Revising Incompletely Specified Convex Probabilistic Belief Bases

Gavin Rens
CAIR*,
University of KwaZulu-Natal,
School of Mathematics, Statistics and Comp. Sci.
CSIR Meraka, South Africa
Email: gavinrens@gmail.com

Thomas Meyer
CAIR,
University of Cape Town,
Dept. of Comp. Sci.
CSIR Meraka, South Africa
Email: tmeyer@cs.uct.ac.za

Giovanni Casini
University of Luxembourg,
Comp. Sci. and Communication Research Unit
Luxembourg
Email: giovanni.casini@uni.lu

Abstract

We propose a method for an agent to revise its incomplete probabilistic beliefs when a new piece of propositional information is observed. In this work, an agent's beliefs are represented by a set of probabilistic formulae – a belief base. The method involves determining a *representative* set of 'boundary' probability distributions consistent with the current belief base, revising each of these probability distributions and then translating the revised information into a new belief base. We use a version of Lewis Imaging as the revision operation. The correctness of the approach is proved. The expressivity of the belief bases under consideration are rather restricted, but has some applications. We also discuss methods of belief base revision employing the notion of optimum entropy, and point out some of the benefits and difficulties in those methods. Both the boundary distribution method and the optimum entropy method are reasonable, yet yield different results.

Suppose an agent represents its probabilistic knowledge with a set of statements; every statement says something about the probability of some features the agent is aware of. Ideally, the agent would want to have enough information to, at least, identify one probability distribution over all the situations (worlds) it deems possible. However, if the agent could not gather sufficient data or if it was not told or given sufficient information, it would not be able to pinpoint exactly one probability distribution. An agent with this sort of ignorance, can be thought of as having beliefs compatible with a *set of* distributions. Now, this agent might need to revise its beliefs when new (non-probabilistic) information is received, even though the agent's beliefs do not characterize a *particular* probability distribution over its current possible worlds.

Several researchers argue that using a single probability distribution requires the agent to make unrealistically precise uncertainty distinctions (Grove and Halpern, 1998; Voorbraak, 1999; Yue and Liu, 2008).[1] "One widely-used approach to dealing with this has been to consider sets of probability measures as a way of modeling uncertainty," (Grove and Halpern, 1998). However, simply applying standard probabilistic conditioning to each of the measures/distributions in the set individually and then combining the results is also not recommended. The framework presented in this paper proposes two ways to go from one 'probabilistically incomplete' belief base to another when new information is acquired.

Both belief revision methods presented, essentially follow this process: From the original belief base, determine a relatively small set of belief states / probability distributions 'compatible' with the belief base which is, in a sense, representative of the belief base. (We shall use the terms *belief state*, *probability distribution*, *probability function* and *distribution* interchangeably). Then revise every belief state in this representative set. Finally, induce a new, revised belief base from the revised representative set.

We shall present two approaches to determine the representative set of belief states from the current belief base: (i) The approach we focus on involves finding belief states which, in a sense, are at the boundaries of the constraints implied by the belief base. These 'boundary belief states' can be thought of as drawing the outline of the convex space of beliefs. This outline is then revised to form a new outline shape, which can be translated into a new belief base. (ii) As a possible alternative approach, the representative set is a *single* belief state which can be imagined to be at the center of the outline of the first approach. This 'central' belief state is found by determining the one in the space of beliefs which is least biased or most entropic in terms of information theory (Jaynes, 1978; Cover and Thomas, 1991).

For approach (i) – where the canonical set is the set of boundary belief states – we shall prove that the revised canonical set characterizes the set of all belief states which would have resulted from revising all (including interior) belief states compatible with the original belief base.

The relevant background theory and notations are now introduced.

We shall work with classical propositional logic. Let $\mathcal{P}$ be the finite set of atomic propositional variables (*atoms*, for short). Formally, a *world* is a unique assignment of truth values to all the atoms in $\mathcal{P}$. There are thus 2^n conceivable worlds. An agent may consider some non-empty subset W

---

[1]See also the references in these cited papers concerning criticisms against traditional probability theory.

of the conceivable worlds called the possible worlds. Often, in the exposition of this paper, a world will be referred to by its truth vector. For instance, if the vocabulary is placed in order $\langle q, r \rangle$ and $w_3 \Vdash \neg q \land r$, then w_3 may be referred to as 01.[2] Let L be all propositional formulae which can be formed from $\mathcal{P}$ and the logical connectives $\land$ and $\neg$, with $\top$ abbreviating tautology and $\bot$ abbreviating contradiction.

Let β be a sentence in L. $[\beta]$ denotes the set of β-worlds, that is, the elements of W satisfying β. The worlds satisfying all sentences in a set of sentences K are denoted by $[K]$.

We define the probabilistic language $L^{prob} = \{(\alpha) \bowtie x \mid \alpha \in L, \bowtie \in \{\le, =, \ge\}, x \in [0, 1]\}$. Sentences with strict inequalities $(<, >)$ are excluded from the language for now. Such sentences are more challenging to deal with and their inclusion is left for future work. We propose a belief base (BB) to be a consistent (logically satisfiable) subset of L^{prob}. A BB specifies an agent's knowledge.

The basic semantic element of an agent's beliefs is a probability distribution or a *belief state*

$$b = \{(w_1, p_1), (w_2, p_2), \ldots, (w_n, p_n)\},$$

where p_i is the probability that w_i is the actual world in which the agent is. $\sum_{(w,p) \in b} p = 1$. We may also use c to refer to a belief state. For parsimony, let $b = \langle p_1, \ldots, p_n \rangle$ be the probabilities that belief state b assigns to $w_1, \ldots, w_n$ where $\langle w_1, w_2, w_3, w_4 \rangle = \langle 11, 10, 01, 00 \rangle$, and $\langle w_1, w_2, \ldots, w_8 \rangle = \langle 111, 110, \ldots, 000 \rangle$. Let Π be the set of all belief states over W.

$b(\alpha)$ abbreviates $\sum_{w \in W, w \Vdash \alpha} b(w)$. b *satisfies* formula $(\alpha) \bowtie x$ (denoted $b \Vdash (\alpha) \bowtie x$) iff $b(\alpha) \bowtie x$. If B is a set of formulae, then b *satisfies* B (denoted $b \Vdash B$) iff $\forall \gamma \in B, b \Vdash \gamma$. If B and B' are sets of formulae, then B *entails* B' (denoted $B \models B'$) iff for all $b \in \Pi$, $b \Vdash B'$ whenever $b \Vdash B$. If $B \models \{\gamma\}$ then we simply write $B \models \gamma$. B is logically equivalent to B' (denoted $B \equiv B'$) iff $B \models B'$ and $B' \models B$.

Instead of an agent's beliefs being represented by a single belief state, a BB B represents a *set* of belief-states: Let $\Pi^B := \{b \in \Pi \mid b \Vdash B\}$. A BB B is *satisfiable* (*consistent*) iff $\Pi^B \ne \emptyset$.

The technique of *Lewis imaging* for the revision of belief states, requires a notion of distance between worlds to be defined. We use a pseudo-distance measure between worlds, as defined by Lehmann, Magidor, and Schlechta (2001) and adopted by Chhogyal et al. (2014).

We add a 'faithfulness' condition, which we feel is lacking from the definition of Lehmann, Magidor, and Schlechta (2001): without this condition, a pseudo-distance measure would allow all worlds to have zero distance between them. Boutilier (1998) mentions this condition, and we use his terminology: "faithfulness".

Definition 1. *A pseudo-distance function* $d : W \times W \to \mathbb{Z}$ *satisfies the following four conditions: for all worlds* $w, w', w'' \in W$,

1. $d(w, w') \ge 0$ *(Non-negativity)*

---

[2]$w \Vdash \alpha$ is read 'w is a model for/satisfies α'.

2. $d(w, w) = 0$ *(Identity)*

3. $d(w, w') = d(w', w)$ *(Symmetry)*

4. $d(w, w') + d(w', w'') \ge d(w, w'')$ *(Triangular Inequality)*

5. *if* $w \ne w'$*, then* $d(w, w') > 0$ *(Faithfulness)*

Presently, the foundation theory, or paradigm, for studying belief change operations is commonly known as AGM theory (Alchourrón, Gärdenfors, and Makinson, 1985; Gärdenfors, 1988). Typically, belief change (in a static world) can be categorized as expansion, revision or contraction, and is performed on a belief set, the set of sentences K closed under logical consequence. Expansion (denoted $+$) is the logical consequences of $K \cup \{\alpha\}$, where α is new information and K is the current belief set. Contraction of α is the removal of some sentences until α cannot be inferred from K. It is the reduction of beliefs. Revision is when α is (possibly) inconsistent with K and K is (minimally) modified so that the new K remains consistent and entails α. In this view, when the new information is consistent with the original beliefs, expansion and revision are equivalent.

The next section presents a generalized imaging method for revising probabilistic belief states. Then we describe the application of generalized imaging in our main contribution; revising boundary belief states instead of all belief states. The subsequent section explain another approaches of revising our belief bases, which prepares us for discussions in the rest of the paper. The latter method finds a single representative belief state through maximum entropy inference. Both the boundary belief state method and the maximum entropy method are reasonable, yet yield different results – a seeming paradox is thus uncovered. Then future possible directions of research are discussed. We end with a section on the related work and the concluding section.

Generalized Imaging

It is not yet universally agreed what revision means in a probabilistic setting. One school of thought says that probabilistic expansion is equivalent to Bayesian conditioning. This is evidenced by Bayesian conditioning (BC) being defined only when $b(\alpha) \ne 0$, thus making BC expansion equivalent to BC revision. In other words, one could define expansion (restricted revision) to be

$$b \, \mathsf{BC} \, \alpha = \{(w, p) \mid w \in W, p = b(w \mid \alpha), b(\alpha) \ne 0\}.$$

To accommodate cases where $b(\alpha) = 0$, that is, where α contradicts the agent's current beliefs and its beliefs need to be revised in the stronger sense, we shall make use of *imaging*. Imaging was introduced by Lewis (1976) as a means of revising a probability function. It has also been discussed in the work of, for instance, Gärdenfors (1988); Dubois and Prade (1993); Chhogyal et al. (2014); Rens and Meyer (2015). Informally, Lewis's original solution for accommodating contradicting evidence α is to move the probability of each world to its closest, α-world. Lewis made the strong assumption that every world has a *unique* closest α-world. More general versions of imaging allows worlds to have *several*, equally proximate, closest worlds.

Gärdenfors (1988) calls one of his generalizations of Lewis's imaging *general imaging*. Our method is also a generalization. We thus refer to his as *Gärdenfors's general imaging* and to our method as *generalized imaging* to distinguish them. It should be noted that all three these imaging methods are general revision methods and can be used in place of Bayesian conditioning for expansion. "Thus imaging is a more general method of describing belief changes than conditionalization," (Gärdenfors, 1988, p. 112).

Let $Min(\alpha, w, d)$ be the set of α-worlds closest to w with respect to pseudo-distance d. Formally,

$$Min(\alpha, w, d) :=$$
$$\{w' \in [\alpha] \mid \forall w'' \in [\alpha], d(w', w) \leq d(w'', w)\},$$

where $d(\cdot)$ is some pseudo-distance measure between worlds (e.g., Hamming or Dalal distance).

Example 1. *Let the vocabulary be* $\{q, r, s\}$. *Let* α *be* $(q \wedge r) \vee (q \wedge \neg r \wedge s)$. *Suppose* d *is Hamming distance. Then*

$$Min((q \wedge r) \vee (q \wedge \neg r \wedge s), 111, d) = \{111\}$$
$$Min((q \wedge r) \vee (q \wedge \neg r \wedge s), 110, d) = \{110\}$$
$$Min((q \wedge r) \vee (q \wedge \neg r \wedge s), 101, d) = \{101\}$$
$$Min((q \wedge r) \vee (q \wedge \neg r \wedge s), 100, d) = \{110, 101\}$$
$$Min((q \wedge r) \vee (q \wedge \neg r \wedge s), 011, d) = \{111\}$$
$$Min((q \wedge r) \vee (q \wedge \neg r \wedge s), 010, d) = \{110\}$$
$$Min((q \wedge r) \vee (q \wedge \neg r \wedge s), 001, d) = \{101\}$$
$$Min((q \wedge r) \vee (q \wedge \neg r \wedge s), 000, d) = \{110, 101\}$$

$\square$

Definition 2 (GI). *Then* generalized imaging *(denoted* GI*) is defined as*

$$b \text{ GI } \alpha := \{(w, p) \mid w \in W, p = 0 \text{ if } w \notin [\alpha],$$
$$else \; p = \sum_{\substack{w' \in W \\ w \in Min(\alpha, w', d)}} b(w')/|Min(\alpha, w', d)|\}.$$

In words, $b \text{ GI } \alpha$ is the new belief state produced by taking the generalized image of b with respect to α. Notice how the probability mass of non-α-worlds is shifted to their closest α-worlds. If a non-α-world $w^{\times}$ with probability p has n closest α-worlds (equally distant), then each of these closest α-worlds gets p/n mass from $w^{\times}$.

We define $b_{\alpha}^{\circ} := b \circ \alpha$ so that we can write $b_{\alpha}^{\circ}(w)$, where $\circ$ is a revision operator.

Example 2. *Continuing on Example 1: Let* $b = \langle 0, 0.1, 0, 0.2, 0, 0.3, 0, 0.4 \rangle$.
$(q \wedge r) \vee (q \wedge \neg r \wedge s)$ *is abbreviated as* α.

$$b_{\alpha}^{\text{GI}}(111) = \sum_{\substack{w' \in W \\ 111 \in Min(\alpha, w', d)}} b(w')/|Min(\alpha, w', d)|$$
$= b(111)/|Min(\alpha, 111, d)| + b(011)/|Min(\alpha, 011, d)| =$
$0/1 + 0/1 = 0.$

$$b_{\alpha}^{\text{GI}}(110) = \sum_{\substack{w' \in W \\ 110 \in Min(\alpha, w', d)}} b(w')/|Min(\alpha, w', d)|$$
$= b(110)/|Min(\alpha, 110, d)| + b(100)/|Min(\alpha, 100, d)| +$

$b(010)/|Min(\alpha, 010, d)| + b(000)/|Min(\alpha, 000, d)| =$
$0.1/1 + 0.2/2 + 0.3/1 + 0.4/2 = 0.7.$

$$b_{\alpha}^{\text{GI}}(101) = \sum_{\substack{w' \in W \\ 101 \in Min(\alpha, w', d)}} b(w')/|Min(\alpha, w', d)|$$
$= b(101)/|Min(\alpha, 101, d)| + b(100)/|Min(\alpha, 100, d)| +$
$b(001)/|Min(\alpha, 001, d)| + b(000)/|Min(\alpha, 000, d)| =$
$0/1 + 0.2/2 + 0/1 + 0.4/2 = 0.3.$

And $b_{\alpha}^{\text{GI}}(100) = b_{\alpha}^{\text{GI}}(011) = b_{\alpha}^{\text{GI}}(010) = b_{\alpha}^{\text{GI}}(001) = b_{\alpha}^{\text{GI}}(000) = 0.$ $\square$

Revision via GI and boundary belief states

Perhaps the most obvious way to revise a given belief base (BB) B is to revise every individual belief state in Π^B and then induce a new BB from the set of revised belief states. Formally, given observation α, first determine a new belief state b^{α} for every $b \in \Pi^B$ via the defined revision operation:

$$\Pi^{B^{\alpha}} = \{b^{\alpha} \in \Pi \mid b^{\alpha} = b \text{ GI } \alpha, b \in \Pi^B\}.$$

If there is more than only a single belief state in Π^B, then Π^B contains an infinite number of belief states. Then how can one compute $\Pi^{B^{\alpha}}$? And how would one subsequently determine B^{α} from $\Pi^{B^{\alpha}}$?

In the rest of this section we shall present a finite method of determining $\Pi^{B^{\alpha}}$. What makes this method possible is the insight that Π^B can be represented by a finite set of 'boundary' belief states – those belief states which, in a sense, represent the limits or the convex hull of Π^B. We shall prove that the set of revised boundary belief states defines $\Pi^{B^{\alpha}}$. Inducing B^{α} from $\Pi^{B^{\alpha}}$ is then relatively easy, as will be seen.

Let W^{perm} be every permutation on the ordering of worlds in W. For instance, if $W = \{w_1, w_2, w_3, w_4\}$, then $W^{perm} = \{\langle w_1, w_2, w_3, w_4 \rangle, \langle w_1, w_2, w_4, w_3 \rangle, \langle w_1, w_3, w_2, w_4 \rangle, \ldots, \langle w_4, w_3, w_2, w_1 \rangle\}$. Given an ordering $W^{\#} \in W^{perm}$, let $W^{\#}(i)$ be the i-th element of $W^{\#}$; for instance, $\langle w_4, w_3, w_2, w_1 \rangle(2) = w_3$. Suppose we are given a BB B. We now define a function which, given a permutation of worlds, returns a belief state where worlds earlier in the ordering are assigned maximal probabilities according to the boundary values enforced by B.

Definition 3. $MaxASAP(B, W^{\#})$ *is the* $b \in \Pi^B$ *such that for* $i = 1, \ldots, |W|$, $\forall b' \in \Pi^B$, *if* $b' \neq b$, *then* $\sum_{j=1}^{i} b(W^{\#}(j)) \geq \sum_{k=1}^{i} b'(W^{\#}(k))$.

Example 3. *Suppose the vocabulary is* $\{q, r\}$ *and* $B_1 = \{(q) \geq 0.6\}$. *Then, for instance,* $MaxASAP(B_1, \langle 01, 00, 11, 10 \rangle) = \{(01, 0.4), (00, 0), (11, 0.6), (10, 0)\} = \{(11, 0.6), (10, 0), (01, 0.4), (00, 0)\}.$ $\square$

Definition 4. *We define the boundary belief states of BB* B *as the set*

$$\Pi_{bnd}^B := \{b \in \Pi^B \mid$$
$$W^{\#} \in W^{perm}, b = MaxASAP(B, W^{\#})\}$$

Note that $|\Pi_{bnd}^B| \leq |W^{perm}|$.

Example 4. *Suppose the vocabulary is $\{q, r\}$ and $B_1 = \{(q) \geq 0.6\}$. Then*

$$
\begin{aligned}
\Pi^{B_1}_{bnd} = \{ & \{(11, 1.0), (10, 0.0), (01, 0.0), (00, 0.0)\}, \\
& \{(11, 0.0), (10, 1.0), (01, 0.0), (00, 0.0)\}, \\
& \{(11, 0.6), (10, 0.0), (01, 0.4), (00, 0.0)\}, \\
& \{(11, 0.6), (10, 0.0), (01, 0.0), (00, 0.4)\}, \\
& \{(11, 0.0), (10, 0.6), (01, 0.4), (00, 0.0)\}, \\
& \{(11, 0.0), (10, 0.6), (01, 0.0), (00, 0.4)\}\}.
\end{aligned}
$$

$\square$

Next, the revision operation is applied to every belief state in Π^B_{bnd}. Let $(\Pi^B_{bnd})^{\mathsf{GI}}_\alpha := \{b' \in \Pi \mid b' = b^{\mathsf{GI}}_\alpha, \ b \in \Pi^B_{bnd}\}$.

Example 5. *Suppose the vocabulary is $\{q, r\}$ and $B_1 = \{(q) \geq 0.6\}$. Let α be $(q \wedge \neg r) \vee (\neg q \wedge r)$. Then*

$$
\begin{aligned}
(\Pi^{B_1}_{bnd})^{\mathsf{GI}}_\alpha = \{ & \{(11, 0.0), (10, 0.5), (01, 0.5), (00, 0.0)\}, \\
& \{(11, 0.0), (10, 1.0), (01, 0.0), (00, 0.0)\}, \\
& \{(11, 0.0), (10, 0.3), (01, 0.7), (00, 0.0)\}, \\
& \{(11, 0.0), (10, 0.6), (01, 0.4), (00, 0.0)\}, \\
& \{(11, 0.0), (10, 0.8), (01, 0.2), (00, 0.0)\}\}.
\end{aligned}
$$

(Two revision operations produce $\{(11, 0), (10, 0.5), (01, 0.5), (00, 0)\}$.) $\square$

To induce the new BB B^α_{bnd} from $(\Pi^B_{bnd})^{\mathsf{GI}}_\alpha$, the following procedure is executed. For every possible world, the procedure adds a sentence enforcing the upper (resp., lower) probability limit of the world, with respect to all the revised boundary belief states. Trivial limits are excepted.

For every $w \in W$, $(\phi_w) \leq \overline{y} \in B^\alpha$, where $\overline{y} = \max_{b \in (\Pi^B_{bnd})^{\mathsf{GI}}_\alpha} b(w)$, except when $\overline{y} = 1$, and $(\phi_w) \geq \underline{y} \in B^\alpha$, where $\underline{y} = \min_{b \in (\Pi^B_{bnd})^{\mathsf{GI}}_\alpha} b(w)$, except when $\underline{y} = 0$.

The intention is that the procedure specifies B^α to represent the upper and lower probability envelopes of the set of revised boundary belief states – B^α thus defines the entire revised belief state space (cf. Theorem 1).

Example 6. *Continuing Example 5, using the translation procedure just above, we see that $B^\alpha_{1bnd} = \{(\phi_{11}) \leq 0, (\phi_{10}) \geq 0.3, (\phi_{01}) \leq 0.7, (\phi_{00}) \leq 0.0\}$.*

Note that if we let $B' = \{((q \wedge \neg r) \vee (\neg q \wedge r)) = 1, (q \wedge \neg r) \geq 0.3\}$, then $\Pi^{B'} = \Pi^{B^\alpha_{1bnd}}$. $\square$

Example 7. *Suppose the vocabulary is $\{q, r\}$ and $B_2 = \{(\neg q \wedge \neg r) = 0.1\}$. Let α be $\neg q$. Then*

$$
\begin{aligned}
\Pi^{B_2}_{bnd} = \{ & \{(11, 0.9), (10, 0), (01, 0), (00, 0.1)\}, \\
& \{(11, 0), (10, 0.9), (01, 0), (00, 0.1)\}, \\
& \{(11, 0), (10, 0), (01, 0.9), (00, 0.1)\}\},
\end{aligned}
$$

$$
\begin{aligned}
(\Pi^{B_2}_{bnd})^{\mathsf{GI}}_\alpha = \{ & \{(11, 0), (10, 0), (01, 0.9), (00, 0.1)\}, \\
& \{(11, 0), (10, 0), (01, 0), (00, 1)\}\} \text{ and}
\end{aligned}
$$

$B^\alpha_{2bnd} = \{(\phi_{11}) \leq 0, (\phi_{10}) \leq 0, (\phi_{01}) \leq 0.9, (\phi_{00}) \geq 0.1\}$.

Note that if we let $B' = \{(\neg q) = 1, (\neg q \wedge r) \leq 0.9\}$, then $\Pi^{B'} = \Pi^{B^\alpha_{2bnd}}$. $\square$

Let $W^{Min(\alpha, d)}$ be a partition of W such that $\{w^i_1, \ldots, w^i_{ni}\}$ is a block in $W^{Min(\alpha,d)}$ iff $|Min(\alpha, w^i_1, d)| = \cdots = |Min(\alpha, w^i_{ni}, d)|$. Denote an element of block $\{w^i_1, \ldots, w^i_{ni}\}$ as w^i, and the block of which w^i is an element as $[w^i]$. Let $i = |Min(\alpha, w^i, d)|$, in other words, the superscript in w^i indicates the size of $Min(\alpha, w^i, d)$. Let $m := \max_{w \in W} |Min(\alpha, w, d)|$.

Observation 1. *Let $\delta_1, \delta_2, \ldots, \delta_m$ be positive integers such that $i < j$ iff $\delta_i < \delta_j$. Let $\nu_1, \nu_2, \ldots, \nu_m$ be values in $[0, 1]$ such that $\sum^m_{k=1} \nu_k = 1$. Associate with every ν_i a maximum value it is allowed to take: $most(\nu_i)$. For every ν_i, we define the* assignment value

$$
av(\nu_i) := \begin{cases} most(\nu_i) & \text{if } \sum^i_{k=1} \leq 1 \\ 1 - \sum^{i-1}_{k=1} & \text{otherwise} \end{cases}
$$

Determine first $av(\nu_1)$, then $av(\nu_2)$ and so on. Then

$$
\frac{av(\nu_1)}{\delta_1} + \cdots + \frac{av(\nu_m)}{\delta_m} > \frac{\nu'_1}{\delta_1} + \cdots + \frac{\nu'_m}{\delta_m}
$$

whenever $\nu'_i \neq av(\nu_i)$ for some i. $\square$

For instance, let $\delta_1 = 1$, $\delta_2 = 2$, $\delta_3 = 3$, $\delta_4 = 4$. Let $most(\nu_1) = 0.5$, $most(\nu_2) = 0.3$, $most(\nu_3) = 0.2$, $most(\nu_4) = 0.3$. Then $av(\nu_1) = 0.5$, $av(\nu_2) = 0.3$, $av(\nu_3) = 0.2$, $av(\nu_4) = 0$ and

$$
\frac{0.5}{1} + \frac{0.3}{2} + \frac{0.2}{3} + \frac{0}{4} = 0.716.
$$

But

$$
\frac{0.49}{1} + \frac{0.3}{2} + \frac{0.2}{3} + \frac{0.01}{4} = 0.709.
$$

And

$$
\frac{0.5}{1} + \frac{0.29}{2} + \frac{0.2}{3} + \frac{0.01}{4} = 0.714.
$$

Lemma 1 essentially says that the belief state in Π^B which causes a revised belief state to have a maximal value at world w (w.r.t. all belief states in Π^B), will be in Π^B_{bnd}.

Lemma 1. *For all $w \in W$, $\arg\max_{b_X \in \Pi^B} \sum_{\substack{w' \in W \\ w \in Min(\alpha, w', d)}} b_X(w')/|Min(\alpha, w', d)|$ is in Π^B_{bnd}.*

Proof. Note that

$$
\sum_{\substack{w' \in W \\ w \in Min(\alpha, w', d)}} b(w')/|Min(\alpha, w', d)|
$$

can be written in the form

$$
\frac{\sum_{\substack{w' \in [w^1] \\ w \in Min(\alpha, w', d)}} b(w')}{1} + \cdots + \frac{\sum_{\substack{w' \in [w^m] \\ w \in Min(\alpha, w', d)}} b(w')}{m}.
$$

Observe that there must be a $W^\# \in W^{perm}$ such that $W^\# = \langle w^1_1, \ldots, w^1_{n1}, \ldots, w^m_1, \ldots, w^m_{nm} \rangle$. Then by the

definition of the set of boundary belief states (Def. 4), $MaxASAP(B, W^{\#})$ will assign maximal probability mass to $[w^1] = \{w_1^1, \ldots, w_{n1}^1\}$, then to $[w^2] = \{w_1^2, \ldots, w_{n2}^m\}$ and so on.

That is, by Observation 1, for some $b_x \in \Pi_{bnd}^B$, $b_x(w) = max_{b_X \in \Pi^B} \sum_{\substack{w' \in W \\ w \in Min(\alpha, w', d)}} b_X(w')/|Min(\alpha, w', d)|$ for all $w \in W$. Therefore, $\arg\max_{b_X \in \Pi^B} \sum_{\substack{w' \in W \\ w \in Min(\alpha, w', d)}} b_X(w')/|Min(\alpha, w', d)|$ is in Π_{bnd}^B. $\quad\square$

Let
$$\overline{x}^w := \max_{b \in \Pi_{bnd}^B} b(w) \qquad \overline{X}^w := \max_{b \in \Pi^B} b(w)$$
$$\overline{y}^w := \max_{b \in (\Pi_{bnd}^B)_\alpha^{GI}} b(w) \qquad \overline{Y}^w := \max_{b \in (\Pi^B)_\alpha^{GI}} b(w)$$
$$\underline{x}^w := \min_{b \in \Pi_{bnd}^B} b(w) \qquad \underline{X}^w := \min_{b \in \Pi^B} b(w)$$
$$\underline{y}^w := \min_{b \in (\Pi_{bnd}^B)_\alpha^{GI}} b(w) \qquad \underline{Y}^w := \min_{b \in (\Pi^B)_\alpha^{GI}} b(w)$$

Lemma 2 states that for every world, the upper/lower probability of the world with respect to Π_{bnd}^B is equal to the upper/lower probability of the world with respect to Π^B. The proof requires Observation 1 and Lemma 1.

Lemma 2. *For all $w \in W$, $\overline{y}^w = \overline{Y}^w$ and $\underline{y}^w = \underline{Y}^w$.*

Proof. Note that if $w \notin [\alpha]$, then $\overline{y}^w = \overline{Y}^w = 0$ and $\underline{y}^w = \underline{Y}^w = 0$.

We now consider the cases where $w \in [\alpha]$.

$$\overline{y}^w = \overline{Y}^w$$

iff

$$\max_{b \in (\Pi_{bnd}^B)} b(w) = \max_{b \in (\Pi^B)} b(w)$$

iff

$$\max_{b_x \in \Pi_{bnd}^B} \sum_{\substack{w' \in W \\ w \in Min(\alpha, w', d)}} b_x(w')/|Min(\alpha, w', d)|$$
$$= \max_{b_X \in \Pi^B} \sum_{\substack{w' \in W \\ w \in Min(\alpha, w', d)}} b_X(w')/|Min(\alpha, w', d)|$$

if
$$\overline{b}_x(w) = \overline{b}_X(w), \text{ where}$$

$$\overline{b}_x(w) := \max_{b_x \in \Pi_{bnd}^B} \sum_{\substack{w' \in W \\ w \in Min(\alpha, w', d)}} b_x(w')/|Min(\alpha, w', d)|$$

and

$$\overline{b}_X(w) := \max_{b_X \in \Pi^B} \sum_{\substack{w' \in W \\ w \in Min(\alpha, w', d)}} b_X(w')/|Min(\alpha, w', d)|.$$

Note that
$$\sum_{\substack{w' \in W \\ w \in Min(\alpha, w', d)}} b(w')/|Min(\alpha, w', d)|$$

can be written in the form

$$\frac{\sum_{\substack{w' \in [w^1] \\ w \in Min(\alpha, w', d)}} b(w')}{1} + \cdots + \frac{\sum_{\substack{w' \in [w^m] \\ w \in Min(\alpha, w', d)}} b(w')}{m}.$$

Then by Observation 1, $\overline{b}_X(w)$ is in Π_{bnd}^B. And also by Lemma 1, the belief state in Π_{bnd}^B identified by $\overline{b}_X(w)$ must be the one which maximizes

$$\sum_{\substack{w' \in W \\ w \in Min(\alpha, w', d)}} b_x(w')/|Min(\alpha, w', d)|,$$

where $b_x \in \Pi_{bnd}^B$. That is, $\overline{b}_x = \overline{b}_X$.

With a symmetrical argument, it can be shown that $\underline{y}^w = \underline{Y}^w$. $\quad\square$

In intuitive language, the following theorem says that the BB determined through the method of revising boundary belief states captures exactly the same beliefs and ignorance as the belief states in Π^B which have been revised. This correspondence relies on the fact that the upper and lower probability envelopes of Π^B can be induce from Π_{bnd}^B, which is what Lemma 2 states.

Theorem 1. *Let $(\Pi^B)_\alpha^{GI} := \{b_\alpha^{GI} \in \Pi \mid b \in \Pi^B\}$. Let B_{bnd}^α be the BB induced from $(\Pi_{bnd}^B)_\alpha^{GI}$. Then $\Pi^{B_{bnd}^\alpha} = (\Pi^B)_\alpha^{GI}$.*

Proof. We show that $\forall b' \in \Pi$, $b' \in \Pi^{B_{bnd}^\alpha} \iff b' \in (\Pi^B)_\alpha^{GI}$.

($\Rightarrow$) $b' \in \Pi^{B_{bnd}^\alpha}$ implies $\forall w \in W$, $\underline{y}^w \leq b'(w) \leq \overline{y}^w$ (by definition of B_{bnd}^α). Lemma 2 states that for all $w \in W$, $\overline{y}^w = \overline{Y}^w$ and $\underline{y}^w = \underline{Y}^w$. Hence, $\forall w \in W$, $\underline{Y}^w \leq b'(w) \leq \overline{Y}^w$ Therefore, $b'(w) \in (\Pi^B)_\alpha^{GI}$.

($\Leftarrow$) $b'(w) \in (\Pi^B)_\alpha^{GI}$ implies $\forall w \in W$, $\underline{Y}^w \leq b'(w) \leq \overline{Y}^w$. Hence, by Lemma 2, $\forall w \in W$, $\underline{y}^w \leq b'(w) \leq \overline{y}^w$. Therefore, by definition of B_{bnd}^α, $b' \in \Pi^{B_{bnd}^\alpha}$. $\quad\square$

Revising via a Representative Belief State

Another approach to the revision of a belief base (BB) is to determine a representative of Π^B (call it b_{rep}), change the representative belief state via the the defined revision operation and then induce a new BB from the revised representative belief state. Selecting a representative probability function from a family of such functions is not new (Goldszmidt, Morris, and Pearl, 1990; Paris, 1994, e.g.). More formally, given observation α, first determine $b_{rep} \in \Pi^B$, then compute its revision b_{rep}^α, and finally induce B^α from b_{rep}^α.

We shall represent Π^B (and thus B) by the single 'least biased' belief state, that is, the belief state in Π^B with *highest entropy*:

Definition 5 (Shannon Entropy)**.**

$$H(b) := -\sum_{w \in W} b(w) \ln b(w),$$

where b is a belief state.

Definition 6 (Maximum Entropy)**.** *Traditionally, given some set of distributions* Π*, the most entropic distribution in* Π *is defined as*

$$b^H := \arg\max_{b \in \Pi} H(b).$$

Suppose $B_2 = \{(\neg q \wedge \neg r) = 0.1\}$. Then the belief state $b \in \Pi^{B_2}$ satisfying the constraints posed by B_2 for which $H(b)$ is maximized is $b_{rep} = b^H = \langle 0.3, 0.3, 0.3, 0.1 \rangle$.

The above distribution can be found directly by applying the principle of maximum entropy: The true belief state is estimated to be the one consistent with known constraints, but is otherwise as unbiased as possible, or "Given no other knowledge, assume that everything is as random as possible. That is, the probabilities are distributed as uniformly as possible consistent with the available information," (Poole and Mackworth, 2010). Obviously world 00 must be assigned probability 0.1. And the remaining 0.9 probability mass should be uniformly spread across the other three worlds.

Applying GI to b_{rep} on evidence $\neg q$ results in $b_{rep}^{\neg q} = \langle 0, 0, 0.6, 0.4 \rangle$.

Example 8. *Suppose the vocabulary is* $\{q, r\}$*,* $B_1 = \{(q) \geq 0.6\}$ *and* α *is* $(q \wedge \neg r) \vee (\neg q \wedge r)$*. Then* $b_{rep} = \arg\max_{b \in \Pi^{B_1}} H(b) = \langle 0.3,\ 0.3,\ 0.2,\ 0.2 \rangle$*. Applying* GI *to* b_{rep} *on* α *results in* $b_{rep}^{\alpha} = \langle 0,\ 0.61,\ 0.39,\ 0 \rangle$*.* b_{rep}^{α} *can be translated into* B_{1rep}^{α} *as* $\{(q \wedge \neg r) = 0.61, (\neg q \wedge r) = 0.39\}$*.* □

Still using $\alpha = (q \wedge \neg r) \vee (\neg q \wedge r)$, notice that $\Pi^{B_{1rep}^{\alpha}} \neq \Pi^{B_{1bnd}^{\alpha}}$. But how different *are* $B_{1rep}^{\alpha} = \{(q \wedge \neg r) = 0.61, (\neg q \wedge r) = 0.39\}$ and $B_{1bnd}^{\alpha} = \{(q \wedge r) \leq 0, (q \wedge \neg r) \geq 0.3, (\neg q \wedge r) \leq 0.7, (\neg q \wedge \neg r) \leq 0.0\}$? Perhaps one should ask, how different B_{1rep}^{α} is from the representative of B_{1bnd}^{α}: The least biased belief state satisfying B_{1bnd}^{α} is $\langle 0, 0.5, 0.5, 0 \rangle$. That is, How different are $\langle 0, 0.61, 0.39, 0 \rangle$ and $\langle 0, 0.5, 0.5, 0 \rangle$?

In the case of B_2, we could compare $B_{2bnd}^{\neg q} = \{(\phi_{11}) \leq 0, (\phi_{10}) \leq 0, (\phi_{01}) \leq 0.9, (\phi_{00}) \geq 0.1\}$ with $b_{rep}^{\neg q} = \langle 0, 0, 0.6, 0.4 \rangle$. Or if we take the least biased belief state satisfying $B_{2bnd}^{\neg q}$, we can compare $\langle 0, 0, 0.5, 0.5 \rangle$ with $\langle 0, 0, 0.6, 0.4 \rangle$.

It has been extensively argued (Jaynes, 1978; Shore and Johnson, 1980; Paris and Vencovsk, 1997) that maximum entropy is a reasonable inference mechanism, if not the most reasonable one (w.r.t. probability constraints). And in the sense that the boundary belief states method requires no compression / information loss, it also seems like a very reasonable inference mechanism for revising BBs as defined here. Resolving this misalignment in the results of the two methods is an obvious task for future research.

Future Directions

Some important aspects still missing from our framework are the representation of conditional probabilistic information such as is done in the work of Kern-Isberner, and the association of information with its level of entrenchment. On the latter point, when one talks about probabilities or likelihoods, if one were to take a frequentist perspective, information observed more (less) often should become more (less) entrenched. Or, without considering observation frequencies, an agent could be designed to have, say, one or two sets of deeply entrenched background knowledge (e.g., domain constraints) which does not change or is more immune to change than 'regular' knowledge.

Given that we have found that the belief base resulting from revising via the boundary-belief-states approach differs from the belief base resulting from revising via the representative-belief-state approach, the question arises, When is it appropriate to use a representative belief state defined as the most entropic belief state of a given set Π^B? This is an important question, especially due to the popularity of employing the Maximum Entropy principle in cases of undespecified probabilistic knowledge (Jaynes, 1978; Goldszmidt, Morris, and Pearl, 1990; Hunter, 1991; Voorbraak, 1999; Kern-Isberner, 2001; Kern-Isberner and Rdder, 2004) and the principle's well-behavedness (Shore and Johnson, 1980; Paris, 1994; Kern-Isberner, 1998).

Katsuno and Mendelzon (1991) modified the eight AGM belief revision postulates (Alchourrón, Gärdenfors, and Makinson, 1985) to the following six (written in the notation of this paper), where $*$ is some revision operator.[3]

- $B_*^{\alpha} \models (\alpha) = 1$.
- If $B \cup \{(\alpha) = 1\}$ is satisfiable, then $B_*^{\alpha} \equiv B \cup \{(\alpha) = 1\}$.
- If $(\alpha) = 1$ is satisfiable, then B_*^{α} is also satisfiable.
- If $\alpha \equiv \beta$, then $B_*^{\alpha} \equiv B_*^{\beta}$.
- $B_*^{\alpha} \cup \{(\beta) = 1\} \models B_*^{\alpha \wedge \beta}$.
- If $B_*^{\alpha} \cup \{(\beta) = 1\}$ is satisfiable, then $B_*^{\alpha \wedge \beta} \models B_*^{\alpha} \cup \{(\beta) = 1\}$.

Testing the various revision operations against these postulates is left for a sequel paper.

An extended version of maximum entropy is *minimum cross-entropy* (MCE) (Kullback, 1968; Csiszár, 1975):

Definition 7 (Minimum Cross-Entropy)**.** *The 'directed divergence' of distribution* c *from distribution* b *is defined as*

$$R(c, b) := \sum_{w \in W} c(w) \ln \frac{c(w)}{b(w)}.$$

$R(c, b)$ *is undefined when* $b(w) = 0$ *while* $c(w) > 0$*; when* $c(w) = 0$*,* $R(c, b) = 0$*, because* $\lim_{x \to 0} \ln(x) = 0$*. Given new evidence* $\phi \in L^{prob}$*, the distribution* c *satisfying* ϕ *diverging least from current belief state* b *is*

$$\arg\min_{c \in \Pi, c \Vdash \phi} R(c, b).$$

Definition 8 (MCI)**.** *Then MCE inference (denoted (*MCI*)) is defined as*

$$b\ \mathsf{MCI}\ \alpha := \arg\min_{b' \in \Pi, b' \Vdash (\alpha) = 1} R(b', b).$$

In the following example, we interpret revision as MCE inference.

---

[3]In these postulates, it is sometimes necessary to write an observation α as a BB, i.e., as $\{(\alpha) = 1\}$ – in the present framework, observations are regarded as certain.

Example 9. *Suppose the vocabulary is $\{q, r\}$ and $B_1 = \{(q) \geq 0.6\}$. Let α be $(q \wedge \neg r) \vee (\neg q \wedge r)$. Then*

$$
\begin{aligned}
\Pi_{bnd}^{B_1} = \ &\{\{(11, 1.0), (10, 0.0), (01, 0.0), (00, 0.0)\}, \\
&\{(11, 0.0), (10, 1.0), (01, 0.0), (00, 0.0)\}, \\
&\{(11, 0.6), (10, 0.0), (01, 0.4), (00, 0.0)\}, \\
&\{(11, 0.6), (10, 0.0), (01, 0.0), (00, 0.4)\}, \\
&\{(11, 0.0), (10, 0.6), (01, 0.4), (00, 0.0)\}, \\
&\{(11, 0.0), (10, 0.6), (01, 0.0), (00, 0.4)\}\},
\end{aligned}
$$

$$
\begin{aligned}
(\Pi_{bnd}^{B_1})_\alpha^{\mathsf{MCI}} = \ &\{\{(11, 0), (10, 0), (01, 1), (00, 0)\}, \\
&\{(11, 0), (10, 1), (01, 0), (00, 0)\}, \\
&\{(11, 0), (10, 0.6), (01, 0.4), (00, 0)\}\} \ and
\end{aligned}
$$

$B_{1bnd}^\alpha = \{(\phi_{11}) \leq 0, (\phi_{00}) \leq 0\}$.

Note that if we let $B' = \{((q \wedge \neg r) \vee (\neg q \wedge r)) = 1\}$, then $\Pi^{B'} = \Pi^{B_{1bnd}^\alpha}$. □

Recall from Example 6 that B' included $(q \wedge \neg r) \geq 0.3$. Hence, in this particular case, combining the boundary belief states approach with MCI results in a less informative revised belief base than when GI is used. The reason for the loss of information might be due to $R(\cdot, \{(11, 1.0), (10, 0.0), (01, 0.0), (00, 0.0)\})$ and $R(\cdot, \{(11, 0.6), (10, 0.0), (01, 0.0), (00, 0.4)\})$ being undefined: Recall that $R(c, b)$ is undefined when $b(w) = 0$ while $c(w) > 0$. But then there is no belief state c for which $c \Vdash \alpha$ and $R(\cdot)$ is defined (with these two belief states as arguments). Hence, there are no revised counterparts of these two belief states in $(\Pi_{bnd}^{B_1})_\alpha^{\mathsf{MCI}}$. We would like to analyse MCI more within this framework. In particular, in the future, we would like to determine whether a statement like Theorem 1 holds for MCI too.

In MCE inference, b-consistency of evidence ϕ is defined as: There exists a belief state c such that $c \Vdash \phi$ and c is *totally continuous* with respect to b (i.e., $b(w) = 0$ implies $c(w) = 0$). MCE is undefined when the evidence is not b-consistent. This is analogous to Bayesian conditioning being undefined for $b(\alpha) = 0$. Obviously, this is a limitation of MCE because some belief states may not be considered as candidate revised belief states. Admittedly, we have not searched the literature on this topic due to it being out of the present scope.

As far as we know, imaging for belief change has never been applied to (conditional) probabilistic evidence. Due to issues with many revision methods required to be consistent with prior beliefs, and imaging not having this limitation, it might be worthwhile investigating.

The translation from the set of belief states back to a belief base is a mapping from every belief state to a probability formula. The size of the belief base is thus in the order of $|W^{perm}|$, where $|W|$ is already exponential in the size of $\mathcal{P}$, the set of atoms. As we saw in several examples in this paper, the new belief base often has a more concise equivalent counterpart. It would be useful to find a way to consistently determine more concise belief bases than our present approach does.

The computational complexity of the process to revise a belief base is at least exponential. This work focused on theoretical issues. If the framework presented here is ever used in practice, computations will have to be optimized.

The following example illustrates how one might deal with strict inequalities.

Example 10. *Suppose the vocabulary is $\{q, r\}$ and $B_3 = \{(q) > 0.6\}$. Let α be $(q \wedge \neg r) \vee (\neg q \wedge r)$. Let ϵ be a real number which tends to 0. Then $\Pi_{bnd}^{B_3} =$*

$$
\begin{aligned}
&\{\{(11, 1.0), (10, 0.0), (01, 0.0), (00, 0.0)\}, \\
&\{(11, 0.0), (10, 1.0), (01, 0.0), (00, 0.0)\}, \\
&\{(11, 0.6 + \epsilon), (10, 0.0), (01, 0.4 - \epsilon), (00, 0.0)\}, \\
&\{(11, 0.6 + \epsilon), (10, 0.0), (01, 0.0), (00, 0.4 - \epsilon)\}, \\
&\{(11, 0.0), (10, 0.6 + \epsilon), (01, 0.4 - \epsilon), (00, 0.0)\}, \\
&\{(11, 0.0), (10, 0.6 + \epsilon), (01, 0.0), (00, 0.4 - \epsilon)\}\},
\end{aligned}
$$

$(\Pi_{bnd}^{B_3})_\alpha^{\mathsf{GI}} =$

$$
\begin{aligned}
&\{\{(11, 0.0), (10, 0.5), (01, 0.5), (00, 0.0)\}, \\
&\{(11, 0.0), (10, 1.0), (01, 0.0), (00, 0.0)\}, \\
&\{(11, 0.0), (10, 0.3 + \epsilon), (01, 0.7 - \epsilon), (00, 0.0)\}, \\
&\{(11, 0.0), (10, 0.6 + \epsilon), (01, 0.4 - \epsilon), (00, 0.0)\}, \\
&\{(11, 0.0), (10, 0.8 + \epsilon), (01, 0.2 - \epsilon), (00, 0.0)\} \ and
\end{aligned}
$$

$B_{3bnd}^\alpha = \{(\phi_{11}) \leq 0, (\phi_{10}) \geq 0.3 + \epsilon, (\phi_{01}) \leq 0.7 - \epsilon, (\phi_{00}) \leq 0.0\}$.

Note that if we let $B' = \{((q \wedge \neg r) \vee (\neg q \wedge r)) = 1, (q \wedge \neg r) > 0.3\}$, then $\Pi^{B'} = \Pi^{B_{3bnd}^\alpha}$. □

It has been suggested by one of the reviewers that GI could be an *affine map* (i.t.o. geometry), thus allowing the proof of Theorem 1 to refer to existing results in the study of affine maps to significantly simplify the proof. The authors are not familiar with affine maps and thus leave investigation of the suggestion to other researchers.

Related Work

Voorbraak (1999) proposed the partial probability theory (PTT), which allows probability assignments to be partially determined, and where there is a distinction between probabilistic information based on (i) hard background evidence and (ii) some assumptions. He does not explicitly define the "constraint language", however, from his examples and discussions, one can infer that he has something like the language L^{PTT} in mind: it contains all formulae which can be formed with sentences in our L^{prob} in combination with connectives $\neg$, $\wedge$ and $\vee$. A "belief state" in PTT is defined as the quadruple $\langle \Omega, \mathcal{B}, \mathcal{A}, \mathcal{C} \rangle$, where Ω is a sample space, $\mathcal{B} \subset L^{PTT}$ is a sets of probability constraints, $\mathcal{A} \subset L^{PTT}$ is a sets of assumptions and $\mathcal{C} \subseteq W$ "represents specific information concerning the case at hand" (an observation or evidence).[4] Our epistemic state can be expressed as a restricted PTT "belief state" by letting $\Omega = W$, $\mathcal{B} = B$, $\mathcal{A} = \emptyset$ and

---

[4]Voorbraak (1999)'s "belief state" would rather be called and *epistemic state* or *knowledge structure* in our language.

$\mathcal{C} = \{w \in W \mid w \Vdash \alpha\}$, where B is a belief base and α is an observation in our notation.

Voorbraak (1999) mentions that he will only consider conditioning where the evidence does not contradict the current beliefs. He defines the set of belief states corresponding to the conditionalized PPT "belief state" as $\{b(\cdot \mid C) \in \Pi \mid b \in \Pi^{\mathcal{B} \cup \mathcal{A}}, b(C) > 0\}$. In our notation, this corresponds to $\{(b \,\mathsf{BC}\, \alpha) \in \Pi \mid b \in \Pi^B, b(\alpha) > 0\}$, where α corresponds to C.

Voorbraak (1999) proposes *constraining* as an alternative to conditioning: Let $\phi \in L^{prob}$ be a probability constraint. In our notation, constraining Π^B on ϕ produces $\Pi^{B \cup \{\phi\}}$.

Note that expanding a belief set reduces the number of models (worlds) and expanding a PPT "belief state" with extra constraints also reduces the number of models (belief states / probability functions).

> In the context of belief sets, it is possible to obtain any belief state from the ignorant belief state by a series of expansions. In PPT, constraining, but not conditioning, has the analogous property. This is one of the main reasons we prefer to constraining and not conditioning to be the probabilistic version of expansion. (Voorbraak, 1999, p. 4)

But Voorbraak does not address the issue that C and ϕ are different kinds of observations, so constraining, as defined here, cannot be an alternative to conditioning. C cannot be used directly for constraining and ϕ cannot be used directly for conditioning.

W.l.o.g., we can assume C is represented by α. If we take $b \,\mathsf{GI}\, \alpha$ to be an expansion operation whenever $b(\alpha) > 0$, then one might ask, Is it possible to obtain any belief base B' from the ignorant belief base $B = \emptyset$ by a series of expansions, using our approach? The answer is, No. For instance, there is no observation or series of observations which can change $B = \{\}$ into $B' = \{(q) \geq 0.6\}$. But if we *were* to allow sentences (constraints) in L^{prob} to be observations, then we *could* obtain any B' from the ignorant B.

Grove and Halpern (1998) investigate what "update" (incorporation of an observation with current beliefs, such that the observation does not contradict the beliefs) means in a framework where beliefs are represented by a set of belief states. They state that the main purpose of their paper is to illustrate how different the set-of-distributions framework can be, "technically", from the standard single-distribution framework. They propose six postulates characterizing what properties an update function should have. They say that some of the postulates are obvious, some arguable and one probably too strong. Out of seven (families of) update functions only the one based on conditioning ($Upd_{cond}(\cdot)$) and the one based on constraining ($Upd_{constrain}(\cdot)$) satisfy all six postulates, where $Upd_{cond}(\Pi^B, \alpha) := \{(b \,\mathsf{BC}\, \alpha) \in \Pi \mid b \in \Pi^B, b(\alpha) > 0\}$ and where they interpret Voorbraak's (1999) constraining as $Upd_{constrain}(\Pi^B, \alpha) := \{b \in \Pi^B \mid b(\alpha) = 1\}$. Grove and Halpern (1998) do not investigate the case when an observation must be incorporated while it is (possibly) inconsistent with the old beliefs (i.e., revision).

Kern-Isberner (2001) develops a new perspective of probabilistic belief change. Based on the ideas of Alchourrón,

Gärdenfors, and Makinson (1985) and Katsuno and Mendelzon (1991) (KM), the operations of revision and update, respectively, are investigated within a probabilistic framework. She employs as basic knowledge structure a belief base $(b, \mathcal{R})$, where b is a probability distribution (belief state) of background knowledge and $\mathcal{R}$ is a set of probabilistic conditionals of the form $A \rightsquigarrow B[x]$ meaning 'The probability of B, given A, is x. A universal inference operation – based on the techniques of optimum entropy – is introduced as an "adequate and powerful method to realize probabilistic belief change".

By having a belief state available in the belief base, minimum cross-entropy can be used. The intention is then that an agent with belief base $(b, \mathcal{T})$ should always reason w.r.t. belief state $b^{\mathcal{T}} := \arg\min_{c \in \Pi, c \Vdash \mathcal{T}} R(c, b)$. Kern-Isberner (2001) then defines the probabilistic belief revision of $(b, \mathcal{R})$ by evidence $\mathcal{S}$ as $(b, \mathcal{R} \cup \mathcal{S})$. And the probabilistic belief update of $(b, \mathcal{R})$ by evidence $\mathcal{S}$ is defined as $(b^{\mathcal{R}}, \mathcal{S})$.[5] She distinguishes between revision as a knowledge adding process, and updating as a change-recording process. Kern-Isberner (2001) sets up comparisons of maximum cross-entropy belief change with AGM revision and KM update. Cases where, for update, new information $\mathcal{R}$ is inconsistent with the prior distribution b, or, for revision, is inconsistent with b or the context $\mathcal{R}$, are not dealt with (Kern-Isberner, 2001, p. 399, 400).

Having a belief state available for modification when new evidence is to be adopted is quite convenient. As Voorbraak (1999) argues, however, an agent's ignorance can hardly be represented in an epistemic state where a single belief state must always be chosen.

The reader may also refer to a later paper (Kern-Isberner, 2008) in which many of the results of the work just reviewed are generalized to belief bases of the form $(\Psi, \mathcal{R})$, where Ψ denotes a general *epistemic state*. In that paper, she considers two instantiations of Ψ, namely as a *probability distribution* and as an *ordinal conditional function* (first introduced by Spohn (1988)).

Yue and Liu (2008) propose a probabilistic revision operation for imprecise probabilistic beliefs in the framework of Probabilistic Logic Programming (PLP). New evidence may be a probabilistic (conditional) formula and needs not be consistent with the original beliefs. Revision via imaging (e.g., GI) also overcomes this consistency issue.

Essentially, their *probabilistic epistemic states* Ψ are induced from a PLP program which is a set of formulae, each formula having the form $(\psi \mid \phi)[l, u]$, meaning that the probability of the conditional $(\psi \mid \phi)$ lies in the interval $[l, u]$.

The operator they propose has the characteristic that if an epistemic state Ψ represents a single probability distribution, revising collapses to Jeffrey's rule and Bayesian conditioning.

They mention that it is required that the models (distributions) of Ψ is a convex set. There might thus be an opportunity to employ their revision operation on a representative set of boundary distributions as proposed in this paper.

---

[5]This is a very simplified version of what she presents. Please refer to the paper for details.

Conclusion

In this paper, we propose an approach how to generate a new probabilistic belief base from an old one, given a new piece of non-probabilistic information, where a belief base is a finite set of sentences, each sentence stating the likelihood of a proposition about the world. In this framework, an agent's belief base represents the set of belief states compatible with the sentences in it. In this sense, the agent is able to represent its knowledge *and* ignorance about the true state of the world.

We used a version of the so-called *imaging* approach to implement the revision operation.

Two methods were proposed: revising a finite set of 'boundary belief states' and revising a least biased belief state. We focussed on the former and showed that the latter gives different results.

There were two main contribution of this paper. The first was to prove that the set of belief states satisfying B_{new} is exactly those belief states satisfying the original belief base, revised. The second was to uncover an interesting conflict in the results of the two belief base revision methods. It is worth further understanding the reasons behind such a difference, as such an investigation could give more insight about the mechanisms behind the two methods and indicate possible pros and cons of each.

Acknowledgements

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(2):510–530.

Boutilier, C. 1998. A unified model of qualitative belief change: a dynamical systems perspective. *Artificial Intelligence* 98(1–2):281–316.

Chhogyal, K.; Nayak, A.; Schwitter, R.; and Sattar, A. 2014. Proceedings of the thirteenth pacific rim international conference on artificial intelligence (pricai 2014). In Pham, D., and Park, S., eds., *Proc. of PRICAI 2014*, volume 8862 of *LNCS*, 694–707. Springer-Verlag.

Cover, T., and Thomas, J. 1991. *Elements of Information Theory*. New York: Wiley.

Csiszár, I. 1975. I-divergence geometry of probability distributions and minimization problems. *Annals of Probability* 3:146–158.

Dubois, D., and Prade, H. 1993. Belief revision and updates in numerical formalisms: An overview, with new results for the possibilistic framework. In *Proceedings of the 13th International Joint Conference on Artifical Intelligence*, volume 1 of *IJCAI'93*, 620–625. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Gärdenfors, P. 1988. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. Massachusetts/England: MIT Press.

Goldszmidt, M.; Morris, P.; and Pearl, J. 1990. A maximum entropy approach to nonmonotonic reasoning. In *Proceedings of the Eighth Natl. Conf. on Artificial Intelligence (AAAI-90)*, 646–652. AAAI Press.

Grove, A., and Halpern, J. 1998. Updating sets of probabilities. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, UAI'98, 173–182. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Hunter, D. 1991. Maximum entropy updating and conditionalization. In Spohn, W.; Van Fraassen, B.; and Skyrms, B., eds., *Existence and Explanation*, volume 49 of *The University of Western Ontario Series in Philosophy of Science*. Springer Netherlands. 45–57.

Jaynes, E. 1978. Where do we stand on maximum entropy? In *The Maximum Entropy Formalism*. MIT Press. 15–118.

Katsuno, H., and Mendelzon, A. 1991. On the difference between updating a knowledge base and revising it. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, 387–394.

Kern-Isberner, G., and Rdder, W. 2004. Belief revision and information fusion on optimum entropy. *International Journal of Intelligent Systems* 19(9):837–857.

Kern-Isberner, G. 1998. Characterizing the principle of minimum cross-entropy within a conditional-logical framework. *Artificial Intelligence* 98(12):169 – 208.

Kern-Isberner, G. 2001. Revising and updating probabilistic beliefs. In Williams, M.-A., and Rott, H., eds., *Frontiers in Belief Revision*, volume 22 of *Applied Logic Series*. Kluwer Academic Publishers, Springer Netherlands. 393–408.

Kern-Isberner, G. 2008. Linking iterated belief change operations to nonmonotonic reasoning. In *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning*, 166–176. Menlo Park, CA: AAAI Press.

Kullback, S. 1968. *Information theory and statistics*, volume 1. New York: Dover, 2nd edition.

Lehmann, D.; Magidor, M.; and Schlechta, K. 2001. Distance semantics for belief revision. *Journal of Symboloc Logic* 66(1):295–317.

Lewis, D. 1976. Probabilities of conditionals and conditional probabilities. *Philosophical Review* 85(3):297–315.

Paris, J., and Vencovsk, A. 1997. In defense of the maximum entropy inference process. *International Journal of Approximate Reasoning* 17(1):77–103.

Paris, J. 1994. *The Uncertain Reasoner's Companion: A Mathematical Perspective*. Cambridge: Cambridge University Press.

Poole, D., and Mackworth, A. 2010. *Artificial Intelligence: Foundations of Computational Agents*. New York, USA: Cambridge University Press.

Rens, G., and Meyer, T. 2015. A new approach to probabilistic belief change. In Russell, I., and Eberle, W., eds., *Proceedings of the International Florida AI Research Society Conference (FLAIRS)*, 582–587. AAAI Press.

Shore, J., and Johnson, R. 1980. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *Information Theory, IEEE Transactions on* 26(1):26–37.

Spohn, W. 1988. Ordinal conditional functions: A dynamic theory of epistemic states. In Harper, W., and Skyrms, B., eds., *Causation in Decision, Belief Change, and Statistics*, volume 42 of *The University of Western Ontario Series in Philosophy of Science*. Springer Netherlands. 105–134.

Voorbraak, F. 1999. Partial Probability: Theory and Applications. In *Proceedings of the First International Symposium on Imprecise Probabilities and Their Applications*, 360–368. url: decsai.ugr.es/ smc/isipta99/proc/073.html.

Yue, A., and Liu, W. 2008. Revising imprecise probabilistic beliefs in the framework of probabilistic logic programming. In *Proceedings of the Twenty-third AAAI Conf. on Artificial Intelligence (AAAI-08)*, 590–596.