Uncertainty representation and combination: new results with applications to nuclear safety issues

Sebastien Destercke [1,2]

[1]Institut de Radioprotection et de Sûreté Nucléaire

[2]Institut de Recherche en Informatique de Toulouse

Thesis defense - 29 Octobre 2008
Directeur de thèse: Didier Dubois
Co-Directeur de thèse: Eric Chojnacki

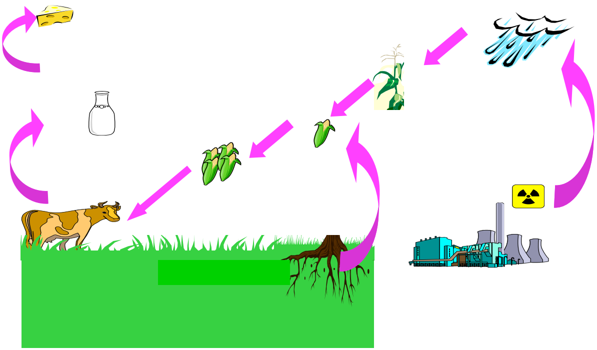
Context

Thesis

S. Destercke

Introduction
Context
Representation
Generic tools
State of art
Gen. p-boxes
Clouds
Mutliple sources
Inf. fusion
Rel. assess.
Application
Propagation
Independence
Prac. prop
Eval and Dec.
Conclusions & perspectives

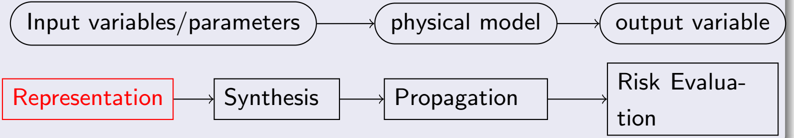
Risk analysis → many uncertainties



Example: environmental protection

Overview

Classical situation



Overview

- Representation
- Synthesis
  - Information fusion
  - Reliability assessment
- Propagation
  - Independence assumptions
  - Practical propagation
- Risk evaluation and decision making

Basic setting

Situation

Describe our uncertainty about the value assumed by a variable X on a domain $\mathscr{X}$ (e.g. temperature in a room, state of a sensor, ...). Here, the domain $\mathscr{X}$ is either:

- finite
- the real line $\mathbb{R}$ with associated borel σ-field

In the latter case, when considering discrete representations, we can come back to a finite domain by taking a suitable partition of $\mathbb{R}$.

Why imprecise probability frameworks?

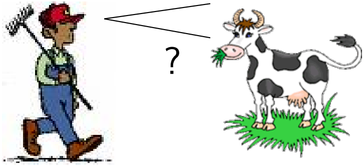
Two basic models

- Intervals or sets: no event is more likely to occur than another, complete imprecision (worst-case analysis)
- Probability distributions: precise estimation of the confidence of the occurence of an event

In practice, often more information than an interval, but not enough to identify a precise probability.

Why imprecise probability frameworks? (Example)

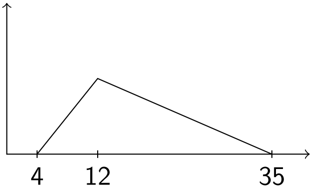
How much grass per day ?



Answer: *usually around 12 Kg, but can go from 4 to 35 Kg*

$\Rightarrow$ interval $[4, 35]$: less information than available

$\Rightarrow$ triangular probability density with mode 12 and support $[4, 35]$) $\rightarrow$ more information than really available



Solutions and approaches

Coping with imprecision

Three main formal frames (denoted $\mathscr{F}$) propose to cope with intermediary states of

- **lower/upper probabilities**
- **random sets**
- **possibility theory**

↑ **Generality**

↓ **Simplicity**

$\rightarrow$ understanding their links, similarities,differences is important to achieve an unified handling of uncertainties.

Generic representation tool

Thesis
S. Destercke
Introduction
Context
Representation
Generic tools
State of art
Gen. p-boxes
Clouds
Mutliple sources
Inf. fusion
Rel. assess.
Application
Propagation
Independence
Prac. prop
Eval and Dec.
Conclusions & perspectives

Capacity

a capacity on $\mathscr{X}$ is a function μ, defined on the power set $\wp(\mathscr{X})$ of $\mathscr{X}$, such that:

- $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$ (monotonicity)
- $\mu(\emptyset) = 0, \mu(\mathscr{X}) = 1$ (boundary conditions)

Generic representation tool

Modeling imprecision and uncertainty

3 state of knowledge $\rightarrow$ need of two measures $\underline{\mu} \leq \overline{\mu}$:

- Certainty of event A: $\underline{\mu}(A) = 1, \overline{\mu}(A) = 1$
- Impossibility of event A: $\underline{\mu}(A) = 0, \overline{\mu}(A) = 0$
- Ignorance about event A: $\underline{\mu}(A) = 0, \overline{\mu}(A) = 1$

$\underline{\mu} \leq \overline{\mu}$ related by conjugacy relation such that, for any event E,

$$\underline{\mu}(E) = 1 - \overline{\mu}(E^c)$$

Generic framework: Lower probabilities (Walley, 91)

Associated set of probabilities

To $\underline{\mu}$ correspond a convex set (Credal set) of probabilities $\mathscr{P}_{\underline{P}}$ s.t.

$$\mathscr{P}_{\underline{\mu}} := \{P \in \mathbb{P}_{\mathscr{X}} | (\forall A \subseteq \mathscr{X})(P(A) \geq \underline{\mu}(A))\},$$

with $\mathbb{P}_{\mathscr{X}}$: set of all probability measures on $\mathscr{X}$.

Consistence/coherence

- $\underline{\mu}$ is said **consistent** if $\mathscr{P}_{\underline{\mu}} \neq \emptyset$
- $\underline{\mu}$ is said **coherent** if $\underline{\mu}(A) = \inf_{P \in \mathscr{P}_{\underline{\mu}}} P(A)$
- If $\underline{\mu}$ coherent, $\underline{\mu} = \underline{P}$

Representation problem statement

Simple representations

General models: hardly tractable in practice → need for simpler representations, easier to deal with → many of them proposed and still proposed.

Problem

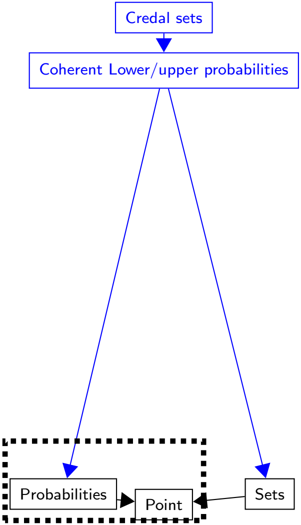
Recent representations (p-boxes, clouds) have not yet been related thoroughly to others.

Why such a study?

Both theoretical and practical issues

- need to know how they settle in existing frameworks
- gain insights about their expressiveness, easiness of use and other features.

A first summary



A → B

B is particular case of A

.... Classic. Proba.

2-monotone lower probabilities

Definition

A lower probability $\underline{P}$ is $2-monotone$ if, for every $A, B \subset \mathscr{X}$, the inequality

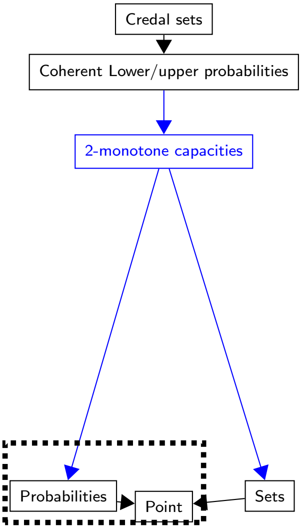
$$\underline{P}(A \cup B) + \underline{P}(A \cap B) \geq \underline{P}(A) + \underline{P}(B)$$

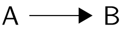
holds

Properties

- Always coherent lower probability
- Simplify many mathematical operations

The scheme continued (again)



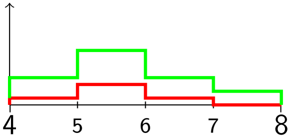


B is particular case of A

Classic. Proba.

Probability intervals (De Campos, Huete, Moral, 94)

Expert providing his opinion about the potential value of pH in a given field



Also correspond to: Imprecise histograms, small multinomial samples

Definition

Set $L = \{[l(x), u(x)] | x \in \mathscr{X}\}$ of bounds on elements of $\mathscr{X}$ verifying inducing the credal set

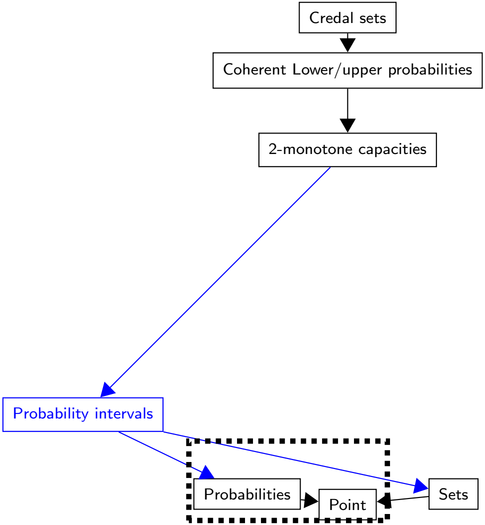
$$\mathscr{P}_L = \{P \in \mathbb{P}_{\mathscr{X}} | \forall x, \; l(x) \leq p(x) \leq u(x)\}.$$

We assume bounds L to be consistent and coherent

Lower (2-monotone) probability s.t.:

$$\underline{P}(A) = \max(\sum_{x \in A} l(x), 1 - \sum_{x \in A^c} u(x))$$

The summary continued (once again)

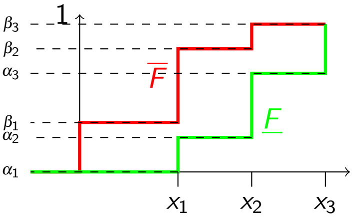


A → B

B is particular case of A

.... Classic. Proba.

P-box: imprecise cumulative distribution (Ferson, 03, Williamson & Downs, 90)



Discrete p-box

Expert opinions expressed through percentiles, small interval with confidence band (Kolmogorov-Smirnov distance)

Definition

Pair of cumulative distribution $[\underline{F}, \overline{F}]$ on $\mathbb{R}$.
Induced Lower probability consistent if $\underline{F}$ stochastically dominate $\overline{F}$

$$\underline{F}(x) \leq \overline{F}(x) \ \forall x \in \mathbb{R}$$

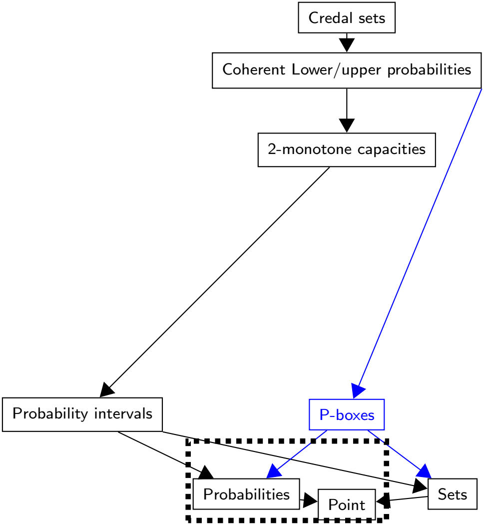
Induced credals set

$$\mathscr{P}_{[\underline{F},\overline{F}]} = \{P | \forall x, \ \underline{F}(x) \leq P((-\infty, x]) \leq \overline{F}(x)\}$$

In practice, discrete p-box induced by a finite set of n constraints

$$i = 1, \ldots, n, \ \alpha_i \leq P((-\infty, x_i]) \leq \beta_i$$

Yet another summary



A → B

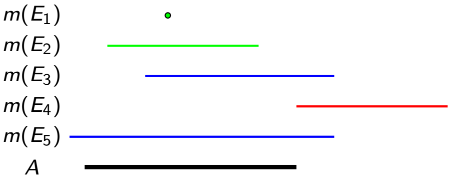
B is particular case of A

.... Classic. Proba.

Second framework: random sets (Shafer, 76), (Dempster, 67), (Smets, 94)

Definition

A (discrete) mass distribution is a mapping $m : \wp(\mathscr{X}) \to [0,1]$ such that $\sum_{E \subseteq \mathscr{X}} m(E) = 1$, and a set with masses > 0 is called focal. $m(E)$ is a probabilistic mass to allocate to elements of E



$Bel(A) = m(E_1) + m(E_2)$

$Pl(A) = m(E_1) + m(E_2) + m(E_3) + m(E_5)$

$$Bel(A) = \sum_{E \subseteq A} m(E) \qquad \text{(Masses necessarily} \in A)$$

$$Pl(A) = \sum_{E \cap A \neq \emptyset} m(E) = 1 - Bel^c(A) \qquad \text{(Masses potentially} \in A\text{)}$$

Link with lower probabilities (Dempster, 67)

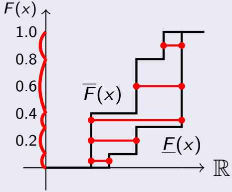
A belief function *Bel* induce the credal set

$$\mathscr{P}_{Bel} := \{P \in \mathbb{P}_{\mathscr{X}} | (\forall A \subseteq \mathscr{X})(P(A) \geq Bel(A))\},$$

Practical usefulness: simulating $\mathscr{P}_{Bel}$ by sampling m

P-boxes

P-boxes are special cases of random sets (Kriegler & Held, 05).

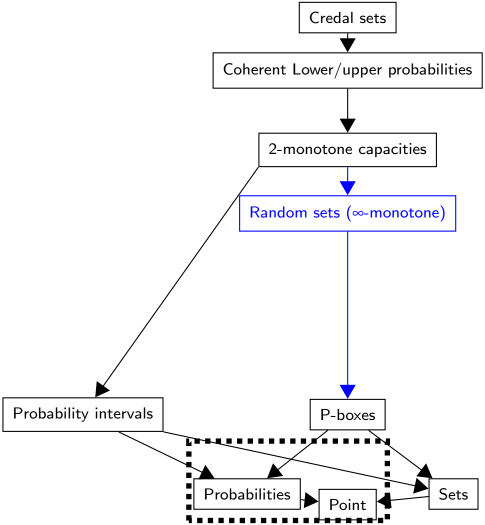


Probability Intervals

No particular links between random sets and probability intervals.

Authors have studied mapping of a prob. int. L into a random set (Lemmer & Kyburg, Denoeux)

This is not a summary



A → B

B is particular case of A

.... Classic. Proba.

Third framework: possibility theory

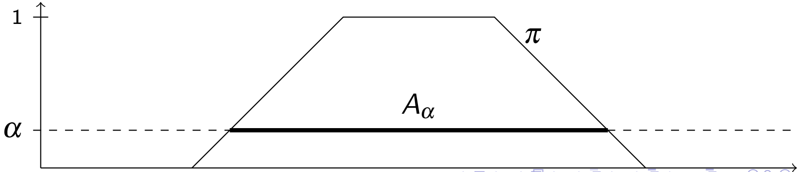
Definition

A possibility distribution π is a mapping $\pi : \mathscr{X} \to [0,1]$ such that $\exists x,\ \pi(x) = 1$, and a set with masses > 0 is called focal. Given $A \subseteq \mathscr{X}$, two measures are defined:

$$\Pi(A) = \sup_{x \in A} \pi(x) \qquad \text{(Possibility)}$$

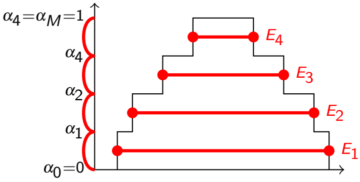
$$N(A) = 1 - \Pi(A^c) \qquad \text{(Necessity)}$$

And an α-cut is defined as $A_\alpha = \{x \in \mathscr{X} | \pi(x) \geq \alpha\}$ (strict if the inequality is strict)



Possibilities as random sets

Possibility distribution is a particular case of random sets with nested realisations (Shafer, 76)

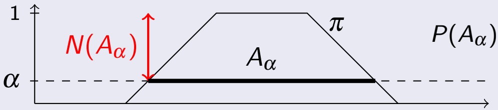


Possibilities as Credal sets

A necessity measure: special case of lower probability (Dubois & Prade, 92), (de Cooman & Aeyels, 99) inducing

$$\mathscr{P}_\pi = \{P \in \mathbb{P}_{\mathscr{X}} | \forall A \subseteq \mathscr{X}, \, P(A) \geq N(A)\}$$

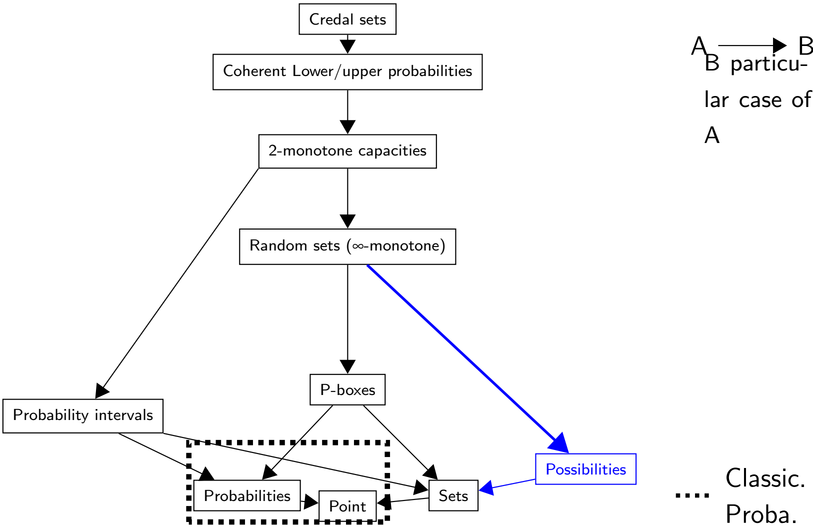
Characterization by constraints on α-cuts (Dubois et al., 04), (Couso et al., 01)



$P(A_\alpha) \geq 1 - \alpha \; (\forall \alpha \in [0,1])$

N.B. upper bounds of $P(A_\alpha)$ always trivial (i.e. =1)

State of the art: summary



Generalized p-boxes: introduction

Thesis

S. Destercke

Introduction
Context
Representation
Generic tools
State of art
Gen. p-boxes
Clouds
Mutliple sources
Inf. fusion
Rel. assess.
Application
Propagation
Independence
Prac. prop
Eval and Dec.
Conclusions & perspectives

Why studying such a model?

- Possibility distributions: nested sets with lower confidence bounds
- (Discrete) P-boxes: lower and upper probabilistic bounds on (nested) sets $(-\infty, x_i]$

Both, even if poorly expressive, are very useful tools in many applications

Basic idea

Extend them both by studying a model where we give lower and upper probabilistic bounds on a collection of nested sets.

Generalized p-boxes: introduction

Constraints

Let $\emptyset \subset A_1 \subset \ldots \subset A_n \subseteq \mathscr{X}$ be a collection and nested sets. A Generalised p-box represent constraints

$$\begin{array}{c} \alpha_i \leq P(A_i) \leq \beta_i \qquad i = 1, \ldots, n \\ 0 \leq \alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_n \leq 1 \\ 0 \leq \beta_1 \leq \beta_2 \leq \ldots \leq \beta_n \leq 1 \end{array}$$

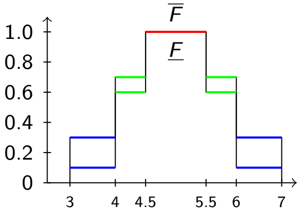
$\rightarrow$ study the induced lower probability and credal set, and its link to previous representations.

An example

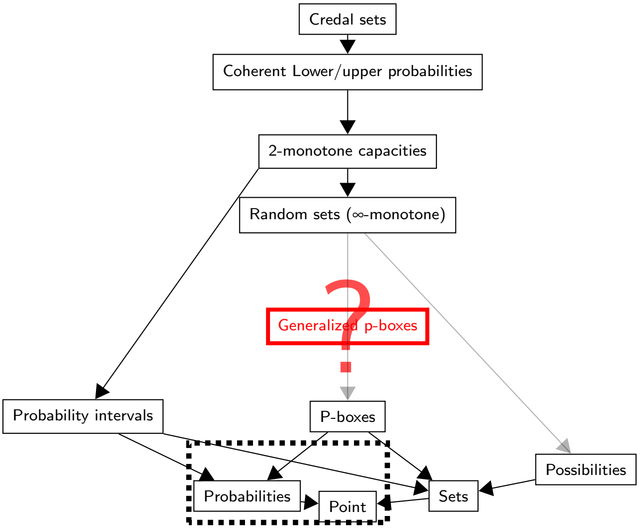
Evaluating impact of radionuclides inhalation on workers (e.g. in Uranium mines) → key parameter: mean diameter of particles (AMAD)

Expert opinion translated in constraints:

- $0.3 \leq P([4.5, 5.5]) \leq 0.6$
- $0.7 \leq P([4, 6]) \leq 0.9$
- $1 \leq P([3, 7]) \leq 1$



Generalised p-boxes enter the picture



A → B

B particular case of A

.... Classic. Proba.

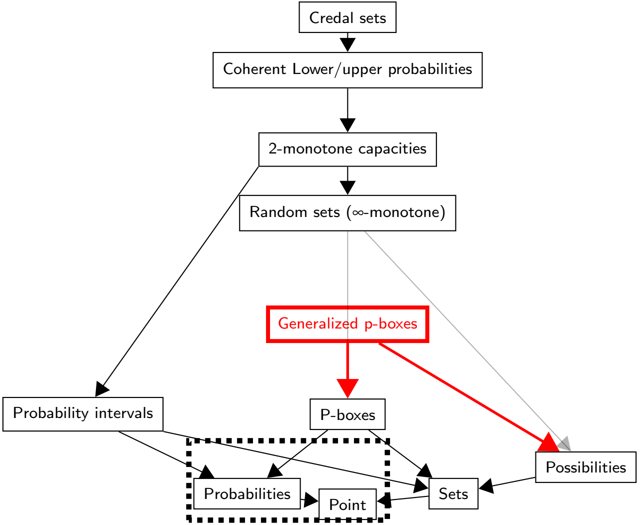
Generalized p-boxes: first results

First links with previous representations

$$\alpha_i \leq P(A_i) \leq \beta_i \qquad i = 1, \dots, n$$
$$0 \leq \alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_n \leq 1$$
$$0 \leq \beta_1 \leq \beta_2 \leq \dots \leq \beta_n \leq 1$$

- We retrieve possibility distributions when $\beta_i = 1$, $i = 1, \dots, N$
- We retrieve p-boxes when $\mathscr{X} = R$ and $A_i = (-\infty, x_i)$

Generalised p-boxes get more involved in the picture



A → B
B particular case of A

.... Classic.

Generalized p-boxes: formal definition (Destercke et al., 08)

Construction

Nested sets $\emptyset \subset A_1 \subset \ldots \subset A_n = \mathscr{X}$ $\rightarrow$ Sets $A_i \setminus A_{i-1}$ partition of $\mathscr{X}$.

Define $[\underline{F}, \overline{F}]$ such that, if $x \in A_i \setminus A_{i-1}$, $\overline{F} = \beta_i, \underline{F} = \alpha_i$

Two mappings f, f' from $\mathscr{X} \rightarrow \mathbb{R}$ are comonotone iff $\forall x, y \in \mathscr{X}$, $f(x) < f(y) \rightarrow f'(x) \leq f(y')$

Definition

A generalized p-box is a pair of comonotone mappings $\overline{F} : \mathscr{X} \rightarrow [0,1]$ and $\underline{F} : \mathscr{X} \rightarrow [0,1]$ s.t. $\exists x, \overline{F}(x) = \underline{F}(x) = 1$

Generalized p-boxes: formal definition (Destercke et al., 08)

Induced credal set

The credal set $\mathscr{P}_{[\underline{F},\overline{F}]}$ induced by a gen. p-box $[\underline{F},\overline{F}]$ is defined as

$$\mathscr{P}_{[\underline{F},\overline{F}]} = \{P \in \mathbb{P}_{\mathscr{X}} | \forall A_i, \; \alpha_i \leq P(A_i) \leq \beta_i\}$$

Generalized p-boxes: links with other representations

Thesis
S. Destercke
Introduction
Context
Representation
Generic tools
State of art
Gen. p-boxes
Clouds
Mutliple sources
Inf. fusion
Rel. assess.
Application
Propagation
Independence
Prac. prop
Eval and Dec.
Conclusions & perspectives

Theorem (Destercke et al., 08)

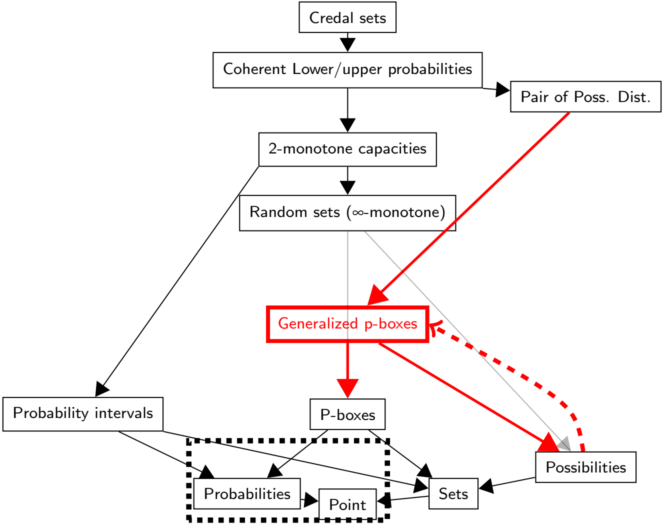
From any generalized p-box $[\underline{F}, \overline{F}]$, we can define two possibility distributions $\pi_{\overline{F}}, \pi_{\underline{F}}$ on $\mathscr{X}$ such that

$$\mathscr{P}_{[\underline{F},\overline{F}]} = \mathscr{P}_{\pi_{\overline{F}}} \cap \mathscr{P}_{\pi_{\underline{F}}}$$

holds

$\Rightarrow$ Generalized p-boxes are representable by pairs of possibility distributions.

Generalised p-boxes get comfortable in the picture



A → B
B particular case of A

A - - - → B
A represent B

···· Classic.

Generalized p-boxes: links with other representations

Theorem (Destercke et al., 08)

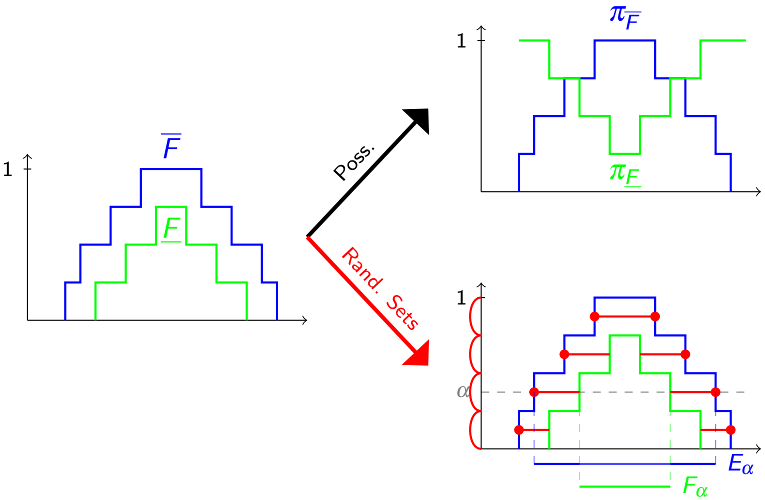
Any generalized p-box $[\underline{F}, \overline{F}]$ can be represented as a particular random set for which, to every level $\alpha \in [0,1]$, we associate the focal element

$$E_\alpha \setminus F_\alpha$$

with E_α:α-cut of $\pi_{\overline{F}}$ and F_α: α-cut of $1 - \pi_{\underline{F}}$

$\Rightarrow$ Calculus used for generic random sets can be directly applied to generalized p-boxes

Generalised p-boxes and links: illustration



Generalized p-boxes: links with other representations

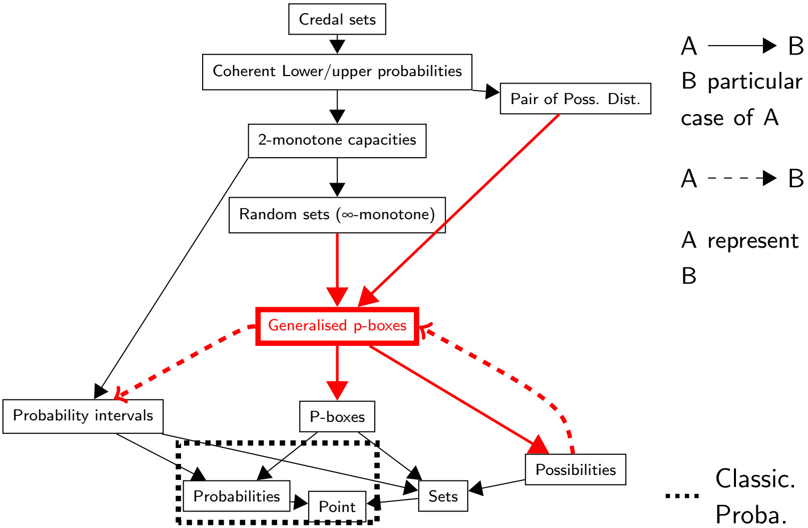
Theorem (Destercke et al., 08)

From a probability interval L, it is possible to build $|\mathscr{X}|/2$ generalized p-boxes $[\underline{F},\overline{F}]_1,\ldots,[\underline{F},\overline{F}]_{|\mathscr{X}|/2}$ such that

$$\mathscr{P}_L = \cap_{i=1}^{|\mathscr{X}|/2} \mathscr{P}_{[\underline{F},\overline{F}]_i}$$

$\Rightarrow$ Probability intervals representable by generalized p-boxes

Generalised p-boxes in the picture: the end



Clouds: introduction and definition

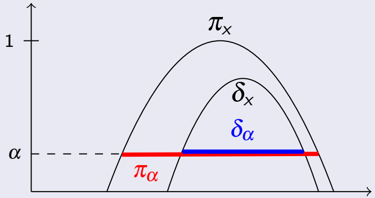
Introduced (Neumaier, 04) to deal with imprecision in high dimensions

Definition

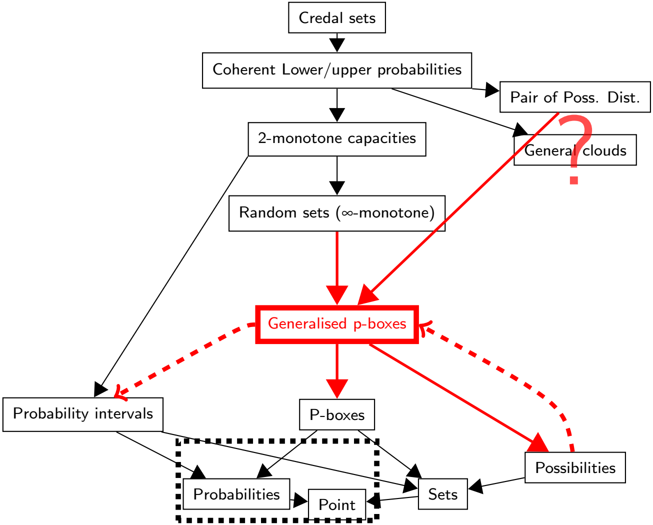
Cloud $[\pi, \delta]$: pair of mappings $\delta : \mathscr{X} \to [0,1]$, $\pi : \mathscr{X} \to [0,1]$, with $\delta \leq \pi$, $\pi(x) = 1$ for at least one element x in $\mathscr{X}$, and $\delta(y) = 0$ for at least one element y in $\mathscr{X}$.

Induced credal set (Neumaier, 04)

$$\mathscr{P}_{[\pi,\delta]} = \{P \in \mathbb{P}_{\mathscr{X}} | P(\delta_\alpha) \leq 1 - \alpha \leq P(\pi_\alpha)\}$$



Now clouds want to get in



A → B
B particular case of A

A - - -→ B

A represent B

.... Classic. Proba.

Clouds: links with other representation

Theorem (Destercke et al., 08)

The two following statements are equivalent:

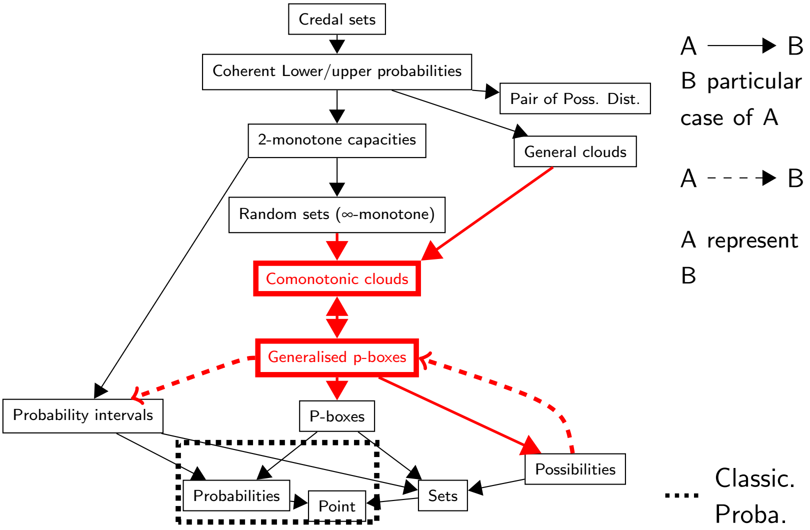
(i) The cloud $[\pi, \delta]$ can be encoded as a generalised p-box $[\underline{F}, \overline{F}]$ such that $\mathscr{P}_{[\pi,\delta]} = \mathscr{P}_{[\underline{F},\overline{F}]}$

(ii) δ and π are comonotonic $(\delta(x) < \delta(y) \Rightarrow \pi(x) \leq \pi(y))$

and a cloud is said comonotonic if δ and π are comonotonic.

$\Rightarrow$ **comonotonic clouds** and **generalised p-boxes: equivalent representations**

They already fit in quite well



Clouds: links with other representations

Thesis
S. Destercke
Introduction
Context
Representation
Generic tools
State of art
Gen. p-boxes
Clouds
Mutliple sources
Inf. fusion
Rel. assess.
Application
Propagation
Independence
Prac. prop
Eval and Dec.
Conclusions & perspectives

Theorem (Destercke et al., 08)

A cloud $[\pi, \delta]$ is representable by the pair of possibility distributions $1 - \delta$ and π, in the following sense:

$$\mathscr{P}_{[\pi,\delta]} = \mathscr{P}_{\pi} \cap \mathscr{P}_{1-\delta}$$

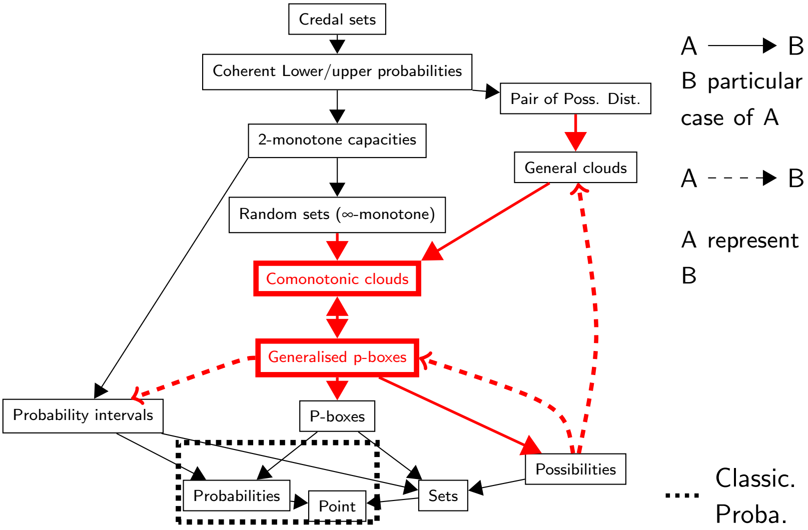
Clouds: links with other representations

Theorem (Destercke et al., 08)

There are families of non-comonotonic clouds $[\pi, \delta]$ such that the lower probability induced by the credal set $\mathscr{P}_{[\pi,\delta]}$ is not even 2-monotone

$\Rightarrow$ clouds not special cases of random sets, and non-comonotonic clouds appears of less practical interest.

Finally


Credal sets
Coherent Lower/upper probabilities
Pair of Poss. Dist.
2-monotone capacities
General clouds
Random sets (∞-monotone)
Comonotonic clouds
Generalised p-boxes
Probability intervals
P-boxes
Possibilities
Probabilities
Point
Sets
A → B
B particular case of A
A - - - → B
A represent B
Classic. Proba.

Overview

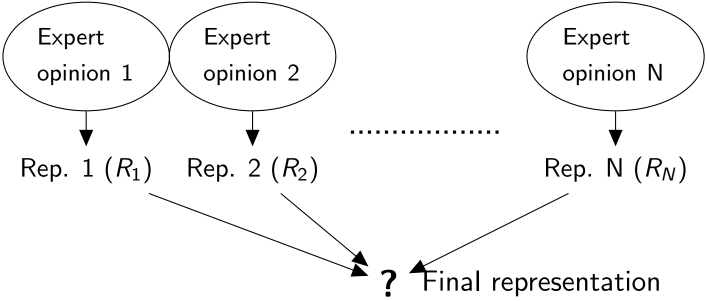
Thesis
S. Destercke
Introduction
Context
Representation
Generic tools
State of art
Gen. p-boxes
Clouds
Mutliple sources
Inf. fusion
Rel. assess.
Application
Propagation
Independence
Prac. prop
Eval and Dec.
Conclusions & perspectives

- Representation
- Synthesis
  - Information fusion
  - Reliability assessment
- Propagation
  - Independence assumptions
  - Practical propagation
- Risk evaluation and decision making

Information fusion: setting

Receiving and representing Information from multiple sources (e.g., experts, physical models) → summarise this information into a single representation

Example: expert opinions on the same variable (e.g., AMAD)



$$\varphi(R_1, R_2, \ldots, R_N)$$

Behaviours of φ

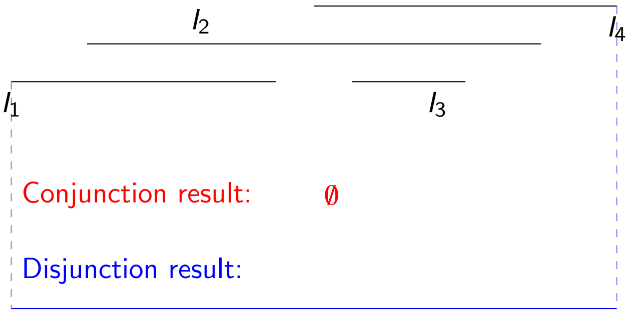
Choice of φ

Can be guided by the presence/absence of conflict between sources

φ can follow three main kinds of behaviour:

- **Conjunctive ($\cap$)**: $\varphi(R_1, \ldots, R_N) \subseteq R_i$ for $i = 1, \ldots, N$. Result is more informative than each source. Assume reliability of all sources and no conflict between them.
- **Disjunctive ($\cup$)**: $\varphi(R_1, \ldots, R_N) \supseteq R_i$ for $i = 1, \ldots, N$. Result is not more informative than each source. Assume reliability of at least one sources.
- **Compromise**: result between conjunctive and disjunctive behaviours.

Conjunction/disjunction: illustration



$\Rightarrow$ Conjunction not reliable.

$\Rightarrow$ Disjunction too imprecise.

$\rightarrow$ inadequate to cope with partial conflict.

a solution

Adaptive fusion rules

Goes from conjunction when there is no conflict towards disjunction when conflict increase

use of maximal coherent subsets as a general approach (Walley, 82), (Dubois & Prade, 90)

maximal coherent subsets: principles

Original idea from logic (Rescher & Manor, 70)

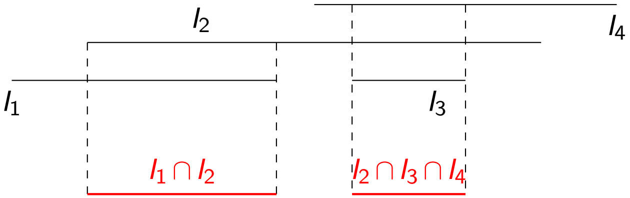
Resolve inconsistencies in knowledge bases :

- extract maximal subsets of consistent formulas (conjunction)
- proposition true if true in every subsets (disjunction)

Application to uncertainty representations

- extract k maximal subsets $K_i \subseteq \{R_1, \ldots, R_N\}$ of representations having non-empty conjunction
- take the disjunction of all conjunctions.

Maximal coherent subsets: illustration (Dubois, Fargier, Prade, 00)



Maximal coherent subsets: $K_1 = \{I_1, I_2\}$ and $K_2 = \{I_2, I_3, I_4\}$

Final result: $(I_1 \cap I_2) \cup (I_2 \cap I_3 \cap I_4)$

MCS: practical issue

Problem

Maximal coherent subsets theoretically and conceptually attractive, but

Extracting MCS → NP-complete problem in boolean logic: **computational intractability!**

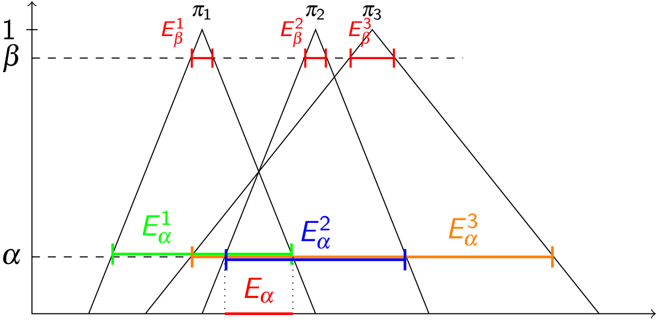
Solutions

- use heuristics and approximations
- work in a restricted but tractable framework: intervals on the real line → polynomial complexity (Dubois, Fargier, Prade, 00)

Level-wise MCS with possibility distributions

Our proposition

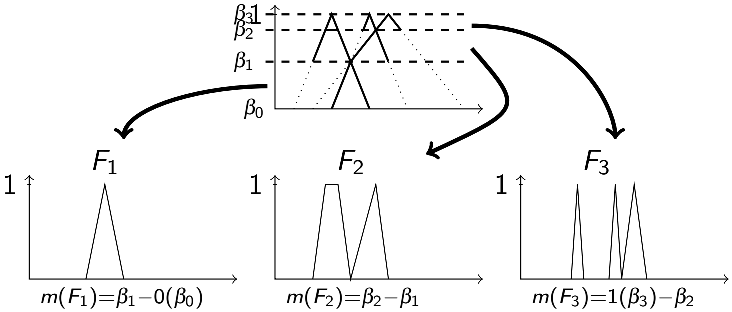
N distributions π_i: apply MCS to each level $\alpha \in [0,1]$.



Results for $\neq$ levels $\rightarrow$ not necessarily nested

Level-wise MCS with possibility distributions

Finite set of values β_i $i = 0, 1, \ldots, n$ such that sets E_α resulting from MCS for $\alpha \in (\beta_i, \beta_{i+1}]$ are nested

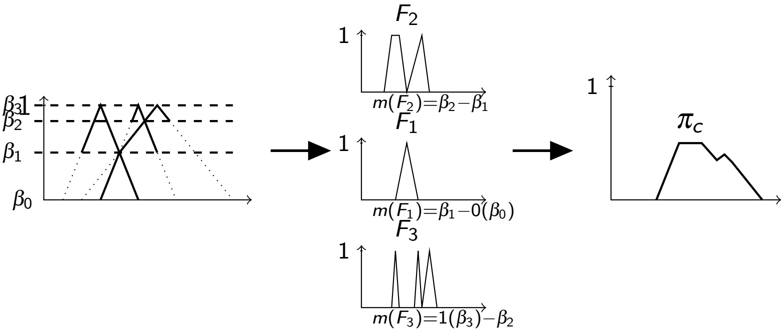


Result: n possibility distributions with weights $(\sum m(F_i) = 1)$

Level-wise MCS with possibility distributions

Summarizing the information

$m(F_i)$ Complex structure $\rightarrow$ compute contour function π_c as an interpretable summary (weighted average of F_i)



Fusion rules for clouds ?

Definition

Let $[\pi, \delta]_1, \ldots, [\pi, \delta]_N$ be N clouds, we propose the following fusion rules:

- Conjunction: $[\pi, \delta]_{\cap} = [\pi_{\cap}, \delta_{\cap}] = [\min_{i=1}^{N}(\pi_i), \max_{i=1}^{N}(\delta_i)]$.
- Disjunction: $[\pi, \delta]_{\cup} = [\pi_{\cup}, \delta_{\cup}] = [\max_{i=1}^{N}(\pi_i), \min_{i=1}^{N}(\delta_i)]$

$\longrightarrow$ conjunction and disjunction defined, maximal coherent subsets follow.

Overview

- Representation
- Synthesis
  - Information fusion
  - Reliability assessment
- Propagation
  - Independence assumptions
  - Practical propagation
- Risk evaluation and decision making

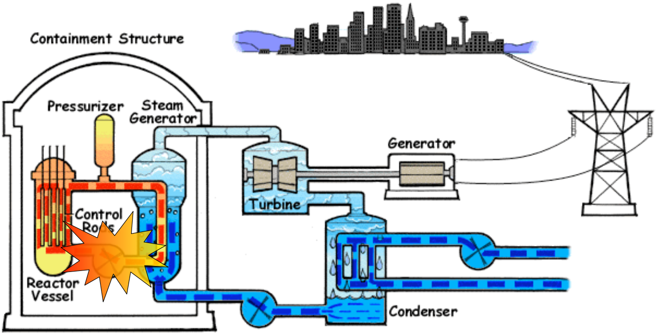
Evaluation of source reliability

Principle (Cooke, 91), (Sandri et al., 95)

Evaluate sources from past performance. Two quantitative values:

- **Precision** of information delivered by source. The more precise the information, the more useful it is $\Rightarrow$ proposition of a general criteria based on cardinality
- **Accuracy**: consistency between delivered information and observed (experimental) values $\Rightarrow$ proposition of a general criteria based on inclusion index
- **Global**: global score=precision $\times$ accuracy

Application to result of OCDE project BEMUSE



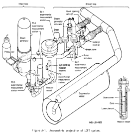
Ten different institutes use their own models and experts to reproduce a simulated accident → **use fusion rules and information evaluation technics to analyse information, with the help of SUNSET software**

Application to result of OCDE project BEMUSE

Experiment to reproduce

10 $\neq$ modelling of this experiment

10 $\neq$ Results to compare

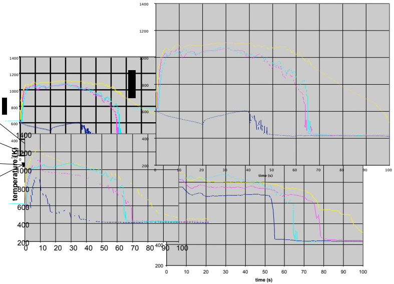


CATHARE

ATHLET

RELAP5

MARS



Application to result of OCDE project BEMUSE

Introduction
Context
Representation
Generic tools
State of art
Gen. p-boxes
Clouds
Mutliple sources
Inf. fusion
Rel. assess.
Application
Propagation
Independence
Prac. prop
Eval and Dec.
Conclusions & perspectives

Result

- Detection of participants overestimating (bad precision, good accuracy) or underestimating (good precision, bad accuracy) their uncertainty
- Quantified evaluation of conflict between subgroups of sources
- Generic tool to validate computer codes

Interest of non-experts

- Results added to final report
- Price at $\lambda\mu$ conference (high number of participants from industry)

Overview

Thesis

S. Destercke

Introduction
Context

Representation
Generic tools
State of art
Gen. p-boxes
Clouds

Mutliple sources
Inf. fusion
Rel. assess.
Application

Propagation
Independence
Prac. prop

Eval and Dec.

Conclusions & perspectives

- Representation
- Synthesis
  - Information fusion
  - Reliability assessment
- Propagation
  - Independence assumptions
  - Practical propagation
- Risk evaluation and decision making

Problem setting

Propagate uncertainty through a model $f(X_1, \ldots, X_N) = Y$ to evaluate uncertainty on Y.

- Often, information given separately for $X_1, \ldots, X_N$
- Then propagate through f with independence assumptions between
- **Many** different notions of independence when using imprecise probabilistic frameworks

$\rightarrow$ need to make some sense of them, to relate them and to understand their respective usefulness

Our contribution

Preliminary work

First classification of independence notions based on:

- **Informative** vs **non-informative**
- **Symmetric** vs **Asymmetric**
- **Objective** vs **Subjective**

Practical results:

- using more tractable independence notions as conservative approximation of less tractable ones
- relating notions of independence to imprecise probabilistic trees (work with G. de Cooman)

Overview

- Representation
- Synthesis
  - Information fusion
  - Reliability assessment
- Propagation
  - Independence assumptions
  - Practical propagation
- Risk evaluation and decision making

Starting point

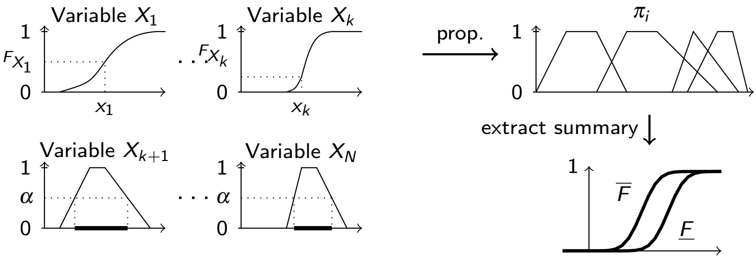
Thesis

S. Destercke

Introduction
Context
Representation
Generic tools
State of art
Gen. p-boxes
Clouds
Mutliple sources
Inf. fusion
Rel. assess.
Application
Propagation
Independence
Prac. prop
Eval and Dec.
Conclusions & perspectives

Hybrid propagation

Propagate by differentiating aleatory uncertainty (probabilistic calculus) from epistemic uncertainty (possibilistic calculus)

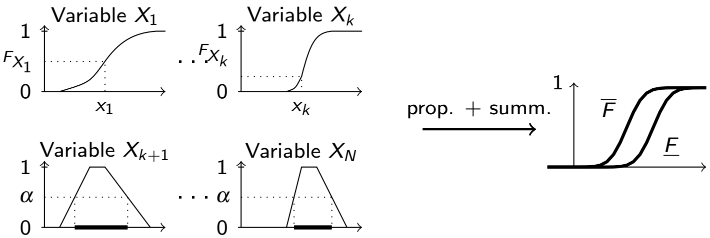


High computational cost to concentrate on specific summary → sometimes unaffordable

Improving efficiency

"RaFu" method (implemented in SUNSET software)

Use hybrid propagation → sample from distributions only values needed to compute desired result.



Reduce number of computations ($\sim$ 10 to 20 times less) by concentrating on desired result ⇒ currently applied in BEMUSE propagation

Overview

Thesis
S. Destercke
Introduction
Context
Representation
Generic tools
State of art
Gen. p-boxes
Clouds
Mutliple sources
Inf. fusion
Rel. assess.
Application
Propagation
Independence
Prac. prop
Eval and Dec.
Conclusions & perspectives

- Representation
- Synthesis
  - Information fusion
  - Reliability assessment
- Propagation
  - Independence assumption
  - Practical propagation
- Risk evaluation and decision making

Computing expectations

With probabilities

Decision making based on the computation of expected value $\mathbb{E}_P(u)$ of a function $u : \mathscr{X} \to \mathbb{R}$, given a probability measure P:

$$\mathbb{E}_P(u) = \sum_{x \in \mathscr{X}} u(x) P(\{x\}) \text{ if } \mathscr{X} \text{ finite}$$

$$\mathbb{E}_P(u) = \int_{\mathbb{R}} u(x) \, \mathrm{d}P \text{ if } \mathscr{X} = \text{ real line}$$

Computing expectations

With imprecise probabilities

Expected values become imprecise → compute $[\underline{E}_{\mathscr{P}}(u), \overline{E}_{\mathscr{P}}(u)]$

- When $\mathscr{X}$ finite → efficient algorithms to compute them (Utkin & Augustin, 05)
- When $\mathscr{X} = \mathbb{R}$ → hard problem in general

→ start from simple representations → p-boxes

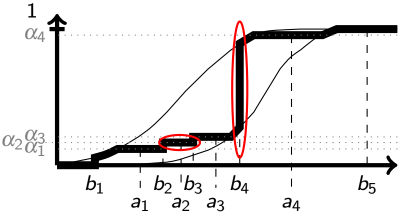
Computing expectations

With P-boxes (work with L. Utkin)

Given a (cont.) function u on $\mathbb{R}$ and a (classical) P-box $[\underline{F}, \overline{F}]$, find

- $\underline{\mathbb{E}}_{[\underline{F},\overline{F}]}(u) = \inf_{F \in [\underline{F},\overline{F}]} \int_{\mathbb{R}} u(x) \mathrm{d}F(x)$,
- $\overline{\mathbb{E}}_{[\underline{F},\overline{F}]}(u) = \sup_{F \in [\underline{F},\overline{F}]} \int_{\mathbb{R}} u(x) \mathrm{d}F(x)$.

$\rightarrow$ Find F inside $[\underline{F}, \overline{F}]$ reaching $[\underline{E}_{[\underline{F},\overline{F}]}(u), \overline{E}_{[\underline{F},\overline{F}]}(u)]$



F for which **lower expectation** is reached with a_i: local maxima, b_i: local minima

Conclusions

New results and new methodologies regarding the problems of

- Representing uncertainty: **Gen. P-boxes, relations with clouds.**
- Dealing with multiple sources: **MCS method on possibilities**
- Propagating uncertainties: **improving IRSN algorithm**
- Making decision under uncertainty: **computation of expectations on p-boxes**

Keeping in mind the three frameworks we chose to work in and that successful applications need:

1. Theoretically sound methods
2. Tractable methods

Next challenges and perspectives

Theoretical

As we have done for uncertainty representations, there is a need to provide a unified framework for the problems of

- Information fusion (e.g., study idempotent rules in random set theory)
- Independence modelling (e.g., how to model both source dependencies and variable dependencies)
- Conditioning our knowledge on some event (e.g., compare the notions of focusing on a particular subfamily, revising my information and learning from new information)

Next challenges and perspectives

Practical

- uncertainty representations:
  - build sound elicitation methods
- multiple sources treatment:
  - propose efficient algorithm to fuse information using maximal coherent subsets approach in general frames
- propagation
  - algorithmic work on the combined use of MC simulation + interval analysis + heuristic approaches
  - design efficient methods to simulate credal sets
- decision making
  - explore the computation of lower/upper expectations for other representations and for multiple variables

Next challenges and perspectives

Thesis
S. Destercke
Introduction
Context
Representation
Generic tools
State of art
Gen. p-boxes
Clouds
Mutliple sources
Inf. fusion
Rel. assess.
Application
Propagation
Independence
Prac. prop
Eval and Dec.
Conclusions & perspectives

Applications

With the help of SUNSET software, applications in perspective encompass:

- Evaluation of environmental impacts of radioactive wastes on river populations (few data available)
- Similar study as the one in BEMUSE programme to study/validate the results provided by computer codes simulating fires
- Expert system using MCS approach in dosimetry (monitoring of exposed workers)