Supposition and desire in a non-classical setting

J Robert G Williams

(October 18, 2010)

Revising semantics and logic has consequences for the theory of mind. Standard formal treatments of rational belief and desire make classical assumptions. If we are to challenge the presuppositions, we need to indicate what is kind of theory is going to take their place. Consider probability theory interpreted as an account of ideal partial belief. But if some propositions are neither true nor false, or are half true, or whatever—then it's far from clear that our degrees of belief in it and its negation should sum to 1, as classical probability theory requires (cf. Field, 2001). There are extant proposals in the literature for generalizing (categorical) probability theory to a non-classical setting, and we will use these below. But subjective probabilities themselves stand in functional relations to other mental states, and we need to trace the knock-on consequences of revisionism for the relation between belief and other attitudes.

This note looks at two crucial steps to an overall non-classical theory of mind. One is the move from categorical to conditional beliefs—needed to link our theory of belief to an account of suppositional belief, belief-revision and to feed into decision theory. We outline a particular proposed account of conditional probability, and give a (synchronic) dutch book argument for it. The second is to link the account of categorical and conditional belief to desire (and hence to action). Decision theory can be viewed as a theory of the rational interrelation between belief and desire. We show how generalized desirability can be characterized in a way parallel to the characterization of generalized probabilities, and that the central equation of Jeffrey's decision theory (suitably interpreted for the general setting) then holds.

1 Generalized probabilities

We suppose there are a finite set of worlds W and of propositions P. A proposition takes a truth value (in the range $[1,0]$) at each world; we write $|q|_w$ for the truth value of the proposition q at w. We will think of the propositions as structured out of 'atoms', via conjunction, arrows, and negations on the model of propositional languages. Various different proposals can be entertained for the patterns that hold between the truth values of molecular propositions and their constituents, but (except where noted) we shall not take a stance on this here.

Let a base-credence c be an assignment of a single number in $[1,0]$ to each world, such that $\sum_{w\in W} c(w) = 1$. A belief state is an assignment of a single number in $[1,0]$ to each member of P. We call a belief state b a (generalized) probability if there is some base-credence c such that following holds:

$$b(q) = \sum_{w \in W} c(w)|q|_w$$

Note that if the truth-value distribution is classical, then generalized probabilities are just classical probabilities; more generally one may think of them as 'expectations of truth value' of the proposition, relative to the base-credence over worlds.

Why define 'generalized probability' this way? One reason is that common rationalizations for subjective probability carry over. Paris (2001) shows that Dutch book arguments can be formulated against belief states that are not generalized probabilities in a non-classical setting, just as they can be against belief states that are not classical probabilities in a standard setting. Likewise, Joyce's accuracy-domination theorem (Joyce, 1998) can be generalized to motivate interest in the same class of functions (see Williams, manuscript).

Relative to a certain type of non-classical truth-value distribution, local constraints that are satisfied by any generalized probability can be formulated, in ways that are closely related to standard probability axioms (specifying a non-classical 'logic' in terms of the class of truth-value assignments in question is vital to this exercise). The converse—i.e. that every function satisfying the relevant local constraints is a generalized probability—-is typically harder. But in various cases, it has been done.[1]

Think of these generalized probabilities as our candidates for what (ideally) rational belief should be like. Notice the following: what we have so far tells us only about belief; and not, for example, about desire and intentions. Standard probability theory is often embedded in a broader theory that covers desire, belief and their interactions with ration action or intention: decision theory. But there is a prior issue. What we have above is an static characterization of rational belief—we have not yet identified an analogue of conditional probability within the generalized framework. Such quantities are vital to broader theoretical deployments of probabilities both in modelling the idea of belief *under a certain supposition* and being the basis for many accounts of rational update of belief. Conditional probabilities are at the heart of at least one leading formulation of decision theory—that of Jeffrey (1965). So before we look at generalizing decision theory, we look at how we might generalize conditional probability.

2 Generalizing conditional probability

Thought of dynamically, the basic idea of conditional probability is to start with a probability b, and see what probability b' should result by a process that can be sensibly thought of as 'updating on the information that p'. Thus the desiderata would be to describe a mapping f, such that $f(b,p) = b'$ for arbitrary b, p. To do this one would seek to describe the values of b' for each proposition q, so an equivalent task would be to describe a three-place function $g(b,p,q) = f(b,p)(q)$. It's standard to write $b(p|q)$ for $g(b,p,q)$ (the "g" is in effect represented by the line). The cost of doing so is that we use b ambiguously—once as picking out a belief state (a unary function from propositions to numbers) and once, apparently, as a

[1] See in particular Paris (op cit).

summary of a dyadic function from pairs of propositions to numbers. That said, I'll use the familiar line-notation henceforth.

Bearing in mind the basic conception of a conditional probability as moving us around the space of rational belief states (generalized probabilities) via 'updating on a proposition', the basic constraint on giving a characterization should be that the belief state defined by $b_q(x) = b(x|q)$ should always live within the space of generalized probabilities. I assume that a similar constraint will be in place for all deployments of conditional probability (for example, in decision theory or as a synchronic state of 'supposing that q'). This will give us crucial traction on the notion.

In the classical case, conditional probabilities are taken to satisfy the 'ratio formula' (I'll leave qualifications about not-dividing-by-zero implicit throughout):

$$b(p|q) := b(p \wedge q)/b(q)$$

It'd be nice if we could carry this characterization over, and show it had nice properties. Unfortunately, doing so (at least with certain familiar treatments of nonclassical conjunction) will violate the basic constraint just mentioned. We'll see this below.

Rather than try to generalize the ratio formula, we might think about various constraints on the behaviour of conditional probability, and see what would satisfy them in the general setting. Perhaps the most basic thought to have is that, whatever else happens, $b(p|p) = 1$. After all, what would 'updating on p' be, if not, first, treating p as something that has been established?

But notice how strong a constraint this is. What it tells us is the updated probability, b_p, assigns p a value of 1. Think this through in the case of Lukasiewicz truth value assignments. At various worlds, p gets intermediate truth values. One's belief in p is a mixture of p's truth value at situations, weighted by one's credences in the respective situations. But if one gives non-zero weight to a situation where p has a truth-value less than 1, then it's clear that one's belief in p will be less than 1. Moral; if b_p meets the constraint given above, then it must 'wipe out' any weight that b originally assigned to any situation in which p has truth value other than 1 (even situations where p has truth value 0.99999, for example). This isn't crazy as a way to think of 'updating on p'; but it's very strong. The natural candidate for such a relation would be to construe 'updating on p' as 'updating on p having truth value 1'. Which is a possible way to go, but surely not the only one.

3 Generalized conditional probabilities characterized

There is a natural modification of the above, however, which also generalizes the classical case. Here's one description of classical conditionalization on p: one first sets the credence in all $\neg p$ worlds to zero; leaving the credence in p-worlds untouched. This, however, won't give you something that's genuinely a probability (for example, the 'base credences' no longer sum to 1). So one renormalizes the credences to ensure we do have a probability. The above recipe for conditionalizion follows this recipe, with the adjustment that at the initial stage one wipes out all credence in any world where the truth value of p falls below 1.

Here's a different generalization, that fails to meet the $b(p|p) = 1$ constraint, but is interesting in other ways. Think of that first step in the classical case: we wipe out credence in worlds where the proposition is false (truth value 0) and leave alone credence in worlds where the proposition is true (truth value 1). Another way to put this is that the updated credence in w (prior to renormalization) is $c_p(w) := c(w)|p|_w$—i.e. the result of multiplying the prior credence in w by the truth value of p at that possibility. Whereas the previous proposal wiped out credence in worlds where p is less than fully true, this proposal scales the credence in proportion to how true p is at that possibility. Renormalizing (it turns out) is achieved by dividing by the prior credence in p (to see this, note that the result of the process is to give a 'base credence' over worlds which may add to less than 1. The sum total is given by $\sum_{w\in W} c_p(w) = \sum_{w\in W} c(w)|p|_w$. But by construction this is exactly $b(p)$. Hence dividing $b(p)$ will renormalize the base credence, making it sum to 1, after the procedure described above).

What does the above give us? Let's first consider point-propositions: propositions that are truth value 1 at w and 0 everywhere else (we'll equivocate by using w both for the world and the corresponding point-proposition). We get the following:

$$b(w|p) = b_p(w) = \sum_{u\in W} c_p(u)|w|_u/b(p) = (1/b(p)) \sum_{u\in W} c(w)|p|_u|w|_u$$

Now, since $|w|_u$ is 1 when $w = u$ and 0 otherwise, this becomes simply $c(w)|p|_w$.

But there's another way to present this. Note that $|p|_u|w|_u$ is either $|p|_u$ (when $|w|_u$ is 1) or 0 (when $|w|_u$ is 0). Either way, $|p|_u|w|_u = \min(|p|_u, |w|_u)$. Let's think this through again for a Lukasiewicz logic. In that account of truth values, the standard min-conjunction has $|p \wedge q|_u = \min(|p|_u, |q|_u)$. For the particular conjunction we are focussed on (where one conjunct is a point proposition), $|p|_u|w|_u = |p \wedge w|_u$. Thus the above becomes:

$$b(w|p) = (1/b(p)) \sum_{u\in W} c(w)|p \wedge w|_u = b(p \wedge w)/b(p)$$

So we've derived a version of the ratio formula, for the special case of credences (under conditionalization) in point-propositions.

Do we also get a ratio formula for conditional credences in arbitrary propositions? We do, but not exactly of this form. To begin with, note conditional credences at worlds fix everything else:

$$b_p(q) = (1/b(p)) \sum_{u\in W} c_p(u)|q|_u = (1/b(p)) \sum_{u\in W} c(u)|p|_u|q|_u$$

But we do not have $|p|_u|q|_u = |p \wedge q|_u$ in general, for min-conjunction. But there is a well-studied fuzzy conjunction that by definition satisfies this equation—product conjunction.

For this conjunction (written with a bullet) $|p|_u|q|_u = |p \bullet q|_u$. This is in fact the defining condition for product conjunction—hence the name! (What we earlier saw is that product conjunction and min conjunction take the same values when one of the propositions in question is a point-proposition). From this point, we can argue just as above:

$$b(q|p) = (1/b(p)) \sum_{u \in W} c(u)|p|_u|q|_u = (1/b(p)) \sum_{u \in W} c(u)|p \bullet q|_u = b(p \bullet q)/b(p)$$

So we have a general ratio characterization of conditional probability, but it uses product rather than min conjunction.

The adaption of the recipe we've been following, and the description in terms of product-conjunction ratio, is by no means specific to the fuzzy setting. For any setting, if we follow that recipe, and have a connective that works like product conjunction over propositions, then we have the result. So for example, in supervaluational or Kleene settings (interpreting both gap and falsity as truth value 0), as in classical settings, ordinary conjunction is a product conjunction. But sometimes we can't have a product conjunction: in a finite Lukasiewicz setting with truth values $1, 1/2, 0$, the product conjunction of a middle-truth value A with itself should take truth value $1/4$, but that isn't an admissible truth value. (The process of conditionalization described above is still well-defined—it's just that we do not have the resources to characterize it in terms of a ratio of categorical probabilities—the same thing would happen in an infinite Lukasiewicz setting, relative to a language that lacked a product conjunction).

We have a candidate notion of conditional probability, and by construction it meets the minimal adequacy constraint of mapping generalized probabilities to generalized probabilities. And given earlier discussion, it is no surprise when we recognize $P(p|p) = 1$ doesn't hold in general for this notion of conditional probability. But is there any reason to be interested in this notion of conditional probability? The next section provides a (synchronic) dutch book argument, on the assumption that conditional probabilities are fair betting odds for a kind of conditional bet to be described.

4 A dutch book argument for generalized conditional probabilities

In our generalized setting, the interpretation of bets needs to be studied. Take a simple categorical case: one where we bet on A, with prize 1 unit, and price α which we will summarize as a bet on A at $[1, \alpha]$. In the classical case, there are just two possibilities: A is perfectly true, or perfectly false. Accordingly, we either are owed the prize, or are owed nothing (in either case, losing the price). But in the general case, A may have an intermediate truth value. Following Paris (2001) we'll take it that categorical bets on A at $[1, \alpha]$ have a prize that is proportional to the truth-value of $|A|_w$ at world w. The reward is scaled by truth value; the cost is fixed. Crucially, we take it that your credence in A gives your fair price for such a unit bet. Under these assumptions, Paris' result is that there is a dutch book against a credal state iff it is not a generalized probability as characterized above.

To extend this to conditional probabilities, it is standard to introduce the notion of a conditional bet. With classical presuppositions, the assumption is that the bet on A conditional on C with prize 1 and price β will return the prize if AC is the case; will return nothing if $\bar{A}C$ is the case; and the bet is called off (with a return of the initial stake) if the condition C is false. In the general case, we need to consider what happens to the bet in situations when C is partially true. Here is the stipulation: a conditional bet on A given C at $[1,\beta]$ are progressively 'more cancelled' as C gets more false. That is, part of the stake is returned, and the potential prize decreased, in proportion to the falsity of C. Modulo this, the returns depend on A's truth value as in the categorical case. The overall return of the unit bet above is therefore $|C|_w(|A|_w - \beta)$ at w. The philosophical premise we need is that the fair price for a conditional bet so construed is exactly the conditional probability of A on C.

In what follows, we write AC for the product conjunction of A and C—so that $|AC|_w = |A|_w|C|_w$ for all w. We shall also assume, for now, that $|\neg X|_w = 1 - |X|_w$—this will hold in some, but not all, non-classical settings.

At this point, we can set up the following triple of bets (generalizing the standard conditional dutch book as given in e.g. Jeffrey (2004)): we sell a bet on AC at $[1,\alpha]$, where we regard the price α as exactly fair (i.e. it matches our credence in AC). We buy a conditional bet on A given C at $[1,\beta]$ again with the price β as fair—matching our conditional credence in A given C. And we sell a bet on $\neg C$ at $[\beta,\beta\gamma]$ again at fair price. Since this bet does not have unit prize, but instead prize β, $\beta\gamma$ is a fair price only if γ matches our credence in $\neg C$.

The net return from each bet depends on the truth values of the propositions in question. Let's write $|A|_w = a$ and $|C|_w = c$. The first bet returns $\alpha - ac$ (i.e. the price we sell at minus the prize we have to give out). The second bet (which we buy) is conditional, and by the earlier characterization equals $c(a-\beta)$. The third bet returns $\beta\gamma - (1-c)\beta$—again, the price minus the prize we have to give out. (Note in this last case that the scaling of the prize by (1-c) reflects the assumption we made about the truth value of $\neg C$).

Summing, the total return from the book of bets is therefore: $\alpha - ac + c(a-\beta) + \beta\gamma - (1-c)\beta = \alpha - \beta + \beta\gamma$. Notice that the only thing this depends on is α, β and γ—our credence in A, A given C, and C respectively. This is therefore a bet that delivers the same return at every world—so it's a dutch book iff this uniform return is negative. If it happens to be positive, switching to selling rather than buying the bets at fair price will be a dutch book). If we are not to be dutch booked, then the above sum better equal zero. Rearranging, this condition gives us that: $\beta = \alpha/(1-\gamma)$.

We assumed that the fair price for categorical bet on B is $c(B)$; and the fair price for a conditional bet on $C|A$ is $b_C(A)$. So we can in particular interpret the prices above as the fair prices for a given credal state. If that credal state is to avoid a dutch book, we therefore have: $b_C(A) = b(AC)/(1-b(\neg C)) = b(AC)/b(C)$.[2] This is exactly the ratio formula, in the generalized setting. And of course, just as predicted earlier, the notion of conjunction we are forced to use is product conjunction.

It would be nice to generalize this argument to get rid of the assumption that the truth value of

[2]Notice that $b(C) = 1 - b(\neg C)$ follows by our characterization of generalized probabilities, and the assumption that $|\neg X| = 1 - |X|$.

the negation of a proposition is one minus the truth value of that proposition—an assumption that some supervaluationists, for example, would deny. We can do this by introducing the notion of a negative bet. A negative bet on X with prize 1 will return $1-|X|_w$ at world w. Thus, if there is a proposition $n(X)$ whose truth value is equal to $1-|X|$ at each world, then a negative bet on X is equivalent to an ordinary bet on $n(X)$—but we do not need to assume that such an n operator exists for the idea of negative betting to make sense. As is natural, we assume that the fair price, in belief state b, for a negative bet of this kind is $1-b(X)$. If we run the argument above replacing the positive bet on $\neg C$ with a negative bet on C, the argument goes through for the same conclusion about conditional credence.

5 Properties of generalized conditionalization

With our notion of conditional probability characterized, we can investigate its properties. The results below give a couple of crucial properties of generalized conditionalization:

1. Lemma. Assume that $|\neg A| = |1-A|$. Then we have that $b(C) = b(C \bullet A) + b(C \bullet \neg A)$.
   Proof. First note that at arbitrary worlds,

   $$|C| = |C|(|A|+1-|A|) = |C|(|A|+|\neg A| = |C||A|+|C||\neg A|) = |C \bullet A| + |C \bullet \neg A|$$

   For arbitrary generalized probability b, there's an underlying credence-over-worlds c such that $b(p) = \sum_w c(w)|p|_w$. So in particular

   $$b(C) = \sum_w c(w)|C|_w = \sum_w c(w)(|C \bullet A| + |C \bullet \neg A|)$$

   But this in turn is equal to:

   $$\sum_w c(w)|C \bullet A| + \sum_w c(w)|C \bullet \neg A| = b(C \bullet A) + b(C \bullet \neg A)$$

   as required.

2. Corollary. $b(C) = b(C|A)b(A) + b(C|\neg A)b(\neg A)$. Follows immediately from the above by the ratio formula.

There is a generalization of this. Let's call a set of propositions Γ a partition if in each world, the sum of the truth values of the propositions in this set is 1 (thus our assumption that $|\neg A| = 1-|A|$ ensured that $A, \neg A$ was a partition). Then replicating the above proof delivers:

1. Generalized Lemma. $b(C) = \sum_{\gamma \in \Gamma} b(C \bullet \gamma)$, so long as Γ is a partition.

2. Generalized Corollary. $b(C) = \sum_{\gamma \in \Gamma} b(C|\gamma)b(\gamma)$, so long as Γ is a partition.

It's nice to have this general form since there are some settings (supervaluational semantics for example) where the truth values of A and $\neg A$ don't sum to 1; the partition-form is still applicable even though the first result is not. The non-trivial question is what kind of partitions exist, and how we write them down.

Another result used below is that $b_C(A|B) = b(A|B \bullet C)$. This follows straightforwardly from the ratio formula. For:

$$b_C(A|B) = b_C(A \bullet B)/b_C(B) = \frac{\frac{b(A \bullet B \bullet C)}{b(C)}}{\frac{b(B \bullet C)}{b(C)}} = \frac{b(A \bullet B \bullet C)}{b(B \bullet C)} = b(A|B \bullet C)$$

6 Jeffrey-style Decision theory

What is a decision theory? From one point of view, it is an axiomatic theory—-a set of qualitative axioms, perhaps, for which we can prove a representation theorem. From another point of view, it consists of a recipe for assigning utilities to arbitrary propositions (perhaps describing options for action) on the basis of utilities of other propositions (describing possible outcomes, perhaps), modulo one's credal state. It is in this second sense that I want to formulate versions of Jeffrey's 'decision theory'. More specifically, I want to obtain a version of the rule Jeffrey uses in practical decision problems, but one that can be used in a generalized setting. I leave the study of the (perhaps qualitative) underpinnings, representation theorems etc appropriate to generalized decision theory for another occasion.[3]

Instead of talking about 'utilities', I will talk of *desirabilities* of propositions, to emphasize the intended continuity with belief-desire psychology. Even those sceptical of Jeffrey's decision theory as capturing the essence of decision-making, may still agree that it captures rational constraints on desire—Lewis (1988, 1996) endorses Jeffrey's account under this interpretation. So construed, a decision theory like Jeffrey's gives us a proposal for quantitative rationality constraints between degrees of belief and desirability. If one thinks further that, given a set of options for action, the choiceworthy one (by one's own lights) coincides that that option one most desires to perform, then the decision theory also gives a recipe for how to resolve practical decision problems (it is on this lst point where Lewis and Jeffrey part ways).

The central equation of Jeffrey-style decision theory is the following (in a classical setting, where Γ is an arbitrary partition of propositions, and b a probability):

$$D(p) = \sum_{q \in \Gamma} b(q|p) D(p \wedge q)$$

An adaption of this makes sense in the general case, using the general notion of partition and conditional probability introduced above, except that we must be clear that the relevant conjunction is product conjunction:

[3]For a comparative discussion of decision theories that focuses on the quantitative equations, see (Lewis, 1981). Lewis and others famously argued on the basis of Newcomb cases that the Jeffrey approach was incorrect, and a causal decision theory was needed (for an overview, see Joyce (1999).

$$D(p) = \sum_{q \in \Gamma} b(q|p) D(p \bullet q)$$

To see this, introduce a valuation function from worlds to reals, v—intuitively, a measure of how much we'd like the world in question to obtain. We use this to provide a model of D defined across all propositions satisfying the above equation. Define D as follows, relative to generalized probability b and valuation v:

$$D(p) := \sum_{w} b(w|p) v(w)$$

Now note the following:

$$b(w|p) = b_p(w) = \sum_{q \in \Gamma} b_p(q) b_p(w|q)$$

This is simply the generalized form of the partition theorem proved at the end of the last section (using the fact that $b_p(w)$ is a generalized probability). Using the final note of the last section, we can rewrite this as follows:

$$b(w|p) = b_p(w) = \sum_{q \in \Gamma} b(q|p) b(w|q \bullet p)$$

We can substitute this into the definition of D:

$$D(p) = \sum_{w} \sum_{q \in \Gamma} b(q|p) b(w|q \bullet p) v(w)$$

Rearranging gives:

$$D(p) = \sum_{q \in \Gamma} b(q|p) \sum_{w} b(w|q \bullet p) v(w)$$

But the embedded sum here is by construction equal to $D(p \bullet q)$. Thus we have:

$$D(p) = \sum_{q \in \Gamma} b(q|p) D(q \bullet p)$$

So valuations over worlds allow us to define a notion of desirability that satisfies the generalized form of Jeffrey's equation.

7 Fetishizing propositions

Let Γ be a partition. Call it a value partition iff for all q, if q is consistent with some $\gamma_k \in \Gamma$, then $D(q \bullet \gamma_k) = D(\gamma_k) = k$.

Let us say that an agent fetishizes p if p, p' is a value partition, where $D(p) = 1$, $D(p') = 0$.

The desires of an agent who fetishizes p can be represented in the following very simple form:

$$D(x) = B(p|x)D(p \bullet x) + B(p'|x)D(p' \bullet x) = B(p|x)$$

This is what David Lewis ((Lewis, 1996) see also (Broome, 1991)) calls "desire as conditional belief". Intuitively, p is the sole good by the light of the agent, and each option is desired by the agent exactly to the extent that it is evidence for p being the case.

We could have odd fetishes (a fetish that everything be red); or more sensible ones (that one's loved one has a happy life). Notice that fetishizing one and the same proposition might lead to very different overall desire states (with different base-valuations over worlds) depending on the truth value distributions. If w is a world that makes it indeterminate whether one's loved one is happy, on a view on which that proposition at that world has truth value 0.5, the desirability of w may be 0.5; but on a view on which it has truth value 0 in indeterminate cases such as these, the world would have desirability 0. This gives one way to get a fix on the psychological differences between various non-classical takes on indeterminacy. One may assess such views by introspecting one's credences about indeterminate propositions and comparing the results to the predictions of theory. But we also have a second introspective test available: to introspect the desirability of the various scenarios featuring indeterminate instances of properties we fetishize.

Of course, the applicability of the test is limited if we can't identify a proposition that we fetishize. However, I think we can imagine fetishizing contents that we in fact intrinsically desire—and hence we can in imagination deploy this test. Indeed, it is not implausible that the basic desirability of a situation is determined by which properties it instantiates, where the properties are drawn from some list of factors we take to be intrinsically valuable—and which themselves can be vague. Thus, Harry's values might include the happiness of loved ones, sufficient equality, personal pleasure, and so forth. What subjective value a given situation has (thus what behaviours are rationalized) will vary depending on what value is inherited by a situation in virtue being a borderline case of some valuable property. Thus, for example, one determinate of value for future possibilities for me, is whether I get good things and avoid pain in that scenario. Suppose I knew that a certain course of action would lead to a situation where someone gets vast arrays of goods, but it's indeterminate whether that person is me. To what extent should I hope/desire that this possibility comes about, just focusing on the goods obtained, and my own benefit? Some might say: not at all. Others say: not as much as in determinate cases of my getting goods; but not as little as in determinate cases of my not getting goods. Holding fixed the assumed source of value, we can evaluate different semantic theories of borderline propositions by how well they predict the reactions one favours.

8 Conclusion

Proposals for revising logic should be accompanied by a specification of how that revision will impact on the theory of mind. Generalizing (categorical) probability theory is one step in this direction. But subjective probabilities themselves stand in functional relations to other mental states, and we need to trace the knock on consequences for this interrelationship (arguably, degrees of belief only count as kinds of *belief* in virtue of standing in these functional relationships).

This note has looked at two crucial steps to an overall non-classical theory of mind. One is the move from categorical to conditional probabilities—needed to link our theory of belief to an account of supposition, revision and to feed into decision theory. The motivation for the particular account of conditional probability specified above, other than the natural way it generalizes the classical case, is given by a dutch book argument (though of course, this simply shifts the locus of philosophical concern to the issue of whether the conditional bets we have described should be linked to suppositional belief).

Building on this, we've shown that one can generalize the core equation of Jeffrey's decision theory. These results are important in themselves, as it is not obvious prior to investigation that anything as neat as this would arise (nor was it at all obvious that product conjunction would turn out to play a major role). I've argued the generalizations given here are motivated, whether they are the best available depends on what their competitors might be.

One final observation. Note that if belief and desirability are as above, then any rational belief and desire state (defined over propositions) can be fixed by specifying (i) truth value distributions for propositions over worlds; and (ii) base-credence and base-valuation functions over worlds. Suppose we think that the basic psychological reality is carried by credences and valuation assignments to possible worlds (a la Lewis, Stalnaker). For all we've said, nothing non-classical need enter the specification of psychological attitudes, at this level of description. We may then see a (classical or non-classical) assignment of truth-values at worlds to propositions (or sentences) as a vehicle for describing psychological attitudes indirectly. On this model, different views on the truth-value distributions of propositions can lead one and the same set of attitudes (specified in terms of worlds) to induce different degrees of belief/desire on propositions—each satisfying the relevant probabilistic and Jeffrey constraints. On this conception of the explanatory order, we could think of rival assignments of degrees of belief and desire to propositions—superficially inconsistent—as projecting one and the same underlying psychological structure. For all we've said, it may be indeterminate which style of semantics is right; and so indeterminate what an agents degree of belief in the proposition p is, while the agents underlying doxastic attitudes were perfectly determinate (i.e. he/she has a definite level of credence in each worldly situation).

Suppose, on the other hand, attitudes to propositions (or mentalese sentences) were the primary psychological reality, and that rational psychology had to satisfy the generalized probabilistic and decision-theoretic constraints. We've seen how we could represent these in terms of a credence-valuation pair over worlds—and the induced 'attitudes' to worlds may be a merely derivative specification of the propositional psychological reality—convenient perhaps for certain theoretical purposes. This leaves open the possibility, moreover, that two theorists could disagree on the degree of belief and desirability appropriate to propositions (and so disagree on the real psychological state), but nevertheless in a strong sense be postulating 'equivalent' psychologies. The sense of 'equivalence' would be that the belief-desire states share a common representation in terms of base-credences and valuations over worlds. This raises an intriguing possibility: If the base credence/value distribution is sufficient for fixing which actions are rational, then there will be a sense in which the two (ex hypothesi genuinely distinct) psychologies are underdetermined by total action-theoretic data.

References

Broome, J. 1991. Desire, Belief and Expectation. *Mind*, **100**(398), 265.

Field, Hartry H. 2001. *Truth and the absence of fact*. Oxford: Oxford University Press.

Jeffrey, Richard. 1965. *The logic of decision*. 2nd edn. Chicago and London: University of Chicago Press. Second edition published 1983.

Jeffrey, Richard. 2004. *'Subjective probability: the real thing'*. Cambridge: Cambridge University Press.

Joyce, James M. 1998. 'A non-pragmatic vindication of probabilism'. *Philosophy of science*, **65**, 575–603.

Joyce, James M. 1999. *The foundations of causal decision theory*. Cambridge, UK: Cambridge University Press.

Lewis, David K. 1981. 'Causal decision theory'. *Australasian journal of philosophy*, **59**, 5–30. Reprinted with postscript in Lewis, *Philosophical Papers II* (Oxford University Press, 1986) 305-36.

Lewis, David K. 1988. 'Desire as belief'. *Mind*, **97**, 323–332. Reprinted in Lewis, *Papers on Ethics and Social Philosophy* (Cambridge University Press, 2000) 42–54.

Lewis, David K. 1996. 'Desire as belief II'. *Mind*, **105**, 303–313. Reprinted in Lewis, *Papers on Ethics and Social Philosophy* (Cambridge University Press, 2000) 55–67.

Paris, J.B. 2001. A note on the Dutch Book method. *Pages 301–306 of: Proceedings of the second international symposium on imprecise probabilities and their applications, isipta*. Ithaca, NY: Shaker.

Williams, J. Robert G. manuscript. 'Gradational accuracy and non-classical logic'.