Outclassing Indices
and the Problem of Coordinating Plain Belief with Probability

David Makinson, LSE

Abstract

In recent work Hannes Leitgeb proposed a principle, which he calls the Humean thesis, to coordinate the plain beliefs of an ideal (rational, logically omniscient) agent with its degrees of belief, where plain beliefs are represented simply by a set of propositions while degrees of belief are represented by a probability function satisfying the usual Kolmogorov postulates. The Humean thesis states that for such agents the two kinds of credence should jointly satisfy a certain equivalence condition. We discuss a feature of the proposed condition, namely the scarcity of probability functions that can enter into it, which appears to make it difficult to play the desired bridging role.

Keywords: Humean thesis, Humean equivalence, stability theory of belief, outclassing.

1. Stability schemes and the Humean equivalence

Consider a Boolean language which, for simplicity, is assumed throughout this paper to be finite, and for convenience of formulation treat propositions of the language as sets, forming the field of all subsets of some finite set $W = \{w_1,\ldots,w_n\}$ (known variously according to one's upbringing as the underlying set of possible worlds, universe, state-space, or set of states). Propositions, for which we use the variables $A,B,..,X,Y,\ldots$, are thus subsets of W; while belief sets $\mathbf{B}$ are sets of propositions, hence subsets of 2^W. Degrees of belief, on the other hand, are probability functions p: $2^W \rightarrow [0,1]$ on the propositions of the language, satisfying the usual (finitary) Kolmogorov postulates.

Leitgeb's *Humean thesis* proposes that for rational agents, plain belief and degrees of belief should together satisfy a certain coordinating principle, known as the *Humean equivalence*. [1] To explain the latter, it is convenient to see it as an instance of a very general stability scheme that says, roughly, that a proposition is believed iff its probability remains above a given threshold under all suitable conditionalizations.

Stated formally, the stability scheme expresses a relationship between belief sets $\mathbf{B}$ and probability functions p with two further parameters: a non-empty collection $\mathbf{Y} \subseteq 2^W$ and a real number r with $0.5 \leq r < 1$. Intuitively, $\mathbf{Y}$ plays the role of as a set of 'seriously entertained eventualities' on which we may be tempted to conditionalize the probability function p, while r serves as a threshold that those conditionalized probabilities should exceed. The *stability scheme* says:

> For every $X \subseteq W$, the following conditions are equivalent: (i) $X \in \mathbf{B}$, (ii) $p(X|Y) > r$ for all $Y \in \mathbf{Y}$ with $p(Y) > 0$.

Leitgeb reviews several candidates for the parameter $\mathbf{Y} \subseteq 2^W$ in this scheme, settling on one obtained by taking it to be a subset of 2^W that is defined in terms of $\mathbf{B}$ itself, namely $\mathbf{Y} := \{Y \subseteq W\text{: } -Y \notin \mathbf{B}\}$, where $-Y$ is the usual shorthand for $W \setminus Y$, that is, for $\{w \in W\text{: } w \notin Y\}$. Intuitively, $\{Y \subseteq W\text{: } -Y \notin \mathbf{B}\}$ may be understood as the collection of all subsets of W that the belief set $\mathbf{B}$ does not exclude; Leitgeb also writes it as $\mathbf{Poss_B}$ as it may be seen as representing a kind of 'epistemic possibility' given a belief set $\mathbf{B}$.

The *Humean equivalence* HE(r) is the instance of the general stability scheme above that is obtained by choosing its parameter $\mathbf{Y}$ in this manner, that is:

> HE(r): For every $X \subseteq W$, the following conditions are equivalent: (i) $X \in \mathbf{B}$, (ii) $p(X|Y) > r$ for all $Y \subseteq W$ with $p(Y) > 0$ and $-Y \notin \mathbf{B}$.

Here r remains as parameter (along with variables $\mathbf{B}$,p) and it is required that $0.5 \leq r < 1$. The Humean equivalence HE(r) thus expresses a relationship between belief sets $\mathbf{B}$ and probability functions p, modulo a parameter r.

2. Some immediate consequences of the Humean equivalence

As noted by Leitgeb (chapter 2) it is easily verified that for any tuple ($\mathbf{B}$,p,$\mathbf{Y}$,r) satisfying the general stability scheme above we have:

$$\text{(i) } S \in \mathbf{B}, \text{ (ii) if } X \in \mathbf{B} \text{ and } X \subseteq Y \text{ then } Y \in \mathbf{B}, \text{ (iii) if } X \in \mathbf{B} \text{ then } -X \notin \mathbf{B}.$$

From (i) and (iii) it further follows that (iv) $\varnothing \notin \mathbf{B}$. In logical parlance $\mathbf{B}$: (i) contains every logically true proposition; (ii) is closed under classical consequence of its individual members, (iii) contains no contradictory pairs of propositions, (d) contains no logically false proposition.

If more specifically ($\mathbf{B}$,p,r) satisfies the Humean equivalence HE(r), then the set $\mathbf{B}$ of beliefs is closed under meet:

$$\text{(v) if both } X,Y \in \mathbf{B} \text{ then } X \cap Y \in \mathbf{B}.$$

In algebraic parlance, given HE(r) the set $\mathbf{B}$ of beliefs becomes a non-empty proper filter over 2^W; in logical terminology it is a consistent non-empty set of propositions closed under classical consequence;

Taken together, this is an attractive collection of features. The Humean equivalence not only coordinates degrees of belief (measured by probability functions) with belief simpliciter (registered by sets of propositions), via a threshold. It also retains a bundle of properties for plain belief which, on the one hand, epistemologists have traditionally seen as desirable but which, on the other hand have also been treated with suspicion in the light of the notorious lottery paradox, and sometimes taken to be render coordination with degrees of belief unattainable.

As we are working in a finite context, the belief set $\mathbf{B}$ will always be a finite collection of propositions and so its meet $B_0 = \cap\mathbf{B}$ will be a proposition – indeed, given HE(r) and the finite context in which we are working, it will be in $\mathbf{B}$ and thus the strongest of all the believed propositions. Moreover, $\mathbf{B}$ may conversely be expressed in terms of B_0, for it follows that $\mathbf{B} = \{X \subseteq W: B_0 \subseteq X\}$.

The Humean thesis thus says: for rational agents, plain belief and degrees of belief should together satisfy the following formulation of the Humean equivalence, with $B_0 = \cap\mathbf{B}$:

> HE(r): For every $X \subseteq W$, the following conditions are equivalent: (i) $B_0 \subseteq X$, (ii) $p(X|Y) > r$ for all $Y \subseteq W$ with $p(Y) > 0$ and $B_0 \nsubseteq -Y$.

This presentation will be particularly useful for our work in what follows.

3. Three questions: definability, fragility, scarcity

The Humean thesis immediately prompts three questions, each with a formal side and a philosophical aspect. One is: Can the Humean equivalence be used as a normative *definition* or epistemological *reduction* of plain belief in terms of a probability function and a threshold? A second is: *how robust/fragile* is the equivalence? In particular, is its satisfaction independent of aspects of the probability space that should arguably be open to variation without affecting qualitative beliefs? Third: *how plentiful/scarce* are the belief sets and thresholds that can satisfy the equivalence given an arbitrary probability function? Can we always find them where we need them? Our focus in this paper is on the third question, but we begin by making brief remarks on the first two.

3.1. Definability.

For definability, a suspicious feature of the Humean equivalence is that **B** (or B_0, according to formulation) occurs on both sides of the equivalence condition HE(r). Moreover, it is easy to show that for some values of p,r there may be more than one **B** (indeed, more than one $B_0 = \cap\mathbf{B}$) satisfying the equivalence, so that the probability function and threshold do not, in general, uniquely determine a belief set satisfying the equivalence.

However, those facts by themselves do not settle the question of definability. For it can also be shown that for any values of p,r there is a *unique strongest* belief set **B** (a unique strongest meet $B_0 = \cap\mathbf{B}$) that satisfies the equivalence.

Thus, from a purely formal point of view, and despite initial appearances, it would be possible to view the Humean equivalence as providing a definition of plain belief from degrees of belief accompanied by a threshold. The questions of whether such a definition would be philosophically adequate and whether it can be regarded as a reduction of one concept to the others, are separate issues into which we do not enter. Leitgeb, after an initial inclination in his 2013 paper to see them positively, has taken a firm negative stance in his subsequent publications.

3.2. Robustness/fragility

As Leitgeb recognizes, satisfaction of the Humean equivalence is not robust under partitions of the state-space W. In other words, if we take a triple **B**,p,r satisfying it, and form ($\mathbf{B}'$,p',r) by partitioning W and lifting p to a probability function on the partition, the latter may fail to satisfy the equivalence.

An example of this phenomenon arises in Kyburg's well-known lottery. Consider a fair lottery of n tickets (with n large, the rules requiring that exactly one ticket wins, and equi-probability between tickets). Take any threshold r with $0.5 < r \leq 1$, and choose a specific ticket k with $1 \leq k \leq n$. What does the Humean equivalence tell us about the proposition $-X_k$ = 'ticket k will lose'? The answer depends on the state-space chosen. On the one hand, if we take W to consist of the n propositions X_i = 'ticket i will win', then the Humean equivalence tells us *not* to believe $-X_k$ (nor, of course, its singleton complement); the only belief authorized is that exactly one ticket will win, that is, all but one unspecified ticket will lose. On the other hand, if we partition W into, say, just the two cells $\{X_k\}$ and $\{-X_k\}$, i.e. 'ticket k will win' and 'some other ticket will win' then the Humean equivalence (mod r) says that we *should* believe $-X_k$, provided $r < (n-1)/n$.

Some see this dependence on partition as a major shortcoming of the Humean thesis, while Leitgeb argues that it is just what we should expect, both in the case of the lottery example and more generally. He suggests that such dependence is an almost inevitable feature of any epistemic rule expressed in probabilistic terms, and not in itself a reason for abandoning such a rule.

3.3. Plentitude/scarcity

We do not discuss either of the above issues here, but focus on the third questions mentioned: *how plentiful or scarce* are Humean tuples? We draw attention to a formal property that they have, which seems to indicate that they are rather too scarce for the role that the Humean thesis asks them to play. The formal property itself is recognized and discussed by Leitgeb (e.g. 2015b chapter 2 appendix B, chapter 3 section 3.2, and chapter 4 section 4.4.5) but without, we suggest, fully appreciating its inconvenience.

4. The outclassing property

Consider any (finite) state space $W = \{s_1,\ldots,s_n\}$, probability function p: $2^W \rightarrow [0,1]$ and threshold r with $0.5 \leq r < 1$. We say that a subset $B \subseteq W$ has the *basic outclassing property* wrt p iff the following condition is satisfied.

> *Out*: B is non-empty and for all $w \in B$, $p(w) > p(-B)$. That is, $p(w) > \Sigma\{p(w'): w' \in W$ but $w' \notin B\}$.

Here we write $p(w)$ for $p(\{w\})$ to simplify notation and, to cover the limiting case that $-B = W \backslash B$ is empty, we understand $\Sigma\varnothing$ to be 0. The rationale for the name is that each state in B has probability large enough to 'outclass' the whole of $-B$, in the sense that it is not only more probable than each of the states outside B considered individually, but is *more so than their sum*.

This property is a particular instance of a more general one that also takes r into account. We say that a subset $B \subseteq W$ has the *outclassing property* (mod r) wrt p iff:

> *Out*(r): B is non-empty and for all $w \in B$, $p(w) > p(-B) \cdot r/(1-r)$,

where $0.5 \leq r < 1$. In other words, outclassing (mod r) requires that each state in B is more probable than the sum of all the probabilities of states outside B, multiplied by $r/(1-r)$.

In the limiting case that $r = 0.5$ we have $r/(1-r) = 1$ so that outclassing property (mod 0.5) reduces to the basic version. As r rises from 0.5 towards 1, the value of $r/1-r$ also rises without finite limit and the property becomes more and more demanding. For example, when $r = 0.9$, a non-empty set B will have the outclassing property mod r iff for all $w \in B$, $p(w) > 9 \cdot p(-B)$.,

It may be asked whether there is any intuitive reason for choosing the multiplier $r/(1-r)$ in this definition, rather than some other monotone function rising from 1 at $r = 0.5$. The rationale is a formal one: as noticed by Leitgeb, *Out*(r) so formulated is equivalent to the Humean equivalence HE(r). That is, for any probability function p: $2^W \rightarrow [0,1]$, belief set $\mathbf{B}$, and r with $0.5 \leq r < 1$, HE(r) holds between them iff *Out*(r) holds between them with B chosen as $B_0 = \cap\mathbf{B}$. A concise verification is given in the Appendix.

For any choice of r with $0.5 \leq r < 1$, the outclassing property for B (mod r) immediately implies that when $w \in B$ and $w' \in W$ with $p(w) \leq p(w')$, then $w' \in B$. Thus, if we write $W = \{w_1,\ldots,w_n\}$ in (weakly) increasing order of probabilities then B is a non-empty right segment of the list, that is, $B = \{w_k,\ldots,w_n\}$ for some $k \leq n$.

This facilitates comparison with a property for probability functions that was considered by Snow 1994, 1996, also by Benferhat, Dubois & Prade 1997. They call a probability function p on a finite state-space W 'atomic bounded' (Snow) or more graphically 'big-stepped' (Benferhat et al) iff we may write the elements of W as $w_1,\ldots,w_n$ with $p(w_i) > p(w_1)+\ldots+p(w_{i-1})$ for all $i \leq n$. Thus big-stepping concerns W and p alone, and we may speak of it simply as a property of the probability function p. We could also generalize the notion mod r where $0.5 \leq r < 1$, by requiring $p(w_i) > (p(w_1)+\ldots+p(w_{i-1})) \cdot r/(1-r)$ for all $i \leq n$.

Clearly, big-stepping (mod r) asks more of p than does outclassing (mod r). In particular, it requires that each $p(w_i) \neq p(w_{i+1})$ so that p is injective on W. It also requires that the inequalities $p(w_i) > (p(w_1)+\ldots+p(w_{i-1}))\cdot r/(1-r)$ hold for all $w_i \in W$, whereas outclassing does so only when w_i is the first in the given ascending enumeration of elements of B.

5. Outclassing indices

So much is essentially concise exposition of material of Leitgeb 2015b and preceding papers. To discuss the question of scarcity, it is convenient to introduce the notion of the *outclassing index* of a probability function.

Consider any probability function p: $W \to [0,1]$ with $\#(W) = n$, and choose any r with $0.5 \leq r < 1$; hold them both fixed in what follows. We say that p has outclassing index u/n (mod r) iff there are exactly u subsets $B \subseteq W$ with the outclassing property (wrt p mod r). When no confusion can arise, we speak briefly of the *index* u/n, leaving p and r understood. Clearly, the lower the index, the scarcer are belief sets that can be coordinated with p via the Humean equivalence; the higher the index, the more flexible is p for coordination with belief sets of different levels of boldness.

Since $B = \{w \in W: p(w) > 0\}$ always has the outclassing property, every probability function on W has index at least $1/n$, irrespective of the choice of r in the interval $0.5 \leq r < 1$; this set $B \subseteq W$ is termed *trivial* in the sense that $p(B) = 1$. The index can never be more than $n/n = 1$ since, as we have noted, every outclassing set must be a non-empty upper segment of a weakly increasing ordering of W by probability, and elements of W with the same probability cannot be separated by an outclassing set, so that permutations of equiprobable elements in the ordering give nothing new. Thus for any index u/n we have $1 \leq u \leq n$.

The top index 1 is achieved by suitably constructed big-stepped probability functions. An example of index only $1/n$ for a large $n = \#(W)$ is given by the probability function in the Kyburg lottery (section 3), taking W to be the set of n propositions X_i = 'ticket i will win', and the only belief countenanced by the Humean thesis (no matter how we choose r in the interval $0.5 \leq r < 1$) is that exactly one ticket will win, that is, all but one will lose.

It might be thought that this outcome is a peculiarity of the uniform function and others close to it. But quite far-from-uniform probability functions can also have index $1/n$, as in the following example where $n = 5$. Put $W = \{w_1,\ldots,w_5\}$ with $p(w_1) = p(w_2) = 0.1$, $p(w_3) = 0.15$, $p(w_4) = 0.3$, $p(w_5) = 0.35$ (thus summing to unity). This is quite lumpy (the probability of w_5 is more than thrice that of w_1), although it is not quite big-stepped (the first two elements of W have the same probability). But again there are no proper subsets of W with the outclassing property, as is immediate. Examples of index $1/n$ for larger values of $n = \#(W)$ are easily constructed.

Leitgeb 2013 (cf 2015 chapter 3 section 3.2) uses measure theory to observe that for 'almost all' probability functions p on a finite state-space W, there is at least one non-trivial $B \subseteq W$ with the basic outclassing property under p; in our terminology, for $r = 0.5$ almost all probability functions are of index at least $2/n$. This can also be given a short and transparent proof that merely skims the surface of measure theory, as follows.

Consider any state-space $W = \{w_1,\ldots,w_n\}$ with $\#(W) \geq 2$, its elements listed in weakly ascending order, and let p: $W \to [0,1]$ be a probability function on W. Let k be the least integer with $p(w_k) > 0$. We claim that $B = \{w_{k+1},\ldots,w_n\}$ will non-trivially have the basic outclassing property wrt p in all cases except two: where $k = n$ and when $p(w_k) = p(w_{k+1})$. For since $p(w_k)$

> 0 we have $p(B) < 1$, so when $k < n$ then $B \neq \varnothing$. And when $p(w_k) < p(w_{k+1})$ then for all $w \in B$ we have $p(w) \geq p(w_{k+1}) > p(w_k) = p(-B)$, completing the verification.

Note, however, that this verification does not go through for thresholds $r > 0.5$. For then the second exception becomes $p(w_k)\cdot r/(1-r) \geq p(w_{k+1})$ rather than $p(w_k) = p(w_{k+1})$, and so can no longer be described as 'almost never' happening in a measure-theoretic sense.

Moreover, even for $r = 0.5$, we cannot strengthen the argument to get an index of at least $3/n$. If one tries to run it for an upper segment of W other than $\{w_{k+1},\ldots,w_n\}$ where k is the least integer with $p(w_k) > 0$, the verification snags: again, the exceptions will no longer be at points but over intervals. For example, when $B = \{w_{k+2},\ldots,w_n\}$ it lacks the basic outclassing property whenever $p(w_k)+p(w_{k+1}) \geq p(w_{k+2})$.

Leitgeb recognizes the second limitation (in a footnote of appendix B for chapter 2 of 2015b), but does not discuss its significance, nor mention explicitly the first limitation. Yet they would both appear to be bad news for the Humean thesis. They tells us that a great many probability functions have index only $1/n$ (if r is taken greater than 0.5) or $2/n$ (if one sets $r = 0.5$), that is, have zero or only one non-trivial belief set satisfying the Humean equivalence. For any such function there is thus little, if any, flexibility in the boldness of the belief set that we may choose in accordance with the Humean proposal.

In epistemic practice we would presumably like to have a criterion for belief that allows us more elbow-room. We would like to be able to modulate somewhat how much we believe, depending on the confidence level that we feel is suitable for the context we are in. The choice of a confidence level depends not only on our probability function p (if we have one) but also on such contextual matters as the benefits of getting a right belief, the costs of getting a wrong one or of suspending belief, the running costs of working things out, and the possibility of timely error correction. It seems reasonable for agents to form at least a general idea of the standards that they want to live up to in a given context, and bring their qualitative beliefs into rough agreement with that standard without changing their probability function. There is little room for this when the outclassing index is $2/n$ or less.

6. Trying to cope with outclassing index ≤2/*n*

How can the stability theory of belief cope with situations in which p has a low outclassing index? Apart from just biting the bullet and living with epistemic inflexibility, two main responses suggest themselves. They both massage the probability configuration, but in different ways. One *merges elements* of the state space, the other *infinitesimally bends* the probability function. We explain and discuss them in turn.

6.1. Merging elements of the state space

This kind of massage capitalizes on the already-noticed sensitivity of the outclassing property to partitioning. The idea is to merge the elements of a suitable set $X \subseteq W$ that does not already have the outclassing property into a single item whose singleton does have it, without disturbing any of the positive instances of that property.

For example, in the lottery with W consisting of the n propositions X_i = 'ticket i will win', for any $k \leq$ we can merge all X_j for $j \neq k$ into a single item, creating the two-cell partition under which (section 3.1 above) the Humean equivalence (mod r) tells us to believe that ticket k will lose, provided $r < (n-1)/n$, raising the index (mod r) to $2/2 = 1$. Again, in the lumpy example of section 5, we can put $W' = \{w_1,w'\}$ where w' is $\{w_2,\ldots,w_5\}$, to ensure that $\{w'\} \subseteq W$ is a basic outclassing set in the quotient structure, likewise raising the index to $2/2 = 1$.

This is a very flexible procedure indeed – indeed too flexible. Unfortunately, we have no formal criterion, or even a semi-formal one, to tell us when a partition of a given state-space is reasonable or on the contrary discards so much relevant information that it becomes inappropriate. In real life, the answer to that question will depend a great deal on the particularities of the problem one is trying to model. Unless constrained by criteria for what partitions are reasonable in a given context, the proposal appears quite unprincipled.

6.2. Infinitesimally bending the probability function

Is there a more controlled way of raising the outclassing index of a probability function? Leitgeb proposes an interesting one, closely related to the 'almost all' theorem discussed in section 5 and, like it, formulated using Lebesgue measure for sets of probability functions. Here too one can express the idea more simply, in terms of arbitrarily small deviations in a given probability function.

The procedure may be expressed as an algorithm to ensure a basic outclassing index of at least $2/n$ on any state-space S with at least two elements. Writing the elements of W as $w_1,\ldots,w_n$ in weakly ascending order, we take the first element w_k of non-zero probability. Two cases present themselves.

Case 1: w_k is the only element of non-zero probability. Then that element must be w_n, with probability 1. Since by supposition W has at least two elements, w_{n-1} exists and has probability zero, so we can bend p by giving w_{n-1} an arbitrarily small positive positive probability (so long as it is less than half that of w_n) and reducing that of w_n by the same amount, thus ensuring the outclassing property (modulo any given $0.5 \leq r < 1$) for both $\{w_n\}$ and $\{w_{n-1},w_n\}$ under the massaged probability function and bringing the index from $1/n$ to $2/n$.

Case 2: There is another element of non-zero probability. Then w_{k+1} must exist with $p(w_{k+1}) \geq p(w_k)$. Two subcases present themselves.

Subcase 2.1: $p(w_k) < p(w_{k+1})$. Then we already have a non-trivial B with the basic outclassing property, namely $B = \{w_{k+1},\ldots,w_n\}$, and the basic index of p is already at least $2/n$.

Subcase 2.2: $p(w_k) = p(w_{k+1})$. Then we bend p by making an arbitrarily small reduction in the probability of w_k with an equal increase in w_n, so that $p'(w_k) < p'(w_{k+1})$ under the new p'. Then the set $B = \{w_{k+1},\ldots,w_n\}$ has the basic outclassing property under the massaged function p' without taking that property away from any sets that already had it; and hence in this subcase the index of p' (for $r = 0.5$) goes up by (at least) $1/n$.

In the lumpy example of section 5 we are in subcase 2.2, $B = \{w_2,\ldots,w_5\}$, and p' reduces the probability of w_1 by a hair while raising that of w_5 accordingly. In the lottery example, $B = \{w_2,\ldots,w_n\}$, and p' shaves w_1 while raising w_n, bringing $p'(B)$ a tad closer to 1 and ensuring that B has the basic outclassing property.

Does the infinitesimal bending procedure have any shortcomings? The main one appears to be that it does not take us very far. Indeed, there are two quite visible road-blocks – essentially the same as those blocking extension of the 'almost all' result (section 5) to thresholds $r > 0.5$ or to index $3/n$ for $r = 0.5$.

- For thresholds $r > 0.5$ the infinitesimal bending procedure does not take us anywhere. For then, in addition to the subcases 2.1 and 2.2 we have to face a further subcase 2.3, that $p(w_{k+1})$ lies somewhere in the strict interval from $p(w_k)$ to the larger (for high r, much larger) $p(w_k)\cdot r/(1-r)$. We cannot get out of that case by an arbitrarily small deviation in p.

- In the limiting case that $r = 0.5$, where the infinitesimal bending does succeed in getting us to index $2/n$, it does not in general take us to $3/n$ or beyond, even when W is large. This can be seen by attempting to re-run the argument used for subcase 2.2 for a third element w_{k+2} of S with non-zero probability under p, considering the segment $B = \{w_{k+2},\ldots,w_n\}$. We may need more than an arbitrarily small deviation from p to ensure that the probability of w_{k+2} is greater than *the sum of* the probabilities of w_k and w_{k+1}. We were able to do this for w_{k+1} because the only item with non-zero probability preceding it was w_k so there was no proper summing to be done. The infinitesimal bending trick will work for w_{k+2} only when $p(w_{k+2}) = p(w_k)+p(w_{k+1})$.

Infinitesimal bending thus does little to increase epistemic flexibility, and usually nothing when $r > 0.5$. To illustrate, consider again the lumpy example of section 5, where the index was 1/5, and imagine that we are quite happy with a threshold of around 0.6. If we bend p infinitesimally into p' in order to bring the index up to 2/5, then the new belief set B that we get has probability $p'(B) > 0.9$, which is rather too timid for our epistemic inclinations. Perhaps it would be better to take $B = \{w_3,w_4,w_5\}$ with $p(B) = 0.80$ or even boldly put $B = \{s_4,s_5\}$ with $p(B) = 0.65$. In either case, our beliefs are all the $X \subseteq W$ with $B \subseteq X$, and the belief set is closed under conjunction. Likewise in the lottery example with say $n = 10^6$ lottery tickets, we may be perfectly satisfied with a level of 90% or 99% rather than $(10^6-1)/10^6$ and seek a belief set that corresponds, at least roughly, to that level. Such epistemic options are not obtainable by the infinitesimal bending algorithm.

7. General perspectives

We have contended that the epistemological coordinating principle known as the Humean thesis has difficulty coping with probability functions with a low outclassing index. Two ideas for massaging such functions within the spirit of the theory are: merging elements of the state space, and infinitesimally bending the probability function. But the former appears to lack guidance to the point of allowing unprincipled repartitioning, and the latter has very limited effect.

However, this is not to say that the Humean thesis faces more difficulties than rivals in the literature on the coordination of plain belief with degrees of belief. It may still be as good as any of them. There may be no way of fitting together the qualitative and quantitative conceptual frameworks without at least *some* tell-tale lines at the joins.

If that is the case, comparing the merits of rival coordination principles becomes a matter of weighing their respective costs and benefits, rather than of finding a ‘correct’ theory. Moreover, the weights going into the cost-benefit balance may vary with context, so that agents might be entitled to use one or another coordination rule according to which appears more suitable and rewarding for their particular situations. Figuratively speaking, epistemic coordination may be more of like making a patchwork with fragile seams than weaving a single cloth.

Notes

1. The Humean thesis evolves through a number of papers Leitgeb 2013, 2014a,b, 2015a, all of which are integrated into the book 2015b. The material discussed in the present paper may be found in chapters 1-4 of that book, their appendices, and their end-notes.

Appendix

To make this paper self-contained, we give a short proof that HE(r) and *Out*(r) are equivalent. Consider any finite state space $W = \{w_1,\ldots,w_n\}$ with probability function p: $2^W \rightarrow [0,1]$ belief set $\mathbf{B} \subseteq 2^W$, and r with $0.5 \leq r < 1$. We claim that HE(r) holds for them iff *Out*(r) holds for p, B, r with B chosen as $B_0 = \cap\mathbf{B}$.

For left to right, suppose that HE(r) holds of p, $\mathbf{B}$, r. We need to show that $B_0 = \cap\mathbf{B}$ is non-empty and for all $w \in B_0$, $p(w) > p(-B_0)\cdot r/(1-r)$. We know (section 2) that $B_0 \in \mathbf{B}$ while $\varnothing \notin \mathbf{B}$, so B_0 is non-empty. Let $w \in B_0$; we need to show that $p(w) > p(-B)\cdot r/(1-r)$. Now $p(w) > 0$: otherwise we would have $p(B_0) = p(B_0\backslash\{w\})$ so that $B_0\backslash\{w\} \in \mathbf{B}$ contrary to $B_0 = \cap\mathbf{B}$. Put $Y = -B_0\cup\{w\}$. Then $p(Y) \geq p(w) > 0$. Also, $B_0 \not\subseteq -Y$, since $w \in B_0$ while $w \in Y$. Hence since $B_0 \in \mathbf{B}$, HE(r) tells us that $p(B_0|Y) > r$. But $p(B_0|Y) = p(Y\cap B_0)/p(Y) = p(w)/p(Y) = p(w)/p(-B_0\cup\{w\}) = p(w)/(p(-B_0)+p(w))$, so $p(w)/(p(-B_0)+p(w)) > r$ and thus:

$$p(w) > r\cdot(p(-B_0)+p(w)) = r\cdot p(-B_0)+r\cdot p(w)).$$

Suppose for reductio that $p(w) \leq p(-B)\cdot r/1-r$. Then $p(-B) \geq p(w)\cdot(1-r)/r$ so, substituting for $p(-B)$ in the display above and simplifying, we have $p(w) > (r\cdot p(w)\cdot(1-r)/r)+(r\cdot p(w)) = p(w)(r+(1-r)) = p(w)$ so that $p(w) > p(w)$: contradiction.

For right to left, suppose that *Out*(r) holds of p, B_0, r where $B_0 = \cap\mathbf{B}$, and let $X \subseteq W$. It suffices to show (see end of section 2) that the following conditions are equivalent: (i) $B_0 \subseteq X$, (ii) $p(X|Y) > r$ for all $Y \subseteq W$ with $p(Y) > 0$ and $B_0 \not\subseteq -Y$.

First, suppose (i) and let $Y \subseteq W$ with $p(Y) > 0$ and $B \not\subseteq -Y$; we need to show that $p(X\cap Y)/p(Y) > r$, that is, that $p(Y) < p(X\cap Y)/r$. Since $B_0 \subseteq X$ it suffices to show $p(Y) < p(B_0\cap Y)/r$. Since $B_0 \not\subseteq -Y$ there is a $w \in B_0\cap Y$, so by *Out*(r) we have $p(-B_0\cap Y) \leq p(-B_0) < p(w)\cdot(1-r)/r \leq p(B_0\cap Y)\cdot(1-r)/r$. But $p(Y) = p(B_0\cap Y)+p(-B_0\cap Y)$ so $p(Y) \leq p(B_0\cap Y)+p(B_0\cap Y)\cdot(1-r)/r = p(B_0\cap Y)/r$, and we are done.

Conversely, suppose (i) fails, so $B_0 \not\subseteq X$ so there an $w \in B$, $w \notin X$. Put $Y = \{w\}$. By *Out*(r) since $w \in B_0$ we have $p(Y) > 0$. By construction, $B_0 \not\subseteq -Y$, but $p(X|Y) = p(X\cap Y)/p(Y) = 0 < r$ so that (ii) fails, completing the verification.

References

Benferhat, Salem, Didier Dubois & Henri Prade 1997. 'Possibilistic and standard probabilistic semantics of conditional knowledge' *Journal of Logic and Computation* 9: 873-895.

Leitgeb, Hannes 2013. 'Reducing belief simpliciter to degrees of belief' *Annals of Pure and Applied Logic* 164: 1138-1389.

Leitgeb, Hannes 2014a. 'The stability theory of belief' *Philosophical Review* 124: 131-171.

Leitgeb, Hannes 2014b. 'A lottery paradox for counterfactuals without aggregation' *Philosophy and Phenomenological Research* 89: 605-636.

Leitgeb, Hannes 2015a. 'The Humean thesis on belief' forthcoming in the *Proceedings of the Aristotelian Society*.

Leitgeb, Hannes 2015b. *The Stability of Belief: an Essay on Stability and Coherence* (to appear, Oxford University Press).

Snow, P. 1994. 'The emergence of ordered belief from initial ignorance' pages 281–286 of *Proceedings of the 12th National Conference on Artificial Intelligence* (AAAI'94), Seattle, WA, July 31-Aug. 4, pp. 281–286, 1994.

Snow, P. 1996. ‘Warranting non-demonstrative reasoning upon the common ground of contending uncertainty approaches’ pages 1381–1386 of *Proceedings of the 6th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems* (IPMU-96), Vol. II, Granada (Spain).

Acknowledgements

Thanks to Hannes Leitgeb, Wlodek Rabinowicz, and participants in the LSE PH500 reading group Michaemas term 2014-15 (Jonathan Burton, Alexandru Marcoci, Samuel Martinet, Silvia Milano, James Nguyen, Aron Vallinder) for stimulating discussions and penetrating criticisms.

Last revised 2015-02-18