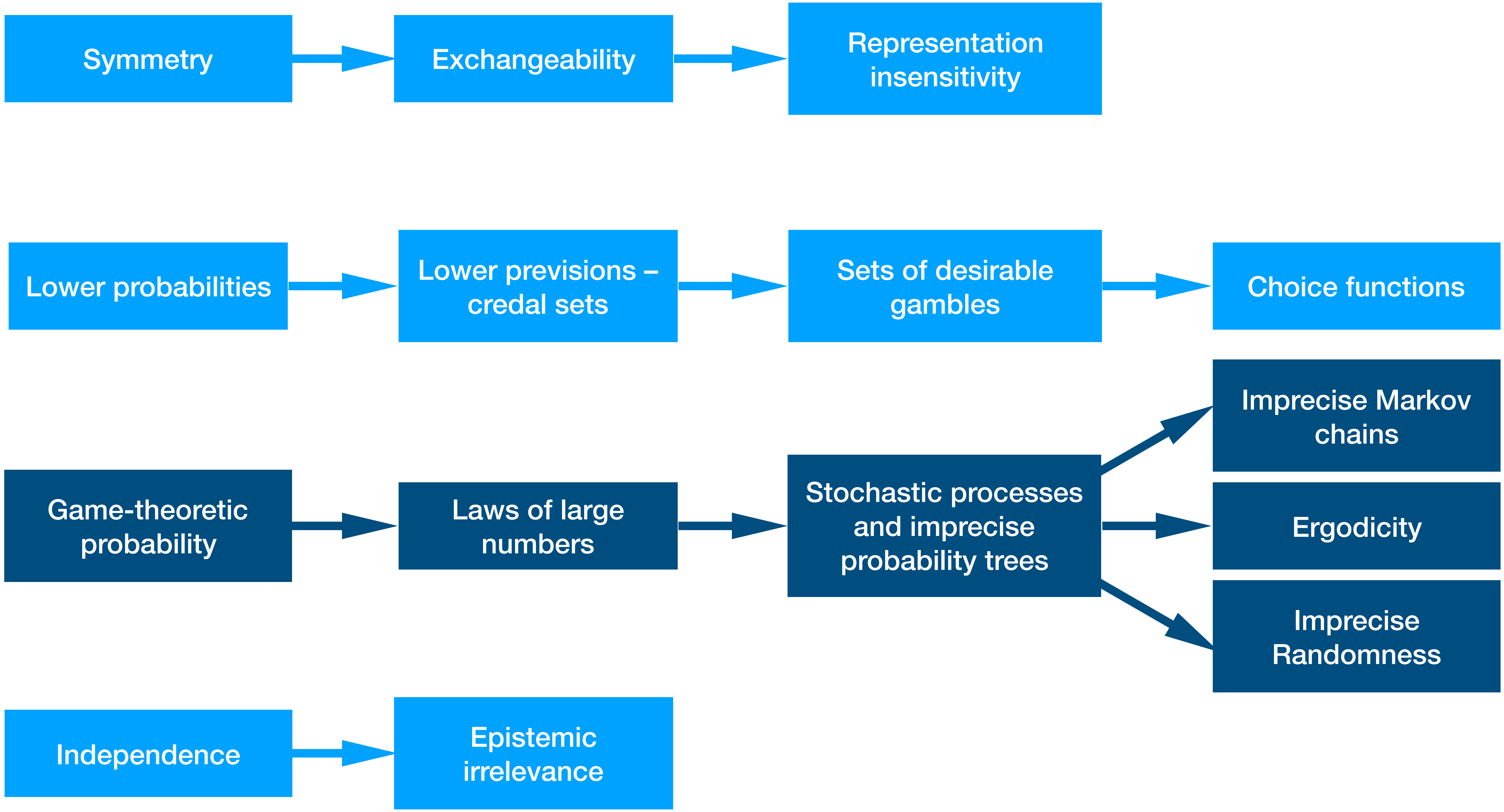
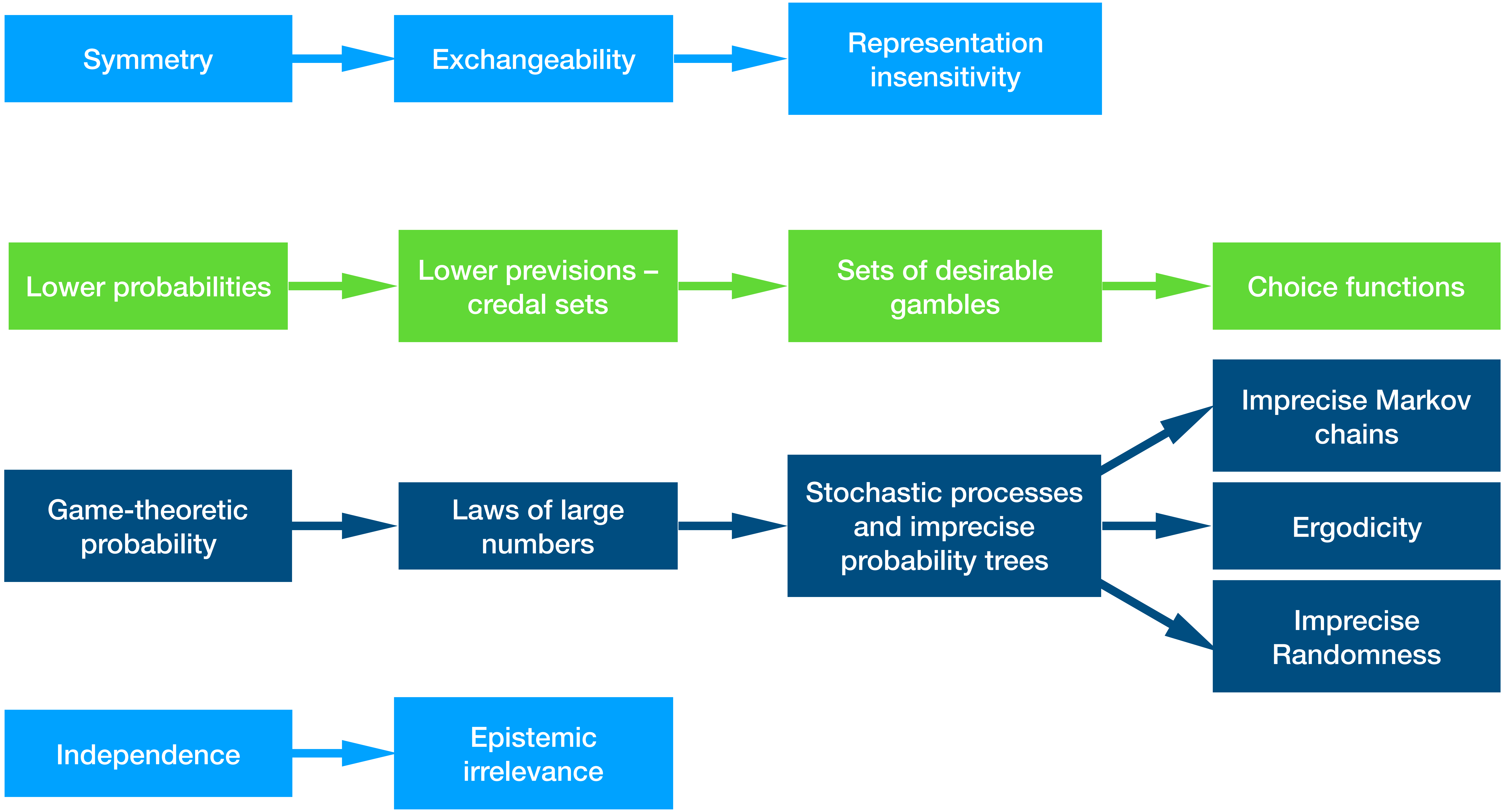
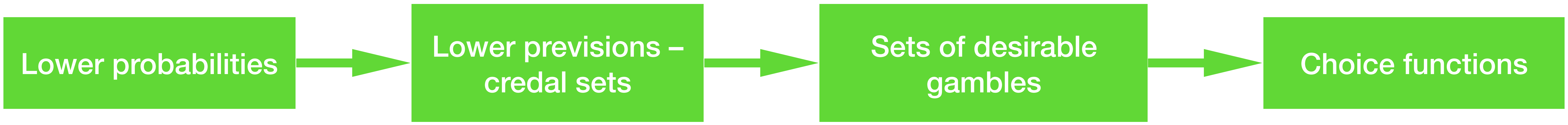
Imprecision in stochastic processes: combining IP and GTP

Gert de Cooman
Foundations Lab – Ghent University
Update@ISIPTA
4 July 2019


Symmetry
Exchangeability
Representation insensitivity
Lower probabilities
Lower previsions – credal sets
Sets of desirable gambles
Choice functions
Imprecise Markov chains
Game-theoretic probability
Laws of large numbers
Stochastic processes and imprecise probability trees
Ergodicity
Imprecise Randomness
Independence
Epistemic irrelevance

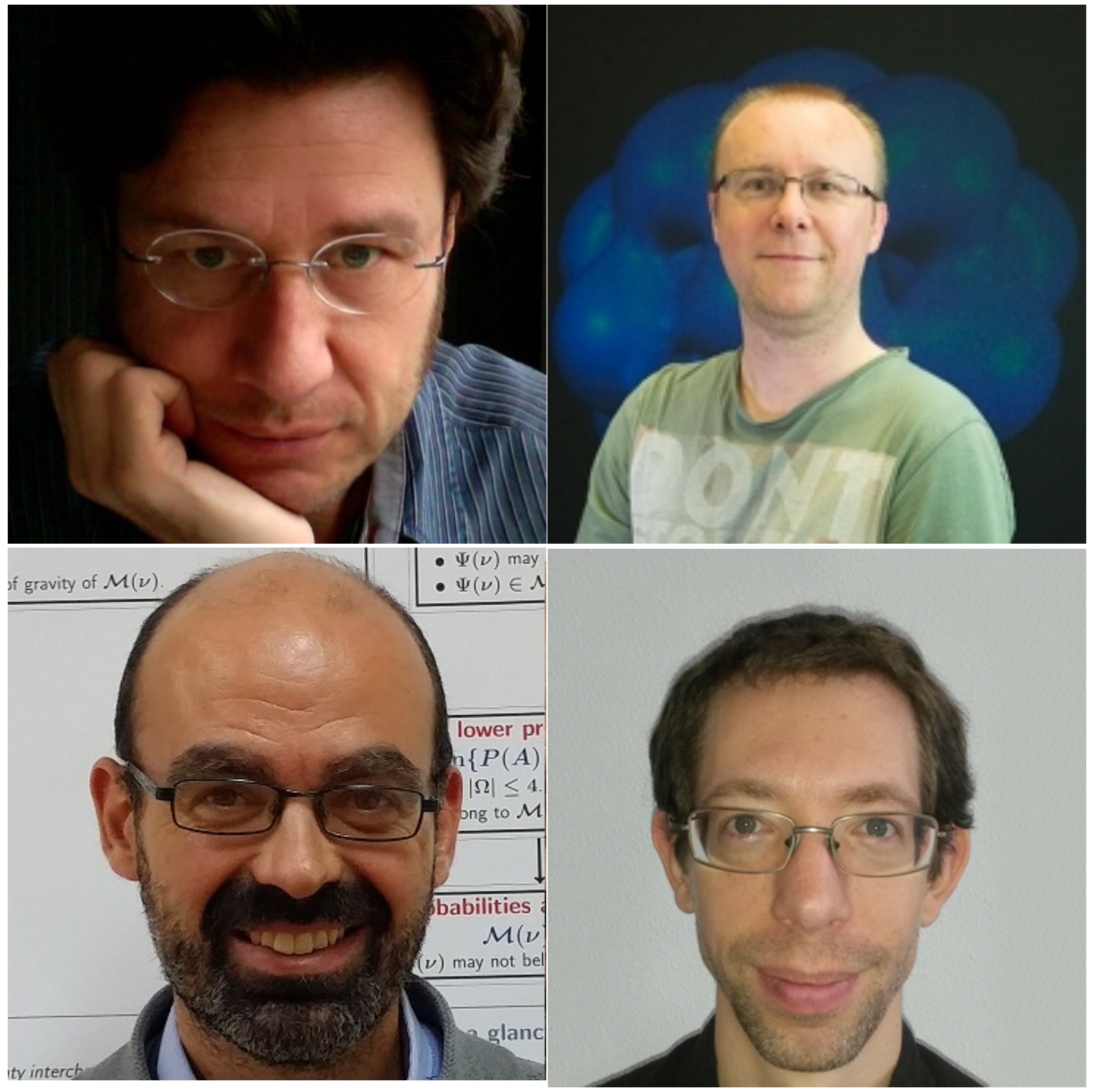

Symmetry
Exchangeability
Representation insensitivity
Lower probabilities
Lower previsions – credal sets
Sets of desirable gambles
Choice functions
Imprecise Markov chains
Game-theoretic probability
Laws of large numbers
Stochastic processes and imprecise probability trees
Ergodicity
Imprecise Randomness
Independence
Epistemic irrelevance


Lower probabilities
Lower previsions – credal sets
Sets of desirable gambles
Choice functions

Lower probabilities → Lower previsions – credal sets → Sets of desirable gambles → Choice functions

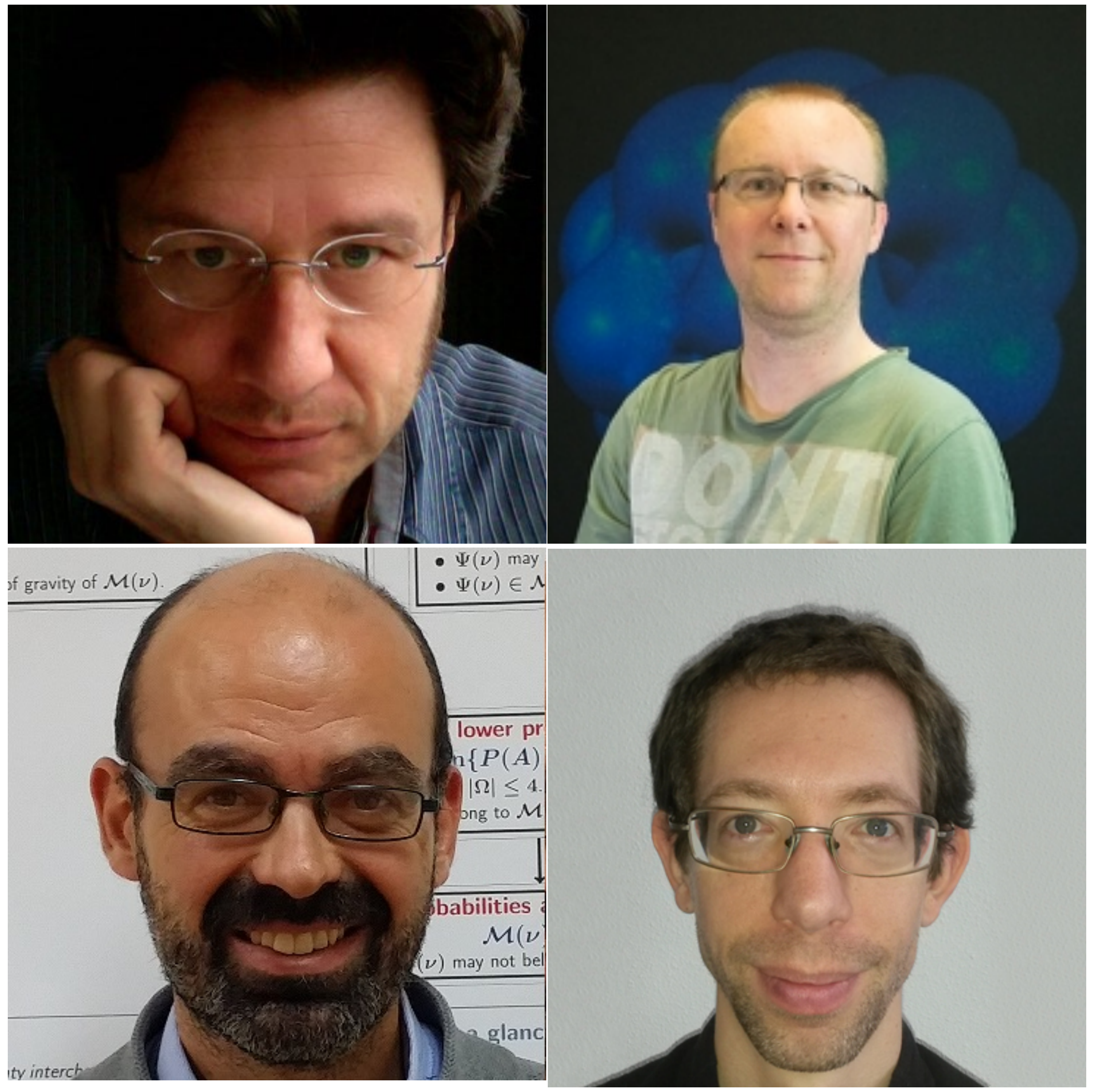
$$\underline{P}(A)$$

Lower probabilities → Lower previsions – credal sets → Sets of desirable gambles → Choice functions

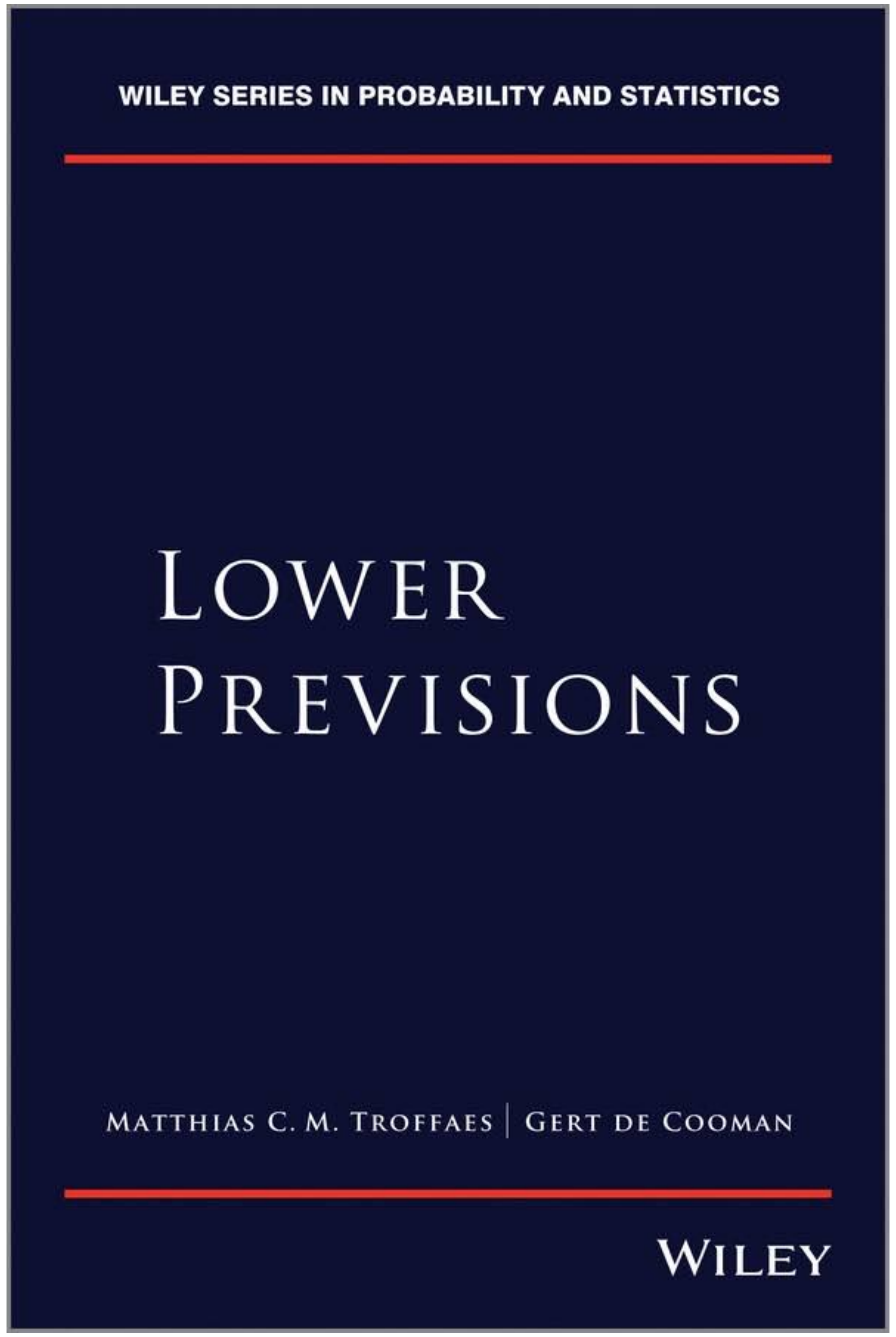


$$\underline{P}(f)$$

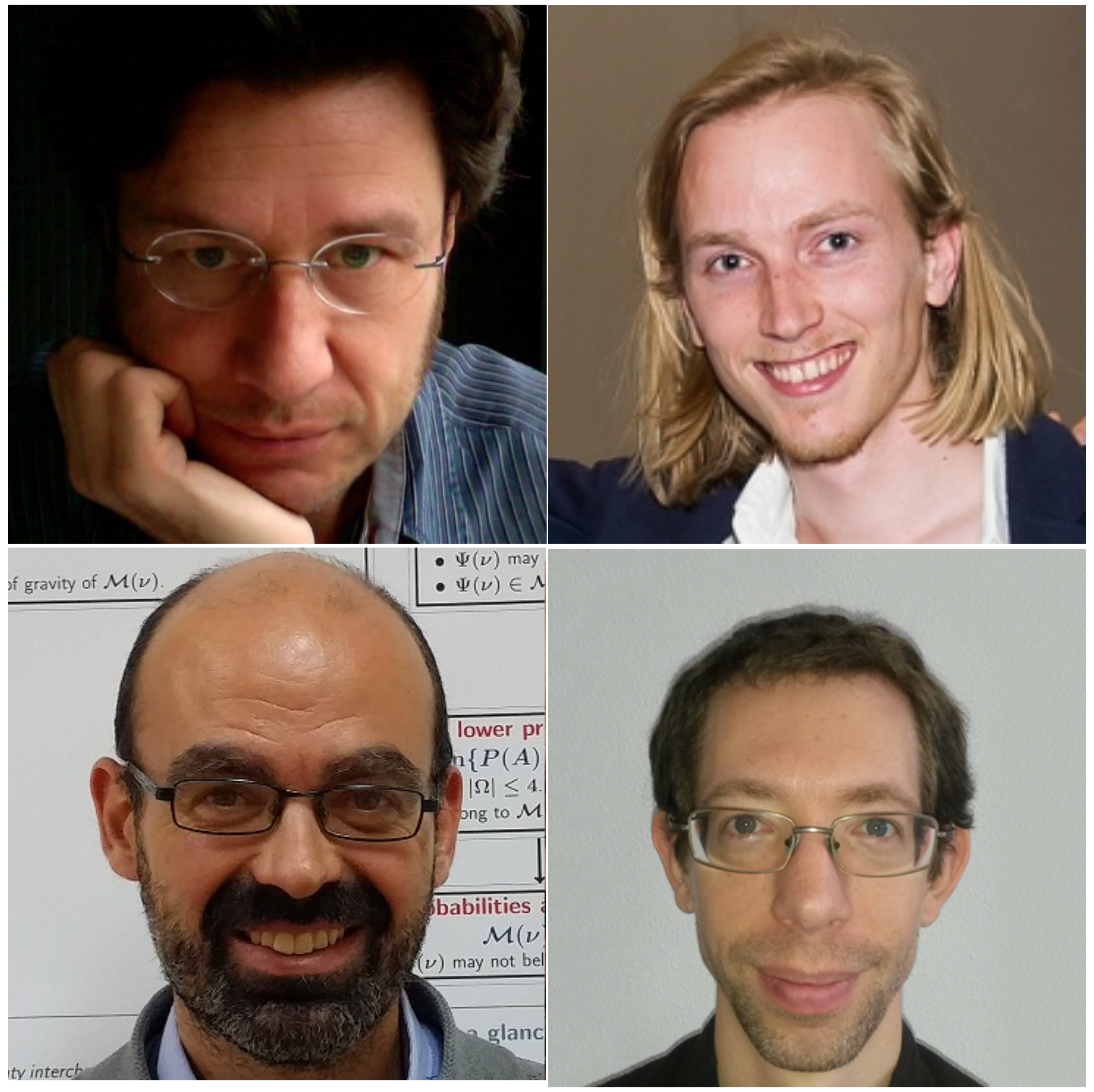
Lower probabilities → Lower previsions – credal sets → Sets of desirable gambles → Choice functions



$$\underline{P}(f)$$

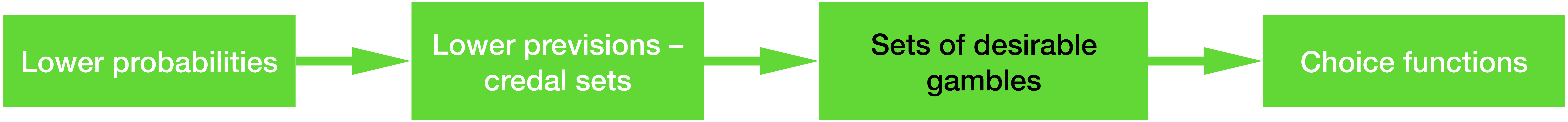


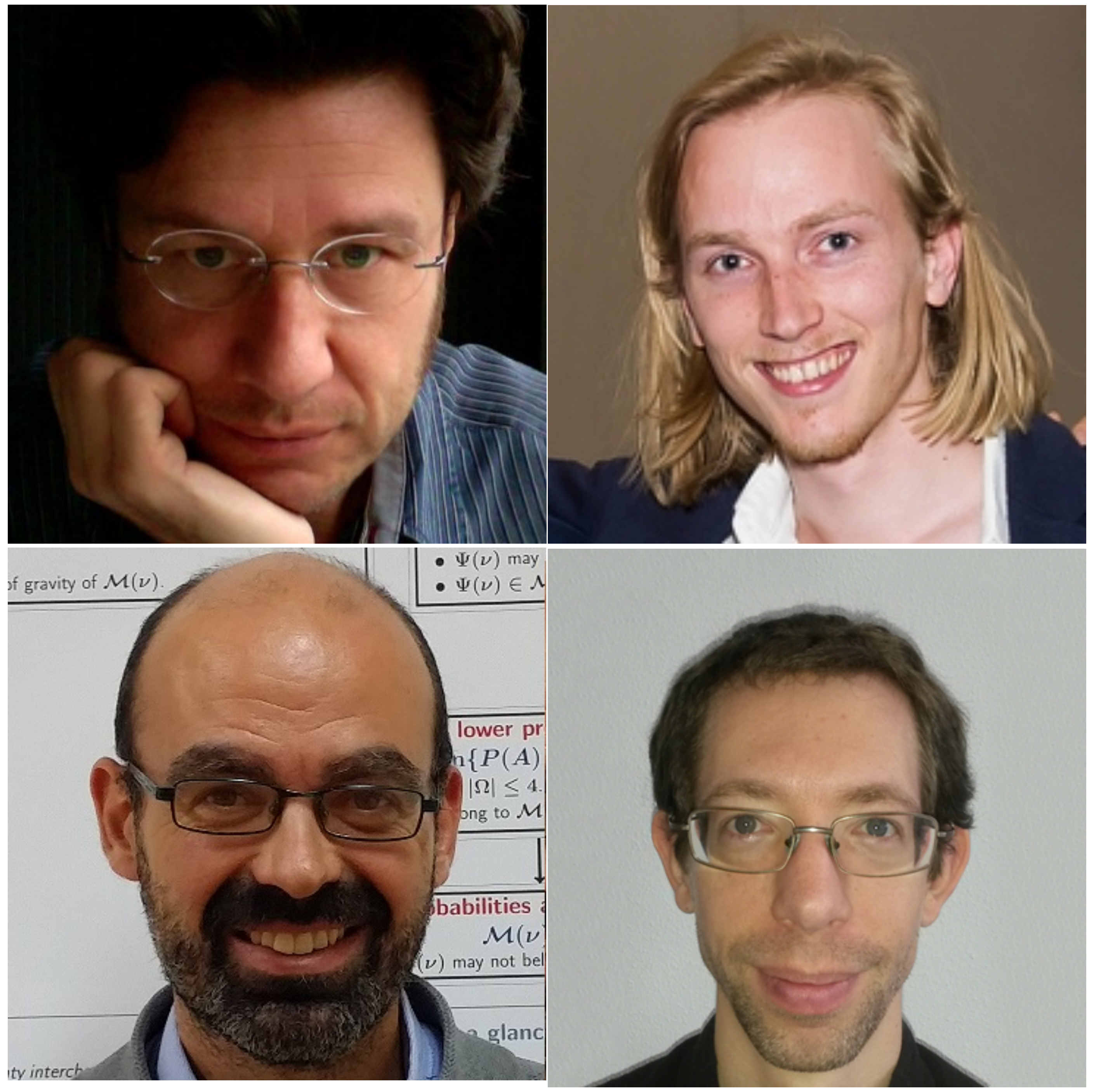
Lower probabilities → Lower previsions – credal sets → Sets of desirable gambles → Choice functions



$$\mathscr{D}$$

$$\underline{P}(f) = \sup\{\alpha : f - \alpha \in \mathscr{D}\}$$





$\mathcal{D}$

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

Received 1 September 1999; accepted 1 December 1999

Abstract

Coherent upper and lower probabilities, Choquet capacities of order 2, belief functions and possibility measures are amongst the most popular mathematical models for uncertainty and partial ignorance. Examples are given to show that these models are not sufficiently general to represent some common types of uncertainty. In particular, they are not sufficiently informative about expectations and conditional probabilities. Coherent lower previsions and sets of probability measures are considerably more general, but they may not be sufficiently informative for some purposes. Two other models for uncertainty, which involve partial preference orderings and sets of desirable gambles, are discussed. These are more informative and more general than the previous models, and they may provide a suitable mathematical foundation for a unified theory of imprecise probability. © 2000 Elsevier Science Inc. All rights reserved.

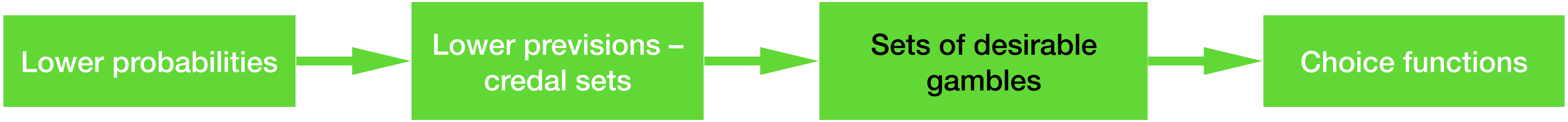
Keywords: Choquet capacity; Coherence; Comparative probability; Credal sets; Desirable gambles; Foundations of probability; Interval-valued probability; Lower prevision; Lower probability; Partial preference ordering; Uncertainty measures

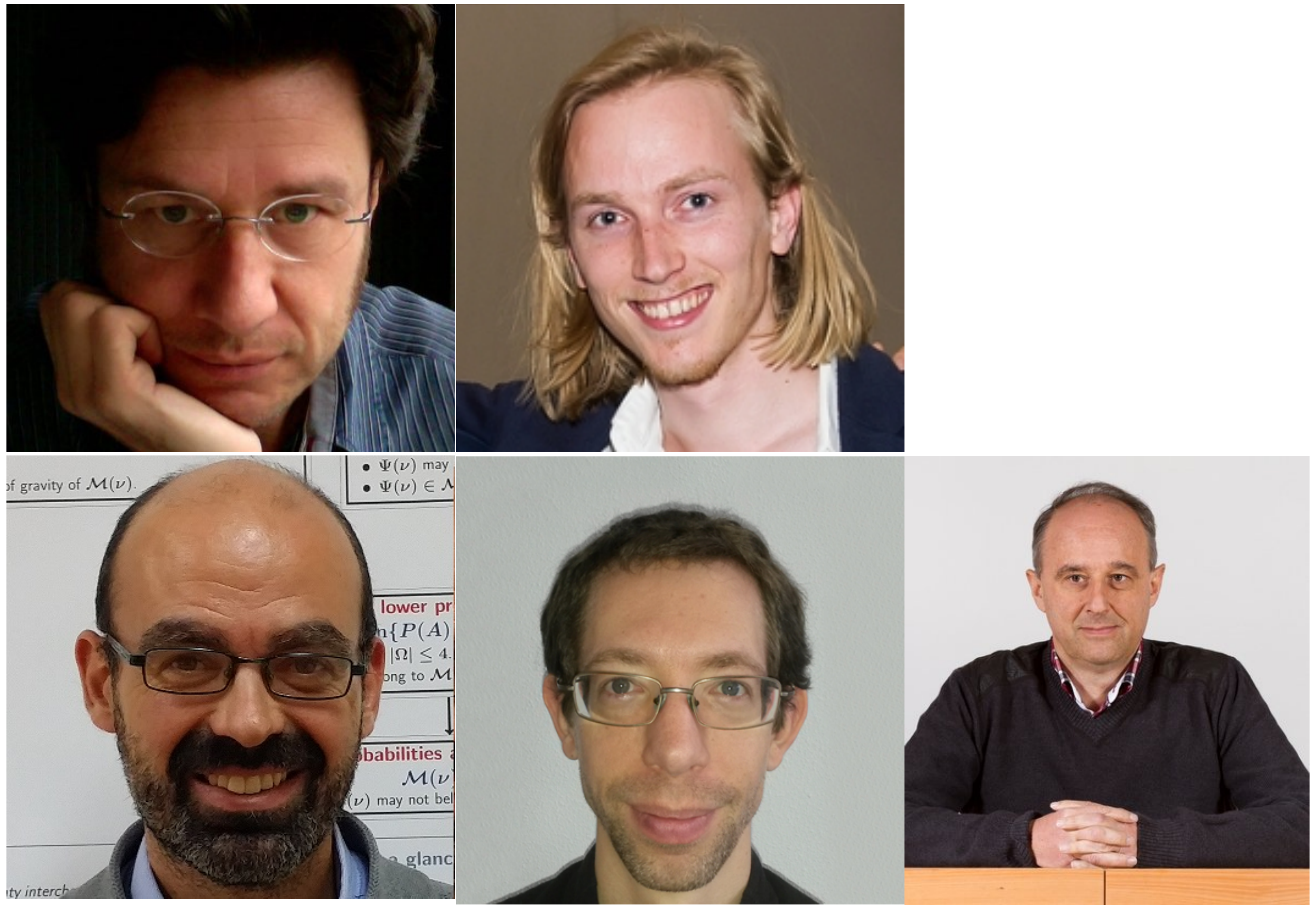
1. Introduction

Can there be a unified theory of imprecise probability? At present there are numerous mathematical models, interpretations and applications of imprecise probabilities. The various articles in this volume and in [3,4] give some idea of

E-mail address: pwalley@endoramail.com (P. Walley).

0888-613X/00/$ - see front matter © 2000 Elsevier Science Inc. All rights reserved.
PII: S 0 8 8 8 - 6 1 3 X (0 0) 0 0 0 3 1 - 1





Annals of Mathematics and Artificial Intelligence (2005) 45: 197–214
DOI: 10.1007/s10472-005-9011-0

Epistemic irrelevance on sets of desirable gambles

Serafín Moral

Departamento de Ciencias de la Computación e IA, Universidad de Granada, 18071 Granada, Spain
E-mail: smc@decsai.ugr.es

This paper studies graphoid properties for epistemic irrelevance in sets of desirable gambles. For that aim, the basic operations of conditioning and marginalization are expressed in terms of variables. Then, it is shown that epistemic irrelevance is an asymmetric graphoid. The intersection property is verified in probability theory when the global probability distribution is positive in all the values. Here it is always verified due to the handling of zero probabilities in sets of gambles. An asymmetrical D-separation principle is also presented, by which this type of independence relationships can be represented in directed acyclic graphs.

Keywords: Desirable gambles, imprecise probabilities, conditioning, epistemic independence, epistemic irrelevance

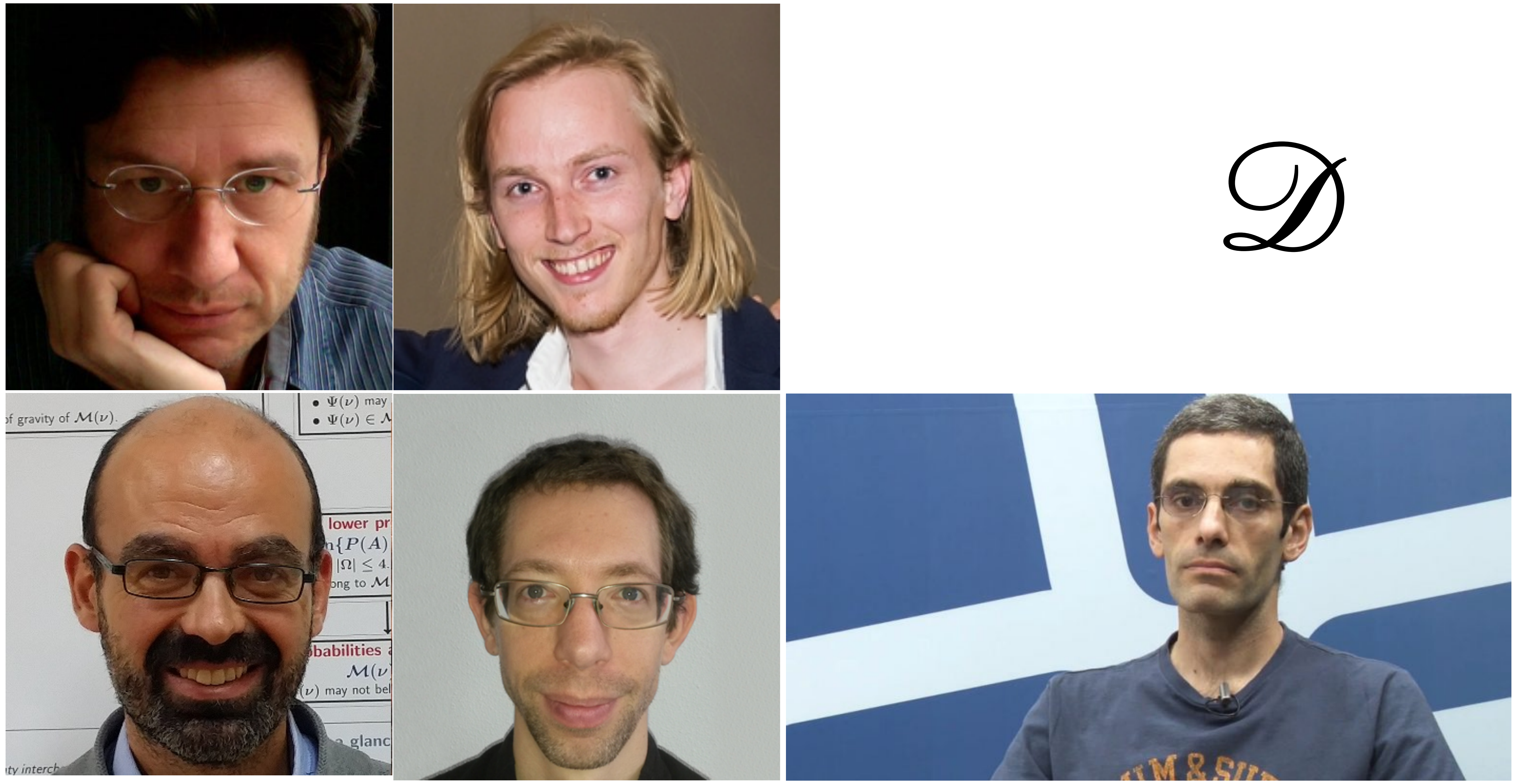
1. Introduction

Coherent sets of desirable gambles [15, 16] are a very general model for imprecise probability. They are more informative than convex sets of probability measures as they can provide the behavior conditioned to events of probability zero. Despite its generality, this model has a simple mathematical formulation, and it allows for the expression of such concepts as conditioning, combination and marginalization in a very simple way. Perhaps, its main drawback may be the difficulty of elicitation from an expert due to the implications of considering as desirable or not the gambles in the frontier.

Independence is one of the key concepts for every theory of uncertainty. In imprecise probability this concept is richer than in the classical theory of probability, and it allows a number of different interpretations [1, 3]. In the model of desirable sets of gambles the most natural definition is epistemic irrelevance: adding information about one event, then the information about the other does not change. As in credal sets [3, 14], this concept is not symmetrical. If we impose irrelevance in both directions we obtain epistemic independence.

This paper tries to investigate the graphoid properties for epistemic irrelevance in sets of desirable gambles, extending the work by [3] in which this concept was studied for credal sets of probabilities. As in that setting, symmetry is not verified. However, if we consider the epistemic independence (symmetrical irrelevance), then contraction is

Lower probabilities → Lower previsions – credal sets → Sets of desirable gambles → Choice functions



$\mathscr{D}$

Journal of Artificial Intelligence Research 52 (2015) 1-95 Submitted 07/14; published 01/15

Coherent Predictive Inference under Exchangeability with Imprecise Probabilities

Gert de Cooman GERT.DECOOMAN@UGENT.BE
Jasper De Bock JASPER.DEBOCK@UGENT.BE
Ghent University, SYSTeMS Research Group
Technologiepark–Zwijnaarde 914
9052 Zwijnaarde, Belgium

Márcio Alves Diniz MARCIO.ALVES.DINIZ@GMAIL.COM
Federal University of São Carlos, Department of Statistics
Rod. Washington Luis, km 235
São Carlos, Brazil

Abstract

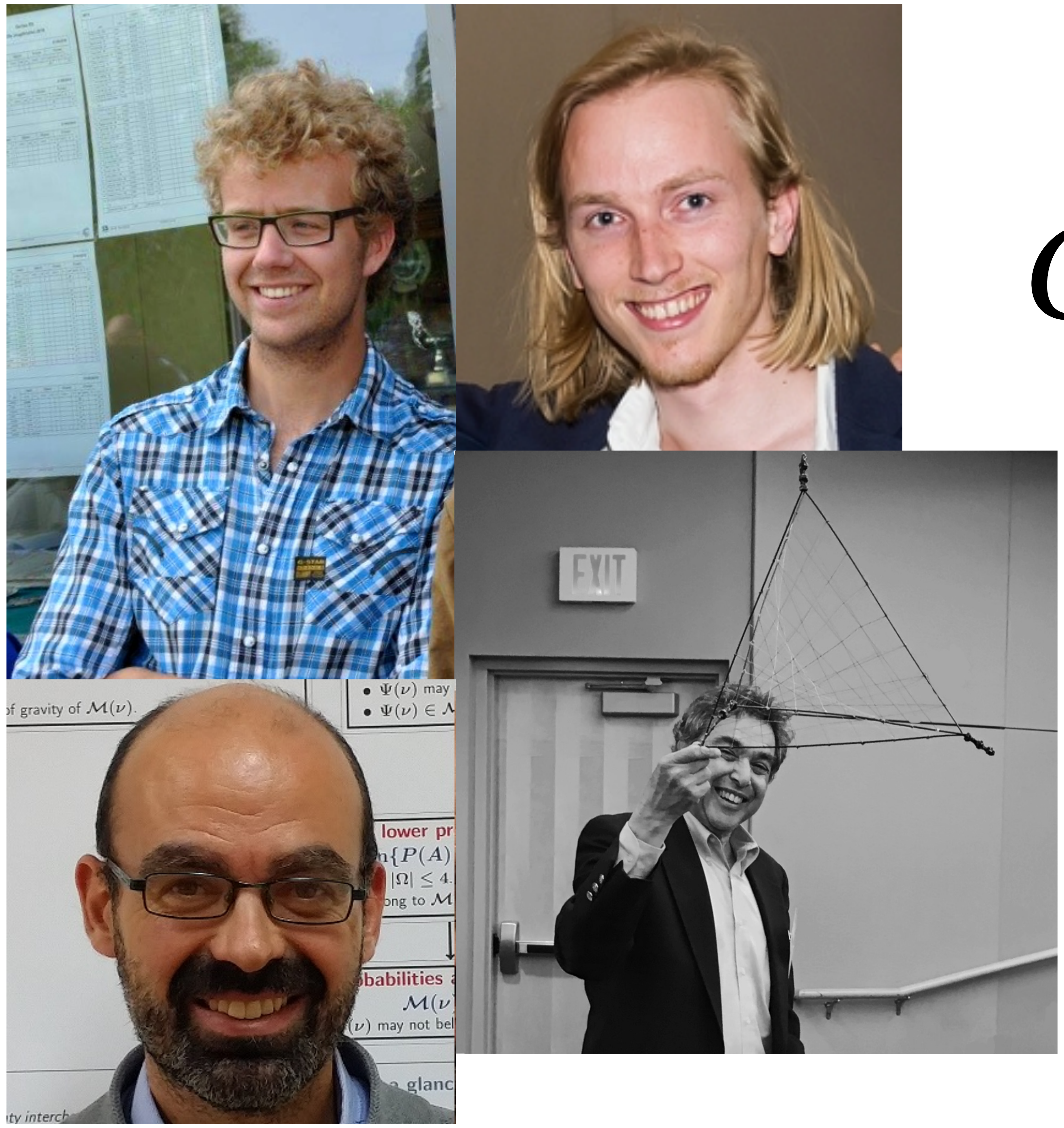
Coherent reasoning under uncertainty can be represented in a very general manner by coherent sets of desirable gambles. In a context that does not allow for indecision, this leads to an approach that is mathematically equivalent to working with coherent conditional probabilities. If we do allow for indecision, this leads to a more general foundation for coherent (imprecise-)probabilistic inference. In this framework, and for a given finite category set, coherent predictive inference under exchangeability can be represented using Bernstein coherent cones of multivariate polynomials on the simplex generated by this category set. This is a powerful generalisation of de Finetti's Representation Theorem allowing for both imprecision and indecision.

We define an inference system as a map that associates a Bernstein coherent cone of polynomials with every finite category set. Many inference principles encountered in the literature can then be interpreted, and represented mathematically, as restrictions on such maps. We discuss, as particular examples, two important inference principles: representation insensitivity—a strengthened version of Walley's representation invariance—and specificity. We show that there is an infinity of inference systems that satisfy these two principles, amongst which we discuss in particular the skeptically cautious inference system, the inference systems corresponding to (a modified version of) Walley and Bernard's Imprecise Dirichlet Multinomial Models (IDMM), the skeptical IDMM inference systems, and the Haldane inference system. We also prove that the latter produces the same posterior inferences as would be obtained using Haldane's improper prior, implying that there is an infinity of proper priors that produce the same coherent posterior inferences as Haldane's improper one. Finally, we impose an additional inference principle that allows us to characterise uniquely the immediate predictions for the IDMM inference systems.

1. Introduction

This paper deals with predictive inference for categorical variables. We are therefore concerned with a (possibly infinite) sequence of variables X_n that assume values in some finite set of categories A. After having observed a number $\check{n}$ of them, and having found that, say $X_1 = x_1$, $X_2 = x_2$, ..., $X_{\check{n}} = x_{\check{n}}$, we consider some subject's belief model for the next $\hat{n}$ variables $X_{\check{n}+1}, \dots X_{\check{n}+\hat{n}}$. In the probabilistic tradition—and we want to build on this tradition in the

Lower probabilities → Lower previsions – credal sets → Sets of desirable gambles → Choice functions



$$C(\{f_1, f_2, \ldots, f_n\})$$

Synthese (2010) 172:157–176
DOI 10.1007/s11229-009-9470-7

Coherent choice functions under uncertainty

Teddy Seidenfeld · **Mark J. Schervish** · **Joseph B. Kadane**

Received: 10 June 2008 / Accepted: 13 August 2008 / Published online: 7 March 2009

Abstract We discuss several features of *coherent choice functions*—where the admissible options in a decision problem are exactly those that maximize expected utility for some probability/utility pair in fixed set ***S*** of probability/utility pairs. In this paper we consider, primarily, normal form decision problems under uncertainty—where only the probability component of ***S*** is indeterminate and utility for two privileged outcomes is determinate. Coherent choice distinguishes between each pair of sets of probabilities regardless the "shape" or "connectedness" of the sets of probabilities. We axiomatize the theory of choice functions and show these axioms are necessary for coherence. The axioms are sufficient for coherence using a set of probability/almost-state-independent utility pairs. We give sufficient conditions when a choice function satisfying our axioms is represented by a set of probability/state-independent utility pairs with a common utility.

Keywords Choice functions · Coherence · Γ-*Maximin* · *Maximality* · Uncertainty · State-independent utility

1 Introduction

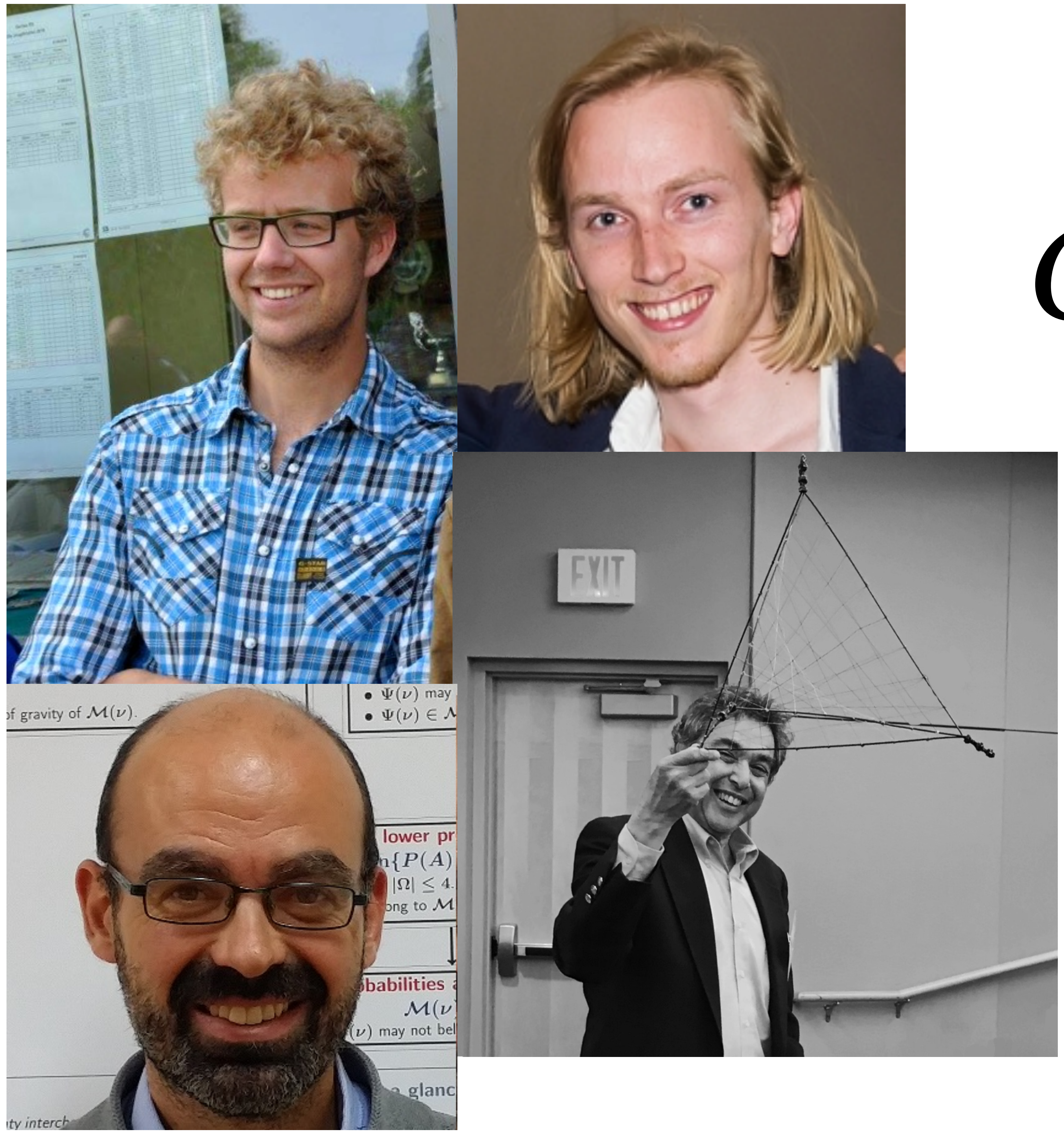
In this paper we continue our study of coherent choice functions, which we started in our (Kadane et al. 2004) "Rubinesque" theory of decision. Let ***O*** be a set of feasible

T. Seidenfeld (✉) · M. J. Schervish · J. B. Kadane
Department of Philosophy, Carnegie Mellon University, Baker Hall, 135, Pittsburgh, PA 15213, USA
e-mail: teddy@stat.cmu.edu

M. J. Schervish
e-mail: mark@stat.cmu.edu

J. B. Kadane
e-mail: kadane@stat.cmu.edu

Lower probabilities → Lower previsions – credal sets → Sets of desirable gambles → Choice functions



$$C(\{f_1, f_2, \ldots, f_n\})$$

Keuzefuncties als onzekerheidsmodellen

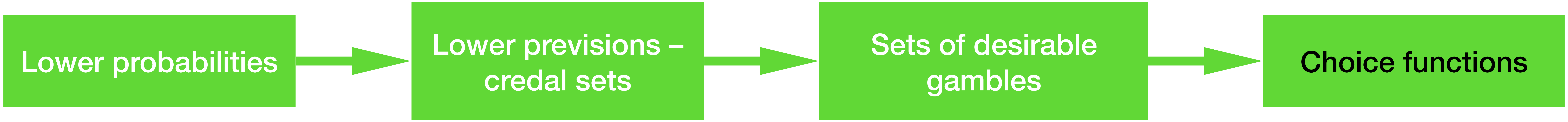
Choice Functions as a Tool to Model Uncertainty

Arthur Van Camp

Promotoren: prof. dr. ir. G. De Cooman, prof. dr. E. Miranda
Proefschrift ingediend tot het behalen van de graad van
Doctor in de ingenieurswetenschappen: wiskundige ingenieurstechnieken

Vakgroep Elektronica en Informatiesystemen
Voorzitter: prof. dr. ir. K. De Bosschere
Faculteit Ingenieurswetenschappen en Architectuur
Academiejaar 2017 - 2018

UNIVERSITEIT GENT



$$C(\{f_1, f_2, \ldots, f_n\})$$

Proceedings of Machine Learning Research 103:125–134, 2019 ISIPTA 2019

Interpreting, Axiomatising and Representing Coherent Choice Functions in Terms of Desirability

Jasper De Bock JASPER.DEBOCK@UGENT.BE
Gert de Cooman GERT.DECOOMAN@UGENT.BE
FLip, ELIS, Ghent University, Belgium

Abstract

Choice functions constitute a simple, direct and very general mathematical framework for modelling choice under uncertainty. In particular, they are able to represent the set-valued choices that appear in imprecise-probabilistic decision making. We provide these choice functions with a clear interpretation in terms of desirability, use this interpretation to derive a set of basic coherence axioms, and show that this notion of coherence leads to a representation in terms of sets of strict preference orders. By imposing additional properties such as totality, the mixing property and Archimedeanity, we obtain representation in terms of sets of strict total orders, lexicographic probability systems, coherent lower previsions or linear previsions.

Keywords: choice functions, coherence, desirability, representation, non-binary choice models.

1. Introduction

Choice functions provide an elegant unifying mathematical framework for studying set-valued choice: when presented with a set of options, they generally return a subset of them. If this subset is a singleton, it provides a unique optimal choice or decision. But if the answer contains multiple options, these are incomparable and no decision is made between them. Such set-valued choices are a typical feature of decision criteria based on imprecise-probabilistic uncertainty models, which aim to make reliable decisions in the face of severe uncertainty. Maximality and E-admissibility are well-known examples. When working with a choice function, however, it is immaterial whether it is based on such a decision criterion. The primitive objects on this approach are simply the set-valued choices themselves, and the choice function that represents all these choices serves as an uncertainty model in and by itself.

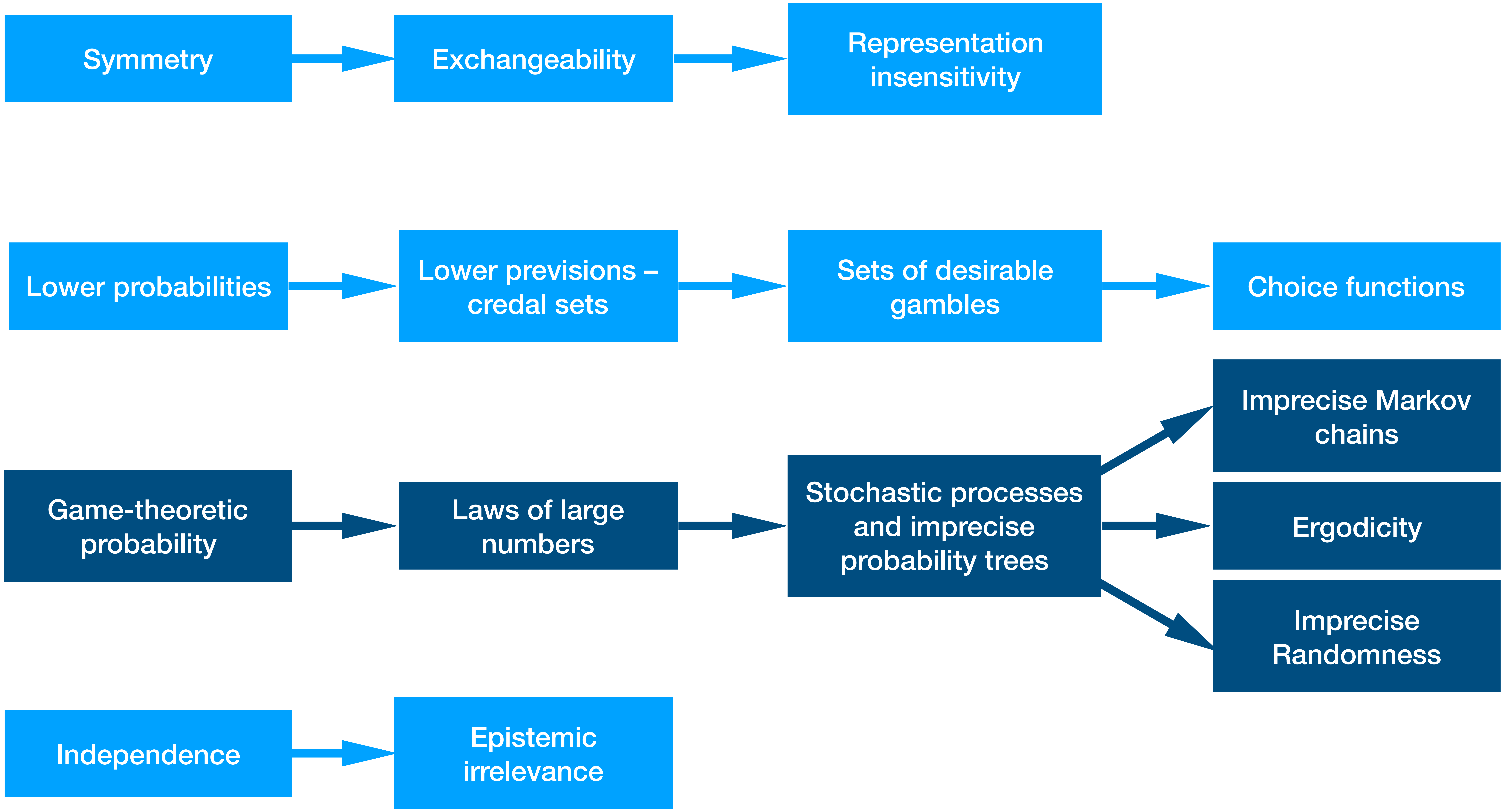
The seminal work by Seidenfeld et al. [17] has shown that a strong advantage of working with choice functions is that they allow us to impose axioms on choices, aimed at characterising what it means for choices to be rational and internally consistent. This is also what we want to do here, but we believe our angle of approach to be novel and unique: rather than think of choice intuitively, we provide it with a concrete interpretation in terms of desirability [4, 8, 9, 25] or binary preference [15]. Another important feature of our approach is that we consider a very general setting, where the options form an abstract real vector space; horse lotteries and gambles correspond to special cases.

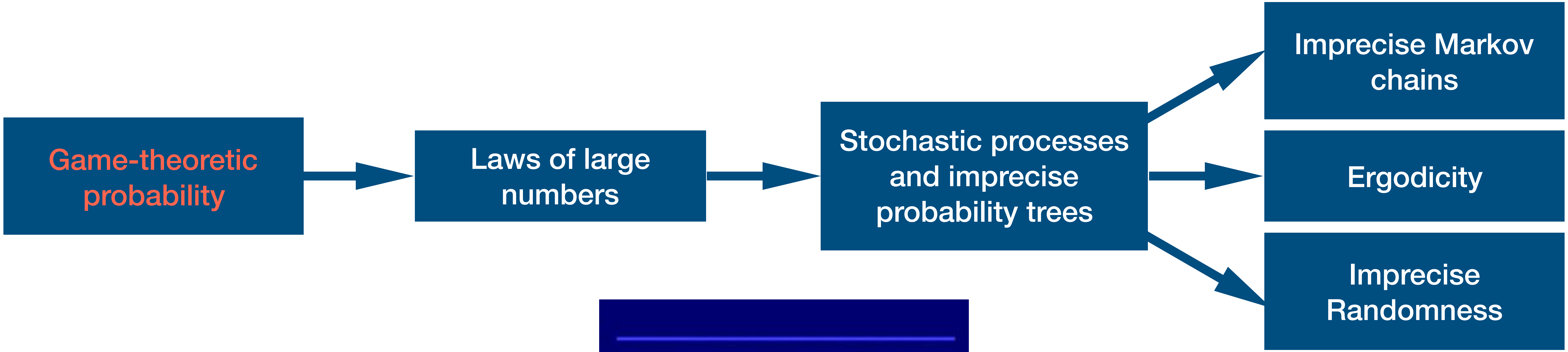
The basic structure of our paper is as follows. We start in Section 2 by introducing choice functions and our interpretation for them. Next, in Section 3, we develop an equivalent way of describing these choice functions: sets of desirable option sets. We use our interpretation to suggest and motivate a number of rationality, or coherence, axioms for such sets of desirable option sets, and show in Section 4 what are the corresponding coherence axioms for choice (or rejection) functions. Section 5 deals with the special case of binary choice, and its relation to the theory of sets of desirable options [4, 8, 9, 25] and binary preference. This is important because our main result in Section 6 shows that any coherent choice model can be represented in terms of sets of such binary choice models. In the remaining Sections 7–9, we consider additional axioms or properties, such as totality, the mixing property, and an Archimedean property, and prove corresponding representation results. This includes representations in terms of sets of strict total orders, sets of lexicographic probability systems, sets of coherent lower previsions and sets of linear previsions.

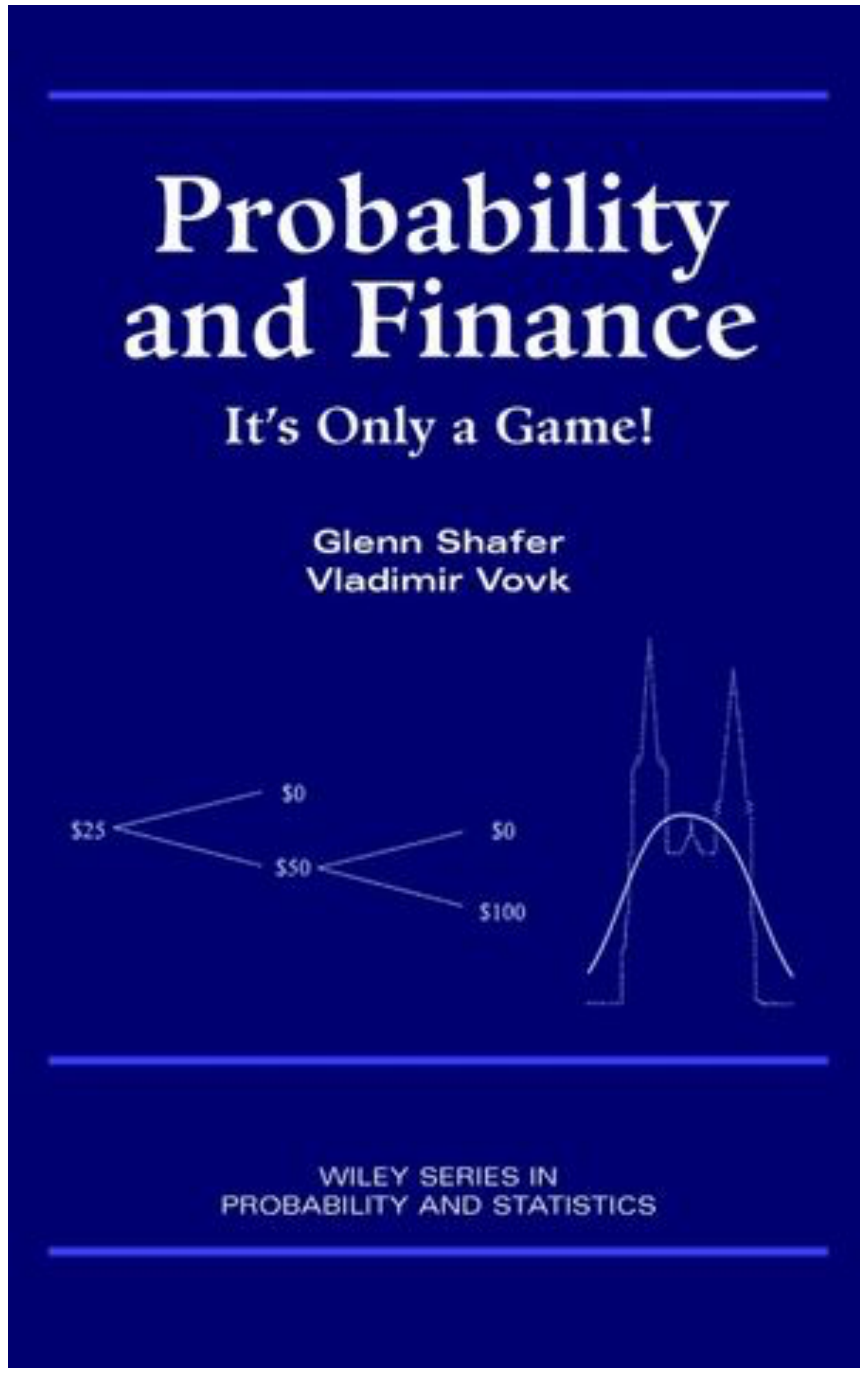
Proofs have been relegated to the appendix of an extended arXiv version [7].

2. Choice Functions and Their Interpretation

A choice function C is a set-valued operator on sets of options. In particular, for any set of options A, the corresponding value of C is a subset $C(A)$ of A. The options themselves are typically actions amongst which a subject wishes to choose. We here follow a very general approach where these options constitute an abstract real vector space $\mathscr{V}$ provided with a—so-called *background*—vector ordering $\preceq$ and a strict version $\prec$. The elements u of $\mathscr{V}$ are called *options* and $\mathscr{V}$ is therefore called the *option space*. We let $\mathscr{V}_{\succ 0} := \{u \in \mathscr{V} : u \succ 0\}$. The purpose of a choice function is to represent our subject's choices between such options.


Symmetry
Exchangeability
Representation insensitivity
Lower probabilities
Lower previsions – credal sets
Sets of desirable gambles
Choice functions
Game-theoretic probability
Laws of large numbers
Stochastic processes and imprecise probability trees
Imprecise Markov chains
Ergodicity
Imprecise Randomness
Independence
Epistemic irrelevance


Game-theoretic probability
Laws of large numbers
Stochastic processes and imprecise probability trees
Imprecise Markov chains
Ergodicity
Imprecise Randomness


Probability and Finance
It's Only a Game!
Glenn Shafer
Vladimir Vovk
$25
$0
$50
$0
$100
WILEY SERIES IN PROBABILITY AND STATISTICS

Game-theoretic probability → Laws of large numbers → Stochastic processes and imprecise probability trees → Imprecise Markov chains / Ergodicity / Imprecise Randomness

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Statistical Planning and Inference 138 (2008) 2409–2432

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Weak and strong laws of large numbers for coherent lower previsions[☆]

Gert De Cooman[a,*], Enrique Miranda[b]

[a]*Systems Research Group, Ghent University, Technologiepark-Zwijnaarde 914, 9052 Zwijnaarde, Belgium*
[b]*Department of Statistics and Operations Research, Rey Juan Carlos University, C-Tulipán, s/n, 28933 Móstoles, Spain*

Received 27 March 2006; received in revised form 13 March 2007; accepted 25 October 2007
Available online 22 November 2007

Abstract

We prove weak and strong laws of large numbers for coherent lower previsions, where the lower prevision of a random variable is given a behavioural interpretation as a subject's supremum acceptable price for buying it. Our laws are a consequence of the rationality criterion of coherence, and they can be proven under assumptions that are surprisingly weak when compared to the standard formulation of the laws in more classical approaches to probability theory.

MSC: 60A99; 60F05; 60F15

Keywords: Imprecise probabilities; Coherent lower previsions; Law of large numbers; Epistemic irrelevance; 2-Monotone capacities

1. Introduction

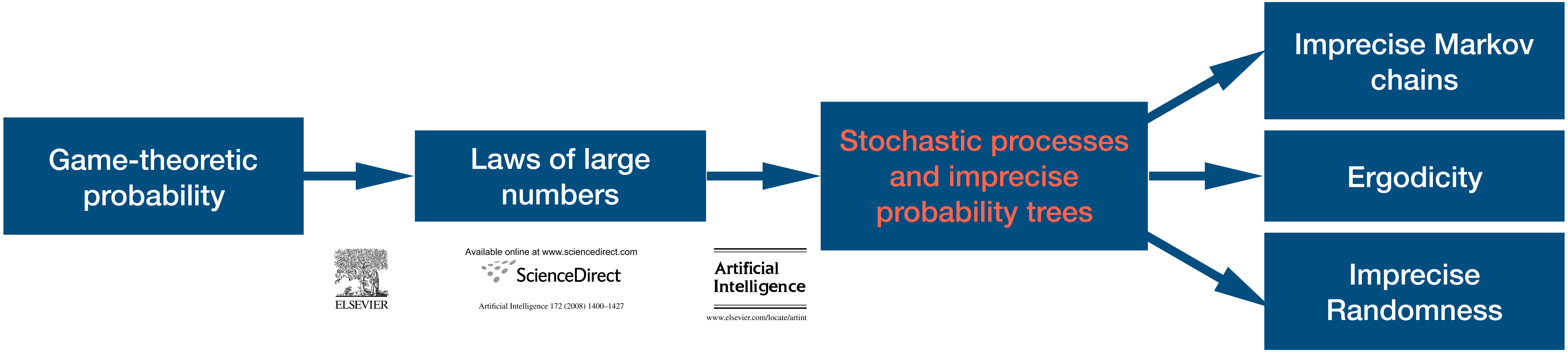
In order to set the stage for this paper, let us briefly recall a simple derivation for Bernoulli's weak law of large numbers. Consider N successive tosses of the same coin. The outcome for the kth toss is denoted by X_k, $k=1,\ldots,N$. This is a random variable, taking values in the set $\{-1, 1\}$, where -1 stands for 'tails' and $+1$ for 'heads'. We denote by p the probability for any toss to result in 'heads'. The common expected value μ of the outcomes X_k is then given by $\mu=2p-1$, and their common variance σ^2 by $\sigma^2=4p(1-p)\leqslant 1$. We are interested in the *sample mean*, which is the random variable $S_N=(1/N)\sum_{k=1}^{N} X_k$ whose expectation is μ. If we make the *extra assumption that the successive outcomes X_k are independent*, then the variance σ_N^2 of S_N is given by $\sigma_N^2=\sigma^2/N\leqslant 1/N$, and if we use Chebychev's inequality, we find for any $\varepsilon>0$ that the probability that $|S_N-\mu|>\varepsilon$ is bounded as follows:

$$P(\{|S_N-\mu|>\varepsilon\})\leqslant\frac{\sigma_N^2}{\varepsilon^2}\leqslant\frac{1}{N\varepsilon^2}. \tag{1}$$

This tells us that for any $\varepsilon>0$, the probability $P(\{|S_N-\mu|>\varepsilon\})$ tends to zero as the number of observations N goes to infinity, and we say that the sample mean S_N *converges in probability* to the expectation μ. If we let $Y_k=(1+X_k)/2$,

[☆] Supported by Research Grant G.0139.01 of the Flemish Fund for Scientific Research (FWO), and by MCYT, Projects MTM2004-01269, TSI2004-06801-C04-01.
[*] Corresponding author.
E-mail addresses: gert.decooman@ugent.be (G. De Cooman), enrique.miranda@urjc.es (E. Miranda).

0378-3758/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2007.10.020



Available online at www.sciencedirect.com
ScienceDirect
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

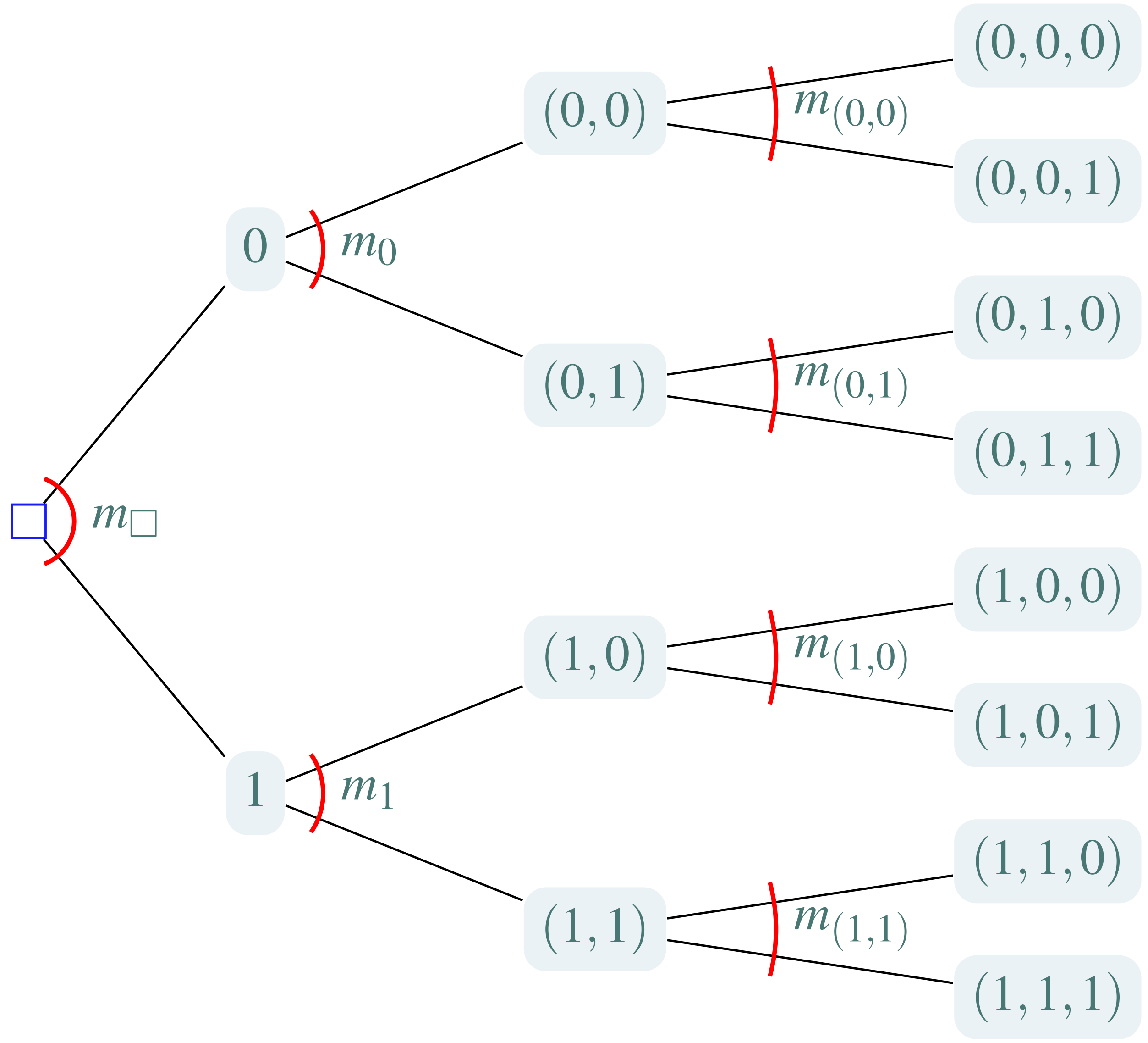
1. Introduction

In recent years, we have witnessed the growth of a number of theories of uncertainty, where imprecise (lower and upper) probabilities and previsions, rather than precise (or point-valued) probabilities and previsions, have a central part. Here we consider two of them, Glenn Shafer and Vladimir Vovk's game-theoretic account of probability [29], which is introduced in Section 2, and Peter Walley's behavioural theory [33], outlined in Section 3. These seem to have a rather different interpretation, and they certainly have been influenced by different schools of thought: Walley follows the tradition of Frank Ramsey [22], Bruno de Finetti [11] and Peter Williams [39] in trying to establish a rational model for a subject's beliefs in terms of her behaviour. Shafer and Vovk follow an approach that has many other influences as well, and is strongly coloured by ideas about gambling systems and martingales. They use Cournot's Principle to interpret lower and upper probabilities (see [28]; and [29, Chapter 2] for a nice historical overview), whereas on Walley's approach, lower and upper probabilities are defined in terms of a subject's betting rates.

* Corresponding author.
E-mail addresses: gert.decooman@UGent.be (G. de Cooman), filip.hermans@UGent.be (F. Hermans).

0004-3702/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.artint.2008.03.001

Precise probability trees



Game-theoretic probability → Laws of large numbers → Stochastic processes and imprecise probability trees → Imprecise Markov chains / Ergodicity / Imprecise Randomness

ELSEVIER

Available online at www.sciencedirect.com
ScienceDirect

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

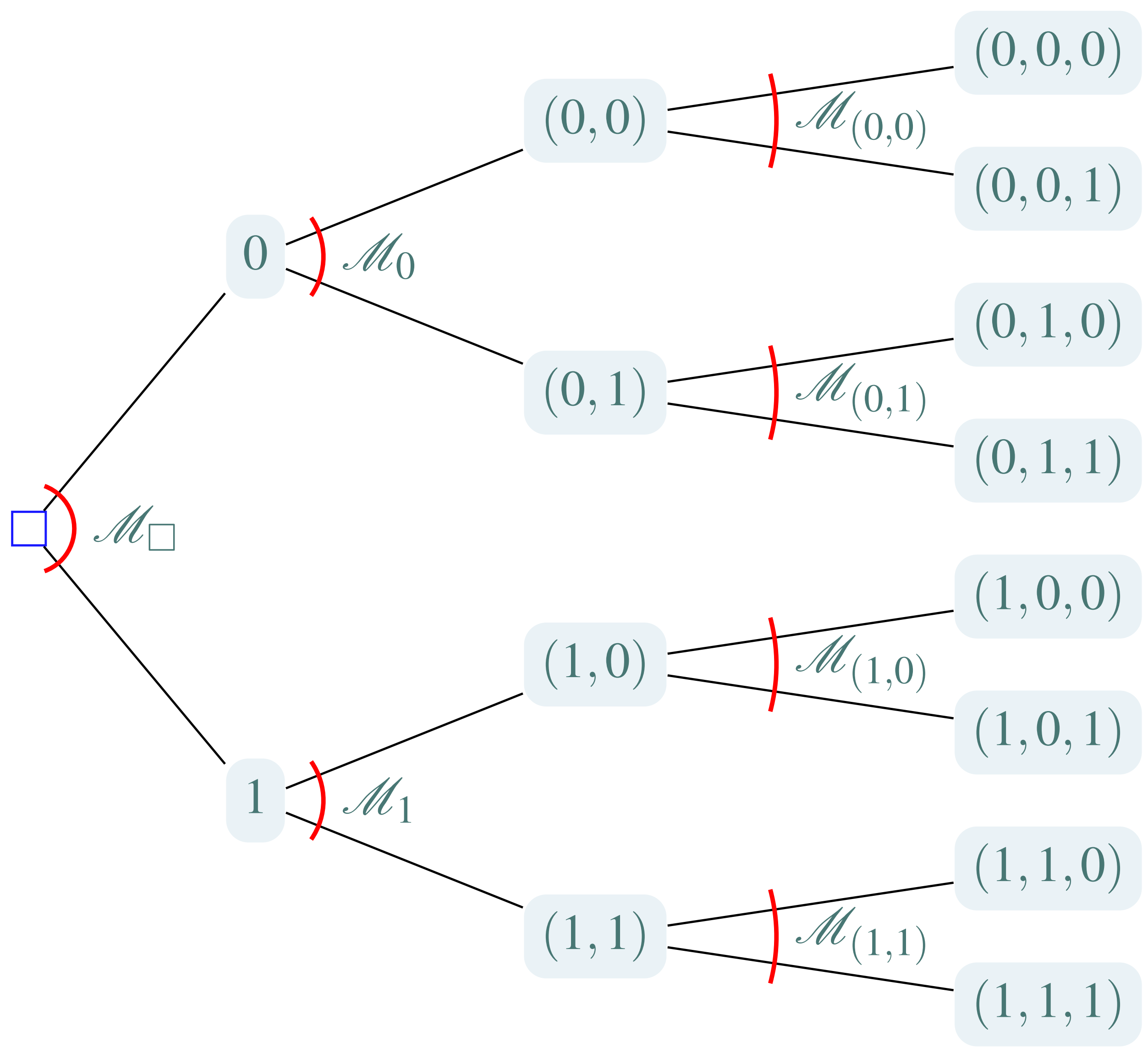
1. Introduction

In recent years, we have witnessed the growth of a number of theories of uncertainty, where imprecise (lower and upper) probabilities and previsions, rather than precise (or point-valued) probabilities and previsions, have a central part. Here we consider two of them, Glenn Shafer and Vladimir Vovk's game-theoretic account of probability [29], which is introduced in Section 2, and Peter Walley's behavioural theory [33], outlined in Section 3. These seem to have a rather different interpretation, and they certainly have been influenced by different schools of thought: Walley follows the tradition of Frank Ramsey [22], Bruno de Finetti [11] and Peter Williams [39] in trying to establish a rational model for a subject's beliefs in terms of her behaviour. Shafer and Vovk follow an approach that has many other influences as well, and is strongly coloured by ideas about gambling systems and martingales. They use Cournot's Principle to interpret lower and upper probabilities (see [28]; and [29, Chapter 2] for a nice historical overview), whereas on Walley's approach, lower and upper probabilities are defined in terms of a subject's betting rates.

* Corresponding author.
E-mail addresses: gert.decooman@UGent.be (G. de Cooman), filip.hermans@UGent.be (F. Hermans).

0004-3702/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.artint.2008.03.001

Definition and interpretation



Game-theoretic probability → Laws of large numbers → Stochastic processes and imprecise probability trees

Stochastic processes and imprecise probability trees → Imprecise Markov chains

Stochastic processes and imprecise probability trees → Ergodicity

Stochastic processes and imprecise probability trees → Imprecise Randomness

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

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

1. Introduction

In recent years, we have witnessed the growth of a number of theories of uncertainty, where imprecise (lower and upper) probabilities and previsions, rather than precise (or point-valued) probabilities and previsions, have a central part. Here we consider two of them, Glenn Shafer and Vladimir Vovk's game-theoretic account of probability [29], which is introduced in Section 2, and Peter Walley's behavioural theory [33], outlined in Section 3. These seem to have a rather different interpretation, and they certainly have been influenced by different schools of thought: Walley follows the tradition of Frank Ramsey [22], Bruno de Finetti [11] and Peter Williams [39] in trying to establish a rational model for a subject's beliefs in terms of her behaviour. Shafer and Vovk follow an approach that has many other influences as well, and is strongly coloured by ideas about gambling systems and martingales. They use Cournot's Principle to interpret lower and upper probabilities (see [28]; and [29, Chapter 2] for a nice historical overview), whereas on Walley's approach, lower and upper probabilities are defined in terms of a subject's betting rates.

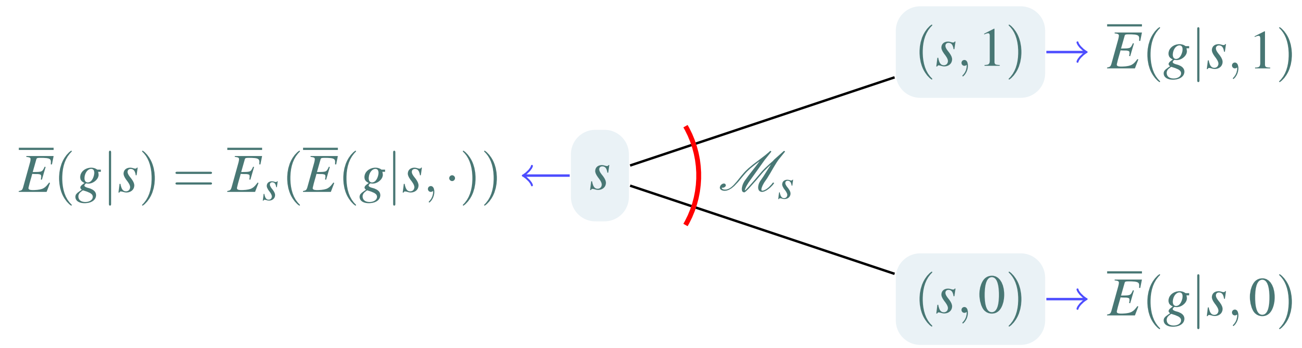
* Corresponding author.
E-mail addresses: gert.decooman@UGent.be (G. de Cooman), filip.hermans@UGent.be (F. Hermans).

0004-3702/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.artint.2008.03.001

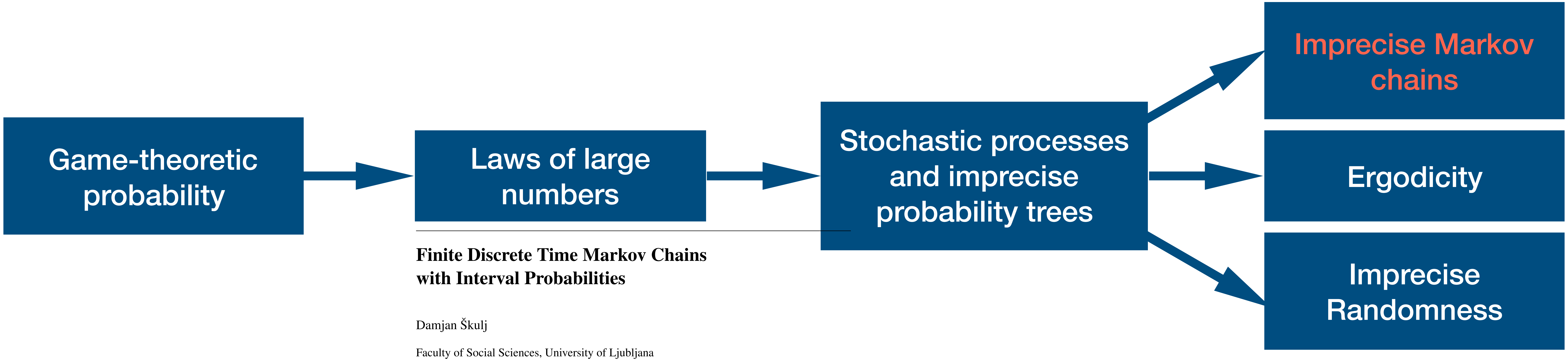
The Law of Iterated Expectation

Theorem (Law of Iterated Expectation)
Suppose we know $\overline{E}(g|s,x)$ *for all* $x \in \mathscr{X}$*, then we can calculate* $\overline{E}(g|s)$ *by backwards recursion using the local model* $\overline{E}_s$*:*

$$\overline{E}(g|s) = \underbrace{\overline{E}_s}_{\text{local}}(\overline{E}(g|s,\cdot)) = \max_{m_s \in \mathscr{M}_s} \sum_{x \in \mathscr{X}} m_s(x)\overline{E}(g|s,x).$$



The complexity of calculating the $\overline{E}(g|s)$, as a function of n, is therefore essentially the same as in the precise case!



Finite Discrete Time Markov Chains with Interval Probabilities

Damjan Škulj

Faculty of Social Sciences, University of Ljubljana
Kardeljeva ploščad 5, 1000 Ljubljana, Slovenia
Damjan.Skulj@Fdv.Uni-Lj.Si

Summary. A Markov chain model in generalised settings of interval probabilities is presented. Instead of the usual assumption of constant transitional probability matrix, we assume that at each step a transitional matrix is chosen from a set of matrices that corresponds to a structure of an interval probability matrix. We set up the model and show how to obtain intervals corresponding to sets of distributions at consecutive steps. We also state the problem of invariant distributions and examine possible approaches to their estimation in terms of convex sets of distributions, and in a special case in terms of interval probabilities.

1 Introduction

Interval probabilities present a generalised probabilistic model where classical single valued probabilities of events are replaced by intervals. In our paper we refer to Weichselberger's theory [4]; although, several other models also allow interval interpretation of probabilities.

An approach to involve interval probabilities to the theory of Markov chains was proposed by Kozine and Utkin [1]. They assume a model where transitional probability matrix is constant but unknown. Instead of that, only intervals belonging to each transitional probability are known.

In this paper we attempt to relax this model. We do this in two directions. First, we omit the assumption of the transitional probability matrix being constant, and second, instead of only allowing intervals to belong to single atoms, we allow them to belong to all subsets.

Allowing non-constant transitional probability matrix makes Markov chain model capable of modeling real situations where in general it is not reasonable to expect exactly the same transitional probabilities at each step. They can, however, be expected to belong to some set of transitional probabilities. In interval probability theory such sets are usually obtained as structures of interval probabilities. Our assumption is thus that transitional probability at each step is an arbitrary member of a set of transitional probability matrices generated by an interval probability matrix.

D. Škulj: *Finite Discrete Time Markov Chains with Interval Probabilities*, Advances in Soft Computing **6**, 299–306 (2006)
www.springerlink.com

Game-theoretic probability → Laws of large numbers → Stochastic processes and imprecise probability trees → Imprecise Markov chains / Ergodicity / Imprecise Randomness

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

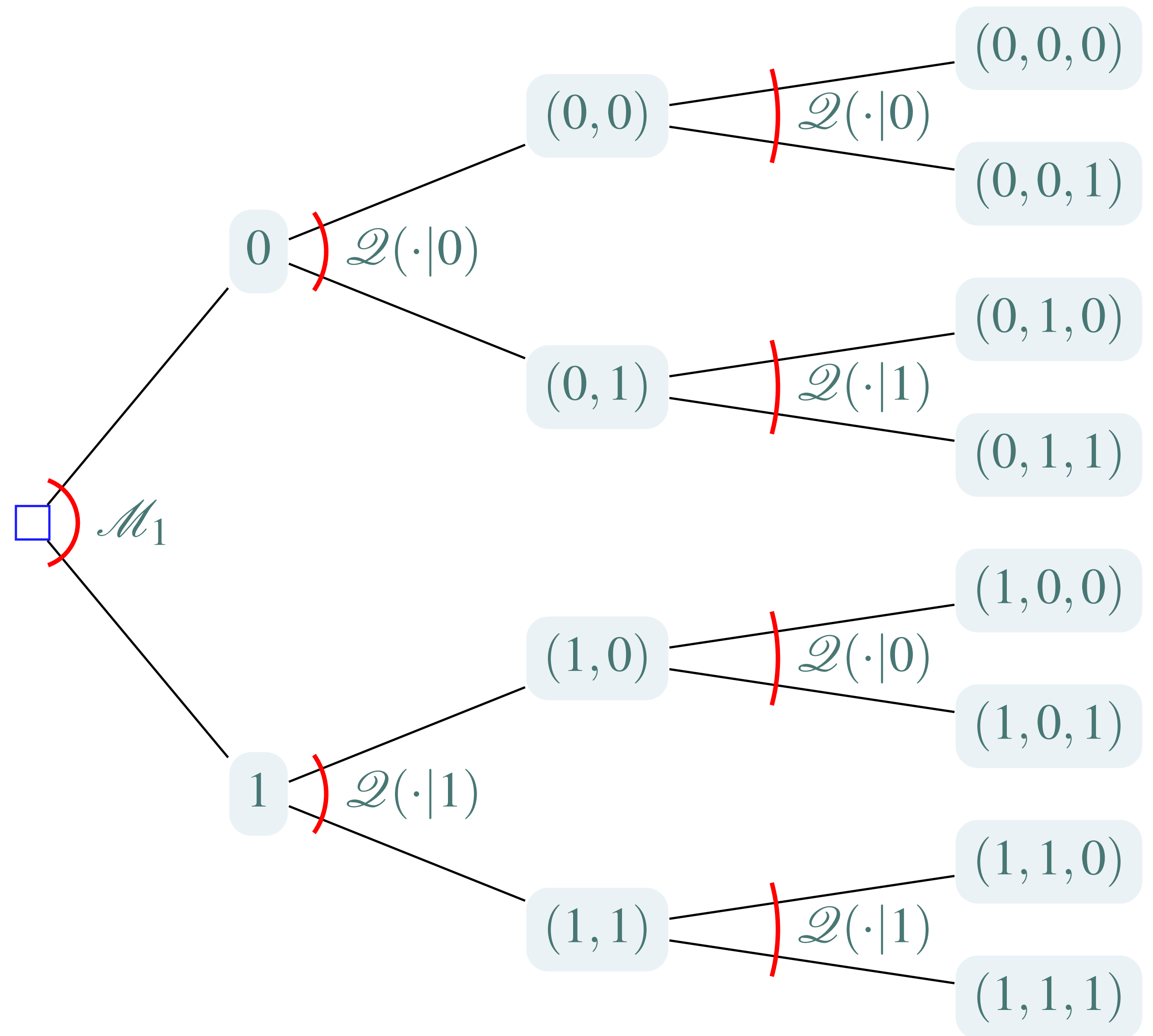
When the initial and transition probabilities of a finite Markov chain in discrete time are not well known, we should perform a sensitivity analysis. This can be done by considering as basic uncertainty models the so-called *credal sets* that these probabilities are known or believed to belong to and 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.

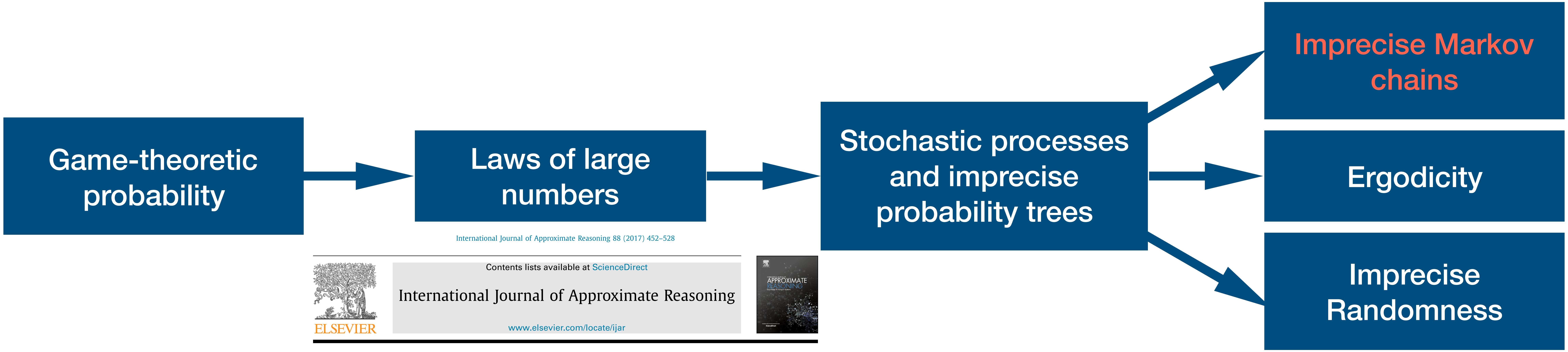
1. INTRODUCTION

One convenient way to model uncertain dynamical systems is to describe them as Markov chains. These have been studied in great detail, and their properties are well known. However, in many practical situations, it remains a challenge to accurately identify the transition probabilities in the Markov chain: The available information about physical systems is often imprecise and uncertain. Describing a real-life dynamical system as a Markov chain will therefore often involve unwarranted precision and might lead to conclusions not supported by the available information.

For this reason, it seems quite useful to perform probabilistic robustness studies, or sensitivity analyses, for Markov chains. This is especially relevant in decision-making applications. Many researchers in Markov Chain Decision Making

Imprecise Markov chains: definition





International Journal of Approximate Reasoning 88 (2017) 452–528

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Imprecise continuous-time Markov chains

Thomas Krak [a,*], Jasper De Bock [b], Arno Siebes [a]

[a] Utrecht University, Department of Information and Computing Sciences, Princetonplein 5, De Uithof, 3584 CC Utrecht, The Netherlands
[b] Ghent University, Department of Electronics and Information Systems, Technologiepark – Zwijnaarde 914, 9052 Zwijnaarde, Belgium

ARTICLE INFO

Article history:
Received 17 November 2016
Received in revised form 20 June 2017
Accepted 22 June 2017
Available online 4 July 2017

Keywords:
Continuous-time markov chain
Imprecise probability
Model uncertainty
Lower and upper expectation
Lower transition operator

ABSTRACT

Continuous-time Markov chains are mathematical models that are used to describe the state-evolution of dynamical systems under stochastic uncertainty, and have found widespread applications in various fields. In order to make these models computationally tractable, they rely on a number of assumptions that—as is well known—may not be realistic for the domain of application; in particular, the ability to provide exact numerical parameter assessments, and the applicability of time-homogeneity and the eponymous Markov property. In this work, we extend these models to *imprecise continuous-time Markov chains* (ICTMC's), which are a robust generalisation that relaxes these assumptions while remaining computationally tractable.

More technically, an ICTMC is a *set* of "precise" continuous-time finite-state stochastic processes, and rather than computing expected values of functions, we seek to compute *lower expectations*, which are tight lower bounds on the expectations that correspond to such a set of "precise" models. Note that, in contrast to e.g. Bayesian methods, all the elements of such a set are treated on equal grounds; we do not consider a distribution over this set. Together with the conjugate notion of *upper expectation*, the bounds that we provide can then be intuitively interpreted as providing best- and worst-case scenarios with respect to all the models in our set of stochastic processes.

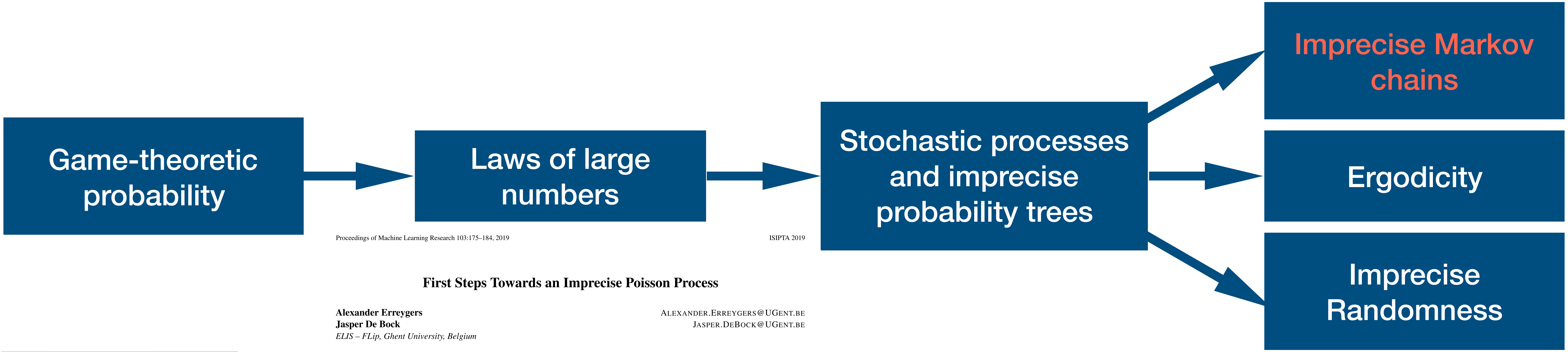
The first part of this paper develops a formalism for describing continuous-time finite-state stochastic processes that does not require the aforementioned simplifying assumptions. Next, this formalism is used to characterise ICTMC's and to investigate their properties. The concept of lower expectation is then given an alternative operator-theoretic characterisation, by means of a *lower transition operator*, and the properties of this operator are investigated as well. Finally, we use this lower transition operator to derive tractable algorithms (with polynomial runtime complexity w.r.t. the maximum numerical error) for computing the lower expectation of functions that depend on the state at any finite number of time points.

1. Introduction

Continuous-time Markov chains are mathematical models that can describe the behaviour of dynamical systems under stochastic uncertainty. In particular, they describe the stochastic evolution of such a system through a discrete state space and over a continuous time-dimension. This class of models has found widespread applications in various fields, includ-

* Corresponding author.
E-mail addresses: t.e.krak@uu.nl (T. Krak), jasper.debock@ugent.be (J. De Bock), a.p.j.m.siebes@uu.nl (A. Siebes).

http://dx.doi.org/10.1016/j.ijar.2017.06.012



Proceedings of Machine Learning Research 103:175–184, 2019 ISIPTA 2019

First Steps Towards an Imprecise Poisson Process

Alexander Erreygers ALEXANDER.ERREYGERS@UGENT.BE
Jasper De Bock JASPER.DEBOCK@UGENT.BE
ELIS – FLip, Ghent University, Belgium

Abstract

The Poisson process is the most elementary continuous-time stochastic process that models a stream of repeating events. It is uniquely characterised by a single parameter called the rate. Instead of a single value for this rate, we here consider a rate interval and let it characterise two nested sets of stochastic processes. We call these two sets of stochastic process imprecise Poisson processes, explain why this is justified, and study the corresponding lower and upper (conditional) expectations. Besides a general theoretical framework, we also provide practical methods to compute lower and upper (conditional) expectations of functions that depend on the number of events at a single point in time.

Keywords: Poisson process, counting process, continuous-time Markov chain, imprecision

1. Introduction

The *Poisson process* is arguably one of the most basic stochastic processes. At the core of this model is our subject, who is interested in something specific that occurs repeatedly over time, where time is assumed to be continuous. For instance, our subject could be interested in the arrival of a customer to a queue, to give an example from queueing theory. For the sake of brevity, we will call such a specific occurrence a *Poisson-event*,[1] whence our subject is interested in a stream of Poisson-events. The time instants at which subsequent Poisson-events occur are uncertain to our subject, hence the need for a probabilistic model. This set-up is not exclusive to queueing theory; it is also used in renewal theory and reliability theory, to name but a few applications.

There is a plethora of alternative but essentially equivalent characterisations of this Poisson process. Some of the more well-known and basic characterisations are as the limit of the Bernoulli process [5, Chapter VI, Sections 5 and 6] or as a sequence of mutually independent and exponentially distributed inter-Poisson-event times [6, Chapter 5, Section 3.A]. An alternative way to look at the Poisson process is as a random dispersion of points in some general space—that need not be the real number line—see for instance [1, Sections 2.1 and 2.2] or [8, Chapter 2]. More theoretically involved characterisations that are relevant to our set-up are as a counting process or as a continuous-time Markov chain, see for example [5, Chapter XVII, Section 2], [7, Section 1], [10, Section 2.4], [12, Section 2.1] or [13, Section 3].

Broadly speaking, these characterisations all make the same three assumptions: (i) *orderliness*, in the sense that the probability that two or more Poisson-events occur at the same time is zero; (ii) independence, more specifically the absence of after-effects or *Markovianity*; and (iii) *homogeneity*. It is essentially well-known that these three assumptions imply the existence of a parameter called the rate, and that this rate uniquely characterises the Poisson process. We here weaken the three aforementioned assumptions. First and foremost, we get rid of the implicit assumption that our subject's beliefs can be accurately modelled by a single stochastic process; instead, we assume that her beliefs only allow us to consider a *set* of stochastic processes. Specifically, we consider a rate interval instead of a precise value for the rate, and examine two distinct sets: (i) the set of all Poisson processes whose rate belongs to this rate interval; and (ii) the set of all processes that are orderly and "consistent" with the rate interval. We then define lower and upper conditional expectations as the infimum and supremum of the conditional expectations with respect to the stochastic processes in these respective sets. Aside from this general theoretical framework, we focus on computing the lower and upper expectation of functions that depend on the number of occurred Poisson-events at a single future time point. For the set of Poisson processes, we show that this requires the solution of a one-parameter optimisation problem; for the second set, we show that this can be computed using backwards recursion. Furthermore, we argue that both sets can be justifiably called imprecise Poisson processes: imprecise because their lower and upper expectations are not equal, and Poisson because their lower and upper expectations satisfy imprecise versions of the defining properties of the (precise) Poisson process. The interested reader can find proofs for all our results in the Appendix of the extended pre-print of this contribution [4], which is available on arXiv.

1. We use the term "Poisson-event" rather than just "event" to avoid confusion with the standard usage of event in probability theory, where event refers to a subset of the sample space; we are indebted to an anonymous reviewer for pointing out this potential confusion, and to Gert de Cooman for suggesting the adopted terminology.

Game-theoretic probability → Laws of large numbers → Stochastic processes and imprecise probability trees → Imprecise Markov chains / Ergodicity / Imprecise Randomness

International Journal of Approximate Reasoning 76 (2016) 18–46

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Imprecise stochastic processes in discrete time: global models, imprecise Markov chains, and ergodic theorems

Gert de Cooman, Jasper De Bock*, Stavros Lopatatzidis

Ghent University, Data Science Lab, Technologiepark–Zwijnaarde 914, 9052 Zwijnaarde, Belgium

A R T I C L E   I N F O

Article history:
Received 9 December 2015
Received in revised form 18 April 2016
Accepted 20 April 2016
Available online 29 April 2016

Keywords:
Imprecise stochastic process
Lower expectation
Game-theoretic probability
Law of iterated expectations
Imprecise Markov chain
Point-wise ergodic theorem

A B S T R A C T

We justify and discuss expressions for joint lower and upper expectations in imprecise probability trees, in terms of the sub- and supermartingales that can be associated with such trees. These imprecise probability trees can be seen as discrete-time stochastic processes with finite state sets and transition probabilities that are imprecise, in the sense that they are only known to belong to some convex closed set of probability measures. We derive various properties for their joint lower and upper expectations, and in particular a law of iterated expectations. We then focus on the special case of imprecise Markov chains, investigate their Markov and stationarity properties, and use these, by way of an example, to derive a system of non-linear equations for lower and upper expected transition and return times. Most importantly, we prove a game-theoretic version of the strong law of large numbers for submartingale differences in imprecise probability trees, and use this to derive point-wise ergodic theorems for imprecise Markov chains.

1. Introduction

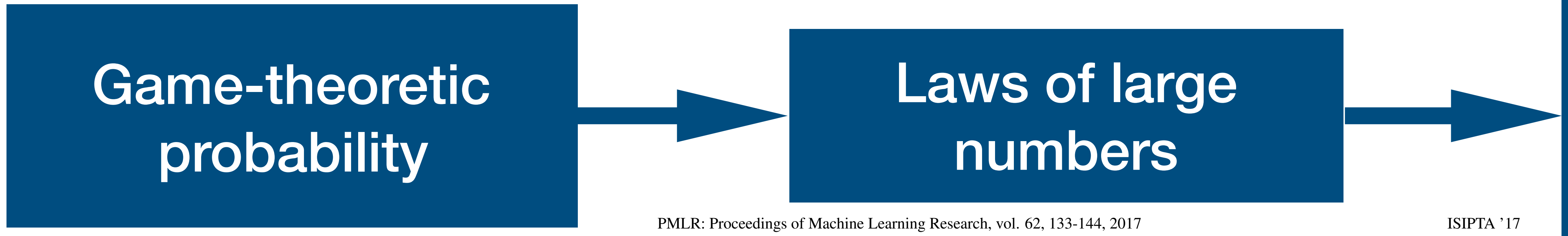
In Ref. [3], De Cooman and Hermans made a first attempt at laying the foundations for a theory of discrete-event (and discrete-time) stochastic processes that are governed by sets of, rather than single, probability measures. They showed how this can be done by connecting Walley's [23] theory of coherent lower previsions with ideas and results from Shafer and Vovk's [17] game-theoretic approach to probability theory. In later papers, De Cooman et al. [7] applied these ideas to finite-state discrete-time Markov chains, inspired by the work of Hartfiel [11]. They showed how to perform efficient inferences in, and proved a Perron–Frobenius-like theorem for, so-called imprecise Markov chains, which are finite-state discrete-time Markov chains whose transition probabilities are imprecise, in the sense that they are only known to belong to a convex closed set of probability measures—typically due to partial assessments involving probabilistic inequalities. This work was later refined and extended by Hermans and De Cooman [12] and Škulj and Hable [22].

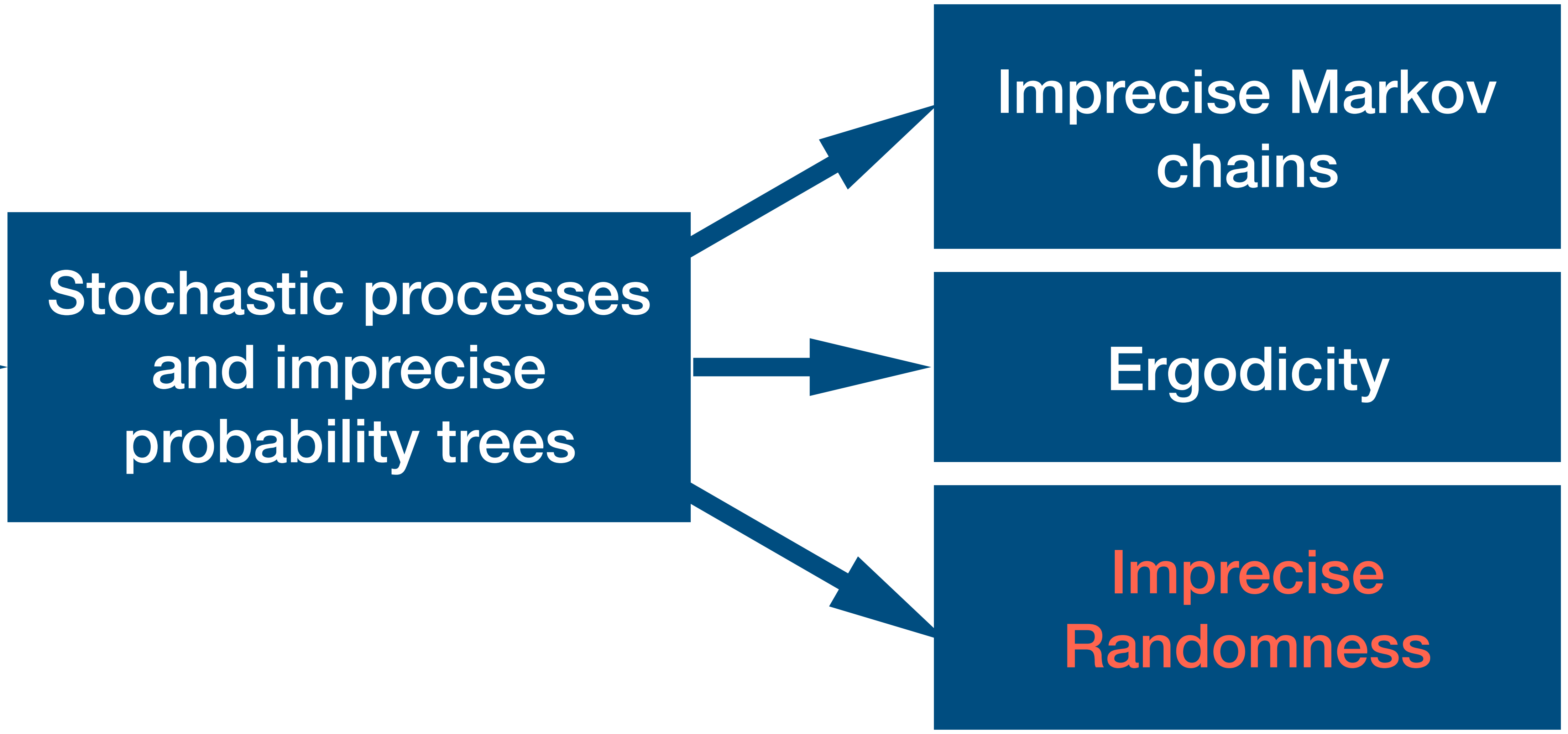
The Perron–Frobenius-like theorems in these papers give equivalent necessary and sufficient conditions for the uncertainty model—a set of probabilities—about the state X_n to converge, for $n \to +\infty$, to an uncertainty model that is independent of the uncertainty model for the initial state X_1.

In Markov chains with 'precise' transition probabilities, this convergence behaviour is sufficient for a point-wise ergodic theorem to hold, namely that:

* Corresponding author.
E-mail addresses: Gert.deCooman@UGent.be (G. de Cooman), Jasper.DeBock@UGent.be (J. De Bock), Stavros.Lopatatzidis@UGent.be (S. Lopatatzidis).

http://dx.doi.org/10.1016/j.ijar.2016.04.009





PMLR: Proceedings of Machine Learning Research, vol. 62, 133-144, 2017 ISIPTA '17

Computable Randomness is Inherently Imprecise

Gert de Cooman GERT.DECOOMAN@UGENT.BE
Jasper De Bock JASPER.DEBOCK@UGENT.BE
Ghent University, IDLab
Technologiepark – Zwijnaarde 914, 9052 Zwijnaarde (Belgium)

Abstract

We use the martingale-theoretic approach of game-theoretic probability to incorporate imprecision into the study of randomness. In particular, we define a notion of computable randomness associated with interval, rather than precise, forecasting systems, and study its properties. The richer mathematical structure that thus arises lets us better understand and place existing results for the precise limit. When we focus on constant interval forecasts, we find that every infinite sequence of zeroes and ones has an associated filter of intervals with respect to which it is computably random. It may happen that none of these intervals is precise, which justifies the title of this paper. We illustrate this by showing that computable randomness associated with non-stationary precise forecasting systems can be captured by a stationary interval forecast, which must then be less precise: a gain in model simplicity is thus paid for by a loss in precision.

Keywords: computable randomness; imprecise probabilities; game-theoretic probability; interval forecast; supermartingale; computability.

1. Introduction

This paper documents the first steps in our attempt to incorporate indecision and imprecision into the study of randomness. Consider a infinite sequence $\omega = (z_1, \dots, z_n, \dots)$ of zeroes and ones; when do we call it *random*? There are many notions of randomness, and many of them have a number of equivalent definitions (Ambos-Spies and Kucera, 2000; Bienvenu et al., 2009). We focus here on *computable randomness*, mainly because its focus on computability—rather than, say, the weaker lower semicomputability—has allowed us in this first attempt to keep the mathematical nitpicking at arm's length. Randomness of a sequence ω is typically associated with a probability measure on the sample space of all infinite sequences, or—what is equivalent—with a *forecasting system* γ that associates with each finite sequence of outcomes $(x_1, \dots, x_n)$ the (conditional) expectation $\gamma(x_1, \dots, x_n)$ for the next (as yet unknown) outcome X_{n+1}. The sequence ω is then called *computably* random when it passes a (countable) number of *computable* tests of randomness, where the collection of randomness tests depends of the forecasting system γ. An alternative but equivalent definition, going back to Ville (1939), sees each forecast $\gamma(x_1, \dots, x_n)$ as a fair price for—and therefore a commitment to bet on—the as yet unknown next outcome X_{n+1}. The sequence ω is then computably random when there is no computable strategy for getting infinitely rich by exploiting the bets made available by the forecasting system γ along the sequence, without borrowing. Technically speaking, all computable non-negative supermartingales should remain bounded on ω, and the forecasting system γ determines what a supermartingale is.

It is this last, martingale-theoretic approach which seems to lend itself most easily to allowing for imprecision in the forecasts, and therefore in the definition of randomness. As we explain in Sections 2 and 3, an 'imprecise' forecasting system γ associates with each finite sequence of outcomes