The Kinematics of Belief and Desire[*]

Richard Bradley
Department of Philosophy, Logic and Scientific Method
London School of Economics
Houghton Street
London WC2A 2AE

March 3, 2010

Abstract

Richard Jeffrey regarded the version of Bayesian decision theory he floated in 'The Logic of Decision' and the idea of a probability kinematics - a generalisation of Bayesian conditioning to contexts in which the evidence is 'uncertain' - as his two most important contributions to philosophy. This paper aims to connect them by developing kinematical models for the study of preference change and practical deliberation. Preference change is treated in a manner analogous to Jeffrey's handling of belief change: not as mechanical outputs of combinations of intrinsic desires plus information, but as a matter of judgement and of making up one's mind. In the first section Jeffrey's probability kinematics is motivated and extended to the treatment of changes in conditional belief. In the second, analogous kinematical models are developed for preference change and in particular belief-induced change that depends on an invariance condition for conditional preference. The two are the brought together in the last section in a tentative model of pratical deliberation.

1 Introduction

Richard Jeffrey's writings spanned epistemology, the logic and philosophy of science, and both individual and social decision theory. Of his many new ideas he regarded two to be the most important. The first is the version of Bayesian decision theory that he developed with the help of Ethan Bolker and is floated in his book *The Logic of Decision* [10]. The second is the idea of a probability kinematics, a generalisation of Bayesian conditioning to contexts in which the evidence is 'uncertain'. In this paper I will attempt to develop some of the connections between these two ideas, by showing how kinematical models can be used to model practical deliberation involving changes in both belief and desire or preference. In the first section, I summarise the motivation for Jeffrey's probability kinematics and its extension to the treatment of changes in conditional belief. In the second, analogous kinematical models will developed for preference change. The two will then brought together in the last section in a tentative model of practical deliberation.

Throughout this paper prospects - the objects of agents' partial degrees of belief and desire - are individuated by propositions $A, B, C, \ldots$. We assume that the relevant set of prospects is closed under the operations of conjunction, disjunction and negation, denoted by $\wedge$, $\vee$ and $\neg$, and contains F and T, respectively the logically false and true propositions. An agent's preferences are represented by a binary relation, $\geq$, on the set of prospects, with $X \geq Y$ saying that the agent does not prefer Y to X. The relations of strict preference and of indifference are defined in terms of $\geq$ in the usual way and respectively denoted by $>$ and $\approx$. Unless otherwise indicated, the proof of any formal claims is given in the appendices.

2 Bayesian Belief Revision

Bayesianism has two faces: it is both an epistemological doctrine and a theory of rational valuation and decision making. In its former role it tells us how we should change our partial beliefs in response to new evidence, where evidence is to be thought of as information generated by interactions with the environment e.g. from observation, experimentation or the reliable testimony of others. Suppose an agent's current degrees of belief are represented by probability measure p, defined on a Boolean algebra Ω of propositions representing all possibilities of concern to her, and that as a result of some such interaction with the environment she learns that A. Classical Bayesianism says that, in these circumstances, your new degrees of belief, q, should be obtained from your old by conditioning on the truth of A, where such (classical) conditioning is defined as follows.

Definition 1 *A probability measure q is obtained from another p by (classical)* **conditioning** *on the truth of A just in case:*

$$q = p(\cdot|A)$$

Now, as Richard Jeffrey was fond of noting, adopting your old conditional degrees of belief given A, $p(\cdot|A)$, as your new degrees of belief is demonstrably the correct thing to do just in case, for all propositions $B \in \Omega$, both:

1. *Certainty*: $q(A) = 1$

2. *Rigidity*: $q(B|A) = p(B|A)$.

So much is just a matter of probability theory. The important question is whether and when we can expect these two conditions be satisfied. The contention of classical Bayesianism in this regard is that they will be satisfied just in case A describes all and everything that is learnt by the agent as a result of the interaction with the environment. One qualification: it is commonly recognised that the Rigidity condition is liable to be violated when the manner in which the agent learns that A has some non-rational effect on the agent's attitudes; an effect that is, so to speak, independent of the propositional content of A. If, for instance, one learns of the consequences of excessive alcohol consumption by doing the drinking oneself or of the presence of a poisonous snake in the house by standing on it, there is every possibility that other beliefs (as well as one's non-credal attitudes) will be altered in the process, but not as a result of conditioning on what has been learnt. So what the Bayesian model claims to describe is just the *rational* effects of learning something; what is implied by what you have learnt in conjunction with your current beliefs.

The Bayesian claim is typically defended by means of a dynamic Dutch Book argument which purports to show that an agent who commits herself to any policy for revising her beliefs other than Bayesian conditioning is vulnerable to sure loss from acceptance of a finite series of bets, all of which are fair by the lights of her degrees of belief. Just how must is established by such arguments is a matter of some controversy.[1] What is important for our purposes however, is the recognition of the fact that, even if the Bayesian contention is true, its scope is narrower than it might at first appear. People are often not aware of all that they have learnt or they fail to adequately represent it, and it is only the failure of the Rigidity condition that alerts us to this. More importantly someone's degree of belief in a particular proposition may change with reason without their being sure of either its truth or falsehood. This was first noted by Ramsey in his criticism of Keynes' interpretation of probability as a logical relation between propositions:

> "I think I perceive or remember something but am not sure; this would seem to give me some ground for believing it, ... He [Keynes] cannot justify a probable belief founded not on argument but on direct inspection." [13, p.86]

For cases where one acquires uncertain evidence - when memory or perception may be mistaken - Jeffrey offers a rule of conditioning that generalises the

---

[1] The original dynamic Dutch book argument for classical conditioning is due to David Lewis, reported in Teller [18]. For criticism of these arguments see Earman [7], Maher [12], Howson [8] and Bradley [5].

Bayesian one. Specifically suppose that $\{A_i\}$ is a partition, a set of mutually exclusive and exhaustive propositions, and that as a result of interaction with the environment (or indeed reflection or deliberation), the agent's probabilities for each A_i changes from $p(A_i)$ to $q(A_i)$. Then:

Definition 2 *An agent is said to obtain her new degrees of belief, q, by **Jeffrey conditioning** on the partition $\{A_i\}$ just in case q is related to p by the 'kinematical' formula:*

$$q(B) = \sum_i p(B|A_i).q(A_i)$$

Jeffrey conditioning on a partition $\{A_i\}$ is appropriate whenever redistribution of belief across $\{A_i\}$ leaves the agent's degrees of conditional belief unchanged i.e. when the Rigidity condition applies to all the conditional degrees of belief given the A_i. This too is just a matter of probability theory. But it is evident that not just our partial beliefs but our conditional beliefs too can change as a result of a number of different kinds of interaction with our environment, including observation, experimentation and testimony.[2] Such changes are not, on the whole, naturally represented as consequences of learning the truth of some set of evidence propositions. For example, I was recently informed that if a magpie is disturbed with young in the nest, it may dive-bomb the intruder. As a result, my conditional degrees of belief for a magpie attacking someone approaching it, conditional on it having young in its nest, rose considerably. But although the effect of the testimony of my informant on my conditional beliefs is clear enough, it is difficult to identify a proposition which would serve as the basis for an explanation of this change in terms of the conditioning model. What I have come to believe about magpies is no doubt supported by facts relating to their biological and cognitive make-up, but I am largely ignorant of them. And although what I was told implies that either the magpie does not have young in its nest or it will attack anyone who comes too close, my new probabilities need not equal my prior conditional probabilities given the truth of this latter proposition. I might, for instance, judge that what I have learnt gives me no reason to change my beliefs about the likelihood that magpie have young (I form my beliefs about this from the statistics say).

This example belongs to an particularly interesting set of cases in which the interaction with the environment gives us cause to change one or more of our conditional beliefs, given some possibility, without it giving us cause to change our probabilities for the possibility itself. In these cases the salient form of revision is what I have labelled updating by Adams conditioning, because of its affinities with Ernst Adams' theory of conditionals and their probabilities.

Definition 3 *Let $\{B_i\}$ be a partition of propositions such that $1 > p(B_i|A) > 0$ and suppose that the agent is caused to change her conditional degrees of belief for the B_i given A from $p(B_i|A)$ to $q(B_i|A)$. Then her new partial beliefs,*

---

[2]Here I am summarising the discussion in Bradley [5], which contains a lengthier explanation of these claims.

*q, are said to be obtained from p by **Adams conditioning** on this change in conditional probabilities just in case:*

$$q(X) = \sum_i (p(AB_iX).\frac{q(B_i|A)}{p(B_i|A)}) + p(\neg AX)$$

Corollary 4 *If $q(B|A) = 1$ for some $B \in \{B_i\}$ and revision is by Adams conditioning then:*

$$q(X) = p(X|AB).p(A) + p(X|\neg A).p(\neg A)$$

What makes Adams conditioning particularly salient is the fact that, in a certain sense, it is the exact complement of Jeffrey conditioning. For in Adams conditioning it is the conditional probabilities with respect to elements of a partition that change while the probabilities of the elements themselves remain rigid, rather than the other way round. Consequently study of this kind of revision offers the possibility of extending kinematical modelling to cases where interaction with the environment affects both the agent's unconditional beliefs and her conditional ones, by representing them in terms of combinations of Jeffrey and Adams conditioning. In later sections we will see just how powerful these models can be.

When is Adams conditioning the right way to change your beliefs? When the total effect of the interaction with the environment on your conditional beliefs for the B_i given A is representable by a redistribution of probability over the partition $\{AB_i, \neg A\}$ satisfying:

1. *Independence*: $q(A) = p(A)$

2. *Rigidity*: $q(\cdot|AB_i) = p(\cdot|AB_i), q(\cdot|\neg A) = p(\cdot|\neg A)$

So much is just a matter of probability theory.[3] But under what conditions should we expect these conditions to apply? Here one might invert the Bayesian argument for classical and Jeffrey conditioning and argue that just as changes in one's degrees of belief concerning some partition $\{A_i\}$ do not in themselves give one reason to change one conditional beliefs for prospects given the A_i, so too if what one learns from an interaction with the environment is appropriately represented by a shift in one's conditional degrees of belief for the B_i given A, and by nothing more and nothing less than these shifts, then the interaction has furnished no reason for a change in one's unconditional degrees of belief for A. Or else the effect of the interaction was not properly represented in the first place.

More than this is needed however to support the claim that either Jeffrey or Adams conditioning are universally valid forms of revision. The argument for both forms of conditioning presupposes that an agent's degrees of belief in some A and her conditional degrees of belief given A are epistemically independent

---

[3] For a proof of both the necessity and sufficiency of Independence and Rigidity for Adams conditioning see Bradley [5].

of one another, *irrespective of what else the agent believes*.[4] I doubt that this is generally the case. The most common counter-examples are those involving actions, when learning that the action has particular effects makes its performance more or less probable and when learning that it is likely to be performed raises the conditional probability of it having certain effects. An increase, for instance, in the degree to which I believe that I will be late for an appointment, given that I take the bus, is liable to make it less likely that I will in fact take it and thereby to motivate a decrease in the degree to which I believe that I will take it. Likewise, if I gather that someone is going to take the bus, then I might infer that he has knowledge about the reliability of the bus that makes the probability of him being late for an appointment, given that he takes the bus, rather low (lower, anyway, than I previously believed). Examples like these show that the various forms of conditioning presented in this section should be regarded not some much as rules of inductive logic, but as tools in what Richard Jeffrey termed 'the art of judgement'.

3 Bayesian Preference Revision

I now want to consider the question of how we should revise our preferences and desires in the light of experience. It is not a question that Richard Jeffrey gave much direct attention to, but many of the tools we will require for our investigation are present in his work. Broadly speaking changes in preference have two sorts of causes: changes in beliefs and what might be called changes in tastes. I shall confine attention to the former and in particular, since it leads on most naturally from what we have discussed thus far, to cases where preference change is induced by a redistribution of belief across some particular partition of the possibility space.

Intuitively there are two kinds of belief change that are especially relevant to preferences. The first is the effect on the desirability of some prospect A of a change in the conditional probability, given A, of prospects that matter to the agent. Here the desirability of A changes because what one expects to be the case in the event that A is, has changed. Thus if learn that drinking red wine, but not white, reduces the chances of a heart attack, I may as a result come to prefer drinking red wine to white. And more generally we would expect that for any prospects A and B:

Conjecture 5 *If the probability of B given A rises, then the desirability of A will rise iff $AB > A\neg B$.*

The second kind of belief change relevant to preference is when a change in the probability of some possibility A makes the prospect of some possibility B more attractive, not because of any probabilistic dependence between the two,

---

[4]By the epistemic independence of an agent's degrees of belief for A and her conditional degrees of belief given A, I mean that the evidential grounds for each are independent. So acquisition of evidence in favour of, say, A will provide no evidential grounds for believing more or less strongly that B, given A.

but because of the desirabilistic dependence of B on A. Thus if I have planned to take my children to the park if I can get away from work early enough, then learning that no rain is forecast for later in the day will make the prospect of getting off work early more attractive. This is not because the forecast affects the likelihood of getting off work, but because I prefer not to go to the park in the rain. More generally we would expect that, for all prospects A and B:

Conjecture 6 *If B is probabilistically independent of A and the probability of A rises then the desirability of B will rise iff $AB > A\neg B$.*

To give a more formal foundation to these conjectures, I will make use of the decision theory developed by Richard Jeffrey in his *Logic of Decision*. In his framework the state of mind of a maximally opinionated rational agent is represented by a pair of functions, $\langle p, v \rangle$, defined on a Boolean algebra of propositions, Ω, and such that p is a probability measure of her degrees of belief and v a real-valued (desirability) measure of her degrees of preference satisfying:

Axiom 7 *(Desirability) If $XY = \bot$, then:*

$$v(X \vee Y) = \frac{v(X).p(X) + v(Y).p(Y)}{p(X \vee Y)}$$

The states of minds of less opinionated agents are represented by sets of pairs of probability and desirability functions: intuitively the set of maximal sharpenings of their opinions consistent with their actual state of mind. It will be simpler, however, to work with maximally opinionated agents and hope that a theory of preference change for real agents can be derived from it.

The axiom of desirability says that the desirability of any prospect X is a weighted average of the possible ways X can be true, where the weighting on each possible way is its conditional probability of truth, given that X. Jeffrey's theory is often labelled a 'news-value' theory of preference, because the desirability of a prospect is a measure of how good it would be to *learn* that something was true, rather than how good it would be to *make* it so. As such, the desirability of any prospect depends not just on how much good stuff it promises, but also on how expected it is. In particular, the tautology is neither desirable nor undesirable because it is certain: its truth is 'no news'. For this reason the tautology is typically assigned a desirability value of zero, so that prospects take positive desirability iff they are good ones (ranked above T). Indeed there is a lot to be said for building this normalisation with respect to the tautology into the formal concept of desirability itself in exactly the same way that it is built into the formal concept of probability; specifically by taking $v(T) = 0$ to be an additional axiom of desirability corresponding to the probability axiom $p(T) = 1$. (I don't do this only because it would settle 'by stealth' the dispute aired in the next section).

A representation theorem for Jeffrey's decision theory was first proved by Ethan Bolker [4], giving necessary and sufficient conditions on an agent's preferences for them to be representable by a pair of probability and desirability

7

functions, $\langle p, v \rangle$, in the sense that $X \geq Y \Leftrightarrow v(X) \geq v(Y)$. A formal statement of his theorem is given in the first appendix and we will need to refer to it in proving some of the formal claims made below. Throughout we will assume that the set of prospects and preference relations on them satisfy the assumptions of Bolker's theorem. However, to avoid needless complication associated with the lack of uniqueness of probability representations of preference in the Jeffrey-Bolker framework we will also assume that the agent's preferences admit only of unbounded desirability representations (we say her preferences are unbounded in this case). In this case it follows from Bolker's theorem that subjective probabilities are uniquely determined by rational preference and desirabilities are determined up to a choice of scale.

3.1 Generalised Conditioning

Suppose that an agent's initial state of mind is represented by the pair of probability and desirability functions $\langle p, v \rangle$. Suppose also that as a result of interaction with the environment the agent's degrees of belief and preference change so that they are now represented by the pair of functions, $\langle q, w \rangle$. Our problem in its most general form is to state the relationship between the pairs $\langle p, v \rangle$ and $\langle q, w \rangle$ in terms of the effects of the interaction on the agent's degrees of belief and desire, where these effects are represented as constraints on the agent's new attitudes. Clearly lots of different kinds of effects can be represented in this way e.g. by $q(A) = q(B)$, $q(B|A) = 1$, $w(A) \geq w(B)$, etc., but I do not propose to study any more than a few salient varieties.

Let us start with the case where there is some proposition A representing all that the agent has learnt, so that the effect of interaction with the environment is given by the constraint that $q(A) = 1$. Here the salient type of updating is classical conditioning, when the agent's new degrees of belief and desire equal her old conditional degrees of belief and desire given A.

Definition 8 *The pair of probability and desirability functions $\langle q, w \rangle$ is said to be obtained from the pair $\langle p, v \rangle$ by **classical Bayesian conditioning** on the truth of A just in case, for all propositions $X \in \Omega$:*

$$
\begin{aligned}
q(X) &= p(X|A) \\
&= \frac{p(AX)}{p(A)} & \text{if } p(A) \neq 0 \\
w(X) &= v(X|A) \\
&= v(AX) - v(A) + v(T) & \text{if } p(AX) \neq 0
\end{aligned}
$$

The functions $p(\cdot|A)$ and $v(\cdot|A)$ appearing in this definition are, respectively, a probability measure and a desirability measure of the agent's degrees of conditional belief and conditional desire given the truth of A.[5] An agent's conditional attitude to any prospect, given that A, is *not* the attitude she will have to it

---

[5] That $v(\cdot|A)$ is a desirability function is proved as Theorem 1 in Bradley [6, p. 32].

in the event that the condition is realised or found to be true, but her current attitude to it on the supposition that A is true. So what Bayesian conditioning on the truth of A essentially consists in, is the adoption of one's current conditional attitudes, on the supposition that A, as one's new attitudes. Thus, if I currently prefer taking white wine to red to a dinner party, conditional on the supposition that fish is to be served, and red to white on the supposition that meat is to be, then once I have phoned my hosts and settled the issue of what we will be eating, I should simply adopt as my preferences over wine, the conditional preference that was based on a correct supposition. So long, of course, as the phone call does not influence my preferences in some manner other than via the information it yields about what is to be served for dinner (in which case we are no longer faced with the circumstances presupposed by the definition).

One important qualification. As Joyce [11] points out, there is more than one way in which a condition can be supposed true. We might suppose that as a *matter of fact* A is true, such as when I suppose, to help with my financial planning, that I won't have enough money at the end of the month to pay the rent. Suppositions of this kind should respect to as greater degree as possible current unconditional beliefs: I should not, for instance, adopt the belief that I will secure a large inheritance to cover the rent. Things are quite different when we suppose or imagine that, *contrary to the facts*, A is true. A supposition of this kind may well be best accommodated by giving up some of one's beliefs not contradicted by A, to allow retention of well-entrenched ideas about the way that the world works. For example, when supposing that it rained yesterday, in order to think about what I would have done had this been the case, I might have to give up my belief that I went for a walk in the park that day, even if I did in fact do so (and have sore feet to prove it). Suppositions of this latter kind are *not* what we hope to capture by conditional probability and conditional desirability functions. What they measure are degrees of belief and desire on the supposition that, as a matter of fact, some condition is true.

The expression for degrees of conditional belief appearing in Definition 8 is well known, that for conditional degrees of desire is likely to be less so. So perhaps a few words of explanation and justification are in order. Intuitively what you learn by getting the news that X and Y is more than just the sum of what you learn from the news that A and the news that B taken separately. Rather it is the sum of what you learn from getting the news that A and of getting the news that B, given that you already know that A.[6] News-value follows news content in this regard. For example, the value of the news that a picnic has been arranged with some friends and that it is going to sunny is greater than the sum of the news value of each taken separately, because sun makes the picnic better and the picnic makes the sun better. What it does equal is the value of the news that it is going to be sunny plus the value of the news of the picnic, given that we already know it will be sunny. (Here we implicitly treat the value of the news that T as zero, since its truth is 'no news').

---

[6] There is rather more to this claim than intuition. Information Theory tells us that the conditional information value of X given A, $I_A(X)$ equals the difference in the information contained in AX and A, $I(AX) - I(A)$. See, for instance, Applebaum [1, chapter 6].

Why the sum? Essentially, because of the way we use desirability to represent preference. Suppose that the utility of money was linear. Then we could define the desirability of a prospect as the fair price for its truth and the conditional desirability of a prospect, given some A, as the fair price for its truth on the supposition that A is true. Then if your conditional desirability for X given A did not satisfy our expression you would be vulnerable to a money pump. Note that since the fair price for T is zero, it follows that $v(T) = 0$. Now suppose that $v(AX) > v(X|A) + v(A)$ i.e. that you are prepared to pay more for the truth of AX than the sum of what you are prepared to pay for A and for X, on the supposition that A is true. Then someone could sell you the truth of A for $v(A)$ and then sell you the truth of X for $v(X|A)$. Finally they could buy the truth of A and X back from you for $v(AX)$ - at a profit! A similar argument will establish that you face sure loss if $v(AX) < v(X|A) + v(A)$.

This argument is necessarily rough, since truths cannot easily be bought and sold and the utility of money is not linear. But I will leave further discussion of conditional desirability to an appendix to section 3. The more important issue is: When is classical Bayesian conditioning, so defined, the right way to change one's attitudes to prospects? To this we can give an precise answer: Whenever interaction with the environment leaves your conditional desirabilities for these prospects, given the truth of A, unchanged. Formally, this invariance of conditional desire is captured by the following rigidity condition. Suppose that an agent's prior and posterior states of mind are respectively represented by the pairs $\langle p, v \rangle$ and $\langle q, w \rangle$. Then:

Condition 9 *(Rigidity of Conditional Desire given A)* $w(\cdot|A) = v(\cdot|A)$

Theorem 10 *If $\langle p, v \rangle$ and $\langle q, w \rangle$ satisfy Certainty and Rigidity of Conditional Desire given A, then $w = v(\cdot|A)$ and $q = p(\cdot|A)$.*

As always the interest of the theorem depends on the scope of the two conditions it invokes. The important one is that of the rigidity of conditional desire for it implies the rigidity of conditional belief (this is proved as Lemma 21 in the appendix). And rigidity of conditional belief we already know to be sufficient for the validity of Bayesian belief revision, be it of the classical variety or the more general form developed by Jeffrey. We have already discussed some of the contexts in which the rigidity condition for conditional belief can fail, including cases where A's turning out to be true has some 'non-rational' effect on her attitudes. In the light of Theorem 10 we can expect that these will be cases in which the rigidity condition on conditional desire fails as well, an expectation borne out by the earlier examples of learning about the effects of alcohol by drinking and of the presence of a snake by standing on it.

As for the Certainty assumption, the requirement that the agent's new degrees of belief for A equal one does little more than formally encapsulate the postulated effect of interaction with the environment. When experience delivers a change in the probabilities of elements of some partition $\{A_i\}$ of the possibility space, without any element achieving probability one, but the rigidity condition holds for conditional desire with respect to the A_i, then the appropriate rule

for updating preferences is a kinematical generalisation of classical conditioning. We state the rule by expressing the agent's new degrees of belief and desire in terms of the old plus her new degrees of belief and desire for the A_i:

Definition 11 *The pair of probability and desirability functions $\langle q, w \rangle$ is said to obtained from the pair $\langle p, v \rangle$ by **generalised conditioning** on a partition $\{A_i\}$ just in case, for all propositions $X \in \Omega$:*

$$w(X) = \sum_i [v(XA_i) + w(A_i) - v(A_i)].q(A_i|X) \tag{1}$$

$$q(X) = \sum_i p(X|A_i).q(A_i) \tag{2}$$

The expression for the agent's new degrees of belief is just Jeffrey's generalisation of classical conditioning; the expression for her new desires is the proposed equivalent generalisation of Bayesian desire revision. To see that it *is* a genuine generalisation, note that if, for some $A \in \{A_i\}$, $q(A) = 0$ and $w(A) = w(T)$, then by 1 $w(X) = v(XA) + w(A) - v(A) = v(X|A)$. Note also that, formally, the relation between the pairs $\langle q, w \rangle$ and $\langle p, v \rangle$ is symmetric: the former is obtained from the latter by generalised conditioning iff the latter is so obtained from the former (this is proved as part of Theorem 13 below). Intuitively, of course, we think of one pair being obtained from the other only when it represents the agent's mind at a later point in time. What is important here is that generalised conditioning is reversible, whenever the probabilities of the A_i remain strictly positive. So it is possible to change one's mind about the evidence and undo the attitude changes premised on it.

Now it is straightforward to establish that:

1. w and q are respectively a desirability and a probability measure on the set of prospects.

2. $w(T) = v(T)$

3. $w(\cdot|A_i) = v(\cdot|A_i)$

Conversely, the equations (1) and (2) appearing in the above definition of generalised conditioning are easily derived from the axioms of probability and desirability, the normalisation of the desirability of T and the assumption that the conditional desirabilities of prospects, given the A_i, are rigid. This is proved as Lemma 22 in the appendix, which in effect establishes that generalised conditioning is indeed a generalisation of classical Bayesian conditioning to cases in which the Certainty condition fails. The first condition is essential, since no method of revision should counsel the adoption of incoherent attitudes. The second is a matter of convenience, with the considerations aired in the previous section in our discussion of news-value applying with equal force here. It is the final property, that of the rigidity of conditional desire given the A_i that marks out Bayesian conditioning as a method of revision.

To deepen our understanding of this condition and hence of the scope of generalised Bayesian conditioning, we can turn to considerations of conditional preference. Conditional preferences are preference judgements regarding ordinary prospects made on the supposition that, as a matter of fact, some condition is true. As we show below, it turns out that invariance of conditional preference is both necessary and, jointly with the Bolker-Jeffrey axioms of preference, sufficient for rigidity of conditional desire and hence for the validity of generalised Bayesian conditioning.

Let $\geq$ and $\geq^*$ respectively represent the agent's preferences before and after her changes in attitude regarding the A_i and let $\geq_{A_i}$ and $\geq^*_{A_i}$ be her corresponding conditional preferences on the supposition that A_i is true. Then we postulate:

Axiom 12 *(Rigidity of Conditional Preference)* $\forall X, Y : XA_i \neq F, YA_i \neq F$,

$$X \geq_{A_i} Y \Leftrightarrow X \geq^*_{A_i} Y$$

Theorem 13 *(Representation of Bayesian Conditioning) Assume that $\geq$ and $\geq^*$ are unbounded and satisfy both the Jeffrey-Bolker axioms of preference and the axiom of Rigidity of Conditional Preference. Let $\langle p, v \rangle$ be a pair of probability and desirability measures that represents $\geq$. Then:*
(i) there exists a pair $\langle q, w \rangle$ of probability and desirability measures on Ω, that represents $\geq^*$ and is obtained from $\langle p, v \rangle$ by generalised conditioning over $\{A_i\}$;
(ii) if the pair $\langle \bar{q}, \bar{w} \rangle$ also represents $\geq^*$, then there exists a pair of functions $\langle \bar{p}, \bar{v} \rangle$ that represents $\geq$ and is such that $\langle \bar{q}, \bar{w} \rangle$ is obtained from $\langle \bar{p}, \bar{v} \rangle$ by generalised conditioning on $\{A_i\}$.

Theorem 13 tells us that if conditional preferences with respect to the A_i are rigid then, on the pain of inconsistency, attitude revision is by generalised conditioning on $\{A_i\}$. But should an agent's conditional preferences given the A_i be rigid? It is tempting to think that they must be whenever the total effect of the interaction with the environment is adequately modelled by a redistribution of probability and desirability over $\{A_i\}$. For suppose that this is the case, but that the agent's conditional preferences, given some $A \in \{A_i\}$, between prospects X and Y change. Then it must be the case that her relative degrees of desire for XA and $\neg XA$ and for YA and $\neg YA$ have changed as a result of the interaction. But then, contrary to assumption, the effects of this interaction are *not* adequately represented by the redistribution of probability and desirability over $\{A_i\}$, since there has also been a redistribution of desirability over the partitions $\{XA, \neg XA, \neg A\}$ and $\{YA, \neg YA, \neg A\}$.

But this argument is clearly not enough as it stands. For what needs to be established is that such a change in her conditional preferences cannot be a rational effect of the postulated changes in beliefs and desire. The thought in favour of the claim that it cannot is as follows. It will generally be the case that the truth of A will be of significance for the agent's preferences for X and Y. But whatever the desirabilistic effects of A on these prospects, this will be

12

accommodated by the supposition of it truth. While it does make a difference to our evaluation of the prospect of X or of Y if we suppose that A is true, the degree to which we believe that A is true should make no difference to our evaluation of these prospects, *conditional* on the supposition that A is true. So while we should *not* expect that $X \geq Y \Leftrightarrow X \geq_A Y$, we should expect that $X \geq_A Y \Leftrightarrow X \geq_A^* Y$.

This argument too is not conclusive. It establishes, I think, that there can be no changes to conditional preferences due to desirabilistic dependency relations between the A_i and other prospects. What it cannot guarantee is that there are no changes in conditional preferences due to changes in the conditional probabilities of prospects, given the A_i. We still need an argument for the rigidity of conditional belief. There are several different kinds on offer, including symmetry arguments (see van Fraassen [20, pp. 334-337]), arguments based on the minimisation of distance between probability functions relative to various measures of closeness, including that of relative entropy (see van Fraassen [19] and Diaconis and Zabell [3]) and, most common of all, dynamic Dutch Book arguments (see, for instance, Armendt [2] and Skyrms [15]). But there are also apparent counter-examples to the rigidity condition and more generally to the claim that conditional degrees of belief given A are epistemically independent of degrees of belief for A (see, for instance, Maher [12] and Bradley [5]). If these counter-examples are real, then we should read Theorem 13 not so much as justifying generalised conditioning but as characterising its domain of application.

3.2   Desirability Kinematics

The characterisation of generalised conditioning given in the previous section does not completely solve the problem we set ourselves; namely to represent the rational effects on preference of a change in the distribution of degrees of belief across a partition $\{A_i\}$ of propositions. What we need to do now is express the agent's new degrees of desire - the $w(A_i)$ - in terms of her prior degrees of belief and desire and her new degrees of belief for the A_i - respectively p, v and the $q(A_i)$. We are guided in this by the principle that changes in the probabilities of the elements of a partition should not affect the desirability of these elements relative to one another, at least in the conditions characteristic of Bayesian conditioning; namely when conditional probabilities relative to these elements are rigid. (Such changes will, of course, affect the relative desirability of propositions that are probabilistically dependent upon them to different degrees). For instance, suppose that I strongly prefer to teach an advanced course in decision theory than an introductory course in logic. Then getting wind of the Head of Department's intentions regarding teaching allocations should not make any difference to the degree to which my preference for the former alternative exceeds the latter.

This principle will not suffice to determine a unique expression for the $w(A_i)$ in terms of p, v and the $q(A_i)$: given the fact that measures of desirability are only determined up to a choice of scale, this would be too much to ask for. But I would contend that the simplest and most natural expression satisfying them,

given the constraint that $w(T) = v(T)$, is the following. Let:

$$w(A_i) = v(A_i) - k$$

where:

$$k = \sum_i v(A_i).q(A_i) - v(T)$$

The k-term is essentially a normalising term and expresses the desirability gain to the agent as result of the change in probabilities. Informally, we can say that it expresses the amount by which the world has been proved, by the experience inducing the belief-change, to be a better (or worse) place than initially believed. Now substituting back into equation 1, we get:

$$\begin{aligned} w(X) &= \Sigma_i[v(XA_i) - k].q(A_i|X) \\ &= \Sigma_i v(XA_i).q(A_i|X) - k \end{aligned} \tag{3}$$

Equation 3 is what we want: an expression for the agent's new degrees of desire in terms of her old plus her new degrees of belief. Note that classical Bayesian conditioning is a special case of it, for when there is some $A \in \{A_i\}$ such that $q(A) = 1$, then it follows from 3 that $w(X) = v(XA) - v(A) + v(T) = v(X|A)$. Let us now illustrate how this all works by recalling the two cases of preference revision that we introduced informally at the beginning of the last section and by deriving the two corresponding revision principles - expressed as Conjectures 5 and 6 - from our conditioning models.

Case 1: Change of Conditional Belief

Suppose that your prior degrees of belief are given by p and that as a result of a change in your degrees of belief for B given that A, you adopt new probabilities q such that $1 > q(B|A) > p(B|A)$, but $q(A) = p(A)$. We are now on the native ground of Adams conditioning. Recall that belief changes of this kind can be represented by a redistribution of probability over the partition $\{AB, A\neg B, \neg A\}$. So by expression 3:

$$w(A) = v(AB).q(B|A) + v(A\neg B).q(\neg B|A) - k$$

where:

$$\begin{aligned} k &= v(AB).q(AB) + v(A\neg B).q(A\neg B) + v(\neg A).q(\neg A) - v(T) \\ &= v(AB).q(AB) + v(A\neg B).q(A\neg B) - v(A).p(A) \end{aligned}$$

by virtue of the fact that $q(A) = p(A)$ and the fact that it follows from the axiom of desirability that $v(\neg A).p(\neg A) = v(T) - v(A).p(A)$. Again by the axiom of desirability $v(A).p(A) = v(AB).q(AB) + v(A\neg B).q(A\neg B)$. Hence:

$$\begin{aligned} k &= v(AB).q(AB) + v(A\neg B).q(A\neg B) - v(AB).p(AB) - v(A\neg B).p(A\neg B) \\ &= v(AB)(q(AB) - p(AB)) + v(A\neg B)(q(A\neg B) - p(A\neg B)) \end{aligned}$$

and so:

$$w(A) = v(AB)(q(B|A) - q(AB) + p(AB)) + v(A\neg B)(q(\neg B|A) - q(A\neg B) + p(A\neg B))$$

14

But by the axiom of desirability $v(A) = v(AB).p(B|A) + v(A\neg B)p(\neg B|A)$. Then since $q(B|A) - q(AB) + p(AB) > p(B|A)$ and $q(\neg B|A) - q(A\neg B) + p(A\neg B) < p(\neg B|A)$ it follows that:

$$w(A) > v(A) \Leftrightarrow AB > A\neg B$$

This establishes Conjecture 5 - that a rise in the probability of B given A will result in a rise/fall in the desirability of A whenever AB is preferred/less preferred to $A\neg B$ - as a consequence of the theory of generalised conditioning.

A rise or fall in the conditional probability of B given A will often lead to preference reversals. Consider, in particular, the case when $q(B|A) = 1$ and $A > \neg A$. Then from 3 it follows that:

$$\begin{aligned} w(A) &= v(AB) - k = w(AB) \\ w(\neg A) &= v(\neg A) - k \end{aligned}$$

Hence in this special case, $w(A) \geq w(\neg A)$ iff $AB \geq \neg A$.

Example 14 *I am considering the prospect (A) of an invitation to lunch at a friend. From past experience I know that if we are invited he will serve either (B) take-away pizza, which I rather like, or (C) homemade Lasagne, which I can barely stomach. I am unsure as to whether to accept or not, but then I am reminded by my wife that our friend served Lasagne last time we went for lunch and this makes it certain that pizza will be served if we go. As a result, the prospect of an invitation appears a good deal more attractive. Schematically we have the following reversal of preference between A and ¬A:*

$\langle p, v \rangle$	$\langle q, w \rangle$
AB	$AB \approx A$
$\neg A$	
T	T
A	
AC	$\neg A$

Case 2: Desirabilistic Dependence

Suppose that some interaction with the environment motivates a redistribution of probability across the partition $\{A_i\}$. So as to isolate the effect of this redistribution on preference due only to relations of desirabilistic dependence with other propositions, let us restrict attention to a proposition B that is probabilistically independent of the A_i. Then $q(B) = p(B)$ and from 3 and the definition of conditional desirability it follows that:

$$\begin{aligned} w(B) &= \Sigma_i v(BA_i).q(A_i|B) - k \\ &= \Sigma_i[(v(B|A_i) + v(A_i) - v(T)].q(A_i) - (\Sigma_i v(A_i).q(A_i) - v(T)) \\ &= \Sigma_i v(B|A_i).q(A_i) \end{aligned}$$

But since B is probabilistically independent of the A_i, it follows from the axiom of desirability and the definition of conditional desirability that:

$$
\begin{aligned}
v(B) &= \Sigma_i v(BA_i).p(A_i) \\
&= \Sigma_i v(B|A_i).p(A_i) + \Sigma_i v(A_i).p(A_i) - v(T) \\
&= \Sigma_i v(B|A_i).p(A_i)
\end{aligned}
$$

since $\Sigma_i v(A_i).p(A_i) = v(T)$, by application of the axiom of desirability. So the desirability of B rises or falls as a result of the interaction depending on whether or not probability shifts to those elements A_j of the partition with respect to which the conditional desirability of B is higher than average i.e. such that $v(B|A_j) > v(B)$. But these are just the elements A_j such that $BA_j > B$. This accords with our initial intuition as to how a shift in the probability in some prospect would affect the desirabilities of prospects probabilistically independent of it.

Example 15 *My friend sometimes serves lunch outside in summer, depending on whether he can be bothered to set the table outside but independently of the amount of pollen in the air. In principle I enjoy eating outside, but since I suffer from hay fever and there is often a lot of pollen at this time of year, I am hoping that he will not be bothered. As it happens, on the day of the lunch the pollen count turns out to be much lower than I expected. Let B be 'We eat outside' and A be 'low pollen count'. Probability has shifted from the state $(\neg A)$ which makes B unattractive to the one (A) which does. So by application of the above we have a reversal of my preference between B and $\neg B$: I now prefer to eat outside.*

$\langle p, v \rangle$: $p(A)$ is low	$\langle q, w \rangle$: $q(A)$ is high
AB	AB
$\neg A \neg B$	$\neg A \neg B$
$\neg B$	B
T	T
B	$\neg B$
$A \neg B$	$A \neg B$
$\neg AB$	$\neg AB$

3.3   Appendix: Conditional Desirability

In this section I will look more closely at the justification for the expression for conditional desirability that has been employed here. To turn the rough money-pump argument given before into something more rigorous, we can look first to another standard weapon in the Bayesians justificatory armoury: a representation theorem for conditional preferences. As before, let us denote an agent's conditional preference for X over Y on the supposition that A is true by $X \geq_A Y$. Such conditional preferences are governed, I claim, by the following rationality constraint linking preferences over prospects given some A

to preferences for these prospects in conjunction with A. Suppose, for example, you prefer the combination of white wine and fish to that of red wine and fish. Then you should prefer white wine to red on the supposition that fish is to be served. And vice versa. More generally:

Axiom 16 *(Conditional Preference)* $X \geq_A Y \Leftrightarrow AX \geq AY$

Now if an agent's conditional preferences satisfy both Bolker's preference axioms and the axiom of Conditional Preference, then it can be shown that there exists a pair of conditional probability and desirability functions, $\langle p(\cdot|A), v(\cdot|A) \rangle$ satisfying Definition 8 and such that, for all prospects X and Y, $v(X|A) \geq v(Y|A) \Leftrightarrow X \geq_A Y$. Furthermore any other desirability function representing the agent's conditional preferences, given A, can be obtained by affine transformation of $v(\cdot|A)$ and is itself a conditional desirability function.[7] In short, rational conditional preference is adequately represented by conditional probability and desirability as we have defined them.

Although considerations of rational conditional preference support our expression for conditional desirability, it does not uniquely pick out this expression, since any expression that amounts to a affine transformation (or rescaling) of ours will receive equal support. In particular it is neutral between my proposal and another common one, which equates the conditional desirability of X given A with the desirability of XA. This proposal is to be found, for instance, in Jeffrey [9], Sobel [17] and, most recently, Joyce [11]. Since I believe that there are no further rationality constraints on conditional preference, I do not think that a choice between the two proposals can be settled by considerations of this kind. We must instead appeal to considerations of conceptual adequacy and representational economy and elegance. There are a number that favour mine.

The first concerns the question of normalisation. On both proposals $v(T|A) = v(A|A)$, which makes good sense from the news-value perspective since the news that A, given that A is already known or supposed to be true, is no news at all. But in addition my proposal fixes the desirability of the tautology across suppositions, so that $v(T|A) = v(T)$. This too is a natural choice from the news-value perspective: the tautology is no news and remains such whatever one supposes true.[8] The alternative proposal, which yields $v(A|A) = v(A)$, fails to reflect the way in which the news-value of A depends on what the agent already knows or supposes to be true. It also thereby rules out the normalisation of desirability with respect to the tautology that was promoted earlier on. And whether or not one wants to adopt it as an axiomatic constraint on desirability, such a normalisation should not be rendered impossible by the framework.

---

[7] This a consequence of the representation theorem for conditional preferences proven in Bradley [6, p. 41], given the simplifying assumptions made in this paper.

[8] Once again, appeal can be made to Information Theory in this regard, since $I_X(T) = I(T) = 0$. Indeed information value can be regarded as a special case of desirability; one in which the value of a prospect depends only on how surprising it is. Given this it to be expected the theory of conditional information should be consistent with the theory of conditional desirability. It is on our version of the latter; not on the alternative.

The second issue concerns the treatment of independence in the desirability of prospects. Intuitively the desirability of some prospect X is independent of the truth of another prospect Y just in case the truth or falsity of Y makes no difference to the desirability of X. Independence of this kind will not be directly manifest in an agent's preferences but, in analogy to probabilistic independence, it would be natural to define it as follows.

Definition 17 X *is desirabilistically independent of* Y *iff* $v(X|Y) = v(X)$.

It is straightforward to establish that it follows from this definition that X is desirabilistically independent of Y iff $v(XY) = v(X) + v(Y) - v(T)$ and hence that:

1. All prospects are desirabilistically independent of the logical truth T.

2. T is desirabilistically independent of all prospects.

3. Only T is desirabilistically independent of itself.

4. Desirabilistic independence is a symmetric relation.

Each of these is what one would expect from a news-value theory of desirability. All prospects already imply the truth of T so the supposition that it is true cannot affect their desirability. Likewise, as noted above, the news that T is valueless whatever else is supposed true. Similarly the supposition that A is true renders the news that A valueless. Finally, symmetry is to be expected, for if the desirability of X depends on Y then it is part of Y's value that it conditions X in this way and so the desirability of Y must depend on the truth of X.

On the alternative proposal, on the other hand, $v(X|Y) = v(X) \Leftrightarrow v(XY) = v(X)$. But XY should be ranked with X only when the desirability of Y is *neutralised* by the truth of X, either because it is implied by X or because Y is a matter of indifference to the agent given X. The alternative proposal thus conflates two quite different notions: the desirabilistic independence of X and Y and the neutrality of Y given X. As a result only the first of the four properties follow from it. T is desirabilistically independent only of prospects ranked with it, while every prospect is categorised as desirabilistically independent of itself. And desirabilistic independence is not symmetric, for it clearly does not follow from the fact that $v(XY) = v(X)$ that $v(XY) = v(Y)$.

The third issue concerns the notion of invariance or rigidity of conditional desire. Intuitively an agent's degrees of desire for some prospect, given the truth of A, are rigid if they do not vary in response to changes in the probability of A. Thus an agent who decides to get fit might change her attitude to the prospect of (A) a long walk to work each day, without changing her conditional attitude, on the supposition that she takes the long walk, to, say, (B) speaking on the phone to her mother. In this case we would expect that $w(AB) \neq v(AB)$, simply because of the change in attitude to A, even though by assumption $w(B|A) = v(B|A)$. But on the alternative account of conditional desirability this is impossible. I conclude that this account must be wrong.

4 A Note on Practical Reasoning

4.1 A Paradox in Preference Theory?

Jeffrey's discussion of preference change is largely confined to a small paper "*The Kinematics of Preference*" penned in response to a problem set for him by Wolfgang Spohn and Ethan Bolker.[9] Suppose that an agent judges some A to be desirable (i.e. $A > T$), has decided to make A true and believes that it is within her power to do so. Then her probability for A should equal one. But then it follows immediately that the desirability of A equals that of the tautology, for $v(T) = v(A \vee \overline{A}) = v(A).p(A) + v(\overline{A}).p(\overline{A}) = v(A)$. So, contrary to assumption, the agent must be indifferent between A and T.

This paradox is generated by a simple conflation of the agent's state of mind before and after practical deliberation. In the Spohn-Bolker example practical deliberation yields the judgement that A is desirable. It is supposed that this judgement motivates certain belief in A, which in turn renders the desirability of A equal to that of T. The relative desirability of A and T changes as a result of deliberation, but at no time is A both strictly preferred and not strictly preferred to T. Only if it is supposed that an agent who finds A desirable must already believe that she will perform it, will the paradox arise. But to suppose this is to misuse the kind of equilibrium models of the states of minds of rational agents that decision theory employs. These models represent the mind 'at rest' in the sense that all logical inconsistencies have been resolved by appropriate adjustments to degrees of belief or preference. But the states so represented should not be assumed to be robust with respect to processes such as perception or practical deliberation. The latter should be regarded as potential perturbers of the equilibrium, causes of changes in beliefs or desires that can only be accommodated by a shift to a new equilibrium state. The Spohn-Bolker paradox arises when one strengthens the equilibrium conditions so as to include all effects of practical deliberation. As this is tantamount to making it a requirement of rationality that an agent knows what they will do, practical deliberation is in effect rendered irrelevant.

The lesson is that if we want to understand practical deliberation and its effects on the agent's state of mind, then we should treat the outcome of deliberation as a new constraint on the agent's probabilities and desirabilities that must be satisfied by our models of her state of mind. (The point is, of course, not specific to practical deliberation; by appropriate adjustment to the equilibrium conditions for the models of minds, all sorts of reasoning can be handled in this way). When deliberation yields a judgement concerning matters on which the agent was previously undecided (e.g. that one course of action is preferable to another), then the model must be made more specific e.g. by shrinking the set of probability and desirability functions representing the agent's state of mind. Where the judgement conflicts with a prior attitude, then it must be accommodated by appropriate revision along the lines developed in the previous sections.

---

[9] The paper is reproduced in [9].

Let us now apply our model of preference revision to the example at hand in order to dispel the lingering worry that the fact the agent comes to be indifferent between A and T undermines her grounds for acting. Suppose that the agent's current state of mind is represented by the pair $\langle p, v \rangle$ and that her options are simply the set $\{A, \neg A\}$ where A denotes the performance of some action. First it should be noted that in general, the judgement that A is preferable to $\neg A$ will not push the agent's probability for A all the way to one, at least while the possibility that the action will not be performed remains open. But the judgement will generally motivate a rise in the probability of A; say to $q(A)$. Then guided by the suggested normalisation of the last section, her new desirabilities w must be such that $w(T) = v(T)$ and $w(A) = v(A) - [v(A).q(A) + v(\neg A).q(\neg A)]$. But $v(A).q(A) + v(\neg A).q(\neg A) > v(T) = v(A).p(A) + v(A).p(A)$, so the desirability of A moves closer to that of T. Simultaneously, however, the desirability of the alternative - $\neg A$ - moves further away from T (in a negative direction), so the grounds for the agent's decision are not undone by her confidence that she will perform it. The worry disappears when we notice that it follows from expression 3 that:

$$
\begin{aligned}
w(A) - w(\neg A) &= v(A) - k - v(\neg A) + k \\
&= v(A) - v(\neg A)
\end{aligned}
$$

So although the agent's judgement of the desirability of both A *relative to that of T* changes, her judgement of the desirability of A *relative to $\neg A$* does not.

What about when the probability of A goes all the way to one, so that the prospect of A's performance is indeed neutral for the agent. Here there is no desirability of an alternative to compare to that of A's for there are no alternatives to A that remain epistemically open. Could the agent's indifference between A and T now abort her action?

> "Of course not: That indifference is predicated on the very fact that you are enacting A (and in a way that you are sure cannot fail)"
> - Jeffrey [9, p. 241]

4.2   Rigidity and Decision

The discussion above directs us to a phenomenon of general interest; namely that the results of practical deliberation can give us reason to revise our beliefs. Discovering that it would be desirable for some agent to act in a certain way, for instance, may give one reason to attach greater probability to any possibilities which actions of this kind tend to realise. Sometimes, of course, it will not; canonically when we do not believe that evaluation of the desirability of the action is one shared by the agent because they do not have access to the information supporting our assessment. But when the agent in question is oneself, for instance, it is normally to be expected that judging that it would be desirable to perform an action is a reason to believe that one will perform it, or at least to raise the probability that one will. There is no rule for doing

so to be offered here, for one's confidence that one will end up performing the action one currently considers to be the most desirable depends on one's beliefs about the possibility that other alternatives will present themselves, that one's perspective will change, that reconsideration will yield a different conclusion, that the option of performing the action will not be removed for reasons beyond one's control, and so on.

Considerations of this kind were previously used to illustrate limitations on the scope of Adams' conditioning. I now want to show how, nonetheless, we can build kinematical models involving practical reasoning of this kind by combining the elements of the last few sections. For simplicity, suppose that an agent has two options O and $\neg O$ which she evaluates in terms of their expected consequences and that our degrees of belief for her picking any particular option depend only on what we consider her evaluation of them to be. Let $\{C_i\}$ be the set of possible consequences of O and $\neg O$. Then how likely the agent is to choose O depends only on the relative desirability of the C_i and their probability given that she decides to O. Suppose now we receive information which changes *our* conditional degrees of belief for the C_i given that she picks O. If we have no reason to believe that such information is available to the agent we can safely Adams condition. If not, or indeed if for any reason we think our own state of mind is an indicator of the agent's, then the Independence condition cannot be assumed to hold. In such cases, nonetheless, correct revision can be approximated by a multi-stage procedure involving conditioning in its various forms. The idea is to proceed as follows:

1. Adams condition on your new conditional probabilities for the C_i given O.

2. Revise the conditional degrees of belief for C_i given O that you impute to the agent and use generalised conditioning on them to revise any other relevant imputed degrees of belief and desire.

3. Calculate the new subjective desirability of O for the agent relative to their (imputed) new degrees of belief and desire.

4. Revise your degree of belief for O in the light of the agent's new preference ordering over them.

5. Condition on your new probabilities for O on the assumption that the Rigidity condition for conditional belief holds relative to them, thereby obtaining your final degrees of belief and desire.

Example 18 *Suppose that some friends phone to say that they are thinking of going to the seaside and ask whether would I like to join them. I decline: although I enjoy both the seaside and their company, my friends are keen sailors and, conditions permitting, they will no doubt go boating. And I would rather stay at home than go sailing. A little later, I hear on the radio that no boats are launching because of high waves. I wonder whether I should change my mind and go with them. I phone, but they are not at home and it is likely that they*

have already left. I could still drive down and join them, but should I? Let B be 'my friends go boating', S_f be 'my friends go to the seaside' and S_I be 'I go to the seaside'. Assume that $S_I S_f \neg B > \neg S_I > S_I S_f \neg B > S_I \neg S_f$. I start by revising my degrees of belief, p, by Adams conditioning on my new conditional probability for B on the assumption that S_f, $q(B|S_f) = 1$. By Corollary 4, this gives:

$$q(X) = p(X|S_f \neg B).p(S_f) + p(X|\neg S_f).p(\neg S_f)$$

In particular, $q(S_f \neg B) = p(S_f)$ while $q(S_f B) = 0$. So from my new perspective, the option of going to the seaside looks very attractive: my most preferred outcome - $S_I S_f \neg B$ - would be the most likely to obtain, were I to exercise it (on the plausible assumption that S_f is causally independent of S_I). But I am not done, for if I have heard the news about the difficulties with launching boats, then in all likelihood so have my friends. Putting myself in their shoes, I evaluate the desirability of going to the seaside using the degrees of belief and desire that I impute to them. From their perspective, the option of going to the seaside looks much less attractive than before. So I considerably reduce my probability for them taking this option. Now I revise by Jeffrey conditioning on the change from $q(S_f)$ to $r(S_f)$, giving new degrees of belief, r, such that:

$$
\begin{aligned}
r(X) &= q(X|S_f).r(S_f) + q(X|\neg S_f).r(\neg S_f) \\
&= p(X|S_f \neg B).r(S_f) + p(X|\neg S_f).r(\neg S_f)
\end{aligned}
$$

Now $r(S_f \neg B) = r(S_f) < r(\neg S_f)$, so the most likely outcome, were I to go to the seaside, is my least preferred one: $S_I \neg S_f$. I decide to stay at home after all.

Reasoning of this kind is critical in strategic contexts when my assessment of the desirability of the options from the other person's point of view is crucial to predicting what they will do and hence in determining what I should. So there is some hope that kinematical models can be put to work in analysing and guiding strategic reasoning. What it does not do, however, is give an universally applicable algorithm for reasoning of this kind. I doubt that steps 2 and 4, for instance, can be made more precise by giving a general rule for revising one's probabilities in the light of desirability calculations of the type given or for revising the degrees of belief one imputes to others.[10] The role of judgement is irreducible here. The elimination of judgement from practical reasoning was not, in any case, my aim here. On the contrary, it was to show how the kinematical tools presented in this paper can be brought together and applied in complex situations in a way that supports, but does not eliminate, judgement.

---

[10]Which is not to deny the interest in studying revision rules of this kind for particular contexts. See for instance Skyrms [16].

5 Appendices

5.1 Bolker's Representation Theorem

Assume that the set of prospects with associated operations of negation, conjunction and disjunction forms a complete atomless Boolean algebra with the logical falsehood, F, removed. An algebra of propositions is atomless if for every proposition $X \neq F$ there exists a proposition $Y \neq F$ that implies X, but which is not implied by it. It is complete if it closed under disjunctions of arbitrary sets of mutually inconsistent propositions.

Let $\geq$ be a complete and transitive (preference) relation on the set of propositions Ω, that is continuous in the sense that if $\{Y_n\}$ is a chain (i.e. a countable, increasing sequence) in Ω, $Y = \vee\{Y_n\}$ and $X \geq Y \geq Z$, then $X \geq Y_n \geq Z$ for all large n. Assume:

Axiom 19 *(Averaging) If $XY = F$, then:*
$$X > Y \Leftrightarrow X > X \vee Y > Y$$
$$X \approx Y \Leftrightarrow X \approx X \vee Y \approx Y$$

Axiom 20 *(Impartiality) If $XZ = YZ = F$, $X \approx Y$ and $X \vee Z \approx Y \vee Z$, then for all $Z' \in \Omega$ such that $XZ' = YZ' = F$, it is the case that $X \vee Z' \approx Y \vee Z'$.*

Under these conditions there exists a probability measure p and signed measure u on Ω such that, $\forall (X, Y \in \Omega - \{F\})$, $p(X) \neq 0$ and:

$$X \geq Y \Leftrightarrow \frac{u(X)}{p(X)} \geq \frac{u(Y)}{p(Y)}$$

Furthermore p' and u' are another such pair of measures on Ω iff there exists real numbers a, b, c and d such that:

$$
\begin{aligned}
ad - bc &> 0 \\
cu(T) + d &= 1 \\
cu + dp &> 0
\end{aligned}
$$

and:

$$
\begin{aligned}
p' &= cu + dp \\
u' &= au + bp
\end{aligned}
$$

The function $v =_{def} \frac{u}{p}$ both satisfies the axiom of desirability and represents preferences in the sense that $v(X) \geq v(Y) \Leftrightarrow X \geq Y$, for all prospects X. From above it follows that v is unique up to positive fractional linear transformation i.e. if v' is a another such measure then $v' = \frac{av+b}{cv+d}$ for real numbers a, b, c and d satisfying the equations above.

5.2   Proofs

Proof of Corollary 4.   Suppose that $q(B|A) = 1$, so that $q(B_i|A) = 0$ for all $B_i \neq B$. Then by the definition of Adams conditioning:

$$
\begin{aligned}
q(X) &= \frac{p(ABX)}{p(B|A)} + p(\neg AX) \\
&= p(X|AB).P(A) + p(X|\neg A).P(\neg A)
\end{aligned}
$$

∎

Lemma 21 *If $\langle p, v \rangle$ and $\langle q, w \rangle$ are such that $w(\cdot|A) = v(\cdot|A)$ then:*
(1) $w(T) = v(T)$
(2) $q(\cdot|A) = p(\cdot|A)$

Proof of Lemma 21.   (1) By assumption $w(T|A) = v(T|A)$. So $w(TA) - w(A) + w(T) = v(TA) - v(A) + w(T)$. But $TA = A$; hence $w(T) = v(T)$.
(2) If $p(X) = 1$ then it follows immediately that $p(X|A) = q(X|A) = 1$. So suppose that $p(X) \neq 1$. By assumption $w(X|A) = v(X|A)$. So by the definition of conditional desirability $w(AX) - w(A) = v(AX) - v(A)$. But by use of the axiom of desirability to expand $w(A)$:

$$
\begin{aligned}
w(AX) - w(A) &= w(AX) - w(AX).q(X|A) - w(A\neg X).q(\neg X|A) \\
&= w(AX).q(\neg X|A) - w(A\neg X).q(\neg X|A) \\
&= q(\neg X|A).(w(AX) - w(A\neg X)) \\
&= q(\neg X|A).(w(X|A) - w(\neg X|A)) \\
&= q(\neg X|A).(v(X|A) - v(\neg X|A))
\end{aligned}
$$

by the rigidity assumption.   Similarly $v(AX) - v(A) = p(X|A)(v(X|A) - v(\neg X|A))$. But $q(\neg X|A).(v(X|A) - v(\neg X|A)) = p(\neg X|A).(v(X|A) - v(\neg X|A))$ iff $p(X|A) = q(X|A)$ or $v(X|A) - v(\neg X|A) = 0$. Suppose that the latter is the case. Then $v(XA) - v(\neg XA) = 0$ and hence $AX \approx A\neg X \approx A$. Let Y be any proposition in $\Omega - F$ such that $XY = F$ and $AY \not\approx A$.[11] Then by the axiom of desirability $AX \vee AY \not\approx A$. Hence, from above, $p(X \vee Y|A) = q(X \vee Y|A)$. But $p(X \vee Y|A) = p(X|A) + p(Y|A)$ and $q(X \vee Y|A) = q(X|A) + q(Y|A) = q(X|A) + p(Y|A)$. Hence $q(X|A) = p(X|A)$.
∎

Proof of Theorem 10.   By the desirability axiom $w(A).q(A) + w(\neg A).q(\neg A) = w(T)$. But by Certainty $q(A) = 1$. So $w(A) = w(T)$. Now again by the desirability axiom, $w(X) = w(AX).q(A|X) + w(\neg AX).q(\neg A|X)$. But $q(A|X) = 1$ and $q(\neg A|X) = 0$. So $w(X) = w(AX) = w(X|A) - w(A) + w(T) = w(X|A)$. But by rigidity of conditional desire, $w(X|A) = v(X|A)$. So $w(X) = v(X|A)$. Equally $q(X) = q(A).q(X|A) + q(\neg A).q(X|\neg A) = q(X|A)$. But by Lemma 21, $q(X|A) = p(X|A)$. So $q(X) = p(X|A)$.   ∎

---

[11] In the event that Ω is atomless, then the existence of such a proposition Y is guaranteed by Bolker's Lemma 3.1 [4, p. 300]. If not its existence must be postulated separately.

Lemma 22 *If the pairs $\langle p, v \rangle$ and $\langle q, w \rangle$ satisfy the condition of Rigidity of Conditional Desire with respect to elements of the partition $\{A_i\}$, then $\langle q, w \rangle$ is obtained from $\langle p, v \rangle$ by generalised conditioning on $\{A_i\}$.*

Proof. Suppose that Rigidity of Conditional Desire obtains. Then by the axiom of desirability and the definition of conditional desirability:

$$
\begin{aligned}
w(X) &= \sum_i w(XA_i).q(A_i|X) \\
&= \sum_i [w(X|A_i) + w(A_i) - w(T)].q(A_i|X)
\end{aligned}
$$

But by assumption $w(T) = v(T)$ and $w(X|A_i) = v(X|A_i)$. It follows that:

$$
\begin{aligned}
w(X) &= \sum_i [v(X|A_i) + w(A_i) - v(T)].q(A_i|X) \\
&= \sum_i [v(XA_i) + w(A_i) - v(A_i)].q(A_i|X)
\end{aligned}
$$

by application of the definition of conditional desirability. Similarly, $q(X) = \sum_i q(X|A_i).q(A_i)$. But from the Rigidity condition and Lemma 21(2) it follows that $q(X|A_i) = p(X|A_i)$. So $q(X) = \sum_i p(X|A_i).q(A_i)$.
∎

Proof of Theorem 13. Let A be any member of $\{A_i\}$. By Theorem 5 of Bradley [6, p.36], $\langle p(\cdot|A), v(\cdot|A) \rangle$ represents $\geq_A$ over Ω. But by the Rigidity condition, $\geq_A \equiv \geq_A^*$. So $\langle p(\cdot|A), v(\cdot|A) \rangle$ represents $\geq_A^*$ over Ω. Then in view of the assumption that $\geq$ and $\geq^*$ are unbounded it follows from Theorem 7 of Bradley [6, p. 38] that there exists a pair of probability and desirability measures, $\langle q, w \rangle$, that represents $\geq^*$ over Ω and such that $w(\cdot|A) = v(\cdot|A)$. It then follows immediately from Lemma 22 that $\langle q, w \rangle$ is obtained from $\langle p, v \rangle$ by generalised conditioning on $\{A_i\}$. (ii) follows from (i) in virtue of the fact that generalised conditioning is symmetric. To see this note that since $q(\cdot|A_i) = p(\cdot|A_i)$ it follows that:

$$
p(X) = \sum_i p(X|A_i).p(A_i) = \sum_i q(X|A_i).p(A_i)
$$

Similarly, by rigidity of conditional desire, $w(X|A_i) = v(X|A_i)$. Hence by the definition of conditional desirability, $w(XA_i) - w(A_i) + w(T) = v(XA_i) - v(A_i) + v(T)$. But $w(T) = v(T)$. So $v(XA_i) = w(XA_i) + v(A_i) - w(A_i)$ and hence:

$$
v(X) = \sum_i v(XA_i).p(A_i|X) = \sum_i [w(XA_i) - w(A_i) + v(A_i)].p(A_i|X)
$$

∎

References

[1] Applebaum, D. (1996) *Probability and Information*, Cambridge: Cambridge University Press

[2] Armendt, B. (1980) "Is there a Dutch Book Theorem for Probability Kinematics?", *Philosophy of Science* 47: 563-588

[3] Diaconis, P. and Zabell, S. (1982) "Updating Subjective Probability", *Journal of the American Statistical Association* 77: 822-30

[4] Bolker, E. (1966) "Functions Resembling Quotients of Measures", *Transactions of the American Mathematical Society* 124: 292-312

[5] Bradley, R. (forthcoming) "Radical Probabilism and Bayesian Conditioning", *Philosophy of Science*

[6] Bradley, R. (1999) "Conditional Desirability", *Theory and Decision* 47: 23-55, Kluwer Academic Press

[7] Earman, J. (1992) *Bayes or Bust: A Critical Examination of Bayesian Confirmation Theory*, Cambridge, Mass: MIT Press

[8] Howson, C. (1996) "Bayesian Rules of Updating", *Erkenntnis* 45: 195-208

[9] Jeffrey, R. (1992) *Probability and the Art of Judgement*, Cambridge: Cambridge University Press

[10] Jeffrey, R. C. (1983) *The Logic of Decision*, 2nd ed, Chicago: University of Chicago Press

[11] Joyce, J. (1999) *The Foundations of Causal Decision Theory*, Cambridge: Cambridge University Press

[12] Maher, P. (1993) *Betting on Theories*, Cambridge: Cambridge University Press

[13] Ramsey, F. P. (1926) "Truth and Probability" in D. H. Mellor (ed) *Philosophical Papers*, Cambridge: Cambridge University Press, 1990.

[14] Savage, L. J. (1972). *The Foundations of Statistics*, 2nd ed, Dover: New York

[15] Skyrms, B. (1987) "Dynamic Coherence and Probability Kinematics", *Philosophy of Science* 54: 1-20

[16] Skyrms, B. (1990) *The Dynamics of Rational Deliberation*, Cambridge, Mass: Harvard University Press

[17] Sobel, J.H. (1989) "Partition-theorems for Causal Decision Theories", *Philosophy of Science* 56: 70-93

[18] Teller, P. (1973) "Conditionalization and Observation", *Synthese* 26: 218-258

[19] van Fraassen, B. (1980) "Rational Belief and Probability Kinematics", *Philosophy of Science* 47: 165-87

[20] van Fraassen, B. (1989) *Laws and Symmetry*, Oxford: Clarendon