

Belief functions over Belnap–Dunn logic

Marta Bílková¹, Sabine Frittella², Daniil Kozhemiachenko²,
Ondrej Majer³ and Sajad Nazari²

¹Institute of Computer Science, ³Institute of Philosophy, Czech Academy of Sciences
²INSA Centre Val de Loire, Univ. Orléans

LATD + MOSAIC 2022

Contradictory information ...

We have to deal with inconsistent information al the time
(media, databases, scientific information, ...)

Contradictory information ...

We have to deal with inconsistent information al the time
(media, databases, scientific information, ...)

- get ride of inconsistencies (belief revision, ...)
- develop some tools for reasoning in these situations

Contradictory information ...

We have to deal with inconsistent information all the time
(media, databases, scientific information, ...)

- get rid of inconsistencies (belief revision, ...)
- develop some tools for reasoning in these situations

Inconsistencies: shortcomings of the available information
(Dunn: "... too much of a good thing")

We have to deal with inconsistent information all the time
(media, databases, scientific information, ...)

- get rid of inconsistencies (belief revision, ...)
- develop some tools for reasoning in these situations

Inconsistencies: shortcomings of the available information
(Dunn: "... too much of a good thing")

- issue that has to be lived with until they can be resolved
- many attempts in the literature to develop paraconsistent logics for these scenarios
- we build our work on the Belnap-Dunn four valued logic

Uncertainty

- probability – the most prominent representation of uncertainty
- defined on a boolean algebra
- each value precisely known (subjective probability)
- values of complex events are determined by the values of simple ones

Uncertainty

- probability – the most prominent representation of uncertainty
- defined on a boolean algebra
- each value precisely known (subjective probability)
- values of complex events are determined by the values of simple ones

... too perfect!!

- probability – the most prominent representation of uncertainty
- defined on a boolean algebra
- each value precisely known (subjective probability)
- values of complex events are determined by the values of simple ones

... too perfect!!

Alternatives:

- avoiding precise values (interval probabilities, ...)
- probability measures on weaker structures
- weaker uncertainty measures (inner outer measures, belief functions,...)

- 1 Representing incomplete/contradictory probabilistic information
 - Belnap-Dunn Logic
 - Non-standard probabilities

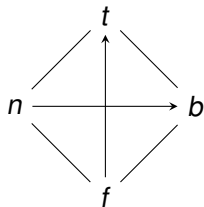
- 2 Dempster-Shafer theory
 - Mass functions, belief functions and plausibility functions
 - Representation of evidence

- 3 Dempster-Shafer theory and BD logic
 - Belief functions
 - Plausibility?
 - Combination of evidence

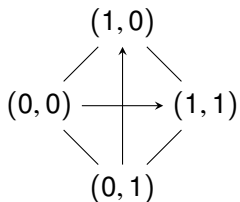
- 1 Representing incomplete/contradictory probabilistic information
 - Belnap-Dunn Logic
 - Non-standard probabilities
- 2 Dempster-Shafer theory
 - Mass functions, belief functions and plausibility functions
 - Representation of evidence
- 3 Dempster-Shafer theory and BD logic
 - Belief functions
 - Plausibility?
 - Combination of evidence

Belnap-Dunn square $(\mathbf{4}, \wedge, \vee, \neg)$ is a de Morgan algebra.

- $(\mathbf{4}, \wedge, \vee)$ is a lattice
- each element represents the available positive and/or negative information
 - n : no information
 - f : false (is bottom)
 - t : true (is top)
 - b : contradictory information
- \neg is an involutive de Morgan negation.



Belnap-Dunn square 4



Independence of positive and negative information.

Language. $L_{BD} \ni \varphi := p \in \text{Prop} \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \neg \varphi$

BD Models. $M = \langle W, v^+, v^- : \text{Prop} \rightarrow \mathcal{P}(W) \rangle$

$v^+(p)$: states containing information supporting p

$v^-(p)$: states containing information refuting p

Language. $L_{BD} \ni \varphi := p \in \text{Prop} \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \neg \varphi$

BD Models. $M = \langle W, v^+, v^- : \text{Prop} \rightarrow \mathcal{P}(W) \rangle$

$v^+(p)$: states containing information supporting p

$v^-(p)$: states containing information refuting p

Semantics. Two satisfaction relations \vDash^+, \vDash^-

$w \vDash^+ p$ iff $w \in v^+(p)$

$w \vDash^- p$ iff $w \in v^-(p)$

$w \vDash^+ \neg \phi$ iff $w \vDash^- \phi$

$w \vDash^- \neg \phi$ iff $w \vDash^+ \phi$

$w \vDash^+ \phi \wedge \phi'$ iff $w \vDash^+ \phi$ and $w \vDash^+ \phi'$

$w \vDash^- \phi \wedge \phi'$ iff $w \vDash^- \phi$ or $w \vDash^- \phi'$

$w \vDash^+ \phi \vee \phi'$ iff $w \vDash^+ \phi$ or $w \vDash^+ \phi'$

$w \vDash^- \phi \vee \phi'$ iff $w \vDash^- \phi$ and $w \vDash^- \phi'$

Language. $L_{BD} \ni \varphi := p \in \text{Prop} \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \neg \varphi$

BD Models. $M = \langle W, v^+, v^- : \text{Prop} \rightarrow \mathcal{P}(W) \rangle$

$v^+(p)$: states containing information supporting p

$v^-(p)$: states containing information refuting p

Semantics. Two satisfaction relations \vDash^+, \vDash^-

$w \vDash^+ p$ iff $w \in v^+(p)$

$w \vDash^- p$ iff $w \in v^-(p)$

$w \vDash^+ \neg \phi$ iff $w \vDash^- \phi$

$w \vDash^- \neg \phi$ iff $w \vDash^+ \phi$

$w \vDash^+ \phi \wedge \phi'$ iff $w \vDash^+ \phi$ and $w \vDash^+ \phi'$

$w \vDash^- \phi \wedge \phi'$ iff $w \vDash^- \phi$ or $w \vDash^- \phi'$

$w \vDash^+ \phi \vee \phi'$ iff $w \vDash^+ \phi$ or $w \vDash^+ \phi'$

$w \vDash^- \phi \vee \phi'$ iff $w \vDash^- \phi$ and $w \vDash^- \phi'$

Can we introduce probability in this framework?

Non-standard probabilities

Frame semantics (Klein, Majer, Rafiee Rad 2021)

- independence of positive and negative probabilistic information
- BD model extended with a (classical) probability measure.

Non-standard probabilities

Frame semantics (Klein, Majer, Rafiee Rad 2021)

- independence of positive and negative probabilistic information
- BD model extended with a (classical) probability measure.

A **probabilistic BD model** is a tuple $M = \langle W, v^+, v^-, m \rangle$, s.t. $\langle W, v^+, v^- \rangle$ is a BD model and $\mu : \mathcal{P}(W) \rightarrow [0, 1]$ is a probability measure on W

Positive probability

$$p^+(\varphi) := \mu(\|\varphi\|^+)$$

Negative probability of φ : $p^-(\varphi) := \mu(\|\varphi\|^-)$

Remark $p^-(\varphi) = p^+(\neg\varphi)$, otherwise $p^+(\varphi)$ and $p^-(\varphi)$ independent,

Theorem [Klein et al]

The following axiomatization of non-standard probabilities is complete with respect to the class of probabilistic frames.

(A1) normalization $0 \leq p^+(\varphi) \leq 1$

(A2) monotonicity if $\varphi \vdash_{BD} \psi$ then $p^+(\varphi) \leq p^+(\psi)$

(A3) import-export $p^+(\varphi \wedge \psi) + p^+(\varphi \vee \psi) = p^+(\varphi) + p^+(\psi)$.

Theorem [Klein et al]

The following axiomatization of non-standard probabilities is complete with respect to the class of probabilistic frames.

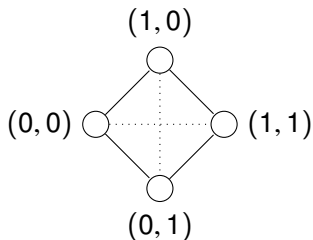
- (A1) normalization $0 \leq p^+(\varphi) \leq 1$
(A2) monotonicity if $\varphi \vdash_{BD} \psi$ then $p^+(\varphi) \leq p^+(\psi)$
(A3) import-export $p^+(\varphi \wedge \psi) + p^+(\varphi \vee \psi) = p^+(\varphi) + p^+(\psi)$.

Remarks

- $p^-(\varphi) = p^+(\neg\varphi)$
- weaker than classical Kolmogorovian axioms.
(A3 instead of additivity).
- In general $p^+(\neg\varphi) \neq 1 - p^+(\varphi)$
- one can have $0 < p^+(\varphi \wedge \neg\varphi), p^+(\varphi \vee \neg\varphi) < 1$

Continuous extension of Belnap-Dunn square, which we can see as the product bilattice $\mathbf{L}_{[0,1]} \odot \mathbf{L}_{[0,1]}$ with $\mathbf{L}_{[0,1]} = ([0, 1], \min, \max)$.

- $(p^+(\varphi), p^-(\varphi))$: positive and negative probabilistic support of φ .
- $(0, 0)$: no information concerning φ is available
- $(1, 1)$: maximally conflicting information
- vertical dashed line: “classical” case



- 1 Representing incomplete/contradictory probabilistic information
 - Belnap-Dunn Logic
 - Non-standard probabilities
- 2 Dempster-Shafer theory
 - Mass functions, belief functions and plausibility functions
 - Representation of evidence
- 3 Dempster-Shafer theory and BD logic
 - Belief functions
 - Plausibility?
 - Combination of evidence

Dempster Shafer belief functions

- $\text{bel}(A)$ represents **total evidence** supporting A
- defined on a boolean algebra of events
- weaker than probability
- complex formulas are not determined by the simpler ones
 $\text{bel}(A \vee B) \geq \text{bel}(A) + \text{bel}(B)$ for A, B disjoint
- provides a lower bound for 'true' probability

Dempster Shafer belief functions

- $\text{bel}(A)$ represents **total evidence** supporting A
- defined on a boolean algebra of events
- weaker than probability
- complex formulas are not determined by the simpler ones
 $\text{bel}(A \vee B) \geq \text{bel}(A) + \text{bel}(B)$ for A, B disjoint
- provides a lower bound for ‘true’ probability

Example

Scenario. A patient has disease a , b or c .

A doctor says “given the evidence she has disease a or b with a high certainty (e.g. 0.7).” However the doctor gives only a low certainty to each of a, b (e.g. 0.1).

Mass function

– evidence assigned *exactly* to a particular event

.

Definition

A function $m : \mathcal{P}(S) \rightarrow [0, 1]$ is a mass function if

- $m(\emptyset) = 0$
- $\sum_{A \in \mathcal{P}(S)} m(A) = 1$.

Evidence via mass functions

Mass function

– evidence assigned *exactly* to a particular event

.

Definition

A function $m : \mathcal{P}(S) \rightarrow [0, 1]$ is a mass function if

- $m(\emptyset)$
- $\sum_{A \in \mathcal{P}(S)} m(A) = 1.$

Belief function via mass function:

$$\text{bel}(A) = \sum_{B \subseteq A} m(B)$$

Evidence via mass functions

Mass function

– evidence assigned *exactly* to a particular event

.

Definition

A function $m : \mathcal{P}(S) \rightarrow [0, 1]$ is a mass function if

- $m(\emptyset)$
- $\sum_{A \in \mathcal{P}(S)} m(A) = 1.$

Belief function via mass function:

$$\text{bel}(A) = \sum_{B \subseteq A} m(B)$$

Probability via mass function:

$$p(A) = \sum_{B \subseteq A} m(B) = \sum_{s \in A} m(\{s\})$$

All information is encoded in singletons.

Definition

$f : \mathcal{P}(S) \rightarrow [0, 1]$ is a belief function if $f(\emptyset) = 0$, $f(S) = 1$ and

$$f\left(\bigvee_{1 \leq i \leq k} A_i\right) \geq \sum_{\substack{J \subseteq \{1, \dots, k\} \\ J \neq \emptyset}} (-1)^{|J|+1} \cdot f\left(\bigwedge_{j \in J} A_j\right). \quad (1)$$

holds for every $k \geq 1$, and for every $A_1, \dots, A_k \in \mathcal{P}(S)$.

Definition

$f : \mathcal{P}(S) \rightarrow [0, 1]$ is a belief function if $f(\emptyset) = 0$, $f(S) = 1$ and

$$f\left(\bigvee_{1 \leq i \leq k} A_i\right) \geq \sum_{\substack{J \subseteq \{1, \dots, k\} \\ J \neq \emptyset}} (-1)^{|J|+1} \cdot f\left(\bigwedge_{j \in J} A_j\right). \quad (1)$$

holds for every $k \geq 1$, and for every $A_1, \dots, A_k \in \mathcal{P}(S)$.

Theorem

For every belief function bel there is a mass function $m_{\text{bel}} : \mathcal{P}(S) \rightarrow [0, 1]$ such that, for every $A \in \mathcal{P}(S)$,

$$\text{bel}(A) = \sum_{B \leq A} m_{\text{bel}}(B)$$

Plausibility functions

- dual to belief
- $pl(A)$ represents evidence which is *compatible* with A
- gives an upper bound for 'true' probability

Plausibility from belief

$$pl(A) = 1 - bel(\neg A)$$

Plausibility via mass function:

$$bel(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

Explicit definition

$$f\left(\bigwedge_{1 \leq i \leq k} A_i\right) \leq \sum_{\substack{J \subseteq \{1, \dots, k\} \\ J \neq \emptyset}} (-1)^{|J|+1} \cdot f\left(\bigvee_{j \in J} A_j\right). \quad (2)$$

Example

Scenario. A patient has disease a , b or c .

A doctor says “the patient has disease a or b with certainty 0.7.”

The doctor gives no information about disease c .

- $m : \mathcal{P}(S) \rightarrow [0, 1]$ is computed based on the evidence
- $\text{bel}(A) = \sum_{B \leq A} m(B)$: the evidence supporting a
- $\text{pl}(A) = 1 - \text{bel}(\neg A) = \sum_{B \cap A \neq \emptyset} m(B)$: the evidence not contradicting A
- $\text{bel}(A) \leq \text{pl}(A)$.

Representation of evidence. An example

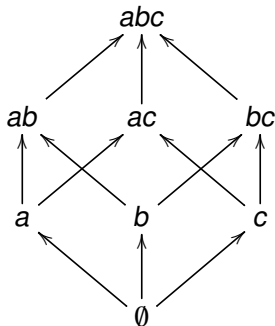
Scenario. A patient has disease a , b or c .

A doctor says “the patient has disease a or b with certainty 0.7.”

It is assumed it is impossible for the patient to have two of them.

Representation

- $S = \{a, b, c\}$ and $m, \text{bel}, \text{pl} : \mathcal{P}(S) \rightarrow [0, 1]$
- $m(\{a, b\}) = 0.7$ and $m(S) = 0.3$.



An example

Scenario. A patient has disease a , b or c .

A doctor says “the patient has disease a or b with certainty 0.7.”

The doctor gives no information about disease c .

Representation

- $S = \{a, b, c\}$ and $m, \text{bel}, \text{pl} : \mathcal{P}(S) \rightarrow [0, 1]$
- $m(\{a, b\}) = 0.7$ and $m(S) = 0.3$.

We get:

$$\text{bel}(\{a\}) = \text{bel}(\{b\}) = \text{bel}(\{c\}) = 0$$

$$\text{bel}(\{a, b\}) = \sum_{X \subseteq \{a, b\}} m(X) = 0.7 \quad \text{pl}(\{a, b\}) = 1 - \text{bel}(\{c\}) = 1$$

$$\text{pl}(\{a\}) = \text{pl}(\{b\}) = 1$$

$$\text{pl}(\{c\}) = 1 - \text{bel}(\{a, b\}) = 0.3$$

- $m(\{a, b\})$: the ‘probability’ that the disease is in the set $\{a, b\}$ without being able to say to which subset it belongs.

Dempster-Shafer combination rule

Let m_1 and m_2 be two mass functions on $\mathcal{P}(S)$. Dempster-Shafer combination rule computes their aggregation

$m_{1\oplus 2} : \mathcal{P}(S) \rightarrow [0, 1]$ as follows.

$$m_{1\oplus 2}(X) \mapsto \begin{cases} 0 & \text{if } X = \emptyset \\ \frac{\sum \{m_1(X_1) \cdot m_2(X_2) \mid X_1 \cap X_2 = X\}}{N} & \text{otherwise.} \end{cases}$$

Normalization factor:

$$\begin{aligned} N &= \sum \{m_1(X_1) \cdot m_2(X_2) \mid X_1 \cap X_2 \neq \emptyset\} \\ &= 1 - \sum \{m_1(X_1) \cdot m_2(X_2) \mid X_1 \cap X_2 = \emptyset\} \end{aligned}$$

What happens with contradictory evidence?

Scenario

A patient has disease a , b or c .

Doctor 1:

“the patient has a with certainty 0.9 and b with certainty 0.1.”

Doctor 2:

“the patient has c with certainty 0.9 and b with certainty 0.1.”

Representation

$S = \{a, b, c\}$

$m_1(\{a\}) = 0.9$ and $m_1(\{b\}) = 0.1$.

$m_2(\{c\}) = 0.9$ and $m_2(\{b\}) = 0.1$.

DS combination rule ignores contradictory information

$$m_{1\oplus 2}(\{b\}) = 1, m_{1\oplus 2}(\{a\}) = m_{1\oplus 2}(\{c\}) = 0$$

because $\{a\} \cap \{b\} = \{a\} \cap \{c\} = \emptyset$.

- 1 Representing incomplete/contradictory probabilistic information
 - Belnap-Dunn Logic
 - Non-standard probabilities
- 2 Dempster-Shafer theory
 - Mass functions, belief functions and plausibility functions
 - Representation of evidence
- 3 Dempster-Shafer theory and BD logic
 - Belief functions
 - Plausibility?
 - Combination of evidence

Extending belief with a (classical) belief function

Definition belief on BD-models

$M = \langle W, v^+, v^-, Bel \rangle$, a BD model plus $Bel : \mathcal{P}(W) \rightarrow [0, 1]$.

$$\text{bel}^+(\phi) := Bel(|\phi|^+) \quad \text{and} \quad \text{bel}^-(\phi) := Bel(|\phi|^-)$$

bel^+ : belief function on the associated Lindenbaum algebra \mathcal{L}_{BD} .

bel^- : belief function on $\mathcal{L}_{\text{BD}}^{\text{op}}$.

Remark. if \perp and \top are not in the language bel^+ (resp. pl^+) are general belief (resp. plausibility) functions.

Non-standard probabilities

Models: $(W, v^+, v^-, m : W \rightarrow [0, 1])$

$$p^+(\phi) = \sum_{s \in |\phi|^+} m(s) \text{ and } p^-(\phi) = \sum_{s \in |\phi|^-} m(s)$$

Immediate generalisation for belief.

Non-standard beliefs

Models: $(W, v^+, v^-, m : \mathcal{P}(W) \rightarrow [0, 1])$

$$\text{bel}^+(\phi) = \sum_{X \subseteq |\phi|^+} m(X) \text{ and } \text{bel}^-(\phi) = \sum_{X \subseteq |\phi|^-} m(X)$$

- $\text{bel}^+(\phi)$: belief that ϕ is true
- $\text{bel}^-(\phi)$: belief that ϕ is false
- bel^+ satisfies the axioms of belief functions

Two-dimensional interpretation

Two dimensional reading allows for various combinations of positive/negative belief/plausibility:

Two-dimensional interpretation

Two dimensional reading allows for various combinations of positive/negative belief/plausibility:

- classically $\text{bel}(\phi) = 1 - \text{pl}(\neg\phi)$.
 $\text{pl}(\neg\phi)$ maximum evidence against ϕ we can consider
in two dimensional reading it corresponds to $\text{pl}^-(\phi)$

$$(\text{bel}^+(\phi), \text{pl}^-(\phi))$$

Two-dimensional interpretation

Two dimensional reading allows for various combinations of positive/negative belief/plausibility:

- classically $\text{bel}(\phi) = 1 - \text{pl}(\neg\phi)$.
 $\text{pl}(\neg\phi)$ maximum evidence against ϕ we can consider
in two dimensional reading it corresponds to $\text{pl}^-(\phi)$

$$(\text{bel}^+(\phi), \text{pl}^-(\phi))$$

- consider both belief and plausibility independently:

$$(\text{bel}^+(\phi), \text{bel}^-(\phi)), (\text{pl}^+(\phi), \text{pl}^-(\phi))$$

Two-dimensional interpretation

Two dimensional reading allows for various combinations of positive/negative belief/plausibility:

- classically $\text{bel}(\phi) = 1 - \text{pl}(\neg\phi)$.
 $\text{pl}(\neg\phi)$ maximum evidence against ϕ we can consider in two dimensional reading it corresponds to $\text{pl}^-(\phi)$

$$(\text{bel}^+(\phi), \text{pl}^-(\phi))$$

- consider both belief and plausibility independently:

$$(\text{bel}^+(\phi), \text{bel}^-(\phi)), (\text{pl}^+(\phi), \text{pl}^-(\phi))$$

- If we require $\text{bel}(X) \leq \text{pl}(X)$, for $X \in \mathcal{P}(W)$, then bel and pl must be defined from different mass functions. a piece of evidence might support belief and plausibility in a different way \rightarrow gives rise to two mass functions (e.g., circumstantial evidence vs. direct evidence)

Combination of evidence

Let \mathcal{L} be a finite distributive lattice.

Without \perp and \top

$$m_{1\oplus 2} : \mathcal{L} \rightarrow [0, 1]$$
$$x \mapsto \sum \{m_1(x_1) \cdot m_2(x_2) \mid x_1 \wedge x_2 = x\}.$$

With \perp and \top

$$m_{1\oplus 2} : \mathcal{L} \rightarrow [0, 1]$$
$$x \mapsto \begin{cases} 0 & \text{if } x = \perp \\ \frac{\sum \{m_1(x_1) \cdot m_2(x_2) \mid x_1 \wedge x_2 = x\}}{\sum \{m_1(x_1) \cdot m_2(x_2) \mid x_1 \wedge x_2 \neq \perp\}} & \text{otherwise.} \end{cases}$$

Examples. The two doctors

Scenario. A patient has disease a , b or c .

Doctor 1: a with certainty 0.9 and b with certainty 0.1.

Doctor 2: c with certainty 0.9 and b with certainty 0.1.

Representation. $m_1, m_2 : \mathcal{DM}_3 \rightarrow [0, 1]$

$$m_1(x) = \begin{cases} 0.9 & \text{if } x = a \\ 0.1 & \text{if } x = b \\ 0 & \text{otherwise.} \end{cases} \quad m_2(x) = \begin{cases} 0.9 & \text{if } x = c \\ 0.1 & \text{if } x = b \\ 0 & \text{otherwise.} \end{cases}$$

Dempster-Shafer combination rule gives

$$m_{1 \oplus 2}(x) = \begin{cases} 0.81 & \text{if } x = a \wedge c \\ 0.09 & \text{if } x = a \wedge b \text{ or } x = b \wedge c \\ 0.01 & \text{if } x = b \\ 0 & \text{otherwise.} \end{cases}$$

$$\text{bel}_{1 \oplus 2}(a) = \text{bel}_{1 \oplus 2}(c) = 0.9 \text{ and } \text{bel}_{1 \oplus 2}(b) = 0.19$$

Examples. The two doctors

Representation. $m_1, m_2 : \mathcal{DM}_3 \rightarrow [0, 1]$

$$m_1(x) = \begin{cases} 0.9 & \text{if } x = a \wedge \neg b \wedge \neg c \\ 0.1 & \text{if } x = \neg a \wedge b \wedge \neg c \\ 0 & \text{otherwise.} \end{cases}$$

$$m_2(x) = \begin{cases} 0.9 & \text{if } x = \neg a \wedge \neg b \wedge c \\ 0.1 & \text{if } x = \neg a \wedge b \wedge \neg c \\ 0 & \text{otherwise.} \end{cases}$$

Dempster-Shafer combination rule gives

$$m_{1 \oplus 2}(x) = \begin{cases} 0.81 & \text{if } x = a \wedge \neg a \wedge \neg b \wedge c \wedge \neg c \\ 0.09 & \text{if } x = a \wedge \neg a \wedge b \wedge \neg b \wedge \neg c \\ & \text{or } x = \neg a \wedge b \wedge \neg b \wedge c \wedge \neg c \\ 0.01 & \text{if } x = \neg a \wedge b \wedge \neg c \\ 0 & \text{otherwise.} \end{cases}$$

$$\text{bel}_{1 \oplus 2}(a) = \text{bel}_{1 \oplus 2}(c) = 0.9 \text{ and } \text{bel}_{1 \oplus 2}(b) = 0.19$$

Reasoning with inconsistent / incomplete uncertain information

- other uncertainty measures (upper/lower probabilities, ...)
- qualitative probability
- various aggregation methods
- two layered framework

References

- [Dunn 76] Dunn, *Intuitive semantics for first-degree entailments and 'coupled trees'*. *Philosophical Studies* 29(3), 149–168, 1976
- [Belnap 19] Belnap, *How a computer should think*, *New Essays on Belnap-Dunn Logic*, 2019.
- [Halpern 17] J.Y. Halpern. *Reasoning about uncertainty*. The MIT Press, 2nd edition, 2017.
- [Klein et al] D. Klein, O. Majer, and S. Rafiee Rad. *Probabilities with gaps and gluts*. *Journal of Philosophical Logic*, 50(5):1107–1141, October 2021
- [Shafer 76] G. Shafer. *A mathematical theory of evidence*. Princeton university press, 1976.
- [Zhou 13] C. Zhou. *Belief functions on distributive lattices*. *Artificial Intelligence*, 201:1–31, 2013.