Belief function theory 101

Sébastien Destercke

Heudiasyc, CNRS Compiegne, France

ISIPTA 2018 School

Lecture goal/content

What you will find in this talk

- An overview of belief functions and how to obtain them
- Short discussion on comparing informative contents
- Discussion about conditioning and fusion
- Pointers to additional topics (statistical learning, preference handling, ...)

What you will not find in this talk

- A deep and exhaustive study of a particular topic

How this will go

- Exercices along the lecture
- You are encouraged to ask questions during the lecture!

Plan

Generic vs singular quantity

A quantity of interest S can be

- **Generic**, when it refers to a population, or a set of situations.

Generic quantity example

The distribution of height within french population

- **Singular**, when it refers to an individual or a peculiar situation

Singular quantity example

My own, personal height

Ontic and epistemic information [10]

An item of information $\mathcal{I}$ possessed by an agent about S can be

- **Ontic**, if it is a faithful, perfect representation of S

Ontic information example

A set S representing the exact set of languages spoken by me
e.g.: $S = \{French, English, Spanish\}$

- **Epistemic**, if it is an imperfect representation of S

Epistemic information example

A set E containing my mother tongue
e.g., $E = \{French, English, Spanish\}$

- $\rightarrow$ same mathematical expression, different interpretation

Everything is possible

We can have

- **Ontic** information about a **singular** quantity: the hair colour of a suspect; the mother tongue of someone
- **Epistemic** information about a **singular** quantity: the result of the next dice toss; the set of possible mother tongues of someone
- **Ontic** information about a **generic** quantity: the exact distribution of pixel colours in an image
- **Epistemic** information about a **generic** quantity: an interval about the frequency of French persons higher than 1.80 m

Uncertainty definition

Uncertainty: when our information $\mathcal{I}$ does not characterize the quantity of interest S with certainty

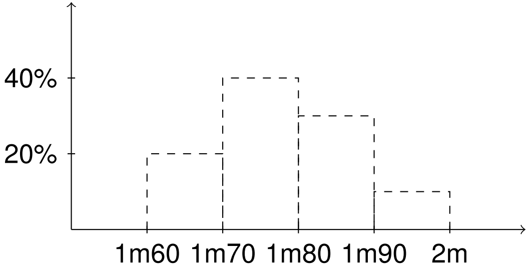
$\rightarrow$ **In this view, uncertainty is necessarily epistemic, as it reflect an imperfect knowledge of the agent**

Can concern both:

- Singular information
  - items in a data-base, values of some logical variables, time before failure of **a** component
- Generic information
  - parameter values of classifiers/regression models/probability distributions, time before failure of component**s**, truth of a logical sentence ("birds fly")

The room example

Heights of people in a room: generic quantity



- Generic question: are 90% of people in room less than 1m80?
  $\Rightarrow$ No, with **full certainty**
- Specific question: is the last person who entered less than 1m80?
  $\Rightarrow$ Probably, with 60% chance (**uncertain answer**)

Uncertainty main origins [6, Ch. 3]

- **Variability** of a population applied to a peculiar, singular situation

Variability example

The result of one dice throw when knowing the probability of each face

- **Imprecision and incompleteness** due to partial information about the quantity S

Imprecision example

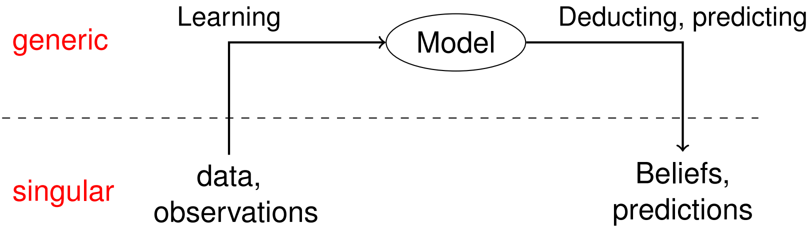
Observing limited sample of the population, describing suspect as "young", limited sensor precision

- **Conflict** between different sources of information (data/expert)

Conflict example

Two redundant data base entries with different information for an attribute, two sensors giving different measurements of a quantity

Handling uncertainty



Common problems in one sentence

- **Learning**: use singular information to estimate generic information (induction in logical sense)
- **Prediction**: interrogate model and observations to deduce information on quantity of interest ($\sim$ inference/deduction in logical sense)
- **Information revision**: merge new information with old one
- **Information fusion**: merge multiple information pieces about same quantity

Plan

1 Introductory elements

2 Belief function: basics, links and representation
- Less general than belief functions
- Belief functions
- More general than belief functions

3 Comparison, conditioning and fusion
- Information comparison
- The different facets of conditioning
- Information fusion
  - Basic operators
  - Rule choice:set/logical approach
  - Rule choice: performance approach

Section goals

- Remind basic ideas of uncertainty modelling
- Introduce main ideas about belief functions
- Provide elements linking belief functions and other approaches
- Illustrate practical representations of belief functions

Outline

Basic framework

Quantity S with possible **exclusive** states $\Omega = \{\omega_1, \dots, \omega_n\}$

▷ S: data feature, model parameter, ...

Basic tools

A confidence degree $P : 2^\Omega \to [0, 1]$ is such that

- $P(A)$: confidence $S \in A$
- $P(\emptyset) = 0$, $P(\Omega) = 1$
- $A \subseteq B \Rightarrow P(A) \leq P(B)$

Uncertainty modelled by 2 degrees $\underline{P}, \overline{P} : 2^\Omega \to [0, 1]$:

- $\underline{P}(A) \leq \overline{P}(A)$ (monotonicity)
- $\underline{P}(A) = 1 - \overline{P}(A^c)$ (duality)

Probability

Basic tool

A probability distribution $p : \Omega \rightarrow [0, 1]$ from which

- $\underline{P}(A) = \overline{P}(A) = P(A) = \sum_{s \in A} p(s)$
- $P(A) = 1 - P(A^c)$: auto-dual

Main interpretations

- **Frequentist [54] :** $P(A)$= number of times A observed in a population
  - ▷ only applies to generic quantities (populations)
- **Subjectivist [36] :** $P(A)$= price for gamble giving 1 if A happens, 0 if not
  - ▷ applies to both singular and generic quantities

Sets

Basic tool

A set $E \subseteq \Omega$ with true value $S \in E$ from which

- $E \subseteq A \rightarrow \underline{P}(A) = \overline{P}(A) = 1$ (certainty truth in A)
- $E \cap A \neq \emptyset, E \cap A^c \neq \emptyset \rightarrow \underline{P}(A) = 0, \overline{P}(A) = 1$ (ignorance)
- $E \cap A = \emptyset \rightarrow \underline{P}(A) = \overline{P}(A) = 0$ (truth cannot be in A)

$\underline{P}, \overline{P}$ are binary $\rightarrow$ limited expressiveness

Classical use of sets:

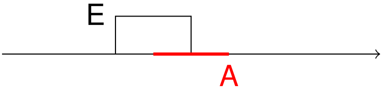
- Interval analysis [40] (E is a subset of $\mathbb{R}$)
- Propositional logic (E is the set of models of a KB)

Other cases: robust optimisation, decision under risk, ...

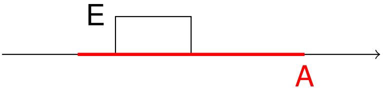
Example

Assume that it is known that pH value $E \in [4.5, 5.5]$, then

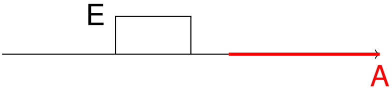
- if $A = [5, 6]$, then $\underline{P}(A) = 0, \overline{P}(A) = 1$



- if $A = [4, 7]$, then $\underline{P}(A) = \overline{P}(A) = 1$



- if $A = [6, 9]$, then $\underline{P}(A) = \overline{P}(A) = 0$



In summary

Probabilities . . .

- (+) very informative quantification (do we need it?)
- (-) need lots of information (do we have it?)
- (-) if not enough, requires a choice (do we want to do that?)
- use probabilistic calculus (convolution, stoch. independence, . . .)

Sets . . .

- (+) need very few information
- (-) very rough quantification of uncertainty (Is it sufficient for us?)
- use set calculus (interval analysis, Cartesian product, . . .)

$\rightarrow$ Need for **frameworks bridging these two**

Possibility theory [27]

Basic tool

A distribution $\pi : \Omega \rightarrow [0,1]$, usually with ω such that $\pi(\omega) = 1$, from which

- $\overline{P}(A) = \max_{\omega \in A} \pi(\omega)$ (Possibility measure)
- $\underline{P}(A) = 1 - \overline{P}(A^c) = \min_{\omega \in A^c}(1 - \pi(\omega))$ (Necessity measure)

Sets E captured by $\pi(\omega) = 1$ if $\omega \in E$, 0 otherwise

Interval/set as special case

The set E can be modelled by the possibility distribution π_E such that

$$\pi_E(\omega) = \begin{cases} 1 & \text{if } \omega \in E \\ 0 & \text{else} \end{cases}$$

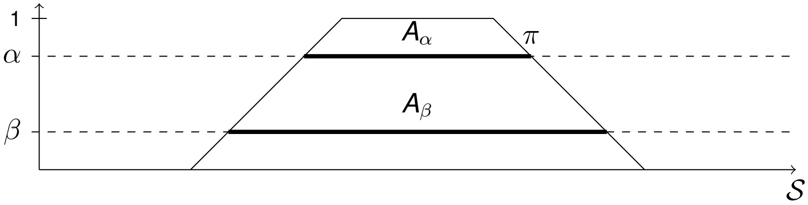
A nice characteristic: Alpha-cut [9]

Definition

$$A_\alpha = \{\omega \in \Omega | \pi(\omega) \geq \alpha\}$$

- $\underline{P}(A_\alpha) = 1 - \alpha$
- If $\beta \leq \alpha$, $A_\alpha \subseteq A_\beta$

Simulation: draw $\alpha \in [0, 1]$ and associate A_α



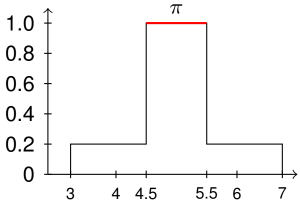
$\Rightarrow$ Possibilistic approach ideal to model **nested structures**

A basic distribution: simple support

- A set E of most plausible values
- A confidence degree $\alpha = \underline{P}(E)$
- Two interesting cases:
  - Expert providing most plausible values E
  - E set of models of a formula ϕ
- Both cases extend to multiple sets $E_1, \ldots, E_p$:
  - confidence degrees over nested sets [49]
  - hierarchical knowledge bases [29]

pH value $\in [4.5, 5.5]$ with

$\alpha = 0.8$ ($\sim$ "quite probable")



A basic distribution: simple support

A set E of most plausible values

A confidence degree $\alpha = \underline{P}(E)$

Two interesting cases:

- Expert providing most plausible values E
- E set of models of a formula ϕ

Both cases extend to multiple sets $E_1, \ldots, E_p$:

- confidence degrees over nested sets [49]
- hierarchical knowledge bases [29]

variables p, q

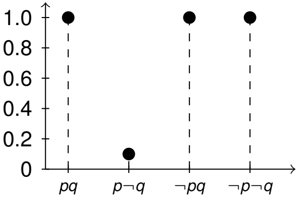
$$\Omega = \{pq, \neg pq, p\neg q, \neg p \neg q\}$$

$$\underline{P}(p \Rightarrow q) = 0.9$$

($\sim$ "almost certain")

$$E = \{pq, p\neg q, \neg p \neg q\}$$

- $\pi(pq) = \pi(p\neg q) = \pi(\neg p \neg q) = 1$
- $\pi(\neg pq) = 0.1$

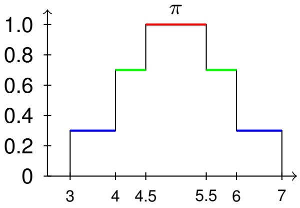


Nested confidence intervals: expert opinions

Expert providing nested intervals + conservative confidence degree

A *pH* degree

- $0.3 \leq P([4.5, 5.5])$
- $0.7 \leq P([4, 6])$
- $1 \leq P([3, 7])$

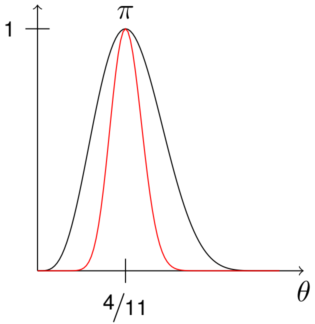


Normalized likelihood as possibilities [24] [7]

$$\pi(\theta) = {}^{\mathcal{L}(\theta|x)}/_{\max_{\theta \in \Theta} \mathcal{L}(\theta|x)}$$

Binomial situation:

- θ = success probability
- x number of observed successes
- x= 4 succ. out of 11
- x= 20 succ. out of 55



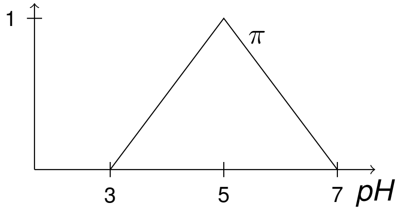
Partially specified probabilities [3] [23]

Triangular distribution: $[\underline{P}, \overline{P}]$ encompass all probabilities with

- mode/reference value M
- support domain $[a, b]$.

Getting back to pH

- $M = 5$
- $[a, b] = [3, 7]$



Other examples

- Statistical inequalities (e.g., Chebyshev inequality) [23]
- Linguistic information (fuzzy sets) [12]
- Approaches based on nested models

Possibility: limitations

$$\underline{P}(A) > 0 \Rightarrow \overline{P}(A) = 1$$
$$\overline{P}(A) < 1 \Rightarrow \underline{P}(A) = 0$$

$\Rightarrow$ interval $[\underline{P}(A), \overline{P}(A)]$ with one trivial bound

Does not include probabilities as special case:

$\Rightarrow$ possibility and probability at odds

$\Rightarrow$ respective calculus hard (sometimes impossible?) to reconcile

Going beyond

Extend the theory

$\Rightarrow$ by complementing π with a lower distribution δ ($\delta \leq \pi$) [30], [21]

$\Rightarrow$ by working with interval-valued possibility/necessity degrees [4]

$\Rightarrow$ by working with sets of possibility measures [32]

Use a more general model

$\Rightarrow$ **Random sets and belief functions**

Outline

Belief functions

The history

- First used by Dempster to make statistical reasoning about imprecise observations, mostly with frequentist view
- Popularized by Shafer as a generic way to handle imprecise evidences
- Used by Smets (in TBM) with a will to not refer at all to probabilities

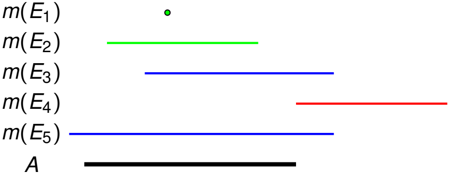
$\rightarrow$ evolved as a uncertainty theory of its own ($\exists \neq$ with IP, Possibility or p-boxes)

Random sets and belief functions

Basic tool

A positive distribution $m : 2^{\Omega} \to [0,1]$, with $\sum_E m(E) = 1$ and usually $m(\emptyset) = 0$, from which

- $\overline{P}(A) = \sum_{E \cap A \neq \emptyset} m(E)$ (Plausibility measure)
- $\underline{P}(A) = \sum_{E \subseteq A} m(E) = 1 - \overline{\mu}(A^c)$ (Belief measure)



$\underline{P}(A) = m(E_1) + m(E_2)$

$\overline{P}(A) = m(E_1) + m(E_2) + m(E_3) + m(E_5)$

$[\underline{P}, \overline{P}]$ as

- subjective confidence degrees of **evidence theory** [50], [51], [13]
- bounds of an **ill-known probability** measure $\mu \Rightarrow \underline{P} \leq \mu \leq \overline{P}$

A characterisation of belief functions

Complete monotonicity

If $\underline{P}$ is a belief measure if and only if it satisfies the inequality

$$\underline{P}(\cup_{i=1}^{n} A_i) \geq \sum_{\mathcal{A} \subseteq \{A_1,\ldots,A_n\}} (-1)^{|\mathcal{A}|+1} \underline{P}(\cap_{A_i \in \mathcal{A}} A_i)$$

for any number n.

Simply the exclusion/inclusion principle with an equality

Another characterisation of belief functions

Möbius inverse: definition

Let $\underline{P}$ be a measure on 2^Ω, its Möbius inverse $m_{\underline{P}} : 2^\Omega \to \mathbb{R}$ is

$$m_{\underline{P}}(E) = \sum_{A \subseteq E} -1^{|E \setminus A|} \underline{P}(E).$$

It is bijective, as $\underline{P}(A) = \sum_{E \subseteq A} m(E)$, and can be applied to any set-function.

Belief characterisation

$m_{\underline{P}}$ will be non-negative for all E **if and only if** $\underline{P}$ is a belief function.

Yet another characterisation: commonality functions

Definition

Given a mass function m, commonality function $Q : 2^{\Omega} \to [0,1]$ defined as

$$Q(A) = \sum_{E \supseteq A} m(E)$$

and express how unsurprising it is to see A happens.

Back to m

Given Q, we have

$$m(A) = \sum_{B \supseteq A} -1^{|B \setminus A|} Q(B)$$

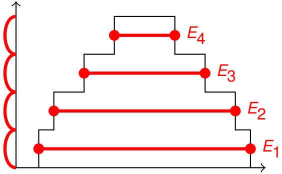
Some notes

- Instrumental to define "complement" of information m
- In possibility theory, equivalent to guaranteed possibility
- In imprecise probability, no equivalent (?)

special cases

Measures $[\underline{P}, \overline{P}]$ include:

- Probability distributions: mass on atoms/singletons
- Possibility distributions: mass on nested sets



$\rightarrow$ "simplest" theory that includes both sets and probabilities as special cases!

Frequencies of imprecise observations

Imprecise poll: "Who will win the next Wimbledon tournament?"

∘ N(adal) ∘ F(ederer) ∘ D(jokovic) ∘ M(urray) ∘ O(ther)

60 % replied $\{N, F, D\} \rightarrow m(\{N, F, D\}) = 0.6$

15 % replied "I do not know" $\{N, F, D, M, O\} \rightarrow m(\mathcal{S}) = 0.15$

10 % replied Murray $\{M\} \rightarrow m(\{M\}) = 0.1$

5 % replied others $\{O\} \rightarrow m(\{O\}) = 0.05$

...

P-box [35]

A pair $[\underline{F}, \overline{F}]$ of cumulative distributions

Bounds over events $[-\infty, x]$

- Percentiles by experts;
- Kolmogorov-Smirnov bounds;

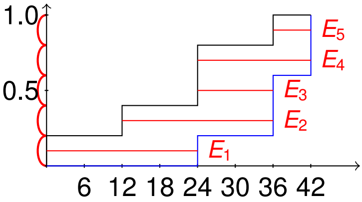
Can be extended to any pre-ordered space [20], [53] $\Rightarrow$ multivariate spaces!

Expert providing percentiles

$$0 \leq P([-\infty, 12]) \leq 0.2$$

$$0.2 \leq P([-\infty, 24]) \leq 0.4$$

$$0.6 \leq P([-\infty, 36]) \leq 0.8$$



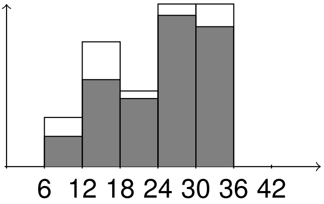
Other means to get random sets/belief functions

- Extending modal logic: probability of provability [52]
- Parameter estimation using pivotal quantities [43]
- Statistical confidence regions [14]
- Modify source information by its reliability [47]
- ...

Outline

Limits of random sets

- Not yet **fully** satisfactory extension of Bayesian/subjective approach
- Still some natural items of information it cannot easily model:
  - probabilistic bounds over atoms ω (imprecise histograms, ...) [11] ;
  - comparative assessments such as $2P(B) \leq P(A)$ [45], ...



Imprecise probabilities

Basic tool

A set $\mathcal{P}$ of probabilities on Ω or an equivalent representation

- $\overline{P}(A) = \sup_{P \in \mathcal{P}} P(A)$ (Upper probability)
- $\underline{P}(A) = \inf_{P \in \mathcal{P}} P(A) = 1 - \overline{P}(A^c)$ (Lower probability)

Reminder: lower/upper bounds on events alone cannot model any convex $\mathcal{P}$

$[\underline{P}, \overline{P}]$ as

- subjective lower and upper betting rates [55]
- bounds of an **ill-known probability measure** $P \Rightarrow \underline{P} \leq P \leq \overline{P}$ [5] [56]

Some basic properties

Avoiding sure loss and coherence

Given some bounds $\underline{P}(A)$ over every event $A \subseteq \Omega$, we say that

- $\underline{P}$ avoids sure loss iff

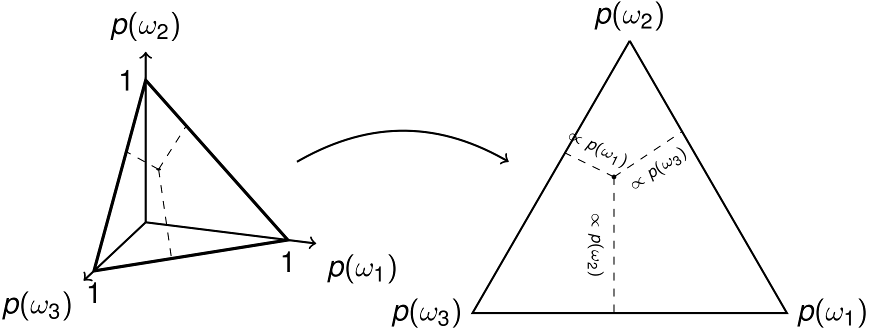
$$\mathcal{P}(\underline{P}) = \{P : \underline{P} \leq P \leq \overline{P}\} \neq \emptyset$$

- $\underline{P}$ is coherent iff for any A, we have

$$\inf_{P \in \mathcal{P}(\underline{P})} P(A) = \underline{P}(A)$$

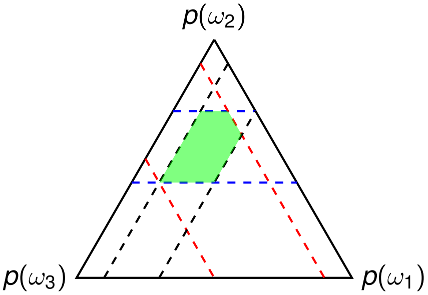
Illustrative example

$$p(\omega_1) = 0.2,\ p(\omega_2) = 0.5,\ p(\omega_3) = 0.3$$



A first exercise

$p(\omega_1) \in [0.1, 0.3]$, $p(\omega_2) \in [0.4, 0.7]$, $p(\omega_3) = [0.1, 0.5]$

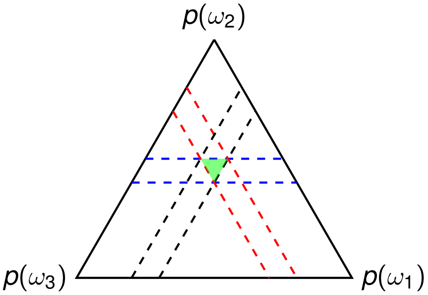


→ Show that these induce a belief function

	$\{\omega_1\}$	$\{\omega_2\}$	$\{\omega_3\}$	$\{\omega_1, \omega_2\}$	$\{\omega_1, \omega_3\}$	$\{\omega_2, \omega_3\}$
$\underline{P}$						

A second exercise

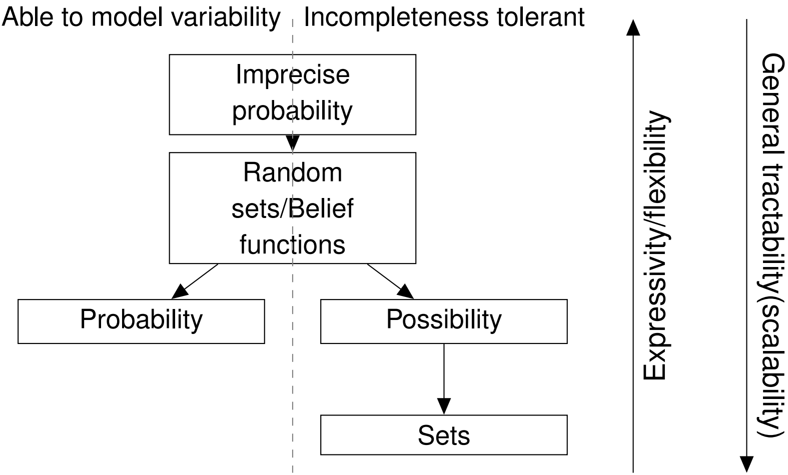
$p(\omega_1) \in [0.2, 0.3]$, $p(\omega_2) \in [0.4, 0.5]$, $p(\omega_3) = [0.2, 0.3]$



$\rightarrow$ Show that these do not induce a belief function

	$\{\omega_1\}$	$\{\omega_2\}$	$\{\omega_3\}$	$\{\omega_1, \omega_2\}$	$\{\omega_1, \omega_3\}$	$\{\omega_2, \omega_3\}$
$\underline{P}$						

A not completely accurate but useful picture



Why belief functions?

Why not?

- You need more (to model properly/not approximate your results)
- You cannot afford it (computationally)

Why?

They offer a fair compromise

- Embed precise probabilities and sets in one frame
- Can use simulation of m + Set computation
- Extreme points/natural extension easy to compute (Choquet Integral, . . .)

Or, you want to use tools proper to BF theory.

Plan

Outline

Introduction

Main question

Given two pieces of information $\underline{P}_1, \underline{P}_2$, is one more informative than the others? How can we answer?

Examples of use

- **Least commitment principle**: given multiple models satisfying given constraints, pick the most conservative one
  - Partial elicitation,
  - Revision,
  - Inverse Pignistic,
  - Natural extension, ...
- **(Outer)-approximation**: Pick a model $\underline{P}_2$ simpler than $\underline{P}_1$ (e.g., generic belief mass into possibility), ensuring that $\underline{P}_2$ does not add information to $\underline{P}_1$.

A natural notion: set inclusion

A set $A \subseteq \mathcal{S}$ is **more informative** than $B \subseteq \Omega$ if

$$A \subseteq B \Leftrightarrow A \sqsubseteq B$$

- Propositional logic: A more informative if A entails B
- Intervals: A includes all values of B, is more precise than B

$\Rightarrow$ extends this notion to other uncertainty theories

Extensions to other models

Denoting $\underline{P}_A, \underline{P}_B$ the uncertainty models of sets A, B, we do have

$$A \sqsubseteq B \Leftrightarrow \underline{P}_A(C) \leq \underline{P}_B(C) \text{ for any } C \subseteq \mathcal{S}$$

Derivations of $\underline{P}_1 \leq \underline{P}_2$ in different frameworks

- Possibility distributions: $\pi_1 \sqsubseteq \pi_2 \Leftrightarrow \pi_1 \geq \pi_2$
- Belief functions: $m_1 \sqsubseteq m_2 \Leftrightarrow \underline{P}_1 \sqsubseteq \underline{P}_2$ (plausibility inclusion, there are others [25])
- Probability sets: $\underline{P}_1 \sqsubseteq \underline{P}_2 \Leftrightarrow \mathcal{P}_1 \subseteq \mathcal{P}_2$ ($\underline{P}_i$ lower previsions)

Inclusion: interest and limitations

- +: very natural way to compare informative content
- -: only induces a partial order between information models

Example

Consider the space $\Omega = \{a, b, c\}$ and the following mass functions:

$$m_1(\{b\}) = 0.3, m_1(\{b, c\}) = 0.2, m_1(\{a, b, c\}) = 0.5$$

$$m_2(\{a\}) = 0.2, m_2(\{b\}) = 0.3, m_2(\{c\}) = 0.3, m_2(\{a, b, c\}) = 0.2$$

$$m_3(\{a, b\}) = 0.3, m_3(\{a, c\}) = 0.3, m_3(\{a\}) = 0.4$$

We have $m_2 \sqsubseteq m_1$, but m_3 incomparable with $\sqsubseteq$ **(side-exercise: show it)**

$\Rightarrow$ ok theoretically, but not always lead to non-uniqueness of solutions

Numerical assessment of informative content [57, 1, 26]

- For probabilities, distinct μ_1 and μ_2 always incomparable by previous definition
- A solution, associate to each μ a number $I(\mu)$, i.e., entropy

$$I(\mu) = -\sum_{\omega \in \Omega} p(\omega) ln(p(\omega))$$

and declare that $\mu_1 \sqsubseteq \mu_2$ if $I(\mu_1) \leq I(\mu_2)$.

- This can be extended to other theories, where we can ask

$$\underline{P}_1 \leq \underline{P}_2 \Rightarrow I(\underline{P}_1) \geq I(\underline{P}_2)$$

Measure I should be consistent with inclusion

Outline

Three use of conditional and conditioning [39, 41]

Focusing: from generic to singular

- P: generic knowledge (usually about population)
- $P(|C)$: what we know from P in the singular context C

Revising: staying either generic or singular

- P: knowledge or belief (generic or singular)
- $P(|C)$: we learn that C is certainly true $\rightarrow$ how should we modify our knowledge/belief

Learning: from singular to generic (not developed here)

- P: beliefs about the parameter
- $P(|C)$: modified beliefs once we observe C ($\simeq$ multiple singular observations)

Focusing and revising in probabilities [28]

In probability, upon learning C, the revised/focused knowledge is

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{P(A \cap C)}{P(A \cap C) + P(A^c \cap C)}$$

coming down to the use of Bayes rule of conditioning in both cases.

Focusing

- Observing C does not modify our generic knowledge/beliefs
- We may lose information $\rightarrow$ the more C is specific, the less our general knowledge applies to it (cf. dilation in IP)
- The consistency of generic knowledge/beliefs should be preserved (C cannot contradict it, only specify to which case it should apply)
- If we observe later $A \subseteq C$, we should start over from generic knowledge

Focusing in uncertainty theories [34]

Focusing with belief functions

- Given initial belief function $\underline{P}$, this gives

$$\underline{P}(A||C) = \frac{\underline{P}(A \cap C)}{\underline{P}(A \cap C) + \overline{P}(A^c \cap C)}$$

$$\overline{P}(A||C) = \frac{\overline{P}(A \cap C)}{\overline{P}(A \cap C) + \underline{P}(A^c \cap C)}$$

We can have $\underline{P}(A||C) < \underline{P}(A) \leq \overline{P}(A) < \overline{P}(A||C)$ ("loss" of information).

- Can be interpreted as a sensitivity analysis of Bayes rule:

$$\underline{P}(A||C) = \inf\{P(A|C) : P \in \mathcal{P}, P(C) > 0\}$$

- $\simeq$ regular extension in imprecise probability

Revision

- Observing C modifies our knowledge and belief
- Observing C refines our beliefs and knowledge, that should become more precise
- If we observe later $A \subseteq C$, we should start from the modified knowledge (we may ask for operation to be order-insensitive)
- C is a new knowledge, that may be partially inconsistent with current belief/knowledge

Revision in uncertainty theories

Revising with belief functions

- Given initial plausibility function $\overline{P}$, this gives

$$\overline{P}(A|C) = \frac{\overline{P}(A \cap C)}{\overline{P}(C)} \Rightarrow \underline{P}(A|C) = 1 - \overline{P}(A^c|C)$$

- If $\overline{P}(C) = 1$, then
  - no conflict between old and new information (no **incoherence**)
  - we necessarily have $\overline{P}(A|C) < \overline{P}(A)$ (refined information)
- Can be interpreted Bayes rule applied to most plausible situations:

$$\underline{P}(A||C) = \inf\{P(A|C) : P \in \mathcal{P}, P(C) = \overline{P}(C)\}$$

- Similarly to fusion, not studied a lot within IP setting (because of incoherence?)

Revision as prioritized fusion

When $\overline{P}(C) = 1$ and C precise observation

- $\overline{P}(A|C)$= result of conjunctive combination rule
- $\mathcal{P}_{|C} = \mathcal{P} \cap \{P : P(C) = 1\}$

$\rightarrow$ can be interpreted as a fusion rule where C has priority. If $\overline{P}(C) < 1$, interpreted as new information inconsistent with the old $\rightarrow$ conditioning as a way to restore consistency.

Case where observation C is uncertain and inconsistent with knowledge.

- Minimally change $\underline{\mu}$ to be consistent with $C \rightarrow$ in probability, Jeffrey's rule (extensions to other theories exist [42])
- Not a symmetric fusion process, new information usually has priority ($\neq$ from usual belief fusion rules)!

A small exercice: focusing

The hotel provides the following plates for breakfast

a=Century egg, b=Rice, c=Croissant, d=Raisin Muffin

In a survey about their choices, respondents gave the reply

$$m(\{a,b\}) = \alpha,\ m(\{c,d\}) = 1-\alpha$$

Applying focusing

We learn that customer C does not like eggs nor raisins ($C = \{b, c\}$), what can we tell about him choosing Rice?

A small exercice: revision

The hotel provides the following plates for breakfast

a=Century egg, b=Rice, c=Croissant, d=Raisin Muffin

In a survey about their choices, respondent gave the reply

$$m(\{a, b\}) = \alpha, \; m(\{c, d\}) = 1 - \alpha$$

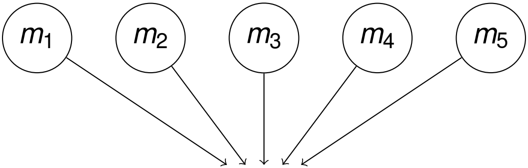
Applying revision

We learn that suppliers no longer have eggs nor raisins ($C = \{b, c\}$), what is the proportion of rice we should buy to satisfy customers?

Outline

An illustration of the issue

Information fusion



$$m^* = h(m_1, m_2, m_3, m_4, m_5)$$

- Information on the same level
- No piece of information has priority over the other (a priori)
- Makes sense to combine multiple pieces of information at once
- Main question: "How to choose h . . . "
  - To obtain a more reliable and informative result?
  - When items m_i's disagree?

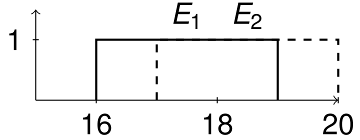
Conjunction

Main Assumption

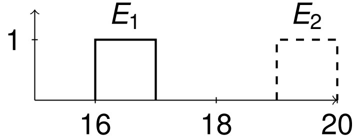
- Information items $E_1, \ldots, E_n$ are **all** fully reliable
- If one source consider ω impossible, then ω impossible

$$\rightarrow h(E_1, \ldots, E_n)(\omega) = \min(E_1(\omega), \ldots, E_n(\omega)) = \bigcap E_i$$

$E_1 = [16, 19]$ and $E_2 = [17, 20]$



$E_1 = [16, 17]$ and $E_2 = [19, 20]$



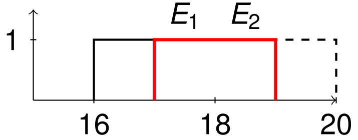
Conjunction

Main Assumption

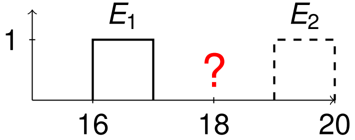
- Information items $E_1, \ldots, E_n$ are **all** fully reliable
- If one source consider ω impossible, then ω impossible

$$\rightarrow h(E_1, \ldots, E_n)(\omega) = \min(E_1(\omega), \ldots, E_n(\omega)) = \bigcap E_i$$

$E_1 = [16, 19]$ and $E_2 = [17, 20]$



$E_1 = [16, 17]$ and $E_2 = [19, 20]$



Pros and Cons

- +: very informative results, logically interpretable
- -: cannot deal with conflicting/unreliable information

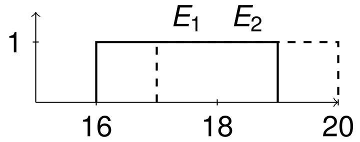
Disjunctive principle

Main Assumption

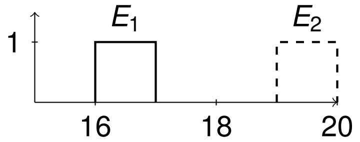
- **At least one** information item among $E_1, \dots, E_n$ is reliable
- ω possible as soon as one source considers it possible

$$\rightarrow h(E_1, \dots, E_n)(\omega) = \max(E_1(\omega), \dots, E_n(\omega)) = \bigcup E_i$$

$E_1 = [16, 19]$ and $E_2 = [17, 20]$



$E_1 = [16, 17]$ and $E_2 = [19, 20]$



Disjunctive principle

Main Assumption

- **At least one** information item among $E_1, \ldots, E_n$ is reliable
- ω possible as soon as one source considers it possible

$$\rightarrow h(E_1, \ldots, E_n)(\omega) = \max(E_1(\omega), \ldots, E_n(\omega)) = \bigcup E_i$$

$E_1 = [16, 19]$ and $E_2 = [17, 20]$ $\qquad$ $E_1 = [16, 17]$ and $E_2 = [19, 20]$

Pros and Cons

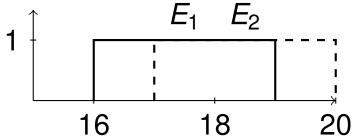
- +: no conflict, logically interpretable
- -: poorly informative results

Average

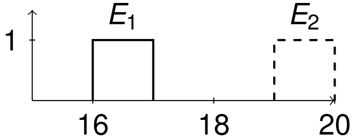
Main Assumption

Sources are statistically independent and in majority reliable

$E_1 = [16, 19]$ and $E_2 = [17, 20]$



$E_1 = [16, 17]$ and $E_2 = [19, 20]$

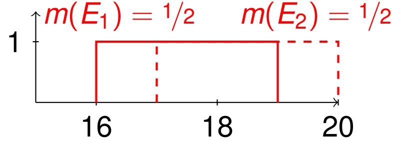


Average

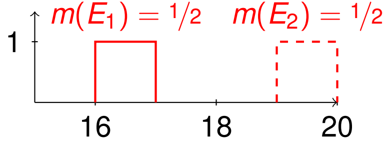
Main Assumption

Sources are statistically independent and in majority reliable

$E_1 = [16, 19]$ and $E_2 = [17, 20]$



$E_1 = [16, 17]$ and $E_2 = [19, 20]$



Pros and Cons

- +: result not conflicting, counting process (statistics)
- -: no logical interpretation, not applicable to sets

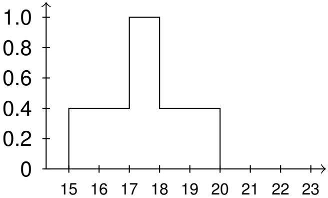
Limits of sets in information fusion

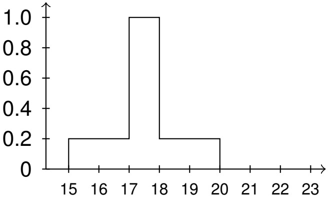
- Very basic information (what is possible/what is impossible)
- Very basic (binary) evaluation of conflict, either:
  - present if $\bigcap E_i = \emptyset$
  - absent if $\bigcap E_i \neq \emptyset$
- Limited number of fusion operators (only logical combinations)
- Limited operation on information items to integrate reliability scores, source importance, ...

$\rightarrow$ how to extend fusion operators to belief functions

Extending conjunction

Consider the two following information





$$m_1([17,18]) = 0.6$$

$$m_1([15,20]) = 0.4$$

Cautious source

$$m_2([20.5,21.5]) = 0.8$$

$$m_2([19.5,22.5]) = 0.2$$

Bold source

Extending conjunction: steps

		m_1	
		$[17, 18] = 0.6$	$[15, 20] = 0.4$
m_2	$[20.5, 21.5] = 0.8$		
	$[19.5, 22.5] = 0.2$		

Extending conjunction: steps

		m_1	
		$[17, 18] = 0.6$	$[15, 20] = 0.4$
m_2	$[20.5, 21.5] = 0.8$	$\emptyset$	$\emptyset$
	$[19.5, 22.5] = 0.2$	$\emptyset$	$[19.5, 20]$

- Step 1: take intersection (sources reliable)

Extending conjunction: steps

	m_1: $[17, 18] = 0.6$	m_1: $[15, 20] = 0.4$
m_2: $[20.5, 21.5] = 0.8$	$\emptyset$ 0.48	$\emptyset$ 0.24
m_2: $[19.5, 22.5] = 0.2$	$\emptyset$ 0.12	$[19.5, 20]$ 0.08

- Step 1: take intersection (sources reliable)
- Step 2: give product of masses (sources independent)

Extending conjunction: steps

		m_1	
		$[17,18]=0.6$	$[15,20]=0.4$
m_2	$[20.5,21.5]=0.8$	$\emptyset$ 0.48	$\emptyset$ 0.24
	$[19.5,22.5]=0.2$	$\emptyset$ 0.12	$[19.5,20]$ 0.08

- Step 1: take intersection (sources reliable)
- Step 2: give product of masses (sources independent)

$$m(\emptyset)=0.92 \rightarrow \text{high conflict evaluation, unsatisfying}$$

Extending conjunction

		m_1	
		$[17, 18] = 0.6$	$[15, 20] = 0.4$
m_2	$[17.5, 18.5] = 0.8$	$[17.5, 18]$ 0.48	$[17.5, 18.5]$ 0.24
	$[16.5, 19.5] = 0.2$	$[17, 18]$ 0.12	$[16.5, 19.5]$ 0.08

- Step 1: take intersection (sources reliable)
- Step 2: give product of masses (sources independent)

$$m(\emptyset) = 0 \rightarrow \text{no conflict, sources consistent}$$

Extending disjunction: steps

		m_1	
		$[17,18] = 0.6$	$[15,20] = 0.4$
m_2	$[20.5,21.5] = 0.8$		
	$[19.5,22.5] = 0.2$		

Extending disjunction: steps

		m_1	
		$[17,18] = 0.6$	$[15,20] = 0.4$
m_2	$[20.5,21.5] = 0.8$	$[17,18] \cup [20.5,21.5]$	$[15,20] \cup [20.5,21.5]$
	$[19.5,22.5] = 0.2$	$[17,18] \cup [19.5,22.5]$	$[15,22.5]$

- Step 1: take union (at least one reliable source)

Extending disjunction: steps

		m_1	
		$[17, 18] = 0.6$	$[15, 20] = 0.4$
m_2	$[20.5, 21.5] = 0.8$	$[17, 18] \cup [20.5, 21.5]$ 0.48	$[15, 20] \cup [20.5, 21.5]$ 0.24
	$[19.5, 22.5] = 0.2$	$[17, 18] \cup [19.5, 22.5]$ 0.12	$[15, 22.5]$ 0.08

- Step 1: take union (at least one reliable source)
- Step 2: give product of masses (sources independent)

Extending disjunction: steps

		m_1	
		$[17, 18] = 0.6$	$[15, 20] = 0.4$
m_2	$[20.5, 21.5] = 0.8$	$[17, 18] \cup [20.5, 21.5]$ 0.48	$[15, 20] \cup [20.5, 21.5]$ 0.24
	$[19.5, 22.5] = 0.2$	$[17, 18] \cup [19.5, 22.5]$ 0.12	$[15, 22.5]$ 0.08

- Step 1: take union (at least one reliable source)
- Step 2: give product of masses (sources independent)

$m(\emptyset) = 0 \rightarrow$ no conflict, but very imprecise result

More formally

Given informations $m_1, \ldots, m_n$

Conjunctive (Dempster's unnormalized) rule

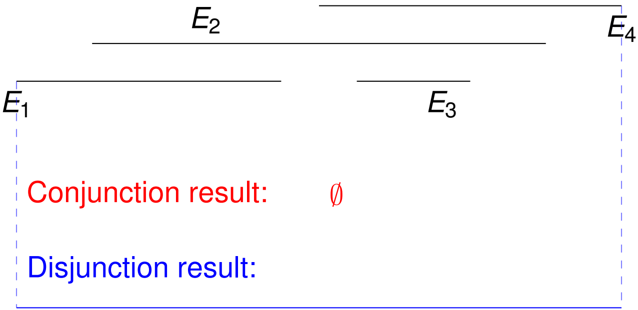
$$m_{\cap}(A) = \sum_{E_1 \cap \ldots \cap E_n = A} \prod_{i=1}^{n} m(E_i)$$

$\rightarrow$ a gradual way to estimate conflict [22]

Disjunctive rule

$$m_{\cup}(A) = \sum_{E_1 \cup \ldots \cup E_n = A} \prod_{i=1}^{n} m(E_i)$$

Conflict management: beyond conjunction and disjunction



$\Rightarrow$ Conjunction poorly reliable/false

$\Rightarrow$ Disjunction very imprecise and inconclusive

$\rightarrow$ A popular solution: choose a logical combination between the two

A simple idea [19]

- Get maximal subsets $M_1, \ldots, M_\ell$ of sources having non-empty intersection
- Take their intersection, then the union of those intersections

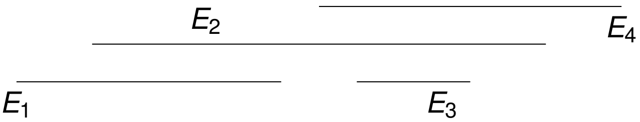
$$h(E_1, \ldots, E_n) = \cup_{M_\ell} \cap_{E_i \in M_\ell} E_i$$

An old idea ...

- In logic, to resolve knowledge base inconsistencies [31]
- In mathematical programming, to solve non-feasible problems [8]
- In interval analysis ...

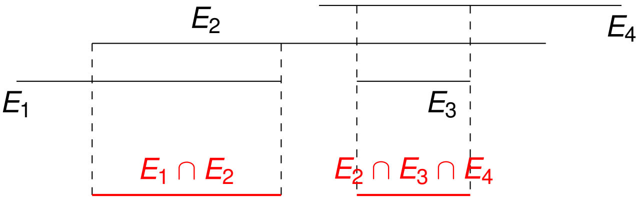
Illustrative exercice

Four sources provide you with basic items of information (sets)



- What are the maximal consistent subsets?
- What is the final result of applying the SMC rule to it?

Illustrative exercice:solution



SMC: $K_1 = \{E_1, E_2\}$ et $K_2 = \{E_2, E_3, E_4\}$

Final result: $(E_1 \cap E_2) \cup (E_2 \cap E_3 \cap E_4)$

- If all agree $\rightarrow$ conjunction
- if every pair is in disagreement (disjoint) $\rightarrow$ disjunction

MCS on belief: illustration

		m_1	
		$[17,18] = 0.6$	$[15,20] = 0.4$
m_2	$[20.5,21.5] = 0.8$	$[17,18] \cup [20.5,21.5]$ 0.48	$[15,20] \cup [20.5,21.5]$ 0.24
	$[19.5,22.5] = 0.2$	$[17,18] \cup [19.5,22.5]$ 0.12	$[15,20] \cap [19.5,22.5]$ 0.08

Set and logical view

Why?

- You want an interpretation to the combination
- You have relatively few information items
- You cannot "learn" your rule

Why not?

- You do not really care about interpretability
- You need to "scale up"
- You have means to learn your rule

Learning fusion rule: rough protocol

- A set of observed values $\hat{\omega}^1, \ldots, \hat{\omega}^o$
- for each $\hat{\omega}^i$, information $m_1^i, \ldots, m_n^i$ provided by n sources
- a decision rule $d : \mathcal{M} \to \Omega$ mapping m to a decision in Ω
- from set $\mathcal{H}$ of possible rules, choose

$$h^* = \arg\max_{h \in \mathcal{H}} \sum_i \mathbb{I}_{d(h(m_1^i, \ldots, m_n^i)) = \hat{\omega}^i}$$

How to choose $\mathcal{H}$?

- $\mathcal{H}$ should be easy to navigate, i.e., based on few parameters
- Maximization optimization problem should be made easy if possible (convex? Linear?)
- In particular, if m_j^i have peculiar forms (possibilities, Bayesian, ...), there is a better hope to find efficient methods

Two examples

- Weighted averaging rules (parameters to learn: weights)
- Denoeux T-(co)norm rules based on canonical decomposition (parameters to learn: parameters of the chosen t-norm family)

The case of averaging rule

- Parameters $\mathbf{w} = (w_1, \ldots, w_n)$ such that $\sum_i w_i = 1$ and $w_i > 0$
- Set $\mathcal{H} = \{h_\mathbf{w} | w \in [0,1]^n, \sum_i w_i = 1\}$ with

$$h_\mathbf{w} = \sum_i w_i m_i$$

- Decision rule d?

$$d(m) = \arg\max_{\omega \in \Omega} \overline{P}(\{\omega\})$$

- **maximum of plausibility**
  $\rightarrow$ use plausibility of average = average of plausibilities at your advantage, i.e.,

$$\overline{P}_\Sigma(\omega) = \sum w_i \overline{P}_i(\omega)$$

Exercice 7: walking dead

A zombie apocalypse has happened, and you must recognize possible threats/supports

The possibilities Ω

- Zombie (Z)
- Friendly Human (F)
- Hostile Human (H)
- Neutral Human (N)

The sources S_i

- Half-broken heat detector (S_1)
- Paranoid Watch guy 1 (S_2)
- Half-borken Motion detector (S_3)
- Sleepy Watch guy 2 (S_4)

Exercice 7: which rule?

Given this table of contour functions, a weighted average and a decision based on maximal plausibility

	$\hat{\omega}^1 = Z$				$\hat{\omega}^2 = H$				$\hat{\omega}^3 = F$			
	Z	F	H	N	Z	F	H	N	Z	F	H	N
S_1	1	0,5	0,5	0,5	1	0,5	0,5	0,5	0,5	1	1	1
S_2	1	0,2	0,8	0,2	0	0,3	1	0,3	0	0,4	1	0,4
S_3	1	0,5	0,5	0,5	0,5	0,7	0,8	0,7	1	0,5	0,5	0,5
S_4	1	1	1	1	0,2	0,2	1	0,5	0,2	1	0,4	0,8
$\mathbf{w}_1 = (0.5, 0.5, 0, 0)$												
$\mathbf{w}_2 = (0, 0, 0.5, 0.5)$												

Choose $h_{\mathbf{w}_1}$ or $h_{\mathbf{w}_2}$? Given the data, can we find a strictly better weight vector?

Some on-going research topics within BF

Or what could you go for if you're interested in BF

Statistical estimation/machine learning

- Extending frequentist approaches [16]
- Embedding BF with classical ML [48, 15]
- BF for recent ML problems (ranking, multi-label) [18, 44]

Inference over large/combinatorial spaces

- Efficient handling over lattices (preferences, etc.) [17]
- Inferences over Boolean formulas [2, 38]
- BF and (discrete) Operations Research [37]

Specific fusion settings

- Decentralized fusion [33]
- Large spaces (2D/3D maps, images) [46]

As a conclusion

Belief functions as specific IP . . .

Many common points

- Specific setting including many important aspects
- May offer tools that facilitate handling/understanding to non-specialist (random set, Mobius inverse, Monte-Carlo + set computation)
- BF theory share strong similarities with IP

. . . but not only

Yet important differences:

- Admit incoherence when needed → may be useful sometimes
- Important notions in BF have no equivalent in IP → commonality function, specialisation notion, fusion rules, . . .

References I

[1] J. Abellan and S. Moral.
A non-specificity measure for convex sets of probability distributions.
International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 8:357–367, 2000.

[2] Felipe Aguirre, Sebastien Destercke, Didier Dubois, Mohamed Sallak, and Christelle Jacob.
Inclusion–exclusion principle for belief functions.
International Journal of Approximate Reasoning, 55(8):1708–1727, 2014.

[3] C. Baudrit and D. Dubois.
Practical representations of incomplete probabilistic knowledge.
Computational Statistics and Data Analysis, 51(1):86–108, 2006.

[4] Salem Benferhat, Julien Hué, Sylvain Lagrue, and Julien Rossit.
Interval-based possibilistic logic.
In *IJCAI*, pages 750–755, 2011.

[5] J. O. Berger.
An overview of robust Bayesian analysis.
Test, 3:5–124, 1994.
With discussion.

[6] Denis Bouyssou, Didier Dubois, Henri Prade, and Marc Pirlot.
Decision Making Process: Concepts and Methods.
John Wiley & Sons, 2013.

[7] M. Cattaneo.
Likelihood-based statistical decisions.
In *Proc. 4th International Symposium on Imprecise Probabilities and Their Applications*, pages 107–116, 2005.

References II

[8] John W Chinneck and Erik W Dravnieks.
Locating minimal infeasible constraint sets in linear programs.
ORSA Journal on Computing, 3(2):157–168, 1991.

[9] I. Couso, S. Montes, and P. Gil.
The necessity of the strong alpha-cuts of a fuzzy set.
Int. J. on Uncertainty, Fuzziness and Knowledge-Based Systems, 9:249–262, 2001.

[10] Inés Couso and Didier Dubois.
Statistical reasoning with set-valued information: Ontic vs. epistemic views.
International Journal of Approximate Reasoning, 55(7):1502–1518, 2014.

[11] L.M. de Campos, J.F. Huete, and S. Moral.
Probability intervals: a tool for uncertain reasoning.
I. J. of Uncertainty, Fuzziness and Knowledge-Based Systems, 2:167–196, 1994.

[12] G. de Cooman and P. Walley.
A possibilistic hierarchical model for behaviour under uncertainty.
Theory and Decision, 52:327–374, 2002.

[13] A.P. Dempster.
Upper and lower probabilities induced by a multivalued mapping.
Annals of Mathematical Statistics, 38:325–339, 1967.

[14] T. Denoeux.
Constructing belief functions from sample data using multinomial confidence regions.
I. J. of Approximate Reasoning, 42:228–252, 2006.

[15] Thierry Denoeux.
Logistic regression, neural networks and dempster-shafer theory: a new perspective.
arXiv preprint arXiv:1807.01846, 2018.

References III

[16] Thierry Denoeux and Shoumei Li.
Frequency-calibrated belief functions: Review and new insights.
International Journal of Approximate Reasoning, 92:232–254, 2018.

[17] Thierry Denœux and Marie-Hélène Masson.
Evidential reasoning in large partially ordered sets.
Annals of Operations Research, 195(1):135–161, 2012.

[18] Thierry Denœux, Zoulficar Younes, and Fahed Abdallah.
Representing uncertainty on set-valued variables using belief functions.
Artificial Intelligence, 174(7-8):479–499, 2010.

[19] S. Destercke, D. Dubois, and E. Chojnacki.
Possibilistic information fusion using maximal coherent subsets.
IEEE Trans. on Fuzzy Systems (in press), 2008.

[20] S. Destercke, D. Dubois, and E. Chojnacki.
Unifying practical uncertainty representations: I generalized p-boxes.
Int. J. of Approximate Reasoning, 49:649–663, 2008.

[21] S. Destercke, D. Dubois, and E. Chojnacki.
Unifying practical uncertainty representations: II clouds.
Int. J. of Approximate Reasoning (in press), pages 664–677, 2008.

[22] Sébastien Destercke and Thomas Burger.
Toward an axiomatic definition of conflict between belief functions.
Cybernetics, IEEE Transactions on, 43(2):585–596, 2013.

[23] D. Dubois, L. Foulloy, G. Mauris, and H. Prade.
Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities.
Reliable Computing, 10:273–297, 2004.

References IV

[24] D. Dubois, S. Moral, and H. Prade.
A semantics for possibility theory based on likelihoods,.
Journal of Mathematical Analysis and Applications, 205(2):359 – 380, 1997.

[25] D. Dubois and H. Prade.
A set-theoretic view on belief functions: logical operations and approximations by fuzzy sets.
Int. J. of General Systems, 12:193–226, 1986.

[26] D. Dubois and H. Prade.
Properties of measures of information in evidence and possibility theory.
Fuzzy sets and systems, 24:161–182, 1987.

[27] D. Dubois and H. Prade.
Possibility Theory: An Approach to Computerized Processing of Uncertainty.
Plenum Press, New York, 1988.

[28] Didier Dubois and Henri Prade.
Focusing vs. belief revision: A fundamental distinction when dealing with generic knowledge.
In *ECSQARU 97*, pages 96–107. Springer, 1997.

[29] Didier Dubois and Henri Prade.
Possibilistic logic: a retrospective and prospective view.
Fuzzy Sets and Systems, 144(1):3 – 23, 2004.

[30] Didier Dubois and Henri Prade.
An overview of the asymmetric bipolar representation of positive and negative information in possibility theory.
Fuzzy Sets and Systems, 160(10):1355–1366, 2009.

[31] Didier Dubois and Henri Prade.
Being consistent about inconsistency: Toward the rational fusing of inconsistent propositional logic bases.
In *The Road to Universal Logic*, pages 565–571. Springer, 2015.

References V

[32] Didier Dubois, Henri Prade, and Agnès Rico.
Representing qualitative capacities as families of possibility measures.
International Journal of Approximate Reasoning, 58:3–24, 2015.

[33] Bertrand Ducourthial and Véronique Cherfaoui.
Experiments with self-stabilizing distributed data fusion.
In *IEEE 35th Symposium on Reliable Distributed Systems (SRDS 2016)*, pages 289–296, 2016.

[34] Ronald Fagin and Joseph Halpern.
A new approach to updating beliefs.
In P. P. Bonissone, M. Henrion, L. N. Kanal, and J. F. Lemmer, editors, *Uncertainty in Artificial Intelligence*, volume 6, pages 347–374. North-Holland, Amsterdam, 1991.

[35] S. Ferson, L. Ginzburg, V. Kreinovich, D.M. Myers, and K. Sentz.
Constructing probability boxes and dempster-shafer structures.
Technical report, Sandia National Laboratories, 2003.

[36] B.de Finetti.
Theory of probability, volume 1-2.
Wiley, NY, 1974.
Translation of 1970 book.

[37] Nathalie Helal, Frédéric Pichon, Daniel Porumbel, David Mercier, and Éric Lefèvre.
The capacitated vehicle routing problem with evidential demands.
International Journal of Approximate Reasoning, 95:124–151, 2018.

[38] Christelle Jacob, Didier Dubois, and Janette Cardoso.
Evaluating the uncertainty of a boolean formula with belief functions.
In *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 521–531. Springer, 2012.

References VI

[39] Jean-Yves Jaffray.
Bayesian updating and belief functions.
IEEE Transactions on Systems, Man and Cybernetics, 22:1144–1152, 1992.

[40] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter.
Applied Interval Analysis.
London, 2001.

[41] Henry E. Kyburg Jr.
Bayesian and non-bayesian evidential updating.
Artif. Intell., 31(3):271–293, 1987.

[42] Jianbing Ma, Weiru Liu, Didier Dubois, and Henri Prade.
Bridging jeffrey's rule, agm revision and dempster conditioning in the theory of evidence.
International Journal on Artificial Intelligence Tools, 20(04):691–720, 2012.

[43] Ryan Martin, Jianchun Zhang, and Chuanhai Liu.
Dempster-shafer theory and statistical inference with weak beliefs.
Technical report, 2008.

[44] Marie-Hélène Masson, Sébastien Destercke, and Thierry Denoeux.
Modelling and predicting partial orders from pairwise belief functions.
Soft Computing, 20(3):939–950, 2016.

[45] Enrique Miranda and Sébastien Destercke.
Extreme points of the credal sets generated by comparative probabilities.
Journal of Mathematical Psychology, 64:44–57, 2015.

[46] Nicola Pellicanò, Sylvie Le Hégarat-Mascle, and Emanuel Aldea.
2cobel: An efficient belief function extension for two-dimensional continuous spaces.
arXiv preprint arXiv:1803.08857, 2018.

References VII

[47] Frédéric Pichon, Didier Dubois, and Thierry Denoeux.
Relevance and truthfulness in information correction and fusion.
Int. J. Approx. Reasoning, 53(2):159–175, 2012.

[48] Benjamin Quost, Thierry Denœux, and Marie-Hélène Masson.
Pairwise classifier combination using belief functions.
Pattern Recognition Letters, 28(5):644–653, 2007.

[49] S.A. Sandri, D. Dubois, and H.W. Kalfsbeek.
Elicitation, assessment and pooling of expert judgments using possibility theory.
IEEE Trans. on Fuzzy Systems, 3(3):313–335, August 1995.

[50] G. Shafer.
A mathematical Theory of Evidence.
Princeton University Press, New Jersey, 1976.

[51] P. Smets.
The transferable belief model and other interpretations of dempster-shafer's model.
In *Proc. of the Sixth Annual Confernce on Uncertainty in Artifical Intelligence*, pages 375–384, 1990.

[52] Philippe Smets.
Probability of provability and belief functions.
Logique et Analyse, 133-134:177–195, 1991.

[53] Matthias C. M. Troffaes and Sébastien Destercke.
Probability boxes on totally preordered spaces for multivariate modelling.
Int. J. Approx. Reasoning, 52(6):767–791, 2011.

[54] R. von Mises.
Probability, Statistics and Truth.
Dover books explaining science. Dover Publications, 1981.

References VIII

[55] P. Walley.
Statistical reasoning with imprecise Probabilities.
Chapman and Hall, New York, 1991.

[56] Kurt Weichselberger.
The theory of interval-probability as a unifying concept for uncertainty.
International Journal of Approximate Reasoning, 24(2–3):149 – 170, 2000.

[57] R.R. Yager.
Entropy and specificity in a mathematical theory of evidence.
I. J. of Systems, 9:249–260, 1983.