Martingale-theoretic approach to discrete-time stochastic processes with imprecise probabilities

Gert de Cooman

Ghent University, SYSTeMS

gert.decooman@UGent.be
http://users.UGent.be/~gdcooma
gertekoo.wordpress.com

CiMi Toulouse
29 May 2015

My boon companions

FILIP HERMANS

ENRIQUE MIRANDA

JASPER DE BOCK

Jean Ville and martingales

Jean-André Ville, 1910–1989

The original definition of a martingale

DÉFINITION 1. — *Soit* $X_1, X_2, \ldots, X_n, \ldots$ *une suite de variables aléatoires, telle que les probabilités*

$$\Pr. \{ X_1 < x_1, X_2 < x_2, \ldots, X_n < x_n \} \qquad (n = 1, 2, 3, \ldots)$$

soient bien définies et que les X_i *ne puissent prendre que des valeurs finies.*

Soit une suite de fonctions $s_0, s_1(x_1), s_2(x_1, x_2), \ldots$ *non négatives telles que*

$$(14) \quad \begin{cases} s_0 = 1, \\ \mathfrak{M}_{x_1, x_2, \ldots, x_{n-1}} \{ s_n(x_1, x_2, \ldots, x_{n-1}, X_n) \} = s_{n-1}(x_1, x_2, \ldots, x_{n-1}), \end{cases}$$

où $\mathfrak{M}_X \{ Y \}$ *représente d'une manière générale la valeur moyenne conditionnelle de la variable* Y *quand on connaît la position du point aléatoire* X, *au sens indiqué par M. P. Lévy.*

Dans ces conditions, nous dirons que la suite $\{ s_n \}$ *définit une martingale ou un jeu équitable.*

Étude critique de la notion de collectif, 1939, p. 83

In a (perhaps) more modern notation

Ville's definition of a martingale

A martingale s is a sequence of real functions s_o, $s_1(X_1)$, $s_2(X_1,X_2)$, ... such that

1. $s_o = 1$;
2. $s_n(X_1,\dots,X_n) \geq 0$ for all $n \in \mathbb{N}$;
3. $E(s_{n+1}(x_1,\dots,x_n,X_{n+1})|x_1,\dots,x_n) = s_n(x_1,\dots,x_n)$ for all $n \in \mathbb{N}_0$ and all $x_1,\dots,x_n$.

It represents the outcome of a fair betting scheme, without borrowing (or bankruptcy).

The definition uses only the local models $E(\cdot|x_1,\dots,x_n)$ for X_{n+1}.

A few results

A precursor to Doob's martingale inequality

Let s be a martingale, then

$$P\left(\limsup_{n\to+\infty} s_n(X_1,\ldots,X_n) \geq \lambda\right) \leq \frac{1}{\lambda} \text{ for all } \lambda > 1.$$

A few results

A precursor to Doob's martingale inequality

Let s be a martingale, then

$$P\left(\limsup_{n\to+\infty} s_n(X_1,\ldots,X_n) \geq \lambda\right) \leq \frac{1}{\lambda} \text{ for all } \lambda > 1.$$

Ville's theorem

The convex collection of all (locally defined) martingales determines the probability P on the sample space Ω:

$$\begin{aligned} P(A) &= \sup\{\lambda \in \mathbb{R} \colon s \text{ martingale and } \limsup_{n\to+\infty} \lambda s_n(X_1,\ldots,X_n) \leq \mathbb{I}_A\} \\ &= \inf\{\lambda \in \mathbb{R} \colon s \text{ martingale and } \liminf_{n\to+\infty} \lambda s_n(X_1,\ldots,X_n) \geq \mathbb{I}_A\} \end{aligned}$$

Consequences of Ville's theorem

$$P(A) = \sup\{\lambda \in \mathbb{R} \colon s \text{ martingale and } \limsup_{n\to+\infty} \lambda s_n(X_1,\dots,X_n) \leq \mathbb{I}_A\}$$
$$= \inf\{\lambda \in \mathbb{R} \colon s \text{ martingale and } \liminf_{n\to+\infty} \lambda s_n(X_1,\dots,X_n) \geq \mathbb{I}_A\}$$

Null events

If there is a martingale s that converges to $+\infty$ on an event A, so

$$\lim_{n\to\infty} s_n(x_1,\dots,x_n) = +\infty \text{ for all } (x_1,\dots,x_n,\dots) \in A,$$

then $P(A) = 0$.

This suggest a 'constructive' method for proving almost sure results.

Consequences of Ville's theorem

$$P(A) = \sup\{\lambda \in \mathbb{R} \colon s \text{ martingale and } \limsup_{n\to+\infty} \lambda s_n(X_1,\dots,X_n) \leq \mathbb{I}_A\}$$
$$= \inf\{\lambda \in \mathbb{R} \colon s \text{ martingale and } \liminf_{n\to+\infty} \lambda s_n(X_1,\dots,X_n) \geq \mathbb{I}_A\}$$

Turning things around

Ville's theorem suggests that we could take a convex set of martingales as a primitive notion, and probabilities and expectations as a derived notion.

That we need an convex set of them, elucidates that martingales are examples of partial probability assessments.

Imprecise probabilities: dealing with partial probability assessments

Partial probability assessments

lower and/or upper bounds for the probabilities of a number of events, of the expectations of a number of random variables

Partial probability assessments

lower and/or upper bounds for the probabilities of a number of events, of the expectations of a number of random variables

Imprecise probability models

A partial assessment generally does not determine a probability measure uniquely, only a convex closed set of them.

Partial probability assessments

lower and/or upper bounds for the probabilities of a number of events, of the expectations of a number of random variables

Imprecise probability models

A partial assessment generally does not determine a probability measure uniquely, only a convex closed set of them.

IP Theory

systematic way of dealing with, representing, and making conservative inferences based on partial probability assessments

Why work with sets of desirable gambles?

Working with sets of desirable gambles:

- is simple, intuitive and elegant
- is more general and expressive than probability and expectation bounds
- gives a geometrical flavour to probabilistic inference
- shows that probabilistic inference and Bayes' Rule are 'logical' inference
- includes precise probability as one special case
- includes classical propositional logic as another special case
- avoids problems with conditioning on sets of probability zero

First steps: Williams (1977)

Available online at www.sciencedirect.com
ScienceDirect
International Journal of Approximate Reasoning
44 (2007) 366–383

ELSEVIER

INTERNATIONAL JOURNAL OF APPROXIMATE REASONING
www.elsevier.com/locate/ijar

Notes on conditional previsions ☆

P.M. Williams

Department of Informatics, The University of Sussex, Brighton BN1 9QH, UK

Received 15 December 2005; received in revised form 30 June 2006; accepted 31 July 2006
Available online 25 September 2006

Abstract

The personalist conception of probability is often explicated in terms of betting rates acceptable to an individual. A common approach, that of de Finetti for example, assumes that the individual is willing to take either side of the bet, so that the bet is "fair" from the individual's point of view. This can sometimes be unrealistic, and leads to difficulties in the case of conditional probabilities or previsions. An alternative conception is presented in which it is only assumed that the collection of acceptable bets forms a convex cone, rather than a linear space. This leads to the more general conception of an upper conditional prevision. The main concerns of the paper are with the extension of upper conditional previsions. The main result is that any upper conditional prevision is the upper envelope of a family of additive conditional previsions.

Keywords: Conditional prevision; Imprecise probabilities; Coherence; de Finetti

```
@ARTICLE{williams2007,
   author = {Williams, Peter M.},
   title = {Notes on conditional previsions},
   journal = {International Journal of Approximate Reasoning},
   year = 2007,
   volume = 44,
   pages = {366--383}
}
```

First steps: Walley (2000)

ELSEVIER International Journal of Approximate Reasoning 24 (2000) 125–148

INTERNATIONAL JOURNAL OF APPROXIMATE REASONING

www.elsevier.com/locate/ijar

Towards a unified theory of imprecise probability

Peter Walley

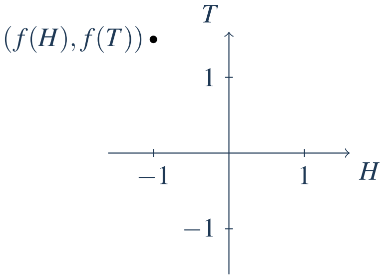
```
@ARTICLE{walley2000,
   author = {Walley, Peter},
   title = {Towards a unified theory of imprecise probability},
   journal = {International Journal of Approximate Reasoning},
   year = 2000,
   volume = 24,
   pages = {125--148}
}
```

Set of desirable gambles as a belief model

A subject is uncertain about the value that a variable X assumes in $\mathscr{X}$.

Gambles:

A gamble $f\colon \mathscr{X} \to \mathbb{R}$ is an uncertain reward whose value is $f(X)$.

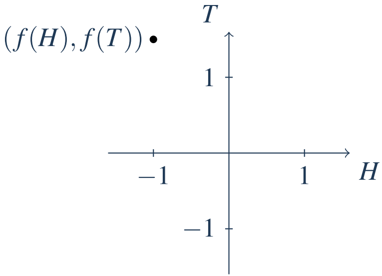


Set of desirable gambles as a belief model

A subject is uncertain about the value that a variable X assumes in $\mathscr{X}$.

Gambles:

A gamble $f\colon \mathscr{X} \to \mathbb{R}$ is an uncertain reward whose value is $f(X)$.



Set of desirable gambles:

$\mathscr{D} \subseteq \mathscr{G}(\mathscr{X})$ is a set of gambles that a subject strictly prefers to zero.

Coherence for a set of desirable gambles

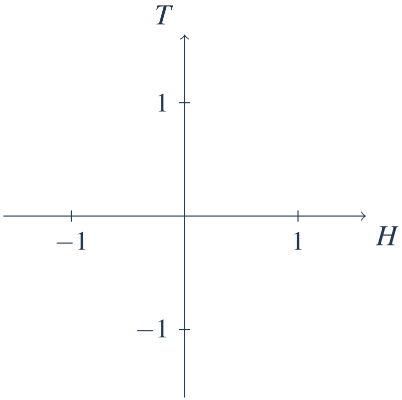
A set of desirable gambles $\mathscr{D}$ is called coherent if:

D1. if $f \leq 0$ then $f \notin \mathscr{D}$ [not desiring non-positivity]

D2. if $f > 0$ then $f \in \mathscr{D}$ [desiring partial gains]

D3. if $f, g \in \mathscr{D}$ then $f + g \in \mathscr{D}$ [addition]

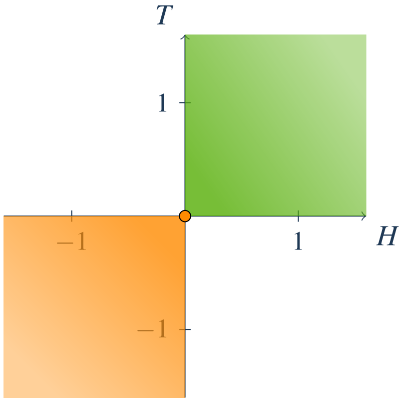
D4. if $f \in \mathscr{D}$ then $\lambda f \in \mathscr{D}$ for all real $\lambda > 0$ [scaling]



Coherence for a set of desirable gambles

A set of desirable gambles $\mathscr{D}$ is called coherent if:

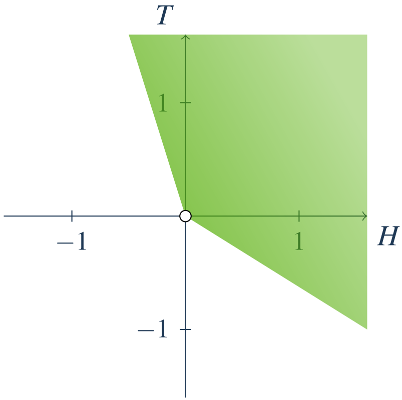
- D1. if $f \leq 0$ then $f \notin \mathscr{D}$ [not desiring non-positivity]
- D2. if $f > 0$ then $f \in \mathscr{D}$ [desiring partial gains]
- D3. if $f, g \in \mathscr{D}$ then $f + g \in \mathscr{D}$ [addition]
- D4. if $f \in \mathscr{D}$ then $\lambda f \in \mathscr{D}$ for all real $\lambda > 0$ [scaling]



Coherence for a set of desirable gambles

A set of desirable gambles $\mathscr{D}$ is called coherent if:

- D1. if $f \leq 0$ then $f \notin \mathscr{D}$ [not desiring non-positivity]
- D2. if $f > 0$ then $f \in \mathscr{D}$ [desiring partial gains]
- D3. if $f, g \in \mathscr{D}$ then $f + g \in \mathscr{D}$ [addition]
- D4. if $f \in \mathscr{D}$ then $\lambda f \in \mathscr{D}$ for all real $\lambda > 0$ [scaling]



Coherence for a set of desirable gambles

A set of desirable gambles $\mathscr{D}$ is called coherent if:

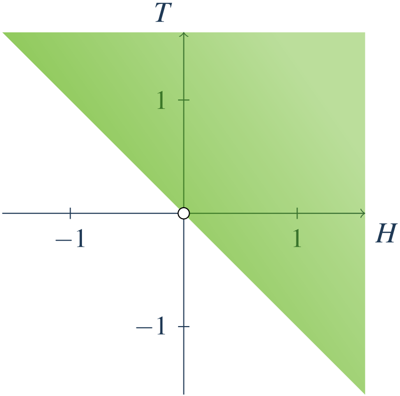
D1. if $f \leq 0$ then $f \notin \mathscr{D}$ [not desiring non-positivity]

D2. if $f > 0$ then $f \in \mathscr{D}$ [desiring partial gains]

D3. if $f, g \in \mathscr{D}$ then $f + g \in \mathscr{D}$ [addition]

D4. if $f \in \mathscr{D}$ then $\lambda f \in \mathscr{D}$ for all real $\lambda > 0$ [scaling]

Precise models correspond to the special case that the convex cones $\mathscr{D}$ are actually halfspaces!



Coherence for a set of desirable gambles

A set of desirable gambles $\mathscr{D}$ is called coherent if:

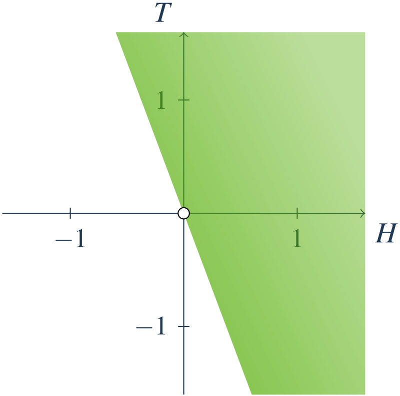
D1. if $f \leq 0$ then $f \notin \mathscr{D}$ [not desiring non-positivity]

D2. if $f > 0$ then $f \in \mathscr{D}$ [desiring partial gains]

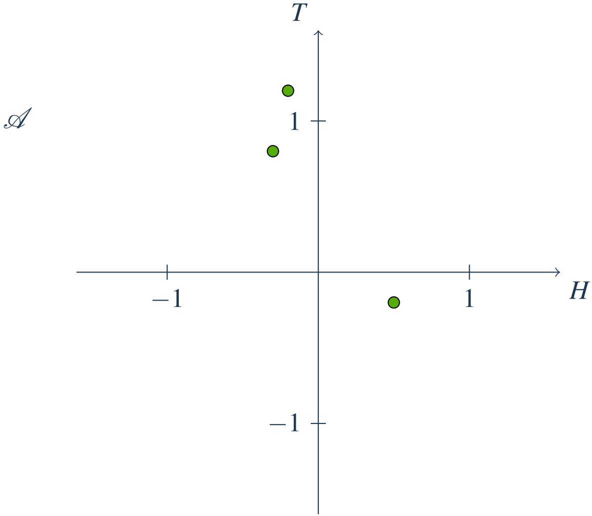
D3. if $f, g \in \mathscr{D}$ then $f + g \in \mathscr{D}$ [addition]

D4. if $f \in \mathscr{D}$ then $\lambda f \in \mathscr{D}$ for all real $\lambda > 0$ [scaling]

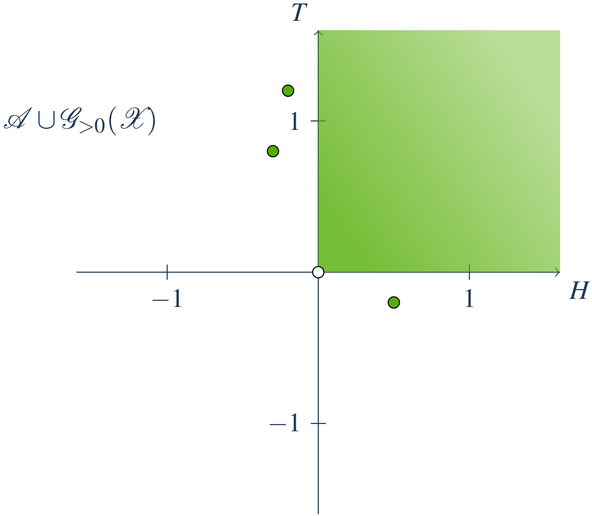
Precise models correspond to the special case that the convex cones $\mathscr{D}$ are actually halfspaces!



Inference: natural extension

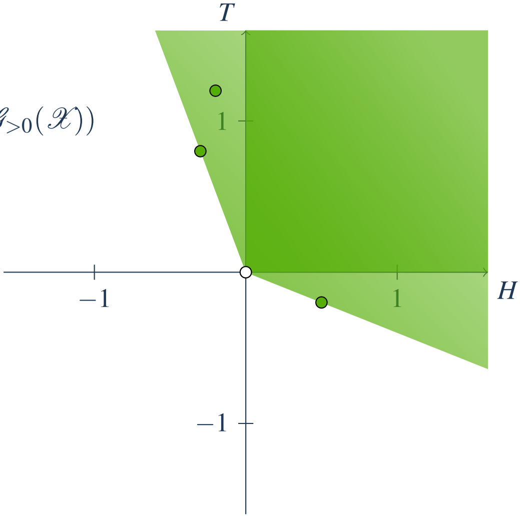


Inference: natural extension



Inference: natural extension

$$\mathscr{E}(\mathscr{A}) := \operatorname{posi}(\mathscr{A} \cup \mathscr{G}_{>0}(\mathscr{X}))$$



$$\operatorname{posi}(\mathscr{K}) := \left\{ \sum_{k=1}^{n} \lambda_k f_k : f_k \in \mathscr{K}, \lambda_k > 0, n > 0 \right\}$$

Inference: conditioning

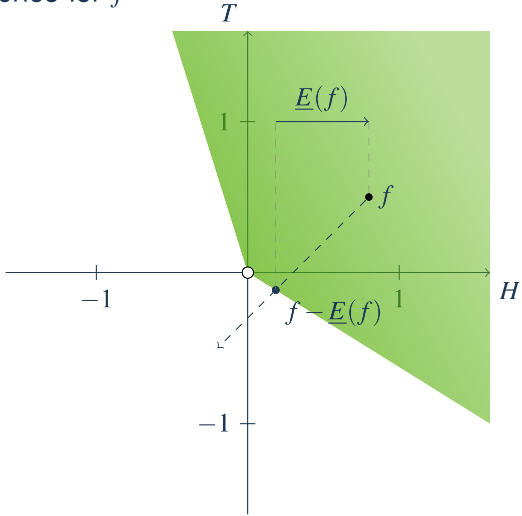
Additional information that $X \in A$ leads to a conditioned set of desirable gambles $\mathscr{D}\rfloor A$ on A:

$$f \in \mathscr{D}\rfloor A \Leftrightarrow f\mathbb{I}_A \in \mathscr{D}$$

Connection with lower expectations

$\underline{E}(f) = \sup\{\alpha \in \mathbb{R} \colon f - \alpha \in \mathscr{D}\}$

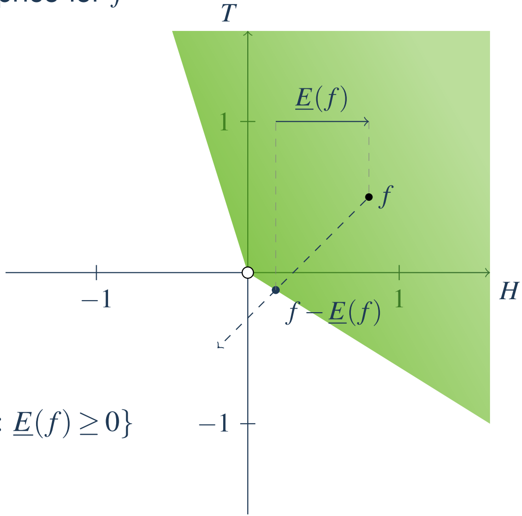
supremum buying price for f



Connection with lower expectations

$\underline{E}(f) = \sup\{\alpha \in \mathbb{R} \colon f - \alpha \in \mathscr{D}\}$

supremum buying price for f

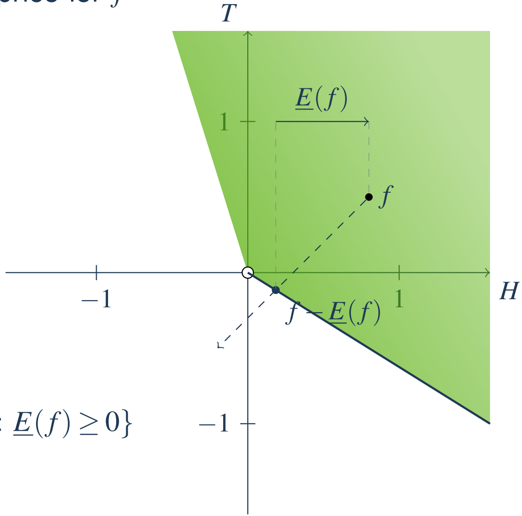


$\mathrm{cl}(\mathscr{D}) = \{f \in \mathscr{G}(\mathscr{X}) \colon \underline{E}(f) \geq 0\}$

Connection with lower expectations

$\underline{E}(f) = \sup\{\alpha \in \mathbb{R} \colon f - \alpha \in \mathscr{D}\}$

supremum buying price for f

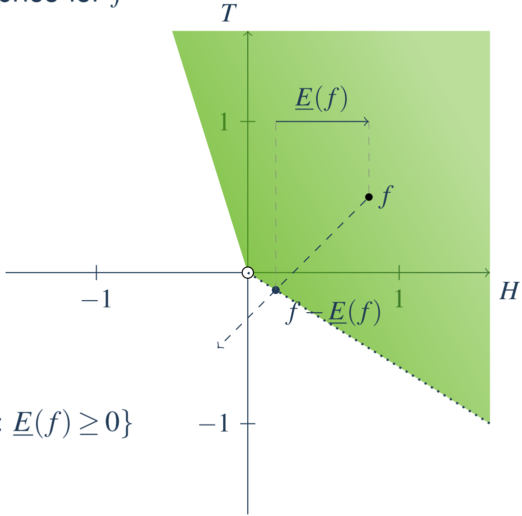


$\mathrm{cl}(\mathscr{D}) = \{f \in \mathscr{G}(\mathscr{X}) \colon \underline{E}(f) \geq 0\}$

Connection with lower expectations

$\underline{E}(f) = \sup\{\alpha \in \mathbb{R} \colon f - \alpha \in \mathscr{D}\}$

supremum buying price for f



$\mathrm{cl}(\mathscr{D}) = \{f \in \mathscr{G}(\mathscr{X}) \colon \underline{E}(f) \geq 0\}$

Conditional lower and upper expectations

Conditional lower expectation:

$$\underline{E}(f|A) = \sup\{\alpha \in \mathbb{R}\colon f - \alpha \in \mathscr{D}\rfloor A\}$$
$$= \sup\{\alpha \in \mathbb{R}\colon [f - \alpha]\mathbb{I}_A \in \mathscr{D}\}$$

Conditional upper expectation:

$$\overline{E}(f|A) = \inf\{\alpha \in \mathbb{R}\colon \alpha - f \in \mathscr{D}\rfloor A\}$$
$$= \inf\{\alpha \in \mathbb{R}\colon [\alpha - f]\mathbb{I}_A \in \mathscr{D}\}$$

Conjugacy:

$$\overline{E}(f|A) = \inf\{\alpha \in \mathbb{R}\colon [\alpha - f]\mathbb{I}_A \in \mathscr{D}\} = \inf\{-\alpha \in \mathbb{R}\colon [-\alpha - f]\mathbb{I}_A \in \mathscr{D}\}$$
$$= -\sup\{\alpha \in \mathbb{R}\colon [-f - \alpha]\mathbb{I}_A \in \mathscr{D}\}$$
$$= -\underline{E}(-f|A)$$

Recall: A few results

A precursor to Doob's martingale inequality

Let s be a martingale, then

$$P\left(\limsup_{n\to+\infty} s_n(X_1,\dots,X_n) \geq \lambda\right) \leq \frac{1}{\lambda} \text{ for all } \lambda > 1.$$

Ville's theorem

The collection of all (locally defined!) martingales determines the probability P on the sample space Ω:

$$\begin{aligned} P(A) &= \sup\{\lambda \in \mathbb{R} \colon s \text{ martingale and } \limsup_{n\to+\infty} \lambda s_n(X_1,\dots,X_n) \leq \mathbb{I}_A\} \\ &= \inf\{\lambda \in \mathbb{R} \colon s \text{ martingale and } \liminf_{n\to+\infty} \lambda s_n(X_1,\dots,X_n) \geq \mathbb{I}_A\} \end{aligned}$$

Back to (sub- and super)martingales

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Artificial Intelligence 172 (2008) 1400–1427

Artificial Intelligence

www.elsevier.com/locate/artint

Imprecise probability trees: Bridging two theories of imprecise probability

Gert de Cooman *, Filip Hermans

Ghent University, SYSTeMS Research Group, Technologiepark – Zwijnaarde 914, 9052 Zwijnaarde, Belgium

Received 30 March 2007; received in revised form 12 February 2008; accepted 3 March 2008

Available online 18 March 2008

Abstract

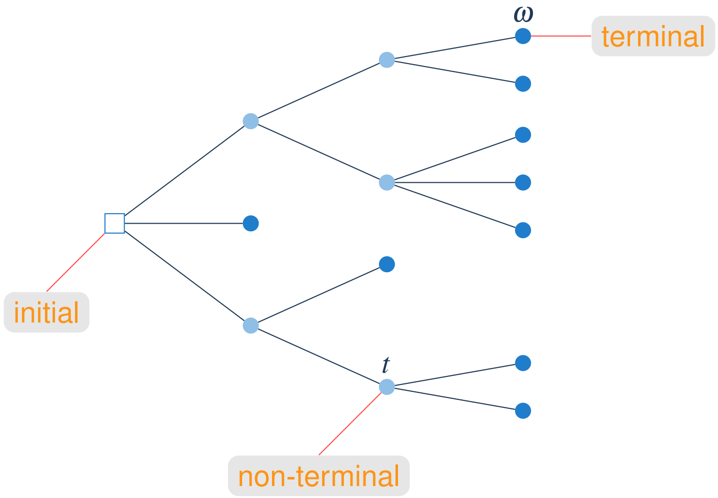
We give an overview of two approaches to probability theory where lower and upper probabilities, rather than probabilities, are used: Walley's behavioural theory of imprecise probabilities, and Shafer and Vovk's game-theoretic account of probability. We show that the two theories are more closely related than would be suspected at first sight, and we establish a correspondence between them that (i) has an interesting interpretation, and (ii) allows us to freely import results from one theory into the other. Our approach leads to an account of probability trees and random processes in the framework of Walley's theory. We indicate how our results can be used to reduce the computational complexity of dealing with imprecision in probability trees, and we prove an interesting and quite general version of the weak law of large numbers.

Keywords: Game-theoretic probability; Imprecise probabilities; Coherence; Conglomerability; Event tree; Probability tree; Imprecise probability tree; Lower prevision; Immediate prediction; Prequential Principle; Law of large numbers; Hoeffding's inequality; Markov chain; Random process

```
@ARTICLE{cooman2008,
   author = {de Cooman, Gert and Hermans, Filip},
   title = {Imprecise probability trees: Bridging two theories of imprecise probability},
   journal = {Artificial Intelligence},
   year = {2008},
   volume = {172},
   pages = {1400–1427},
   number = {11},
   doi = {10.1016/j.artint.2008.03.001}
}
```

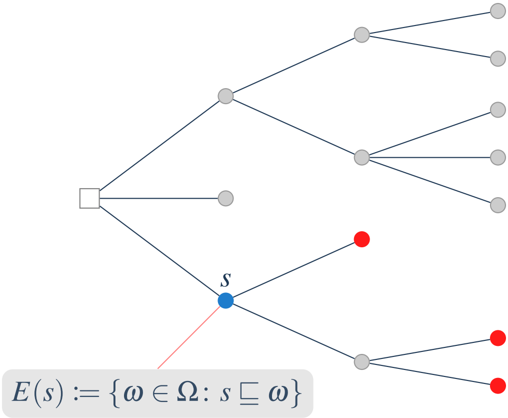
An event tree and its situations

Situations are nodes in the event tree, and the sample space Ω is the set of all terminal situations:

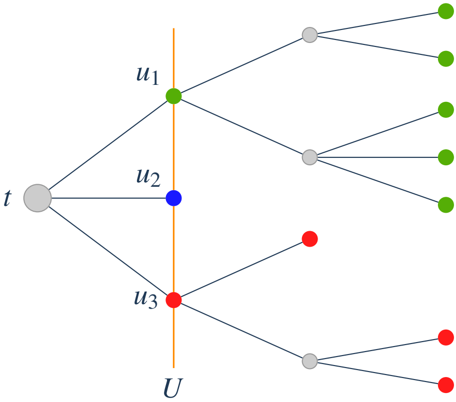


Events

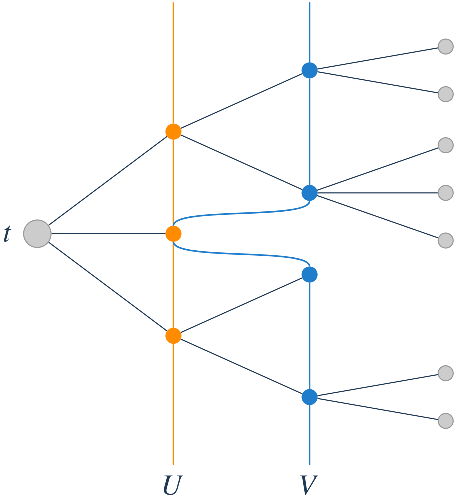
An event A is a subset of the sample space Ω:



Cuts of a situation



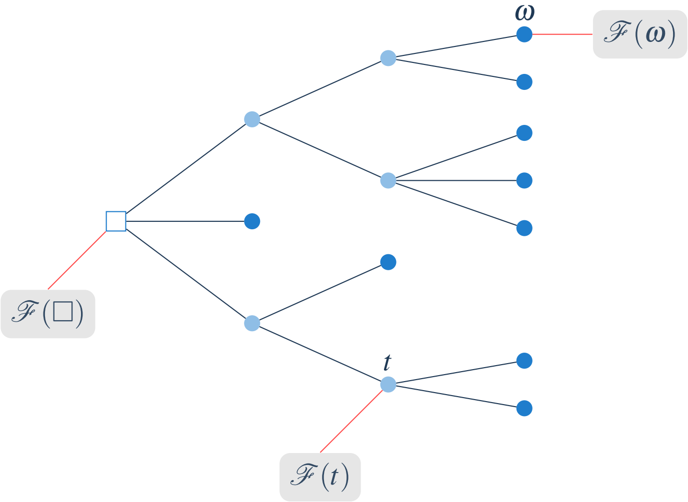
A directed set of cuts



U precedes V: $U \sqsubseteq V$

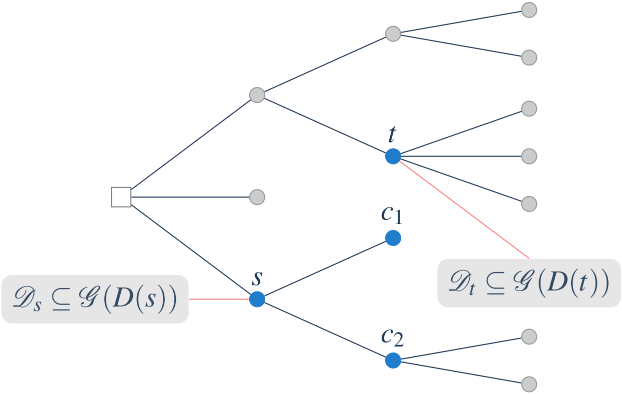
A process

A process $\mathscr{F}$ is a way of populating the event tree with real numbers:



Local, or immediate prediction, models

In each non-terminal situation s, Subject has a belief model $\mathscr{D}_s$, satisfying D1–D4. This also leads to local lower expectations $\underline{Q}(\cdot|s)$

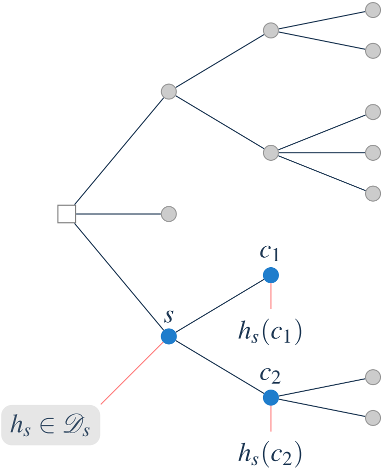


$D(s) = \{c_1, c_2\}$ is the set of daughters of s.

From a local to a global model

How to combine the local information into a coherent global model:

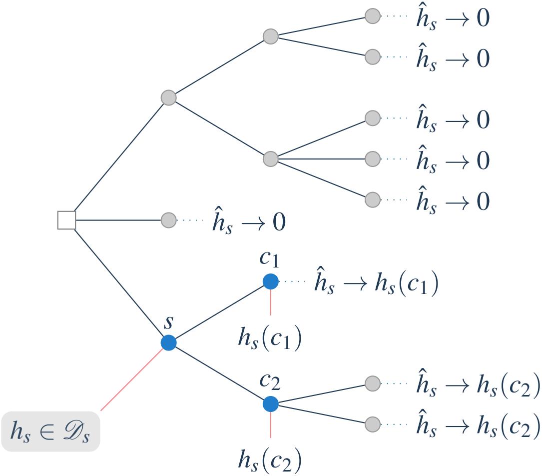
Subject accepts which gambles f on the entire sample space Ω?



From a local to a global model

How to combine the local information into a coherent global model:

Subject accepts which gambles f on the entire sample space Ω?



Natural extension

So the Subject accepts all gambles in the set:

$$\mathscr{A} \coloneqq \{\hat{h}_s \colon h_s \in \mathscr{D}_s \text{ and } s \text{ non-terminal}\}.$$

We found an expression for the natural extension $\mathscr{E}(\mathscr{A})$ of $\mathscr{A}$:

Natural extension

So the Subject accepts all gambles in the set:

$$\mathscr{A} \coloneqq \{\hat{h}_s \colon h_s \in \mathscr{D}_s \text{ and } s \text{ non-terminal}\}.$$

We found an expression for the natural extension $\mathscr{E}(\mathscr{A})$ of $\mathscr{A}$:

this is the smallest subset $\mathscr{D}$ of $\mathscr{G}(\Omega)$ that includes $\mathscr{A}$, is coherent (satisfies D1–D4) and satisfies

D5. bounded cut conglomerability: for all bounded cuts U:

$$(\forall u \in U)(\mathbb{I}_{E(u)} f \in \mathscr{D} \cup \{0\}) \Rightarrow f \in \mathscr{D} \cup \{0\}.$$

D6. bounded cut continuity: for any real process $\mathscr{F}$ such that $\limsup_{U \text{ bounded}} \mathscr{F}_U \in \mathscr{G}(\Omega)$, and such that $\mathscr{F}_V - \mathscr{F}_U \in \mathscr{D} \cup \{0\}$ for all bounded cuts $U \sqsubseteq V$: $\limsup_{U \text{ bounded}} \mathscr{F}_U - \mathscr{F}(\square) \in \mathscr{D} \cup \{0\}$.

Observe that $\limsup_{U \text{ bounded}} \mathscr{F}_U = \limsup \mathscr{F}$.

Sub- and supermartingales

We can use the local lower expectations $\underline{Q}(\cdot|s)$ associated with the local modes $\mathscr{D}_s$ to define sub- and supermartingales:

A submartingale $\underline{\mathscr{M}}$

is a real process such that in all non-terminal situations s:

$$\underline{Q}(\underline{\mathscr{M}}(s\,\cdot)|s) \geq \underline{\mathscr{M}}(s)$$

or in other words $\Delta\underline{\mathscr{M}}(s) = \underline{\mathscr{M}}(s\,\cdot) - \underline{\mathscr{M}}(s) \in \mathrm{cl}(\mathscr{D}_s)$.

A supermartingale $\overline{\mathscr{M}}$

is a real process such that in all non-terminal situations s:

$$\overline{Q}(\overline{\mathscr{M}}(s\,\cdot)|s) \leq \overline{\mathscr{M}}(s).$$

Natural extension, sub- and supermartingales

Conditional lower and upper expectations:

$$\underline{E}(f|s) := \sup\{\alpha \in \mathbb{R} \colon [f-\alpha]\mathbb{I}_{E(s)} \in \mathscr{E}(\mathscr{A})\}$$
$$= \sup\{\underline{\mathscr{M}}(s) \colon \liminf\sup\underline{\mathscr{M}} \leq f \text{ on } E(s)\}$$

$$\overline{E}(f|s) := \inf\{\alpha \in \mathbb{R} \colon [\alpha-f]\mathbb{I}_{E(s)} \in \mathscr{E}(\mathscr{A})\}$$
$$= \inf\{\overline{\mathscr{M}}(s) \colon \lim\inf\overline{\mathscr{M}} \geq f \text{ on } E(s)\}$$

Natural extension, sub- and supermartingales

Conditional lower and upper expectations:

$$\underline{E}(f|s) := \sup\{\alpha \in \mathbb{R} \colon [f-\alpha]\mathbb{I}_{E(s)} \in \mathscr{E}(\mathscr{A})\}$$
$$= \sup\{\underline{\mathscr{M}}(s) \colon \limsup \underline{\mathscr{M}} \leq f \text{ on } E(s)\}$$

$$\overline{E}(f|s) := \inf\{\alpha \in \mathbb{R} \colon [\alpha-f]\mathbb{I}_{E(s)} \in \mathscr{E}(\mathscr{A})\}$$
$$= \inf\{\overline{\mathscr{M}}(s) \colon \liminf \overline{\mathscr{M}} \geq f \text{ on } E(s)\}$$

Recall Ville's Theorem

$$P(A) = \sup\{\lambda \in \mathbb{R} \colon s \text{ martingale and } \limsup_{n\to+\infty} \lambda s_n(X_1,\dots,X_n) \leq \mathbb{I}_A\}$$
$$= \inf\{\lambda \in \mathbb{R} \colon s \text{ martingale and } \liminf_{n\to+\infty} \lambda s_n(X_1,\dots,X_n) \geq \mathbb{I}_A\}$$

Test supermartingales and strictly null events

A test supermartingale $\mathscr{T}$

is a non-negative supermartingale with $\mathscr{T}(\square) = 1$.
(Very close to Ville's definition of a martingale.)

An event A is strictly null

if there is some test supermartingale $\mathscr{T}$ that converges to $+\infty$ on A:

$$\lim \mathscr{T}(\omega) = \lim_{n\to\infty} \mathscr{T}(\omega^n) = +\infty \text{ for all } \omega \in A.$$

If A is strictly null then

$$\overline{P}(A) = \overline{E}(\mathbb{I}_A) = \inf\{\overline{\mathscr{M}}(\square)\colon \liminf \overline{\mathscr{M}} \geq \mathbb{I}_A\} = 0.$$

A few basic limit results

Supermartingale convergence theorem [Shafer and Vovk, 2001]

A supermartingale $\overline{\mathscr{M}}$ that is bounded below converges strictly almost surely to a real number:

$$\liminf \overline{\mathscr{M}}(\omega) = \limsup \overline{\mathscr{M}}(\omega) \in \mathbb{R} \text{ strictly almost surely.}$$

A few basic limit results

Strong law of large numbers for submartingale differences [De Cooman and De Bock, 2013]

Consider any submartingale $\underline{\mathscr{M}}$ such that its difference process

$$\Delta\underline{\mathscr{M}}(s) = \underline{\mathscr{M}}(s\,\cdot) - \underline{\mathscr{M}}(s) \in \mathscr{G}(D(s)) \text{ for all non-terminal } s$$

is uniformly bounded. Then $\liminf \langle\underline{\mathscr{M}}\rangle \geq 0$ strictly almost surely, where

$$\langle\underline{\mathscr{M}}\rangle(\omega^n) = \frac{1}{n}\underline{\mathscr{M}}(\omega^n) \text{ for all } \omega \in \Omega \text{ and } n \in \mathbb{N}$$

A few basic limit results

Lévy's zero–one law [Shafer, Vovk and Takemura, 2012]

For any bounded real gamble f on Ω:

$$\limsup_{n\to+\infty} \underline{E}(f|\omega^n) \leq f(\omega) \leq \liminf_{n\to+\infty} \overline{E}(f|\omega^n) \text{ strictly almost surely.}$$

Imprecise Markov chains

Probability in the Engineering and Informational Sciences, **23**, 2009, 597–635. Printed in the U.S.A.
doi:10.1017/S0269964809990039

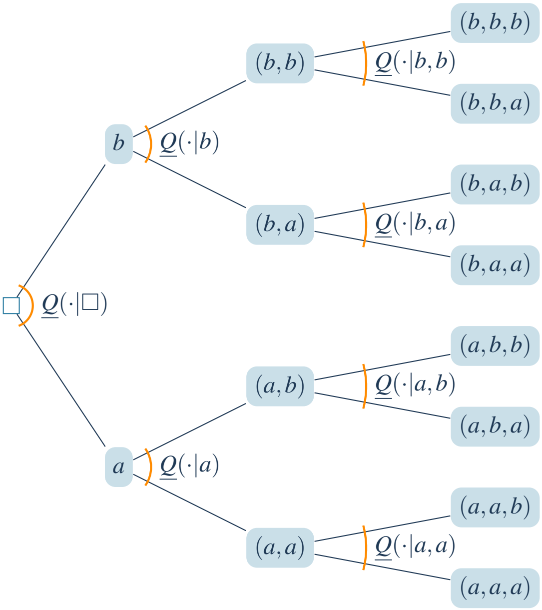
IMPRECISE MARKOV CHAINS AND THEIR LIMIT BEHAVIOR

Gert de Cooman, Filip Hermans, and Erik Quaeghebeur
SYSTeMS Research Group
Ghent University
Technologiepark–Zwijnaarde 914, 9052 Zwijnaarde, Belgium
E-mail: {gert.decooman, filip.hermans, erik.quaeghebeur}@ugent.be

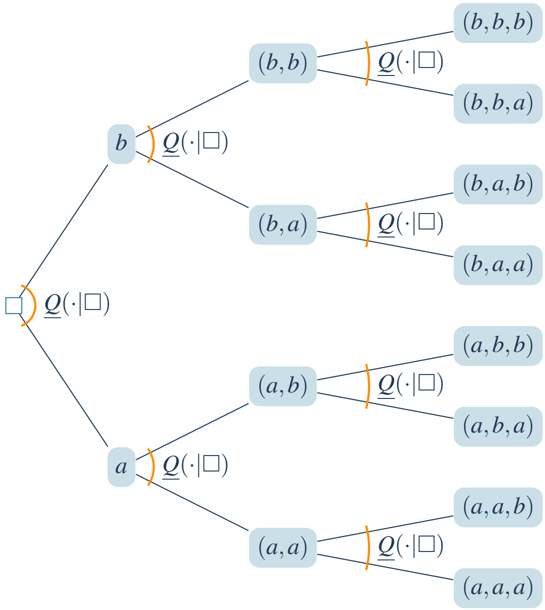
When the initial and transition probabilities of a finite Markov chain in discrete time are not well known, we should perform a sensitivity analysis. This can be done by considering as basic uncertainty models the so-called *credal sets* that these probabilities are known or believed to belong to and by allowing the probabilities to vary over such sets. This leads to the definition of an *imprecise Markov chain*. We show that the time evolution of such a system can be studied very efficiently using so-called *lower* and *upper expectations*, which are equivalent mathematical representations of credal sets. We also study how the inferred credal set about the state at time n evolves as $n \to \infty$: under quite unrestrictive conditions, it converges to a uniquely invariant credal set, regardless of the credal set given for the initial state. This leads to a non-trivial generalization of the classical Perron–Frobenius theorem to imprecise Markov chains.

```
@ARTICLE{cooman2009,
  author = {{d}e Cooman, Gert and Hermans, Filip and Quaegehebeur, Erik},
  title = {Imprecise {M}arkov chains and their limit behaviour},
  journal = {Probability in the Engineering and Informational Sciences},
  year = 2009,
  volume = 23,
  pages = {597--635},
  doi = {10.1017/S0269964809990039}
}
```

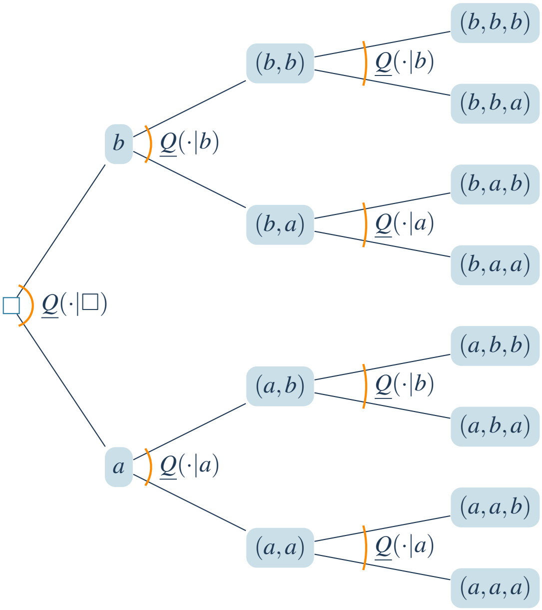
A simple discrete-time finite-state stochastic process



An imprecise IID model



An imprecise Markov chain



Lower transition operators

The lower expectation $\underline{E}_n$ for the state X_n at time n:

$$\underline{E}_n(f) = \underline{E}(f(X_n)) \text{ for all } f \in \mathscr{G}(\mathscr{X}).$$

Lower transition operators

The lower expectation $\underline{E}_n$ for the state X_n at time n:

$$\underline{E}_n(f) = \underline{E}(f(X_n)) \text{ for all } f \in \mathscr{G}(\mathscr{X}).$$

Lower transition operator

It follows from the SVV-formulas for lower expectations in an imprecise Markov tree that:

$$\underline{E}_n(f) = \underline{E}_1(\underline{\mathrm{T}}^{n-1} f) \text{ for all } f \in \mathscr{G}(\mathscr{X}).$$

where $\underline{\mathrm{T}} \colon \mathscr{G}(\mathscr{X}) \to \mathscr{G}(\mathscr{X})$ is defined by

$$\underline{\mathrm{T}} f(x) = \underline{Q}(f|x) \text{ for all } x \in \mathscr{X}.$$

Compare this with precise case: $p_n = p_1 M^{n-1}$.

Stationarity and ergodicity

The imprecise Markov chain is Perron–Frobenius-like if for all marginal models $\underline{E}_1$ and all f:

$$\underline{E}_n(f) = \underline{E}_1(\underline{\mathrm{T}}^{n-1} f) \to \underline{E}_\infty(f).$$

and if $\underline{E}_1 = \underline{E}_\infty$ then $\underline{E}_n = \underline{E}_\infty$, and the imprecise Markov chain is stationary.

Stationarity and ergodicity

The imprecise Markov chain is Perron–Frobenius-like if for all marginal models $\underline{E}_1$ and all f:

$$\underline{E}_n(f) = \underline{E}_1(\underline{\mathrm{T}}^{n-1} f) \to \underline{E}_\infty(f).$$

and if $\underline{E}_1 = \underline{E}_\infty$ then $\underline{E}_n = \underline{E}_\infty$, and the imprecise Markov chain is stationary.

In any Perron–Frobenius-like imprecise Markov chain:

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} \underline{E}_n(f) = \underline{E}_\infty(f)$$

and

$$\underline{E}_\infty(f) \leq \liminf_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k) \leq \limsup_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k) \leq \overline{E}_\infty(f)$$

strictly almost surely.

What’s next?

So much still to be done:

- mathematical foundations: continuous time, alternative formulae, other definitions of sub- and supermartingales, …
- allowing for robust modelling in stochastic processes
- other processes than Markov chains
- applications in finance and economics
- foundations of uncertain inference and time: dynamic coherence
- …