloor e \mid \phi_i \rfloor \le m$ or $\lfloor e \mid \phi_i \rfloor = \infty$ (**), suppose there is a "hole" in the plausibility ranking, that is, there is some p and i such that $p + 1 = \lfloor e \mid \phi_i \rfloor \ne \infty$, and $\lfloor e \mid \phi_j \rfloor \ne p$ for all j. Then $w \in e_p$ iff (by (*)) $w \models \bigwedge_{k:\lfloor e \mid \phi_k \rfloor \ge p}(\phi_k \supset \psi_k)$ iff (since p is a hole) $w \models \bigwedge_{k:\lfloor e \mid \phi_k \rfloor \ge p+1}(\phi_k \supset \psi_k)$ iff $w \in e_{p+1}$. Then $w \models \phi_i$ for some $w \in e_{p+1} = e_p$, which contradicts the assumption $p + 1 = \lfloor e \mid \phi_i \rfloor$. By (*) and (**), e satisfies Rule S10. □

The proof reveals a close relationship between only-believing and System Z [57]: the definition of e reflects Pearl's labeling procedure which produces his Z-ranking. Only-believing orders the $\phi_i \Rightarrow \psi_i$ by their plausibilities $\lfloor e \mid \phi_i \rfloor$; this ordering thus matches the Z-ranking. (System Z is only defined for propositional logic and requires the conditionals to be consistent, which in our terminology means that $\lfloor e \mid \phi_i \rfloor \ne \infty$ for all rules. However, these restrictions are not essential for the Z-ranking. A full analysis can be found in [56].)

Together, Theorems 18 and 19 constitute the unique-model property of only-believing.

Corollary 20. *Let $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ be objective. Then there is one unique e with $e \models \mathbf{O}\Gamma$.*

Intuitively, only-believing expresses *all* that is believed by maximizing the epistemic state. More precisely, for some formula α involving beliefs, we say e is maximal with $e \models \alpha$ when no worlds can be added to any plausibility level without falsifying α, that is, $e' \not\models \alpha$ for all e' with $e'_p \supseteq e_p$ for all $p \in \mathbb{N}$ and $e'_{p'} \supsetneq e_{p'}$ for some $p' \in \mathbb{N}$.

Theorem 21. *Let $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ be objective. Then $e \models \mathbf{O}\Gamma$ iff $e \models \bigwedge_i \mathbf{B}(\phi_i \Rightarrow \psi_i)$ and e is maximal.*

Proof. For the *if* direction suppose $e \models \mathbf{B}(\phi_i \Rightarrow \psi_i)$ for all i and e is maximal with that property. Let $p \in \mathbb{N}$ and w be a world. By Rule S9, if $p \le \lfloor e \mid \phi_i \rfloor$ and $w \in e_p$, then $w \models \phi_i \supset \psi_i$ for all i. Since e is maximal, if $p \le \lfloor e \mid \phi_i \rfloor$ and $w \notin e_p$, then $w \not\models \phi_i \supset \psi_i$ for some i. Therefore $w \models \bigwedge_{i:\lfloor e \mid \phi_i \rfloor \ge p}(\phi_i \supset \psi_i)$ iff $w \in e_p$. Hence $e \models \mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$.

For the *only-if* direction suppose $e \models \mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$. Then from Rules S9 and S10 immediately follows that $e \models \mathbf{B}(\phi_i \Rightarrow \psi_i)$ for all i. To see that e is maximal, suppose $e' \models \mathbf{B}(\phi_i \Rightarrow \psi_i)$ for all i and $e_p \subseteq e'_p$ for all $p \in \mathbb{N}$. Then by the *if* direction $e' \models \mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$, and by Theorem 18 $e = e'$. □

The theorem illustrates the connection between only-believing and Levesque's only-knowing [26]. Only-knowing α means that *all* the agent *knows* is α, where knowledge differs from belief in that it cannot be revised. In Levesque's semantics, this means the epistemic state satisfies $\mathbf{K}\alpha$ and is maximal. The very same effect is achieved in our logic by $\mathbf{O}\{\neg\alpha \Rightarrow \textsc{false}\}$. Only-believing can hence be considered to generalize only-knowing to the case of conditional beliefs. For the formal proof that $\mathcal{ESB}$ subsumes Levesque's logic of only-knowing we refer to [56].

Only-believing a set of conditionals aims to express that these conditionals are believed, but nothing else is. This idea resembles a closed-world assumption, and indeed Levesque's only-knowing can be simulated by a closed-world assumption on knowledge to some extent: Reiter [4] proposes the infinite knowledge base $\{\mathbf{K}\phi\} \cup \{\neg\mathbf{K}\psi \mid \mathbf{K}\phi \not\models \mathbf{K}\psi\}$ to capture that ϕ is all that is known. For queries without nested $\mathbf{K}$, this closed-world assumption on knowledge is equivalent to Levesque's only-knowing. (It fails to deal with introspective queries like $\mathbf{K}(\mathit{InBox}(\mathit{gift}) \land \neg\exists y \mathbf{K}\mathit{gift} = y)$, though.)

So the question may arise: does Reiter's closed-world assumption work similarly for conditional belief? Perhaps surprisingly, the answer is *no*. As a matter of fact, making a closed-world assumption on belief does not even preserve satisfiability. Analogous to Reiter's definition [4], let us define the closed-world assumption on $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ as

$$\mathrm{CWA}(\Gamma) = \{\textstyle\bigwedge_i \mathbf{B}(\phi_i \Rightarrow \psi_i)\} \cup \{\neg\mathbf{B}(\phi \Rightarrow \psi) \mid \textstyle\bigwedge_i \mathbf{B}(\phi_i \Rightarrow \psi_i) \not\models \mathbf{B}(\phi \Rightarrow \psi)\}.$$

To see that this closed-world assumption does not preserve satisfiability, consider $\Gamma = \{R_1 \Rightarrow R_2, R_3 \Rightarrow R_4\}$: CWA($\Gamma$) contains the negative beliefs $\neg\mathbf{B}(R_1 \Rightarrow (R_3 \supset R_4))$ and $\neg\mathbf{B}(R_3 \Rightarrow (R_1 \supset R_2))$, which are inconsistent with the positive beliefs $\mathbf{B}(R_1 \Rightarrow R_2)$ and $\mathbf{B}(R_3 \Rightarrow R_4)$ in CWA(Γ) (as can be easily verified by confirming that R_1 and R_3 must be equally plausible, and at least one of the most-plausible R_1-worlds must falsify $(R_3 \supset R_4)$, but all of them must satisfy $(R_3 \supset R_4)$, which is a contradiction). A closed-world assumption on conditional belief is therefore no viable alternative to only-believing, as it adds too many negative beliefs. In particular, there seems to be no obvious way of capturing a conditional knowledge base in formalisms like Shapiro et al.'s [12].

However, only-believing is not free from counterintuitive behavior either. In the above example, it is somewhat odd that $\mathbf{O}\{R_1 \Rightarrow R_2, R_3 \Rightarrow R_4\} \models \neg\mathbf{B}(R_1 \wedge \neg R_2 \wedge R_3 \Rightarrow R_4)$, although there is no reason why evidence of $R_1 \wedge \neg R_2$ should defeat that R_3 usually implies R_4. The same problem occurs in System Z. It is noteworthy that c-representations [61] avoid this problem.

4. The belief projection problem

The belief projection problem is to decide if a certain belief holds true after a sequence of actions. Logically this is expressed as an entailment problem: given a knowledge base about the domain's dynamics and the agent's initial (conditional) beliefs, does a certain formula about actions and beliefs follow?

Besides the agent's initial beliefs, such a knowledge base needs to represent the dynamics of the worlds. For this purpose, we adopt Reiter's solution of the frame problem through *successor-state axioms* [4, 24]. The idea is to specify one axiom per fluent that specifies its truth value after the next action in terms of some formula before that action. In our language, a successor-state axiom for a fluent F has the form $\square[a]F(x_1, \dots, x_k) \equiv \gamma_F$, where γ_F is static The axiom says that after any sequence of actions, the fluent $F(x_1, \dots, x_k)$ is true after another action a iff the formula γ_F is true before that action a. It is very common to have γ_F be of the form $\gamma_F^+ \vee F(x_1, \dots, x_k) \wedge \neg\gamma_F^-$, where γ_F^+ represents the positive effect condition, that is, the condition under which $F(x_1, \dots, x_k)$ is set to true, and γ_F^- accordingly is the negative effect condition. Such an axiom $\square[a]F(x_1, \dots, x_k) \equiv \gamma_F^+ \vee F(x_1, \dots, x_k) \wedge \neg\gamma_F^-$ then intuitively means that $F(x_1, \dots, x_k)$ holds after a iff it is either turned on (γ_F^+) or not turned off ($\neg\gamma_F^-$) by a.

While successor-state axioms take care of the physical effects of actions, another kind of axiom is needed to capture their epistemic effects. Recall that any action a induces a belief revision by $IF(a)$. In order to assign a deeper meaning to this new information, a knowledge base shall contain an axiom of the form $\square IF(a) \equiv \varphi$, which we call an *informed-fluent axiom*. A customary pattern for φ is a conjunction of formulas $((a = A(x_1, \dots, x_k)) \supset \varphi_A)$: when an action matches the antecedent of one of the conjuncts, it revises φ_A; if it matches none of the conjuncts, then φ evaluates to TRUE, which effectively means that the action has no epistemic effect.

Based on these ideas, and in the spirit Reiter's situation calculus [4, 24], the knowledge bases we consider in this paper adhere to the following definition.

Definition 22. Let $\mathcal{F}$ be a finite set of fluent predicate symbols and $IF \notin \mathcal{F}$. A formula is *fluent* when it is objective, static, and all fluent predicate symbols are from $\mathcal{F}$. A *basic action theory* over $\mathcal{F}$ consists of two sets Σ_{dyn} and Σ_{bel}, where

- Σ_{dyn} contains dynamic axioms, namely:
  - a sentence $\square[a]F(x_1, \dots, x_k) \equiv \gamma_F$ for every fluent predicate symbol $F \in \mathcal{F}$ where γ_F is a fluent formula;
  - a single sentence $\square IF(a) \equiv \varphi$ where φ is a fluent formula;[3]
- Σ_{bel} contains finitely many conditionals $\phi \Rightarrow \psi$ where ϕ and ψ are fluent sentences.

We identify Σ_{dyn} with the conjunction of its elements and let $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$ stand for $\mathbf{O}(\{\neg\Sigma_{\text{dyn}} \Rightarrow \text{FALSE}\} \cup \Sigma_{\text{bel}})$. Then the *belief projection problem* is to decide entailments of the form

$$\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models [t_1] \dots [t_k]\mathbf{B}(\alpha \Rightarrow \beta).$$

[3] Usually, basic action theories also feature a precondition axiom of the form $\square Poss(a) \equiv \pi$ for a fluent formula π. It is generally treated very similar to the $\square IF(a) \equiv \varphi$ axiom. For simplicity, we omit it here.

Notice that in the conditional $\neg\Sigma_{\text{dyn}} \Rightarrow \text{FALSE}$ in $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$ expresses that the agent knows the dynamic axioms Σ_{dyn}, that is, the successor-state axioms and the informed-fluent axiom are unequivocal.

In Sections 5 and 6 we provide two solutions for the belief projection problem, one by regression and another by progression. Briefly said, regression takes the dynamics out of the projection problem by rewriting the query to undo the actions. Progression, on the other hand, eliminates the actions by applying their effects to the initial beliefs.

4.1. A model of being informed

The intuition behind our model of receiving new information is that actions *inform* the agent that some information is likely true. The informed-fluent axiom $\Box IF(a) \equiv \varphi$ of a basic action theory is to specify this, and before the action a is performed, the beliefs are revised by $IF(a)$. Hence $IF(a)$ can also be considered a (plausible) precondition of a being executable; so whenever a is executed, it is plausible that $IF(a)$ is true, and the revision by $IF(a)$ ensures that this is believed. Let us briefly discuss the advantages of informing over the classical sensing model for our purposes.

The traditional approach to perception in the situation calculus is that an action *senses* whether an associated formula holds in the actual world [30, 24]. The intuitive difference between classical sensing and our approach is as follows. Classical sensing answers questions based on the real world's *ground truth* (for example, "is the gift broken?"); such an answer is definitive and cannot be retracted. In our model, on the other hand, actions *inherently* carry information which is assumed to be likely true (for example, "the gift is broken" or "the gift is not broken"); that information is not checked back with the real world and may thus very well be wrong.

The view of informing as opposed to sensing enables us to handle contradicting information in a reasonable way. Whereas the classical approach then gets caught in an inconsistent state (where "everything" is known), our logic displaces the old information by belief revision.

The intuition behind an informing action is however somewhat different from a classical sensing action. The classical view is that the action is under the control of the agent. Informing actions are often better seen as *exogenous* actions, meaning not the agent executes them but *nature*. For example, the clinking noise in our introductory example is not performed by the agent, but is an event that occurs outside of the agent's control. (We refer to [4] for how to model the emergence of exogenous actions in the situation calculus.)

Nevertheless we can mimic classical sensing in our approach using an additional parameter for the sensing outcome. Suppose A should sense whether or not ϕ holds in the actual world. We can simulate this with a strong-revision action $A(outcome)$, where *outcome* takes a binary value to represent whether ϕ holds in the real world (for example, $outcome = {}^{\#}1$ iff ϕ is true), and $IF(A(y))$ is defined as ϕ or $\neg\phi$ depending on the value of y. Then $A(outcome)$ informs the agent about the real-world value of ϕ, and the revision promotes the worlds that accord with this value. Thus, when the agent performs a sequence of sensings of this form, she believes their outcomes afterwards (provided they are consistent): if she first senses ϕ_1 and then $\neg\phi_2$, say, then after strong revision by ϕ_1 and then by $\neg\phi_2$, the most-plausible worlds satisfy $\phi_1 \wedge \neg\phi_2$ (this follows from the NPP3 and AGM2 postulates shown in Section 7). This modeling is not limited to binary sensing. For instance, a sonar sensor that senses a distance to some obstacle can be represented with an action $sonar(y)$, where y is the sensed distance. When the basic action theory stipulates $IF(sonar(y)) \equiv distance = y$, then $\forall y((distance = y) \supset [sonar(y)]\mathbf{B}(distance = y))$ holds, that is, the agent believes the correct distance.

4.2. An example

We recap the gift box example from Section 1 and model it with a basic action theory. The scenario comprises a single box that may contain items, which we represent by a fluent predicate $InBox(n)$. The box can be dropped by action *dropbox*, and an item n can be taken out of the box by action $unbox(n)$. Dropping the box breaks all fragile items in it, which is formalized using a rigid predicate $Fragile(n)$ and another fluent predicate $Broken(n)$. A clinking noise, represented by the action *clink*, indicates that something in the box seems to be broken: $\exists y (InBox(y) \wedge Broken(y))$. Intuitively *clink* is exogenous, that is, it is not under the agent's control but she observes (nature executing) a clink. Unboxing an item n through action $unbox(n)$ tells us that this item was in the box and is not broken: $InBox(n) \wedge \neg Broken(n)$. The successor-state axioms for *InBox* and *Broken* and the informed-fluent axiom constitute the dynamic axioms:

$$\begin{aligned}\Sigma_{\text{dyn}} = \{&\Box[a]\mathit{InBox}(y) \equiv \mathit{InBox}(y) \land a \neq \mathit{unbox}(y),\\ &\Box[a]\mathit{Broken}(y) \equiv \mathit{Broken}(y) \lor \mathit{InBox}(y) \land \mathit{Fragile}(y) \land a = \mathit{dropbox},\\ &\Box \mathit{IF}(a) \equiv (a = \mathit{clink} \supset \exists y\, (\mathit{InBox}(y) \land \mathit{Broken}(y))) \land \forall y\, (a = \mathit{unbox}(y) \supset \mathit{InBox}(y) \land \neg \mathit{Broken}(y))\}.\end{aligned}$$

We still need to decide of which revision sort the actions are. As pointed out earlier, there is no definite answer to the question which operator is most appropriate in certain circumstances. Since a clinking noise is a rather unreliable hint that something is broken, we decide to make *clink* a weak-revision action. By contrast, when one takes an object out of the box, that object must indeed have been in the box and be in one piece (otherwise one probably hallucinates), so *unbox*(n) shall be a strong-revision action. We let *dropbox* be a strong-revision action, too; since *IF*(*dropbox*) is vacuously true (as $\Sigma_{\text{dyn}} \models \mathit{IF}(\mathit{dropbox})$) the revision has no effect anyway.

Our agent believes that most likely the box is empty; but taking the possibility into account that she may be wrong about that, she believes that in this case only the gift would be in the box. We use the object constant *gift* to refer to the gift that may or may not be in the box. Note that it is not a standard name, so the agent might have no clue what the gift actually is. She moreover believes that if there was something in the box, it would not be broken yet. Thus we define the initial beliefs as

$$\begin{aligned}\Sigma_{\text{bel}} = \{&\textsc{true} \Rightarrow \forall y \neg \mathit{InBox}(y),\\ &\exists y\, \mathit{InBox}(y) \Rightarrow \forall y\, (\mathit{InBox}(y) \equiv y = \mathit{gift}),\\ &\exists y\, \mathit{InBox}(y) \Rightarrow \forall y\, (\mathit{InBox}(y) \supset \neg \mathit{Broken}(y))\}.\end{aligned}$$

Let us conclude the example for now by investigating a few queries.

1. Initially the agent believes the box is empty: $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models \mathbf{B} \forall y \neg \mathit{InBox}(y)$.
2. After dropping the box, she still believes the box is empty, but also that if something fragile is in the box, then presumably it is broken: $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models [\mathit{dropbox}](\mathbf{B}(\forall y \neg \mathit{InBox}(y)) \land \forall y\, \mathbf{B}(\mathit{InBox}(y) \land \mathit{Fragile}(y) \Rightarrow \mathit{Broken}(y)))$.
3. When a clink occurs after dropping the box, she comes to believe that the gift is in the box, but she has no idea what the gift is: $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models [\mathit{dropbox}][\mathit{clink}]\mathbf{B}(\mathit{InBox}(\mathit{gift}) \land \mathit{Broken}(\mathit{gift}) \land \neg \exists y\, \mathbf{B} \mathit{gift} = y)$.
4. When the object #5 is taken out of the box, she believes that this must be the gift, and that it is not broken after all:[4] $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models [\mathit{dropbox}][\mathit{clink}][\mathit{unbox}(\#5)] \exists y\, \mathbf{B}(\mathit{gift} = y \land \neg \mathit{InBox}(\mathit{gift}) \land \neg \mathit{Broken}(\mathit{gift}))$.

We will use the latter query to illustrate the techniques developed in the upcoming sections of the paper to automate the reasoning task. But for now let us verify the entailments through semantical proofs.

Firstly we need to determine the epistemic state $e \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$. By Corollary 20, e is unique, and using the idea from the proof of Theorem 19 we generate $e = \langle e_1, e_2, e_3 \rangle$. The first level e_1 contains all worlds that satisfy Σ_{dyn} and the material-implication-versions of all conditionals in Σ_{bel}, which simplifies to

$$e_1 = \{w \mid w \models \Sigma_{\text{dyn}} \land \forall y \neg \mathit{InBox}(y)\}.$$

Thus $\lfloor e \mid \textsc{true} \rfloor = 1$ and $\lfloor e \mid \exists y\, \mathit{InBox}(y) \rfloor > 1$, so the next level e_2 contains all worlds that satisfy Σ_{dyn} and $\exists y\, \mathit{InBox}(y) \supset \forall y\, (\mathit{InBox}(y) \equiv y = \mathit{gift})$ as well as $\exists y\, \mathit{InBox}(y) \supset \forall y\, (\mathit{InBox}(y) \supset \neg \mathit{Broken}(y))$, which simplifies to

$$e_2 = \{w \mid w \models \Sigma_{\text{dyn}} \land \forall y\, (\mathit{InBox}(y) \supset y = \mathit{gift} \land \neg \mathit{Broken}(y))\}.$$

Hence $\lfloor e \mid \exists y\, \mathit{InBox}(y) \rfloor = 2$, so all following levels contain all worlds that satisfy Σ_{dyn}, that is,

$$e_3 = \{w \mid w \models \Sigma_{\text{dyn}}\}.$$

[4]To ease the presentation we use *unbox*(#5) directly instead of using *unbox*(*item*) where *item* is a variable or a constant to represent a sensing outcome that depends on the real world, as sketched in Section 4.1.

The first query, $\mathbf{B}\forall y \neg \mathit{InBox}(y)$, obviously holds because $w \models \forall y \neg \mathit{InBox}(y)$ for all $w \in e_1$. For the other queries we need to progress e. The action *dropbox* makes each $\mathit{Broken}(n)$ true when $\mathit{Fragile}(n)$ and $\mathit{InBox}(n)$ are true. Since *dropbox* makes a (strong) revision by the vacuously true $\mathit{IF}(\mathit{dropbox})$, there effectively is no revision. The progression $e \gg \mathit{dropbox}$ is thus

$$(e \gg \mathit{dropbox})_1 = \{w \mid w \models \Sigma_{\text{dyn}} \wedge \forall y \neg \mathit{InBox}(y)\};$$
$$(e \gg \mathit{dropbox})_2 = \{w \mid w \models \Sigma_{\text{dyn}} \wedge \forall y (\mathit{InBox}(y) \supset y = \mathit{gift} \wedge (\mathit{Broken}(y) \equiv \mathit{Fragile}(y)))\};$$
$$(e \gg \mathit{dropbox})_3 = \{w \mid w \models \Sigma_{\text{dyn}} \wedge \forall y (\mathit{InBox}(y) \wedge \mathit{Fragile}(y) \supset \mathit{Broken}(y))\}.$$

It is then easy to see that the second query holds. Clearly $[\mathit{dropbox}]\mathbf{B}\forall y \neg \mathit{InBox}(y)$ is true. And for all n, there is some $w \in (e \gg \mathit{dropbox})_2$ such that $w \models \mathit{InBox}(n) \wedge \mathit{Fragile}(n)$, and then also $w \models \mathit{Broken}(n)$, so $[\mathit{dropbox}]\forall y \mathbf{B}(\mathit{InBox}(y) \wedge \mathit{Fragile}(y) \Rightarrow \mathit{Broken}(y))$ holds as well.

In the third query the agent hears a clink after dropping the box. The action *clink* does not change the truth value of any fluents, but it triggers a weak revision by $\exists y (\mathit{InBox}(y) \wedge \mathit{Broken}(y))$, that is, the most-plausible worlds from $e \gg \mathit{dropbox}$ satisfying this formula constitute the first plausibility level in $(e \gg \mathit{dropbox}) * \mathit{IF}(\mathit{clink})$. It is therefore easy to see that $(e \gg \mathit{dropbox}) \gg \mathit{clink} = (e \gg \mathit{dropbox}) * \mathit{IF}(\mathit{clink})$ is defined by

$$((e \gg \mathit{dropbox}) * \mathit{IF}(\mathit{clink}))_1 = \{w \mid w \models \Sigma_{\text{dyn}} \wedge \forall y (\mathit{InBox}(y) \equiv y = \mathit{gift}) \wedge \mathit{Broken}(\mathit{gift}) \wedge \mathit{Fragile}(\mathit{gift})\};$$
$$((e \gg \mathit{dropbox}) * \mathit{IF}(\mathit{clink}))_2 = \{w \mid w \models \Sigma_{\text{dyn}} \wedge \forall y (\mathit{InBox}(y) \supset y = \mathit{gift} \wedge \mathit{Broken}(\mathit{gift}) \wedge \mathit{Fragile}(\mathit{gift}))\};$$
$$((e \gg \mathit{dropbox}) * \mathit{IF}(\mathit{clink}))_3 = \{w \mid w \models \Sigma_{\text{dyn}} \wedge \forall y (\mathit{InBox}(y) \supset y = \mathit{gift} \wedge (\mathit{Broken}(y) \equiv \mathit{Fragile}(y)))\};$$
$$((e \gg \mathit{dropbox}) * \mathit{IF}(\mathit{clink}))_4 = \{w \mid w \models \Sigma_{\text{dyn}} \wedge \forall y (\mathit{InBox}(y) \wedge \mathit{Fragile}(y) \supset \mathit{Broken}(y))\}.$$

We therefore have $w \models \mathit{InBox}(\mathit{gift}) \wedge \mathit{Broken}(\mathit{gift})$ for all $w \in (e \gg \mathit{dropbox} \gg \mathit{clink})_1$. Moreover, the worlds do not agree on the denotation of *gift*, so there is no standard name n such that $w \models \mathit{gift} = n$ for all $w \in (e \gg \mathit{dropbox} \gg \mathit{clink})_1$. Thus the third query, $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models [\mathit{dropbox}][\mathit{clink}]\mathbf{B}(\mathit{InBox}(\mathit{gift}) \wedge \mathit{Broken}(\mathit{gift}) \wedge \neg\exists y \mathbf{B} \mathit{gift} = y))$, meaning that the gift is believed to be in the box and broken but the agent has no clue what the gift is, comes out true.

For the last query, we need to make another progression by $\mathit{unbox}(^{\#}5)$. Firstly, the state is strongly revised by $\mathit{IF}(\mathit{unbox}(^{\#}5))$, which is equivalent to $\mathit{InBox}(^{\#}5) \wedge \neg \mathit{Broken}(^{\#}5)$. The first two levels of the revised state thus contain the $\mathit{IF}(\mathit{unbox}(^{\#}5))$-worlds from $(e \gg \mathit{dropbox} \gg \mathit{clink})_3$ and $(e \gg \mathit{dropbox} \gg \mathit{clink})_4$. For space reasons we only consider the first plausibility level, which is

$$((e \gg \mathit{dropbox} \gg \mathit{clink}) * \mathit{IF}(\mathit{unbox}(^{\#}5)))_1 =$$
$$\{w \mid w \models \Sigma_{\text{dyn}} \wedge \forall y (\mathit{InBox}(y) \equiv y = {^{\#}5}) \wedge \mathit{gift} = {^{\#}5} \wedge \neg \mathit{Broken}(^{\#}5) \wedge \neg \mathit{Fragile}(^{\#}5)\},$$

and when we then apply the physical effect of $\mathit{unbox}(^{\#}5)$, namely make $\mathit{InBox}(^{\#}5)$ false, we obtain

$$(e \gg \mathit{dropbox} \gg \mathit{clink} \gg \mathit{IF}(\mathit{unbox}(^{\#}5)))_1 =$$
$$\{w \mid w \models \Sigma_{\text{dyn}} \wedge \forall y \neg \mathit{InBox}(y) \wedge \mathit{gift} = {^{\#}5} \wedge \neg \mathit{InBox}(^{\#}5) \wedge \neg \mathit{Broken}(^{\#}5) \wedge \neg \mathit{Fragile}(^{\#}5)\}.$$

The fourth query, $[\mathit{dropbox}][\mathit{clink}][\mathit{unbox}(^{\#}5)]\exists y \mathbf{B}(\mathit{gift} = y \wedge \neg \mathit{InBox}(\mathit{gift}) \wedge \neg \mathit{Broken}(\mathit{gift}))$, is thus true, because all worlds at the first plausibility level agree on *gift* being $^{\#}5$.

It is interesting that the fourth query would not have come out true if *clink* was a strong-revision action. Then the agent would have rather believed that there were two (or more) items in the box than that the clink was due to something other than an object in the box breaking. That is quite reasonable: a strong-revision action *clink* would have made the agent very reluctant to give up the belief that something inside the box broke.

5. Belief projection by regression

The first solution we offer for the belief projection problem is by regression. Regression rewrites a formula about future situations to a formula about the initial situation. The idea, due to Reiter [3, 4], is to successively replace subformulas $[t]F(t_1, \ldots, t_k)$ with the right-hand side of F's successor-state axiom γ_F. Intuitively this is sound because the successor-state axioms ensure that actions have deterministic effects. As we shall see in this section, we can regress beliefs after actions in a similar way. Our regression operator can thus handle formulas with no $\square$ or **O**, provided the fluent predicates are taken from $\mathcal{F} \cup \{IF\}$. We call such a formula *regressable*.

To ease the technical treatment we assume that the formula to be regressed adheres to the following form:

- it is rectified: quantifiers use distinct variables, and none of the variables occurs in the basic action theory;
- action terms $A(t_1, \dots, t_n)$ have only standard names or variables as arguments t_i.

It is easy to see that any formula can be rewritten to satisfy these constraints. For example, $IF(unbox(gift))$ is transformed to $\exists y (y = gift \wedge IF(unbox(y)))$. The first restriction is needed because otherwise scopes of variables may collide during regression. The second one will allow us to push action operators inside **B**, which would be inappropriate for action terms like *unbox*(*gift*) because the denotation of *gift* shall be determined by the real world.

Section 5.1 is only concerned with objective formulas, whose regression works the same way as in $\mathcal{ES}$ [24]. Section 5.2 then extends the regression operator to also handle beliefs as well. In Section 5.3 we discuss an example.

5.1. Regression of objective formulas

For objective regressable formulas, our regression operator follows the one presented by Lakemeyer and Levesque [24] for the logic $\mathcal{ES}$. Intuitively, it works by pushing action operators $[t]$ inwards and replacing subformulas of the form $[t]F(t_1, \dots, t_k)$ for $F \in \mathcal{F}$ with the right-hand side of F's successor-state axiom γ_F as well as replacing atoms $IF(t)$ with the right-hand side of the informed-fluent axiom φ. That way, eventually all $[t]$ operators are eliminated.

Definition 23. The regression of an objective regressable formula α after actions r with respect to a basic action theory over fluents $\mathcal{F}$ is defined as follows:

R1. $\mathcal{R}[r, R(t_1, \dots, t_k)] = R(t_1, \dots, t_k)$ for rigid predicate symbols R;

R2. $\mathcal{R}[r, F(t_1, \dots, t_k)]$ for fluent predicate symbols $F \in \mathcal{F}$ is defined inductively on r:

- $\mathcal{R}[\langle\rangle, F(t_1, \dots, t_k)] = F(t_1, \dots, t_k)$;
- $\mathcal{R}[r \cdot t, F(t_1, \dots, t_k)] = \mathcal{R}[r, {\gamma_F}^{x_1 \dots x_k\, a}_{t_1 \dots t_k\, t}]$;

R3. $\mathcal{R}[r, IF(t)] = \mathcal{R}[r, \varphi^a_t]$;

R4. $\mathcal{R}[r, (t_1 = t_2)] = (t_1 = t_2)$;

R5. $\mathcal{R}[r, \neg\alpha] = \neg\mathcal{R}[r, \alpha]$;

R6. $\mathcal{R}[r, (\alpha_1 \vee \alpha_2)] = (\mathcal{R}[r, \alpha_1] \vee \mathcal{R}[r, \alpha_2])$;

R7. $\mathcal{R}[r, \exists x\, \alpha] = \exists x \mathcal{R}[r, \alpha]$;

R8. $\mathcal{R}[r, [t]\alpha] = \mathcal{R}[r \cdot t, \alpha]$.

We write $\mathcal{R}[\alpha]$ for $\mathcal{R}[\langle\rangle, \alpha]$.

Theorem 24. *Let Σ_{dyn} be the dynamic axioms of a basic action theory, ϕ be a fluent sentence, and ψ be an objective regressable sentence. Then*

$$\Sigma_{\text{dyn}} \wedge \phi \models \psi \quad \textit{iff} \quad \phi \models \mathcal{R}[\psi].$$

The theorem is very similar to one in [24]. The proof is in Appendix A.

5.2. Regression of non-objective formulas

The key to extending regression to beliefs is the relationship between beliefs after an action and the (conditional) beliefs before that action. We will use this correspondence to regress beliefs similar to how we use successor-state axioms to regress fluent atoms.

The next two lemmas relate the plausibilities of sentences in e to $e \gg n$. These relations are not surprising as they mirror the definitions of $e *_{\mathrm{w}} IF(n)$ and $e *_{\mathrm{s}} IF(n)$, respectively, from Definition 8.

Lemma 25. *Let n be a weak-revision action standard name and* $\lfloor e \,|\, IF(n) \rfloor \neq \infty$.

(i) If $\lfloor e \,|\, IF(n) \rfloor = \lfloor e \,|\, IF(n) \land [n]\alpha \rfloor$, *then* $\lfloor e \gg n \,|\, \alpha \rfloor = 1$.

(ii) If $\lfloor e \,|\, IF(n) \rfloor \neq \lfloor e \,|\, IF(n) \land [n]\alpha \rfloor$, *then* $\lfloor e \gg n \,|\, \alpha \rfloor = \lfloor e \,|\, [n]\alpha \rfloor + 1$.

Proof. (i) By assumption $e, w \models IF(n) \land [n]\alpha$ for some $w \in e_{\lfloor e \,|\, IF(n) \rfloor}$. Thus $e, w \models [n]\alpha$ for some $w \in (e *_{\mathrm{w}} IF(n))_1$, and so by Rule S7, $e \gg n, w \models \alpha$ for some $w \in (e \gg n)_1$. Therefore $\lfloor e \gg n \,|\, \alpha \rfloor = 1$.

(ii) By assumption, $e, w \not\models IF(n) \land [n]\alpha$ for all $w \in e_{\lfloor e \,|\, IF(n) \rfloor}$. Thus $e, w \not\models [n]\alpha$ for all $w \in (e *_{\mathrm{w}} IF(n))_1$ (*). Thus $e \gg n, w \not\models \alpha$ for all $w \in (e \gg n)_1$, and hence $\lfloor e \gg n \,|\, \alpha \rfloor > 1$. Now let $p \in \mathbb{N}$. First suppose $p < \lfloor e \,|\, [n]\alpha \rfloor$. Then $e, w \not\models [n]\alpha$ for all $w \in e_p$. Hence and by (*), $e, w \not\models [n]\alpha$ for all $w \in (e *_{\mathrm{w}} IF(n))_{p+1}$. Thus $e \gg n, w \not\models \alpha$ for all $w \in (e \gg n)_{p+1}$, and therefore $p + 1 < \lfloor e \gg n \,|\, \alpha \rfloor$. Now suppose $p \geq \lfloor e \,|\, [n]\alpha \rfloor$. Then $e, w \models [n]\alpha$ for some $w \in e_p \subseteq (e *_{\mathrm{w}} IF(n))_{p+1}$. Thus $e \gg n, w \models \alpha$ for some $w \in (e \gg n)_{p+1}$, and hence $p + 1 \geq \lfloor e \gg n \,|\, \alpha \rfloor$. □

Lemma 26. *Let n be a strong-revision action standard name and* $\lfloor e \,|\, IF(n) \rfloor \neq \infty$.

(i) If $\lfloor e \,|\, IF(n) \land [n]\alpha \rfloor \neq \infty$, *then* $\lfloor e \gg n \,|\, \alpha \rfloor = \lfloor e \,|\, IF(n) \land [n]\alpha \rfloor - \lfloor e \,|\, IF(n) \rfloor + 1$.

(ii) If $\lfloor e \,|\, IF(n) \land [n]\alpha \rfloor = \infty$ *and* $\lfloor e \,|\, \neg IF(n) \rfloor \neq \infty$, *then* $\lfloor e \gg n \,|\, \alpha \rfloor = \lfloor e \,|\, [n]\alpha \rfloor + \lceil e \rceil - \lfloor e \,|\, IF(n) \rfloor - \lfloor e \,|\, \neg IF(n) \rfloor + 2$.

Proof. (i) Suppose $\lfloor e \,|\, IF(n) \rfloor \leq p < \lfloor e \,|\, IF(n) \land [n]\alpha \rfloor$. Then $e, w \not\models IF(n) \land [n]\alpha$ for all $w \in e_p$. Thus $e, w \not\models [n]\alpha$ for all $w \in (e *_{\mathrm{s}} IF(n))_{p - \lfloor e \,|\, IF(n) \rfloor + 1}$. By Rule S7, $e \gg n, w \not\models \alpha$ for all $w \in (e \gg n)_{p - \lfloor e \,|\, IF(n) \rfloor + 1}$. Thus $p - \lfloor e \,|\, IF(n) \rfloor + 1 < \lfloor e \gg n \,|\, \alpha \rfloor$. Analogously $p \geq \lfloor e \,|\, IF(n) \land [n]\alpha \rfloor$ implies $p - \lfloor e \,|\, IF(n) \rfloor + 1 \geq \lfloor e \gg n \,|\, \alpha \rfloor$.

(ii) Suppose $\lfloor e \,|\, \neg IF(n) \rfloor \leq p < \lfloor e \,|\, [n]\alpha \rfloor$. Then $e, w \not\models [n]\alpha$ for all $w \in e_p$. So $e, w \not\models [n]\alpha$ for all $w \in (e *_{\mathrm{s}} IF(n))_{p^*}$ where $p^* = p + \lceil e \rceil - \lfloor e \,|\, IF(n) \rfloor + 1 - \lfloor e \,|\, \neg IF(n) \rfloor + 1$, because by the same argument as in (i), $e \gg n, w \not\models \alpha$ for all $w \in (e \gg n)_{p'}$ and $p' \leq \lceil e \rceil - \lfloor e \,|\, IF(n) \rfloor + 1$. By Rule S7, $e \gg n, w \not\models \alpha$ for all $w \in (e \gg n)_{p^*}$. Thus $p^* < \lfloor e \gg n \,|\, \alpha \rfloor$. Analogously $p \geq \lfloor e \,|\, [n]\alpha \rfloor$ implies $p^* \geq \lfloor e \gg n \,|\, \alpha \rfloor$. □

We are now ready to establish the relationship between beliefs after and before an action. Theorem 27 considers weak-revision actions, Theorem 28 is about strong-revision actions.

Theorem 27. *Let a be a weak-revision action variable. Then*

$$\begin{aligned} \models \Box[a]\mathbf{B}(\alpha \Rightarrow \beta) \equiv\; & \neg\mathbf{B}(IF(a) \Rightarrow \neg[a]\alpha) \land \mathbf{B}(IF(a) \land [a]\alpha \Rightarrow [a]\beta) \lor \\ & \mathbf{B}(IF(a) \Rightarrow \neg[a]\alpha) \land \mathbf{B}([a]\alpha \Rightarrow [a]\beta) \lor \\ & \mathbf{B}(IF(a) \Rightarrow \textsc{false}). \end{aligned}$$

Intuitively the disjunction on the right-hand side considers three different cases. Action a triggers a revision, which promotes certain worlds to the first plausibility level. In the first case, at least one of these worlds satisfies α after a, and therefore we need to consider information learned by a in the antecedent. In the second case, none of them satisfies α after a, and therefore the revision is not relevant to the belief. The third case deals with revision by inconsistent information. The formal proof follows that intuition.

Proof. We prove that the equivalence holds in any epistemic state e for any weak-revision action n substituted for a. We distinguish three cases. The first case supposes $e \not\models \mathbf{B}(IF(n) \Rightarrow \neg[n]\alpha)$. The second one supposes the opposite plus $\lfloor e \,|\, IF(n) \rfloor \neq \infty$. The third case supposes $\lfloor e \,|\, IF(n) \rfloor = \infty$. For each case we show the equivalence. Since the cases are exhaustive, the theorem follows.

First suppose $e \not\models \mathbf{B}(IF(n) \Rightarrow \neg[a]\alpha)$. Then also $e \not\models \mathbf{B}(IF(n) \Rightarrow \text{FALSE})$. Hence the equivalence to be shown reduces to $e \models [n]\mathbf{B}(\alpha \Rightarrow \beta) \equiv \mathbf{B}(IF(n) \wedge [n]\alpha \Rightarrow [n]\beta)$. Notice that by assumption $\lfloor e \mid IF(n) \rfloor = \lfloor e \mid IF(n) \wedge [n]\alpha \rfloor \neq \infty$ (*), and by Lemma 25 $\lfloor e \gg n \mid \alpha \rfloor \neq \infty$ (**). Now we prove the equivalence: $e \models \mathbf{B}(IF(n) \wedge [n]\alpha \Rightarrow [n]\beta)$ iff (by Theorem 13 and (*)) $e, w \models IF(n) \wedge [n]\alpha \supset [n]\beta$ for all $w \in e_{\lfloor e \mid IF(n) \rfloor}$ iff $e, w \models [n]\alpha \supset [n]\beta$ for all $w \in (e *_{\mathrm{w}} IF(n))_1$ iff (by (**)) $e \gg n, w \models \alpha \supset \beta$ for all $w \in (e \gg n)_{\lfloor e \gg n \mid \alpha \rfloor}$ iff (by Theorem 13 and (**)) $e \models [n]\mathbf{B}(\alpha \Rightarrow \beta)$.

Now suppose $\lfloor e \mid IF(n) \rfloor \neq \infty$ and $e \models \mathbf{B}(IF(n) \Rightarrow \neg[n]\alpha)$. Then $e \not\models \mathbf{B}(IF(n) \Rightarrow \text{FALSE})$. Similar to the previous case, the remaining equivalence is $e \models [n]\mathbf{B}(\alpha \Rightarrow \beta) \equiv \mathbf{B}([n]\alpha \Rightarrow [n]\beta)$. Notice that by assumption, $e, w \models IF(n) \supset \neg[n]\alpha$ for all $w \in e_{\lfloor e \mid IF(n) \rfloor}$, so $e, w \not\models [n]\alpha$ for all $w \in (e *_{\mathrm{w}} IF(n))_1$ (*). Now we prove the equivalence: $e \models \mathbf{B}([n]\alpha \Rightarrow [n]\beta)$ iff $e, w \models [n]\alpha \supset [n]\beta$ for all $w \in e_p$ for all $p \in \mathbb{N}$ with $p \leq \lfloor e \mid [n]\alpha \rfloor$ iff (by (*)) $e, w \models [n]\alpha \supset [n]\beta$ for all $w \in (e *_{\mathrm{w}} IF(n))_p$ for all $p \in \mathbb{N}$ with $p \leq \lfloor e \mid [n]\alpha \rfloor + 1$ iff (by Lemma 25) $e \gg n, w \models \alpha \supset \beta$ for all $w \in (e \gg n)_p$ for all $p \in \mathbb{N}$ with $p \leq \lfloor e \gg n \mid \alpha \rfloor$ iff $e \models [n]\mathbf{B}(\alpha \Rightarrow \beta)$.

Finally suppose $\lfloor e \mid IF(n) \rfloor = \infty$. Then $e, w \not\models IF(n)$ for all $p \in \mathbb{N}$ and $w \in e_p$, and so $e \models \mathbf{B}(IF(n) \Rightarrow \text{FALSE})$. Since $\lfloor e \mid IF(n) \rfloor = \infty$, $(e \gg n)_p = \{\}$ for all $p \in \mathbb{N}$, and so $e \gg n, w \models \alpha \supset \beta$ for all $w \in (e \gg n)_p$. Thus $e \models [n]\mathbf{B}(\alpha \Rightarrow \beta)$. □

Theorem 28. *Let a be a strong-revision action variable. Then*

$$\begin{aligned} \models \Box[a]\mathbf{B}(\alpha \Rightarrow \beta) \equiv\ & \neg\mathbf{B}(IF(a) \wedge [a]\alpha \Rightarrow \text{FALSE}) \wedge \mathbf{B}(IF(a) \wedge [a]\alpha \Rightarrow [a]\beta) \vee \\ & \mathbf{B}(IF(a) \wedge [a]\alpha \Rightarrow \text{FALSE}) \wedge \mathbf{B}([a]\alpha \Rightarrow [a]\beta) \vee \\ & \mathbf{B}(IF(a) \Rightarrow \text{FALSE}). \end{aligned}$$

The three cases on the right-hand side are similar to the ones for weak revision in Theorem 27. The strong revision caused by a promotes all $IF(a)$-worlds over all $\neg IF(a)$-worlds. In case some of the former worlds satisfy α after a, some of them make up the most-plausible α-worlds after a, so the belief must also be conditioned on $IF(a)$. This is covered by the first case. Otherwise, if none of the promoted worlds satisfies α after a, the revision is irrelevant for that particular conditional belief. The third case deals with revision by inconsistent information. The formal argument follows this intuition and proceeds generally similar to the one of Theorem 27.

Proof. We prove that the equivalence holds in any epistemic state e for any strong-revision action n substituted for a. We distinguish three cases. The first case supposes $e \not\models \mathbf{B}(IF(n) \wedge \neg[n]\alpha \Rightarrow \text{FALSE})$. The second one supposes the opposite plus $\lfloor e \mid IF(n) \rfloor \neq \infty$. The third case supposes $\lfloor e \mid IF(n) \rfloor = \infty$. For each case we show the equivalence. Since the cases are exhaustive, the theorem follows.

First suppose $e \not\models \mathbf{B}(IF(n) \wedge [n]\alpha \Rightarrow \text{FALSE})$. Then also $e \not\models \mathbf{B}(IF(n) \Rightarrow \text{FALSE})$. Hence the equivalence to be proved reduces to $e \models [n]\mathbf{B}(\alpha \Rightarrow \beta) \equiv \mathbf{B}(IF(n) \wedge [n]\alpha \Rightarrow [n]\beta)$. Notice that by assumption $\lfloor e \mid IF(n) \wedge [n]\alpha \rfloor \neq \infty$ (*), and by Lemma 25 $\lfloor e \gg n \mid \alpha \rfloor \neq \infty$ (**). Now we can prove the equivalence: $e \models \mathbf{B}(IF(n) \wedge [n]\alpha \Rightarrow [n]\beta)$ iff (by Theorem 13 and (*)) $e, w \models IF(n) \wedge [n]\alpha \supset [n]\beta$ for all $w \in e_{\lfloor e \mid IF(n) \wedge [n]\alpha \rfloor}$ iff $e, w \models [n]\alpha \supset [n]\beta$ for all $w \in (e *_{\mathrm{s}} IF(n))_{\lfloor e \mid IF(n) \wedge [n]\alpha \rfloor - \lfloor e \mid IF(n) \rfloor + 1}$ iff (by Lemma 26) $e \gg n, w \models \alpha \supset \beta$ for all $w \in (e \gg n)_{\lfloor e \gg n \mid \alpha \rfloor}$ iff (by Theorem 13 and (**)) $e \models [n]\mathbf{B}(\alpha \Rightarrow \beta)$.

Now suppose $e \models \mathbf{B}(IF(n) \wedge [n]\alpha \Rightarrow \text{FALSE})$ and $\lfloor e \mid IF(n) \rfloor \neq \infty$. Then $e \not\models \mathbf{B}(IF(n) \Rightarrow \text{FALSE})$. Hence the equivalence left to be shown is $e \models [n]\mathbf{B}(\alpha \Rightarrow \beta) \equiv \mathbf{B}([n]\alpha \Rightarrow [n]\beta)$. Notice that by assumption, $e, w \not\models IF(n) \wedge [n]\alpha$ for all $w \in e_p$ and $p \in \mathbb{N}$ (*). Thus also $e \gg n, w \not\models \alpha$ for all $w \in (e \gg n)_p$ for all $p \in \mathbb{N}$ with $p \leq \lceil e \rceil - \lfloor e \mid IF(n) \rfloor + 1$ (**). Now we prove the equivalence. If $\lfloor e \mid \neg IF(n) \rfloor = \infty$, then there are no $\neg IF(n)$-worlds in e, so $e \models \mathbf{B}([n]\alpha \Rightarrow [n]\beta)$ holds by (*) and $e \models [n]\mathbf{B}(\alpha \Rightarrow \beta)$ holds by (**). Otherwise the equivalence is shown as follows: $e \models \mathbf{B}([n]\alpha \Rightarrow [n]\beta)$ iff $e, w \models [n]\alpha \supset [n]\beta$ for all $w \in e_p$ for all $p \in \mathbb{N}$ with $p \leq \lfloor e \mid [n]\alpha \rfloor$ iff (by (**)) $e, w \models [n]\alpha \supset [n]\beta$ for all $w \in (e *_{\mathrm{s}} IF(n))_p$ for all $p \in \mathbb{N}$ with $p \leq \lfloor e \mid [n]\alpha \rfloor + \lceil e \rceil - \lfloor e \mid IF(n) \rfloor - \lfloor e \mid \neg IF(n) \rfloor + 2$ iff (by Lemma 26) $e \gg n, w \models \alpha \supset \beta$ for all $w \in (e \gg n)_p$ for all $p \in \mathbb{N}$ with $p \leq \lfloor e \gg n \mid \alpha \rfloor$ iff $e \models [n]\mathbf{B}(\alpha \Rightarrow \beta)$.

Finally suppose $\lfloor e \mid IF(n) \rfloor = \infty$. Then $e, w \not\models IF(n)$ for all $p \in \mathbb{N}$ and $w \in e_p$, and so $e \models \mathbf{B}(IF(n) \Rightarrow \text{FALSE})$. Since $\lfloor e \mid IF(n) \rfloor = \infty$, $(e \gg n)_p = \{\}$ for all $p \in \mathbb{N}$, and so $e \gg n, w \models \alpha \supset \beta$ for all $w \in (e \gg n)_p$. Thus $e \models [n]\mathbf{B}(\alpha \Rightarrow \beta)$. □

Theorems 27 and 28 resemble successor-state axioms in that the action a occurs outside of the belief at the left-hand side, but not at the right-hand side of the equivalence. We can use them in a similar fashion to Rule R2 to push inside the scope of $\mathbf{B}$. Once that is done, regression proceeds with the antecedent and consequent in $\mathbf{B}$.

Definition 29. The regression of a regressable formula α is defined as in Definition 23 plus the following rule:

R9. $\mathcal{R}[r, \mathbf{B}(\alpha \Rightarrow \beta)]$ is defined inductively on r:

- $\mathcal{R}[\langle\rangle, \mathbf{B}(\alpha \Rightarrow \beta)] = \mathbf{B}(\mathcal{R}[\alpha] \Rightarrow \mathcal{R}[\beta])$;
- $\mathcal{R}[r \cdot t, \mathbf{B}(\alpha \Rightarrow \beta)] = \mathcal{R}[r, \sigma_t^a]$ where σ is the right-hand side of Theorem 27 or 28 depending on t's sort.

The following theorem states the correctness of regression:

Theorem 30. *Let $\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}$ be a basic action theory and let α be a regressable sentence. Then*

$$\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models \alpha \quad \textit{iff} \quad \mathbf{O}\Sigma_{\text{bel}} \models \mathcal{R}[\alpha].$$

The proof can be found in Appendix A.

5.3. An example

Let us illustrate regression using the gift-giving example. More precisely, we use the fourth example query from Section 4.2:

$$\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models [\mathit{dropbox}][\mathit{clink}][\mathit{unbox}(\#5)]\exists y\, \mathbf{B}(\mathit{gift} = y \wedge \neg \mathit{InBox}(\mathit{gift}) \wedge \neg \mathit{Broken}(\mathit{gift})).$$

We first regress $[\mathit{unbox}(\#5)]\exists y\, \mathbf{B}(\mathit{gift} = y \wedge \neg \mathit{InBox}(\mathit{gift}) \wedge \neg \mathit{Broken}(\mathit{gift}))$, and then show that the regressed sentence is satisfied by $e \gg \mathit{dropbox} \gg \mathit{clink}$, the progression of the model of $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$, which we determined in Section 4.2. We do not regress by *clink* and *dropbox* here for space reasons; we will handle them in the next section by progression. After rewriting the formula to adhere to the normal form required for regression, the task is to determine

$$\mathcal{R}[[\mathit{unbox}(\#5)]\exists y\, \mathbf{B}\exists y' (\mathit{gift} = y' \wedge \mathit{gift} = y \wedge \neg \mathit{InBox}(y') \wedge \neg \mathit{Broken}(y'))].$$

Regression then moves inside the existential and the action $\mathit{unbox}(\#5)$ and we obtain

$$\exists y \mathcal{R}[\mathit{unbox}(\#5), \mathbf{B}\exists y' (\mathit{gift} = y' \wedge \mathit{gift} = y \wedge \neg \mathit{InBox}(y') \wedge \neg \mathit{Broken}(y'))].$$

The action $\mathit{unbox}(\#5)$ is then pushed inside of the belief modalities and we obtain, after minor simplifications,

$$\begin{aligned}
&\exists y (\mathcal{R}[\neg \mathbf{B}(\mathit{IF}(\mathit{unbox}(\#5)) \Rightarrow \text{FALSE}) \wedge \mathbf{B}(\mathit{IF}(\mathit{unbox}(\#5)) \Rightarrow \psi)] \vee \\
&\quad \mathcal{R}[\mathbf{B}(\mathit{IF}(\mathit{unbox}(\#5)) \Rightarrow \text{FALSE}) \wedge \mathbf{B}\psi)] \vee \\
&\quad \mathcal{R}[\mathbf{B}(\mathit{IF}(\mathit{unbox}(\#5)) \Rightarrow \text{FALSE})]) \\
&\text{where } \psi = [\mathit{unbox}(\#5)]\exists y' (\mathit{gift} = y' \wedge \mathit{gift} = y \wedge \neg \mathit{InBox}(y') \wedge \neg \mathit{Broken}(y')).
\end{aligned}$$

Now regression proceeds inside the belief modalities with the antecedents and consequents. In particular, regressing ψ substitutes $\mathit{InBox}(y')$ and $\mathit{Broken}(y')$ with the right-hand sides of the successor-state axioms:

$$\begin{aligned}
\mathcal{R}[\psi] = \exists y' (&\mathit{gift} = y' \wedge \mathit{gift} = y \wedge \\
&\neg(\mathit{InBox}(y') \wedge \mathit{unbox}(\#5) \neq \mathit{unbox}(y')) \wedge \\
&\neg(\mathit{Broken}(y') \vee \mathit{InBox}(y') \wedge \mathit{Fragile}(y') \wedge \mathit{unbox}(\#5) = \mathit{dropbox})).
\end{aligned}$$

After some trivial simplifications, the final regressed formula is equivalent to

$$\begin{aligned}
&(\neg \mathbf{B}(\mathit{InBox}(\#5) \wedge \neg \mathit{Broken}(\#5) \Rightarrow \text{FALSE}) \wedge \exists y \mathbf{B}(\mathit{InBox}(\#5) \wedge \neg \mathit{Broken}(\#5) \Rightarrow \mathit{gift} = \#5 \wedge \mathit{gift} = y \wedge \neg \mathit{Broken}(\#5))) \vee \\
&\mathbf{B}(\mathit{InBox}(\#5) \wedge \neg \mathit{Broken}(\#5) \Rightarrow \text{FALSE}).
\end{aligned}$$

Finally we need to prove that $e \gg \mathit{dropbox} \gg \mathit{clink}$ satisfies this formula. Note that there are $w \in (e \gg \mathit{dropbox} \gg \mathit{clink})_3$ with $w \models \mathit{InBox}(\#5) \wedge \neg \mathit{Broken}(\#5)$. Consider any such w. Since such worlds do exist, $\neg \mathbf{B}(\mathit{InBox}(\#5) \wedge \neg \mathit{Broken}(\#5) \Rightarrow \text{FALSE})$ is true, and therefore we need to prove that $\exists y \mathbf{B}(\mathit{InBox}(\#5) \wedge \neg \mathit{Broken}(\#5) \Rightarrow \mathit{gift} = \#5 \wedge \mathit{gift} = y \wedge \neg \mathit{Broken}(\#5))$ is true as well. We substitute $\#5$ for the existentially quantified y. Since $w \models \mathit{InBox}(\#5)$ by assumption and nothing but *gift* is in the box at level $(e \gg \mathit{dropbox} \gg \mathit{clink})_3$, w also satisfies the consequent, namely $w \models \mathit{gift} = \#5 \wedge (\mathit{gift} = y)^y_{\#5} \wedge \neg \mathit{Broken}(\#5)$. Thus $e \gg \mathit{dropbox} \gg \mathit{clink}$ satisfies the regressed formula. We will finish the proof of the query in the next section where we also handle *dropbox* and *clink*.

6. Belief projection by progression

Another way to solve the projection problem is by progression. Progression means to update the knowledge base to account for the effects of actions. It is commonly considered necessary, particularly in long-lived systems, so that the system can move on and discard historic states. Unfortunately, progression in general requires second-order logic due to results by Lin and Reiter [6] and Vassos and Levesque [7]. Lin and Reiter showed that progression is closely related to logical forgetting: one needs to forget the obsolete values of the predicates. The syntactic way to forget a predicate in a knowledge base is to substitute it with an existentially quantified second-order variable.

In this section, we show how beliefs can be progressed. Given a basic action theory and an action n, we aim to project the epistemic and the physical effects of n onto the initial beliefs. The result of this operation shall be a new, progressed basic action theory. At first, we extend the only-believing operator by a mechanism to forget predicates, and show that all former results carry over to that extended operator. Said extension is not full second-order logic but allows for existential quantification of predicates within only-believing for reasons we elaborate below. Then we proceed by first investigating how revision of a conditional knowledge base can be captured by only-believing in Section 6.2. This allows us to account for the *epistemic* effects of actions, and in Section 6.3 we then add the *physical* effects and obtain the progression theorem. Section 6.4 continues our running example.

6.1. Only-believing with forgetting

Following Lin and Reiter [6], we use logical forgetting to represent the effects of actions on the knowledge base. As mentioned above, forgetting a predicate in a knowledge base boils down to making that predicate an existentially quantified second-order variable. For the purposes of forgetting in $\mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$, what should the scope of that second-order variable be? It clearly should be quantified within $\mathbf{O}$, yet its scope should encompass all ϕ_i and ψ_i, so that all occurrences in ϕ_i, ψ_i refer to the same variable. Adding full support of second-order quantifiers and allowing them to appear between the $\mathbf{O}$ and its arguments $\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ requires a quite cumbersome semantics, though. On the other hand, full second-order logic is not even required for forgetting; existentials between $\mathbf{O}$ and its arguments suffice. As it permits a much simpler semantics, we parameterize the only-believing operator with a set of function and predicate symbols, which are taken to be existentially quantified inside $\mathbf{O}$.

Definition 31. The set of well-formed formulas is the least set formed from the rules from Definition 3 and

- $\mathbf{O}_{\mathcal{S}}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ is a formula where the ϕ_i and ψ_i are objective formulas and $\mathcal{S}$ is a finite set of object function and predicate symbols.

We say α is *$\mathcal{S}$-free* when it mentions no object function or predicate symbol from $\mathcal{S}$.

Due to the relationship between existential quantification and forgetting, we read $\mathbf{O}_{\mathcal{S}}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ as "before everything about $\mathcal{S}$ is forgotten, the conditionals $\phi_i \Rightarrow \psi_i$ are all that is believed."

To characterize the semantics of existential quantification, we use the following relation to say that two worlds agree on everything except perhaps certain symbols.

Definition 32. For a set of object function and predicate symbols $\mathcal{S}$, we define $w \approx_{\mathcal{S}} w'$ iff

- $w[g(n_1, \ldots, n_k)] = w'[g(n_1, \ldots, n_k)]$ for all object function symbols $g \notin \mathcal{S}$;
- $w[R(n_1, \ldots, n_k)] = w'[R(n_1, \ldots, n_k)]$ for all rigid predicate symbols $R \notin \mathcal{S}$;
- $w[F(n_1, \ldots, n_k), z] = w'[F(n_1, \ldots, n_k), z]$ for all fluent predicate symbols $F \notin \mathcal{S}$ and action sequences z.

For a set of worlds W and an epistemic state e, we let $W_{\mathcal{S}} = \{w' \mid w \approx_{\mathcal{S}} w' \text{ for some } w \in W\}$ and $e_{\mathcal{S}} = \langle (e_1)_{\mathcal{S}}, \ldots, (e_{\lceil e \rceil})_{\mathcal{S}} \rangle$.

Intuitively, $w \approx_{\mathcal{S}} w'$ means that w and w' agree on everything except perhaps $\mathcal{S}$. Notice that $w \approx_{\{\}} w'$ iff $w = w'$. Hence, $e_{\mathcal{S}}$ intuitively denotes the epistemic state that results from forgetting everything about $\mathcal{S}$ in the original epistemic state e. For example, consider an epistemic state e with $e_1 = \{w \mid w \models R \land (R \equiv R')\}$. Then $w[R] = w[R'] = 1$ for all $w \in e_1$. Belief in R is then lost in $e_{\{R\}}$, while R' is retained: for every $w \in e_1$, not only $w \in (e_{\{R\}})_1$, but there also is a $w' \in (e_{\{R\}})_1$ which agrees with w on everything except that $w'[R] = 0$, so we have $(e_{\{R\}})_1 = \{w \mid w \models R'\}$.

Definition 33. The semantics of the new only-believing operator is defined using standard only-believing:

S11. $e, w \models \mathbf{O}_{\mathcal{S}}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ iff for some e', $e', w \models \mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ and $e = e'_{\mathcal{S}}$.

Note that extended only-believing subsumes standard only-believing, as their semantics coincide for $\mathcal{S} = \{\}$. It is not surprising that earlier results such as the unique-model property from Corollary 20 and the regression result Theorem 30 carry over to the extended operator.

Corollary 34. *Let $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ be objective. Then there is one unique e with $e \models \mathbf{O}_{\mathcal{S}}\Gamma$.*

Proof. By Corollary 20, there is a unique e' such that $e' \models \mathbf{O}\Gamma$. By Rule S11, $e'_{\mathcal{S}}$ is unique such that $e'_{\mathcal{S}} \models \mathbf{O}_{\mathcal{S}}\Gamma$. □

Theorem 35. *Let $\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}$ be a basic action theory with $\mathcal{S}$-free Σ_{dyn} and let α be a regressable sentence. Then*

$$\mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models \alpha \quad \textit{iff} \quad \mathbf{O}_{\mathcal{S}}\Sigma_{\text{bel}} \models \mathcal{R}[\alpha].$$

For the proof see Appendix A.

6.2. Revision of only-believing

Semantically, performing an action brings along a revision of the epistemic state, which promotes certain worlds over others. In this subsection we examine how the semantic revision can be matched syntactically. More precisely, we are looking for a set of conditionals $\Gamma * \upsilon$ which is only-believed when Γ was only-believed before revising by υ.

Recall that by Theorem 15, $\mathbf{B}(\alpha \vee \beta \Rightarrow \neg\beta)$ asserts that α is more plausible than β or both are considered impossible. We use this to define Γ_α as the set of conditionals whose material implication holds up to the plausibility level of α.

Definition 36. For $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$, we define $\Gamma_\alpha = \{\phi \Rightarrow \psi \in \Gamma \mid \mathbf{O}\Gamma \models \mathbf{B}(\alpha \vee \neg(\phi \supset \psi) \Rightarrow (\phi \supset \psi))\}$.

A syntactic representation of weak revision by υ needs to reflect that the most-plausible υ-worlds are promoted to the first level. To this end we use a new dummy predicate R to partition the worlds: the R-worlds represent those which are promoted, and the $\neg R$-worlds represent the beliefs before the revision.

Definition 37. Let R be a rigid predicate symbol. Let $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ and υ be objective and $\{R\}$-free. Then the *weak revision* of Γ by υ is

$$\begin{aligned}
\Gamma *_{\text{w}} \upsilon = \{&\textsc{true} && \Rightarrow R\} && \cup \\
\{&\neg(R \supset \upsilon) && \Rightarrow \textsc{false}\} && \cup \\
\{&\neg(R \supset (\phi \supset \psi)) && \Rightarrow \textsc{false} \mid \phi \Rightarrow \psi \in \Gamma_\upsilon\} && \cup \\
\{&(\neg R \wedge \phi) && \Rightarrow \psi \quad\ \mid \phi \Rightarrow \psi \in \Gamma\}.
\end{aligned}$$

We now prove that the revised set of conditionals (after forgetting R) matches semantic weak revision. To begin with, we prove a few technical lemmas that will also be helpful for the strong revision theorem.

Lemma 38. *If $\lfloor e \mid \alpha \rfloor > \lceil e \rceil$, then $\lfloor e \mid \alpha \rfloor = \infty$.*

Proof. Suppose $\lfloor e \mid \alpha \rfloor \neq \infty$. Then $e, w \models \alpha$ for some $w \in e_p$ and $p \in \mathbb{N}$. Since $e_p \subseteq e_{\lceil e \rceil}$, $\lfloor e \mid \alpha \rfloor \leq \lceil e \rceil$. □

Lemma 39. *Let $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ be objective, let e be such that $e \models \mathbf{O}\Gamma$, and let $p \in \mathbb{N} \cup \{\infty\}$ such that $p \geq \lfloor e \mid \alpha \rfloor$. Then $\models \bigwedge_{i:\lfloor e \mid \phi_i \rfloor \geq p}(\phi_i \supset \psi_i) \equiv \bigwedge_{\phi \Rightarrow \psi \in \Gamma_\alpha \text{ with } \max\{\lfloor e \mid \phi \rfloor, \lfloor e \mid \alpha \rfloor\} \geq p}(\phi \supset \psi)$.*

Proof. By Corollary 20 and Theorem 15, $\phi \Rightarrow \psi \in \Gamma_\alpha$ iff $\lfloor e \mid \alpha \rfloor = \lfloor e \mid \neg(\phi \supset \psi) \rfloor = \infty$ or $\lfloor e \mid \alpha \rfloor < \lfloor e \mid \neg(\phi \supset \psi) \rfloor$ (*).

For the *only-if* direction, suppose $w \models \bigwedge_{i:\lfloor e \mid \phi_i \rfloor \geq p}(\phi_i \supset \psi_i)$ and let $\phi \Rightarrow \psi \in \Gamma_\alpha$ with $\max\{\lfloor e \mid \phi \rfloor, \lfloor e \mid \alpha \rfloor\} \geq p$. Then $w \in e_{\min\{p, \lceil e \rceil\}}$ by Rule S10. Note that $w' \models (\phi \supset \psi)$ for all $w' \in e_{p'}$ and $p' \in \mathbb{N}$ with $p' \leq \lfloor e \mid \phi \rfloor$ by Rule S10. Likewise, $w' \models (\phi \supset \psi)$ for all $w' \in e_{p'}$ and $p' \in \mathbb{N}$ with $p' \leq \lfloor e \mid \alpha \rfloor$ by (*). Hence, since $p \leq \max\{\lfloor e \mid \phi \rfloor, \lfloor e \mid \alpha \rfloor\}$, $w \models (\phi \supset \psi)$.

Conversely, suppose $w \not\models (\phi_i \supset \psi_i)$ for some i with $\lfloor e \mid \phi_i \rfloor \geq p$. Then trivially $\max\{\lfloor e \mid \phi_i \rfloor, \lfloor e \mid \alpha \rfloor\} \geq p$, so we only need to show that $\phi_i \Rightarrow \psi_i \in \Gamma_\alpha$. By Rule S10, $w' \models (\phi_i \supset \psi_i)$ for all $w' \in e_{p'}$ and $p' \in \mathbb{N}$ with $p' \leq p$. Hence $p < \lfloor e \mid \neg(\phi_i \supset \psi_i) \rfloor$. If $\lfloor e \mid \alpha \rfloor \leq \lceil e \rceil$, then $\lfloor e \mid \alpha \rfloor \leq p < \lfloor e \mid \neg(\phi_i \supset \psi_i) \rfloor$. Otherwise, if $\lceil e \rceil < \lfloor e \mid \alpha \rfloor$, then $\lceil e \rceil \leq p < \lfloor e \mid \neg(\phi_i \supset \psi_i) \rfloor$, and by Lemma 38, $\lfloor e \mid \alpha \rfloor = \lfloor e \mid \neg(\phi_i \supset \psi_i) \rfloor = \infty$. In both cases, by (*), $\phi_i \Rightarrow \psi_i \in \Gamma_\alpha$. □

Lemma 40. *Let ϕ be objective and $\mathcal{S}$-free, and $w \approx_{\mathcal{S}} w'$. Then $w \models \phi$ iff $w' \models \phi$.*

Proof. Follows by a trivial induction on the length of ϕ.

Lemma 41. *Let ϕ be objective and $\mathcal{S}$-free. Then $(e * \phi)_{\mathcal{S}} = e_{\mathcal{S}} * \phi$.*

Proof. By Lemma 40, $\lfloor e \,|\, \phi \rfloor = \lfloor e_{\mathcal{S}} \,|\, \phi \rfloor$ and $(e_{\mathcal{S}} \,|\, \phi)_p = ((e \,|\, \phi)_p)_{\mathcal{S}}$. Thus by Definition 8 the lemma follows. ◻

Theorem 42. *Let $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ and υ be objective and $\mathcal{S}$-free. Let R be the nullary rigid predicate newly introduced in $\Gamma *_{\mathrm{w}} \upsilon$. If $e \models \mathbf{O}_{\mathcal{S}}\Gamma$, then*

$$e *_{\mathrm{w}} \upsilon \models \mathbf{O}_{\mathcal{S} \cup \{R\}} \Gamma *_{\mathrm{w}} \upsilon.$$

Proof. We show the theorem for the case $\mathcal{S} = \{\}$ first. Let $e \models \mathbf{O}\Gamma$. We construct an e' such that $e' \models \mathbf{O}\Gamma *_{\mathrm{w}} \upsilon$ and $e'_{\{R\}} = e *_{\mathrm{w}} \upsilon$, which gives $e *_{\mathrm{w}} \upsilon \models \mathbf{O}_{\{R\}}\Gamma *_{\mathrm{w}} \upsilon$. If $\lfloor e \,|\, \upsilon \rfloor = \infty$, we let $e'_p = \{\}$ for all $p \in \mathbb{N}$. Otherwise, we let $e'_1 = ((e \,|\, \upsilon) \,|\, R)_{\lfloor e \,|\, \upsilon \rfloor}$ and $e'_p = e'_1 \cup (e \,|\, \neg R)_{p-1}$ for $p > 1$.

First suppose $\lfloor e \,|\, \upsilon \rfloor = \infty$. Then $e *_{\mathrm{w}} \upsilon = \langle\{\}\rangle$. Clearly, $\lfloor \langle\{\}\rangle \,|\, \phi \rfloor = \infty$ for all ϕ, so $\langle\{\}\rangle \models \mathbf{O}\Gamma *_{\mathrm{w}} \upsilon$ if (by Rule S10) $\bigwedge_{\phi \Rightarrow \psi \in \Gamma_\upsilon} (R \supset (\phi \supset \psi)) \wedge (R \supset \upsilon) \wedge (\textsc{true} \supset R)$ is unsatisfiable iff $\bigwedge_{\phi \Rightarrow \psi \in \Gamma_\upsilon} (\phi \supset \psi) \wedge \upsilon \wedge R$ is unsatisfiable if (by Lemma 39) $\bigwedge_{i : \lfloor e \,|\, \phi_i \rfloor = \infty} (\phi_i \supset \psi_i) \wedge \upsilon$ is unsatisfiable, which holds by Rule S10 and $\lfloor e \,|\, \upsilon \rfloor = \infty$.

Now suppose $\lfloor e \,|\, \upsilon \rfloor \neq \infty$. We show that $e' \models \mathbf{O}\Gamma *_{\mathrm{w}} \upsilon$ for the following plausibilities of the conditionals in $\Gamma *_{\mathrm{w}} \upsilon$:

- $\lfloor e' \,|\, \textsc{true} \rfloor = 1$; because by assumption $\lfloor e \,|\, \upsilon \rfloor \neq \infty$ and thus $e'_1 \neq \{\}$.
- $\lfloor e' \,|\, \neg(R \supset \upsilon) \rfloor = \infty$; because $w \models \upsilon$ for all $w \in e'_1$, and $w \not\models R$ for all $w \in e'_p \setminus e'_1$.
- $\lfloor e' \,|\, \neg(R \supset (\phi \supset \psi)) \rfloor = \infty$ for all $\phi \Rightarrow \psi \in \Gamma_\upsilon$; because by Lemma 39, $w \models (\phi \supset \psi)$ for all $w \in e_{\lfloor e \,|\, \upsilon \rfloor} \supseteq e'_1$, and $w \not\models R$ for all $w \in e'_p \setminus e'_1$.
- $\lfloor e' \,|\, (\neg R \wedge \phi) \rfloor = \lfloor e \,|\, \phi \rfloor + 1$ for all $\phi \Rightarrow \psi \in \Gamma$; because $w \models R$ for all $w \in e'_1$, and so for all $p \in \mathbb{N}$ we have $p + 1 \geq \lfloor e' \,|\, (\neg R \wedge \phi) \rfloor$ iff $w \models (\neg R \wedge \phi)$ for some $w \in e'_{p+1}$ iff $w \models (\neg R \wedge \phi)$ for some $w \in (e \,|\, \neg R)_p$ iff (since Γ, υ are $\{R\}$-free and by Rule S10 and Lemma 40) $w \models \phi$ for some $w \in e_p$ iff $p \geq \lfloor e \,|\, \phi \rfloor$.

Then $w \in e'_1$ iff $w \in ((e \,|\, \upsilon) \,|\, R)_{\lfloor e \,|\, \upsilon \rfloor}$ iff $w \models \bigwedge_{i : \lfloor e \,|\, \phi_i \rfloor \geq \lfloor e \,|\, \upsilon \rfloor} (\phi_i \supset \psi_i) \wedge \upsilon \wedge R$ iff (by Lemma 39) $w \models \bigwedge_{\phi \Rightarrow \psi \in \Gamma_\upsilon} (\phi \supset \psi) \wedge \upsilon \wedge R$ iff $w \models \bigwedge_{\phi \Rightarrow \psi \in \Gamma_\upsilon} (R \supset (\phi \supset \psi)) \wedge (R \supset \upsilon) \wedge (\textsc{true} \supset R) \wedge \bigwedge_{i : \lfloor e \,|\, \phi_i \rfloor + 1 \geq 1} ((\neg R \wedge \phi_i) \supset \psi_i)$. For $p > 1$, $w \in e'_p$ iff $w \in ((e \,|\, \upsilon) \,|\, R)_{\lfloor e \,|\, \upsilon \rfloor}$ or $w \in (e \,|\, \neg R)_{p-1}$ iff $w \models \bigwedge_{i : \lfloor e \,|\, \phi_i \rfloor \geq \lfloor e \,|\, \upsilon \rfloor} (\phi_i \supset \psi_i) \wedge \upsilon \wedge R$ or $w \models \bigwedge_{i : \lfloor e \,|\, \phi_i \rfloor \geq p-1} (\phi_i \supset \psi_i) \wedge \neg R$ iff (by Lemma 39) $w \models \bigwedge_{\phi \Rightarrow \psi \in \Gamma_\upsilon} (\phi \supset \psi) \wedge \upsilon \wedge R$ or $w \models \bigwedge_{i : \lfloor e \,|\, \phi_i \rfloor + 1 \geq p} (\phi_i \supset \psi_i) \wedge \neg R$ iff $w \models \bigwedge_{\phi \Rightarrow \psi \in \Gamma_\upsilon} (R \supset (\phi \supset \psi)) \wedge (R \supset \upsilon) \wedge \bigwedge_{i : \lfloor e \,|\, \phi_i \rfloor + 1 \geq p} ((\neg R \wedge \phi_i) \supset \psi_i)$. Thus the right-hand side of Rule S10 holds, and hence $e' \models \mathbf{O}\Gamma *_{\mathrm{w}} \upsilon$.

Since Γ and υ are $\{R\}$-free, for each $w \in (e \,|\, \upsilon)_{\lfloor e \,|\, \upsilon \rfloor}$ there is a $w' \in e'_1$ with $w \approx_{\{R\}} w'$ by Lemma 40 and Rule S10. Likewise, for each $w \in e_p$ there is a $w' \in (e \,|\, \neg R)_p$ with $w \approx_{\{R\}} w'$. Thus $e *_{\mathrm{w}} \upsilon = e'_{\{R\}}$, and hence $e *_{\mathrm{w}} \upsilon \models \mathbf{O}_{\{R\}}\Gamma *_{\mathrm{w}} \upsilon$.

Now let $\mathcal{S} \neq \{\}$. Let $e \models \mathbf{O}_{\mathcal{S}}\Gamma$ and $e' \models \mathbf{O}\Gamma$. By Rule S11 and Corollary 34, $e = e'_{\mathcal{S}}$. By the case for $\mathcal{S} = \{\}$, $e' *_{\mathrm{w}} \upsilon \models \mathbf{O}_{\{R\}}\Gamma *_{\mathrm{w}} \upsilon$. By Rule S11, $(e' *_{\mathrm{w}} \upsilon)_{\mathcal{S}} \models \mathbf{O}_{\mathcal{S} \cup \{R\}}\Gamma *_{\mathrm{w}} \upsilon$. By Lemma 41, $e *_{\mathrm{w}} \upsilon \models \mathbf{O}_{\mathcal{S} \cup \{R\}}\Gamma *_{\mathrm{w}} \upsilon$. ◻

Strong revision changes the ranking of the worlds more profoundly than weak revision, and representing this change is hence more complex. Strong revision by υ promotes all υ-worlds over all $\neg\upsilon$-worlds. We therefore duplicate the conditionals from Γ twice using new predicates, and require υ to be true in the first copy. The revised truth values are then set through additional conditionals based on the dummies' truth values.

To ease the presentation, we restrict our consideration of strong revision to static formulas. They are sufficient for our purposes of progressing a basic action theories, since the initial beliefs Σ_{bel} are static as well. Extending the below definition and theorem for non-static conditionals is straightforward.

Definition 43. Let $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ and υ be objective and static. Let $\mathcal{S}'$ be the object function and predicate symbols in Γ, and let $\mathcal{S}''$ be just as many object function and rigid predicate symbols of corresponding arity which do not occur in Γ or υ. For any formula β, let β^* be the formula obtained from β by replacing each symbol from

$\mathcal{S}'$ with the corresponding symbol from $\mathcal{S}''$. Let $\Delta = \{\phi \Rightarrow \psi \in \Gamma_\upsilon \mid \mathbf{O}\Gamma \not\models \mathbf{K}(\phi \supset \psi)\}$. Then the *strong revision* of Γ by υ is defined as

$$\begin{aligned}
\Gamma *_{\mathrm{s}} \upsilon = \{&\phi^* \Rightarrow \psi^* \mid \phi \Rightarrow \psi \in \Gamma_\upsilon\} \cup \\
&\{(\phi^* \wedge \neg\upsilon) \Rightarrow \psi^* \mid \phi \Rightarrow \psi \in \Gamma_{\neg\upsilon}\} \cup \\
&\{\text{TRUE} \Rightarrow \upsilon\} \cup \{\neg(\phi^* \supset \psi^*) \vee \neg\upsilon \Rightarrow \upsilon \mid \phi \Rightarrow \psi \in \Delta\} \cup \\
&\{\neg((\upsilon \wedge \neg\upsilon^*) \supset (\phi \supset \psi)) \Rightarrow \text{FALSE} \mid \phi \Rightarrow \psi \in \Gamma_\upsilon\} \cup \\
&\{\neg((\neg\upsilon \wedge \upsilon^*) \supset (\phi \supset \psi)) \Rightarrow \text{FALSE} \mid \phi \Rightarrow \psi \in \Gamma_{\neg\upsilon}\} \cup \\
&\{\neg((\upsilon \equiv \upsilon^*) \supset (\phi \equiv \phi^*) \wedge (\psi \equiv \psi^*)) \Rightarrow \text{FALSE} \mid \phi \Rightarrow \psi \in \Gamma\}.
\end{aligned}$$

The first and the third line account for the promoted υ-worlds in the revised epistemic state. In particular, the third line asserts that there is the right number of levels where υ holds. The material implications in the last three lines set the original predicates according to the values of the dummy predicates.

As with weak revision, the syntactic strong revision $\Gamma *_{\mathrm{s}} \upsilon$ matches its semantic counterpart.

Theorem 44. *Let $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ and υ be objective and static. Let $\mathcal{S}$ be a set of object function and predicate symbols, and let υ be $\mathcal{S}$-free. Let $\mathcal{S}''$ be the symbols newly introduced in $\Gamma *_{\mathrm{s}} \upsilon$. If $e \models \mathbf{O}_{\mathcal{S}}\Gamma$, then*

$$e *_{\mathrm{s}} \upsilon \models \mathbf{O}_{\mathcal{S}\cup\mathcal{S}''}\Gamma *_{\mathrm{s}} \upsilon.$$

The proof is similar to the proof of Theorem 42. It can be found in Appendix B.

6.3. Progression of only-believing

We are now ready to define the progression of a basic action theory $\Sigma_{\mathrm{dyn}}, \Sigma_{\mathrm{bel}}$. Given an action standard name n, we first revise the theory by n's information and then apply n's effects on fluents. The revision is captured by $\Sigma_{\mathrm{bel}} * \varphi^a_n$ where φ is from the informed-fluent axiom $\Box IF(a) \equiv \varphi \in \Sigma_{\mathrm{dyn}}$, and the type of revision corresponds to the subsort of n. (The reason for taking φ^a_n instead of $IF(n)$ is to keep the belief conditionals fluent.) In this subsection we show how the physical effects of n are handled.

For two sets of predicate symbols $\mathcal{F} = \{F_1, \ldots, F_k\}$ and $\mathcal{R} = \{R_1, \ldots, R_k\}$ of corresponding arity we denote by $\alpha^{\mathcal{F}}_{\mathcal{R}}$ the formula obtained by replacing each F_i with R_i in α. Recall that, in the context of a basic action theory, γ_F denotes the right-hand side of the successor-state axiom for a fluent F and φ is the right-hand side of the informed-fluent axiom.

Definition 45. Let $\Sigma_{\mathrm{dyn}}, \Sigma_{\mathrm{bel}}$ be a basic action theory over fluents $\mathcal{F} = \{F_1, \ldots, F_k\}$, and let $\mathcal{R} = \{R_1, \ldots, R_k\}$ be rigid predicates of corresponding arity which do not otherwise occur in Σ_{dyn} or $\Sigma_{\mathrm{bel}} * \varphi^a_n$. Let n be a action standard name. Then the *progression* of Σ_{bel} by n is defined as

$$\Sigma_{\mathrm{bel}} \gg n = (\Sigma_{\mathrm{bel}} * \varphi^a_n)^{\mathcal{F}}_{\mathcal{R}} \cup \{\neg\forall x_1 \ldots \forall x_l (F(x_1, \ldots, x_l) \equiv {\gamma_F}^{a\,\mathcal{F}}_{n\,\mathcal{R}}) \Rightarrow \text{FALSE} \mid F \in \mathcal{F}\}.$$

The intuition behind the definition is as follows. When n is executed, the beliefs are first revised by the information φ^a_n produced by n, which leads to $\Sigma_{\mathrm{bel}} * \varphi^a_n$. The beliefs $(\Sigma_{\mathrm{bel}} * \varphi^a_n)^{\mathcal{F}}_{\mathcal{R}}$ represent the same conditionals belief as $\Sigma_{\mathrm{bel}} * \varphi^a_n$, except that each F_i is renamed to R_i. Intuitively, R_i memorizes the value of F_i before the physical effect of n. The additional conditionals in $\Sigma_{\mathrm{bel}} \gg n$ initialize each fluent $F(x_1, \ldots, x_l)$ with its value after doing n, that is, ${\gamma_F}^{a\,\mathcal{F}}_{n\,\mathcal{R}}$. Notice that the progression of a basic action theory again is a basic action theory over $\mathcal{F}$, so progression can iterate.

The following two results establish the correctness of progression. The first theorem says that, if all that is believed is a basic action theory, then after doing action n all that is believed is the progressed basic action theory:

Theorem 46. *Let $\Sigma_{\mathrm{dyn}}, \Sigma_{\mathrm{bel}}$ be a basic action theory. Let $\mathcal{S}'$ be the symbols newly introduced by $\Sigma_{\mathrm{bel}} \gg n$. Then*

$$\models \mathbf{O}_{\mathcal{S}}(\Sigma_{\mathrm{dyn}}, \Sigma_{\mathrm{bel}}) \supset [n]\mathbf{O}_{\mathcal{S}\cup\mathcal{S}'}(\Sigma_{\mathrm{dyn}}, \Sigma_{\mathrm{bel}} \gg n).$$

The second theorem says that the same beliefs are entailed by a basic action theory after doing action n and the progression by n of that basic action theory:

Theorem 47. *Let $\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}$ be a basic action theory. Let $\mathcal{S}'$ be the symbols newly introduced by $\Sigma_{\text{bel}} \gg n$. Then*

$$\mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models [n]\alpha \quad \textit{iff} \quad \mathbf{O}_{\mathcal{S}\cup\mathcal{S}'}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n) \models \alpha.$$

The proofs of both theorems are in Appendix B. They proceed in two steps. First, we show that $\Sigma_{\text{bel}} * \varphi_n^a$ and $(\Sigma_{\text{bel}} * \varphi_n^a)_{\mathcal{R}}^{\mathcal{F}}$ determine the same conditional beliefs modulo the substitution of $\mathcal{F}$ by $\mathcal{R}$, where $\mathcal{R} \subseteq \mathcal{S}'$ is the set of rigid predicates from Definition 45. Second, we see that after progressing the individual worlds in the model of $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} * \varphi_n^a)$ by n, the resulting epistemic state agrees with the model of $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n)$ on everything except perhaps $\mathcal{R}$. We therefore have that, when using standard only-believing, the semantic progression and the syntactic progression agree on everything except $\mathcal{R}$. With the extended only-believing operator, the differences in $\mathcal{R}$ vanish as well.

6.4. An example

Let us proceed with the query from the gift-giving example which we investigated in Section 5.3 already. The query involves the action sequence *dropbox*, *clink*, *unbox*(#5). Since *dropbox* is a physical action with no epistemic effect, let us take an abbreviation instead of doing it by the definitions: it is easy to see that $e \gg \textit{dropbox}$ from Section 4.2 satisfies $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma'_{\text{bel}})$ where

$$\begin{aligned}
\Sigma'_{\text{bel}} = \{&\textsc{true} \Rightarrow \forall y \neg \textit{InBox}(y),\\
&\exists y\, \textit{InBox}(y) \Rightarrow \forall y (\textit{InBox}(y) \equiv y = \textit{gift}),\\
&\exists y\, \textit{InBox}(y) \Rightarrow \forall y (\textit{InBox}(y) \supset (\textit{Broken}(y) \equiv \textit{Fragile}(y))),\\
&\neg \forall y (\textit{InBox}(y) \wedge \textit{Fragile}(y) \supset \textit{Broken}(y)) \Rightarrow \textsc{false}\}.
\end{aligned}$$

We focus on the progression of $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma'_{\text{bel}})$ by the weak-revision action *clink*. According to Definition 37, the revision $\Sigma'_{\text{bel}} * \varphi^a_{\textit{clink}}$ contains the conditionals

- $\textsc{true} \Rightarrow R$;
- $\neg(R \supset \varphi^a_{\textit{clink}}) \Rightarrow \textsc{false}$;
- $\neg(R \supset (\phi \supset \psi)) \Rightarrow \textsc{false}$ for each $\phi \Rightarrow \psi \in \Sigma'_{\text{bel}}$ such that $\mathbf{O}\Sigma'_{\text{bel}} \models \mathbf{B}(\varphi^a_{\textit{clink}} \vee \neg(\phi \supset \psi) \Rightarrow (\phi \supset \psi))$;
- $\neg R \wedge \phi \Rightarrow \psi$ for each $\phi \Rightarrow \psi \in \Sigma'_{\text{bel}}$.

This amounts to

$$\begin{aligned}
\Sigma'_{\text{bel}} * \varphi^a_{\textit{clink}} = \{&\textsc{true} \Rightarrow R,\\
&\neg(R \supset \exists y (\textit{InBox}(y) \wedge \textit{Broken}(y))) \Rightarrow \textsc{false},\\
&\neg(R \supset (\exists y\, \textit{InBox}(y) \supset \forall y (\textit{InBox}(y) \equiv y = \textit{gift}))) \Rightarrow \textsc{false},\\
&\neg(R \supset (\exists y\, \textit{InBox}(y) \supset \forall y (\textit{InBox}(y) \supset (\textit{Broken}(y) \equiv \textit{Fragile}(y))))) \Rightarrow \textsc{false},\\
&\neg(R \supset (\neg \forall y (\textit{InBox}(y) \wedge \textit{Fragile}(y) \supset \textit{Broken}(y)) \supset \textsc{false})) \Rightarrow \textsc{false},\\
&\textsc{true} \wedge \neg R \Rightarrow \forall y \neg \textit{InBox}(y),\\
&\neg R \wedge \exists y\, \textit{InBox}(y) \Rightarrow \forall y (\textit{InBox}(y) \equiv y = \textit{gift}),\\
&\neg R \wedge \exists y\, \textit{InBox}(y) \Rightarrow \forall y (\textit{InBox}(y) \supset (\textit{Broken}(y) \equiv \textit{Fragile}(y))),\\
&\neg R \wedge \neg \forall y (\textit{InBox}(y) \wedge \textit{Fragile}(y) \supset \textit{Broken}(y)) \Rightarrow \textsc{false}\}.
\end{aligned}$$

In the progression $\Sigma'_{\text{bel}} \gg \textit{clink}$ the fluents *InBox* and *Broken* are renamed $R_{\textit{InBox}}$ and $R_{\textit{Broken}}$, respectively, and two conditionals are added to set *InBox* and *Broken* to its correct value:

$$\begin{aligned}
\Sigma'_{\text{bel}} \gg \textit{clink} = {} & (\Sigma'_{\text{bel}} * \varphi^a_{\textit{clink}})^{\textit{InBox}\ \textit{Broken}}_{R_{\textit{InBox}}\ R_{\textit{Broken}}} \cup\\
&\{\neg(\textit{InBox}(y) \equiv R_{\textit{InBox}}(y) \wedge \textit{clink} \neq \textit{unbox}(y)) \Rightarrow \textsc{false},\\
&\ \neg(\textit{Broken}(y) \equiv R_{\textit{Broken}}(y) \vee R_{\textit{InBox}}(y) \wedge \textit{Fragile}(y) \wedge \textit{clink} = \textit{dropbox}) \Rightarrow \textsc{false}\}.
\end{aligned}$$

Let us consider the fourth example projection problem from Section 4.2 another time:

$$\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models [\mathit{dropbox}][\mathit{clink}][\mathit{unbox}(\#5)]\exists y\,\mathbf{B}(\mathit{gift} = y \wedge \neg \mathit{InBox}(\mathit{gift}) \wedge \neg \mathit{Broken}(\mathit{gift})).$$

In this section, we have determined the progression of Σ_{bel} by *dropbox* and *clink*, and in Section 5.3 we regressed the query by $\mathit{unbox}(\#5)$. Together, this projection problem can therefore be recast as the purely static entailment problem

$$\begin{aligned}&\mathbf{O}_{\{R,R_{\mathit{InBox}},R_{\mathit{Broken}}\}}(\Sigma_{\text{dyn}}, \Sigma'_{\text{bel}} \gg \mathit{clink}) \models \\ &\quad \neg\mathbf{B}(\mathit{InBox}(\#5) \wedge \neg\mathit{Broken}(\#5) \Rightarrow \text{FALSE}) \wedge \mathbf{B}(\mathit{InBox}(\#5) \wedge \neg\mathit{Broken}(\#5) \Rightarrow \mathit{gift} = \#5 \wedge \neg\mathit{Broken}(\#5)) \vee \\ &\quad \mathbf{B}(\mathit{InBox}(\#5) \wedge \neg\mathit{Broken}(\#5) \Rightarrow \text{FALSE}).\end{aligned}$$

The equivalence of both entailment problems follows from the progression and regression results of Theorems 47 and 35. Let us confirm for this example that the (regressed) query indeed is a logical consequence of the (progressed) theory. In Section 4.2 we proved that $e \gg \mathit{dropbox} \gg \mathit{clink}$ satisfies the regressed query. So it suffices to show that $e \gg \mathit{dropbox} \gg \mathit{clink}$ also is the model of $\mathbf{O}_{\{R,R_{\mathit{InBox}},R_{\mathit{Broken}}\}}(\Sigma_{\text{dyn}}, \Sigma'_{\text{bel}} \gg \mathit{clink})$. Following the procedure from the proof of Theorem 19, $e' \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma'_{\text{bel}} \gg \mathit{clink})$ iff $e' = \langle e'_1, e'_2, e'_3, e'_4\rangle$ where

$$\begin{aligned}
e'_1 = \{w \mid w \models\; &\Sigma_{\text{dyn}} \wedge R \wedge (R \supset \forall y(R_{\mathit{InBox}}(y) \equiv y = \mathit{gift}) \wedge R_{\mathit{Broken}}(\mathit{gift}) \wedge \mathit{Fragile}(\mathit{gift})) \wedge \\
&\forall y(\mathit{InBox}(y) \equiv R_{\mathit{InBox}}(y)) \wedge \forall y(\mathit{Broken}(y) \equiv R_{\mathit{Broken}}(y))\};\\
e'_2 = \{w \mid w \models\; &\Sigma_{\text{dyn}} \wedge (R \supset \forall y(R_{\mathit{InBox}}(y) \equiv y = \mathit{gift}) \wedge R_{\mathit{Broken}}(\mathit{gift}) \wedge \mathit{Fragile}(\mathit{gift})) \wedge \\
&(\neg R \supset \forall y \neg R_{\mathit{InBox}}(y)) \wedge \\
&\forall y(\mathit{InBox}(y) \equiv R_{\mathit{InBox}}(y)) \wedge \forall y(\mathit{Broken}(y) \equiv R_{\mathit{Broken}}(y))\};\\
e'_3 = \{w \mid w \models\; &\Sigma_{\text{dyn}} \wedge (R \supset \forall y(R_{\mathit{InBox}}(y) \equiv y = \mathit{gift}) \wedge R_{\mathit{Broken}}(\mathit{gift}) \wedge \mathit{Fragile}(\mathit{gift})) \wedge \\
&(\neg R \supset \forall y(R_{\mathit{InBox}}(y) \supset y = \mathit{gift} \wedge (R_{\mathit{Broken}}(y) \equiv \mathit{Fragile}(y)))) \wedge \\
&\forall y(\mathit{InBox}(y) \equiv R_{\mathit{InBox}}(y)) \wedge \forall y(\mathit{Broken}(y) \equiv R_{\mathit{Broken}}(y))\};\\
e'_4 = \{w \mid w \models\; &\Sigma_{\text{dyn}} \wedge (R \supset \forall y(R_{\mathit{InBox}}(y) \equiv y = \mathit{gift}) \wedge R_{\mathit{Broken}}(\mathit{gift}) \wedge \mathit{Fragile}(\mathit{gift})) \wedge \\
&(\neg R \supset \forall y(R_{\mathit{InBox}}(y) \wedge \mathit{Fragile}(y) \supset R_{\mathit{Broken}}(y))) \wedge \\
&\forall y(\mathit{InBox}(y) \equiv R_{\mathit{InBox}}(y)) \wedge \forall y(\mathit{Broken}(y) \equiv R_{\mathit{Broken}}(y))\};
\end{aligned}$$

and then it is easy to verify that $e'_{\{R,R_{\mathit{InBox}},R_{\mathit{Broken}}\}} = e \gg \mathit{dropbox} \gg \mathit{clink}$: the R-worlds in e' represent the worlds from $(e \gg \mathit{dropbox} \gg \mathit{clink})_1$, and the $\neg R$-worlds represent the additional worlds in $(e \gg \mathit{dropbox} \gg \mathit{clink})_p$. Since there is no other model by Corollary 34, the action-less entailment indeed holds.

7. Belief revision postulates

In this section we relate $\mathcal{ESB}$ to the most well known accounts of belief change: the postulates by Alchourrón, Gärdenfors, and Makinson (AGM) for single revision [41, 42], Darwiche and Pearl's (DP) postulates for iterated revision [43], and the alternative proposal by Nayak, Pagnucco, and Peppas (NPP) [21]. We will see that the AGM postulates are satisfied and a slight weakening of the DP postulates hold. For strong revision furthermore all but the first NPP postulate are satisfied as well. The divergences from these postulate systems only concern the special case of revision by an inconsistent formula: this leads to the inconsistent epistemic state $\langle\{\}\rangle$, and our semantics provides no escape from the inconsistent state once it is reached.

We rephrase and prove the postulates in our *semantic* framework here, that is, with respect to an epistemic state e. The corresponding results for *theory* revision follow by the theorems from Section 6.2. Our translation of the postulates to $\mathcal{ESB}$ is similar to Shapiro et. al's [12]. Perhaps the most notable translation is that belief expansion is modeled as material implication: $e \models \mathbf{B}(\phi \supset \psi)$ represents that ψ is in the belief set after it is expanded with ϕ. It may also be worth noting that we phrase conditions like consistency or equivalence of objective formulas with respect to knowledge in the epistemic state, such as $e \not\models \mathbf{K}\neg\phi$ instead of $\not\models \neg\phi$, and $e \models \mathbf{K}(\phi \equiv \psi)$ instead of $\models (\phi \equiv \psi)$. We remark that this is no effective restriction, as we prove the below theorems for arbitrary epistemic states, which particularly includes those where precisely the valid sentences are known.

For the remainder of this section let e be an arbitrary epistemic state and ϕ, ψ, υ be objective sentences.

Theorem 48. *The AGM postulates are satisfied:*

1. *If* $e * \phi \models \mathbf{B}\psi$ *and* $e * \phi \models \mathbf{B}(\psi \supset \upsilon)$*, then* $e * \phi \models \mathbf{B}\upsilon$*.*
2. $e * \phi \models \mathbf{B}\phi$*.*
3. *If* $e * \phi \models \mathbf{B}\upsilon$*, then* $e \models \mathbf{B}(\phi \supset \upsilon)$*.*
4. *If* $e \not\models \mathbf{B}\neg\phi$ *and* $e \models \mathbf{B}(\phi \supset \upsilon)$*, then* $e * \phi \models \mathbf{B}\upsilon$*.*
5. *If* $e \not\models \mathbf{K}\neg\phi$*, then* $e * \phi \not\models \mathbf{B}\textsc{false}$*.*
6. *If* $e \models \mathbf{K}(\phi \equiv \psi)$*, then* $e * \phi \models \mathbf{B}\upsilon$ *iff* $e * \psi \models \mathbf{B}\upsilon$*.*
7. *If* $e * (\phi \wedge \psi) \models \mathbf{B}\upsilon$*, then* $e * \phi \models \mathbf{B}(\psi \supset \upsilon)$*.*
8. *If* $e * \phi \not\models \mathbf{B}\neg\psi$ *and* $e * \phi \models \mathbf{B}(\psi \supset \upsilon)$*, then* $e * (\phi \wedge \psi) \models \mathbf{B}\upsilon$*.*

Proof. We suppose $e_p \neq \{\}$ for some $p \in \mathbb{N}$, for otherwise the postulates hold trivially as $(e * \delta)_p = \{\}$ and thus $e * \delta \models \mathbf{B}\textsc{false}$. Since the postulates refer only to a single revision and $(e *_{\mathrm{w}} \delta)_1 = (e *_{\mathrm{s}} \delta)_1$ by Lemma 11, the proof does not need to distinguish between weak and strong revision.

1. Follows from Property 7 of Theorem 17.
2. If there is no ϕ-world, $(e * \phi)_p = \{\}$ for all $p \in \mathbb{N}$; else for all $w \in (e * \phi)_1 \neq \{\}$, $w \models \phi$. In either case, $e * \phi \models \mathbf{B}\phi$.
3. Let $e * \phi \models \mathbf{B}\upsilon$. Suppose $e \not\models \mathbf{B}(\phi \supset \upsilon)$. Then for some $w \in e_{\lfloor e \mid \textsc{true} \rfloor}$, $w \models \phi \wedge \neg\upsilon$. Then $w \in (e * \phi)_1$, and therefore $e * \phi \not\models \mathbf{B}\upsilon$, which contradicts the assumption. Thus $e \models \mathbf{B}(\phi \supset \upsilon)$.
4. Let $e \not\models \mathbf{B}\neg\phi$ and $e \models \mathbf{B}(\phi \supset \upsilon)$. Then for some $w \in e_{\lfloor e \rfloor}$, $w \models \phi$, and for all $w \in e_{\lfloor e \rfloor}$, $w \models \phi \supset \upsilon$. Therefore $(e * \phi)_1 \subseteq e_p$, and for all $w \in (e * \phi)_1$, $w \models \phi \wedge \upsilon$, so $e * \phi \models \mathbf{B}\upsilon$.
5. Let $e \not\models \mathbf{K}\neg\phi$. Then for some $w \in e_{\lfloor e \rfloor}$, $w \models \phi$. Hence $(e * \phi)_1 \neq \{\}$, so $e * \phi \not\models \mathbf{B}\textsc{false}$.
6. Let $e \models \mathbf{K}(\phi \equiv \psi)$. Then $w \models \phi$ iff $w \models \psi$ for all $p \in \mathbb{N}$ and $w \in e_p$. Thus $e * \phi = e * \psi$.
7. If there is no ϕ-world, $e * \phi \models \mathbf{B}(\psi \supset \upsilon)$ holds trivially. Otherwise, consider $w \in (e * \phi)_1$ with $w \models \psi$. Then $w \in (e * (\phi \wedge \psi))_1$ and by assumption, $w \models \upsilon$. Hence $e * \phi \models \mathbf{B}(\psi \supset \upsilon)$.
8. Let $e * \phi \not\models \mathbf{B}\neg\psi$ and $e * \phi \models \mathbf{B}(\psi \supset \upsilon)$. Then for some $w \in (e * \phi)_1$, $w \models \psi$. Therefore $(e * (\phi \wedge \psi))_1 \subseteq (e * \phi)_1$, and for all $w \in (e * (\phi \wedge \psi))_1$, $w \models \psi \wedge (\psi \supset \upsilon)$, and so $w \models \upsilon$. Hence $e * (\phi \wedge \psi) \models \mathbf{B}\upsilon$. □

A slightly restricted version of the DP postulates for iterated revision holds as well. The restriction concerns the special case of revision by an inconsistent formula. Since $\mathcal{ESB}$ provides no escape from the empty epistemic state, DP2 holds in case the first revision is by an inconsistent formula only if the second revision is by an inconsistent formula as well. We hence require $e \models \mathbf{K}\neg\phi \supset \mathbf{K}\neg\psi$ in our variant of DP2. We remark that the restricted postulate is still slightly stronger than NPP4 (see below).

Theorem 49. *The DP postulates are satisfied for a restricted second postulate:*

1. *If* $e \models \mathbf{K}(\psi \supset \phi)$*, then* $(e * \phi) * \psi \models \mathbf{B}\upsilon$ *iff* $e * \psi \models \mathbf{B}\upsilon$*.*
2. *If* $e \models \mathbf{K}(\psi \supset \neg\phi)$ *and* $e \models \mathbf{K}\neg\phi \supset \mathbf{K}\neg\psi$*, then* $(e * \phi) * \psi \models \mathbf{B}\upsilon$ *iff* $e * \psi \models \mathbf{B}\upsilon$*.*
3. *If* $e * \psi \models \mathbf{B}\phi$*, then* $(e * \phi) * \psi \models \mathbf{B}\phi$*.*
4. *If* $e * \psi \not\models \mathbf{B}\neg\phi$*, then* $(e * \phi) * \psi \not\models \mathbf{B}\neg\phi$*.*

Proof.

1. Let $e \models \mathbf{K}(\psi \supset \phi)$. Then $(e \mid \psi)_p \subseteq (e \mid \phi)_p$ for all $p \in \mathbb{N}$ (*). Therefore $w \in ((e * \phi) * \psi)_1$ iff $\lfloor e * \phi \mid \psi \rfloor \neq \infty$ and $w \in (e * \phi \mid \psi)_{\lfloor e * \phi \mid \psi \rfloor}$ iff (by (*)) $\lfloor e \mid \psi \rfloor \neq \infty$ and $w \in (e \mid \psi)_{\lfloor e \mid \psi \rfloor}$ iff $w \in (e * \psi)_1$.
2. The proof is very similar to 1. Let $e \models \mathbf{K}(\psi \supset \neg\phi)$ and $e \models \mathbf{K}\neg\phi \supset \mathbf{K}\neg\psi$. If $\lfloor e \mid \psi \rfloor = \infty$, then $\lfloor e * \phi \mid \psi \rfloor = \infty$, and therefore $e * \psi = (e * \phi) * \psi = \langle\{\}\rangle$, so the postulate holds. Now suppose $\lfloor e \mid \psi \rfloor \neq \infty$. By the second assumption, $e \not\models \mathbf{K}\neg\phi$, so $\lfloor e \mid \phi \rfloor \neq \infty$. By the first assumption, $w \not\models \psi$ for all $w \in (e \mid \phi)_p$ for all $p \in \mathbb{N}$. Thus $(e \mid \phi)_p \cap (e \mid \psi)_p = \{\}$ (*), so neither weak nor strong revision by ϕ affects the relative order of the ψ-worlds. Hence $w \in ((e * \phi) * \psi)_1$ iff $w \in (e * \phi \mid \psi)_{\lfloor e * \phi \mid \psi \rfloor}$ iff (by (*)) $w \in (e \mid \psi)_{\lfloor e \mid \psi \rfloor}$ iff $w \in (e * \psi)_1$.

3. Let $e * \psi \models \mathbf{B}\phi$. If $\lfloor e \mid \phi \rfloor = \infty$, then $e * \phi = (e * \phi) * \psi = \langle\{\}\rangle$, so the postulate holds. If $\lfloor e \mid \psi \rfloor = \infty$, then $\lfloor e * \phi \mid \psi \rfloor = \infty$, and therefore $(e * \phi) * \psi = \langle\{\}\rangle$, so the postulate holds. Now suppose $\lfloor e \mid \phi \rfloor \neq \infty$ and $\lfloor e \mid \psi \rfloor \neq \infty$. By assumption, $(e \mid \psi)_{\lfloor e \mid \psi \rfloor} \subseteq (e \mid \phi)_{\lfloor e \mid \psi \rfloor}$. Therefore the most-plausible ψ-worlds remain most plausible after weak or strong revision by ϕ, so $(e * \phi \mid \psi)_{\lfloor e * \phi \mid \psi \rfloor} \subseteq (e * \phi \mid \phi)_{\lfloor e * \phi \mid \psi \rfloor}$. Thus, if $w \in ((e * \phi) * \psi)_1$, then $w \in (e * \phi \mid \psi)_{\lfloor e * \phi \mid \psi \rfloor}$, and $w \in (e * \phi \mid \phi)_{\lfloor e * \phi \mid \psi \rfloor}$, so $w \models \phi$.
4. Let $e * \psi \not\models \mathbf{B}\neg\phi$. Then $\lfloor e * \psi \mid \phi \rfloor \neq \infty$, and so $\lfloor e \mid \phi \rfloor \neq \infty$ and $\lfloor e \mid \psi \rfloor \neq \infty$. By assumption, for some $(e \mid \psi \wedge \phi)_{\lfloor e \mid \psi \rfloor} \neq \{\}$. The ϕ-worlds among the most-plausible ψ-worlds remain most plausible after weak or strong revision by ϕ, so $(e * \phi \mid \psi \wedge \phi)_{\lfloor e * \phi \mid \psi \rfloor} \neq \{\}$. Hence there is some $w \in (e * \phi \mid \psi \wedge \phi)_{\lfloor e * \phi \mid \psi \rfloor}$, and therefore also $w \in (e * \phi \mid \psi)_{\lfloor e * \phi \mid \psi \rfloor}$. Thus $w \models \phi$ for some $w \in ((e * \phi) * \psi)_1$. □

The Nayak–Pagnucco–Peppas postulates hold with two exceptions: for one thing, the absurdity postulate NPP1 does not hold; for another, the conjunction postulate NPP3 only holds for strong revision. The absurdity postulate NPP1 facilitates recovery from an inconsistent state: it says that after revising an inconsistent state by ϕ, ϕ shall be all that is believed. In our language, the absurdity postulate could be written as

1. If $e \models \mathbf{K}\text{FALSE}$, then $e * \phi \models \mathbf{O}\{\text{TRUE} \Rightarrow \phi\}$.

Only-believing ϕ alone intuitively does not suffice, though, because we would lose any indefeasible *knowledge* we might have had already before reaching the inconsistent e (such as the dynamic axioms of a basic action theory).

Theorem 50. *The NPP postulates except for NPP1 and except for NPP3 for weak revision hold:*

2. *AGM1–6 hold.*
3. *If $e \not\models \mathbf{K}\neg(\phi \wedge \psi)$ then $(e *_s \phi) *_s \psi \models \mathbf{B}\upsilon$ iff $e *_s (\phi \wedge \psi) \models \mathbf{B}\upsilon$.*
4. *If $e \models \mathbf{K}(\psi \supset \neg\phi)$ and $e \not\models \mathbf{K}\neg\phi$, then $(e * \phi) * \psi \models \mathbf{B}\upsilon$ iff $e * \psi \models \mathbf{B}\upsilon$.*

Proof. We only need to prove NPP3, as we have shown the AGM1–6 in Theorem 48 already and NPP4 is a special case of DP2. Suppose $w \in ((e *_s \phi) *_s \psi)_1$. Then $w \in (e *_s \phi)_{\lfloor e *_s \phi \mid \psi \rfloor}$. By assumption, $\lfloor e *_s \phi \mid \psi \rfloor < \lfloor e *_s \phi \mid \neg\phi \rfloor$, so $w \models \phi$ and, since the revision by ϕ did not affect the relative ordering of the $(\phi \wedge \psi)$-worlds, also $w \in e_{\lfloor e \mid \phi \wedge \psi \rfloor}$. Thus $w \in (e *_s \phi \wedge \psi)_1$. Conversely, suppose $w \in (e *_s \phi \wedge \psi)_1$. Then $w \in e_{\lfloor e \mid \phi \wedge \psi \rfloor}$. Therefore $w \in (e *_s \phi)_{\lfloor (e *_s \phi) \mid \psi \rfloor}$, and thus $w \in ((e *_s \phi) *_s \psi)_1$. □

8. Conclusion

We have studied defeasible beliefs in a modal variant of the situation calculus. The first-order language takes after Lakemeyer and Levesque's epistemic situation calculus, but replaces the notion of indefeasible knowledge with the more expressive concept of conditional belief. Only-believing allows for representing all that is conditionally believed; it thus generalizes Levesque's only-knowing to the case of beliefs. Newly obtained information is incorporated by two classical belief revision schemes, namely natural or lexicographic revision, and is thus itself defeasible as well. Introspection of beliefs and quantifying-in both work as expected. Meanwhile, the stratified-possible-worlds semantics is relatively simple and easy to work with compared to other approaches. The unique-model property of only-believing also contributes to simplicity of the formalism. Within this framework, we investigated the belief projection problem for Reiter-style basic action theories, and gave solutions by regression as well as by progression.

As for future work, a very interesting problem is when and how belief progression is first-order definable. The particular challenge is to capture the *revision* of a conditional knowledge base without second-order logic. For natural revision we have already presented a first-order representation in [11]. Whether this is possible for lexicographic revision remains open. Another avenue of future work is to investigate other belief revision operators. Particularly revision operators using Spohn's ranking functions [34] seem challenging because of their use of arithmetic. The main goal of this work is to integrate this theory into a practical reasoning system. While its first-order nature prohibits implementing of the full theory, we are currently working towards a limited-belief version, which retains the first-order features but sacrifices completeness to make reasoning decidable. The system is based on the theory of limited conditional belief in [22] and the reasoner from [23]. We hope to use this system for epistemic planning in robotics.

Acknowledgments

This work was partially supported by the DAAD Go8 project 56266625 and by the DFG Research Unit FOR 1513. The first author was supported by a DAAD doctoral scholarship.

Appendix A. Proof of belief projection by regression

In this appendix we prove the regression results. We begin with Theorems 24 and 30 from Section 5. Then we generalize Theorem 30 for the extended only-believing operator to show Theorem 35.

Proving Theorems 24 and 30 follows a scheme similar to the knowledge regression proof sketched in [24]. Namely, we show that every world and epistemic state can be converted to one that adheres to the dynamic axioms Σ_{dyn} without changing its initial truth values. In Lemma 60 we show that an epistemic state and a world satisfy a regressed sentence iff their Σ_{dyn}-compliant counterparts satisfy the non-regressed sentence. The regression theorem is then an easy consequence.

For the rest of this section, let $\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}$ be a basic action theory over fluents $\mathcal{F}$. Recall that Σ_{dyn} contains the successor state axioms $\Box[a]F(x_1,\ldots,x_k) \equiv \gamma_F$ for $F \in \mathcal{F}$, and the informed-fluent axiom $\Box IF(a) \equiv \varphi$.

Definition 51. For a world w, $w_{\Sigma_{\text{dyn}}}$ is a world such that $w_{\Sigma_{\text{dyn}}} \approx_{\mathcal{F}\cup\{IF\}} w$ and

- $w_{\Sigma_{\text{dyn}}}[F(n_1,\ldots,n_k),\langle\rangle] = w[F(n_1,\ldots,n_k),\langle\rangle]$ for all $F \in \mathcal{F}$;
- $w_{\Sigma_{\text{dyn}}}[F(n_1,\ldots,n_k), z\cdot n] = 1$ iff $w_{\Sigma_{\text{dyn}}} \gg z \models \gamma_F{}^{x_1\,\ldots\,x_k\,a}_{n_1\,\ldots\,n_k\,n}$ for all $F \in \mathcal{F}$, action sequences z, and actions n;
- $w_{\Sigma_{\text{dyn}}}[IF(n), z] = 1$ iff $w_{\Sigma_{\text{dyn}}} \gg z \models \varphi^a_n$ for all action sequences z.

For a set of worlds W and an epistemic state e, we let $W_{\Sigma_{\text{dyn}}} = \{w_{\Sigma_{\text{dyn}}} \mid w \in W\}$ and $e_{\Sigma_{\text{dyn}}} = \langle (e_1)_{\Sigma_{\text{dyn}}}, \ldots, (e_{\lceil e\rceil})_{\Sigma_{\text{dyn}}}\rangle$.

Lemma 52. *$w_{\Sigma_{\text{dyn}}}$ is uniquely defined.*

Proof. Intuitively, once all values except for IF are fixed after z, the truth of γ_F and φ after z is uniquely determined as they are fluent formulas, and thus by definition also the value of F after $z \cdot n$ and of IF after z are uniquely determined. The formal proof is by straightforward induction on z and subinduction on the length of γ_F and φ. □

Lemma 53. *$w_{\Sigma_{\text{dyn}}} \models \Sigma_{\text{dyn}}$.*

Proof. By definition, $w_{\Sigma_{\text{dyn}}}[F(n_1,\ldots,n_k), z\cdot n] = 1$ iff $w_{\Sigma_{\text{dyn}}} \gg z \models \gamma_F{}^{x_1\,\ldots\,x_k\,a}_{n_1\,\ldots\,n_k\,n}$, so $w_{\Sigma_{\text{dyn}}} \models \Box[a]F(x_1,\ldots,x_k) \equiv \gamma_F$ for all $F \in \mathcal{F}$. Analogously, $w_{\Sigma_{\text{dyn}}} \models \Box IF(a) \equiv \varphi$. Hence $w_{\Sigma_{\text{dyn}}} \models \Sigma_{\text{dyn}}$. □

Lemma 54. *If $w \models \Sigma_{\text{dyn}}$, then $w_{\Sigma_{\text{dyn}}} = w$.*

Proof. Suppose $w \models \Sigma_{\text{dyn}}$. Then $w \models \Box[a]F(x_1,\ldots,x_n) \equiv \gamma_F$ and $w \models \Box IF(a) \equiv \varphi$. Thus, w satisfies the conditions from Definition 51: $w[F(n_1,\ldots,n_k), z\cdot n] = 1$ iff $(w \gg z\cdot n)[F(n_1,\ldots,n_k),\langle\rangle] = 1$ iff $w \gg z \models \gamma_F{}^{x_1\,\ldots\,x_k\,a}_{n_1\,\ldots\,n_k\,n}$ for all $F \in \mathcal{F}$; analogously, $w[IF(n), z] = 1$ iff $(w \gg z)[IF(n),\langle\rangle] = 1$ iff $w \gg z \models \varphi^a_n$. By Lemma 52 $w_{\Sigma_{\text{dyn}}}$ is unique, so $w_{\Sigma_{\text{dyn}}} = w$. □

Lemma 55. *Let ϕ be a fluent sentence. Then $w \models \phi$ iff $w_{\Sigma_{\text{dyn}}} \models \phi$.*

Proof. By an easy induction on the length of ϕ since $w, w_{\Sigma_{\text{dyn}}}$ agree on all initial values except perhaps for IF. □

Lemma 56. *If $e \models \mathbf{O}\Sigma_{\text{bel}}$, then $e_{\Sigma_{\text{dyn}}} \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$.*

Proof. Let $\Sigma_{\text{bel}} = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ and $e \models \mathbf{O}\Sigma_{\text{bel}}$. We show that $e_{\Sigma_{\text{dyn}}} \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$. Note that by Lemma 55, $\lfloor e \mid \phi_i\rfloor = \lfloor e_{\Sigma_{\text{dyn}}} \mid \phi_i\rfloor$ (*). Suppose $w \in (e_{\Sigma_{\text{dyn}}})_p$. Then there is some $w' \in e_p$ such that $w'_{\Sigma_{\text{dyn}}} = w$, and $w' \models \bigwedge_{i:\lfloor e\mid\phi_i\rfloor \geq p}(\phi_i \supset \psi_i)$ iff (by Lemmas 53 and 55 and (*)) $w \models \Sigma_{\text{dyn}} \wedge \bigwedge_{i:\lfloor e_{\Sigma_{\text{dyn}}}\mid\phi_i\rfloor \geq p}(\phi_i \supset \psi_i)$. Conversely, suppose $w \models \Sigma_{\text{dyn}} \wedge \bigwedge_{i:\lfloor e_{\Sigma_{\text{dyn}}}\mid\phi_i\rfloor \geq p}(\phi_i \supset \psi_i)$. Then $w \in e_p$ by Rule S10 and (*). By Lemma 54, $w = w_{\Sigma_{\text{dyn}}} \in (e_{\Sigma_{\text{dyn}}})_p$. □

For induction proofs about regression we introduce the following non-standard measure. Intuitively, $\|\alpha\|$ measures the length of the regressed formula $\mathcal{R}[\alpha]$ plus how many "calls" to the regression operator it takes to determine $\mathcal{R}[\alpha]$ (not counting Rule R8).

Definition 57. Let α be a regressable formula and $k \geq 0$. We define the measure $\|\alpha\|$ with respect to a basic action theory with dynamic axioms Σ_{dyn} as

- $\|[t_1]\ldots[t_k]R(t'_1,\ldots,t'_l)\| = 1$ for rigid R;
- $\|[t_1]\ldots[t_k]F(t'_1,\ldots,t'_l)\| = \begin{cases} 1 & \text{if } k = 0 \\ 1 + \|[t_1]\ldots[t_{k-1}]\gamma_F\| & \text{if } k > 0 \end{cases}$ for fluent $F \in \mathcal{F}$;
- $\|[t_1]\ldots[t_k]IF(t)\| = 1 + \|[t_1]\ldots[t_k]\varphi\|$;
- $\|[t_1]\ldots[t_k](t'_1 = t'_2)\| = 1$;
- $\|[t_1]\ldots[t_k]\neg\alpha\| = 1 + \|[t_1]\ldots[t_k]\alpha\|$;
- $\|[t_1]\ldots[t_k](\alpha \vee \beta)\| = 1 + \|[t_1]\ldots[t_k]\alpha\| + \|[t_1]\ldots[t_k]\beta\|$;
- $\|[t_1]\ldots[t_k]\exists x\,\alpha\| = 1 + \|[t_1]\ldots[t_k]\alpha\|$;
- $\|[t_1]\ldots[t_k]\mathbf{B}(\alpha \Rightarrow \beta)\| = \begin{cases} 1 + \|(\alpha \supset \beta)\| & \text{if } k = 0 \\ 1 + \|[t_1]\ldots[t_{k-1}]\sigma\| & \text{if } k > 0 \end{cases}$ where σ is the right-hand side of Theorem 27 or 28 depending on the sort of t_k.

Observe that $\|[t_1]\ldots[t_k]\alpha\|$ reflects the regression operator $\mathcal{R}[\langle t_1,\ldots,t_k\rangle,\alpha]$ from Definitions 23 and 29. For example, $\|[t]F(t')\| = 1 + \|\gamma_F\|$ corresponds to $\mathcal{R}[[t]F(t')] = \mathcal{R}[\gamma_F{}^{x\,a}_{t'\,t}]$; similarly for the other cases. This makes $\|\cdot\|$ very useful for induction proofs involving regression: the base cases are $\|[t_1]\ldots[t_k]R(t'_1,\ldots,t'_l)\|$ for rigid R, $\|F(t'_1,\ldots,t'_l)\|$ for fluent $F \in \mathcal{F}$, and $\|[t_1]\ldots[t_k](t = t')\|$, whose regression is trivial; all other cases are proved by induction.

We first need to show that $\|\cdot\|$ is a well-defined function from the regressable formulas to the natural numbers. Intuitively this is true because the right-hand sides of $\|\cdot\|$ for fluent atoms and beliefs eliminate an action or push it inside the belief, respectively, and the right-hand side for IF mentions no IF itself. Given the construction of $\|\cdot\|$ it then follows immediately that the measure for expressions on the left-hand side of the equations in Definition 57 is always bigger than the measure of expressions on the right-hand side.

Lemma 58. *$\|\cdot\|$ is a well-defined function from the regressable formulas to the natural numbers.*

Proof. Let $|\alpha|_B$ be the nesting depth of $\mathbf{B}$ operators: $|R(t_1,\ldots,t_k)|_B = |F(t_1,\ldots,t_k)|_B = |(t = t')|_B = 0$ for rigid R and fluent F; $|\neg\alpha|_B = |\exists x\,\alpha|_B = |[t]\alpha|_B = |\alpha|_B$; $|(\alpha \vee \beta)|_B = \max\{|\alpha|_B, |\beta|_B\}$; and $|\mathbf{B}(\alpha \Rightarrow \beta)|_B = 1 + \max\{|\alpha|_B, |\beta|_B\}$.

Let $|\alpha|_A$ be as follows: $|R(t_1,\ldots,t_k)|_A = |F(t_1,\ldots,t_k)|_A = |(t = t')|_A = 0$ for rigid R and fluent F; $|\neg\alpha|_A = |\exists x\,\alpha|_A = |\alpha|_A$; $|(\alpha \vee \beta)|_A = |\mathbf{B}(\alpha \Rightarrow \beta)|_A = \max\{|\alpha|_A, |\beta|_A\}$; and $|[t]\alpha|_A = 2^{|\alpha|_B} + |\alpha|_A$. Note that for objective ϕ, $|\phi|_A$ is just the number of nested action operators in ϕ. In subjective formulas every action is additionally penalized with $|\cdot|_B$.

First we show that $|[t_1]\ldots[t_k]\mathbf{B}(\alpha \Rightarrow \beta)|_A > |[t_1]\ldots[t_{k-1}]\sigma|_A$ for $k > 0$ where σ is the right-hand side of Theorem 27 or 28 (*). Let $|\mathbf{B}(\alpha \Rightarrow \beta)|_B = n$. Then $|[t_1]\ldots[t_k]\mathbf{B}(\alpha \Rightarrow \beta)|_A = k \cdot 2^n + \max\{|\alpha|_A, |\beta|_A\}$. We also have that $|\sigma|_B = n$. Hence $|[t_1]\ldots[t_{k-1}]\sigma|_A = (k-1)\cdot 2^n + |\sigma|_A$. It is immediate from Theorems 27 and 28 that $|\sigma|_A = \max\{|[t_k]\alpha|_A, |[t_k]\beta|_A\}$. Since $|\alpha|_B \leq n-1$ and $|\beta|_B \leq n-1$, we have $\max\{|[t_k]\alpha|_A, |[t_k]\beta|_A\} \leq 2^{n-1} + \max\{|\alpha|_A, |\beta|_A\}$. Thus $|[t_1]\ldots[t_{k-1}]\sigma|_A \leq (k-1)\cdot 2^n + 2^{n-1} + \max\{|\alpha|_A, |\beta|_A\}$. Since $k \cdot 2^n > (k-1)\cdot 2^n + 2^{n-1}$, (*) holds.

Now we prove the lemma by induction on $|\alpha|_A$. For the base case, consider regressable α with $|\alpha|_A = 0$. We show that $\|\alpha\|$ is well-defined by subinduction on the length of α, where we take the length of $IF(t)$ to be the length of φ plus 1 (which is well-behaved because φ contains no IF), and the length of $\mathbf{B}(\alpha \Rightarrow \beta)$ to be the length of $(\alpha \supset \beta)$ plus 1. The subinduction base cases $\|R(t'_1,\ldots,t'_l)\| = 1$ for rigid R, $\|F(t'_1,\ldots,t'_l)\| = 1$ for fluent $F \in \mathcal{F}$, and $\|(t = t')\| = 1$ are obviously well-defined. For the subinduction steps, $\|IF(t)\|$ is well-defined iff $\|\varphi\|$ is well-defined, $\|\neg\alpha\|$ is well-defined iff $\|\alpha\|$ is well-defined, $\|(\alpha \vee \beta)\|$ is well-defined if $\|\alpha\|$ and $\|\beta\|$ are well-defined, $\|\exists x\,\alpha\|$ is well-defined iff $\|\alpha\|$ is well-defined, $\|\mathbf{B}(\alpha \Rightarrow \beta)\|$ is well-defined iff $\|(\alpha \supset \beta)\|$ is well-defined, all of which is the case by subinduction.

For the induction step consider α with $|\alpha|_A = m > 0$ and suppose that $\|\beta\|$ is well defined for all regressable β with $|\beta|_A < m$. We show that $\|\alpha\|$ is well-defined by a subinduction in the same vein as in the main base case. The first base case $\|[t_1]\ldots[t_m]F(t'_1,\ldots,t'_l)\|$ for fluent $F \in \mathcal{F}$ is well-defined iff $\|[t_1]\ldots[t_{m-1}]\gamma_F\|$ is well-defined, which holds by induction since γ_F is fluent and thus mentions neither actions, beliefs, nor IF, so $|[t_1]\ldots[t_{m-1}]\gamma_F|_{\mathrm{A}} = m-1$. The other base cases $\|[t_1]\ldots[t_m]R(t'_1,\ldots,t'_l)\| = 1$ for rigid R and $\|[t_1]\ldots[t_m](t = t')\| = 1$ are immediate. For the first subinduction step, $\|[t_1]\ldots[t_m]IF(t)\|$ is well-defined iff $\|[t_1]\ldots[t_m]\varphi\|$ is well-defined, which holds by subinduction. For disjunction with $|[t_1]\ldots[t_k](\alpha\vee\beta)|_{\mathrm{A}} = m$, $\|[t_1]\ldots[t_k](\alpha\vee\beta)\|$ is well-defined iff $\|[t_1]\ldots[t_k]\alpha\|$ and $\|[t_1]\ldots[t_k]\beta\|$ are well-defined, which for α holds by induction in case $|[t_1]\ldots[t_k]\alpha|_{\mathrm{A}} < m$ and otherwise by subinduction, and likewise for β. The subinduction steps $\|[t_1]\ldots[t_k]\neg\alpha\|$ and $\|[t_1]\ldots[t_k]\exists x\,\alpha\|$ trivially hold by subinduction. For the subinduction step for beliefs, let $|[t_1]\ldots[t_k]\mathbf{B}(\alpha \Rightarrow \beta)|_{\mathrm{A}} = m$; then $\|[t_1]\ldots[t_k]\mathbf{B}(\alpha \Rightarrow \beta)\|$ is well-defined iff $\|[t_1]\ldots[t_{k-1}]\sigma\|$ is well-defined where σ is the right-hand side of Theorem 27 or 28 depending on the sort of t_k, which holds by induction since $|[t_1]\ldots[t_k]\mathbf{B}(\alpha \Rightarrow \beta)|_{\mathrm{A}} > |[t_1]\ldots[t_{k-1}]\sigma|_{\mathrm{A}}$ by (*). □

Since $\|\alpha\|$ is a natural number for every regressable α, we can prove properties of α by induction over $\|\alpha\|$, as we will do in the next two lemmas. For example, for the induction step for a fluent atom $[t_1]\ldots[t_k]F(t'_1,\ldots,t'_l)$ after $k \geq 1$ actions, we have $\mathcal{R}[[t_1]\ldots[t_k]F(t'_1,\ldots,t'_l)] = \mathcal{R}[[t_1]\ldots[t_{k-1}]\gamma_F{}^{x_1 \ldots x_l\, a}_{t'_1 \ldots t'_l\, t_k}]$ by the definition of $\mathcal{R}$, and then use the induction assumption since clearly $\|[t_1]\ldots[t_k]F(t'_1,\ldots,t'_l)\| > \|[t_1]\ldots[t_{k-1}]\gamma_F\|$ by Definition 57.

Lemma 59. *Let α be a regressable sentence. Then $\mathcal{R}[\alpha^x_n] = \mathcal{R}[\alpha]^x_n$.*

Proof. By induction on $\|\alpha\|$. For the base case let $\|\alpha\| = 1$. For rigid R and $k \geq 0$, $\mathcal{R}[([t_1]\ldots[t_k]R(t'_1,\ldots,t'_l))^x_n] = R(t'_1,\ldots,t'_l)^x_n = \mathcal{R}[[t_1]\ldots[t_k]R(t'_1,\ldots,t'_l)]^x_n$; analogously for $[t_1]\ldots[t_k](t = t')$. For fluent $F \in \mathcal{F}$, $\mathcal{R}[F(t'_1,\ldots,t'_l)^x_n] = F(t'_1,\ldots,t'_l)^x_n = \mathcal{R}[F(t'_1,\ldots,t'_l)]^x_n$.

For the induction step let $\|\alpha\| = m > 1$ and suppose the lemma holds for all β with $\|\beta\| < m$. For fluent $F \in \mathcal{F}$ and $k \geq 1$, $\mathcal{R}[([t_1]\ldots[t_k]F(t'_1,\ldots,t'_l))^x_n]$ = (by Rules R2 and R8) $\mathcal{R}[[t_1{}^x_n]\ldots[t_{k-1}{}^x_n]\gamma_F{}^{x_1 \ldots x_l\, a}_{t'_1{}^x_n \ldots t'_l\, t_k{}^x_n}]$ = (since γ_F does not mention x due to rectification) $\mathcal{R}[([t_1]\ldots[t_{k-1}]\gamma_F{}^{x_1 \ldots x_l\, a}_{t'_1 \ldots t'_l\, t_k})^x_n]$ = (by induction since $\|[t_1]\ldots[t_{k-1}]\gamma_F\| < m$) $\mathcal{R}[[t_1]\ldots[t_{k-1}]\gamma_F{}^{x_1 \ldots x_l\, a}_{t'_1 \ldots t'_l\, t_k}]^x_n$ = (by Rules R2 and R8) $\mathcal{R}[[t_1]\ldots[t_k]F(t'_1,\ldots,t'_l)]^x_n$. Similarly for the induction step for IF, $\mathcal{R}[([t_1]\ldots[t_k]IF(t))^x_n] = \mathcal{R}[[t_1{}^x_n]\ldots[t_k{}^x_n]IF(t^x_n)]$ = (by Rules R3 and R8) $\mathcal{R}[[t_1{}^x_n]\ldots[t_k{}^x_n]\varphi^a_{t^x_n}]$ = (since x does not occur in φ) $\mathcal{R}[([t_1]\ldots[t_k]\varphi^a_t)^x_n]$ = (by induction since $\|[t_1]\ldots[t_k]\varphi\| < m$) $\mathcal{R}[[t_1]\ldots[t_k]\varphi^a_t]^x_n$ = (by Rules R3 and R8) $\mathcal{R}[IF(t)]^x_n$.

For a quantifier, $\mathcal{R}[([t_1]\ldots[t_k]\exists x'\,\alpha)^x_n]$ = (since x', x are distinct due to rectification and by Rules R7 and R8) $\exists x'\,\mathcal{R}[[t_1{}^x_n]\ldots[t_k{}^x_n]\alpha^x_n]$ = (by induction since $\|[t_1]\ldots[t_k]\alpha\| < m$) $\exists x'\,\mathcal{R}[[t_1]\ldots[t_k]\alpha]^x_n$ = (since x', x are distinct and by Rules R8 and R7) $\mathcal{R}[[t_1]\ldots[t_k]\exists x'\,\alpha]^x_n$. We omit the similar induction steps for $[t_1]\ldots[t_k]\neg\alpha$ and $[t_1]\ldots[t_k](\alpha \vee \beta)$.

For beliefs after no actions, $\mathcal{R}[\mathbf{B}(\alpha \Rightarrow \beta)^x_n] = \mathcal{R}[\mathbf{B}(\alpha^x_n \Rightarrow \beta^x_n)]$ = (by Rule R9) $\mathbf{B}(\mathcal{R}[\alpha^x_n] \Rightarrow \mathcal{R}[\beta^x_n])$ = (by induction since $\|\alpha\| < m$ and $\|\beta\| < m$) $\mathbf{B}(\mathcal{R}[\alpha]^x_n \Rightarrow \mathcal{R}[\beta]^x_n) = \mathbf{B}(\mathcal{R}[\alpha] \Rightarrow \mathcal{R}[\beta])^x_n$ = (by Rule R9) $\mathcal{R}[\mathbf{B}(\alpha \Rightarrow \beta)]^x_n$. Finally, for beliefs after $k \geq 1$ actions, $\mathcal{R}[([t_1]\ldots[t_k]\mathbf{B}(\alpha \Rightarrow \beta))^x_n]$ = (by Rules R9 and R8, where σ is the right-hand side of Theorem 27 or 28 depending on the sort of t_k) $\mathcal{R}[[t_1{}^x_n]\ldots[t_{k-1}{}^x_n]\sigma^{x\,a}_{n\,t_k{}^x_n}] = \mathcal{R}[([t_1]\ldots[t_{k-1}]\sigma^a_{t_k})^x_n]$ = (by induction since $\|[t_1]\ldots[t_{k-1}]\sigma\| < m$) $\mathcal{R}[[t_1]\ldots[t_{k-1}]\sigma^a_{t_k}]^x_n$ = (by Rules R9 and R8) $\mathcal{R}[[t_1]\ldots[t_k]\mathbf{B}(\alpha \Rightarrow \beta)]^x_n$. □

Lemma 60. *Let α be a regressable sentence. Then $e, w \models \mathcal{R}[\alpha]$ iff $e_{\Sigma_{\mathrm{dyn}}}, w_{\Sigma_{\mathrm{dyn}}} \models \alpha$.*

Proof. By induction on $\|\alpha\|$. For the base case let $\|\alpha\| = 1$. For rigid R and $k \geq 0$, $w_{\Sigma_{\mathrm{dyn}}} \models [t_1]\ldots[t_k]R(t'_1,\ldots,t'_l)$ iff $w_{\Sigma_{\mathrm{dyn}}}[R(n'_1,\ldots,n'_l)] = 1$ where $n'_i = w_{\Sigma_{\mathrm{dyn}}}(t'_i)$ iff $w[R(n'_1,\ldots,n'_l)] = 1$ where $n'_i = w(t'_i)$ iff $w \models R(t'_1,\ldots,t'_l)$ iff (by Rule R1) $w \models \mathcal{R}[R(t'_1,\ldots,t'_l)]$; similarly for $[t_1]\ldots[t_k](t = t')$. For fluent $F \in \mathcal{F}$, $w_{\Sigma_{\mathrm{dyn}}} \models F(t'_1,\ldots,t'_l)$ iff $w_{\Sigma_{\mathrm{dyn}}}[F(n'_1,\ldots,n'_l),\langle\rangle] = 1$ where $n'_i = w_{\Sigma_{\mathrm{dyn}}}(t'_i)$ iff (by definition of $w_{\Sigma_{\mathrm{dyn}}}$) $w[F(n'_1,\ldots,n'_l),\langle\rangle] = 1$ where $n'_i = w(t'_i)$ iff $w \models F(t'_1,\ldots,t'_l)$ iff (by Rule R2) $w \models \mathcal{R}[F(t'_1,\ldots,t'_l)]$.

For the induction step let $\|\alpha\| = m > 1$ and suppose the lemma holds for all β with $\|\beta\| < m$. For fluent $F \in \mathcal{F}$ and $k \geq 1$, $w_{\Sigma_{\mathrm{dyn}}} \models [t_1]\ldots[t_k]F(t'_1,\ldots,t'_l)$ iff (by Rule S7) $(w_{\Sigma_{\mathrm{dyn}}} \gg n_1 \gg \ldots \gg n_k)[F(n'_1,\ldots,n'_l),\langle\rangle] = 1$ where $n_i = w_{\Sigma_{\mathrm{dyn}}}(t_i)$ and $n'_i = w_{\Sigma_{\mathrm{dyn}}}(t'_i)$ iff (by definition of $w_{\Sigma_{\mathrm{dyn}}}$ and Rule S7) $w_{\Sigma_{\mathrm{dyn}}} \models [t_1]\ldots[t_{k-1}]\gamma_F{}^{x_1 \ldots x_l\, a}_{t'_1 \ldots t'_l\, t_k}$ iff (by induction since $\|[t_1]\ldots[t_{k-1}]\gamma_F\| < m$) $w \models \mathcal{R}[[t_1]\ldots[t_{k-1}]\gamma_F{}^{x_1 \ldots x_l\, a}_{t'_1 \ldots t'_l\, t_k}]$ iff (by Rules R2 and R8) $w \models \mathcal{R}[[t_1]\ldots[t_k]F(t'_1,\ldots,t'_l)]$. Similarly for IF, $w_{\Sigma_{\mathrm{dyn}}} \models [t_1]\ldots[t_k]IF(t)$ iff (by Rule S7) $(w_{\Sigma_{\mathrm{dyn}}} \gg n_1 \gg \ldots \gg n_k)[IF(n),\langle\rangle] = 1$ where $n_i = w_{\Sigma_{\mathrm{dyn}}}(t_i)$ and $n = w_{\Sigma_{\mathrm{dyn}}}(t)$ iff (by definition of $w_{\Sigma_{\mathrm{dyn}}}$ and Rule S7) $w_{\Sigma_{\mathrm{dyn}}} \models [t_1]\ldots[t_k]\varphi^a_t$ iff (by induction since $\|[t_1]\ldots[t_k]\varphi\| < m$) $w \models \mathcal{R}[[t_1]\ldots[t_k]\varphi^a_t]$ iff (by Rules R3 and R8) $w \models \mathcal{R}[[t_1]\ldots[t_k]IF(t)]$.

For a quantifier, $e_{\Sigma_{\mathrm{dyn}}}, w_{\Sigma_{\mathrm{dyn}}} \models [t_1]\ldots[t_k]\exists x\,\alpha$ iff (by Rules S6 and S7) $e_{\Sigma_{\mathrm{dyn}}}, w_{\Sigma_{\mathrm{dyn}}} \models ([t_1]\ldots[t_k]\alpha)^x_n$ for some n iff (by induction since $\|[t_1]\ldots[t_k]\alpha\| < m$) $e, w \models \mathcal{R}[([t_1]\ldots[t_k]\alpha)^x_n]$ for some n iff (by Lemma 59) $e, w \models \mathcal{R}[[t_1]\ldots[t_k]\alpha]^x_n$

for some n iff (by Rule S6) $e, w \models \exists x \mathcal{R}[[t_1] \ldots [t_k]\alpha]$ iff (by Rules R7 and R8) $e, w \models \mathcal{R}[[t_1] \ldots [t_k]\exists x \alpha]$. We omit the similar induction steps for $[t_1] \ldots [t_k]\neg\alpha$ and $[t_1] \ldots [t_k](\alpha \vee \beta)$.

For beliefs after no actions, $e_{\Sigma_{\text{dyn}}} \models \mathbf{B}(\alpha \Rightarrow \beta)$ iff (by Theorem 13) $\lfloor e_{\Sigma_{\text{dyn}}} | \alpha \rfloor = \infty$ or $e_{\Sigma_{\text{dyn}}}, w \models (\alpha \supset \beta)$ for all $w \in (e_{\Sigma_{\text{dyn}}})_{\lfloor e_{\Sigma_{\text{dyn}}} | \alpha \rfloor}$ iff (by definition of $e_{\Sigma_{\text{dyn}}}$) $\lfloor e_{\Sigma_{\text{dyn}}} | \alpha \rfloor = \infty$ or $e_{\Sigma_{\text{dyn}}}, w_{\Sigma_{\text{dyn}}} \models (\alpha \supset \beta)$ for all $w \in e_{\lfloor e_{\Sigma_{\text{dyn}}} | \alpha \rfloor}$ iff (by induction since $\|(\alpha \supset \beta)\| < m$) $\lfloor e | \mathcal{R}[\alpha] \rfloor = \infty$ or $e, w \models \mathcal{R}[(\alpha \supset \beta)]$ for all $w \in e_{\lfloor e | \mathcal{R}[\alpha] \rfloor}$ iff (by Rules R5 and R6) $\lfloor e | \mathcal{R}[\alpha] \rfloor = \infty$ or $e, w \models (\mathcal{R}[\alpha] \supset \mathcal{R}[\beta])$ for all $w \in e_{\lfloor e | \mathcal{R}[\alpha] \rfloor}$ iff (by Theorem 13) $e \models \mathbf{B}(\mathcal{R}[\alpha] \Rightarrow \mathcal{R}[\beta])$ iff (by Rule R9) $e \models \mathcal{R}[\mathbf{B}(\alpha \Rightarrow \beta)]$. Finally, for beliefs after $k \geq 1$ actions, $e_{\Sigma_{\text{dyn}}}, w_{\Sigma_{\text{dyn}}} \models [t_1] \ldots [t_k]\mathbf{B}(\alpha \Rightarrow \beta)$ iff (by Rule S7 and, depending on the sort of t_k, Theorem 27 or 28, where σ is that theorem's right-hand side) $e_{\Sigma_{\text{dyn}}}, w_{\Sigma_{\text{dyn}}} \models [t_1] \ldots [t_{k-1}]\sigma^a_{n_k}$ where $n_k = w_{\Sigma_{\text{dyn}}}(t_k)$ iff (since by assumption action terms in formulas to be regressed only have variables or names as arguments) $e_{\Sigma_{\text{dyn}}}, w_{\Sigma_{\text{dyn}}} \models [t_1] \ldots [t_{k-1}]\sigma^a_{t_k}$ iff (by induction since $\|[t_1] \ldots [t_{k-1}]\sigma\| < m$) $e, w \models \mathcal{R}[[t_1] \ldots [t_{k-1}]\sigma^a_{t_k}]$ iff (by Rules R9 and R8) $e, w \models \mathcal{R}[[t_1] \ldots [t_k]\mathbf{B}(\alpha \Rightarrow \beta)]$. □

Theorem 24. *Let ϕ be a fluent sentence and ψ be an objective regressable sentence. Then $\Sigma_{\text{dyn}} \wedge \phi \models \psi$ iff $\phi \models \mathcal{R}[\psi]$.*

Proof. For the *only-if* direction suppose $\Sigma_{\text{dyn}} \wedge \phi \models \psi$ and $w \models \phi$. By Lemma 55 and the assumption, $w_{\Sigma_{\text{dyn}}} \models \psi$, and by Lemma 60, $w \models \mathcal{R}[\psi]$. Conversely, suppose $\phi \models \mathcal{R}[\psi]$ and $w \models \Sigma_{\text{dyn}} \wedge \phi$. Then $w \models \mathcal{R}[\psi]$ by assumption, and thus $w_{\Sigma_{\text{dyn}}} \models \psi$ by Lemma 60. By Lemma 54, $w_{\Sigma_{\text{dyn}}} = w$, so $w \models \psi$. □

Theorem 30. *Let α be a regressable sentence. Then $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models \alpha$ iff $\mathbf{O}\Sigma_{\text{bel}} \models \mathcal{R}[\alpha]$.*

Proof. For the *only-if* direction suppose $\mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models \alpha$ and $e \models \mathbf{O}\Sigma_{\text{bel}}$. By Lemma 56, $e_{\Sigma_{\text{dyn}}} \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$. By assumption, $e_{\Sigma_{\text{dyn}}} \models \alpha$. By Lemma 60, $e \models \mathcal{R}[\alpha]$. Conversely, suppose $\mathbf{O}\Sigma_{\text{bel}} \models \mathcal{R}[\alpha]$ and $e \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$. Let $e' \models \mathbf{O}\Sigma_{\text{bel}}$, which exists by Theorem 19. By assumption, $e' \models \mathcal{R}[\alpha]$. By Lemma 60, $e'_{\Sigma_{\text{dyn}}} \models \alpha$. By Lemma 56, $e'_{\Sigma_{\text{dyn}}} \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$. By Theorem 18, $e'_{\Sigma_{\text{dyn}}} = e$, so $e \models \alpha$. □

To generalize the regression result for the extended only-believing operator from Section 6.1, suppose $\mathcal{S}$ is a finite set of object function and predicate symbols and that Σ_{dyn} is $\mathcal{S}$-free from now on. It suffices to generalize Lemma 56, as the other lemmas involved do not refer to only-believing.

Lemma 61. *If $e \models \mathbf{O}_{\mathcal{S}}\Sigma_{\text{bel}}$, then $e_{\Sigma_{\text{dyn}}} \models \mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$.*

Proof. Let $e \models \mathbf{O}\Sigma_{\text{bel}}$. By Lemma 56, $e_{\Sigma_{\text{dyn}}} \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$. Then $e_{\mathcal{S}} \models \mathbf{O}_{\mathcal{S}}\Sigma_{\text{bel}}$, and $(e_{\Sigma_{\text{dyn}}})_{\mathcal{S}} \models \mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$. We need to show that $(e_{\mathcal{S}})_{\Sigma_{\text{dyn}}} = (e_{\Sigma_{\text{dyn}}})_{\mathcal{S}}$. Let $w \in ((e_{\mathcal{S}})_{\Sigma_{\text{dyn}}})_p$. Then for some $w' \in (e_{\mathcal{S}})_p$, $w'_{\Sigma_{\text{dyn}}} = w$. Then for some $w'' \in e_p$, $w'' \approx_{\mathcal{S}} w'$, and therefore $w''_{\Sigma_{\text{dyn}}} \approx_{\mathcal{S}} w'_{\Sigma_{\text{dyn}}} = w$. Since $w''_{\Sigma_{\text{dyn}}} \in (e_{\Sigma_{\text{dyn}}})_p$, $w \in ((e_{\Sigma_{\text{dyn}}})_{\mathcal{S}})_p$. Conversely, let $w \in ((e_{\Sigma_{\text{dyn}}})_{\mathcal{S}})_p$. Then for some $w' \in (e_{\Sigma_{\text{dyn}}})_p$, $w' \approx_{\mathcal{S}} w$. Then $w' \in e_p$, and so $w \in (e_{\mathcal{S}})_p$. By Lemma 53, $w' \models \Sigma_{\text{dyn}}$. Thus by Lemma 40 and since Σ_{dyn} is $\mathcal{S}$-free, $w \models \Sigma_{\text{dyn}}$. By Lemma 54, $w = w_{\Sigma_{\text{dyn}}}$, so $w \in ((e_{\mathcal{S}})_{\Sigma_{\text{dyn}}})_p$. □

Theorem 35. *Let α be a regressable sentence. Then $\mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models \alpha$ iff $\mathbf{O}_{\mathcal{S}}\Sigma_{\text{bel}} \models \mathcal{R}[\alpha]$.*

Proof. Proceeds by the exact same argument as the proof of Theorem 30, with $\mathbf{O}$ replaced by $\mathbf{O}_{\mathcal{S}}$, Lemma 56 replaced by Lemma 61, and Theorems 19 and 18 replaced by Corollary 34. □

Appendix B. Proof of belief projection by progression

In this appendix we first prove Theorem 44, which claims correctness of our strong revision of a conditional knowledge base. Then we proceed with the actual progression results, Theorems 46 and 47.

Definition 62. We say $*$ is a *symbol involution* when it maps object function symbols to object function symbols and predicate symbols to predicate symbols of corresponding arities, and $S = S^{**}$ for all object function or predicate symbols S. We denote by α^* the formula obtained from α by simultaneously replacing each object function or predicate symbol S with S^*. For a world w, we define the world w^* such that

- for all object function symbols g, $w^*[g^*(n_1, \ldots, n_k)] = w[g(n_1, \ldots, n_k)]$;

- for all rigid predicate symbols R and action sequences $z \neq \langle\rangle$,
  - $w^*[R^*(n_1, \ldots, n_k)] = w[R(n_1, \ldots, n_k)]$ if R^* is rigid;
  - $w^*[R^*(n_1, \ldots, n_k), \langle\rangle] = w[R(n_1, \ldots, n_k)]$ and $w^*[R^*(n_1, \ldots, n_k), z] = w[R^*(n_1, \ldots, n_k), z]$ if R^* is fluent;
- for all fluent predicate symbols F and action sequences z,
  - $w^*[F^*(n_1, \ldots, n_k)] = w[F(n_1, \ldots, n_k), \langle\rangle]$ if F^* is rigid;
  - $w^*[F^*(n_1, \ldots, n_k), z] = w[F(n_1, \ldots, n_k), z]$ if F^* is fluent.

For a set of worlds W and an epistemic state e, we let $W^* = \{w^* \mid w \in W\}$ and $e^* = \langle e^*_1, \ldots, e^*_{\lceil e \rceil}\rangle$.

We use symbol involutions to rename the symbols of a formula: when $\mathcal{S}'$ contains all symbols of α, and $*$ maps each of them to a new symbol from a set $\mathcal{S}''$ disjoint with $\mathcal{S}'$, then α^* is just the result of replacing every symbol from $\mathcal{S}'$ with the corresponding symbol from $\mathcal{S}''$.

Lemma 63. *Let $*$ be a symbol involution. Then $w = w^{**}$, $W = W^{**}$, and $e = e^{**}$.*

Proof. For $w = w^{**}$, the only non-trivial cases are when a fluent F is mapped to rigid F^* or the other way around. If R is rigid and R^* is fluent, then $w^{**}[R(n_1, \ldots, n_k)] = w^{**}[R^{**}(n_1, \ldots, n_k)] = w^*[R^*(n_1, \ldots, n_k), \langle\rangle] = w[R(n_1, \ldots, n_k)]$. If F is fluent and F^* is rigid, then $w^{**}[F(n_1, \ldots, n_k), \langle\rangle] = w^{**}[F^{**}(n_1, \ldots, n_k), \langle\rangle] = w^*[F^*(n_1, \ldots, n_k)] = w[F(n_1, \ldots, n_k), \langle\rangle]$, and $w^{**}[F(n_1, \ldots, n_k), z] = w^{**}[F^{**}(n_1, \ldots, n_k), z] = w^*[F^*(n_1, \ldots, n_k), z] = w[F(n_1, \ldots, n_k), z]$ for all $z \neq \langle\rangle$. Thus $w = w^{**}$. $W = W^{**}$ and $e = e^{**}$ then follow immediately from the definition. □

Lemma 64. *Let ϕ be objective static and let $*$ be a symbol involution. Then $w \models \phi$ iff $w^* \models \phi^*$.*

Proof. By induction on the length of ϕ. We show the base case only for fluent predicate symbols F with rigid image F^* (the other cases are analogous): $w \models F(t_1, \ldots, t_k)$ iff $w[F(n_1, \ldots, n_k), \langle\rangle] = 1$ where $n_i = w(t_i)$ iff $w^*[F^*(n_1, \ldots, n_k)] = 1$ where $n_i = w^*(t^*_i)$ iff $w^* \models (F(t_1, \ldots, t_k))^*$. The induction steps for $\neg\phi$, $(\phi \vee \beta)$, and $\exists x\, \phi$ are trivial. □

Lemma 65. *Let ϕ be objective static and let $*$ be a symbol involution. Then $\{w \mid w \models \phi\}^* = \{w \mid w \models \phi^*\}$.*

Proof. Let $w^* \in \{w' \mid w' \models \phi\}^*$. Then $w \models \phi$, and by Lemma 64, $w^* \models \phi^*$, so $w^* \in \{w' \mid w' \models \phi^*\}$. Conversely, let $w \in \{w' \mid w' \models \phi^*\}$. Then $w \models \phi^*$. By Lemma 64, $w^{**} \models \phi$, and by Lemma 63, $w = w^{**} \in \{w' \mid w' \models \phi\}^*$. □

Lemma 66. *Let $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ be objective and static, and let $*$ be a symbol involution. Then $e \models \mathbf{O}\Gamma$ iff $e^* \models \mathbf{O}\Gamma^*$.*

Proof. Let $e \models \mathbf{O}\Gamma$ and $e' \models \mathbf{O}\Gamma^*$. Then $e_p = \{w \mid w \models \bigwedge_{i:\lfloor e \mid \phi_i \rfloor \geq p}(\phi_i \supset \psi_i)\}$, and $e'_p = \{w' \mid w' \models \bigwedge_{i:\lfloor e' \mid \phi^*_i \rfloor \geq p}(\phi^*_i \supset \psi^*_i)\}$ by Rule S10. We show by induction on p that $e^*_p = e'_p$, and $\lfloor e \mid \phi_i \rfloor > p$ iff $\lfloor e' \mid \phi^*_i \rfloor > p$. For $p = 1$, this follows from Lemmas 65 and 64, respectively. For $p > 1$, by induction $\lfloor e \mid \phi_i \rfloor \geq p$ iff $\lfloor e' \mid \phi^*_i \rfloor \geq p$. Thus, just like in the base case, $e^*_p = e'_p$ by Lemma 65, and $\lfloor e \mid \phi_i \rfloor > p$ iff $\lfloor e' \mid \phi^*_i \rfloor > p$ by Lemma 64. Thus $e^* = e'$. The *only-if* direction of the lemma thus holds. Conversely, the *if* direction holds because $*$ is an involution: if $e^* \models \mathbf{O}\Gamma^*$ for some e, then $e^{**} \models \mathbf{O}\Gamma^{**}$ by the *only-if* direction, and since $e^{**} = e$ by Lemma 63 and $\mathbf{O}\Gamma^{**} = \mathbf{O}\Gamma$, $e \models \mathbf{O}\Gamma$. □

Lemma 67. *Let ϕ_1, ϕ_2 be objective sentences over object function and predicate symbols $\mathcal{S}_1, \mathcal{S}_2$ and let $\mathcal{S}_1, \mathcal{S}_2$ be disjoint. Suppose $w_1 \models \phi_1$, $w_2 \models \phi_2$. Then there is some w such that $w \models \phi_1 \wedge \phi_2$ and $w \approx_{\mathcal{S}_2} w_1$ and $w \approx_{\mathcal{S}_1} w_2$.*

Proof. Since $\mathcal{S}_1$ and $\mathcal{S}_2$ are disjoint, there clearly is a w with $w \approx_{\mathcal{S}_2} w_1$ and $w \approx_{\mathcal{S}_1} w_2$. By a simple induction on ϕ_1, $w \models \phi_1$ because ϕ_1 does not mention any symbol from $\mathcal{S}_2$. Analogously, $w \models \phi_2$. □

Theorem 44. *Let $\Gamma = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ and υ be objective and static. Let $\mathcal{S}$ be a set of object function and predicate symbols, and let υ be $\mathcal{S}$-free. Let $\mathcal{S}''$ be the symbols newly introduced in $\Gamma *_{\mathrm{s}} \upsilon$. If $e \models \mathbf{O}_{\mathcal{S}}\Gamma$, then $e *_{\mathrm{s}} \upsilon \models \mathbf{O}_{\mathcal{S}\cup\mathcal{S}''}\Gamma *_{\mathrm{s}} \upsilon$.*

Proof. Let $\mathcal{S}'$ be the object function and predicate symbols that occur in Γ and $*$ be the symbol involution that maps each symbol from $\mathcal{S}'$ to the corresponding symbol from $\mathcal{S}''$. Recall that $\Delta = \{\phi \Rightarrow \psi \in \Gamma_\upsilon \mid \mathbf{O}\Gamma \not\models \mathbf{K}(\phi \supset \psi)\}$. Then $\Gamma *_{\mathrm{s}} \upsilon = \Gamma'_\upsilon \cup \Gamma'_{\neg\upsilon} \cup \Delta' \cup \Lambda_1 \cup \Lambda_2 \cup \Lambda_3$, where

$$\begin{aligned}
\Gamma'_\upsilon &= \{\phi^* \Rightarrow \psi^* \mid \phi \Rightarrow \psi \in \Gamma_\upsilon\};\\
\Gamma'_{\neg\upsilon} &= \{(\phi^* \wedge \neg\upsilon) \Rightarrow \psi^* \mid \phi \Rightarrow \psi \in \Gamma_{\neg\upsilon}\};\\
\Delta' &= \{\textsc{true} \Rightarrow \upsilon\} \cup \{\neg(\phi^* \supset \psi^*) \vee \neg\upsilon \Rightarrow \upsilon \mid \phi \Rightarrow \psi \in \Delta\};\\
\Lambda'_1 &= \{\neg((\upsilon \wedge \neg\upsilon^*) \supset (\phi \supset \psi)) \Rightarrow \textsc{false} \mid \phi \Rightarrow \psi \in \Gamma_\upsilon\};\\
\Lambda'_2 &= \{\neg((\neg\upsilon \wedge \upsilon^*) \supset (\phi \supset \psi)) \Rightarrow \textsc{false} \mid \phi \Rightarrow \psi \in \Gamma_{\neg\upsilon}\};\\
\Lambda'_3 &= \{\neg((\upsilon \equiv \upsilon^*) \supset (\phi \equiv \phi^*) \wedge (\psi \equiv \psi^*)) \Rightarrow \textsc{false} \mid \phi \Rightarrow \psi \in \Gamma\}.
\end{aligned}$$

We show the theorem for the case $\mathcal{S} = \{\}$ first. Let $e \models \mathbf{O}\Gamma$. We show that $e *_{\mathrm{s}} \upsilon \models \mathbf{O}_{\mathcal{S}''}\Gamma *_{\mathrm{s}} \upsilon$.

First suppose $\lfloor e \mid \upsilon \rfloor = \infty$. Then $e *_{\mathrm{s}} \upsilon = \langle\{\}\rangle$. Clearly, $\lfloor \langle\{\}\rangle \mid \phi \rfloor = \infty$ for all ϕ, so $\langle\{\}\rangle \models \mathbf{O}\Gamma *_{\mathrm{s}} \upsilon$ if (by Rule S10) $\bigwedge_{\phi\Rightarrow\psi\in\Gamma_\upsilon}(\phi^* \supset \psi^*) \wedge \bigwedge_{\phi\Rightarrow\psi\in\Gamma_\upsilon} \neg(((\upsilon \wedge \neg\upsilon^*) \supset (\phi \supset \psi)) \supset \textsc{false}) \wedge \upsilon$ is unsatisfiable. By Rule S10 and $\lfloor e \mid \upsilon \rfloor = \infty$, $\bigwedge_{i:\lfloor e \mid \phi\rfloor=\infty}(\phi \supset \psi) \wedge \upsilon$ is unsatisfiable, and by Lemma 39, $\bigwedge_{\phi\Rightarrow\psi\in\Gamma_\upsilon}(\phi \supset \psi) \wedge \upsilon$ is unsatisfiable, too, and by Lemma 64, $\bigwedge_{\phi\Rightarrow\psi\in\Gamma_\upsilon}(\phi^* \supset \psi^*) \models \neg\upsilon^*$. Hence, the formula mentioned before is indeed inconsistent.

Now suppose $\lfloor e \mid \upsilon \rfloor \neq \infty$. To show $e *_{\mathrm{s}} \upsilon \models \mathbf{O}_{\mathcal{S}''}\Gamma *_{\mathrm{s}} \upsilon$ we construct a model of $\mathbf{O}\Gamma *_{\mathrm{s}} \upsilon$ and show that forgetting $\mathcal{S}''$ in this epistemic state yields $e *_{\mathrm{s}} \upsilon$. The proof proceeds in three steps. Step 1 is to show $e' \models \mathbf{O}\Gamma'_\upsilon \cup \Gamma'_{\neg\upsilon} \cup \Delta'$ where

$$e' = \langle (e^* \mid \upsilon)_{\lfloor e \mid \upsilon \rfloor}, \ldots, (e^* \mid \upsilon)_{\lceil e \rceil}, (e^* \mid \neg\upsilon)_{\lfloor e \mid \neg\upsilon \rfloor} \cup (e^* \mid \upsilon)_{\lceil e \rceil}, \ldots, (e^* \mid \neg\upsilon)_{\lceil e \rceil} \cup (e^* \mid \upsilon)_{\lceil e \rceil} \rangle.$$

Step 2 is to show $(e' \mid \lambda) \models \mathbf{O}\Gamma *_{\mathrm{s}} \upsilon$, where $\lambda = \bigwedge_{\phi\Rightarrow\psi\in\Lambda'_1\cup\Lambda'_2\cup\Lambda'_3}(\phi \supset \psi)$. Step 3 is to prove $(e' \mid \lambda)_{\mathcal{S}''} = e *_{\mathrm{s}} \upsilon$.

Step 1. We now prove that $e' \models \mathbf{O}\Pi$ where $\Pi = \Gamma'_\upsilon \cup \Gamma'_{\neg\upsilon} \cup \Delta'$ for the following plausibilities of the conditionals in Π:

- $\lfloor e' \mid \phi^* \rfloor = \max\{\lfloor e \mid \upsilon \rfloor, \lfloor e \mid \phi \rfloor\} - \lfloor e \mid \upsilon \rfloor + 1$ for all $\phi \Rightarrow \psi \in \Gamma_\upsilon$; because for all $p \in \mathbb{N}$ with $p \geq \lfloor e \mid \upsilon \rfloor$ we have $p \geq \lfloor e \mid \phi \rfloor$ iff $w \models \phi$ for some $w \in e_p$ iff (by Lemma 64) $w \models \phi^*$ for some $w \in e^*_p$ iff (by Rule S10 and Lemma 67) $w \models \phi^* \wedge \upsilon$ for some $w \in e^*_p$ iff $w \models \phi^*$ for some $w \in e'_{p-\lfloor e \mid \upsilon \rfloor+1}$ iff $p - \lfloor e \mid \upsilon \rfloor + 1 \geq \lfloor e' \mid \phi^* \rfloor$.
- $\lfloor e' \mid \phi^* \wedge \neg\upsilon \rfloor = \max\{\lfloor e \mid \neg\upsilon \rfloor, \lfloor e \mid \phi \rfloor\} + \lceil e \rceil - \lfloor e \mid \upsilon \rfloor - \lfloor e \mid \neg\upsilon \rfloor + 2$ for all $\phi \Rightarrow \psi \in \Gamma_{\neg\upsilon}$; because, very similarly to the above, for all $p \in \mathbb{N}$ with $p \geq \lfloor e \mid \neg\upsilon \rfloor$ we have $p \geq \lfloor e \mid \phi \rfloor$ iff $w \models \phi$ for some $w \in e_p$ iff (by Lemma 64) $w \models \phi^*$ for some $w \in e^*_p$ iff (by Rule S10 and Lemma 67) $w \models \phi^* \wedge \neg\upsilon$ for some $w \in e^*_p$ iff (since $w \models \upsilon$ for all $w \in e'_{p'}$ and $p' \leq \lceil e \rceil - \lfloor e \mid \upsilon \rfloor + 1$) $w \models \phi^* \wedge \neg\upsilon$ for some $w \in e'_{p-\lfloor e \mid \neg\upsilon \rfloor+1+\lceil e \rceil-\lfloor e \mid \upsilon \rfloor+1}$ iff $p + \lceil e \rceil - \lfloor e \mid \upsilon \rfloor - \lfloor e \mid \neg\upsilon \rfloor + 2 \geq \lfloor e' \mid \phi^* \wedge \neg\upsilon \rfloor$.
- $\max(\{\lfloor e' \mid \textsc{true} \rfloor\} \cup \{\lfloor e' \mid \neg(\phi^* \supset \psi^*) \vee \neg\upsilon \rfloor \mid \phi \Rightarrow \psi \in \Delta\}) = \lceil e \rceil - \lfloor e \mid \upsilon \rfloor + 1$; for the following reason. If $\Delta = \{\}$, then $\lfloor e \mid \upsilon \rfloor = \lceil e \rceil$, and since $\lfloor e \mid \upsilon \rfloor \neq \infty$, $\lfloor e' \mid \textsc{true} \rfloor = 1$. Now suppose $\phi \Rightarrow \psi \in \Delta$. Then $\lfloor e \mid \neg(\phi \supset \psi) \rfloor \neq \infty$, and by Lemma 38, $\lfloor e \mid \neg(\phi \supset \psi) \rfloor \leq \lceil e \rceil$. Moreover, $w \models (\phi \supset \psi)$ for all $w \in e_{\lfloor e \mid \phi \rfloor}$, so $\lfloor e \mid \neg(\phi \supset \psi) \rfloor \geq \lfloor e \mid \phi \rfloor + 1$. In particular, there is some $\phi \Rightarrow \psi \in \Delta$ with $\lfloor e \mid \phi \rfloor = \lceil e \rceil - 1$, for otherwise $e_{\lceil e \rceil - 1} = e_{\lceil e \rceil}$ by Rule S10; hence $\lfloor e \mid \neg(\phi \supset \psi) \rfloor = \lceil e \rceil$. By Lemma 64 and since $w \not\models \neg\upsilon$ for all $w \in e'_p$ and $p \leq \lceil e \rceil - \lfloor e \mid \upsilon \rfloor + 1$, the equality follows.

First consider $p \leq \lceil e \rceil - \lfloor e \mid \upsilon \rfloor + 1$. Then $w \in e'_p$ iff $w \in (e^* \mid \upsilon)_{p+\lfloor e \mid \upsilon \rfloor-1}$ iff (by Rule S10 and Lemma 64) $w \models \bigwedge_{i:\lfloor e \mid \phi_i \rfloor \geq p+\lfloor e \mid \upsilon \rfloor-1}(\phi^*_i \supset \psi^*_i) \wedge \upsilon$ iff (by Lemma 39) $w \models \bigwedge_{\phi\Rightarrow\psi\in\Gamma_\upsilon \text{ with } \max\{\lfloor e \mid \upsilon \rfloor, \lfloor e \mid \phi \rfloor\} \geq p+\lfloor e \mid \upsilon \rfloor-1}(\phi^* \supset \psi^*) \wedge \upsilon$ iff (since $\models \upsilon \supset (\phi^* \wedge \neg\upsilon \supset \psi^*)$ as well as $\models \upsilon \equiv ((\neg(\phi^* \supset \psi^*) \vee \neg\upsilon) \supset \upsilon)$ and by the above plausibilities)

$$w \models \bigwedge_{\substack{\phi\Rightarrow\psi\in\Gamma_\upsilon\\ \lfloor e' \mid \phi^* \rfloor \geq p}}(\phi^* \supset \psi^*) \wedge \bigwedge_{\substack{\phi\Rightarrow\psi\in\Gamma_{\neg\upsilon}\\ \lfloor e' \mid \phi^* \wedge \neg\upsilon \rfloor \geq p}}(\phi^* \wedge \neg\upsilon \supset \psi^*) \wedge \bigwedge_{\substack{\phi\Rightarrow\psi\in\Delta\\ \lfloor e' \mid \neg(\phi^*\supset\psi^*)\vee\neg\upsilon \rfloor \geq p}}((\neg(\phi^* \supset \psi^*) \vee \neg\upsilon) \supset \upsilon) \wedge \bigwedge_{\lfloor e' \mid \textsc{true} \rfloor \geq p}(\textsc{true} \supset \upsilon)$$

iff $w \models \bigwedge_{\phi\Rightarrow\psi\in\Pi \text{ with } \lfloor e' \mid \phi \rfloor \geq p}(\phi \supset \psi)$.

Now consider $p > \lceil e \rceil - \lfloor e \mid \upsilon \rfloor + 1$. Then $w \in e'_p$ iff $w \in (e^* \mid \upsilon)_{\lceil e \rceil}$ or $w \in (e^* \mid \neg\upsilon)_{p-\lceil e \rceil+\lfloor e \mid \upsilon \rfloor+\lfloor e \mid \neg\upsilon \rfloor-2}$ iff (by Lemma 64) $w \models \bigwedge_{i:\lfloor e \mid \phi_i \rfloor=\infty}(\phi^*_i \supset \psi^*_i) \wedge \upsilon$ or $w \models \bigwedge_{\lfloor e \mid \phi_i \rfloor \geq p-\lceil e \rceil+\lfloor e \mid \upsilon \rfloor+\lfloor e \mid \neg\upsilon \rfloor-2}(\phi^*_i \supset \psi^*_i) \wedge \neg\upsilon$ iff (by Lemma 39 and by the above plausibilities)

$$w \models \bigwedge_{\substack{\phi\Rightarrow\psi\in\Gamma_{\neg\upsilon}\\ \lfloor e' \mid \phi^* \rfloor=\infty}}(\phi^* \supset \psi^*) \wedge \upsilon \text{ or } w \models \bigwedge_{\substack{\phi\Rightarrow\psi\in\Gamma_{\neg\upsilon}\\ \lfloor e' \mid \phi^* \wedge \neg\upsilon \rfloor \geq p}}(\phi^* \supset \psi^*) \wedge \neg\upsilon \quad \text{iff} \quad w \models \bigwedge_{\substack{\phi\Rightarrow\psi\in\Gamma_{\neg\upsilon}\\ \lfloor e' \mid \phi^* \rfloor=\infty}}(\phi^* \supset \psi^*) \wedge \bigwedge_{\substack{\phi\Rightarrow\psi\in\Gamma_{\neg\upsilon}\\ \lfloor e' \mid \phi^* \wedge \neg\upsilon \rfloor \geq p}}(\phi^* \wedge \neg\upsilon \supset \psi^*)$$

iff $w \models \bigwedge_{\phi \Rightarrow \psi \in \Pi \text{ with } \lfloor e' \mid \phi \rfloor \geq p}(\phi \supset \psi)$. Therefore $e \models \mathbf{O}\Pi$ and Step 1 is completed.

Step 2. The second step of the proof is to show $(e' \mid \lambda) \models \mathbf{O}\Gamma *_s \upsilon$. The conditionals added in $\Gamma *_s \upsilon$ over Π are those from $\Lambda'_1 \cup \Lambda'_2 \cup \Lambda'_3$, which simply assert knowledge of λ as they have unsatisfiable consequents. Having shown $e' \models \mathbf{O}\Pi$ in Step 1, it is immediate from Rule S10 that $(e' \mid \lambda) \models \mathbf{O}\Gamma *_s \upsilon$ if $\lfloor e' \mid \tau \rfloor = \lfloor (e' \mid \lambda) \mid \tau \rfloor$ for the antecedents $\tau \in \{\phi^*, \phi^* \wedge \neg\upsilon, (\phi^* \wedge \psi^*) \vee \neg\upsilon\}$ of the conditionals in Π. Clearly, $\lfloor e' \mid \tau \rfloor \leq \lfloor (e' \mid \lambda) \mid \tau \rfloor$. To show the converse, we let $w \in e'_p$ be arbitrary and show that there is some $w' \in (e' \mid \lambda)_p$ such that $w' \models \tau$ iff $w \models \tau$. First suppose $w \models \neg\upsilon \wedge \upsilon^*$. Since $w \models \neg\upsilon$, by definition of e', $\lfloor e \mid \neg\upsilon \rfloor \neq \infty$. Hence $\bigwedge_{\phi \Rightarrow \psi \in \Gamma_{\neg\upsilon}}(\phi \supset \psi) \wedge \neg\upsilon$ is satisfiable by Lemma 39. By Lemma 67 there is some w' with $w' \approx_{\mathcal{S}'} w$ with $w' \models \bigwedge_{\phi \Rightarrow \psi \in \Gamma_{\neg\upsilon}}((\neg\upsilon \wedge \upsilon^*) \supset (\phi \supset \psi)) \wedge \neg\upsilon \wedge \upsilon^*$, so $w' \models \lambda$. Since $w' \models \upsilon$ iff $w \models \upsilon$ and by Lemma 40 we have that for all $\phi \Rightarrow \psi \in \Pi$, $w' \models \phi$ iff $w \models \phi$ as well as $w' \models \psi$ iff $w \models \psi$, so $w' \in e'_p$ by Rule S10. Therefore $w' \in (e' \mid \lambda)_p$, and $w' \models \tau$ iff $w \models \tau$. The case for $w \models \upsilon \wedge \neg\upsilon^*$ is analogous. Now suppose $w \models \upsilon \equiv \upsilon^*$. By Lemma 64, $\bigwedge_{\phi \Rightarrow \psi \in \Gamma}(\pm\phi \wedge \pm\psi) \wedge \pm\upsilon$ is satisfiable where $\pm\beta$ stands for β if $w \models \beta^*$ and for $\neg\beta$ otherwise. By Lemma 67 there is some w' with $w \approx_{\mathcal{S}'} w'$ with $w' \models \bigwedge_{\phi \Rightarrow \psi \in \Gamma}(\pm\phi \wedge \pm\psi) \wedge \pm\upsilon$. Since $w' \models \upsilon$ iff $w \models \upsilon$ and by Lemma 40 we have that for all $\phi \Rightarrow \psi \in \Pi$, $w' \models \phi$ iff $w \models \phi$ as well as $w' \models \psi$ iff $w \models \psi$, so $w' \in e'_p$ by Rule S10. Therefore $w' \in (e' \mid \lambda)_p$, and $w' \models \tau$ iff $w \models \tau$.

Step 3. Lastly, we need to show that $(e' \mid \lambda)_{\mathcal{S}''} = e *_s \upsilon$. Let $w \in ((e' \mid \lambda)_{\mathcal{S}''})_p$. Then there is some $w' \in (e' \mid \lambda)_p$ with $w \approx_{\mathcal{S}''} w'$. Thus $w' \models \bigwedge_{\phi \Rightarrow \psi \in \Gamma *_s \upsilon \text{ with } \lfloor e' \mid \phi \rfloor \geq p}(\phi \supset \psi)$. First suppose $p \leq \lceil e \rceil - \lfloor e \mid \upsilon \rfloor + 1$. Then $w \models \upsilon$. If $w' \models \upsilon \equiv \upsilon^*$, then by Lemma 40 and Λ_3, $w \models \bigwedge_{i : \lfloor e \mid \phi_i \rfloor \geq p + \lfloor e \mid \upsilon \rfloor - 1}(\phi_i \supset \psi_i) \wedge \upsilon$, and thus $w \in (e \mid \upsilon)_{p + \lfloor e \mid \upsilon \rfloor - 1} = (e *_s \upsilon)_p$. Otherwise, $w' \models \upsilon \wedge \neg\upsilon^*$, and then by Lemma 40 and Λ_1, $w \models \bigwedge_{i : \lfloor e \mid \phi_i \rfloor \geq \lfloor e \mid \upsilon \rfloor}(\phi_i \supset \psi_i) \wedge \upsilon$, and thus $w \in (e \mid \upsilon)_{\lfloor e \mid \upsilon \rfloor} \subseteq (e \mid \upsilon)_{p + \lfloor e \mid \upsilon \rfloor - 1} = (e *_s \upsilon)_p$. Now suppose $p > \lceil e \rceil - \lfloor e \mid \upsilon \rfloor + 1$. The cases for $w' \models \upsilon \equiv \upsilon^*$ and $w' \models \upsilon \wedge \neg\upsilon^*$ are analogous. If $w' \models \neg\upsilon \wedge \upsilon^*$, then by Lemma 40 and Λ_2, $w \models \bigwedge_{i : \lfloor e \mid \phi_i \rfloor \geq \lfloor e \mid \neg\upsilon \rfloor}(\phi_i \supset \psi_i) \wedge \neg\upsilon$, and thus $w \in (e \mid \neg\upsilon)_{\lfloor e \mid \neg\upsilon \rfloor} \subseteq (e \mid \neg\upsilon)_{p - \lceil e \rceil + \lfloor e \mid \upsilon \rfloor + \lfloor e \mid \neg\upsilon \rfloor - 2} \subseteq (e *_s \upsilon)_p$. For the converse direction, let $w' \in (e *_s \upsilon)_p$. Since $*$ swaps the (initial) values of $\mathcal{S}'$ and $\mathcal{S}''$, there clearly is a w such that $w \approx_{\mathcal{S}''} w'$ and $w^* = w$. Thus $w \models \upsilon$ iff $w^* \models \upsilon$, and therefore $w \in e'_p$. For all $\phi \Rightarrow \psi \in \Gamma$, $w \models \phi$ iff (by Lemma 64) $w^* \models \phi^*$ iff $w \models \phi^*$, and likewise for ψ. Thus $w \models \lambda$, and hence $w \in (e' \mid \lambda)_p$. As $w \approx_{\mathcal{S}''} w'$, we have $w' \in ((e' \mid \lambda)_{\mathcal{S}''})_p$.

Now let $\mathcal{S} \neq \{\}$. Let $e \models \mathbf{O}_{\mathcal{S}}\Gamma$ and $e' \models \mathbf{O}\Gamma$. By Rule S11 and Corollary 34, $e = e'_{\mathcal{S}}$. By the case for $\mathcal{S} = \{\}$, $e' *_s \upsilon \models \mathbf{O}_{\mathcal{S}''}\Gamma *_s \upsilon$. By Rule S11, $(e' *_s \upsilon)_{\mathcal{S}} \models \mathbf{O}_{\mathcal{S} \cup \mathcal{S}''}\Gamma *_s \upsilon$. By Lemma 41, $e *_s \upsilon \models \mathbf{O}_{\mathcal{S} \cup \mathcal{S}''}\Gamma *_s \upsilon$. □

In the rest of this appendix we prove the correctness of progression, that is, Theorems 46 and 47. We use the following assumptions throughout the rest of the appendix.

- Let $\Sigma_{\text{bel}}, \Sigma_{\text{dyn}}$ be a basic action theory over fluents $\mathcal{F} = \{F_1, \ldots, F_l\}$, and let n be an action standard name. Recall that Σ_{dyn} contains the successor state axioms $\square[a]F(x_1, \ldots, x_k) \equiv \gamma_F$ for $F \in \mathcal{F}$, and the informed-fluent axiom $\square IF(a) \equiv \varphi$.
- Let $\mathcal{S}'$ be the symbols newly introduced in $\Sigma_{\text{bel}} \gg n$, which is partitioned into two subsets: $\mathcal{R} = \{R_1, \ldots, R_l\} \subseteq \mathcal{S}'$ contains the rigid predicates for the physical progression as in Definition 45; $\mathcal{S}' \setminus \mathcal{R}$ contains the rigid symbols introduced by the revision as in Definition 37 or 43.
- Let $*$ be the symbol involution that maps F_i to R_i and leaves the rest unchanged.

Definition 68. For a world w and an action n, w^n is a world such that $w^n \approx_{\mathcal{F}} (w \gg n)$ and

- $w^n[F(n_1, \ldots, n_k), \langle\rangle] = 1$ iff $w^n \models (\gamma_F{}^{x_1 \ldots x_k\, a}_{n_1 \ldots n_k\, n})^*$ for all $F \in \mathcal{F}$;
- $w^n[F(n_1, \ldots, n_k), z] = w[F(n_1, \ldots, n_k), z]$ for all $F \in \mathcal{F}$ and action sequences $z \neq \langle\rangle$.

For a set of worlds W and an epistemic state e, we let $W^n = \{w^n \mid w \in W\}$ and $e^n = \langle e^n_1, \ldots, e^n_{\lceil e \rceil}\rangle$.

Intuitively, w^n sets the initial values of every fluent $F \in \mathcal{F}$ to the its value after n when the values before n are memorized in $\mathcal{R}$: $w^n \models F(n_1, \ldots, n_k)$ iff $w^n \models (\gamma_F{}^{x_1 \ldots x_k\, a}_{n_1 \ldots n_k\, n})^*$, where $*$ replaces all fluents in γ_F with the corresponding rigid predicates from $\mathcal{R}$.

Lemma 69. *w^n is uniquely defined.*

Proof. Let $w' \approx_{\mathcal{F}} (w \gg n)$ be such that for every $F \in \mathcal{F}$, $w'[F(n_1, \ldots, n_k), \langle\rangle] = 1$ iff $w \gg n \models (\gamma_F{}^{x_1 \ldots x_k\, a}_{n_1 \ldots n_k\, n})^*$, and moreover $w'[F(n_1, \ldots, n_k), z] = w[F(n_1, \ldots, n_k), z]$ for all $z \neq \langle\rangle$. Clearly such a w' exists and is uniquely defined. By Lemma 40 and since γ^*_F is $\mathcal{F}$-free, $w' \models (\gamma_F{}^{x_1 \ldots x_k\, a}_{n_1 \ldots n_k\, n})^*$ iff $w \gg n \models (\gamma_F{}^{x_1 \ldots x_k\, a}_{n_1 \ldots n_k\, n})^*$. Thus $w' = w^n$. □

Lemma 70. *Let ϕ be fluent. Then $w \models \phi^*$ iff $w^n \models \phi^*$.*

Proof. Since ϕ is fluent and by definition of $*$, ϕ^* mentions only rigid predicates. By Lemma 69, w^n is uniquely defined, and since w and w^n agree on all rigids, a simple induction on the length of ϕ shows that the lemma holds. □

Lemma 71. *If $e \models \mathbf{O}(\Sigma_{\text{bel}} * \varphi_n^a)^*$, then $e^n \models \mathbf{O}\Sigma_{\text{bel}} \gg n$.*

Proof. Let $(\Sigma_{\text{bel}} * \varphi_n^a)^* = \{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ and $e \models \mathbf{O}(\Sigma_{\text{bel}} * \varphi_n^a)^*$. We show that $e^n \models \mathbf{O}\Sigma_{\text{bel}} \gg n$ for the same plausibilities and plausibility ∞ for the added conditionals. We first show that $\lfloor e \,|\, \phi_i \rfloor = \lfloor e^n \,|\, \phi_i \rfloor$ for all i and $\lfloor e^n \,|\, \phi \rfloor = \infty$ for the newly added conditionals $\phi \Rightarrow \psi \in (\Sigma_{\text{bel}} \gg n) \setminus (\Sigma_{\text{bel}} * \varphi_n^a)$. If $p < \lfloor e \,|\, \phi_i \rfloor$, then $w \not\models \phi_i$ for all $w \in e_p$, and by Lemma 70, $w \not\models \phi_i$ for all $w \in e_p^n$, so $p < \lfloor e^n \,|\, \phi_i \rfloor$. Analogously, if $p \geq \lfloor e \,|\, \phi_i \rfloor$, then $p \geq \lfloor e^n \,|\, \phi_i \rfloor$. Thus $\lfloor e \,|\, \phi_i \rfloor = \lfloor e \,|\, \phi_i \rfloor$. All other conditionals from $\Sigma_{\text{bel}} \gg n$ are of the form $\neg(F(x_1, \ldots, x_k) \equiv (\gamma_F{}^{x_1 \dots x_k\, a}_{n_1 \dots n_k\, n})^*) \Rightarrow \text{FALSE}$ for some $F \in \mathcal{F}$, and by definition of w^n, $\lfloor e^n \,|\, \neg(F(x_1, \ldots, x_k) \equiv (\gamma_F{}^{x_1 \dots x_k\, a}_{n_1 \dots n_k\, n})^*) \rfloor = \infty$.

Now we prove $w \in e_p^n$ iff $w \models \bigwedge_{i:\lfloor e^n | \phi_i \rfloor \geq p} (\phi_i \supset \psi_i) \wedge \bigwedge_{F \in \mathcal{F}} (\neg(F(x_1, \ldots, x_k) \equiv (\gamma_F{}^{x_1 \dots x_k\, a}_{n_1 \dots n_k\, n})^*) \supset \text{FALSE})$, that is, e^n satisfies Rule S10. For the *only-if* direction suppose $w' \in e_p^n$. Then for some $w \in e_p$, $w^n = w'$. By Rule S10 for e we have $w \models \bigwedge_{i:\lfloor e | \phi_i \rfloor \geq p} (\phi_i \supset \psi_i)$. Since $w' = w^n$ and by $\lfloor e \,|\, \phi_i \rfloor = \lfloor e^n \,|\, \phi_i \rfloor$ and Lemma 70, $w' \models \bigwedge_{i:\lfloor e^n | \phi_i \rfloor \geq p} (\phi_i \supset \psi_i)$, and by of w^n also $w' \models F(x_1, \ldots, x_k) \equiv (\gamma_F{}^{x_1 \dots x_k\, a}_{n_1 \dots n_k\, n})^*$ for every $F \in \mathcal{F}$. Thus the right-hand side holds. Conversely, suppose $w' \notin e_p^n$. If there is some w with $w^n = w'$, then $w \notin e_p$, and hence by $\lfloor e \,|\, \phi_i \rfloor = \lfloor e^n \,|\, \phi_i \rfloor$, $w' \not\models \bigwedge_{i:\lfloor e^p | \phi_i \rfloor \geq p} (\phi_i \supset \psi_i)$. Otherwise, for all w, $w^n \neq w'$, and hence $w' \not\models F(x_1, \ldots, x_k) \equiv (\gamma_F{}^{x_1 \dots x_k\, a}_{n_1 \dots n_k\, n})^*$ for some $F \in \mathcal{F}$. In either case the right-hand side is false. □

Lemma 72. $(w_{\Sigma_{\text{dyn}}} \gg n) \approx_{\mathcal{R}} ((w^*)^n)_{\Sigma_{\text{dyn}}}$.

Proof. The cases for object function symbols and rigid predicate symbols $R \notin \mathcal{R}$ are trivial as they are left unchanged by the involved Definitions 9, 51, 62, 68. For $F \in \mathcal{F}$ and $z = \langle\rangle$, $(w_{\Sigma_{\text{dyn}}} \gg n)[F(n_1, \ldots, n_k), \langle\rangle] = 1$ iff $w_{\Sigma_{\text{dyn}}}[F(n_1, \ldots, n_k), n] = 1$ iff $w_{\Sigma_{\text{dyn}}} \models \gamma_F{}^{x_1 \dots x_k\, a}_{n_1 \dots n_k\, n}$ iff (by Lemma 55) $w \models \gamma_F{}^{x_1 \dots x_k\, a}_{n_1 \dots n_k\, n}$ iff (by Lemma 64) $w^* \models (\gamma_F{}^{x_1 \dots x_k\, a}_{n_1 \dots n_k\, n})^*$ iff $(w^*)^n[F(n_1, \ldots, n_k), \langle\rangle] = 1$ iff $((w^*)^n)_{\Sigma_{\text{dyn}}}[F(n_1, \ldots, n_k), \langle\rangle] = 1$. The cases for $F \in \mathcal{F}$ with $z \neq \langle\rangle$ and for IF follow from the definition of $w_{\Sigma_{\text{dyn}}}$ and φ, γ_F being $\mathcal{R}$-free and fluent. Finally, $F \notin \mathcal{F} \cup \{IF\}$ follows because both $w \gg n$ and w^n progress F by n. □

Lemma 73. *Let $e \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} * \varphi_n^a)$ and $e' \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n)$. Then for all $p \in \mathbb{N}$, $(e_p \gg n) = (e'_{\mathcal{R}})_p$.*

Proof. Let $e'' \models \mathbf{O}\Sigma_{\text{bel}} * \varphi_n^a$, which exists by Theorem 19. By Lemma 56, $e''_{\Sigma_{\text{dyn}}} \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} * \varphi_n^a)$, and by Theorem 18, $e = e''_{\Sigma_{\text{dyn}}}$ (*). By Lemma 66, $e''^* \models \mathbf{O}(\Sigma_{\text{bel}} * \varphi_n^a)^*$, and by Lemma 71, $(e''^*)^n \models \mathbf{O}\Sigma_{\text{bel}} \gg n$, and by Lemma 56, $((e''^*)^n)_{\Sigma_{\text{dyn}}} \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n)$, and by Theorem 18, $((e''^*)^n)_{\Sigma_{\text{dyn}}} = e'$ (**).

For the $\subseteq$ direction let $w' \in (e_p \gg n)$. By (*) there is some $w \in e''_p$ such that $(w_{\Sigma_{\text{dyn}}} \gg n) = w'$. Also, by (**), $((w^*)^n)_{\Sigma_{\text{dyn}}} \in e'_p$. By Lemma 72, $(w_{\Sigma_{\text{dyn}}} \gg n) \approx_{\mathcal{R}} ((w^*)^n)_{\Sigma_{\text{dyn}}}$. Thus $w' = (w_{\Sigma_{\text{dyn}}} \gg n) \in (e'_{\mathcal{R}})_p$, so $(e_p \gg n) \subseteq (e'_{\mathcal{R}})_p$.

Conversely, let $w \in (e'_{\mathcal{R}})_p$. Then $w \approx_{\mathcal{R}} w'$ for some $w' \in e'_p$. By (**) there is some $w'' \in e''_p$ such that $((w''^*)^n)_{\Sigma_{\text{dyn}}} = w'$. By Lemma 72, $(w''_{\Sigma_{\text{dyn}}} \gg n) \approx_{\mathcal{R}} w'$, and thus $(w''_{\Sigma_{\text{dyn}}} \gg n) \approx_{\mathcal{R}} w$. Hence $w \in ((e''_p)_{\Sigma_{\text{dyn}}} \gg n)_{\mathcal{R}} = (e_p \gg n)_{\mathcal{R}}$ with (*). Thus $(e'_{\mathcal{R}})_p \subseteq (e_p \gg n)_{\mathcal{R}}$. Since $\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} * \varphi_n^a$ are $\mathcal{R}$-free and by Lemma 40, $(e_p)_{\mathcal{R}} = e_p$, and so $(e_p \gg n)_{\mathcal{R}} = (e_p \gg n)$. Thus $(e'_{\mathcal{R}})_p \subseteq (e_p \gg n)$. □

Lemma 74. *If $e \models \mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$, then $e * IF(n) \models \mathbf{O}_{\mathcal{S} \cup (\mathcal{S}' \setminus \mathcal{R})}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} * \varphi_n^a)$.*

Proof. Since $w \models \Box IF(n) \equiv \varphi_n^a$ for all $w \in e_p$ and $p \in \mathbb{N}$, we have $e * IF(n) = e * \varphi_n^a$. Hence by Theorems 42 and 44, $e * IF(n) \models \mathbf{O}_{\mathcal{S} \cup (\mathcal{S}' \setminus \mathcal{R})}((\{\neg\Sigma_{\text{dyn}} \Rightarrow \text{FALSE}\} \cup \Sigma_{\text{bel}}) * \varphi_n^a)$. Let $\Theta = ((\{\neg\Sigma_{\text{dyn}} \Rightarrow \text{FALSE}\} \cup \Sigma_{\text{bel}}) * \varphi_n^a) \setminus (\Sigma_{\text{bel}} * \varphi_n^a)$. It is easy to see that $\bigwedge_{\phi \Rightarrow \psi \in \Theta} (\phi \supset \psi)$ is equivalent to $\neg\Sigma_{\text{dyn}} \supset \text{FALSE}$. Therefore $e * IF(n) \models \mathbf{O}_{\mathcal{S} \cup (\mathcal{S}' \setminus \mathcal{R})}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} * \varphi_n^a)$. □

Theorem 46. $\models \mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \supset [n]\mathbf{O}_{\mathcal{S} \cup \mathcal{S}'}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n)$.

Proof. Suppose $e \models \mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$. Let $e' \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} * \varphi_n^a)$ and $e'' \models \mathbf{O}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n)$, which exist by Theorem 19. By Lemma 74, $e * IF(n) \models \mathbf{O}_{\mathcal{S} \cup (\mathcal{S}' \setminus \mathcal{R})}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} * \varphi_n^a)$. By Rule S11 and Theorem 18, $e * IF(n) = e'_{\mathcal{S} \cup (\mathcal{S}' \setminus \mathcal{R})}$. By Lemma 73, $(e'_p \gg n) = (e''_{\mathcal{R}})_p$. Thus $(e'_{\mathcal{S} \cup (\mathcal{S}' \setminus \mathcal{R})})_p \gg n = (e''_{\mathcal{S} \cup \mathcal{S}'})_p$, and so $e \gg n = e''_{\mathcal{S} \cup \mathcal{S}'}$. Moreover by assumption and Rule S11, $e''_{\mathcal{S} \cup \mathcal{S}'} \models \mathbf{O}_{\mathcal{S} \cup \mathcal{S}'}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n)$. Thus by Rule S7, $e \models [n]\mathbf{O}_{\mathcal{S} \cup \mathcal{S}'}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n)$. □

Theorem 47. $\mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models [n]\alpha$ *iff* $\mathbf{O}_{\mathcal{S}\cup\mathcal{S}'}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n) \models \alpha$.

Proof. For the *only-if* direction suppose $\mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}}) \models [n]\alpha$ and let $e \models \mathbf{O}_{\mathcal{S}\cup\mathcal{S}'}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n)$. By Corollary 34, there is some $e' \models \mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$. By assumption, $e', w \models [n]\alpha$ for all w, and thus $e' \gg n, w \models \alpha$ for all w. By Theorem 46, $e' \gg n \models \mathbf{O}_{\mathcal{S}\cup\mathcal{S}'}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n)$, and again by Corollary 34, $e = e' \gg n$. Thus $e, w \models \alpha$ for all w.

For the *if* direction, suppose $\mathbf{O}_{\mathcal{S}\cup\mathcal{S}'}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n) \models \alpha$ and let $e \models \mathbf{O}_{\mathcal{S}}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}})$. By Theorem 46, $e \gg n \models \mathbf{O}_{\mathcal{S}\cup\mathcal{S}'}(\Sigma_{\text{dyn}}, \Sigma_{\text{bel}} \gg n)$. By assumption, $e \gg n, w \gg n \models \alpha$ for all w. By Rule S7, $e, w \models [n]\alpha$ for all w. □

References

[1] J. McCarthy, Programs with common sense, in: Proceedings of the Symposium on Mechanization of Thought Processes, Her Majesty's Stationary Office, 1959.

[2] J. McCarthy, Situations, actions, and causal laws, Technical Report AI Memo 2, AI Lab, Stanford University (1963).

[3] R. Reiter, The frame problem in the situation calculus: A simple solution (sometimes) and a completeness result for goal regression, in: Artificial Intelligence and Mathematical Theory of Computation, Academic Press, San Diego, CA, 1991, pp. 359–380.

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

[5] P. Peppas, Belief revision, in: Handbook of Knowledge Representation, Elsevier, 2008, pp. 317–359.

[6] F. Lin, R. Reiter, How to progress a database, Artificial Intelligence 92 (1) (1997) 131–167.

[7] S. Vassos, H. J. Levesque, How to progress a database III, Artificial Intelligence 195 (2013) 203–221.

[8] H. J. Levesque, G. Lakemeyer, Cognitive robotics, in: Handbook of Knowledge Representation, Elsevier, 2008, pp. 869–886.

[9] C. Schwering, G. Lakemeyer, A semantic account of iterated belief revision in the situation calculus, in: Proceedings of the Twenty-First European Conference on Artificial Intelligence, 2014, pp. 801–806.

[10] C. Schwering, G. Lakemeyer, Projection in the epistemic situation calculus with belief conditionals, in: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015, pp. 1583–1589.

[11] C. Schwering, G. Lakemeyer, M. Pagnucco, Belief revision and progression of knowledge bases in the epistemic situation calculus, in: Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015, pp. 3124–3220.

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

[13] R. Demolombe, M. d. P. Pozos Parra, Belief revision in the situation calculus without plausibility levels, in: Proceedings of the Sixteenth International Symposium on Methodologies for Intelligent Systems, 2006, pp. 504–513.

[14] J. P. Delgrande, H. J. Levesque, Belief revision with sensing and fallible actions, in: Proceedings of the Thirteenth International Conference on Principles of Knowledge Representation and Reasoning, 2012, pp. 148–157.

[15] L. Fang, Y. Liu, Multiagent knowledge and belief change in the situation calculus, in: Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, 2013, pp. 304–312.

[16] F. P. Ramsey, General Propositions and Causality, Kegan Paul, Trench & Trubner, 1931.

[17] P. Gärdenfors, Belief revisions and the ramsey test for conditionals, The Philosophical Review 95 (1) (1986) 81–93.

[18] C. Boutilier, Revision sequences and nested conditionals, in: Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence, 1993, pp. 519–525.

[19] C. Boutilier, Iterated revision and minimal change of conditional beliefs, Journal of Philosophical Logic 25 (3) (1996) 263–305.

[20] A. C. Nayak, Iterated belief change based on epistemic entrenchment, Erkenntnis 41 (3) (1994) 353–390.

[21] A. C. Nayak, M. Pagnucco, P. Peppas, Dynamic belief revision operators, Artificial Intelligence 146 (2003) 193–228.

[22] C. Schwering, G. Lakemeyer, Decidable reasoning in a first-order logic of limited conditional belief, in: Proceedings of the Twenty-Second European Conference on Artificial Intelligence, 2016, pp. 1379–1387.

[23] C. Schwering, A reasoning system for a first-order logic of limited belief, in: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, 2017, to appear.

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

[25] H. J. Levesque, Foundations of a functional approach to knowledge representation, Artificial Intelligence 23 (2) (1984) 155–212.

[26] H. J. Levesque, G. Lakemeyer, The Logic of Knowledge Bases, MIT Press, 2001.

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

[28] R. Kowalski, M. Sergot, A logic-based calculus of events, in: Foundations of knowledge base management, 1989, pp. 23–55.

[29] M. Gelfond, V. Lifschitz, Action languages, Electronic Transactions on AI 3 (16).

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

[31] G. Lakemeyer, H. J. Levesque, A semantical account of progression in the presence of defaults, in: Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence, 2009, pp. 842–847.

[32] S. Vassos, G. Lakemeyer, H. J. Levesque, First-order strong progression for local-effect basic action theories, in: Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning, 2008, pp. 662–672.

[33] L. Fang, Y. Liu, X. Wen, On the progression of knowledge and belief for nondeterministic actions in the situation calculus (2015) 2955–2963.

[34] W. Spohn, Ordinal conditional functions: A dynamic theory of epistemic states, in: Causation in Decision, Belief Change, and Statistics, 1988, pp. 105–134.

[35] F. Bacchus, J. Y. Halpern, H. J. Levesque, Reasoning about noisy sensors and effectors in the situation calculus, Artificial Intelligence 111 (1–2) (1999) 171–208.

[36] H. van Ditmarsch, W. van der Hoek, B. P. Kooi, Dynamic epistemic logic, Springer, 2007.
[37] G. Aucher, A combined system for update logic and belief revision, in: Proceedings of the Seventh Pacific Rim International Workshop on Multi-Agents, Springer, 2005, pp. 1–17.
[38] A. Baltag, S. Smets, A qualitative theory of dynamic interactive belief revision, Texts in logic and games 3 (2008) 9–58.
[39] H. van Ditmarsch, Prolegomena to dynamic logic for belief revision, Synthese 147 (2) (2005) 229–275.
[40] J. van Benthem, Dynamic logic for belief revision, Journal of Applied Non-Classical Logics 17 (2) (2007) 129–155.
[41] C. E. Alchourrón, P. Gärdenfors, D. Makinson, On the logic of theory change: Partial meet contraction and revision functions, Journal of Symbolic Logic 50 (2) (1985) 510–530.
[42] P. Gärdenfors, Knowledge in flux: Modeling the dynamics of epistemic states, The MIT press, 1988.
[43] A. Darwiche, J. Pearl, On the logic of iterated belief revision, Artificial Intelligence 89 (1) (1997) 1–29.
[44] Y. Jin, M. Thielscher, Iterated belief revision, revised, Artificial Intelligence 171 (1) (2007) 1–18.
[45] R. Booth, T. Meyer, Admissible and restrained revision, Journal of Artificial Intelligence Research 26 (2006) 127–151.
[46] J. P. Delgrande, Y. Jin, Parallel belief revision: Revising by sets of formulas, Artificial Intelligence 176 (1) (2012) 2223–2245.
[47] K. Segerberg, Irrevocable belief revision in dynamic doxastic logic, Notre Dame Journal of Formal Logic 39 (3) (1998) 287–306.
[48] N. Friedman, J. Y. Halpern, Belief revision: A critique, Journal of Logic, Language and Information 8 (4) (1999) 401–420.
[49] H. Rott, Coherence and conservatism in the dynamics of belief ii: Iterated belief change without dispositional coherence, Journal of Logic and Computation 13 (1) (2003) 111–145.
[50] H. Rott, Shifting priorities: Simple representations for twenty-seven iterated theory change operators, in: Towards Mathematical Philosophy, Vol. 28, Springer Netherlands, 2009, pp. 269–296.
[51] H. Katsuno, A. O. Mendelzon, On the difference between updating a knowledge database and revising it, 1990.
[52] C. Boutilier, A unified model of qualitative belief change: A dynamical systems perspective, Artificial Intelligence 98 (1) (1998) 281–316.
[53] D. Lewis, Counterfactuals, John Wiley & Sons, 1973.
[54] A. Grove, Two modellings for theory change, Journal of Philosophical Logic 17 (2) (1988) 157–170.
[55] I. Levi, Iteration of conditionals and the ramsey test, Synthese 76 (1) (1988) 49–81.
[56] C. Schwering, Conditional beliefs in action, Ph.D. thesis, RWTH Aachen University (2016).
[57] J. Pearl, System Z: A natural ordering of defaults with tractable applications to nonmonotonic reasoning, in: Proceedings of the Third Conference on Theoretical Aspects of Reasoning about Knowledge, 1990, pp. 121–135.
[58] C. Boutilier, Inaccessible worlds and irrelevance: Preliminary report, in: Proceedings of the Twelfth International Joint Conference on Artificial Intelligence, 1991, pp. 413–418.
[59] S. A. Kripke, Is there a problem about substitutional quantification?, in: Truth and Meaning, Oxford University Press, 1976, pp. 324–419.
[60] R. Fagin, J. Y. Halpern, Y. Moses, M. Y. Vardi, Reasoning about knowledge, Vol. 4, MIT press Cambridge, 1995.
[61] G. Kern-Isberner, Conditionals in Nonmonotonic Reasoning and Belief Revision, Springer, 2001.