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

ARTICLE INFO

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

ABSTRACT

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:

☆ This paper is part of the virtual special issue on ISIPTA 2015, edited by T. Augustin, S. Doria, M. Marinacci.

* 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
0888-613X/

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n} f(X_k)=E_\infty(f) \text{ almost surely} \tag{1}$$

for all real functions f on the finite state set $\mathscr{X}$, where E_∞ is the limit expectation operator that the expectation operators E_n for the state X_n at time n converge to point-wise, independently of the initial model E_1 for X_1, according to the classical Perron–Frobenius Theorem.[1]

One of the aims of the present paper is to extend this result to a version for imprecise Markov chains; see Theorem 32 further on. In contradistinction with the so-called *Markov set-chains* more commonly encountered in the literature [10,9,11], our imprecise Markov chains are not merely collections of (precise) Markov chains—incidentally, for such Markov set-chains, proving an ergodic theorem would be a fairly trivial affair, as it would amount to applying the classical point-wise ergodic theorem to each of the Markov chains in the collection. Rather, as we will explain in Section 6, our imprecise Markov chains correspond to a collection of stochastic processes that need not satisfy the Markov property. They are only 'superficially Markov', in the sense that their *sets* of transition probabilities satisfy a Markov condition, whereas the individual members of those sets need not. In other words, imprecise Markov chains are *not* simply collections of precise Markov chains, but rather correspond to collections of general stochastic processes whose transition models belong to sets that satisfy a Markov condition.

We are aware of one other ergodicity result for imprecise probability models, which is the subject of a very recent paper by Cerreia-Vioglio et al. [8]. It is at once more general and more restricted than our result: the context is not restricted to shift invariance in Markov chains, but extends to invariance under arbitrary transformations on arbitrary sample spaces. But, on the other hand, the (continuity) conditions imposed on the imprecise probability models are much more stringent than what we will require here. In summary, then, our results cannot be obtained as a special case of theirs.

How do we mean to go about proving our ergodicity result? In Section 2, we explain what we mean by imprecise probability models: we extend the notion of an expectation operator to so-called lower (and upper) expectation operators, and explain how these can be associated with (convex and closed) sets of expectation operators.

In Section 3, we explain how these generalised uncertainty models can be combined with event trees to form so-called imprecise probability trees, to produce a simple theory of discrete-time stochastic processes. We show in particular how to combine local uncertainty models associated with the nodes in the tree into global uncertainty models (global conditional lower expectations) about the paths in the tree, and how this procedure is related to sub- and supermartingales. We also indicate how it extends and subsumes the (precise-)probabilistic approach.

In Section 4 we prove a very general strong law of large numbers for submartingale differences in our imprecise probability trees. Our point-wise ergodic theorem will turn out to be a consequence of this in the particular context of imprecise Markov chains. Section 5 is more technical, and is devoted to extending the joint lower and upper expectations to extended real variables, and to proving a number of important properties for them, such as generalisations of well-known coherence properties, and a version of the law of iterated (lower) expectations.

We explain what imprecise Markov chains are in Section 6: how they are special cases of imprecise probability trees, how to do efficient inference for them, and how to define Perron–Frobenius-like behaviour. We generalise existing results [7] about global lower expectations in such imprecise Markov trees from a finite to an infinite time horizon, and from bounded real argument functions to extended real-valued ones. We also explore the influence of time shifts on the global (conditional) lower expectations, investigate their Markov properties, prove various corollaries of the law of iterated lower expectations, and discuss stationarity and its relation with Perron–Frobenius-like behaviour. As an illustration of the power of our approach, we derive in Section 7 a system of non-linear equations for lower and upper expected transition and return times, and solve it in special case.

In Section 8 we show that there is an interesting identity between the time averages that appear in our strong law of large numbers, and the ones that appear in the point-wise ergodic theorem. The discussion in Section 9 first focusses on a number of terms in this identity, and investigates their convergence for Perron–Frobenius-like imprecise Markov chains. This allows us to use the identity to prove two versions of the point-wise ergodic theorem: one for functions of a single state (Theorem 32) and its extension (Corollary 34) to functions of a finite number of states. We briefly discuss their significance in Section 10.

Some of the results in this paper have already been discussed—without proofs—in an earlier conference version [6]. This paper significantly extends the earlier version.

2. Basic notions from imprecise probabilities

Let us begin with a brief sketch of a few basic definitions and results about imprecise probabilities. For more details, we refer to Walley's [23] seminal book, as well as more recent textbooks [1,20].

Suppose a subject is uncertain about the value that a variable Y assumes in a non-empty set of possible values $\mathscr{Y}$. He is therefore also uncertain about the value $f(Y)$ a so-called *gamble*—a bounded real-valued function—$f : \mathscr{Y} \to \mathbb{R}$ on the set

[1] Actually, much more general results can be proved, for functions f that do not depend on a single state only, but on the entire sequence of states; see for instance Ref. [13, Chapter 20]. In this paper, we will focus on the simpler version, but we will show that it can be extended to functions on a finite number of states.