Of judges, aliens and total preorders

Heltweg, Philip
University of Hagen
pheltweg@gmail.com

March 14, 2021

Adapting one's world view in the light of new information is a central skill of intelligent agents. Total preorders are a common tool to model plausibility orderings over possible worlds in the research field of belief change. In their paper "How to Revise a Total Preorder", Booth and Meyer present an approach to revising preorders for iterated belief revision. Their operator is based on assigning abstract intervals of plausibility to worlds, depending on new evidence supporting them or not.

This synopsis presents part of their work in tpo-revision operators and their properties with the help of an accompanying example and additional visualisation.

Contents

1 Introduction

A core process for an intelligent agent dealing with uncertain knowledge is updating their worldview when new information becomes available, also called belief change. As an agent can be human or machine, work in belief change has impact on the fields of both philosophy and artificial intelligence [7]. Different solutions to handling belief change have been discussed, with notable approaches including nonmonotonic logic, probabilistic reasoning and belief revision [6].

In belief revision, belief change is modelled using an operator that produces an updated set of beliefs from the state of an agent and new evidence. Research in this area discusses appropriate formalisations of agent state and constrains on individual operators or families of operators. Depending on if the suggested operator handles only one new evidence or consecutive evidence, the approaches are called one-step or iterated belief revision.

A common tool used to encode plausibility assumptions of an agent about possible worlds are total preorders (tpos) [3], also discussed in [12]. In "How to Revise a Total Preorder", Booth and Meyer describe how to model the change of these total preorders when new evidence becomes available. Being able to derive a new tpo from additional evidence is a way to handle consecutive revision steps. It therefore places the article in the field of iterated belief revision.

To support iterated belief revision it is necessary to research how tpos can be changed to adapt to an agent's changing world view. Booth and Meyer suggest an approach based on the idea of assigning an interval of additional metadata to propositional worlds and iterate the tpo of an agent using that structure. They define and discuss properties of operators for tpo-revision as well as relate their approach to other fields of research. This synopsis presents parts of their work to a broader audience. An accompanying example will guide the discussion and visualize changes in the agent's state. The focus will be on the tpo-revision operators discussed in the primary paper and explicitly avoid the addition of strict preference hierarchies and relations to other research areas.

The structure of the synopsis will be as follows: Establish the context of the paper by presenting the research area of iterated belief revision and how it relates to other areas of artificial intelligence and philosophy. To create a shared understanding of notation and definitions section 2 discusses formal background and other previous approaches to belief change. It also introduces the main contribution of this synopsis, the use of an accompanying example demonstrating tpo-revision as described by Booth and Meyer. After building this foundation, section 4 and 5 present selected parts of the work from the original paper "How to Revise a Total Preorder" [3]. Section 4 establishes and defines functions $*$, used to revise tpos. Properties of these functions are presented in section 5. To provide an outlook into the additional work done by Booth and Meyer, section 6 will show their proposal of a concrete operator for tpo-revision. It includes several examples and adapted visualisations of the operator being applied in the context of the accompanying example. Finally, section 7 closes the synopsis with a summary and commentary on Booth and Meyer's approach to iterated belief revision with interval orderings.

2 Formal background

The following notation will be used, largely aligned to Booth and Meyer [3]: A propositional language L generated from finitely many propositional variables. Lower case greek letters represent formulae in L, and $\top$ and $\bot$ represent tautology and contradiction respectively. $\models$ denotes classical logical consequence, $\equiv$ classical logical equivalence, and Cn is used to denote the deductive closure of a formula or set of formulae in L.

W is the set of propositional worlds (also called propositional interpretations in classical logic [11]). Given a formula $\alpha \in L$, the set of worlds that satisfy α is denoted by $[\![\alpha]\!]$.

For any set of worlds $S \subseteq W$, $Th(S)$ is the set of sentences true in all worlds in S.

2.1 Total preorders

Tpos are a common tool to handle preference orderings over propositional worlds in literature about belief change [3].

A total preorder is a binary relation $\leq$ (in this context over the set of propositional worlds W) that is total, reflexive and transitive (i.e. for all $x, y, z \in W$: either $x \leq y$ or $y \leq x$, $x \leq x$ and if $x \leq y$ and $y \leq z$, then $x \leq z$).

The symbol $<$ is used to denote the strict part of $\leq$ while $\sim$ represents the symmetric closure of $\leq$ (i.e. $x \sim y$ iff $x \leq y$ and $y \leq x$).

A helpful visualisation for tpos is described in [4]. It uses the fact that tpos can be represented as a linearly ordered set of ranks. Each rank of a tpo $\leq$ is defined as the equivalence classes modulo the symmetric closure of $\leq$: $[[x]]_\sim = \{y \mid y \sim x\}$. These equivalence classes are then ordered by the relation $[[x]] \leq [[y]]$ iff $x \leq y$.

Example 1. As an accompanying example, consider the following situation from a courtroom, closely aligned to Booth and Meyer [3], as well as to Darwiche and Pearl [6]:

Our agent is a judge in a murder trial. John and Mary are suspects. p represents "John is the murderer", q represents "Mary is the murderer" and r represents "the victim is an alien". Possible propositional worlds will be denoted as triplets of 0s and 1s denoting p, q and r to be true or false. For example the 101-world stands for John being the murderer and the victim being an alien.

To start the agent believes it is reasonable to assume the murderer acted alone (but is not ruling out both conspiring). In addition they consider it extremely unlikely, but not impossible, for the victim to be an alien.

Consider the propositional worlds $W = \{000, 001, 010, 011, 100, 101, 110, 111\}$. A tpo representing the judges assumptions would be $\leq$ with $010 \sim 100 < 000 \sim 110 < 011 \sim 101 < 001 \sim 111$. The equivalent representation as a linearly ordered set of ranks is shown in Table 1.

R_1	R_2	R_3	R_4
010	000	011	001
100	110	101	111

Table 1: Visualizing a tpo as a linearly ordered set of ranks, as done in [4]

.

2.2 Belief sets and epistemic states

A belief set is the deductively closed set of propositions an agent accepts as true at any given point in time [7]. Alchourrón, Gärdenfors, and Makinson (AGM theory, [1]) define postulates for belief sets and their expansion, contraction or revision with new evidence. Belief set expansion is the process of incorporating a new piece of evidence that is not inconsistent with currently held beliefs. Contraction refers to giving up a belief that has become questionable. The process of belief revision refers to keeping an agent's set of beliefs consistent, while incorporating new information that can be inconsistent with the current belief set [12, 8].

Darwiche and Pearl [6] argue that working with belief sets is not expressive enough for satisfying results in iterated belief revision. They make the distinction between propositional beliefs (beliefs the agent accepts and are part of the belief set) and conditional beliefs as beliefs the agent is prepared to adopt with new evidence.

While the AGM postulates define restrictions on revising propositional beliefs (a core one being the principle of minimal change), they don't restrict changes in conditional beliefs. In their paper, Darwiche and Pearl argue for epistemic states, abstract entities that contain all information an agent needs for their reasoning. This includes especially their strategy for belief revision in addition to their belief set. As conditional beliefs are represented by these strategies, it is necessary to define how to modify the strategy itself when encountering new information [6].

Darwiche and Pearl show that their version of belief revision can be modelled with a tpo $\leq_{\mathbb{E}}$, associated with an epistemic state $\mathbb{E}$. A mapping of a belief state $\mathbb{E}$ to a tpo $\leq_{\mathbb{E}}$ that maintains the belief set is called a faithful assignment. As Booth and Meyer describe an additional structure from which a tpo $\leq_{\mathbb{E}}$ can be determined faithful assignment is mentioned here only for completeness. For more information, refer to Katsuno and Mendelzon [12].

Belief sets can be extracted from epistemic states. The belief set extracted from an epistemic state $\mathbb{E}$ is denoted as $B(\mathbb{E})$. Extraction of $B(\mathbb{E})$ is achieved by considering the set of lowest ranked worlds (i.e. the most plausible interpretations) under $\leq_{\mathbb{E}}$. The set of propositional sentences that holds in all those worlds is defined to be the belief set of the agent. For notation, let $min(\alpha, \leq_{\mathbb{E}})$ denote the set of minimal models for the propositional formula

α under $\leq_{\mathbb{E}}$. Then $[\![B(\mathbb{E})]\!] = min(\top, \leq_{\mathbb{E}})$ is the set of worlds that are models of the belief set $B(\mathbb{E})$ associated with $\mathbb{E}$. The belief set of the agent with this notation is $Th(min(\top, \leq_{\mathbb{E}}))$.

Example 2. Continuing Example 1, it is possible to model an epistemic state $\mathbb{E}$ with the tpo $\leq_{\mathbb{E}}$.

$min(\top, \leq_{\mathbb{E}})$ is the set of worlds on the lowest rank, $\{010, 100\}$. A possible belief set is $B(\mathbb{E}) = \{p \vee q, \neg p \vee \neg q, \neg r\}$ as $[\![B(\mathbb{E})]\!] = \{010, 100\} = min(\top, \leq_{\mathbb{E}})$. Intuitively this makes sense, with the initial assumption that John or Mary are the suspects $(p \vee q)$, have acted alone $(\neg p \vee \neg q)$ and the victim is not an alien $(\neg r)$.

2.3 Belief revision postulates

A well established set of postulates for revision are the AGM postulates [1]. In this text, a reformulation for epistemic states by Darwiche and Pearl [6] is used, presented in the following. $\mathbb{E}$ denotes an epistemic state, $B(\mathbb{E})$ its associated belief set and $B(\mathbb{E}) + \alpha$ is the expansion of $B(\mathbb{E})$ by α, with $*$ being a belief change operator on epistemic states.

($\mathbb{E}$*1) $\quad B(\mathbb{E} * \alpha) = Cn(B(\mathbb{E} * \alpha))$

($\mathbb{E}$*2) $\quad \alpha \in B(\mathbb{E} * \alpha)$

($\mathbb{E}$*3) $\quad B(\mathbb{E} * \alpha) \subseteq B(\mathbb{E}) + \alpha$

($\mathbb{E}$*4) $\quad$ If $\neg\alpha \notin B(\mathbb{E})$ then $B(\mathbb{E}) + \alpha \subseteq B(\mathbb{E} * \alpha)$

($\mathbb{E}$*5) $\quad$ If $\mathbb{E} = \mathbb{F}$ and $\alpha \equiv \beta$ then $B(\mathbb{E} * \alpha) = B(\mathbb{F} * \beta)$

($\mathbb{E}$*6) $\quad \bot \in B(\mathbb{E} * \alpha)$ iff $\models \neg\alpha$

($\mathbb{E}$*7) $\quad B(\mathbb{E} * (\alpha \wedge \beta)) \subseteq B(\mathbb{E} * \alpha) + \beta$

($\mathbb{E}$*8) $\quad$ If $\neg\beta \notin B(\mathbb{E} * \alpha)$ then $B(\mathbb{E} * \alpha) + \beta \subseteq B(\mathbb{E} * (\alpha \wedge \beta))$

To guarantee the ability to extract unique belief sets from epistemic states after a revision by α, Booth and Meyer require consistent epistemic inputs [3]. That means dropping ($\mathbb{E}*6$) and considering only ($\mathbb{E}*1$)-($\mathbb{E}*5$) and ($\mathbb{E}*7$)-($\mathbb{E}*8$) for their belief revision postulates, named DP-AGM.

The DP-AGM postulates put restrictions on how the belief set of an agent changes after revision with new evidence α. As outlined before, belief sets can be extracted from epistemic states by taking the set of formulae true in all most plausible worlds after revision by α.

Darwiche and Pearl also define postulates that restrict how the rest of the new ordering changes. This extends the constrains to not only the changes in the currently held belief of an agent (the belief set), but also to the conditional beliefs it is prepared to accept depending on future evidence. Since this synopsis only focused on the revision of tpos, the semantic (i.e. in terms of how the ordering of worlds undergoes change) versions are shown here. For more details and sentential versions, refer to Darwiche and Pearl [6].

(CR1) If $v \in [\![\alpha]\!], w \in [\![\alpha]\!]$ then $v \leq_{\mathbb{E}} w$ iff $v \leq_{\mathbb{E}*\alpha} w$

(CR2) If $v \in [\![\neg\alpha]\!], w \in [\![\neg\alpha]\!]$ then $v \leq_{\mathbb{E}} w$ iff $v \leq_{\mathbb{E}*\alpha} w$

(CR3) If $v \in [\![\alpha]\!], w \in [\![\neg\alpha]\!]$ then $v <_{\mathbb{E}} w$ only if $v <_{\mathbb{E}*\alpha} w$

(CR4) If $v \in [\![\alpha]\!], w \in [\![\neg\alpha]\!]$ then $v \leq_{\mathbb{E}} w$ only if $v \leq_{\mathbb{E}*\alpha} w$

(CR1) and (CR2) mean that the relative ordering of worlds following a revision by α stays the same if the worlds are either both α- or $\neg\alpha$- worlds. (CR3) and (CR4) require that α-worlds that are strictly/weakly more plausible than $\neg\alpha$-worlds are still strictly/weakly more plausible than them after an α-revision.

3 Additional metadata for tpo revision

After discussing notation and previous research, the following sections present an excerpt of the new work done by Booth and Meyer in [3]. To support their approach to tpo revision, Booth and Meyer propose inserting additional metadata in epistemic states.

3.1 Enriching epistemic states

Based on DP-AGM and the postulates ($CR1$)-($CR4$), Booth and Meyer assume a fixed tpo $\leq$ over a set of worlds W that acts as a plausibility ordering. The goal of their paper is the discussion of functions $*$ that return a new ordering $\leq^*_\alpha$ for every $\alpha \in L$. These functions are referred to as revision operators for $\leq$.

Booth and Meyer's [3] approach is to include an additional structure with metadata for every world $w \in W$. This new structure is denoted as $W^{\pm} = \{x^{\epsilon} \mid x \in W \text{ and } \epsilon \in \{+, -\}\}$. In this notation, a world $w \in W$ is represented twice: When new evidence α arrives that makes w more plausible (because $w \in [\![\alpha]\!]$), the agent can assume its positive representation as $w^{+} \in W^{\pm}$. In contrast, if the new evidence makes w less plausible ($w \in [\![\neg\alpha]\!]$), then the negative representation $w^{-} \in W^{\pm}$ is the agent's view of w. Booth and Meyer call the pair (w^{+}, w^{-}) the positive/negative representation of the world w. They are an abstract interval, representing metadata about plausibility assumptions for w.

3.2 $\leq$-faithful tpos

For the ordering on $W^{\pm}$, Booth and Meyer suppose an additional relation, denoted $\preceq$ over $W^{\pm}$, that is added to the epistemic state of an agent.

Example 3. Recall that tpos can be equivalently thought of as a linearly ordered set of ranks (Example 1). Therefore, a tpo over W can be visualized by displaying every $w \in W$ in a table that has one column for every rank (like in Table 1). To visualize the ordering introduced by $\preceq$, Booth and Meyer use an interval with the endpoints defined as w^{+}/w^{-} respectively. This type of visualisation was first introduced in a previous paper by Booth and Meyer themselves [2] and is shown in Figure 1.

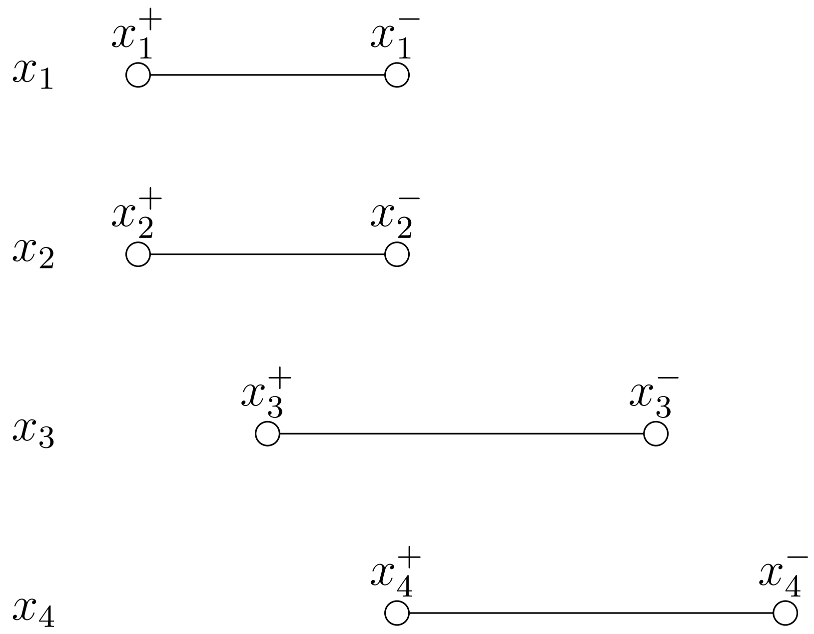

Figure 1: Representation of $\preceq$ over $W^{\pm}$ using intervals

The worlds on lower ranks (more on the left) are preferred and assumed to be more plausible. Here for example $x_1^+ \prec x_3^+$ and $x_1^- \sim x_4^+$. Note, even though all intervals have the same length in this example, that is not required.

To characterise the relations between $\preceq$ and $\leq$, Booth and Meyer define a list of conditions.

($\preceq$ 1) $\quad\preceq$ is a tpo over $W^{\pm}$

($\preceq$ 2) $\quad x^+ \preceq y^+$ iff $x \leq y$

($\preceq$ 3) $\quad x^- \preceq y^-$ iff $x \leq y$

($\preceq$ 4) $\quad x^+ \prec x^-$

The choice between positive or negative representations of two worlds should be the same as under $\leq$ (due to ($\preceq$ 2) and ($\preceq$ 3)). According to ($\preceq$ 4), there has to be a difference between a positive representation and a negative representation of the same world. Given the choice between both, the positive representation has to be chosen.

Definition 1 ($\leq$-faithful tpo over $W^{\pm}$ [3])**.** Let $\preceq \subseteq W^{\pm} \times W^{\pm}$. If $\preceq$ satisfies ($\preceq$ 1)-($\preceq$ 4), we say $\preceq$ is a $\leq$-faithful tpo (over $W^{\pm}$).

Because of ($\preceq$ 2) and ($\preceq$ 3), it is sufficient to only include $\preceq$ in the epistemic state. The tpo $\leq$ over W can be determined from $\preceq$ by restricting it to only $\{x^+ \mid x \in W\}$ or $\{x^- \mid x \in W\}$, respectively.

Example 4. Consider the following tpo $\preceq$ over $W^{\pm}$.

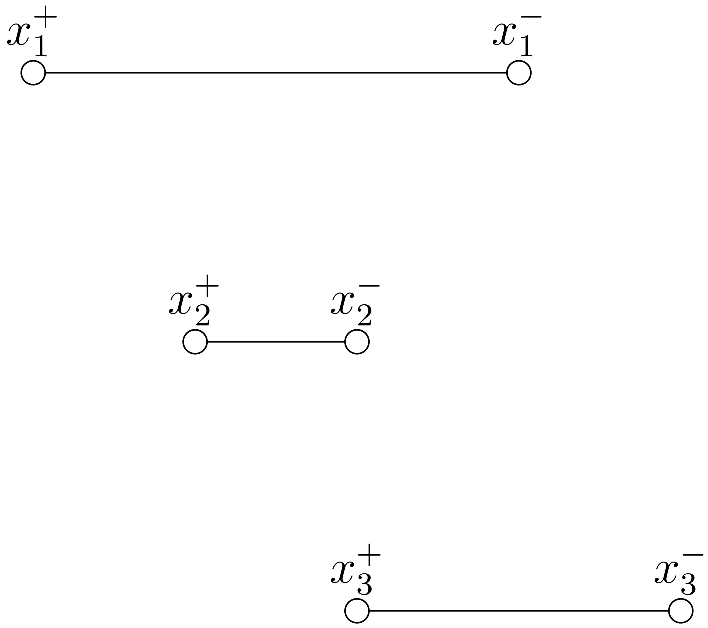

Figure 2: A tpo $\preceq$ over $W^{\pm}$

Even though ($\preceq$ 1) and ($\preceq$ 4) are satisfied, $\preceq$ is not a $\leq$-faithful tpo: From ($\preceq$ 2) and $x_1^+ \prec x_2^+$, follows $x_1 < x_2$. That means ($\preceq$ 3) can not hold as $x_2^- \prec x_1^-$ which would require $x_2 < x_3$ to be true.

As demonstrated here, even though the intervals do not need to be the same size for all worlds (they just must exist due to ($\preceq$ 4)), they need to be the same size for worlds that share a rank for their corresponding representations. They can not overlap other intervals completely to continue to satisfy ($\preceq$ 2) and ($\preceq$ 3).

Example 5. Continuing the courtroom demonstration, introduced in Example 1, of a judge deciding on a verdict on the suspects John and Mary,

Figure 3 shows a possible version of a $\leq$-faithful tpo $\preceq$ over $W^{\pm}$ from which the tpo $\leq$ from Table 1 can be reconstructed.

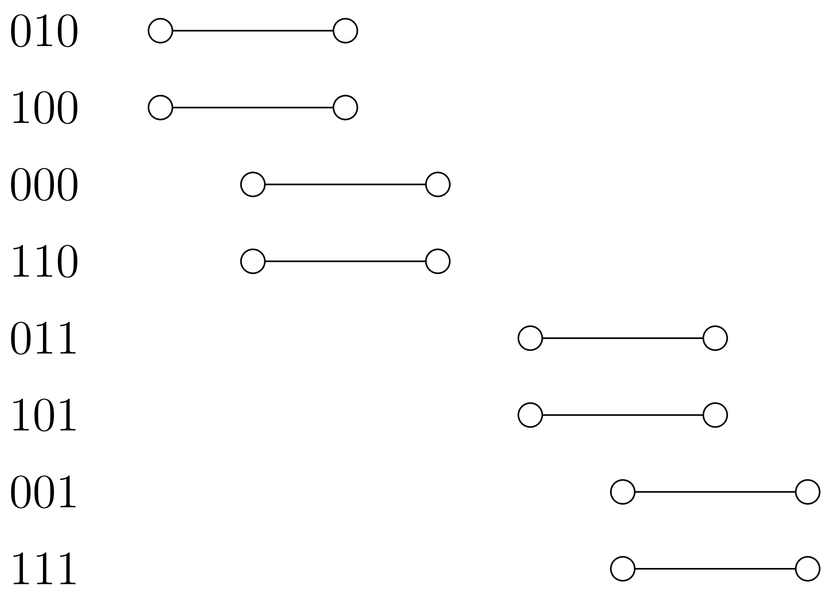

Figure 3: Representation of $\preceq$ over $W^{\pm}$ for the courtroom Example 1

Notice the "gap" between the $[\![r]\!]$-worlds that include the victim being an alien. Even in their positive representations they are not as plausible as negative representations of $[\![\neg r]\!]$-worlds. That fact will be discussed more in section 4.1 and Example 8.

4 Booth and Meyer tpo-revision operators

Booth and Meyer discuss tpo revision operators, functions $*$ that return a new tpo $\leq^*_\alpha$ for every $\alpha \in L$ [3]. They define how to use the additional information from a $\leq$-faithful tpo $\preceq$ over $W^{\pm}$ to create a revision operator $*_\preceq$ (from now on referred to as BM tpo-revision operators).

Definition 2 (Revision operator $*_\preceq$ for $\leq$ generated by $\preceq$ [3])**.** For each $\leq$-faithful tpo $\preceq$ over $W^{\pm}$, refer to $*_\preceq$ as the revision operator for $\leq$ generated by $\preceq$, defined by:

Set for any $\alpha \in L$ and $x \in W$:

$$r_\alpha(x) = \begin{cases} x^+ \text{ if } x \in [\![\alpha]\!] \\ x^- \text{ if } x \in [\![\neg\alpha]\!] \end{cases}$$

The revised tpo $\leq^*_\alpha$ is defined by setting, for each $x, y \in W$,

$$x \leq^*_\alpha y \text{ iff } r_\alpha(x) \preceq r_\alpha(y)$$

Intuitively, new evidence α makes worlds that satisfy α more plausible and worlds that do not satisfy α less plausible. Therefore, the agent associates worlds $x \in [\![\alpha]\!]$ with their positive representation x^+ and worlds $x \in [\![\neg\alpha]\!]$ with their negative representation x^-.

Example 6. Reconsider the $\leq$-faithful tpo $\preceq$ from Example 3. When a new piece of evidence α has to be considered, each world gets mapped to one end of the interval assigned to it, depending on if it is an α-world or not. In Figure 4, that mapping is indicated by the filled out dot. For this example, x_1, x_2 and x_3 are $[\![\neg\alpha]\!]$-worlds, while x_4 is an $[\![\alpha]\!]$-world.

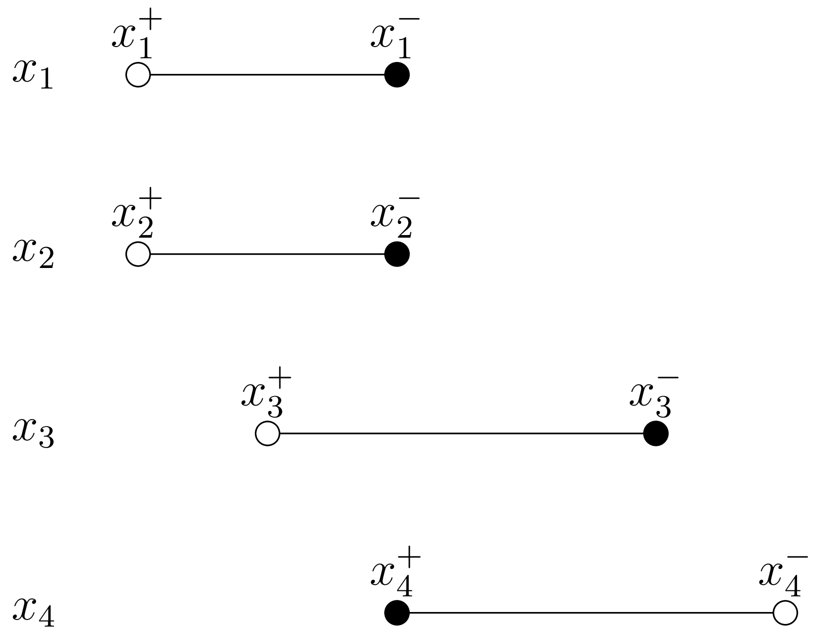

Figure 4: Associating worlds with their positive/negative representations

The updated tpo $\leq^*_\alpha$ can now be inferred from Figure 4. It is the ordering of the filled dots, representing the new assignment for the respective world: $x_1 \sim x_2 \sim x_4 < x_3$.

Example 7. A more complex example is the revision of $\leq$ to $\leq^*_\alpha$ (of the courtroom Example 1), using the $\leq$-faithful tpo $\preceq$ introduced in Example 5 and shown in Figure 5. In this case, the new evidence received points to John being the murderer so $\alpha = p$.

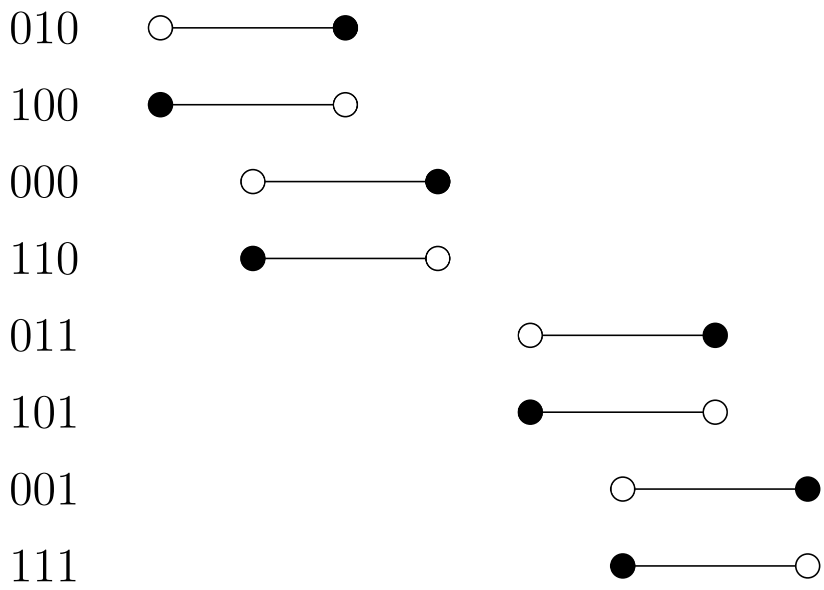

Figure 5: Associating positive and negative representations of worlds after receiving evidence $\alpha = p$

For the worlds $010, 110 \in W$, $010 < 110$ was true. To rank them with the revised tpo $\leq^*_\alpha$, calculate:

$$010 \in [\![\neg\alpha]\!] : r_\alpha(010) = 010^-$$

$$110 \in [\![\alpha]\!] : r_\alpha(110) = 110^+$$

and, as $110^+ \prec 010^-$ is true, set $110 <^*_\alpha 010$.

Repeating for every world in W, the new tpo $\leq^*_\alpha$ is: $100 <^*_\alpha 110 <^*_\alpha 010 <^*_\alpha 000 <^*_\alpha 101 <^*_\alpha 111 <^*_\alpha 011 <^*_\alpha 001$. Unsurprisingly, the belief set associated with the new epistemic state is $Th(min(\top, \leq_{\mathbb{E}})) = Th(\{100\})$: "John is the only murderer and the victim is not an alien".

As $\leq^*_\alpha$ is a representation of the conditional beliefs of the agent, it is also interesting to look at how they changed: Initially the judge thought both suspects being the murderers (110) was less plausible than only Mary being the murderer (010). Now they have updated their conditional beliefs to think only Mary being the murderer less plausible than both conspiring.

4.1 Non-prioritised revision

Belief revision with the AGM postulates [1] or the reformulation by Darwiche and Pearl [6] always includes the new information α in the belief set after revision. This feature is explicit in ($\mathbb{E}$∗2) of DP-AGM, given as $\alpha \in B(\mathbb{E} * \alpha)$.

The operators characterised by Booth and Meyer do not require new evidence α to be part of the revised belief set, which makes them part of non-prioritised revision operators [9].

A demonstration is already provided with Example 6 and Figure 4. The most plausible worlds in the new epistemic state $\mathbb{E}$, $min(\top, \leq^*_\alpha)$, include the $[\![\neg\alpha]\!]$-worlds x_1 and x_2. Therefore, $Th(min(\top, \leq^*_\alpha))$ does not include the new evidence α.

Example 8. Using the $\leq$-faithful tpo $\preceq$ already established for the courtroom Example 1 the revision by $\alpha = r$ ("the victim was an alien") is shown in Figure 6.

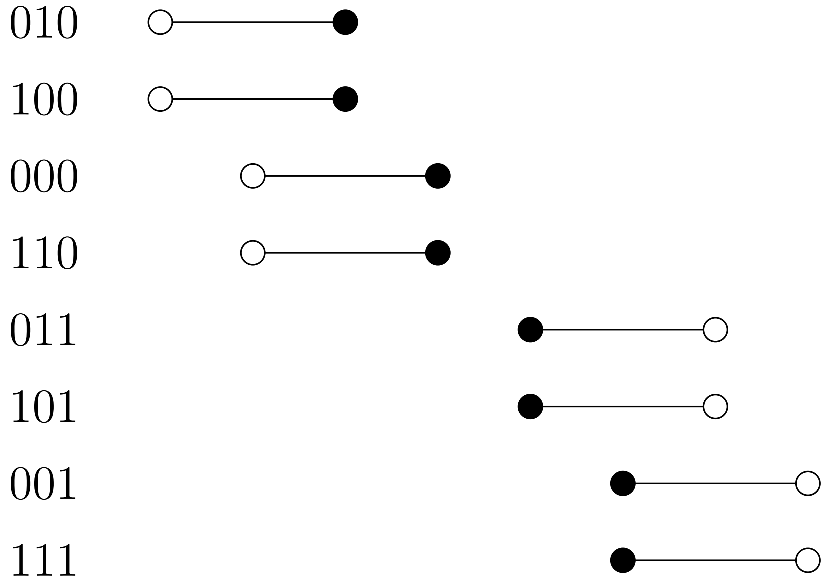

Figure 6: Non-prioritised revision by $\alpha = r$

Because the agent considered the victim being an alien very unlikely, even the positive representations of worlds in $[\![\alpha]\!]$ are on higher ranks (so considered less plausible) than negative representations of worlds in $[\![\neg\alpha]\!]$. The belief set $B(\mathbb{E})$ of the agent is unchanged under $\leq^*_\alpha$, since $\{010, 100\}$ are still the most plausible worlds, and does not include r "the victim is an alien".

5 Properties of BM tpo-revision operators

Definition 2 allows Booth and Meyer to discuss what properties any operator in that family must have [3].

The list of properties they consider to be a complete axiomatisation of the family of operators they describe is presented below:

(∗1) $\leq^*_\alpha$ is a tpo over W

(∗2) $\alpha \equiv \gamma$ implies $\leq^*_\alpha = \leq^*_\gamma$

(∗3) If $x, y \in [\![\alpha]\!]$ then $x \leq^*_\alpha y$ iff $x \leq y$

(∗4) If $x, y \in [\![\neg\alpha]\!]$ then $x \leq^*_\alpha y$ iff $x \leq y$

(∗5) If $x \in [\![\alpha]\!], y \in [\![\neg\alpha]\!]$ and $x \leq y$ then $x <^*_\alpha y$

(∗6) If $x \in [\![\alpha]\!], y \in [\![\neg\alpha]\!]$ and $y \leq^*_\alpha x$ then $y \leq^*_\gamma x$

(∗7) If $x \in [\![\alpha]\!], y \in [\![\neg\alpha]\!]$ and $y <^*_\alpha x$ then $y <^*_\gamma x$

(∗1) and (∗2) are considered basic properties: A revision of a tpo $\leq$ over W must generate another tpo over W. In addition, the operator must return an equal revised tpo for semantically equivalent sentences, the so called syntax-irrelevance property.

The next group of properties, (∗3)-(∗5) are common rules in iterated belief change. Booth and Meyer themselves consider them characteristic for admissible revision operators, a class of operators that requires new evidence to not be ignored completely [13]. This is in contrast to revision strategies like natural revision by Boutilier [5], which minimises change in conditional beliefs.

If revising a tpo by a sentence α, the relative ordering of worlds that are either both $[\![\alpha]\!]$- or $[\![\neg\alpha]\!]$-worlds, must stay the same ((∗3)/(∗4)). These properties were already part of the semantic postulates (CR1) and (CR2), by Darwiche and Pearl [6], discussed in section 2.3.

In a similar way, (∗5) is a stronger requirement than the other two Darwiche and Pearl postulates (CR3) and (CR4). It means the following: When a world x is considered at least as plausible as a world y before a revision by new information α, if α makes x more plausible ($x \in [\![\alpha]\!]$) and y less plausible ($y \in [\![\neg\alpha]\!]$), then after revision by α, x should be considered strictly more plausible than y. This property was not only proposed by Booth and Meyer [13], but also by Jin und Thielscher in [10] (as the postulate of independence).

These properties define tpo-revision with one input sentence. Booth and Meyer relate them to the AGM postulates for belief set revision [1] as basic postulates for tpo-revision [3].

With (∗6) and (∗7), Booth and Meyer aim to add supplementary rationality properties to keep revision with different sentences coherent. These properties are largely ignored in other literature on iterated belief change [3]. (∗6) means that if, after revision by evidence α that makes a world x more plausible than y, y is still considered at least as plausible as x, then for revision by any possible evidence γ, y must still be at least as plausible as x. (∗7) is the equivalent property for strict preference.

With these properties for $*$ outlined, Booth and Meyer define the family of operators they are discussing with Theorem 1.

Theorem 1. Let $*$ be any revision operator for $\leq$. Then $*$ is generated from some $\leq$-faithful tpo $\preceq$ over $W^{\pm}$ iff $*$ satisfies $(*1)$-$(*7)$. [3]

6 Iterating: $\preceq$-revision

The operator $*$, based on $\preceq$, allows the revision of a tpo $\leq$ associated with an epistemic state to $\leq^*_\alpha$. To be used in iterated belief revision, the newly introduced structure $\preceq$ also needs to be revised. Otherwise, as Booth and Meyer put it, the problem has just "re-emerged 'one level up'" [3]. In their paper, they discuss strict preference hierarchies (SPHs), an equivalent structure to $\preceq$, their properties and how to revise them.

Since this synopsis focuses on the revision of tpos itself, only the revision of a $\leq$-faithful tpo $\preceq$ using the example operator from the original paper is included here. For a more detailed explanation of the operator, SPHs or their properties, refer to Booth and Meyer [3].

6.1 A concrete operator

Booth and Meyer define their operator to revise $\preceq$ with a function $p : W^{\pm} \mapsto \mathbb{R}$. For all $x \in W$ the interval between its positive and negative representation is a real number a:

$$p(x^-) - p(x^+) = a > 0$$

The smaller $p(x^\epsilon)$ is, the more plausible the world is. The interval (x^+, x^-) represents a world $x \in W$ using its representations $x^+, x^- \in W^{\pm}$. For revision on the basis of p, it becomes an interval of real numbers $(p(x^+), p(x^-))$ with length a.

Using p, a $\leq$-faithful tpo $\preceq_p$ satisfying $(\preceq 1)$-$(\preceq 4)$ can be defined as:

$$x^\epsilon \preceq_p y^\delta \text{ iff } p(x^\epsilon) \leq p(y^\delta)$$

Revision of $\preceq$ becomes a two step process: First, choose a function p so that $\preceq_p = \preceq$; Then revise p to get a new assignment $p * \alpha$ that defines a new

$\leq$-faithful tpo $\preceq_{p*\alpha}$.

Booth and Meyer propose an operator that keeps worlds that satisfy α constant and moves $\neg\alpha$-worlds "back" by a. Other operators are possible and an area of further research [3]. They achieve this by setting p, for every $x^\epsilon \in W^\pm$, as:

$$(p * \alpha)(x^\epsilon) = \begin{cases} p(x^\epsilon) \text{ if } x \in [\![\alpha]\!] \\ p(x^\epsilon) + a \text{ if } x \in [\![\neg\alpha]\!] \end{cases}$$

By this definition, the interval representing α-worlds stays the same as $(p(x^+), p(x^-))$, while intervals of $\neg\alpha$-worlds become $(p(x^-), p(x^-) + a)$.

Example 9. For the courtroom Example 1, the revision of $\leq$ to $\leq^*_\alpha$ for $\alpha = p$ ("John is the murderer") was already done in Example 7. The next step is revising $\preceq$ by α. Figure 7 shows the revised tpo $\preceq_\alpha$ with $\neg\alpha$-worlds moved back by a. The position of their previous positive representation is displayed as a cross.

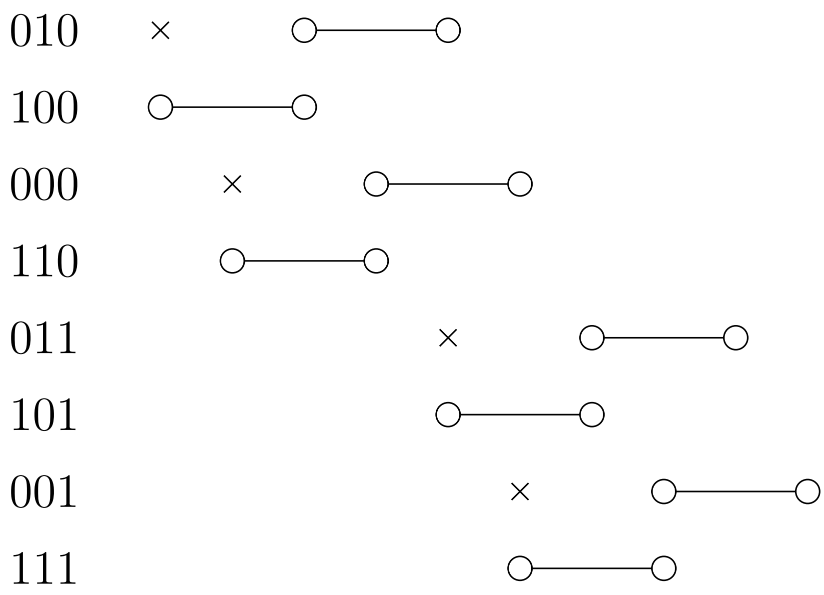

Figure 7: $\preceq_{p*\alpha}$ for $\alpha = p$

As first step, a function p is chosen so that $\preceq_p = \preceq$ and the intervals for all worlds $x \in W^\pm$ can be defined using p.

For example, the interval representing world 010 is $(010^+, 010^-)$ and becomes $(p(010^+), p(010^-)) = (0, a)$. Initially, the world 100 is represented by the same interval $(100^+, 100^-)$, which now becomes $(p(100^+), p(100^-)) = (0, a)$.

In the second step, p is revised by α to $p * \alpha$. The $\neg\alpha$-world 010 is moved back by a, while the α-world 100 is unchanged:

$$010 \in [\![\neg\alpha]\!] : (p(010^-), p(010^-) + a) = (a, 2a)$$

$$100 \in [\![\alpha]\!] : (p(100^+), p(100^-)) = (0, a)$$

Example 10. For an example that demonstrates the impact that the revision of $\preceq$ has on the acceptance of conditional beliefs, consider again the courtroom Example 1. The fact that the victim is an alien ($r =$ true) seems very unlikely, which is modelled by the gap between $[\![r]\!]$-worlds and $[\![\neg r]\!]$-worlds mentioned in Example 5.

When evidence $\alpha = r$ arrives, there seems to be little change in the epistemic state of the judge: $\leq^*_\alpha$ is unchanged, as, even considering the negative representations of $[\![\neg r]\!]$-worlds, they rank lower than $[\![r]\!]$-worlds. An unchanged tpo $\leq^*_\alpha$ also means the believe set $B(\mathbb{E})$ stays the same and does not include r (a property of non-prioritised revision, as described in item 4.1). What did change is the willingness of the agent to accept conditional beliefs about r being true in light of future evidence. This is reflected by the revised structure $\preceq_{p*\alpha}$ shown in Figure 8.

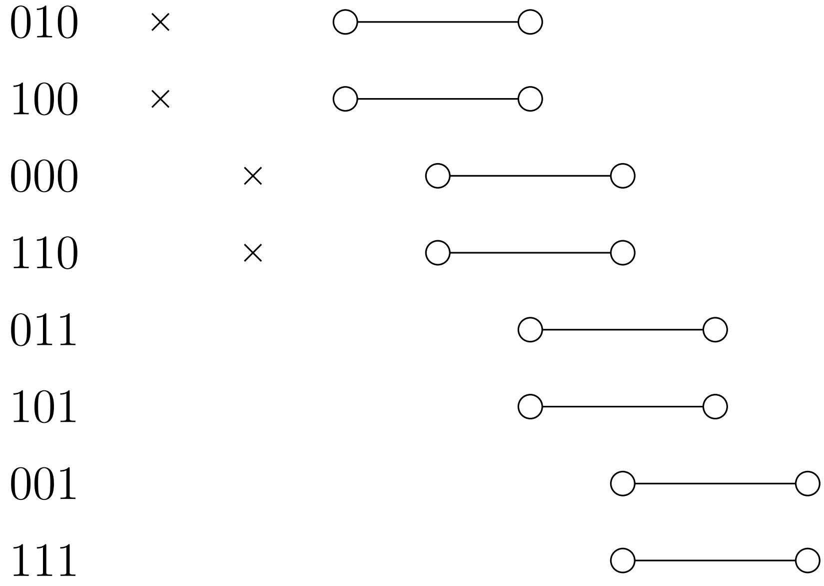

Figure 8: $\preceq_{p*\alpha}$ for $\alpha = r$

Notice how the gap between $[\![\neg r]\!]$- and $[\![r]\!]$-worlds has now closed. After revising $\preceq_{p*\alpha}$ for a second piece of evidence $\beta = (p \vee q) \wedge (\neg p \vee \neg q)$ ("the murderer was either John or Mary, not both"), the new tpo $\leq^*_{\alpha*\beta}$ can be read from the filled dots in Figure 9.

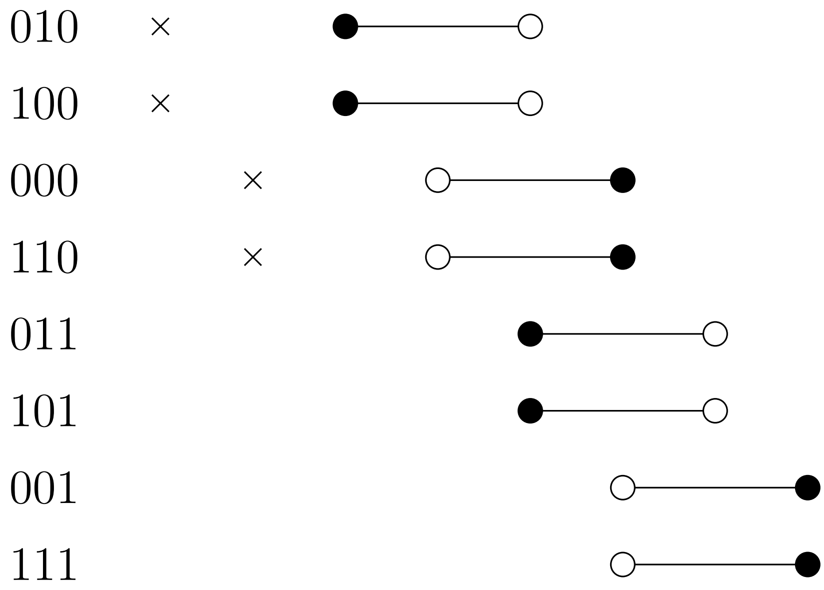

Figure 9: $\leq^*_{\alpha*\beta}$ for $\beta = (p \vee q) \wedge (\neg p \vee \neg q)$

For the initial tpo $\leq$ there was no possible evidence that would change the plausibility ordering between worlds where the victim is an alien and worlds where they are not. That was reflected by the new tpo $\leq^*_\alpha$ that, even after direct evidence for alien life, did not change the judges plausibility orderings. But because of the revision of $\preceq$ to $\preceq_{p*\alpha}$, the evidence α changed the acceptance of the judge to belief we are not alone in the universe. After receiving β, the judge even considers the $[\![r]\!]$-worlds $\{011, 101\}$ more likely than the $[\![\neg r]\!]$-worlds $\{000, 110\}$!

(Luckily, due to the existence of $[\![\beta]\!]$-worlds $\{010, 100\}$, the belief set of the judge still does not include r and rationality prevails.)

7 Final remarks

The featured paper "How to Revise a Total Preorder" by Booth and Meyer [3] is a comprehensive work on using interval orderings for belief change. This synopsis focused on presenting the research context and previous work in iterated belief change, as well as the concrete tpo-revision based on the additional structure of $\leq$-faithful tpos. A consistent example including a trial situation (adapted from [6] and [3]) was used to accompany the discussion of the paper. In addition to the provided example, the synopsis also introduced an adapted visualisation for $\leq$-faithful tpo-revision based on the interval visualisation used by Booth and Meyer themselves.

To make the presented content easier to follow, some parts of the original paper had to be excluded. Booth and Meyer discuss multiple additional properties of their proposed operators, as well as edge cases. They also provide a different way to express the $\preceq$-structure using strict preference hierarchies, and discuss how they relate to other research areas. For those topics, as well as detailed proofs, it is highly advised to read the primary paper [3] itself.

The idea to define an interval of plausibility for worlds depending on if evidence supports them or not does make intuitive sense. The proposed concrete operator for the revision of $\leq$-faithful tpos allows consecutive revisions by the same evidence α or semantically equivalent evidence $\beta \equiv \alpha$. Those would consistently lower the acceptance of $[\![\neg\alpha]\!]$-worlds without presenting new evidence. It seems like an interesting question to ask if tpo-revision for iterated belief change could be idempotent in regards to already known information.

Considering the quality and impact of new evidence also raises the question if $(*6)$ and $(*7)$ (properties of BM tpo-revision operators, section 5) should be weakened. Especially high quality evidence for a very unlikely type of world might change an agent's plausibility ordering, while low quality evidence that is very expected might not. Consider the following situation in a world described by the set of propositional variables $\{x_1, x_2, x_3, y_1\}$: New evidence $\alpha = x_1$ arrives that makes a world x, with x_1 being true, more likely. But, because α is just one piece of fairly unspecific evidence among many, the agent still considers a world y to be more plausible. This means there can be no evidence γ that would make the agent consider the world x more likely than y according to $(*7)$:

$$\text{If } x \in [\![\alpha]\!], y \in [\![\neg\alpha]\!] \text{ and } y <^*_\alpha x \text{ then } y <^*_\gamma x$$

But what if very specific evidence γ is submitted, for example $\gamma = x_1 \wedge x_2 \wedge x_3$? Because γ holds in less worlds ($|[\![\gamma]\!]| < |[\![\alpha]\!]|$) a human might consider it semantically stronger in favour of the worlds in $[\![\gamma]\!]$. Because all worlds in

$[\![\gamma]\!]$ are also in $[\![\alpha]\!]$, and the agent has already established no evidence can change their mind due to $(*7)$, even this very specific evidence can not change the relative ordering of worlds. Different operators might weaken $(*6)$ and $(*7)$ to solve this apparent contradiction by modelling different strengths of evidence.

References

[1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 6 1985.

[2] Richard Booth and Thomas Meyer. On the dynamics of total preorders: Revising abstract interval orders. *Springer*, 4724 LNAI:42–53, 2007.

[3] Richard Booth and Thomas Meyer. How to revise a total preorder. *Journal of Philosophical Logic*, 40:193–238, 4 2011.

[4] Richard Booth, Thomas Meyer, and Ka-Shu Wong. A bad day surfing is better than a good day working: How to revise a total preorder, 2006.

[5] Craig Boutilier. Iterated revision and minimal change of conditional beliefs. *Journal of Philosophical Logic*, 25:263–305, 1996.

[6] Adnan Darwiche and Judea Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89:1–29, 1997.

[7] Eduardo Fermé and Sven Ove Hansson. Agm 25 years: Twenty-five years of research in belief change article in journal of philosophical logic. *J Philos Logic*, 40:295–331, 2011.

[8] Peter Gärdenfors. Epistemic importance and minimal changes of belief. *Australasian Journal of Philosophy*, 62(2):136–157, jun 1984.

[9] Sven Ove Hansson. A survey of non-prioritized belief revision, 1999.

[10] Yi Jin and Michael Thielscher. Iterated belief revision, revised. *Artificial Intelligence*, 171:1–18, 1 2007.

[11] Sauerwald Kai, Kern-Isberner Gabriele, and Beierle Christoph. A conditional perspective for iterated belief contraction. *Frontiers in Artificial Intelligence and Applications*, 325:889–896, 2020.

[12] Hirofumi Katsuno and Alberto O. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52(3):263–294, dec 1991.

[13] Thomas Meyer and Richard Booth. Admissible and restrained revision. *Journal of Artificial Intelligence Research (JAIR)*, 26:127–151, 2006.