

Connecting Two Theories of Imprecise Probability

Gert de Cooman

SYSTeMS Research Group

Ghent University

Technologiepark – Zwijnaarde 914, 9052 Zwijnaarde, Belgium

`gert.decooman@UGent.be`

1 Introduction

In recent years, we have witnessed the growth of a number of theories of uncertainty, where imprecise (lower and upper) probabilities, or probability intervals, rather than precise (or point-valued) probabilities, have a central part. Here we consider two of them, Peter Walley’s behavioural theory [8], and Glenn Shafer and Vladimir Vovk’s game-theoretic account of probability [7]. These seem to have a completely different interpretation, and they certainly stem from quite different schools of thought: Walley follows the tradition of Frank Ramsey [6], Bruno de Finetti [3] and Peter Williams [10] 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 bridge to interpret lower and upper probabilities (see [7, 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.

What we set out to do here, is show that in many practical situations, both approaches are very strongly connected. This means that results, valid in one theory, can automatically be converted and reinterpreted in terms of the other.

2 Shafer and Vovk’s game-theoretic approach to probability

In their game-theoretic approach to probability [7], Shafer and Vovk consider a game with two players, World and Skeptic, who play according to a certain *protocol*. They obtain the most interesting results for a special type of protocol, called a *coherent probability protocol*. This section is devoted to explaining what this means.

G1. The first player, World, can make a number of moves, where the possible next moves may depend on the previous moves he has made, but do not in any way depend on the previous moves made by Skeptic.

This means that we can represent his game-play by a (decision) tree. We restrict ourselves here to the discussion of *bounded protocols*, where World can only make a finite and bounded number of moves, whatever happens. But we do not exclude the possibility that at some point in the tree, World has the choice between an infinite number of next moves.

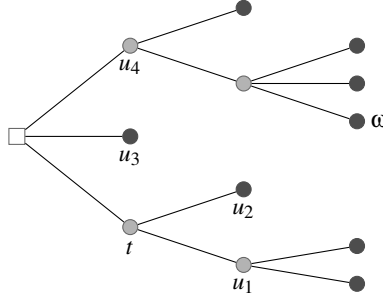


Fig. 1. A simple decision tree for World, displaying the initial situation \square , other non-final situations (such as t) as grey circles, and paths, or final situations, (such as ω) as black circles. Also depicted is a cut of \square , consisting of the situations u_1, \dots, u_4 .

Let us establish some terminology related to World's decision tree. A *path* in the tree represents a possible sequence of moves for World from the beginning to the end of the game. We denote the set of all possible paths ω by Ω , the *sample space* of the game. A *situation* t is some connected segment of a path that is *initial*, i.e., starts at the root of the tree. It identifies the decisions or moves World has made up to a certain point. We denote the set of all situations by Ω^\diamond . It includes the set Ω of *final* situations, or paths. All other situations are called *non-final*; among them is the *initial* situation \square , which represents the empty initial segment. See Figure 1 for a graphical example explaining these notions.

World's *move space* in a non-final situation t is the set \mathbf{W}_t of those moves \mathbf{w} that World can make in t : $\mathbf{W}_t = \{\mathbf{w} : t\mathbf{w} \in \Omega^\diamond\}$.

If for two situations s and t , s is an (initial) segment of t , then we say that s *precedes* t or that t *follows* s , and write $s \sqsubseteq t$. If ω is a path and $t \sqsubseteq \omega$ then we say that the path ω *goes through* situation t . We write $s \sqsubset t$ if $s \sqsubseteq t$ and $s \neq t$.

A function on Ω^\diamond is called a *process*, and a partial process whose domain includes all situations that follow a situation t is called a *t-process*. Similarly, a function on Ω is called a *variable*, and a partial variable on Ω whose domain includes all paths that go through a situation t is called a *t-variable*. If we restrict a *t-process* \mathcal{F} to the final situations that follow t , we obtain a *t-variable*, which we denote by \mathcal{F}_Ω .

We now turn to the other player, Skeptic. His moves may be influenced by the previous moves that World has made, in the following sense. In each situation t , he has a set \mathbf{S}_t of moves \mathbf{s} available to him, called Skeptic's *move space* in t .

G2. In each non-final situation t , there is a (positive or negative) gain for Skeptic associated with each of the possible moves \mathbf{s} in \mathbf{S}_t that World can make. This gain depends only on the situation t and the next move \mathbf{w} that World will make.

This means that for each non-final situation t there is a *gain function* $\lambda_t : \mathbf{S}_t \times \mathbf{W}_t \rightarrow \mathbb{R}$, such that $\lambda_t(\mathbf{s}, \mathbf{w})$ represents the change in Skeptic's capital in situation t when he makes move \mathbf{s} and World makes move \mathbf{w} .

Let us introduce some further notions and terminology related to Skeptic's gameplay. A *strategy* \mathcal{P} for Skeptic is a partial process defined on the set $\Omega^\diamond \setminus \Omega$ of non-final

situations, such that $\mathcal{P}(t) \in \mathbf{S}_t$ is the move that Skeptic will make in each non-final situation t . With each such strategy \mathcal{P} there corresponds a *capital process* $\mathcal{K}^\mathcal{P}$, whose value in each situation t gives us Skeptic's capital accumulated so far, when he starts out with zero capital and plays according to the strategy \mathcal{P} . It is given by the recursion relation

$$\mathcal{K}^\mathcal{P}(t\mathbf{w}) = \mathcal{K}^\mathcal{P}(t) + \lambda_t(\mathcal{P}(t), \mathbf{w}),$$

with initial condition $\mathcal{K}^\mathcal{P}(\square) = 0$. Of course, when Skeptic starts out (in \square) with capital α and uses strategy \mathcal{P} , his corresponding accumulated capital is given by the process $\alpha + \mathcal{K}^\mathcal{P}$. In the final situations, his accumulated capital is then given by the variable $\alpha + \mathcal{K}_\Omega^\mathcal{P}$.

If we start in a non-final situation t , rather than in \square , then we can consider t -strategies \mathcal{P} that tell Skeptic how to move starting from t , and the corresponding capital process $\mathcal{K}^\mathcal{P}$ is then also a t -process, that tells us how much capital Skeptic has accumulated since starting with zero capital in situation t and using t -strategy \mathcal{P} .

Assumptions G1 and G2 determine so-called *gambling protocols*. They are sufficient for us to be able to define upper and lower prices for variables. Consider a non-final situation t and a t -variable f . Then the *upper price* $\overline{\mathbb{E}}_t(f)$ for f in t is defined as the infimum capital α that Skeptic has to start out with in t in order that there would be some t -strategy \mathcal{P} such that his accumulated capital $\alpha + \mathcal{K}^\mathcal{P}$ allows him, at the end of the game, to buy f , whatever moves World makes after t :

$$\overline{\mathbb{E}}_t(f) := \inf \left\{ \alpha : \text{there is some } t\text{-strategy } \mathcal{P} \text{ such that } \alpha + \mathcal{K}_\Omega^\mathcal{P} \geq f \right\}, \quad (1)$$

where $\alpha + \mathcal{K}_\Omega^\mathcal{P} \geq f$ is taken to mean that $\alpha + \mathcal{K}^\mathcal{P}(\omega) \geq f(\omega)$ for all final situations ω that go through t . Similarly, for the *lower price* $\underline{\mathbb{E}}_t(f)$ for f in t :

$$\underline{\mathbb{E}}_t(f) := \sup \left\{ \alpha : \text{there is some } t\text{-strategy } \mathcal{P} \text{ such that } \alpha - \mathcal{K}_\Omega^\mathcal{P} \leq f \right\},$$

so $\underline{\mathbb{E}}_t(f) = -\overline{\mathbb{E}}_t(-f)$. If we start from the initial situation $t = \square$, we simply get the *upper and lower prices* for a variable f , which we also denote by $\overline{\mathbb{E}}(f)$ and $\underline{\mathbb{E}}(f)$.

A gambling protocol is called a *probability protocol* when besides S1 and S2, two more requirements are satisfied.

- P1. For each non-final situation t , Skeptic's move space \mathbf{S}_t is a convex cone in some linear space: $a_1\mathbf{s}_1 + a_2\mathbf{s}_2 \in \mathbf{S}_t$ for all non-negative real numbers a_1 and a_2 and all \mathbf{s}_1 and \mathbf{s}_2 in \mathbf{S}_t .
- P2. For each non-final situation t , Skeptic's gain function λ_t has the following linearity property: $\lambda_t(a_1\mathbf{s}_1 + a_2\mathbf{s}_2, \mathbf{w}) = a_1\lambda_t(\mathbf{s}_1, \mathbf{w}) + a_2\lambda_t(\mathbf{s}_2, \mathbf{w})$ for all non-negative real numbers a_1 and a_2 , all \mathbf{s}_1 and \mathbf{s}_2 in \mathbf{S}_t and all \mathbf{w} in \mathbf{W}_t .

Finally, a probability protocol is called *coherent* when moreover

- C. For each non-final situation t , and for each \mathbf{s} in \mathbf{S}_t there is some \mathbf{w} in \mathbf{W}_t such that $\lambda_t(\mathbf{s}, \mathbf{w}) \leq 0$.

It is clear what this requirement means: for each non-final situation, World has a strategy for playing from t onwards such that Skeptic cannot (strictly) increase his capital from t onwards, whatever t -strategy he uses.

For such coherent probability protocols, Shafer and Vovk prove a number of interesting properties for the corresponding upper (and lower) prices. We list a number of them here. Call a *cut* U of a non-final situation t any set of situations that (i) follow t , and (ii) such that for all paths ω through t [$t \sqsubseteq \omega$], there is a unique $u \in U$ such that ω goes through u [$u \sqsubseteq \omega$]; see also Figure 1. For any t -variable f , we can associate with such a cut U another special t -variable \mathbb{E}_U by $\mathbb{E}_U(f)(\omega) = \mathbb{E}_u(f)$, for all paths ω through t , where u is the unique situation in U that ω goes through. For any two t -variables f_1 and f_2 , $f_1 \leq f_2$ is taken to mean that $f_1(\omega) \leq f_2(\omega)$ for all paths ω that go through t .

Proposition 1 (Properties of prices in a coherent probability protocol [7]). *Consider a coherent probability protocol, let t be a non-final situation, f , f_1 and f_2 t -variables, and U a cut of t . Then*

1. $\inf \{f(\omega) : \omega \in \Omega, t \sqsubseteq \omega\} \leq \mathbb{E}_t(f) \leq \mathbb{E}_t(f) \leq \sup \{f(\omega) : \omega \in \Omega, t \sqsubseteq \omega\}$ [positivity];
2. $\mathbb{E}_t(f_1 + f_2) \leq \mathbb{E}_t(f_1) + \mathbb{E}_t(f_2)$ [sub-additivity];
3. $\mathbb{E}_t(\lambda f) = \lambda \mathbb{E}_t(f)$ for all real $\lambda \geq 0$ [non-negative homogeneity];
4. $\mathbb{E}_t(f + \alpha) = \mathbb{E}_t(f) + \alpha$ for all real α [constant additivity];
5. $\mathbb{E}_t(\alpha) = \alpha$ for all real α [normalisation];
6. $f_1 \leq f_2$ implies that $\mathbb{E}_t(f_1) \leq \mathbb{E}_t(f_2)$ [monotonicity];
7. $\mathbb{E}_t(f) = \mathbb{E}_t(\mathbb{E}_U(f))$ [law of iterated expectation].

What is more, Shafer and Vovk use specific instances of such coherent probability protocols to prove various limit theorems (such as the law of large numbers, the central limit theorem, the law of the iterated logarithm), from which they can derive the well-known measure-theoretic versions. We shall come back to this in Section 5.

3 Walley's behavioural approach to probability

In his book on the behavioural theory of imprecise probabilities [8], Walley considers many different types of related models. We shall restrict ourselves here to the most general and most powerful one, which also turns out to be the easiest to explain, namely coherent sets of desirable gambles; see also [9].

Consider a non-empty set Ω of possible alternatives ω , only one which actually obtains (or will obtain); we assume that it is possible, at least in principle, to determine which alternative does so. Also consider a subject who is uncertain about which possible alternative actually obtains (or will obtain). A *gamble* on Ω is a real-valued function on Ω , and it is interpreted as an uncertain reward, expressed in units of some predetermined linear utility scale: if ω actually obtains, then the reward is $f(\omega)$, which may be positive or negative. If a subject *accepts* a gamble f , this means that she is willing to engage in the transaction, where (i) first it is determined which ω obtains, and then (ii) she receives the reward $f(\omega)$. We can try and model the subject's beliefs about Ω by considering which gambles she accepts.

Suppose our subject specifies some set \mathcal{D} of gambles she accepts, called a *set of desirable gambles*. Such a set is called *coherent* if it satisfies the following *rationality requirements*:

- D1. if $f < 0$ then $f \notin \mathcal{D}$ [avoiding partial loss];
- D2. if $f \geq 0$ then $f \in \mathcal{D}$ [accepting partial gain];
- D3. if f_1 and f_2 belong to \mathcal{D} then their (point-wise) sum $f_1 + f_2$ also belongs to \mathcal{D} [combination];
- D4. if f belongs to \mathcal{D} then its (point-wise) scalar product λf also belongs to \mathcal{D} for all non-negative real numbers λ [scaling].

Here ' $f < 0$ ' means ' $f \leq 0$ and not $f = 0$ '. Walley has also argued that sets of desirable gambles should satisfy an additional axiom:

- D5. \mathcal{D} is \mathcal{B} -conglomerable for any partition \mathcal{B} of Ω : if $I_B f \in \mathcal{D}$ for all $B \in \mathcal{B}$, then also $f \in \mathcal{D}$ [full conglomerability].

Full conglomerability is a very strong requirement, and it is not without controversy. If a model \mathcal{D} is \mathcal{B} -conglomerable, this means that certain inconsistency problems when conditioning on elements B of \mathcal{B} are avoided; see [8] for more details. Conglomerability of belief models was not required by forerunners of Walley, such as Williams [10], or de Finetti [3]. While I agree with Walley that conglomerability is a desirable property for sets of desirable gambles, I do not believe that *full* conglomerability is always necessary: it seems that we need only require conglomerability with respect to those partitions that we actually intend to condition our model on. This is the path I shall follow in Section 4.

Given a coherent set of desirable gambles, we can define *conditional upper and lower previsions* as follows: for any gamble f and any non-empty subset B of Ω , with indicator I_B ,

$$\begin{aligned}\bar{P}(f|B) &:= \inf \{ \alpha : I_B(\alpha - f) \in \mathcal{D} \} \\ \underline{P}(f|B) &:= \sup \{ \alpha : I_B(f - \alpha) \in \mathcal{D} \}\end{aligned}$$

so $\underline{P}(f|B) = -\bar{P}(-f|B)$, and $\bar{P}(f|B)$ is the infimum price α for which the subject will sell the gamble f , i.e., accept the gamble $\alpha - f$, contingent on the occurrence of B . For any event A , we define the conditional lower probability $\underline{P}(A|B) := \underline{P}(I_A|B)$, i.e., the subject's supremum rate for betting on the event A , contingent on the occurrence of B , and similarly for $\bar{P}(A|B) := \bar{P}(I_A|B)$.

If \mathcal{B} is a partition of Ω , then we define $\bar{P}(f|\mathcal{B})$ as the gamble that in any element ω of Ω assumes the value $\bar{P}(f|B)$, where B is the unique element of \mathcal{B} such that $\omega \in B$.

The following properties of conditional upper and lower previsions associated with a coherent set of desirable gambles were (essentially) proven by Walley.

Proposition 2 (Properties of conditional upper and lower previsions [8]). *Consider a coherent set of desirable gambles, let B be any non-empty subset of Ω , f , f_1 and f_2 gambles on Ω . Then¹*

¹ Here, as in Proposition 1, we assume that whatever we write down is well-defined, meaning that for instance no sums of $-\infty$ and $+\infty$ appear, and that the function $\bar{P}(\cdot|B)$ is real-valued, and nowhere infinite. Shafer and Vovk do not seem to mention this.

1. $\inf \{f(\omega) : \omega \in B\} \leq \underline{P}(f|B) \leq \overline{P}(f|B) \leq \sup \{f(\omega) : \omega \in B\}$ [positivity];
2. $\overline{P}(f_1 + f_2|B) \leq \overline{P}(f_1|B) + \overline{P}(f_2|B)$ [sub-additivity];
3. $\overline{P}(\lambda f|B) = \lambda \overline{P}(f|B)$ for all real $\lambda \geq 0$ [non-negative homogeneity];
4. $\overline{P}(f + \alpha|B) = \overline{P}(f|B) + \alpha$ for all real α [constant additivity];
5. $\overline{P}(\alpha|B) = \alpha$ for all real α [normalisation];
6. $f_1 \leq f_2$ implies that $\overline{P}(f_1|B) \leq \overline{P}(f_2|B)$ [monotonicity];
7. if \mathcal{B} is a partition of Ω that refines the partition $\{B, B^c\}$ and \mathcal{D} is \mathcal{B} -conglomerable, then $\overline{P}(f|B) \leq \overline{P}(\overline{P}(f|\mathcal{B})|B)$ [conglomerative property].

The analogy between Propositions 1 and 2 is too striking to be coincidental. The fact that there is an equality in Proposition 1.7, where we have only an inequality in Propositions 2.7, seems to indicate moreover that Shafer and Vovk's approach leads to a less general type of model.² We now set out to identify the exact correspondence between the two models.

4 Connecting the two approaches

In order to lay bare the connections between the game-theoretic and the behavioural approach, we enter Shafer and Vovk's world, and consider another player, called Subject, who has certain *piece-wise* beliefs about what moves World will make.

More specifically, for each non-final situation $t \in \Omega^\diamond \setminus \Omega$, she has beliefs about which move \mathbf{w} World will choose next from the set \mathbf{W}_t of moves available to him in t . We suppose she represents those beliefs in the form of a coherent³ set \mathcal{D}_t of desirable gambles on \mathbf{W}_t . These beliefs are conditional, in the sense that they represent Subject's beliefs about what World will do *immediately after he gets to situation t* . We call any specification of such coherent \mathcal{D}_t , $t \in \Omega^\diamond \setminus \Omega$, a *coherent conditional assessment* for Subject.

We can now ask ourselves what the behavioural implications of these conditional assessments are. For instance, what do they tell us about whether or not Subject should accept certain gambles on Ω , the set of possible paths for World? In other words, how can these beliefs about which next move World will make in each non-final situation t be combined rationally into beliefs about World's complete sequence of moves?

In order to investigate this, we use Walley's very general and powerful method of *natural extension*, which is just *conservative coherent reasoning*. We shall construct, using the pieces of information \mathcal{D}_t , a set of desirable gambles on Ω that is (i) coherent, and (ii) as small as possible, meaning that no more gambles should be accepted than is actually required by coherence.

First, we collect the pieces. Consider any non-final situation $t \in \Omega^\diamond \setminus \Omega$ and any gamble h_t in \mathcal{D}_t . Just as for variables, we can define a t -gamble as a partial gamble

² This also shows that the claim on p. 186 in [7] to the effect that "[de Finetti, Williams and Walley] also considered the relation between unconditional and conditional prices, but they were not working in an dynamic framework and so did not formulate [Shafer and Vovk's equivalent of our Proposition 1.7]", at least needs some qualification.

³ Since we do not envisage conditioning this model on subsets of \mathbf{W}_t , we impose no extra conglomerability requirements here, only the coherence conditions D1–D4.

whose domain contains all paths ω that go through t . Then with each h_t we can associate a t -gamble, also denoted by h_t , and defined by

$$h_t(\omega) := h_t(\omega(t)),$$

for all $t \sqsubseteq \omega$, where we denote by $\omega(t)$ the unique element of \mathbf{W}_t such that $t\omega(t) \sqsubseteq \omega$. If we consider the set $\uparrow t := \{\omega \in \Omega : t \sqsubseteq \omega\}$ of all paths that go through t , then $I_{\uparrow t}h_t$ represents the gamble on Ω that is called off unless World ends up in situation t , and which, when it is not called off, depends only on World's move immediately after t , and gives the same value $h_t(\mathbf{w})$ to all paths ω that go through $t\mathbf{w}$. The fact that Subject accepts h_t on \mathbf{W}_t contingent on World's getting to t , translates immediately to the fact that Subject accepts the gamble $I_{\uparrow t}h_t$ on Ω . We thus end up with a set of gambles on Ω

$$\mathcal{D} := \bigcup_{t \in \Omega^\diamond \setminus \Omega} \{I_{\uparrow t}h_t : h_t \in \mathcal{D}_t\}$$

that Subject accepts. The only thing left to do now, is to find the smallest coherent set $\mathcal{E}_{\mathcal{D}}$ of desirable gambles that includes \mathcal{D} (if there is such a coherent set). Here we take coherence to refer to conditions D1–D4, together with D5', a variation on D5 which refers to conglomerability with respect to those partitions that we actually intend to condition on, as discussed in Section 3.

These partitions are what we call cut partitions. Consider any *non-final* cut $U \subseteq \Omega^\diamond \setminus \Omega$ of the initial situation \square . Then the set of events $\mathcal{B}_U := \{\uparrow u : u \in U\}$ is a partition of Ω , called the U -partition. D5' requires that our set of desirable gambles should be *cut conglomerable*, i.e., conglomerable with respect to every cut partition \mathcal{B}_U .

Because we require cut conglomerability, it follows that $\mathcal{E}_{\mathcal{D}}$ will contain the sums of gambles $\sum_{u \in U} I_{\uparrow u}h_u$ for all non-final cuts U of \square and all choices of $h_u \in \mathcal{D}_u$, $u \in U$. Because $\mathcal{E}_{\mathcal{D}}$ should be a convex cone [by D3 and D4], any sum of such sums $\sum_{u \in U} I_{\uparrow u}h_u$ over a finite number of non-final cuts U should also belong to $\mathcal{E}_{\mathcal{D}}$. But, since in the case of bounded protocols we are discussing here, World can only make a bounded and finite number of moves, $\Omega^\diamond \setminus \Omega$ is a finite union of such non-final cuts, and therefore the sums $\sum_{u \in \Omega^\diamond \setminus \Omega} I_{\uparrow u}h_u$ should belong to $\mathcal{E}_{\mathcal{D}}$ for all choices $h_u \in \mathcal{D}_u$, $u \in \Omega^\diamond \setminus \Omega$.

Call therefore, for any initial situation t , a t -selection any partial process \mathcal{S} defined on the non-final situations $s \sqsupseteq t$ such that $\mathcal{S}(s) \in \mathcal{D}_s$. With such a t -selection, we can associate a t -process, called a *gamble process* $\mathcal{G}^{\mathcal{S}}$, with value

$$\mathcal{G}^{\mathcal{S}}(s) = \sum_{t \sqsubseteq u, u \sqsubseteq s} I_{\uparrow u}\mathcal{S}(u)(s(u))$$

in all situations s that follow t . Alternatively, $\mathcal{G}^{\mathcal{S}}$ is given by the recursion relation $\mathcal{G}^{\mathcal{S}}(s\mathbf{w}) = \mathcal{G}^{\mathcal{S}}(s) + \mathcal{S}(s)(\mathbf{w})$ for all non-final $s \sqsupseteq t$, with initial value $\mathcal{G}^{\mathcal{S}}(t) = 0$. In particular, this leads to the t -gamble $\mathcal{G}_\Omega^{\mathcal{S}}$ defined on all final situations ω that follow t , by letting

$$\mathcal{G}_\Omega^{\mathcal{S}} = \sum_{t \sqsubseteq u, u \in \Omega^\diamond \setminus \Omega} I_{\uparrow u}\mathcal{S}(u).$$

Then we have just shown that the gambles $\mathcal{G}_\Omega^{\mathcal{S}}$ should belong to $\mathcal{E}_{\mathcal{D}}$ for all non-final situations t and all t -selections \mathcal{S} . As before for strategy and capital processes, we call

a \square -selection \mathcal{S} simply a *selection*, and a \square -gamble process simple a *gamble process*. It is now but a small step to prove the following result.

Proposition 3. *The smallest set of gambles that satisfies D1–D4 and D5' and includes \mathcal{D} , or in other words, the natural extension of \mathcal{D} , is given by*

$$\mathcal{E}_{\mathcal{D}} := \left\{ g : \text{there is some selection } \mathcal{S} \text{ such that } g \geq \mathcal{G}_{\Omega}^{\mathcal{S}} \right\}.$$

Moreover, for any non-final situation t and any t -gamble g , we have that $I_{\uparrow t}g \in \mathcal{E}_{\mathcal{D}}$ if and only if $g \geq \mathcal{G}_{\Omega}^{\mathcal{S}}$ for some t -selection \mathcal{S} , where as before, $g \geq \mathcal{G}_{\Omega}^{\mathcal{S}}$ is taken to mean that $g(\omega) \geq \mathcal{G}_{\Omega}^{\mathcal{S}}(\omega)$ for all final situations ω that follow t .

We now use the coherent set of desirable gambles $\mathcal{E}_{\mathcal{D}}$ to define upper (and lower) previsions conditional on the cut partitions \mathcal{B}_U as indicated in Section 3. We then get, using Proposition 3, that for any cut U of \square and any situation u in U :

$$\begin{aligned} \bar{P}(f|\uparrow u) &:= \inf \{ \alpha : I_{\uparrow u}(\alpha - f) \in \mathcal{E}_{\mathcal{D}} \} \\ &= \inf \left\{ \alpha : \text{there is some } u\text{-selection } \mathcal{S} \text{ such that } \alpha - \mathcal{G}_{\Omega}^{\mathcal{S}} \geq f \right\}. \end{aligned} \quad (2)$$

There seems to be a close correspondence between the expressions [such as (1)] for upper prices $\bar{\mathbb{E}}_t(f)$ associated with coherent probability protocols and those [such as (2)] for the conditional upper previsions $\bar{P}(f|\uparrow t)$ based on a coherent conditional assessments. This correspondence is made explicit in the following theorem. Say that a given coherent probability protocol and given coherent conditional assessment *match* whenever they lead to identical corresponding upper prices $\bar{\mathbb{E}}_t$ and conditional upper previsions $\bar{P}(\cdot|\uparrow t)$ for all non-final $t \in \Omega^{\diamond} \setminus \Omega$.

Theorem 1 (Matching Theorem). *For every coherent probability protocol there is a coherent conditional assessment such that both match, and conversely, for every coherent conditional assessment there is a coherent probability protocol such that both match.*

The proof of this result is quite technical, but the underlying ideas should be clear. If we have a coherent probability protocol with move spaces \mathbf{S}_t and gain functions λ_t for Skeptic, define the conditional assessment for Subject to be (essentially) $\mathcal{D}_t := \{ -\lambda(\mathbf{s}, \cdot) : \mathbf{s} \in \mathbf{S}_t \}$. If, conversely, we have a coherent conditional assessment for Subject consisting of the sets \mathcal{D}_t , define the move spaces for Skeptic by $\mathbf{S}_t := \mathcal{D}_t$, and his gain functions by $\lambda_t(h, \cdot) := -h$ for all h in \mathcal{D}_t .

5 Interpretation

The Matching Theorem has a very interesting interpretation. In Shafer and Vovk's approach, World is sometimes decomposed into two players, Reality and Forecaster. It is Reality whose moves are characterised by the above-mentioned decision tree, and it is Forecaster who determines in each non-final situation t what Skeptic's move space

\mathbf{S}_t and gain function λ_t is. We now go beyond Shafer and Vovk's model, by adding something to it.

Suppose that Forecaster has certain beliefs, in each non-final situation t , about what move Reality will make next, and suppose she models those beliefs by specifying a coherent set \mathcal{D}_t of desirable gambles on \mathbf{W}_t . In other words, we identify Forecaster with Subject.

When Forecaster specifies such a set, she is making certain behavioural commitments. In fact, she is committing herself to accepting any gamble in \mathcal{D}_t , and to accepting any combination of such gambles according to the combination axioms D3, D4 and D5'. This implies that we can derive conditional upper previsions $\bar{P}(\cdot|\uparrow t)$, with the following interpretation: in situation t , $\bar{P}(f|\uparrow)$ is the infimum price for which Forecaster can be made to sell the t -gamble f for on the basis of the commitments she has made.

What Skeptic can now do, is take Forecaster up on her commitments. This means that in each situation t , he can select a gamble (or equivalently, any non-negative linear combination of gambles) h_t in \mathcal{D}_t and offer it to Forecaster. If Reality's next move in situation t is $\mathbf{w} \in \mathbf{W}_t$, this means that Skeptic can increase his capital by (the positive or negative amount) $-h_t(\mathbf{w})$, by exploiting Forecaster's commitments. In other words, his move space \mathbf{s}_t can then be identified with the convex set of gambles \mathcal{D}_t and his gain function λ_t is then given by $\lambda_t(h_t, \cdot) = -h_t$. But then Theorem 1 tells us that this leads to a coherent probability protocol, and that the corresponding upper prices $\bar{\mathbb{E}}_t$ for Skeptic coincide with Forecaster's conditional upper previsions $\bar{P}(\cdot|\uparrow t)$.

This is of particular relevance to the laws of large numbers that Shafer and Vovk derive in their game-theoretic framework, because such laws now can be given a behavioural interpretation in terms of Forecaster (or any Subject's) (conditional) lower and upper previsions. To give an example, let us consider the following game.

FINITE-HORIZON BOUNDED FIXED LOWER FORECASTING GAME

Parameters: $N, B > 0, \varepsilon > 0, \alpha > 0$

Players: Reality, Forecaster, Skeptic

Protocol: Forecaster announces $m \in [-B, B]$

$\mathcal{K}_0 := \alpha$

FOR $n = 1, \dots, N$:

Skeptic announces $\lambda_n \geq 0$

Reality announces $x_n \in [-B, B]$

$\mathcal{K}_n = \mathcal{K}_{n-1} + \lambda_n(x_n - m)$.

Winner: Skeptic wins if \mathcal{K}_n is never negative and either $\mathcal{K}_N \geq 1$ or $\frac{1}{N} \sum_{n=1}^N x_n < m - \varepsilon$.
Otherwise Reality wins.

Then Enrique Miranda and I have proven elsewhere [1], amongst other things, that Skeptic has a strategy that guarantees that he wins the game if he starts out with capital $\alpha \geq \exp(-\frac{N\varepsilon^2}{16B^2})$. But this means that for the event

$$\Delta_{N,\varepsilon} := \left\{ (x_1, \dots, x_N) : \frac{1}{N} \sum_{n=1}^N x_n < m - \varepsilon \right\}$$

we have that $\bar{\mathbb{E}}(\Delta_{N,\varepsilon}) \leq \exp(-\frac{N\varepsilon^2}{16B^2})$. We are now able to import this result into the behavioural theory of imprecise probabilities, using Theorem 1. Consider a number of bounded random variables X_1, \dots, X_N , where $X_k \in [-B, B]$, whose values will be revealed successively. Assume that some Subject models her beliefs about the values that these variables assume by specifying, on beforehand, a common lower prevision m for each of them, meaning that she accepts to buy each X_k for any price that is at least m .⁴ Then coherence requires her to bet on the event that the sample mean $\frac{1}{N} \sum_{n=1}^N X_k$ will be at least $m - \varepsilon$ at rates that are higher than $1 - \exp(-\frac{N\varepsilon^2}{16B^2})$, so these rates go to one as N increases, for any $\varepsilon > 0$. This is a weak law of large numbers for bounded random variables.

6 Additional Remarks

We have proven the correspondence between the two approaches only for decision trees with a bounded horizon. For games with infinite horizon, the correspondence becomes less immediate, because Shafer and Vovk implicitly make use of coherence axioms that are stronger than D1–D4 and D5', leading to upper prices that are dominated by the corresponding conditional upper previsions. Exact matching would be restored of course, provided we can argue that these additional requirements are rational for any subject to comply with. This could be an interesting topic for further research.

References

1. G. de Cooman and E. Miranda. Weak and strong laws of large numbers for coherent lower previsions. *Journal of Statistical Planning and Inference*, 2006. Submitted for publication.
2. B. de Finetti. *Teoria delle Probabilità*. Einaudi, Turin, 1970.
3. B. de Finetti. *Theory of Probability*. John Wiley & Sons, Chichester, 1974–1975. English translation of [2], two volumes.
4. P. Gärdenfors and N.-E. Sahlin. *Decision, Probability, and Utility*. Cambridge University Press, Cambridge, 1988.
5. H. E. Kyburg Jr. and H. E. Smokler, editors. *Studies in Subjective Probability*. Wiley, New York, 1964. Second edition (with new material) 1980.
6. F. P. Ramsey. Truth and probability (1926). In R. B. Braithwaite, editor, *The Foundations of Mathematics and other Logical Essays*, chapter VII, pages 156–198. Kegan, Paul, Trench, Trubner & Co., London, 1931. Reprinted in [5] and [4].
7. G. Shafer and V. Vovk. *Probability and Finance: It's Only a Game*. Wiley, New York, 2001.
8. P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.
9. P. Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24:125–148, 2000.
10. P. M. Williams. Notes on conditional previsions. Technical report, School of Mathematical and Physical Science, University of Sussex, UK, 1975.

⁴ Observe that this implies that Subject is not learning about the value of X_k from previous observations of X_1, \dots, X_{k-1} , because she uses the same m , independently of what has happened before. This means that there is some assessment of independence, which we have called [1] forward irrelevance, which is much weaker than the usual independence assessment found in more common weak laws of large numbers.