Logic and Probabilistic Models of Belief Change

Eric Pacuit

Department of Philosophy
University of Maryland, College Park
pacuit.org

March 11, 2016

- Accuracy, credences and full beliefs
- Two important distinctions
- AGM, credences and full beliefs
- More on credences and full beliefs

R. Pettigrew. *Accuracy and the Credence-Beleif Connection*. Philosophers Imprint, 15:16, 2015.

(1) **Veritism** Accuracy is the sole source of epistemic value for a doxastic state.

(2) **Strict Propriety** The inaccuracy of a credence function is measured by a continuous, strictly proper scoring rule.

(3) **Dominance** It is irrational to adopt an option that is strictly dominated by an alternative option that is not itself even weakly dominated.

(4) **Predd et al./de Finnetti Theorem**.

Therefore, (5) **Probabilism** An agent is rational only if her credence function is probabilistic.

$\mathsf{q} : \{0,1\} \times [0,1] \to [0,1]$ $\mathsf{q}(i,x) = (i-x)^2$.

$\mathfrak{B}(c,w) := \sum_{X \in \mathcal{F}} \mathsf{q}(w(X), c(X))$

$c_{Cleo}(X) = 0.7, c_{Cleo}(\overline{X}) = 0.6$

$\mathfrak{B}(c_{Cleo}, w_X) = (1-0.7)^2 + (0-0.6)^2 = 0.45$
$\mathfrak{B}(c_{Cleo}, w_{\overline{X}}) = (0-0.7)^2 + (1-0.6)^2 = 0.65$

$c^*_{Cleo}(X) = 0.55, c^*_{Cleo}(\overline{X}) = 0.45$

$\mathfrak{B}(c^*_{Cleo}, w_X) = (1-0.55)^2 + (0-0.45)^2 = 0.405$
$\mathfrak{B}(c^*_{Cleo}, w_{\overline{X}}) = (0-0.55)^2 + (1-0.45)^2 = 0.605$

	w_X	$w_{\overline{X}}$
c_{Cleo}	$\mathfrak{B}(c_{Cleo}, w_X) = 0.45$	$\mathfrak{B}(c_{Cleo}, w_{\overline{X}}) = 0.65$
c^*_{Cleo}	$\mathfrak{B}(c^*_{Cleo}, w_X) = 0.405$	$\mathfrak{B}(c^*_{Cleo}, w_{\overline{X}}) = 0.605$

A function $s : \{0,1\} \times [0,1] \to [0,\infty]$ is a **proper scoring rule** in case:

1. $ps(1,x) + (1-p)s(0,x)$ si uniquely minimized at $x = p$ for $p \in [0,1]$.
2. s is continuous. For $i \in \{0,1\}$, $\lim_{n\to\infty} s(i,x_n) = s(i,x)$ for any sequence x_n from $[0,1]$ converging to x.

$$\mathfrak{I}_{\mathsf{s}}(c, w) := \sum_{X \in \mathcal{F}} \mathsf{s}(w(X), c(X))$$

Logarithmic Scoring Rule:

- $\mathsf{l}(1, x) = -\log x$,
- $\mathsf{l}(0, x) = -\log(1 - x)$

Spherical Scoring Rule:

- $\mathsf{h}(1, x) = -\frac{r}{\sqrt{r^2+(1-r)^2}}$,
- $\mathsf{h}(0, x) = -\frac{1-r}{\sqrt{r^2+(1-r)^2}}$

Perfectionism The accuracy of a credence function at a world is its proximity to the ideal credence function at that world.

Let S be a set. A **distance** is a function $D : S \times S \to \mathbb{R}_0^+$ such that

- D is non-negative: $D(c, c') \geq 0$ with equality iff $c = c'$
- D is symmetric: $D(c, c') = D(c', c)$
- D satisfies triangle inequality: $D(c, c'') \leq D(c, c') + D(c', c'')$

We will assume that D satisfies non-negativity. I.e., D is a **divergence**.

Perfectionism If I is a legitimate inaccuracy measure, there is a **divergence** D such that $I(w, c) = D(i_w, c)$. Recall: i_w is the ideal or vindicated credence function at w. We say that D generates I (relative to that notion of vindication).

Brier Accuracy

Alethic Vindication The ideal credence function at world w is the omniscient credence function at w, namely, v_w.

Squared Euclidean Distance Distance between credence functions is measured by squared Euclidean distance.

Additivity

If I is a legitimate (global) measure of inaccuracy, then there is a local measure of inaccuracy $\mathfrak{s}$ such that

$$I(w, c) = \sum_{X \in \mathcal{F}} \mathfrak{s}(i_w(X), c(X))$$

where i_w is the ideal credence at w.

When we say that we represent an agent by her credence function, it can sound as if we're representing her as having a single, unified doxastic state.

When we say that we represent an agent by her credence function, it can sound as if we're representing her as having a single, unified doxastic state.

Really, we are just representing her as having an agglomeration of individual doxastic states, namely, the individual credences she assigns to the various propositions about which she has an opinion. A credence function is simply a mathematical way of representing this agglomeration; it is a way of collecting together these individual credences into a single object.

I is a legitimate inaccuracy measure, then there is a divergence D such that

- $I(w, c) = D(i_w, c)$ (in such a case, we write $I = I_D$)
- There is a function $\mathsf{s} : [0, 1] \times [0, 1] \to [0, \infty]$ such that
  - for all $x, y \in [0, 1]$, $\mathsf{s}(x, y) \geq 0$ with equality if $x = y$.
  - $D(c, c') = \sum_{X \in \mathcal{F}} \mathsf{s}(c(X), c'(X))$

Continuity If I is a legitimate inaccuracy measure and there is a divergence D generated by s such that

$$I(w,c) = I_D(w,c) = D(i_w, c) = \sum_{X \in \mathcal{F}} \mathsf{s}(i_w(X), c(X))$$

then $\mathsf{s}(x,y)$ is continuous in both its arguments.

To demand that s is continuous in its second argument is to say that there are no ‘jumps’ in inaccuracy as credences change - that is, small changes in credence will give rise to small changes in inaccuracy; there can be no small shift in credence that is accompanied by a large jump in inaccuracy.

Calibration

> Granting that [an agent] is going to think always in the same way about all yellow toadstools, we can ask what degree of confidence it would be best for him to have that they are unwholesome. And the answer is that it will in general be best for his degree of belief that a yellow toadstool is unwholesome to be equal to the proportion of yellow toadstools that are unwholesome. (This follows from the meaning of degree of belief.)
>
> (Ramsey, 1931, 195)

Decomposition If I is a legitimate inaccuracy measure generated by a divergence D, then there are α, β such that

$$D(v_w, c) = \alpha D(c^w, c) + \beta D(v_w, c^w)$$

c^w is the ideally calibrated credence in w.

Theorem (Pettigrew). Suppose Alethic Vindication, Perfectionism, Divergence Additivity, Divergence Continuity and Decomposition. Then, if I is a legitimate inaccuracy measure, there is an additive **Bregman divergence** D such that $I(w, c) = D(v_w, c)$.

Symmetry If I is a legitimate inaccuracy measure generated by a divergence D, then D is symmetric: that is, $D(c, c') = D(c', c)$ for all c, c'

Theorem (Pettigrew). Suppose Alethic Vindication, Perfectionism, Divergence Additivity, Divergence Continuity, Decomposition, and Symmetry. Then, if I is a legitimate inaccuracy measure, then I is the Brier score or some linear transformation of it.

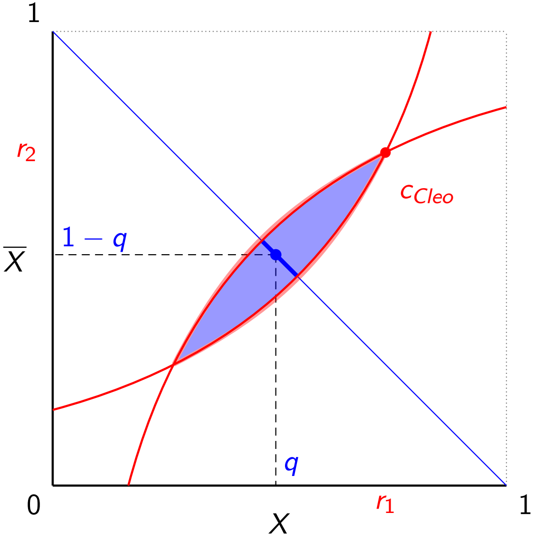
(1) **Veritism** Accuracy is the sole source of epistemic value for a doxastic state.

(2) **Strict Propriety** The inaccuracy of a credence function is measured by a continuous, strictly proper scoring rule.

(3) **Dominance** It is irrational to adopt an option that is strictly dominated by an alternative option that is not itself even weakly dominated.

(4) **Predd et al./de Finnetti Theorem**.

Therefore, (5) **Probabilism** An agent is rational only if her credence function is probabilistic.


1
r_2
c_{Cleo}
$1-q$
$\overline{X}$
q
0
r_1
1
X

A Lacuna in the Argument

There might be credences that violate Probabilism of which the following hold:

- There are alternative credences in the same propositions that accuracy-dominate those credences.
- There is no **total** doxastic state - which includes categorical attitudes as well as credal attitudes - that accuracy-dominates the total doxastic state to which those credences belong.

A Lacuna in the Argument

Her imagined response has three parts: the first is a claim

1. there is a connection between credal attitudes and categorical attitudes, such as full belief, full disbelief, and suspension of judgment.

2. the inaccuracy of an agent's *total* doxastic state ought to take into account both the credences and full beliefs.

3. Cleo's total doxastic state, which includes her credence function c_{Cleo}, as well as her categorical doxastic attitudes, is not accuracy-dominated given the inaccuracy measure for total doxastic states usesd in the second part.

$$b : \mathcal{F} \to \{B, D, S\}$$

- $b(X) = B$ means the agent **believes** that X
- $b(X) = D$ means the agent **disbelieves** that X
- $b(X) = S$ means the agent **suspends judgement** on X

Total doxastic state (b, c)

Bridge Principles

Probability 1: $b(A) = B$ iff $c(A) = 1$

The Lockean Thesis: $b(A) = B$ iff $c(A) > t$

$$
\begin{aligned}
c(X) > t &\implies b(X) = B \\
1 - t < c(X) < t &\implies b(X) = S \\
c(X) > 1 - t &\implies b(X) = D \\
c(X) = t &\implies b(X) = B \text{ or } S \\
c(X) = 1 - t &\implies b(X) = D \text{ or } S
\end{aligned}
$$

- $LT[t]$ (Analytic) It is analytic that $LT[t]$.
- $LT[t]$ (Metaphysical) It is metaphysically necessary that $LT[t]$.
- $LT[t]$ (Normative) It is normatively required that $LT[t]$.

$$\mathfrak{i} : \{0,1\} \times \{B, D, S\} \to [0,1]$$

- $\mathfrak{i}(1, B) = \mathfrak{i}(0, D) = R$
- $\mathfrak{i}(0, B) = \mathfrak{i}(1, D) = W$
- $\mathfrak{i}(1, S) = \mathfrak{i}(0, S) = N$

E.g., $R = 0$, $W = 1$ (and $R < N < W$)

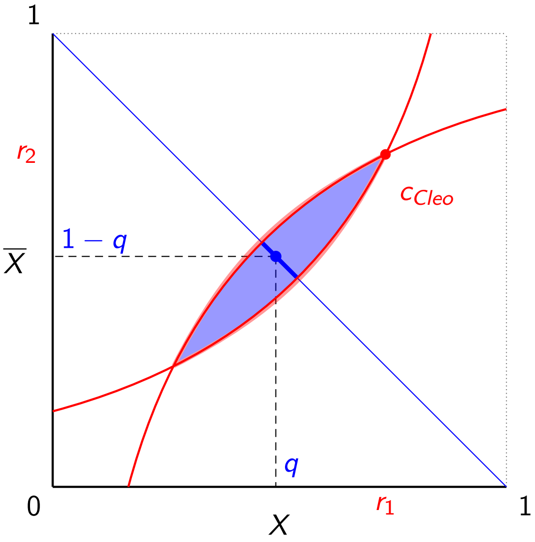
$$\mathfrak{I}_{\mathfrak{i}}(b, w) = \sum_{X \in \mathcal{F}} \mathfrak{i}(w(X), b(X))$$

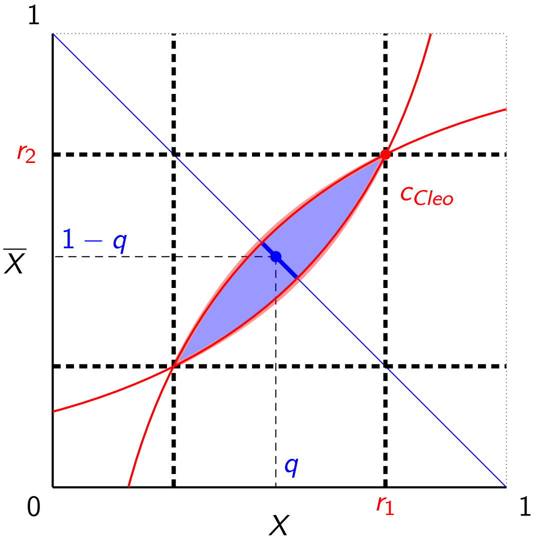
$$\mathfrak{I}_{\mathsf{i},\mathsf{s}}((b,c),w)=\mathfrak{I}_{\mathsf{i}}(b,w)+\mathfrak{I}_{\mathsf{s}}(c,w)$$

Theorem

- Let $t = 0.7$. That is, the Lockean threshold for belief is 0.7 and the threshold for disbelief is 0.3.
- Let the inaccuracy of a credence be measured by the quadratic scoring rule q; and let the inaccuracy of a credence function be measured by the Brier score it generates $\mathfrak{B}$.
- Let the inaccuracy of a categorical doxastic state be measured by $\mathfrak{i}$, where $R = 0$, $N = 0.3$, and $W = 1$; and let the inaccuracy of a belief function be measured by the inaccuracy measure it generates, namely, $\mathfrak{I}_{\mathfrak{i}}$.

Then there is no Lockean$_t$ state (b^*, c^*) that accuracy-dominates (b_{Cleo}, c_{Cleo}) when accuracy is measured by $\mathfrak{I}_{\mathfrak{i},\mathsf{q}}$


1
r_2
c_{Cleo}
$1-q$
$\overline{X}$
q
0
r_1
1
X


1
r_2
c_{Cleo}
$1-q$
$\overline{X}$
q
0
r_1
1
X

Is partial dominance irrational?

	w_1	w_2
Phil	\$5	\$5
Rachel	\$10	\$10

Option 1

	w_1	w_2
Phil	\$10	\$10
Rachel	\$20	\$2

Option 2

Is partial dominance irrational?

	w_1	w_2
Phil	\$5	\$5
Rachel	\$10	\$10

Option 1

	w_1	w_2
Phil	\$10	\$10
Rachel	\$20	\$2

Option 2

...it is clear that my choice of Option 1 is not irrational. The upshot: being dominated in part is not sufficient for irrationality, even when the option that dominates in part is not itself dominated in part.

- $LT[t]$ (Analytic) It is analytic that $LT[t]$.
- $LT[t]$ (Metaphysical) It is metaphysically necessary that $LT[t]$.
- $LT[t]$ (Normative) It is normatively required that $LT[t]$.

Hempel, Easwaran

An agent ought to adopt a belief function b such that, for all belief functions b' , we have

$$\sum_{w} c(w)\mathfrak{i}(b, w) \leq \sum_{w} c(w)\mathfrak{i}(b', w))$$

Hempel, Easwaran

An agent ought to adopt a belief function b such that, for all belief functions b' , we have

$$\sum_{w} c(w)\mathfrak{i}(b,w) \leq \sum_{w} c(w)\mathfrak{i}(b',w))$$

As Hempel and Easwaran have shown, if we assume that $R = 0$ and $W = 1$ and N is closer to R than to W, then b minimizes expected inaccuracy by the lights of c iff (b, c) is a Lockean$_{1-N}$ state.

- We know that Cleo's state (b_{Cleo}, c_{Cleo}) is dominated by (b_{Cleo}, c^*).

- We know that Cleo's state (b_{Cleo}, c_{Cleo}) is dominated by (b_{Cleo}, c^*).
- We also know that (b_{Cleo}, c^*) violates Lockean_{1-N} (Normative) when $N = 0.3$.

- We know that Cleo's state (b_{Cleo}, c_{Cleo}) is dominated by (b_{Cleo}, c^*).
- We also know that (b_{Cleo}, c^*) violates Lockean$_{1-N}$ (Normative) when $N = 0.3$.
- Thus, we know that there is a belief function $b^* \neq b_{Cleo}$ such that c^* expects b to be more accurate than c^* expects b_{Cleo} to be.

- We know that Cleo's state (b_{Cleo}, c_{Cleo}) is dominated by (b_{Cleo}, c^*).
- We also know that (b_{Cleo}, c^*) violates Lockean$_{1-N}$ (Normative) when $N = 0.3$.
- Thus, we know that there is a belief function $b^* \neq b_{Cleo}$ such that c^* expects b to be more accurate than c^* expects b_{Cleo} to be.
- For dominance reasons, we ought to prefer (b_{Cleo}, c^*) to (b_{Cleo}, c_{Cleo}).

- We know that Cleo's state (b_{Cleo}, c_{Cleo}) is dominated by (b_{Cleo}, c^*).
- We also know that (b_{Cleo}, c^*) violates Lockean$_{1-N}$ (Normative) when $N = 0.3$.
- Thus, we know that there is a belief function $b^* \neq b_{Cleo}$ such that c^* expects b to be more accurate than c^* expects b_{Cleo} to be.
- For dominance reasons, we ought to prefer (b_{Cleo}, c^*) to (b_{Cleo}, c_{Cleo}).
- For expected accuracy reasons, we ought to prefer (b^*, c^*) to (b_{Cleo}, c^*).

- We know that Cleo's state (b_{Cleo}, c_{Cleo}) is dominated by (b_{Cleo}, c^*).
- We also know that (b_{Cleo}, c^*) violates Lockean$_{1-N}$ (Normative) when $N = 0.3$.
- Thus, we know that there is a belief function $b^* \neq b_{Cleo}$ such that c^* expects b to be more accurate than c^* expects b_{Cleo} to be.
- For dominance reasons, we ought to prefer (b_{Cleo}, c^*) to (b_{Cleo}, c_{Cleo}).
- For expected accuracy reasons, we ought to prefer (b^*, c^*) to (b_{Cleo}, c^*).
- By the transitivity of preference, we ought to prefer (b^*, c^*) to (b_{Cleo}, c_{Cleo}). Thus, Cleo is irrational, since there is an available option that ought to be preferred to the one she has adopted.

However, there is an illegitimate move in this argument.
It is true that, from the point of view of part of *part of* (b_{Cleo}, c^*) — name, the credal part c^* — (b^*, c^*) is better than (b_{Cleo}, c^*)).
But it does not follow from this that the same is true from the point of view of *the total state* (b_{Cleo}, c^*) After all, if we instead ask which state is optimal from the point of view of b_{Cleo} — that is, the other part of the total state — we would most likely receive as an answer a state that includes b_{Cleo} as its belief function.

What we have just encountered is an instance of a more general problem that arises when one seeks guidance from a doxastic state that doesn't present a consistent attitude to the world: different parts of such a state often give rise to different and incompatible preference orderings.

(Ellsberg example)

Theorem 4 For all Lockean thresholds t, there is a Lockean$_t$ state (b, c) such that:

- c is non-probabilistic;
- There is no Lockean$_t$ state (b^*, c^*) that accuracy-dominates (b, c) when accuracy is measured by $\mathfrak{I}_{\mathfrak{i},\mathfrak{q}}$ where $\mathfrak{i}$ is given by $R = 0$, $N = 1 - t$ and $W = 1$.

- We say that a scoring rule is **normalised** if $\mathsf{s}(0,0) = \mathsf{s}(1,1) = 0$ and $\mathsf{s}(0,1) = \mathsf{s}(1,0) = 1$.
- We say that a scoring rule generates a **connected** inaccuracy measure $\mathfrak{I}_{\mathsf{s}}$ if, for each credence function, the set of credence functions that dominates it is either empty or connected (in the topological sense).

Theorem 5 Suppose $\mathfrak{s}$ is a continuous, normalised strictly proper scoring rule that generates a connected inaccuracy measure $\mathfrak{I}_{\mathfrak{s}}$. Then there is $0.5 < r \leq 1$ such that, for all Lockean thresholds $t \geq r$, there is a Lockean$_t$ state (b, c) such that

- c is non-probabilistic;
- There is no Lockean$_t$ state (b^*, c^*) that accuracy-dominates (b, c) when accuracy is measures by $\mathfrak{I}_{\mathfrak{i},\mathfrak{s}}$, where $\mathfrak{i}$ is given by $R = 0$, $N = 1 - t$ and $W = 1$.

In sum: We have seen that, if categorical doxastic states such as full belief, full disbelief, and suspension of judgment are distinct existences, and if they are connected by the metaphysical or normative versions of the Lockean Thesis, then the Accuracy Dominance Argument for Probabilism fails: there are total doxastic states that include non-probabilistic credence functions that are not accuracy dominated when their total inaccuracy is considered.

In sum: We have seen that, if categorical doxastic states such as full belief, full disbelief, and suspension of judgment are distinct existences, and if they are connected by the metaphysical or normative versions of the Lockean Thesis, then the Accuracy Dominance Argument for Probabilism fails: there are total doxastic states that include non-probabilistic credence functions that are not accuracy dominated when their total inaccuracy is considered.

On the other hand, if they are not distinct existences and are instead connected by the analytic version of the Lockean Thesis, then the Accuracy Dominance Argument for Probabilism succeeds.

- Accuracy, credences and full beliefs
- Two important distinctions
- AGM, credences and full beliefs
- More on credences and full beliefs

1. If Shakespeare had not written Hamlet, it would never have been written.
2. If Shakespeare didn't write Hamlet, someone else did.

1. is a causal counterfactual, and 2. is an expression of a belief revision policy.

1. General Smith is a shrewd judge of character—he knows (better than I) who is brave and who is not.
2. The general sends only brave men into battle.
3. Private Jones is cowardly.

I believe that (1) Jones would run away if he were sent into battle and (2) if Jones *is* sent into battle, then he won't run away.

1. Ann cheats — she has seen her opponent's cards.
2. Ann has a losing hand, since I have seen both her hand and her opponent's.
3. Ann is rational.

So, I conclude that she will not bet. But how should I revise my beliefs if I learn that Ann did bet?

1. Ann cheats — she has seen her opponent's cards.
2. Ann has a losing hand, since I have seen both her hand and her opponent's.
3. Ann is rational.

So, I conclude that she will not bet. But how should I revise my beliefs if I learn that Ann did bet?

It may be perfectly reasonable for me to be disposed to give up 2.

1. Ann cheats — she has seen her opponent's cards.
2. Ann has a losing hand, since I have seen both her hand and her opponent's.
3. Ann is rational.

So, I conclude that she will not bet. But how should I revise my beliefs if I learn that Ann did bet?

It may be perfectly reasonable for me to be disposed to give up 2.

I believe that (1) I Ann *were* to bet, she would lose (since she has a losing hand) and (2) If I were to learn that she *did* bet, I would conclude she will win.

Updating vs. Revising

Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

An important distinction:

- If φ describes facts about the current state of affairs
- If φ describes facts that have possible become true only after the original beliefs were formed.

Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

An important distinction:

- If φ describes facts about the current state of affairs
- If φ describes facts that have possible become true only after the original beliefs were formed.

Revising by $\neg p$ ($K * \neg p$) vs. Updating by $\neg p$ ($K \diamond \neg p$)

H. Katsuno and A. O. Mendelzon. *Propositional knowledge base revision and minimal change*. Artificial Intelligence, 52, pp. 263 - 294 (1991).

The logic of updating differs from that of revision. This can be seen from the following example:

To begin with, the agent knows that there is either a book on the table (p) or a magazine on the table (q), but not both.

- Case 1: The agent is told that there is a book on the table. She concludes that there is no magazine on the table. This is revision.
- Case 2: The agent is told that after the first information was given, a book has been put on the table. In this case she should not conclude that there is no magazine on the table. This is updating.

J. Lang. *Belief Update Revisited*. Proceedings of IJCAI-07.

N. Friedman and J. Halpern. *Modeling Belief in Dynamics Systems Part II: Revision and Update*. Journal of Artificial Intelligence Research, 10, pp. 117 - 167 (1999).

A. Herzig. *Belief Change Operations: A shorty history of nearly everything, told in dynamic logic of propositional assignments*. AAAI, 2014.

In the literature on belief change the distinction between static and dynamic environment has become important....

In the literature on belief change the distinction between static and dynamic environment has become important....it seems right to say that belief change due to new information in an unchanging environment has come to be called belief revision (the static case, in the sense that the "world" remains unchanged), while it is fairly generally accepted to use the term belief update for belief change that is due to reported changes in the environment itself (the dynamic case, in the sense that the "world" changes; compare our analysis in the last subsection).

In the literature on belief change the distinction between static and dynamic environment has become important....it seems right to say that belief change due to new information in an unchanging environment has come to be called belief revision (the static case, in the sense that the "world" remains unchanged), while it is fairly generally accepted to use the term belief update for belief change that is due to reported changes in the environment itself (the dynamic case, in the sense that the "world" changes; compare our analysis in the last subsection). It has been held for some time that these cases support different logics (...)

In the literature on belief change the distinction between static and dynamic environment has become important....it seems right to say that belief change due to new information in an unchanging environment has come to be called belief revision (the static case, in the sense that the "world" remains unchanged), while it is fairly generally accepted to use the term belief update for belief change that is due to reported changes in the environment itself (the dynamic case, in the sense that the "world" changes; compare our analysis in the last subsection). It has been held for some time that these cases support different logics (...) The established tradition notwithstanding, it would be interesting to see a really convincing argument for tying AGM revision to static environments.

Hannes Leitgeb and Krister Segerberg. *Dynamic doxastic logic: why, how, and where to?*. Synthese, 155, pp. 167 - 190 (2007).

KM Postulates

KM 1: $K \diamond \varphi = Cn(K \diamond \varphi)$

KM 2: $\varphi \in K \diamond \varphi$

KM 3: If $\varphi \in K$ then $K \diamond \varphi = K$

KM 4: $K \diamond \varphi$ is inconsistent iff φ is inconsistent

KM 5: If φ and ψ are logically equivalent then $K \diamond \varphi = K \diamond \psi$

KM 6: $K \diamond (\varphi \wedge \psi) \subseteq Cn(K \diamond \varphi \cup \{\psi\})$

KM 7: If $\psi \in K \diamond \varphi$ and $\varphi \in K \diamond \psi$ then $K \diamond \varphi = K \diamond \psi$

KM 8: If K is complete then $K \diamond (\varphi \wedge \psi) \subseteq K \diamond \varphi \cap K \diamond \psi$

KM 9: $K \diamond \varphi = \bigcap_{M \in Comp(K)} M \diamond \varphi$, where $Comp(K)$ is the class of all complete theories containing K.

Updating and Revising

$$K \diamond \varphi = \bigcap_{M \in Comp(K)} M * \varphi$$

H. Katsuno and A. O. Mendelzon. *On the difference between updating a knowledge base and revising it*. *Belief Revision*, P. Gärdenfors (ed.), pp 182 - 203 (1992).

T. Shear, J. Weisberg and B. Fitelson. *Two Approaches to Belief Revision*. manuscript, 2016.

$$u(B(p), w) = \begin{cases} \mathsf{r} & \text{if } p \text{ is true at } w \\ -\mathsf{w} & \text{if } p \text{ is false at } w \end{cases}$$

$$1 \geq \mathsf{w} > \left(\frac{1+\sqrt{5}}{2}\right) \cdot \mathsf{r} > 0$$

$$EEU(B(p), b) := \sum_{w \in W} b(w)u(B(p), w)$$

$$EEU(B, b) := \sum_{p \in B} EEU(B(p), w)$$

Theorem (Dorst). An agent's belief set B maximizes EEU from the point of view of her credence function b if and only if, for every $p \in B$

$$b(p) > \frac{\mathsf{w}}{\mathsf{r} + \mathsf{w}}$$

$$B \divideontimes E = \{p \mid b(p \mid E) > \frac{\mathsf{w}}{\mathsf{r} + \mathsf{w}}\}$$

(P2) If an agent initially believes X (i.e., if $X \in B$), then updating B on X should *not change* B. [More formally, $X \in B$ implies that $B' = B \star X = B$

Proposition. Suppose $b(p) > \frac{\mathrm{w}}{\mathrm{r}+\mathrm{w}}$ and $b(q) > \frac{\mathrm{w}}{\mathrm{r}+\mathrm{w}}$ (i.e., that our deductive cogent EUT agent believes both p and q). And, following the constraint, suppose that $\varphi - 1 < \frac{\mathrm{w}}{\mathrm{r}+\mathrm{w}} \leq 1$. Then $b(p \mid q) > \frac{\mathrm{w}-\mathrm{r}}{\mathrm{w}}$ and $b(p \mid q) - b(p) < \frac{\mathrm{r}^2}{\mathrm{rw}+\mathrm{w}^2}$

Cogency. An agent's belief set B should (at any given time) be deductively cogent, i.e., B should be both deductively consistent and closed under logic.

Foley

...if the avoidance of recognizable inconsistency were an absolute prerequisite of rational belief, we could not rationally believe each member of a set of propositions and also rationally believe of this set that at least one of its members is false. But this in turn pressures us to be unduly cautious. It pressures us to believe only those propositions that are certain or at least close to certain for us, since otherwise we are likely to have reasons to believe that at least one of these propositions is false. At first glance, the requirement that we avoid recognizable inconsistency seems little enough to ask in the name of rationality. It asks only that we avoid certain error. It turns out, however, that this is far too much to ask.

AGM Postulates

Closure $\quad B * E = Cn(B * E)$

AGM Postulates

Closure $B * E = Cn(B * E)$

Success $E \in B * E$

AGM Postulates

Closure $B * E = Cn(B * E)$

Success $E \in B * E$

Inclusion $B * E \subseteq Cn(B \cup \{E\})$

AGM Postulates

Closure $B * E = Cn(B * E)$

Success $E \in B * E$

Inclusion $B * E \subseteq Cn(B \cup \{E\})$

Vacuity If E is consistent with B, then $B * E \supseteq Cn(B \cup \{E\})$

AGM Postulates

Closure $B * E = Cn(B * E)$

Success $E \in B * E$

Inclusion $B * E \subseteq Cn(B \cup \{E\})$

Vacuity If E is consistent with B, then $B * E \supseteq Cn(B \cup \{E\})$

Consistency If E is not self-contradictory, then $B * E$ is consistent

AGM Postulates

Closure $B * E = Cn(B * E)$

Success $E \in B * E$

Inclusion $B * E \subseteq Cn(B \cup \{E\})$

Vacuity If E is consistent with B, then $B * E \supseteq Cn(B \cup \{E\})$

Consistency If E is not self-contradictory, then $B * E$ is consistent

Extensionality If $X \equiv Y \in Cn(\emptyset)$, then $B * X = B * Y$

AGM Postulates

Closure $B * E = Cn(B * E)$

Success $E \in B * E$

Inclusion $B * E \subseteq Cn(B \cup \{E\})$

Vacuity If E is consistent with B, then $B * E \supseteq Cn(B \cup \{E\})$

Consistency If E is not self-contradictory, then $B * E$ is consistent

Extensionality If $X \equiv Y \in Cn(\emptyset)$, then $B * X = B * Y$

Superexpansion $B * (X \wedge Y) \subseteq Cn((B * X) \cup \{Y\})$

AGM Postulates

Closure $B * E = Cn(B * E)$

Success $E \in B * E$

Inclusion $B * E \subseteq Cn(B \cup \{E\})$

Vacuity If E is consistent with B, then $B * E \supseteq Cn(B \cup \{E\})$

Consistency If E is not self-contradictory, then $B * E$ is consistent

Extensionality If $X \equiv Y \in Cn(\emptyset)$, then $B * X = B * Y$

Superexpansion $B * (X \wedge Y) \subseteq Cn((B * X) \cup \{Y\})$

Subexpansion If Y is consistent with $Cn(B * X)$, then
$B * (X \wedge Y) \supseteq Cn((B * X) \cup \{Y\})$

Claim. (P2) follows from the AGM postulates Closure, Inclusion and Vacuity.

Claim. (P2) follows from the AGM postulates Closure, Inclusion and Vacuity.

1.	B is consistent.	Assumption

Claim. (P2) follows from the AGM postulates Closure, Inclusion and Vacuity.

1.	B is consistent.	Assumption
2.	B is closed, i.e., $B = Cn(B)$.	Assumption

Claim. (P2) follows from the AGM postulates Closure, Inclusion and Vacuity.

1.	B is consistent.	Assumption
2.	B is closed, i.e., $B = Cn(B)$.	Assumption
3.	$X \in B$.	Assumption

Claim. (P2) follows from the AGM postulates Closure, Inclusion and Vacuity.

1.	B is consistent.	Assumption
2.	B is closed, i.e., $B = Cn(B)$.	Assumption
3.	$X \in B$.	Assumption
4.	X is consistent with B.	(1), (3), Logic

Claim. (P2) follows from the AGM postulates Closure, Inclusion and Vacuity.

1.	B is consistent.	Assumption
2.	B is closed, i.e., $B = Cn(B)$.	Assumption
3.	$X \in B$.	Assumption
4.	X is consistent with B.	(1), (3), Logic
5.	$B * X = Cn(B \cup \{X\})$.	(4), Vacuity, Inclusion

Claim. (P2) follows from the AGM postulates Closure, Inclusion and Vacuity.

1.	B is consistent.	Assumption
2.	B is closed, i.e., $B = Cn(B)$.	Assumption
3.	$X \in B$.	Assumption
4.	X is consistent with B.	(1), (3), Logic
5.	$B * X = Cn(B \cup \{X\})$.	(4), Vacuity, Inclusion
6.	$B * X = Cn(B)$.	(5), (3), Logic

Claim. (P2) follows from the AGM postulates Closure, Inclusion and Vacuity.

1.	B is consistent.	Assumption
2.	B is closed, i.e., $B = Cn(B)$.	Assumption
3.	$X \in B$.	Assumption
4.	X is consistent with B.	(1), (3), Logic
5.	$B * X = Cn(B \cup \{X\})$.	(4), Vacuity, Inclusion
6.	$B * X = Cn(B)$.	(5), (3), Logic
7.	$B * X = Cn(B * X)$	Closure

Claim. (P2) follows from the AGM postulates Closure, Inclusion and Vacuity.

1.	B is consistent.	Assumption
2.	B is closed, i.e., $B = Cn(B)$.	Assumption
3.	$X \in B$.	Assumption
4.	X is consistent with B.	(1), (3), Logic
5.	$B * X = Cn(B \cup \{X\})$.	(4), Vacuity, Inclusion
6.	$B * X = Cn(B)$.	(5), (3), Logic
7.	$B * X = Cn(B * X)$	Closure
8.	$Cn(B * X) = Cn(B)$	(6), (7), Logic

Claim. (P2) follows from the AGM postulates Closure, Inclusion and Vacuity.

1.	B is consistent.	Assumption
2.	B is closed, i.e., $B = Cn(B)$.	Assumption
3.	$X \in B$.	Assumption
4.	X is consistent with B.	(1), (3), Logic
5.	$B * X = Cn(B \cup \{X\})$.	(4), Vacuity, Inclusion
6.	$B * X = Cn(B)$.	(5), (3), Logic
7.	$B * X = Cn(B * X)$	Closure
8.	$Cn(B * X) = Cn(B)$	(6), (7), Logic
9.	$B * X = B$	(7), (8), (2), Logic

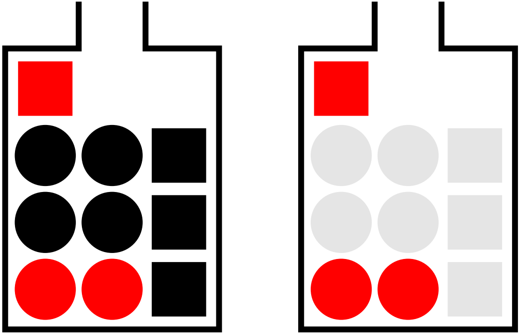
Theorem. (Gärdenfors) Suppose $r = 0$, $w = 1$, B is synchronically coherent in the EUT sense, and that for all propositions X and Y that our agent might learn, $b(X \mid Y) > 0$. Then $\divideontimes$ satisfies all eight of the AGM postulates above.

Proposition 1 *EUT* Revision (Generally) Satisfies Success.
Proposition 2 *EUT* Revision (Generally) Satisfies Inclusion.
Proposition 3 *EUT* Revision (Generally) Satisfies Extensionality.
Proposition 4 *EUT* Revision (Generally) Satisfies Superexpansion.

Proposition 5 Non-Extremal EUT Revision Violates Consistency and Closure.

Proposition 6 Non-Extremal *EUT* Revision Violates Vacuity - even if it is restricted to deductively cogent agents.

TED SHEAR, JONATHAN WEISBERG, AND BRANDEN FITELSON

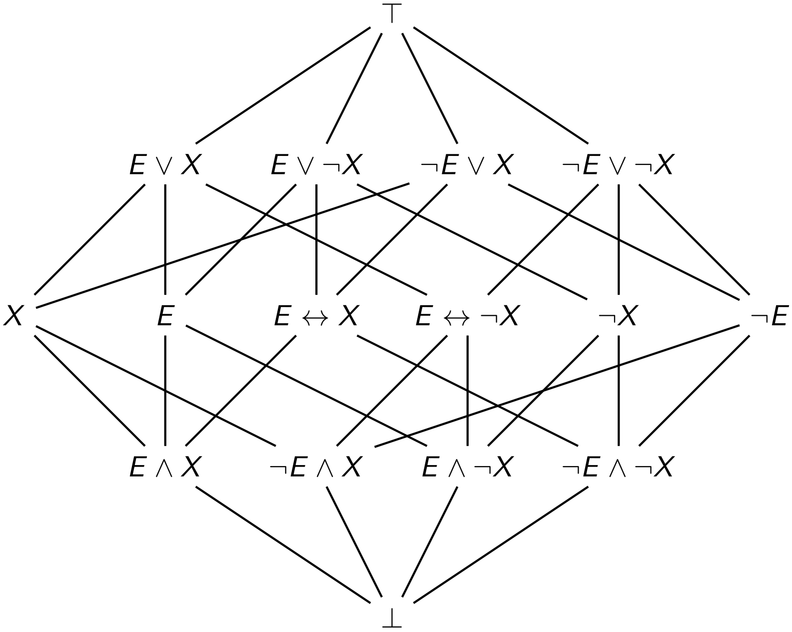


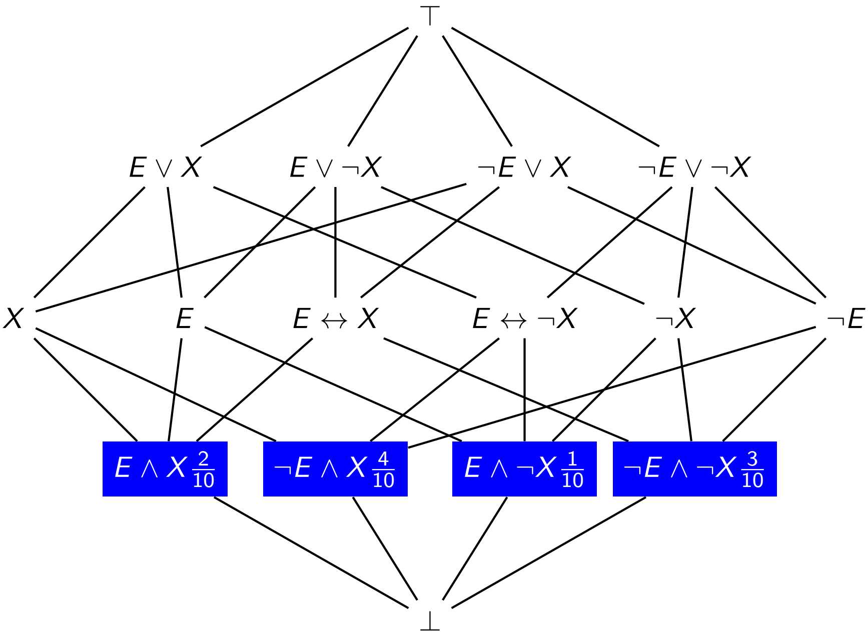
(A) Prior distribution (B) Posterior (given E)

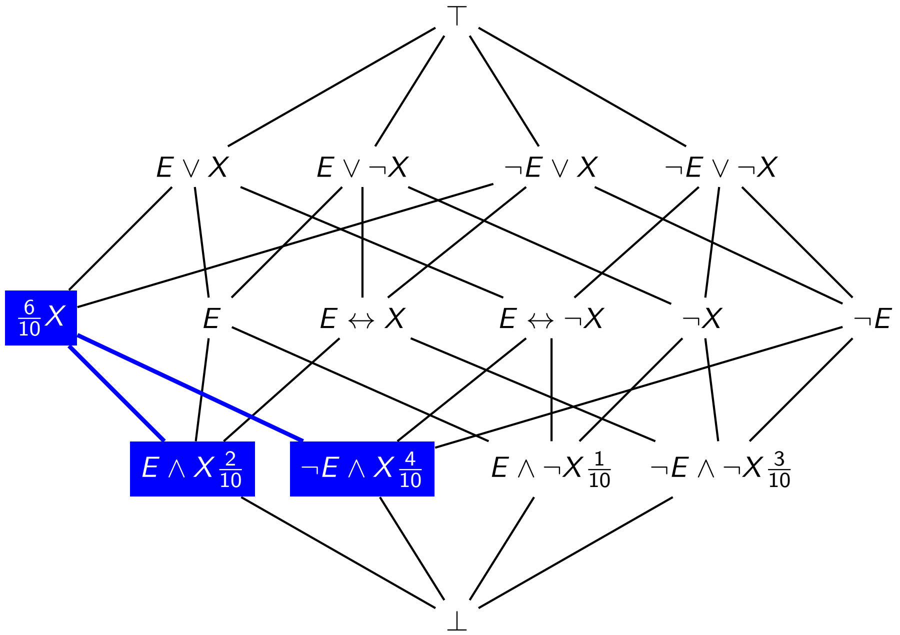
FIGURE 2. Visualization of counterexample to **Vacuity** for EUT Revision

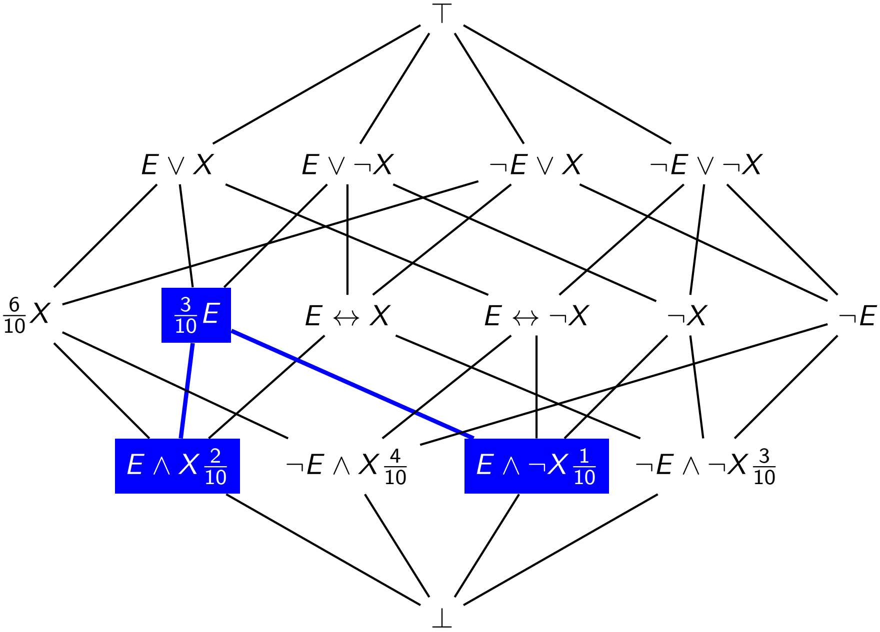
E :='The object sampled from the urn is red'

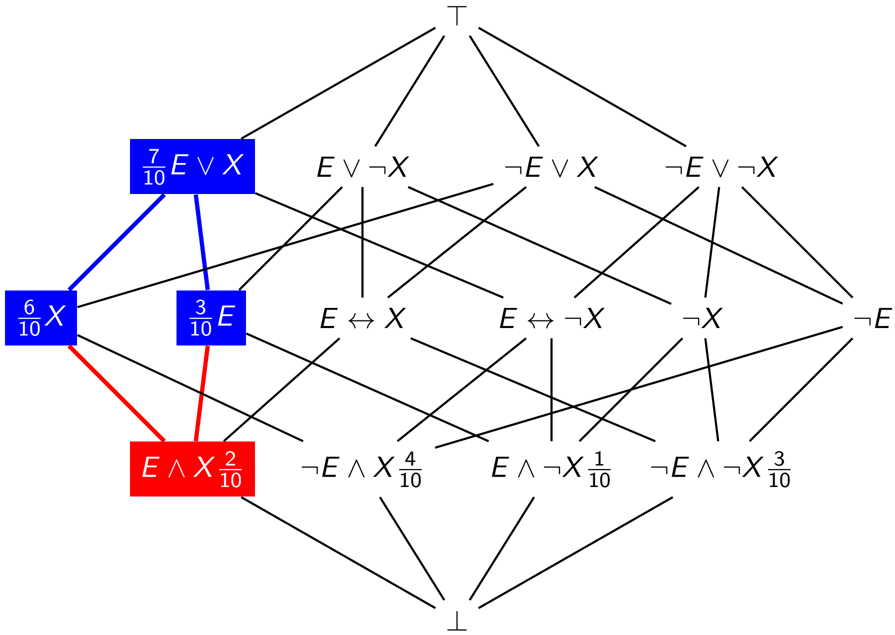
X := 'The object sampled from the urn is a circle'.

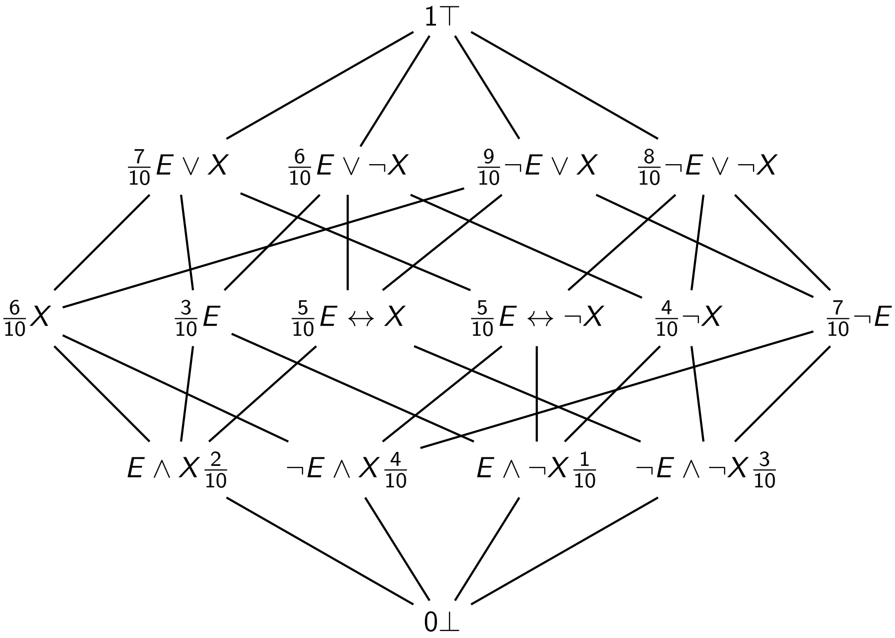

$\top$
$E \vee X$
$E \vee \neg X$
$\neg E \vee X$
$\neg E \vee \neg X$
X
E
$E \leftrightarrow X$
$E \leftrightarrow \neg X$
$\neg X$
$\neg E$
$E \wedge X$
$\neg E \wedge X$
$E \wedge \neg X$
$\neg E \wedge \neg X$
$\bot$


⊤
E ∨ X
E ∨ ¬X
¬E ∨ X
¬E ∨ ¬X
X
E
E ↔ X
E ↔ ¬X
¬X
¬E
E ∧ X 2/10
¬E ∧ X 4/10
E ∧ ¬X 1/10
¬E ∧ ¬X 3/10
⊥

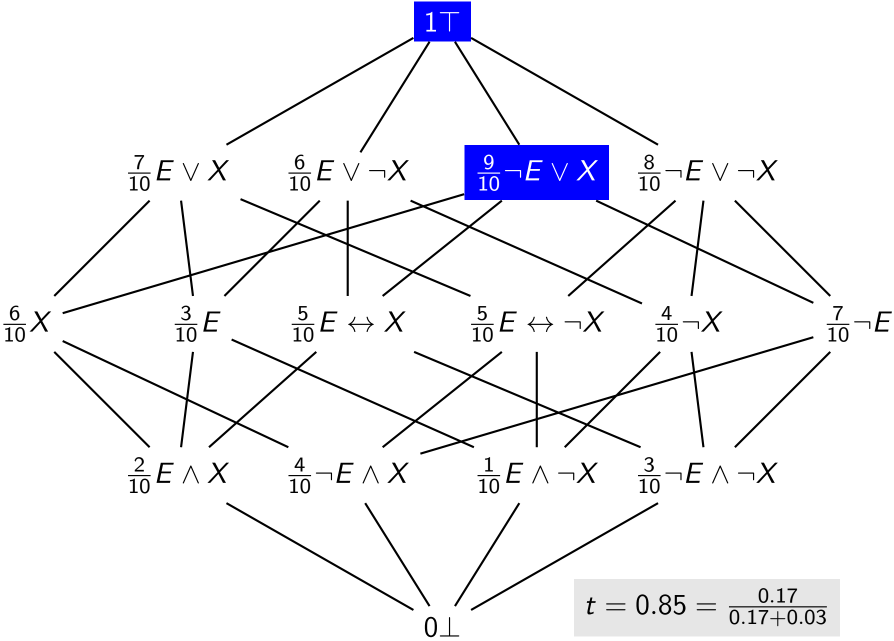

⊤
E ∨ X
E ∨ ¬X
¬E ∨ X
¬E ∨ ¬X
6/10 X
E
E ↔ X
E ↔ ¬X
¬X
¬E
E ∧ X 2/10
¬E ∧ X 4/10
E ∧ ¬X 1/10
¬E ∧ ¬X 3/10
⊥


$\top$
$E \vee X$
$E \vee \neg X$
$\neg E \vee X$
$\neg E \vee \neg X$
$\frac{6}{10} X$
$\frac{3}{10} E$
$E \leftrightarrow X$
$E \leftrightarrow \neg X$
$\neg X$
$\neg E$
$E \wedge X \, \frac{2}{10}$
$\neg E \wedge X \, \frac{4}{10}$
$E \wedge \neg X \, \frac{1}{10}$
$\neg E \wedge \neg X \, \frac{3}{10}$
$\bot$

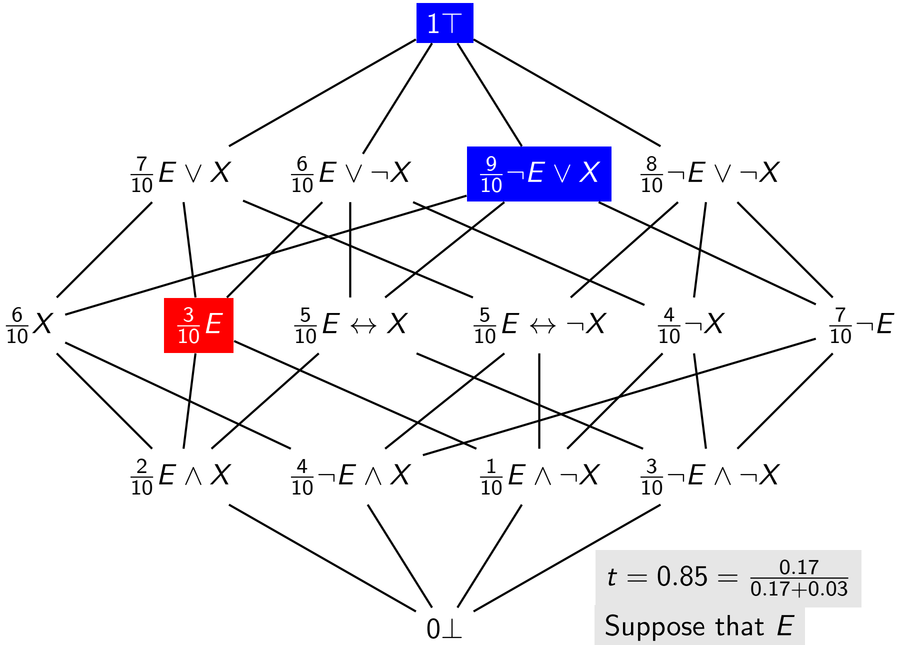

$\top$
$\frac{7}{10} E \vee X$
$E \vee \neg X$
$\neg E \vee X$
$\neg E \vee \neg X$
$\frac{6}{10} X$
$\frac{3}{10} E$
$E \leftrightarrow X$
$E \leftrightarrow \neg X$
$\neg X$
$\neg E$
$E \wedge X \, \frac{2}{10}$
$\neg E \wedge X \, \frac{4}{10}$
$E \wedge \neg X \, \frac{1}{10}$
$\neg E \wedge \neg X \, \frac{3}{10}$
$\bot$



$1 \top$
$\frac{7}{10} E \vee X$
$\frac{6}{10} E \vee \neg X$
$\frac{9}{10} \neg E \vee X$
$\frac{8}{10} \neg E \vee \neg X$
$\frac{6}{10} X$
$\frac{3}{10} E$
$\frac{5}{10} E \leftrightarrow X$
$\frac{5}{10} E \leftrightarrow \neg X$
$\frac{4}{10} \neg X$
$\frac{7}{10} \neg E$
$E \wedge X \frac{2}{10}$
$\neg E \wedge X \frac{4}{10}$
$E \wedge \neg X \frac{1}{10}$
$\neg E \wedge \neg X \frac{3}{10}$
$0 \bot$

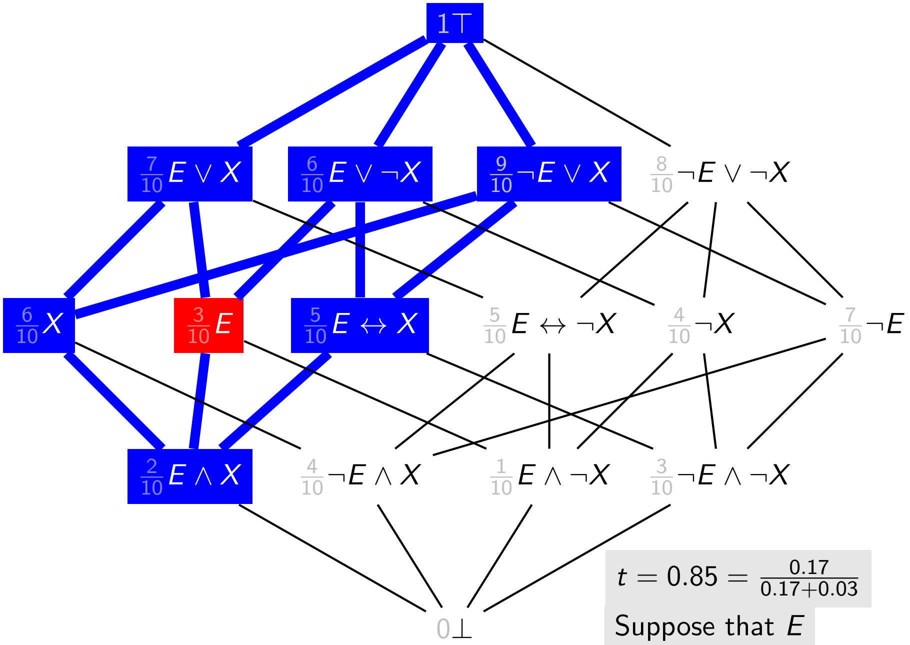


$t = 0.85 = \frac{0.17}{0.17+0.03}$



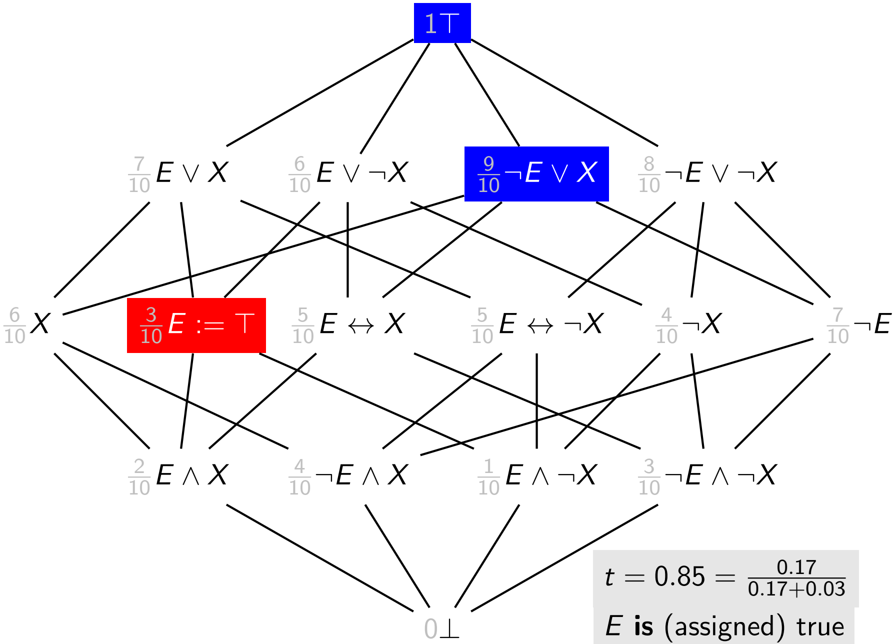
$t = 0.85 = \frac{0.17}{0.17+0.03}$

Suppose that E



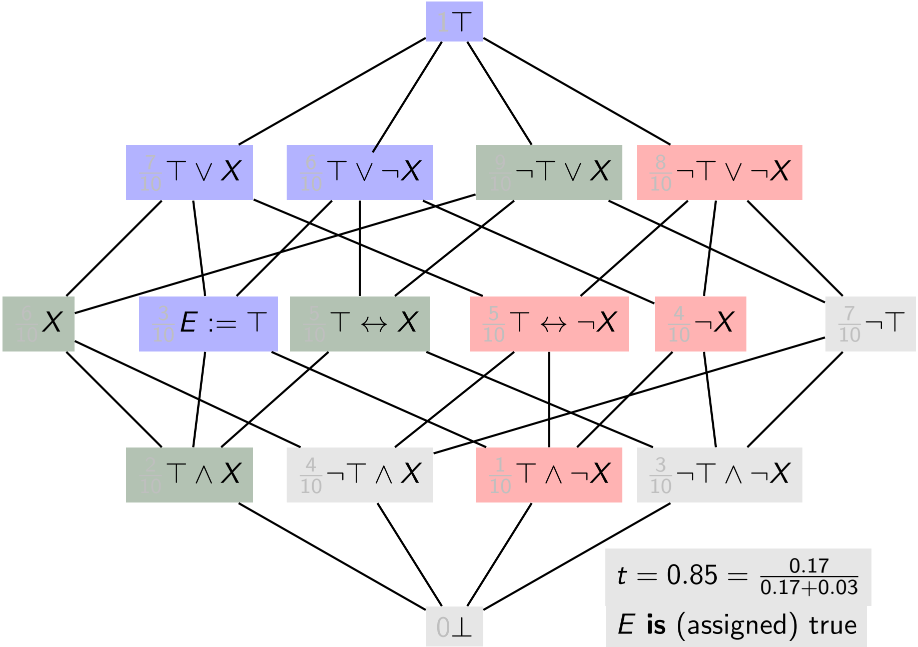
$t = 0.85 = \frac{0.17}{0.17+0.03}$

Suppose that E



$t = 0.85 = \frac{0.17}{0.17+0.03}$

E **is** (assigned) true



$t = 0.85 = \frac{0.17}{0.17+0.03}$

E **is** (assigned) true

$E \vee X \quad E \vee \neg X \quad E \quad \top$

$\neg E \vee X \quad E \leftrightarrow X \quad E \wedge X \quad X \qquad \neg X \quad E \leftrightarrow \neg X \quad E \wedge \neg X \quad \neg E \vee \neg X$

$\neg E \wedge \neg X \quad \neg E \wedge X \quad \neg E \quad \bot$

E **is** (assigned) true

$E \vee X$ $\quad E \vee \neg X$ $\quad E$ $\quad \top$

$\neg E \vee X$ $\quad E \leftrightarrow X$ $\quad E \wedge X$ $\quad X$ $\qquad \neg X$ $\quad E \leftrightarrow \neg X$ $\quad E \wedge \neg X$ $\quad \neg E \vee \neg X$

$\neg E \wedge \neg X$ $\quad \neg E \wedge X$ $\quad \neg E$ $\quad \bot$

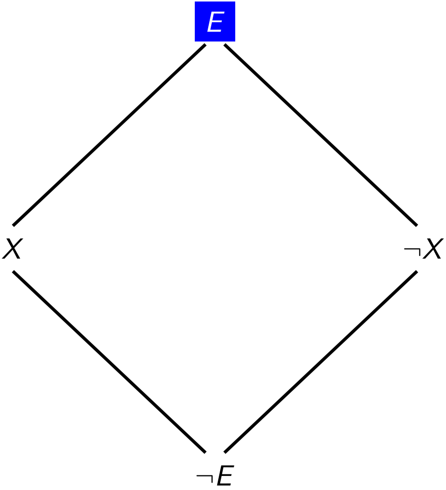
E **is** (assigned) true

$\boxed{E \vee X}$ $\boxed{E \vee \neg X}$ $\boxed{E}$ $\boxed{\top}$

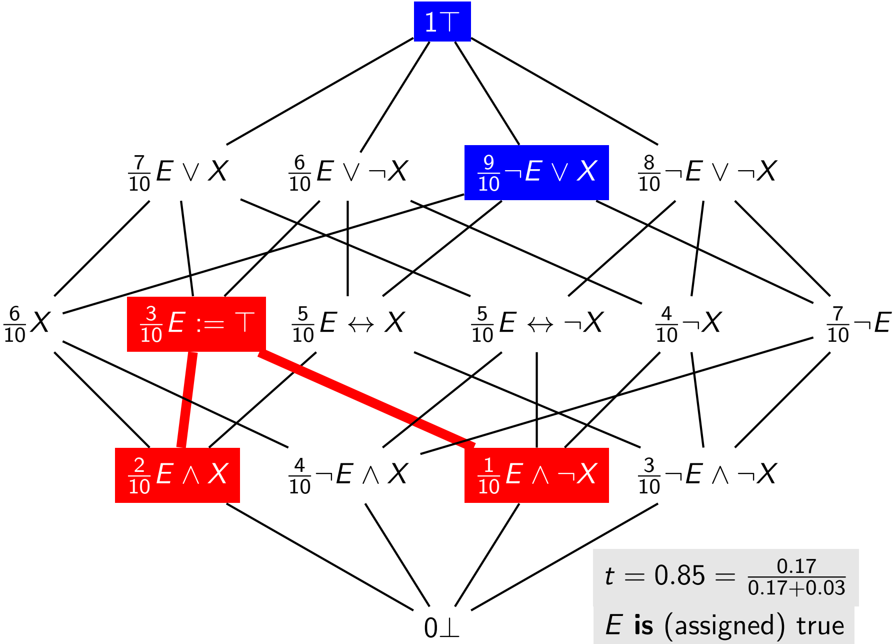
$\neg E \vee X$ $E \leftrightarrow X$ $E \wedge X$ X $\neg X$ $E \leftrightarrow \neg X$ $E \wedge \neg X$ $\neg E \vee \neg X$

$\neg E \wedge \neg X$ $\neg E \wedge X$ $\neg E$ $\bot$

E **is** (assigned) true

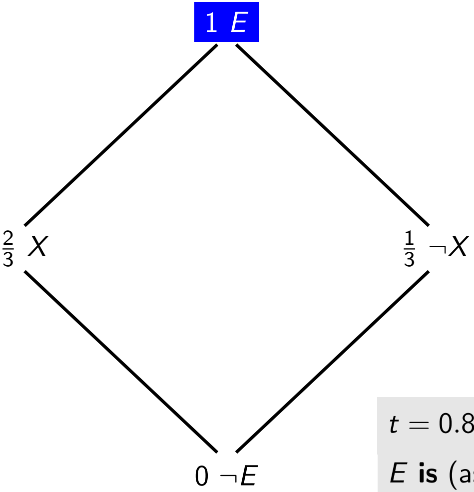


E **is** (assigned) true



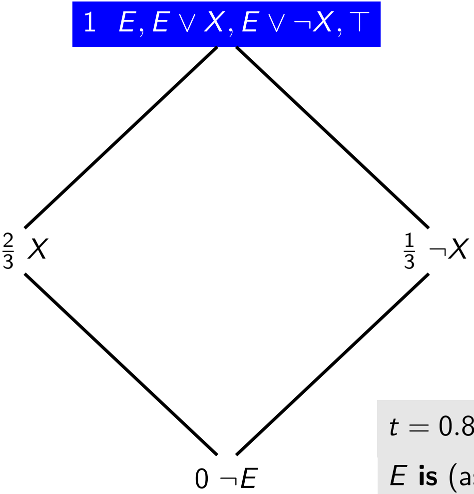
$t = 0.85 = \frac{0.17}{0.17+0.03}$

E **is** (assigned) true



$t = 0.85 = \frac{0.17}{0.17+0.03}$

E **is** (assigned) true



$t = 0.85 = \frac{0.17}{0.17+0.03}$

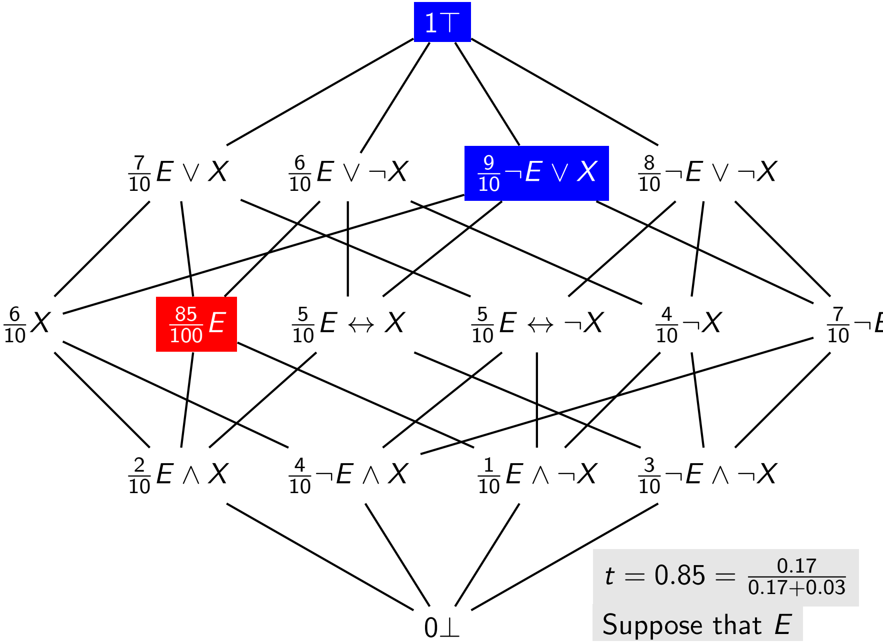
E **is** (assigned) true

Vacuity If E is consistent with B, then $B * E \supseteq Cn(B \cup \{E\})$

$B = Cn(\{\neg E \vee X\})$

$B \nvdash \neg E$, $B * E = Cn(B \cup \{E\}) = Cn(\{E \wedge X\})$ So, $X \in B * E$.

$B \divideontimes E = Cn(\{E, E \vee X, E \vee \neg X\})$, so $X \notin B \divideontimes E$.



$t = 0.85 = \frac{0.17}{0.17+0.03}$

Suppose that E

Leitgeb & Segerberg: Belief Update vs. Belief Revsions

...given new evidence, we find that in the case of belief revision the agent tries to change his beliefs in a manner such that the worlds that he subsequently believes to be in comprise the *subjectively most plausible deviation* from the worlds he originally believed to inhabit.

Leitgeb & Segerberg: Belief Update vs. Belief Revsions

...given new evidence, we find that in the case of belief revision the agent tries to change his beliefs in a manner such that the worlds that he subsequently believes to be in comprise the *subjectively most plausible deviation* from the worlds he originally believed to inhabit.

However, when confronted with the same evidence in belief update, the agent tries to change his beliefs in a way such that the worlds that he subsequently believes to be in are as **objectively similar as possible** to the worlds he originally believed to be the most plausible candidates for being the actual world.

Leitgeb & Segerberg: Belief Update vs. Belief Revsions

It is tempting to relate these different views on belief change to the traditional distinction of indicative and subjective conditionals. Using the stock example: everyone considers the indicative 'If Oswald did not kill Kennedy somebody else did' as acceptable, but many regard the subjunctive 'If Oswald had not killed Kennedy somebody else would have' as false.

Note that in this setting the difference between supposing and updating is mathematically clearcut. In a typical Bayesian updating situation one is uncertain about the chances, and so ones subjective probability distribution on the outcome space is a mixture of the possible chance distributions. Updating is an operation which typically takes one from one point in the interior of the convex closure of the chance distributions to another; supposing moves from one chance distribution to another.

B. Skyrms. *Updating, Supposing and MAXENT*. Theory and Decision, 22, pp. 225 - 246, 1987.

Such normative virtues suggest a psychological question. One way of formulating (1) is that *supposing* an event B should have the same impact on the credibility of an event A as *learning* B. Is this true for typical assessments of chance? For example, is the judged probability of a Democratic victory in 2012 supposing that Hilary Clinton is the vice presidential candidate the same as the judged probability of a Democratic victory in 2012 after learning that Clinton, as a matter of fact, is the vice presidential candidate?

Jiaying Zhao, Vincenzo Crupi, Katya Tentori, Branden Fitelson, and Daniel Osherson. *Updating: Learning versus supposing*. Cognition 124 (2012) 373378.

Non-Extremal EUT revision is more conservative than AGM revision (when the two approaches interestingly) diverge:

Theorem EUT violates Vacuity (wrt B, E) if and only if E is consistent with B and $B \divideontimes E \subset B * E$

Non-Extremal EUT revision is more conservative than AGM revision (when the two approaches interestingly) diverge:

Theorem EUT violates Vacuity (wrt B, E) if and only if E is consistent with B and $B \divideontimes E \subset B * E$

In other words, when EUT and AGM (interestingly) diverge, AGM will be more demanding on an agents beliefs (insofar as they are maintained via revision). Since AGM will require agents to maintain beliefs in the face of counter-evidence (such as in our counter-example to Vacuity), it may be seen as an epistemically risk-seeking policy for belief revision. On the other hand, EUT will recommend that agents suspend belief in many cases and so it may be seen as epistemically risk-averse.

Proposition 8. If an EUT/Lockean agent is deductively cogent (at all times), then they can only violate Vacuity (via learning some E that they do not already believe) if their Lockean threshold is on the half-open interval $[\varphi - 1, 1)$.

- Accuracy, credences and full beliefs
- Two important distinctions
- AGM, credences and full beliefs
- More on credences and full beliefs