Game-Theoretic Probability Project Working Paper #3

A New Understanding of Subjective Probability and Its Generalization to Lower and Upper Prevision

Glenn Shafer[1]

Peter R. Gillett[2]

Richard B. Scherl[3]

October 17, 2002

ABSTRACT

This article introduces a new way of understanding subjective probability and its generalization to lower and upper prevision. Instead of asking whether a person is willing to pay given prices for given risky payoffs, we ask whether the person believes he can make a lot of money at those prices. If not—if the person is convinced that no strategy for exploiting the prices can make him very rich in the long run—then the prices measure his subjective uncertainty about the events involved.

This new understanding justifies Peter Walley's updating principle, which applies when new information is anticipated exactly. It also justifies a weaker principle that is more useful for planning because it applies even when new information is not anticipated exactly. This weaker principle can serve as a basis for flexible probabilistic planning in event trees.

[1]Glenn Shafer is on the faculty of the Rutgers Business School—Newark and New Brunswick. He may be contacted at gshafer@andromeda.rutgers.edu.

[2]Peter R. Gillett is on the faculty of the Rutgers Business School—Newark and New Brunswick. He may be contacted at gillett@business.rutgers.edu.

[3]Richard B. Scherl is on the faculty of Monmouth University's Department of Computer Science. He may be contacted at rscherl@monmouth.edu.

Contents

This article introduces a new way of understanding subjective probability and its generalization to lower and upper prevision. Instead of asking whether a person is willing to pay given prices for given risky payoffs, as it is conventional to do [30, 33, 5], we ask whether the person believes he can make a lot of money at those prices. If not—if the person is convinced that no strategy for exploiting the prices can make him very rich in the long run—then the prices measure his subjective uncertainty about the events involved.

This new understanding justifies Peter Walley's updating principle, which applies when new information is anticipated exactly [33]. It also justifies a weaker principle that is more useful for planning because it applies even when new information is not anticipated exactly.

Our analysis relies on Shafer and Vovk's work on the foundations of probability [26] and on Shafer's work on event trees [22]. Shafer and Vovk's explicit game-theoretic protocols impose our distinction between a player who offers prices and a player who tries to beat them, and their version of Cournot's principle provides the basis for our analysis of updating. But event trees are more general and more flexible than Shafer and Vovk's protocols, and their generality and flexibility are needed in planning.

We begin, in §1, by using Shafer and Vovk's protocols and Cournot's principle to study subjective probabilities. In §2, we turn to the more general and more complex theory of subjective lower and upper previsions derived from limited gambling offers. Readers familiar with subjective probability and with Walley's theory of lower and upper previsions will find that these sections cover much familiar ground. This replowing of well-tilled fields has proven necessary for a clear presentation of the differences between our understanding of subjective probability and prevision and the established understanding.

In §3, we generalize lower and upper prevision from Shafer and Vovk's protocols to event trees. In §4, we summarize the message of this article and touch on some other perspectives on that message. Three appendixes deal with some tangential topics: the origin of Cournot's principle, Walley's own presentation of his updating principle, and "incoherence" as a technical term.

Although we state a few mathematical results and point towards practical problems of planning, the purpose of this article is conceptual. It clarifies the informational assumptions that underlie current theories of subjective probability, and it shows how these assumptions can be relaxed and adapted to accommodate tasks where we can foresee only some of the new information that will come our way.

1 Subjective Probability

According to the established understanding of subjective probability, set out by Bruno de Finetti [7] and his followers, a person's beliefs are revealed by the bets he is willing to make. The odds at which he is willing to bet define his probabilities. As time passes, these probabilities change according to the rule of conditional probability: his later probability for an event is his initial conditional probability for it—the condition being what he has learned in the interim.

In this section, we develop a somewhat different understanding of subjective probability, using Shafer and Vovk's game-theoretic framework [26]. In this framework, probability is understood in terms of two players: one who offers bets, and one to whom the bets are offered. We call these two players *House* and *Gambler*, respectively. The established understanding seems to be concerned with House's uncertainty, since he is the one stating odds and offering to bet. But following Shafer and Vovk, we take Gambler's point of view. Gambler is trying to beat the odds, and Shafer and Vovk's work suggests that what makes odds expressions of a person's uncertainty is his conviction that he cannot beat them.

To forestall confusion, we hasten to add that Shafer and Vovk name the two players differently than we do here. In the first part of their book, where they are primarily concerned with an objective conception of probability, Shafer and Vovk call the player who offers bets *Forecaster* and the player to whom the offers are made *Skeptic*. Skeptic's role is to test putative objective probabilities put forward by a theory or a forecasting method. In the second part of their book, where they are concerned with applications to finance, Shafer and Vovk call the player who offers bets *Market* and the player who to whom the offers are made *Investor*. Our names, *House* and *Gambler*, are tailored more to the subjective conception of probability with which we are concerned in this article. Shafer and Vovk do not discuss this subjective conception of probability, and they do not use the names *House* and *Gambler*.

We begin this section by reviewing the established understanding (§1.1) and restating it in terms of explicit protocols (§1.2). As we explain, the established understanding relies heavily on de Finetti's principle that House should avoid sure loss to Gambler.

We then turn to Cournot's principle, which Shafer and Vovk use to give empirical content to their protocols. Roughly speaking, Cournot's principle asserts that Gambler does not have a chance of winning heavily over the long run. As we explain in §1.3, this principle can be used as the basis for both objective and subjective interpretations of probabilities. An objective interpretation is set up by claiming that Gambler cannot win heavily over the long run no matter what he knows. A subjective interpretation is set up by claiming that Gambler is convinced that he cannot win heavily over the long run with the information he actually has.

In §1.4, we turn to the rule of conditional probability. We explain how it is easily justified by the principle that House should avoid sure loss in a protocol in which House must state in advance a rule for changing his probabilities (§1.5) and then why this easy justification fails if no such advance statement is required (§1.6). Then we explain how Cournot's principle can be used to justify the use of conditional probability for updating exactly anticipated new information (§1.7). The new information here is Gambler's, not House's. We show that if Gambler can beat conditional probabilities with his new information, then he would have been able to beat initial probabilities with his initial information. Thus the assumption that initial probabilities are valid (Gambler cannot beat them) implies that the conditional probabilities become equally valid when the new information is received.

In a concluding subsection (§1.8), we summarize our new understanding of subjective probability.

1.1 Offering to Bet

Suppose House announces a subjective probability p for an event E. What does this announcement mean? De Finetti's answer is that House is willing, or at least disposed, to take either side of a bet on E at the odds $p : (1-p)$.

Suppose House does state the odds $p : (1-p)$ and does offer Gambler the opportunity to bet any amount he chooses for or against E at these odds. This means that House offers Gambler the payoff

$$\begin{cases} \alpha(1-p) & \text{if } E \text{ happens} \\ -\alpha p & \text{if } E \text{ fails} \end{cases} \tag{1}$$

for any real number α, which Gambler must choose immediately, before any other information becomes available. The absolute value of α is the total stakes for the bet, and the sign of α indicates which side Gambler is taking:

- If α is positive, then Gambler is betting on E happening. Gambler puts up αp, which he loses to House if E fails, while House puts up $\alpha(1-p)$, which he loses to Gambler if E happens. The total stakes are $\alpha p + \alpha(1-p)$, or α.

- If α is negative, then Gambler is betting against E happening. Gambler puts up $-\alpha(1-p)$, which he loses to House if E happens, while House puts up $-\alpha p$, which he loses to Gambler if E happens. The total stakes are $-\alpha(1-p) - \alpha p$, or $-\alpha$.

No principle of logic requires House to state odds at which Gambler can take either side. But mathematical probability has earned our attention by its practical successes over several centuries, and if we follow de Finetti in rejecting as defective all past attempts to provide objective interpretations of probability, then we seem to be left with (1) as the only viable way of interpreting this successful mathematical theory.

In his publications, spanning more than five decades in the middle of the twentieth century, de Finetti developed this interpretation from the viewpoint of the player we are calling House. The principle that House should avoid sure loss to Gambler was fundamental to this development.

If we agree that House should offer Gambler (1) for some p, then the principle that House should avoid sure loss leads immediately to the conclusion that p should be unique. If House offers (1) for both p_1 and p_2, where $p_1 < p_2$, then Gambler can accept the p_1-offer with $\alpha = 1$ and the p_2-offer with $\alpha = -1$, and this produces a sure gain of $p_2 - p_1$ for Gambler, no matter whether E happens or fails.

Before turning to our alternative understanding of subjective probability, we will further explore, within the Shafer-Vovk formalism, the implications of de Finetti's principle that House should avoid sure loss.

1.2 Protocols

From a thoroughly game-theoretic point of view, the game between House and Gambler also involves a third player, who decides the outcomes on which they are betting. Calling this third player *Reality*, we can lay out an explicit protocol for the game in the style of Shafer and Vovk [26].

Probability Forecasting
House announces $p \in [0, 1]$.
Gambler announces $\alpha \in \mathbb{R}$.
Reality announces $x \in \{0, 1\}$.
$\mathcal{K}_1 := \mathcal{K}_0 + \alpha(x - p)$.

This is a perfect-information protocol; the players move in the order indicated (not simultaneously), and each player sees the other players' moves as they are made. We have written $\mathcal{K}_0$ for Gambler's initial capital and $\mathcal{K}_1$ for his final capital. Reality's announcement indicates the happening or failure of E: $x = 1$ means E happens, and $x = 0$ means E fails. Thus $\alpha(x - p)$ is the same as (1). This is Gambler's net gain, which we can think of as the result of his paying αp for αx ; Gambler buys α units of x for p per unit.

Perfect-information protocols facilitate the exposition of some standard arguments in de Finetti's theory of subjective probability. We now illustrate this point with two of these arguments: de Finetti's argument for the additivity of probability, and an argument for the rule that relates the expected value of a payoff to the probabilities of events that determine the payoff. In both cases, we use de Finetti's principle of House's avoiding sure loss: House should choose his probabilities and other prices so that no strategy for Gambler guarantees Gambler a strictly positive gain no matter how Reality moves.

1.2.1 De Finetti's Argument for Additivity

Consider the following protocol, where House gives probabilities for the three events E, F, and $E \cup F$:

Multiple Probability Forecasting
House announces $p_E, p_F, p_{E\cup F} \in [0, 1]$.
Gambler announces $\alpha_E, \alpha_F, \alpha_{E\cup F} \in \mathbb{R}$.
Reality announces $x_E, x_F, x_{E\cup F} \in \{0, 1\}$.
$\mathcal{K}_1 := \mathcal{K}_0 + \alpha_E(x_E - p_E) + \alpha_F(x_F - p_F) + \alpha_{E\cup F}(x_{E\cup F} - p_{E\cup F})$.
Constraint on Reality: Reality must make $x_{E\cup F} = x_E + x_F$ (this expresses the assumptions that E and F are disjoint and that $E \cup F$ is their disjunction).

The constraint on Reality is part of the rules of the game. Like the other rules, it is known to the players at the outset.

To see that House must make $p_{E\cup F} = p_E + p_F$ in order to avoid sure loss in this protocol, set

$$\delta := \begin{cases} 1 & \text{if } p_{E\cup F} > p_E + p_F \\ 0 & \text{if } p_{E\cup F} = p_E + p_F \\ -1 & \text{if } p_{E\cup F} < p_E + p_F \end{cases}$$

and consider the strategy for Gambler in which α_E and α_F are equal to δ and $\alpha_{E\cup F}$ is equal to $-\delta$. Gambler's net gain with this strategy is

$$\delta(x_E - p_E) + \delta(x_F - p_F) - \delta(x_{E\cup F} - p_{E\cup F}) = \delta(p_{E\cup F} - (p_E + p_F)),$$

which is positive unless $p_E + p_F = p_{E\cup F}$.

1.2.2 Expected Value

The preceding argument generalizes to an argument for determining prices for payoffs. Suppose $E_1, \ldots, E_n$ are disjoint events, and suppose x is a payoff that depends on the outcomes of these events. We write E_0 for the event that none of the $E_1, \ldots, E_n$ happen, so that

$$x = \sum_{j=0}^{n} a_j x_j,$$

where x_j is the indicator variable for E_j, and a_j is the value of x when E_j happens.

The following protocol says that House announces probabilities for $E_0, E_1, \ldots, E_n$ and also a price at which Gambler can buy or sell x:

Pricing a Payoff

House announces $p_0, p_1, \ldots, p_n \in [0, 1]$ and $p \in \mathbb{R}$.
Gambler announces $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $\alpha \in \mathbb{R}$.
Reality announces $x_0, x_1, \ldots, x_n \in \{0, 1\}$ and $x \in \mathbb{R}$.
$\mathcal{K}_1 := \mathcal{K}_0 + \sum_{j=0}^{n} \alpha_j(x_j - p_j) + \alpha(x - p)$.

Constraint on Reality: Reality must make $x = \sum_{j=0}^{n} a_j x_j$.

In this protocol, House must satisfy

$$p = \sum_{j=0}^{n} a_j p_j \tag{2}$$

in order to avoid sure loss. To state the strategy for Gambler that will make money for sure if House violates (2), we set

$$\delta := \begin{cases} 1 & \text{if } p > \sum_{j=0}^{n} a_j p_j \\ 0 & \text{if } p = \sum_{j=0}^{n} a_j p_j \\ -1 & \text{if } p < \sum_{j=0}^{n} a_j p_j. \end{cases}$$

Setting each α_j equal to δa_j and α equal to $-\delta$, we obtain

$$\sum_{j=0}^{n} \delta a_j (x_j - p_j) - \delta(x - p) = \delta(p - \sum_{j=0}^{n} a_j p_j),$$

as Gambler's net gain, and this is positive unless (2) holds.

The number p, the price at which Gambler can buy or sell x, is called the *expected value* or the *prevision* of x. Equation (2) tells us how this price is determined by the probabilities of the disjoint events $E_0, E_1, \ldots, E_n$.

1.3 Trying to Beat the Odds

Shafer and Vovk argue that perfect-information protocols provide a framework in which to understand a broad range of applications of mathematical probability. In many of these applications, the role of House is played by a theory or a model, which gives probabilities that should hold in various situations. A statistician or scientist who wants to test the theory can play the role of Gambler, trying to find a strategy that refutes the theory by making a lot of money.

At first glance, this kind of testing by Gambler might seem relevant only to an objective concept of probability. Indeed, our review of de Finetti's theory suggests that subjective probability is probability from the viewpoint of House, not Gambler. The rules of probability result, it would seem, from House's motivation to avoid sure loss. But as we now show, a clear understanding of how probabilities should change over time requires that we shift to Gambler's viewpoint and invoke Cournot's principle, thus bringing our concept of subjective probability closer to the concept of objective probability. When we assert that certain numbers are valid as objective probabilities, we are asserting that they do not offer anyone any opportunity to get very rich. When we advance them as our subjective probabilities, we are saying something only a little different: we are asserting that they do not offer us, with the knowledge we have, any opportunity to get very rich. When we say this, we put ourselves in the role of Gambler, not in the role of House. The point is not how we got the numbers; perhaps we got them from a theory or from a different person. The point is what we think we can do with them.

1.3.1 Cournot's Principle

We now consider in more detail how Gambler can test probabilities.

A probability for a single event, if it is not equal to 0 or 1, can hardly be refuted. Even if Gambler chooses the winning side, with stakes high enough to make a lot of money, we will hesitate to conclude that the probability was wrong. Gambler may simply have been lucky. On the other hand, if House announces probabilities for a sequence of events, and Gambler consistently manages to make money, then the validity of the probabilities will be cast in doubt.

Shafer and Vovk [26] have shown that we can make this notion of testing precise within the following protocol, where House announces probabilities $p_1, p_2, \dots$ for a series of events $E_1, E_2, \dots$ with indicator variables $x_1, x_2, \dots$:

Sequential Probability Forecasting
$\mathcal{K}_0 := 1$.
For $n = 1, 2, \dots$:
House announces $p_n \in [0, 1]$.
Gambler announces $\alpha_n \in \mathbb{R}$.
Reality announces $x_n \in \{0, 1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + \alpha_n(x_n - p_n)$.

In this protocol, Gambler can test House's probabilities by trying to get infinitely rich ($\lim_{n\to\infty} \mathcal{K}_n = \infty$) without ever risking bankruptcy (without giving Reality an opportunity to make $\mathcal{K}_n$ negative for any n). If Gambler succeeds in doing this, he has refuted the probabilities.

We ask the reader not to take the infinitary aspects of this formulation too seriously. Instead of talking about Gambler multiplying his capital by an infinite factor in an infinite number of trials, we can instead talk about his multiplying it by a large factor in a large number of trials. Because the details of the finitary formulation are relatively complicated ([26], Chapter 1), we leave them aside here. But we will sometimes speak of Gambler "winning heavily" or "becoming very rich" instead of "becoming infinitely rich".

Why do we insist that Gambler begin with limited initial capital $\mathcal{K}_0$ (it is important only that $\mathcal{K}_0$ be positive and finite, not that it have the particular positive value 1) and require that he avoid risking bankruptcy in order for his getting very rich to be a refutation of House? One reason for this formulation is to rule out Gambler's consistently making money by doubling his bet following every loss until he scores a large gain ([26], p. 51). But its real justification is Shafer and Vovk's demonstration that it provides a new and more general foundation for the classical limit theorems of probability. For example, instead of proving that the convergence of empirical frequency to probability occurs "except on a set of measure zero" (this is the textbook formulation of the law of large numbers), Shafer and Vovk prove that the convergence occurs unless Reality permits Gambler (or Skeptic, as they call him) to become infinitely rich without risking bankruptcy. More precisely, Gambler has a strategy (a rule for moving based on House's and Reality's previous moves) that does make him infinitely rich without risking bankruptcy unless Reality's moves converge as required.

Shafer and Vovk use the name *Cournot's principle* for the hypothesis that Reality will not allow Gambler to become infinitely rich without risking bankruptcy (see Appendix A). This principle says that no matter what bankruptcy-free strategy for Gambler we specify (in addition to House's and Reality's previous moves, such a strategy may also use other information available to Gambler), we can be confident that Reality will move in such a way that the strategy will not make Gambler infinitely rich. This is an empirical hypothesis—a hypothesis about how Reality will behave, not a rule of the game.

If given probabilities satisfy Cournot's principle for any potential gambler, no matter how much information that gambler has, then we might call them *objective* or *causal* probabilities [22, 24]. On the other hand, if they satisfy Cournot's principle only for gamblers with a certain level of information, then we might call them *subjective* probabilities for that level of information. An individual who believes that given probabilities do not permit any bankruptcy-free strategy to make him very rich—whether he made up the probabilities or obtained them from a theory or another person—might reasonably call them his personal subjective probabilities.

Under this interpretation, a person with subjective probabilities is not merely saying that he does not know how to get very rich betting at these probabilities. He is saying much more. He is saying that he is convinced that no bankruptcy-free strategy that he might try can make him very rich. He will at best more or less break even.

The appeal of Cournot's principle is strengthened by a result of A. P. Dawid, which he calls *Jeffreys's Law* in honor of the applied mathematician and probabilist Harold Jeffreys. Roughly speaking, Jeffreys's Law says that if two different systems of probabilities satisfy Cournot's principle, then they will be asymptotically equal [4].

1.3.2 Suppressing House

A person might have a rule for determining his subjective probabilities based on what has happened in the world so far. In this case, the person can express the meaning of his probabilities by putting himself in the role of Gambler in a game in which the rule replaces House.

Suppose the rule gives p_n as a function of Reality's moves so far:

$$p_n := \mathcal{P}(x_1, \dots, x_{n-1}),$$

where $\mathcal{P}$ is a function that assigns a number in the interval $[0, 1]$ to every finite sequence of 0s and 1s, including the empty sequence. Such a function $\mathcal{P}$ can be thought of as a strategy for House in the sequential probability forecasting protocol. Once it is fixed, House has no decisions to make, and we no longer need to consider him as a player in the game. Writing $\{0, 1\}^*$ for the set of all finite sequences of 0s and 1, we can then describe the protocol like this:

SEQUENTIAL PROBABILITY WITHOUT HOUSE
Parameter: $\mathcal{P} : \{0, 1\}^* \mapsto [0, 1]$
  $\mathcal{K}_0 := 1$.
  For $n = 1, 2, \dots$:
    Gambler announces $\alpha_n \in \mathbb{R}$.
    Reality announces $x_n \in \{0, 1\}$.
    $\mathcal{K}_n := \mathcal{K}_{n-1} + \alpha_n(x_n - \mathcal{P}(x_1, \dots, x_{n-1}))$.

The parameter $\mathcal{P}$, being part of the rules of the game, is known to both players at the outset.

Our discussion of Cournot's principle applies to this simplified protocol just as well as to the protocol in which House plays a free role. See Chapter 3 of [26].

In the context of Shafer and Vovk's general theory, the protocol without House is fairly special. It moves us closer, however, to the problem of planning using subjective probabilities. Planning requires that we have some advance information about how future events will affect future possibilities and probabilities.

1.3.3 Generalizing to Event Trees

For clarity, we have introduced Cournot's principle using a relatively simple protocol, in which Reality has a binary choice at each step. The principle can also be adopted, however, when Reality sometimes has more than two choices, and when the choices available to her may depend on what she has done previously. This brings us to the generality of an event tree [19].

Figure 1 illustrates the idea of an event tree, which we study in more detail in §3. In this example, Gambler is watching the actions of a youngster. The youngster may deliberate about his actions, but from Gambler's point of view, these actions are moves by Reality. Gambler somehow knows in advance that the youngster will first either watch television, call a friend, or play his saxophone. What the youngster may do next depends on this first choice.

Gambler may have at the outset probabilities for each step, as indicated on the right-hand side of the figure. In this case, he is playing a game with Reality without House. Alternatively, he may wait until the first step is taken before finding out or deciding his probabilities for the next step; in this case, House is in the game. In either case, Gambler may adopt the probabilities as his subjective probabilities by subscribing to Cournot's principle, provided only that the tree actually continues indefinitely rather than ending after two steps as in the figure.

1.3.4 Summary

The established understanding of subjective probability, associated with the name of Bruno de Finetti, attributes subjective probabilities to a person when he is willing to bet at those probabilities. But we attribute subjective probabilities to a person when he believes that he cannot win heavily when he has the opportunity to use them as betting rates. We formalize this idea by putting the person in the role of Gambler, who plays a game against Reality. The person's belief about his own inability to beat the probabilities is then

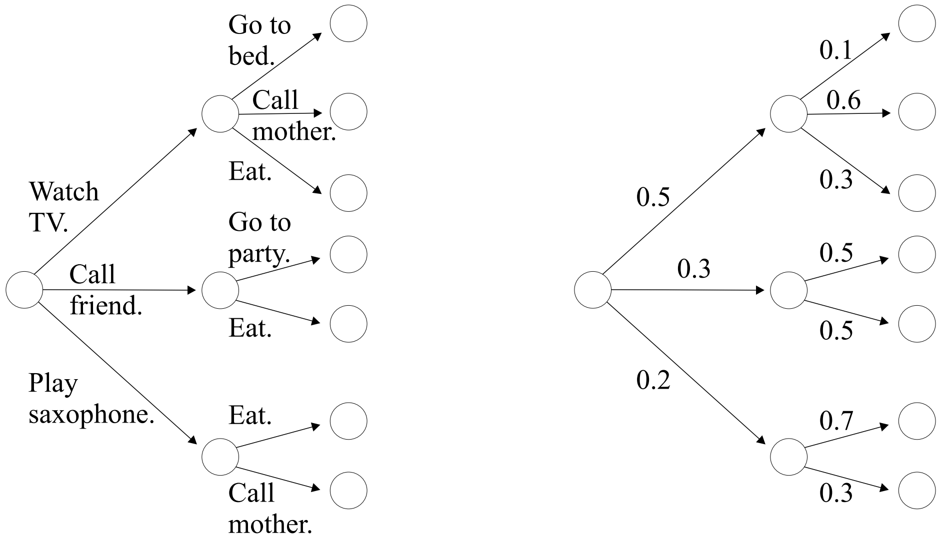


Figure 1: An event tree. The steps are possible moves by Reality. On the right, probabilities are assigned to Reality's moves, making the event tree into a probability tree. The probabilities shown are probabilities for steps. Later we discuss also probabilities for histories. The probability for the history "watch TV, then go to bed", for example, is 0.5×0.1, or 0.05.

---

formalized by Cournot's principle, which states expectations about Reality's behavior.

There are many ways a person might obtain the numbers he adopts as subjective probabilities. He may have a rule for calculating these numbers. He may rely on someone else, such as a weather forecaster, to provide them. Or he may make them up himself as he goes along. Our sequential protocol without House formalizes the case where he has a rule. Our sequential protocol with House formalizes the case where someone chooses and announces the numbers as events proceed. Both types of protocol generalize to event trees. The protocol without House is the most interesting for planning; it generalizes to probability trees—event trees with probabilities assigned to the steps.

1.4 Conditional Probability

The usual mathematical theory of probability [2, 8, 10, 11, 27] uses the concept of conditional probability to deal with changes in probability over time. If we write $\mathbb{P}_0$ for probabilities at time 0 and $\mathbb{P}_t$ for probabilities at time t, then the theory says that

$$\mathbb{P}_0[A_t \cap E] = \mathbb{P}_0[A_t]\,\mathbb{P}_t[E], \tag{3}$$

where A_t represents what has happened by time t. Equation (3) is often called the *rule of compound probability*.

Although our notation does not make the dependence explicit, the probability $\mathbb{P}_t[E]$ depends on A_t (what has happened by time t), not merely on t. It is the *conditional probability* of E given A_t. If $\mathbb{P}_0[A_t] \neq 0$, then we can rewrite (3) to express $\mathbb{P}_t[E]$ in terms of probabilities at time 0:

$$\mathbb{P}_t[E] = \frac{\mathbb{P}_0[A_t \cap E]}{\mathbb{P}_0[A_t]}. \tag{4}$$

This equation, the *rule of conditional probability*, is a rule of updating: it tells us how probabilities at time t are determined by initial probabilities and what has happened by time t.

The rule of conditional probability plays a particularly important role in the theory of subjective probability. It seems quite remarkable, in fact, that a rule of this type exists for subjective probabilities. Its existence suggests that once a person has announced a complete set of initial subjective probabilities, he has no future work to do; his future subjective probabilities are determined for him. We might be able to wiggle out of this conclusion when $\mathbb{P}_0[A_t] = 0$ [3], but most scholars who study subjective probability do not try to do so; instead, they glory in the coherence of new with old beliefs represented by (3) and (4) [1, 13, 18].

Why should (3) hold for subjective probabilities? The usual answer to this question, which goes back to the eighteenth-century work of De Moivre and Bayes [20, 21], relies on constructing a bet on E at time t from bets on A_t and $A_t \cap E$ at time 0. The probability (4), it is argued, can be justified by the cost of this construction. Spelling this argument out in detail involves dealing with the question of timing. When does House offer bets, and when does Gambler have to accept them? Different assumptions about the timing lead to different versions of the argument, some more convincing than others.

We now take a careful look at several of the different versions. First (§1.5 and §1.6) we look at what can be said when we take House's point of view and rely only on de Finetti's principle that House should avoid sure loss, and then (§1.7) we look at what can be said when we shift to Gambler's point of view and invoke Cournot's principle.

1.5 Announcing Future Probabilities in Advance

Updating by conditional probability is most inescapable under the assumption that House announces at time 0 how the probability he will announce for E at time t depends on what happens by then.

Say $A_t^1, \ldots, A_t^k$ are the possibilities House foresees for what will happen by time t. In order to write protocols that make assumptions about timing explicit, we adopt the following notation:

- p_j is House's probability for A_t^j at time 0,
- q_j is House's probability for E at time t if A_t^j happens,
- r_j is House's probability for $A_t^j \cap E$ at time 0,
- i is the index for which A_t^i actually happens, and
- x is 1 if E happens and 0 if it fails.

In the following protocol, House commits himself to the q_j in advance.

Advance Probability Forecasting
 At time 0:
  House announces $p_1, \ldots, p_k, r_1, \ldots, r_k, q_1, \ldots, q_k \in [0,1]$.
  Gambler announces $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k \in \mathbb{R}$.
  Reality announces $i \in \{1, 2, \ldots, k\}$.
 At time t:
  Gambler announces $\gamma \in \mathbb{R}$.
  Reality announces $x \in \{0, 1\}$.
 $\mathcal{K}_t := \mathcal{K}_0 + (\alpha_i - \sum_{j=1}^k \alpha_j p_j) + (\beta_i x - \sum_{j=1}^k \beta_j r_j) + \gamma(x - q_i)$.

As usual, the Greek letters represent stakes for the different bets: α_j is the total stakes for A_t^j, β_j the total stakes for $A_t^j \cap E$, and γ the total stakes for E.

Gambler's net gain, $\mathcal{K}_t - \mathcal{K}_0$, can be written in the form

$$(\alpha_i - \gamma q_i) + x(\beta_i + \gamma) - \sum_{j=1}^{k} (\alpha_j p_j + \beta_j r_j). \qquad (5)$$

This expression makes it easy to show that House must obey (3) in order to avoid sure loss. First we set

$$\delta_j := \begin{cases} 1 & \text{if } r_j > p_j q_j \\ 0 & \text{if } r_j = p_j q_j \\ -1 & \text{if } r_j < p_j q_j. \end{cases}$$

for $j = 1, \ldots, k$. Then we prescribe for Gambler the strategy given by

$$\alpha_j := \delta_j q_j, \qquad \beta_j := -\delta_j, \qquad \text{and} \qquad \gamma := \delta_i. \qquad (6)$$

With this strategy, Gambler's net gain, (5), becomes

$$(\delta_i q_i - \delta_i q_i) + x(-\delta_i + \delta_i) - \sum_{j=1}^{k} (\delta_j q_j p_j - \delta_j r_j) = \sum_{j=1}^{k} \delta_j (r_j - p_j q_j).$$

By the definition of δ_j, the product $\delta_j(r_j - p_j q_j)$ is always nonnegative and is positive unless $r_j = p_j q_j$. So House must make $r_j = p_j q_j$ for all j in order to keep Gambler from making a sure gain with this strategy. In particular, he must make $r_i = p_i q_i$ hold, and this is merely another way of writing (3), the rule of compound probability.

Let us summarize. We made these assumptions:

1. House knows at time 0 the possibilities for what he will have learned by time t.

2. House announces at time 0 joint probabilities for these possibilities and the event E.

3. House also announces at time 0 how his new probability for E at time t will depend on how what he has learned by then.

We deduced from these assumptions that House's advance announcements must conform with (3), the rule of compound probability, if he is to avoid sure loss. Assuming further that House did not assign probability zero at time

0 to what actually happened by time t, it follows that his announcements must also conform with (4), the rule of conditional probability. *If House announces in advance a rule for updating his probabilities, then he can avoid sure loss only if it is the usual rule of conditional probability.*

This argument generalizes readily to an event tree. If Gambler is given at the outset probabilities for each step in the tree (this gives him a probability tree, as in Figure 1) and also initial probabilities for histories (sequences of steps), then the probabilities for the steps should be related to the probabilities for the histories by the rule of conditional probability; otherwise Gambler will have a way to make money for sure.

In another important direction, however, the argument does not generalize. It depends crucially on the assumption that Gambler can freely switch the sign of any payoff he is offered—i.e., that he can take either side of any bet he is offered. When we drop this assumption, so that we have only lower and upper probabilities and previsions instead of subjective probabilities and exact expected values, we will have to turn to an alternative argument based on Cournot's principle (see §2.4).

1.6 Updating When the Time Comes

The preceding argument does not apply when House does not announce a rule of updating in advance. This becomes clear when we think about the following protocol, in which House announces only at time t what he is willing to do at time t:

TWO-STAGE PROBABILITY FORECASTING
At time 0:
House announces $p_1, \dots, p_k, r_1, \dots, r_k \in [0, 1]$.
Gambler announces $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k \in \mathbb{R}$.
Reality announces $i \in \{1, 2, \dots, k\}$.
At time t:
House announces $q \in [0, 1]$.
Gambler announces $\gamma \in \mathbb{R}$.
Reality announces $x \in \{0, 1\}$.
$\mathcal{K}_t := \mathcal{K}_0 + (\alpha_i - \sum_{j=1}^k \alpha_j p_j) + (\beta_i x - \sum_{j=1}^k \beta_j r_j) + \gamma(x - q)$.

The symbols in this protocol are the same as in the protocol for advance probability forecasting, except that we now write simply q for House's probability for E at time t; he does not announce probabilities that he would have had at time t had events gone differently between time 0 and time t.

It is intuitively clear that House's announcement of q at time t in this two-stage protocol should not be constrained by his earlier announcements at time 0. In addition to Reality's announcement of i, House and Gambler might learn any number of things between time 0 and time t, and so there is no reason to suppose that House's new probability for E should depend only on i and his own previous announcements. This intuition is confirmed by the fact that the strategy (6), which enforces the rule of conditional probability in the advance protocol, is not available to Gambler in the two-stage protocol. House does not announce probabilities q_j at time 0 in the two-stage protocol, and (6) depends on having such announcements.

1.6.1 Advance Offers

Although House can do what he wants at time t without risking sure loss, it remains true that his offers at time 0 include what can be interpreted as offers to agree at time 0 to bets at time t. Indeed, from the offers House makes at time 0, Gambler can construct a payoff that will come out the same as the payoff of a bet on E made at time t. To do this, Gambler fixes α and sets

$$\alpha_j := -\frac{r_j}{p_j}\alpha \qquad \text{and} \qquad \beta_j := \alpha \tag{7}$$

for $j = 1, \ldots, k$. Gambler's net gain from these moves will be

$$\alpha(x - \frac{r_i}{p_i}), \tag{8}$$

where i is Reality's announcement. This is the same as the net gain Gambler would have on a bet on E at probability r_i/p_i and total stakes α, made at time t after Reality announces i.

If House's announcement q at time t is different from r_i/p_i, we can say he has changed his mind about offering (8). He offered (8) at time 0, but now he is offering $\alpha(x-q)$. But because Gambler did not know at time 0 whether q would be greater than r_i/p_i or less than r_i/p_i, Gambler cannot exploit this change to inflict a sure loss on House.

1.6.2 A Single Advance Offer

Gambler can also choose a particular value j_0 at time 0 and construct a payoff that looks like a bet on E at time t only if Reality chooses j_0. To do this, he fixes α, sets

$$\alpha_{j_0} := -\frac{r_{j_0}}{p_{j_0}}\alpha \qquad \text{and} \qquad \beta_{j_0} := \alpha, \tag{9}$$

and sets $\alpha_j = \beta_j = 0$ for $j \neq j_0$. His net gain from these moves will be

$$\begin{cases} \alpha(x - \frac{r_{j_0}}{p_{j_0}}) & \text{if } i = j_0 \\ 0 & \text{if } i \neq j_0. \end{cases} \tag{10}$$

This is the same as the payoff from a contingent bet on E made at time 0—an agreement to bet with stakes α on E at the conditional probability if Reality chooses j_0 at time t but not to bet at all if Reality does not choose j_0.

1.6.3 Walley's Updating Principle

In his treatise *Statistical Reasoning with Imprecise Probabilities* [33], Peter Walley states a general principle about how a person's betting offers should change when he obtains new information. Using our own terminology rather than Walley's, we may state his principle by saying that House should offer at time 0 a particular payoff that pays nothing when B fails if and only if he intends to continue to offer this payoff if and when he learns of B's happening and nothing more. We state the principle more carefully in §2.3, and we quote Walley's own statement of it in Appendix B. Here we merely note that in the case at hand, where House is offering (10) at time 0, the principle implies that at time 0 House should intend to offer at time t to pay r_{j_0}/p_{j_0} for x—i.e., to use r_{j_0}/p_{j_0} as his new probability for E—if Reality announces j_0 and this is House's only new information.

Aside from observing that the payoff (10) is the same whether one bets at time 0 or time t, Walley gives little argument for his principle. He appears to regard it as a relatively self-evident principle of rationality that underlies the widespread acceptance of conditional probability. As we will now explain, Cournot's principle leads to a clear argument for Walley's principle, an argument that makes clear why the caveat "this is House's only new information" is needed.

1.7 Updating with Exact Information

So far (§1.5–§1.6), our attempts to justify the rule of conditional probability have remained within de Finetti's understanding, which attributes subjective probabilities to House, the player who announces them. We have been looking for arguments that constrain House to obey the rule of conditional probability. We now turn to look at the matter from the viewpoint we developed in §1.3—the viewpoint of Gambler, who thinks he cannot beat the probabilities. It is this viewpoint that permits a clear argument for Walley's updating principle.

Consider a sequence of events $E_1, E_2, \ldots$. These events may be substantively very different, but for simplicity let us suppose that House makes announcements about each of them as in the two-stage protocol of §1.6. First, at time n, House announces probabilities for k events and for their conjunctions with E_n, and Reality decides which of the k events happens. A little later, say at time $n+1/2$, House gives a new probability for E_n, and Reality decides whether E_n happens. This produces the following protocol:

Sequential Two-Stage Probability Forecasting
$\mathcal{K}_0 := 1$.
For $n = 1, 2, \ldots$
At time n:
House announces $p_{n1}, \ldots, p_{nk} \in (0,1]$ and $r_{n1}, \ldots, r_{nk} \in [0,1]$.
Gambler announces $\alpha_{n1}, \ldots, \alpha_{nk}, \beta_{n1}, \ldots, \beta_{nk} \in \mathbb{R}$.
Reality announces $i_n \in \{1, 2, \ldots, k\}$.
At time $n+1/2$:
House announces $q_n \in [0,1]$.
Gambler announces $\gamma_n \in \mathbb{R}$.
Reality announces $x_n \in \{0,1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + (\alpha_{ni_n} - \sum_{j=1}^k \alpha_{nj} p_{nj}) + (\beta_{ni_n} x_n - \sum_{j=1}^k \beta_{nj} r_{nj}) + \gamma_n (x_n - q_n)$.

For simplicity, we have required House to make the p_{ni} nonzero, so that we can freely use them as divisors.

Let us make the following assumptions:

1. House's first-stage announcements (his $p_{n1}, \ldots, p_{nk}$ and $r_{n1}, \ldots, r_{nk}$ for $n = 1, 2, \ldots$) satisfy Cournot's principle: Reality will not allow Gambler to become infinitely rich following a bankruptcy-free strategy for betting at these probabilities.

2. House agrees in advance that his second-stage announcements will obey the rule of conditioning: he will always set q_n equal to r_{ni_n}/p_{ni_n}.

3. The only new information Gambler acquires between his move at time n and his move at time $n+1/2$ is Reality's choice of of i_n. (By the preceding assumption, he already knows how this will determine House's move q_n.)

4. Reality pays no attention to how Gambler moves when she chooses her moves.

Will all of House's announcements (the $p_{n1}, \ldots, p_{nk}$, the $r_{n1}, \ldots, r_{nk}$, and the q_n) satisfy Cournot's principle as a group under these assumptions? It is reasonable to conclude that they will. If they did not, then Gambler would have a bankruptcy-free strategy $\mathcal{S}$ for choosing $\alpha_{n1}, \ldots, \alpha_{nk}, \beta_{n1}, \ldots, \beta_{nk}$ at time n and γ_n at time $n+1/2$ that would make him infinitely rich. Because Reality's moves do not depend on what Gambler does (Assumption 4) and House will set q_n equal to r_{ni_n}/p_{ni_n} (Assumption 2), Gambler has a strategy $\mathcal{S}'$ for choosing $\alpha_{n1}, \ldots, \alpha_{nk}, \beta_{n1}, \ldots, \beta_{nk}$ alone that makes his capital grow exactly as $\mathcal{S}$ does: to duplicate the effect of $\mathcal{S}$'s move γ_n, he adds $-\gamma_n r_{nj}/p_{nj}$ to α_{nj} and γ_n to β_{nj}, for $j = 1, \ldots, k$ as suggested by (7). This strategy does not require knowledge of i_n, and so Gambler would have the information needed to implement it (Assumption 3). So $\mathcal{S}'$ would also make Gambler infinitely rich, contradicting Assumption 1.

This result is a new justification of the rule of conditional probability. It tells us that if $\mathbb{P}_0[A_t^j]$ and $\mathbb{P}_0[A_t^j \cap E]$ are valid probabilities for Gambler, in the sense that he has insufficient information to beat them, then the conditional probability $\mathbb{P}_0[A_t^i \cap E]/\,\mathbb{P}_0[A_t^i]$ will be equally valid as his probability for E in a later situation where his only additional information is the observation of A_t^i. We must add, of course, that this is a long-run justification. It does not really apply to the single case but instead assumes that there will be a sequence of similar updating problems. It is a justification for using the rule of conditional probability as a policy in such problems.

Insofar as we have shown that the payoff (10) defines appropriate beliefs for Gambler after Reality announces $i = j_0$, provided this announcement is Gambler's only new information, we can also say that we have justified a version of Walley's updating principle. We should not, however, overlook the differences in formulation. Our argument is concerned with the long run, and it is concerned with odds Gambler knows he will not be able to beat, not with odds House should intend to offer.

1.7.1 The Concept of Exact Information

The crucial assumption in our argument for Walley's updating principle is Assumption 3: Reality's move at time n, i_n, is the only new thing Gambler learns before he makes his next move. Adapting the term *exact event*, introduced by Shafer [21], we may say that i_n is Gambler's *exact information.*

From a thoroughly subjective point of view, it is anodyne to say that probabilities should be updated using exact information. Certainly a person should update using *all* his information, and this is the same as saying that he should use *exactly* what he has learned. It is truly daunting, however, to plan ahead on the assumption that we will update our probabilities by the rule of conditional probability using all our information. In order to condition on all our information, we must have initial probabilities not merely for future possibilities that interest us but for all the possibilities for exactly what we will learn.

1.7.2 Generalizing to Event Trees

Our derivation of Walley's updating principle from Cournot's principle readily generalizes from the relatively rigid sequential two-stage protocol that we have used here to an event tree, in which Reality's choices on the next step depend on what she has done so far.

One generalization is to an event tree in which every second step represents exact information for Gambler. (Assumption 3 for our derivation of Walley's principle from Cournot's principle was that every second step in the two-stage sequential protocol represented exact information for Gambler.) It is also natural, however, to consider event trees in which all steps represent exact information for Gambler. In such trees, Gambler knows at the outset the exact possibilities for the future step-by-step development of his knowledge. If House states at the outset initial probabilities for all these possibilities (i.e., he states probabilities not only for the first step but also for the histories—the complete paths through the tree), and Gambler adopts Cournot's principle for these initial probabilities, then our argument yields the conclusion that Cournot's principle will also hold for conditional probabilities as future probabilities. If we place these conditional probabilities on the steps of the tree, then we again obtain a probability tree, as in Figure 1.

The lesson drawn here for event trees should be contrasted with the lesson drawn in §1.5 for event trees. The argument in §1.5 was based on the avoidance of sure loss by House, while the argument here is based on the adoption of Cournot's principle by Gambler. There is also a difference in the conclusion. In §1.5, we concluded something about consistency of advance probabilities. If House gives at the outset both probabilities for steps and probabilities for histories, then the principle of sure loss demands that they be related by the rule of conditional probability. Here, in contrast, we have assumed only that probabilities of histories are given at the outset; Cournot's principle then justifies the use of the rule of conditioning to obtain probabilities for steps.

1.8 Summary

We have arrived at a new understanding of subjective probability. According to the de Finetti's understanding, a person's subjective probabilities are rates at which he is willing to bet. According to our new understanding, a person's subjective probabilities are two-sided betting rates at which he believes he will not win heavily, no matter what strategy for betting he follows. Our new understanding provides a clear justification for using conditional probabilities as new probabilities when new information is exact—i.e., when one knows in advance a set $A_t^1, \ldots, A_t^k$ of possibilities for exactly what all one's new information will be.

The clarity of our new understanding encourages some questions. Does a person always know two-sided betting rates at which he is confident he cannot win heavily? And even if he does know such rates for some events, why should these events include a list of possibilities for exactly what he will learn between time 0 and time t? These questions inspire the generalizations considered in the remainder of this article.

In its most extreme form, the theory of subjective probability assumes that a decision maker begins with knowledge of exact possibilities for the future development of his knowledge, together with probabilities for each of these possibilities. We will now explore how this extreme picture can be relaxed in order to obtain a framework that is more useful for planning. In the next section (§2) we develop protocols that relax the demand for two-sided betting rates, as well as the demand for exact information. In a later section (§3), we extend this more flexible approach to event trees.

2 Subjective Lower and Upper Prevision

In recent decades, there has been great interest in supplementing subjective probability with more flexible representations of uncertainty. Some of the representations studied emphasize evidence rather than gambling [19, 28, 31]; others use a concept of partial possibility [9]. But many scholars prefer to generalize the story about betting that underlies subjective probability. The first step of such a generalization is obvious. Instead of requiring a person to set odds at which he will take either side of a bet, allow him to set separate odds for the two sides. This leads to lower and upper probabilities and lower and upper previsions rather than additive probabilities and expected values. See the early work of C. A. B. Smith [29, 30] and Peter Williams [37, 38, 39], the influential work of Peter Walley [32, 33, 34, 35], and the recent work of the imprecise probabilities project [5].

In this section, we look at lower and upper previsions from the point of view developed in the preceding section. This leads to a better understanding of how these measures of subjective uncertainty should change with new information, both when the new information is exact and when it is not.

We begin, in §2.1, by generalizing §1.2's protocol for subjective probability to lower and upper previsions. In §2.2, we formulate and study an abstract protocol from which lower and upper previsions can be derived. Then we turn to the problem of updating lower and upper previsions after the passage of time and the acquisition of new information. In §2.3, we state Walley's updating principle in the context of a precise protocol. Then we derive our own principles for updating under different assumptions. In §2.4 we consider the case where future beliefs are stated in advance but the information that is anticipated need not be exact; this is the assumption we consider most appropriate in planning. In §2.5 we consider the case where new information is anticipated exactly; in this case we obtain Walley's principle. We summarize our results in §2.6.

2.1 Pricing Events and Payoffs

Lower and upper probabilities arise when we relax the requirement that House announce odds for an event and offer to bet on either side. We instead allow him to offer one set of odds for betting on the event and another for betting against it.

Whereas probabilities for events determine expected values for payoffs that depend on those events (see §1.2), lower and upper probabilities are not so informative. The rates at which a person is willing to bet for or against given events do not necessarily determine the prices at which he is willing to buy or sell payoffs depending on those events. So we need more than a theory of lower and upper probabilities for events; we also need a theory of lower and upper previsions for payoffs.

2.1.1 Lower and Upper Probabilities

Suppose House expresses his uncertainty about E by specifying two numbers, p_1 and p_2. He offers to pay Gambler

$$-\alpha_1(x-p_1) = \begin{cases} -\alpha_1(1-p_1) & \text{if } E \text{ happens} \\ \alpha_1 p_1 & \text{if } E \text{ fails} \end{cases} \tag{11}$$

for any $\alpha_1 \geq 0$, and he also offers to pay Gambler

$$\alpha_2(x-p_2) = \begin{cases} \alpha_2(1-p_2) & \text{if } E \text{ happens} \\ -\alpha_2 p_2 & \text{if } E \text{ fails} \end{cases} \tag{12}$$

for any $\alpha_2 \geq 0$. In (11), Gambler sells α_1 units of x for p_1 per unit, while in (12), he buys α_2 units of x for p_2 per unit.

Here is the protocol:

FORECASTING WITH LOWER AND UPPER PROBABILITIES
House announces $p_1, p_2 \in [0,1]$.
Gambler announces $\alpha_1, \alpha_2 \in [0,\infty)$.
Reality announces $x \in \{0,1\}$.
$\mathcal{K}_1 := \mathcal{K}_0 - \alpha_1(x-p_1) + \alpha_2(x-p_2)$.

To avoid sure loss, House must make $p_1 \leq p_2$. If $p_1 > p_2$, then Gambler can make money for sure by making α_1 and α_2 strictly positive and equal (buying at p_2 what he sells at p_1).

House would presumably be willing to increase his own payoffs by decreasing p_1 in (11) and increasing p_2 in (12). The natural question is how high House will make p_1 and how low he will make p_2. We may call p_1 and p_2 *House's lower and upper probabilities*, respectively, if House will not offer (11) for any value higher than p_1 and will not offer (12) for any value lower than p_2.

When we model our beliefs by putting ourselves in the role of House, we have some flexibility in the meaning we give our refusal to offer higher values of p_1 or lower values of p_2. Perhaps we are certain that we do not want to make additional offers, perhaps we are hesitating, or perhaps we are providing merely an incomplete model of our beliefs (Walley [33], pp. 61–63).

When we instead model our beliefs by putting ourselves in the role of Gambler, the question is what values of p_1 and p_2 we believe will satisfy Cournot's principle. In the context of a sequence of forecasts, we might call p_1 and p_2 *Gambler's lower and upper probabilities* when (1) Gambler believes that no strategy for buying and selling will make him very rich in the long run when he can sell x for p_1 or buy it for p_2 but (2) Gambler is not confident about this in the case where he is allowed to sell x for more than p_1 or buy it for less than p_2. Here again clause (2) can be made precise in more than one way. Gambler might be unsure about whether he can get very rich with more advantageous values of p_1 or p_2, or he might believe that a strategy available to him would succeed with such values.

2.1.2 Lower and Upper Previsions

In order to price a payoff x that depends on the outcome of more than one event, we can generalize directly the protocol for forecasting with lower and upper probabilities:

Forecasting with Lower and Upper Previsions
House announces $p_1, p_2 \in \mathbb{R}$.
Gambler announces $\alpha_1, \alpha_2 \in [0, \infty)$.
Reality announces $x \in \mathbb{R}$.
$\mathcal{K}_1 := \mathcal{K}_0 - \alpha_1(x - p_1) + \alpha_2(x - p_2)$.

Again, Gambler is allowed to sell x for p_1 and buy it for p_2. If p_1 is the highest price at which Gambler can sell x (either the highest price House will offer or the highest price at which Gambler believes Cournot's principle, depending on which viewpoint we adopt), we may call it the *lower prevision* of x. Similarly, if p_2 is the lowest price at which Gambler can buy x, we may call it the *upper prevision* of x. (The terms *lower* and *upper prevision* appear to be due to Peter Williams [39].)

Although its meaning is clear, this protocol is not ideal for a discussion of House's or Gambler's uncertainty about x. House may have more to say about x than the lower and upper previsions p_1 and p_2, and even the statement that these are lower and upper previsions is not exactly a statement about the protocol itself. We now turn to a more abstract approach, better suited to general discussion.

2.2 Forecasting in General

Consider a set $\mathbf{R}$, and consider a set $\mathbf{H}$ of real-valued functions on $\mathbf{R}$. We call $\mathbf{H}$ a *belief cone* on $\mathbf{R}$ if it satisfies these two conditions:

1. If $\mathbf{g}$ is a real-valued function on $\mathbf{R}$ and $\mathbf{g}(\mathbf{r}) \leq 0$ for all $\mathbf{r} \in \mathbf{R}$, then $\mathbf{g}$ is in $\mathbf{H}$.

2. If $\mathbf{g}_1$ and $\mathbf{g}_2$ are in $\mathbf{H}$ and a_1 and a_2 are nonnegative numbers, then $a_1\mathbf{g}_1 + a_2\mathbf{g}_2$ is in $\mathbf{H}$.

We write $\mathcal{C}_{\mathbf{R}}$ for the set of all belief cones on $\mathbf{R}$.

Intuitively, a belief cone is a set of payoffs that House might offer Gambler. Condition 1 says that House will offer any contract that does not require him to risk a loss. Condition 2 says House will allow Gambler to combine any two of his offers, in any amounts.

The following abstract protocol is adapted from p. 90 of [26].

FORECASTING
Parameters: $\mathbf{R}$ and $\mathcal{C} \subseteq \mathcal{C}_{\mathbf{R}}$
Protocol:
House announces $\mathbf{H} \in \mathcal{C}$.
Gambler announces $\mathbf{g} \in \mathbf{H}$.
Reality announces $\mathbf{r} \in \mathbf{R}$.
$\mathcal{K}_1 := \mathcal{K}_0 + \mathbf{g}(\mathbf{r})$.

We call any protocol obtained by a specific choice of $\mathbf{R}$ and $\mathcal{C}$ a *forecasting protocol*. We call $\mathbf{R}$ the *sample space*.

We call a real-valued function on the sample space $\mathbf{R}$ a *variable*. House's move $\mathbf{H}$, itself a set of variables, determines lower and upper previsions for all variables. The *lower prevision* for a variable x is

$$\underline{\mathbb{E}}_{\mathbf{H}}\, x := \sup\{\alpha \mid (\alpha - x) \in \mathbf{H}\}, \tag{13}$$

and the *upper prevision* is

$$\overline{\mathbb{E}}_{\mathbf{H}}\, x := \inf\{\alpha \mid (x - \alpha) \in \mathbf{H}\}. \tag{14}$$

These definitions are similar to those given by Walley ([33], pp. 64–65). There is a difference in sign, however, because Walley considers a collection $\mathcal{D}$ of payoffs that House is willing to accept for himself rather than a collection $\mathbf{H}$ that House offers to Gambler.

The condition $(\alpha - x) \in \mathbf{H}$ in (13) means that Gambler can sell x for α. So roughly speaking, the lower prevision $\underline{\mathbb{E}}_{\mathbf{H}}\, x$ is the highest price at which Gambler can sell x. We say "roughly speaking" because (13) tells us only that Gambler can obtain $\alpha - x$ for α arbitrarily close to $\underline{\mathbb{E}}_{\mathbf{H}}\, x$, not that he can obtain $(\underline{\mathbb{E}}_{\mathbf{H}}\, x) - x$. Similarly, the upper prevision $\overline{\mathbb{E}}_{\mathbf{H}}\, x$ is roughly the lowest price at which Gambler can buy x.

Once we know lower previsions for all variables, we also know upper previsions for all variables, and vice versa, because

$$\overline{\mathbb{E}}_{\mathbf{H}}\, x = -\,\underline{\mathbb{E}}_{\mathbf{H}}(-x)$$

for every variable x. For additional general properties of lower and upper previsions, see Chapter 2 of Walley [33] and Chapters 1 and 8 of [26].

2.2.1 Coherence

We call a forecasting protocol *coherent* if $\mathcal{C}$ contains at least one belief cone $\mathbf{H}$ such that for every $\mathbf{g} \in \mathbf{H}$ there exists $\mathbf{r} \in \mathbf{R}$ for which $\mathbf{g}(\mathbf{r}) \leq 0$. This means that House can avoid sure loss.

Our use of "coherent" follows Shafer and Vovk [26]. In cases where House's moves are fixed, so that the game reduces to a game between Gambler and Reality (as in §3.2 of [26] or §1.3.2 of this article), our formulation simplifies; in these cases, the protocol's being "coherent" means simply that Gambler cannot make money for sure. Appendix C discusses related uses of "coherent" and "incoherent".

If $\mathbf{H}$ satisfies the condition that for every $\mathbf{g} \in \mathbf{H}$ there exists $\mathbf{r} \in \mathbf{R}$ with $\mathbf{g}(\mathbf{r}) \leq 0$, then

$$\underline{\mathbb{E}}_{\mathbf{H}}\, x \leq \overline{\mathbb{E}}_{\mathbf{H}}\, x$$

for every variable x. See §8.3 of [26].

2.2.2 Regular Protocols

Given $\mathbf{H} \in \mathcal{C}_{\mathbf{R}}$, set

$$\mathbf{H}^* := \{x : \mathbf{R} \mapsto \mathbb{R} \mid \overline{\mathbb{E}}_{\mathbf{H}}\, x \leq 0\}.$$

The following facts can be verified straightforwardly:

- $\mathbf{H}^*$ is also a belief cone ($\mathbf{H}^* \in \mathcal{C}_{\mathbf{R}}$),
- $\mathbf{H} \subseteq \mathbf{H}^*$,
- $\overline{\mathbb{E}}_{\mathbf{H}}\, x = \overline{\mathbb{E}}_{\mathbf{H}^*}\, x$ and $\underline{\mathbb{E}}_{\mathbf{H}}\, x = \underline{\mathbb{E}}_{\mathbf{H}^*}\, x$ for every variable x, and
- $(\mathbf{H}^*)^* = \mathbf{H}^*$.

Intuitively, if House offers Gambler all the payoffs in $\mathbf{H}$, then he might as well also offer the other payoffs in $\mathbf{H}^*$, because for every payoff in $\mathbf{H}^*$, there is one in $\mathbf{H}$ that is arbitrarily close to being at least as good.

We call a forecasting protocol *regular* if $\mathbf{H} = \mathbf{H}^*$ for every $\mathbf{H}$ in $\mathcal{C}$. Because any forecasting protocol can be replaced with a regular one with the same lower and upper previsions (enlarge each $\mathbf{H}$ in $\mathcal{C}$ to $\mathbf{H}^*$), little generality is lost when we assume regularity. This assumption allows us to remove the "roughly speaking" from the statements that the lower prevision of x is the highest price at which Gambler can sell x and the upper prevision the lowest price at which he can buy it. It also allows us to say that $\mathbf{H}$ is completely determined by its upper previsions (and hence also by its lower previsions):

$$x \in \mathbf{H} \text{ if and only if } \overline{\mathbb{E}}_{\mathbf{H}}\, x \leq 0.$$

The condition $x \in \mathbf{H}$ says that House will give x to Gambler. The condition $\overline{\mathbb{E}}_{\mathbf{H}}\, x \leq 0$ says that House will sell x to Gambler for 0 or less.

2.2.3 Bayesian Protocols

The protocols for subjective probability we studied in §1.2 have the property that if $\mathbf{H}$ avoids sure loss by House, then $\underline{\mathbb{E}}_{\mathbf{H}}\, x = \overline{\mathbb{E}}_{\mathbf{H}}\, x$ for every variable x. The common value of $\underline{\mathbb{E}}_{\mathbf{H}}$ and $\overline{\mathbb{E}}\, x_{\mathbf{H}}$ is, of course, x's expected value. We call a forecasting protocol with this property *Bayesian*. A regular coherent forecasting protocol is Bayesian if and only if for each $\mathbf{H} \in \mathcal{C}$, either x or $-x$ is in $\mathbf{H}$.

When the sample space $\mathbf{R}$ is infinite and endowed with a σ-algebra, some readers may prefer to require $\underline{\mathbb{E}}_{\mathbf{H}}\, x = \overline{\mathbb{E}}_{\mathbf{H}}\, x$ only for measurable x. We will not explore this issue, because we draw our motivation from the particular protocols in §1.2, not from the general concept of a Bayesian protocol.

2.2.4 Interpretation

Both interpretations of lower and upper previsions we discussed in §2.1 generalize to forecasting protocols in general. We can put ourselves in the role of House and say that our beliefs are expressed by the prices we are willing to pay—our lower and upper previsions. Or, as we prefer, we can put ourselves in the role of Gambler and subscribe to these prices in the sense of believing that they will not allow us to become very rich in the long run, no matter what strategy we follow.

The reference to the long run in the second interpretation must be understood in terms of a sequential version of our abstract protocol. If we suppose, for simplicity, that Reality and House have the same choices on every move, this sequential protocol can be written as follows:

SEQUENTIAL FORECASTING
Parameters: $\mathbf{R}$ and $\mathcal{C} \subseteq \mathcal{C}_{\mathbf{R}}$
Protocol:
  $\mathcal{K}_0 := 1$.
  For $n = 1, 2, \ldots$:
    House announces $\mathbf{H}_n \in \mathcal{C}$.
    Gambler announces $\mathbf{g}_n \in \mathbf{H}_n$.
    Reality announces $\mathbf{r}_n \in \mathbf{R}$.
    $\mathcal{K}_n := \mathcal{K}_{n-1} + \mathbf{g}_n(\mathbf{r}_n)$.

The ambiguities we discussed in §2.1 also arise here. If we take House's point of view, we may or may not be categorical about our unwillingness to offer riskier payoffs than those in $\mathbf{H}_n$. If we take Gambler's point of view, we may be more or less certain about whether larger $\mathbf{H}_n$ would also satisfy Cournot's principle.

2.3 Walley's Updating Principle

We turn now to Peter Walley's updating principle. As we saw in §1.6, this principle entails the rule of conditional probability when it is applied to subjective probability. Now we apply it to our abstract framework for lower and upper previsions.

We begin by generalizing the two-stage protocol we considered in §1.6.

TWO-STAGE FORECASTING
Parameters: $\mathbf{R}$, a disjoint partition $\mathbf{B}_1, \ldots, \mathbf{B}_k$ of $\mathbf{R}$, $\mathcal{C} \subseteq \mathcal{C}_{\mathbf{R}}$
Protocol:
At time 0:
House announces $\mathbf{H}_0 \in \mathcal{C}$.
Gambler announces $\mathbf{g}_0 \in \mathbf{H}_0$.
Reality announces $i \in \{1, 2, \ldots, k\}$.
At time t:
House announces $\mathbf{H}_t \in \mathcal{C}_{\mathbf{B}_i}$.
Gambler announces $\mathbf{g}_t \in \mathbf{H}_t$.
Reality announces $\mathbf{r} \in \mathbf{B}_i$.
$\mathcal{K}_t := \mathcal{K}_0 + \mathbf{g}_0(\mathbf{r}) + \mathbf{g}_t(\mathbf{r})$.

Because we are considering how House should make his second move, we leave this move unconstrained by the protocol. House can choose any belief cone on the reduced sample space $\mathbf{B}_i$.

Walley's updating principle says that if House knows at time 0 that Reality's announcement of i will be House's only new information when he moves at time t, then at time 0, as he makes his move $\mathbf{H}_0$, House should intend for his move $\mathbf{H}_t$ to be the belief cone $\mathbf{w}_t^i$ on $\mathbf{B}_i$ given by

$$\mathbf{w}_t^i := \{\mathbf{g} : \mathbf{B}_i \mapsto \mathbb{R} \mid \mathbf{g}^{\uparrow} \in \mathbf{H}_0\}, \tag{15}$$

where $\mathbf{g}^{\uparrow}$ is defined by

$$\mathbf{g}^{\uparrow}(\mathbf{r}) := \begin{cases} \mathbf{g}(\mathbf{r}) & \text{if } \mathbf{r} \in \mathbf{B}_i \\ 0 & \text{if } \mathbf{r} \notin \mathbf{B}_i. \end{cases} \tag{16}$$

In words: House should intend to offer a given payoff at the second stage after Reality announces i if and only if he is already offering that payoff at the first stage contingent on that value of i. This produces simple formulae relating the new lower and upper previsions to the old ones:

$$\underline{\mathbb{E}}_{\mathbf{w}_t^i}\, x = \sup\{\alpha \mid \underline{\mathbb{E}}_{\mathbf{H}_0}(x - \alpha)^{\uparrow} \geq 0\} \tag{17}$$

and

$$\overline{\mathbb{E}}_{\mathbf{w}_t^i}\, x = \inf\{\alpha \mid \overline{\mathbb{E}}_{\mathbf{H}_0}(x - \alpha)^{\uparrow} \leq 0\} \tag{18}$$

for every variable x on the reduced sample space $\mathbf{B}_i$.

A comparison of (16) with (10) confirms that the statement of Walley's updating principle given here agrees with the statement we gave in §1.6. See Appendix B for Walley's own statement of his updating principle.

2.4 Announcing Future Beliefs in Advance

As we learned in §1.5, the rule of conditional probability is mandated by the principle of House's avoiding sure loss when he announces future subjective probabilities in advance. What can we say when House announces in advance future beliefs that determine only lower and upper previsions?

Advance Forecasting
Parameters: $\mathbf{R}$, a disjoint partition $\mathbf{B}_1, \ldots, \mathbf{B}_k$ of $\mathbf{R}$, $\mathcal{C} \subseteq \mathcal{C}_{\mathbf{R}}$.
Protocol:
At time 0:
House announces $\mathbf{H}_0 \in \mathcal{C}$ and $\mathbf{H}_t^j \in \mathcal{C}_{\mathbf{B}_j}$ for $j = 1, \ldots, k$.
Gambler announces $\mathbf{g}_0 \in \mathbf{H}_0$.
Reality announces $i \in \{1, 2, \ldots, k\}$.
At time t:
Gambler announces $\mathbf{g}_t \in \mathbf{H}_t^i$.
Reality announces $\mathbf{r} \in \mathbf{B}_i$.
$\mathcal{K}_t := \mathcal{K}_0 + \mathbf{g}_0(\mathbf{r}) + \mathbf{g}_t(\mathbf{r})$.

The strategy for enforcing the rule of conditional probability that we studied in §1.5 exploited the two-sided nature of the betting offers in the probability protocol; Gambler could switch the signs of his payoffs as he pleased. Because the protocol we are now studying is not necessarily Bayesian ($x \in \mathbf{H}_0$ does not necessarily imply $-x \in \mathbf{H}_0$), this flexibility is not available. We can, however, make an argument from Gambler's point of view, relying on Cournot's principle.

Consider House's $\mathbf{H}_0$ and his $\mathbf{H}_t^j$ for some particular j. Suppose the variable $\mathbf{g}$ is in $\mathbf{H}_t^j$, but $\mathbf{g}^\uparrow$ is not in $\mathbf{H}_0$. Then it would make no difference in what Gambler can do if House were to enlarge $\mathbf{H}_0$ by adding $\mathbf{g}^\uparrow$ to it. He can already get the effect of $\mathbf{g}^\uparrow$ at time 0 by planning in advance to announce $\mathbf{g}$ at time t.

So we can assume, without changing what Gambler can accomplish, that if $\mathbf{g} \in \mathbf{H}_t^j$, then $\mathbf{g}^{\uparrow} \in \mathbf{H}_0$. This assumption implies $\mathbf{w}_t^j \subseteq \mathbf{H}_t^j$ by (15) and then

$$\underline{\mathbb{E}}_{\mathbf{H}_t^j} \leq \underline{\mathbb{E}}_{\mathbf{w}_t^j} \tag{19}$$

by (13). *The lower prevision at time t that is foreseen and announced at time* 0 *should not exceed the lower prevision given by Walley's updating principle.* Writing simply $\underline{\mathbb{E}}_0\, x$ for $\underline{\mathbb{E}}_{\mathbf{H}_0}\, x$ and $\underline{\mathbb{E}}_t\, x$ for $\underline{\mathbb{E}}_{\mathbf{H}_t^i}\, x$ (the lower previsions that House's time-0 announcements imply for time 0 and t, respectively) and recalling (17), we can write (19) in the form

$$\underline{\mathbb{E}}_t\, x \leq \sup\{\alpha \mid \underline{\mathbb{E}}_0 (x - \alpha)^{\uparrow} \geq 0\}, \tag{20}$$

where x is a variable on the reduced sample space $\mathbf{B}_i$.

The argument for (20) relies on the new viewpoint developed in this article, according to which a person's uncertainty is measured by prices he believes he cannot beat, not by prices he is disposed to offer. We expect (20) to hold because if it did not, the time 0 lower previsions would need to be increased to reflect stronger betting offers that Gambler cannot beat. Strictly speaking, of course, talk about Gambler not being able to beat given prices is talk about the long run, and so a complete exposition of the argument would involve a sequential protocol, in which advance forecasting is repeated a large or infinite number of times. We leave this further elaboration of the argument to the reader.

The argument does *not* rely on any assumption about exact information. Possibly House and Gambler will learn more than $\mathbf{B}_i$ by time t. We should keep in mind, however, that $\underline{\mathbb{E}}_t\, x$, in (20), is not necessarily the lower prevision at time t. It is merely the lower prevision at time t to which House commits himself at time 0. This commitment does not exclude the possibility that House and Gambler will acquire additional unanticipated information and that House will consequently offer Gambler more variables at time t than those to which he committed himself at time 0. In this case, the actual lower prevision for x at time t may come out higher than $\underline{\mathbb{E}}_{\mathbf{H}_t^i}\, x$ and even higher than $\underline{\mathbb{E}}_{\mathbf{w}_t^i}\, x$.

For planning at time 0, we are interested in what we can count on already at time 0. This is why the upper bound in (20) is interesting. When time t comes around, positive unanticipated information may lead us to give x a lower prevision exceeding this upper bound, but there is also the possibility of negative unanticipated information, and the upper bound can be thought of as telling us how conservative we need to be in our advance commitments in order to hedge against the possible negative information.

The complexity and subtlety of this analysis contrasts with the simplicity of our analysis of the advance protocol for subjective probabilities in §1.5. We may explain the contrast by pointing to the strength of the assumption that House can set two-sided betting rates in advance. The argument of §1.5 does not assume explicitly that new information is anticipated exactly, but from our new point of view, the assumption that House can set two-sided betting rates in advance based on i alone is not sensible unless we do indeed know in advance that there will be no other new information, or at least no other relevant new information. Otherwise, some kind of hedge against the unanticipated is in order, and this leads away from two-sided advance offers and subjective probabilities, to unequal preannounced lower and upper previsions.

2.5 Updating with Exact Information

Although the case we have just analyzed, where commitments are made in advance in the face of possible unanticipated new information, seems to us to have greater practical importance, it is also of interest to consider the case where new information is anticipated exactly. This is the case where Walley's principle applies, and as we now show, the derivation of Walley's principle from Cournot's principle that we presented for subjective probabilities in §1.7 does generalize to lower and upper previsions.

Extending the two-stage protocol of §2.3 just as we extended the two-stage probability forecasting protocol of §1.6 in §1.7, we obtain the following sequential protocol:

Sequential Two-Stage Forecasting
$\mathcal{K}_0 := 1$.
For $n = 1, 2, \ldots$
At time n:
House announces $\mathbf{H}_{n0} \in \mathcal{C}$.
Gambler announces $\mathbf{g}_{n0} \in \mathbf{H}_{n0}$.
Reality announces $i_n \in \{1, 2, \ldots, k\}$.
At time $n + 1/2$:
House announces $\mathbf{H}_{n1} \in \mathcal{C}_{\mathbf{B}_{i_n}}$.
Gambler announces $\mathbf{g}_{n1} \in \mathbf{H}_{n1}$.

Reality announces $\mathbf{r_n} \in \mathbf{B}_{i_n}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + \mathbf{g}_{n0}(\mathbf{r_n}) + \mathbf{g}_{n1}(\mathbf{r_n})$.

We reason just as in §1.7. First, we make the following assumptions:

1. House's $\mathbf{H}_{n0}$ satisfy Cournot's principle.

2. House agrees in advance to follow Walley's updating principle: $\mathbf{H}_{n1} = \mathbf{w}_n^{i_n}$, where $\mathbf{w}_n^j := \{\mathbf{g} : \mathbf{B}_j \mapsto \mathbb{R} \mid \mathbf{g}^\uparrow \in \mathbf{H}_{n0}\}$.

3. The only new information Gambler acquires between his move at time n and his move at time $n + 1/2$ is Reality's choice of of i_n. (By the preceding assumption, he already knows House's move $\mathbf{H}_{n1}$.)

4. Reality pays no attention to how Gambler moves when she chooses her moves.

Will all of House's announcements (the $\mathbf{H}_{n0}$ and $\mathbf{H}_{n1}$) satisfy Cournot's principle as a group? It is reasonable to conclude that they will. If they did not, then Gambler would have a bankruptcy-free strategy $\mathcal{S}$ that would make him infinitely rich. This strategy would specify $\mathbf{g}_{n0} \in \mathcal{C}$ for $n = 1, 2, \ldots$ and $\mathbf{g}_{n1}^j \in \mathbf{w}_n^j$ for $n = 1, 2, \ldots$ and $j = 1, \ldots, k$. Because Reality's moves do not depend on what Gambler does (Assumption 4) and House will follow Walley's recommendation for $\mathbf{H}_{n1}$ (Assumption 2), Gambler has a strategy $\mathcal{S}'$ for choosing the $\mathbf{g}_{n0}$ alone that makes his capital grow exactly as $\mathcal{S}$ does: to duplicate the effect of $\mathcal{S}$'s move $\mathbf{g}_{n1}$, he adds $(\mathbf{g}_{n1}^j)^\uparrow$ to $\mathcal{S}$'s $\mathbf{g}_{n0}$ for $j = 1, \ldots, k$. This strategy does not require knowledge of i_n, and so Gambler would have the information needed to implement it (Assumption 3). So $\mathcal{S}'$ would also make Gambler infinitely rich, contradicting Assumption 1.

This result is a long-run justification for Walley's updating principle in its full generality.

2.6 Summary

In this section we have used Gambler's viewpoint to understand lower and upper previsions and their updating.

Here as in the case of subjective probability (see §1.8), the proper handling of updating depends on whether we can exactly anticipate new information.

- We learned in §2.5 that if we can exactly anticipate new information—i.e., if we have an exhaustive advance list $\mathbf{B}_1, \ldots, \mathbf{B}_k$ of possibilities for exactly what all our new information will be, then we can follow Walley's updating principle, deriving new lower previsions from old ones using the formula

$$\underline{\mathbb{E}}_t\, x = \sup\{\alpha \mid \underline{\mathbb{E}}_0 (x - \alpha)^{\uparrow} \geq 0\}. \tag{21}$$

- We learned in §2.4 that if we cannot exactly anticipate new information, but we do know that we will learn which of the mutually exclusive events $\mathbf{B}_1, \ldots, \mathbf{B}_k$ has happened, and we commit ourselves in advance to lower previsions that depend on which $\mathbf{B}_i$ happens, then these pre-announced lower previsions should satisfy the upper bound

$$\underline{\mathbb{E}}_t\, x \leq \sup\{\alpha \mid \underline{\mathbb{E}}_0 (x - \alpha)^{\uparrow} \geq 0\}. \tag{22}$$

The requirement of exact new information is very strong. The inequality (22) depends only on the weaker condition that we learn which of the $\mathbf{B}_1, \ldots, \mathbf{B}_k$ happens. There is no requirement that this be all we learn. On the other hand, the inequality only bounds the new lower prevision that can be guaranteed at the outset, at the planning stage. Unanticipated information may actually produce a higher lower prevision.

3 Subjective Uncertainty in Event Trees

We now sketch a theory of lower and upper previsions in event trees. As we explained in §1.3.3, the framework provided by an event tree is more general than the framework provided by the protocols we have been studying, because the moves available to Reality in an event tree may depend on her previous moves. On the other hand, we simplify at the outset by suppressing House. As in §1.3.2, we assume that Gambler is told at the outset what variables will be offered to him in each possible situation in the tree. This assumption is appropriate for planning, in which we must make assumptions at the outset not only about our current uncertainty but also about our uncertainty in future situations.

We do *not* assume that our event tree provides a protocol for exact anticipation of new information. In other words, we build on the argument of §2.4 rather than the argument of §2.5. What Gambler learns following a particular situation in the tree is not necessarily represented exactly by one of the steps to the right of that situation. We assume that Reality will take one of these steps, and that Gambler will know which step Reality takes when she takes it, but this might not be all that Reality does, and it might not be all that Gambler learns. When Gambler is in a situation in the tree, he knows it, but he may also know more. He may know more and believe more than he anticipated at the planning stage.

We begin this section by characterizing event trees mathematically (§3.1). Then we use an example to clarify our interpretation of event trees (§3.2). After explaining how uncertainty in an event tree can be described by a collection of belief cones—a belief structure, as we call it (§3.3), we look at the lower and upper previsions determined by belief structures (§3.4).

An important aspect of our interpretation of event trees is our acknowlegement of the possibility of refinement. As we explain in §3.2.3, Gambler may know both that he is in a situation S in one event tree and also that he is in a more detailed version of S in a more refined event tree. The beliefs specified for S by the belief structure on the less refined tree will not be contradicted by the more detailed belief structure on the more refined tree.

3.1 What is an Event Tree?

Formally, an *event tree* is a set $\mathcal{T}$ of objects (*situations*) partially ordered by time. We write $\leq$ for the partial order, and we assume that $\leq$ has the following properties:

1. $S \leq S$ for all $S \in \mathcal{T}$.
2. If $T \leq S$ and $S \leq R$, then $T \leq R$.
3. If $S \leq R$ and $R \leq S$, then $R = S$.
4. If $T \leq S$ and $T \leq R$, then $R \leq S$ or $S \leq R$.

Properties 1–3 are the usual rules for partial order. Property 4 makes the partial order a tree.

When $S \leq R$, we say that S *follows* R. This means that S, if it happens, happens after or at the same time as R. Writing $S \leq R$ when S follows R clashes with the convention that later times are represented by larger numbers, but it is imposed by our need to keep our notation consistent with the theory of event spaces, in which $S \leq R$ has a more general meaning, encompassing both ordering by time and ordering by specificity [25].

When $R \leq S$ or $S \leq R$, we say that S and R are *ordered*. So we can state Property 4 by saying that any two situations that follow a third situation are themselves ordered.

As Shafer [19] explains, the abstract concept of an event tree is very flexible. It allows for the possibility that Reality might sometimes have infinitely many choices for her next step. It also allows for the possibility that there is no initial situation in the tree; there might instead be infinite sequences of earlier and earlier situations. When we use an event tree for planning, however, we assume that there is an initial situation and that we are in it.

3.1.1 Sample Spaces and Variables

A *history* (perhaps we should say a *possible future history*; we are looking towards the future, not the past) is a complete path through the tree—a complete account of how the events the tree tracks might unfold through time. Formally, a history is a maximal set of ordered situations—a subset $\mathbf{r}$ of $\mathcal{T}$ such that every pair of situations in $\mathbf{r}$ is ordered and this is true of no larger subset of $\mathcal{T}$ that contains $\mathbf{r}$. When a situation S is in a history $\mathbf{r}$, we say that $\mathbf{r}$ *goes through* S.

We call the set of all histories the *sample space*, and we designate it by $\mathbf{R}$. We call the set of histories that go through S the *reduced sample space* for S, and we designate it by $\mathbf{R}_S$. We call a real-valued function on $\mathbf{R}$ a *variable*. We call a real-valued function on $\mathbf{R}_S$ a *variable on* S. This is the kind of variable that might be offered to Gambler when he is in S.

If $S \leq R$ and x is a variable on S, then we write $x^{\uparrow R}$ for the variable on R given by

$$x^{\uparrow R}(\mathbf{r}) := \begin{cases} x(\mathbf{r}) & \text{if } \mathbf{r} \text{ goes through } S \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

for all $\mathbf{r}$ that go through R. An offer of the variable $x^{\uparrow R}$ in R can be thought of as a contingent offer—an offer of x provided Reality later arrives in S.

If $S \leq R$ and x is a variable on R, then we write $x^{\downarrow S}$ for the variable on S given by

$$x^{\downarrow S}(\mathbf{r}) = x(\mathbf{r})$$

for all $\mathbf{r}$ that go through S.

3.1.2 Clades and Cuts

We call a subset $\mathcal{U}$ of an event tree $\mathcal{T}$ a *clade* if no two distinct situations in $\mathcal{U}$ are ordered. This is equivalent to saying that no history goes through more than one situation in $\mathcal{U}$. We say that a variable x is *measurable* with respect to a clade $\mathcal{U}$ if for any situation S in $\mathcal{U}$ and for any pair of histories $\mathbf{r}_1$ and $\mathbf{r}_2$ that go through S, $x(\mathbf{r}_1) = x(\mathbf{r}_2)$. In words: x is constant on the histories that go through a given situation in $\mathcal{U}$.

We call a nonempty clade $\mathcal{U}$ a *cut* of a situation R if every history that goes through R goes through exactly one situation in $\mathcal{U}$ at the same time or later.

3.2 How to Interpret an Event Tree

The event trees that we consider in this section must be understood in relation to our two players, Reality and Gambler. We now explain how we see this relationship.

We can say that an event tree represents possibilities for what Reality will do. Each step is a possible move by Reality. But as we explain in §3.2.2, it is better to think of the tree as a collection of assertions about what Reality will not do—what is impossible. This is the real empirical meaning of the tree, inasmuch as it can be refuted.

We assume that Gambler sees Reality's steps as they are taken, and thus the steps also represent possibilities for what Gambler will learn. We may make this vivid by saying that Gambler moves from situation to situation with Reality. In Figure 2, for example, Gambler and Reality move to situation S_1 when Bill removes the teakettle from the fire. Gambler knows when he is in a given situation in the tree, although we do not rule out the possibility that he also knows more. This is the epistemic meaning of the tree.

Standing back a step, we can think of an event tree as a tool for planning by a person who places himself in the role of Gambler. It expresses the person's assumptions about certain events whose happening or failing he expects to follow, but it does not purport to make exhaustive predictions about what else the person will observe, even in relation to these events. Some steps on the tree may be determined by decisions that the person himself makes, either in the planning process or later. In this respect, the person is part of Reality.

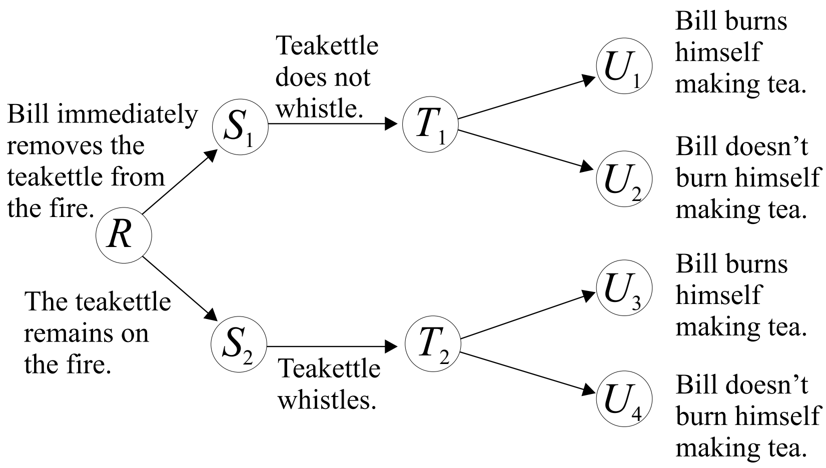


Figure 2: Another event tree. We call the nodes *situations* or *instantaneous events*.

3.2.1 Situations as Instantaneous Events

Each situation in an event tree is to be understood as a situation at a precise instant of time. In Figure 2, the timing might be tied down as follows:

- R might be defined by Gambler's actual initial state of knowledge. He is standing in Bill's kitchen watching him, and R is what Gambler knows of the situation in the kitchen as he points to Bill.
- S_1 is the situation where Bill has just removed the teakettle from the fire.
- S_2 is the situation where the teakettle has remained on the fire just long enough that its whistling is inevitable.
- T_1 and T_2 are situations where Bill has just picked up the teakettle to pour the hot water into his teapot.
- U_1 and U_3 are situations where Bill has just burned himself.
- U_2 and U_4 are situations where Bill has just finished giving his guests their tea without having burned himself.

Because the situations are instantaneous, we can also think of them as events. A situation is the same as the instantaneous event that the world (or Reality, as we have been saying) arrives in that situation.

The statement that the situations in the tree are instantaneous should not be interpreted as meaning that they have a clock time assigned to them in advance. Figure 2 does not tell us, for example, exactly what time Bill might burn himself.

3.2.2 Impossibility

According to Figure 2, it is impossible for the teakettle to whistle (T_2) unless it is left on the fire (S_2). Moreover, the teakettle will remain on the fire unless Bill removes it (either S_1 or S_2 must happen). It will not move or disintegrate of its own accord, and no man-made or natural catastrophe is about to destroy the entire kitchen. These two examples illustrate two general rules of interpretation for an event tree:

- If there is an arrow from S to T, then T can only happen if S happens first. It is impossible for T to happen without S happening first.
- If there are k arrows from R, where $k > 0$, say arrows to $S_1, \ldots, S_k$, then exactly one of these k situations must happen after R happens. It is impossible that $S_1, \ldots, S_k$ should all fail to happen or that more than one of them should happen.

The impossibilities that follow from these rules constitute the immediate empirical meaning of an event tree. We say the tree is *empirically valid* if these impossibilities are correct—i.e., if they are not refuted by subsequent events.

An event tree makes assertions about what is possible as well as assertions about what is impossible. Only the assertions of impossibility, however, can be directly tested. Who is to say, after that fact, whether something that did not happen had really been possible? It is for this reason that we treat the assertions of impossibility as the empirical meaning of the tree. We treat possibility as an epistemic matter—a fact about what one knows from being in a situation rather than a fact about the world. We hasten to add that we consider this epistemic treatment of possibility appropriate only for a subjective theory of probability. A theory of objective probability and causality based on event trees, such as the theory in [22], must treat

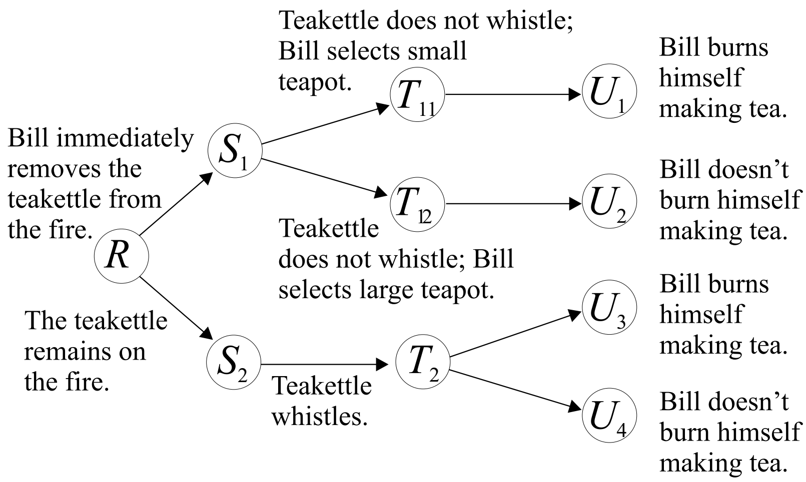


Figure 3: A refinement of Figure 2. It gives more information about what might cause Bill to burn himself if he removes the teakettle from the fire before it whistles.

---

possibility as an aspect of the world—an aspect that persists as one learns more.

3.2.3 Refinement

No event tree can show more than a fragment of what happens in the world—a fragment of what Reality does. There is always more to be said about the current state of the world and possibilities for the future. Figure 3 shows a bit of additional detail that might lie beneath Figure 2's assertions about Bill's teamaking. This figure *refines* Figure 2 by splitting T_1 into two situations, T_{11} and T_{12}, depending on whether Bill is pouring the hot water into the small or large teapot. A person who knows that he is in T_{11} and knows Figure 3 knows more at this point than a person who knows only that he is in T_1: he knows Bill is going to burn himself. But nothing he knows contradicts the impossibilities announced by Figure 2. The two event trees can be simultaneously empirically valid.

In addition to splitting situations, we can also refine by interpolating situations to represent intermediate events. In general, this involves refining what looks like an individual step into chains, or multiple chains, which branch according to the outcome of the intermediate events. Examples are given in [22], [23], and [25]. (The notion of refinement used here differs, however from the objective concept of refinement used in [22], which does not permit possibilities to be ruled out by refinement. See also the discussion in [36].)

We say that an event tree is *epistemically valid* if Gambler knows at the outset that he will know when and if he arrives in a given situation in the tree. A tree can fail to be epistemically valid even though it is empirically valid, and even though it refines an epistemically valid event tree.

We will assume that the event tree we are working with is both empirically and epistemically valid. This assumption is appropriate for planning, because we it is reasonable to make plans about what to do in future situations—what information to gather and what actions to take—only if we can assume we will know when to carry out these plans.

The need for epistemic validity underlies our insistence that a situation in our event tree might not be assigned a determinate clock time. Presumably we can always specify a range of possible times for each situation in the tree, and presumably we could refine the tree by splitting each situation into separate situations, each labeled by one of these possible times. But if we are not always watching the clock, then this refinement may fail to be epistemically valid for us. It may also be quite unnecessarily cumbersome.

3.3 Belief Structures on Event Trees

In the protocols of §2, Gambler expressed his beliefs about what would happen next by adopting Cournot's principle with respect to given belief cones. We now generalize this idea to event trees.

Because an event tree can be refined by the interpolation of intermediate events, we do not want to lean on the notion of what happens next. So we allow the variables in the belief cone adopted by Gambler for a given situation to depend on any number of future steps by Reality, not just on the next step. (This fits the definition of variable we gave in §3.1.2.) We call the collection of these belief cones a belief structure.

3.3.1 Definition

Consider a set $\mathcal{B}$ of pairs of the form (x, S), where S is a situation in our event tree and x is a variable on S. We write $\langle x \mid S\rangle$ to indicate that the pair (x, S) is in $\mathcal{B}$. We call $\mathcal{B}$ a *belief structure* if the following principles are satisfied:

1. RATIONALITY: If x is a nonpositive variable on S (this means that $x(\mathbf{r}) \leq 0$ for all $\mathbf{r}$ through S), then $\langle x \mid S \rangle$.

2. ADDITIVITY: If $\langle x_1 \mid S \rangle$ and $\langle x_2 \mid S \rangle$, then $\langle x_1 + x_2 \mid S \rangle$.

3. SCALING: If α is a nonnegative real number and $\langle x \mid S \rangle$, then $\langle \alpha x \mid S \rangle$.

4. CONTINGENCY: If $S \leq R$ and $\langle x \mid S \rangle$, then $\langle x^{\uparrow R} \mid R \rangle$.

We suppose that $\mathcal{B}$ is adopted by Gambler at the outset, before he accompanies Reality through the event tree. His adoption of $\mathcal{B}$ means that he thinks the offers in it (an offer of x if and when he is in S for every (x, S) in $\mathcal{B}$) are not enough to allow him to get very rich.

The first three principles, RATIONALITY, ADDITIVITY, and SCALING, are the principles underlying our definition of belief cone in §2.2. They require the set of x for which (x, S) is in $\mathcal{B}$ to be a belief cone for each fixed S.

The fourth principle, CONTINGENCY, is justified by the argument we learned in §2.4. Because Gambler will know if and when he arrives in S, he can plan in R to accept the offer of x that he knows he will have in S if he arrives in S. This plan can be adopted at the same time and has the same effect as accepting an offer of $x^{\uparrow R}$ in R. So if Cournot's principle is valid for the offer of x in S, then it is also valid for an offer of $x^{\uparrow R}$ in R, where Gambler knows less.

All four principles can be thought of ways of enlarging a set of variables offered to Gambler. We can start with an arbitrary set of variables and use each of the four principles as a rule for adding others. If x in a nonpositive variable on S, then we can add $\langle x \mid S \rangle$. If $\langle x_1 \mid S \rangle$ and $\langle x_2 \mid S \rangle$ are already in the set, then we can add $\langle x_1 + x_2 \mid S \rangle$, and so on. The set formed by all variables that can be obtained in this way in a finite number of steps may be called the *closure* of the initial set of variables. It will satisfy all four conditions and thus qualify as a belief structure.

Figure 4 illustrates CONTINGENCY. The event tree at the top of the figure shows that if it rains on Monday, then Gambler can bet on it raining again on Tuesday. The event tree at the bottom shows the same offer, made on a contingent basis on Sunday evening. The two bets have the same payoffs, and if Gambler cannot get rich exploiting offers on Monday evening, then

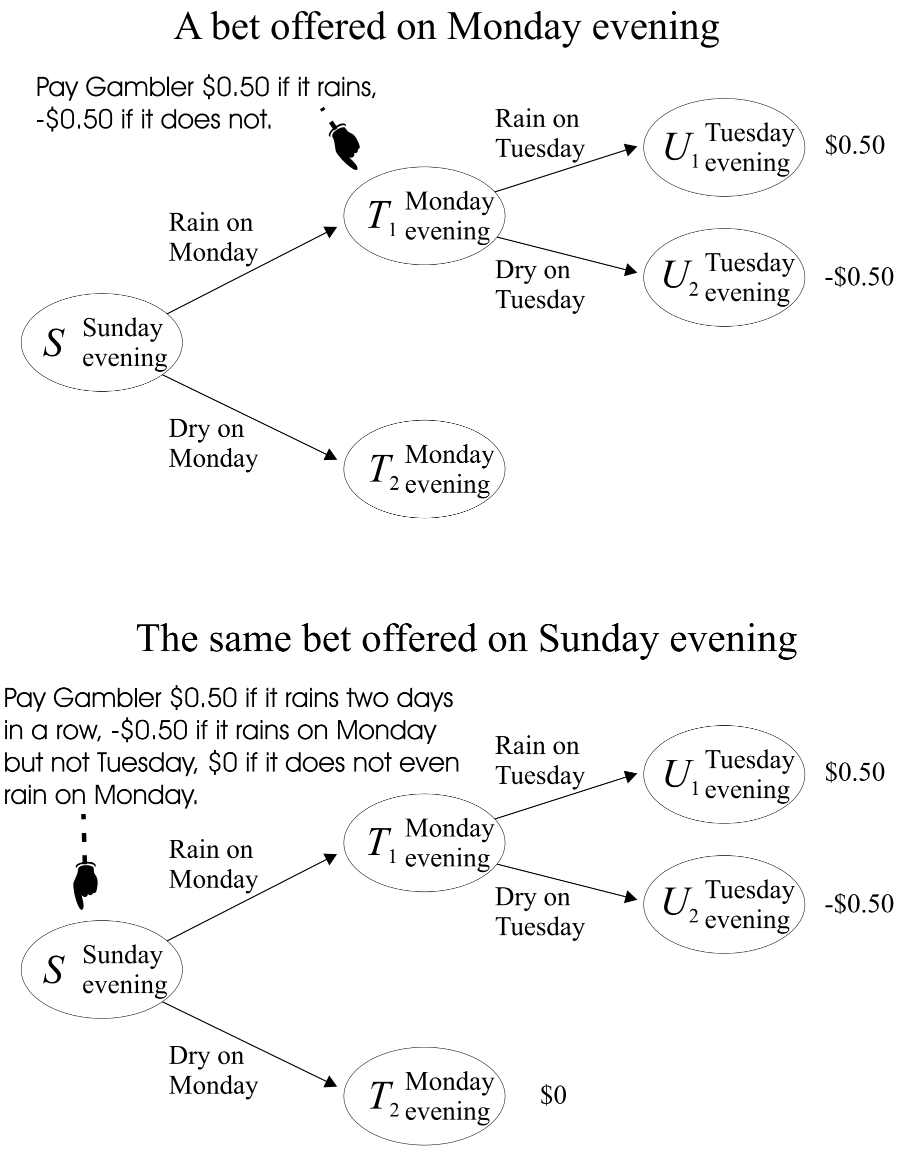


Figure 4: Offering a bet contingently.

he cannot get rich exploiting the corresponding contingent offers on Sunday evening, when he knows less about the future.

Figure 5 illustrates how a belief structure might extend to a valid refinement. At the top we have that same tree that we saw in Figure 2, with one variable offered in S_1 and another offered in T_2. At the bottom we have the refinement we saw in Figure 3, with the same two offers. For both trees, we consider the belief structure obtained by taking the closure of the two gambles offered. By CONTINGENCY, the variable offered in T_2 is also offered in S_2. But the variable offered in S_1 is not offered in T_1 in the tree at the top. The only variables offered in T_1 are those introduced by RATIONALITY. Gambler is not offered any bet on whether Bill will burn himself. Our assumption that the refinement is epistemically valid makes this quite appropriate, for when Gambler is in T_1 he is also in either T_{11} or T_{12}, and hence he knows whether Bill will burn himself or not.

In general, refinement can reveal more variables available to Gambler. But neither these additional variables nor the additional knowledge indicated by the refinement should invalidate Gambler's adoption of Cournot's principle for the belief structure on the less refined event tree.

3.3.2 Other Properties

Here some properties that a belief structure may or may not have:

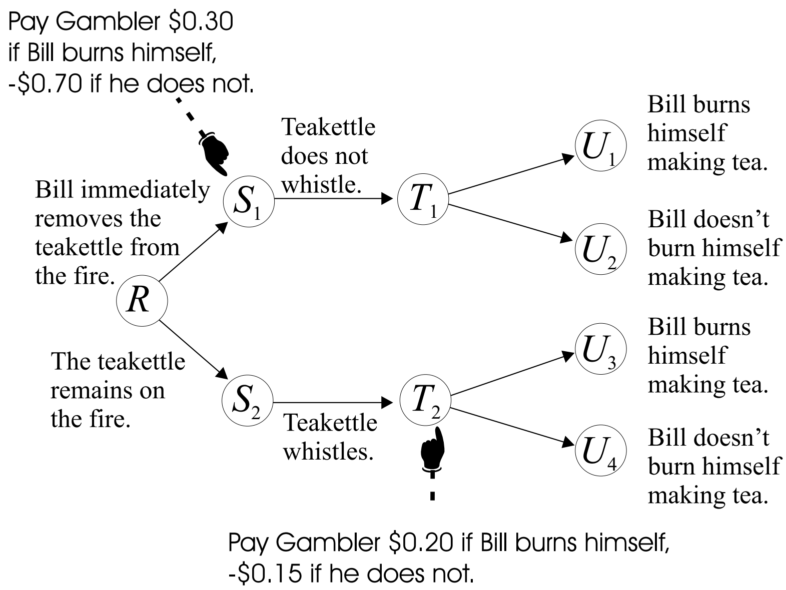
- COHERENCE: The structure is *coherent* if whenever $\langle x \mid S\rangle$ there is a path $\mathbf{r}$ going through S such that $x(\mathbf{r}) \le 0$.

- REGULARITY: The structure is *regular* if $\langle x \mid S\rangle$ whenever $\langle x - \epsilon \mid S\rangle$ for every $\epsilon > 0$.

- WALLEY UPDATE: The structure is *Walley updated* if $\langle x \mid S\rangle$ whenever x is a variable on S, $S \le R$, and $\langle x^{\uparrow R} \mid R\rangle$.

- TEMPORAL DECOMPOSABILITY: The structure is *temporally decomposable* if whenever $\langle x \mid R\rangle$ and $\mathcal{U}$ is cut of R, there exists a variable y such that (1) y is measurable with respect to $\mathcal{U}$, (2) $\langle y \mid R\rangle$, and (3) $\langle (x-y)^{\downarrow S} \mid S\rangle$ for each $S \in \mathcal{U}$.

We discussed COHERENCE for forecasting protocols in §2.2.1. We certainly want out belief structures to be coherent. If we discover, in the course of enlarging an initial set of offers using the five defining conditions, that



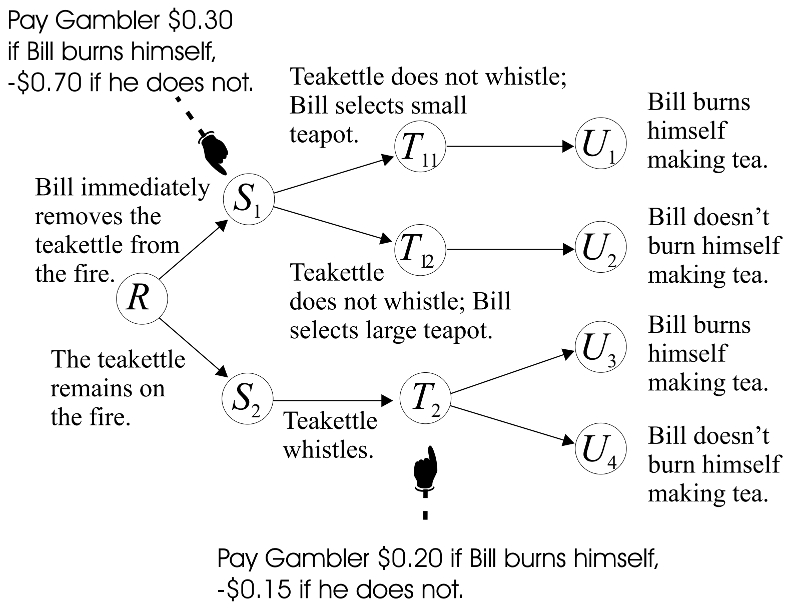


Figure 5: Belief structures for Figures 2 and 3.

Gambler is offered a sure gain in S, then we will have refuted Gambler's adoption of Cournot's principle for the structure. But this incoherence might be difficult to discover, and so it is convenient to leave COHERENCE out of the definition of belief structure.

We discussed REGULARITY for forecasting protocols in §2.2.2. As we noted there, it is a convenient property. Moreover, it has a virtue in common with the four principles we have assumed for belief structures: it can be interpreted as an instruction for enlarging a set of offers. But we are interested in implementing the ideas in this article in a practical logic, and it is not clear how such a logic could make use of an infinite number of premises, one for each $\epsilon > 0$. So we also leave REGULARITY as an auxiliary condition.

WALLEY UPDATE is the converse of our principle of CONTINGENCY. The two together constitute Walley's updating principle. As Walley himself would agree, it is appropriate to demand WALLEY UPDATE only when the event tree is interpreted as a protocol for exactly anticipated new information. So it certainly should not be included in our definition of belief structure. Walley [33] studies the implications of WALLEY UPDATE at length. We study the concept further from our own point of view in [14].

TEMPORAL DECOMPOSABILITY says that Gambler can always decompose an offer in R into two successive offers, the first of which is settled in $\mathcal{U}$. The variable x offered in R may still be unsettled in $\mathcal{U}$, but there must be another variable y offered in R that is settled in $\mathcal{U}$, such that in each situation S in $\mathcal{U}$, Gambler can buy x for the amount he gets there from y. Gambler can accept y in R and then perhaps wait until $\mathcal{U}$ to decide whether to buy x for the payoff he has obtained from y. The sequential protocols that we studied in §2 define temporally decomposable belief structures, because the variables offered there are always immediately settled. We single out the concept of temporal decomposability here only in order to point out that it is not required by our general concept of a belief structure on an event tree. Figure 6 shows a belief structure that is not temporally decomposable.

3.4 Lower and Upper Prevision

A belief structure determines lower and upper previsions in an event tree just as a belief cone does in a forecasting protocol. In the case of the belief structure in an event tree, the lower and upper previsions are relative to the situation.

The lower and upper previsions in S for a variable x on S are defined, of

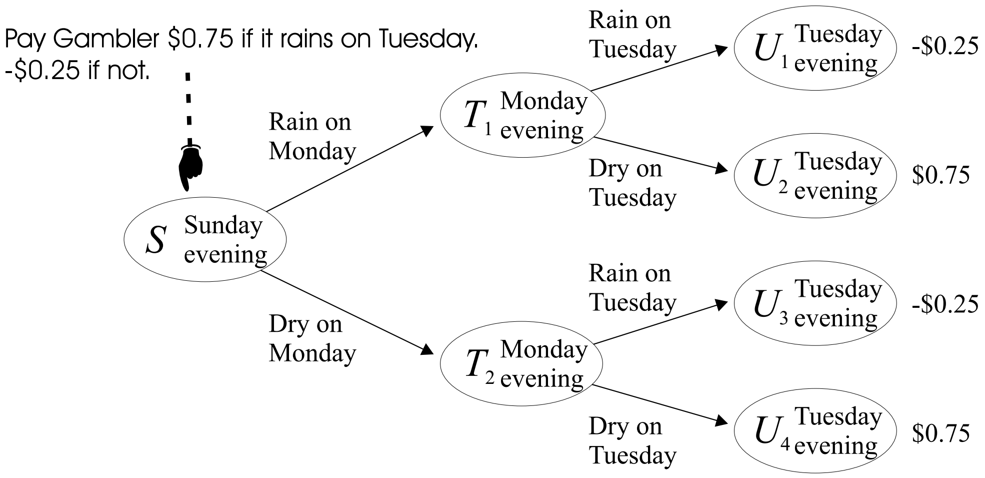


Figure 6: A weather forecaster offers to gamble on whether it will rain the day after tomorrow but not on whether it will rain tomorrow. On Sunday evening, he offers $0.25 for a return of $1 if it rains during the day on Tuesday. The offer is only for Sunday evening; he will no longer gamble on Tuesday's weather after he has seen Monday's. The belief structure he offers to Gambler is not temporally decomposable.

---

course by

$$\underline{\mathbb{E}}_S\, x := \sup\{\alpha \mid \langle \alpha - x \mid S\rangle\}$$

and

$$\overline{\mathbb{E}}_S\, x := \inf\{\alpha \mid \langle x - \alpha \mid S\rangle\}.$$

These quantities have the properties we would expect from our work in §2. For example,

$$\overline{\mathbb{E}}_S\, x = -\underline{\mathbb{E}}_S(-x)$$

whenever x is a variable on S, and

$$\underline{\mathbb{E}}_S\, x \leq \sup\{\alpha \mid \underline{\mathbb{E}}_R(x-\alpha)^{\uparrow R} \geq 0\}$$

whenever x is a variable on S and $S \leq R$.

If the belief structure is coherent, we obtain

$$\underline{\mathbb{E}}_S\, x \leq \overline{\mathbb{E}}_S\, x.$$

If it is regular, we obtain

$$\langle x \mid S\rangle \text{ if and only if } \overline{\mathbb{E}}_S\, x \leq 0.$$

If it is Walley updated, we obtain

$$\underline{\mathbb{E}}_S\, x = \sup\{\alpha \mid \underline{\mathbb{E}}_R(x-\alpha)^{\uparrow R} \geq 0\}$$

whenever x is a variable on S and $S \leq R$.

In order to see the effect of temporal decomposability on lower and upper previsions, we need another definition. Given a situation R, a variable x on R, and a cut $\mathcal{U}$ of R, write $\underline{\mathbb{E}}_{\mathcal{U}}\, x$ for the variable on R given by

$$(\underline{\mathbb{E}}_{\mathcal{U}}\, x)(\mathbf{r}) := \underline{\mathbb{E}}_{S(\mathbf{r})}\, x^{\downarrow S(\mathbf{r})},$$

where $S(\mathbf{r})$ is the situation in $\mathcal{U}$ that $\mathbf{r}$ goes through, and define $\overline{\mathbb{E}}_{\mathcal{U}}\, x$ analogously. In general, we have

$$\underline{\mathbb{E}}_R\, x \geq \underline{\mathbb{E}}_R[\underline{\mathbb{E}}_{\mathcal{U}}\, x] \text{ and } \overline{\mathbb{E}}_R\, x \leq \overline{\mathbb{E}}_R[\overline{\mathbb{E}}_{\mathcal{U}}\, x],$$

but when the belief structure is temporally decomposable, this can be strengthened to

$$\underline{\mathbb{E}}_R\, x = \underline{\mathbb{E}}_R[\underline{\mathbb{E}}_{\mathcal{U}}\, x] \text{ and } \overline{\mathbb{E}}_R\, x = \overline{\mathbb{E}}_R[\overline{\mathbb{E}}_{\mathcal{U}}\, x].$$

See [26], pp. 184–185.

For additional discussion of the properties of lower and upper previsions, see [33] and §8.3 of [26].

4 Further Perspectives

We have developed a simple mathematical framework for representing subjective uncertainty through time. This framework has many traditional elements, but it relates subjective uncertainty to gambling in a novel way. Instead of emphasizing that prices for uncertain payoffs should avoid sure loss, it emphasizes a person's conviction that he cannot get rich at these prices.

The framework is also distinguished by the flexbility of its assumptions about future information. The beliefs it specifies for a future situation where we have given knowledge are not invalidated if it turns out, when we arrive in that situation, that we know even more.

In this concluding section, we offer some additional perspectives on this new framework. We emphasize the simplicity of its principles (§4.1). We explain how it is motivated by the problem of planning (§4.2). We discuss how it might deal with the distinction between updating and revision (§4.3). And finally, we discuss how it can be further developed using the concept of an event space (§4.4).

4.1 Two Principles

Early in the twentieth century, the French mathematicians Jacques Hadamard and Paul Lévy suggested that probability theory is founded on two principles ([26], p. 44). The first principle says that if a gambler can make either of two bets, he can make them both. This is a mathematical principle, and it leads to the fundamental property of mathematical probability: probabilities add. The second principle is not a mathematical principle; rather it is a principle that connects mathematical probability with reality. This is the principle that an event with very small probability will not happen.

At its core, our framework is merely an elaboration of Hadamard and Lévy's two principles. Our definition of a belief structure formalizes the principle of additivity. Our version of Cournot's principle is a generalization of the principle that an event with very small probability will not happen.

4.2 Uncertain Reasoning

Event trees and the belief structures on them are purely mathematical objects. We have designed them, however, to serve as part of a semantics for languages for planning under uncertainty.

A language for planning under uncertainty must name events, describe possible actions and their outcomes in terms of these events, and describe the planner's uncertainty both about the behavior of the world and the outcomes his own actions. We propose that instantaneous events or situations should be taken as the fundamental concept in such a language, and that statements of uncertainty about such events should expressed by gambles be interpreted using Cournot's principle. Our four properties for event trees and our four principles for belief structures can be translated into rules for reasoning about assertions in such a planning language.

This article has not done any of the work that will be involved in developing such languages. But the modular character of our semantics should contribute to make this work practical. Our notion of refinement will facilitate the description of events, because it authorizes us describe different events at different levels of detail. Our principles for belief structures will allow us to construct belief structures from individual offers to gamble.

4.3 Updating and Revision

There is a sizeable literature, going back at least to work of Isaac Levi [17] and Peter Gärdenfors [12] in the 1980s, that considers how knowledge should be revised with experience. Many recent contributions to this literature, including an influential contribution by Katsuno and Mendelzon [15], emphasize the distinction between bringing a knowledge base up to date when the world it describes changes ("updating") and revising a knowledge base to incorporate new information about a static world ("revision").

The analysis in this article suggests that this distinction between updating and revision is not simple for subjective probability or subjective lower and upper prevision. Our steps in our protocols and trees include changes in the world (moves by Reality), but these changes are also learning steps for the person who is cast in the role of Gambler. The two aspects cannot be disentangled, and the abstract theory applies equally well to examples in which the changes are substantial changes external to the person who plays Gambler and to examples where the changes are little more than changes in Gambler's knowledge.

On the other hand, a belief structure on an event tree can only be an imperfect plan, drawn up by an imperfect planner and subject to revision. Several kinds of revision are possible and important. The most innocuous revision is refinement—the addition of more detail to the event tree or the addition of more gambles to the belief structure, raising some of our lower previsions. The most severe is empirical refutation. And as we learned in §2.4, we may increase our lower previsions in the course of events event when nothing happens to refute our belief structure.

A belief structure is best thought of, perhaps, as a plan. Plans are seldom followed for very long. Even when a plan is working well, we usually quickly obtain enough new insight to make its revision worthwhile. So no matter how long-term a belief structure on an event tree is, we may expect to change it before we have followed Reality for many steps.

4.4 Event Spaces

Although event trees are more flexible than protocols for representing uncertainty, a single event tree cannot represent all the future situations we will want to consider in a planning problem. Our reasoning in such problems will be based on rules (rules about what Reality is allowed to do next, rules that specify probabilities or gambles for what Reality will do next, rules for gathering information, perhaps even rules for taking substantive actions) that are triggered when a situation satisfies certain premises. Some rules will require more specificity about the situation than others. Yet an event tree is limited to one level of specificity; it cannot include situations at different levels of specificity such as T_1 and T_{11} in Figures 2 and 3.

In [25], we propose a formalism that treats situations in abstraction from any particular event tree in which they might be represented. In this formalism, we treat on an equal basis not only situations that are ordered in time, but also situations that are ordered by specificity. These situations form an *event space*.

The principles that we have stated for a belief structure in an event tree readily extend to an event space. We can also add a principle that formalizes our comments in connection with Figure 5 on p. 47: any variable that is included in the belief structure for a given situation should also be included for any refinement of that situation.

A Cournot's Principle

Antoine Augustin Cournot (1801-1877) was one of the first proponents of what came to be called the frequentist conception of probability. He also enunciated what has sometimes been called Cournot's principle: an event with probability zero will not happen. This principle, together with the law of large numbers, implies that probabilities will be matched in the world by frequencies.

In this article, we have followed Shafer and Vovk [26] in using "Cournot's principle" to name a more general principle, which applies to sequential gambling offers that fall short of determining probabilities: Gambler cannot take advantage of the offers to become infinitely rich.

The concept of probability zero is not necessarily applicable to our protocols, because they do not necessarily allow us to say whether a given event has probability zero or not. The protocol for sequential probability forecasting in §1.3, for example, requires only that House give a probability for each E_n after $E_1, E_2, \dots, E_{n-1}$ have been settled; it does not require House to give a joint probability measure for $E_1, E_2, \dots$ and hence does not require him to say whether a given event defined by $E_1, E_2, \dots$ has probability zero. But if House did give such a joint probability measure, then Shafer and Vovk's version of Cournot's principle could be considered a consequence of the principle Cournot advocated, because when Gambler followed a strategy that did not risk bankruptcy, his capital would be a nonnegative martingale with respect to the probability measure, and the event that a nonnegative martingale diverges to infinity has probability zero.

We must distinguish Cournot's principle from the principle of House's avoiding sure loss. Both principles limit Gambler's ability to make money. But they differ sharply in other ways:

- The principle of House's avoiding sure loss is a constraint on House. It forbids House from choosing probabilities that will permit Gambler to arrange to make money from House no matter how Reality moves. This is a hard constraint; it constrains House's probabilities in specific and precise ways.

- Formally, Cournot's principle is a constraint on Reality. It forbids Reality from moving in such a way that Gambler can make too much money in the long run. This is a soft constraint; it does not constrain Reality much on any single move.

Of course, House must avoid sure loss in order for Reality to be able to obey Cournot's principle. Moreover, we do not expect Reality to obey Cournot's principle unless House chooses his probabilities properly.

B Peter Walley on Updating

In *Statistical Reasoning with Imprecise Probabilities* [33], Peter Walley paints a picture in which a person is disposed to agree to certain gambles. Following de Finetti, Walley calls this person "You", but "You" is analogous to our "House", inasmuch as he expresses beliefs by offering gambles. So far as we have noticed, Walley does not discuss a counter party; there is no Gambler in the picture.

Walley states his updating principle on p. 287 in these words: "Any gamble Z is B-desirable if and only if BZ is desirable." The terms in this statement are defined as follows:

- A gamble (payoff) is desirable, roughly speaking, if You are disposed to accept it (p. 615).
- B is an event that You will observe to happen or fail.
- A payoff Z is B-desirable if You "intend to accept" Z provided You observe "just the event B" (p. 287).
- The payoff BZ is equal to Z if B happens and 0 otherwise.

If we identify Walley's You with our House and identify You being disposed to accept Z with House's offering $-Z$ to Gambler, then this becomes the principle that we stated in §2.3.

Walley argues for his updating principle with the following comments

- "...the time at which gambles are accepted does not affect their value." (p. 294)
- "The two dispositions mentioned in the updating principle have the same effect." (p. 288)

In the last paragraph of §6.1.6 on p. 288, Walley argues that in the Bayesian case violation of his updating principle may commit a person to accepting two gambles that together cannot produce a gain and may produce a net loss. This is not confirmed, however, by our analysis of the two-stage Bayesian protocol of §1.6. There Gambler cannot make House suffer a loss if House does not disclose in advance how he will violate the updating principle (whether his new probability for E will be greater or less than the value implied by his betting offers at time 0).

C "Incoherence"

The use of "coherent" and "incoherent" in probability theory seems to have originated with Bruno de Finetti. In a seminal article first published in French in 1937, [6], de Finetti considered a person—let us call him House—who offers odds at which he will bet for or against various events. De Finetti suggested that these odds should fit together with each other, or cohere, in such a way that House would not be vulnerable to Gambler's selecting a portfolio of bets that would produce a net loss for House no matter how the events come out. If the odds did not have this property, they would be called incoherent. De Finetti showed that coherence is sufficient to establish the classical properties of probability. In later work [7], de Finetti reformulated these ideas in terms of prices a person sets for buying or selling uncertain payoffs.

It would be unfair to call a person incoherent for refusing to play de Finetti's game—i.e., for refusing to offer all comers prices at which he will both buy and sell. But once he does agree to the game, it relatively reasonable to call his prices incoherent when he opens himself up to a sure loss.

In the looser games of §2, where House may offer distinct buying and selling prices, criticism of these prices becomes more complicated. Here are two distinct ways we might criticize them:

- **Vulnerability to Sure Loss.** If House merely states prices at which he will buy certain variables, then perhaps we can combine various of his offers in such a way that he pays more than he will get back no matter how events come out. In this case, his prices are incoherent in de Finetti's original sense.

- **Understatement.** If House not only states prices at which he will buy certain variables but also states that these are his maximum prices for these variables, then we may be able to refute one of his statements—say the statement that α is the most he will pay for x, by combining various of his other buying offers in such a way that what he is buying adds up to x and what he pays for it adds up more than α. In this case, we may say that his statements about what he is willing to pay are incorrect.

In this article, we say that prices avoid sure loss when they are coherent in de Finetti's original sense. This leaves the words "coherent" and "incoherent" available for other roles. So we apply them to protocols: a protocol is coherent if House is permitted to avoid sure loss. This is also the way the term is used by Shafer and Vovk [26].

Walley and some other authors use "incoherent" to criticize systems of prices for understatement. They call a system of buying prices incoherent whenever the buying price given for one variable is less than the buying price that can be derived from buying prices given for other variables. Moreover, they allow the derivation to include not only the combination of offers but also the application of other principles to which they subscribe, such as Walley's updating principle. This leads to the unpleasant situation where someone who considers the updating principle inapplicable to a particular problem risks having his beliefs labeled "incoherent" for this reason.

We suggest that this use of "incoherent" be discontinued. It is impolite to label systems of prices that some people find reasonable incoherent, and it is disingenuous, in those circumstances, to pretend that the word is an innocuous technical term.

Acknowledgements

Research for this article was supported by NSF grant SES-9819116. The authors have benefited from discussions with Nic Wilson and Volodya Vovk and from helpful comments by several referees on a very flawed earlier version of the article.

References

[1] José M. Bernardo and Adrian F. M. Smith. *Bayesian Theory*. Wiley, Chichester, England, 1994.

[2] Yuan S. Chow and Henry Teicher. *Probability Theory: Independence, Interchangeability, Martingales*. Springer, New York, third edition, 1997.

[3] Giulianella Coletti and Romano Scozzafava. Characterization of coherent conditional probabilities as a tool for their assessment and extension. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 4:103–127, 1996.

[4] A. P. Dawid. Statistical theory: The prequential approach. *Journal of the Royal Statistical Society, Series A*, 147:278–292, 1984.

[5] Gert de Cooman and Peter Walley. The imprecise probabilities project. `http://ippserv.rug.ac.be/home/ipp.html`.

[6] Bruno de Finetti. La prévision, ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 7:1–68, 1937. An English translation by Henry E. Kyburg, Jr., is included in both editions of [16].

[7] Bruno de Finetti. *Teoria Delle Probabilità*. Einaudi, Turin, 1970. An English translation, by Antonio Machi and Adrian Smith, was published as *Theory of Probability* by Wiley (London, England) in two volumes in 1974 and 1975.

[8] Joseph L. Doob. *Stochastic Processes*. Wiley, New York, 1953.

[9] Didier Dubois and Henri Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, 1988.

[10] Richard Durrett. *Probability: Theory and Examples*. Wadsworth, Belmont, California, second edition, 1996.

[11] William Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. Wiley, New York, second edition, 1971. The first edition of this volume appeared in 1966.

[12] Peter Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press, Cambridge, Massachusetts, 1988.

[13] Andrew Gelman, John B. Carlin, Hal S. Stern, and Donald B. Rubin. *Bayesian Data Analysis*. Chapman and Hall, London, 1995.

[14] Peter R. Gillett, Richard B. Scherl, and Glenn Shafer. A probabilistic logic based on the acceptability of gambles. Unpublished draft, 2002.

[15] Hirofumi Katsuno and Alberto O. Mendelzon. On the difference between updating a knowledge base and revising it. In Peter Gärdenfors, editor, *Belief Revision*, Cambridge Tracts in Theoretical Computer Science 29, pp. 183–203. Cambridge University Press, Cambridge, 1992.

[16] Henry E. Kyburg Jr. and Howard E. Smokler, editors. *Studies in Subjective Probability*. Wiley, New York, 1964. A selection of readings, ranging chronologically from John Venn in 1888 to Leonard J. Savage in 1961. A second edition, with a slightly different selection, was published by Krieger, New York, in 1980.

[17] Isaac Levi. *The Enterprise of Knowledge*. MIT Press, Cambridge, Massachusetts, 1980.

[18] Christian P. Robert. *The Bayesian Choice*. Springer, New York, second edition, 2001.

[19] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, New Jersey, 1976.

[20] Glenn Shafer. Bayes's two arguments for the rule of conditioning. *Annals of Statistics*, 10:1075–1089, 1982.

[21] Glenn Shafer. Conditional probability. *International Statistical Review*, 53:261–277, 1985.

[22] Glenn Shafer. *The Art of Causal Conjecture*. The MIT Press, Cambridge, Massachusetts, 1996.

[23] Glenn Shafer. Mathematical foundations for probability and causality. In Frederick Hoffman, editor, *Mathematical Aspects of Artificial Intelligence*, volume 55 of *Symposia in Applied Mathematics*, pp. 207–270. American Mathematical Society, Providence, Rhode Island, 1998.

[24] Glenn Shafer. Causality and responsibility. *Cardozo Law Review*, 22:101–123, 2001.

[25] Glenn Shafer, Peter R. Gillett, and Richard Scherl. The logic of events. *Annals of Mathematics and Artificial Intelligence*, 28:315–389, 2000.

[26] Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's Only a Game!* Wiley, New York, 2001. For more information on this book, see `http://www.cs.rhul.ac.uk/~vovk/book/`.

[27] Albert N. Shiryaev. *Probability*. Springer, New York, second edition, 1996.

[28] Philippe Smets. Varieties of ignorance. *Information Sciences*, 57–58:135–144, 1991.

[29] C. A. B. Smith. Consistency in statistical inference and decision. *Journal of the Royal Statistical Society, Series B*, 23:1–25, 1961.

[30] C. A. B. Smith. Personal probabilities and statistical analysis. *Journal of the Royal Statistical Society, Series A*, 128:469–499, 1965.

[31] Rajendra P. Srivastava and Theodore J. Mock, editors. *Belief Functions in Business Decisions.* Springer Verlag, New York, 2002.

[32] Peter Walley. Coherent lower (and upper) probabilities. Technical Report 22, Department of Statistics, University of Warwick, Coventry, UK, 1981.

[33] Peter Walley. *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall, London, 1991.

[34] Peter Walley. Inferences from multinomial data: learning about a bag of marbles (with discussion). *Journal of the Royal Statistical Society, Series B*, 58:3–57, 1996.

[35] Peter Walley. Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning*, 24:125–148, 2000.

[36] Harald Wiese. On interpreting and changing decision trees. Note de recherche 99-03, GRID, École Normale Supérieure de Cachan, France, March 1999.

[37] Peter M. Williams. Coherence, strict coherence and zero probabilities. In *Fifth International Congress of Logic, Methodology and Philosophy of Science*, pp. VI, 29–33, 1975.

[38] Peter M. Williams. Notes on conditional previsions. Research Report, School of Mathematical and Physical Sciences, University of Sussex, February 1975.

[39] Peter M. Williams. Indeterminate probabilities. In M. Przełęcki, K. Szaniawski, and R. Wojcíckí, editors, *Formal Methods in the Methodology of Empirical Sciences*, pp. 229–246. Ossolineum & Reidel, 1976.