Epistemic utility theory

Richard Pettigrew

March 29, 2010

One of the central projects of formal epistemology concerns the formulation and justification of epistemic norms. The project has three stages: First, the formal epistemologist produces a mathematical model of an agent's epistemic states—call this the *descriptive stage*. Next, she formulates, in terms of this model, putative norms that she claims govern these states—call this the *normative stage*. Finally, she provides a justification for these norms—call this the *justificatory stage*. It is one of the great virtues of formal epistemology that the final justificatory stage can be made mathematically precise. The strategy is this: the formal epistemologist states an epistemic norm that is taken to be more general and fundamental than those given in the normative stage; this norm is then formulated in terms of the mathematical model provided by the descriptive stage; and the norms posited in the normative stage are derived from this fundamental norm by means of a mathematical theorem.

1 Introducing epistemic utility theory

In this paper, I wish to give a survey of a branch of formal epistemology that I will call *epistemic utility theory*. In epistemic utility theory, the descriptive and normative stages are not novel. We follow Bayesianism and other theories of partial beliefs in modelling an agent's epistemic state at a given time t by a belief function b_t, which takes each proposition A about which the agent has an opinion and returns a real number $b_t(A)$ such that $0 \leq b_t(A) \leq 1$. We take $b_t(A)$ to measure the agent's degree of belief in A at time t.

Throughout, we represent propositions as sets of possible worlds. We denote by W the set of all possible worlds about which our agent has an opinion, and we assume that W is finite. We let $\mathcal{F}$ be the algebra over W that represents the set of all propositions about which our agent has an opinion. We denote by $\mathcal{B}$ the set of possible belief functions on this algebra: that is, $\mathcal{B} = \{b : \mathcal{F} \to [0,1]\}$. We denote by $\mathcal{P}$ the set of belief functions that satisfy the axioms of finitely additive probability, and by $\mathcal{N}$ the set of belief functions that do not satisfy those axioms. Thus, $\mathcal{P}$ and $\mathcal{N}$ partition $\mathcal{B}$.

On the whole, the norms that we seek to justify in epistemic utility theory are those endorsed by the Bayesian. For instance:

Probabilism For any time t, it ought to be the case that $b_t \in \mathcal{P}$; that is, b_t ought to be a finitely additive probability function on $\mathcal{F}$.

Conditionalization If, between t and t', the agent learns the proposition E with certainty and nothing more, and if $b_t(E) > 0$, then it ought to be the case that

$$b_{t'}(\cdot) = b_t(\cdot|E) =_{df.} \frac{b_t(\cdot \cap E)}{b_t(E)}$$

Thus, the novelty of epistemic utility lies in the justificatory stage. Before I explain in what this novelty consists, let me review the traditional moves made by Bayesians in the justificatory stage.

Bayesians tend to appeal to one of two putative fundamental norms:

Undutchbookable An agent ought not to have a belief function that would lead her to consider as fair each of a series of bets that would, if taken together, be sure to lose her money—such a series of bets is known as a *Dutch book*.

Consistent preferences An agent ought not to have a belief function that gives rise to an inconsistent set of preferences when combined with her utility function in the standard way.

From the first, by means of a mathematical result known as a *Dutch book theorem*, Bayesians conclude that an agent ought to obey Probabilism ([16], [2]). From slight amendments to the first, and again by means of Dutch book theorems, they conclude that an agent ought to obey Conditionalization ([12]) together with a host of extensions to the Bayesian norms, such as Regularity ([18]), Reflection Principle ([21]), and Jeffrey Conditionalization ([19]). From the second putative fundamental norm, by means of a mathematical result known as a *representation theorem*, Bayesians conclude that an agent ought to obey Probabilism ([7], [17], [13]).

The literature is teeming with objections to these approaches to the justificatory stage, as well as with increasingly sophisticated versions of these approaches that hope to avoid these objections. However, one objection stands out for its simplicity and power. According to this objection, the sort of justification just considered fails because it fails to identify what we really think is irrational about someone whose degrees of belief violate the axioms of the probability calculus, or who updates in the face of new evidence in some way other than by conditionalizing. If Paul believes that Linda is both a bank teller and a political activist more strongly than he believes that she is a bank teller, we regard him as irrational. But this is not because his partial beliefs will lead him to consider a Dutch book fair, or because the preferences to which his beliefs will give rise when combined with his utility function will be inconsistent. These latter facts hold, and they are presumably undesirable for Paul; but they are not relevant to the irrationality that we ascribe

2

to him. Intuitively, what is irrational about his partial beliefs is something purely epistemic; it is not even partly pragmatic. This is the first observation that motivates epistemic utility theory.

The second observation that motivates epistemic utility theory is that epistemic states can be treated as epistemic acts. That is, we can treat an agent who is in a particular epistemic state as an agent who has performed a particular sort of act, namely, the act of adopting that epistemic state. Putting these two observations together gives epistemic utility theory. Since an epistemic state is treated as a kind of act, we can assess the rationality of being in a particular epistemic state at a particular time using the apparatus of utility theory, which we traditionally use to assess the rationality of a particular sort of non-epistemic act. In utility theory, we appeal to an agent's utility function U, which takes an action A from the set $\mathcal{A}$ of possible actions that the agent might perform, together with a possible world $w \in W$, and returns a real number or $-\infty$ or ∞, which we denote $U(A, w)$, that measures the degree to which the agent values the outcome of act A at world w. And we state norms that govern which act she should choose to perform either in terms only of her utility function, or in terms of both her utility and belief functions. Typically, of course, the agent will be represented as valuing the pragmatic, non-epistemic features of these possible outcomes, such as the level of well-being it entails for her, and this will be reflected in the utility function. However, if we are assessing the rationality of epistemic acts in which a particular epistemic state is adopted, there is no reason why the agent could not be represented as valuing the purely epistemic features of the outcomes of these epistemic acts at different worlds. All of this would then be reflected in an epistemic utility function EU, which would take a belief function $b \in \mathcal{B}$, together with a possible world w, and return a real number or $-\infty$ or ∞ that measures the degree to which the agent would value b at w. With this in hand, we could then appeal to the same norms that govern which non-epistemic act an agent should choose to perform to give the norms that govern which epistemic states an agent should adopt. This is the strategy of epistemic utility theory. In the rest of the paper, I review the results it has yielded so far, and I suggest work that needs to be done in the future. First, however, we must answer a possible objection to the general strategy.

It might be complained that epistemic states cannot be treated as epistemic acts, since they are not something over which we have control, unlike acts: we cannot choose what to believe, but we can choose how to act. Of course, both conjuncts of this objection have been denied: *voluntarism* denies that we cannot choose our epistemic states; and *scepticism about free will* denies that we can choose our acts. But we would not wish our epistemology to depend on such controversial claims. Fortunately, we do not have to. Decision theory has two purposes: prior to a decision over which the agent has control, it can be used to help her to decide rationally what to do—call this the *prescriptive use*; and after an act has been performed, it can be used to determine whether that act was rational—call this the *evaluative use*. When we

appeal to decision theory in epistemic utility theory, we wish to appeal only to its evaluative use, and that is available even when the act under consideration was not within the control of the agent.

2 Probabilism, propriety, and act-type dominance

We begin with a collection of arguments for Probabilism that share a similar structure: each appeals to a version of the decision-theoretic norm Dominance; and each assumes amongst its premises a permissive version of Probabilism, the mandatory version of which they seek to establish. The arguments are due to Joyce ([9]) and Predd, et al. ([15]), though Predd, et al. do not explicitly endorse this normative reading of their results.

2.1 The arguments

To state the results, we require some terminology.

Definition 2.1 (Weak and strong dominance) *Suppose $A, A' \in \mathcal{A}$; and suppose that U is a utility function. Then*

- A' weakly dominates A relative to U if

  (i) $U(A', w) \geq U(A, w)$ *for all $w \in W$; and*

  (ii) $U(A', w) > U(A, w)$ *for some $w \in W$.*

- A' strongly dominates A relative to U if

  (i) $U(A', w) > U(A, w)$ *for all $w \in W$.*

Definition 2.2 (Weak and strong act-type dominance) *Now suppose that $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{A}$ together partition $\mathcal{A}$. Then*

- $\mathcal{A}_1$ weakly act-type dominates $\mathcal{A}_2$ relative to U if

  (i) *Every act in $\mathcal{A}_2$ is weakly dominated by an act in $\mathcal{A}_1$ relative to U; and*

  (ii) *No act in $\mathcal{A}_1$ is weakly dominated by any other act relative to U.*

- $\mathcal{A}_1$ strongly act-type dominates $\mathcal{A}_2$ relative to U if

  (i) *Every act in $\mathcal{A}_2$ is strongly dominated by an act in $\mathcal{A}_1$ relative to U; and*

  (ii) *No act in $\mathcal{A}_1$ is weakly dominated by any other act relative to U.*

With this terminology in hand, we can state two act-type versions of Dominance:

> Weak Act-Type Dominance If $\mathcal{A}_1$ weakly act-type dominates $\mathcal{A}_2$ relative to U, then the agent ought to perform an agent in $\mathcal{A}_1$.

4

Strong Act-Type Dominance If $\mathcal{A}_1$ strongly act-type dominates $\mathcal{A}_2$ relative to U, then the agent ought to perform an agent in $\mathcal{A}_1$.

The various arguments we are considering in this section attempt to justify Probabilism by appealing to Weak or Strong Act-Type Dominance. They begin by dividing an agent's possible epistemic acts—that is, the belief functions she might adopt at a particular time—into those that satisfy Probabilism and those that do not. They then present a list of features that they claim an epistemic utility function ought to boast. And finally they prove that, for any epistemic utility function EU that has these features, the set $\mathcal{P}$ of epistemic acts that satisfy Probabilism either weakly or strongly act-type dominates the set $\mathcal{N}$ of epistemic acts that violate it relative to EU. They conclude that Probabilism is correct.

Before I present the mathematical theorems upon which these arguments rely, I survey the features that have been proposed as necessary for a legitimate epistemic utility function. The first three putative necessary conditions on an epistemic utility function EU each say that EU should not rule out as irrational prior to any evidence those belief functions that satisfy Probabilism. In this sense, they state weak permissive versions of Probabilism: where Probabilism states that having a probabilistic belief function is mandatory, the following three conditions on EU demand that the epistemic utility function should make it the case that such belief functions are at least permitted prior to any evidence.

Definition 2.3 (Propriety) *An epistemic utility function EU is* proper *if, for all $p \in \mathcal{P}$ and $b \in \mathcal{B}$, if $b \neq p$, then, prior to any evidence, p expects itself to have at least as great epistemic utility relative to EU as it expects b to have.*

That is, for all $p \in \mathcal{P}$ and $b \in \mathcal{B}$, if $b \neq p$, then

$$\mathrm{Exp}_p^W(p) = \sum_{w \in W} p(w)EU(p, w) \geq \sum_{w \in W} p(w)EU(b, w) = \mathrm{Exp}_p^W(b)$$

where we abuse notation and write $p(w)$ for $p(\{w\})$.

Definition 2.4 (Strict Propriety) *An epistemic utility function EU is* strictly proper *if, for all $p \in \mathcal{P}$ and $b \in \mathcal{B}$, if $b \neq p$, then, prior to any evidence, p expects itself to have greater epistemic utility relative to EU than it expects b to have.*

That is, for all $p \in \mathcal{P}$ and $b \in \mathcal{B}$, if $b \neq p$, then

$$\mathrm{Exp}_p^W(p) = \sum_{w \in W} p(w)EU(p, w) > \sum_{w \in W} p(w)EU(b, w) = \mathrm{Exp}_p^W(b)$$

Definition 2.5 (Coherent Admissibility) *An epistemic utility function EU is* coherent admissible *if, for all $p \in \mathcal{P}$ and $b \in \mathcal{B}$, if $b \neq p$, then b does not weakly dominate p relative to EU.*

We will have much to say about these three properties of *EU* in section 2.2 below. Before that, however, we turn to the two further features that are demanded of *EU* by the arguments we are considering in this section.

Definition 2.6 (Truth-Directedness) *An epistemic utility function EU is truth-directed if, for all belief functions $b, b' \in \mathcal{B}$, and all worlds $w \in W$, if*

(i) $|b(A) - \chi_A(w)| \leq |b'(A) - \chi_A(w)|$ *for all propositions $A \in \mathcal{F}$ and*

(ii) $|b(A) - \chi_A(w)| < |b'(A) - \chi_A(w)|$ *for some proposition $A \in \mathcal{F}$*

then

$$EU(b', w) < EU(b, w)$$

where $\chi_A : W \to \{0, 1\}$ is the characteristic function of A: that is, $\chi_A(w) = 1$, if $w \in A$; $\chi_A(w) = 0$, if $w \notin A$.

Thus, an epistemic utility function is truth-directed if, whenever b is always at least as close to the truth as b' and sometimes closer, the epistemic utility of b is greater than the epistemic utility of b'.

Definition 2.7 (Additivity) *An epistemic utility function EU is additive if, for each $A \in \mathcal{F}$, there is $u_A : [0, 1] \times \{0, 1\} \to [0, \infty]$ such that*

$$EU(b, w) = \sum_{A \in \mathcal{F}} u_A(b(A), \chi_A(w))$$

Thus, an epistemic utility function is additive if the epistemic utility of b at w is obtained by taking, for each proposition $A \in \mathcal{F}$, a measure u_A of the 'local' epistemic utility of the degree of belief $b(A)$ at w, and then summing together all of these 'local' epistemic utilities.

We are now in a position to state the three mathematical theorems that are taken to justify Probabilism on the basis of the versions of Dominance stated above.

Theorem 2.8 (Predd, et al.) *Suppose*

(i) *EU is proper;*

(ii) *EU is additive;*

(iii) *For all $A \in \mathcal{F}$, $u_A(x, 0)$ and $u_A(x, 1)$ are continuous on $[0, 1]$.*

Then $\mathcal{P}$ weakly act-type dominates $\mathcal{N}$ relative to EU.

Theorem 2.9 (Predd, et al.) *Suppose*

(i) *EU is strictly proper;*

(ii) *EU is additive.*

(iii) *For all $A \in \mathcal{F}$, $u_A(x, 0)$ and $u_A(x, 1)$ are continuous on $[0, 1]$.*

Then $\mathcal{P}$ strongly act-type dominates $\mathcal{N}$ relative to EU.

Theorem 2.10 (Joyce) *Suppose*

(i) *EU is truth-directed;*

(ii) *EU is coherent admissible;*

(iii) *For $b \in \mathcal{B}$ and $w \in W$, $EU(b,w)$ is finite;*

(iv) *For all $w \in W$, $EU(b,w)$ is continuous on $\mathcal{B}$.*

Then $\mathcal{P}$ strongly act-type dominates $\mathcal{N}$ relative to EU.

2.2 Propriety, strict propriety, and coherent admissibility

I turn now to consider an objection to these arguments. It concerns the claims that propriety, strict propriety, or coherent admissibility is a necessary feature of a legitimate epistemic utility function. The literature contains two sorts of argument for these claims. I follow Gibbard ([4]) in calling the first sort the *arguments from immodesty*; and I follow Oddie ([14]) in calling the second sort the *arguments from conservativism*.

The objection I wish to raise is not directed against the claims that an epistemic utility function must be proper, or strictly proper, or coherent admissible. Instead, I object that these claims cannot form the premise of an argument that seeks to delimit the set of legitimate epistemic utility functions; rather, if they are true at all, they must be corollaries of such an argument.

2.2.1 The arguments from immodesty

First, some terminology:

Definition 2.11 (Grades of modesty) *Suppose b is a belief function. Then*

- *b is extremely modest if there is a belief function b' such that b expects b' to have* greater *expected epistemic utility than b expects itself to have.*

- *b is moderately modest if there is a belief function b' such that b expects b' to have* at least as great *expected epistemic utility as b expects itself to have.*

- *b is quite modest if there is a belief function b' that weakly dominates b.*

Then the argument from immodesty in favour of demanding propriety runs as follows ([4], [9]):

    Argument from Weak Immodesty to Propriety

    (1) Permissive Probabilism Prior to any evidence, each probabilistic belief function is rationally permitted.

    (2) Weak Immodesty No extremely modest belief function is rationally permitted at any time.

(3) Therefore, *EU* must be proper.

The argument is valid, so we consider the premises. The assumption of Permissive Probabilism is shared by all the arguments we will consider, so we postpone its discussion until later (section 2.2.3). Thus, let us consider Weak Immodesty. As Joyce notes, this premise is a probabilistic version of the norm for full beliefs known as *Moore's paradox*, which says that it is irrational for someone to believe '*A*, but I do not believe *A*'. The problem with such a belief is that it undermines itself. Similarly, a belief function that expects a different belief function to be better from a purely epistemic point of view than it expects itself to be undermines itself. Therefore, it is never rationally permitted. This seems reasonable.

Of course, this version of the argument from immodesty will only deliver the demand of propriety required for this first argument of Predd, et al. It will not support the demand of strict propriety required by their second argument, nor Joyce's demand of coherent admissibility. To obtain the conclusion that *EU* must be strictly proper, we must replace Weak Immodesty by

(2′) Strong Immodesty 1 No moderately modest belief function is rationally permitted at any time.

Unfortunately, this is considerably less plausible than Weak Immodesty. No analogue of Moore's paradox threatens here because a moderately modest belief function does not undermine itself; rather, it merely expects another belief function to be as good, and it is far from clear that this is an epistemic defect.

A similar problem arises if we wish to obtain the conclusion that *EU* must be coherent admissible. To obtain this conclusion, we must replace Weak Immodesty by

(2″) Strong Immodesty 2 No quite modest belief function is rationally permitted at any time.

This may seem more plausible than Strong Immodesty. Surely any weakly dominated belief function undermines itself in exactly the way that a probabilistic version of Moore's paradox rules out. But this is not necessarily so. Suppose that b is a quite modest belief function that is weakly dominated by b'; and suppose further that, for all those worlds $w \in W$ such that $EU(b', w) > EU(b, w)$, we have that $b(w) = 0$. Then b will expect b' to have the same epistemic utility that it expects itself to have. Moreover, it seems that b does not undermine itself in this case. After all, in all the worlds that it considers possible (that is, those for which $b(w) > 0$) it has no less epistemic utility than b'. So it does not undermine itself. This is just the familiar point that an act that weakly dominates all other acts is not necessarily the only rational act to choose, since the agent's belief function might rule out as impossible those worlds at which the domination is strict.

8

2.2.2 The argument from conservatism

Let us see whether the argument from conservatism fares any better. It runs as follows:

Argument from Conservatism

(1) Permissive Probabilism Prior to any evidence, each probabilistic belief function is rationally permitted.

(2) Conservatism If b is rationally permitted, then it is not rational to abandon b in favour of an alternative belief function b' in the absence of any new evidence.

(3) Maximize Expected Utility It is rationally permitted to perform an act if, and only if, that act maximizes expected utility.

(4) Maximal Epistemic Expected Utility Exists For any belief function b, there is a belief function b' for which $\mathrm{Exp}_b^W(\cdot)$ is maximal.

(5) Therefore, *EU* must be strictly proper, and thus proper and coherent admissible.

The argument is valid. But what of Conservatism? I suspect that this principle seems plausible only when we restrict our attention to the familiar example of an agent who arbitrarily abandons her original epistemic state in favour of another without evidence: for instance, the religious convert who, without any new evidence, suddenly shifts from a low to a high degree of belief in the existence of God. However, when we broaden our view and consider the less familiar example of an agent who, in the absence of new evidence, shifts from one epistemic state to another because the original one expects the other to be just as good, the intuitive force of Conservatism disappears. Of such an agent, the proponent of Conservatism would have to say that her original epistemic state was not rationally permitted, and this seems too quick.

2.2.3 Permissive probabilism

Finally, we turn to the assumption of Permissive Probabilism. All existing arguments in favour of this claim proceed along similar lines. They identify a particular feature **X** of belief functions; they argue that **X** is a desirable feature for a belief function to have; they show that each probabilistic belief function has feature **X**; and they conclude that it is rationally permitted to have any probabilistic belief function prior to accumulating evidence. For instance, the desirable feature **X** might be the feature had by a belief function just in case there is a world in which the objective chances, or the long-run relative frequencies, match the degrees of belief assigned by the belief function (§8, [9]).

However, if Permissive Probabilism is established in this way, the arguments from immodesty and conservatism become invalid because they equivocate on the notion of *rational permission* that they employ. Consider, for instance, the argument from weak immodesty to propriety. Suppose that its

first premise relies on the sort of argument I have just outlined. Then that premise should read:

(1′) Permissive Probabilism**ₓ** Prior to any evidence, each probabilistic belief function is rationally permitted *by an epistemic utility function that values only* **X**.

But, in that case, the argument is invalid. Its premises only entail that an epistemic utility function *that values only* **X** must be proper. But the conclusion is that *any* legitimate epistemic utility function must be proper. That is what is required in order to mobilize the arguments for Probabilism by Joyce and Predd, et al. But, absent further argument, we have no reason to think that only epistemic utility functions that value only **X** are legitimate. Thus, it seems that these arguments fail.

Further scepticism about propriety, strict propriety, and coherent admissibility comes from appreciating just how strong Permissive Probabilism really is. It amounts to an extreme form of *subjectivism*, for it states that, at the beginning of an agent's epistemic life, prior to the accumulation of any evidence, any probabilistic belief function is rationally permitted. But many philosophers have wished to deny that. For instance, those who favour Regularity deny that a belief function for which $b(A) = 0$ for some $A \neq \varnothing$ is rational at any time in an agent's epistemic life. Of course, this contradicts Conditionalization, but the tension is removed if the demand is made only for the first point in an agent's epistemic life, and this would be enough to refute Permissive Probabilism if it were correct. Similarly, the *objectivist Bayesians* claim that there is only *one* probabilistic belief function that is rationally permitted prior to any evidence ([6], [22]). Thus, it seems again that, if Permissive Probabilism is true, then it should be a corollary of an argument from epistemic utility, not a premise in that argument. To include it as a premise begs too many of the questions that we wish epistemic utility theory to answer.

3 Conditionalization, strict propriety, and maximizing expected epistemic utility

The arguments considered in the previous section sought to establish Probabilism by appealing to epistemic utility functions that do not rule out any probabilistic belief functions as irrational. In this section, we consider an argument due to Greaves and Wallace that seeks to establish Conditionalization by appealing to epistemic utility functions that do not rule just one probabilistic belief function as irrational ([5]). While the arguments of the previous section appealed to act-type versions of Dominance, this argument appeals to Maximize Expected Utility. Throughout, we assume that Probabilism has been established.

At time t, an agent has a belief function b_t such that $b_t(E) > 0$. Between t and t' she learns the proposition E (and nothing stronger) with certainty. She

is thus faced with a range of epistemic acts from which she must choose: she must choose which belief function to adopt. The natural norm that governs this choice is the following version of Maximize Expected Utility:

Maximize Expected Utility In Light Of E If, between t and t', an agent obtains evidence that restricts the set of epistemically possible worlds to $E \in \mathcal{F}$, then she ought to adopt a belief function $b_{t'}$ at time t such that, for all b,

$$\text{Exp}_{b_t}^E(b_{t'}) = \sum_{w \in E} b_t(w)EU(b_{t'}, w) > \sum_{w \in E} b_t(w)EU(b, w) = \text{Exp}_{b_t}^E(b)$$

Note that the sum ranges only over the set of worlds that are epistemically possible for the agent at t'. And the weightings are provided by the original belief function b_t.

Suppose now that we demand that our epistemic utility function satisfy the following local version of strict propriety:

Definition 3.1 (Local strict propriety for $b_t(\cdot|E)$) *An epistemic utility function EU is* locally strictly proper *for $b_t(\cdot|E)$ if, for all $b \neq b_t(\cdot|E)$,*

$$\text{Exp}_{b_t(\cdot|E)}^W(b_t(\cdot|E)) = \sum_{w \in W} b_t(w|E)EU(b_t(\cdot|E), w) > \sum_{w \in W} b_t(w|E)EU(b, w) = \text{Exp}_{b_t(\cdot|E)}^W(b)$$

Then the following theorem shows that Conditionalization follows:

Theorem 3.2 (Greaves and Wallace) *Suppose EU is locally strictly proper for $b_t(\cdot|E)$. Then, for all $b \neq b_t(\cdot|E)$,*

$$\text{Exp}_{b_t}^E(b_t(\cdot|E)) = \sum_{w \in E} b_t(w)EU(b_t(\cdot|E), w) > \sum_{w \in E} b_t(w)EU(b, w) = \text{Exp}_{b_t}^E(b)$$

Thus, if *EU* is locally strictly proper for $b_t(\cdot|E)$, then one ought to update by conditionalization. Unfortunately, of course, the demand of local strict propriety for $b_t(\cdot|E)$ is vulnerable to the same objections as the global version considered in the previous section. Arguments to establish it will make use of a similarly localized version of Permissive Probabilism, as well as Strong Immodesty 1. And we have seen the fate of these already.

4   The virtue of accuracy

In this final section, I consider two arguments in epistemic utility theory that do not appeal to a permissive version of Probabilism. The first is due to Joyce ([8]); the second to Leitgeb and Pettigrew ([10], [11]).

Both arguments follow the same strategy. They begin with the claim that the ultimate epistemic virtue is *accuracy* or *closeness to the truth value*. They then attempt to characterize the epistemic utility functions that measure accuracy. Finally, they employ decision-theoretic norms of the sort we have met

above to derive certain epistemic norms using these epistemic utility functions.

Joyce's characterization of the accuracy-measuring utility functions allows him to show that, for every non-probabilistic belief function b, there is a probabilistic belief function p that strongly dominates it. That is, he proves the first half of what is required to show that the set $\mathcal{P}$ of probabilistic belief functions strongly act-type dominates the set $\mathcal{N}$ of non-probabilistic belief functions. However, he does not prove the second half, which requires that there is no probabilistic belief function p that is even weakly dominated by another belief function. Thus, as Joyce admits, his argument fails (264, [9]).

Leitgeb and Pettigrew's conditions on a measure of accuracy narrow down the legitimate epistemic utility functions to a unique (up to positive linear transformation) function, called the *global quadratic accuracy measure*:

$$QG(b, w') = 1 - \sum_{w \in W} |b(w) - \chi_{\{w'\}}(w)|^2$$

Unfortunately, QG does not discriminate between belief functions that agree on degrees of belief they assign to individual worlds, but disagree on the degrees of belief they assign to more general propositions. Thus, in particular, it cannot be used to establish that belief functions ought to be finitely additive, as Probabilism demands. To rectify this shortcoming, Leitgeb and Pettigrew also present arguments in favour of a particular *local epistemic utility function*: whereas a (global) epistemic utility function measures the epistemic utility of a whole belief function at a world, a local epistemic utility function measures the epistemic utility of a particular degree of belief in a particular proposition at a world. In particular, they argue that the only local epistemic utility function (up to positive linear transformation) that measures the accuracy of degree of belief x in proposition A at world w is the *local quadratic accuracy measure*:

$$QL(x, A, w) = 1 - |x - \chi_A(w)|^2$$

They use this to establish Probabilism. They assume the following norm:

> Weak Local Immodesty about Accuracy Suppose b is a belief function. And suppose there exists a proposition A and a possible degree of belief r such that b expects the accuracy of degree of belief r in A to be greater than it expects the accuracy of degree of belief $b(A)$ in A to be. Then b is irrational.

And they prove the following theorem:

Theorem 4.1 (Leitgeb and Pettigrew) *Suppose $E \in \mathcal{F}$. Then the following two propositions are equivalent:*

(i) *b is a probability function and $b(E) = 0$.*

(ii) *b is a belief function such that $b(w) = 0$ for all $w \in E$ and, for all $A \in \mathcal{F}$,*

$$\sum_{w \in E} b(w) QL(x, A, w)$$

is maximal for $x = b(A)$.

Having done that, they are able to use QG to establish their other conclusions. They establish Conditionalization using an argument similar to that used by Greaves and Wallace. And they show that there are situations in which it is irrational by the lights of QG to update using Richard Jeffrey's generalization of Conditionalization, known as Jeffrey Conditionalization. They propose an alternative updating norm, which they establish by showing that it follows from the relevant version of Maximize Expected Utility and the characterization of QG.

By their own admission, the central problem with Leitgeb and Pettigrew's account is that it relies on certain geometric assumptions that seem stronger than is warranted by purely epistemic considerations. It would be preferable to excise these assumptions, yet retain the conclusions.

5   Conclusion

In sum, epistemic utility theory has so far furnished us with a number of arguments for some of the central norms governing partial beliefs. Of course, some are stronger than others, and it seems that none is yet decisive; each relies on a premise that we might reasonably question. To conclude, I present questions that need to be answered to improve these arguments, and to extend them to further norms that have not yet been considered.

- Is it legitimate to employ the notion of *expected utility* when the belief function by the lights of which the expected utility is calculated is not a probability function? (589, [8])

- To what extent do Leitgeb and Pettigrew's results rely on the particular geometrical assumptions they make?

- Can Joyce's geometrical characterization of accuracy-measuring epistemic utility functions be improved to allow a proof that, for all such functions EU, the probabilistic belief functions strongly act-type dominate the non-probabilistic belief functions relative to EU?

- Can we exploit the theorems considered in section 2.2 without assuming propriety, strict propriety, or coherent admissibility by instead demanding that our epistemic utility function have certain properties that together entail these features?

- What light does epistemic utility theory shed on the more controversial questions about partial beliefs? For instance, Elga's Sleeping Beauty problem [3], van Fraassen's Judy Benjamin problem [20], the Doomsday Argument and the Anthropic Principle [1], and van Fraassen's Reflection Principle [21].

References

[1] Nick Bostrom. *Anthropic Bias: Observation Selection Effects in Science and Philosophy*. Routledge, New York, 2002.

[2] Bruno de Finetti. Sul significato soggettivo della probabilita. *Fundamenta Mathematicae*, 17:298–329, 1931.

[3] Adam Elga. Self-locating belief and the Sleeping Beauty problem. *Analysis*, 60(2):143–147, 2000.

[4] Allan Gibbard. Rational Credence and the Value of Truth. In T. Gendler and J. Hawthorne, editors, *Oxford Studies in Epistemology*, volume 2. Oxford University Press, 2008.

[5] Hilary Greaves and David Wallace. Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility. *Mind*, 115(459):607–632, 2006.

[6] E. T. Jaynes. *Probability Theory: The Logic of Science.* Cambridge University Press, Cambridge, UK, 2003.

[7] Richard Jeffrey. *Logic of Decision*. McGraw-Hill, New York, 1965.

[8] James M. Joyce. A Nonpragmatic Vindication of Probabilism. *Philosophy of Science*, 65(4):575–603, 1998.

[9] James M. Joyce. Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief. In F. Huber and C. Schmidt-Petri, editors, *Degrees of Belief*. Springer, 2009.

[10] Hannes Leitgeb and Richard Pettigrew. An Objective Justification of Bayesianism I: Measuring Inaccuracy. *Philosophy of Science*, 2010.

[11] Hannes Leitgeb and Richard Pettigrew. An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy. *Philosophy of Science*, 2010.

[12] David Lewis. Why Conditionalize? In *Papers in Metaphysics and Epistemology*. Cambridge University Press, Cambridge, UK, 1999.

[13] Patrick Maher. *Betting on Theories*. Cambridge Studies in Probability, Induction, and Decision Theory. Cambridge University Press, Cambridge, 1993.

[14] Graham Oddie. Conditionalization, Cogency, and Cognitive Value. *British Journal for the Philosophy of Science*, 48:533–41, 1997.

[15] Joel Predd, Robert Seiringer, Elliott H. Lieb, Daniel Osherson, Vincent Poor, and Sanjeev Kulkarni. Probabilistic Coherence and Proper Scoring Rules. ms.

[16] Frank P. Ramsey. Truth and Probability. *The Foundations of Mathematics and Other Logical Essays*, pages 156–198, 1931.

[17] Leonard J. Savage. *The Foundations of Statistics*. John Wiley & Sons, 1954.

[18] Abner Shimony. Coherence and the Axioms of Confirmation. *Journal of Symbolic Logic*, 20:1–28, 1955.

[19] Brian Skyrms. Dynamic Coherence and Probability Kinematics. *Philosophy of Science*, 54(1):1–20, 1987.

[20] Bas C. van Fraassen. A Problem for Relative Information Minimizers. *The British Journal for the Philosophy of Science*, 32(4):375–379, 1981.

[21] Bas C. van Fraassen. Belief and the Will. *Journal of Philosophy*, 81:235–56, 1984.

[22] Jon Williamson. Motivating Objective Bayesianism: From Empirical Constraints to Objective Probabilities. In W. L. Harper and G. R. Wheeler, editors, *Probability and Inference: Essays in Honor of Henry E. Kyburg Jr.*, pages 155–183. College Publications, London, 2007.