Introduction[1]

This is not the dissertation that in 1990 I set out to write. That one would have been called "Two Dogmas of Conditional Probability". Realizing that it was overly ambitious, I settled for "A Dogma of Conditional Probability" instead—until about three months ago, when I realized that it too was overly ambitious. What remains could be aptly titled "A Hypothesis About Conditional Probability". I decided that I had better submit it before it devolved to "A Modest Proposal Concerning Conditional Probability", or worse.

This is not to say that I have abandoned the original project. On the contrary, I've become more firmly committed to it, if anything. It strikes me that, by and large, conditional probability has been regarded by philosophers as a relatively uncontroversial and well understood aspect of probability theory. "Conditional probability is simply given by the usual ratio formula, and we know how to understand *that*." This encapsulates, I think, an attitude that is prevalent. And I think that it embodies two dogmas. The first is that 'the conditional probability of A, given B' is to be univocally analyzed by the ratio formula, $P(A|B) = P(A\&B)/P(B)$. The second is that there is no difficulty in interpreting that ratio; and if pressed for details, people will

---

[1] Large parts of this dissertation will represent my contribution to a paper of the same title, co-authored by Ned Hall, and forthcoming in Eells and Skyrms (eds.) (1993). They are §1 through §5, and §9, although they may receive some editing before publication. I am omitting certain parts of that paper that were written by Hall, but several ideas that I am including (for the sake of smoother exposition and discussion, and for their own interest) are due solely or partly to him; these will all be acknowledged in the main text, or in footnotes. I thank him for letting me use them here. My various other debts to him are also detailed in the main text, and in footnotes.

Everything that I have included in the dissertation, with the exception of §5.2, and part of one paragraph in §5.3.1 (and direct quotation), was written by me, although much of it has profited from editorial and stylistic suggestions from Ned Hall, David Lewis, and Bas van Fraassen. As I will indicate again in footnotes at the appropriate places, much of the terminology of §5.2, and part of one paragraph in §5.3.1, is taken from the joint paper, and was written almost entirely by Hall. Again, I thank him for letting me include them here.

offer interpretations such as 'the probability of A, upon the minimal revision of P to accommodate B', or 'the conditional betting odds on A, where the bet is conditional on B', as if these were unproblematic.

I will not discuss the second dogma any further here; and the first dogma will only get a brief airing at the very end of the dissertation. Rather, my concern will be with a hypothesis involving conditional probability that I think is closely connected with the dogmas—what I call 'the conditional construal of conditional probability'. Very roughly, it says that the probability of the conditional 'if A, then B' equals P(B|A). My main purposes are to hone this rough statement down to various precise versions of the Hypothesis, as I call it, and to argue that virtually none of them is tenable. (The versions are collected together in an appendix at the end for easy reference.)

The dissertation has ten chapters—or ten sections, I should perhaps say, since some of them are quite short. The broad outline is this: in the first five sections I set up the problem, and give a critical review of some of the relevant literature; in the subsequent four sections I present various negative results of my own, and draw some philosophical consequences from them; and in the final section I offer a positive proposal.

In slightly more detail: the conditional construal of conditional probability *somehow* equates conditional probability with the probability of a conditional—that much is clear. But characterizing this carefully requires some work. §1 attempts to do that, distinguishing four versions of the Hypothesis. The following four sections are largely an opinionated historical survey, tracing the motivations for and origins of the Hypothesis, and its fluctuating fortunes. By the end of §5, the first version has been shown to be refuted, and the second version moribund.

My own negative results against the Hypothesis begin in §6. I first generalize Lewis' so-called 'second triviality result', adding insult to injury as far as the second version is concerned. §7 refutes the third version of the Hypothesis, and casts serious

doubt on the fourth, or so I argue. I then imagine four ways in which the Hypothesis could be resurrected. In §8, I refute the first of these ways, and strengthen some old results; and in §9 I argue against the other three ways.

Thus I believe that by §10, essentially all salient versions of the Hypothesis have met their demise. This, I think, opens up interesting new avenues of research. I argue that philosophers have been in the grip of what I have called the first dogma of conditional probability, briefly summarizing a dissertation that might have been. I offer a positive proposal, which I call 'the ambiguity thesis': 'conditional probability' is ambiguous between the usual ratio formula, and the probability of a conditional. The demise of the Hypothesis shows that these two disambiguations cannot be identified.

1. What is the Hypothesis?

"Conditional probability is the probability of a conditional." This slogan, while catchy, leaves underspecified the exact content of the hypothesis that I wish to discuss. In fact, it is not easy to characterize this content, since my concern is not with some mathematical hypothesis (although to be sure, it does have mathematical import), but rather with a hypothesis that is supposed to be able to explain, among other things, certain features of ordinary language. We begin with the equation that captures the leading idea behind the *conditional construal of conditional probability:*

(0) $P(A \rightarrow B) = P(B|A)$,

where P is a probability function, $\rightarrow$ is an interpreted conditional connective, and P(B|A) is given by the usual ratio formula for conditional probability.[2] There are three types of variable in this formulation: the probability function, the $\rightarrow$, and the propositions to appear on either side of the $\rightarrow$. To make a genuine statement, we need to specify appropriate quantifiers, and domains of quantification (the '0' is supposed to be suggestive of the absence of these). Let's first make clear our quantification over propositions:

(CCCP) $P(A \rightarrow B) = P(B|A)$ for all A, B in the domain of P, with $P(A) > 0$.

---

[2] At the very end of this work, I will come to question the ratio formula's cogency as a univocal analysis of conditional probability; but until then I think that confusion will be avoided, and none created, if I follow convention in this respect.

Now let's quantify over the probability functions and the conditionals—not in all the ways that one might conceive of, but rather so as to distinguish four particularly important versions of the hypothesis:

Universal version: There is some $\rightarrow$ such that for all P, (CCCP) holds.

Belief function version: There is some $\rightarrow$ such that for all P that could represent a rational agent's system of beliefs, (CCCP) holds.

Universal tailoring version: For each P there is some $\rightarrow$ such that (CCCP) holds.

Belief function tailoring version: For each P that could represent a rational agent's system of beliefs, there is some $\rightarrow$ such that (CCCP) holds.

The first version has been refuted a number of times, as we will see, and I will offer in §7 a result that refutes the third. The second and fourth versions are, of course, still underspecified: who is the agent? A rational human being? An ideally rational being? To some extent it doesn't matter since, as I will show, the second version is untenable however it is construed, and I will argue that even the fourth version is in a precarious state (despite a related positive result of van Fraassen's that I will discuss in §7.4). Nevertheless, I will refine them further at the appropriate time.

In what follows, when I talk about 'the Hypothesis', I will often leave open exactly which version I mean, except when it matters.

2. Why care about the Hypothesis?

The idea that the probability of a conditional equals the corresponding conditional probability goes back at least to Jeffrey (1964), who thought that it could illuminate a theory of confirmation, and Ellis (1969), who assumed it when arguing that truth logic is a special case of probability logic. In a sense, it finds a precursor as far back even as de Finetti (1936), as we will see later. However, its best known presentation is due to Stalnaker (1970). Thus, it came to be known as 'Stalnaker's Hypothesis', at once an unhappy and a happy name. Unhappy, because Stalnaker himself no longer holds the Hypothesis, and indeed has provided cogent arguments against it, which I will discuss; happy, because the Hypothesis has been a rich source of philosophical debate, and Stalnaker should take much of the credit for this. Nowadays it is more Adams' name that we should associate with it, since a certain variant of it—the so-called 'Adams' Thesis', which I will also discuss—is still held by him, and has much currency.

Both Stalnaker and Adams (1965 and 1975) had similar motivations: they intended the Hypothesis to illuminate the semantics of the conditional. Stalnaker's idea was roughly that the Hypothesis would serve as a criterion of adequacy for a truth-conditional account of the conditional. Stalnaker had already provided such an account (1968), but Lewis was working at the same time on his (which was to appear in (1971), and more fully in (1973)), and—as Stalnaker knew—the two accounts disagreed. The Hypothesis provided Stalnaker with a forceful argument for his side of the debate, since as we will see, it entails (a version of) conditional excluded middle, and acceptance of this is the main point of disagreement between the two accounts. In the 1970 paper, Stalnaker's strategy is this. Rather than looking directly for the truth conditions of the conditional, he considers its belief conditions—that is, conditions under which it would be reasonable for a rational agent to believe a given conditional. As is common

practice, he models the belief system of a rational agent as an assignment P of subjective probabilities to worlds in some probability space; the agent believes a conditional A→B exactly when his or her subjective probability for it is sufficiently high. Now the Hypothesis enters, identifying the subjective probability for A→B with P(B|A), which in turn is straightforwardly determined by the agent's assignment of probabilities to the unproblematic propositions A and A&B, provided that P(A) > 0.[3] So the agent believes the conditional just in case P(B|A) is sufficiently high. While such belief conditions for the →do not fix its truth conditions, they do put constraints on them—sufficient to rule against the Lewis account, for example.

Throughout a good part of the 60's and 70's, Adams was engaged in the project of supplementing the traditional truth-conditional notion of validity of arguments with a 'probabilistic soundness' criterion of his own—it finds its fullest treatment in Adams (1975), upon which I will focus my discussion. Adams believed that conditionals do not have truth conditions, and thus he found the traditional approach inadequate for arguments in which conditionals appear; but he was happy to speak of 'probabilities' attaching to conditionals. Roughly, a probabilistically sound argument is one for which it is impossible for the premises to be probable while the conclusion is improbable. He invoked the Hypothesis to govern the assignment of probabilities to indicative conditionals, and argued that the resulting scheme respected intuitions about which inferences were reasonable, and which not. While Adams appealed to the same hypothesis (at least superficially) as Stalnaker did, unlike Stalnaker he did not regard the probability of a conditional as probability of its *truth*. Further, Adams'

---

[3] Stalnaker extends the classical probability calculus to one in which conditional probabilities are primitive, so that conditional probabilities are defined even when the condition has probability zero; however, this falls outside the scope of my (CCCP).

Strictly speaking, we should distinguish propositions from the sentences that express them, and make clear whether we are thinking of probability as attaching to the former or to the latter. I will not fuss about this distinction when nothing turns on it, and I take this to be common practice.

'probabilities' of conditionals do not conform to the usual probability calculus—thus the suggestion by Lewis (1976) that they be called "assertabilities" instead, a practice that has been widely adopted subsequently. So the left hand side of (CCCP) is perhaps best read as "the assertability of B, if A" on the Adams view. This conditional assertibility then goes by P(B|A).

Thus it seems that Stalnaker proposed the belief function version of the Hypothesis, and that Adams proposed a near variant of it. Call a probability function P that conforms to (CCCP) a *CCCP-function for* →(although I will sometimes drop the reference to the →, when it is clear which one I mean). Stalnaker's original proposal, then, was that there exists an → such that

1. → can plausibly be construed as a conditional;

2. the set of CCCP-functions for this →includes all probability functions that could represent a rational agent's system of beliefs (for short: all *belief functions*).

Indeed, he believed that the Stalnaker conditional was just such an →.

This proposal has since been dashed by an impressive battery of results that count against the truth of the Hypothesis—although not decisively against all its versions. I will provide a survey of these results here. I will also explore in §9.3 the extent to which Adams' variant of the Hypothesis is immune to them.

The situation is interesting whatever the significance of the results turns out to be. If the Hypothesis (on any of its versions) is false, then seemingly synonymous locutions are not in fact synonymous: surprisingly, "the probability of B, given A" does not mean the same thing as "the probability of: B if A". If the Hypothesis (on any of its versions) is true, then it establishes important links between logic and probability theory, as Stalnaker and Adams hoped it would. As Stalnaker (1970) observes, "although the interpretation of probability is controversial, the abstract calculus is a relatively well defined and well established mathematical theory. In contrast to this, there is little agreement about the logic of conditional sentences... Probability theory could be a

source of insight into the formal structure of conditional sentences" (p. 107[4]). For example, the material conditional does not conform to (CCCP), as we will see (at least when the 'P' in the equation is genuine probability, rather than assertability); this in turn legislates against the Hypothesis applying to the indicative conditional, according to one major account of that conditional (favored by Jackson (1987) and Lewis (1976) among others). Or again, the Hypothesis entails (a probabilistic version of) conditional excluded middle; this in turn legislates against Lewis' counterfactual logic, and in favor of Stalnaker's.

And the Hypothesis could similarly enrich probability theory, since it would assist in the interpretation of conditional probability. Recall de Finetti's (1972) lament that the usual ratio gives the formula, but not the meaning, of conditional probability. The Hypothesis could serve to characterize more fully what the ratio means, and what its use is.

Finally, the Hypothesis would solve what van Fraassen (1989) calls "the Judy Benjamin problem", a problem in probability kinematics.[5] The general problem for probability kinematics is: given a prior probability function P, and the imposition of some constraint on the posterior probability function, what should this posterior be? This problem has a unique solution for certain constraints—for example:

1.   Assign probability 1 to some proposition E (while preserving the odds of all propositions that imply E). (Solution: conditionalize P on E.)

2.   Assign probabilities $p_1, ... p_n$ to the cells of the partition $\{E_1, ... , E_n\}$ (while preserving the odds of all propositions within each cell).   (Solution: Jeffrey conditionalize P on this partition, according to the specification.[6])

---

[4] Page references are to Harper et al. (1981) for all articles that are cited as appearing there.

[5] I owe this point to Ned Hall.

[6] Jeffrey conditioning is first introduced in Jeffrey (1965) under the name "probability kinematics".

But consider the constraint:

3. Assign conditional probability p to B, given A.

The Judy Benjamin problem is that of finding a rule for transforming a prior, subject to this third constraint.

Van Fraassen provides arguments for three distinct such rules, and surmises that this raises the possibility that such uniqueness results "will not extend to more broadly applicable rules in general probability kinematics. In that case rationality will not dictate epistemic procedure even when we decide that it shall be rule governed" (1989, p. 343). But if the Hypothesis were true, a particularly simple solution would present itself. After all, constraint 3 would then be equivalent to:

3'. Assign probability p to $P(A \rightarrow B)$,

and this is uniquely met by a simple Jeffrey conditioning, on the partition $\{A \rightarrow B, \neg(A \rightarrow B)\}$ (assuming that the odds of propositions within each cell are to remain the same).

One conclusion of this dissertation will be that the first three versions of the Hypothesis are not tenable. Nevertheless, it is noteworthy that it is surprisingly hard to produce an intuitive counterexample to (CCCP). This suggests that we should consider the fourth version, weak though it is. And yet even this can only be held at considerable cost.

3. Why believe the Hypothesis?

Before arguing against the Hypothesis, I want to make the case in favor of it as strong as possible. Here are some reasons for believing it.

It sounds right

Both sides of (CCCP) seem to 'say' the same thing. As van Fraassen (1976) writes, "The English statement of a conditional probability sounds exactly like that of the probability of a conditional. What is the probability that I throw a six if I throw an even number, if not the probability that: if I throw an even number, it will be a six?" (pp. 272-3). Many assertions of conditional probability sound like assertions *of probability*, within the scope of which is something that seems best analyzed as a conditional. Thus, the Hypothesis certainly does justice to the grammar of conditional probability statements. And case-by-case evidence, such as van Fraassen's example here, seems to support it. This surely explains the initial appeal of the Hypothesis.

Ramsey's test

Ramsey (1965) suggests that you evaluate the conditional 'if A, then B' as follows: first, hypothetically add A to your system of beliefs, minimally revising what you currently believe in order to do so; second, evaluate B on the basis of your revised body of beliefs. $P(A \rightarrow B)$ measures how well the conditional performs on Ramsey's test. But apparently $P(B|A)$ does too. For conditioning on A *prima facie* seems to capture the notion of 'minimally revising what you currently believe in order to accommodate A'; and your evaluation of B in your new belief state $P(-|A)$ is just $P(B|A)$.

Adams' Thesis[7]

Assertability is said to go by subjective probability of truth. (Or at least it usually does, though not always: for example, the assertability of 'A but B' differs from that of 'A and B', yet the two have the same subjective probability, namely P(A&B).) At first sight, the indicative conditional appears to provide a counterexample to this dictum. After all, according to Adams' Thesis, the assertability of A→B equals P(B|A)—and even someone who, unlike Adams, believes that the indicative conditional has truth conditions, may find Adams' Thesis about its assertability compelling. (Of course, someone *like* Adams thinks that the indicative conditional provides a counterexample to the dictum, since according to him, there is no such thing as the probability of its *truth*.) It would be nice for proponents of the truth-conditional view of the indicative conditional if it were no exception to the dictum. If the Hypothesis were true, it would then explain Adams' Thesis admirably: the assertability of A→B goes by P(A→B) (as per the dictum), which in turn equals P(B|A). We infer the truth of the Hypothesis from Adams' Thesis since the Hypothesis provides a good—perhaps the best—explanation of it.

Stalnaker validity

Stalnaker (1970) deployed the Hypothesis in support of his C2 logic for the conditional by proving the coincidence of two sorts of validity: the sentences that are valid under his truth conditions for the conditional turn out to be exactly those which receive conditional probability of one, given every condition, assuming the truth of (CCCP).

---

[7] This is the motivation for the hypothesis discussed in Lewis (1976) and (1986).

Adams' probabilistic soundness

As I indicated earlier, Adams found that his variant of the Hypothesis, coupled with his probabilistic soundness criterion for arguments, gives verdicts that accord with our intuitions on the acceptability of arguments. In particular, it classifies as fallacious the various notorious 'paradoxes of material implication', such as the inference from 'not-A' to 'if A, then B'.

Independence of the conditional from its antecedent

(CCCP) is equivalent to A's being probabilistically independent of A→B according to P, for all A and B, under a plausible enough assumption about the logic of the →. The assumption, common to Stalnaker's and Lewis' logics, is that (A→B) & A is equivalent to A&B —a combination of modus ponens, and the principle that A&B implies A→B:[8]

A & (A→B) = A & B.[9]

It follows from this that

P(A & (A→B)) = P(A & B),

and thus

---

[8] The assumption is plausible enough, though not uncontroversial. McGee (1985) disputes the unrestricted use of modus ponens; and a number of authors question the principle that A&B implies A→B. Here is my reason for being suspicious about the latter principle, at least for counterfactuals (although I admit that intuitions vary here). Consider some indeterministic process that is certain to occur—as it might be, a particular tossing of a fair coin (or if you doubt that coin tosses are genuinely indeterministic, a simulation of a coin toss by a quantum randomizer). What is the probability of the counterfactual 'if the coin were tossed, it would lands heads'? *Zero*, I am inclined to say. After all, probability is probability of truth, and the conditional seems to be certainly false, since the following 'might' counterfactual seems to be certainly true: 'if the coin were tossed, it might land tails'. But the principle of which I am suspicious requires the probability of the 'would' counterfactual to be at least as great as the probability of the conjunction: coin is tossed & coin lands heads. And since the coin is certain to be tossed, that probability is surely 1/2.

[9] The '=' is the identity sign between terms that denote propositions.

$$\frac{P(A \& (A \varnothing B))}{P(A)} = \frac{P(A \& B)}{P(A)},$$

that is,

$$P(A \rightarrow B | A) = P(B | A).$$

So given the assumption about the logic, (CCCP) hinges on whether or not $P(A \rightarrow B | A)$ = $P(A \rightarrow B)$. Furthermore, it might seem that a conditional *is* always independent of its antecedent. *Prima facie*, this is plausible at least for the counterfactual conditional. After all, whether or not A is actually true might seem to be quite irrelevant to what *would be* the case *if* A *were* true.

It would be nice if the Hypothesis were true. Those with a pragmatist bent might think that, lacking reasons to the contrary, this is in itself a reason to think that it *is* true. But there are reasons to the contrary.

4. Why disbelieve the Hypothesis? Part I
Sources of suspicion

There were bad omens for the Hypothesis—some more serious than others—even before Lewis started the industry of 'triviality results' against it. In this section, I will consider some intuitive reasons to be suspicious about the Hypothesis. They fall short of precise refutations, but set the stage for them.

The Hypothesis fails for the material conditional

The probability of the material conditional,

$A \supset B = \neg A \vee (A\&B)$

is given by

$P(A \supset B) = P(\neg A) \times 1 + P(A) \times P(B|A)$,

which, as Jeffrey (1991) points out, is a weighted average of $P(B|A)$ and 1, with respective weights $P(A)$ and $P(\neg A)$. So provided $P(A) \neq 0$,

$P(A \supset B) = P(B|A)$ iff $P(A) = 1$ or $P(B|A) = 1$,

and the latter condition obtains only in trivial cases.

This need not alarm a proponent of the Hypothesis who either does not think the material conditional captures any important conditional of natural language, or takes the conditional appearing in the statement of the Hypothesis to be some other one. However, as I have noted, the Hypothesis clearly cannot apply non-trivially to the indicative conditional, on one major account of that conditional. That in turn undermines to some extent one of the reasons for believing the Hypothesis: the nice explanation that it would provide of Adams' Thesis about the indicative conditional.

Independence of the conditional from its antecedent again: causal decision theory

Stalnaker's own suspicions about the Hypothesis arose from considerations of situations in which an outcome, or act, is evidentially relevant to, or *stochastically dependent on,* another outcome without *causally influencing* that outcome. Stalnaker (1976) draws this distinction as follows:

A is stochastically dependent on B iff $P(B|A) \neq P(B)$; whereas

A is causally independent of B iff $P(A \rightarrow B) = P(B)$.

If there really are cases such as Stalnaker imagines, then we immediately have the failure of (CCCP), at least for the $\rightarrow$ in terms of which causal independence is defined.[10]   The conviction that there are such cases—Newcomb's problem, for instance—led to the development of various causal decision theories.[11]   Of course, these are cases of conditionals which are *not* independent of their antecedents, contrary to the intuition expressed in the previous section. For instance, the truth value of

I choose both boxes $\rightarrow$ I get just the \$1000

in Newcomb's problem is thought to depend on whether or not I choose both boxes.

Conditional excluded middle

The distinctive feature of Stalnaker's logic for conditionals is its adoption of conditional excluded middle as an axiom:

---

[10] Stalnaker believes that this $\rightarrow$ is in fact the Stalnaker conditional. There is really no issue for him of *which* conditional is involved here, since it is part of his position that his is the only one—although to be sure he does regard the indicative/subjunctive distinction of grammar as playing a role in determining a certain contextually determined parameter, the selection function, which in turn partly determines the conditional's truth conditions.

[11] Newcomb's problem, in which the outcome of an agent's choice depends on what a certain forecaster of her behavior predicted that choice to be, was introduced in Nozick (1969). For an exposition of various versions of causal decision theory, see Lewis (1981a). What unifies them, in Lewis' view, is roughly that they prescribe acting so as to maximize expectation over a partition of *dependency hypotheses*—maximally specific propositions about how outcomes do and do not causally depend on the agent's actions.

(CEM)   (A→B) v (A→¬B)

It is still controversial whether this feature is a virtue or a vice. (CCCP) implies something that is tantamount to (CEM). To see that this is the case, assume (CCCP), and conditional non-contradiction (that is, (A→B) & (A→¬B) is a contradiction, where A is possible—a principle common to both the Stalnaker and Lewis logics). Then we have

P(A→B) = P(B|A)

and

P(A→¬B) = P(¬B|A).

Adding these equations, we have

(1)   P(A→B) + P(A→¬B) = 1.

Now

P((A→B) v (A→¬B)) = P(A→B) + P(A→¬B),

by conditional non-contradiction

$$= 1 \text{ by (1).}$$

So we have P((A→B) v (A→¬B)) = 1, the probabilistic counterpart of (CEM).

Now we see that putative counterexamples to (CEM) will serve as putative counterexamples to (CCCP). See for instance Lewis' (1973, p. 80) critique of (CEM), using Quine's (1950, p. 14) famous example with A taken to be 'Bizet and Verdi are compatriots', and B taken to be 'Bizet and Verdi are Italian'. On Lewis' view, at least one of P(A→B) and P(A→¬B) must be less than (CCCP) predicts. Similarly, indeterminism may cast doubt on (CEM) (although intuitions vary here). Suppose coin tosses are genuinely indeterministic processes. Then both the conditionals 'if the coin were tossed, it would land heads' and 'if the coin were tossed, it would not land heads' appear to be false.

The bad omens become sharp arguments in the triviality results that follow.

5. Why disbelieve the Hypothesis? Part II
A selective survey of previous triviality results

5.1 Preliminaries

The slogan form of the Hypothesis with which I began gives rise to the misconception that it suffices to exhibit, for any $\rightarrow$, a single probability function and a pair of propositions in its domain for which the equation fails, in order to refute the Hypothesis. To be sure, that refutes the universal version of it; but doing *that* is easy and it leaves the weaker versions untouched. To see just how easy it is, note that (apart from special cases) the family of CCCP-functions for a given $\rightarrow$ is not closed under mixing.[12] Suppose that

$P_1(A \rightarrow B) = P_1(B|A)$

and

$P_2(A \rightarrow B) = P_2(B|A)$

and let $P_3 = \frac{1}{2} P_1 + \frac{1}{2} P_2$. Then in general

$P_3(A \rightarrow B) \neq P_3(B|A)$.[13]

Thus P_3 is generally not a CCCP-function for this $\rightarrow$.

---

[12] I am indebted here to Bas van Fraassen.

[13] Here's why:

$P_3(A \rightarrow B)$

$= \frac{1}{2} P_1(A \rightarrow B) + \frac{1}{2} P_2(A \rightarrow B)$

$= \frac{1}{2} \frac{P_1(A \ \& \ B)}{P_1(A)} + \frac{1}{2} \frac{P_2(A \ \& \ B)}{P_2(A)}$

$\neq \dfrac{\frac{1}{2}P_1(A \ \& \ B) + \frac{1}{2}P_2(A \ \& \ B)}{\frac{1}{2}P_1(A) + \frac{1}{2}P_2(A)}$

except in the special cases that $P_1(A) = P_2(A) \neq 0$, or $\frac{P_1(A \ \& \ B)}{P_2(A \ \& \ B)} = \frac{P_1(A)}{P_2(A)}$; but the final line is equal to $P_3(B|A)$.

Since refuting the universal version is so easy, I will give expositions only of results that attack also one of the weaker versions (with the exception of Stalnaker's result, the interest of which should be self-evident).

The discussion of each result will be divided into two sub-sections: technical and philosophical. The reader who is impatient with formalities may want to skip or skim the following sub-section, and the technical discussions, and to proceed quickly to the philosophical discussions.

5.2 Terminology[14]

Some preliminary comments are needed to facilitate the classification of some of the triviality results. It will help to describe probability functions in a more fine-grained manner. We therefore introduce the term *probability space* to denote a triple <W,F,P>, and the term *model* to denote a quadruple <W,F,P,→>; we will call <W,F,→> a *conditional space*. W is here understood to be a set (heuristically: a set of possible worlds); F is a σ-field of subsets of W (that is, it is closed under the Boolean operations, including countable union—heuristically: a set of significant propositions); P is a finitely additive probability measure on F; and → is a binary operator on F, defined for all A, B $\in$ F.

We will say that (CCCP) *holds for* a given model <W,F,P,→> iff for all A, B $\in$ F where P(A) > 0, P(A→B) = P(B|A). In that case, say that →is a *CCCP-conditional for* P, or else as before, that P is a *CCCP-function for* →. Call a function that is not a CCCP-function for → a *non-CCCP function for* →. If →is a CCCP-conditional for every probability function in some class of of probability functions, call → a *CCCP-*

---

conditional for the class. A model for which (CCCP) holds we will call a *CCCP-model*.

Call <W,F,P,→>, and also P itself, *trivial* if P has at most four conditional probability values.[15] It is easy to construct an → so that a trivial model <W,F,P,→> will be a CCCP-model.[16] Hence, each triviality result can now be seen as demonstrating some limitation on the class of CCCP-models that are not trivial—the name 'triviality' derives from the conclusion of these results, namely, that models obeying certain constraints are CCCP-models only if they are *trivial*.

The triviality results come in two varieties. The first kind, which we will call *no-go* results, demonstrate that, in a class of CCCP-models with such-and-such features, there are no non-trivial models. The second kind, which we will call *limitation* results, demonstrate that, in a class of CCCP-models with such-and-such features, the non-trivial models make up only a very restricted subset of all the models with those features.

As is customary, we will say that a probability function P_C is derived from P by *conditioning* if there is some proposition C such that $P_C(-) = P(-|C)$—that is, for all X ∈ F, $P_C(X) = P(X|C)$. We will say that a class of probability functions is *closed under conditioning* if any probability function derived from a function in the class by conditioning is itself in the class.

---

[15]It follows that P has at most four *un*conditional values. However, the reverse implication does not hold: if W contains just three worlds, each given weight $\frac{1}{3}$, then P has as unconditional values the set {0, $\frac{1}{3}$, $\frac{2}{3}$, 1}, but has as conditional values the set {0, $\frac{1}{3}$, $\frac{1}{2}$, $\frac{2}{3}$, 1}. As we shall see in §7, no model equipped with such a probability function can conform to (CCCP).

Of course, there are other models and probability functions that one could justly call "trivial". We will meet some of them in §8.

[16] Here's how: For all X,Y ∈ F, let X→Y = ∅ if X∩Y=∅; let X→Y = W if X∩Y≠∅ and X is a subset of Y; let X→Y = Y if X∩Y≠∅ and Y is a proper subset of X.

Finally, a proposition A ∈ F is a *P-atom* iff, for all X ∈ F, either P(AX) = 0 or P(AX) = P(A) > 0. (I suppress here, and at various points subsequently, the conjunction sign.)

My purpose in each of the ensuing technical sub-sections is mainly to clarify what it is that needs to be analyzed and interpreted. But I caution against supposing too quickly that any of these results decisively settles the question of whether conditional probability is the probability of some conditional—or, more generally, whether a study of the relationship between these two quantities might illuminate the semantics of conditionals. *These* questions require further philosophical analysis and interpretation, which I will attempt after each result is presented.

5.3 *Lewis' first three triviality results (1976 and 1986)*

5.3.1 Technical results

David Lewis has three closely related limitation results, the first two appearing in (1976), the third in (1986). They make no assumptions about the logic of the →. Lewis' results reveal the following:

First triviality result: There is no CCCP-conditional for the class of all probability functions.

Second triviality result: There is no CCCP-conditional for any class of probability functions closed under conditioning, unless the class consists entirely of trivial functions.

Third triviality result: There is no CCCP-conditional for any class of probability functions closed under conditioning restricted to the propositions in a single finite partition, unless the class consists entirely of trivial functions.

Lewis' results deserve careful scrutiny, particularly because they are so well known and widely cited. First of all, note that the third result implies the second, which in turn

implies the first; so from a logical point of view there is really only one result here. Actually, the proof of the second result really shows something more than Lewis claims (although it is slightly more cumbersome to state):

if $<W,F,P,\rightarrow>$ is a CCCP-model, then for any C, at least one of $<W,F,P_C,\rightarrow>$ and $<W,F,P_{\neg C},\rightarrow>$ isn't.

Likewise, the proof of the third result really shows:

if $<W,F,P,\rightarrow>$ is a CCCP-model, then for a partition $\{E_1, E_2, ...\}$ of evidence propositions, at least one of the $<W,F,P_{E_i},\rightarrow>$'s isn't.

(Lewis assumes that the partition is finite; it could be countably infinite if we assume countable additivity, although he doesn't.) More importantly, using an idea of Lewis', we can derive an even stronger result, which has his first three triviality results as corollaries.

The key maneuver in Lewis' proofs is to demonstrate the following: if $<W,F,P,\rightarrow>$ and $<W,F,P_C,\rightarrow>$ are both CCCP-models (where both employ the same $\rightarrow$), and P_C is derived from P by conditioning on C, then for all A, B $\in$ F such that $P(AC) \neq 0$, $P(A\rightarrow B|C) = P(B|AC)$. This is easy to show: $P(A\rightarrow B|C) = P_C(A\rightarrow B) = P_C(B|A) = P(B|AC)$. The equality of the second and third terms requires the uniform interpretation of the conditional; for suppose that the interpretations differed, so that the second CCCP-model was not $<W,F,P_C,\rightarrow>$ but rather $<W,F,P_C,\rightarrow'>$, where $\rightarrow' \neq \rightarrow$. Then we would have $P(A\rightarrow B|C) = P_C(A\rightarrow B)$ and $P_C(A \rightarrow' B) = P_C(B|A) = P(B|AC)$, but not necessarily the needed $P_C(A\rightarrow B) = P_C(A \rightarrow' B)$. But I will now show that uniformity of the $\rightarrow$ over just *these two* CCCP-models is sufficient to derive triviality[17] (and of course this falls far short of uniformity over an entire class of models, whose probability functions are closed under conditioning, or restricted conditioning). It was Ned Hall who first showed me a proof of the following theorem, so I will refer to it as

---

[17] This paragraph up to this point was written almost entirely by Ned Hall, and is taken directly from our joint paper.

'Hall's strengthening of Lewis' result'; however, I offer a simplification of Hall's original proof.

Theorem. If $\langle W,F,P,\rightarrow\rangle$ and $\langle W,F,P_C,\rightarrow\rangle$ are distinct non-trivial models, with P_C derived from P by conditioning on C, then at least one of them is not a CCCP-model.

Proof.[18] Suppose for *reductio* that $\langle W,F,P,\rightarrow\rangle$ and $\langle W,F,P_C,\rightarrow\rangle$ are distinct non-trivial CCCP-models. Their distinctness implies that P(C) < 1. Non-triviality of the latter guarantees that C is not a P-atom—for conditioning on a P-atom yields a probability function with just the values 0 and 1. Since C is not a P-atom, there is some D properly contained in C such that 0 < P(D) < P(C).

Let $E = D \cup \neg C$. We have

$$P(E\rightarrow\neg C) = P(E\rightarrow\neg C|C)P(C) + P(E\rightarrow\neg C|\neg C)P(\neg C)$$

$$= P_C(\neg C|E)P(C) + P(E\rightarrow\neg C|\neg C)P(\neg C), \text{ applying (CCCP) to } P_C$$

$$= P(\neg C|EC)P(C) + P(E\rightarrow\neg C|\neg C)P(\neg C) \text{ (thus completing Lewis' key maneuver)}$$

$$= 0 + P(E\rightarrow\neg C|\neg C)P(\neg C).$$

Applying (CCCP) to the left-hand side, we therefore have

$$P(\neg C|E) = P(E\rightarrow\neg C|\neg C)P(\neg C),$$

but since $\neg C \subset E$, and thus $\neg CE = \neg C$, this becomes

$$\frac{P(\neg C)}{P(E)} = P(E\rightarrow\neg C|\neg C)P(\neg C).$$

Multiplying both sides by $\frac{P(E)}{P(\neg C)}$ (as we can, since P(¬C) > 0), this yields

$$1 = P(E\rightarrow\neg C|\neg C)P(E),$$

which implies that P(E) = 1. But

---

[18] The first six lines of the proof (through "Let $E = D \cup \neg C$") come directly from Hall's original proof. Line 7 is reminiscent of the first step of Lewis' proof; it is my choice of $P(E\rightarrow\neg C)$ as the thing to be expanded there that simplifies Hall's original proof from then on.

$P(E) = P(\neg C) + P(D) < P(\neg C) + P(C) = 1,$

so $P(E) < 1$, completing the *reductio*.

Thus, the technique that Lewis employs to prove his first three triviality results can be easily adapted to establish something more telling: given any class of non-trivial CCCP-models employing the same $\rightarrow$, no probability function in the class is the conditionalization of any other.

5.3.2 Philosophical discussion

Context dependence

Lewis conclusively refutes the universal version of the Hypothesis here, and certainly scathes the belief function version. However, his results invite the response that there might be a CCCP-conditional for *some* and perhaps *most* of the members of the classes of probability spaces in question—enough, at any rate, for (CCCP) to retain its interest. In his 1976 paper, Lewis takes his first two results to be conclusive, giving the following argument (read "CCCP-conditional" for "probability conditional"):

> Even if there is a probability conditional for each probability function in a class, it does not follow that there is one probability conditional for the entire class. Different members of the class might require different interpretations of $\rightarrow$ to make the probabilities of conditionals and the conditional probabilities come out equal. But presumably our indicative conditional has a fixed interpretation, the same for speakers with different beliefs, and for one speaker before and after a change in his beliefs. Else how are disagreements about a conditional possible, or changes of mind? (p. 133)

Thus, Lewis believes that the conditional is not indexical to the beliefs of its utterer: he assumes that $\rightarrow$ means the same for speakers with different beliefs.

Lewis' assumption that the $\rightarrow$ has uniform interpretation, irrespective of the probability function in whose scope it appears, is dubbed "metaphysical realism" by Stalnaker (1976)—an allusion to the Lewis metaphysical framework, no doubt, although the assumption could be found plausible without it. Van Fraassen (1976) surmises that Lewis' framework—his realism about possible worlds, and about the similarity relations thereon—affords him an argument for the uniformity of $\rightarrow$. The

idea is that there is an objective fact about what the worlds are, and about what the similarity relations among them are, and that together these determine the truth values of all propositions. The rational agent's uncertainty about which world she inhabits is represented by her probability assignment to the worlds; but which worlds constitute a given proposition—even a proposition involving conditionals—is insensitive to this assignment.

I hasten to point out that van Fraassen is no realist about the Lewisian framework, preferring as he does to "locate all of reality in the actual world and the representing subject, seeing nothing but manipulable fictions in the possible world menagerie!" (p. 275). Note, though, that van Fraassen's formulation of the argument seems to assume that Lewis must regard $\rightarrow$ as depending on the constituents of his metaphysical framework; yet Lewis' interest in the Hypothesis really concerns the indicative conditional, which he believes to be simply the material conditional. Note also that van Fraassen is a little unfair to Lewis when he writes that his assumption of the uniformity of the $\rightarrow$ through a change of probability function is "justified only by metaphysics" (p. 275)—after all, Lewis' argument quoted above makes no appeal to his metaphysics, and could be found persuasive by someone who does not subscribe to that metaphysics.

It would suffice to evade Lewis' results if the $\rightarrow$ were not stable in interpretation throughout an entire class of probability functions closed under conditioning, or restricted conditioning—but this could still be achieved by an $\rightarrow$ that was uniform throughout *most* of such a class. This might prompt the hope that such a near-uniform $\rightarrow$ might do well as an approximation to a conditional of natural language, and be amenable to the Hypothesis at the same time. However, Hall's strengthening of Lewis' results rules against even this modest hope. Uniformity of the $\rightarrow$ across just a pair of probability functions, related by conditioning, is uniformity enough to give the stronger result a toehold.

Perhaps a more radical response is in order. Van Fraassen gives us such a response: "Would it not seem rather, that our probabilities are inextricably involved in the way we represent the possibilities, and nearness relations among them, to ourselves ... if our ideas about the one change, will we not revise our modelling of the other?" (p. 274). (Again, this is not relevant to the indicative conditional, according to Lewis, although to be sure it is relevant to various conditionals, such as the Lewis counterfactual.) It appears that on van Fraassen's view the conditional is so intimately tied to opinion that *any* change in opinion entails corresponding changes in which propositions are picked out by conditional utterances. We should pause to note just *how* radical this is. For example, you previously assigned probability 1/2 to the coin landing heads; but now that you have seen that it landed tails, you assign that proposition probability 0. Van Fraassen believes that this very fact implies a revision in your nearness relation, with consequent revisions in which proposition gets expressed by a given conditional.

Radical though it is, van Fraassen has a striking positive result that shows that such context dependence *will* ensure the unassailability of a certain form of the Hypothesis—not one of the versions that we have identified so far, but another 'tailoring' version, whose details we will see in §7.4. I should emphasize that van Fraassen does *not* commit himself to the truth of any of the versions of the Hypothesis that we have seen, and if anything plays the role of devil's advocate; still, he provides a clear statement of an important challenge to Lewis' triviality results (and indeed to several of the other ones), one which must be met if they are to retain any philosophical interest.

So let us examine Lewis' argument for the fixed interpretation of →, and van Fraassen's radical response to it. Certain indexicals might seem at first sight to provide a counterexample to Lewis' position that (spoken) disagreement between speakers with different beliefs is possible only if the sentences that they utter have a fixed interpretation. Take the case of the indexical 'here'. 'Here' often means something

different coming from your mouth to what it does coming from your friend's, because of your different beliefs about where you are located, and these beliefs often matter.[19] Nevertheless, you can disagree over a sentence that contains the word, as when your friend calls you from a distant town and says "it's boring here", and you beg to differ. Similarly for changes of mind. When you first visited that town, you too used to say "it's boring here"; since then, however, you have learned of its many riches.

But look again, and we see that indexicals provide no such counterexample to Lewis, and indeed it is more van Fraassen for whom they appear to be problematic. For disagreement between two people, over a sentence S with an indexical that varies between them, does not take the form simply of one person saying S, and the other saying not-S.[20] You express your disagreement with your friend, or with your former self, by saying "it's not boring *there*"—the indexical has to change appropriately when it's your (present self's) turn to speak (and incidentally, this is just as true for *agreement*, too). In order for you to be able to negate what your friend said, or what you used to say, when uttering a sentence with such an indexical in it, you can't simply utter the same sentence with a 'not' thrown in—unless both utterances are made at the same index, or close enough to it for the purposes of the context, in which case the indexical nature of the word doesn't surface.

This in turn suggests where we should look for an argument against radical context dependence of the conditional. If conditionals are indexical, with the speaker's probability function being the index, then we should expect to see a similar phenomenon—a suitable adjustment of what is said to take account of a change in index. But do we? Apparently not. Jenny says: "if Lyle comes, we'll need more

---

[19]   I say "different beliefs about where you are located", rather than just "different location" because of my intuition that is sympathetic to a 'speaker meaning'. Of course, given the vagueness of 'here', there are contexts in which you mean close enough to the same thing by it, and your different beliefs don't matter. Those are not the cases at issue.

[20]   Here I am indebted to David Lewis (in conversation).

gazpacho"; you can express your disagreement by simply negating her very words: "it's not true that if Lyle comes, we'll need more gazpacho." If the probability function of the speaker really plays a role in determining what was said, then it's surprising that you do not need to adjust Jenny's words, to make it clear that it is your probability function, and not hers, that is operative.

Without wanting to enter too deeply into stormy waters in the philosophy of language, let us for the moment recognize with Kaplan the distinction between a sentence's character, and its content (though much of what I will say will not hinge on this distinction). And let us suppose that there is *some* important aspect of understanding a sentence that involves its content, and not merely its character. In *that* sense of understanding, you do not understand Jenny's utterance of a sentence containing the word 'here' until you know her location—you know the character of the sentence, without knowing the content. The parallel, on van Fraassen's radically context dependent view of the conditional, should then be this: in *that* sense of understanding, you do not understand her utterance of a conditional until you know her probability function. (In fact, sometimes we may not even understand (in that sense) our *own* utterances of conditionals, since our own degrees of belief are not always known to us.) This I find implausible. You can know the content of her utterance (or your own) without knowing her (your) degree of belief in it, or in anything else.

And so it would appear to be in general: degrees of belief just don't seem to be relevant to what is said by utterances of conditionals. Furthermore, given the sensitivity of what is said to the degree of belief of the sayer, on this view, it is surprising how impoverished are the resources of ordinary English to indicate such degrees of belief. Compare again the case of 'here', where we have rich resources for guaranteeing that our interlocutors know exactly where we are.

You might think that our resources for conveying degrees of belief are rich enough after all, since they are to some extent telegraphed pragmatically—that typically, we

utter something only when our degree of belief in it is high. But this won't help. Firstly, high degrees of belief still come in many shades: fairly high, high, very high, certain—let alone exact numerical values—and pragmatics does not distinguish among these. Yet according to the van Fraassen picture, every small difference in degree of belief can make a difference in the proposition expressed.

More importantly, pragmatics telegraphs (to some extent) only a speaker's degrees of belief regarding the subject matter under discussion, and that falls far short of telegraphing his or her entire probability function. Yet the proposition expressed by the utterance of a conditional is supposed to depend on the probability function, and we don't yet know exactly how. So for all we know, to understand what you meant when you said "If the printer breaks down, I'm quitting philosophy", one has to know your opinions about quite irrelevant things, such as how many cane toads there are in Goondiwindi. And in that case, communication between two people would seem to be either unintelligible when it involves conditionals (because they do not know each other's probability functions), or else pointless (because they do).

Maybe the change in meaning of a given conditional as the probability function changes is so small that it doesn't matter—maybe the conditional is a function of opinion, but it approximates a constant function. Put another way: maybe we rarely know *exactly* what someone has said when uttering a conditional, because it is dependent on that person's probability function, just as van Fraassen says; but the dependence is slight, so whatever we take the probability function to be we will know well enough what was expressed. Maybe. However, until we are given a story about what the character (in Kaplan's sense) of conditional utterances is, we don't know just how the proposition expressed depends on the probability function of the speaker. The burden of proof is on the proponent of this view to show that the change in content, as a function of the degree of belief, is small.

This will be no easy task, partly because it is unclear just what 'small change of meaning' of a conditional consists in.[21] In this respect also, conditionals are unlike indexicals such as 'here'. It is easy to spell out what a small change in the meaning of 'here' amounts to: roughly, a small change in spatial location of the speaker. But what does a small change in the proposition picked out by a conditional amount to? A small change in which worlds constitute it? This makes no sense unless we have a measure on worlds, one which tells us when a set of worlds is 'small'. However, different measures will give different verdicts on 'smallness'. Which one is to be privileged? The Carnapian 'ur-distribution', or rational prior?[22] Not according to van Fraassen, who has expressed elsewhere (e.g. 1989, p. 120-125) his conviction that there is no such distribution—and this view is widely endorsed.

Van Fraassen's position also has curious consequences when it comes to reports of utterances.[23] Jenny says to Mike: "if Lyle comes, we'll need more gazpacho". Mike says: "Jenny told me that if Lyle comes, we'll need more gazpacho". Since Mike does not know exactly what Jenny's probability for that conditional is (let alone her entire probability function), and yet the content of the conditional is supposed to depend on that probability, we are not guaranteed that Mike's report is correct. But surely Mike's repetition of the exact conditional that Jenny said to him, prefaced by the words "Jenny told me that", is a true utterance if anything is.[24]

Finally, how can van Fraassen make sense of the following utterance of mine: "I still believe that very conditional that I asserted yesterday; in fact, I believe it more strongly now"?[25] A possible response, on his behalf: that sentence is to be paraphrased

---

[21] I am indebted at this point to Ned Hall.

[22] I thank David Lewis for bringing this suggestion to my attention.

[23] I owe this point to Jennifer Saul.

[24] This also suggests that conditionals express propositions: could anything but a proposition meaningfully follow the words "Jenny told me that..."?

[25] I owe this argument to Jennifer Saul, and thank David Lewis for suggesting the answer to the response that follows.

as "yesterday I believed a proposition, which then I picked out by uttering a certain conditional, but which isn't picked out by that conditional now; and now I believe that proposition more strongly." But that doesn't seem right—when asked to remind others what it was that I believed, I just repeat what I said yesterday.

Regularity

So much for problems with radical context dependence. Returning now to Hall's strengthening of Lewis' result, we have seen that the →cannot be uniform across even a pair of non-trivial CCCP-models, whose probability functions are related by conditioning. "Too bad for even that modicum of uniformity", says the friend of the Hypothesis who sees radical context dependence as a loophole. It should be clear by now that I do not find this a plausible response. But there is another response, one due to Appiah (1985). It is to insist that the probability function of a rational agent is *regular*—it assigns probability 0 only to the empty proposition—and to observe that the result of conditioning is always an irregular probability function. If he is correct, no two probability functions in the epistemic history of a rational agent are related by conditioning, so for all we know they may all conform to (CCCP) non-trivially. Indeed (CCCP) may yet hold, for a single →, throughout the class of all belief functions. Or so the argument goes.

Lewis (1986) moves a little swiftly in his discussion of this point, in my opinion. He writes: "It's one thing to say ... that an irregular probability function cannot represent a reasonable system of belief; it's something else to say that it cannot represent a system of belief at all. The latter is what you need if, despite my triviality results so far, you still say that [(CCCP)] holds throughout the class of all belief functions" (p. 585). That is, according to Lewis, anyone who accepts:

1. the first three triviality theorems;

and

2. (CCCP) holds throughout the class of all belief functions;

must deny

3. an irregular function can be a belief function.

But you *can* consistently accept all three of these—for example, if you see the situation as follows: "Any class closed under conditioning obviously contains many irregular functions. *Some* of these irregular functions are belief functions, (although to be sure, some of them are not). (CCCP) holds throughout the class of all belief functions; *a fortiori*, it holds for all the irregular belief functions. But what prevents (CCCP) from holding throughout a (non-trivial) class closed under conditioning, or restricted conditioning, is that such a class always contains some irregular functions that are not belief functions; and we may concede that (CCCP) fails to hold for *them*."

Setting that aside now, Lewis (1986) has a two-pronged reply to Appiah's argument. Firstly, he contends that at least the ideally rational agent *does* conditionalize herself into irregularity, since to update in any other way would be to render herself susceptible to a Dutch book, given that she never mistakes the evidence. Note that this is *not* to say that she must have her sensory apparatus in perfect working order. Rather, it is to say that there is always some proposition which encapsulates the total content of her experience—imperfect though that experience might be—and it is this that she conditionalizes on.

The second prong is:

5.4 Lewis' fourth triviality result (1986)

5.4.1 Technical result

This further limitation result, also free of assumptions about the logic of the →, has force for the agent who updates by Jeffrey conditioning, and who thus avoids irregularity. We will say that P_x is derived from P by *non-degenerate two-celled*

Jeffrey conditioning if there is a proposition C, and an x with $0 < x < P(\neg C)$, such that for all $B \in F$,

$$P_x(B) = P(B) + x[P(B|C) - P(B|\neg C)].$$

The restriction on x is there to guarantee that we do not simply have a case of conditioning, a degenerate case of Jeffrey conditioning. (Note that this formulation, which is the one that Lewis uses, is equivalent to

$$P_x(B) = \alpha P(B|C) + (1-\alpha)P(B|\neg C),$$

with $\alpha = P(C) + x$. This may be a more familiar schema for Jeffrey conditioning over a two-celled partition; α represents the new probability assigned to C, after a learning experience that makes it more probable by an amount x than it was previously.)

Lewis reveals that:

Fourth triviality result: There is no CCCP-conditional for any class of probability functions closed under non-degenerate two-celled Jeffrey conditioning, unless the class consists entirely of trivial functions.

However, Lewis again assumes that the $\rightarrow$ is interpreted uniformly throughout a class of CCCP-models. And pending a strengthening of this result analogous to Hall's strengthening of the earlier results, a small amount of non-uniformity of the $\rightarrow$ might save at least the spirit of the belief function version, if not the letter.

Such a strengthening is ready to hand—indeed, it was implicit in Lewis' work all along. His own proof really shows that if P is a certain 'parent' distribution which is a CCCP-function for $\rightarrow$, and P_1 and P_2 are two different distributions descended from P by non-degenerate two-celled Jeffrey conditioning, then at most one of P_1 and P_2 are CCCP-functions for $\rightarrow$. Let's see why.

Lewis assumes for *reductio* that P and all functions P_x are CCCP-functions for some $\rightarrow$, where P_x is as above. Each different value of x corresponds to a different Jeffrey conditioning. He goes on to derive a long equation, whose form is:

$$x.k_1 = k_2,$$

where k_1 and k_2 are constants. Given Lewis' assumptions, this equation must hold for a range of values of x. This implies that $k_1 = k_2 = 0$, from which he proves the triviality of P. But that conclusion follows even if x takes on just *two* different values, which correspond to just two different Jeffrey conditionings. So only one Jeffrey conditioning is compatible with the non-triviality of P, namely the one that arises from an increase in the probability of C of $x = k_2/k_1$. Among all the ways of Jeffrey conditioning a non-trivial CCCP-function on a two-celled partition, at most one can be a CCCP-function for the $\rightarrow$.

5.4.2 Philosophical discussion

This is obviously bad news for the belief function version of the Hypothesis, since it seems to imply that there is only one way that an agent, whose original (non-trivial) probability function is P_0, can change his mind while employing the same $\rightarrow$ (assuming for now that belief revision takes place by two-celled Jeffrey conditioning), *irrespective of his course of experience*. But of course, there are countless possible courses of experience that he could have, which should lead to countless possible changes of mind. It is an unacceptable consequence of the conformity to (CCCP) that all of these but one are ruled out *a priori*.

It might be tempting to try to revive a version of an objection that Lewis himself gives (1986) to his second triviality result, one which prompted his third result. In the proof of the second result, Lewis conditionalizes on a proposition, and on its negation. The objection is that not just any proposition can be an evidence proposition, and the class of belief functions should only be closed under conditioning on those that can. In particular, "a proposition that could be someone's total evidence must be, in certain respects, highly specific. But to the extent that a proposition is specific, its negation is unspecific." (pp. 582-3). So Lewis goes on in his third result to consider a certain finite partition of evidence propositions, and derives triviality from the adherence to (CCCP),

on the assumption that the class of belief functions is closed under conditioning on those propositions. But in the fourth result, it is a two-celled partition over which the Jeffrey conditioning takes place: a proposition and its negation. Now it might be tempting to object that Jeffrey conditioning should take place over a partition of evidence propositions, and a two-celled partition cannot be such.

Looking more closely, however, we see that Lewis is immune to this objection now. We will say that P' is derived from P by *n-celled Jeffrey conditioning* if there is a partition $\{E_1, E_2, ..., E_n\}$ of (evidence) propositions, such that for all $B \in F$,

$$P'(B) = P(B|E_1)P'(E_1) + P(B|E_2)P'(E_2) + ... + P(B|E_n)P'(E_n).$$

It turns out that we have as a corollary to Lewis' result:

There is no CCCP-conditional for any class of probability functions closed under n-celled Jeffrey conditioning, unless the class consists entirely of trivial functions, *for each n.*

The reason is this. As he shows us, any class S of CCCP-functions for a given $\rightarrow$, at least one of which is non-trivial, is not closed under two-celled Jeffrey conditioning; that means that there is a two-celled Jeffrey conditioning on a function in S that takes us out of S. But for each n, that two-celled Jeffrey conditioning can be mimicked by an n-celled Jeffrey conditioning: the first cell gets the same weight as it did in the two-celled case, and the other (n–1) cells are all given equal weight (and thus jointly play the role of the second cell in the two-celled case). So that must be an n-celled Jeffrey conditioning on a function in S that takes us out of S; hence S's probability functions are not closed under n-celled Jeffrey conditioning, for each n.

5.5 *Stalnaker (1976)*

5.5.1 Technical result

Robert Stalnaker's (1976) triviality result is a no-go theorem for models $<W,F,P,\rightarrow>$, where $\rightarrow$ is the Stalnaker conditional. Thus, it makes certain specific

assumptions about the logic of the →. Stalnaker assumes that → conforms to the following constraints, characteristic of his C2 logic, for all A, B, C ∈ F:

(i) $[(A \rightarrow B) \cap (A \rightarrow C)] \subseteq [A \rightarrow (B \cap C)]$

(ii) $[A \cap (A \rightarrow B)] = (A \cap B)$

(iii) $(A \rightarrow B) \cup (A \rightarrow \neg B) = W$

(iv) $((A \cup B) \rightarrow A)) \cup ((A \cup B) \rightarrow B) = W$

(v) $[(A \rightarrow B) \cap (B \rightarrow A) \cap (A \rightarrow C)] \subseteq (B \rightarrow C).$

(Of course, these could equally be presented as logical axioms—for example, we recognize (iii) as the set-theoretic analogue of the axiom of conditional excluded middle, $(A \rightarrow B) \lor (A \rightarrow \neg B).$)

Let **C₂** be the sub-class of the CCCP-models satisfying these constraints. We are now in a position to state Stalnaker's triviality result:

*There are no non-trivial **C₂** models.*

Here is a proof, somewhat different from Stalnaker's own, and more along the lines of Gibbard's (1981) reformulation of it, which I think is more intuitive:

Assume that → is the Stalnaker conditional. As is well known, we can interpret this as meaning that X→Y is true at world w iff the nearest X world to w is a Y world. Also, since the Stalnaker → obeys modus ponens and the principle that A&B implies A→B (see constraint (ii)), we recall from §3 that $P(X \rightarrow Y) = P(Y|X)$ iff X is probabilistically independent of X→Y; but this is the case iff

(*) $P(X \rightarrow Y | \neg X) = P(Y|X).$

Let P be any non-trivial probability function. I will show that P can't be a CCCP-function for this →, by constructing a conditional proposition whose probability according to P must differ from the corresponding conditional probability.

By the non-triviality of P, we can find three disjoint and jointly exhaustive propositions that are assigned non-zero probability by P (for otherwise P would have at most four conditional probability values). Call them A & B, A & ¬B, and ¬A.

Let C = A v (A → ¬B), and hence ¬C = ¬A & (A→B), by conditional excluded middle. (Here I follow Stalnaker's rather than Gibbard's choice of proposition.) Note that P(¬C) > 0. For suppose otherwise. Then by (ii), P(A→B) = P(A & B), so by (CCCP), P(A & B)/P(A) = P(A & B), contradicting our assumptions that P(A & B) > 0 and P(A) < 1.

I will show that C → (A & ¬B) is a proposition for which (CCCP) must fail: Suppose otherwise. Then, this proposition must conform to (*)—that is,

P(C → (A & ¬B)|¬C) = P(A & ¬B|C), (with both sides being defined) or equivalently,

$$\frac{P(\neg C \ \& \ [C \ \varnothing \ (A \ \& \ \neg B)])}{P(\neg C)} = P(A \ \& \ \neg B | C).$$

Now the right-hand side is defined since P(C) > 0, and greater than 0 since its numerator is

P(A & ¬B & C) = P(A & ¬B) > 0.

Thus, the left-hand side must also be greater than 0, which implies that

¬C & [C → (A & ¬B)] is not a contradiction.

So there is a ¬C world (call it w), whose nearest C world (call it x) is an A & ¬B world. Then we have A, ¬B and C all true at x: so x must also be w's nearest A&C world.

It follows from this that

(1)　　w is not an A→B world.

For if it were, its nearest A world (call it y) would be a B world—and since C is true throughout A, we would have A, B and C all true at y, whereby y would be w's nearest A & C world, contradicting x being so. (Clearly x and y must differ, since they disagree on the truth value of B.)

However, since w is a ¬C world, it follows from the definition of ¬C that

(2)　　w is an A→B world,

which contradicts (1). Q.E.D.

5.5.2 Philosophical discussion

Since Stalnaker's ingenious result does not speak to any of the versions of the Hypothesis that I have countenanced, I will keep my discussion of it brief. However, it is interesting in its own right; and ironic, in so far as it shows that Stalnaker's original hypothesis concerning his own $\rightarrow$ is not only false, but actually is not satisfied by any (non-trivial) probability function.

All of the strength of Stalnaker's assumptions is packed into the logic of the $\rightarrow$. (It is the only result that I will canvass that makes no assumptions about the probability functions, other than their non-triviality.) Among other principles, the C2 logic includes conditional excluded middle (constraint (iii) above), and weakened transitivity ((v) above). Lewis (1973) objects to the former, as we've seen; van Fraassen (1976 and 1981) objects to the latter. The adherent of the Hypothesis can respond to Stalnaker that one or more of the constraints simply fail to hold for the relevant $\rightarrow$.

Stalnaker's sure-fire counterexample to (CCCP) is seen to be quite complex when written out fully: [A v (A $\rightarrow$ ¬B)] $\rightarrow$ (A & ¬B). In particular, it involves embedded conditionals. Against it, it's plausible to say that such conditionals are never uttered in natural language, and that it is no disgrace to the Hypothesis that it does not cover such arcane cases. And while such nestings might make sense (even if they are never uttered), perhaps we should not be surprised that the Hypothesis cannot handle them, because conditionals can behave strangely in such situations. Indeed, McGee (1985) argues that even modus ponens fails for nested indicative conditionals.

6. Why disbelieve the Hypothesis? Part III
Generalizing Lewis' second triviality result

A probability revision rule typically takes as input an initial probability distribution and a proposition, and yields as output a (new) probability distribution. We will regard an agent's system of beliefs as being represented by a probability function, and a change in the belief system to accommodate certainty about the truth of some proposition as taking place according to some revision rule, such that the proposition gets probability one after the revision.

In his second triviality result, as we have seen, Lewis proves that there is no CCCP-conditional for any class of probability functions closed under conditioning (unless the class consists entirely of trivial functions). We can picture this in terms of inverted 'family trees' of probability functions. At the top of a given tree is a certain 'parent' probability function (the 'prior'); below it are all the 'children', functions descended from it by all possible conditionings; below them are all the 'grandchildren', functions descended from the children by all possible conditionings; and so on. (Of course, a sequence of conditionings can be mimicked by a shorter sequence of conditionings, so all the grandchildren also make appearances among the children; all the great-grandchildren also make appearances among the children and grandchildren; and so on.) We get different family trees, also related by conditioning, by starting with different parents. Lewis shows that for any →, any forest of family trees, all of which have conditioning as the revision rule relating parent and child, must contain a non-CCCP-function for that →(unless all the functions in the forest are trivial).

This is bad news for the Hypothesis, since it appears that a rational agent could have an epistemic history in which her opinion is eventually represented by a function P that violates (CCCP): her initial probability function is the ancestor at the top of P's

tree, and she repeatedly updates by conditioning until she reaches P. (Again, she really only needs to conditionalize once, since P already appears among the children.)

However, there are other well-known revision rules, and hence other ways of generating interesting family trees. Let me mention three such types of rule here. *Imaging* on some proposition E is the rule which moves all the probability from each ¬E world to its nearest neighbor inside E—'nearest', as determined by the relevant similarity relation. (This assumes, as Stalnaker (1968) does, that there is always a unique such neighbor.) A *blurred* imaging on E removes all the probability from a ¬E world, but spreads it over more than one E world. Various ways of doing this have been proposed—see Lewis (1986a, p. 318) for more details and references. Finally, MAXENT is the rule which Skyrms (1987) explains as follows:

> Given an initial probability measure and a constraint on possible final probability measures one moves to a final probability by the rule of MAXENT if one chooses from among the final probabilities which satisfy the constraint, the one which has minimum information (or equivalently) maximum entropy relative to the initial probability. (p. 225)

Defining 'entropy' for the general case requires some work; but to indicate the idea, if a finite number n of worlds receive positive probabilities $p_1, ..., p_n$, then the entropy of this distribution is $-\Sigma_i \, p_i \log(p_i)$, and the information is $\Sigma_i \, p_i \log(p_i)$. If the constraint on possible final probability measures is 'the probability of E equals 1', then we have the case of accommodating E by MAXENT; the distribution that is produced is uniform over E.

I will show here that similar results to Lewis' are available for imaging, various blurred imagings, MAXENT, and other familiar revision rules. Lewis' second result will fall out as a corollary, but my result will be more general than it is.

6.1. Fearlessness, conservativeness, moderation

Suppose we have an initial probability distribution P, and want to revise it in order to accommodate a proposition E. We use some revision rule to derive a distribution P_E. Call the rule *fearless* if for any P and for any E, $P_E(E) = 1$. 'Fearless', for two reasons: the rule is prepared to take as input *any* initial probability function and proposition; and the function produced as output by the rule is not afraid to commit itself fully, giving probability 1 to the proposition (as it must, in order for the proposition to be genuinely accommodated). Fearless revision rules take us from some initial probability distribution to a new distribution that is fully concentrated on any proposition that we specify.[26]

Conditioning is clearly fearless; so is imaging, the varieties of blurred imaging, and MAXENT (provided the constraint on final probability measures is of the form 'the probability of E is 1'). Jeffrey conditioning is fearless, in so far as it does allow one to shift all probability onto a proper subset of a partition; however, those are not the distinctive cases of Jeffrey conditioning, wherein the probability function arrived at after the revision is regular. (That's why I do not claim to cover Lewis' fourth result as well, since that concerns closure under the sort of Jeffrey conditioning that preserves regularity.)

Furthermore, some revision rules are what I will call *conservative*. Such a rule takes a function P and a proposition E to a function P' with the following property: for any A that implies E, $P'(A) \geq P(A)$. A conservative rule never decreases the probabilities of propositions that imply the proposition that is accommodated.

----

[26] Those who insist that belief functions must always be regular will deny that belief revision ever takes place by a fearless rule (indeed, they deny that genuine accommodation of propositions ever happens). To them, I recommend instead my perturbation result of §8.2, or Lewis' fourth triviality result. Similarly, those who think that belief revision takes place by conditioning *restricted to the members of some finite partition* deny that it takes place by a fearless rule (since such restricted conditioning is not prepared to take just any proposition as input). To them, I recommend Lewis' third triviality result, or Hall's strengthening of it; see also footnotes 27 and 29.

Conditioning is conservative; so is imaging and various blurred imagings, assuming that the similarity relation obeys centering (each world is closer to itself than any other world is to it). Still more rules are what I will call *moderate:* for any A that implies E, if $P(A) > 0$, then $P'(A) > 0$. Moderate rules can decrease the probabilities of propositions that imply the proposition that is accommodated, but never all the way to 0. Moderation is surely a desideratum of a revision rule: your belief system would be fragile indeed if you could suddenly *fully* disbelieve something, to which previously you gave some credence, when nothing compelled such a radical change. All conservative rules are moderate. Moreover, so is MAXENT (at least in the finite case that I have defined above); and various blurred imagings, even when the similarity relation obeys merely weak centering (any world is one of the closest worlds to itself).

The result that I will prove shows that for any $\rightarrow$, the following is true: any forest of family trees, all of which have a particular fearless and moderate revision rule relating parent and child, must contain a non-CCCP-function for that $\rightarrow$. Again, scrutiny of the proof reveals that a non-CCCP-function appears already in the first set of offspring. Even if, *pace* Lewis, probability functions in an ideally rational epistemic history are not related by conditioning, a negative result analogous to his holds, provided that they are related by *some* particular rule that is fearless and moderate.

6.2 Generalizing Lewis' result

Here, then, is the promised limitation result that generalizes Lewis' second triviality result.

Theorem. There is no CCCP-conditional for any class of probability functions closed under a fearless[27] and moderate rule, unless the class consists entirely of trivial functions.

Proof.[28] Let $\mathfrak{R}$ be a fearless and moderate rule. Suppose for *reductio* that $\rightarrow$ is a CCCP-conditional for a class of probability functions closed under $\mathfrak{R}$, and that the class contains at least one non-trivial function. Take any non-trivial function P in the class: then there are propositions A and B such that P(A) < 1, P(A&B) > 0 and P(A & ¬B) > 0. There are two cases:

Case 1. (A&B) −(A→B) is non-empty
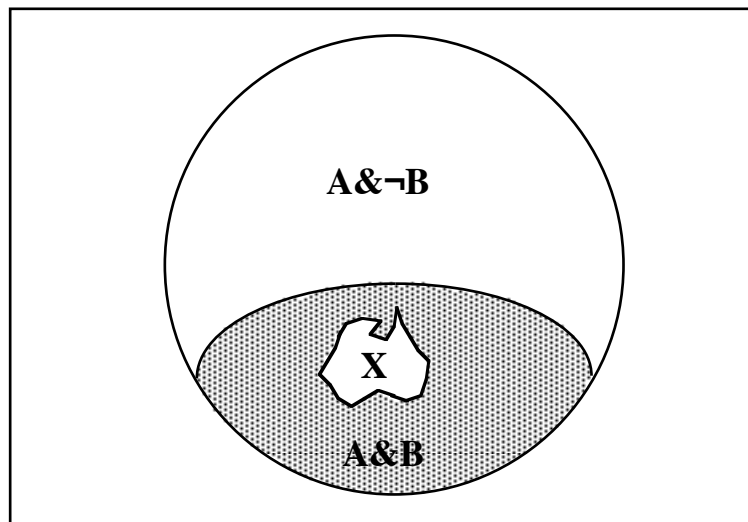
Let X = (A&B) −(A→B). Use $\mathfrak{R}$ to revise P to accommodate X, (as we know we can, by $\mathfrak{R}$'s fearlessness) thus producing P_X. Clearly $P_X(A\rightarrow B) = 0$, since we have

---

[27] I don't actually need the full strength of the assumption of fearlessness in the proof. All I need is that, for any non-trivial probability function P, and propositions A and B as described in the proof, the rule can revise P so as to give probability one to (A&B) −(A→B), or so as to give probability one to ¬(A&B). That is certainly weaker than fearlessness, but I don't know a neat way of characterizing it.

[28] I thank Ned Hall, David Lewis, Richard Jeffrey, Bas van Fraassen and Lyle Zynda for helping streamline the proof of the theorem.

accommodated X, and X is incompatible with A→B. But $P_X(B|A) = 1$, since $P_X(X) = 1$, and X implies A&B. So P_X violates (CCCP).

Case 2. (A&B) − (A→B) is empty
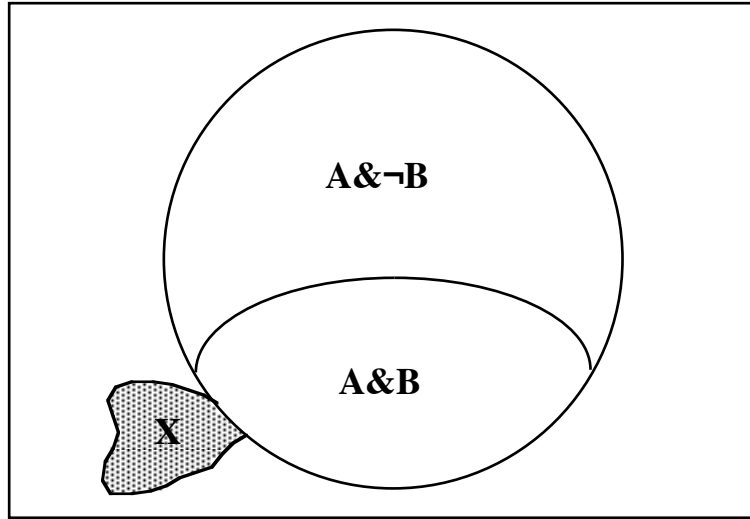
Let X = (A→B) − (A&B). P(X) > 0 (for otherwise P(A→B) = P(A&B), but then by (CCCP), P(B|A) = P(A&B), contradicting our assumption that 0 < P(A&B) and P(A) < 1). Use $\Re$ to revise P to accommodate ¬(A&B) (as we know we can, by $\Re$'s fearlessness), thus producing $P_{¬(A\&B)}$, which I will call P' for short. Since X implies ¬(A&B) and $\Re$ is moderate, P'(X) > 0, and so P'(A→B) > 0. But P'(B|A) = 0, since P'(A&B) = 0 (after all, we have used $\Re$ to fearlessly remove all probability from A&B), while P'(A) > 0 (since we assumed that P(A&¬B) > 0, and that $\Re$ is moderate), guaranteeing that the conditional probability does not go undefined. So P' violates (CCCP).

This exhausts the cases. Either way, we find that we can use $\Re$ to derive from P a non-CCCP-function for →. Thus, →is not a CCCP-conditional for a class of

probability functions closed under $\Re$, or else the class consists entirely of trivial functions. That completes the *reductio*.[29] Q.E.D.

You may have wondered why I didn't use the same trick in case 2 as I did in case 1: simply use $\Re$ to move all probability onto X (instead of onto ¬(A&B)). The reason is that we do not want the conditional probability to go undefined, as it would if the probability of A became 0. We have no guarantee that X is compatible with A— indeed, we have good reason to think that it is not, if the → is a conditional worth its salt.

---

[29] As Bas van Fraassen has pointed out to me, we could modify the proof so as to get a result which has Lewis' third triviality result as a corollary, provided we could assume that contained within X in case 1, and within ¬(A&B) in case 2, there is some member of a finite partition, upon which we can conditionalize.

7. Why disbelieve the Hypothesis? Part IV
Finite models

7.1 Preliminaries

Let us take stock. The first two versions of the Hypothesis that I distinguished at the outset are, I think, decisively refuted. (I hope that the previous section has brought around anyone who wasn't already convinced of this.) We've seen nothing so far, however, that impacts upon the third and fourth versions. Recall what they are:

Universal tailoring version: For each P there is some → such that (CCCP) holds.

Belief function tailoring version: For each P that could represent a rational agent's system of beliefs, there is some → such that (CCCP) holds.

In this section I will prove a result that refutes the universal tailoring version. (An earlier version of the result appeared in Hájek (1989), but the version that I will prove here is slightly stronger.) I will then argue that it also casts serious doubt on the belief function tailoring version. More precisely: I will distinguish two plausible readings of that version, the first in terms of an ideally rational agent's system of beliefs, the second in terms of a rational human being's system of beliefs. I will contend that my result casts serious doubt on both of them.

It is a no-go result that, like Lewis' results, assumes nothing about the logic of the →. In a sense, it lies at the opposite end of the spectrum from Stalnaker's result: it makes an assumption *only* about the individual probability functions, and it is the only result with that feature that I canvass in the dissertation. Some motivating remarks may be helpful.

Call a model <W,F,P,→> *finite* if the range of P is finite. The result is that no finite, non-trivial model is a CCCP-model.[30] The strategy in the proof is to show that,

---

[30] The original, slightly weaker result in Hájek (1989) amounts to this:

for any non-trivial finite model with probability function P, there are more distinct values of P(B|A) than there are distinct values of P(A→B), for all the possible substitution instances of A and B.

The proof begins with an obvious fact about the cardinalities of three sets. Let #(S) denote the cardinality of set S. Then

#(range of P(– →–)) ≤ #(range of P) ≤ #(range of P(–|–)).

After all, anything in the range of P(– →–) is in the range of P (for P assigns values to all conditionals formed from members of F, and to other things besides); and anything in the range of P is in the range of P(–|–) (for any value P(X) of the former appears as P(X|W) in the latter).

What is less obvious, but also true, is that the second inequality is strict, and it is here that the assumption of the finiteness of the range of P is crucial. The bulk of the proof consists in establishing this. The upshot is that for each non-trivial finite model, there is some conditional probability value that equals none of the unconditional probability values, and so *a fortiori* equals none of the P(– →–) values. Suppose that P(Y|Z) is such an unmatched value—that is, nothing of the form P(– → –) has the same value. Then *a fortiori* P(Z→Y) fails to match it, and we have a violation of (CCCP).

We may picture the situation colorfully as follows. Take any non-trivial probability function P with finite range. Imagine a dance, for which various men and women have tickets. Suppose that for each distinct value of P(–|–), there is exactly one man with that value written on his ticket, and that for each distinct value of P(– → –), there is exactly one woman with that value written on hers. There are no other men or women at the dance. It is a rule that for any couple that dances, the woman must have the same

---

There is no non-trivial CCCP-model <W,F,P,→> for which W is finite.
But my small amendment in §7.2 to the original proof will show that the finiteness of the range of P suffices to derive the result, and while this is implied by the finiteness of W, the converse is not true. Indeed, certain models with infinitely many worlds also fail to sustain (CCCP) for the reasons given in the new proof. I thank Ned Hall for helpful discussion on this point.

number on her ticket as her partner does. (I assume here that each couple consists of a woman and a man.) (CCCP) amounts to the promise that everyone has a partner to dance with. The proof establishes that this is not so—there is at least one 'unmatched' man who must remain a wall-flower. (The picture was inspired by a 'Waltz Night' at Princeton's Graduate College.)

As an example, suppose P assigns probability 1/3 to three different worlds, w_1, w_2 and w_3. All unconditional probability values are a multiple of 1/3, so *a fortiori* all (unconditional) probabilities of conditionals are multiples of 1/3. But $P(w_1| w_1 \vee w_2) = 1/2$. The poor man at this dance who has 1/2 written on his ticket must remain a wall-flower. (This shows just how easy it is to find a counterexample to the universal version of the Hypothesis.) Or consider a model of the result of tossing a fair die, in which probability 1/6 is assigned to six worlds (corresponding to the six possible results). P(6 shows up|some number other than 1 shows up) = (1/6)/(5/6) = 1/5. But all unconditional probability values are a multiple of 1/6, and 1/5 is not a multiple of 1/6, so the unconditional probability

P(some number other than 1 shows up $\rightarrow$ 6 shows up) $\neq$ 1/5.

Thus, there is a wall-flower with 1/5 written on his ticket. The proof of the theorem establishes that this is not due merely to some unfortunate feature of the P's chosen here—*all* non-trivial probability functions with finite range suffer the same plight.

Note that this result does not apply to models with only one or two worlds, and hence at most four distinct probability values—(CCCP) can hold for such models, albeit trivially. In the one world case, the unconditional and conditional probabilities are just 0 and 1, so these can be identified. In the two world case, with probability p_1 assigned to one world, and p_2 assigned to the other, the only distinct unconditional probabilities are 0, p_1, p_2, and 1, so these are the only possible values for the probabilities of conditionals. Likewise, these are the only distinct values that the conditional probabilities can assume (for anything of the form $\frac{P(A\&B)}{P(B)}$ must have

either the value 0, $\frac{p_1}{p_1 + p_2} = p_1$, $\frac{p_2}{p_1 + p_2} = p_2$, or 1), so the two sets of probabilities can be paired off with each other. These, however, are the only finite models for which (CCCP) can hold, (as I will show), and they are degenerate cases. Enough preamble; let us move on, then, to more interesting cases.

7.2 Technical result

Theorem. No finite, non-trivial model is a CCCP-model.

Proof. Let $\langle W,F,P,\rightarrow\rangle$ be a finite, non-trivial model. Since P takes only finitely many values in the interval $(0,1]$, we can let p_1 be the minimum such value. So no proposition receives a positive probability less than p_1. This means that there are at most $1/p_1$ P-atoms, a finite number. Furthermore, the probabilities of all the P-atoms must sum to 1. (For suppose they fall short of 1 by some positive amount. Then there is a proposition X with positive probability, and yet with no P-atoms contained within it. Since P takes only finitely many positive values on propositions contained within X, there must be a minimum such value, assigned to proposition A, say. Then A is a P-atom contained within X. Contradiction.) It may help to think of P as being an assignment of probabilities to a finite number of indivisible 'blobs'.

Let $p_1, p_2, \ldots, p_n$ be the probabilities assigned to these P-atoms. All of the p_i's are greater than zero, n is finite and $\Sigma p_i = 1$. Non-triviality ensures that $n \geq 3$. Assume, without loss of generality, that the probabilities are ordered so that

$p_1 \leq p_2 \leq \ldots \leq p_n$.

Notice firstly that there can be *at most* as many distinct values of probabilities of the form $P(A \rightarrow B)$ as there are distinct unconditional probability values—since each $P(A \rightarrow B)$ *is* an unconditional probability, for any A, B $\in$ F. Secondly, there must be *at least* as many distinct conditional probability values as there are distinct unconditional probability values—to any unconditional probability $P(X)$, there corresponds the

conditional probability P(X|W) that has the same value. I will show that in fact there are *more* distinct conditional probability values than distinct unconditional probability values.

Suppose for *reductio* that every conditional probability equals some unconditional probability—i.e. that there is no conditional probability that is 'unmatched'. There are two cases:

Case 1. $p_n < 1/2$

Then[31]

$$p_1 < \frac{p_1}{1-p_2} \leq \frac{p_1}{1-p_3} \leq ... \leq \frac{p_1}{1-p_n} < 2p_1 \leq p_1 + p_2$$

(since $p_i \leq p_{i+1}$).

We thus have these $n-1$ conditional probabilities of the form $p_1/(1-p_i)$ to be matched with unconditional probabilities strictly between p_1 and $p_1 + p_2$. As the only unconditional probabilities that could be in this interval are p_2, p_3, ..., p_n, this forces:

$$\frac{p_1}{1-p_2} = p_2, \ \frac{p_1}{1-p_3} = p_3, \ ..., \ \frac{p_1}{1-p_n} = p_n.$$

Thus

$$\frac{p_1}{1-p_i} = p_i \qquad \text{for } i = 2, ..., n$$

and so

$$p_i^2 - p_i + p_1 = 0 \qquad \text{for } i = 2, ..., n,$$

which are quadratic equations that we can solve for p_i. Therefore

$$p_i = \frac{1 \pm (1 - 4p_1)}{2} \qquad \text{for } i = 2, ..., n.$$

Now, we have assumed that $p_n < 1/2$, and so $p_i < 1/2$ for all i. Thus, we can disregard the positive square root.

---

[31] For the result and proof that, by considering Case 1 separately from Case 2, this set of inequalities forces the equality of p_2, ..., p_n in Case 1 (and for the similar result and proof in Case 2), I am indebted to Mike Larsen.

Therefore,

$$p_i = \frac{1 - |(1 - 4p_1)}{2} \quad \text{(a constant) for } i = 2, ..., n.$$

that is,

$$p_2 = p_3 = ... = p_n.$$

This implies that there is no unconditional probability value strictly between p_2 and $p_1 + p_2$. However, there *is* a conditional probability in this interval, namely $p_2/(1 - p_1)$. To see this, notice that

$$\frac{p_2}{1 - p_1} < \frac{p_1 + p_2}{p_1 + 1 - p_1} = p_1 + p_2$$

(the inequality holding since, for $n \geq 3$, $p_2 < p_2 + ... + p_n = 1 - p_1$, so the numerator of the first fraction is less than the denominator; thus, adding the same positive constant to both yields a greater fraction). This contradicts our hypothesis that every conditional probability has a match among the unconditional probabilities. So now consider:

Case 2. $p_n \geq 1/2$

Then $p_{n-1} < 1/2$, so

$$p_1 < \frac{p_1}{1 - p_2} \leq \frac{p_1}{1 - p_3} \leq ... \leq \frac{p_1}{1 - p_{n-1}} < 2p_1 \leq p_n$$

(the last of these inequalities holding since, with $n \geq 3$ and $p_n \geq 1/2$, we must have $p_1 \leq 1/4$).

Similarly to Case 1, matching conditional probabilities of the form $p_1/(1 - p_i)$ with unconditional probabilities strictly between p_1 and p_n forces

$$\frac{p_1}{1 - p_2} = p_2, \frac{p_1}{1 - p_3} = p_3, ..., \frac{p_1}{1 - p_{n-1}} = p_{n-1}$$

whence

$$p_i^2 - p_i + p_1 = 0 \quad \text{for } i = 2, ..., n-1$$

which implies

$$p_i = \frac{1 \pm \sqrt{(1 - 4p_1)}}{2} \qquad \text{for } i = 2, ..., n{-}1.$$

However, since $p_{n-1} < 1/2$, we can disregard the positive square root.

Therefore,

$$p_i = \frac{1 - \sqrt{(1 - 4p_1)}}{2} \qquad \text{(a constant) for } i = 2, ..., n{-}1$$

and thus

$$p_2 = ... = p_{n-1}.$$

Once again, this implies that there can be no unconditional probability value strictly between p_2 and $p_1{+}p_2$ (p_n is ruled out, since $p_1{+}p_2 \le 1/2$, but $p_n \ge 1/2$, by hypothesis of Case 2). As we saw in Case 1, this means that the conditional probability $p_2/(1{-}p_1)$ has no match among the unconditional probabilities, contradicting the assumption of the *reductio*.

This completes the *reductio*.

Thus, we have found that there must be a conditional probability that has no match among the unconditional probabilities, hence *a fortiori* no match among the (unconditional) probabilities of the form $P(- \rightarrow -)$.

Thus $<W,F,P,\rightarrow>$ is not a CCCP-model. Q.E.D.

7.3 Philosophical discussion

This refutes again the universal version of the Hypothesis, and drives a further nail in the coffin of the already buried belief function version. More importantly, it is the first result we have seen that refutes the universal tailoring version: no CCCP-conditional can be tailored to a probability function with finite range. Of the versions of the Hypothesis that I distinguished at the outset, this just leaves the belief function tailoring version, and it casts considerable doubt on that too, as my discussion will show.

The central assumption of the theorem is the finiteness of the range of the probability functions in question, and it is here that the defender of the Hypothesis must balk. He might question the usefulness of finite models along the following lines: "In realistic cases, a probability function with infinite range is required. Take, for instance, the die example above. A less simple-minded model of a die toss would spread probability over uncountably many worlds, each one corresponding to a distinct set of initial conditions of the toss—the force and angular momentum which is imparted to the die, the configuration of air molecules in the region of the toss, ... We recover the probability of each outcome by integrating the probabilities over the worlds that yield that outcome; but the range of the probability function is a continuum, and hence it is not susceptible to the theorem." So the challenge is that it is only infinite models that are really of interest. For the theorem to have bite, this sort of challenge must be countered.

I think it can be. Here are two reasonable construals of the Hypothesis:

Construal 1: it is an idealization, a hypothesis about the conditional used by some ideally rational agent;

Construal 2: it is a hypothesis about a conditional that finds its place in human language.

Thus, now is the time to make good my early promise to refine the Hypothesis further, making it clear just who the agent is supposed to be:

Ideal belief function tailoring version: For each P that could represent an ideally rational agent's system of beliefs, there is some $\rightarrow$ such that (CCCP) holds.

Human belief function tailoring version: For each P that could represent a rational human agent's system of beliefs, there is some $\rightarrow$ such that (CCCP) holds.

These versions presuppose that the system of beliefs of a rational human agent, or of an ideally rational agent, can be represented by a *single* probability function. When I come to my discussion of vague opinion, I will have reason to drop this presupposition,

and consider instead systems of belief that require *sets* of probability functions. But I want to add the complications in stages, and there is much to be said already at this stage.

My strategy will be this. I will argue that on either construal, an insistence on infinite models lacks cogency; thus I believe that the theorem essentially curtails the prospects for both these refined versions. These versions lie at opposite ends of a spectrum, from extreme idealization at one end, to none (or little) at the other end. To be sure, there are versions intermediate between these (although not all of them are so easily spelled out). However, I think that any intermediate version will inherit the difficulties of the end that it more closely approximates, or perhaps even difficulties from both ends, so it will not provide a refuge for the Hypothesis either. Thus I hope that my arguments, while not conclusive, will place an uncomfortable onus on the supporter of the Hypothesis.

Ideally rational beings

Take Construal 1 first. The insistence on infinite models imposes a strange constraint on ideally rational beings. It requires roughly that such a being not be very opinionated about where in logical space it lives. Less roughly: we can think of the agent's probability assignment to a given world as representing its degree of belief that the world it actually inhabits is that one; to uphold (CCCP), the agent must not know within a finite set which world it inhabits—unless it can zero in on a set with at most two members (the trivial cases that can never be ruled out by the triviality results). To be sure, gods such as Lewis imagines in (1979), who lack no propositional knowledge, can happily conform to (CCCP), albeit trivially—they've narrowed down to a single member the set of worlds which, for all they know, they inhabit (although it is unclear what use it could be to such a god to have a conditional in his language). But imagine a god-like being, who is on his way to becoming like Lewis' gods, but who hasn't quite

made it yet: he has acquired enough knowledge to rule out all but finitely many worlds as being the actual world, but not quite enough knowledge to narrow them down to just one or two. By the theorem, he violates (CCCP). Yet this fact alone should hardly count against his rationality. If he has already done such an impressive job of locating himself in logical space, more power to him.

Or consider the dilemma of the god-like being's younger brother, who currently spreads non-zero probability over infinitely many worlds, but who realizes that if he conditionalizes on a certain evidence proposition E, his probability function will then have finite (but non-trivial) support. In order to uphold (CCCP), he effectively must say to himself: "I had better never believe E, even if I learn it to be true"! This is absurd. (Saying instead "I will never learn E, even though I think it is possible" is little better—E then turns out to be a proposition fit for Moore's paradox, for no apparent reason.) Moreover, it is absurd to think that, at the point of coming to believe that he inhabits one of a certain finite set of worlds (perhaps thanks to learning E), he suddenly switches from being rational to being irrational, and that he remains so unless he eliminates all but two of the members of that set—at which point he suddenly switches back to being rational again. And of course, his period of 'irrationality' may well begin even sooner than that, namely at the point when his probability function first has only finitely many values.

In short, I believe the ideal belief function tailoring version of the Hypothesis should be laid to rest. The human belief function tailoring version is the only version left standing of those that I have distinguished so far.

Human beings[32]

This brings us to Construal 2 of the Hypothesis, which sees it as a hypothesis about a conditional that humans actually use. Here, then, we are supposed not to idealize

---

[32] I thank Fiona Cowie for a very helpful discussion on this section.

away our limitations, but to presuppose them. This may be problematic, because probability is already an idealized model of human belief (and Harman (1989), for example, doesn't think much of it as a model). Thus the Hypothesis, construed this way, may not even get off the ground. I'll set that aside for now (although I will discuss the prospects for the Hypothesis cast in terms of vague probability in §9.2, and for Adams' Thesis, a variant of the Hypothesis that eschews genuine probabilities, in §9.3). Supposing, then, that at least the most rational of us have belief states that are well described by probability functions (at least on our good days), the question is whether those probability functions must have infinite range.

Let me begin by arguing for a much stronger view than I actually need to dispose of the human belief function tailoring version. I think that there is a sense in which finite beings like ourselves are not accurately represented by *anything but* finite models—any infinite model must add structure that isn't there in the opinion of the human agent that is supposedly being modeled. For I think that any human actually discriminates among only finitely many evidence propositions[33], and distinguishes only finitely many propositions in thought, so modeling him or her accurately requires a probability function over just those propositions (plus those arrived at by closing under the Boolean operations). Hence, only a probability function with finite domain accurately models him or her. But a probability function with finite domain has finite range. Thus, I think that any infinite model distinguishes propositions that aren't so distinguished by the agent being modeled, and indeed distinguishes at least some of them in probability.

Certain 'language of thought'-ists probably don't need any convincing of this—there surely can be only finitely many sentences in the 'belief box' of an actual person. Furthermore, considerations such as the ones that I have outlined, and others, might even incline one to be a finitist about mathematics itself—a true Wittgensteinian would

---

[33] On this point, see also Lewis (1986).

presumably have no quarrel with finite models here. However, I don't think I need to join these camps in order to make my point. Given our finite processing times, finite memories, finite lifetimes, and so on, I take it to be just an empirical fact that none of us actually entertains more than finitely propositions in any belief state—and that includes the best of us, on our best days. Note that so far I have been careful to steer clear of modally charged notions, such as the propositions that we are *capable* of entertaining. And while we may speculate about what we would be like if we were suitably endowed with infinite computational powers, and so on, that's not playing fair *here*. The idealizer had his chance before.

You may think that I am ignoring important middle ground here—that we should entertain certain modest counterfactuals about what we would be like if, for example, our lives went differently. But if we keep fixed the important psychological facts about us (our finite computational power, and so on), then the people that we are imagining still have finite-ranged probability functions. And here we *must* keep fixed the important psychological facts about us—again, the place for entertaining more extravagant counterfactuals, in which those facts are suspended, was in the previous sub-section. Counterfactuals that don't idealize away our limitations pose no problem for finite models, in my opinion.

This is still controversial, I admit, and soon I will consider an argument that controverts it; but to repeat, note that it's actually a much stronger view than I really need to defend. For the proponent of the human belief function tailoring version must insist that a probability function with finite range *could not represent* a rational human's system of beliefs. So he denies even a far more moderate position than the one I have expressed above—one that insists only that finite models *sometimes suffice* to model rational human agents. I find it hard to see how this more moderate position can be plausibly denied.

Incidentally, it is worthwhile recalling the objector who challenged the usefulness of finite models, appealing to the die toss example to argue that realistic models of opinion should be infinite. It is curious that this objector to finite models at once insists on strict faithfulness to the facts about tosses of a die, and yet plays fast and loose with the facts about human agents.[34] He is quick to point out the fine details of dice tossing that require an infinite model in order to be captured accurately (actually, even that is questionable in the light of some recent physics, but let that pass), and yet he glosses over the fine details of what *we* are like—that we have finite minds, finite lifetimes, and so on. Indeed, even a coarse description of us should recognize *that* much. He is trying to have it both ways—to abstract away from our finitude, and yet to shun such abstraction about dice tossing.

To sum up: if the Hypothesis concerns human beings as we actually are—and for now we are assuming that it does—I believe that it concerns beings whose belief systems happen not to require infinite models. But there is an argument that I must answer.

Implicit beliefs

You might say that while your finitude implies that you have only finitely many *explicit* beliefs, you nevertheless have infinitely many *implicit* beliefs—namely, everything that is implied by your explicit beliefs. Now, we may leave open for the moment just how implicit beliefs should be analyzed. But whatever implicit beliefs are, statements like this are true: since you explicitly believe that there are nine planets, you implicitly believe that the number of planets is less than 10, less than 11, and so on. So counting implicit beliefs as beliefs, you have infinitely many beliefs about just the planets! Furthermore, modeling your system of explicit *and* implicit beliefs will

---

[34] I am indebted here to discussion with Mike Thau.

require a probability function with infinitely many distinct values. Or so this argument runs.

Firstly, the step from 'infinitely many implicit beliefs' to 'infinitely many distinct probability values' is too quick. For simplicity, imagine someone who has exactly one explicit belief—that there are nine planets, say—and he believes this firmly. Then we may represent him as assigning probability one to there being nine planets. We may concede for sake of argument that he implicitly believes all the infinitely many consequences of this—but they all get probability one as well. So despite his having infinitely many (implicit) beliefs, his probability function takes only the values 0 and 1.[35]

It may also help to see why the 'infinitely many implicit beliefs, therefore infinitely many distinct probability values' inference is invalid, if we note that the inference is *clearly* invalid if we replace 'implicit' by 'explicit'. Consider again one of Lewis' gods, who knows the truth values of all propositions. He has infinitely many beliefs[36]; yet his probability function takes only the values 0 and 1, fully concentrated as it is on the one world he knows to be actual. And to a considerable extent, the more opinionated one is, the more closely one's probability distribution approximates a highly 'peaked' distribution like this. The invalid inference seems to rest on the mistaken rule of thumb that—roughly—the greater the stock of opinions, the richer the range of the probability function. If anything, the rule of thumb should be more: the greater the stock of opinions, the *less* rich the range of the probability function.

Perhaps there are humans whose putative implicit beliefs *might* require a probability function with infinite range—for example, a poor dart-thrower who explicitly believes that a dart thrown by him is equally likely to land anywhere on a dart board, and so

---

[35] Even insisting that he has vague opinions about anything that is more finely grained, this will only add the single interval of vagueness [0,1] to his stock of probabilities.

[36] Lewis himself might prefer this to be worded: "infinitely many things are true according to his belief system", but I hope nonetheless that the point here is well taken.

implicitly assigns probability 1/2 to it landing in the lower 1/2, probability 1/3 to it landing in the lower 1/3, and so on. But we are not all like that. Furthermore, it would be odd indeed if the dart-thrower were to say: "Just as well that I'm so bad at darts, because it's what keeps me rational—it's what keeps my probability function infinite-ranged"! The appeal to implicit beliefs in order to save the Hypothesis seems to presuppose that *every* rational human being must have something that plays the role that the dart board does for the dart-thrower—something that generates implicit beliefs of infinitely many different probabilities. I find this hard to believe.

Secondly, the argument presupposes a debatable view of what implicit beliefs are: roughly, they are whatever is implicit *in* your (explicit) beliefs—that is, whatever follows logically and mathematically from them. But do we really want to say that a two-year old believes all truths of mathematics (albeit implicitly)? There is another view of what implicit beliefs are, which I find at least as plausible: they are propositions that are "easily inferable from one's explicit beliefs", as Harman (1989, p. 13) suggests. On this view, an implicit belief could easily be made an explicit belief, provided the agent gives it some reflection. But understood this way, it is doubtful that any human has infinitely many implicit beliefs—and this includes the dart-thrower— for all but finitely many of the implications of what one explicitly believes are too complicated to be comprehended, let alone believed.

In short, I believe that the human belief function version, as I have stated it, is not tenable. (Further versions in the spirit of it, however, will be discussed in §9, as I have indicated.)

7.4. *Are there CCCP-models of greater cardinality?*

I have shown, with no assumptions about the logic of the →, that there are no non-trivial finite CCCP-models. But my result does not rule out the possibility of constructing out of a given finite model, a new, *more finely grained* CCCP-model. This

raises the natural question: are there CCCP-models of higher cardinality? Here, two results are relevant—one negative, the other positive.

First, the negative result, due to Ned Hall (1993). Call a probability space <W,F,P> *full* if, for each element A of F, P takes every value in [0, P(A)] on the elements of F which are subsets of A. Hall shows that, with minimal assumptions about the logic of the →, every CCCP-model is full. It follows immediately from this that every CCCP-model's probability function has uncountable range. Thus, Hall extends my 'no finite CCCP-models' result to a 'no countable CCCP-models' result. The extra strength of his conclusion is bought by his minimal logical assumptions: that the → obeys the following two constraints, for all A, B, C ∈ F:

$$[(A \rightarrow B) \cap (A \rightarrow C)] \subseteq [A \rightarrow (B \cap C)]$$

$$A \cap (A \rightarrow B) \subseteq A \cap B$$

These assumptions are certainly reasonable, agreed to by both Stalnaker and Lewis. The first is a certain distributivity condition; the second is modus ponens.

The natural question, then, is whether even Hall's result can be extended—whether even full CCCP-models can be eliminated, thus effectively refuting any 'tailoring' version of the Hypothesis. The answer is 'no'—and here we appeal to van Fraassen's positive result in (1976). Following van Fraassen, call a *proposition algebra* any field of propositions with the following further constraints:

(i) $[(A \rightarrow B) \cap (A \rightarrow C)] = [A \rightarrow (B \cap C)]$

(ii) $[(A \rightarrow B) \cup (A \rightarrow C)] = [A \rightarrow (B \cup C)]$

(iii) $[A \cap (A \rightarrow B)] = (A \cap B)$

(iv) $A \rightarrow A = W$.

An → that obeys these principles conforms to the logic *CE*. Van Fraassen proves what amounts to the following theorem:

Any probability space can be extended to a CCCP-model, whose → conforms to CE.

In the proof, he shows that any probability space can be extended to a full probability space; and for every full probability space <W,F,P>, there is an →such that <W,F,P,→> is a CCCP-model. So give van Fraassen the probability space, and he will embed it in a full probability space, and then construct a CE →tailored to fit. The guiding insight behind van Fraassen's construction is simple enough to state. Since CE includes modus ponens and the principle that A & (A→B) is equivalent to A&B,[37] for any A and B, A→B consists of none of the A&¬B worlds, all of the A&B worlds, plus (perhaps) some of the ¬A worlds. Let P, the probability function on a full probability space, be given. The probability proportion of A→B worlds among the A worlds is P(A&B)/P(A), so (CCCP) is automatically correct if we restrict our attention to the A worlds. The truth of (CCCP) then hinges on what the proportion of A→B worlds is among the ¬A worlds. The proportion is right iff P(A→B|¬A) = P(B|A) also. So the trick is to choose (A→B) & ¬A judiciously so that this is the case. Given the fullness of the space, it turns out that there is always a candidate for such an →.

It is not surprising that the CCCP space that van Fraassen constructs out of the original probability space is full, given Hall's negative result for smaller spaces (which makes no logical assumption that van Fraassen's result doesn't). As Hall shows us, nothing less than fullness will do.

So van Fraassen guarantees the immortality of a certain version of the Hypothesis, which we can call the:

Full probability space tailoring version: For each full <W,F,P>, there is some → such that <W,F,P,→> is a CCCP-model.

In fact, as I've said, the → that is tailored conforms to the CE logic. Whether that means that it deserves the name 'conditional' is a further question, about which I don't have a strong opinion. I will just mention, however, that CE lacks the principle of weakened transitivity:

---

[37] Following van Fraassen, I use the sentential and the set theoretic notation interchangeably.

$$[(A \to B) \cap (B \to A) \cap (A \to C)] \subseteq (B \to C),$$

which is common to both the Stalnaker and Lewis logics. (As I noted earlier, van Fraassen (1976) and (1981) says that he rejects this principle—but he doesn't elaborate.) On the other hand, constraint (ii) is controversial. Among other things, it entails conditional excluded middle (which should come as no surprise, given the close connection between (CCCP) and conditional excluded middle, which I have discussed at some length).

Let's return to the original versions of the Hypothesis. The universal tailoring version has been refuted; and in my opinion, the belief function tailoring version has been essentially refuted too. I have argued that belief functions can have finite range, and they are susceptible to my negative result. I imagined the proponent of the belief function tailoring version saying in response that all belief functions have infinite range. Thanks to Hall's result (accepting his assumptions about the logic of the $\to$, as I think we should) we see now that he must say more: that all belief functions have uncountable range, and indeed that belief can only be represented by a *full* model. I don't think that van Fraassen's positive result is much consolation here—I think that this is too big a bullet to bite.

8. Why disbelieve the Hypothesis? Part V
Triviality on the cheap?[38]

8.1 Preliminaries

The Hypothesis is not dead yet, even if the four versions of it that I distinguished at the outset are. Suitably qualified, it can thread its way through the loopholes left open by the results of Lewis, Hall and Stalnaker, and the results that I have just offered. I have said enough about the first four versions—I have shown the first and third of these to be refuted, and I have argued that the second and fourth are untenable. However, there are other reasonable construals of the Hypothesis. Of course, I cannot hope to exhaust every imaginable construal that might be considered 'reasonable'—I would exhaust my reader and myself long before that. But four strategies open to the friend of the Hypothesis strike me as particularly salient. Three of them I will discuss in §9: to settle for an approximate rather than an exact version of the Hypothesis; to cast it in terms of vague opinion; and to loosen the requirement that it concerns genuine probability functions. In this section I want to discuss—and thwart—another strategy.

There is a familiar line of response that philosophers are fond of employing when some concept that we cherish appears to be incoherent: admit the incoherence of that exact concept, but argue that a close relative to it, which can play most of the important roles in practice of the original concept, *is* coherent. Lewis (1992a), for instance, takes such a line with respect to chance. It seems that nothing can perfectly deserve the name 'chance', as he argues; nevertheless, "an imperfect candidate may deserve the name quite well enough". Similarly, the triviality results may be taken to show that nothing

---

[38] This section (slightly edited) is to appear as "Triviality on the Cheap?" in Eells and Skyrms (eds.) (1993).

can perfectly deserve the name 'CCCP-conditional'; nevertheless, an imperfect candidate may deserve the name quite well enough.

One way to have an imperfect CCCP-conditional for a given probability function is to have it satisfy (CCCP) for most arguments of the probability function, but not for all. We saw at the outset that we might want to weaken the universal version of the Hypothesis by shrinking a certain domain of quantification: from all probability functions, on the one hand, to all probability functions that could represent a rational agent's system of beliefs, on the other. Now, we might try shrinking another domain of quantification: the domain of *propositions*, which appears in the very statement of (CCCP). After all, domain-shrinking is all the rage nowadays, and despite the demise of the full (CCCP), we may yet be optimistic about the prospects for a reduced version of it.

The spirit of Stalnaker's and Adams' original proposals could be upheld by a reduced version of (CCCP). Perhaps it was unreasonable to hope that conditional probability equals the probability of a conditional for *all* possible antecedents and consequents. That lets in propositions that may have no simple expression in a given language—for example, multiply iterated conditionals, extravagant Boolean combinations of conditionals, and so on—and such sentences are never uttered in natural language. And it is reasonable to think that all the linguistic data that we can garner is unable to distinguish between the full-blown Hypothesis, and the Hypothesis with some suitable restriction on the domain—thus, the latter could suffice to serve our purposes in illuminating the semantics of the conditional, and so on. Until now, we have been requiring that, for a given model $<W,F,P,\rightarrow>$,

(CCCP) $\quad$ $P(A{\rightarrow}B) = P(B|A)$ *for all* $A, B \in F$ where $P(A) > 0$.

But perhaps it is appropriate to retreat to some more restricted formulation of (CCCP):

(Restricted CCCP) $\quad$ $P(A{\rightarrow}B) = P(B|A)$ $\quad$ for all $A, B \in \mathbf{S}$

where $\mathbf{S}$ is some proper subset of F. The version that I want to consider, then, is the:

Restricted universal version: There is some → such that for all P, (Restricted CCCP)

 holds.

Of course, it is really more a hypothesis schema, since **S** is not properly specified. This will not matter, since any non-trivial version of the restricted universal version must fail. My main purpose in this section is to show this.

That will mean that with the exception of trivial cases, for any →, A and B, there is a probability function P (defined on the relevant conditional space, which I will take for granted from now on), such that P(A→B) ≠ P(B|A). Restricting the propositions that can appear on either side of the → simply isn't enough. It might be tempting, then, to demand not that (Restricted CCCP) holds for *all* probability functions, but rather that it holds for some privileged subset of them. For example:

Restricted belief function version: There is some → such that for all P that could

 represent a rational agent's system of beliefs, (Restricted CCCP) holds.

Harder though it is to refute this doubly qualified Hypothesis, I will argue that it too is untenable (for any non-trivial restriction).[39]

My attack on these restricted versions of the Hypothesis will begin with an intuitive argument, which I will then recast in the form of a more rigorous proof. I will make no assumptions about the logic of the →, and I will show that, for a given A and B, there is *no* relation that A→B can bear to A and B that allows all probability functions (or even all belief functions, as I will argue) to conform to the conditional construal of conditional probability.

This limitation result sharpens an argument due to Carlstrom and Hill (1978), as I will show. Appiah (1985) prefers their argument to those of Lewis (1976), since Lewis assumes that a belief function can be irregular, while they do not. (To repeat: an irregular probability function assigns probability zero to other propositions besides the

---

[39] You may be expecting me to introduce restricted analogues to the universal tailoring and belief function tailoring versions. I won't do this, since I have no negative results against them.

empty one.) Their argument has a small flaw in it, and it has some unnecessary, and some unnecessarily strong, assumptions. I believe that my result corrects the flaw, dispenses with the unnecessary and weakens the unnecessarily strong assumptions, while still avoiding irregular probability functions.

The proof will also avoid various assumptions distinctive of earlier triviality results—assumptions about the closure properties of sets of CCCP-functions for a given → (à la Lewis, and §6 above); about the logic of the → (à la Stalnaker); and about the cardinality of the probability functions' range (à la §7 above). Thus, I believe that triviality results can be had quite cheaply. Just how cheaply is to be gauged by the strength of the one assumption that I do make concerning the uniformity (in a certain sense) of the →, and I will argue that even this assumption can be considerably weakened: where Lewis (and §6) assumes the uniformity of the → throughout certain classes of probability functions, here I will need only some constraints on what non-uniformities there could be.

8.2 A method for generating counterexamples to (CCCP)

The method that I will employ is best displayed initially via operations on probability functions that I will call *perturbations,*[40] but it will be easily seen to generalize to other interesting operations, namely conditionings and Jeffrey conditionings. The trick is similar in each case. We begin by noting a simple home truth. Suppose we have some P, A and B for which

$$P(A{\rightarrow}B) = P(B|A).$$

Now suppose that some other probability function P' assigns a different probability to the conditional:

$$P'(A{\rightarrow}B) \neq P(A{\rightarrow}B)$$

Then if P' assigns the *same* conditional probability as P does:

---

[40] This term was suggested to me by David Lewis.

P'(B|A) = P(B|A),

we have immediately

P'(A→B) ≠ P'(B|A).

Similarly, if P and P' agree on the probability of the conditional, but disagree on the conditional probability, then they cannot both equate the two, for this particular choice of A and B. I will show that it is easy to find such pairs of probability functions. Now, there are countless ways that two probability functions can be related to one another. I will show that there are important ways that P and P' can be related that yield this negative result for the restricted universal version of the Hypothesis.

8.2.1 Perturbations

The intuitive argument

It is best to present the argument intuitively first. Afterwards, I will provide a more rigorous presentation of it; and after that, I will derive some philosophical consequences from it. I prefer the intuitive version of the argument to the rigorous one, because it not only shows *that* probabilities of conditionals come apart from conditional probabilities, but also shows more perspicuously *why* this is so.

I will picture probability in terms of van Fraassen's 'muddy Venn diagram' (1989, p. 161). Think of an agent assigning subjective probability to worlds. Propositions are represented by regions on a Venn diagram (and are thought of as sets of worlds). Infinitely divisible mud is imagined to be heaped on the diagram, in such a way that the total amount of mud is always taken to be 1, and the amount of mud on any region is just the probability of the proposition represented by that region. Different probability distributions over a fixed algebra are represented by different distributions of mud over the same diagram.

Assume nothing about the logic of →, aside from its being a two-place connective. If you like, you may replace '→' in what follows by some unfamiliar symbol that is free

of any associations. Suppose that there is an →, a pair of propositions A and B, and a probability function P such that P assigns positive probability to A&¬B, A&B and ¬A, and P(A→B) = P(B|A). P will be our initial distribution of mud. The trick is to find certain movements of the mud which, if allowed, would represent changes in the probability of the conditional, without corresponding changes in the conditional probability, or vice versa.[41] If P' is the distribution of mud that results, then P'(A→B) ≠ P'(B|A).

Either A→B is not a Boolean combination of A and A&B, or it is. (It is important to remember that A and B are not variables here—I am saying that either the unique proposition which is A→B is not a Boolean combination of the particular A and A&B that we are considering, or it is. To say that this particular A→B is a Boolean combination of A and A&B is *not* to say that → is truth functional, since X→Y (where X and Y are not A and B) might be some quite different Boolean, or even non-Boolean combination of X and X&Y.)

Case 1. A→B is not a Boolean combination of A and A&B

Then either

i.   A→B takes both truth values within A&¬B; or

ii. A→B takes both truth values within A&B; or

iii. A→B takes both truth values within ¬A.

We will consider these sub-cases in turn:

---

[41] During a discussion of our joint paper, Ned Hall pointed out to me that mud movements over the ¬A & (A→B) boundary have this property; he is thus an important source of inspiration for my result.

i. We may illustrate this sub-case as follows:
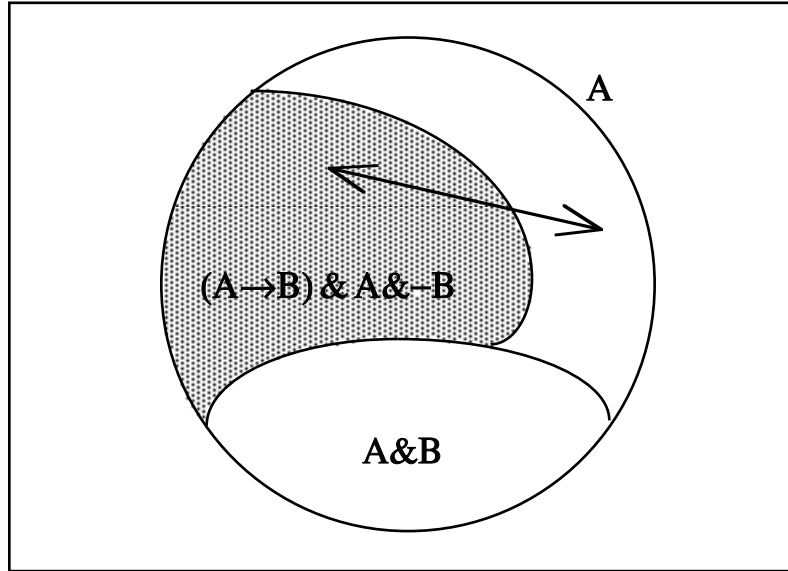
If A→B takes both truth values within A&¬B, then we can move mud between (A→B) & A&¬B and ¬(A→B) & A&¬B (as I have indicated with an arrow in the diagram). This changes the total amount of A→B mud—that is, the probability of A→B changes. However, the conditional probability of B, given A, remains unchanged. After all,

$$P(B|A) = \frac{P(A\&B)}{P(A)} \ ,$$

and the terms of this ratio are unaffected by moving mud from one part of A&¬B to another. So starting with a probability distribution P for which P(A→B) = P(B|A), we can produce others for which this is not so.

ii. We may illustrate this sub-case as follows:
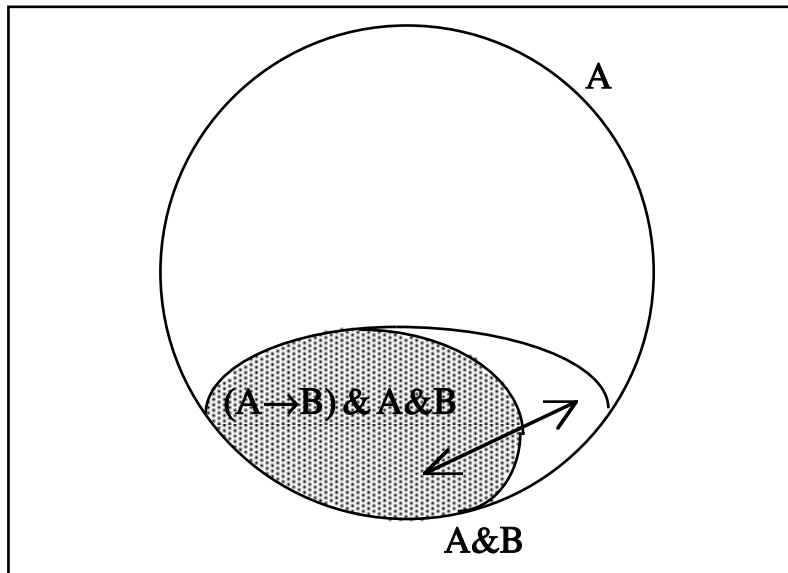
If A→B takes both truth values within A&B, then we can move mud between (A→B) & A&B and ¬(A→B) & A&B (as I have indicated with an arrow in the diagram). This changes the probability of A→B; however, the conditional probability of B, given A, remains unchanged, since the terms of the ratio are unaffected by moving mud from one part of A&B to another. So starting with a probability distribution P for which P(A→B) = P(B|A), we can produce others for which this is not so.

iii. We may illustrate this sub-case as follows:
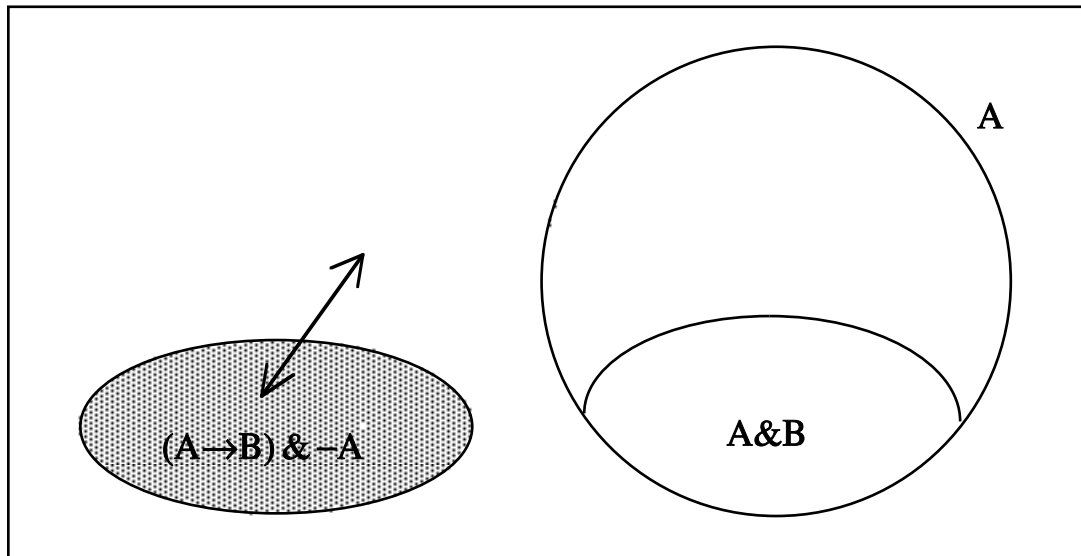
If A→B takes both truth values within ¬A, then we can move mud between (A→B) & ¬A and ¬(A→B) & ¬A (as I have indicated with an arrow on the diagram). This changes the probability of A→B; however, the conditional probability of B, given A, remains unchanged, since the terms of the ratio are unaffected by moving mud from one part of ¬A to another. So starting with a probability distribution P for which P(A→B) = P(B|A), we can produce others for which this is not so.

Thus we see that if A→B is not a Boolean combination of A and A&B, we can find functions whose probabilities for A→B differ from their corresponding conditional probabilities. This completes case 1.[42]

Case 2. A→B is a Boolean combination of A and A&B

Since P assigns positive probability to A&¬B, A&B and ¬A, and P(A→B) = P(B|A), we can rule out the sub-cases in which A→B is a tautology or a contradiction.

---

[42] Actually, there are other mud movements that have the same effect. For instance, one could do two or three of the mud movements described in i–iii at once, provided there is a net change of mud in A→B.

That leaves six sub-cases, which we can group in pairs:

iv. A→B = A&B, or A→B = ¬(A&B). Then moving mud between ¬A and A&¬B will change the conditional probability, without changing the probability of the conditional. So we can find functions for which the two come apart. See the following diagram:
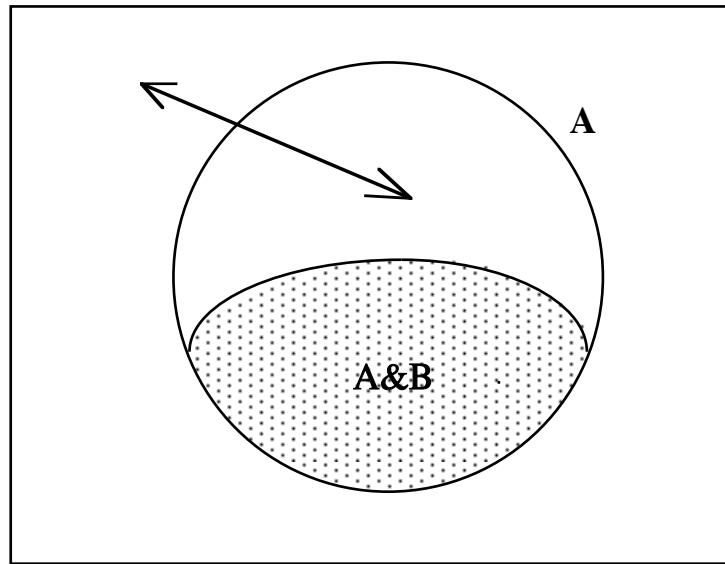
v. A→B = A, or A→B = ¬A. Then moving mud between A&B and A&¬B will change the conditional probability, without changing the probability of the conditional. So we can find functions for which the two come apart. See the following diagram:
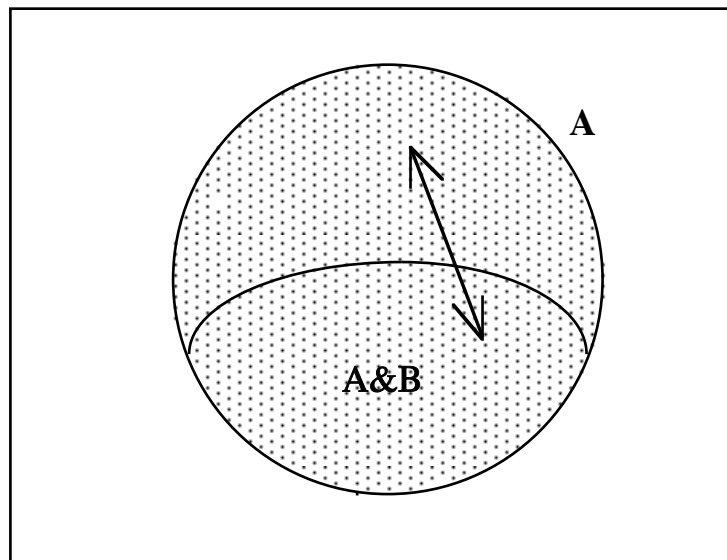
vi. A→B = A⊃B, or A→B = ¬(A⊃B). Then moving mud between ¬A and A&B will change the conditional probability, without changing the probability of the conditional. So we can find functions for which the two come apart. See the following diagram:
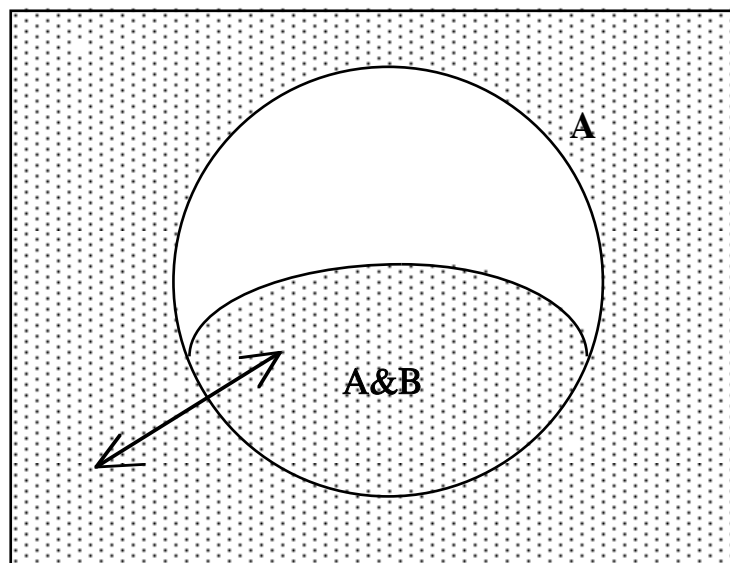
So much for the intuitive argument. Sub-case i is perhaps not a real possibility if the $\rightarrow$ is a genuine conditional, for it is reasonable to think that A→B should be uniformly false throughout A&¬B. Still, it is worth including this step in the argument, since I do not want it to presuppose anything about the logic of the $\rightarrow$ —not even something as innocent as this. As a result, the argument embraces even probabilistic, or graded 'conditionals' such as Lewis discusses in §8 of (1973a), or Halpin's (1991) $\rightarrow_p$, for which A$\rightarrow_p$B can be true even when A is true and B false. (I have also mentioned McGee's argument against modus ponens.) To repeat, it embraces *all* two-place connectives, irrespective of their logic.

Thinking again of $\rightarrow$ as a genuine conditional: those who favor the principle that A&B implies A→B will deny that sub-case ii is a real possibility—they think that A→B does not distinguish among the A&B worlds, since it is uniformly true throughout them. And in general, X→Y will not be a Boolean combination of X and X&Y (for it is well known that no truth function can allow non-trivial conformity to the Hypothesis), so sub-cases iv through vi will in general not obtain (although one of them may obtain for the particular A and B that is being considered). I think the real damage is done in sub-case iii, for I think that it is the normal case.

It may be instructive to note why I had to assume that P assigned non-zero probability to each of A&¬B, A&B, and ¬A. I needed a guarantee that there was some mud to start with in *each* of the three key regions, so that I could make good my threat to move it.

A more rigorous argument

Now, as promised, I will give a more rigorous reformulation of the argument above.

Let <W,F,→> be given, and let A ∈ F and B ∈ F be given. Let P and P' be two probability functions defined on this conditional space. Let A♥B be any one of these combinations of A and B:

   i. A&¬B

   ii. A&B

   iii. ¬A.

It's a shorthand that will often save me the trouble of writing out all three cases (and you the trouble of reading them), when they can be treated uniformly.

Definition. Call P' *a (→, A, B)-perturbation of* P if any of the following is the case:

   i-iii. P' and P agree everywhere outside A♥B,[43] but disagree on the probability of (A→B) & A♥B;

   iv. A→B = A&B, or A→B = ¬(A&B), and P'(A&B) = P(A&B) but P'(A) ≠ P(A);

   v. A→B = A, or A→B = ¬A, and P'(A) = P(A) but P'(A&B) ≠ P(A&B);

   vi. A→B = A⊃B, or A→B = ¬(A⊃B), and P'(A⊃B) = P(A⊃B) but P'(A&B) ≠ P(A&B).

Definition. Let A and B be given. Call P *non-trivial relative to A and B* if P(A&¬B), P(A&B) and P(¬A) are all positive.

Lemma 1. For all →, A and B, and P non-trivial relative to A and B, there exists a (→, A, B)-perturbation of P.

Proof. Let →, A and B be given, and let P be non-trivial relative to A and B. There are two cases.

Case 1. A→B is not a Boolean combination of A and A&B

Then either:

i.   A→B takes both truth values within A & ¬B, or

ii.   A→B takes both truth values within A&B, or

---

[43] In case it is not already obvious what this means: for all propositions X ∈ F which imply ¬(A♥B), P(X) = P'(X).

iii. A→B takes both truth values within ¬A.

In each sub-case, there exists a (→, A, B)-perturbation of P:

Consider sub-case iii (I choose it only because it is notationally the simplest, even simpler than the A♥B notation is). By non-triviality of P relative to A and B, we have P(¬A) > 0, and so we have at least one of P((A→B) & ¬A) and P(¬(A→B) & ¬A) being positive. Then there are infinitely many ways that we can choose a (→, A, B)-perturbation of P, which we will call P'. Here's just one. If both the probabilities just mentioned are positive, let X be (A→B) & ¬A; otherwise, let X be ¬(A→B) & ¬A. So P(X) > 0. P' agrees with P throughout A, but P' halves all of P's probabilities for propositions contained in X, and adjusts all probabilities of propositions contained in ¬X & ¬A proportionally.[44] That is, for all Y that imply X, P'(Y) = P(Y)/2; and for all Z that imply ¬X & ¬A,

$$P'(Z) = P(Z).\frac{P(\neg X \& A) + P(X)/2}{P(\neg X \& A)} \ .$$

We reason analogously for sub-cases i and ii. In all three sub-cases, there exists a (→, A, B)-perturbation of P.

Case 2. A→B is a Boolean combination of A and A&B

By non-triviality of P relative to A and B, we can rule out the sub-case in which A→B = T or A→B = F.

Consider sub-case iv, in which A→B = A&B, or A→B = ¬(A&B).

---

[44] There is another, even simpler perturbation, which moves all of P's probability for X to ¬X & ¬A. That choice of P' would be irregular, whereas mine isn't if P isn't. Given that some authors insist on regularity of belief functions, my choice is better strategically—it shows that I need no assumption of irregularity.

As Bas van Fraassen has pointed out to me, I assume that the family of admissible probability functions here is not constrained by any special conditions relating the probabilities of propositions that imply A&B, A&¬B, or ¬A (as might happen if the language contained some other connector besides the usual Boolean ones, and the →).

Again we have $P(\neg A) > 0$. Then there are infinitely many ways that we can choose a $(\rightarrow, A, B)$-perturbation of P, which we will call P'. Here's just one. P' agrees with P throughout A&B, but P' halves all of P's probabilities for propositions that imply $\neg A$, and adjusts all probabilities of propositions that imply A&$\neg$B proportionally. That is, for all Y that imply $\neg A$, $P'(Y) = P(Y)/2$; and for all Z that imply A&$\neg$B,

$$P'(Z) = P(Z).\frac{P(A\&\neg B) + P(\neg A)/2}{P(A\&\neg B)} \; .$$

Sub-cases v and vi are analogous: P' can be chosen in each case to agree with P where required, and to disagree with P where required. In all three sub-cases, there exists a $(\rightarrow, A, B)$-perturbation of P. Q.E.D.

Lemma 2. For all $\rightarrow$, A and B, if P' is a $(\rightarrow, A, B)$-perturbation of P, then at least one of P' and P violate (CCCP) for this particular choice of $\rightarrow$, A and B.

Proof. Assume P' is a $(\rightarrow, A, B)$-perturbation of P. Suppose that $P(A{\rightarrow}B) = P(B|A)$. Then (at least) one of sub-cases i–vi obtains.

<u>In each of sub-cases i–iii</u>, $P'(A{\rightarrow}B) \neq P(A{\rightarrow}B)$, but $P'(B|A) = P(B|A)$. I will show this for sub-case iii, for which P' and P agree everywhere outside $\neg$A, but disagree on the probability of $(A{\rightarrow}B)$ & $\neg$A. (i and ii are analogous, but slightly more tedious to state.)

Now

$$P'(A{\rightarrow}B) \;\; = P'((A{\rightarrow}B) \& \neg A) + P'((A{\rightarrow}B) \& A)$$

$$= P'((A{\rightarrow}B) \& \neg A) + P((A{\rightarrow}B) \& A)$$

since P' and P agree outside $\neg$A

$$\neq P((A{\rightarrow}B) \& \neg A) + P((A{\rightarrow}B) \& A)$$

since P' and P disagree on the probability of $(A{\rightarrow}B)$ & $\neg$A

$$= P(A{\rightarrow}B).$$

Hence

(1)   P'(A→B) ≠ P(A→B).

However,

P'(A&B) = P(A&B),

and

P'(A) = P(A),

since P' and P agree outside ¬A. Hence,

(2)   P'(B|A) = P(B|A).

Since

P(A→B)   = P(B|A),

we have by (1) and (2),

P'(A→B)   ≠ P'(B|A),

a violation of (CCCP) for this particular choice of A and B.

   In each of sub-cases iv-vi, P'(A→B) = P(A→B), but P'(B|A) ≠ P(B|A). Consider sub-case iv, for which A→B = A&B, or A→B = ¬(A&B), and P'(A&B) = P(A&B) but P'(A) ≠ P(A). (v and vi are analogous, so I won't give the details for them.) If A→B = A&B, the result is virtually immediate; so suppose A→B = ¬(A&B).

   Since

P'(A&B) = P(A&B),

we have

1–P'(A→B) = 1–P(A→B),

so

P'(A→B) = P(A→B).

However, since

P'(A) ≠ P(A),

we have

P'(B|A)   ≠ P(B|A)

(the numerators of the two fractions agreeing, but the denominators not). So

P'(A→B)   ≠ P'(B|A),

violating (CCCP) for this particular choice of A and B. Q.E.D.

Definition. Call a proposition A *intermediate* if there is a probability function P such that $0 < P(A) < 1$.

Theorem. There is no →, and intermediate, consistent but distinct propositions A and B such that for all P, P(A→B) = P(B|A).

Proof. By lemmas 1 and 2, and the fact that for any such A and B, there is a P that is non-trivial relative to them. Q.E.D.

This refutes all (non-trivial) instances of the restricted universal version of the Hypothesis.

It will be handy to have a more general notion of perturbation at the ready. Let → be given.

Definition. Call P' a *perturbation of* P *relative to* → if there is an A and B such that P' is a $(→, A, B)$-perturbation of P.

I will sometimes suppress the reference to the →, when it is obvious which one is meant. Intuitively, a perturbation of P relative to → is a probability function which, for some A and B, agrees with P everywhere except on some Boolean combination of A and A&B. Their disagreement over that region involves one of the sides of (CCCP); their agreement elsewhere ensures their agreement on the other side of (CCCP). We could equally speak of perturbation as an *operation*, taking a probability function as input and producing a function that is a specified perturbation of it as output.

As a corollary to lemma 2, we have what I will call the

Perturbation Theorem.   If P' is a perturbation of P relative to a given →, then at most one of P and P' is a CCCP-function for →.

Proof. Assume the hypothesis, so there is an A and B such that P' is a $(\rightarrow, A, B)$-perturbation of P. Suppose that P is a CCCP-function for $\rightarrow$. Then $P(A\rightarrow B) = P(B|A)$ for this A and B. But then by lemma 2, P' violates (CCCP) for this particular choice of $\rightarrow$, A and B, so P' is not a CCCP-function for $\rightarrow$. Q.E.D.

I showed in §5.1 that (trivial cases aside) the family of CCCP-functions for a given $\rightarrow$ is not closed under mixing; Lewis showed us that it is not closed under conditioning, restricted conditioning, or Jeffrey conditioning; and in §6 I showed that it is not closed under a fearless and moderate revision rule. Now we see that (trivial cases aside) it is not closed under perturbation either. Indeed, we see something stronger still: the family of functions that conform to (Restricted CCCP), for a given $\rightarrow$, is not closed under perturbation (whatever non-trivial restriction we impose). This will have important ramifications for certain further attempts to rescue the Hypothesis, as I will show in §9; and for the restricted belief function version, as I will show shortly.

Comparison to Carlstrom and Hill's argument

Before that, however, I want to compare my result here to that of Carlstrom and Hill in their (1978) review of Adams—one which Appiah (1985) applauds as being superior to Lewis' (1976) results. Carlstrom and Hill give the following argument, based on an argument by Adams (1975) to show that the probability of a conditional can come apart from the corresponding conditional probability:

Let A and B be contingent (by which I take them to mean 'intermediate' in my sense) and logically independent. Assume that A$\rightarrow$B can be true when both A and B are: let w_1 be a world at which all three are true. Assume that A$\rightarrow$B is not a truth function of A and B. Let w_2 and w_3 be worlds that agree on the truth values of A and A & B, and such that A$\rightarrow$B is false at w_2 but true at w_3. Now consider two probability assignments, P and P', such that P divides almost all the probability roughly equally

between w_1 and w_2, and P' divides almost all the probability roughly equally between w_1 and w_3.

It follows from this that

$P(A{\rightarrow}B) \approx 1/2$, and $P'(A{\rightarrow}B) \approx 1$,

so

(1)   it is not the case that $P(A{\rightarrow}B) \approx P'(A{\rightarrow}B)$.

But

(2) $P(A \ \& \ B) \approx P'(A \ \& \ B)$, and

(3) $P(A) \approx P'(A)$, (with both of these positive),

and so

(4) $P(B|A) \approx P'(B|A)$.

Thus, from (1) and (4) we have that at least one of P and P' is a non-CCCP-function for $\rightarrow$.

First, a quibble. The deduction of (4) from (2) and (3) is too quick. Two fractions may have numerators that approximately agree and denominators that approximately agree, without themselves approximately agreeing. For example, if we read 'x $\approx$ y' as 'differs by no more than ε from', for some suitably small $\varepsilon > 0$ that has been chosen, then it is clear that we can have (2) and (3) without (4): simply let $P(A \ \& \ B) = P'(A \ \& \ B) + \varepsilon$, and $P(A) = P'(A) - \varepsilon$, so that the numerators of the conditional probabilities differ, but in the opposite way to the denominators, thus exacerbating the difference in their ratios.[45] Similarly if we read 'x $\approx$ y' as 'x = ky', for some k sufficiently close to 1.

---

[45] If more details are wanted:

Suppose that $P(A \ \& \ B) = P'(A \ \& \ B) + \varepsilon$, and $P(A) = P'(A) - \varepsilon$; then

$P(A \ \& \ B) \approx P'(A \ \& \ B)$, and

$P(A) \approx P'(A)$, but

$P(B|A) \ = \ \dfrac{P'(A \ \& \ B) + \varepsilon}{P'(A) - \varepsilon}$

$> \ \dfrac{P'(A \ \& \ B) + \varepsilon}{P'(A)}$

To be sure, this small flaw could be patched up—and it is only a quibble, because it is not sufficient to cast doubt on the fact that the probability of the conditional must come apart from the conditional probability for at least one of P and P', and of course it is that which really matters. The reservations that I have about the argument stem more from its employment of several unnecessarily strong assumptions; the result which is proved is correspondingly unnecessarily weak. Carlstrom and Hill show how a particular assignment of probabilities to worlds leads to trouble for the Hypothesis. They seem to be content to establish that a non-CCCP-function for a given →exists— but *that* is old news (not that they claim otherwise). They do not show just how restricted the family of CCCP-functions for this →must be.

Nevertheless, their argument can be recast in more general form, and then it begins to look like my intuitive argument. Carlstrom and Hill essentially give us an example of a perturbation relative to →. Start with P's distribution of mud, which heaps roughly half the mud on a world w_2 outside A→B, and move all of this mud (or, more in the spirit of their argument, 'roughly' all of it) to a corresponding world w_3 inside A→B— 'corresponding', in the sense that it agrees with w_2 on the truth value of A♥B, where as before this stands for one of A&¬B, A&B, or ¬A. This produces P', a particular perturbation of P relative to →. They prove that two probability functions, related by this specific sort of perturbation, cannot both be CCCP-functions.

The proof lacks generality in so far as it makes what amount to the following unnecessarily strong assumptions:

A1. A→B is not a truth function of A and B.

$$> \frac{P'(A \ \& \ B) + \varepsilon P'(A)}{P'(A)}$$

$$= P'(B|A) + \varepsilon,$$

so P(B|A) > P'(B|A) + ε, violating (4).

A2. There is a world w_1 at which A & B and A$\rightarrow$B are true, which gets probability 1/2 (roughly) from both P and P'.

A3. w_2 is a world inside $\neg$(A$\rightarrow$B) & A$\heartsuit$B, w_3 is a world inside (A$\rightarrow$B) & A$\heartsuit$B, and P assigns probability 1/2 (roughly) to w_2.

A4. All of w_2's mud (or roughly all of it) is shifted to w_3.

We may simply drop A1: we need no assumptions about the logic of the $\rightarrow$, as I have shown. Furthermore, w_1 actually does no work in the argument, and so A2 should also be dropped—in its place, it suffices simply to assume that P and P' both give A positive probability (to guarantee that the conditional probability does not go undefined). In place of A3, we may assume that there is a non-empty proposition W_2 that implies $\neg$(A$\rightarrow$B) & A$\heartsuit$B, a non-empty proposition W_3 that implies (A$\rightarrow$B) & A$\heartsuit$B, and that P assigns *some* positive probability to W_2 (or to W_3). Finally, in place of A4, we may assume simply that *some* non-zero amount of W_2's mud is moved to W_3—or vice versa.

Weakening assumptions A1–A4 in the ways that I have indicated, and checking all the cases, we essentially reproduce my argument of §8.2.1. I say "essentially"— I do assume that P and P' agree *exactly* outside A$\heartsuit$B, whereas Carlstrom and Hill only assume that they agree approximately. But this only renders my argument immune to the quibble that I had with their argument—and in any case, the notion of approximate agreement is too vague as it stands to be used in a rigorous argument.

The virtue of Carlstrom and Hill's argument over Lewis', in Appiah's opinion, is that it obviates the need for irregular probability functions (recall our discussion at the end of §5.3.2), although Carlstrom and Hill don't raise the issue themselves. In their argument, they concentrate *roughly* all of the probability on certain pairs of worlds— but this leaves a little bit left over that presumably can be spread over the rest of the worlds. Clearly, my result also gives Appiah what he wants in this regard.

Commentary; a plea for perturbation

I said earlier that I would show that various *important* operations on CCCP-functions produce non-CCCP-functions. Why is perturbation important? A perturbation transforms a probability function to another one that is, in a certain sense, similar to it—one that agrees with it everywhere except on some region. (Just how similar the two functions are depends on how much they differ on this region.) This means that whatever the class of CCCP-functions for a given $\rightarrow$ might be, no two functions in that class are similar in this sense. It also suggests a certain 'gappiness' of the CCCP-functions for the $\rightarrow$ among probability functions in general: 'around' each CCCP-function, there is a vast cluster of non-CCCP-functions that are otherwise similar to it.

The perturbation theorem has unpleasant consequences for any agent who wants to conform to (CCCP). For example, his probabilities for propositions that imply A (whatever A may be) fully determine all probabilities of propositions of the form $(A\rightarrow\_\_) \& \neg A$, for all ways of filling in the blank, despite the fact that *these propositions lie totally outside A.* So once we know his distribution over A, we can derive his probabilities for all such propositions. Suppose we are told his probabilities for all propositions that imply A, but that's all. Of course, it's no surprise that, given this knowledge, we can derive his probability for $\neg A$ (namely, $1-P(A)$). What *is* surprising is that we can derive his probabilities for all conditionals that have A as antecedent, when these surely depend (at least to some extent) on his distribution outside A, about which we are given no details. Furthermore, he cannot change the probabilities that these conditionals have outside A, without changing his distribution inside A.

I hope that the notion of perturbation has already proved its worth in enabling the refutation above of the restricted universal version of Hypothesis. However, I think

that it has many other uses. Here I want to discuss some other ramifications of the perturbation theorem: the argument that it provides against the restricted belief function version of the Hypothesis, its application to sensitivity analysis, perturbations of the boundaries of propositions, and the unification that it provides of Lewis' triviality results. (As I have said, I will postpone until §9 a discussion of two of its most important applications: to *approximate* versions of the Hypothesis, and to the Hypothesis cast in terms of vague opinion.)

<u>An argument against the restricted belief function version</u>

Firstly, as I just foreshadowed, there is the doubt that the perturbation theorem casts on the belief function version of the Hypothesis, and more importantly, on the restricted belief function version (and since the former has been shown in §5–§7 to be untenable, I will imagine the advocate of the Hypothesis as defending the latter, for some suitable restriction). On this view, no two probability functions that represent the states of mind in the epistemic history of a rational agent can be perturbations of each other, unless the $\rightarrow$ changes appropriately. Nor can the probability functions of two different rational agents be perturbations of each other relative to some $\rightarrow$. To save the restricted belief function version, we need to be convinced that at most one of any two probability functions related by perturbation can be a belief function. But this will be no easy task. A perturbation of a probability function can differ ever so slightly from that function—for example, agreeing with it virtually everywhere, and elsewhere disagreeing with it only beyond the hundredth decimal place. It is hard to see how one of these functions can be a belief function, and the other not.

The friend of the Hypothesis may insist that we follow Lewis' suggestion that "the content of a total mental state is the system of belief and desire that best rationalizes the behavior to which that state would tend to dispose one" (1986, p. 585), and that while a

non-CCCP-function may fit an agent's mental state, it could not *rationalize* it.[46] (After all, he may remind us, the restricted belief version does see (Restricted CCCP) as a constraint on rationality.) Thus it is no embarrassment for the Hypothesis that one function is a belief function, while a very near neighbor of it isn't—the extent to which they fit a certain agent's mental state is roughly the same, but while one rationalizes the behavior to which that state would tend to dispose one, the other doesn't. Or so the argument goes.

This sort of move is always open, although it is one that would raise the hackles of anyone with Popperian sympathies—little surprise that (CCCP) holds throughout the class of all belief functions (on an appropriately restricted domain), because any function that fails to conform to it is thereby disqualified from being a belief function. In the absence of good arguments to support this, it seems that the restricted belief function version is being rescued, as the old joke goes, simply by stipulating that there are no counterexamples to it!

In any case, functions related by perturbation don't *have to* be very near neighbors. For example, they may both agree that A should be given a certain very tiny probability, but disagree wildly on how probability should be distributed within ¬A. We need to be convinced once again that at most one of those functions can be a belief function, and frankly, I don't see how that can be done in a non-question-begging way, without an appeal to the Hypothesis itself. For I see no reason to doubt that two such functions can be belief functions, apart from a prior faith in the Hypothesis. To be sure, there may be a *consistent* package deal of the Hypothesis plus elevated standards of what it takes to be belief function. But consistency is no great virtue of a position (we learned from Quine that you can hang onto virtually any cherished belief if you are

---

[46] Lewis offers the friend of regularity such a response in the face of his third triviality result.

prepared to twist and turn enough elsewhere). What this position lacks, in my opinion, is *plausibility*.

Robustness, sensitivity analysis[47]

Let's move on to other applications of the perturbation theorem. Following Walley (1991, p. 5), for example, we may call the conclusion of a statistical analysis *robust* if that same conclusion is reached by the use of each member of a wide class of probability functions. Those who are uncomfortable about the Achilles heel of Bayesianism that the choice of priors is arbitrary, but who are otherwise sympathetic to the Bayesian enterprise, may find robustness to be a useful notion. The robustness of an inference or decision can be checked by performing a so-called 'sensitivity analysis'. Here's the idea. Take a large class of priors, and derive posterior distributions from each of them (by conditioning, say, or by a specified sequence of conditionings). If the same inference or decision is reached on the basis of all of these posterior distributions, then it is robust: the choice of prior was not critical. (See Walley for further discussion and references.)

Suppose that included among the modifications of a given prior are perturbations of that prior (relative to some →). While a specific choice of prior may, for all we have learned so far, be a CCCP-function for that →, these modifications will not be. Conformity to (CCCP) forces one to be very careful that the prior, and its modifications, are just right.

Perturbations of the boundaries of propositions

Here is another use for perturbations. Keeping in mind the muddy Venn diagram, notice that perturbations involve movement of mud *relative to the boundaries of propositions*. I assumed above that the boundaries of the propositions were fixed, and

---

[47] I thank Bas van Fraassen for drawing this application to my attention.

that it was the mud that moved; but the same effect could be achieved instead by keeping the distribution of mud fixed, and moving the boundaries, so to speak. Here's one way that this might come about. Suppose that the → depends partly on the similarity relation on worlds. Imagine an agent who changes her mind about the similarity relation, without actually learning anything about the world in which she lives. She might, for example, read Lewis' *Counterfactuals*, and become persuaded that sameness of laws is a more important respect of similarity of worlds than she previously thought; or she might be in some context in which a similarity relation that is finer grained than before becomes appropriate. In neither case does she learn anything about *her* world—her assignment of probabilities to individual worlds is as it was before. But for at least some A and B, A→B will be a different set of worlds to what it was previously. This we might picturesquely call *a revision in the boundary of* A→B (after all, that's how we would picture it on the muddy Venn diagram).

Certain revisions in the boundaries of propositions of this sort would have the same effect as perturbations of her old probability assignment. We could even extend our usage of the word, and call such revisions 'perturbations' of the propositions concerned. For example, a perturbation of (A→B) & ¬A involves a revision in the boundary of that proposition, such that the new assignment of probabilities to propositions is a type iii (→, A, B)-perturbation of the old. Note that there is always a conditional that is involved in a perturbation—and it is propositions involving conditionals, above all, that we should reasonably expect to undergo occasional revision in the course of an agent's deliberation (since their truth depends partly on what the agent takes the similarity relation to be, something that is not fixed simply by the facts about the worlds themselves).

There is yet another way to achieve the same effect of relative mud movement: move *both* the mud, *and* the boundaries of propositions at the same time. Indeed, someone who believes that the conditional is radically context-dependent, changing

with every change of mind of the agent who entertains it, believes that this is how things really are. Again, certain of these joint movements will have the same effect as perturbations of probability functions (and we could express their kinship by giving them the same name). I will discuss them again at the end of this section.

Unifying Lewis' triviality results

There is an interesting connection between conditioning and perturbation. Consider the special case of perturbation in which all of the mud, x say, is removed from a region W that implies ¬(A→B) & A♥B; and then all of it is deposited in some way on a region W' that implies (A→B) & A♥B, and that previously had no mud on it. (This could also be done in the reverse direction, with suitable changes in the discussion that follows.) Let P be the distribution before the amount x of mud is shifted, and P' the distribution afterwards. Then there is an ur-distribution P_0, such that both P and P' are derived from P_0 by conditioning. Here's the idea: P_0 is a certain distribution that gives equal probability to W and W'. Punch a hole through W' and renormalize, and you have P; punch a hole through W and renormalize, and you have P'.[48] Precisely: P_0 is the distribution that gives the same odds as P and P' do to all propositions that are incompatible with both W and W', the same odds to propositions that imply W as P does, and the same odds to propositions that imply W' as P' does, but which gives probability $\frac{x}{x+1}$ to both W and W'. P is derived from P_0 by conditioning on ¬W'; P' is derived from P_0 by conditioning on ¬W.

Suppose that the set of CCCP-functions for some → contains P_0. Then we have found that it is not closed under conditioning—for it doesn't contain both P and P'. Now suppose instead that the set *is* closed under conditioning (as Lewis (1976) does in order to derive his second triviality result). Then we know that it doesn't contain any probability function that gives the same non-zero probability to a proposition inside

---

[48] It was David Lewis who pointed out (a close relative of) this fact to me.

(A→B) & A♥B as it does to a proposition inside ¬(A→B) & A♥B —for if it did, we could take that function to play the role of P_0, and use the above argument and the perturbation theorem to show that one can derive a non-CCCP-function from it by conditioning, *contra* the closure assumption.

The trouble is that we have good reasons for thinking that the set of CCCP-functions for a given → *does* contain a probability function that gives the same non-zero probability to some proposition inside (A→B) & A♥B as it does to some proposition inside ¬(A→B) & A♥B, for some A, B, and ♥. Indeed, this would seem to be the norm for CCCP-functions, since it is known that they have very rich domains. I showed in §7.2 that, without any assumptions about the logic of the →, all CCCP-functions have infinite domains; and I mentioned in §7.4 that Hall (1993) shows that, with minimal assumptions about the logic, all CCCP-functions have uncountable domains—indeed, that any probability space whose probability function is a CCCP-function is *full*.

So let P_0 be a CCCP-function that gives positive probability to both (A→B) & A♥B, and ¬(A→B) & A♥B, for some A, B, and ♥ —it would surely trivialize the set of CCCP-functions for this → if there were no such function. Since its probability space is full (making Hall's assumption about the →'s logic), there is a proposition W that implies ¬(A→B) & A♥B, and a proposition W' that implies (A→B) & A♥B, such that $P_0(W) = P_0(W') > 0$.[49] However, $P = P_0(–|¬W')$, and $P' = P_0(–|¬W)$ are perturbations of each other, so they cannot both be CCCP-functions. Thus, the set of CCCP-functions for this → is not closed under conditioning. Our study of perturbations has enabled us to produce what is basically a new proof of Lewis' second triviality result.[50]

---

[49] In case this is not already obvious: suppose $P_0(¬(A→B)$ & $A♥B) \leq P_0((A→B)$ & $A♥B)$. Then take W to be ¬(A→B) & A♥B; by the fullness of the space, we are guaranteed the existence of an appropriate W'; and similarly if the inequality goes the other way.

[50] Only "basically", because "triviality" in this case does not mean the same thing as it does in Lewis' result.

Conditioning is a special case of Jeffrey conditioning, which more generally involves a change in the probability weights given to the members of some finite partition, while preserving the odds of all propositions within each of the cells of the partition. Various perturbations are Jeffrey conditionings. Of course, any perturbation for a model that has finitely many worlds is, for the cheap reason that *any* mud movement whatsoever in such a model is: take the partition to consist of the individual worlds themselves, so that any mud movement will involve such a reassignment of weights to the members of the partition. Similarly, in models with infinitely many worlds, there are also perturbations that are trivially Jeffrey conditionings: for example, a perturbation involving only finitely many worlds of that model, with the partition consisting of those worlds, and the union of everything else.

But there are less trivial examples of perturbations that are Jeffrey conditionings. Suppose that for some A, B and ♥, and for some W that implies ¬(A→B) & A♥B, and some W' that implies (A→B) & A♥B, we shift an amount of mud from W to W', in such a way that the odds of all propositions that imply W and that imply W' remain the same. This is a perturbation which is a Jeffrey conditioning over the partition {W, W', ¬(W v W')}, with the weights of only the first two cells changing. Such perturbations are useful in modeling belief change that is not global (the way that conditioning is), but local—when an agent learns something that gives probabilistic information about a specific region of logical space, without this information having effects that spread across the whole space.

And we can generalize further. Consider an m-celled partition of (A→B) & (A♥B), giving us an (m+2)-celled partition of the whole space—the (m+1)st cell being ¬(A→B) & (A♥B), and the (m+2)nd cell being ¬(A♥B). Suppose we shift various amounts of mud from cells 1 through m, to cell m+1 (as we can in appropriately non-trivial cases), in such a way that the odds of all propositions within these cells remain the same. This is a Jeffrey conditioning over the (m+2)-celled partition; the same is

true if we move the mud in the opposite direction. And again, since it is a perturbation, if the prior function (before the shift) is a CCCP function for the →, the posterior function (after the shift) isn't.

I have thus given an alternative proof of the corollary to Lewis' fourth triviality result that I discussed at the end of §5.4.2, for m+2 = n ≥ 3:[51] there is no CCCP-conditional for a class of probability functions closed under n-celled Jeffrey conditioning, for each n ≥ 3, unless that class consists entirely of trivial functions.

I have said enough about perturbations—the illustration they provide of a general method for refuting the restricted universal version of Hypothesis, their role in generalizing the Carlstrom and Hill argument, their interest in their own right, and their relation to important forms of belief revision. I promised that the method for refuting the restricted universal version generalizes to other interesting operations on probability functions. Let's see how.
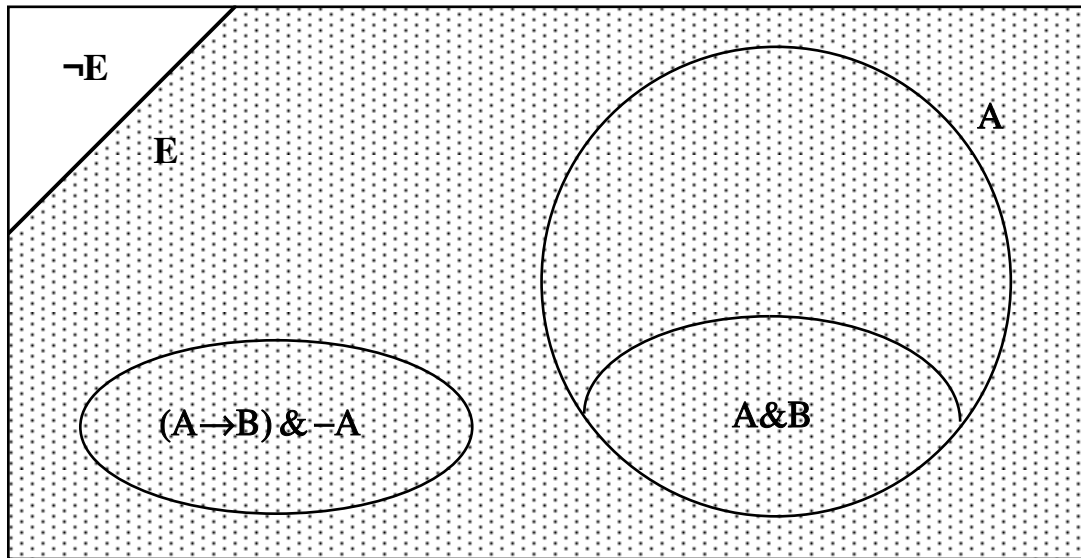

8.2.2 Conditioning[52]

I began my discussion of conditioning in the previous sub-section and it is now time to say more. Start with propositions A, B, and A→B. Suppose P is a CCCP-function for this →, and conditionalize it on a proposition E, arriving at a new probability function P_E. Then there are many choices of E such that $P_E(A{\to}B) \neq P(A{\to}B)$, and yet $P_E(B|A) = P(B|A)$. Thus, P_E will be a non-CCCP-function—and indeed it will violate the equation in (CCCP) for this very choice of A→B.

---

[51] Again, modulo the different meaning of 'triviality' here.

[52] I thank David Lewis and Lyle Zynda for helpful discussion.

1. Here is a simple example of such an E:

E is implied by (A→B) v A, with $0 < P(E) < 1$. (It is not essential that (A→B) & ¬A be non-empty—the diagram may lack the ellipse.) Conditioning on E amounts to what we might call 'trimming' the original diagram, and renormalizing what remains so that it once again has a total of one unit of mud. The probability of A→B after conditioning is greater than it was before, because A→B receives a greater proportion of the new distribution of the mud than it did of the old. But the corresponding conditional probability is unchanged by the conditioning. This is obvious enough when it is noted that the proportion of A mud that falls inside A&B remains unchanged, but I should prove it:

Suppose that E is implied by (A→B) v A, that $P(E) < 1$, and that $P(A→B) = P(B|A)$. Then:

$P_E(A→B) = P(A→B|E)$

$$= \frac{P((A\varnothing B) \,\&\, E)}{P(E)}$$

$$= \frac{P(A \oslash B)}{P(E)} , \text{ since } A \to B \text{ implies } E$$

$$> P(A \to B), \text{ since } P(E) < 1.$$

Thus

$$P_E(A \to B) \; > P(A \to B).$$

But

$$P_E(B|A) \; = \frac{P_E(A \& B)}{P_E(A)}$$

$$= \frac{P(A \& B|E)}{P(A|E)}$$

$$= \frac{P(A \& B \& E)}{P(A \& E)}$$

$$= \frac{P(A \& B)}{P(A)} , \text{ since } A \& B \text{ and } A \text{ both imply } E$$

$$= P(B|A).$$

Thus

$$P_E(B|A) = P(B|A),$$

and hence

$$P_E(A \to B) > P_E(B|A).$$

Now that we have seen the trick, it is easy to come up with other examples of propositions that a CCCP-function can't be conditionalized on without producing a non-CCCP-function—indeed, one which violates (CCCP) for this very choice of A→B. (I assume that these propositions all have positive probability, so that they *can* be conditionalized on.) I don't see any need to give the proofs for the next examples— they are equally straightforward, and easily read off the diagrams:

2. Let E be any proposition whose probability is less than 1, and whose negation implies:

i.   ¬(A→B) & ¬A,

ii. ¬(A→B) & A&B,

or the conjunction of any such propositions.

These are like 'hole-punches':
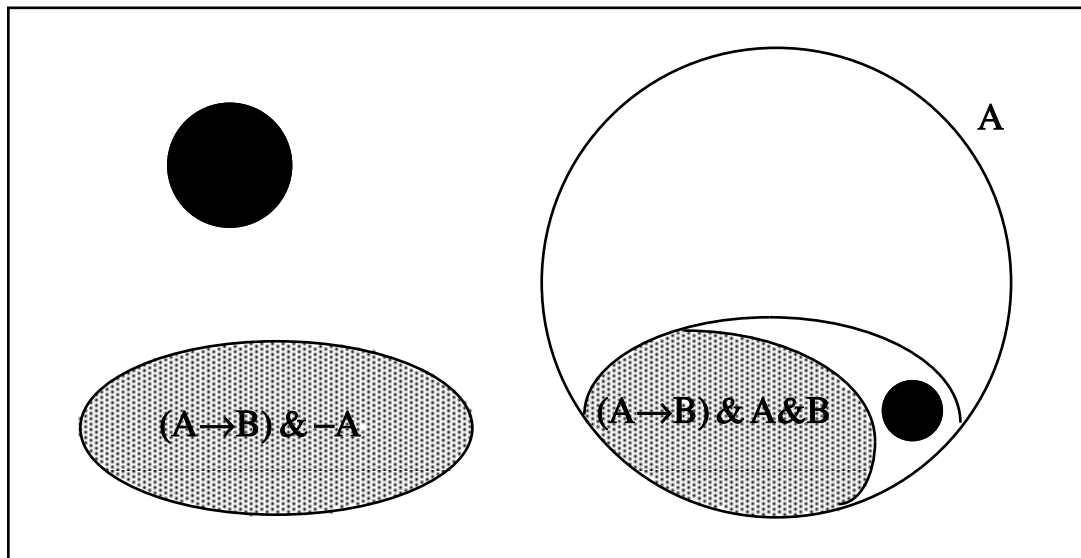
3. Let E be any proposition whose probability is less than 1, and whose negation implies:

i.   (A→B) & ¬A,

ii. (A→B) & A & ¬B,

or the conjunction of any such propositions.

More hole-punches:
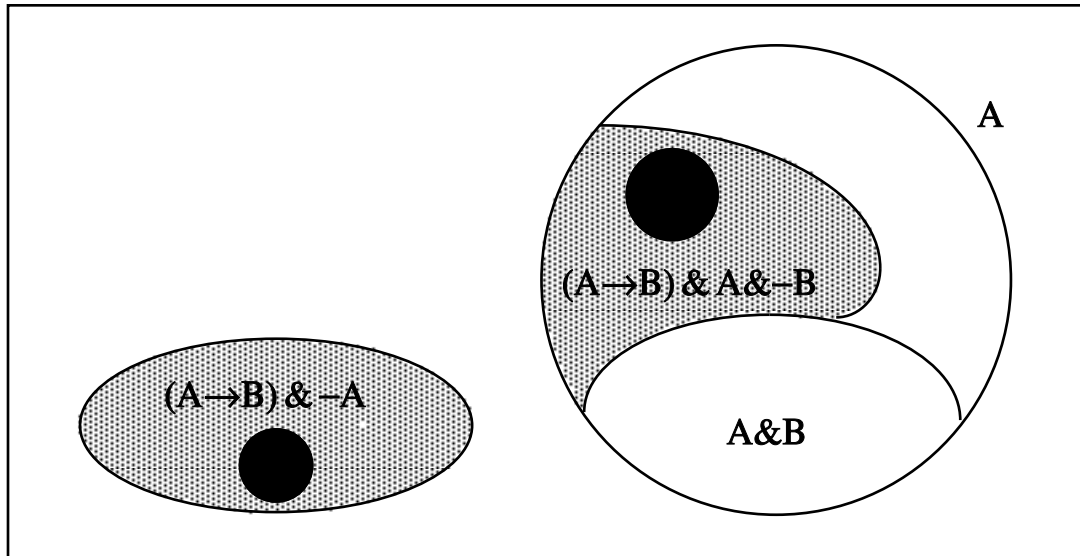
It is worth noting that cases 2.ii and 3.ii—if such there be—do not exemplify the strategy of increasing or decreasing one side of (CCCP) while keeping the other constant. Rather, they're examples of something *worse still*, as far as (CCCP) is concerned. Conditioning on a proposition whose negation implies ¬(A→B) & A&B increases the probability of the conditional while decreasing the conditional probability; conditioning on a proposition whose negation implies (A→B) & A&¬B decreases the former, while increasing the latter. I say "if such there be", because case 2.ii is ruled out if A&B implies A→B, as Stalnaker and Lewis among others think; and very plausibly, A&¬B implies ¬(A→B), and this rules out case 3.ii. The other cases, however, are unproblematic. Apart from cases that we can justly call trivial, given →, A and B, such examples of E can always be found.

We see yet again that for any →, the family of CCCP-functions is not closed under conditioning (apart from trivial cases). In fact, we see something stronger: for any →, A and B, the set of functions for which the probability for A→B equals the corresponding conditional probability is not closed under conditioning (apart from

trivial cases). Thus, we have proven yet another strengthening[53] of Lewis' second triviality result:

Theorem. There is no → such that (Restricted CCCP) holds throughout a class of probability functions closed under conditioning, irrespective of the (non-trivial) restriction that we impose, unless the class consists entirely of trivial functions.

It appears that conditioning is a dicey business if you are an agent who wants to conform to (CCCP). You must be careful that, for any E that you conditionalize on, there is *no* pair of propositions such that E falls into one of the outlawed categories with respect to them. This *might* have seemed manageable if your probability space weren't very rich (so that there wouldn't be that many pairs of propositions to worry about). But your probability space *has to be* very rich—certainly at least infinite, as we've seen in §7, and plausibly even uncountable.

Recalling our discussion of §7.3, these are further examples of propositions fit for Moore's paradox—ones that you regard as possible, but which you nevertheless cannot conditionalize on without violating (CCCP). Van Fraassen (1984) calls propositions that you can't conditionalize on without violating such a structural constraint "Moore propositions". Regarding conditioning as the model of ideal learning, this means that the ideally rational agent could not learn such propositions while adhering to the constraint. That a certain proposition should be a Moore proposition is not surprising if it is 'reflexive' in some appropriate sense (which I won't try to spell out here)—something about the agent's own beliefs, say, or about a future history that a theory of chance says would undermine that very theory (see Lewis (1981)). But what's disturbing for the Hypothesis is that the preservation of (CCCP) turns so *many* propositions into Moore propositions, seemingly without explanation.

In one sense, this is overkill. After all, Hall's strengthening of Lewis' result showed us that conditioning a CCCP-function P on anything other than a P-atom will produce a

---

[53] Again, modulo the different meaning of 'triviality' here.

non-CCCP-function. In other words, the preservation of (CCCP) implies that every proposition that is not a P-atom is a Moore proposition for an agent whose state of mind is represented by P. So the existence of the various Moore propositions outlined above follows immediately.

But enough concession. Having said this, I should now say why my discussion is overkill only *in one sense*. Firstly, I think my method gives us a way of seeing intuitively *why* the preservation of (CCCP) turns various propositions into Moore propositions, whereas this may not be so obvious from the proof of Hall's result, which is more algebraic.

Secondly, and more importantly, recall that Hall's result assumes that the → is uniform across a pair of probability functions: its interpretation is the same, whether it appears within the scope of a certain probability function, or another one reached by conditioning that function. My argument can go through without this assumption, as I will show in my discussion in §8.3. To briefly foreshadow that discussion: radical context dependence is not sufficient to evade my argument, the way that it is sufficient to evade Lewis' and Hall's. In terms of the muddy Venn diagram, radical context dependence means that the boundary of A→B changes with each change in the distribution of mud. To be sure, *some* movements of the boundary do exactly counterbalance the deleterious effect (as far as (CCCP) is concerned) of the mud movements that I have described, preserving the equality between the probability of A→B and the conditional probability of B, given A throughout such a movement. However, some do not counterbalance it enough, some do not counterbalance it at all, and some actually accentuate it. The proponent of the Hypothesis must assume *not only* radical context dependence, but also that the dependence is of the right sort.

8.2.3 Jeffrey conditioning

Conditioning is a special case of Jeffrey conditioning, as I have said, so the examples in §8.2.2 could equally be described as Jeffrey conditionings that take a CCCP-function to a non-CCCP-function. In fact, there are many other Jeffrey conditionings with this property. We saw some of them in §8.2.1, and I will mention a few more here.

To get the idea, consider again the case of conditioning on an E which 'trimmed' the diagram. We removed all the ¬E mud, and renormalized—or equivalently, redistributed the original mud to E—preserving the odds of all propositions that imply E. Now suppose that we 'shave' the diagram instead, moving some but not all of the ¬E mud to E, in such a way that the odds of all propositions that imply ¬E, and of all propositions that imply E, are preserved. This is a Jeffrey conditioning, on the partition {¬E, E}. It is also incompatible with the preservation of (CCCP). For this mud shift increases the probability of A→B, without increasing the corresponding conditional probability. (Obviously, a Jeffrey conditioning in the other direction, which decreases the probability of A→B, would equally violate (CCCP).)

More precisely, let →, A and B be given, let P be a CCCP-function for this →, let E be implied by (A→B) ∨ A, and let P' be a function derived from P by a Jeffrey conditioning on the partition {¬E, E} that increases the probability of E. We have:

$$P'(A{\to}B) = P(A{\to}B|E)P'(E) + P(A{\to}B|¬E)P'(¬E)$$

$$= \frac{P(A\oslash B)}{P(E)} \cdot P'(E), \text{ since } A{\to}B \text{ implies } E$$

$$> P(A{\to}B), \text{ since by hypothesis } P'(E) > P(E).$$

Hence

(1)   P'(A→B) > P(A→B).

However,

$$P'(B|A) = \frac{P'(A \& B)}{P'(A)}$$

$$= \frac{P(A \& B|E)P'(E) + P(A \& B|\neg E)P'(\neg E)}{P(A|E)P'(E) + P(A|\neg E)P'(\neg E)}$$

$$= \frac{\dfrac{P(A \& B)}{P(E)}P'(E)}{\dfrac{P(A)}{P(E)}P'(E)} \text{, since A, and hence A \& B, implies E}$$

$$= \frac{P(A \& B)}{P(A)}$$

$$= P(B|A).$$

Hence

$$P'(B|A) = P(B|A),$$

and so by (1), and the fact that P is a (CCCP) function for →,

$$P'(A{\to}B) > P'(B|A).$$

So such a Jeffrey conditioning is incompatible with the preservation of (CCCP), and indeed any function so produced violates (CCCP) for this very choice of A and B.

Clearly, any of the Moore propositions listed in §8.2.2 will play the role of such an E here—giving us various {¬E, E} partitions on which we cannot Jeffrey conditionalize.

This in turn implies that for any →, A and B, the class of functions for which the probability for A→B equals the corresponding conditional probability is not closed under two-celled Jeffrey conditioning (apart from trivial cases). But then an argument similar to the one that I gave for the corollary to Lewis' fourth triviality result shows that it is not closed under n-celled Jeffrey conditioning, for each n. Thus, we have a result that strengthens[54] that corollary:

---

[54] Again, modulo the different meaning of 'triviality' here.

Theorem. For each n, there is no →such that (Restricted CCCP) holds throughout a class of probability functions closed under n-celled Jeffrey conditioning, irrespective of the (non-trivial) restriction that we impose, unless the class consists entirely of trivial functions.

8.3 Assessment: a loophole?

We have found that, assuming nothing about the logic of the conditional, all perturbations, some conditionings, and some (further) Jeffrey conditionings performed on a CCCP-function produce non-CCCP-functions (trivial cases aside). Moreover, specify any particular A, B and → that you like, and a probability function P for which P(A→B) = P(B|A), and we can produce a P' by application of any of these operations on P, such that P' violates (CCCP) for this very choice of A, B and →.

Now let us return to the argument of §8.2.1, which is a good example of the method that I have employed—the intuitive argument involving perturbation (or if you prefer, its more rigorous reformulation). I have discussed how it generalizes the Carlstrom and Hill result, and how it can be used to unify the results of Lewis. Now I want to draw further comparisons to the latter. Like Lewis, I assumed nothing about the logic of →. Lewis showed that any class of CCCP-functions closed under conditioning (restricted conditioning, Jeffrey conditioning) consists entirely of trivial functions. I derived analogous results—for example, that any class of CCCP-functions closed under perturbation consists entirely of trivial functions. However, as Hall's strengthening of Lewis' result shows, conditioning a CCCP-function P on a proposition that is not a P-atom produces a non-CCCP-function. And this parallels my result too (replacing 'conditioning ... on a proposition that is not a P-atom' by 'perturbing').

To resume our discussion of §5.3.2 and §8.2.2: the assumption of Lewis' that has met with the most criticism is the fixed interpretation of the →—the assumption that Stalnaker calls 'metaphysical realism'. It may be thought—and van Fraassen (1976), as

we have seen, does think—that the → is inextricably tied to the probability function in whose scope it appears: change the function, and you change the →. Radical context dependence is an antidote to Lewis' results, even after Hall's strengthening—not an antidote that I can swallow myself, but others can.

It might seem that radical context dependence is also an antidote to my intuitive result from perturbation—and hence to the results that follow it. To a certain extent, this must be right. After all, van Fraassen has shown us how (CCCP) can be saved in just such a manner. Let's see how context dependence can disarm my intuitive argument, by considering the following diagram.
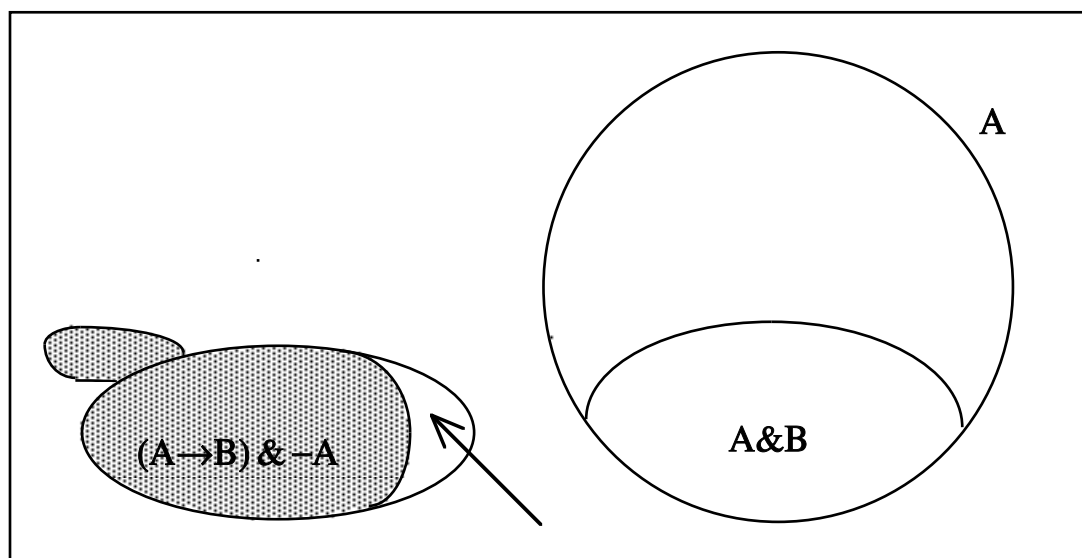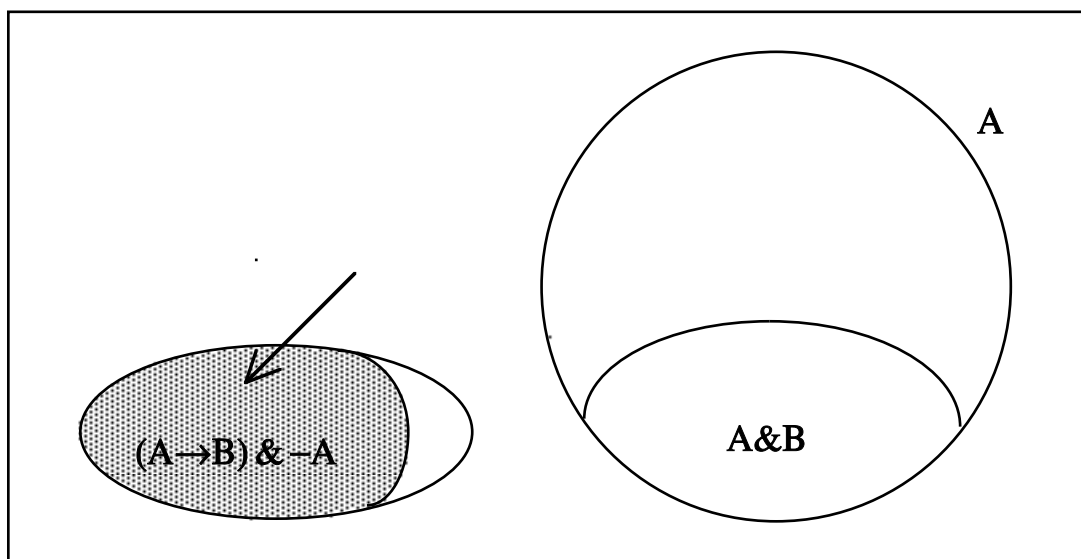

Dodging the mud

Suppose that some mud that currently lies outside (A→B) & ¬A is to move to certain worlds that are currently inside (A→B) & ¬A. A perturbation threatens. However, it can be prevented if the boundary of (A→B) & ¬A moves too, in just such a way as to 'dodge' the approaching mud.

But not just any movement of the boundary will do. The probability of A→B must not change in any way despite the movement—that would be equally unpropitious for

the Hypothesis, for the conditional probability still remains unchanged. So while the boundary retreats to avoid the scene of the mud movement, it must compensate elsewhere, invading territory that it did not occupy previously.

Alternatively, A→B could 'accept' the moved mud, provided its boundary shrank so as to keep the total amount of mud inside constant:

Accepting the mud

Finally, there are intermediate cases, in which A→B changes so as to dodge some of the moved mud, while accepting the rest. (In these cases, not all of the mud is moved to a single world.)

And so it is with the various other forms of perturbation: if the boundary of the conditional always moves in just the right ways, in tandem with the movement of mud—which is to say that the conditional is always reinterpreted in just the right ways, in tandem with the change in probability distribution—then all is well for the Hypothesis. However, this should not necessarily provide the friend of the Hypothesis much solace. As I have already pointed out, the effect of a perturbation can be achieved *even if* the → is radically context-dependent—a joint mud/boundary shift can

still be trouble. In fact, certain boundary shifts only aggravate the effect of a given mud movement. (A shift that invades new territory *and* accepts all of the moved mud is an example of this.) And certain other boundary shifts, while not aggravating the effect of a given mud movement, do not counterbalance it enough. The only hope for the upholding of (CCCP) is if the boundary shifts are always of just the right kind.

This is as true for the cases of conditioning and Jeffrey conditioning as it is for the cases of perturbation. Consider the case of 'trimming' the diagram, which raises the probability of A→B, without raising the corresponding conditional probability. Suppose that the → also changes with the change in probability distribution, in such a way that the new A→B includes all the worlds that it did previously, *plus some new worlds* that receive non-zero probability. This is doubly bad news for (CCCP): the probability of A→B is raised both by the effect of the conditioning, *and* by the growth in its boundary, while the conditional probability remains unchanged. In a similar vein, my argument can allow the boundary of A→B to change simultaneously with a conditioning on any of the other Moore propositions catalogued above, if the change accentuates the deleterious effect (as far as (CCCP) is concerned) of the conditioning; or the effect of the change in the boundary can even counteract the effect of the change due to the conditioning, provided the two do not exactly cancel.

In this respect, I think that my result buys triviality more cheaply than Lewis' results, and more cheaply than even Hall's strengthening of Lewis' result: the fixed interpretation of → across (at least) a pair of probability functions is necessary for their results, but not for mine. The upholder of (CCCP) can appeal to radical context-dependence of the → to elude Lewis' results (although I argued in §5.3.2 that it's not a plausible appeal); but he has to appeal to something even stronger to elude mine. What might that be? Why, the upholding of (CCCP) of course! After all, it is the upholding of (CCCP) that would guarantee that the → is sensitive to the probability function in just such a way that (CCCP) is upheld. But I do not see how *else* he could justify the

assumption that the $\rightarrow$ not only depends on the probability function, but that it depends on the probability function in just the right way. I concede once again that his position is consistent—van Fraassen's positive result assures us of that—but again, *that* is scant praise for it.

9. Can the Hypothesis be saved?

This is where my exposition of the triviality results will stop. At some point short of complete demolition of the Hypothesis, *all* triviality results have to stop—the search for knock-down refutations of all versions of the Hypothesis would be futile, because as van Fraassen has shown us, suitably construed it *is* tenable. However, for reasons that I have given, I don't think that the suitable construal upholds the spirit of the Hypothesis as it was initially conceived—a hypothesis that illuminates the semantics of the conditional of natural language.

I suggest to the diehard supporter of the Hypothesis that the thing to do is to loosen the statement of (CCCP). Then, some of the various versions of the Hypothesis, suitably weakened, become live options again. In the previous section, we saw one unsuccessful attempt to do this: weakening (CCCP) to (Restricted CCCP). I will now look at some further strategies along these lines.

I argued at length against the radical context dependence of the conditional in §5.3.2. The arguments mainly turned on the point that if a conditional that we used were radically context dependent, we should expect to see various phenomena that we don't in fact see. No natural language conditional is context dependent in that way, as I argued. Thus, I won't be discussing any proposals that assume radical context dependence, or indeed any context dependence that would predict similar phenomena.

For example, Jeffrey and Stalnaker (1993), developing an idea by de Finetti (1936), offer a 'conditional' that conforms to a variant of (CCCP). The idea is to regard propositions as random variables—functions from worlds to real numbers—that take the values 0 or 1, and the conditional as expressing a random variable that takes values in the interval [0,1]. The random variable takes the same value as the consequent (0 or 1) when the antecedent is true, and the same value as the corresponding conditional probability when the antecedent is false. It follows from this that the expectation value

of a conditional is equal to the corresponding conditional probability. (I call this a 'variant' of (CCCP), because expectation is a generalization of probability.)

This is a very interesting result, and Jeffrey and Stalnaker synthesize a lot of the literature in their paper. However, for reasons that parallel those in §5.3.2, I doubt that the de Finetti-Jeffrey-Stalnaker conditional is a conditional that we use in natural language (not that Jeffrey and Stalnaker make any such claims for it).

Let's move on, then, to the three strategies that I want to discuss in some detail.

9.1 Retreat to an approximate version of (CCCP)

Instead of demanding the exact equality of conditional probability with the probability of a conditional, we might demand only that they *approximately* equal each other:

(Approximate CCCP) $P(A{\rightarrow}B) \approx P(B|A)$ for all A, B in the domain of P, with

$P(A) > 0.$

The idea is then to amend the original versions of the Hypothesis, replacing the exact (CCCP) with (Approximate CCCP). Call these versions of 'the Approximate Hypothesis'. For example, the universal version of the Approximate Hypothesis is:

There is some $\rightarrow$ such that for all P, (Approximate CCCP) holds.

I have not yet said how 'approximately' is to be understood—how the '$\approx$' is to be read. Here's one reading:

Reading 1. $P(A{\rightarrow}B) \approx P(B|A)$ iff $P(A{\rightarrow}B) = kP(B|A)$, where $k \approx 1$.

Here's another:

Reading 2. $P(A{\rightarrow}B) \approx P(B|A)$ iff

$P(B|A) - \varepsilon < P(A{\rightarrow}B) < P(B|A) + \varepsilon$, where $\varepsilon \approx 0$.

I'll stop there, even though there may be other reasonable construals of the Approximate Hypothesis. If we're not careful, new versions of the Hypothesis could proliferate out of control: we have various orders of quantifiers, various restrictions on

the domains of quantification, and now various meanings of 'approximate' that we could canvass. I don't think doing so would be profitable. I will limit myself to the two readings of '$\approx$' that I have suggested, and restrict my attention to a small number of salient versions, and rest content with that. Nonetheless, given the sorry fates that these construals will meet, I think that this will shift an unwelcome burden onto the proponent of the Hypothesis to provide one that fares better.

The retreat to the Approximate Hypothesis could be principled. Echoing words that we heard in §8.1, its proponent might insist that all the linguistic data that we can garner is unable to distinguish between the (exact) Hypothesis, and the Approximate Hypothesis—at least if the approximation is sufficiently good. In practice, the probability of a conditional is always so close to the corresponding conditional probability that we cannot tell the two apart. It was a mistake to think that the two coincide in general, as the various triviality results have taught us; but approximate coincidence will save the phenomena (our intuitions about what 'sounds right', about the independence of conditionals from their antecedents, and so on) equally well, and that much may well be tenable. Or so the argument goes.

The two readings of 'approximate equality' that I have offered are still somewhat vague (how close does k have to be to 1, or ε to 0, to be considered 'close enough'?). Still, as I hope to make clear, under reasonable resolutions of the vagueness the Approximate Hypothesis will face some serious problems.

Take reading 1 first. The Approximate Hypothesis, read that way, will founder on an appropriately revised version of Hall's strengthening of Lewis' result:

Suppose $<W,F,P,\to>$ and $<W,F,P_C,\to>$ are distinct non-trivial models, with P_C derived from P by conditioning on C, and that both of them conform to (Approximate CCCP) (reading 1). Their distinctness implies that $P(C) < 1$. Non-triviality of the latter guarantees that C is not a P-atom. Since C is not a P-atom, there is some D properly contained in C such that $0 < P(D) < P(C)$.

Let $E = D \cup \neg C$. We have

$$P(E \to \neg C) \quad = P(E \to \neg C | C)P(C) + P(E \to \neg C | \neg C)P(\neg C)$$

$$= P_C(E \to \neg C)P(C) + P(E \to \neg C | \neg C)P(\neg C)$$

$$= k_1 P_C(\neg C | E)P(C) + P(E \to \neg C | \neg C)P(\neg C), \text{ for some } k_1 \approx 1$$

$$= 0 + P(E \to \neg C | \neg C)P(\neg C)$$

Applying (Approximate CCCP) to the left-hand side, we therefore have

$$k_2 P(\neg C | E) \quad = P(E \to \neg C | \neg C)P(\neg C), \text{ for some } k_2 \approx 1$$

but since $\neg C \subset E$, and thus $\neg CE = \neg C$, this becomes

$$k_2 \cdot \frac{P(\neg C)}{P(E)} \quad = P(E \to \neg C | \neg C)P(\neg C).$$

Multiplying both sides by $\frac{P(E)}{P(\neg C)}$ (as we can, since $P(\neg C) > 0$), this yields

$$k_2 \quad = P(E \to \neg C | \neg C)P(E),$$

and hence

$$1 \quad \approx P(E \to \neg C | \neg C)P(E).$$

Hence $P(E) \approx 1$.

But all we assumed about E was that it had the form $D \cup \neg C$, where $D \subset C$ and $0 < P(D) < P(C)$. So *anything* of this form has probability approximately equal to 1. This is embarrassing for the Approximate Hypothesis.

It becomes more than embarrassing if we assume that $<W,F,P,\to>$ is full (and Hall (1993) makes this appear to be the case, as I have remarked). In that case, we can find a $D \subset C$ with any probability less than $P(C)$ that we like—in particular, with as small a probability as we like. In that case, with $D \cup \neg C$ having probability approximately equal to 1, $\neg C$ itself must have probability approximately equal to 1. So if we begin with a probability function P that conforms to (Approximate CCCP), and conditionalize it on a proposition C, then P must give C probability approximately equal to 0 if P_C is also to conform. The agent whose state of mind is to obey (Approximate CCCP), and who learns by conditioning, can only learn things that he considers extremely unlikely!

This is absurd. Put another way: P, while perhaps not quite a trivial function, certainly approximates one.

If we further assume that the same $\rightarrow$ is employed in $\langle W, F, P_{\neg C}, \rightarrow \rangle$ as well, and that it too conforms to (Approximate CCCP), then we can run a very similar argument to the one above to show that P must give $\neg C$ very low probability as well. Worse than absurd, this is an outright contradiction.

Now take reading 2 of '$\approx$'. A similar argument reveals the problem with the Approximate Hypothesis, construed that way. Let C, D and E be as before.

$$
\begin{aligned}
P(E \rightarrow \neg C) &= P(E \rightarrow \neg C | C)P(C) + P((E \rightarrow \neg C) \,\&\, \neg C)^{55} \\
&= P_C(E \rightarrow \neg C)P(C) + P((E \rightarrow \neg C) \,\&\, \neg C) \\
&< P_C(\neg C | E)P(C) + P(C)\varepsilon_1 + P((E \rightarrow \neg C) \,\&\, \neg C), \text{ for some } \varepsilon_1 \approx 0 \\
&= P(C)\varepsilon_1 + P((E \rightarrow \neg C) \,\&\, \neg C)
\end{aligned}
$$

Applying (Approximate CCCP) to the left-hand side, we therefore have

$$P(\neg C | E) - \varepsilon_2 < P(C)\varepsilon_1 + P((E \rightarrow \neg C) \,\&\, \neg C), \text{ for some } \varepsilon_2 \approx 0$$

(writing down only the inequality that is most favorable for the Hypothesis—the one that exacerbates the difference between the two sides), and thus

$$\frac{P(\neg C)}{P(E)} - \varepsilon_2 \;< P(C)\varepsilon_1 + P((E \rightarrow \neg C) \,\&\, \neg C).$$

Since $P((E \rightarrow \neg C) \,\&\, \neg C) \leq P(\neg C)$, it follows from this that

$$\frac{P(\neg C)}{P(E)} - \varepsilon_2 - P(C)\varepsilon_1 < P(\neg C),$$

or equivalently

$$\frac{P(\neg C)}{P(D) + P(\neg C)} - (\varepsilon_2 + P(C)\varepsilon_1) < P(\neg C).$$

---

[55] Choosing to express the expansion this way, and later noting that $P((E \rightarrow \neg C) \,\&\, \neg C) \leq P(\neg C)$ makes the proof here go more smoothly, and Ned Hall should get the credit for this.

The fraction will approximate 1 if P(D) is sufficiently small; and $\varepsilon_2 + P(C)\varepsilon_1$ is very small. If we assume that $\langle W,F,P_C,\rightarrow\rangle$ is full, then we can make the fraction approximate 1 as closely as we like. Again, we then have $P(\neg C)$ very high, and thus $P(C)$ very low.

If we further assume that the same $\rightarrow$ is employed in $\langle W,F,P_{\neg C},\rightarrow\rangle$ as well, and that it too conforms to (Approximate CCCP), then we can run a similar argument to the one above to show that $P(\neg C)$ is very low as well—a contradiction.

At this point, the proponent of the Approximate Hypothesis might want to exploit the vagueness in its formulation: the less strict the degree of approximation that is demanded, the less these arguments count against it. Unfortunately for the proponent, however, the less strict the degree of approximation that is demanded, the less interesting the Approximate Hypothesis becomes. When the standards of strictness drop too low, the Hypothesis becomes virtually contentless, and can no longer do the work it was originally intended to do—to illuminate the semantics of the conditional, and so on. In order to hang onto a hypothesis that has a chance of being true, the proponent must sacrifice its informativeness. This recalls the dilemma of Carnap's weather forecaster, who is to predict tomorrow's temperature to within some interval: if the interval is too narrow, the prediction is likely to be mistaken; but if the interval is too wide, the prediction is uninformative.

In any event, the perturbation theorem of §8.2.1 shows us that there's a sense in which *any* degree of approximation to (CCCP) is too strict a demand. I showed that any perturbation of a CCCP-function for a certain $\rightarrow$ will yield a non-CCCP-function for that $\rightarrow$. The idea was that a perturbation changes the probability of the conditional, while keeping the conditional probability constant, or vice versa. While small perturbations of a CCCP-function do conform to (Approximate CCCP), not all perturbations are small. After all, we can *significantly* change the probability of a certain conditional, without any change in the corresponding conditional probability, or

vice versa, by a mud movement of the right sort. The more significant the perturbation, the more poorly the resulting function will approximate a CCCP-function for that →.

In fact, for any →, A, and B, (apart from trivial cases) there must be a probability function P, such that P(A→B) and P(B|A) differ as dramatically as they could—one being 0, the other 1. This is what happens if we begin with a CCCP-function P' for this →which concentrates all of its probability on (A→B) & A♥B, say. In that case, P'(A→B) = 1, and hence P'(B|A) = 1. But if P agrees with P' everywhere outside A♥B, and concentrates all of its probability on ¬(A→B) & A♥B, then it is a perturbation of P' that provides one such dramatic counterexample to (CCCP): P(A→B) = 0, but P(B|A) = 1.

This is disastrous for the Approximate Hypothesis, because there's *no* reasonable standard by which a probability equal to 0 approximates a probability equal to 1. And it shows that even *combining* the strategies of restricting the domain and retreating to (Approximate CCCP) does not ensure a smooth ride for the Hypothesis.

As I have said, this does not purport to be a complete discussion of all ways that (Approximate CCCP) might come to the rescue of the Hypothesis. For example, nothing that I have said counts against the universal tailoring version of the Approximate Hypothesis—I have recast two of the limitation results here, but not the no-go results. (The 'no finite CCCP-models' result of §7.2 can't be recast, as far as I can tell—while I showed that conditional probabilities outnumber probabilities of conditionals in finite models, and thus they can't all be identified, that tells us nothing about how much they must differ by.) Still, I hope that my discussion has made it plausible that any defender of the Approximate Hypothesis has his work cut out for him.

representor contains all probability functions that assign a sharp value in the interval [0.7, 0.9] to Collingwood winning, and that are compatible with your other judgments.

We could equally adopt a Lewisian view of what goes into the representor of your mental state: all those functions which belong to systems of belief and desire that best rationalize the behavior to which that state would tend to dispose one. Again, when your opinion is vague, there will be a multiplicity of such functions. There will be no clear front runner among such systems of belief and desire, one that fits your dispositions markedly better than any other, so your opinion is represented instead by a set of such systems, and hence a set of probability functions. Or it may be that one system rationalizes your dispositions well, but does not fit so well, while another fits well, but does not rationalize so well since it is unreasonable in some way (assigning probability zero to empirical propositions, especially those that the agent still regards as possible, is one way discussed by Lewis (1986 and 1992)). If the exchange rate between fit and rationalizing is to some extent up for grabs, these two systems—and also systems intermediate between the two—might run neck-and-neck. In that case again, the probability functions of all such systems deserve a place in the representor.

There are other stories that can be told about what it is that makes a probability function yours. Jeffrey (1965 and 1983) tells his in terms of preferences: your representor ("probasition", in his terminology) is the set of all probability functions that are consistent with *them*. Again, we face the possibility of representors with many— indeed, infinitely many—probability functions, since the propositions in your preference ranking may not be unbounded above and below in desirability, and (as Jeffrey shows), it is only then that the probability function is uniquely determined.

The versions of the Hypothesis concerning belief that I've distinguished so far are clearly inappropriate for opinion represented by a set of probability functions. Nevertheless, we can modify them appropriately:

Precisified belief function version: There is some → such that for all P that could appear in the representor of a rational agent's system of beliefs, (CCCP) holds.

Precisified belief function tailoring version: For each P that could appear in the representor of a rational agent's system of beliefs, there is some → such that (CCCP) holds.

I don't want to labor my discussion of these versions—they are not the only versions appropriate for vague opinion (as we'll see shortly), and they perhaps look too much like the versions that they are meant to replace, since they insist on function-by-function adherence to (CCCP). Nevertheless, it is worth pointing out how they inherit some of the difficulties that beset the original versions, before I proceed further.

Firstly, there are difficulties posed by the 'no finite CCCP-models' result of §7.2. The proponent of these versions of the Hypothesis will have to convince us that only probability functions with infinite range can precisify a rational agent's system of beliefs. Thus, he adds an additional constraint to van Fraassen's on what gets into the agent's representor: namely, anything that conforms to the agent's judgments, *and to (CCCP)*; and the latter constraint automatically filters out any probability function with finite range. Or he may say that any finite-ranged function, while perhaps fitting the agent's dispositions, could not *rationalize* them, for only a CCCP-function can do that—and that only rationalizing functions should find their way into a representor. Of course, the representor is just a theoretical entity, and its constitution is a matter of stipulation rather than of discovery (it is not imprinted on the agent's brain in microscopic ink), so such a reply is always open. Nevertheless, it seems that *only* a commitment to the Hypothesis could justify this additional constraint. And that game can be played both ways. Someone opposed to the Hypothesis can equally say that all probability functions with infinite range are weeded out of the representor, leaving behind only *non*-CCCP-functions—perhaps arguing along the same lines as I did in §7.3. At best, it's a stalemate.

And faced with this stalemate, it seems that at best there's simply no fact of the matter as to the truth of these precisified versions of the Hypothesis—and this is something that their advocate cannot say. Not all stalemates are like this, of course. Mathematicians may be stalemated as to the truth or falsity of some conjecture, while agreeing that there *is* a fact of the matter. Could the constitution of a rational agent's representor be like that? Is there a brute fact about which functions it contains, and which it doesn't? I find this implausible. We should not reify the representor in this way. Van Fraassen's model is just that: *a model*. This proposal confuses a representation of reality for reality itself.

Put another way: the stalemate here suggests the presence of higher-order vagueness (and there is surely nothing analogous in the case of the mathematical stalemates that I had in mind): for a given agent, there may be a *set* of representors, each of which is *not determinately not* hers: one representor contains no CCCP-functions, another contains nothing but CCCP-functions, still others contain some of both. To uphold the Hypothesis, one must insist that there is indeed a determinate fact of the matter here: the representor is the one that contains only CCCP-functions, and none other. (Here I assume that the representor is designated as the largest one with that property—presumably the proper subsets of that largest representor are not themselves regarded as candidate representors.) But isn't this like insisting that it is a determinate fact that bald men have exactly 1000 hairs or fewer? Vague language—and, I suggest, vague opinion—just doesn't work like that.

The proponent of the Hypothesis might counter that, unlike the case of the number of hairs that constitutes baldness, the Hypothesis has significant theoretical utility, as we saw in §1–3 —it is not just some inapt stipulation. However, as we saw in §4, it also has some theoretical *dis*utility, and now we see that it has still more. It is not clear to me that, apart from his own theoretical motives, the Hypothesis' friend has any reason to think that only probability functions with infinite range can precisify

someone's belief system. Presumably he befriended the Hypothesis because he thought that it provided a simple explanation for certain linguistic phenomena. But now we see that he has to complicate his theory elsewhere. I want to say that opinion can be precisified by finite probability functions, but this is not something that he can say.

Incidentally, it would appear that many a bizarre 'hypothesis' could be 'defended' by similarly rigging appropriately the standards of admission into the representor. Consider the hypothesis that each probability function that could represent a rational agent's opinion assigns only irrational numbers as values (apart from 0 and 1).[57] How could one 'show' this? Why, simply consider the functions that remain in a rational agent's representor, after all the functions that lack that property are filtered out (which, after all, the hypothesis tells us that they must be). Each one assigns only irrational numbers as values (apart from 0 and 1)! It is hard to know how to refute such a hypothesis, but that's no great credit to it (and a Popperian would call it something worse).

In any case, this has only postponed more serious trouble for the precisified belief function version of the Hypothesis, brought about by the perturbation theorem of §8.2.1.[58] We saw that if P' is a perturbation of P relative to a given $\rightarrow$, then at most one of P and P' is a CCCP-function for $\rightarrow$. It is hard to see in general why one of these functions could appear in a given agent's representor, and yet the other one could not. For example, if a representor contains a certain function, then according to this version of the Hypothesis, it cannot contain any of its very near neighbors that are tiny perturbations of it. Again, it will be hard to make a convincing case for this. Of course, conceiving of representors à la Lewis, the friend of the Hypothesis can insist again that these neighbors are not worthy candidates to appear in the representor, since they cannot rationalize the agent's state of mind. But it's hard to see how this could be

---

[57] I thank Ned Hall for the example.
[58] Here I am indebted to David Lewis.

supported independently of the Hypothesis; and it's not much help to someone who conceives of representors à la van Fraassen (in terms of determinate judgments) or à la Jeffrey (in terms of preferences).

It's worth noting that such filtering out of non-CCCP-functions from a representor has the further consequence that representors are in general not convex sets—for as I noted in §5.1, a mixture of two CCCP-functions is generally a non-CCCP-function. This consequence would be unacceptable to Levi (1980), who argues that representors should always be convex. Still, the friend of the Hypothesis might point out that conformity to various other types of judgments give rise to non-convex representors, so we are stuck with them anyway. Judgments of independence are like that, as Jeffrey (1992, p. 70) shows, and yet it may be reasonable to impose conformity to such judgments as constraints on a representor.

The trouble is that excluding non-CCCP-functions from a representor might exclude too much—what remains may not represent the agent's belief system at all. For the agent whose probability for A→B is vague might nonetheless have a sharp conditional probability P(B|A). It is easy to see how this could happen. For it may be vague what worlds constitute a given conditional—the boundary of the conditional, as it were, may not be determinate. Suppose → depends partly on the similarity relation on worlds, as the Stalnaker or Lewis conditionals do, so that A→B is true iff B is true at the closest A world (or at all the closest A worlds, if there is a tie among them). Then if the similarity relation on worlds is vague, so will be the boundaries of at least some conditionals. So the probability that the agent assigns to A→B may depend partly on just what worlds she takes A→B to consist of—and vagueness enters there. On the other hand, the conditional probability depends only on the probabilities of A&B and A, and these may well be sharp, since they do not depend on the similarity relation. (Of course, this is an argument against the Hypothesis that can stand independently of my perturbation result.)

I have argued that representors of human opinion will have, or at least might have, non-CCCP-functions for any given →: functions with finitely many values, or functions that are perturbations of CCCP-functions for that →. Up till now, I have imagined the defender of the faith denying this. But now suppose that we confront a defender of a different stripe, one who shows little interest in the function-by-function details of the representor. After all, he urges, what we care about is the overall opinion of the agent, not the precisifications themselves.[59] What matters, rather, is how the functions behave collectively. He prefers to see the Hypothesis worded in these ways:

Vague belief version: There is some → such that, for all representors of a rational human agent's system of beliefs, the intervals of vagueness (derived from these representors) of both sides of (CCCP) always agree.

Vague belief tailoring version: For each representor of a rational human agent's system of beliefs, there is some → such that the intervals of vagueness (derived from these representors) of both sides of (CCCP) always agree.

(He might weaken these still further—for example, by requiring only that one interval of vagueness is always contained within the other, or even that they merely always *overlap*, so that (CCCP) is not determinately violated.) He is prepared to admit that various precisifications in the representor of an agent's opinion might not conform to (CCCP); but he is not troubled by this, since he cares only about a *global* feature of the representor.

Furthermore, this defender of the Hypothesis, unlike those who came before him, might allow the possibility of the sort of higher-order vagueness of opinion that I endorsed: *various* representors, some containing non-CCCP-functions, may represent

---

[59] Mike Thau and Bas van Fraassen, both playing the role of devil's advocate, independently suggested this to me.

an agent's opinion equally well. He only asks that they all produce intervals of vagueness that conform to the equation.

None of the triviality results rules this out; however, I think that certain philosophical arguments render it an unsatisfactory defense. Firstly, precisifications of an agent's opinion are like limiting cases, ways that a certain ideally rational agent would be, who has a sharp belief system, but who otherwise is like the agent being represented. If the Hypothesis is a rational constraint, then surely all these limiting cases should also turn out to be rational. Compare: despite the fact that the concepts 'red' and 'orange' are vague, I want to say that red and orange are different colors; but it would be hard to say that convincingly, if it did not come out true on all ways of precisifying the concepts.

More importantly, there is a general problem for *any* recourse to vagueness, one reminiscent of the problem I posed for the objector to finite models in my discussion of §7.3. If the Hypothesis is construed as a hypothesis about a conditional used by some ideally rational agent (or even a human agent) then any such defense regards that agent as being saddled with an odd constraint. It requires that her opinion not be completely sharp. Bad news for a god, or some other ideally rational being, whose opinion is like that! Surely, if she has such a sharp belief function, more power to her. Moreover, even if an agent's current opinion happens to be vague, *learning* becomes an unacceptably precarious business for her. She must make sure that her opinion never becomes completely sharp. Again, certain propositions will inexplicably become Moore propositions—namely, any proposition which, if learned by the agent, would result in the prohibited sharpening of her opinion. Bad news for the god's younger sister, who is on the verge of learning just such a proposition!

Now perhaps it is nomically impossible for human beings to have perfectly precise opinions, so there is never any danger of a human's opinion sharpening completely; or perhaps it is irrational for a human being to have perfectly sharp opinion, because no

evidence that she could acquire would justify that (because of her finite powers of discrimination, and so on). But the triviality results acquire some bite even before the states of mind that are being modeled become completely sharp.[60]   Recall my discussion of the Approximate Hypothesis: it is difficult to sustain even the approximate equality of conditional probability with the probability of a conditional. Take some conditional A→B for which an agent's probability is vague, but only over a small interval, say [x, x+ε]. The vague belief version has it that the corresponding conditional probability is also vague over [x, x+ε]. So every precisification of the agent's opinion assigns A→B a sharp probability in this interval; and also a sharp corresponding conditional probability in this interval. Thus, the two values can differ by at most ε. According to our second reading of 'approximate equality' in the previous sub-section, this is just to say that for each precisification P, $P(A \rightarrow B) \approx P(B|A)$. Now suppose that the agent's probability for all conditionals is like this: vague, but only over small intervals. Then the vague belief version entails that each precisification conforms to (Approximate CCCP). But we've seen what trouble that means—at least if the agent is ever to change her mind.

So even a state of mind which is approximately but not completely sharp would appear not to conform to the Hypothesis—or at least not both before and after various learning experiences. So it looks like the Hypothesis saddles the rational agent with an odder constraint still: her opinion must not become even approximately sharp. Surely not even *human* opinion should be so constrained. To drive the point home: understood as a rational constraint, (CCCP) becomes harder and harder to sustain as the agent's opinion becomes progressively sharper. But it is ridiculous to say that the vaguer that agent's opinion is, the more rational she is—that each step towards sharper opinion is a step in the wrong direction. Total vagueness is not the state of mind of the ideally

---

[60] I thank David Lewis for this observation, and Jennifer Saul for helping me drive the point home in the next paragraph.

rational agent, or even the somewhat rational agent. If anything, it's the state of mind of a moron.

In short, the recourse to vagueness is not enough. What is needed is *substantial* vagueness—vagueness over at least fairly large intervals, and the larger the better. But what sort of a constraint on rationality is *that?*

9.3 *Introduce non-Kolmogorovian 'assertabilities'*

The triviality results assume the usual Kolmogorov axioms of probability (or if not the full strength of countable additivity, then at least finite additivity). They can be blocked by casting the Hypothesis in terms of 'probability functions' that do not conform to these axioms. We have already seen one way that this can be done—in fact, we saw it right at the outset. Adams, I said, proposes a variant of the Hypothesis (to repeat: the so-called 'Adams' Thesis')—but only a *variant*, since his 'probabilities' of conditionals do not conform to the usual probability calculus. To be sure, Adams *calls* them 'probabilities'—but he admits that they do not attach to Boolean combinations of propositions in the usual ways. Indeed, he says that "we should regard the inapplicability of probability to compounds of conditionals as a fundamental limitation of probability, on a par with the inapplicability of truth to simple conditionals" (1975, p. 35).

In that case, why speak in terms of 'probability' at all, as far as conditionals are concerned? Here, I'm in total agreement with Lewis (1976, p. 136) when he writes: "if it be granted that the 'probabilities' of conditionals do not obey the standard laws, I do not see what is to be gained by insisting on calling them 'probabilities'." So I endorse Lewis' suggestion that they be called 'assertibilities' instead, as several authors have, (adopting here, and from now on, the spelling favored by Jackson and Collins, whom I'll be quoting at some length). I will understand Adams' Thesis, then, to be a thesis about the assertibility of the indicative conditional, which I will write in the following

form (to distinguish it from its near relative, the belief function version of the Hypothesis):

For all P that could represent a rational agent's system of beliefs,

(AT)　　$As_P(A{\rightarrow}B) = P(B|A)$, for all A, B in the domain of P, with $P(A) > 0$,[61]

where $As_P(X)$ is the degree of assertibility of X according to the belief state given by the probability function P, and $\rightarrow$ is (for present purposes) the indicative conditional.

Jackson (1987), who probably has said more in defense of (AT) than anyone else, explains the notion of assertibility thus:

> The aspect of a sentence's usage which tells us something about its meaning are the conditions governing when it is justified or warranted—in the epistemological sense, not in a purely pragmatic one—to assert it, or, as this comes in degrees, to what extent it is justified to assert it in various circumstances. (p. 8)

More succinctly, assertibility is "the justifiability of *what* is said" (p. 11). Note that an account of assertibility properly belongs to a theory of semantics, rather than pragmatics—it concerns what is said, rather than the saying of it, so considerations of etiquette and politeness, for example, are irrelevant.[62]

(AT) evades the triviality results, since assertibility does not obey the laws of probability. Indeed, it is hard to imagine how a triviality result could ever be brought to bear against (AT)—not so much because it is obviously correct, but rather because assertibility is such a nebulous thing. Much as this renders (AT) impervious to formal refutation, it may also be a source of suspicion. For (AT) equates something nebulous with something precise.[63] To be sure, explications often—and perhaps typically—do

---

[61]　Strictly speaking, the arguments of As_p are sentences, rather than propositions— presumably, assertibility attaches to items of language.

[62]　Jackson separates the semantic and the pragmatic notions, calling the former 'assert**i**bility', and the latter 'assert**a**bility'. Collins conforms to this spelling.

[63]　Or at least, something that is always nebulous, with something that is occasionally precise. The reason for the qualification is that the conditional probabilities of an agent with vague opinion may be themselves vague.

just that; however, I will argue that there is something peculiar about this particular equation.

What (AT) cannot evade is the *cardinality argument* in the proof of the 'no finite CCCP-models' theorem of §7.2. Since it was shown there that, for finite models, the distinct values of the conditional probability function outnumber the distinct values of the unconditional probability function, (AT) implies, *inter alia*, that the same is true of the assertibility function. I have given my defense of finite models for human opinion in §7.3; presupposing that defense, I want to now consider the plausibility—or lack thereof—of the claim that assertibility values outnumber probability values (for a given agent).

At first sight it may seem reasonable enough. While assertibilities are said to typically coincide with probabilities, they sometimes do not, and this is not due merely to the presence of conditionals (which we already knew to be temperamental critters) in the language. As I have noted before, the assertibilities of 'A and B' and 'A, but B' typically disagree, even though their probabilities agree (since both 'A and B' and 'A, but B' have the same truth conditions, and probability is probability of truth). This may suggest that assertibilities do outnumber probabilities—here, for example, we have two assertibility values corresponding to the one probability value. But it *only suggests* this—after all, for all we know so far, an assertibility value that differs from the corresponding probability value may agree with some other probability value. (While $As_P(A \text{ but } B)$ may differ from $P(A \text{ but } B)$, it may agree with $P(C)$ for some C.)

Now it is time to go more on the offensive. I will draw on an argument by John Collins (1991). He writes:

> I think I understand what it means to say of a proposition that it is or is not assertible relative to a particular belief state. And I think I can make sense too of a comparative notion of assertibility: A is more assertible than B relative to the belief state K. But it is far from clear that there is any way of assigning a precise numerical value between zero and one to a proposition as a measure of its "degree of assertibility". (pp. 92-3)

Given how assertibility is defined, Collins' point seems well taken. Does it really make sense to award precise 'justifiability' numbers to sentences? As Collins goes on to remark, we get a grip on the notion of 'degrees of belief' via Ramsey's analysis in terms of dispositions to bet (and there, the precise explication of a somewhat nebulous concept seems well founded); but we have no analysis that would give us a similar understanding of the notion of 'degrees of assertibility'.

Still less does it make sense to award to sentences 'justifiability' numbers that are even more finely grained than probability values are. And yet this is what (AT) entails.[64] Of course, if 'assertibility' is to be treated merely as a technical term whose meaning is given functionally by (AT), then it is legitimate to simply stipulate that it is exactly as finely grained as one likes. But then Adams' Thesis loses some of its force as a thesis *about natural language*. On the other hand, if 'assertibility' is supposed to be an explication of the intuitive notion that bears the same name—and this is surely the intention of proponents of the Thesis—then it deserves closer scrutiny.

Indeed, when one examines the arguments given for Adams' Thesis, it appears that they support equally either the Hypothesis (which is to say that they show too much if they show anything[65]); or else a much weaker thesis (which is to say that they show too little), say:

(Qualitative AT) $As_P(A \rightarrow B)$ is high iff $P(B|A)$ is high; $As_P(A \rightarrow B)$ is low iff

   $P(B|A)$ is low;

or at most a comparative version of the Thesis:

(Comparative AT) $As_P(A \rightarrow B) > As_P(C \rightarrow D)$ iff $P(B|A) > P(D|C)$.

---

[64] Note that the problem only becomes worse if we allow, or insist on, infinite models—for them, (AT) entails that assertibility assumes infinitely many distinct numerical values, something that I find even harder to believe.

[65] As David Lewis has pointed out to me, the proponent of such an argument may not think that it *shows* the truth of the Thesis, but only that it *supports* it; and that it likewise only supports the Hypothesis, but in that case it is overridden by other considerations.

An argument of Jackson's that supports the Hypothesis at least as much as it does Adams' Thesis is from Ramsey's test: it is the very argument that I gave in §3 in support of the Hypothesis, just replacing 'probability' of A→B by 'assertibility' of A→B. But what does Ramsey's test speak to: probability or assertibility of the conditional? Looking at Ramsey's original text (reprinted in Ramsey (1965)), it seems he had probability in mind, if anything: "If two people are arguing 'If p will q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q; so that in a sense 'If p, q' and 'If p, $\bar{q}$ ' are contradictories. We can say they are fixing their degrees of belief in q given p" (p. 247 fn.). Ramsey's own intentions aside, Jackson has not told us why Ramsey's test is appropriate for determining the assertibility, but not the probability, of the conditional.

An argument of Jackson's that shows too little is from "case-by-case evidence":

> Take a conditional which is highly assertible, say, 'If unemployment drops sharply, the unions will be pleased'; it will invariably be one whose consequent is highly probable given the antecedent. And, indeed, the probability that the unions will be pleased given unemployment drops sharply is very high. Or take a conditional with 0.5 assertibility, say, 'If I toss this fair coin, it will land heads'; the probability of the coin landing heads given it is tossed is 0.5 also. Or take a conditional with very low assertibility, say, 'If I spend this afternoon trying to solve Fermat's last theorem, I will succeed'; the probability of my solving it given I spend this afternoon on it is correspondingly very low. (p. 12)

We may agree that the cases of high assertibility and of low assertibility that he discusses conform to (AT), but no more so than they conform to (Qualitative AT). And as for the case of the coin landing heads, it is questionable whether we have any grounds at all for thinking that the conditional has 0.5 assertibility *apart from* a prior appeal to (AT)—so it is questionable whether that judgment counts as evidence for (AT) at all (as opposed to (AT) being evidence for the judgment). Indeed, Collins argues on independent grounds that the assertibility of the conditional is 0, if anything. (See also footnote 8, p. 13 above: how can a conditional that seems to be certainly false have 0.5 assertability?)

Another argument of Jackson's that shows too little is this: "There is evidence for [(AT)] from our attitude to pairs of 'divergent' conditionals: (A→B), (A→not-B). When A is consistent, there is something quite generally wrong with asserting both (A→B) and (A→not-B)" (p. 12). But this is equally evidence for (Qualitative AT), since it can account for this fact equally well. Jackson continues: "Indeed, ([AT]) explains the further fact that (A→B) and (A→not-B) have a kind of 'see-saw' relationship. As the assertibility of one goes up, the assertibility of the other goes down." True; but (Comparative AT) will suffice to explain the phenomenon.

Finally, we have this argument of Jackson's:

> There is also evidence for [(AT)] from the fact that, by and large, an assertion of a conditional is a conditional assertion in the following sense: to assert 'If A, then B' is to commit oneself *ceteris paribus* to asserting B should one learn A... [(AT)] explains this connection between asserting conditionals and conditional assertions because, by and large, the probability of B given A is high just when learning A makes the probability of B high. (p. 13)

But one needs only (Qualitative AT) to explain this connection between asserting conditionals and conditional assertions.

The weaker versions of the Thesis, then, seem to be just as well supported as (AT) is. Moreover, they do not commit us to assertibilities that are more finely grained than probabilities, and that is to their credit. Finally, they may well be able to do all the semantic work that Adams originally wanted of (AT)—for example, helping to provide a 'probabilistic soundness' criterion for arguments that involve conditionals. It is really this last point that should decide the fate of Adams' Thesis (unless it can be shown that a sufficient purchase on the notion of assertibility can be gained by independent means, and that this vindicates the Thesis). If nothing less than the full-strength, fine-grained account of assertibility that it gives will offer the same explanatory power, then that is a reason to hold onto it. But it's not clear to me that nothing less will do as well.

In any case, it's surprising that such sensitive assertibilities appear not to be detectable in linguistic usage—they are theoretical entities, hidden variables if you will,

postulated to simplify theory. This is a slightly odd view, especially coming from Jackson. After all, usage figures prominently in his account of assertibility. And as he says, "A theory of indicative conditionals is a theory about a fragment of ordinary language. Accordingly, it is—unlike a theory of electrons or of the mind—*peculiarly* responsive to the linguistic intuitions and practices of ordinary speakers" (p. 8). It's ironic, then, that Jackson commits himself to a notion of assertibility that apparently does outrun our intuitions and practices.

10. Where does the demise of the Hypothesis leave us?

10.1 Preliminaries

I have nothing further to add to my critique of the Hypothesis. I have not considered every combination of quantifiers, nor every restriction of the domains of quantification that we could attach to its unadorned formulation (0) on p. 4—not just to spare myself and you the tedium, but also because I think that the versions that I *have* considered more or less exhaust those of philosophical interest. Furthermore, I have not discussed all of Ned Hall's devastating triviality results—partly because I think I have said enough already about the demise of the Hypothesis, partly because some of his work is still in progress as I write (October 1992), and partly because he will be better able than I to give them the exposition that they deserve.[66] I think that this is a good place to stop. I believe that the Hypothesis should be laid to rest. In this section, I want to suggest some consequences that this might have.

To identify the probability of a conditional with conditional probability, understood as a ratio, is to conflate two distinct notions, natural though the conflation is. Furthermore, I think that both notions have important roles to play. Once we have distinguished them, we will no longer make the mistake of asking one concept—as it might be, the ratio—to do double work.

---

[66] They are to appear in Hall (1993).

10.2 A dogma of conditional probability: a manifesto[67]

Who can tell me the formula for conditional probability?

- Television commercial for Twix bars

It is clear what we have asked the ratio to do for us. Philosophers have ritually taken the ratio to just *be* conditional probability, nothing more, nothing less: the conditional probability of A, given B, just *is*

$$P(A|B) = \frac{P(A \& B)}{P(B)} \quad (P(B) \neq 0) ,$$

(where I take P(A|B) to be short-hand for the ratio), or so we've been told. I will call this *the ratio analysis* of conditional probability. Such is the grip that this analysis has had on our thinking about conditional probability that I think it is fair to call it a dogma, at least among philosophers. Throughout this work, I have been speaking in ways that may have suggested that I endorse the dogma, so as not to confuse matters when my concerns lay elsewhere. But now I should briefly state my manifesto (although I must postpone a fuller statement of my views for another occasion).

I question the ratio analysis. This may strike you as absurd—you may think that conditional probability is just a technical term that is *defined* by this formula. If so, you are in good company. For example, Jackson (1987, p. 12) writes: "it is important to take the conditional probability of consequent given antecedent to be *defined* as the probability of the conjunction of the antecedent and the consequent divided by the probability of the antecedent." (The emphasis is his.) This only underscores, to my mind, the extent to which the ratio analysis has become dogma.

To be sure, *if* 'the conditional probability of A, given B' were a purely technical notion, then Jackson could well be right. There is nothing problematic about

---

[67] This is a precis of a dissertation that might have been, and in fact, nearly was.

introducing the symbol '|' into the language of the probability calculus, and baptizing it via the ratio formula. You can even intone the words 'the conditional probability of A, given B' when reading 'P(A|B)' out loud, if you want. If these words are not philosophically loaded, then this is no more contentious than saying 'the probability of A slash B', and taking that as convenient shorthand for the ratio.

But the words *are* philosophically loaded. They are tinged with associations with models of belief revision and decision-making, with analyses of confirmation, and causation; they have even found their way into a lot of the recent literature in ethics, and philosophy of law; and so on. Furthermore, they attempt to capture a notion about which we had intuitions before we donned the philosopher's cap. Conditional probability is not a technical notion that one simply defines, but a familiar concept that requires analysis; our choice of words in reading P(A|B) as 'the conditional probability of A, given B' is not so innocent after all.

Perhaps analogies will help. Choosing a book somewhat at random from my mathematics shelf, and opening the book somewhat at random, I find: "A Banach space is a complete normed linear space." Now *that's* a definition. We have no prior concept of 'Banach space', intuitions about which we could use to tease out a statement like this. Contrast this with the case of the conditional 'if A, then B' in English. It would be specious to claim that it is *defined* via the truth table for $(\neg A \vee B)$; to be sure, the material conditional $A \supset B$ is so defined, but whether or not '$\supset$' adequately analyses the English 'if ... then' is a substantive question. Or simply consider *un*conditional probability. Kolmogorov gave us a certain axiomatization of the concept, but it would be specious to claim that his axioms are *definitive* of the concept (if they were, someone like de Finetti who disputes them would simply appear foolish or ignorant, akin to someone disputing the definition above of a Banach space). So I think that 'conditional probability', like 'if ... then', like 'unconditional probability', and unlike 'Banach space' is a pre-theoretical concept of ours. It is there to be analyzed, not defined.

While the ratio is certainly not to be regarded as a definition of our concept of conditional probability, mightn't it be reasonable to think of it as a good analysis of that concept? I have two main arguments for my thesis that it is inadequate on its own. Here I will merely summarize them, without fanfare, elaboration, or defense against objections. (They get all three in my (1992).)

The problem of the zero denominator

The first argument, which is far from original, is that the analysis is mute whenever the condition has probability zero, and yet conditional probabilities may nevertheless be well defined in such cases. Examples abound, but one will suffice; it is a variant of a paradox due to Borel. Assuming that you assign a uniform probability measure over the Earth's surface (imagine it to be a perfect sphere), what is the probability that a randomly chosen point lies in the western hemisphere, given that it lies on the equator? 1/2, I hope you agree. But the probability that the point lies on the equator is 0, since the equator has no area. The ratio analysis fails to yield an answer for the conditional probability, yet intuitively it is well-defined.[68]

The problem of undefined terms

The second argument concerns cases in which conditional probabilities are defined, but the corresponding unconditional probabilities in the ratio analysis are undefined. One example should suffice to illustrate the point. Suppose that I have two urns, the first containing 50% white balls and 50% black, the second containing only white balls. I will pick an urn, and pick a ball from it at random. What is the probability that I pick

---

[68] In my (1992), I argue against three ways of contending that the ratio analysis does indeed deliver the goods, provided it is used 'correctly'. The first embraces finitism, and insists that the requisite unconditional probabilities are real-valued, and positive; the second assigns them positive infinitesimal values; the third has us take limits of appropriate sequences of conditional probabilities, each of which conforms to the ratio analysis.

a white ball (W), given that I pick from the first urn (F)? 50%, surely. According to the ratio analysis, it is P(W|F), that is:

$$\frac{\text{P(I pick a white ball \& I pick from the first urn)}}{\text{P(I pick from the first urn)}}.$$

However, I have been careful not to give you any information about these unconditional probabilities. They are no part of the specification of the problem. Consider the denominator: what is the probability that I pick from the first urn? Perhaps I have made up my mind to pick from the first urn no matter what. Perhaps I will impulsively lunge for one of the urns for no apparent reason. Perhaps I will use some further chance mechanism, like a coin toss, to guide my choice of urn. But I tell you nothing of that. I give you no information on which you could base your judgment of the probability that I pick from the first urn. Nevertheless, it seems you should have a sharp judgment of the conditional probability, namely 1/2.

I pose this as a challenge to the dogmatic.[69]

10.3 The ambiguity thesis

We come to what is probably the most speculative part of this work. I offer it only tentatively.   However, I like to think that it has some *prima facie* plausibility, and that it ties together various strands in the literature.

Given that conditional probability is not the sort of thing that is merely defined to be such-and-such, we may wonder whether there is a univocal concept there to be analyzed. It shouldn't come as too much of a surprise if there isn't. For instance, the English conditional is not univocal, according to many people who have written about the subject—the indicative conditional and the counterfactual are often given different

---

[69] In my (1992), I imagine three main sorts of objections. The first argues that the requisite unconditional probabilities are defined, by symmetry; the second argues that they are defined implicitly; the third argues that while they are not determinate, their ratio is, using supervaluations as a way of making sense of this. Only the third of these do I really take seriously, but I think that it also falters in the end.

treatments by the same author.[70] Likewise, the word 'probability' is taken by many to be ambiguous between *chance* on the one hand, and *degree of belief* on the other.[71] Now, I'm not shaping up to say here that 'conditional probability' inherits the ambiguities of 'conditional' and 'probability' (although I think that's true too); rather, I merely want us to be psychologically prepared with the thought that familiar concepts need not have univocal analyses. If that's a platitude, all the better.

Let's abbreviate 'the conditional probability of B, given A' as 'P(B, given A)', where this is meant to be neutral as to analysis. My suggestion is that P(B, given A) is ambiguous. What then, might be its disambiguations? I suggest: the two sides of (CCCP), which equal each other according to the Hypothesis, but which don't (in general) according to me. On one disambiguation, it goes by the usual ratio formula, $\frac{P(A \& B)}{P(A)}$ —I have called this 'the ratio analysis'. On another disambiguation, it goes by the probability of a conditional, P(A→B) —let me call this the 'probability–of–a– conditional' analysis, or more succinctly (and I hope appropriately) *the PC analysis*. And the thesis that conditional probability is ambiguous between the ratio analysis and the PC analysis I will call the 'the ambiguity thesis'. I hope that you, like me, have become convinced that the two disambiguations really are *two*, and not simply the same thing said two different ways.

There is a sense in which the Hypothesis goes hand in hand with the dogma of conditional probability—and considering how natural the Hypothesis is, so too the dogma's grip is only natural. If the universal version of the Hypothesis were right, we perhaps could have *afforded* to be dogmatic about the analysis of conditional probability; and if the belief function version were right, we could perhaps have been dogmatic at least about the analysis of subjective conditional probability. Analyzing conditional probability, or subjective conditional probability according to the left or

---

[70] I am thinking of authors like Lewis, Jackson, and Gibbard.
[71] I am thinking, for example, of Lewis (1981).

right hand side of (CCCP) would have come to the same thing. Now we know that it matters which side we choose. The ambiguity is seen to be genuine once we recognize the failure of the Hypothesis. Thus, the failure of the Hypothesis goes hand in hand with the ambiguity thesis, which I offer in place of the dogma.

But I should say more about what I see as motivating the ambiguity thesis.

10.3.1 The 'sounds right' argument, revisited

I said early on in defense of (CCCP) that it "sounds right", and quoted van Fraassen saying something similar. (Of course, at that stage of proceedings I was just playing devil's advocate—I hope I have clearly shown myself to be no friend of the Hypothesis, much as I was imagining myself to be one then.) And while sounds can be deceptive, and we shouldn't give them overriding weight, we shouldn't ignore them either. After all, I think that it is *surprising* that the Hypothesis is false, and I think that most of the surprise stems from the natural ring that it has to the ear when we read it out loud in the way that we've been taught to. I actually suspect that this was what led Stalnaker to formulate the Hypothesis in the first place—and Adams to formulate his near variant of it.

But (CCCP) only sounds right if we take the ratio analysis of conditional probability for granted (as we've been taught to). Indeed, if I read out loud what the equation in (CCCP) literally says, and don't tacitly buy into the dogma, I don't know how it sounds:

"the probability of 'if A, then B', equals the ratio of the probability of A and B to the probability of A".

If anything, it sounds rather strange!

What "sounds right" is that the probability of B, given A, should be analyzed as the probability of: B, if A. As I said at that early stage, many assertions of conditional probability sound like assertions *of probability*, within the scope of which is something

that seems best analyzed as a conditional. But *that* only translates into (CCCP) when you are convinced of the unequivocal truth of the ratio analysis of conditional probability, as I am not. If we want to follow the sounds in the face of the demise of (CCCP), we should make a place for the PC analysis, and this we can do once we are freed of the dogma. It is to PC's credit that it does justice to the surface grammar of conditional probability statements—certainly more so than the ratio analysis does.[72]

To summarize, the friend of the Hypothesis that I have imagined advocates the following argument:

---

[72] David Lewis has suggested to me that 'if' should not always be analyzed as a sentential connective, but at least occasionally as a restriction modifier instead (see, for example, his (1975)). "Often, if it's raining, I catch a cold" is not to be analyzed as 'Often(it's raining→I catch a cold)' —for *that* has the form 'Often, the following conditional is true...', which seems to mean something else (it suggests that there are times when the conditional is true, though other times when the conditional is false). Nor does it have the form 'It's raining→Often I catch a cold'. For suppose that it's raining; by modus ponens we could then derive 'Often I catch a cold', which may not be true (it may rarely rain, and I may only catch colds when it rains). Rather, the right idea seems to be more along these lines: *restricting our attention to times when it's raining* (and ignoring times when it isn't), I often catch a cold.

Maybe 'the conditional probability of B, given A is x' should be analyzed along similar lines: *restricting our attention to the A cases*, B has probability x'. But isn't that just the ratio analysis, redescribed? No. For this 'restriction' analysis of conditional probability, unlike the ratio analysis, makes sense even if A has probability 0, or if its probability is undefined. Restricting our attention to cases in which the point lies on the equator, the probability is 1/2 that it lies in the western hemisphere; restricting our attention to cases in which I pick from the first urn, the probability is 1/2 that I pick a white ball. Nevertheless, this 'restriction' analysis may be a *generalization* of the ratio analysis, one that reduces to the ratio analysis in cases where the terms in the ratio are both defined, and the denominator is non-zero.

I think that there still is a place for the PC analysis. There still is something to the argument that it 'sounds right', as I have said, enough at least to give the Hypothesis its initial plausibility. Moreover, I believe the arguments that I offer in the following sub-sections point to roles that the PC analysis can play, but that even a generalization of the ratio analysis can't—for example, in analyzing causal independence. Perhaps, then, our concept of conditional probability is ambiguous between the PC analysis on the one hand, and the restriction analysis (thought of as a generalization of the ratio analysis) on the other. This is too big a topic to take up here, although I hope to do so in the future.

1. (The dogma): Conditional probability is to be unequivocally analyzed according to the ratio formula.

2. Conditional probability sounds like the probability of a conditional.

Therefore

3. The ratio is to be identified with the probability of a conditional.

I see the bulk of this dissertation as a very long argument against 3, and since I too find 2 plausible, and have independent arguments against 1 (which I could only sketch here), I prefer running this argument in reverse, concluding with the rejection of the dogma.

The ambiguity in conditional probability has also been exploited in the literature to some extent, as I have already begun to indicate.

10.3.2 Independence, revisited

In §3 and §4, I considered various intuitive arguments for and against the Hypothesis. One argument for it was this: given that the $\rightarrow$ obeys a certain plausible logical principle, (CCCP) is true iff the conditional is independent of its antecedent; and there seemed to be intuitions that favored the independence of the conditional from its antecedent. But a few pages later, this turned into an argument *against* the Hypothesis: in Newcombesque examples, intuitions are supposed to favor the conditional's *not* being independent of its antecedent. How can our intuitions go both ways like this? It would be nice to be able to explain this. I suggest (tentatively, again), that perhaps it is because the notion of 'independence' goes both ways—the conditional is independent of its antecedent in one sense, but not another.

I think a good analysis of 'B is independent of A' is: P(B, given A) = P(B). The ambiguity thesis predicts that there should be two senses of independence, corresponding to the two disambiguations of conditional probability:

(Ratio-independence): $\dfrac{P(A \ \& \ B)}{P(A)} = P(B)$

(PC-independence): $P(A{\rightarrow}B) = P(B)$

Could this account for our seemingly conflicting intuitions regarding the independence of conditional and antecedent?

Consider ratio-independence first. As I showed, ratio-independence of conditional and antecedent is equivalent to (CCCP) holding for that conditional (assuming the plausible logical principle). I hope I have convinced you since then of the untenability of (CCCP), in general. Thus, I hope I have convinced you that conditionals cannot be ratio-independent of their antecedents, in general. Count that as one intuition respected.

Now consider PC-independence. $A{\rightarrow}B$ is PC-independent of A just in case $P(A{\rightarrow}(A{\rightarrow}B)) = P(A{\rightarrow}B)$, so this is what I need to argue for. But this follows immediately from the principle:

$A{\rightarrow}(A{\rightarrow}B) = A{\rightarrow}B$,

which is validated by both the Stalnaker and Lewis logics, and which strikes me as very plausible.[73] Count that as the other intuition respected.

Moreover, recalling our discussion of §4, the two notions of independence predicted by the ambiguity thesis are exactly the two notions proposed by Stalnaker—ratio-independence is his stochastic independence, while PC-independence is his causal independence. And I take Stalnaker's distinction, his 'independence ambiguity thesis' as it were, to be well motivated.

Actually, I think that ratio-independence cannot exhaust our concept of independence, for other reasons. I think that it leads to some unintuitive judgments of independence, for sometimes it just happens to be the case that $P(B|A) = P(B)$ —or equivalently, $P(AB) = P(A)P(B)$—even though A and B seem not to be independent on physical, or other, grounds.

---

[73] It does not strike everyone that way, however, as David Lewis and Bas van Fraassen have pointed out to me. It is denied, for example, by most relevant logicians—see Anderson and Belnap (eds.) (1975) section 5.2.1.

<u>Example:</u> Consider a family with three children of different ages, and assume that a given child is equally likely to be a girl (g) or a boy (b), and the sex of a given child is independent of those of the other children in the family. Thus, we have a sample space with eight equiprobable points: ggg, ggb, ..., bbb.

Let A be 'there are at least two girls' and B be 'the eldest two children have the same sex'. My intuitions tell me that A and B should not come out independent, but I will not rest my case on them. Firstly, note that A and B *are* ratio-independent:

$P(AB) = 1/4$; $P(A) = 1/2$; $P(B) = 1/2$, so

$P(AB) = P(A)P(B)$.

You may not share my intuitions; or you may share them but allow them to be overridden, conceding that the two events are independent after all, since the usual formula for independence says that they are. However, I find it strange that my intuition would have been vindicated for *every other family size*. For every family size except 3, A and B are *not* ratio-independent.[74]

It seems to me just a mathematical accident that 3 is the odd one out. I say 'mathematical accident', because I can think of nothing relevant that distinguishes this case from all the rest, and yet the mathematics does so. Exclusive adherence to the ratio analysis makes us read too much into this accident, with its declaration of

---

[74] In case a proof is desired:

For a family of size n,

$P(A) = 1 – P(0 \text{ girls}) – P(1 \text{ girl})$

$\quad = 1 – (1/2)^n – n.(1/2)^n$

$P(B) = 1/2$

$P(AB) = 1/4 + 1/4(1 – (1/2)^{n–2} – (n–2)(1/2)^{n–2})$ (the case where the eldest two are girls + the case where the eldest two are boys).

We want to solve the equation $P(AB) = P(A)P(B)$ for n:

$1/4 + 1/4(1 – (1/2)^{n-2} – (n–2)(1/2)^{n–2}) = 1/2 – (1/2)^{n+1} – n.(1/2)^{n+1}$

or, after rearranging,

$(1/2)^n + (n–2).(1/2)^n – (n+1).(1/2)^{n+1} = 0$

and multiplying both sides by 2^{n+1},

$2 + (n–2).2 – (n+1) = 0$,

from which we derive that $n = 3$.

independence. I think we need a notion of independence that does not distinguish among family sizes in this surprising way.

10.3.3 Joyce's Foundations for Causal Decision Theory

Of course, the two notions of independence go hand in hand with the two decision theories: ratio-independence with evidential decision theory, PC-independence with causal decision theory. Moreover, Joyce (1991) provides a representation theorem for causal decision theory by giving a representation theorem for 'conditional' decision theory, wherein what he calls the 'conditional probability' of outcome, given act, is ambiguous between the ratio P(outcome & act)/P(act), and the PC P(act→outcome). The first can be thought to concern the impact of evidence about the performance of an agent's act on the belief about the outcome; the second can be thought to concern the agent's belief about the causal link between the act and the outcome.

10.3.4. Updating and Supposing

Collins (1991) contends that there are two distinct methods of belief revision: *updating*, which goes by ratio-conditioning, and *supposing*, which goes by imaging.[75] That is, upon updating upon some proposition E, an agent's probability function P(–) changes to the function P(–|E), where this, as usual, is shorthand for the appropriate ratio; however, upon supposing that E, P(–) changes to $P_E(-)$, the function one gets from P by imaging on E. So far this does not look much like the bifurcation that I want. But the crucial link is supplied by a result due to Lewis (1976):

---

[75] Collins presupposes a non-quantitative framework for belief, and belief revision, but he is not unsympathetic to the probabilistic framework, and has indicated to me that he endorses this remark.

The probability of a Stalnaker conditional with a possible antecedent is the probability of the consequent after imaging[76] *on the antecedent.*

Combining Collins' contention with Lewis' result, we have:

When one supposes that E, one revises one's probability assignment according to the schema

$$P_E(-) = P(E \rightarrow -)$$

where $\rightarrow$ *is the Stalnaker conditional.*

(I assume, as Collins must, that one can genuinely suppose only something that is possible.) To help myself to this in support of the ambiguity thesis, I need to commit myself further on the nature of the conditional that appears in the PC analysis: it is at least sometimes the Stalnaker conditional.[77] Now I can state Collins' position on belief revision, combined with Lewis' result, in a way congenial to my own purpose:

Belief revision to accommodate E takes place in accordance with the schema:

$$P_E(-) = P(-, \text{ given } E);$$

if the revision is one of updating on E, then P(–, given E) = P(–|E);

if the revision is one of supposing that E, then P(–, given E) = P(E $\rightarrow$ *–).*

We use conditional probability statements to express revisions of our beliefs; they are ambiguous because belief revision is ambiguous.

---

[76] Since Lewis' paper, various sorts of blurred imaging have been suggested; 'imaging' here and elsewhere should be taken to refer to Lewis' original sense of it: the probability from each ¬E world is shifted to its nearest E world.

[77] Thanks to my ambiguity thesis, am I therefore saddled with conditional excluded middle? No more so than Collins is saddled with it (in fact less so, since I have several independent arguments for the ambiguity thesis, and the appeal to Collins is one that I could drop if it proved too costly). But if belief revision ever goes by imaging, then by Lewis' theorem, we're *all* saddled with conditional excluded middle (to the extent that I am)—for whether we like it or not, imaging on E is just another way of describing a revision that turns probabilities of the form P(X) into probabilities of Stalnaker conditionals of the form P(E→X). And I am certainly not saddled with conditional excluded middle in the way that Stalnaker is, for I do not say that there is only one conditional, as he does (see footnote 10 for clarification). In any case, as I said in §4, it is still controversial whether conditional excluded middle is a virtue or a vice.

10.3.5 Simpson's Paradox

You have two doctors, whom you take to be expert on matters medical. Their probability functions are q and r respectively, and you derive your own probability function P by mixing theirs (so P = wq + (1–w)r, for some 0 < w < 1). When they both tell you that smoking is correlated with lung cancer, surely you should believe that yourself. Yet as van Fraassen (1989) shows, it is possible to have probability functions q, r, and P such that

q(cancer|smoking) > q(cancer)

r(cancer|smoking) > r(cancer)

and yet

P(cancer|smoking) = P(cancer)

where P is a mixture of q and r . That is, while your experts see the correlation, you might not. Indeed, mixing can even reverse a correlation: while both your experts see a positive correlation, you see a negative correlation. You cannot take their common advice, and trust your own probability function, which you derive from theirs by mixing, at the same time. The point generalizes beyond situations of combining expert opinions. Simpson's Paradox, in slogan form, is this: mixing can wash out, and even reverse, correlations.

This is a dilemma for the dogmatic, who insist on analyzing the conditional probabilities that figure in the statements of correlation as ratios. If, however, we use the PC analysis instead, then Simpson's paradox admits of a simple solution. (Not the only solution,[78] I hasten to add, though the ones I know of aren't as simple.) The intuition that we want to respect is this:

if q(X, given Y) > q(X),

and r(X, given Y) > r(X),

---

[78] It is not, for example, the one favored by van Fraassen (1989).

then whenever P is a mixture of q and r,

P(X, given Y) > P(X)

(and likewise when we replace the '>' by '<' or '=').

Suppose that the antecedent of this 4-line conditional is true. According to the PC analysis, this means that q(Y→X) > q(X), and r(Y→X) > r(X). Then

P(X, given Y) = P(Y→X)

$\qquad$ = w.q(Y→X) + (1−w).r(Y→X), by definition of P

$\qquad$ > w.q(X) + (1−w).r(X), by our supposition, and the PC analysis

$\qquad$ = P(X).

Hence P(X, given Y) > P(X), respecting the intuition.

I like to think that the lesson of this is the following: (unconditional) probability is additive, and it thus behaves nicely under mixing, taking expectations, and so on. It would be useful to have a notion of conditional probability that also behaves nicely in this respect. The ratio analysis does not provide such a notion; the PC analysis does.

10.4 Conclusion

So much for motivations for the ambiguity thesis. I believe that like the ratio analysis, the PC analysis is simple and fruitful. Furthermore, it will not founder on the problem of the zero denominator, as the ratio analysis does in my opinion. P(A→B) can be defined, even when P(A) = 0 —indeed, these should be paradigm cases of 'supposing', as per Collins. I offer it as another construal of conditional probability. (As should now be clear, there is a second reading of the title of this dissertation.)

As I have already indicated, I am not firmly committed to the ambiguity thesis. One concern I have is that I am not convinced that it handles the problem of undefined terms—for example, the case of the two urns—any better than the ratio analysis does (although I think it has a better chance of doing so). Perhaps, then, conditional probability is three-way, or even more than three-way ambiguous, with the other

disambiguations taking up the slack when required. For example, a third analysis of 'the probability of B, given A, is x' might be: 'if A, then: the probability of B is x', a conditional with a probabilistic consequent. I suspect that much of our thinking involving conditional probability tends to be confused, and that this confusion may well have bred an even more multifarious concept. That would still accord with my main theses: that the Hypothesis is false (so that the disambiguations don't collapse to just one), and that conditional probability is not unambiguously given by the ratio. This work, however, has concerned the relationship between the ratio and the probability of a conditional, and for present purposes I am happy to stop with just that.

Or perhaps a different sort of revision of our thinking about conditional probability is appropriate (and here I offer a mere sketch). Philosophers have traditionally taken unconditional probability as the primitive, and have analyzed conditional probability in terms of it; but perhaps the demise of the Hypothesis, and my proposed loosening of the stranglehold of the ratio analysis, suggest that we should turn the tables, take conditional probability as the primitive, and analyze unconditional probability in terms of *it*. This is the preferred approach of Popper and Renyi, among others. On this view, probability theory is fundamentally a theory about *two*-place functions, P(A,B). Unconditional probability is recovered by putting a tautology in the second place.

What does this have to do with the demise of the Hypothesis?[79] The Hypothesis, in the final analysis, can be seen as a plausible but ultimately flawed attempt to treat conditional probability as a special case of one-place unconditional probability. According to the ambiguity thesis, that attempt does not fail entirely, and there is an important sense of conditional probability which *is* just a special case of unconditional probability. I hope to pursue this line of thought further in the future. But if the ambiguity thesis is also flawed, I think that we should be prepared to recognize the autonomy of two-place conditional probability, and to follow more the lead of Popper

---

[79] I thank Bas van Fraassen for helping me come to the answer to this question.

and Renyi. (I take van Fraassen (1992), for example, to be doing just that.) So while the demise of the Hypothesis may have stymied certain research programs, it should only give impetus to others.

Appendix: versions of the Hypothesis

Here I gather together for easy reference all of the versions of the Hypothesis that I have distinguished, indicating the section in which they first appear.

§1

(0) $\qquad$ P(A→B) = P(B|A)

(CCCP) $\qquad$ P(A→B) = P(B|A) for all A, B in the domain of P, with P(A) > 0.

Universal version: There is some → such that for all P, (CCCP) holds.

Belief function version: There is some → such that for all P that could represent a rational agent's system of beliefs, (CCCP) holds.

Universal tailoring version: For each P there is some → such that (CCCP) holds.

Belief function tailoring version: For each P that could represent a rational agent's system of beliefs, there is some → such that (CCCP) holds.

§7.3

Ideal belief function tailoring version: For each P that could represent an ideally rational agent's system of beliefs, there is some → such that (CCCP) holds.

Human belief function tailoring version: For each P that could represent a rational human agent's system of beliefs, there is some → such that (CCCP) holds.

§7.4

Full probability space tailoring version: For each full <W,F,P>, there is some → such that <W,F,P,→> is a CCCP-model.

§8.2

(Restricted CCCP)   P(A→B) = P(B|A)   for all A, B ∈ **S**

where **S** is some proper subset of F.

Restricted universal version: There is some → such that for all P, (Restricted CCCP) holds.

Restricted belief function version: There is some → such that for all P that could represent a rational agent's system of beliefs, (Restricted CCCP) holds.

§9.1

(Approximate CCCP)   P(A→B) ≈ P(B|A) for all A, B in the domain of P, with

P(A) > 0.

Reading 1. P(A→B) ≈ P(B|A) iff P(A→B) = kP(B|A), where k ≈ 1.

Reading 2. P(A→B) ≈ P(B|A) iff P(B|A) − ε < P(A→B) < P(B|A) + ε, where ε ≈ 0.

We get versions of the Approximate Hypothesis by replacing (CCCP) by (Approximate CCCP) in the statements of versions of the Hypothesis.

§9.2

Precisified belief function version: There is some → such that for all P that could appear in the representor of a rational agent's system of beliefs, (CCCP) holds.

Precisified belief function tailoring version: For each P that could appear in the representor of a rational agent's system of beliefs, there is some → such that (CCCP) holds.

Vague belief version: There is some → such that, for all representors of a rational human agent's system of beliefs, the intervals of vagueness (derived from these representors) of both sides of (CCCP) always agree.

Vague belief tailoring version: For each representor of a rational human agent's system of beliefs, there is some → such that the intervals of vagueness (derived from these representors) of both sides of (CCCP) always agree.

References

Adams, Ernest (1965): "The Logic of Conditionals", *Inquiry* 8, 166-197.

Adams, Ernest (1975): *The Logic of Conditionals,* Reidel.

Anderson, Alan R. and Nuel Belnap (eds.) (1975): *Entailment*, Princeton University Press.

Appiah, Anthony (1985): *Assertion and Conditionals,* Cambridge University Press.

Carlstrom, Ian F. and Hill, Christopher S. (1978): Review of Adams' *The Logic of Conditionals, Philosophy of Science* 45, 155-158.

Collins, John (1991): *Belief Revision*, Ph.D. dissertation, Princeton University.

de Finetti, Bruno (1936): "La Logique de la Probabilité", in *Induction et Probabilité*, *Actualités Scientifiques et Industrielles* 391, 31-39.

de Finetti, Bruno (1972): *Probability, Induction and Statistics*, Wiley.

Eells, Ellery, and Brian Skyrms (eds.) (1993), *Probability and Conditionals* (working title).

Ellis, Brian (1969): "An Epistemological Concept of Truth", in Robert Brown and C.D. Rollins (eds.), *Contemporary Philosophy in Australia*, London.

Gibbard, Allan (1981): "Two Recent Theories of Conditionals", in Harper et al.

Hájek, Alan (1989): "Probabilities of Conditionals—Revisited", *Journal of Philosophical Logic* 18, 423-428.

Hájek, Alan (1992): "A Dogma of Conditional Probability", unpublished manuscript.

Hall, Ned (1993): "Back in the (CCCP)", to appear in Eells and Skyrms.

Halpin, John (1991): "What is the Logical Form of Probability Assignment in Quantum Mechanics?", *Philosophy of Science* 58, 1991.

Harman, Gilbert (1989): *Change in View*, MIT Press, Cambridge, Massachussetts.

Harper, W.L., Stalnaker, R., and Pearce, G. (eds.) (1981): *Ifs,* Reidel, Dordrecht.

Jackson, Frank (1987): *Conditionals*, Blackwell, Oxford.

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

Jeffrey, Richard (1965): *The Logic of Decision*, University of Chicago.

Jeffrey, Richard (1983): "Bayesianism With a Human Face", in Jeffrey (1992).

Jeffrey, Richard (1991): "Matter-Of-Fact Conditionals", The Supplementary Volume LXV, *The Aristotelian Society*.

Jeffrey, Richard (1992): *Probability and the Art of Judgment*, Cambridge Studies in Probability, Induction, and Decision Theory.

Jeffrey, Richard, and Robert Stalnaker (1993): "Conditionals as Random Variables", to appear in Eells and Skyrms.

Joyce, James (1991): *The Axiomatic Foundations of Bayesian Decision Theory*, Ph.D dissertation, University of Michigan.

Levi, Isaac (1980): *The Enterprise of Knowledge,* MIT Press, Cambridge, Massachussetts.

Lewis, David (1971): "Completeness and Decidability of Three Logics of Counterfactual Conditionals", *Theoria* 37, 74-85.

Lewis, David (1973): *Counterfactuals*, Blackwell and Harvard University Press.

Lewis, David (1973a): "Counterfactuals and Comparative Possibility", in Harper et al.

Lewis, David (1975): "Adverbs of Quantification", in Edward L. Keenan, ed., *Formal Semantics of Natural Language* (Cambridge University Press).

Lewis, David (1976): "Probabilities of Conditionals and Conditional Probabilities", *Philosophical Review* 85, 297-315; reprinted in Harper et al.

Lewis, David (1979): "Attitudes *De Dicto* and *De Se*", *Philosophical Review* 88, 513-43; reprinted in Lewis (1983).

Lewis, David (1981): "A Subjectivist's Guide to Objective Chance", in Harper et al.

Lewis, David (1981a): "Causal Decision Theory", *Australasian Journal of Philosophy* 59, 5-30; reprinted in Lewis (1986a).

Lewis, David (1983): *Philosophical Papers*, Volume I, Oxford University Press.

Lewis, David (1986): "Probabilities of Conditionals and Conditional Probabilities II", *Philosophical Review* 95, 581-589.

Lewis, David (1986a): *Philosophical Papers*, Volume II, Oxford University Press.

Lewis, David (1992): "Meaning Without Use: Reply to Hawthorne", *Australasian Journal of Philosophy* 70, No. 1, 106-110.

Lewis, David (1992a): "Humean Supervenience Debugged" (talk presented at the 1992 Australasian Association of Philosophy Conference, Brisbane, Australia).

McGee, Vann (1985): "A Counterexample to Modus Ponens", *The Journal of Philosophy* 82, 462-470.

Nozick, Robert (1969): "Newcomb's Problem and Two Principles of Choice", in N. Rescher et al., eds., *Essays in Honor of Carl G. Hempel*, Reidel, 1970.

Quine, Willard Van Orman (1950): *Methods of Logic*, Holt, New York.

Ramsey, Frank Plumpton (1965): *The Foundations of Mathematics (and Other Logical Essays)*, Routledge and Kegan Paul.

Skyrms, Brian (1987): "Updating, Supposing, and MAXENT", *Theory and Decision*, 225-246.

Stalnaker, Robert (1968): "A Theory of Conditionals", *Studies in Logical Theory*, American Philosphical Quarterly Monograph Series, No. 2, Blackwell, Oxford.

Stalnaker, Robert (1970): "Probability and Conditionals", *Philosophy of Science* 37, 64-80; reprinted in Harper et al.

Stalnaker, Robert (1976): "Letter to van Fraassen", in Harper and Hooker (eds.), *Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science*, Vol. I, Reidel, 302-306.

van Fraassen, Bas (1976): "Probabilities of Conditionals", in Harper and Hooker (eds.), *Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science*, Vol. I, Reidel, 261-301.

van Fraassen, Bas (1981): "A Temporal Framework for Conditionals and Chance", in Harper et al.

van Fraassen, Bas (1984): "Belief and the Will", *Journal of Philosophy* 81, 235-256.

van Fraassen, Bas (1989): *Laws and Symmetry*, Clarendon Press, Oxford.

van Fraassen, Bas (1990): "Figures in a Probability Landscape", in J.M. Dunn and A. Gupta (eds.), *Truth or Consequences*, Kluwer.

van Fraassen, Bas (1992): "Fine Grained Opinion and the Logic of Full Belief", unpublished manuscript.

Walley, Peter (1991): *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall.