# Relating epistemic irrelevance to event trees 

Sébastien Destercke and Gert de Cooman


#### Abstract

We relate the epistemic irrelevance in Walley's behavioural theory of imprecise probabilities to the event-tree independence due to Shafer. In particular, we show that forward irrelevance is equivalent to event-tree independence in particular event trees, suitably generalised to allow for the fact that imprecise rather than precise probability models are attached to the nodes in the tree. This allows us to argue that in a theory of uncertain processes, the asymmetrical notion of epistemic irrelevance has a more important role to play than its more involved and symmetrical counterpart called epistemic independence.


## 1 Introduction

Assessments of independence between variables are very important and useful in modelling uncertainty, as they allow for a reduction of complexity in many problems (e.g., in building joint models from marginal information, making statistical inferences, etc.). Here, we are interested in the case where beliefs are modelled by lower and upper expectations for random variables or, equivalently [13], by closed convex sets of (finitely additive) probabilities, also called credal sets [6, 7, 8]. In this imprecise probabilities setting, there are many different notions of irrelevance and independence, each with a different interpretation, but which generally coincide for models involving only precise probabilities, i.e., classical Bayesian belief models; see Couso et al. [5] for a review. Starting from given imprecise marginals, these different types of irrelevance and independence assessments will generally lead to

[^0]different joint belief models, whereas they all lead to the classical independent product when marginal beliefs are modelled by precise, or Bayesian, probabilities. A discussion of this phenomenon can also be found in De Cooman and Miranda [3].

As far as we know, there are currently two important approaches to probability theory that involve lower and upper expectations (also called previsions or prices, depending on the interpretation): Walley's [13] behavioural approach, and Shafer and Vovk's [12] game-theoretic framework, where event trees play a central role. De Cooman and Hermans [1, 2] have shown that these two approaches can be related to each other, and they have introduced imprecise probability trees as a bridge between them. By showing that many results can be imported from one theory into the other, they make some progress towards a more unified handling of uncertainty.

Here, we take one more step towards such a unification, by studying, in Sec. 5, how Walley's epistemic irrelevance [13, Chap. 9] can be related to the notion of event-tree independence that is central in Shafer's discussion of causal reasoning [11]. We discuss the relevance of our findings in the Conclusions, where we also argue why in a theory of uncertain processes, (forward) epistemic irrelevance may be more useful than its symmetrical counterpart, epistemic independence. But let us first recall the basic ideas behind Walley's behavioural theory of coherent lower previsions [13] (Sec. 2), Shafer's event and probability trees [11] (Sec. 3), and the imprecise probability trees that form the connection between them [1, 2] (Sec. 4).

## 2 Coherent lower and upper previsions

In Walley's theory, beliefs held by a subject about the actual value of a random variable $X$ on a finite ${ }^{1}$ space $\mathscr{X}$ are modelled by coherent lower and upper previsions. We call gamble a real-valued function $f$ on $\mathscr{X}$, and denote by $\mathscr{L}(\mathscr{X})$ the set of all gambles on $\mathscr{X} . f(X)$ is interpreted as an uncertain reward. A lower prevision $\underline{P}$ is a real-valued map defined on some subset $\mathscr{K}$ of $\mathscr{L}(\mathscr{X})$. Its conjugate upper prevision $\bar{P}$ is then defined on the set of gambles $-\mathscr{K}:=\{-f: f \in \mathscr{K}\}$ by $\bar{P}(f):=-\underline{P}(-f) \cdot \underline{P}(f)$ is interpreted as the subject's supremum buying price for the uncertain reward $f(X)$, i.e., the smallest price $s$ such that the subject accepts to buy $f(X)$ for any price $\mu<s$, meaning he accepts the uncertain transaction $f(X)-\mu$. Given an event $A \subseteq \mathscr{X}$, its lower probability $\underline{P}(A)$ is the lower prevision of its indicator $I_{A}$, a gamble that assumes the value one on $A$ and zero elsewhere. The upper probability $\bar{P}(A)$ is defined likewise in terms of the upper prevision $\bar{P}\left(I_{A}\right)$. With a lower prevision $\underline{P}$ we can associate a closed convex set of (dominating) probability mass functions: $\mathscr{M}(\underline{P}):=\left\{p \in \Sigma_{\mathscr{X}}:(\forall f \in \mathscr{K})\left(E_{p}(f) \geq \underline{P}(f)\right\}\right.$, where $\Sigma_{\mathscr{X}}$ is the set (simplex) of all probability mass functions on $\mathscr{X}$, and $E_{p}(f):=\sum_{x \in \mathscr{X}} f(x) p(x)$. We call $\mathscr{M}(\underline{P})$ the credal set induced by $\underline{P}$. A lower prevision $\underline{P}$ is said to be coherent if and only if $\mathscr{M}(\underline{P}) \neq \emptyset$ and $\underline{P}(f)=\min \left\{E_{p}(f): p \in \mathscr{M}(\underline{P})\right\}$ for all $f$ in $\mathscr{K}$, i.e., if $\underline{P}$ is the lower envelope of $\mathscr{M}(\underline{P})$.

[^1]
## 3 Event trees

An event tree is composed of situations linked together, and it represents what relevant events may possibly happen in what particular order in the world, according to a particular subject. It is formally equivalent to a rooted tree in graph theory. We restrict ourselves to trees with finite depth and width. The notions we are about to introduce are illustrated in Fig. 1. A situation is a node in the tree. The initial situation is the root of the tree. A terminal situation is a leaf of the tree; all other situations, including the initial one, are called non-terminal. A path is a sequence of situations from the initial to a terminal situation. A path goes through a situation $s$ if $s$ belongs to it. The set $\Omega$ of all possible paths, or equivalently, of all terminal situations, is called the sample space. Any set of terminal situations is an event. Situations immediately following a non-terminal situation $s$ are called daughters of $s$, and the set of such daughters is denoted by $d(s)$. The link between a situation $s$ and one of its daughters $t$ is called a move from $s$ to $t$. If a situation $s$ is before a situation $t$ in the tree, we say that $s$ strictly precedes $t$, and denote this as $s<t$; and if a situation $s$ is before or equal to a situation $t$, we say that $s$ precedes $t$, and denote this as $s \leq t$. Two situations are called disjoint if there is no path they both belong to. A cut is a set of disjoint situations, such that every path goes through exactly one situation in the cut. If each situation in a cut $V$ (strictly) precedes some situation in another cut $U$, then $V$ is said to (strictly) precede $U$, and we denote this as $V \leq U(V<U)$.

Fig. 1 Event tree with nonterminal situations (grey), terminal situations (black), and root $\square . U=\left\{u_{1}, \ldots, u_{4}\right\}$ is a cut, $t<u_{1}$ and $d(t)=$ $\left\{u_{1}, u_{2}\right\}$. Also, $u_{4}$ and $t$ are disjoint, but not $u_{4}$ and $\omega$.


## 4 Imprecise probability trees

Branching probabilities $p_{s}$ for a non-terminal situation $s$ are non-negative numbers summing up to one, each of them attached to a different move originating in $s$ : we denote by $p_{s}(t)$ the probability to go from $s$ to its daughter $t ; p_{s}$ is a probability mass function on $d(s)$. A (precise) probability tree is an event tree for which every non-terminal situation has such branching probabilities.

An imprecise probability tree ${ }^{2}$ is an event tree for which each non-terminal situation $s$ has a closed convex set $\mathscr{M}_{s}$ of branching probabilities $p_{s}$, describing a subject's uncertainty about which move is going to be observed just after $s$. With

[^2]an imprecise probability tree, we can associate coherent lower previsions. First of all, for any non-terminal situation $s$, and for any gamble $h$ on $d(s)$, we can consider the lower prevision $\underline{P}_{s}(h)=\min \left\{E_{p_{s}}(h): p_{s} \in \mathscr{M}_{s}\right\} . \underline{P}_{s}$ and $\mathscr{M}_{s}$ are equivalent local predictive models for what is going to be observed immediately after $s$. But we can also consider global predictive models: Let $f$ be a gamble on the set of paths $\Omega$. For every situation $t$, we consider the lower prevision $\underline{P}(f \mid t)$ conditional on $t$ : the subject's supremum buying price for $f$, given that the actual path goes through $t$.

The global models $\underline{P}(\cdot \mid t)$ can be calculated from the local $\underline{P}_{s}$ by backwards recursion, using the Concatenation Formula $[1,2]$ : for any given situation $t, \underline{P}(f \mid t)=$ $\underline{P}_{t}(\underline{P}(f \mid d(t)))$, where $\underline{P}(f \mid d(t))$ is the gamble on $d(t)$ that assumes the value $\underline{P}(f \mid s)$ in each $s \in d(t)$; and for a terminal situation $\omega \in \Omega$, we have $\underline{P}(f \mid \omega)=f(\omega)$.

Example 1. Let us illustrate this with the successive flipping of two coins. In the corresponding event tree:

the labels for the situations are explicit, e.g., $h$, ? means that the first coin has landed 'heads', and the second still has to be flipped. As indicated on the edges of the tree, the subject's beliefs about the first coin are modelled by the imprecise probability assignments $p(h) \in[1 / 4,3 / 4]$ and $p(t) \in[1 / 4,3 / 4]$. If it lands 'heads', we keep the same coin, otherwise the second flip is made with a fair $\operatorname{coin}(p(h)=p(t)=1 / 2)$. We have also indicated the different steps in the calculation of the lower and upper probability of getting 'heads' at least once, using the Concatenation Formula.

## 5 Forward irrelevance in event trees

Let us briefly recall the notion of forward irrelevance, discussed in detail by De Cooman and Miranda [3], before relating it to independence in event trees.

For two random variables $X_{1}$ and $X_{2}$, if a subject says that $X_{1}$ is epistemically irrelevant to $X_{2}$, this means that he assesses that learning the actual value of $X_{1}$ won't change his beliefs about the value of $X_{2}$. For imprecise probability models, this notion is asymmetric: the epistemic irrelevance of $X_{1}$ to $X_{2}$ is not generally equivalent to the epistemic irrelevance of $X_{2}$ to $X_{1}[5,3]$.

Assume that the uncertainty bears on random variables $X_{1}, \ldots, X_{N}$ that assume values in the respective finite sets $\mathscr{X}_{1}, \ldots, \mathscr{X}_{N}$. For $1 \leq k \leq \ell \leq N$, we denote by $\mathscr{X}_{\ell: k}:=\times_{i=\ell}^{k} \mathscr{X}_{i}$ the Cartesian product of the $k-\ell+1$ sets $\mathscr{X}_{\ell}, \ldots, \mathscr{X}_{k}$, and by $X_{\ell: k}:=\left(X_{\ell}, \ldots, X_{k}\right)$ the associated joint random variable taking values in $\mathscr{X}_{\ell: k}$.

Similarly, $x_{\ell: k}:=\left(x_{\ell}, \ldots, x_{k}\right) \in \mathscr{X}_{\ell: k}$ denotes a generic value of $X_{\ell: k}$. The random variables $X_{1}, \ldots, X_{N}$ are assumed to be logically independent, meaning that $X_{\ell: k}$ can assume all values in $\mathscr{X}_{\ell: k}$, for all $1 \leq \ell \leq k \leq N$. A gamble $f$ defined on $\mathscr{X}_{1: N}$ is called $\mathscr{X}_{\ell: k-m e a s u r a b l e}$ if $f\left(x_{1: N}\right)=f\left(y_{1: N}\right)$ for all $x_{1: N}$ and $y_{1: N}$ in $\mathscr{X}_{1: N}$ such that $x_{\ell: k}=y_{\ell: k}$. We denote by $\mathscr{L}\left(\mathscr{X}_{\ell: k}\right)$ the set of all $\mathscr{X}_{\ell: k}$-measurable gambles, and by $f_{\ell: k}$ a generic gamble in this set. Of course, we identify the index ' $k: k$ ' with ' $k$ '.

An important problem is how to build joint belief models from partial ones. Let us consider the specific example where the $X_{k}$ constitute a stochastic process with time variable $k$, implying in particular that the subject knows in advance that the value of random variable $X_{\ell}$ will be revealed to him before that of $X_{\ell+1}$, where $\ell=1,2, \ldots, N-1$. This leads to a special event tree (also called a standard tree [11, Chap. 2]) where the nodes $s$ have the general form $x_{1: k} \in \mathscr{X}_{1: k}, k=0, \ldots, N$. For $k=0$ there is some abuse of notation, as we let $\mathscr{X}_{1: 0}:=\{\square\}$ and $x_{1: 0}:=\square$. The sets $\mathscr{X}_{1: k}$ constitute special cuts of the tree, where the value of $X_{k}$ is revealed. We have $\mathscr{X}_{1: 1}<\mathscr{X}_{1: 2}<\cdots<\mathscr{X}_{1: N}$, and this sequence of cuts is also called a standard filter [11, Chap. 2]. It is clear that $d\left(x_{1: k}\right)=\left\{x_{1: k}\right\} \times \mathscr{X}_{k+1}$ for $k=0,1, \ldots, N-1$. The sample space of such a tree is $\Omega=\mathscr{X}_{1: N}$, and with the variable $X_{k}$ there corresponds a set $\mathscr{L}\left(\mathscr{X}_{k}\right)$ of $\mathscr{X}_{k}$-measurable gambles on this sample space. For instance, in the standard tree of Example 1, gambles characterising the second coin flip are such that $f(t, h)=f(h, h)$ and $f(t, t)=f(h, t)$. Below, we see the first two cuts of another standard tree, with $\mathscr{X}_{1}=\{a, b\}$ and $\mathscr{X}_{2}=\{\alpha, \beta, \gamma\}$.


A natural way to specify partial beliefs consists in attaching, as explained in the previous section, to each of the non-terminal nodes $x_{1: k}$ a (coherent) local predictive lower prevision $\underline{P}_{x_{1: k}}$ on $\mathscr{L}\left(d\left(x_{1: k}\right)\right)$, i.e., on $\mathscr{L}\left(\mathscr{X}_{k+1}\right)$, where $k=0,1, \ldots, N-1$. This represents a subject's beliefs about the value of $X_{k+1}$ given that the $k$ previous variables $X_{1: k}$ assume the values $x_{1: k}$. For standard imprecise probability trees, the Concatenation Formula given above for deriving the global lower previsions $\underline{P}\left(\cdot \mid x_{1: \ell}\right)$ on $\mathscr{L}\left(\mathscr{X}_{1: N}\right)$ from the local models $\underline{P}_{x_{1: k}}$ completely coincides with the formulae for Marginal Extension, derived by Miranda and De Cooman [9].

A subject may make an assessment of forward irrelevance, meaning that for $1 \leq k \leq N-1$, his beliefs about the 'future' random variable $X_{k+1}$ won't be changed by learning new information about the values of the 'past' random variables $X_{1: k}$ : the past random variables $X_{1}, \ldots, X_{k}$ are epistemically irrelevant to the future random variable $X_{k+1}$, for $1 \leq k \leq N-1$. This is expressed by the following condition involving the local models: for all $0 \leq k \leq N-1$, any gamble $f_{k+1}$ in $\mathscr{L}\left(\mathscr{X}_{k+1}\right)$, and all $x_{1: k}$ in $\mathscr{X}_{1: k}$ :

$$
\begin{equation*}
\underline{\underline{P}}_{x_{1: k}}\left(f_{k+1}\right)=\underline{P}_{k+1}\left(f_{k+1}\right), \tag{1}
\end{equation*}
$$

where $\underline{P}_{k+1}$ is the so-called marginal lower prevision on $\mathscr{L}\left(\mathscr{X}_{k+1}\right)$, which expresses the subject's beliefs about the value of $X_{k+1}$, irrespective of the values assumed by the other random variables. Invoking the Concatenation Formula now leads to a very specific way of combining the marginal lower previsions $\underline{P}_{1}, \ldots, \underline{P}_{N}$ into a joint lower prevision, reflecting the assessment of forward irrelevance. This joint lower prevision, called the forward irrelevant product, is studied in detail by De Cooman and Miranda [3], who also use it to prove very general laws of large numbers [4].

We now proceed to show that forward irrelevance is exactly the same thing as Shafer's notion of event-tree independence, when applied to standard imprecise probability trees. In Shafer's [11] terminology, a situation s influences a variable $X$ if there is at least one situation $t \in d(s)$ such that the subject's beliefs about the value of $X$ are modified when moving from $s$ to $t$; for imprecise probability trees, this means that there should be at least one gamble $f$ whose value depends on the outcome of $X$ for which $\underline{P}(f \mid s) \neq \underline{P}(f \mid t)$. Two variables $X$ and $Y$ are called eventtree independent if there is no situation that influences both of them.

In a standard imprecise probability tree, a situation $x_{1: k}$ influences a variable $X_{m}$ if there is at least one situation $x_{1: k+1}$ in $d\left(x_{1: k}\right)$ and at least one gamble $f_{m}$ on $\mathscr{X}_{m}$ such that $\underline{P}\left(f_{m} \mid x_{1: k}\right) \neq \underline{P}\left(f_{m} \mid x_{1: k+1}\right)$. The only situations $x_{1: k}$ that can influence $X_{m}$ are such that $k<m$, since in all other situations, the value of $X_{m}$ has already been revealed 'for some time'. In addition, it is easy to check that $X_{m}$ is always influenced by any situation $x_{1: m-1}$ in the cut $\mathscr{X}_{1: m-1}$ right before the value of $X_{m}$ is revealed.

Theorem 1. Let $X_{1}, \ldots, X_{N}$ be $N$ random variables. Then there is forward irrelevance, or in other words, the random variables $X_{1: k}$ are epistemically irrelevant to $X_{k+1}$ for $1 \leq k \leq N-1$ if and only if the random variables $X_{1}, \ldots, X_{N}$ are event-tree independent in the corresponding standard imprecise probability tree.

Proof. We deal with the 'only if' part first. Suppose the random variables $X_{1: N}$ are forward irrelevant. Consider any $X_{k}$ and $f_{k} \in \mathscr{L}\left(\mathscr{X}_{k}\right)$, where $1 \leq k \leq N$. Then it follows from the forward irrelevance condition (1) and the Concatenation Formula that $\underline{P}_{k}\left(f_{k}\right)=\underline{P}_{x_{1: k-1}}\left(f_{k}\right)=\underline{P}\left(f_{k} \mid x_{1: k-1}\right)$ for all $x_{1: k-1}$ in $\mathscr{X}_{1: k-1}$. Applying the Concatenation Formula again leads to $\underline{P}\left(f_{k} \mid x_{1: k-2}\right)=\underline{P}_{x_{1: k-2}}\left(\underline{P}\left(f_{k} \mid x_{1: k-2}, \cdot\right)\right)=$ $\underline{\underline{P}}_{x_{1: k-2}}\left(\underline{P}_{k}\left(f_{k}\right)\right)=\underline{P}_{k}\left(f_{k}\right)$, and if we continue the backwards recursion, we see that

$$
\underline{P}_{k}\left(f_{k}\right)=\underline{P}\left(f_{k} \mid x_{1: k-1}\right)=\underline{P}\left(f_{k} \mid x_{1: k-2}\right)=\cdots=\underline{P}\left(f_{k} \mid x_{1: 2}\right)=\underline{P}\left(f_{k} \mid x_{1}\right)=\underline{P}\left(f_{k} \mid \square\right)
$$

This implies that the only situations that (may) influence $X_{k}$ are the ones in the cut $\mathscr{X}_{1: k-1}$ immediately before $X_{k}$ is revealed. Therefore, no situation can influence more than one variable, and there is event-tree independence.

Next, we turn to the 'if' part. Assume that all variables are event-tree independent in the standard tree. This implies that no variable $X_{k}$ can be influenced by a situation $x_{1: \ell}$ corresponding to a time $\ell<k-1$ [If $X_{k}$ were influenced by such a situation, then we know that this situation also always influences $X_{\ell+1}$, and $\ell+1<k$, a contradiction]. So for all $x_{1: k-1} \in \mathscr{X}_{1: k-1}$ and all $f_{k} \in \mathscr{L}\left(\mathscr{X}_{k}\right)$ :

$$
\underline{P}\left(f_{k} \mid x_{1: k-1}\right)=\underline{P}\left(f_{k} \mid x_{1: k-2}\right)=\cdots=\underline{P}\left(f_{k} \mid x_{1: 2}\right)=\underline{P}\left(f_{k} \mid x_{1}\right)=\underline{P}\left(f_{k} \mid \square\right) .
$$

Now of course $\underline{P}\left(f_{k} \mid \square\right)=\underline{P}\left(f_{k}\right)=\underline{P}_{k}\left(f_{k}\right)$, where $\underline{P}_{k}$ is the marginal lower prevision for $X_{k}$, and it follows from the Concatenation Formula that $\underline{P}\left(f_{k} \mid x_{1: k-1}\right)=$ $\underline{P}_{x_{1: k-1}}\left(f_{k}\right)$. This shows that (1) is satisfied, so there is forward irrelevance.

## 6 Conclusions

What is the message we want to convey in this paper? In the theory of coherent lower previsions [13], there are essentially two behavioural notions that generalise classical independence: ${ }^{3}$ epistemic irrelevance and the derived notion of epistemic independence. Assessing that two random variables $X_{1}$ and $X_{2}$ are epistemically independent amounts to assessing that (i) $X_{1}$ is epistemically irrelevant to $X_{2}$, meaning that getting to know the value of $X_{1}$ doesn't change our subject's beliefs about $X_{2}$; and (ii) $X_{2}$ is epistemically irrelevant to $X_{1}$.

Suppose we want to consider a theory of uncertain processes where probabilities aren't necessarily precise. What will be the most useful or meaningful counterpart of the important notion of independence in the classical theory of random processes? There are a number of reasons for preferring the asymmetric notion of epistemic irrelevance, and its generalisation to many variables, called forward irrelevance, to that of epistemic independence. We begin with arguments of perhaps less importance, and then go on to present the most compelling one.

First of all, when a notion that is (more or less) automatically symmetrical, breaks apart into two asymmetrical counterparts when using a more powerful language, symmetry becomes something that has to be justified: it can't be imposed without giving it another thought.

Secondly, an assessment of epistemic independence is stronger, and leads to higher joint lower previsions. As lower previsions represent supremum buying prices, higher values represent stronger commitments, and these may be unwarranted when it is only epistemic irrelevance that our subject really wants to model.

Thirdly, joint lower previsions based on an epistemic irrelevance assessment are generally speaking straightforward to calculate, as the discussion of the Concatenation Formula in Sec. 5 testifies. But calculating joint lower previsions from marginals based on an epistemic independence assessment is quite often a very complicated affair [13, Sec. 9.3.2].

Finally, and most importantly, when considering an uncertain process, the subject knows that the values of the random variables $X_{k}$ will be revealed one after the other, and that the value of $X_{k}$ will be revealed before that of $X_{k+1}$. If he states that $X_{k}$ and $X_{k+1}$ are epistemically independent, this amounts to his assessing that (i) getting to know the value of $X_{k}$ won't change his beliefs about $X_{k+1}$ [forward irrelevance]; and (ii) getting to know the value of $X_{k+1}$ won't change his beliefs about $X_{k}$ [backward irrelevance]. But since the subject knows that he will always know the value of $X_{k}$

[^3]before that of $X_{k+1}$, (ii) is effectively a counter-factual statement for him: "if I got to the value of $X_{k+1}$ first, then learning that value wouldn't affect my beliefs about $X_{k}$ ". It's not clear that making such an assessment has any real value, and we feel it is much more natural in such situations context to let go of (ii) and therefore to resort to epistemic (forward) irrelevance.

This line of reasoning can also be related to Shafer's [10] idea that conditioning is never automatic, and must always be associated with a protocol. A subject can only meaningfully condition a probability model on events that he envisages may happen (according to the established protocol). In the specific situation described above, conditioning the belief model about $X_{k}$ on the variable $X_{k+1}$ could only legitimately be done if it were possible to find out the value of $X_{k+1}$ without getting to know that of $X_{k}$, quod non. Therefore, it isn't legitimate to consider the conditional lower prevision $\underline{P}_{k}\left(\cdot \mid X_{k+1}\right)$ expressing the beliefs about $X_{k}$ conditional on $X_{k+1}$, and we therefore can't meaningfully impose (ii), as it requires that $\underline{P}_{k}\left(\cdot \mid X_{k+1}\right)=\underline{P}_{k}$. Again, this leads to epistemic (forward) irrelevance, instead of epistemic independence.

In his book on causal reasoning [11], Shafer seems to propose the notion of an event tree in order to develop and formalise his ideas about protocols and conditioning. We have seen in Theorem 1 that for standard event trees, which correspond to uncertain processes, the general notion of event-tree independence that he develops in his book, is effectively equivalent to the notion of forward irrelevance.

## References

1. De Cooman G, Hermans F (2007) On coherent immediate prediction: Connecting two theories of imprecise probability. In: De Cooman G, Vejnarova J, Zaffalon M (eds) ISIPTA '07Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications, SIPTA, pp 107-116
2. De Cooman G, Hermans F (2008) Imprecise probability trees: Bridging two theories of imprecise probability. Artificial Intelligence DOI 10.1016/j.artint.2008.03.001, in press
3. De Cooman G, Miranda E (2008a) Forward irrelevance. Journal of Statistical Planning and Inference, DOI 10.1016/j.jspi.2008.01.012, in press
4. De Cooman G, Miranda E (2008b) Weak and strong laws of large numbers for coherent lower previsions. Journal of Statistical Planning and Inference, DOI 10.1016/j.jspi.2007.10.020, in press
5. Couso I, Moral S, Walley P (2000) Examples of independence for imprecise probabilities. Risk Decision and Policy 5:165-181
6. Cozman FG (2000) Credal networks. Artificial Intelligence 120:199-233
7. Cozman FG (2005) Graphical models for imprecise probabilities. International Journal of Approximate Reasoning 39(2-3):167-184, DOI 10.1016/j.ijar.2004.10.003
8. Levi I (1980) The Enterprise of Knowledge. MIT Press, London
9. Miranda E, De Cooman G (2007) Marginal extension in the theory of coherent lower previsions. International Journal of Approximate Reasoning 46(1):188-225, DOI 10.1016/j.ijar.2006.12.009
10. Shafer $G$ (1985) Conditional probability. International Statistical Review 53:261-277
11. Shafer G (1996) The Art of Causal Conjecture. The MIT Press, Cambridge, MA
12. Shafer G, Vovk V (2001) Probability and Finance: It's Only a Game! Wiley, New York
13. Walley P (1991) Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London

[^0]:    Sébastien Destercke
    Institut de Radioprotection et de Sureté Nucléaire (IRSN), bat 702, Cadarache, 13115, St Paul lez Durance, France. e-mail: sebastien.destercke@irsn.fr

    Gert de Cooman
    Ghent University, SYSTeMS Research Group, Technologiepark-Zwijnaarde 914, 9052 Zwijnaarde, Belgium. e-mail: gert.decooman@UGent.be

[^1]:    ${ }^{1}$ To make this discussion as simple as possible, we restrict ourselves to finite spaces throughout, but it is straightforward to extend our results to infinite spaces.

[^2]:    ${ }^{2}$ Shafer [11, Chap. 12] uses the term 'martingale tree'.

[^3]:    ${ }^{3}$ There are other generalisations, such as strong independence [5], but these have a sensitivity analysis interpretation, rather than a behavioural one; see also [13, Chap. 9]. Our comments below don't bear on such other types of independence.

