Justifying conditionalization: Conditionalization maximizes expected epistemic utility

Hilary Greaves and David Wallace*

February 26, 2005

Abstract

According to Bayesian epistemology, the epistemically rational agent updates her beliefs by conditionalization: that is, her posterior subjective probability after taking account of evidence X, p_{new}, is to be set equal to her prior conditional probability $p_{old}(\cdot/X)$. Bayesians can be challenged to provide a justification for their claim that conditionalization is recommended by *rationality* — whence the normative force of the injunction to conditionalize?

There are several existing justifications for conditionalization, but none directly addresses the idea that conditionalization will be epistemically rational if and only if it can reasonably be expected to lead to *epistemically good outcomes*. We apply the approach of cognitive decision theory to provide a justification for conditionalization using precisely that idea. We assign epistemic utility functions to epistemically rational agents; an agent's epistemic utility is to depend both upon the actual state of the world and on the agent's credence distribution over possible states. We prove that, under independently motivated conditions, conditionalization is the unique updating rule that maximizes expected epistemic utility.

1 Introduction: Justifying conditionalization

According to Bayesian orthodoxy, the ideal epistemic agent can be modeled as follows. The agent contemplates a set Ω of possible worlds. At every time, the agent's epistemic state can be represented by a probability distribution p over Ω (that is, probabilism holds). A learning event occurs when, for some subset X of Ω, the agent learns that the actual world is a member of X. On learning this, the agent updates her probability distribution by conditionalization on X: that is, she sets $p_{new}(\cdot) = p_{old}(\cdot/X)$.

Hilary Greaves, Department of Philosophy, Rutgers University, 26 Nichol Avenue, New Brunswick, NJ 08901-2882, USA. *Email:* hgreaves@rci.rutgers.edu.
David Wallace, Magdalen College, Oxford OX1 4AU, UK. *Email:* david.wallace@philosophy.ox.ac.uk.

Real epistemic agents are not (at least not quite) like this: Bayesian epistemology is a normative theory, rather than a descriptive one. But then Bayesians face the challenge: whence the normative force of the injunction to conditionalize?

One answer is 'it's just *obvious* that conditionalization is the epistemically rational way to update one's credence distribution'. But it is better if we can justify the obvious. Existing justifications for conditionalization include: a Dutch Book argument (if you update your beliefs non-trivially but other than by conditionalization, then a Dutch Book can be made against you [9]); an appeal to a Principle of Minimum Information (conditionalization gives the posterior credence distributions that are 'closest' to your prior distribution while being consistent with the fact that you have just learnt X [13]); an appeal to a Principle of Reflection (Reflection entails conditionalization [11]); and a symmetry argument ([10]:331-337). While these approaches have their interest and merits, none directly addresses the idea that conditionalization will be epistemically rational if and only if it can reasonably be expected to lead to *epistemically good outcomes*.

This paper applies the approach of cognitive decision theory to provide a new justification for conditionalization, based on precisely that idea. We assume that a rational epistemic agent always chooses that epistemic act that maximizes his expected epistemic utility, and prove that, under independently motivated conditions, conditionalization maximizes expected epistemic utility.

Section 2 is an introduction to the basic ideas of the cognitive decision theory we will use, including that of epistemic utility. We illustrate, by means of a toy example, how the fact that a given agent will always maximize expected epistemic utility is supposed to determine his choice of updating policy.

Section 3 contains the central claims of this paper. We note (section 3.1) that an agent faithfully represented by cognitive decision theory might (depending on the form of his utility function) be forbidden to hold some particular subset of possible credence distributions, on the grounds that those credence distributions fail to maximize expected epistemic utility *calculated with respect to themselves*; that is, it may be that some credence distributions undermine themselves. We then prove (sections 3.2 and 3.3) that, of all possible belief-updating policies, conditionalization maximizes expected epistemic utility provided only that the conditional probability distributions are not self-undermining in this sense.

There are two perspectives one might take on this result. First, *if* we regard it as a rationality constraint that the epistemic utility function must not forbid *a priori* any credence distribution in this way, then we will regard the proof as showing that for any epistemically rational agent, conditionalization maximizes expected epistemic utility. Second, whether or not we accept that constraint as a necessary condition for epistemic rationality, it has been shown *for an arbitrary epistemic utility function* that the EU-maximizing agent conditionalizes whenever his conditional posterior does not undermine itself. Since an agent whose conditional posteriors *do* undermine themselves obviously should *not* conditionalize, this is not only as strong an optimality proof as someone

2

unprepared to accept constraints on utility functions can hope for – it is also as strong an optimality proof as she could want.

Section 4 considers a few particular candidate epistemic utility functions, by way of illustration. Section 4.1 suggests a plausible epistemic utility function, and discusses how this particular utility function encodes a respect for epistemic values such as truth and verisimilitude. Section 4.2 discusses a *prima facie* intuitive, but on reflection less plausible, utility function that has been considered in the literature, according to which epistemic utility is *linearly related* to degree of belief in the truth. We discuss an objection to the approach of this paper: the objection that the possibility of utility functions such as this undermines the whole decision-theoretic approach to probabilist epistemic rationality.

Section 5 is the conclusion.

2   Cognitive decision theory

This section introduces the basic ideas of the cognitive decision theory we will use: states, probability distributions, (epistemic) acts, act availability, epistemic utility and expected epistemic utility. We explicate each of these notions below (section 2.1), and illustrate the theory throughout by means of a toy model of a simple cognitive decision process. Following this exposition, section 2.2 mentions, only to set aside, two closely related issues that we do *not* intend to address: the place (or lack of it) of cognitive decision theory in an 'all-things-considered' decision theory, and the relevance (or lack of it) of epistemic utility to the advisability of *gathering*, as opposed to epistemically responding to, evidence.

2.1   The framework of cognitive decision theory

Some cognitive states are, epistemically speaking, better than others. For example, it is presumably (epistemically) better to have higher credences in truths and lower credences in falsehoods. According to the (cognitive) decision-theoretic approach, epistemic rationality consists in taking steps that can reasonably be expected to bring about (epistemically) good outcomes.

Cognitive decision theory provides a framework in which the ideas of the preceding paragraph can be made precise and quantitative. The decision problems with which we will be concerned take the following form. The agent begins in some fixed belief state (that is, he holds some fixed initial credence distribution). He knows that he is about to receive some new piece of information, from among a fixed range of possibilities. Before receiving the information, he chooses an updating policy: that is, he specifies, for each of the possible pieces of new information, how he will change his credence distribution if that turns out to be the information that he does in fact receive. The decision he has to make is the choice of an updating policy.

> **EXAMPLE.** Mike has a coin. He is unsure as to whether or not it is a fair coin — specifically, he assigns 50% credence to its being fair — but he is (let us

suppose) certain that *either* it is fair *or* it is weighted in such a way that the chances for outcomes (Heads, Tails) on a given toss are $(\frac{1}{4}, \frac{3}{4})$ respectively. The coin is about to be tossed; after observing the result of the toss, Mike will reassess his degrees of belief as to whether or not the coin is fair. He must decide in advance how the reassessment will proceed: which credence distribution he will move to if he sees heads, and which if he sees tails. We want to know how that decision should proceed.

The remainder of section 2.1 spells this out in more detail, in a framework of cognitive decision theory. (Cognitive decision theory is in many respects similar to ordinary, prudential decision theory; our framework is loosely based on that of Savage [8].)

States. The agent contemplates a set $\mathcal{S}$ of (mutually exclusive and jointly exhaustive) possible **states of the world**; he is unsure as to which element of $\mathcal{S}$ obtains. $\mathcal{S}$ can be thought of as a partition of the set of possible worlds.[1]

For our toy example, the states might be as follows:

$$\mathcal{S} = \{s_{FH}, s_{FT}, s_{UH}, s_{UT}\}, \text{where} \quad \begin{array}{l} s_{FH} : \text{coin fair, outcome of toss is H} \\ s_{FT} : \text{coin fair, outcome of toss is T} \\ s_{UH} : \text{coin unfair, outcome of toss is H} \\ s_{UT} : \text{coin unfair, outcome of toss is T.} \end{array}$$

Probability distributions. The agent does, however, have definite subjective degrees of belief as to which state obtains: his doxastic state is, at any time, represented by some probability distribution p over $\mathcal{S}$. We write $\mathcal{P}$ for the set of *all* probability distributions over $\mathcal{S}$.

The agent's prior. One particular probability distribution $p^* \in \mathcal{P}$ represents the agent's *prior* doxastic state — his doxastic state before learning the evidence, when he is making his cognitive decision. For our purposes, p^* is arbitrary but is to be held fixed.

Mike's prior doxastic state is represented by the following probability distribution over $\mathcal{S}$:
$$\begin{array}{l} p^*(s_{FH}) = \frac{1}{4}; \\ p^*(s_{FT}) = \frac{1}{4}; \\ p^*(s_{UH}) = \frac{1}{8}; \\ p^*(s_{UT}) = \frac{3}{8}. \end{array}$$

Acts. The set $\mathcal{A}$ of **acts** is supposed to reflect the possible courses of (epistemic) action among which the agent must choose. An epistemic act a is an assignment of a probability distribution to each state $s \in \mathcal{S}$, with the intended interpretation that if $a(s) = p_s$, then p_s is the probability function that an agent performing act a would adopt as his credence distribution if state s obtained. (We make no assumption that the 'choice' between acts is a voluntary one.)

---

[1] We will assume throughout that $\mathcal{S}$ is finite. This is merely for simplicity of exposition.

Act availability; experiments. Not all acts will be **available** to a given agent at a given time. Specifically, suppose that s_1 and s_2 are empirically indistinguishable for an agent. Then he cannot perform any act which requires him to hold different probability distributions according to which of them obtains.

To identify the subset of acts that are available, we introduce the notion of an **experiment**. An experiment is represented mathematically by a partition $\mathbf{E}$ of $\mathcal{S}$ (that is, $\mathbf{E}$ is a set of mutually exclusive and jointly exhaustive subsets of $\mathcal{S}$); we say that experiment $\mathbf{E}$ is *performed* (for a particular agent) when the elements of $\mathbf{E}$ become epistemically distinguishable for the agent. When a particular experiment $\mathbf{E}$ is performed, the available acts are just those acts that always assign the same probability distribution $p_j \in \mathcal{P}$ to any two states that are members of the same element E_j of $\mathbf{E}$; write $\mathcal{A}_{\mathbf{E}} \subseteq \mathcal{A}$ for the set of such available acts. (We will have $\mathcal{A}_{\mathbf{E}} = \mathcal{A}$ only in the somewhat implausible case in which the maximally fine-grained experiment, which distinguishes between every pair of possible worlds, is performed.) Since such acts are prescriptions for how to revise one's beliefs in the light of new evidence, we also call them **updating policies**. (For convenience, for a given act $a \in \mathcal{A}_{\mathbf{E}}$, we will generally write $a(E_j)$ (with $E_j \in \mathbf{E}$) rather than $a(s)$ (with $s \in \mathcal{S}$) for the probability distribution that the act a assigns to each state in E_j.)

> Mike's experiment is given by $\mathbf{E} = \{H, T\}$, where $H = \{s_{FH}, s_{UH}\}$ and $T = \{s_{FT}, s_{UT}\}$. The acts that are available to him as a result of this experiment are just those that assign one probability distribution over $\mathcal{S}$ to H, and another (not necessarily distinct) to T.
>
> Mike's cognitive decision is whether to commit himself to updating by conditionalization from his prior p^* on the result of the coin toss, or by a particular rival updating policy $\mathbf{R}$ (given below) that has just popped into his head.
>
> Updating by conditionalization from the prior p^* would lead to the following possible posteriors:[2]
>
> $$\mathbf{Cond}(H) = p(\cdot/H) \ =: p_H, \quad \text{where} \quad \begin{aligned} p_H(s_{FH}) &= \tfrac{2}{3} \\ p_H(s_{FT}) &= 0 \\ p_H(s_{UH}) &= \tfrac{1}{3} \\ p_H(s_{UT}) &= 0; \end{aligned}$$
>
> $$\mathbf{Cond}(T) = p(\cdot/T) \ =: p_T, \quad \text{where} \quad \begin{aligned} p_T(s_{FH}) &= 0 \\ p_T(s_{FT}) &= \tfrac{2}{5} \\ p_T(s_{UH}) &= 0 \\ p_T(s_{UT}) &= \tfrac{3}{5}. \end{aligned}$$

---

[2] Recall that we have defined an act (updating policy) as a function from states (or disjunctions of states) to probability distributions. According to this definition, conditionalization from some prior p^* and conditionalization from a distinct prior $q \neq p^*$ count as distinct updating policies, as do conditionalization from prior p^* given an experiment $\mathbf{E}$ and conditionalization from p^* given a distinct experiment $\mathbf{E}'$. Strictly, to pick out a unique act, we should therefore write $\mathbf{Cond}_{\mathbf{E}}^{p^*}$; we will shorten this to $\mathbf{Cond}$ since the prior p^* and experiment $\mathbf{E}$ will be held fixed throughout our discussion. This way of speaking has the advantage that the expected utility (with respect to some fixed probability distribution that may or may not equal the agent's prior) of an act will be independent of what the agent's prior was, and independent of which experiment is performed. It has the disadvantage that conditionalization *simpliciter* does not count as a single updating policy, which is perhaps contrary to ordinary usage of 'updating policy'.

We stipulate that the alternative updating policy **R**, on the other hand, is as follows:

$$\mathbf{R}(H) =: q_H, \quad \text{where} \quad
\begin{aligned}
q_H(s_{FH}) &= \tfrac{1}{2}\\
q_H(s_{FT}) &= 0\\
q_H(s_{UH}) &= \tfrac{1}{2}\\
q_H(s_{UT}) &= 0;
\end{aligned}$$

$$\mathbf{R}(T) =: q_T, \quad \text{where} \quad
\begin{aligned}
q_T(s_{FH}) &= 0\\
q_T(s_{FT}) &= \tfrac{1}{4}\\
q_T(s_{UH}) &= 0\\
q_T(s_{UT}) &= \tfrac{3}{4}.
\end{aligned}$$

Cond and **R** are both *available* acts after the coin-flip, since each assigns the same probability distribution (**Cond**(H) and **R**(H) resp.) to s_{FH} as it does to s_{UH}, and the same probability distribution (**Cond**(T) and **R**(T) resp.) to s_{FT} as it does to s_{FT}. That is, neither **Cond** nor **R** makes the unreasonable requirement that Mike must take a different course of epistemic action depending (say) on whether the world, unbeknownst to Mike, happens to be in some fair-coin state or in some unfair-coin state; both updating rules require him to react only to information that the experiment will provide him with.

We offer no intuitive rationale for the rule **R**, and indeed we have none. The point is not that **R** has any intuitive plausibility whatsoever as a serious rival to conditionalization, but rather that **R** is a course of epistemic action that an agent could in principle adopt. Our aim is to show that considerations of intuitive plausibility need not be invoked in order to outlaw **R**, because the inferiority of that updating policy will follow by calculation from the decision-theoretic model.

Epistemic utility functions. A given agent (we are assuming) holds a particular **epistemic utility function** — a function $U : \mathcal{S} \times \mathcal{P} \to \Re$ assigning a real number to each pair consisting of a state and a probability distribution. $U(s, p)$ represents the epistemic value ("epistemic utility") of holding credence function p when state s in fact obtains.[3]

Note that we allow our notion of utility to be *externalist* in the sense that we allow two pairs $<s, p>$, $<s', p>$, in which the agent is in the same cognitive state but a different state of the world in fact obtains, to be valued differently. This is to be expected since the agent may well want his beliefs to be *true*, over and above the subjective feelings associated with being in that belief state.

Presumably, since he is a responsible epistemic agent, Mike attaches a high epistemic utility to having high degrees of beliefs in truths. In that case, his

---

[3]A different sort of cognitive decision theory (e.g. Levi [4], Maher [5]) focusses on cognitive acts that involve *accepting particular propositions*, rather than holding particular probability distributions. The domain of the epistemic utility function for such acceptance-based cognitive decision theories is the set of pairs $<s, A>$ of states and *propositions*: $<s, A>$ is to be read 'accepting proposition A when state s obtains'. We do not regard such theories as necessarily in competition with our own: they are engaged in a different project.

epistemic utility function might look something like this:

For arbitrary state $s \in \mathcal{S}$ and probability distribution p over $\mathcal{S}$,
$$U(s,p) = -(1 - p(s))^2 - \sum_{s' \neq s} (p(s'))^2. \tag{1}$$

Heuristically, we can see that this utility function 'values truth' because it equates the utility of holding credence function p when state s obtains to the sum of two terms, the first of which increases with increasing degree of belief in the true state s, and the second of which increases with decreasing degrees of belief in false states s'. (Note the minus signs.)

Expected epistemic utility. We assume[4] that the rational epistemic agent always performs that act that has the highest expected epistemic utility (with respect to the agent's prior p^*) of all available acts, where the expected epistemic utility of an act a (with repsect to probability distribution p) is given by

$$EU^p(a) = \sum_{s \in \mathcal{S}} p(s) \cdot U(s, a(s))$$

Using the alternative notation mentioned above, in which acts are defined on experiments rather than directly on $\mathcal{S}$, we can also write the expected utility of a as

$$EU^p(a) = \sum_{E_j \in \mathbf{E}} \sum_{s \in E_j} p(s) \cdot U(s, a(E_j)), \tag{2}$$

where $\mathbf{E}$ is the experiment on which the act a is defined.

Iff an act a maximizes expected epistemic utility given an experiment $\mathbf{E}$ (that is, if it has at least as high an expected epistemic utility as any other act in $\mathcal{A}_{\mathbf{E}}$), say that a is **optimal** (given $\mathbf{E}$). Iff a is the *unique* optimal act, say that a is **strongly optimal** (given $\mathbf{E}$). (We will often leave 'given $\mathbf{E}$' implicit.)

Being an epistemically rational agent, Mike will choose whichever updating policy has the higher expected epistemic utility. To see which policy this is, we evaluate the EU of each policy using Mike's prior p^* and his epistemic utility function U, as follows:

Expected epistemic utility of updating by conditionalization from prior p^* given experiment $\mathbf{E}$,

$$
\begin{aligned}
EU^{p^*}(\mathbf{Cond}) &= p^*(s_{FH}) \cdot U(s_{FH}, p_H) + p^*(s_{FT}) \cdot U(s_{FT}, p_T) \\
&\quad + p^*(s_{UH}) \cdot U(s_{UH}, p_H) + p^*(s_{UT}) \cdot U(s_{UT}, p_T) \\
&= \tfrac{1}{4}(-(1 - \tfrac{2}{3})^2 - (\tfrac{1}{3})^2) + \tfrac{1}{4}(-(1 - \tfrac{2}{5})^2 - (\tfrac{3}{5})^2) \\
&\quad + \tfrac{1}{8}(-(1 - \tfrac{1}{3})^2 - (\tfrac{2}{3})^2) + \tfrac{3}{8}(-(1 - \tfrac{3}{5})^2 - (\tfrac{2}{5})^2) \\
&= -0.0479.
\end{aligned}
$$

---

[4]Without providing a justification. Perhaps the justification is that a representation theorem shows that any agent whose preferences over acts satisfy certain axioms must be representable by some utility function U and some probability distribution p such that he always prefers an act a to another b iff $EU^p(a) > EU^p(b)$, with EU^p as defined in (2). Or perhaps the justification is simply that the injunction to maximize EU seems intuitively plausible and gives intuitively plausible results. We won't go into this issue; we take as a premise, for the purposes of this paper, that there is *some* adequate justification.

On the other hand, the expected epistemic utility of adopting the alternative policy **R** is given by

$$
\begin{aligned}
EU^{p^*}(\mathbf{R}) &= p^*(s_{FH}) \cdot U(s_{FH}, q_H) + p^*(s_{FT}) \cdot U(s_{FT}, q_T) \\
&\quad + p^*(s_{UH}) \cdot U(s_{UH}, q_H) + p^*(s_{UT}) \cdot U(s_{UT}, q_U) \\
&= \tfrac{1}{4}(-(1 - \tfrac{1}{2})^2 - (\tfrac{1}{2})^2) + \tfrac{1}{4}(-(1 - \tfrac{1}{4})^2 \\
&\quad - (\tfrac{3}{4})^2) + \tfrac{1}{8}(-(1 - \tfrac{1}{2})^2 - (\tfrac{1}{2})^2) + \tfrac{3}{8}(-(1 - \tfrac{3}{4})^2 - (\tfrac{1}{4})^2) \\
&= -0.120.
\end{aligned}
$$

Since $EU^{p^*}(\mathbf{Cond}) > EU^{p^*}(\mathbf{R})$, Mike will choose to update by **Cond** rather than by the alternative policy **R**.[5]

2.2   Disclaimers

Before proceeding further, we flag two issues that are closely related to the spirit of the decision-theoretic approach, but that we do not (and need not) tackle in this paper: the place (or lack of it) of cognitive decision theory in an 'all-things-considered' decision theory, and the relevance (or lack of it) of epistemic utility to the advisability of *gathering*, as opposed to epistemically responding to, evidence.

Epistemic vs. all-things-considered utility.   There is a sense in which a particular belief state can have a high 'utility' without it being at all *epistemically* rational to pursue that state. A high degree of belief in the existence of a god, for instance, may make me happy and make my life go better in all sorts of ways — and thus be of high 'utility' in a prudential sense — but yet, if there is no god then such a belief state is, *epistemically* speaking, undesirable, and if the evidence for its truth is slim then it is *epistemically* irrational to hold such beliefs. We fully accept that there is also this prudential sense of 'utility', and that the demand to maximize *prudential* utility may conflict with the demands of epistemic rationality. But this does not entail that considerations of *epistemic* utility cannot hope to account for epistemic rationality. An *epistemic* utility function, such as those we work with in this paper, is concerned only with the *epistemic* desirability of belief states. Where epistemic desirability diverges from 'prudential' or 'all-things-considered' desirability, an epistemic utility function tracks only the former. Whether and how an epistemic utility function plays a role in any integrated theory of the 'all-things-considered' rationality of epistemic acts is an important question, but not one we are concerned with here.

---

[5]In any real epistemic decision problem, of course, the agent will not really be deciding only among *two* updating policies — there is always an infinite number of available candidate updating policies, and the agent is to choose the policy that, of *all* those candidates, gives the highest expected epistemic utility. We have restricted consideration to two candidate rules in our toy model only for purposes of illustration.

Epistemic vs. non-epistemic acts. It has been argued that a decision-theoretic framework — cognitive ([6], [2]:127-9) or otherwise ([1]) — can be invoked to *justify experimentation*. However that may be, that is not the project we are engaged in here. Our application of cognitive decision theory is concerned with *purely epistemic acts*; the act of performing a particular experiment, however epistemically motivated, is a non-epistemic act. That is, we assume only that, *given* that one has received a given piece of evidence (against one's will or otherwise), epistemic rationality requires that one then perform the epistemic act of altering one's doxastic state in the manner that, in the light of that evidence, maximizes expected epistemic utility. Whether and how a cognitive decision theory for epistemic acts could be integrated into a satisfactory decision theory, paying due respect to epistemic goods, for choices among non-epistemic acts (including evidence-gathering acts) is an important question, but, again, not one that we are concerned with here.

3 Conditionalization and maximization of expected epistemic utility

This section contains our claims in defense of conditionalization. We proceed in three steps. In section 3.1 we define the notion of a constant act, and a relation we call *recommendation* between probability distributions. We note an important consequence of our assumption that an epistemically rational agent always chooses the epistemic act that maximizes expected epistemic utility: for some utility functions, there exist probability distributions that the ideal agent is forbidden to hold on the grounds that they fail to 'recommend' themselves. In section 3.2 we use the notion of recommendation to define a class of updating policies, *quasi-conditionalization* ($\mathcal{QC}$), and we prove (for an arbitrary epistemic utility function) that $\mathcal{QC}$ is optimal. In section 3.3 we characterize (Corollary 1) a set of epistemic utility functions for which *conditionalization* is optimal.We also prove (Corollary 2; again for an arbitrary epistemic utility function) that in any case conditionalization is optimal if it is even *coherent*, in the sense that the probabilities conditionalization would have the agent adopt are not ones that his own utility function forbids him ever to hold.

3.1 Constant acts, recommendation, self-recommendation and stable utility functions

Constant acts. We will have particular interest in the **constant acts**: those acts that instruct the agent to adopt the *same* probability distribution as his credence function, regardless of which state obtains.

The expression for the expected utility of a constant act takes a particularly simple form. Let k_q denote the constant act that assigns q to all states, for arbitrary $q \in \mathcal{P}$. The expected epistemic utility of k_q, calculated with respect to the agent's current credence distribution, p, is given by

$$EU^p(k_q) = \sum_{s \in S} p(s) \cdot U(s, q). \tag{3}$$

The recommendation relation; stability. The notion of the epistemic utility of a constant act raises an interesting issue that is important to our present project. All constant acts are, of course, always *available* in the sense explicated in section 2.1. So, at every time, the agent is to regard all constant acts as available options between which he can choose. Therefore, an ideally rational agent is able to hold a probability distribution p as his credence distribution only if p, *by the lights of p itself*, maximizes expected utility — that is, only if $(\forall q \in \mathcal{P})(EU^p(k_p) \geq EU^p(k_q))$. If this condition fails, the minute the agent found himself holding p, he would be compelled to move to some other distribution q that maximized EU calculated with respect to p — which is to say that an ideally rational agent could not hold p in the first place, even for a moment.[6]

We make the following definitions:

- Say that p **recommends** q (write $p \xrightarrow{R} q$) iff, when the only available acts are the constant acts, k_q maximizes expected utility calculated with respect to p — that is, iff $\forall r \in \mathcal{P}, EU^p(k_q) \geq EU^p(k_r)$.

- Iff, in addition, p recommends no distribution distinct from q, say that p **strongly recommends** q.

- Iff p recommends p, say that p **is self-recommending**. Iff, in addition, p recommends no distribution distinct from p, say that p is **strongly self-recommending**. Iff p is not self-recommending, say that p is **self-undermining**.

Clearly, the extension of the recommendation relation depends on the utility function U. We can thus classify utility functions based on the structure of the recommendation relations they induce:

- Say that U is **everywhere stable** iff, according to U, every probability distribution is self-recommending.

- Say that U is **everywhere strongly stable** iff, according to U, every probability distribution is strongly self-recommending.

- Say that U is **partly stable** iff, according to U, some probability distributions are self-recommending and others are self-undermining.

- Say that U is **nowhere stable** iff, according to U, every probability distribution is self-undermining.

---

[6] We are assuming here that the ideal agent deliberates at infinite speed and that time is continuous.

We now consider the following question: which of these types of utility function might an epistemically responsible agent hold?

Utility function that are nowhere stable are pathological. In fact they are worse than pathological — it is actually impossible for an ideally rational agent to hold such a utility function, since there is no probability distribution that an ideally rational agent with a nowhere stable utility function could coherently hold. (Neither, though, can he refrain from holding any probability distribution: that would contradict our assumption of probabilism.)

At first sight, perhaps, *partly stable* utility functions (perhaps) also seem to be pathological for the same reason — the notion of an ideal agent who held a partly stable utility function but also held a credence distribution that that utility function deemed self-undermining would be similarly contradictory. A moment's reflection, however, shows that there is no contradiction here. A partly stable utility function merely amounts to a constraint on the credence distributions that the agent may coherently hold; there is no literal inconsistency in holding a partly stable utility function and being an ideal expected-utility maximizer, provided that one's credence distribution is at all times one that that utility function deems self-recommending. It will turn out that the coherence of partly stable utility functions alone prevents us from claiming *categorically* that conditionalization maximizes expected epistemic utility.

The distinction between (merely) everywhere stable utility functions on the one hand, and those that are everywhere *strongly* stable on the other, also deserves comment. An agent who holds a utility function that is everywhere stable but fails to be everywhere *strongly* stable may find himself with a choice between constant acts that are equally good by his own lights: he currently holds credence distribution p, but there may be some distinct credence distribution q such that, by the lights of p itself, k_p and k_q are of equal (optimal) expected epistemic utility. When this occurs, the agent *can* stick to his current distribution p, but it will be equally consistent with ideal rationality if he chooses to move to q on a whim. An agent whose utility function is everywhere *strongly* stable, on the other hand, never has any such free options; he must always, unless new evidence comes along, stick to his current credence distribution. (This distinction will be of some importance in section 3.3.)

3.2 Theorem: Quasiconditionalization maximizes expected epistemic utility

We give names to two updating policies in which we will have particular interest:

- *Conditionalization from prior p^* given experiment $\boldsymbol{E}$* ($\mathbf{Cond}_{\mathbf{E}}^{p^*}$, or $\mathbf{Cond}$ for short) is defined (as usual) by

$$\mathbf{Cond} : \text{For all } E_j \in \mathbf{E}, \mathbf{Cond}(E_j) = p^*(\cdot/E_j).$$

- *Quasi-conditionalization from prior p^* given experiment $\boldsymbol{E}$* ($\mathcal{QC}_{\mathbf{E}}^{p^*}$, or $\mathcal{QC}$ for short) is, in general, a family of policies rather than a single updating

policy. It is defined by

$$\mathcal{QC} : \text{For all } E_j \in \mathbf{E}, \mathcal{QC}(E_j) = \{q \in \mathcal{P} : \mathbf{Cond}(E_j) \xrightarrow{R} q\}.$$

That is, $\mathcal{QC}$ is the set of acts according to which the agent, on receiving evidence E_j, moves, not necessarily to the conditional probability $p^*(\cdot/E_j)$, but to some probability distribution q that *is recommended by* that conditional probability (where 'recommendation' is as defined in section 3.1 above). (The reason $\mathcal{QC}$ does not in general prescribe a unique policy is that $p^*(\cdot/E_j)$ may in general simultaneously recommend more than one distribution.)

Theorem. *Of all acts that are available given an experiment $\mathbf{E}$, each quasi-conditionalizing updating rule is optimal. That is,*

$$\forall \mathbf{Q} \in \mathcal{QC}, \forall \mathbf{R} \in \mathcal{A_E}, EU^{p^*}(\mathbf{Q}) \geq EU^{p^*}(\mathbf{R}).$$

Proof. The expected utility of adopting an arbitrary updating policy $\mathbf{R} \in \mathcal{A}_E$ is given by

$$
\begin{align}
EU^{p^*}(\mathbf{R}) &\equiv \sum_{s \in S} p^*(s) \cdot U(s, \mathbf{R}(s)) \tag{4}\\
&\equiv \sum_{E_j \in \mathbf{E}} \sum_{s \in E_j} p^*(s) \cdot U(s, \mathbf{R}(E_j)) \tag{5}\\
&\equiv \sum_{E_j \in \mathbf{E}} \sum_{s \in E_j} p^*(s \wedge E_j) \cdot U(s, \mathbf{R}(E_j)) \tag{6}\\
&\equiv \sum_{E_j \in \mathbf{E}} p^*(E_j) \cdot \left(\sum_{s \in E_j} p^*(s/E_j) \cdot U(s, \mathbf{R}(E_j)) \right) \tag{7}\\
&\equiv \sum_{E_j \in \mathbf{E}} p^*(E_j) \cdot \left(\sum_{s \in \mathcal{S}} p^*(s/E_j) \cdot U(s, \mathbf{R}(E_j)) \right) \tag{8}\\
&\equiv \sum_{E_j \in \mathbf{E}} p^*(E_j) \cdot EU^{p^*(\cdot/E_j)}(k_{\mathbf{R}(E_j)}). \tag{9}
\end{align}
$$

Substituting a quasi-conditionalizing act $\mathbf{Q} \in \mathcal{QC}$ for $\mathbf{R}$ in line (9) yields an expression for the expected utility of adopting that $\mathbf{Q}$ as one's updating rule:

$$EU^{p^*}(\mathbf{Q}) = \sum_{E_j \in \mathbf{E}} p^*(E_j) \cdot EU^{p^*(\cdot/E_j)}(\mathbf{Q}(E_j)). \tag{10}$$

But, by the definition of quasi-conditionalization, we have, for all rules $\mathbf{R} \in \mathcal{A}$, all quasi-conditionalizing rules $\mathbf{Q} \in \mathcal{QC}$ and all events $E_j \in \mathbf{E}$,

$$EU^{p^*(\cdot/E_j)}(\mathbf{Q}(E_j)) \geq EU^{p^*(\cdot/E_j)}(\mathbf{R}(E_j)) \tag{11}$$

Combining (9), (10) and (11) and noting that the coefficients $p^*(E_j)$ are all nonnegative, we have

$$\forall \mathbf{Q} \in \mathcal{QC}, \forall \mathbf{R} \in \mathcal{A}_E, EU^{p^*}(\mathbf{Q}) \geq EU^{p^*}(\mathbf{R}). \tag{12}$$

$\square$

It is striking that this theorem involves *no assumptions whatsoever* about the nature of the utility function — yet we seem (at first sight) to have given a name to a particular set of epistemic acts (viz. $\mathcal{QC}$) and proved that every act in that set is optimal.

If this were really what we had done, it should arouse puzzlement in anyone acquainted with 'ordinary' (i.e. prudential, non-cognitive) decision theory — it is a familiar point from that theory that tastes are encoded in the utility function, so that one cannot prove anything about the EU of a given act *without* constraining the utility function. (There is no hope, for instance, of proving *from decision theory alone* that a rational agent tries to avoid electric shocks; I may be perfectly 'rational', in the sense that I satisfy the axioms of decision theory, but happen to *like* electric shocks, and accordingly assign high utility to situations in which I receive such shocks.) But our above 'first-sight' gloss on the content of our theorem is, of course, not quite correct. We have *not* shown, in the absence of any information about the utility function, that *some particular act* is optimal. This is because, in the absence of information about the utility function, we have no idea what the recommended probabilities $\{\mathbf{Q}(E_j)\}_{E_j \in \mathbf{E}}$ are. In other words, while we know (without knowing anything about the utility function) that all acts that meet the definition of $\mathcal{QC}$ are optimal, we do *not* know which acts (i.e., which functions from $\mathbf{E}$ to $\mathcal{P}$) those are.

3.3   Corollaries: When conditionalization is optimal

The circumstances under which *conditionalization* (as opposed to 'mere' quasi-conditionalization) is optimal are brought out by the following two corollaries to our theorem.

Corollary 1. *Conditionalization is optimal for a given experiment $\mathbf{E}$ if the conditional probabilities $\{p*(\cdot/E_j) : E_j \in \mathbf{E}\}$ are self-recommending. If, in addition, at least one of the conditional probabilities is strongly self-recommending, then conditionalization is strongly optimal.*

Proof. As shown above (equation (9)), for an arbitrary rule $\mathbf{R} \in \mathcal{A}_\mathbf{E}$,

$$EU^{p^*}(\mathbf{R}) = \sum_{E_j \in \mathbf{E}} p^*(E_j) \cdot EU^{p^*(\cdot/E_j)}(k_{\mathbf{R}(E_j)}) \tag{13}$$

Therefore, in particular,

$$EU^{p^*}(\mathbf{Cond}) = \sum_{E_j \in \mathbf{E}} p^*(E_j) \cdot EU^{p^*(\cdot/E_j)}(k_{\mathbf{Cond}(E_j)}) \tag{14}$$

Suppose now that the conditional probabilities $\{p^*(\cdot/E_j) : E_j \in \mathbf{E}\}$ are self-recommending. That is, suppose that we have, for all $E_j \in \mathbf{E}$ and all probability distributions $q \in \mathcal{P}$,

$$EU^{p^*(\cdot/E_j)}(k_{\mathbf{Cond}(E_j)}) \equiv EU^{p^*(\cdot/E_j)}(k_{p^*(\cdot/E_j)}) \geq EU^{p^*(\cdot/E_j)}(k_q). \qquad (15)$$

In particular, for all updating policies $\mathbf{R} \in \mathcal{A}_{\mathbf{E}}$,

$$EU^{p^*(\cdot/E_j)}(k_{\mathbf{Cond}(E_j)}) \geq EU^{p^*(\cdot/E_j)}(k_{\mathbf{R}(E_j)}). \qquad (16)$$

Thus for all convex combinations of such quantities, $\mathbf{Cond}$ weakly dominates $\mathbf{R}$, and hence (combining (13), (14) and (16)) we have, for all updating policies $\mathbf{R} \in \mathcal{A}_{\mathbf{E}}$,

$$EU^p(\mathbf{Cond}) \geq EU^p(\mathbf{R}) \qquad (17)$$

— that is, conditionalization is optimal.

If, in addition, at least one of the conditional probabilities is *strongly* self-recommending, then strict inequality holds — that is, conditionalization is strongly optimal.

$\square$

Corollary 1 establishes that conditionalization is optimal whenever the conditional probabilities are self-recommending. Now, one who hoped to justify conditionalization within a decision-theoretic framework really could not want stronger support from the mathematics, for the following reason. If the conditional probabilities are *not* self-recommending, conditionalization is obviously not even a live option for our agent — for then, conditionalization advises him to move to probabilities that he never would be able to hold, 'no matter how he arrived at them', while remaining an expected utility maximizer. (Cf. our discussion of self-recommendation in section 3.1.) It would be somewhat worrying if our proof insisted *even then* that conditionalization was optimal. So, corollary 1 conditionalization is optimal whenever conditionalization is even a live option. We stress that this follows from the decision theory alone, with no constraints on the form of the epistemic utility function. This is our first result in support of the normative status of conditionalization.

Our second corollary concerns a second gloss we might put on our result, if we are prepared to accept normative constraints on the form of the epistemic utility function:

Corollary 2. *If the agent's epistemic utility function U is everywhere stable, then conditionalization is optimal. If U is everywhere* strongly *stable, then conditionalization is strongly optimal.*

Proof. This is an immediate consequence of Corollary 1. $\square$

If it is a *rationality constraint* that one's epistemic utility function be everywhere strongly stable (so that one's utility function alone does not preclude holding any particular probability distribution, and always advises one strictly

to stick to one's current credence distribution until and unless new evidence comes along), then Corollary 2 demonstrates that, for any rational agent, conditionalization is the unique updating policy that maximizes expected epistemic utility. This would be a second statement in favor of conditionalization. We find this rationality constraint plausible, but we offer no argument for it here.[7] (The reader may or may not find that she accepts it without argument.) If the constraint is not accepted, our categorical claims are restricted to those we drew above from Corollary 1.

4   A plausible epistemic utility function

So far, we have for the most part focussed on certain abstract features (everywhere/somewhere and weak/strong stability) of the epistemic utility function; such abstract features have sufficed to state and prove our claims. However, in order better to understand what is going on, we need to consider what a plausible utility function exhibiting some of these features might actually look like. In section 4.1 we take a brief look at one class of plausible (everywhere strongly stable) epistemic utility functions, and consider how functions in that class could encode various epistemic values. In section 4.2 we comment briefly on a particular partly stable utility function, the 'linear utility function', that has appeared in the literature. We answer an objection that the possibility of partly stable utility functions undermines the whole decision-theoretic approach.

4.1   An everywhere strongly stable utility function

Consider the following utility function schema:

General quadratic utility function $: U_{GQ}(s, p) = - \sum_{X \subseteq \mathcal{S}} \lambda_X (\chi_X(s) - p(X))^2,$

where $\chi_X(s)$ is 1 if $s \in X$ and zero otherwise, and the λ_X are constant coefficients. (This is a generalization of the utility function (1) we used in our toy model in section 2.1.) We will briefly discuss how epistemic utility functions of this form do justice to various epistemic norms.

---

[7]Wayne Myrvold has pointed out (personal correspondence) that the fact that we do have several other arguments to the effect that (within the domain of applicability of Bayesian modeling) conditionalization is always rationally required is relevant here. (Some of these arguments were cited in section 1.) The expected-utility approach with no rule against partly stable utility functions, since it permits updating rules other than conditionalization, is in tension with those results. If any of those other arguments is sound (a question we have not addressed), it may also contain the seeds of an explanation *from the perspective of cognitive decision theory* of how and why partly stable utility functions should be disallowed. We have not pursued this line of research.

Stability. For arbitrary choices of the coefficients λ_X, U_{GQ} is everywhere strongly stable. [8] [9]

A concern for truth. In the first instance, U_{GQ} favors placing credence 1 in the true state. *This* is a property of any stable utility function[10], and *a fortiori* of U_{GQ}. More generally (and less rigorously), other things being equal, U_{GQ} favors increasing one's credence in the true state. (We can see the latter by noting that $U_{GQ}(s, p)$ is always an increasing function of $p(s)$, and always a decreasing function of $p(s')$ for $s' \neq s$.)

Discriminating among falsehoods: taking account of verisimilitude. A different stable epistemic utility function (viz. $U(s, p) = \log p(s)$) encodes a sort of *epistemic perfectionism*: according to that utility function, epistemic utility depends *only* on credence in the true state. Such perfectionism may, on occasion, be appropriate. But, often, we will want instead to judge one credence distribution as epistemically better than another even when both assign the same degree of belief to the true state, on the grounds that the first concentrates its remaining credence among (false) states that are *closer to the truth* than does the second. Our sample schema U_{GQ} can take account of the value of verisimilitude, by a judicious choice of the coefficients λ_X: we simply assign high λ_X when X is a set of 'close' states.

Informativeness. Discussions of the epistemic desirability of holding *informative* or *contentful* beliefs are important in *acceptance*-based epistemologies, as opposed to the purely probabilist epistemology under consideration here – given that you're going to accept (say) some true proposition, it is epistemically better to adopt a more informative one, i.e. a stronger one. In the probabilist case, however, the epistemic value of informativeness is already captured by attaching epistemic value to truth and to verisimilitude – an agent will do better in terms of truth-credence and verisimilitude by peaking his probability

---

[8]Proof: in each case, use Lagrange multipliers to extremize the expected utility $\sum_{s \in \mathcal{S}} p(s) U(s, p')$ w.r.t. p', subject to the constraint $\sum_{s \in \mathcal{S}} p'_s = 1$; thence show that extremization occurs at $p' = p$.

[9]Stable utility functions have been discussed in the statistics literature — outside the context of cognitive decision theory, but in a situation with identical mathematical desiderata — where such functions are known as 'proper scoring rules'. See, e.g., Lad [3], Savage [7], for discussion and for lists of other such functions.

[10]Proof: Consider the totally peaked probability distributions, that assign probability 1 to some state $s \in S$ (and, of course, probability zero to all others $s' \neq s$). Relative to such a probability distribution, the expected epistemic utility $EU^p(p')$ of an arbitrary probability distribution p' just is the utility $U(s, p')$ of holding p' when state s obtains. But, if U is stable, then $EU^p(p')$ must be highest when $p' = p$. Thus, $U(s, \cdot)$ must be highest for the probability distribution p that is totally peaked on the state s. That is, if U is to be stable, U must encode the fact that the most epistemically favored probability distribution, when an arbitrary state s obtains, is the probability distribution that assigns credence 1 to the state s. That is, whatever the true state of the world, a maximum degree of belief in the true state is valued higher than any other credence distribution by any stable utility function.

distribution near the true state than he would by having a 'flatter' probability distribution.

4.2   A partly stable utility function/Defense of cognitive decision theory

The following partly stable epistemic utility function has been discussed by Horwich [2]:127-9, Maher [5]:177-9, and Weintraub [12]:

$$\textbf{Linear utility function}: U_L(s, p) = p(s).$$

This utility function has an appealing mathematical simplicity, but, as Maher and Weintraub emphasize, it leads to very odd results. Specifically, the only credence distributions that are self-recommending with respect to this utility function are the totally peaked credence distributions ($p(s) = 1$ for some $s \in S$), and the indifferent distribution ($p(s) = \frac{1}{n}$ for each of n states, $n \leq |\mathcal{S}|$; again we assume that $\mathcal{S}$ has finite cardinality). If one were to hold any other credence distribution, one would maximize EU by shifting to a credence distribution that is assigns credence 1 to some disjunction of states that one currently considers most likely.[11]

What are we to make of this utility function? Maher and Weintraub think that it contains the seeds of an argument by *reductio* against the proposition (CDT), which the approach of present paper has taken as a premise:

CDT   The dynamics of rational credence-distribution updating can be captured by a cognitive decision theory that recommends maximization of expected epistemic utility.

Maher's argument, with which Weintraub agrees, can be reconstructed as follows:

P1   There exist (in logical space) rational agents who hold partly stable utility functions (SSUFs).

P2   If CDT is true, then, for any agent who holds a SSUF, spontaneous shifts of credence distribution are sometimes rational.

C1   If CDT is true, then there exist (in logical space) rational agents who might sometimes perform shifts. (from P1, P2)

P3   Performance of shifts (acceptance of prohibitions) is necessarily irrational: that is, nowhere in logical space are there *rational* agents who might sometimes perform shifts.

---

[11]The following feature of this utility function should also be noted: it is not possible, by conditionalizing on any proposition, to move from a probability distribution that (according to U_L) is self-recommending to one that is not. We have not, in fact, been able to find any partly stable utility functions that do not possess this feature. If it could be shown that none exist, this would obviously strengthen our result (cf. our comments at the end of section 3). We have not been able to find any, but we have no non-existence proof.

C2 CDT is not true. (from C1, P3)

Clearly, if this argument were sound, the central claim of this paper (that Bayesians can justify conditionalization by an appeal to maximization of expected epistemic utility) would be utterly undermined.

The argument is *valid*. Our objection to it is that either P1 or P3 is false, although we are not committed to a view as to which. We choose to insist on the correctness of the cognitive decision-theoretic approach, and argue by dilemma. Either shifts could be rational, or they could not. If they could, P3 is false, and so the argument fails. If they could not, P1 is false, and so the argument fails. In other words, either partly stable utility functions are to be ruled out as irrational, or they and their consequences are acceptable; in neither case is the CDT programme itself impugned.

Maher is aware of the possibility of this response. Since he insists absolutely on the irrationality of shifts (a view with which, as we noted in section 3.3 above, we are not unsympathetic), he gives serious consideration only to the possibility of rejecting P1. Maher's objection to this move is that it is 'completely *ad hoc*', since such a constraint on utility function lacks 'any **prior** plausibility' (ibid, p.179; our emphasis in boldface). Our disagreement with Maher is methodological: we don't see why *prior* plausibility (i.e. prior to thinking through the consequences of adopting a partly stable utility function) should be required for rejection of P1. In any case, the problem with rejecting CDT in response to this argument is that that rejection is no less *ad hoc*: we are left with no convincing explanation of why one should maximize expected utility when choosing whether or not to go for a swim and when choosing whether or not to accept the proposition that humans are descended from apes, but not when choosing which credence distribution to adopt.

Incidentally, we *would* have a paradox for the decision-theoretic approach if we thought both that partly stable utility functions were pathological *and in addition* that no everywhere-stable utility functions existed. But, as we have illustrated by example above, this latter condition does not obtain.

5   Conclusion

By modeling the Bayesian agent's choice of updating policy as a decision problem within a cognitive decision theory, it is possible to provide a justification for conditionalization that appeals directly to the idea that epistemic rationality consists in taking steps that can reasonably be expected to lead to epistemically good outcomes: under independently motivated constraints, conditionalization maximizes expected epistemic utility.

References

[1] I. J. Good. On the principle of total evidence. *British Journal for the Philosophy of Science*, 17(3):319–321, February 1967.

[2] Paul Horwich. *Probability and evidence*. Cambridge University Press, Cambridge, 1982.

[3] Frank Lad. *Operational subjective statistical methods*. John Wiley & Sons, 1996.

[4] Isaac Levi. *Gambling with truth*. Knopf, New York, 1967.

[5] Patrick Maher. *Betting on theories*. Cambridge University Press, 1993.

[6] Graham Oddie. Conditionalization, cogency, and cognitive value. *British Journal for the Philosophy of Science*, 48:533–541, 1997.

[7] Leonard J. Savage. Elicitation of personal probabilities and expectations. *Journal of the Americal Statistical Association*, 66(366):783–801, December 1971.

[8] Leonard J. Savage. *The foundations of statistics*. Dover, New York, 2nd edition, 1972.

[9] Paul Teller. Conditionalization, observation, and change of preference. In William Harper and C. A. Hooker, editors, *Foundations of probability theory, statistical inference, and statistical theories of science*. D. Reidel, Dordrecht, 1976.

[10] Bas C. van Fraassen. *Laws and Symmetry*. Oxford University Press, New York, 1989.

[11] Bas C. Van Fraassen. A new argument for conditionalization. *Topoi*, 18:93–96, 1999.

[12] Ruth Weintraub. A bayesian paradox. *British Journal for the Philosophy of Science*, 52:51–66, 2001.

[13] Paul Williams. Bayesian conditionalisation and the principle of minimum information. *The British Journal for the Philosophy of Science*, 31(2):131–144, June 1980.