Richard Bradley
Department of Philosophy, Logic and Scientific Method
London School of Economics
r.bradley@lse.ac.uk

Multidimensional Possible-world Semantics for Conditionals

Abstract

Adams' Thesis - the claim that the probabilities of indicative conditionals equal the conditional probabilities of their consequents given their antecedents – has proven impossible to accommodate within orthodox possible worlds semantics. This paper considers the approaches to the problem taken by Jeffrey and Stalnaker (1994) and by McGee (1989), but rejects them on the grounds that they imply a false principle, namely that probability of a conditional is independent of any proposition inconsistent with its antecedent. Instead it is proposed that the semantic contents of conditionals be treated as sets of *vectors* of worlds, not worlds, where each co-ordinate of a vector specifies the world that is or would be true under some supposition. It is shown that this treatment implies the truth of Adams' Thesis whenever the mode of supposition is evidential.

0. Introduction

This paper addresses a problem that, if not yet of the 'venerable' status, is certainly as recalcitrant as any on the interface between the semantics and pragmatics of language. The problem is the difficulty (indeed, apparent impossibility) of accommodating Adams' Thesis - the claim that the probabilities of indicative conditionals equal the conditional probabilities of their consequents given their antecedents – within the framework of standard possible-worlds semantics.[1] This is unfortunate because Adams' Thesis is strongly supported by both introspection and by empirical evidence relating to the use of conditionals in hypothetical reasoning.

I will proceed as follows. In the first section I will describe the conflict between possible-world semantics and Adams' Thesis and survey some of the possible responses to it. In the second section I will examine one rather sophisticated attempt to accommodate Adams' Thesis within a modified possible-worlds framework, one which treats conditionals as random variables taking semantic values in the unit interval. Although this attempt fails, the way it does so is instructive and serves as a spring-board for my own proposal, which is developed in the 4th and 5th sections. This proposal involves a somewhat different modification of possible-worlds models, namely one in which the semantic contents of sentences are represented by sets of vectors of possible worlds, rather than by sets of worlds.

In what follows we work with a background language ***L*** and a set $W = \{w_1, w_2, \ldots, w_n\}$ of possible worlds, assumed for simplicity to be finite (nothing of substance depends on this assumption). The power set of W - i.e. the set of all subsets of W - is denoted by Ω and the power set of any subset A of W, by Ω_A. By convention when p is a probability mass function on W, then P will be the corresponding probability function on Ω, such that the measure that P places on any set of worlds is the sum of the masses of its world elements, as measured by p.

---

[1] See Adams (1975) for the canonical statement and defence of this thesis. It was proposed prior to this by Jeffrey (1964).

Throughout I will use non-italic capitals to denote sets of possible worlds and italic capitals as sentence variables, reserving the symbols A, B and C for variables that range over factual sentences only (these being sentences in which the conditional operator introduced below does not occur). When the context makes for clear application of this convention, the set of worlds at which the factual sentence A is true will be denoted by the non-italic letter A, and vice versa. The symbols $\neg$ and $\vee$ will respectively denote the sentential operations of negation and disjunction, while the concatenation of two sentences will denote their conjunction. The symbol $\rightarrow$ will denote the sentential operation performed by the words "If ... then ..." in English conditional sentences. We will restrict attention in this paper to conditionals sentences with factual antecedent and consequent.

The use of a single conditional operator may seem to prejudge an important question in the study of conditional sentences, namely whether or not the grammatical difference between indicative and subjunctive or counterfactual conditionals marks a fundamental semantic difference. Certainly there should be no denying that indicative and subjunctive versions of the same sentence can be evaluated quite differently, as is displayed in Adams' famous example of the difference between "If Oswald didn't kill Kennedy, then someone else did" (which is very probably true) and "If Oswald hadn't killed Kennedy, then someone else would have" (which is probably false).

A good theory of conditionals should be able to explain these differences. Equally it should be able to explain the many similarities in the behaviour of the two kinds of conditional. Some authors do so by postulating two different semantic operations with some common properties, others by postulating a single one with a parameter whose values will differ in the two cases. Although I will develop a theory of the second kind, the use of a single conditional operator is motivated by a concern to keep things simple, rather than to force such a unified treatment. Prejudgement can in any case be avoided by thinking of the arrow as a sentence-operator variable. In keeping with this, when discussion is restricted to indicative conditionals I will use the symbol '>' to denote the specific value taken by this operator in this case (no context will arise in which we restrict attention to counterfactuals).

1. The Reconciliation Problem

Adams' Thesis is a version of a more general theory about conditionals known as the Ramsey Test hypothesis. On this hypothesis your belief (or degree of belief) in a conditional $A \rightarrow B$ should match or equal your belief (or degree of belief) in the consequent B, on the supposition that the antecedent A is true. The idea of using the Ramsey Test hypothesis as a constraint on the semantics of conditionals is an old one. The first attempt to embed the hypothesis within a possible worlds semantics was made by Robert Stalnaker in his 1968 and 1970 papers, and although Lewis' (1976) triviality results seemed to put paid to it, there have been subsequent attempts by, amongst others, van Fraassen (1976), McGee (1989), Jeffrey and Stalnaker (1994), and Kaufman (2005).

To explain the difficulty in effecting this accommodation of the Ramsey Test hypothesis, let me start by sketching out the orthodox possible-worlds model of language meaning and use. For the purposes of this essay it is most usefully captured by four central propositions.

1. *Semantics*

The meanings of sentences are given by the conditions in which they are true, conditions being represented by possible worlds. More precisely, the semantic contents of the L-sentences can be specified by a mapping v from pairs of sentences and possible worlds to the set of permissible semantic values. If we let the semantic value assigned to sentence A at a world w be denoted by $v_w(A)$, then the core of the orthodoxy can be given by two propositions:

(1a) Bivalence: $v_w(A) \in \{0,1\}$.

(1b) Boolean Composition:

$$v_w(AB) = \begin{bmatrix} v_w(B) \text{ if } v_w(A) = 1 \\ 0 \quad\ \text{ if } v_w(A) = 0 \end{bmatrix}$$

$$v_w(\neg A) = 1 - v_w(A)$$

$$v_w(A \vee B) = v_w(A) + v_w(B) - v_w(AB)$$

Bivalence says that sentences can take only one of two possible semantic values – truth (1) or falsity (0) – at each possible world. The meaning of the sentence A can therefore be identified with the set of worlds in which it is true, denoted hereafter by

$[A]$. Boolean Composition, on the other hand, determines the relation between the semantic values (truth-conditions) of compound sentences and those of their constituents.

2. *Pragmatics*

The degree to which a rational agent will believe a sentence is given by her subjective probability for the sentence being true. More formally, let p be a probability mass function on the set of worlds that measures the probability of each world being the actual world. Then the rational agent's degrees of belief in sentences will equal her expectation of their semantic value, $E(v(A))$, i.e. be given by a probability function Pr on L such that for all L-sentences A:

(2) $$\Pr(A) = E(v(A)) = \sum_{w \in W} v_w(A).p(w)$$

Note that because of Bivalence this implies that:

$$\Pr(A) = \sum_{w \in [A]} p(w) = P([A])$$

3. *Logic*

A sentence B is a logical consequence of another sentence A (denoted $A \vdash B$) just in case the truth of A ensures that of B, i.e.:

(3) $A \vdash B$ iff $[A] \subseteq [B]$

Note that (1b) together with (3) ensures that $\vdash$ is a classical consequence relation.

4. *Explanation*

The final claim concerns the relationship between the semantics, pragmatics and logic of a language. Loosely, it is this: what belief attitudes it is rational to take to sentences and what inferences is it correct to make with them is determined by what sentences mean and what beliefs one has about possible worlds, and not the other way round.

To make this more precise, let $\Pi = \{p_j\}$ be the set of all probability mass functions on the set of possible worlds W, interpreted as the set of rationally permissible beliefs. And let $V_L = \{v_i\}$ be the set of all permissible assignments of semantic values to

sentences of L.[2] A possible-worlds model (PW-model for short) of L is a structure <W, v, p> where W is the background set of worlds, v belongs to V_L and p to Π. Such a structure determines both what belief attitudes the speaker can rationally take to L-sentences and what inferences she can rationally make with them. In particular, if *Pr* and $\vdash$ are respectively a probability measure and a logical consequence relation on L-sentences then we can say that a PW-model <W, v, p> explains the pair (Pr, $\vdash$) just in case Pr and $\vdash$ are related to v and p by (2) and (3). That is, it yields explanations of the form '$A \vdash B$ *because* $[A] \subseteq [B]$' and '$Pr(X) = x$ *because* $P([X]) = x$'.

The final assumption underlying standard applications of possible-worlds models can now be made explicit, namely:

(4) For all $v \in V_L$ and $p \in \Pi$, < W, v, p > is a PW-model of L.

The implication of (4) is two-fold. Firstly, the semantic assignment is independent of the agent's belief over worlds, and vice versa. And secondly, there are no constraints on agents' attitudes to sentences other than those contained in the specification of V_L and Π.

Possible-worlds semantics is of course consistent with numerous different specific hypotheses about the truth conditions of conditionals. The attraction of the Ramsey Test hypothesis is that it offers a highly plausible constraint on such theories. To apply it however one must be able to determine what one would believe under a supposition. But, as Joyce (1999) observes, there is more than one way of supposing that something is true. In general, if we suppose that *A* is true, then we should form a set of (suppositional) beliefs that includes that *A* and differs as little as possible from our actual beliefs. One way is to suppose that, as a matter of fact, something is true, such as when we assume that Oswald was not in fact the one who killed Kennedy. This will be called matter-of-fact or *evidential* supposition. Alternatively one might suppose that something is true, contrary to what one knows or believes to be the case, such as when we suppose that, for the sake of argument, Oswald hadn't killed Kennedy. This will be called contrary-to-fact or *counterfactual* supposition.

[2] What assignments are permissible depends on the semantic theory under consideration and in particular what values the theory requires of compounded sentences. Condition (1b) above constrains V_L to contain only those assignments respecting the Boolean laws, but does not constrain the assignment to conditionals in any other way.

(Possibly there are more kinds of supposition, but all that matters to our discussion is that there are at least two distinct kinds).

One of the great strengths of the Ramsey Test hypothesis is that these different ways of supposing something true match the differences in our evaluations of indicative and subjunctive conditionals: roughly, we determine our beliefs in indicative conditionals by evidential supposition of the antecedent, and of subjunctives by counterfactual supposition. Furthermore, it is generally accepted that when belief comes in degrees, evidential supposition is achieved by ordinary Bayesian conditioning on the sentence that is supposed true. That is, when I suppose that as a matter of fact that X, I adopt as my (suppositional) degrees of belief my (pre-suppositional) conditional degrees of belief, given that X. It follows that if my degrees of belief in factual sentences are given by probability measure Pr, then my degree of belief in the simple indicative conditional $A>B$ should be given by:

Adams' Thesis: $\Pr(A>B) = \Pr(B|A)$

A corresponding application of the Ramsey Test hypothesis for degrees of belief to counterfactual conditionals would yield that they should be believed to a degree equal to the probability of truth of their consequent on the supposition that, contrary-to-fact, their antecedent were true. It is much more difficult to say how precisely such supposition works, even if the abstract idea of minimal revision is clear enough. Proposals do nonetheless exist – e.g. that of Skyrms (1981) – and though we will not examine any in detail, there is no reason to think that anything said here would be fundamentally incompatible with them.

Adams' Thesis is widely recognised both to capture our intuitions about rational belief in conditionals and to provide the best explanation for the empirical evidence concerning the role played by conditionals in the inferences that people make. But a series of triviality results show that it is impossible to accommodate Adams' Thesis within the kind of possible-worlds semantic framework described above.[3] Indeed it is not even possible to accommodate a very weak consequence of Adams' Thesis,

[3] In addition to Lewis (1976), see, for instance, Hajek (1989), Edgington (1991), Döring (1994), Bradley (1998) and the papers by Hajek and Hall in Eels and Skyrms (1995).

known as the Preservation condition, without generating highly implausible consequences. This latter condition says that if it is epistemically impossible that B, but possible that A, then it is epistemically impossible that if A then B. Formally:

Preservation Condition: If $\Pr(A) > 0$, but $\Pr(B) = 0$, then $\Pr(A>B) = 0$.

Let us examine more precisely why the Preservation condition, and hence Adams' Thesis, is not non-trivially consist with the orthodoxy as presented here. Let A and B be any factual sentences in L. Then the Preservation condition implies that for all PW-models of L, $<W, v, p>$, such that $P([B]) = 0$ and $P([A]) > 0$, it must be the case that $P([A>B]) = 0$. And this can be so only if for every semantic assignment v, either $[A] \subseteq [B]$ or $[A>B] \subseteq [B]$; else the Preservation condition will be violated by some permissible probability p on worlds. But this implies that the Preservation condition can be satisfied by all PW-models of L only if L is trivial in the sense of containing no sentences A and B such that B is logically independent of both A and $A>B$.

Note that we only made use of claims (1a), (2) and (4) to construct this triviality argument for the Preservation condition. Claim (1a) ensured that the set of possible semantic contents of sentences forms a Boolean algebra, (2) that the probabilities of sentences equalled the probabilities of their contents and (4) that the probabilities of semantic contents were independent of the assignment of these contents to sentences. It follows that if the Preservation Condition is true then either conditionals are not just true or false at a world, or that their probabilities are not probabilities of truth, or that there is some restriction on the co-assignment of meaning to sentences and beliefs to worlds not contained in the standard theory.

2. Routes to Reconciliation

A wide range of responses to this problem has been explored in the literature. Authors as Lewis, Jackson and Douvens argue that the triviality results show that Adams' Thesis is false as a claim about rational belief and that the evidence we are disposed to assert conditionals to a degree equal to the conditional probability of their consequent given their antecedent should be explained by pragmatic principles of one kind or another, not by the semantic content of indicative conditionals (which

they take to be that of the material conditional). But as I have argued elsewhere (Bradley, 2002) these accounts are rather unsatisfactory because that don't extend in a natural way to sentences containing conditionals. For example the sentence "If I try to climb Mt. Everest, then I will succeed" is, on these accounts, very probably true, (because I won't attempt the ascent) but not necessarily assertable. But then why is the sentence "It is probable that if I try to climb Mt. Everest, then I will succeed" also not assertable?[4]

At the other extreme, and a good deal more plausibly, non-factualists such as Edgington (1991, 1995) and Gibbard (1981) argue that the triviality results show that conditionals don't make factual claims and hence do not have (standard) truth conditions. This response implies giving up the possible-worlds framework entirely and adopting Adams' Thesis as a stand-alone hypothesis about rational belief in conditionals. The problem with strategy is that it makes it something of a mystery that we argue over the claims expressed by conditional sentences in much the same way as we argue over factual claims (i.e. by arguing over what is the case, not over what we believe to be the case). Furthermore, without some account of semantic value, it is difficult to explain how we compound conditional sentences with other conditional and factual sentences using the usual sentential connectives and how we can make inferences with conditionals that eventuate in sentences that make factual claims. (Consider Modus Ponens: how can we infer the truth of B from that of A using the hypothesis that $A{>}B$, when the latter makes no truth claim?)

In an earlier paper (Bradley 2002), I argued for an intermediate position, namely that conditionals do take truth values, but only in worlds in which their antecedents are true. Others (e.g. Milne 1997, McDermott 1996) have pursued a similar strategy, dropping Bivalence in favour of a three-valued semantics based on the values of truth, falsity and neither. This approach is able to deliver an explanation of Adams' Thesis by modifying (2) in favour of the hypothesis that probabilities of sentences are their probabilities of truth, conditional on them being candidates for truth i.e. either true or false. Unfortunately, in my opinion, nobody pursuing this strategy has given a convincing account of how the truth-values, and hence probabilities, of compounded sentences depend on those of their constituents.

[4] The obvious answer is because it is false, as Adams' Thesis would imply.

In this paper, I want to look at another class of responses, namely those that involve, in one way or another, some restriction on the co-assignment of meaning to sentences and beliefs to worlds. There are two salient candidates for restrictions. It might be the case that what sentences mean depends on what beliefs one holds, or vice versa. This possibility is explored in the papers of McGee (1989), Jeffrey (1991), and Jeffrey and Stalnaker (1994), for instance. Alternatively it might be the case that there are restrictions on what beliefs we can hold that are not contained in the requirement that degrees of belief be probabilities. Either way (4) would fail: some probability functions on worlds would not be admissible belief measures or some combinations of meaning and belief would be impossible.

Let us start by looking at a theory of the first kind, that of Jeffrey and Stalnaker (1994), in which the contents of sentences are treated, not as bivalent propositions, but as random variables taking values in the interval [0,1]. In particular, the content of a simple conditional $A \rightarrow B$ is represented as a random variable taking the value '1' in all worlds in which both A and B are true, the value '0' in all worlds in which A is true, but B is false, and the conditional probability of B given A in all worlds in which A is false. More formally, a semantic assignment is for them a mapping from sentences to [0,1] satisfying Boolean Compositionality plus:

(**JS-semantics**)

$$v_w(A \rightarrow B) = \begin{bmatrix} v_w(B) & \text{if } v_w(A) = 1 \\ \mathrm{E}(v(B \mid v(A)) = 1 & \text{if } v_w(A) = 0 \end{bmatrix}$$

The subjective probabilities of sentences are still determined by their semantic values in the 'orthodox' way – i.e. in accordance with equation (2) - with the probability of a sentence being its expected semantic value. It then follows from the JS-semantics that:

$$\Pr(A \rightarrow B) = \Pr(AB).1 + \Pr(A\neg B).0 + \Pr(\neg A).\Pr(B \mid A) = \Pr(B \mid A)$$

in accordance with Adams' Thesis. It also follows that a rational agent's beliefs will satisfy the following independence condition:

(**Independence**) If AC is a logical falsehood then:

$$\Pr(C(A \rightarrow B)) = \Pr(C).\Pr(A \rightarrow B)$$

The most notable feature of Jeffrey and Stalnaker's account is that the meanings of conditionals depend on agents' subjective degrees of belief, since the latter determine the semantic value of a conditional in worlds in which its antecedent is false. Hence the contents of conditionals are not strictly random variables (despite the title of their paper), but functions from probability measures to random variables.[5] Because belief restricts meaning, condition (4) is violated, and it is this feature of their account, rather than the dropping of Bivalence, that enables them to satisfy the Preservation condition without running afoul of our triviality result. For, on their account, the element in the Boolean algebra of random variables that is picked out by the sentence $A \rightarrow B$ varies with the agent's beliefs. In particular when $\Pr(B) = 0$, the content of $A \rightarrow B$ is the same random variable as that of AB, so that every probability measure on the algebra of random variables determined by Pr must give measure zero to the random variable associated with the sentence $A \rightarrow B$. Hence the Preservation condition is non-trivially satisfied on their account.

Jeffrey and Stalnaker's account does leave some questions unanswered however. Firstly, since random variables take values other than '0' and '1' (or 'false' and 'true') probabilities of sentences are not ordinary probabilities of truth. How then are these probabilities, and hence Adams' Thesis, is to be interpreted? Secondly, and more importantly, they do not offer any explanation as to why the semantic values of conditionals should be related to the agent's conditional beliefs in the way that they postulate and how this dependence of meaning on belief is to be squared with the fact that we appear to use conditionals to make claims about the way that the world is rather than to express our state of mind.

Partial answers to both these questions can be drawn, I believe, from an earlier paper of Vann McGee (McGee 1991). McGee adopts a modified version of Stalnaker's (1968) semantic theory for conditionals in which a conditional $A \rightarrow B$ is true at a world w iff its consequent, B, is true at the world f(w, A), where f is a selection function picking out for any sentence and world w, the 'nearest' world to w at which the

5 This is somewhat obscured by the fact that Jeffrey and Stalnaker implicitly assume a fixed probability measure in their discussion.

sentence is true.[6] Which world is the nearest is not something that is determined by the facts, but depends in part on the agents' beliefs.

> "purely semantic considerations are … only able to tell us which world is the actual world. Beyond this, to try to say which of the many selection functions that originate at the actual world is the actual selection function, we rely on pragmatic considerations in the form of personal probabilities".[7]

In general these pragmatic considerations will be insufficient to determine which world is the nearest one in which the antecedent of a conditional is true (or which selection function is the right one). Hence they will not determine the truth-value of a conditional at a world. But the agent's partial beliefs will constrain the choice of selection function to the extent of determining what might be called the expected truth-value of a sentence at a world, where the latter is defined as the probability weighted sum of the truth-values ('0' or '1') of a sentence, given particular selection functions, with the weights being determined by the subjective probabilities of the selection functions. More formally, let F={f_i} be the set of selection functions on W×Ω, with f_i(w, A) being the world selected by f_i as the nearest one in which A is true. Let q be a probability mass function on F. Then the semantic value of a conditional at a world w is given by:

$$(1c) \qquad v_w(A \rightarrow B) = \sum_{f_i} v_{f_i(w,A)}(B).q(f_i)$$

The expected truth-values yielded by agents' uncertainty about distance between worlds provide a natural interpretation of the intermediate semantic values postulated by Jeffrey and Stalnaker. However equation (1c) does not by itself imply that conditionals take the specific semantic values claimed by JS-semantics. For this the probabilities of selection functions codifying possible nearness judgements must be correlated with the agent's degrees of conditional belief in the right kind of way. On McGee's account this correlation is secured by means of further constraints on rational partial belief, most notably including the aforementioned Independence principle. The additional constraints on rational belief suffice to determine

6 McGee's modification will not matter here, so for simplicity I will omit it.
7 McGee (1989), p. 518

probabilities for selection functions, identified by McGee with complex conditionals of the form $(A \rightarrow w_A)(B \rightarrow w_B)(C \rightarrow w_C)\ldots$, i.e. explicit descriptions of the world that would be the case for each possible condition supposed true. Equations (1c) and (2) then yield Adams' Thesis.

To summarise: on this interpretation of McGee's theory, the semantic value of a conditional at a world at which its antecedent is false is its expected truth value, calculated relative to probabilities of selection functions. Probabilities of selection functions are in turn determined by the agent's partial beliefs, assumed to conform to both the laws of probability and the Independence principle. Hence it is these latter properties of rational belief that determine what conditionals mean and explain why rational belief in conditionals must satisfy Adams' Thesis.

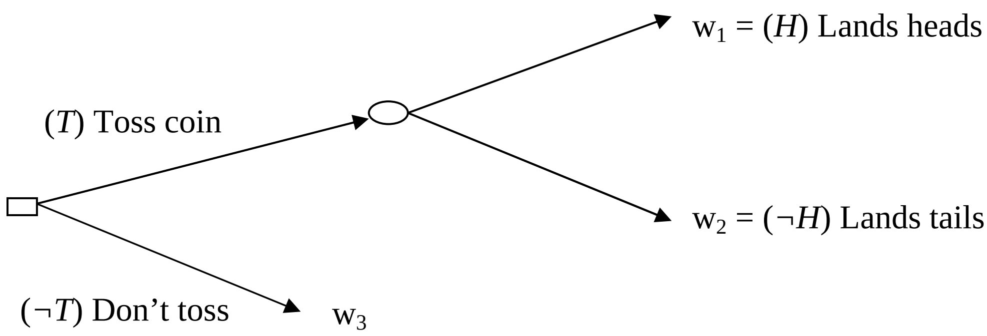
3. Testing McGee's theory

The Independence principle is unsatisfactory in one notable respect: it is stated as a constraint on belief attitudes towards sentences rather than their contents. Ideally, however, our attitudes to sentences should be explained in terms of what these sentences mean or say and what belief attitude it is rational to take to such meanings. Since, on McGee's account, the semantic value of a conditional depends not only on what world is the actual one, but also on what world is picked out by selection functions, it would be expected that such an explanation would make reference to the properties of rational belief attitudes to selection functions. Instead McGee simply argues that the Independence principle has to be true if Adams' Thesis is. For were it not, Adams' Thesis would not survive belief change by Jeffrey conditionalisation.

This argument is far from persuasive. As Richard Jeffrey himself pointed out, rationality does not require belief revision by Jeffrey conditionalisation. It is only rationally mandatory when the shift in probability of some proposition X, does not disturb the conditional probabilities of any other proposition conditional on X. But

this is precisely what must be assumed for McGee's argument to go through.[8] Rather than offering further theoretical considerations for or against the Independence principle, however, I propose to test both it and Adams' Thesis against a couple of examples in which application of the central concepts of McGee's semantic theory – possible worlds and selections functions – is unproblematic.

The first example is a case in which there is uncertainty about what the correct selection function is, but not uncertainty about which world is the actual one. Suppose that we have before us a coin that is known to be fair, so that the chance of it landing heads in the event of being tossed is 0.5. This set-up can be illustrated as follows.



Example 1: Selection uncertainty

Suppose the coin is not tossed. What is the probability that, had it been tossed, it would have landed heads? There are only two plausible selection functions to consider, namely functions f and f′ such that:

$f(w3,T) = w1$

$f'(w3,T) = w2$

If we calculate the probability of $T \rightarrow H$ using (1c′), we see that Adams' Thesis is satisfied, irrespective of the probability of the coin being tossed, since:

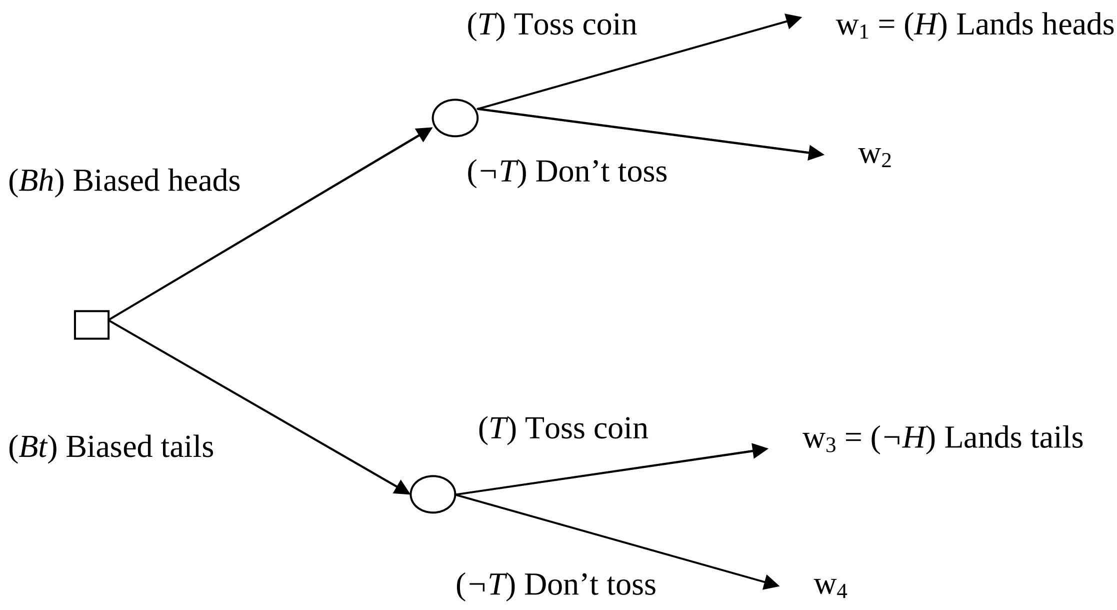
$p(T \rightarrow H) = p(w1) + p(w3).q(f) = 0.5$

$p(H \mid T) = p(w1)/(p(w1) + p(w2)) = 0.5$

In this simple case of selection uncertainty, therefore, his theory offers a plausible explanation for the truth of Adams' Thesis.

8 An associated argument for a more complex version of the Independence principle is made in terms of fair betting arrangements, but this argument assumes the simpler version of the principle.

The second example is a case in which there is uncertainty about which world is the actual one, but no uncertainty about what the correct selection function is. Suppose that we have before us a coin that is known to be biased, but that we do not know whether it is biased in favour of heads or in favour of tails (it is either a two-headed or two-tailed coin say).



Example 2: World Uncertainty

Suppose the coin is tossed. In this case, if it is biased heads it will land heads, and if biased tails it will land tails. Now suppose that it is not tossed. What should we conclude about how it would have landed if it had been tossed? In view of the known fact of bias, there is only one plausible selection function in this case: it is the function f such that:

f(w2,T) = f(w1,T) = w1

f(w4,T) = f(w3,T) = w3

It follows that:

$$P(T \rightarrow H) = p(w_1) + p(w_2) = P(Bh)$$

while:

$$P(H \mid T) = \frac{p(w_1)}{p(w_1) + p(w_3)} = P(Bh \mid T) = P(Bh)$$

on the assumption that whether the coin is biased one way or another is independent of whether it is tossed (which is a reasonable assumption to make in this context. So in this example too Adams' Thesis follows by application of (1c).

On the other hand, the Independence principle fails. To see this note that Independence requires that at every world in which the coin is not tossed, the probability that it would have landed heads had it been tossed, is probabilistically independent of it being tossed. So by Independence:

$$\Pr(T \to H \mid \neg T \wedge Bt) = \Pr(T \to H)$$

But:

$$\Pr(T \to H \mid \neg T \wedge Bt) = 0$$

because, given that the coin is biased towards tails, it is certain that the coin would not have landed heads had it been tossed. On the other hand, as we have seen:

$$\Pr(T \to H) = \Pr(Bh) > 0$$

But this contradicts the above.

The Independence principle is closely related to a claim common to McGee and Jeffrey and Stalnaker's theories, namely that the semantic value of a conditional equals the conditional probability of its consequent given its antecedent in all worlds in which its antecedent is false. This claim too is not satisfied in the second example: the semantic value of $T \to H$ at world w4 is not 0.5 but 0, since at this world it is certain that at the nearest world in which T is true (namely w3), it is false that H. It would seem that JS-semantics upon which these theories rest is false as a claim about the contents of indicative conditionals.

Fortunately for the project of accommodating Adams' Thesis it is not necessary that the semantic value of a conditional be the conditional probability of its consequent given its antecedent at *every* world in which its antecedent is false. All that is required is that *on average* it has this semantic value. Or to put is slightly differently, what is required for Adams' Thesis is not Independence, but the following, logically weaker, condition:

(**Restricted Independence**) $\Pr(\neg A \wedge (A \to B)) = \Pr(\neg A).\Pr(A \to B)$

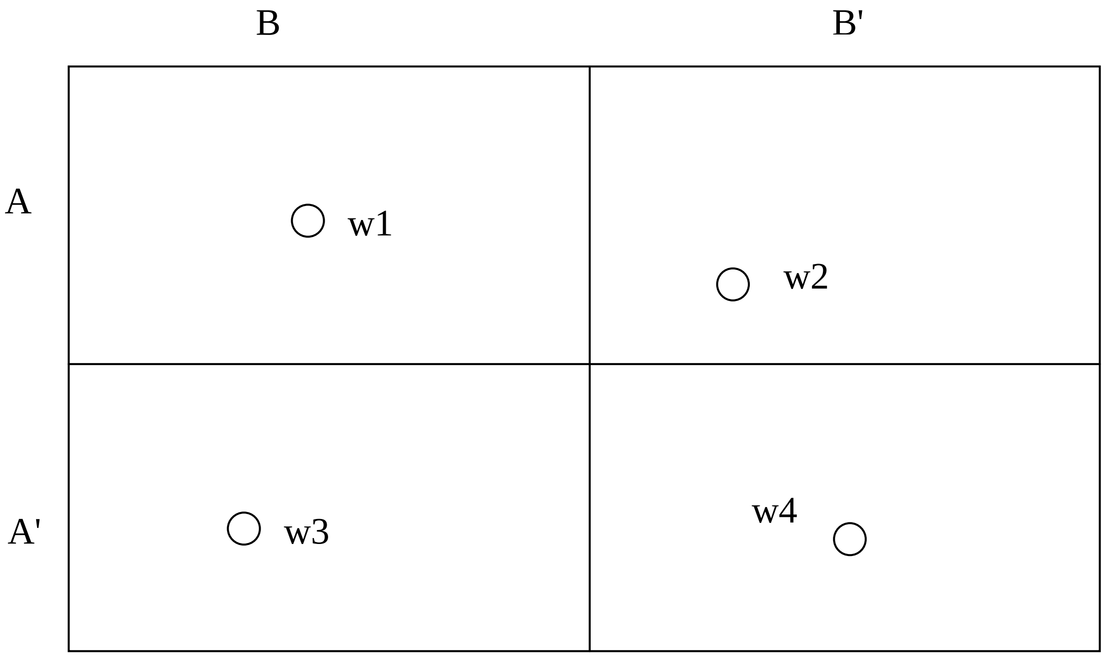
And this condition, like Adams' Thesis (but unlike Independence) does not seem to be undermined by our examples. So in the next section, I would like to build on what we have learnt from these theories to construct a more plausible semantic basis for this restricted kind of independence.

4. Two-dimensional Semantics

The discussion thus far suggests that two types of uncertainty are at play when evaluating conditionals. On the one hand there is the familiar uncertainty about what is the case or about which world is the actual one. On the other hand, there is uncertainty about what is or would be the case if some supposed condition is or were true, or about which world is the nearest one satisfying the condition. I will speak in this latter case of uncertainty as to which world is the counter-actual world under the supposition in question.

McGee represents the first kind of uncertainty in the standard way, by a probability mass function on worlds measuring the probabilities of them being the actual world. The second kind is measured by a mass function on selections functions, each of which represents a hypothesis as to which world is the counter-actual one under each possible supposition. The mass on a selection function gives the probability that it correctly identifies these counter-actual worlds.

To see how this works, consider the following (extremely simple) possible-worlds model shown in the diagram below, exhibiting a set of four possible worlds W = {w1, w2, w3, w4} and 15 associated non-empty propositions, with for instance A={w1, w2} and B={w1, w3}.



Even such a simple model has quite a few possible selection functions associated with it, since each world and proposition pair can take any of the four worlds as values (subject to the usual constraints on selection functions). To derive the probability of the conditional $A \rightarrow B$ in McGee's framework, for instance, we need to determine its expected truth value at each world and then calculate the average of its expected truth values, using the probabilities of worlds as weights. But to work out the expected truth value of $A \rightarrow B$ at w3 say, we need to assign probabilities to function f: f(w3, A) = w1, f(w4, A) = w1; function g: g(w3, A) = w1, g(w4, A) = w3; and so on, and then use these as weights in finding the average truth value of the conditional at that world.

This is unnecessarily complicated. Instead of trying to judge the probability of every complete specification of distance between worlds, we can turn things around and simply judge the probability that each world is the nearest one to the actual world in which the antecedent of the conditional is true. When evaluating the conditional $A \rightarrow B$, for instance, we can confine attention to the two worlds w1 and w2 in which A is true and ask ourselves how probable it is that each is or would be the counter-actual world, if A is or were true. These probabilities may then serve as the weights on the truth-value of $A \rightarrow B$ at w1 and w2 (1 and 0 respectively) that are needed to calculate the expected truth value of the conditional.

Note that how we judge this kind of uncertainty will depend on the kind of supposition that we are engaging it. When evaluating an indicative conditional by engaging in evidential supposition, we need to evaluate how probable it is that one or another A-world (really) is the case, given the truth of *A*. On the other hand when we are engaged in the kind of contrary-to-fact supposition appropriate to the evaluation of counterfactuals, we should evaluate how probable it is that each world would be the case were *A* true (contrary-to-fact).

To summarise: what I am suggesting is that we represent the second kind of uncertainty by a probability function on *worlds*, not on selection functions. This function on worlds measures not their probability of being the actual world but their probability of being the counter-actual world under a supposition. This has the advantage of allowing us to dispense altogether with talk of selection functions and simply work with possible worlds.

With this simplification in place, we can represent in the following way our state of uncertainty (or the part of it relevant to the evaluation of the conditional $A \rightarrow B$) for the simple model we have been using. There are four possible worlds that we need to assess with regard to their probability of being the actual world, and two possible A-worlds which need to be assessed with regard to the probability that they are or would be case if *A* is or were (worlds at which *A* is false are not candidates for this). What we need to examine is how these assessments depend on each other. For this purpose we can tabulate their objects in the following way.

	Supposed A-Worlds	
Possible Worlds	**W1**	**W2**
W1	<w1,w1>	-
W2	-	<w2,w2>
W3	<w3,w1>	<w3,w2>
W4	<w4,w1>	<w4,w2>

Each ordered pair <w_i,w_j> in the table represents a possibility: the event that w_i is the actual world and that w_j is the counter-actual A-world. The blanks in the table mark an important assumption, which I will call Centring in line with the terminology

introduced by Lewis. (I don't insist on the assumption, but it is a natural one to make.) The Centring condition rules that certain combinations of factuality and counter-factuality are impossible. For instance, it is impossible that w2 be the actual world and that w1 be the nearest A-world. For w2 *is* an A-world, and the actual world is always the closest member to itself of any set it belongs to. Likewise if w1 is the actual world, then w2 cannot be the nearest A-world.

Sets of the possibilities represented by ordered pairs of worlds will play the role of semantic values for us. The contents of *factual* sentences are identified by rows of the table. The sentence *A*, for instance, has as its content the first and second rows, i.e. the set {<w1,w1>,<w2,w2>}. The contents of conditional sentences, on the other hand, are identified by columns of the table. The sentence *A*→*B*, for instance, has as its content the first column, i.e. the set {<w1,w1>,<w3,w1>,<w4,w1>}. The contents of conjunctions, disjunctions and negations of sentences, conditional or otherwise, are given by the intersection, union and complements of the contents of their component sentences. So, for instance, if L is the closure of {*A*, *B*, *A*→*B*} under these sentential operations and semantic function [.] an assignment of ordered pairs of worlds to sentences, such that [*A*] = {<w1,w1>,<w2,w2>} and [*B*] = {<w1,w1>,<w3,w1>,<w3,w2>}, then:

[¬*A*] = {<w3,w1>,<w3,w2>,<w4,w1>,<w4,w2>}

[*A*→*B*] = {<w1,w1>,<w3,w1>,<w4,w1>}

[¬*(A*→*B)*] = {<w2,w2>,<w3,w2>,<w4,w2>}

[*A*(*A*→*B*)] = {<w1,w1>}

And so on.

Given this construal of semantic value, it is natural to identify semantic entailment with the subset relation between sets of ordered pairs of worlds. This yields, via the orthodox conception of logical consequence expressed by claim (3), not only the usual laws of propositional logic, but also a number of the familiar rules of conditional logic. For instance, the content of both *A*(*A*→*B*) and *AB* is the set {<w1,w1>}, so this semantic model validates the rule of Modus Ponens for factual sentences. Likewise the content of the sentence ¬(*A*→*B*) is the set {<w2,w2>,<w3,w2>,<w4,w2>}. This is just the content of the sentence *A*→¬*B*, so the semantic model validates the law of conditional excluded middle.

When a pair of worlds $<w_i,w_j>$ is part of the content of a sentence *X* we can say that $<w_i,w_j>$ makes *X* true as a kind of short-hand for the claim that w_i being the actual world and w_j being the counter-actual A-world makes it true that *X*. In this sense we can think of the ordered pairs of worlds as giving the truth conditions of the sentences they belong to, where these conditions are not confined to actuality but extend to counter-actuality under a supposition. But some caution is required: to say that $<w_i,w_j>$ makes *X* true is not to say that *X* is true at either of these worlds. For instance, while the conditional *A→B* is made true by the ordered pair <w3,w1>, we cannot say that *A→B* is made true by the facts (at w3), because these facts alone do not determine the truth of the conditional. In this respect the non-factualists have it right. On the other hand we *can* say that *A→B* is true whenever it is true that w3 *and* true that w1 under the supposition that *A*. And this is enough to support a truth-compositional semantics that accounts for our ability to use conditionals within discourse aimed at establishing facts.

5. Probabilities of Conditionals

I have argued that in order to represent the different uncertainties associated with suppositions, we need not one probability measure, but many: one for each supposition in fact. To examine the implications for conditionals, let *p* be a probability mass function on W that measures the probability that any world is the actual one and *q* be a probability mass function on A that measures the probability that any world is the counter-actual one, on the supposition that *A*. Finally let *pr* be a joint probability mass function on the pairs of worlds that lie in the table cells, measuring the joint probabilities of actuality and counter-actuality under the supposition that *A*. For example, $pr(<w_i,w_j>)$ is the probability that w_i is the actual world and w_j the counter-actual world on the supposition that *A*.

Let me say a word about joint probabilities. For *pr* to be the joint probability formed from *p* and *q* - i.e. if *p* and *q* are the marginal probabilities of the joint probability *pr* - it must be the case that *pr* is defined on the product domain WxA and that the three probability measures are related by the following condition:

(**Marginalisation**):

$$\sum_{w_i \in W} pr(< w_i, w_j >) = q(w_j)$$

$$\sum_{w_j \in A} pr(< w_i, w_j >) = p(w_i)$$

It follows from the Marginalisation property that:

$$pr(< w_i, w_j >) = p(w_i) \times pr(w_j \mid w_i)$$

where $pr(w_j \mid w_i)$ is the probability that w_j is the counter-actual A-world given that w_i is the actual world. Note therefore that $pr(w_j \mid w1)$, unlike $p(w_j \mid w1)$, need not equal zero whenever $i \neq j$. On the other hand it follows from Centring that:

$$pr(w_1 \mid w_1) = 1$$

$$pr(w_2 \mid w_1) = 0$$

Hence our total state of uncertainty can be summarised as follows:

	Supposed A-Worlds	
Possible Worlds	**W1**	**W2**
W1	p(w1)	0
W2	0	p(w2)
W3	p(w3).pr(w1 | w3)	p(w3).pr(w2 | w3)
W4	p(w4).pr(w1 | w4)	p(w4).pr(w2 | w4)

A number of remarks are in order. Firstly, this representation can be used to support the kind of semantic accounts of conditionals looked at in earlier sections in which the contents of sentences are treated as random variables whose values are the expected truth values of the sentences. To do this we simply need to take the semantic value of a sentence at a world w to be given by the sum of conditional probabilities, given w, of the counter-actual worlds at which the sentence is true. Thus the random variable associated with conditional $A \rightarrow B$ would, in our simple model, take the value 1 at w1, 0 at w2, pr(w1 | w3) at w3 and pr(w1 | w4) at w4. (Evidently more assumptions are required in order that these values agree with the JS-semantics).

Secondly, construing these values as expected truth values does not preclude that the expectations in question should be partially or wholly determined by the description

of the relevant world. For instance, if we are willing to assume that the objective chances of prospects are determined by possible worlds, then it might be reasonable to require that the expected truth values of sentences conform to these chances. In particular this would suggest that the expected truth value at a world w of *B* on the supposition that *A* should be given by the conditional chances of B given A at w, as it is on Skyrms' (1981) account of counterfactuals. In the extreme one may even postulate that the properties of a world completely determine which are the counter-actual worlds corresponding to it under each supposition. The result would be to collapse the multidimensional semantics into the orthodox uni-dimensional one.

Finally, and more importantly, this representation of our state of uncertainty in conjunction with the earlier claim that the semantic contents of conditionals are given by columns of the table ensure satisfaction of a probabilistic version of the Ramsey Test hypothesis. For instance, the conditional sentence $A \rightarrow B$ has as its truth conditions the W1 column of the table. This column has probability q(w1), which is the probability that *B* is true on the supposition that *A* is. Hence the probability of $A \rightarrow B$ must be the probability of *B* on the supposition that *A*. This is not just a feature of our very simple example, but is intrinsic to the way in which the multi-dimensional possible-worlds models being advocated here are constructed. Indeed, it is reasonable to say that they are constructed in such a way as to encode the Ramsey Test.

To derive more specific versions of the Ramsey Test hypothesis, appropriate for particular modes of supposition, further constraints need to be placed on the relation between marginal probabilities. Three candidates for constraints appropriate to evidential supposition, are the following:

1. Stochastic independence: For all $w_i \in W$ and $w_j \in A$:

$$pr(w_j \mid w_i) = q(w_j)$$

2. Counterfactual independence: For all $w_i \notin A$ and $w_j \in A$:

$$pr(w_j \mid w_i) = q(w_j)$$

3. Restricted independence: For all $w_j \in A$:

$$pr(w_j \mid A) = q(w_j)$$

These conditions are in decreasing order of strength: stochastic independence implies counterfactual independence which implies restricted independence. Stochastic independence says that the probability of a world being the counter-actual A-world is independent of what world is the actual world, counterfactual independence that it is independent of any world at which *A* is false, and restricted independence that it independent of the truth of *A* itself.

Stochastic independence, the most studied condition on joint probabilities, is in fact ruled out by Centring. Not so Counterfactual independence, which is of interest primarily because it is the counterpart in the multi-dimensional possible worlds space of McGee's Independence condition on sentences. Together Centring and Counterfactual independence are sufficient for the joint probabilities of world-pairs to be completely determined by the marginal measures *p* and *q*. For instance in our four-world example, the relevant uncertainties would now be as follows.

Possible Worlds	*Supposed A-Worlds*	
	W1	**W2**
W1	1.p(w1)	0
W2	0	1.p(w2)
W3	q(w1).p(w3)	q(w2).p(w3)
W4	q(w1).p(w4)	q(w2).p(w4)

A further consequence of the assumption of Counterfactual independence is that uncertainty regarding counter-actuality is reduced to ordinary conditional uncertainty about actuality. To see this, note that it follows from the fact that *q* is a marginal probability that:

$$q(w_1) = 1.p(w_1) + 0.p(w_2) + q(w_1).p(w_3) + q(w_1).p(w_4) = \frac{p(w_1)}{p(w_1) + p(w_2)} = p(w_1 \mid A)$$

$$q(w_2) = 1.p(w_2) + q(w_1).p(w_3) + q(w_1).p(w_4) = \frac{p(w_2)}{p(w_1) + p(w_2)} = p(w_2 \mid A)$$

Hence $q = p(. \mid \text{A})$. Adams' Thesis is in fact an immediate consequence of this reduction, for $pr(A \rightarrow B) = q(\text{w1}) = P(\text{B} \mid \text{A})$.

Counterfactual independence is subject to the same counter-examples as McGee's Independence condition (which it implies). Not so the third and weakest of the independence conditions, Restricted independence, which is the counterpart in the possible worlds space of our eponymous condition on sentences. Unlike Counterfactual Independence, Restricted Independence is a necessary condition for Adams' Thesis to hold; like the former it is sufficient for it do so, given the assumptions built into our framework. This follows directly from the fact that:

$$pr(A \to B) = q(w_1) = pr(w_1 \mid A) = \frac{p(w_1)}{P(A)} = p(w_1 \mid A)$$

Restricted independence thus gives the best characterisation of the dependence of suppositional uncertainty on factual uncertainty validating Adams' Thesis.

Under what conditions should one expect the Restricted Independence condition to hold? Answer: whenever, the conditional under consideration is an indicative conditional. This follows from the fact that the probabilistic Ramsey test is built into our model together with the usual assumption that when dealing with indicative conditionals the relevant mode of supposition is evidential. For in this case, $q(.)$ is just $p(.|\mathrm{A})$. Given this, it is an immediate consequence of the treatment of semantic content and uncertainty presented here that the probability of a simple indicative conditional is indeed just the conditional probability of its consequent, given the truth of its antecedent. The reconciliation of Adams' Thesis with possible-worlds semantics is thus secured.

6. Concluding Remarks

The crucial modification to standard theory proposed here is the representation of semantic content by ordered sets of possible worlds rather than just sets of worlds. Up to this point we have dealt only with a very simple example involving just one supposition, but the treatment can be generalised in a straightforward way to all possible factual suppositions. The further generalisation to conditional suppositions is more complicated and will not be addressed here.

First we need to introduce the idea of a suppositional space, which is just a multidimensional possible-worlds space. We define a suppositional space S formally as follows. Let $\wp(W) = \{X_i\}$ be the set of all subsets of W. Each member of this set, X_i, is a possible factual supposition. S is now the space spanned by all the possible suppositions i.e. $S = W \times X_1 \times \ldots \times X_n$. Any element of S is a vector w = $<w_0, w_{X1}, \ldots, w_{Xn}>$ of worlds with w_0 being a possible actual world and each w_{Xi} being a possible counter-actual world under the supposition that X_i. For any vector w and co-ordinate w_{Xi} corresponding to the supposition that X_i, let $w|X_i$ denote the vector $<w_{Xi}, w_{X1}, \ldots, w_{Xn}>$ obtained by substitution of w_{Xi} for w_0 in the first co-ordinate.

With this modification to the possible-worlds framework in place, we can state corresponding versions of the four propositions characterising the orthodoxy.

Semantics: An interpretation of a language L is a mapping v from pairs of sentences and ordered sets of possible worlds to semantic values satisfying the conditions of (1a*) Bivalence and (1b*) Boolean Compounding, and such that if A = $\{w_0 \in W: v_w(A) = 1\}$ then:

(1c*) Conditionals: $v_w(A \rightarrow B) = v_{w|A}(B)$

Pragmatics: Let *pr* be a joint probability mass function on the set of world vectors, with *pr*(w) measuring the probability that w_0 is the actual world and w_i the counter-actual X_i world. Then rational degrees of belief in L-sentences are measured by a probability function Pr such that for all L-sentences *A*:

(2*) $\Pr(A) = \mathrm{E}(v(A)) = \sum_{w \in S} v_w(A).pr(w)$

Logic: Let [A] denote the content of the sentence *A*, i.e. the sets of world vectors making it true. Then:

(3*) $A \vdash B$ iff $[A] \subseteq [B]$

Explanation: Let $\Pi = \{pr_j\}$ be the set of all permissible joint probability mass functions on the set of world vectors S and $V_L = \{v_i\}$ be the set of all permissible assignments of semantic values to sentence of L. A multidimensional possible-worlds model (MPW-model for short) of L is a structure $<v, pr>$ where v belongs to V_L and p to Π. Then

(4*) Every pair < v, pr > in $V_L \times \Pi$ is an MPW-model.

The permissibility of an assignment of semantic values is determined by the semantic conditions 1a*, 1b* and 1c*. On the other hand, which joint probabilities are permissible depends on the constraints imposed on the relation between joint and marginal probabilities. And these in turn depend on the kind of supposition being modelled; in particular whether it is of the evidential or counterfactual variety. In the case of evidential supposition, I have argued that the condition of Restricted Independence is appropriate, a condition which, jointly with (1c*) and (2*), implies Adams' Thesis.

In the light of the similarity between the multidimensional possible worlds models and the orthodox uni-dimensional ones, one may wonder if the triviality results for Adams' Thesis do not still apply in the new modified framework. Constructing a triviality argument seems simple enough. Let any occurrence of the word 'possible world' in your favourite triviality result be replaced by the phrase 'vector of possible worlds' and you appear to get a triviality result for the theory presented here.

In fact this is not so, for one crucial assumption is no longer satisfied. Condition (4*), unlike the original condition (4), allows for restrictions on permissible belief measures above and beyond the requirement that they be probabilities. In particular, it allows for restrictions on the relation between the joint probabilities on the suppositional space (which determine beliefs in sentences) and marginal probabilities defined on spaces of possible worlds, such as those contained in the various independence principles canvassed above. As such they cannot be framed without the additional structure contained in suppositional spaces, and so cannot be stated within an orthodox possible world model. Crucially, amongst those belief measures ruled out by the independence constraints are just those required for the triviality results to go through. For example, given the Restricted Independence condition, there can be no joint probability pr on W×A with marginals p on W and $q = p(.|A)$ on A that does not satisfy the Preservation condition. This is because if $P(B) = 0$ then $Q(B) = P(B|A) = 0$. So the non-trivial accommodation of the Preservation condition, and indeed of Adams' Thesis, within a modified possible-worlds framework is assured.

References

Adams, E. W. (1975). The Logic of Conditionals. Dordrecht and Boston, Reidel.

Bradley, R. W. (2000) "A Preservation Condition for Conditionals", Analysis 60 (3) 219-22

Bradley, R. W. (2002) "Indicative Conditionals" Erkenntnis 56: 345-378

Döring, F. E. (1997). "The Ramsey Test and Conditional Semantics" Journal of Philosophical Logic **26** (4):359-376,

Döring, F. E. (1994). "Probabilities of Conditionals." The Philosophical Review(103): 689-700.

Edgington, D. 1995: "On Conditionals". Mind 104, pp. 235-329

Edgington, D. (1991) "Do Conditionals Have Truth-Conditions?" in Conditionals. F. Jackson (Ed). Oxford, Oxford University Press**:** 176-201.

Eels, E. and B. Skyrms (Eds.) (1995). Probabilities and Conditionals: Belief Revision and Rational Decision. Cambridge, Cambridge University Press.

Gardenfors, P. (1988). "Knowledge in Flux: modelling the dynamics of epistemic states", Cambridge Mass., MIT Press

Gibbard,A . 1981: "Two Recent Theories of Conditionals" in Harper, Stalnaker and Pearce (eds.) "Ifs" Dordrecht: Reidel, 1981, pp. 211-47.

Hajek, A. (1989). "Probabilities of Conditionals - Revisited." Journal of Philosophical Logic **18**: 423-428.

Jeffrey, R. (1964). "If (abstract)." Journal of Philosophy **61**: 702-703.

Jeffrey, R. C. (1991) "Matter-of-fact Conditionals" Aristotelian Society Supplementary Volume **65**: 161-183

Jeffrey, R. C. and R. Stalnaker (1994) "Conditionals as Random Variables." in Probabilities and Conditionals. E. Eels and B. Skyrms (Eds).

Joyce, J. (1999) "The Foundations of Causal Decision Theory", Cambridge, Cambridge University Press.

Kaufmann, S. (2005) "Conditional predictions: A probabilistic account" Linguistics and Philosophy 28(2):181-231

Lewis, D. K. (1976). "Probabilities of Conditionals and Conditional Probabilities." Philosophical Review **85**: 297-315.

McGee, V. (1989). “Conditional Probabilities and Compounds of Conditionals.” Philosophical Review **48**: 485-541.

McDermott, Michael 1996: "On the Truth Conditions of Certain ‘If’-Sentences". Philosophical Review 105, pp. 1-37

Milne, Peter 1997: "Bruno de Finetti and the Logic of Conditional Events". British Journal for the Philosophy of Science, 48, pp. 195-232

Ramsey, F (1926) “Law and Causality” in D. Mellor (Ed.) (1990). Ramsey: Philosophical Papers. Cambridge, Cambridge University Press.

Skyrms, B. (1981). “The Prior Propensity Account of Subjunctive Conditionals”, in Harper et al, eds. “Ifs: conditionals, belief, decision, chance and time”, Dordrecht, Reidel, 1981

Stalnaker, R. C. (1970). “Probability and Conditionals”, Philosophy of Science **37**: 64-80.

Stalnaker, R. C. (1968). “A Theory of Conditionals.” Studies in Logical Theory. American Philosophical Quarterly Monograph Series **2**: 98-112.