On the Dynamics of Total Preorders: Revising Abstract Interval Orders

Richard Booth[1] and Thomas Meyer[2]

[1] Mahasarakham University, Faculty of Informatics, Mahasarakham 44150, Thailand
richard.b@msu.ac.th
[2] Meraka Institute, CSIR, Pretoria, South Africa
tmeyer@csir.co.za

Abstract. Total preorders (tpos) are often used in belief revision to encode an agent's strategy for revising its belief set in response to new information. Thus the problem of tpo-revision is of critical importance to the problem of iterated belief revision. Booth et al. [1] provide a useful framework for revising tpos by adding extra structure to guide the revision of the initial tpo, but this results in *single-step* tpo revision only. In this paper we extend that framework to consider *double-step* tpo revision. We provide new ways of representing the structure required to revise a tpo, based on *abstract interval orders*, and look at some desirable properties for revising this structure. We prove the consistency of these properties by giving a concrete operator satisfying all of them.

1 Introduction

Total preorders (*tpos* for short) are used to represent preferences in many contexts. In particular, in the area of belief revision [2], a common way to encode an agent's *strategy* for revising its belief set is via a tpo $\leq$ over the set W of possible worlds [3, 4]. The agent's current belief set is identified with the set of sentences true in all the most preferred worlds, while upon receiving new evidence α, its *new* belief set is calculated with the help of $\leq$, typically by taking it to be the set of sentences true in all the most preferred worlds in which α holds. Of course in order to enable a *further* revision, what is needed is not just a new belief set, but also a new *tpo* to go with it. Thus the problem of *tpo-revision* is of critical importance to the problem of *iterated belief revision* [5–7].

In the problem of *belief set* revision, the tpo $\leq$ can be thought of as *extra structure* which is brought in to guide revision of the belief set. This extra structure goes *beyond* that given by the initial belief set, in the sense that the belief set can be extracted from it. Thus one natural way to attack the problem of tpo revision is to call up even more extra structure, let's denote it by X, which similarly goes beyond $\leq$ and can be used to guide revision of $\leq$. This is the approach taken by Booth et al. [1] where X takes the form of a purely qualitative structure (to be described in more detail below). Other, more quantitative forms are also conceivable [8]. Either way, X is used to determine a revised tpo $\leq^*_\alpha$ given any new evidence α. However, there is a problem with this approach regarding *iterated* tpo-revision: While the extra structure X tells us how to determine a new tpo $\leq^*_\alpha$, it tells us nothing about how to determine the *new* extra structure X^*_α to go with $\leq^*_\alpha$ which can then guide the *next* revision. Clearly the problem of iterated belief revision has simply re-emerged "one level up". The purpose of this paper is to investigate this problem in the particular case when the extra structure X takes the form studied by Booth et al. [1].

The intuition behind the family of tpo-revision operators defined by Booth et al. is that *context* ought to play a role when comparing different possible worlds according to preference. The starting point is to assume that to each possible world x are associated two abstract objects x^+ and x^-. Intuitively, x^+ will represent x in contexts favourable to it, while x^- will be the representative of x in those contexts unfavourable to it. Then, along with the initial tpo $\leq$ over W to be revised, it is assumed an agent has a tpo $\preceq$ over this entire set of objects $W^\pm$. This new tpo $\preceq$ represents the additional structure X which is used to encode the agent's strategy for revising $\leq$ in response to new evidence α. The arrival of α is seen as a context favourable to (a "good day" for) those worlds consistent with α, and a context unfavourable to (a "bad day" for) for all the other worlds. Thus the revised tpo $\leq^*_\alpha$ is obtained by setting $x \leq^*_\alpha y$ iff $x^\epsilon \preceq y^\delta$, with the values $\epsilon, \delta \in \{+, -\}$ dependent on whether x, y satisfy α or not. As was shown by Booth et al. [1], the family of tpo-revision operators so generated is characterised exactly by a relatively small list of rules, including several well-known properties which have previously been proposed. The family also includes as special cases several specific, and diverse, operators which have previously been studied [7, 9]. Thus, this framework constitutes an important contribution to single-step tpo-revision.

The plan of the paper is as follows. In Section 2 we recall the framework for single-step tpo-revision described by Booth et al. [1]. We give the formal definition of the orderings $\preceq$ described above and introduce a useful new graphical representation of these orderings in terms of *abstract intervals*. In Section 3 we introduce an alternative way of representing this structure which we call *strict preference hierarchies* (SPHs). We show that these are equivalent to the $\preceq$ orderings. A consequence of this is that the problem mentioned above of determining $\preceq^*_\alpha$ may be equivalently posed as the problem of revising SPHs. In Section 4 we consider a few desirable properties which any good operator for revising SPHs should satisfy, before proving the consistency of these properties in Section 5 by providing an example of a concrete operator which is shown to satisfy them all. We conclude and mention ideas for further research in Section 6.

Preliminaries: We work in a finitely-generated propositional language L. As mentioned above, the set of propositional worlds is denoted by W. Given a sentence $\alpha \in L$, $[\alpha]$ denotes the set of worlds which satisfy α. Classical logical equivalence over L is denoted by $\equiv$. A total preorder is any binary relation $\leq$ (or $\preceq$) which is transitive and connected. For any such relation $<$ (or $\prec$) denotes its strict part ($x < y$ iff both $x \leq y$ and $y \not\leq x$) and $\sim$ its symmetric closure ($x \sim y$ iff both $x \leq y$ and $y \leq x$). For each $\alpha \in L$ it will be useful to define the tpo $\leq^\alpha$ over W *generated by* α by setting $x \leq^\alpha y$ iff $x \in [\alpha]$ or $y \in [\neg\alpha]$.

2 Single-step Revision of Tpos

We let $W^{\pm} = \{x^{\epsilon} \mid x \in W \text{ and } \epsilon \in \{+,-\}\}$, and we assume, for any $x, y \in W$ and $\epsilon, \delta \in \{+,-\}$, that $x^{\epsilon} = y^{\delta}$ only if both $x = y$ and $\epsilon = \delta$. In other words all these abstract objects are distinct. Then we assume a given order $\preceq$ over $W^{\pm}$ satisfying the following conditions:

($\preceq$1) $\preceq$ is a total preorder
($\preceq$2) $x^+ \preceq y^+$ iff $x^- \preceq y^-$
($\preceq$3) $x^+ \prec x^-$

Rule ($\preceq$2) was split by Booth et al. [1] into two separate rules "$x^+ \preceq y^+$ iff $x \leq y$" and "$x^- \preceq y^-$ iff $x \leq y$", which made reference to an explicitly given initial tpo $\leq$ over W which is meant to be revised. However we can clearly recover $\leq$ from $\preceq$ satisfying the above three rules. We just define it by $x \leq y$ iff $x^+ \preceq y^+$ (or $x \leq y$ iff $x^- \preceq y^-$). In this case we say $\leq$ is the tpo over W *associated to* $\preceq$, or that $\preceq$ is $\leq$-*faithful*. From this $\leq$ in turn we can if we wish extract the *belief set* associated to $\preceq$: it is the set of sentences true in all the minimal $\leq$-worlds. However in this paper the dynamics of the belief set is not so much the focus as that of $\leq$, or indeed $\preceq$.

How can we picture these orderings $\preceq$? One way was given by Booth et al. [1], using an assignment of numbers to a $2 \times n$ array, where n is the number of *ranks* according to the tpo associated to $\preceq$. In this paper we would like to suggest an alternative graphical representation which is perhaps more intuitive, and is easier to work with when trying to construct examples. The idea is, for each $x \in W$, to think of the pair (x^+, x^-) as representing an *abstract interval* assigned to x. We can imagine that to each x we assign a "stick" whose left and right endpoints are x^+ and x^- respectively. Condition ($\preceq$1) says the endpoints of all these possible sticks are totally preordered. ($\preceq$2) says the left endpoints of any two of these sticks always stand in exactly same relation to each other as the right endpoints, just as if all the sticks have the same length. ($\preceq$3) demands the stick-lengths are non-zero. We may arrange the sticks in an order such as the one shown in Figure 1, which shows the sticks associated to the five worlds x_1–x_5. The further to the left an endpoint is, the lower, i.e., more preferred, it is according to $\preceq$. Thus we see for example that $x_1^+ \prec x_3^+$ and $x_2^- \sim x_4^+$.

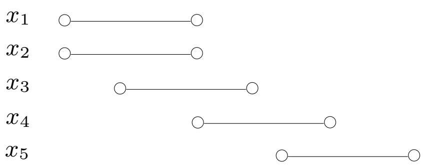


Figure 1 Example interval ordering

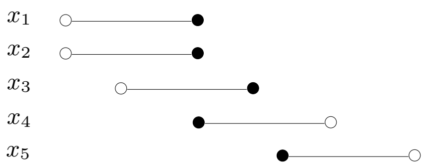


Figure 2 Example revision

We are assuming the sticks as having equal length, but this is mainly for visual convenience. It has no semantic significance in the framework.

Given a tpo $\leq$ over W, we may use a $\leq$-faithful tpo $\preceq$ over $W^{\pm}$ to define a revision operator $*$ for $\leq$. The idea is that when evidence α arrives it casts a more favourable light on worlds satisfying α. So we consider α as signalling a "good" day for the α-worlds, and a "bad day" for the $\neg\alpha$-worlds. This leads us to define a new, revised tpo $\leq^*_\alpha$ by setting, for each $x, y \in W$, $x \leq^*_\alpha y$ iff $r_\alpha(x) \preceq r_\alpha(y)$, where, for any $x \in W$ and $\alpha \in L$,

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

In terms of our new picture, each world gets mapped to one of the endpoints of the stick associated to it – left if it is an α-world and right if it is a $\neg\alpha$-world. From this the new tpo $\leq^*_\alpha$ may be read off. For example in Figure 1 suppose we revise by α such that $x_4, x_5 \in [\alpha]$ and $x_1, x_2, x_3 \in [\neg\alpha]$. Then $\leq^*_\alpha$ may be read off by looking at the black circles in Figure 2. So we see $x_1 \sim^*_\alpha x_2 \sim^*_\alpha x_4 <^*_\alpha x_3 <^*_\alpha x_5$.

Example 1. For a more concrete example (recast from one in [1] which used the old graphical representation) we assume L is generated from just two variables p, q, leading to four worlds each of which we may denote as a pair of digits denoting the truth-values of p, q respectively. The sticks associated to each world are given in Fig. 3. The initial tpo $\leq$ is specified by $10 < 11 \sim 00 < 01$. Revising by $\neg p \wedge q$ leads to Fig. 4 from which we read off $10 \sim^*_{\neg p\wedge q} 01 <^*_{\neg p\wedge q} 11 \sim^*_{\neg p\wedge q} 00$.

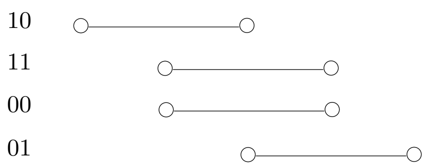


Figure 3 Example interval ordering

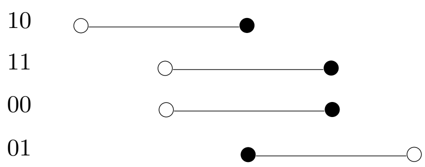


Figure 4 Example revision

If we look at the *belief set* associated to the new tpo $\leq^*_{\neg p\wedge q}$ in this example then we see it does not contain the new evidence $\neg p \wedge q$ due to the presence of world 10 among the minimal worlds in $\leq^*_{\neg p\wedge q}$. Thus we see that, at the level of belief sets, we are in the realm of so-called *non-prioritised* belief revision [10].

Given a fixed initial tpo $\leq$ over W, if the revision operator $*$ for $\leq$ can be defined from some $\leq$-faithful tpo $\preceq$ over $W^{\pm}$ as above then $*$ is said to be *generated by* $\preceq$. Booth et al. [1] characterised the class of revision operators for $\leq$ which can be generated from some $\preceq$. A revision operator $*$ can be generated from a $\leq$-faithful tpo over $W^{\pm}$ iff it satisfies the following properties for any $\alpha, \gamma \in L$:

($*1$) $\leq^*_\alpha$ 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 $y <^*_\gamma x$

Rule ($*1$) just says revising a tpo over W should result in another tpo over W. ($*2$) is a syntax-irrelevance property. ($*3$) and ($*4$) are well-known as (CR1) and (CR2) [6]. They say the relative ordering of the α-worlds, respectively the $\neg\alpha$-worlds, should remain unchanged after receiving α. ($*5$) was introduced independently by Booth et al. [5] and Jin & Thielscher [11]. It says if an α-world x was considered at least as preferred as a $\neg\alpha$-world y *before* receiving α, then *after* receiving α, x should be *strictly* preferred to y. ($*6$) says that if a world x is not more preferred to a world y, *even after* receiving evidence α which clearly points more to x being the case than it does to y, then there can be *no* evidence which will lead to x being more preferred to y. ($*7$) is similar. Rule ($*2$) is actually redundant in this list, since it can be proved from the other rules [1].

3 Strict Preference Hierarchies

A given ordering $\preceq$ over $W^\pm$ satisfying ($\preceq$1)–($\preceq$3) represents the structure required to revise its associated tpo $\leq$ over W. In this section we introduce a way of re-packaging that structure. As observed by Booth et al. [1], from a single $\preceq$ we can extract *three* different notions of *strict preference* over W. First we have the simple one given by $x < y$ iff $x^+ \prec y^+$ (equivalently $x < y$ iff $x^- \prec y^-$), i.e., $<$ is just the strict part of the tpo over W associated to $\preceq$. In terms of our new graphical representation, $x < y$ iff the stick corresponding to x lies to the left of that associated to y, but *possibly with some overlap*. For example in Figure 1 we have $x_1 < x_3$. A second, stronger notion of strict preference can be expressed by: $x \lll y$ iff $x^- \prec y^+$. In other words, $x \lll y$ iff x, *even on a bad day*, is preferred to y or, in terms of the picture, iff the stick associated to x lies *completely to the left* of that associated to y, and furthermore there is "daylight" between them. E.g., in Figure 1 we see $x_2 \lll x_5$. Finally a third case, intermediate between $\lll$ and $<$, can be expressed by: $x \ll y$ iff $x^- \preceq y^+$. In other words $x \ll y$ iff x on a bad day is *at least as* preferred to y. This third case captures a "hesitation" [12] between strong strict preference $\lll$ and mere ordinary strict preference $<$. We will have $x \ll y$ and $x \not\lll y$ precisely when the right endpoint x^- of the x-stick and the left endpoint y^+ of the y-stick are vertically aligned with each other. E.g., in Figure 1 we have $x_1 \not\lll x_4$ but $x_1 \ll x_4$.We are now in a position to define our alternative representation of the structure used by Booth et al.

Definition 1. *The triple $\mathbb{S} = (\lll, \ll, <)$ of binary relations over W is a* strict preference hierarchy (over W) *(SPH for short) iff there is some relation $\preceq$ over $W^\pm$ satisfying ($\preceq$1)–($\preceq$3) such that $\lll, \ll$ and $<$ can all be defined from $\preceq$ as above. We shall sometimes say that $\mathbb{S}$ is* relative to $<$.

Such "interval orderings" like the above have already been studied in the context of temporal reasoning [13], as well as in preference modelling [12]. Indeed, concerning the former case, the relations $\lll$, $\ll$, $<$ could all be defined in terms of the relations *before*, *meets* and *overlaps* between temporal intervals studied by Allen [13].

What are the properties of the three relations $(\lll, \ll, <)$? A couple were already mentioned by Booth et al. [1]. For example we already know from there that $\lll$ and $\ll$ are strict partial orders (i.e., irreflexive and transitive). But what else do they satisfy? In particular how do they *interrelate* with each other? Furthermore, given any *arbitrary* triple $\mathbb{S} = (\lll, \ll, <)$ of binary relations over W, under what conditions on $\mathbb{S}$ can we be sure that $\mathbb{S}$ forms an SPH, i.e., under what conditions can we be sure there is *some* $\preceq$ satisfying ($\preceq$1)–($\preceq$3) such that $\mathbb{S}$ can be derived from $\preceq$ in the above manner. These questions are answered by the following representation result for SPHs. We point out that part *(iii)* of the "only if" part (but not the "if" part) was essentially already proved, in the temporal reasoning context, by Allen [13].

Theorem 1. *Let $\lll, \ll$ and $<$ be three binary relations over W. Then $\mathbb{S} = (\lll, \ll, <)$ is an SPH iff the following conditions hold (where $x \leq y$ iff $y \not< x$):*
(i). *$\leq$ is a total preorder.*
(ii). *$\lll \subseteq \ll \subseteq <$.*
(iii). *The following are satisfied, for all $x, y, z \in W$:*

(SPH1) $z \leq x$ *and* $x \lll y$ *implies* $z \lll y$
(SPH2) $x \lll y$ *and* $y \leq z$ *implies* $x \lll z$
(SPH3) $z \leq x$ *and* $x \ll y$ *implies* $z \ll y$
(SPH4) $x \ll y$ *and* $y \leq z$ *implies* $x \ll z$
(SPH5) $z < x$ *and* $x \ll y$ *implies* $z \lll y$
(SPH6) $x \ll y$ *and* $y < z$ *implies* $x \lll z$

The rules (SPH1)–(SPH6) each represent some sort of transitivity condition across the relations of the SPH.

The "only if" direction of Theorem 1 is quite straightforward to prove, and in fact easy to visualise given our new graphical representation of $\preceq$. For the "if" direction, we may translate any triple $\mathbb{S} = (\lll, \ll, <)$ into a binary relation $\preceq_{\mathbb{S}}$ over $W^{\pm}$ as follows: Given $x^{\epsilon}, y^{\delta} \in W^{\pm}$, if $\epsilon = \delta$ then we set $x^{\epsilon} \preceq_{\mathbb{S}} y^{\epsilon}$ iff $x \leq y$. This ensures $\preceq_{\mathbb{S}}$ satisfies ($\preceq$2). If $\epsilon \neq \delta$ but $x = y$ then we declare $x^{+} \prec_{\mathbb{S}} x^{-}$. This ensures ($\preceq$3) is satisfied. Finally if $\epsilon \neq \delta$ and $x \neq y$ then we set $x^{+} \preceq_{\mathbb{S}} y^{-}$ iff $y \not\lll x$ and $x^{-} \preceq_{\mathbb{S}} y^{+}$ iff $x \ll y$. Then if $\mathbb{S}$ satisfies conditions *(i)*–*(iii)* from the theorem, then $\preceq_{\mathbb{S}}$ satisfies ($\preceq$1) in addition to ($\preceq$2) and ($\preceq$3). Furthermore the SPH corresponding to $\preceq_{\mathbb{S}}$ is precisely $\mathbb{S}$ itself.

Two special limiting cases of SPHs were already mentioned by Booth et al. [1]: Given any tpo $\leq$ over W with strict part $<$, the triples $(\emptyset, \emptyset, <)$ and $(<, <, <)$ each *always* forms an SPH, as can easily be seen by checking conditions *(i)*–*(iii)* of the theorem. In fact these are the SPH forms of the well-known lexicographic tpo-revision operator [7] and Papini's [9] "reverse" lexicographic tpo-revision operator respectively.

SPHs seem closely related to the notion of "PQI interval order" studied by Öztürk et al. [12]. Indeed several representation results in the same spirit as Theorem 1 can be found in their work. The main difference with ours is that PQI interval orders make use of an explicit numerical scale, so the endpoints of the intervals are ordinary real numbers, whereas our intervals are "abstract", having endpoints only in some totally preordered set (but see Section 5 of this paper). Also, with PQI interval orders, different possibilities (i.e., possible worlds for us) may be assigned intervals of different length. It is even possible for the interval assigned to one possibility to be completely *enclosed* in the interval assigned to another. This is something we do not allow. We are currently examining in more detail the relationship between SPHs and PQI interval orders.

To summarise the findings of this section, we now see we have two different, but equivalent ways of describing the structure required to revise a tpo $\leq$:

1. As a $\leq$-faithful tpo $\preceq$ over $W^{\pm}$ satisfying ($\preceq$1)–($\preceq$3).
2. As a triple $(\lll, \ll, <)$ of binary relations over W satisfying conditions *(i)*–*(iii)* from Theorem 1 (with $<$ being the strict part of $\leq$).

Recall that the revision operator $*$ for $\leq$ derived from a $\leq$-faithful tpo $\preceq$ over $W^{\pm}$ is defined by setting $x \leq^*_\alpha y$ iff $r_\alpha(x) \preceq r_\alpha(y)$. The next result shows how we can describe $*$ purely in terms of the SPH corresponding to $\preceq$.

Proposition 1. *Let $\leq$ be a tpo over W and let $\preceq$ be a given $\leq$-faithful tpo over $W^{\pm}$. Let $\mathbb{S} = (\lll, \ll, <)$ be the SPH corresponding to $\preceq$ and let $*$ be the revision operator for $\leq$ derived from $\preceq$. Then, for all $x, y \in W$,*

$$x \leq^*_\alpha y \text{ iff } \begin{cases} x \sim^\alpha y \text{ and } x \leq y \\ \text{or } x <^\alpha y \text{ and } y \not\lll x \\ \text{or } y <^\alpha x \text{ and } x \ll y. \end{cases}$$

Since the class of orderings $\preceq$ and the class of SPHs are equivalent, any way of revising one of these two types of structure will automatically give us a way of revising the other. We are free to use whichever one seems more appropriate at the time. For the purpose of expressing *desirable properties* of revising $\preceq$, it is easier to express such properties in terms of SPHs than $\preceq$.

4 Properties of SPH Revision

Given an SPH $\mathbb{S}$ and a sentence α, we want to determine the new SPH $\mathbb{S} \circledast \alpha$ which is the result of revising the entire SPH $\mathbb{S}$ by α. Assume $\mathbb{S} = (\lll, \ll, <)$ and let's denote $\mathbb{S} \circledast \alpha$ by $(\lll', \ll', <')$. Firstly, we have the following three fundamental properties:

($\circledast$1) $\mathbb{S} \circledast \alpha$ is an SPH
($\circledast$2) $<' = <^*_\alpha$
($\circledast$3) If $\alpha \equiv \gamma$ then $\mathbb{S} \circledast \alpha = \mathbb{S} \circledast \gamma$

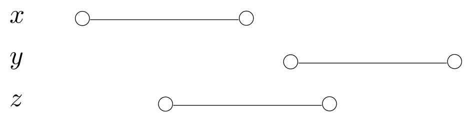
In ($\circledast$2), $<^*_\alpha$ is the strict version of the tpo $\leq^*_\alpha$ determined using $\leq$, $\ll$ and $\lll$

as in Proposition 1. In other words, $\mathbb{S} \circledast \alpha$ should be an SPH relative to $<^*_\alpha$. ($\circledast$3) is a syntax-irrelevance property.

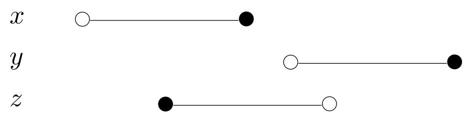
With $<'$ settled, it remains to specify $\lll'$ and $\ll'$. An initial suggestion for the new strong strict preferences $\lll'$ might be to keep it unchanged. That is, to set $\lll'$ equal to $\lll$. This can be seen as a pure application of minimal change to $\lll$. In addition, it is easy to see that $\lll \subseteq <'$ and so such a choice is not at odds with part *(ii)* of Theorem 1. However, the following example shows this can't be done in general. For $\mathbb{S} \circledast \alpha$ to be an SPH it is necessary to satisfy

(SPH1) $z \leq^*_\alpha x$ and $x \lll' y$ implies $z \lll' y$

But if we set $\lll = \lll'$ this might not hold in general. For suppose we are given a portion of the $\preceq$ corresponding to $\mathbb{S}$ as follows:



So $x \lll y$ and $z \not\lll y$. Now suppose we revise by a sentence α such that $z \in [\alpha]$ and $x, y \in [\neg\alpha]$.



Then $z <^*_\alpha x$, thus giving the required counterexample. Note, incidentally, that it is still a counterexample if we assume $y \in [\alpha]$. Thus there are times when the set of strong strict preferences *must* change. In the above counterexample, when we move from $\lll$ to $\lll'$ we must *either* lose $x \lll y$, *or* gain $z \lll y$. How do we decide which? A useful approach is to distinguish between the case $y \in [\neg\alpha]$, as indicated in the counterexample above, and the case $y \in [\alpha]$. In the former case intuition dictates that $x \lll y$ ought to be retained since α does not discriminate between x and y: they are both in $[\neg\alpha]$. Moreover, it is justifiable to gain $z \lll y$ since we have caught z on a good day ($z \in [\alpha]$) and y on a bad day ($y \in [\neg\alpha]$). On the other hand, in the case where $y \in [\alpha]$ it can be argued that the strong preference $x \lll y$ can be lost since we don't have such a strong case to prefer x over y anymore when $x \in [\neg\alpha]$ and $y \in [\alpha]$. Also, note that in this case it seems reasonable to require that the relative ordering of z and y with respect to $<$, $\ll$ and $\lll$ ought to remain unchanged since α does not distinguish between z and y: they are both in $[\alpha]$. This brings us to what can be regarded as the basic postulates for SPH revision, once ($\circledast$1)-($\circledast$3) are included as well:

($\circledast$4a) If $x \sim^\alpha y$ then $x \ll y$ iff $x \ll' y$
($\circledast$4b) If $x \sim^\alpha y$ then $x \lll y$ iff $x \lll' y$
($\circledast$5a) If $x <^\alpha y$ then $x \leq y$ implies $x \ll' y$
($\circledast$5b) If $x <^\alpha y$ then $x < y$ implies $x \lll' y$

Definition 2. *The SPH-revision operator* $\circledast$ *is* admissible *iff it satisfies* ($\circledast$*1*)-($\circledast$*3*)*,* ($\circledast$*4a*)*,* ($\circledast$*4b*)*,* ($\circledast$*5a*) *and* ($\circledast$*5b*)*.*

We refer to this as admissible SPH revision since it corresponds closely to admissible revision as defined by [5]. ($\circledast$4a) and ($\circledast$4b) are versions of Darwiche and Pearl's (CR1) and (CR2) [6], or rules ($*3$) and ($*4$) defined earlier. They require that the ordering of two elements x and y be unchanged, wrt to $\ll$ and $\lll$, provided that the circumstances for x and y are the same (i.e. either both are in $[\alpha]$ or both are in $[\neg\alpha]$). This can be seen as an application of minimal change to $\ll$ and $\lll$. The postulates ($\circledast$5a) and ($\circledast$5b) are versions of rule ($*5$) defined earlier. In fact, in the presence of the fundamental rules ($\circledast$1) and ($\circledast$2), ($\circledast$5a) is a *strengthening* of ($*5$). They ensure that a "widening of the gap" between x and y occurs when x has a good day and y a bad day. This can be viewed as making sure that the evidence α is taken seriously. A world x in $[\alpha]$ will be more preferred with respect to a world y in $[\neg\alpha]$, provided that y was not preferred to x to start with. So, informally, admissible SPH revision effects a "slide to the right" of those worlds in $[\neg\alpha]$ in a manner similar to that described by Booth et al. [5]. The difference here is that, with the aid of $\ll$ and $\lll$, we can specify more precisely how such a slide is allowed to take place.

We now turn to some additional properties and investigate how they square up against admissible SPH revision. The first one we consider is

($\circledast$6) $\mathbb{S} \circledast \top = \mathbb{S}$

which states that everything remains unchanged if we revise by a tautology. And indeed, ($\circledast$6) follows immediately from ($\circledast$2), ($\circledast$4a) and ($\circledast$4b).

Next we consider the pair of properties

($\circledast$7a) If $x \ll y$ and $x \not\ll' y$ then $y <^{\alpha} x$
($\circledast$7b) If $x \lll y$ and $x \not\lll' y$ then $y <^{\alpha} x$

which state that losing a $\ll$-preference or a $\lll$-preference of x over y must be the result of y having a good day ($y \in [\alpha]$) and x a bad day ($x \in [\neg\alpha]$). It's easy to verify that ($\circledast$7a) follows from ($\circledast$4a) and ($\circledast$5a), while ($\circledast$7b) follows from ($\circledast$4b) and ($\circledast$5b).

Next is the pair of properties

($\circledast$8a) If $x \not\ll y$ and $x \ll' y$ then $x <^{\alpha} y$
($\circledast$8b) If $x \not\lll y$ and $x \lll' y$ then $x <^{\alpha} y$

which state that *gaining* a $\ll$-preference or an $\lll$-preference of x over y must be the result of x having a good day ($x \in [\alpha]$) and y a bad day ($y \in [\neg\alpha]$). It turns out that ($\circledast$8a) follows from ($\circledast$1), ($\circledast$2) and ($\circledast$4a), while ($\circledast$8b) follows from ($\circledast$1), ($\circledast$2) and ($\circledast$4b).

Next we mention a property *not* compatible with admissible SPH revision:

($\circledast$9) If $(\lll, \ll \cap <^*_\alpha, <^*_\alpha)$ is an SPH then $\mathbb{S} \circledast \alpha = (\lll, \ll \cap <^*_\alpha, <^*_\alpha)$

Property ($\circledast$9) is an attempt to enforce the principle of minimal change with respect to both $\ll$ and $\lll$. To see that it is incompatible with admissible revision, suppose $\mathbb{S}$ is of the form $(\emptyset, \emptyset, <)$, i.e., $\lll = \ll = \emptyset$. Assume furthermore that $x < y$ and suppose we then revise by α such that $x <^{\alpha} y$. Then $(\lll, \ll \cap <^*_\alpha, <^*_\alpha) =$

$(\emptyset, \emptyset, <^*_\alpha)$ is an SPH and so ($\circledast$9) dictates that $\mathbb{S} \circledast \alpha = (\emptyset, \emptyset, <^*_\alpha)$. But observe that admissible SPH revision, and more specifically ($\circledast$5b), requires that $x \lll' y$, which contradicts $\lll' = \emptyset$.

The difference between the approach advocated by ($\circledast$9) and admissible SPH revision is that ($\circledast$9) requires all three orderings to change as little as possible, while with ($\circledast$5a) and ($\circledast$5b) we are advocating that the new evidence α overrides the principle of minimal change.

Finally we mention a couple of plausible properties which go *beyond* those of admissible revision, in that they relate the results of revising by *different* sentences. We say sentences α, γ *agree on* worlds x, y iff either $[x <^\alpha y$ and $x <^\gamma y]$ or $[x \sim^\alpha y$ and $x \sim^\gamma y]$ or $[y <^\alpha x$ and $y <^\gamma x]$. That is, α and γ both "say the same thing" regarding the relative plausibility of x, y. The next 2 rules express that whether or not $x \ll' y$ and $x \lll' y$ should depend only on $\mathbb{S}$ and on what the input sentence says about the relative plausibility between x, y. They express a principle of "Independence of Irrelevant Alternatives in the Input". Here we are writing $\mathbb{S} \circledast \alpha = (\lll^*_\alpha, \ll^*_\alpha, <^*_\alpha)$ and $\mathbb{S} \circledast \gamma = (\lll^*_\gamma, \ll^*_\gamma, <^*_\gamma)$.

($\circledast$10a) If α and γ agree on x, y then $x \ll^*_\alpha y$ iff $x \ll^*_\gamma y$
($\circledast$10b) If α and γ agree on x, y then $x \lll^*_\alpha y$ iff $x \lll^*_\gamma y$

We omit the case for $<^*_\alpha$, $<^*_\gamma$, since it was already proved to follow from $(*1)$–$(*7)$ from the Section 2 [1]. It is thus already handled by ($\circledast$2). It can be shown that adding these two rules to those for admissible revision leads to the redundancy of ($\circledast$3) and allows ($\circledast$4a) and ($\circledast$4b) to be replaced by the simple rule ($\circledast$6).

5 A Concrete Revision Operator

In the previous section we proposed that any reasonable SPH-revision operator should at the very least be admissible according to Definition 2. In this section we demonstrate that such operators exist by defining a concrete admissible operator for SPH revision. This operator employs yet more structure which goes beyond SPHs and their corresponding orderings $\preceq$ over $W^\pm$, and which is a step closer to the PQI interval orders of Öztürk et al. [12] and also to semi-quantitative representations of epistemic states such as that of Spohn [8]. But we expect there will be other, interesting, admissible revision operators which can still be defined in a purely qualitative fashion. This is a topic for further research.

To decribe our operator it will be useful to switch back to the $\preceq$-representation of our tpo-revising structure rather than work directly with SPHs. The basic idea is to enrich the $\preceq$-representation with numerical information. More precisely we assume we are given upfront some fixed function p which assigns to each element $x^\epsilon \in W^\pm$ a real number $p(x^\epsilon)$ such that for all $x \in W$, $p(x^-) - p(x^+) = a > 0$, where a is some given real number which is also fixed upfront. The idea is that the smaller the number $p(x^\epsilon)$, the more preferred x^ϵ is. To each such assignment p we may associate an ordering $\preceq_p$ over $W^\pm$ given by $x^\epsilon \preceq_p y^\delta$ iff $p(x^\epsilon) \leq p(y^\delta)$. (But note that the mapping is not on-to-one – many different choices for p can yield the same ordering over $W^\pm$.) Essentially we replace our abstract intervals (x^+, x^-) with the real intervals $(p(x^+), p(x^-))$, all of length a. It is obvious that

$\preceq_p$ satisfies ($\preceq$1)–($\preceq$3). (Again, we point out it is not *absolutely* necessary for all the intervals to be of the *same* length a in order for $\preceq_p$ to satisfy ($\preceq$2).)

To revise a given SPH $\mathbb{S}$ by sentence α we will use the following procedure:

1. Convert $\mathbb{S}$ to its corresponding tpo $\preceq$ over $W^{\pm}$
2. Choose some p such that $\preceq = \preceq_p$
3. Revise p to get a new assignment $p * \alpha$
4. Take $\mathbb{S} \circledast \alpha$ to be the SPH corresponding to $\preceq_{p*\alpha}$

Clearly the crucial step here is step 3. How should we determine $p*\alpha$? We propose a very simple method here. We define $p * \alpha$ by setting, for each $x^{\epsilon} \in W^{\pm}$,

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

In other words, the interval $(p(x^+), p(x^-))$ associated to x remains unchanged if x satisfies α, but is "moved back" by amount a to $(p(x^-), p(x^-)+a)$ if x satisfies $\neg\alpha$. Essentially this boils down to nothing more than an operation familiar from the context of Spohn-type rankings known as *L-conditionalisation* [14].

The following result reveals what $\mathbb{S} \circledast \alpha$ will look like.

Proposition 2. *Assume* $\mathbb{S} = (\lll, \ll, <)$ *and let* $\mathbb{S} \circledast \alpha = (\lll', \ll', <')$ *be as defined in the above procedure, for suitable* p *in step 2. Then, for any* $x, y \in W$*,*
(i) $<' = <^*_{\alpha}$*, where* $*$ *is the revision operator corresponding to* $\mathbb{S}$ *as in Prop. 1.*
(ii)

$$x \ll' y \text{ iff } \begin{cases} x \sim^{\alpha} y \text{ and } x \ll y \\ \text{or } x <^{\alpha} y \text{ and } x \leq y \\ \text{or } y <^{\alpha} x \text{ and } p(x^-) + a \leq p(y^+). \end{cases}$$

(iii)

$$x \lll' y \text{ iff } \begin{cases} x \sim^{\alpha} y \text{ and } x \lll y \\ \text{or } x <^{\alpha} y \text{ and } x < y \\ \text{or } y <^{\alpha} x \text{ and } p(x^-) + a < p(y^+). \end{cases}$$

From this result we can see that $\circledast$ satisfies ($\circledast$2), ($\circledast$4a), ($\circledast$4b), ($\circledast$5a) and ($\circledast$5b). We can also see from this that the result of revision depends on $[\alpha]$ rather than α, thus ($\circledast$3) is also satisfied. Meanwhile rule ($\circledast$1) obviously holds. Thus:

Corollary 1. *The SPH-revision operator* $\circledast$ *defined via the above procedure from a given assignment* p *is admissible. Furthermore (*$\circledast$*10a) and (*$\circledast$*10b) also hold.*

6 Conclusion

Motivated by the problem of iterated revision of tpos, we extended the one-step revision framework of Booth et al. [1]. We revise not only the tpo, but also the *structure* required to guide the revision of the tpo. We showed that this structure may be described in terms of strict preference hierarchies (SPHs), and proved the equivalence of this representation with that already described by Booth et al.. We gave some properties which any reasonable SPH-revision operator should satisfy, and proved their consistency by giving a concrete example of an SPH-revision operator which satisfy them.

For future work we plan to investigate more desirable properties, and to examine useful equivalent ways to reformulate the ones we already have. In this paper all our properties are formulated as rules for single-step revision of SPHs. But since an SPH encodes the structure required to revise its associated tpo, these properties correspond to properties for *double-step* revision of tpos. To give an example, property ($\circledast$5a) corresponds to the following rule governing revision of a tpo $\leq$ by α followed by β, which we denote for now by $\leq^*_{\alpha\cdot\beta}$:

$$\text{If } x <^\alpha y \text{ and } x \leq y \text{ then } x \leq^*_{\alpha\cdot\beta} y.$$

As mentioned above we intend to come up with other concrete SPH-revision operators, which perhaps can be described in purely qualitative terms rather than requiring extra numerical information like the operator described in this paper. Finally there seems to be a close connection between our work and the work done on preference modelling by Öztürk et al. [12]. The possible relationships between iterated belief revision and works such as these have, as far as we are aware, not been previously explored. We plan to look more closely at this.

References

1. Booth, R., Meyer, T., Wong, K.S.: A bad day surfing is better than a good day working: How to revise a total preorder. In: Proceedings of KR'06. (2006) 230–238
2. Hansson, S.O.: A Textbook of Belief Dynamics. Kluwer Academic Publishers (1999)
3. Grove, A.: Two modellings for theory change. Journal of Philosophical Logic **17** (1988) 157–170
4. Katsuno, H., Mendelzon, A.: Propositional knowledge base revision and minimal change. Artificial Intelligence **52** (1991) 263–294
5. Booth, R., Meyer, T.: Admissible and restrained revision. Journal of Artificial Intelligence Research **26** (2006) 127–151
6. Darwiche, A., Pearl, J.: On the logic of iterated belief revision. Artificial Intelligence **89** (1997) 1–29
7. Nayak, A., Pagnucco, M., Peppas, P.: Dynamic belief change operators. Artificial Intelligence **146** (2003) 193–228
8. Spohn, W.: Ordinal conditional functions: A dynamic theory of epistemic states. In: Causation in Decision: Belief, Change and Statistics, Kluwer, Dordrecht (1988) 105–134
9. Papini, O.: Iterated revision operations stemming from the history of an agent's observations. In: Frontiers of belief revision. Kluwer, Dordrecht (2001) 281–303
10. Hansson, S.O.: A survey of non-prioritized belief revision. Erkenntnis **50** (1999) 413–427
11. Jin, Y., Thielscher, M.: Iterated belief revision, revised. In: Proceedings of IJCAI'05. (2005) 478–483
12. Öztürk, M., Tsoukiàs, A., Vincke, P.: Preference modelling. In: Multiple Criteria Decision Analysis: State of the Art Surveys. Springer (2005) 27–72
13. Allen, J.: Maintaining knowledge about temporal intervals. Journal of the ACM **26** (1983) 832–843
14. Goldszmidt, M., Pearl, J.: Qualitative probabilities for default reasoning, belief revision and causal modeling. Artificial Intelligence **84** (1996) 57–112