Of judges, aliens and total preorders

Heltweg, Philip
pheltweg@gmail.com

University of Hagen

March 13th, 2021

Table of Contents

Table of Contents

Motivation

How should a judge change their worldview when presented with new information?

Research Context

- Philosophy and artificial intelligence [Fermé and Hansson, 2011]
- Belief change [Darwiche and Pearl, 1997]
  - Nonmonotonic logic
  - Probabilistic reasoning
  - Belief revision
- One-step vs. iterated belief revision

Different types of belief

- Belief set [Alchourrón et al., 1985]
- Conditional beliefs [Darwiche and Pearl, 1997]
- Strategy to change conditional beliefs [Booth and Meyer, 2011]

Research question

- conditional belief revision operators
- axiomatisation of a family of operators by defining properties
- discuss properties and define concrete example

Table of Contents

Courtroom example [1]

- The agent is a judge in a murder trial, "John" and "Mary" are suspects, the victim might be an alien
- $\Sigma = \{p, q, r\}$
  - p = "John is the murderer"
  - q = "Mary is the murderer"
  - r = "The victim is an alien"
- $Int(\Sigma) = W = \{000, 001, 010, 011, 100, 101, 110, 111\}$
- $Mod(p) = [\![p]\!] = \{100, 101, 110, 111\}$, $100 \in [\![p]\!]$
- Lower case greek letters used for formulas α

[1] inspired by [Booth and Meyer, 2011] and [Darwiche and Pearl, 1997]

Belief Sets

- Set of propositions the agent accepts as true at any point in time [Fermé and Hansson, 2011]
- Deductively closed
- Possible for example: $Cn(\{p \vee q, \neg(p \wedge q), \neg r\})$

Belief Set Revision Postulates[2]

- AGM theory for belief set revision
- minimal change for belief set with new information
- no restrictions on the changes in conditional beliefs

[2] AGM postulates [Alchourrón et al., 1985]

Epistemic States

- abstract entity $\mathbb{E}$ that contain all information an agent need for their reasoning [Darwiche and Pearl, 1997]
- strategy for reasoning can be modeled as tpo $\leq_{\mathbb{E}}$ over worlds
- belief sets $B(\mathbb{E})$ can be extracted from epistemic states
  - Set of most plausible worlds $min(\top, \leq_{\mathbb{E}})$
  - All sentences true in those worlds: $Th(min(\top, \leq_{\mathbb{E}}))$

Total preorders

- Common tool to handle preference orderings over propositional worlds [Booth and Meyer, 2011]
- binary relation $\leq$, total, reflexive, transitive
- $<$ strict, $\sim$ symmetric closure

Preorder example

- Judge beliefs
  - Murderer probably acted alone but possible that they conspired
  - Unlikely, but not impossible, for the victim to be an alien
- $\leq$ over W: $010 \sim 100 < 000 \sim 110 < 011 \sim 101 < 001 \sim 111$

Preorder visualisation

- $[[x]]_\sim = \{y \mid y \sim x\}$
- $[[x]] \leq [[y]]$ iff $x \leq y$

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 [Booth et al., 2006]

.

Conditional Belief Revision Postulates[3]

(CR1) If $v \in \llbracket \alpha \rrbracket, w \in \llbracket \alpha \rrbracket$ then $v \leq_{\mathbb{E}} w$ iff $v \leq_{\mathbb{E}*\alpha} w$

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

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

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

[3] by Darwiche and Pearl [Darwiche and Pearl, 1997]

Table of Contents

Enriching Epistemic States

- Additional structure $W^{\pm} = \{x^{\epsilon} \mid x \in W \text{ and } \epsilon \in \{+, -\}\}$
- Interval representing a world (w^{+}, w^{-})
- Worlds are either supported by evidence or not $w \in [\![\alpha]\!]$ / $w \in [\![\neg\alpha]\!]$

$\leq$-faithful tpo

- original tpo $\leq$ was an order over W
- $\preceq$ over new $W^{\pm}$

$\leq$-faithful tpo - definition

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

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

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

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

Definition 1 ($\leq$-faithful tpo over $W^{\pm}$ [Booth and Meyer, 2011])

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

$\leq$-faithful tpo visualisation

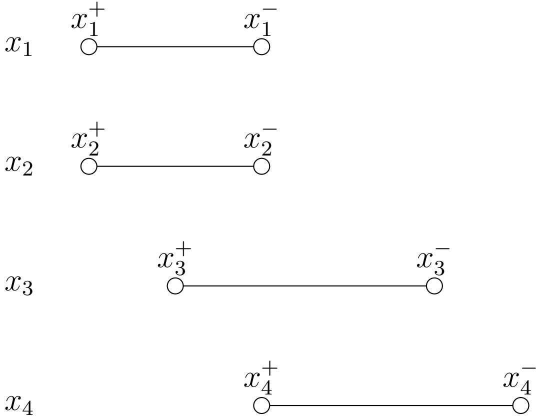

Figure 1: Representation of $\preceq$ over $W^{\pm}$ using intervals

Courtroom example: $\leq$-faithful tpo

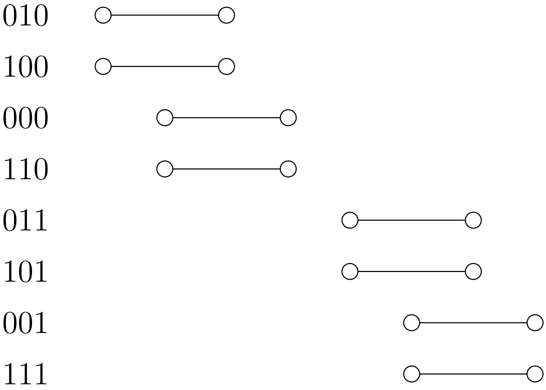

Figure 2: Representation of $\preceq$ over $W^{\pm}$ for the courtroom example

Table of Contents

BM tpo-revision operator

Definition 2 (Revision operator $*_{\preceq}$ for $\leq$ generated by $\preceq$ [Booth and Meyer, 2011])

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 $x, y \in W$,

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

Courtroom example: Revision Visualised

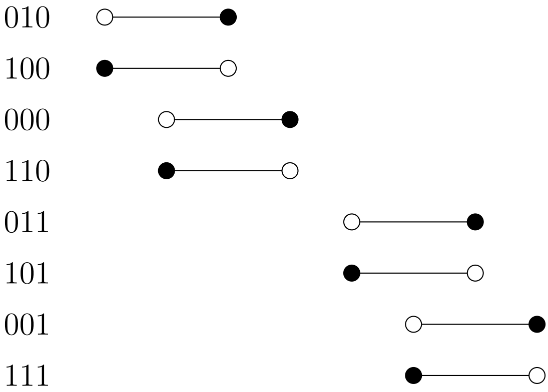

Figure 3: Associating positive and negative representations of worlds after receiving evidence $\alpha = p$

Courtroom example: Revision 1

- for $010, 110 \in W$, $010 < 110$ before, revise by $\alpha = p$

$$010 \in [\![\neg\alpha]\!] : r_\alpha(010) = 010^-$$
$$110 \in [\![\alpha]\!] : r_\alpha(110) = 110^+$$

- $110^+ \prec 010^-$ is true, set $110 <^*_\alpha 010$
- new tpo $\leq^*_\alpha$ is:
$100 <^*_\alpha 110 <^*_\alpha 010 <^*_\alpha 000 <^*_\alpha 101 <^*_\alpha 111 <^*_\alpha 011 <^*_\alpha 001$

Courtroom example: Revision 2

- new tpo $\leq^*_\alpha$ is:
  $100 <^*_\alpha 110 <^*_\alpha 010 <^*_\alpha 000 <^*_\alpha 101 <^*_\alpha 111 <^*_\alpha 011 <^*_\alpha 001$
- $min(\top, \leq^*_\alpha) = \{100\}$: "John is the murderer and the victim is not an alien".
- $\leq^*_\alpha$ as representation of the conditional beliefs
  - before $010 < 110$: "Both suspects being the murderer is less plausible than only Mary being the murderer"
  - now $110 <^*_\alpha 010$: "Only Mary being the murderer less plausible than both conspiring".

Properties of BM Revision Operators: Basic properties

$$(*1) \qquad \leq_\alpha^* \text{ is a tpo over } W$$

$$(*2) \qquad \alpha \equiv \gamma \text{ implies } \leq_\alpha^* = \leq_\gamma^*$$

Properties of BM Revision Operators: Common rules in iterated belief change

(∗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 $x <^*_\alpha y$

Properties of BM Revision Operators: Supplementary rationality properties

(∗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 $y <^*_\gamma x$

Family of BM Revision Operators

Theorem 1

Let $*$ be any revision operator for $\leq$. Then $*$ is generated from some $\leq$-faithful tpo $\preceq$ over $W^{\pm}$ iff $*$ satisfies $(*1)$-$(*7)$.
[Booth and Meyer, 2011]

Non-priotized revision

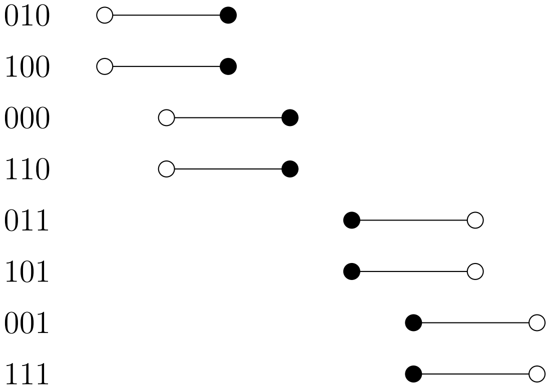

Figure 4: Non-prioritised revision by $\alpha = r$

Table of Contents

A concrete operator [4]: Setup

- Function p mapping worlds to real numbers: $p : W^{\pm} \mapsto \mathbb{R}$
- Interval representing a world x: $(p(x^+), p(x^-))$
- Distance between representations: $p(x^-) - p(x^+) = a > 0$
- Define a tpo from p: $x^\epsilon \preceq_p y^\delta$ iff $p(x^\epsilon) \leq p(y^\delta)$

[4] as shown in [Booth and Meyer, 2011]

A concrete operator: Iteration

- Choose initial p so that $\preceq_p = \preceq$
- Revise p by α to $p * \alpha$, for every $x^\epsilon \in W^\pm$:

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

- Define a revised tpo $\preceq_{p*\alpha}$ from $p * \alpha$:
  $x^\epsilon \preceq_{p*\alpha} y^\delta$ iff $(p * \alpha)(x^\epsilon) \leq (p * \alpha)(y^\delta)$

Courtroom Example: A concrete operator

- $\preceq_{p*\alpha}$ for $\alpha = p$
- Choose initial p so that $\preceq_p = \preceq$
  - $010 : (p(010^+), p(010^-)) = (0, a)$.
  - $100 : (p(100^+), p(100^-)) = (0, a)$.
- Revise p by α to $p * \alpha$
  - $010 \in [\![\neg\alpha]\!] : (p(010^-), p(010^-) + a) = (a, 2a)$
  - $100 \in [\![\alpha]\!] : (p(100^+), p(100^-)) = (0, a)$

Courtroom Example: Visualised

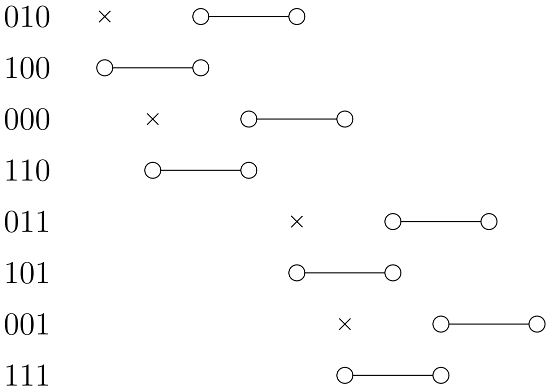

Figure 5: $\preceq_{p*\alpha}$ for $\alpha = p$

Courtroom Example: Conditional beliefs 1

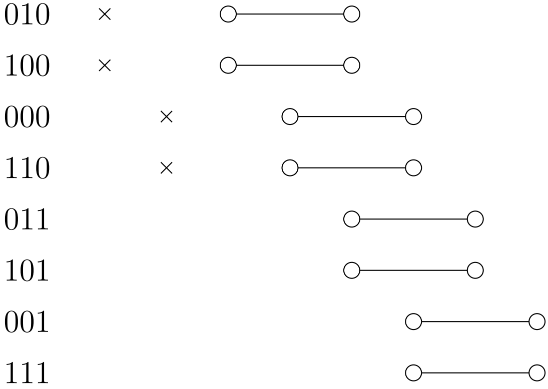

Figure 6: $\preceq_{p*\alpha}$ for $\alpha = r$

Courtroom Example: Conditional beliefs 2

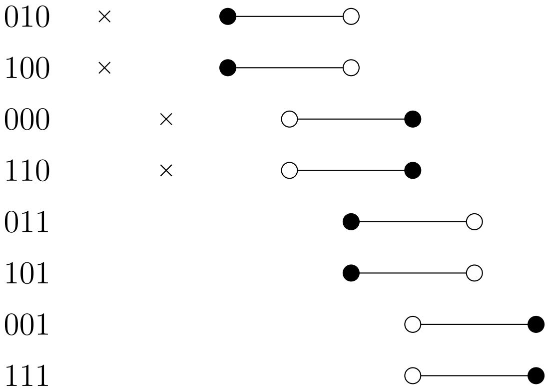

Figure 7: $\preceq_{p*\alpha*\beta}$ for $\beta = (p \vee q) \wedge (\neg p \vee \neg q)$

Thank You

References I

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.

Booth, R. and Meyer, T. (2011).
How to revise a total preorder.
Journal of Philosophical Logic, 40:193–238.

Booth, R., Meyer, T. A., and Wong, K.-S. (2006).
A bad day surfing is better than a good day working: How to revise a total preorder.
In *KR*, pages 230–238.

References II

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

Fermé, E. and Hansson, S. O. (2011).
Agm 25 years: Twenty-five years of research in belief change article in journal of philosophical logic.
J Philos Logic, 40:295–331.