Proceedings of the 23rd International Conference on Principles of Knowledge Representation and Reasoning
Main Track

Truth-Tracking by Iterated Belief Change

Nicolas Schwind[1] , **Patricia Everaere**[2] , **Sébastien Konieczny**[3]
[1] National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan
[2] Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France
[3] CRIL, CNRS - Université d'Artois, France
nicolas-schwind@aist.go.jp, patricia.everaere-caillier@univ-lille.fr, konieczny@cril.fr

Abstract

We investigate the truth-tracking performance of iterated belief change operators. In particular, we show that a class of improvement operators is guaranteed to converge to the truth when the input sequence contains sufficiently many correct pieces of information, and we establish a corresponding convergence theorem. We also report experimental results indicating that this convergence typically occurs with relatively short input sequences.

1 Introduction

Belief revision (Alchourrón, Gärdenfors, and Makinson 1985; Gärdenfors 1988; Katsuno and Mendelzon 1991; Fermé and Hansson 2011) concerns the correction of an agent's current beliefs in light of a more reliable piece of information. Iterated belief revision (Darwiche and Pearl 1997; Jin and Thielscher 2007; Booth and Meyer 2006) (and iterated belief change (Konieczny and Pino Pérez 2008; Konieczny, Medina Grespan, and Pino Pérez 2010; Schwind et al. 2025)) is the required generalization for the standard case where this process is iterated: the agent receives a sequence of new pieces of information.

Note that the rationale for this process is that the new pieces of information are supposed to be more reliable than the previous beliefs of the agent. As the beliefs are (typically) represented by formulas in propositional logic, this means that the agent expects the truth (the true world) to be one of the models of her beliefs. Additionally, because the new information is thought to be more trustworthy than the existing beliefs, the agent considers that the real world is more likely to fit the models of the new information than her previous views.

So the very aim of belief revision (and iterated belief change) is motivated by this notion of truth. It is thus surprising that there is almost no truth-tracking evaluation of these operators. Most studies on belief revision are about logical properties (postulates), representation theorems, and computational complexity issues.

The only exception we are aware of is the work of (Baltag, Gierasimczuk, and Smets 2019). The authors of this work examine the fact that revision operators guarantee discovery of the truth for infinite sequences of revisions by correct pieces of information. They focus on a particular revision operator[1]: Nayak's lexicographic revision (Nayak 1994). They show that this operator is able to find the truth under very restricted hypotheses. Indeed, they assume that the sequence is infinite, that epistemic states are represented by total preorders (despite the fact that some interesting iterated belief change operators cannot be instantiated in this epistemic space (Schwind, Konieczny, and Pino Pérez 2022)), and that the new pieces of information include all and only correct formulas.

This last hypothesis in particular is very unrealistic. There will be very few frameworks where we can guarantee that all new pieces of information will be correct (i.e., that they all contain the true world), and that all possible correct pieces of information will be provided.

In realistic scenarios there will also be incorrect pieces of information (that do not contain the true world). The best we can expect is that the new pieces of information satisfy a reliability condition: the true world is more likely to be included than any other fixed world. This is the scenario that we study in this work. So we will consider sequences of changes where the new piece of information can be correct or incorrect, but where the process is reliable in the sense that the true world is more likely to be included in each generated formula than any other fixed world. Without this hypothesis, one cannot expect to find the truth.

It is important to stress that revision operators are not the appropriate operators for truth-tracking because of the success postulate, which forces the resulting beliefs to imply the last piece of information. The agent's beliefs will therefore be incorrect each time this piece of information is inaccurate, even if the wrong information occurs once after hundreds of correct inputs. For this truth-tracking task, improvement operators (Konieczny and Pino Pérez 2008) are more adequate. Improvement operators aim to ensure that the plausibility of the new piece of information is increased, but do not force it to be believed after the change. If the plausibility of the new piece of information is increased, and if the true world is more likely to be selected as a model than any other fixed world, the process may end with the true

[1] They also study a "conditioning" operator that is basically a conjunction, and thus does not satisfy AGM (nor improvement) postulates, and a "minimal" operator that is Boutilier's natural revision (Boutilier 1996), but they show that it does not allow to correctly track truth in their setting.

world as the most plausible world.

This is the hypothesis that we test in this work. We show that for a class of improvement operators we can state a truth-tracking theorem. Since this is a limit theorem, it could be problematic if convergence required very long sequences of change. So we also perform an empirical study showing that the convergence is obtained quite quickly, and study under which hypotheses convergence is faster or slower.

2 Preliminaries

Let $\mathcal{L}_\mathcal{P}$ be a propositional language built up from a finite set of propositional variables $\mathcal{P}$ and the usual connectives. The set $\mathcal{L}^*_\mathcal{P}$ denotes the set of consistent formulas from $\mathcal{L}_\mathcal{P}$. $\perp$ (resp. $\top$) is the Boolean constant always false (resp. true). A world is a mapping from $\mathcal{P}$ to $\{0,1\}$. We denote by Ω the set of all worlds over $\mathcal{L}_\mathcal{P}$. $\models$ denotes logical entailment, $\equiv$ logical equivalence, and $[\varphi]$ denotes the set of models of the formula φ. When a world ω is a model of a formula φ, we often write $\omega \models \varphi$ instead of $\omega \in [\varphi]$. For a set of worlds $X \subseteq \Omega$, we denote by ψ_X a formula whose models are X.

In iterated belief change, it is standard to assume that the current set of beliefs of an agent is represented by an epistemic state. Unlike a simple belief set, an epistemic state captures both the agent's current beliefs and the conditional information necessary to govern the revision process. Following the seminal work of Darwiche and Pearl (1997), an epistemic state is an abstract object Ψ. The agent's actual beliefs are retrieved via a belief mapping B, so that $B(\Psi)$ is a propositional formula from $\mathcal{L}_\mathcal{P}$. Formally, let U be the set of all epistemic states, then B is a mapping $B : U \to \mathcal{L}_\mathcal{P}$. Schwind, Konieczny, and Pino Pérez (2022) have introduced the notion of epistemic space, which corresponds to the choice of a concrete representation, where an epistemic state is an element of a given epistemic space:

Definition 1 (Epistemic Space)**.** *An* epistemic space *is a pair $\mathcal{E} = \langle U, B\rangle$, where U is a non-empty set of so-called* epistemic states *and B is a mapping $B : U \to \mathcal{L}_\mathcal{P}$.*

An iterated belief change operator maps an initial epistemic state and a change formula to a successor epistemic state. In this work, we restrict our focus to consistent change formulas. Formally, a change operator is defined as a mapping $\circ : U \times \mathcal{L}^*_\mathcal{P} \to U$.

Improvement operators (Konieczny and Pino Pérez 2008; Konieczny, Medina Grespan, and Pino Pérez 2010; Medina Grespan and Pino Pérez 2013) constitute a more general class of operators than the standard iterated revision operators (Darwiche and Pearl 1997). The primary divergence lies in the relaxation of the Success Postulate (R*1), which traditionally requires that the change formula be entailed by the resulting epistemic state. In the context of improvement, (R*1) is replaced by a weaker requirement: the change formula must only be entailed after a finite sequence of repeated improvements, as formalized by property (I1). In the following, given an iterated belief change operator $\circ$, $\Psi \circ^k \varphi$ is inductively defined as $\Psi \circ^1 \varphi = \Psi \circ \varphi$ and for each $k > 1$, $\Psi \circ^k \varphi = (\Psi \circ^{k-1} \varphi) \circ \varphi$. Then $\Psi \star_\circ \varphi$ is defined as $\Psi \circ^n \varphi$, where n is the first integer such that $B(\Psi \circ^n \varphi) \models \varphi$. Note that $\star_\circ$ is undefined if there is no n such that $B(\Psi \circ^n \varphi) \models \varphi$, but for all operators $\circ$ considered in this work (because they all satisfy (I1) below), the associated operator $\star_\circ$ is total, that is, for any pair Ψ, φ there exists n such that $B(\Psi \circ^n \varphi) \models \varphi$.

Here are the postulates that have been proposed for the main classes of improvement operators (Konieczny, Medina Grespan, and Pino Pérez 2010; Konieczny and Pino Pérez 2008):

Definition 2 (Improvement Operators)**.** *An operator $\circ$ is called a* weak improvement operator *if it satisfies postulates (I1-I6). $\circ$ is an* improvement operator *if it satisfies postulates (I1-I9). $\circ$ is a* soft improvement operator *if it satisfies postulates (I1-I10).*

- ***(I1)*** *There exists n such that $B(\Psi \circ^n \varphi) \models \varphi$*
- ***(I2)*** *If $B(\Psi) \wedge \varphi \not\models \perp$, then $B(\Psi \star_\circ \varphi) \equiv B(\Psi) \wedge \varphi$*
- ***(I3)*** *If $\varphi \not\models \perp$, then $B(\Psi \circ \varphi) \not\models \perp$*
- ***(I4)*** *For any positive integer n, if $\varphi_i \equiv \psi_i$ for all $i \leq n$, then $B(\Psi \circ \varphi_1 \circ \ldots \circ \varphi_n) \equiv B(\Psi \circ \psi_1 \circ \ldots \circ \psi_n)$*
- ***(I5)*** *$B(\Psi \star_\circ \varphi) \wedge \psi \models B(\Psi \star_\circ (\varphi \wedge \psi))$*
- ***(I6)*** *If $B(\Psi \star_\circ \varphi) \wedge \psi \not\models \perp$, then $B(\Psi \star_\circ (\varphi \wedge \psi)) \models B(\Psi \star_\circ \varphi) \wedge \psi$*
- ***(I7)*** *If $\varphi \models \psi$, then $B((\Psi \circ \psi) \star_\circ \varphi) \equiv B(\Psi \star_\circ \varphi)$*
- ***(I8)*** *If $\varphi \models \neg\psi$, then $B((\Psi \circ \psi) \star_\circ \varphi) \equiv B(\Psi \star_\circ \varphi)$*
- ***(I9)*** *If $B(\Psi \star_\circ \varphi) \not\models \neg\psi$, then $B((\Psi \circ \psi) \star_\circ \varphi) \models \psi$*
- ***(I10)*** *If $B(\Psi \star_\circ \varphi) \models \neg\psi$, then $B((\Psi \circ \psi) \star_\circ \varphi) \not\models \psi$*

For a detailed justification of these postulates, the reader is referred to (Konieczny and Pino Pérez 2008; Konieczny, Medina Grespan, and Pino Pérez 2010). The postulates (I1-I6) are the most basic ones, corresponding to the classical AGM/KM belief revision postulates under a relaxed success condition. These postulates define the class of weak improvement operators. (I7-I9) adapt the requirements of iterated revision to this generalized setting ((I7) is a translation of (C1) (Darwiche and Pearl 1997), (I8) is a translation of (C2), and (I9) is a translation of the admissibility postulate (P) from (Booth and Meyer 2006; Jin and Thielscher 2007)). These three postulates (together with the basic ones) define the class of improvement operators. Finally, the class of soft improvement operators is distinguished by postulate (I10), which ensures that the increase in plausibility for the change formula remains small at each step.

While Definition 2 provides an axiomatic definition of three classes of improvement operators, each of these classes can be characterized by associating each epistemic state with a total preorder over worlds:

Definition 3 (Faithful Assignments)**.** *A function $\Psi \mapsto \preceq_\Psi$ that maps each epistemic state Ψ to a total preorder[2] over worlds $\preceq_\Psi$ is called a* strong faithful assignment *if and only if:*

1. *If $\omega, \omega' \models B(\Psi)$, then $\omega \simeq_\Psi \omega'$*
2. *If $\omega \models B(\Psi)$ and $\omega' \not\models B(\Psi)$, then $\omega \prec_\Psi \omega'$*
3. *For any positive integer n, if $\varphi_i \equiv \psi_i$ for all $i \leq n$, then $\preceq_{\Psi \circ \varphi_1 \circ \ldots \circ \varphi_n} = \preceq_{\Psi \circ \psi_1 \circ \ldots \circ \psi_n}$*

[2]For each preorder $\preceq$, $\simeq$ denotes the corresponding indifference relation, and $\prec$ the strict part of $\preceq$.

A strong faithful assignment is called a gradual assignment *if and only if:*

(S1) *If $\omega, \omega' \models \varphi$, then $\omega \preceq_\Psi \omega' \Leftrightarrow \omega \preceq_{\Psi\circ\varphi} \omega'$*

(S2) *If $\omega, \omega' \models \neg\varphi$, then $\omega \preceq_\Psi \omega' \Leftrightarrow \omega \preceq_{\Psi\circ\varphi} \omega'$*

(S3) *If $\omega \models \varphi$ and $\omega' \models \neg\varphi$, then $\omega \preceq_\Psi \omega' \Rightarrow \omega \prec_{\Psi\circ\varphi} \omega'$*

A gradual assignment is called a soft gradual assignment *if and only if:*

(S4) *If $\omega \models \varphi$ and $\omega' \models \neg\varphi$, then $\omega' \prec_\Psi \omega \Rightarrow \omega' \preceq_{\Psi\circ\varphi} \omega$*

Proposition 1 ((Konieczny, Medina Grespan, and Pino Pérez 2010)). *An iterated belief change operator $\circ$ is a weak improvement operator (resp. improvement operator, soft improvement operator) if and only if there exists a strong faithful assignment (resp. gradual assignment, soft gradual assignment) that maps each epistemic state Ψ to a total preorder over worlds $\preceq_\Psi$ such that for each epistemic state Ψ and each formula $\varphi \in \mathcal{L}^*_{\mathcal{P}}$:*

$$[B(\Psi \star_\circ \varphi)] = \min([\varphi], \preceq_\Psi).$$

An important remark is that (weak) improvement operators generalize iterated revision operators. More precisely, weak improvement operators satisfying (R*1) are AGM/DP revision operators (Alchourrón, Gärdenfors, and Makinson 1985; Darwiche and Pearl 1997), and improvement operators satisfying (R*1) are admissible iterated revision operators in the sense of (Booth and Meyer 2006).

An important epistemic space in the literature is the epistemic space of Ordinal Conditional Functions (OCFs) (Spohn 1988), also called ranking functions or κ-functions, that are functions κ that associate a non-negative integer[3] $\kappa(\omega)$ to each world ω, such that there is a world $\omega \in \Omega$ with $\kappa(\omega) = 0$. Then, the OCF-*based epistemic space* is the epistemic space $\mathcal{E}^{\mathsf{ocf}} = \langle U^{\mathsf{ocf}}, B^{\mathsf{ocf}} \rangle$, where U^{ocf} is the set of all OCFs, and B^{ocf} is the mapping associating each OCF $\kappa \in U^{\mathsf{ocf}}$ with the formula $B^{\mathsf{ocf}}(\kappa) = \psi_{\{\omega\in\Omega \mid \kappa(\omega)=0\}}$. For every OCF κ and every formula α, we denote $\kappa(\alpha) = \min_{\omega\models\alpha} \kappa(\omega)$, which is the degree of disbelief of α, i.e., how much "effort" it would take to the agent to believe α.

3 +x Improvement Operators

In this section we introduce $+x$ improvement operators, which ensure that the degree of disbelief of the worlds that are not models of the new piece of information is worsened by x in the OCF.

Definition 4 (Free-OCF). *A free-*OCF *is a function κ that associates a non-negative integer $\kappa(\omega)$ to each world ω.*

The difference between an OCF and a free-OCF is that free-OCFs do not require a normalization, i.e., that at least one world is assigned 0. In this case, the beliefs associated to a free-OCF κ are not the worlds assigned to 0 but to the minimal integer: $B^{\mathsf{focf}}(\kappa) = \psi_{\{\omega\in\Omega \mid \kappa(\omega)=\kappa(\top)\}}$, where, similarly to the case of OCFs, for every free-OCF κ and every formula α, we denote $\kappa(\alpha) = \min_{\omega\models\alpha} \kappa(\omega)$. Note that the degree of disbelief of a formula α in a free-OCF now corresponds to $\kappa(\alpha) - \kappa(\top)$. Then, the *free-*OCF*-based epistemic space* is the epistemic space $\mathcal{E}^{\mathsf{focf}} = \langle U^{\mathsf{focf}}, B^{\mathsf{focf}} \rangle$, where U^{focf} is the set of all free-OCFs, and B^{focf} is the mapping defined above.

This epistemic space is interesting because it is strictly more expressive than the OCF-based epistemic space, but we will not develop this point here. We want to emphasize that the definition of several iterated belief change operators is more natural with free-OCFs instead of OCFs, i.e., without the normalization required by OCFs. Let us introduce the $+x$ improvement operators on free-OCFs:

Definition 5 (Operators $\bullet_{+x}$). *Let x be a positive integer. The operator $\bullet_{+x}$ associates with any free-*OCF *κ and any propositional formula $\varphi \in \mathcal{L}^*_{\mathcal{P}}$ the free-*OCF *$\kappa \bullet_{+x} \varphi$ defined for each world $\omega \in \Omega$ by:*

$$(\kappa \bullet_{+x} \varphi)(\omega) = \begin{cases} \kappa(\omega) & \text{if } \omega \models \varphi \\ \kappa(\omega) + x & \text{if } \omega \not\models \varphi \end{cases}$$

Note that the obtained ranking is not an OCF in general since in some cases there may be no world assigned 0 in the result, so it is a free-OCF. So, if one wants to obtain a true OCF operator one needs to add a normalization step:

Definition 6 (Normalization). *Let κ be a free-*OCF*. The normalization of κ is the* OCF *$\kappa_{\downarrow_0}$ defined for each world $\omega \in \Omega$ by:*

$$\kappa_{\downarrow_0}(\omega) = \kappa(\omega) - \kappa(\top)$$

Accordingly, normalizing a free-OCF κ into $\kappa_{\downarrow_0}$ always leads to a true OCF: for all worlds $\omega \models B^{\mathsf{focf}}(\kappa)$, $\kappa_{\downarrow_0}(\omega) = \kappa(\omega) - \kappa(\top) = 0$, and for all worlds $\omega' \in \Omega$, $\kappa_{\downarrow_0}(\omega') \geq \kappa_{\downarrow_0}(\omega) = 0$.

So now we can define the $+x$ improvement operators on OCF:

Definition 7 (Operators $\circ_{+x}$). *Let x be a positive integer. The operator $\circ_{+x}$ associates with any* OCF *κ and any propositional formula $\varphi \in \mathcal{L}^*_{\mathcal{P}}$ the* OCF *$\kappa \circ_{+x} \varphi = (\kappa \bullet_{+x} \varphi)_{\downarrow_0}$.*

We can show that for each positive integer x, the operator $\bullet_{+x}$ on free-OCFs and the operator $\circ_{+x}$ on OCFs represent the same operator. Let us put it formally, following the notion of equivalence between iterated belief revision operators from (Schwind, Konieczny, and Pino Pérez 2022).

Let $\circ$ and $\circ'$ be two iterated belief change operators defined on epistemic spaces $\mathcal{E} = \langle U, B \rangle$ and $\mathcal{E}' = \langle U', B' \rangle$, respectively. A *scenario* is a finite, possibly empty, sequence of formulas $\sigma = (\varphi_1, \ldots, \varphi_m)$. We write $\Psi \circ \sigma$ for the epistemic state obtained from Ψ by successively applying the formulas in σ. Two epistemic states $\Psi \in U$ and $\Psi' \in U'$ are *strongly equivalent* with respect to $(\circ, \circ')$ if, for every scenario σ, $B(\Psi \circ \sigma) \equiv B'(\Psi' \circ' \sigma)$. The operators $\circ$ and $\circ'$ are *equivalent* if there is a one-to-one correspondence between the strong-equivalence classes of U induced by $\circ$ and those of U' induced by $\circ'$, such that any two states belonging to corresponding classes are strongly equivalent with respect to $(\circ, \circ')$.

[3]In (Spohn 1988) ordinals were used instead of integers. For this work, integers are enough. See (Konieczny 2009) for a true use of ordinals.

The proof that for each positive integer x, $\bullet_{+x}$ and $\circ_{+x}$ are equivalent operators takes advantage of a few lemmas.

The first lemma identifies the new minimum value of a free-OCF after one $+x$ improvement. Since the models of φ keep their previous ranks whereas the countermodels of φ are penalized by x, the new global minimum is either the best previous rank among the models of φ, or the old global minimum shifted upward by x:

Lemma 1. *For each free-*OCF *κ and each formula $\varphi \in \mathcal{L}^*_{\mathcal{P}}$,*

$$(\kappa \bullet_{+x} \varphi)(\top) = \min(\{\kappa(\varphi), \kappa(\top) + x\})$$

Proof. We consider two cases.

Case 1: $\kappa(\varphi) < \kappa(\top) + x$. Let us prove that $(\kappa \bullet_{+x} \varphi)(\top) = \kappa(\varphi)$. Let ω be a world such that (i) $\omega \models \varphi$ and (ii) $\kappa(\omega) = \kappa(\varphi)$. From (i) and by definition of $\bullet_{+x}$, we know that (iii) $(\kappa \bullet_{+x} \varphi)(\omega) = \kappa(\omega)$. We intend to show that for all worlds $\omega' \in \Omega$, $(\kappa \bullet_{+x} \varphi)(\omega) \leq (\kappa \bullet_{+x} \varphi)(\omega')$. Let $\omega' \in \Omega$. Assume first that $\omega' \models \varphi$. Then (iv) $(\kappa \bullet_{+x} \varphi)(\omega') = \kappa(\omega')$ by definition of $\bullet_{+x}$. Yet from (ii), we have that (v) $\kappa(\omega) \leq \kappa(\omega')$. So from (iii), (iv) and (v) we get that $(\kappa \bullet_{+x} \varphi)(\omega) \leq (\kappa \bullet_{+x} \varphi)(\omega')$. Assume now that $\omega' \models \neg\varphi$. Then (vi) $(\kappa \bullet_{+x} \varphi)(\omega') = \kappa(\omega') + x$ by definition of $\bullet_{+x}$. Yet the assumption for this case is $\kappa(\varphi) < \kappa(\top) + x$, so in particular (vii) $\kappa(\varphi) < \kappa(\omega') + x$. From (ii), (iii) and (vi) we obtain $(\kappa \bullet_{+x} \varphi)(\omega) < (\kappa \bullet_{+x} \varphi)(\omega')$. We have shown that for all worlds $\omega' \in \Omega$, $(\kappa \bullet_{+x} \varphi)(\omega) \leq (\kappa \bullet_{+x} \varphi)(\omega')$. Thus, $(\kappa \bullet_{+x} \varphi)(\omega) = (\kappa \bullet_{+x} \varphi)(\top)$. From (i) and (ii), this means that $(\kappa \bullet_{+x} \varphi)(\top) = \kappa(\varphi)$.

Case 2: $\kappa(\varphi) \geq \kappa(\top) + x$. Let us prove that $(\kappa \bullet_{+x} \varphi)(\top) = \kappa(\top) + x$. Let ω be a world such that $\omega \models B^{\mathsf{focf}}(\kappa)$. So by definition of $B^{\mathsf{focf}}(\kappa)$, (vii) $\kappa(\omega) = \kappa(\top)$, and since we assumed that $\kappa(\varphi) \geq \kappa(\top) + x$, we know that $\kappa(\varphi) > \kappa(\top)$, so $\omega \models \neg\varphi$, so by definition of $\bullet_{+x}$, $(\kappa \bullet_{+x} \varphi)(\omega) = \kappa(\omega) + x$, which by (vii) can be written as (viii) $(\kappa \bullet_{+x} \varphi)(\omega) = \kappa(\top) + x$. We intend to show that for all worlds $\omega' \in \Omega$, $(\kappa \bullet_{+x} \varphi)(\omega) \leq (\kappa \bullet_{+x} \varphi)(\omega')$. Let $\omega' \in \Omega$. Assume first that $\omega' \models \varphi$. Then by definition of $\bullet_{+x}$, (ix) $(\kappa \bullet_{+x} \varphi)(\omega') = \kappa(\omega')$. Yet (x) $\kappa(\omega') \geq \kappa(\varphi)$ (since $\omega' \models \varphi$), and (xi) $\kappa(\varphi) \geq \kappa(\top) + x$ (our assumption for this case). Hence, by (viii-xi), $(\kappa \bullet_{+x} \varphi)(\omega) \leq (\kappa \bullet_{+x} \varphi)(\omega')$. Assume now that $\omega' \models \neg\varphi$. Then (xii) $(\kappa \bullet_{+x} \varphi)(\omega') = \kappa(\omega') + x$ by definition of $\bullet_{+x}$. Yet $\kappa(\omega') \geq \kappa(\top)$, so (xiii) $\kappa(\omega') + x \geq \kappa(\top) + x$. Hence, by (viii), (xii) and (xiii), $(\kappa \bullet_{+x} \varphi)(\omega) \leq (\kappa \bullet_{+x} \varphi)(\omega')$. We have shown that for all worlds $\omega' \in \Omega$, $(\kappa \bullet_{+x} \varphi)(\omega) \leq (\kappa \bullet_{+x} \varphi)(\omega')$. Thus, $(\kappa \bullet_{+x} \varphi)(\omega) = (\kappa \bullet_{+x} \varphi)(\top)$. By (viii), we obtain $(\kappa \bullet_{+x} \varphi)(\top) = \kappa(\top) + x$.

We have shown that $\kappa(\varphi) < \kappa(\top) + x$ when $(\kappa \bullet_{+x} \varphi)(\top) = \kappa(\varphi)$, and $(\kappa \bullet_{+x} \varphi)(\top) = \kappa(\top) + x$ when $\kappa(\varphi) \geq \kappa(\top) + x$, which concludes the proof that $(\kappa \bullet_{+x} \varphi)(\top) = \min(\{\kappa(\varphi), \kappa(\top) + x\})$. ☐

The second lemma states that the normalization step does not interfere with the effect of a $+x$ improvement. This is the key technical reason why we can work freely with free-OCFs and later recover the corresponding normalized OCF behavior:

Lemma 2. *For each free-*OCF *κ and each formula $\varphi \in \mathcal{L}^*_{\mathcal{P}}$,*

$$(\kappa \bullet_{+x} \varphi)_{\downarrow 0} = (\kappa_{\downarrow 0} \bullet_{+x} \varphi)_{\downarrow 0}$$

Proof. We need to show that for all worlds $\omega \in \Omega$, $(\kappa \bullet_{+x} \varphi)_{\downarrow 0}(\omega) = (\kappa_{\downarrow 0} \bullet_{+x} \varphi)_{\downarrow 0}(\omega)$. So let $\omega \in \Omega$. Assume first that $\omega \models \varphi$. On the one hand, $(\kappa \bullet_{+x} \varphi)_{\downarrow 0}(\omega) = (\kappa \bullet_{+x} \varphi)(\omega) - (\kappa \bullet_{+x} \varphi)(\top) = \kappa(\omega) - (\kappa \bullet_{+x} \varphi)(\top)$ (since $(\kappa \bullet_{+x} \varphi)(\omega) = \kappa(\omega)$ by definition of $\bullet_{+x}$). On the other hand, $(\kappa_{\downarrow 0} \bullet_{+x} \varphi)_{\downarrow 0}(\omega) = (\kappa_{\downarrow 0} \bullet_{+x} \varphi)(\omega) - (\kappa_{\downarrow 0} \bullet_{+x} \varphi)(\top) = \kappa_{\downarrow 0}(\omega) - (\kappa_{\downarrow 0} \bullet_{+x} \varphi)(\top)$ (since $(\kappa_{\downarrow 0} \bullet_{+x} \varphi)(\omega) = \kappa_{\downarrow 0}(\omega)$ by definition of $\bullet_{+x}$). So to show that $(\kappa \bullet_{+x} \varphi)_{\downarrow 0}(\omega) = (\kappa_{\downarrow 0} \bullet_{+x} \varphi)_{\downarrow 0}(\omega)$ when $\omega \models \varphi$, we need to show that $\kappa(\omega) - (\kappa \bullet_{+x} \varphi)(\top) = \kappa_{\downarrow 0}(\omega) - (\kappa_{\downarrow 0} \bullet_{+x} \varphi)(\top)$. Similarly, using again the definition of $\bullet_{+x}$, when $\omega \not\models \varphi$ we need to show that $\kappa(\omega) - (\kappa \bullet_{+x} \varphi)(\top) + x = \kappa_{\downarrow 0}(\omega) - (\kappa_{\downarrow 0} \bullet_{+x} \varphi)(\top) + x$. So in both cases, we need to show that $\kappa(\omega) - (\kappa \bullet_{+x} \varphi)(\top) = \kappa_{\downarrow 0}(\omega) - (\kappa_{\downarrow 0} \bullet_{+x} \varphi)(\top)$, which can be simplified to $(\kappa \bullet_{+x} \varphi)(\top) = \kappa(\top) + (\kappa_{\downarrow 0} \bullet_{+x} \varphi)(\top)$ since $\kappa_{\downarrow 0}(\omega) = \kappa(\omega) - \kappa(\top)$.

Yet on the one hand, from Lemma 1, we have that $(\kappa_{\downarrow 0} \bullet_{+x} \varphi)(\top) = \min(\{\kappa_{\downarrow 0}(\varphi), \kappa_{\downarrow 0}(\top) + x\}) = \min(\{\kappa(\varphi) - \kappa(\top), \kappa(\top) - \kappa(\top) + x\}) = \min(\{\kappa(\varphi), \kappa(\top) + x\}) - \kappa(\top)$. And on the other hand, from Lemma 1 again, $(\kappa \bullet_{+x} \varphi)(\top) = \min(\{\kappa(\varphi), \kappa(\top) + x\})$. This shows that $(\kappa \bullet_{+x} \varphi)(\top) = \kappa(\top) + (\kappa_{\downarrow 0} \bullet_{+x} \varphi)(\top)$, and concludes the proof. ☐

We can now state the following result:

Proposition 2. *For each positive integer x, $\bullet_{+x}$ and $\circ_{+x}$ are equivalent.*

Proof. Let $\sigma = (\varphi_1, \ldots, \varphi_m)$ be any scenario. For a free-OCF κ, define $\kappa^0 = \kappa$ and $\kappa^i = \kappa^{i-1} \bullet_{+x} \varphi_i$ for each $i \in \{1, \ldots, m\}$. Similarly, define $\widehat{\kappa}^0 = \kappa_{\downarrow 0}$ and $\widehat{\kappa}^i = \widehat{\kappa}^{i-1} \circ_{+x} \varphi_i$ for each $i \in \{1, \ldots, m\}$.

We first show, by induction on i, that $(\kappa^i)_{\downarrow 0} = \widehat{\kappa}^i$ for each $i \in \{0, \ldots, m\}$. The case $i = 0$ is immediate. Assume that $(\kappa^{i-1})_{\downarrow 0} = \widehat{\kappa}^{i-1}$. Then, using Lemma 2 and the definition of $\circ_{+x}$, we get $(\kappa^i)_{\downarrow 0} = (\kappa^{i-1} \bullet_{+x} \varphi_i)_{\downarrow 0} = ((\kappa^{i-1})_{\downarrow 0} \bullet_{+x} \varphi_i)_{\downarrow 0} = (\widehat{\kappa}^{i-1} \bullet_{+x} \varphi_i)_{\downarrow 0} = \widehat{\kappa}^i$. Hence, for every scenario σ, $(\kappa \bullet_{+x} \sigma)_{\downarrow 0} = \kappa_{\downarrow 0} \circ_{+x} \sigma$. Moreover, for every free-OCF η, $B^{\mathsf{focf}}(\eta) \equiv B^{\mathsf{ocf}}(\eta_{\downarrow 0})$. Applying this to $\eta = \kappa \bullet_{+x} \sigma$, we obtain, for every free-OCF κ and every scenario σ, $B^{\mathsf{focf}}(\kappa \bullet_{+x} \sigma) \equiv B^{\mathsf{ocf}}((\kappa \bullet_{+x} \sigma)_{\downarrow 0}) = B^{\mathsf{ocf}}(\kappa_{\downarrow 0} \circ_{+x} \sigma)$.

It remains to connect this observation with the definition of operator equivalence. Consider the correspondence that maps each strong-equivalence class of a free-OCF κ with respect to $\bullet_{+x}$ to the strong-equivalence class of $\kappa_{\downarrow 0}$ with respect to $\circ_{+x}$.

First, this correspondence is well-defined. Indeed, suppose that κ and κ' are strongly equivalent with respect to $\bullet_{+x}$. Then, for every scenario σ, $B^{\mathsf{focf}}(\kappa \bullet_{+x} \sigma) \equiv B^{\mathsf{focf}}(\kappa' \bullet_{+x} \sigma)$. Using the equivalence established above, we get $B^{\mathsf{ocf}}(\kappa_{\downarrow 0} \circ_{+x} \sigma) \equiv B^{\mathsf{ocf}}(\kappa'_{\downarrow 0} \circ_{+x} \sigma)$. Hence $\kappa_{\downarrow 0}$ and $\kappa'_{\downarrow 0}$ are strongly equivalent with respect to $\circ_{+x}$.

The correspondence is surjective: it reaches every strong-equivalence class of OCFs, because every OCF λ is the normalization of the free-OCF λ itself. Moreover, the correspondence is injective. Indeed, suppose that two free-OCFs κ and κ' are mapped to the same strong-equivalence class of OCFs. This means that $\kappa_{\downarrow 0}$ and $\kappa'_{\downarrow 0}$ are strongly equivalent with respect to $\circ_{+x}$. Hence, for every scenario σ, $B^{\mathsf{ocf}}(\kappa_{\downarrow 0} \circ_{+x} \sigma) \equiv B^{\mathsf{ocf}}(\kappa'_{\downarrow 0} \circ_{+x} \sigma)$. Using again the equivalence established above, once with κ and once with κ', we obtain $B^{\mathsf{focf}}(\kappa \bullet_{+x} \sigma) \equiv B^{\mathsf{focf}}(\kappa' \bullet_{+x} \sigma)$. Therefore κ and κ' are strongly equivalent with respect to $\bullet_{+x}$.

Hence the correspondence is a bijection between strong-equivalence classes, and so $\bullet_{+x}$ and $\circ_{+x}$ are equivalent. $\square$

Proposition 2 says that the $\bullet_{+x}$ operators on free OCFs and the $\circ_{+x}$ operators on OCFs represent the same operator (for any fixed positive integer x). Since the postulates (I1-I10) are formulated only in terms of beliefs obtained after finite sequences of changes, Proposition 2 also implies that an operator $\bullet_{+x}$ on free-OCFs satisfies any of these postulates if and only if its counterpart $\circ_{+x}$ on OCFs satisfies the same postulate. We can now prove that the $\circ_{+x}$ operators on OCFs are improvement operators, by proving that their free-OCF counterparts $\bullet_{+x}$ are improvement operators:

Proposition 3. *$\circ_{+x}$ operators on $\mathcal{E}^{\mathsf{ocf}}$ are improvement operators, i.e., they satisfy the postulates (I1-I9).*

Proof. We intend to prove that each $\bullet_{+x}$ operator satisfies (I1-I9). By Proposition 2, it is then straightforward to conclude that each $\circ_{+x}$ operator satisfies (I1-I9). The proof unfolds as follows. We first associate with each $\bullet_{+x}$ operator on $\mathcal{E}^{\mathsf{focf}}$ an assignment $\kappa \mapsto \preceq_\kappa$, and show that $\kappa \mapsto \preceq_\kappa$ is a gradual assignment. We then prove that $\bullet_{+x}$ satisfies (I1), and finally that for every formula $\varphi \in \mathcal{L}^*_{\mathcal{P}}$, $[B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \varphi)] = \min([\varphi], \preceq_\kappa)$.

Let us construct such an assignment $\kappa \mapsto \preceq_\kappa$ as follows. We associate each free-OCF κ with a binary relation over worlds $\preceq_\kappa$ defined for all worlds $\omega, \omega' \in \Omega$ as:

$$\omega \preceq_\kappa \omega' \text{ if and only if } \kappa(\omega) \leq \kappa(\omega')$$

By construction, $\preceq_\kappa$ is a total preorder: the fact that $\preceq_\kappa$ is total and transitive is a direct consequence of the totality and transitivity of $\leq$.

To prove that $\kappa \mapsto \preceq_\kappa$ is a strong faithful assignment, we need to prove that it satisfies conditions 1-3 from Definition 3. Yet by definition of B^{focf}, we have for each free-OCF κ that $B^{\mathsf{focf}}(\kappa) = \psi_{\{\omega \in \Omega | \kappa(\omega) = \kappa(\top)\}}$, i.e., $[B^{\mathsf{focf}}(\kappa)] = \arg\min_{\omega \in \Omega} \kappa(\omega)$, so $[B^{\mathsf{focf}}(\kappa)] = \min(\Omega, \preceq_\kappa)$, by definition of $\preceq_\kappa$. This directly shows that conditions 1 and 2 of a strong faithful assignment are satisfied. Condition 3 is direct by definition of $\bullet_{+x}$ (cf. Definition 5), which implies that if $\varphi \equiv \psi$ then $\kappa \bullet_{+x} \varphi = \kappa \bullet_{+x} \psi$. Hence, $\kappa \mapsto \preceq_\kappa$ is a strong faithful assignment.

We now prove that $\kappa \mapsto \preceq_\kappa$ is a gradual assignment, i.e., that it satisfies (S1)-(S3).

Let us prove condition (S1). Assume $\omega, \omega' \models \varphi$. By Definition 5, $\kappa(\omega) = (\kappa \bullet_{+x} \varphi)(\omega)$ and $\kappa(\omega') = (\kappa \bullet_{+x} \varphi)(\omega')$. Then, we have that $\omega \preceq_\kappa \omega'$ if and only if $\kappa(\omega) \leq \kappa(\omega')$ if and only if $(\kappa \bullet_{+x} \varphi)(\omega) \leq (\kappa \bullet_{+x} \varphi)(\omega')$ if and only if $\omega \preceq_{\kappa \bullet_{+x} \varphi} \omega'$. This proves condition (S1).

For condition (S2), assume $\omega, \omega' \models \neg\varphi$. Then by Definition 5, $(\kappa \bullet_{+x} \varphi)(\omega) = \kappa(\omega) + x$ and $(\kappa \bullet_{+x} \varphi)(\omega') = \kappa(\omega') + x$. Then the proof follows the exact same scheme as the proof of condition (S1) above, showing that $\omega \preceq_\kappa \omega'$ if and only if $\omega \preceq_{\kappa \bullet_{+x} \varphi} \omega'$.

For condition (S3), if $\omega \models \varphi$ and $\omega' \models \neg\varphi$, then $(\kappa \bullet_{+x} \varphi)(\omega) = \kappa(\omega)$ and $(\kappa \bullet_{+x} \varphi)(\omega') = \kappa(\omega') + x$. Assume that $\omega \preceq_\kappa \omega'$. Then $\kappa(\omega) \leq \kappa(\omega')$, which implies that $(\kappa \bullet_{+x} \varphi)(\omega) < (\kappa \bullet_{+x} \varphi)(\omega')$, thus $\omega \prec_{\kappa \bullet_{+x} \varphi} \omega'$.

Let us now show that $\bullet_{+x}$ satisfies (I1). Let κ be a free-OCF and φ be a formula. Recall that $\kappa(\varphi) = \min_{\omega \models \varphi} \kappa(\omega)$. Let $\omega \models \varphi$ such that $\kappa(\omega) = \kappa(\varphi)$. Let k be any integer such that $k > \kappa(\omega)/x$. By definition of $\bullet_{+x}$, since $\omega \models \varphi$ we have that $(\kappa \bullet^k_{+x} \varphi)(\omega) = \kappa(\omega)$; and for each world $\omega' \models \neg\varphi$, $(\kappa \bullet^k_{+x} \varphi)(\omega') = \kappa(\omega') + k \cdot x \geq k \cdot x > (\kappa(\omega)/x) \cdot x = \kappa(\omega)$. That is, for each world $\omega' \models \neg\varphi$, $(\kappa \bullet^k_{+x} \varphi)(\omega') > (\kappa \bullet^k_{+x} \varphi)(\omega)$. Yet $[B^{\mathsf{focf}}(\kappa \bullet^k_{+x} \varphi)] = \arg\min_{\omega'' \in \Omega} (\kappa \bullet^k_{+x} \varphi)(\omega'')$, which means that $[B^{\mathsf{focf}}(\kappa \bullet^k_{+x} \varphi)] \subseteq [\varphi]$. This shows that there exists an integer k such that $B^{\mathsf{focf}}(\kappa \bullet^k_{+x} \varphi) \models \varphi$, which proves that $\bullet_{+x}$ satisfies (I1).

What remains to be proved is that the equation from Proposition 1 is satisfied. So let κ be a free-OCF and φ be a formula. We need to prove that $[B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \varphi)] = \min([\varphi], \preceq_\kappa)$.

Let us prove the first inclusion $[B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \varphi)] \subseteq \min([\varphi], \preceq_\kappa)$. So let ω be a world such that $\omega \models B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \varphi)$. Note that by (I1), $\omega \models \varphi$. Then $\omega \in \arg\min_{\omega'' \in \Omega} (\kappa \star_{\bullet_{+x}} \varphi)(\omega'')$, which means that for each world $\omega' \in \Omega$, and in particular for each world $\omega' \models \varphi$, $(\kappa \star_{\bullet_{+x}} \varphi)(\omega) \leq (\kappa \star_{\bullet_{+x}} \varphi)(\omega')$, that is, $(\kappa \bullet^k_{+x} \varphi)(\omega) \leq (\kappa \bullet^k_{+x} \varphi)(\omega')$ for some integer $k \geq 1$, by definition of $\star_{\bullet_{+x}}$. And since $\omega, \omega' \models \varphi$, by definition of $\bullet_{+x}$ we get that $(\kappa \bullet^k_{+x} \varphi)(\omega) = \kappa(\omega)$ and $(\kappa \bullet^k_{+x} \varphi)(\omega') = \kappa(\omega')$, so $\kappa(\omega) \leq \kappa(\omega')$. This shows that for each world $\omega' \models \varphi$, $\kappa(\omega) \leq \kappa(\omega')$, or equivalently, that $\omega \in \min([\varphi], \preceq_\kappa)$.

We now prove the other inclusion $\min([\varphi], \preceq_\kappa) \subseteq [B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \varphi)]$. So let ω be a world such that $\omega \in \min([\varphi], \preceq_\kappa)$, and assume toward a contradiction that $\omega \notin [B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \varphi)]$. Let $\omega' \models B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \varphi)$. Note that by (I1), $\omega' \models \varphi$. And by definition of $\star_{\bullet_{+x}}$, there is an integer k such that $\omega' \models B^{\mathsf{focf}}(\kappa \bullet^k_{+x} \varphi)$. Since we assumed that $\omega \notin [B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \varphi)]$, this means by definition of $B^{\mathsf{focf}}(\kappa \bullet^k_{+x} \varphi)$ that $(\kappa \bullet^k_{+x} \varphi)(\omega') < (\kappa \bullet^k_{+x} \varphi)(\omega)$. Yet $\omega, \omega' \models \varphi$, so by definition of $\bullet_{+x}$ we get that $(\kappa \bullet^k_{+x} \varphi)(\omega) = \kappa(\omega)$ and $(\kappa \bullet^k_{+x} \varphi)(\omega') = \kappa(\omega')$, thus $\kappa(\omega') < \kappa(\omega)$, i.e., $\omega' \prec_\kappa \omega$ by definition of $\preceq_\kappa$. This contradicts the fact that $\omega \in \min([\varphi], \preceq_\kappa)$, and proves that $\min([\varphi], \preceq_\kappa) \subseteq [B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \varphi)]$.

We showed that $\kappa \mapsto \preceq_\kappa$ is a gradual assignment and for every formula $\varphi \in \mathcal{L}^*_{\mathcal{P}}$, $[B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \varphi)] = \min([\varphi], \preceq_\kappa)$. From Proposition 1, this concludes the proof that $\bullet_{+x}$ (and by Proposition 2, $\circ_{+x}$) is an improvement operator, i.e., it satisfies (I1-I9). $\square$

In the following, we refer to the operators $\bullet_{+x}$ on free-

OCFs or $\circ_{+x}$ on OCFs interchangeably as "$+x$ improvement operators", since they are equivalent in the sense of Proposition 2.

Of particular interest is the case when $x = 1$, which is precisely the case when the resulting $+x$ improvement operator is a soft one:

Proposition 4. *A $+x$ improvement operator is a soft improvement operator (i.e., it satisfies the postulate (I10)) if and only if $x = 1$.*

Proof. Let us first prove that the operator $\bullet_{+1}$ on free-OCFs is a soft improvement operator, i.e., it satisfies (I10). From Proposition 1, it is enough to show that the gradual assignment $\kappa \mapsto \preceq_\kappa$ constructed in the proof of the previous proposition for $\bullet_{+1}$ satisfies condition (S4) of a soft gradual assignment. Let κ be any free-OCF, let φ be any formula, and let ω, ω' be two worlds such that $\omega \models \varphi$, $\omega' \models \neg\varphi$, and $\omega' \prec_\kappa \omega$. We need to prove that $\omega' \preceq_{\kappa \bullet_{+1} \varphi} \omega$. Observe by definition of $\bullet_{+1}$ that (i) $(\kappa \bullet_{+1} \varphi)(\omega) = \kappa(\omega)$ and $(\kappa \bullet_{+1} \varphi)(\omega') = \kappa(\omega') + 1$. Then from $\omega' \prec_\kappa \omega$, we get that $\kappa(\omega') < \kappa(\omega)$ (by definition of $\preceq_\kappa$), which we can also write as $\kappa(\omega') + 1 \leq \kappa(\omega)$. Then from (i) we get that $(\kappa \bullet_{+1} \varphi)(\omega') \leq (\kappa \bullet_{+1} \varphi)(\omega)$, which means that $\omega' \preceq_{\kappa \bullet_{+1} \varphi} \omega$. This shows that $\kappa \mapsto \preceq_\kappa$ satisfies (S4), and hence is a soft gradual assignment, which by Proposition 1 means that $\bullet_{+1}$ is a soft improvement operator.

We now prove that if $x > 1$, then the operator $\bullet_{+x}$ on free-OCFs does not satisfy (I10). First, note that by definition of $\bullet_{+x}$, we have that (ii) for each free-OCF κ, $\kappa \star_{\bullet_{+x}} \top = \kappa$. Let φ be any consistent, non-valid formula (i.e., $\varphi \not\models \bot$ and $\neg\varphi \not\models \bot$). Let κ be the free-OCF defined for each world $\omega \in \Omega$ as

$$\kappa(\omega) = \begin{cases} 0 & \text{if } \omega \models \neg\varphi \\ 1 & \text{if } \omega \models \varphi \end{cases}$$

That is, $B^{\mathsf{focf}}(\kappa) \equiv \neg\varphi$. We get that $B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \top) \equiv B^{\mathsf{focf}}(\kappa) \equiv \neg\varphi$. Then, by definition of $\bullet_{+x}$, the free-OCF $\kappa \bullet_{+x} \varphi$ is defined as

$$(\kappa \bullet_{+x} \varphi)(\omega) = \begin{cases} x & \text{if } \omega \models \neg\varphi \\ 1 & \text{if } \omega \models \varphi \end{cases}$$

That is, since $x > 1$, $B^{\mathsf{focf}}(\kappa \bullet_{+x} \varphi) \equiv \varphi$. We get that $B^{\mathsf{focf}}((\kappa \bullet_{+x} \varphi) \star_{\bullet_{+x}} \top) \equiv B^{\mathsf{focf}}(\kappa \bullet_{+x} \varphi) \equiv \varphi$.

We have shown that $B^{\mathsf{focf}}(\kappa \star_{\bullet_{+x}} \top) \models \neg\varphi$ and $B^{\mathsf{focf}}((\kappa \bullet_{+x} \varphi) \star_{\bullet_{+x}} \top) \models \varphi$, which shows that $\bullet_{+x}$ does not satisfy (I10), and this concludes the proof. □

This family of operators is interesting since it produces a homogeneous increase of the degree of disbelief of the countermodels of the new piece of evidence. By choosing an appropriate value of x, one can decide how much the plausibility of the new piece of information is improved at each step. If x is large enough for the new piece of information to be believed after the change, the result will be a revision, that is close to Spohn's conditionalization (Spohn 1988), since there is this similar homogeneity in the way all the countermodels' degrees change, but with a cumulative behavior that conditionalization does not have.

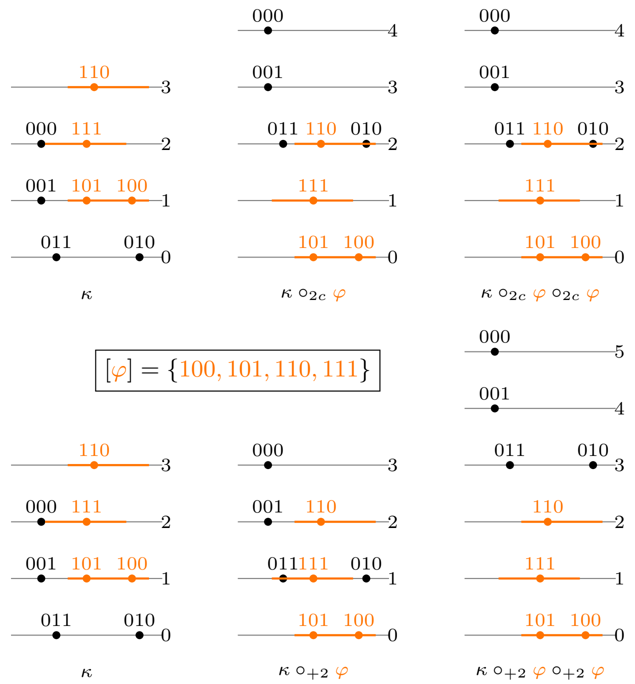

Figure 1: 2-conditionalization vs $+2$ improvement

Consider the situation of Figure 1, where $B^{\mathsf{ocf}}(\kappa) \models \neg\varphi$, and $\kappa(\varphi) = 1$. In the figure, $\circ_{2c}$ denotes 2-conditionalization, i.e., the conditionalization operator that raises the rank of each countermodel of the input formula by 2. Let us see what happens with the sequence $(\kappa \circ \varphi) \circ \varphi$, with the 2-conditionalization ($\circ_{2c}$) and with the $\circ_{+2}$ operator. Note that for one step the behaviors of $\circ_{2c}$ and $\circ_{+2}$ are very similar with a homogeneous worsening of the countermodels of φ. But more interesting is what happens after the second step. Note that the conditionalization is idempotent: $(\kappa \circ_{2c} \varphi) \circ_{2c} \varphi = \kappa \circ_{2c} \varphi$. By contrast, the $\circ_{+x}$ operators are cumulative: the worsening of the non-models of φ will never stop if we iterate by φ. This behavior explains why they seem to be good candidates for truth-tracking: if the process supplies sufficiently many correct inputs, the truth should end up more entrenched. The rest of the paper is an investigation of this hypothesis.

4 A Truth-tracking Theorem

In this section, we give two theoretical results that show that $\circ_{+x}$ operators are truth-tracking operators, i.e., that they make it possible to find the truth after a sufficiently long sequence. Among the worlds of the language, a unique one is distinguished as the actual world ω^*, and truth is evaluated with respect to this world. A belief base is interpreted as a set of worlds in which the agent expects ω^* to be. We make a distinction between correct formulas, which are formulas satisfied by the true world ω^*, and incorrect formulas, which are not. The question we address is whether the truth is found when the formulas used to improve the initial beliefs are generated by a reliable process, so that sufficiently many correct formulas are received compared to incorrect ones.

We consider a stochastic process on the subsets of Ω. At each step i of the process, a formula φ_i is randomly generated through a random selection of worlds and used to improve the beliefs. More precisely, we consider a formula as a set of worlds, and each world is a binary random variable (with value 1 if it is selected as one of the models of the formula, and 0 otherwise). We sometimes simply call the output of this stochastic process a sequence.

We work with free-OCFs in this section, in order not to have to normalize the ranking at each step, but recall that, by Proposition 2, the $+x$ operators on OCFs and free-OCFs are equivalent. Note that the process may generate an inconsistent formula. In that case, we do not apply $\bullet_{+x}$: the next free-OCF is instead defined to be the current one. The next definition specifies the sequence of free-OCFs generated by successively applying $\bullet_{+x}$ to the input formulas.

Definition 8. *Let κ be any free-OCF, $(\varphi_i)_{i\geq 1}$ a sequence of formulas, and $\bullet_{+x}$ be any $+x$ improvement operator. Then, for each world $\omega \in \Omega$:*

- *if $[\varphi_1] = \emptyset$, then $\kappa_1(\omega) = \kappa(\omega)$; otherwise $\kappa_1(\omega) = (\kappa \bullet_{+x} \varphi_1)(\omega)$;*
- *for each $i > 1$, if $[\varphi_i] = \emptyset$, then $\kappa_i(\omega) = \kappa_{i-1}(\omega)$; otherwise $\kappa_i(\omega) = (\kappa_{i-1} \bullet_{+x} \varphi_i)(\omega)$.*

Formally, this means that at each step i we have a set of random variables $\{X^i_{\omega_k}\}_{k\in\{1,\ldots,|\Omega|\}}$ with binary values $\{0,1\}$. $X^i_{\omega_k} = 1$ means that the world ω_k is a model of the formula φ_i by which the beliefs of the agent are improved, $X^i_{\omega_k} = 0$ otherwise. If $X^i_{\omega_k} = 0$ for every world ω_k, then φ_i is the inconsistent formula and the i-th improvement step is skipped. In the following, we denote the event $\{X^i_{\omega_k} = 1\}$ that a model ω_k of φ_i is selected at step i as $\{\omega_k \models \varphi_i\}$.

We now define the reliability and independence assumptions on the stochastic processes used in the truth-tracking results:

Definition 9 (Reliability)**.** *A stochastic process $(\varphi_i)_{i\geq 1}$ is* reliable *if there exists $\varepsilon > 0$ such that, for each $i \geq 1$ and each world $\omega \in \Omega$ such that $\omega \neq \omega^*$,*

$$P(\{\omega^* \models \varphi_i\}) - P(\{\omega \models \varphi_i\}) \geq \varepsilon.$$

The greatest such ε is called the reliability margin *of the process.*

Thus, a stochastic process is reliable if, at each step, the generated formula has a uniformly greater chance of having the true world as a model than any other world.

Definition 10 (Independence)**.** *A stochastic process $(\varphi_i)_{i\geq 1}$ is* independent *if all the random variables $X^i_{\omega_k}$, for $i \geq 1$ and $k \in \{1,\ldots,|\Omega|\}$, are mutually independent, i.e., if for every $n \geq 1$, every choice of distinct pairs $(i_1,k_1),\ldots,(i_n,k_n)$ with $i_j \geq 1$ and $k_j \in \{1,\ldots,|\Omega|\}$ for each $j \in \{1,\ldots,n\}$, and every choice of values $x_1,\ldots,x_n \in \{0,1\}$, we have*

$$P\left(\bigcap_{j=1}^{n}\{X^{i_j}_{\omega_{k_j}} = x_j\}\right) = \prod_{j=1}^{n} P\left(\{X^{i_j}_{\omega_{k_j}} = x_j\}\right)$$

This independence property states that all world selections are mutually independent. Thus, selecting or not selecting any finite subset of worlds at any finite subset of steps has no influence on the other such selections.

Let us now state the main technical result that allows us to show the truth-tracking properties of $+x$ improvement operators. This result states that we can, with sufficiently many formulas in the sequence, achieve any target gap between the degree of the true world and any other world, thereby ensuring that the true world is sufficiently well entrenched:

Theorem 1. *Consider a reliable and independent stochastic process $(\varphi_i)_{i\geq 1}$. For each free-OCF κ, let $(\kappa_i)_{i\geq 1}$ be defined as in Definition 8. Then, for each world $\omega \neq \omega^*$ and for each $c > 0$,*

$$\lim_{i\to+\infty} P\big(\kappa_i(\omega) - \kappa_i(\omega^*) > c\big) = 1.$$

Proof. Let $\varepsilon > 0$ be a reliability margin of the process. Fix a free-OCF κ, a world $\omega \neq \omega^*$, and a real number $c > 0$. Let $d_0 = \kappa(\omega) - \kappa(\omega^*)$.

For each step $k \geq 1$, let Y_k be the difference between the increase of the degree of ω and the increase of the degree of ω^* at step k. If $[\varphi_k] = \emptyset$, then the epistemic state is left unchanged, and therefore $Y_k = 0$. If $[\varphi_k] \neq \emptyset$, then $\bullet_{+x}$ adds x to the countermodels of φ_k and does not change the degree of its models. Hence, in all cases,

$$Y_k = x(X^k_{\omega^*} - X^k_{\omega}).$$

Indeed, if both ω^* and ω are models of φ_k, or if both are countermodels of φ_k, the two degrees receive the same increment. If ω^* is a model and ω is not, the gap increases by x. If ω is a model and ω^* is not, the gap decreases by x.

For every $n \geq 1$, we have

$$\kappa_n(\omega) - \kappa_n(\omega^*) = d_0 + \Sigma_{k=1}^n Y_k.$$

Thus it is enough to prove that, for every real number C,

$$P\big(\Sigma_{k=1}^n Y_k > C\big) \to 1$$

when $n \to +\infty$; the theorem follows by taking $C = c - d_0$.

For each $k \geq 1$, the variable Y_k belongs to $[-x, x]$. Moreover, since Y_k is a function of $X^k_{\omega^*}$ and X^k_{ω}, the variables Y_k are mutually independent by Independence. By Reliability,

$$E(Y_k) = x(P(\omega^* \models \varphi_k) - P(\omega \models \varphi_k)) \geq x\varepsilon.$$

Let $\mu_n = \Sigma_{k=1}^n E(Y_k)$. Then $\mu_n \geq nx\varepsilon$.

Let $S_n = \Sigma_{k=1}^n Y_k$ and let D be a real number. For every n such that $\mu_n > D$, and since $P(S_n \leq D) = P(S_n - \mu_n \leq D - \mu_n)$, Hoeffding's inequality (Hoeffding 1963) gives

$$P(S_n \leq D) \leq \exp\left(-\frac{2(\mu_n - D)^2}{\Sigma_{k=1}^n (2x)^2}\right).$$

Since $\mu_n \geq nx\varepsilon$, for all sufficiently large n we get

$$P(S_n \leq D) \leq \exp\left(-\frac{(nx\varepsilon - D)^2}{2nx^2}\right).$$

The exponent tends to $-\infty$ when $n \to +\infty$. Therefore $P(S_n \leq D) \to 0$, and hence $P(S_n > D) \to 1$. Taking $D = c - d_0$ gives

$$P\big(\kappa_n(\omega) - \kappa_n(\omega^*) > c\big) \to 1,$$

which proves the result. □

The proof gives an explicit upper bound on the probability that a fixed world $\omega \neq \omega^*$ is not separated from ω^* by a target cumulative gap after n steps. Let ε be the reliability margin and let x be the parameter of the $+x$ operator. With the notation of the proof, for every real number C and every n such that $nx\varepsilon > C$, Hoeffding's inequality gives

$$P(S_n \leq C) \leq \exp\left(-\frac{(nx\varepsilon - C)^2}{2nx^2}\right).$$

For $x = 1$, $\varepsilon = 0.2$, and $C = 10$, Table 1 gives this upper bound for several values of n. If the initial free-OCF gives the same degree to ω^* and ω, then $S_n \leq 10$ is exactly the event $\kappa_n(\omega) - \kappa_n(\omega^*) \leq 10$.

n	bound
100	0.6065306597
200	0.1053992246
300	0.0155038536
400	0.0021874911
500	0.0003035391

Table 1: Hoeffding upper bound on $P(S_n \leq 10)$ according to the number n of steps for the $\circ_{+1}$ operator and reliability margin 0.2.

We can now state the truth-tracking result, showing that the probability of identifying the true world converges to 1 for any sequence generated by a reliable and independent stochastic process.

Theorem 2. *Consider a reliable and independent stochastic process $(\varphi_i)_{i\geq 1}$. For each free-OCF κ, let $(\kappa_n)_{n\geq 1}$ be defined as in Definition 8. Then*

$$\lim_{n\to\infty} P([B^{\mathsf{focf}}(\kappa_n)] = \{\omega^*\}) = 1.$$

Proof. Let $(\kappa_n)_{n\geq 1}$ be defined as in Definition 8. By Theorem 1, applied with any $c > 0$, and since all values are integers, for every world $\omega \in \Omega$ such that $\omega \neq \omega^*$, $\lim_{n\to\infty} P(\kappa_n(\omega) > \kappa_n(\omega^*)) = 1$. Since Ω is finite, the probability that at least one world $\omega \neq \omega^*$ satisfies $\kappa_n(\omega) \leq \kappa_n(\omega^*)$ is at most the sum, over all such worlds ω, of the probabilities of the corresponding events: $P(\exists \omega \neq \omega^*, \kappa_n(\omega) \leq \kappa_n(\omega^*)) \leq \sum_{\omega\neq\omega^*} P(\kappa_n(\omega) \leq \kappa_n(\omega^*))$. The right-hand side tends to 0. Thus, $P(\forall \omega \neq \omega^*, \kappa_n(\omega) > \kappa_n(\omega^*)) \to 1$. With probability tending to 1, ω^* is the unique world of minimal rank in κ_n. Therefore, $\lim_{n\to\infty} P([B^{\mathsf{focf}}(\kappa_n)] = \{\omega^*\}) = 1$, which proves the result. □

It is worth noting that this theorem does not impose a minimal value for the probability p itself. It requires a uniform positive lower bound on the difference between the probability of selecting the true world and the probability of selecting any other fixed world.

5 Experiments

We now evaluate empirically the stochastic dynamics induced by the $+x$-improvement operators $\circ_{+x}$ on OCFs. This is an important issue since the results of the previous sections are limit results, so they could require too many iterations to identify the truth to be useful in practice. We show in this section that this is not the case, and how the convergence speed varies with respect to different parameters.

Note that although the theoretical analysis above was carried out using free-OCFs in order to avoid normalizing after each step, the experiments are performed here on standard OCFs. This is without loss of generality: by Proposition 2, the $+x$ operators on free-OCFs and on OCFs are equivalent. In the experiments, rankings are normalized after each step, so the true world has degree 0 whenever it is believed.

Our experimental protocol is as follows. Given an initial OCF κ, an arbitrarily chosen "true" world ω^*, and a sequence of formulas $(\varphi_i)_{i\geq 1}$ randomly generated by a reliable and independent stochastic process, we define $\kappa_0 = \kappa$ and, for each $i \geq 1$:

$$\kappa_i = \begin{cases} \kappa_{i-1} & \text{if } [\varphi_i] = \emptyset, \\ \kappa_{i-1} \circ_{+x} \varphi_i & \text{otherwise.} \end{cases}$$

Accordingly, the degree of entrenchment of the truth at step i is measured as

$$\min\{\kappa_i(\omega) - \kappa_i(\omega^*) \mid \omega \in \Omega,\ \omega \neq \omega^*\},$$

which is the gap between the true world and the most plausible competing world.

Specifically, formula generation follows the scheme described below. For each formula φ_i in any sequence:

- the true world ω^* is selected with probability p, with $1 > p > 0$;
- each world $\omega \in \Omega$ such that $\omega \neq \omega^*$ is selected with probability q, with $p > q > 0$.

This generation scheme defines a reliable and independent stochastic process. Its reliability margin is $p - q$. It is independent because each formula is generated from fresh random choices, independently from the choices used to generate the other formulas.

If no world is selected, the generated formula is inconsistent; in the experiments, this draw is ignored and a new formula is generated for the same step.

We have run a set of experiments on a batch of 1000 sequences, starting from the "empty" OCF $\kappa = \kappa_\top$ defined as $\kappa_\top(\omega) = 0$ for all worlds $\omega \in \Omega$. The value p is fixed for a given batch, while we use two complementary protocols for the choice of q:

- Fixed q: q is also a fixed value with $p > q > 0$ for the full batch;
- Random q: for each sequence in a batch, q is sampled uniformly in $(0, p)$, and is then kept fixed for that sequence (so q varies between different sequences in the same batch).

Across experiments we varied p and q and we tracked:

- the evolution of the degree of entrenchment of the truth over time
- the speed of convergence to the truth, i.e., the first step i for which $[B(\kappa_j)] = \{\omega^*\}$ for $j \in \{i, \ldots, i + 49\}$.

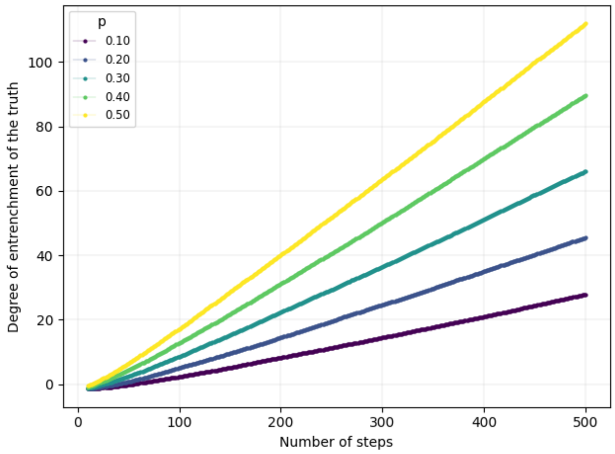

Figure 2: Evolution of the mean degree of entrenchment of the truth over time across 1000 sequences, for $p \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ and a random $q \in (0, p)$.

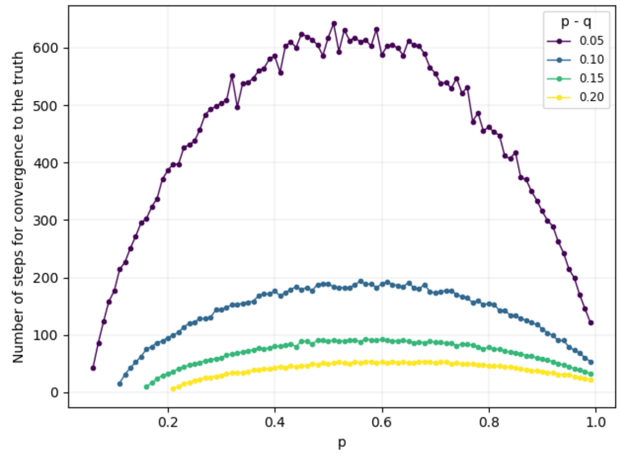

Figure 3: Mean speed of convergence to the truth across 1000 sequences, for $p \in (0, 1)$ and a fixed q, with $p - q \in \{0.05, 0.1, 0.15, 0.2\}$.

We implemented a Python package and CLI (`truth-tracking`) that makes it possible to reproduce all simulations and figures, and to perform additional simulations. It is available online[4]. The code supports single-run simulation and step-by-step traces, multi-run benchmarks, and plotting from CSV outputs.

We selected two experiments for discussion, in which we set the number of propositional variables to $n = 6$ and the operator to be $\circ_{+1}$ (i.e., $x = 1$).

Figure 2 shows the evolution of the mean degree of entrenchment of the truth over time across 1000 sequences, for $p \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ and a random $q \in (0, p)$. This degree of entrenchment is the difference between the minimal rank of any world different from the truth and the rank of the true world, so it encodes how much the agent believes the true world (as soon as the degree is positive, the true world is identified).

We are in the worst case here (since there is no fixed lower bound on the gap between p and q; this point is discussed in the next experiment).

It can be seen that the mean entrenchment grows approximately linearly with time for all selected p-values, and that the larger the value of p, the larger the entrenchment of the truth.

More interesting is Figure 3, which shows the mean speed of convergence to the truth across 1000 independent sequences, for $p \in (0, 1)$ with a fixed q, with $p - q \in \{0.05, 0.1, 0.15, 0.2\}$. In this experiment we consider that convergence is reached when the process first reaches the truth 50 consecutive times.

This experiment is interesting because the difference $p-q$ measures the reliability of the pieces of information provided to the agent. As expected, a larger gap makes the truth easier to distinguish.

Except for the curve $p - q = 0.05$ where the true world is marginally more likely to be a model of the formula than the other worlds, as soon as the difference is more than 10% it never requires more than 200 steps to achieve convergence, and less than 50 steps if the difference is 20%. So the convergence can be really quick if the provided information is highly reliable.

An interesting point is that all these curves are concave down. When p approaches 1 it is natural that convergence is quicker since the true world is in almost all the formulas of the sequence. The quick convergence for low values of p is more surprising. But it can be explained by the fact that in this case q is very close to 0, so the non-true worlds are almost never in the formulas, and much less often than the true world. Moreover, when p and q are low, random draws more often select no world. Since the experiments discard such draws and redraw for the same step, this may add a further bias toward the true world. A plausible explanation for the fact that the difficulty peaks around $p = 0.5$ is that this is where the uncertainty of the world-selection process is highest.

6 Conclusion and Future Work

In this paper, we have introduced the family of $+x$ improvement operators and we have shown that these operators allow truth-tracking, i.e., they make it possible to identify the true world when the agent receives formulas generated by a sufficiently reliable process. We also studied the speed of convergence through experiments, which illustrate the fact that in practice truth-tracking is achieved quite quickly, and that the larger the reliability margin, i.e., the difference $p-q$ between the selection probability of the true world and that of any other fixed world, the faster the convergence.

The only other work on truth-tracking for iterated belief change operators is (Baltag, Gierasimczuk, and Smets 2019). They focus on belief revision and show truth-tracking (for Nayak's lexicographic revision) only on infinite sequences of correct formulas. They also study a case where

[4] https://github.com/nicolas-schwind/truth-tracking

some formulas can be incorrect but they have to be corrected later. This is a very interesting approach on truth-tracking for iterated belief change operators, but it is built on strong hypotheses. In practice we cannot expect all provided information to be correct; the best we can expect is that the information source provides sufficiently many correct inputs. In this setting, our work demonstrates that improvement operators achieve truth-tracking even in noisy environments, by leveraging the accumulation of such inputs.

This study of truth-tracking for belief change can also be related to truth-tracking results obtained for belief merging (Everaere, Konieczny, and Marquis 2010; Karge and Rudolph 2022). As there are some links between belief merging and improvement operators (Schwind and Konieczny 2020), it would be interesting to study the links between these truth-tracking results.

As future work, it would be interesting to obtain truth-tracking results for a wider class of operators. We have proved the results here for $+x$ improvement operators. We know that revision operators cannot be helpful since they make too drastic a change at each step. But an interesting question is to identify the subclass of improvement operators that ensures truth-tracking.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP25K00375, the AI Chair BE4musIA (ANR-20-CHIA-0028), and the THEMIS project (ANR-20-CE23-0018) funded by the French National Research Agency.

AI Declaration

The authors used ChatGPT (OpenAI) for minor English corrections and rephrasing, and OpenAI Codex to assist in polishing code for the simulation software. No AI tools were used for the scientific contributions.

References

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

Baltag, A.; Gierasimczuk, N.; and Smets, S. 2019. Truth-tracking by belief revision. *Studia Logica* 107(5):917–947.

Booth, R., and Meyer, T. A. 2006. Admissible and restrained revision. *Journal of Artificial Intelligence Research* 26:127–151.

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

Darwiche, A., and Pearl, J. 1997. On the logic of iterated belief revision. *Artificial Intelligence* 89(1-2):1–29.

Everaere, P.; Konieczny, S.; and Marquis, P. 2010. The epistemic view of belief merging: Can we track the truth? In *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI'10)*, 621–626.

Fermé, E., and Hansson, S. O. 2011. AGM 25 years: Twenty-five years of research in belief change. *Journal of Philosophical Logic* 40(2):295–331.

Gärdenfors, P. 1988. *Knowledge in flux*. MIT Press.

Hoeffding, W. 1963. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* 58(301):13–30.

Jin, Y., and Thielscher, M. 2007. Iterated belief revision, revised. *Artificial Intelligence* 171(1):1–18.

Karge, J., and Rudolph, S. 2022. The More the Worst-Case-Merrier: A Generalized Condorcet Jury Theorem for Belief Fusion. In *Proceedings of the 19th International Conference on Principles of Knowledge Representation and Reasoning (KR'22)*, 205–214.

Katsuno, H., and Mendelzon, A. O. 1991. Propositional knowledge base revision and minimal change. *Artificial Intelligence* 52:263–294.

Konieczny, S., and Pino Pérez, R. 2008. Improvement operators. In *Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR'08)*, 177–187.

Konieczny, S.; Medina Grespan, M.; and Pino Pérez, R. 2010. Taxonomy of improvement operators and the problem of minimal change. In *Proceedings of the 12th International Conference on Principles of Knowledge Representation and Reasoning (KR'10)*, 161–170.

Konieczny, S. 2009. Using transfinite ordinal conditional functions. In *Proceedings of the 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'09)*, 396–407.

Medina Grespan, M., and Pino Pérez, R. 2013. Representation of basic improvement operators. In *Trends in Belief Revision and Argumentation Dynamics*. College Publications. 195–227.

Nayak, A. 1994. Iterated belief change based on epistemic entrenchment. *Erkenntnis* 41:353–390.

Schwind, N., and Konieczny, S. 2020. Non-prioritized iterated revision: Improvement via incremental belief merging. In *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning (KR'20)*, 738–747.

Schwind, N.; Inoue, K.; Konieczny, S.; and Marquis, P. 2025. Iterated belief change as learning. In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI'25)*, 4669–4677.

Schwind, N.; Konieczny, S.; and Pino Pérez, R. 2022. On the representation of Darwiche and Pearl's epistemic states for iterated belief revision. In *Proceedings of the 19th International Conference on Principles of Knowledge Representation and Reasoning (KR'22)*, 320–330.

Spohn, W. 1988. *Ordinal Conditional Functions: A Dynamic Theory of Epistemic States*. Springer Netherlands. 105–134.