Revisiting McGee's Probabilistic Analysis of Conditionals

John Cantwell[1]

Abstract

This paper calls for a re-appraisal of McGee's analysis of the semantics, logic and probabilities of indicative conditionals presented in his 1989 paper *Conditional probabilities and compounds of conditionals*. The probabilistic measures introduced by McGee are given a new axiomatisation—built on the principle that the antecedent of a conditional is probabilistically independent of the conditional—and a more transparent method of constructing such measures is provided. McGee's Dutch book argument is restructured to more clearly reveal that it introduces a novel contribution to the epistemology of semantic indeterminacy, and shows that its more controversial implications are unavoidable if we want to maintain the Ramsey Test along with the standard laws of probability. Importantly, it is shown that the counterexamples that have been levelled at McGee's analysis—generating a rather wide consensus that it yields 'unintuitive' or 'wrong' probabilities for compounds —fail to strike at their intended target; for to honour the intuitions of the counterexamples one must either give up the Ramsey Test or the standard laws of probability. It will be argued that we need to give up neither if we take the counterexamples as further evidence that the indicative conditional sometimes allows for a non-epistemic 'causal' interpretation alongside its usual epistemic interpretation.

Keywords Probability of conditionals · Compounds of conditionals ·
Sequence semantics

1 Introduction

In 1989 McGee published *Conditional probabilities and compounds of conditionals*—a seminal paper on the semantics, logic and probabilities of indicative conditional sentences and sentences containing compounds of such conditionals.

✉ John Cantwell
cantwell@kth.se

1 Division of Philosophy, KTH - Royal Institute of Technology, 100 44. Stockholm, Sweden

It showed that—somewhat surprisingly—a number of pervasive logical, epistemic and semantic intuitions concerning indicative conditionals can be combined into a coherent formal analysis. However, while the paper is widely cited, the analytical framework introduced in the paper has not been developed and studied in the way one would expect given the richness of the framework and its close conformity to principles that—while controversial—seem relatively well grounded in empirical fact.

This paper presents a series of results intended to bring to further light the depth of McGee's analytical framework, both formally and philosophically. An alternative, more transparent, axiomatisation of the probability measures introduced by McGee (henceforth: *McGee-measures*) is provided, yielding insight into its probabilistic constraints. A more direct and transparent way of constructing such measures from factual probabilities is provided, removing any suspicion that the selection function semantics draws on some notion of similarity between worlds. The Dutch book arguments provided by McGee—that contain an uneasy mix of syntactic and semantic principles—are shown to be grounded in purely seman-tic considerations, and it is shown that their more philosophically controversial implications are unavoidable given the laws of probability (additivity) and the Ramsey Test.

It is shown, finally, that a whole cluster of counterexamples that have been lev-elled specifically at McGee's framework, resulting in a rather wide consensus that its implications are 'unintuitive'—indeed *wrong*—miss their target. For to honour the intuitions one must either give up the laws of probability (additivity) or the Ramsey Test, and so the intuitions behind the counterexamples cannot be honoured without giving up either of these. Moreover, it is argued that at least some of the intuitions are perhaps best explained as involving a non-epistemic causal reading of the indicative conditional, a reading that sometimes coexists with the default epistemic reading (a phenomenon studied by, for instance, [11–13]).

The results call for a reassessment of McGee's analysis, in its favour. Some of these results also have direct implications for a family of related analyses that take the probabilistic semantics of Stalnaker and Jeffrey [33] as their starting point (and which, logically, differ from McGee's analysis only in their treatment of conditionals with conditional consequents). Indeed, the results suggest that beyond these there is very little room for alternative analyses that aspire to maintain both the Ramsey Test and the standard laws of probability.

2 McGee's Framework

2.1 A Formal Semantics for Conditionals

The *factual* language L_F (the set of *factual sentences*) contains atomic sentences p, q, r, and is closed under the boolean connectives: if A and B belong to L_F, then so do $\neg A$, $A \lor B$, $A \land B$, $A \supset B$, and $A \equiv B$. Capital Roman letters at the beginning of the alphabet (A, B, C, D, E) will exclusively be used to denote factual sentences.

The *base language* L_B contains the factual language, is closed under the boolean connectives and, moreover, is closed under the clause: if A, B belong to L_F, then so does $A \rightarrow B$. A sentence of the form $A \rightarrow B$ is called a *base conditional*.

The *full language* L_M contains the factual language, is closed under the boolean connectives and, moreover, is closed under the clause: if A belongs to L_F and φ belongs to L_M, then so does $A \rightarrow \varphi$. Greek letters φ, ψ, χ, etc. will be used to denote arbitrary sentences of L_M. Note that conditionals even in the full language are only allowed to have factual antecedents.

The semantic core of McGee's framework is a standard selection-function semantics in the style of Stalnaker, where the selection functions can be thought of as selecting among the epistemically possible worlds. Let W be a set of worlds. Subsets of W ("worldly propositions") will be denoted by X, Y, Z, and so on. A *selection function f* is a partial function that for some worldly propositions X returns a world $w = f(X)$. Selection functions are supposed to satisfy

1. $f(W)$ is defined.
2. If $f(X)$ is defined and $f(X) \in Y$, then $f(Y)$ is defined.
3. $f(X) \in X$ (Centering).
4. If $f(X) \in Y$, then $f(X \cap Y) = f(X)$.

From the above it follows that each selection function is associated with some non-empty subset of D of W such that $D \cap X$ is non-empty iff $f(X)$ is defined. This set is denoted by $D(f)$. The world $f(W)$ will be said to be the *preferred* world of f and will be denoted by w_f. Define the restriction of f to X, in symbols, f/X as the function:

$$f/X(Y) = f(X \cap Y).$$

Sets of selection functions will be denoted by P, Q, R, etc., and will be used to represent *propositions*. The *factual content* of a proposition P will be denoted by $F(P) = \{w_f : f \in P\}$. Where X is a set of worlds $\mathcal{P}(X)$ will be the proposition $\{f : f_w \in X\}$ ($\mathcal{P}(X)$ is the weakest proposition P such that $F(P) = X$, it has 'purely factual' content). The set of all selection functions (the trivially true proposition) will be denoted by $\mathcal{F}$.

It should be noted that McGee does not explicitly refer to 'propositions' in this respect. Often, propositions are represented as sets of worlds. However, if we take propositions to denote the contents of sentences (as used in a particular context), and to be the objects of propositional attitudes (so that the probability of a sentence in a context becomes the probability that the proposition that the sentence expresses is true), then, when conditionals are involved, we need to represent propositions using a richer structure, in this case as sets of selection functions. (It is widely realised that due to the impossibility results, a conditional satisfying the Ramsey Test cannot be treated as expressing an ordinary factual proposition, more on this below).

A model for the language is a pair $M = (W, V)$ where W is a set of worlds and V is a function that assigns to each atomic sentence some subset of W. We will define a function $[\![\varphi]\!]_M^f$ that has the sentences of L_M as its domain and returns one of the values 0 or 1. Before giving a recursive definition it is useful to specify two auxiliary

notions. The proposition expressed by a sentence is given by (reference to the model will typically be omitted):

$$[\varphi] = \{f : [\![\varphi]\!]^f = 1\}.$$

(For richer languages we would need to relativize the proposition expressed by a sentence to a context, and indeed at least one feature of contexts—what is taken as given in a context—becomes relevant in subsequent discussions, see Section 3.) The factual content of this proposition will be denoted by $F(\varphi)$.

$$[\![p]\!]^f = 1 \quad \text{iff } w_f \in V(p)$$
$$[\![\neg\varphi]\!]^f = 1 \quad \text{iff } [\![\varphi]\!]^f = 0.$$
$$[\![\varphi \vee \psi]\!]^f = 1 \quad \text{iff } [\![\varphi]\!]^f = 1 \text{ or } [\![\psi]\!]^f = 1, \text{ and so on for the other boolean}$$
$$\text{connectives.}$$
$$[\![A \to \psi]\!]^f = 1 \quad \text{iff } [\![\psi]\!]^{f/F(A)} = 1 \text{ or } D(f) \cap F(A) = \emptyset.$$

Note how, in the clause for the conditional, the antecedent of the conditional is used to restrict the domain of the selection function f. McGee's semantics is thus an early instance of a *restrictor* semantics for the conditional: the role of the antecedent is to restrict the domain of possibilities relative to which the consequent is evaluated.[1]

Definition 1 The logic ICL is the consequence relation:

$$\varphi_1, \ldots, \varphi_n \models_{\text{ICL}} \psi \text{ iff for every model } M: [\varphi_1]_M \cap \cdots \cap [\varphi_n]_M \subseteq [\psi]_M.$$

If φ follows from no premises it is said to be *ICL-valid* (or simply *valid*).

McGee provided the first complete axiomatisation of a conditional with a restrictor semantics. ICL is the smallest Tarskian consequence relation[2] that satisfies:

MP$^\supset$ $\varphi, \varphi \supset \psi \models_{\text{ICL}} \psi$.
CL $\models_{\text{ICL}} \tau$, for any classical tautology τ.
CN   If $A \models_{\text{ICL}} \psi$, then $\models_{\text{ICL}} A \to \psi$.
CK   If $\psi_1, \ldots, \psi_n \models_{\text{ICL}} \chi$, then $A \to \psi_1, \ldots, A \to \psi_n \models_{\text{ICL}} A \to \chi$.
LLE   If $\models_{\text{ICL}} A \equiv B$, then $\models_{\text{ICL}} (A \to \varphi) \equiv (B \to \varphi)$.
CEM $\models_{\text{ICL}} (A \to \varphi) \vee (A \to \neg\varphi)$.
Import-Export $\models_{\text{ICL}} (A \to (B \to \varphi)) \equiv ((A \wedge B) \to \varphi)$.
rMP $A, A \to B \models_{\text{ICL}} B$.
rCS $A, B \models_{\text{ICL}} A \to B$.

Note the restrictions to factual sentences in rMP and rCS. The perhaps most distinctive logical property of ICL is Import-Export. As noted by [7] a conditional satisfying both Import-Export and unrestricted modus ponens collapses into the material conditional. Consequently ICL only satisfies a restricted form of modus ponens (restricted

---

[1]Restrictor semantics for conditionals was first introduced by Angelika Kratzer (18, see collection in 19). More recent accounts are found in [16, 34] and [38], while [3] explicitly employs McGee's framework.
[2]A Tarskian consequence relation satisfies (i) $\varphi \models \varphi$, (ii) $\Gamma \models \psi$ implies $\Gamma, \varphi \models \psi$, and (iii) $\Gamma \models \varphi$ and $\Gamma, \varphi \models \psi$ together imply $\Gamma \models \psi$.

to factual consequents). While 'unorthodox', it is supported by both widely shared intuitions and experimental evidence.[3]

2.2 McGee-measures

Definition 2 Given a logic $\mathcal{L}$ on a language L, an $\mathcal{L}$-*measure* is a real-valued function $\Pr(\cdot)$ with domain L such that:

1. $0 \leq \Pr(\varphi) \leq 1$.
2. $\Pr(\varphi) = 1$, if φ is $\mathcal{L}$-valid.
3. $\Pr(\varphi \vee \psi) = \Pr(\varphi) + \Pr(\psi)$, if $\neg(\varphi \wedge \psi)$ is $\mathcal{L}$-valid. (Additivity)

The logic of primary concern is ICL. An ICL-measure on the factual language L_F will be called *factual*, a measure on the base language L_B will be called a *base* measure, while an ICL-measure on the full language L_M will be called *full*. A measure that doesn't satisfy additivity is (probabilistically) non-additive. An ICL-measure will be said to be *regular* if it satisfies:

$$\text{If } \Pr(A) = 0, \text{ then } \Pr(A \rightarrow \psi) = 1.$$

Definition 3 A *McGee-measure* is a full ICL-measure satisfying *Generalised Independence* (McGee simply calls it *Independence*):

(GI) $$\Pr(C \wedge (A_1 \rightarrow B_1) \wedge \cdots \wedge (A_n \rightarrow B_n)) = \\ \Pr(C) \times \Pr((A_1 \rightarrow B_1) \wedge \cdots \wedge (A_n \rightarrow B_n)),$$

where each C, A_i and B_i is factual, $\Pr(A_i) > 0$, and C is truth-functionally incompatible with each A_i.

A core target property for an analysis of the epistemic interpretation of the indicative conditional is the *Ramsey Test*. It relates probabilities of conditionals to conditional probabilities, stating that $\Pr(A \rightarrow \varphi)$ = the probability of φ given A. Writing $\Pr_A(\varphi)$ for the probability of φ given A, the Ramsey Test thus states that $\Pr(A \rightarrow \varphi) = \Pr_A(\varphi)$. The standard *ratio analysis* in turn provides an interpretation of conditional probabilities, stating that, when $\Pr(A) > 0$, $\Pr_A(\varphi)$ (the probability of φ given A) $= \Pr(A \wedge \varphi)/\Pr(A)$ (the *ratio rule*). Combining the Ramsey Test with the ratio rule we get: $\Pr(A \rightarrow \varphi) = \Pr_A(\varphi) = \Pr(A \wedge \varphi)/\Pr(A)$.[4] McGee-measures, however, satisfy only a restricted version of the probabilistic Ramsey Test on its ratio analysis:

Observation 1 Any McGee-measure will satisfy (for factual A and B):

rRT $\Pr(A \rightarrow B) = \Pr(A \wedge B)/\Pr(A)$, if $\Pr(A) > 0$.

---

[3]See [36] for experimental evidence for Import-Export and McGee's(25, p.462) counterexample to modus ponens.

[4]The combination of the unrestricted Ramsey Test and the unrestricted ratio rule is sometimes called *Stalnaker's thesis*. In the restricted forms (where consequents of conditionals are assumed to be factual) it is sometimes called *Adams' thesis*.

Proof By GI $\Pr(\neg A \wedge (A \rightarrow B)) = \Pr(\neg A) \times \Pr(A \rightarrow B)$ when $\Pr(A) > 0$. By standard principles of probability we then have $\Pr(A \wedge (A \rightarrow B)) = \Pr(A) \times \Pr(A \rightarrow B)$. As, by rMP and rCS, $A \wedge (A \rightarrow B)$ is logically equivalent to $A \wedge B$, $\Pr(A \wedge B) = \Pr(A) \times \Pr(A \rightarrow B)$. So $\Pr(A \rightarrow B) = \Pr(A \wedge B)/\Pr(A)$. $\qquad\square$

So, given the background logic, GI implies rRT. The converse direction does not hold:

Observation 2 A regular ICL-measure satisfying rRT need not satisfy GI.

(Longer proofs of Observations and Theorems are found in the Appendix.) So given ICL as the background logic, rRT is strictly weaker than GI.

The restriction to factual consequents in rRT is essential. We know from Lewis' triviality results [21, 22] that when the underlying logic satisfies Import-Export, and conditional probabilities satisfy the ratio rule, no non-trivial measure can satisfy the unrestricted Ramsey Test. McGee-measures, however, provably do *not* collapse into triviality, as witnessed by McGee's central result:

Theorem 1 [26] *For any factual ICL-measure Pr′ there exists exactly one regular McGee-measure Pr such that $Pr(A) = Pr'(A)$ for all factual sentences A.*

The *existence* part of the claim ensures non-triviality. The *uniqueness* part demonstrates just how strong McGee's principle GI is. It means that the probability of any sentence containing conditionals, no matter how complex, is completely determined by the probabilities of the factual sentences (indeed of the probabilities of the boolean combinations of the factual sentences occurring in the sentence).

If we consider only the base language (the language where conditionals only have factual consequents) McGee-measures will coincide with the measures studied by [33] on the corresponding language (see also [12] for a fuller treatment).

Notably, while the empirical support for rRT is wide and well documented (e.g. [4, 6, 30, 31, 37]), the empirical support for the *unrestricted* Ramsey Test—assuming the unrestricted ratio rule for conditional probabilities—is very weak. Indeed, it is easy to generate counterexamples.

Say that a card is to be drawn from a deck of cards. The conditional "If it is not diamonds ($\neg d$), then if it it's red (r) it's hearts (h)" should presumably have probability 1. But $\Pr(\neg d) = 3/4$ and $\Pr(r \rightarrow h) = 1/2$ so $\Pr(\neg d \wedge (r \rightarrow h)) \leq 1/2$ and so, combining the unrestricted Ramsey Test with the unrestricted ratio rule:

$$\Pr(\neg d \rightarrow (r \rightarrow h)) = \frac{\Pr(\neg d \wedge (r \rightarrow h))}{\Pr(\neg d)} \leq \frac{1/2}{3/4} \ (= 2/3).$$

This does not seem right, and it presents a challenge to the contention that the Ramsey Test is the defining feature of the epistemic interpretation of the indicative conditional. However, McGee proposed that the problem resides not in the Ramsey Test, but in the ratio analysis of conditional probabilities when conditionals are involved. Indeed, if we—like McGee—set $\Pr_A(\varphi) = \Pr(A \rightarrow \varphi)$, the resulting analysis of the

conditional probability $\Pr_A$ will satisfy the restricted ratio rule (witness rRT), but not the unrestricted ratio rule, and it will by definition satisfy the unrestricted Ramsey Test. Importantly, on this analysis, the class of McGee-measures is closed under conditionalisation (if Pr is a McGee-measure then so is $\Pr_A$, when $\Pr(A) > 0$), as is the class of regular McGee-measures (if Pr is a regular McGee-measure then so is $\Pr_A$, when $\Pr(A) > 0$).

3 Re-axiomatising McGee-measures

GI is stronger than rRT making the very strength of GI a bit worrisome. How plausible is it? GI is a property governing the probability of a conjunction of conditionals. But it is notoriously difficult to form intuitions about the probability of a conjunction of conditionals. So why should the principle be trusted?

Consider the following principle, here called Independence of Antecedents (IA):

IA $\Pr(A \wedge (A \to \varphi)) = \Pr(A) \times \Pr(A \to \varphi).$

Note that φ here is not restricted to factual sentences.

IA states that the antecedent of a conditional is probabilistically independent of the conditional. This reflects a property of evidential relationships that has been regarded as fundamental, sometimes referred to as *Rigidity* [1, 10]: the extent to which P is evidence for Q is independent of the probability of P. IA is thus the exact property we would expect of a conditional that seeks to track evidential relationships.

Theorem 2 *A regular ICL-measure will satisfy IA if and only if it satisfies GI.*

So we can replace GI by IA and get an alternative axiomatisation of regular McGee-measures. With ICL as the background logic, the class of regular McGee-measures is characterised by the principle that the antecedent of a conditional is probabilistically independent of the conditional.

If we rearrange IA we get:

$$\Pr(A \to \varphi) = \frac{\Pr(A \wedge (A \to \varphi))}{\Pr(A)}, \text{ if } \Pr(A) > 0.$$

Notably, IA in this format has more than a passing resemblance to the Ramsey Test (in its ratio rule format). Indeed, when we restrict the conditional involved to factual consequents (call the resulting principle rIA) then it becomes *equivalent* to the restricted Ramsey Test:

Observation 3 An ICL-measure satisfies rIA if and only if it satisfies rRT.

Proof By rMP and rCS, $A \wedge (A \to B)$ is logically equivalent to $A \wedge B$. So

$\Pr(A \wedge (A \to B)) = \Pr(A \wedge B)$. The equivalence of rIA and rRT follows immediately.   □

So McGee's analysis of conditional probabilities that sets $\text{Pr}_A(\varphi) = \text{Pr}(A \to \varphi)$ can, when $\text{Pr}(A) > 0$, equivalently be given as:

$$\text{Pr}_A(\varphi) = \frac{\text{Pr}(A \wedge (A \to \varphi))}{\text{Pr}(A)}.$$

This is not quite the ratio rule, and assumes a language that contains the conditional. But using McGee's semantics it can be motivated as a 'semantically sophisticated' version of the ratio rule. In two steps. First, as conditionals semantically operate by restricting the domain of possibilities, we can plausibly hold that the proposition expressed by a conditional sentence can vary with what is *assumed* (given) in a context; for what is assumed to hold restrict the contextually relevant set of possibilities to which conditionals are semantically sensitive (we find the idea that the context comes with a constrained set of possibilities in [19], see also [32]; the present idea that an assumption can serve to constrain the set of possibilities and so change the content of a conditional is similar to Mandelkern's [23] idea of a *local* context).[5] Second—plausibly—a semantically sophisticated ratio rule analysis of $\text{Pr}_A(\varphi)$ should be based on the probability of the proposition expressed by φ *in the context where it is given that* A.

Now combine these two ideas. Let $[\varphi]^X$ denote the proposition expressed by φ in the context where the set of possible worlds have been restricted to X (this should be non-empty), and set (where, recall, $\mathcal{P}(X)$ is the proposition with purely factual content X)

$$[\varphi]^X = (\mathcal{P}(X) \cap \{f \in \mathcal{F} \,:\, f/X \in [\varphi]\}) \cup (-\mathcal{P}(X) \cap [\varphi]).$$

So:

$$[\varphi]^{F(A)} = [(A \wedge (A \to \varphi)) \vee (\neg A \wedge \varphi)].$$

We have $[\varphi] = [\varphi]^W$. For any factual sentence A we have $[A]^X = [A]$ (the proposition expressed by a factual sentence does not vary with what is given). Moreover, $[A \wedge \varphi]^{F(A)} = [A \wedge (A \to \varphi)]$ (one of the terms of IA). So, assuming a probability measure pr on propositions (rather than sentences) where $\text{Pr}(\varphi) = \text{pr}([\varphi])$, the analysis we seek becomes:

$$\text{Pr}_A(\varphi) = \frac{\text{pr}([A] \cap [\varphi]^{F(A)})}{\text{pr}([A])} \quad \left(= \frac{\text{Pr}(A \wedge (A \to \varphi))}{\text{Pr}(A)} \right).$$

This is more a proof-of-concept than a full-fledged analysis. A proper analysis would have to motivate the definition of $[\varphi]^X$. It does suggest, however, that we can retain a version of the ratio rule on the level of propositional content, even though we can't retain it on the level of syntax. The problem with the ordinary unrestricted ratio rule is diagnosed as a failure to take into consideration the fact that conditionals are semantically sensitive to what is assumed. Importantly, on the revised analysis of conditional

---

[5]Indeed, the idea that conditionals are semantically context sensitive is itself far from new. For instance, [35] showed that it can be used to block Lewis-style triviality proofs. But van Fraassen's development of the idea has also been criticised (e.g. [9, 21]) as it makes conditionals semantically hyper-sensitive: any two speakers with different degrees of belief will mean different things by the same conditional. The present suggestion is less radical: conditionals are only semantically sensitive to what is taken as given in the context, and not on the credences of individual speakers ([17] calls this *moderate* context sensitivity).

probabilities, IA—the characteristic axiom of regular McGee-measures—turns out to be equivalent to the target principle that conditionals satisfy the unrestricted Ramsey Test:

RT $\quad \mathrm{Pr}(A \rightarrow \varphi) = \mathrm{Pr}_A(\varphi).$

4 Constructing McGee-measures

Any factual measure can be extended to a McGee-measure. McGee proved as much. The proof—which is far from trivial—involves a rather complex way of constructing McGee-measures from factual measures. The construction serves its purpose in the proof, but is less than transparent and so provides only a limited understanding of how the probabilities of complex conditional sentences come to be determined by probabilities of factual sentences.

Consider a finite model $M = (W, V)$. A probability mass on *worlds* is a function m from W to $[0, 1]$ such that $\sum_{w \in W} m(w) = 1$. Where X is a set of worlds let $m(X) = \sum_{w \in X} m(w)$. A probability mass m on worlds lets us in a standard way define a probability function Pr_m on factual sentences:

$$\mathrm{Pr}_m(\varphi) = m(F(\varphi)).$$

A probability mass m on *selection functions* is a function m from $\mathcal{F}$ to $[0, 1]$ such that $\sum_{f \in \mathcal{F}} m(f) = 1$. As before let $m(P) = \sum_{f \in P} m(f)$. Define the following function on sentences of the full language:

$$\mathrm{Pr}_m(\varphi) = m([\varphi]).$$

When thus defined, Pr_m will be an ICL-measure, but it won't necessarily be a McGee-measure.

The task is to show how from a probability mass m on worlds one can construct a probability mass $\hat{m}$ on selection functions such that $\mathrm{Pr}_{\hat{m}}$ is the unique McGee-measure extending the factual measure Pr_m. To this end we can explore the fact that any selection function f with finite domain can be equivalently represented as an *ordered sequence* $s = \langle w_1, \ldots, w_n \rangle$ where $w_1, \ldots, w_n$ are non-identical worlds that jointly make up the domain of f, and where, for any set of worlds X, $f(X)$ selects the element of X that occurs first in the sequence s. Let $s[i]$ denote the i-th element of s.

One can think of a selection function as a strategy for selecting the actual world given different sets of alternatives, keeping in mind that one doesn't *know* which world is the actual world. Different strategies can be more or less likely to be successful—i.e. more or less likely to select the actual world from a given set of alternatives. Say that one can choose from *all* worlds and that the strategy says to choose w_1. Here the probability of choosing the actual world is the probability of w_1 itself. Now say that we instead remove w_1 from the set of alternatives and that the strategy says to pick the world w_2 from $W - \{w_1\}$. The probability of having made the right choice, *assuming* that the actual world is in $W - \{w_1\}$, is the probability of w_2 given $W - \{w_1\}$. And so on. An ordered sequence represents a selection strategy by giving

the order in which worlds would be chosen when alternatives are gradually removed. It turns out that we get McGee-measures by taking the probability of a sequence (a choice strategy) to be the product of the probabilities of the individual choices.[6]

Consider a probability mass m on worlds. Let $D(m)$ be the set of worlds w such that $m(w) > 0$. The probability mass on selection functions (now represented as sequences) generated by m, in symbols $\hat{m}$, is defined as follows. For any sequence s with domain of cardinality n, when $D(s) = D(m)$ set:

$$\hat{m}(s) = \prod_{1 \leq i \leq n} \frac{m(s[i])}{\sum_{i \leq j \leq n} m(s[j])}.$$

When $D(s) \neq D(m)$ set $\hat{m}(s) = 0$.[7]

---

[6]The use of sequences echoes van Fraassen's (35) *Stalnaker-Bernoulli* models, that are also employed by [33] (and in subsequent developments of their framework). See also [12]. In their sequences a possible world may occur countably many times in the sequence (in the present construct it can appear only once), thus even in a model with finitely many possible worlds, there are infinitely many sequences. As Kaufmann puts it "each such sequence corresponds to the outcome of an infinite series of random choices of worlds... with replacement, where the expectation of each trial is independent of the previous outcomes" (p.6). So they employ a different space of sequences, and a very different motivation for their use. Moreover, the semantic clause for conditionals is different. However, for base conditionals the possibility of multiple successive choices is irrelevant (there is only one choice), so, as already mentioned, for languages containing only base conditionals this analysis generates exactly the set of McGee-measures on such languages (so they too satisfy GI).

[7] An anonymous referee has brought to my attention that [15] (p.65) in lecture notes, and more elaborated in [14] (forthcoming) and [8] (unpublished), present a construction using sequences that is similar (but not identical) to the present construction. They place constraints on probability measures on *sets* of sequences (propositions), where sequences do not allow for repetition of worlds. The details of their frameworks differ, but the general format is that they place the following constraints on a measure pr on propositions (relative to some set P of sequences, where m is a probability measure on worlds, and $\langle w_1, \ldots, w_n, - \rangle$ denotes the set of sequences s in P that have $\langle w_1, \ldots, w_n \rangle$ as their initial segment):

$$\text{pr}(\langle w, - \rangle) = m(w).$$
$$\text{pr}(\langle w_1, \ldots, w_n, - \rangle) = \text{pr}(\langle w_1, \ldots, w_{n-1}, - \rangle) \times m(w_n)/m(W - \{w_1, \ldots, w_{n-1}\}).$$

The second clause in particular is similar in structure to the present construct.

Constraints do not guarantee existence. A problem is that the properties of pr depend on the choice of P. If P is not judiciously chosen (e.g., if we let P be the set of all sequences) the constraints do not ensure additivity (the probability of a set of sequences is the sum of the probabilities of the individual sequences), a property one would expect if individual sequences are the carriers of probability. In the construction of $\hat{m}$ in the present paper it is postulated that all and only sequences with domain $D(m)$ will be assigned a positive probability; this ensures that the probabilities add to 1 and the property of Regularity. Indeed if we let P be the set of sequences with domain $D(m)$, we get pr $= \hat{m}$ (e.g., [14] has a different constraint on P that also delivers). So (in this case) the constraints force the same probabilities to propositions as McGee measures. This congruence, however, is at the level of propositions, not of sentences. The three cited accounts that employ the above constraints give three very different semantic clauses for the conditional ([8] comes closest to McGee's semantics by taking the antecedent of a conditional to function as a domain restrictor); so the mapping of sentences to propositions need not coincide with McGee's. As a result it does not follow, for instance, that they satisfy McGee's principle GI (which is not discussed). So while these frameworks clearly are related to McGee's, more work needs to be done to untangle how far this relationship goes.

For instance, let $D(m) = \{w_1, w_2, w_3, w_4\}$, with $m(w_1) = .2$, $m(w_2) = .4$, $m(w_3) = .15$ and $m(w_4) = .25$ and let s be the ordered sequence $s = \langle w_3, w_1, w_4, w_2 \rangle$. Then:

$$\hat{m}(s) = \frac{m(w_3)}{m(\{w_1, w_2, w_3, w_4\})} \times \frac{m(w_1)}{m(\{w_1, w_2, w_4\})} \times \frac{m(w_4)}{m(\{w_2, w_4\})} \times \frac{m(w_2)}{m(w_2)}$$

$$= \frac{.15}{1} \times \frac{.2}{.85} \times \frac{.25}{.65} \times \frac{.4}{.4} = .014$$

Theorem 3 *$Pr_{\hat{m}}$ is the unique regular McGee-measure extending the factual measure Pr_m.*

The result to some extent clarifies the deeply epistemic role of selection functions in McGee's framework. If a world w_1 is more likely than w_2, then, one will have more confidence in selection strategies that prioritize w_1 over w_2 rather than the other way around (the order of the other worlds being equal), with most confidence afforded to those strategies that order worlds by their probability (in descending order). The confidence one has in a selection strategy is entirely based on probabilistic considerations of factual (non-modal) matters, and so does not intrinsically encode any sensitivity to extra-epistemic facts (like 'physical similarity' between worlds).

The construction also dictates how we are to conditionalise a mass $\hat{m}$ on sequences with a factual proposition X. Let m be a factual probability mass and let m_X be the mass (for $m(X) > 0$):

$$m_X(w) = \begin{cases} m(w)/m(X), & \text{if } w \in X; \\ 0, & \text{otherwise.} \end{cases}$$

m_X is the factual probability mass that we get when conditionalising m on X. Let $\hat{m}_X$ be the mass on sequences generated from m_X according to the above construction. The relationship between $\hat{m}$ (the mass on sequences generated from m) and $\hat{m}_X$ (the mass on sequences generated from m_X) is not explicitly given, but is still fixed by the construction. We get $\hat{m}_X$ from $\hat{m}$ by *imaging* $\hat{m}$ on X in a special way, specifically, by transferring the probability of each sequence s to the sequence s/X (where s/X is the sequence we get when removing the non-X worlds from s) and summing (to deal with the cases where $s/X = s'/X$ even though $s \neq s'$):

Theorem 4 *For any probability mass m on worlds such that $m(X) > 0$, and for any sequence s:*

$$\hat{m}_X(s) = \sum_{s'/X=s} \hat{m}(s').$$

A sequence encodes what is possible, and an assumption serves to restrict what is possible; so on conditionalising by X we need to shift the probability of a sequence to its conditionalised counterpart.

From Theorem 3 we know that $Pr_{\hat{m}_X}$ is the McGee-measure that uniquely extends the factual measure Pr_{m_X}. The relationship between $Pr_{\hat{m}_X}$ and $Pr_{\hat{m}}$ is not explicitly given but is also fixed by the construction:

Corollary 1 For any factual sentence A:

$$\mathrm{Pr}_{\hat{m}_{F(A)}}(\varphi) = \mathrm{Pr}_{\hat{m}}(A \to \varphi), \text{ if } \mathrm{Pr}_{\hat{m}}(A) > 0.$$

(The proof is embedded in the proof of Theorem 4). The effect of standard ratio rule conditionalisation on the underlying factual ptobabilities will, given the present construction of probabilities of sequences, result in unrestricted RT when taking the whole language into account.

5 Justification by Dutch Book

McGee provided a synchronic Dutch book argument in favour of McGee-measures (his Theorem 1), and a diachronic Dutch book (his Theorem 5) in favour of the identity $\mathrm{Pr}_A(\varphi) = \mathrm{Pr}(A \to \varphi)$. The synchronic Dutch book establishes that any-one whose credences violates the constraints given by McGee-measures can face a collection of bets that individually seem fair but jointly give a guaranteed loss. The diachronic Dutch book establishes that anyone who learns A for certain and does not update by the given identity can likewise be subject to a collection of bets that individually seem fair but jointly give a guaranteed loss.

Dutch book arguments involve bets, and so presuppose a practice for settling bets. The standard practice is that a bet on a proposition P is *won* if P turns out true, and *lost* if P turns out false. A *conditional* bet on P is also won or lost depending on whether P is true or false, with the proviso that the bet is called off (premium refund) if the condition for the bet is not satisfied. The standard practice for settling bets is thus squarely grounded in the semantic properties of the propositions that are the objects of the bet.

McGee's Dutch book arguments deviate from the standard practice. He stipulates a number of non-semantic rules for settling bets on conditionals and shows that if these rules for settling bets are followed, one can be subjected to a Dutch book if one's subjective degrees of belief do not correspond to McGee-measures. The obvi-ous complaint with such an approach is that if one can set the rules for settling bets without regard to the truth or falsity of the proposition that the bet concerns, one can devise a Dutch book for just about any rationality constraint. So: why did McGee feel the need to write a whole new rulebook for settling bets on conditionals, why should we think that these rules reveal something interesting about probabilities of conditionals, and can they be grounded in purely semantic considerations?

Fair settlement conditions for a bet on P are ultimately codifications of how and to what extent various possible states of affairs count in epistemic favour or disfavour of P; they provide, in this respect, *proxies* for an epistemic account of what is to count as grounds—and the extent to which they are to count as grounds—for or against holding that P is true. The point of this section is not to establish a Dutch book argument that is more persuasive than McGee's. The point is rather to 'fill in the blanks'. McGee provides no explicit link between his settlement conditions for bets on conditional propositions (codifying epistemic constraints) and the semantics that he provides for them; but this can be done. It turns our that certain contentious

consequences of McGee's settlement conditions for bets—and so of the epistemic principles that they encode—become very difficult to avoid.

The focus here will be on two critical rules for settling bets that McGee offered without any obvious semantic justification (the labels are not McGee's):[8]

Bet-additivity When φ and ψ express ICL-incompatible propositions, the payoff from a fair standard bet on $\varphi \vee \psi$ should always be the sum of the payoffs from fair standard bets on φ and ψ.

Conjunction Cancellation A bet on $(A_1 \rightarrow B_1) \wedge \cdots \wedge (A_n \rightarrow B_n)$ is called off (premium refund), if all $A_1, \ldots, A_n$ are false.

The first rule— bet-additivity—imposes a structural property on settlement conditions. Ultimately, the rule guarantees that a measure representing coherent degrees of belief satisfies probabilistic additivity. This is plausible enough, but it is noteworthy precisely because it hard-wires additivity into the very practice of settling bets without any direct reference to the semantic content of the proposition one bets on (by contrast, establishing that additivity is a rationality constraint is one of the core targets of traditional Dutch book arguments, and derive from the semantic properties of the propositions one bets on).

The second rule invokes settlement for bets on a particular class of propositions: conjunctions of base conditionals. It too is noteworthy, not because it is inherently implausible, but because it makes no direct reference to the semantic status (true or false?) of the conjunction, instead stating the settlement condition for the conjunction on the basis of truth-values of sentences that partially compose the conjunction.

So why do we require special rules for settling bets on conditional propositions? Let us go back to McGee's semantic framework. The basic element is that of a selection function. But the structural properties of a selection function do not seem to mirror any particular feature of the world. As McGee puts it:

> Purely semantic considerations are, at best, only able to tell us which world is the actual world... we have neither the need nor the ability to pick out a particular selection function as the actual one... and we count a sentence as genuinely true if it is made true by all the selection functions that are not excluded. (p.518-19)

We are left with a supervaluational account of truth and falsity and so with semantic indeterminacy:

> P is *determinately true* at w iff for all f such that $w_f = w$, $f \in P$.
> P is *determinately false* at w iff for all f such that $w_f = w$, $f \notin P$.
> P has an *indeterminate truth value* at w iff it is neither determinately true nor determinately false at w.

So, for logically independent factual sentences A and B:

---

[8]McGee actually uses a weaker form of bet-additivity that only presupposes classical logic for the boolean connectives. The remaining settlement principles are: bets on sentences that are boolean equivalent should give equal pay-off, and a bet on $A \wedge (A \rightarrow B)$ should pay the same as a bet on $A \wedge B$. In McGee's setting the language in the Dutch book is restricted to compounds containing base conditionals only.

$A \rightarrow B$ is determinately true at w iff $A \wedge B$ is true at w,

$A \rightarrow B$ is determinately false at w iff $A \wedge \neg B$ is true at w, and

$A \rightarrow B$ has an indeterminate truth value at w if A is false at w.

This places McGee's analysis alongside a long tradition of taking conditionals to require a three-valued semantics (e.g. see [28] for an historical overview), but in this case we get there by a supervaluational route, and the three 'truth values' are modal notions.[9] One advantage of the supervaluational approach (of several) is that a logical truth like $A \rightarrow A$ is indeed determinately true at every world, while on a standard three-valued semantics it will lack truth value when A is false.

Semantic indeterminacy explains the need for non-standard settlement conditions. For the standard settlement conditions for bets do not cover the possibility that a proposition may lack a determinate truth value. If, as McGee suggests, a conditional proposition can lack a determinate truth value, we need to fill this gap.

Let us fix some terminology. A bet is characterised by its cost, and its possible pay-offs. Here a bet is represented as a pair (c, p) where c is the cost of the bet, and p is a real-valued function from worlds to pay-offs. The net worth of a bet will thus be $p(w) - c$, if the actual world is w. A *standard* bet pays in the interval \$0 − \$1 where logical truths pay \$1 and logical falsities \$0 (standard bets are logically normalized). It is assumed that the agent's credence in φ can be represented by some real valued measure Pr, and that a standard bet on φ is *fair* for the agent if the agent is indifferent between buying and selling the bet at the price \$Pr(φ). Betting preferences are *coherent* if (1) there exists a fair standard bet for every sentence, and (2) the cost of a fair standard bet is also the agent's factual expectation value for the bet; so if $A_1, \ldots, A_n$ are factual sentences that form a logical partition, and if for all $w, w' \in F(A_i)$, $p(w) = p(w')$ (allowing us to speak of $p(A)$, the pay-off in case A is true), then $c = \sum \Pr(A_i) p(A)$. A measure that can give rise to coherent betting preferences will likewise be said to be coherent.

The *minimal requirements* on a standard bet (c, p) on φ is that $p(w) = \$1$ if φ is determinately true in w, and that $p(w) = \$0$ if φ is determinately false in w. This much is standard.

So, what to do with a bet on φ in the event that it turns out to have an indeterminate truth value? The following condition (the Strong Cancellation condition), modelled on the idea of a conditional bet, suggests itself:

SC A bet on φ is *called off*, (the premium Pr(φ) is repaid), if φ has an indeterminate truth value.

This gives us rRT (in the sense that any coherent measure that gives rise to betting preferences satisfying SC will satisfy rRT), which is encouraging. However McGee observes that the strong cancellation condition will not work when the goal is to

---

[9] The claim that the truth value of a proposition is indeterminate need not be understood as the claim that it has *no* truth value. For on at least one interpretation of a supervaluational semantic framework like the present—an interpretation that McGee [27] has promoted in another context—it is still perfectly legitimate to say that the proposition has a truth value (it is either true or false), it is just that the facts fail to determine whether it is true or whether it is false.

measure credences in compound conditional sentences. For a sentence like $(A \rightarrow B) \wedge (\neg A \rightarrow B)$ cannot be determinately true (if B is not a logical truth), but it can be determinately false, so a bet on the conjunction cannot be won but can be lost, and so the only rational price to pay for the bet would be \$0. Yet it is not hard to find examples where a rational agent could assign such a sentence a non-zero probability.

If we are to take seriously the contention that conditionals allow for indeterminate truth values we must allow that agents take the space of possibilities where a proposition has indeterminate truth value to have a more finely grained epistemic structure than the strong cancellation condition can capture. Moreover, this more finely grained epistemic structure should be grounded in a more finely grained semantic structure.

McGee points us in this direction; for, intuitively, there seems to be additional semantic structure that could and perhaps should have epistemic significance. To illustrate the idea, consider the pair:

(1) $(A \rightarrow B) \wedge (\neg A \rightarrow C)$.
(2) $(\neg A \rightarrow B) \wedge (\neg A \rightarrow C)$.

Assuming that all sentences involved are factual and logically independent, both conjunctions will lack a determinate truth value in the event that $A \wedge B$ is true; for then their second conjunct $\neg A \rightarrow C$ will lack a determinate truth value. But there is a difference between (1) and (2). For in (1) the first conjunct is determinately true when $A \wedge B$ is true, while in (2) *both* conjuncts have indeterminate truth value. We might say, somewhat suggestively (these are not McGee's labels), that when $A \wedge B$ is true, (1) is 'partially true' while (2) is 'strongly indeterminate'.

This appears to be borne out by intuition. Consider an example from [24] concerning the toss of a die:

If it's above three it will be a six, and if it's below three it will be a one.

On the present semantics the conjunction is determinately false if the die shows a two, four or seven. It cannot be determinately true, however, if the die shows a one or a six it will be 'partially' true (as one conjunct is true and the other has indeterminate truth value), while if it shows a three it will be *strongly* indeterminate (both conjuncts have indeterminate truth value). This matches the results reported by McDermott where it seems that people, when given the choice between 'true', 'false' or 'neither', tend to view the conjunction as true if the die shows a one or a six, but to lack truth value if it shows a three.

The idea is that we should limit the strong cancellation condition to apply only to the strongly indeterminate case; for only in such cases is there nothing that points either to the truth or the falsity of a proposition. Indeed, this is in effect what McGee's conjunction cancellation condition tells us: when all antecedents of a conjunction of base conditionals are false (and no conditional is a logical truth or falsity), the conjunction itself is strongly indeterminate and a bet on the conjunction should thereby be cancelled. (We will return to what should happen if it is 'partially' true.)

But the distinction between 'partially true' and 'strongly indeterminate' has been made by a mixture of semantic and syntactic criteria. To see the problem, say that a conjunction is 'half-true' when one conjunct is true and the other false, while a conjunction is 'fully false' if both conjuncts are false. Now say that A and B are both

false. Then $A \wedge B$ is 'fully false'. Meanwhile $A \wedge (\neg A \vee B)$ would be 'half-true' in virtue of the truth of $\neg A \vee B$. Yet the two conjunctions are logically equivalent. Thus 'half true' and 'fully false' are not properties of the proposition expressed by a conjunction, they draw on both semantic *and* syntactic properties. Why would we think that the distinction between conjunctions that are 'partially true' and 'strongly indeterminate' would fare any better? And if we cannot make the distinction on semantic grounds, where does that leave McGee's Conjunction Cancellation condition?

The semantic framework does not contain 'atomic' propositions in any language independent sense. Moreover, factual propositions—represented as sets of worlds—do not come in some 'hierarchy of complexity' that allow us to distinguish, say, 'disjunctive' from 'conjunctive' propositions. Conditional propositions turn out to be different. They exhibit an internal structure that allow them to be placed in a hierarchy of complexity.

For any sets of worlds X and Y define the conditional operator $>$ as follows:

$$X > Y = \{f \; : \; f(X) \in Y \text{ or } f(X) \text{ is undefined}\}.$$

If P is a proposition such that there exist sets of worlds X and Y where $P = X > Y$, then P is a (positive) *simple* proposition, and its complement $-P$ is a (negative) *simple* proposition. A *simple compound* proposition is an intersection of simple propositions. Finally, a *complex compound* proposition is a union of simple compound propositions. Every proposition is a complex compound, but there are complex compounds that are not simple compounds, there are simple compounds that are not simple propositions and there are simple propositions that are not factual, giving us a hierarchy of propositional complexity. It is this additional structure that allows us to identify what 'part' of the content of a proposition that becomes indeterminate at a given world.

A *propositional decomposition* is a set $\mathcal{K}$ of sets K of simple (positive or negative) propositions P. The propositional content of a decomposition $\mathcal{K}$, $C(\mathcal{K})$, is given by:

$$C(\mathcal{K}) = \bigcup_{K \in \mathcal{K}} \bigcap K.$$

A propositional decomposition is a representation of a proposition reminiscent of the 'disjunctive normal form' (though note that a propositional decomposition is not a proposition, it is a set of sets of propositions).

For any proposition P there is a decomposition $\mathcal{K}$ such that $P = C(\mathcal{K})$.[10] However, a proposition always has more than one decomposition; so we cannot automatically take the properties of a particular decomposition to reflect on the proposition it decomposes. But certain properties of a decomposition are *decomposition invariant*: all decompositions of a proposition possess them, and so reveal an

---

[10]For any selection function f we have $\{f\} = \bigcap D_f \cap \bigcap U_f$, where

$$D_f \; = \; \{X > \{w\} \; : \; f(X) \text{ is defined and } f(X) = w\}.$$
$$U_f \; = \; \{X > \emptyset \; : \; f(X) \text{ is undefined}\}.$$

So, any selection function is given by the intersection of a set of positive simple propositions. Hence, any *set* of selection functions can be represented as a set of sets of simple propositions.

underlying structural property of the proposition. In this way they become a stepping stone in extracting the notion of the 'strongly indeterminate' content of a proposition with indeterminate truth value.

When K is a set of simple propositions define:

$$SIC_w(K) = \begin{cases} \mathcal{F}, \text{ if } \bigcap K \text{ is determinately true at } w. \\ \emptyset, \text{ if } \bigcap K \text{ is determinately false at } w. \\ \bigcap\{P \in K : P \text{ lacks determinate truth value at } w\}, \text{ otherwise.} \end{cases}$$

For a decomposition $\mathcal{K}$ define:

$$SIC_w(\mathcal{K}) = \bigcup_{K \in \mathcal{K}} SIC_w(K).$$

$SIC_w(\mathcal{K})$ is a proposition. It is supposed to capture the (s)trongly (i)ndeterminate (c)ontent of $C(\mathcal{K})$ in w if it in w has indeterminate truth value.

At first blush, the analysis seems to fit the bill. For instance, when $A \wedge B$ is true, the strongly indeterminate content of $(A \to B) \wedge (\neg A \to C)$ is expressed by $\neg A \to C$. This is captured by the above definition; for the content of $(A \to B) \wedge (\neg A \to C)$ can be decomposed by $\mathcal{K} = \{\{[A \to B], [\neg A \to C]\}\}$, and $SIC_w(\mathcal{K}) = [\neg A \to C]$ for any w in which $A \wedge B$ is true.

But our function SIC_w can only serve its intended purpose if it is decomposition invariant. Moreover, to check the reasonableness of the analysis we need to minimally ensure that the strongly indeterminate part of a proposition that lacks determinate truth value also lacks determinate truth value. Indeed, we get them both:

Theorem 5

1. If $C(\mathcal{K}) = C(\mathcal{K}')$, then $SIC_w(\mathcal{K}) = SIC_w(\mathcal{K}')$.
2. If $C(\mathcal{K})$ has indeterminate truth value at w, then so does $SIC_w(\mathcal{K})$.

So, we can without qualms define, for any proposition P:

$$SIC_w(P) = SIC_w(\mathcal{K}), \text{ for some decomposition } \mathcal{K} \text{ of } P.$$

Now we can define our target properties:

Definition 4

1. P is *strongly indeterminate* at w iff P has indeterminate truth value at w and $P = SIC_w(P)$.
2. P is *partially true* at w iff P has indeterminate truth value at w and $P \subset SIC_w(P)$.
3. P is *partially false* at w iff P has indeterminate truth value at w and $SIC_w(P) \subset P$.

So, for instance $(A \to B) \wedge (\neg A \to C)$ will be partially true when $A \wedge B$ is true as it entails its strongly indeterminate content $\neg A \to C$. Meanwhile, $(A \to$

$\neg B) \vee (\neg A \rightarrow C)$ will be partially false when $A \wedge B$ is true as it is entailed by its strongly indeterminate content $\neg A \rightarrow C$.

Note, however, that the trichotomy is not exhaustive, there are complex compound propositions that have an indeterminate truth value at w but are neither strongly indeterminate, partially true, nor partially false. For instance $\neg B \rightarrow C$ expresses the strongly indeterminate content of $\neg A \vee ((\neg B \rightarrow C) \wedge D)$ in the case that $A \wedge B \wedge D$ is true. But $\neg A \vee ((\neg B \rightarrow C) \wedge D)$ neither entails nor is entailed by $\neg B \rightarrow C$. Propositions that are neither partially true, partially false nor strongly indeterminate will be said to be *semantically neutral*.

Having established the semantic credentials of the notion of a strongly indeterminate proposition we can state the *weak* cancellation condition:

WC   A bet on φ is called off (premium refund) if φ is strongly indeterminate.

In combination with the other requirements the weak cancellation condition turns out to be equivalent to McGee's conjunction cancellation condition.

So let us turn to the remaining cases: the pay-off for a bet on φ if it turns out to have indeterminate truth value without being strongly indeterminate. It is not easy to see how this case should be handled. All we know is that we cannot simply cancel such bets; for then we in effect have the strong cancellation condition. However, McGee shows that there is a sense in which we don't *need* to explicitly state the settlement conditions for this case. This is where his assumption of bet-additivity comes in: combined with the other conditions, it will fill in the blanks. And bet-additivity has its own appeal: its epistemic counterpart is the assumption that rational credences should be probabilistically additive. On this basis McGee secures a Dutch book argument for the claim that rational credences should be representable as McGee-measures. We get a corresponding result with the present settlement conditions.

Theorem 6 *The following claims are equivalent:*

1. *Pr is a McGee-measure.*
2. *Pr is coherent and the class of fair standard bets for Pr satisfies the minimal requirements, the weak cancellation condition and bet-additivity.*

For instance, as McGee shows, the settlement conditions jointly imply that a fair standard bet on $(A \rightarrow B) \wedge (C \rightarrow D)$ should pay:

1. \$1 if $A \wedge B \wedge C \wedge D$.
2. \$0 if $(A \wedge \neg B) \vee (C \wedge \neg D)$.
3. $\$Pr(A \rightarrow B \wedge C \rightarrow D)$ (the bet is cancelled, premium repaid), if $\neg A \wedge \neg C$.
4. $\$Pr(C \rightarrow D)$, if $\neg C \wedge A \wedge B$.
5. $\$Pr(A \rightarrow B)$, if $\neg A \wedge C \wedge D$.

We recognize the minimal requirements in (1) and (2) and the weak cancellation condition in (3). It is (4) and (5) that are special; for they do not give the same pay-off as as a clear win, a clear loss, or a cancelled bet: they are *partial* compensations. These 'partial' payoffs are derived from the other conditions together with bet-additivity.

The fact that the settlement conditions give rise to such partial compensations is arguably the most contentious implication of McGee's settlement conditions.

To get a feeling for what is going on, take the case when $\neg C \wedge A \wedge B$ is true. In this case $A \rightarrow B$ is determinately true and $C \rightarrow D$ is strongly indeterminate; so their conjunction is partially true. The pay-off for a bet on $(A \rightarrow B) \wedge (C \rightarrow D)$ will then be $\$\Pr(C \rightarrow D)$, where $C \rightarrow D$ is its strongly indeterminate content. Except in the extreme case where the conditionals have probability 0 or 1, this pay-off will be higher than a premium refund, but less than \$1: the pay-off will be proportionate to the probability of its strongly indeterminate content. A standard fair bet on a partially *true* proposition is thus *partially* won, and the pay-off is the probability of its strongly indeterminate content. (Correspondingly, a standard fair bet on a partially *false* proposition will be partially *lost*, the pay-off in non-extreme cases will be less than a premium refund but greater than \$0, and the pay-off is still the probability of its strongly indeterminate part.)

We can generalise this. Consider the *partial compensation* condition:

PC   The pay-off for a fair standard bet on φ at w is the probability of its strongly indeterminate content at w.

(Given the definition of SIC_w it is straightforward to show that for any sentence φ there are factual sentences $A_1^\varphi, \ldots, A_n^\varphi$ that form a logical partition, and sentences $\psi_1^\varphi, \ldots, \psi_n^\varphi$, such that $SIC_w([\varphi]) = [\psi_i^\varphi]$ if $w \in F(A_i^\varphi)$. So, the partial compensation condition states that a fair standard bet on φ will pay $\$Pr(\psi_i^\varphi)$ if A_i^φ is true.)

The partial compensation condition provides a single settlement condition for all bets. The conditions under which a bet is settled are stated purely in terms of the semantic properties of the sentence one bets on, and are stated without any side rules like bet-additivity or conjunction cancellation. Note, however, that the pay-off for a bet under a given condition is not fixed by factual or semantic properties alone, but by the credences of the agent.

Theorem 7 *The following claims are equivalent for any measure Pr:*

1. *Pr is a McGee-measure.*
2. *Pr is coherent and the class of fair standard bets for Pr satisfies the partial compensation condition.*

When we take away the betting context the partial compensation condition is ultimately an epistemic thesis about the evidential weight provided by the different possible ways in which a proposition can have an indeterminate truth value. The weight can be either positive or negative, but will not categorically establish the truth or falsity of the proposition. Now, in at least some domains where we (or some) are ready to speak of indeterminacy—e.g. in making claims about the outcome of a quantum experiment or about the position of a sub-atomic particle—we are still ready to differentiate those indeterminate propositions that have a high probability of being true from those that do not (all indeterminate propositions are not epistemically on par). The partial compensation condition makes sense if the indeterminacy

of conditional propositions follow suite and also allows for such differentiation: the possibility that a conditional proposition has indeterminate truth value can rationally count as evidence for or against the proposition.

The point is not that the partial compensation condition provides a more compelling Dutch book argument than the conditions given by McGee. To the contrary, once we make explicit the role of partial compensations, one would expect more (not less) worries about the plausibility of the argument. But we have unveiled the semantic underpinning for McGee's settlement conditions and the interplay between semantic and epistemic properties that they require.

Indeed, it turns out that the underlying problem of how to epistemically deal with what has here been diagnosed as semantic indeterminacy provides a significant hurdle for any account: the space of factual possibilities is not sufficiently fine-grained to yield, for each conditional proposition, a determinate true/false verdict, whether one is comfortable of talking about semantically indeterminate propositions or not. For say that one is sceptical of partial compensations in bets (as I expect many would be). This means that one must give up on one of McGee's settlement conditions: the minimal requirements, bet-additivity or the weak cancellation condition. The obvious strategy would be to reject the weak cancellation condition (rejecting the minimal requirements or bet-additivity would require us to drop the standard laws of probability). However, as it turns out, this will not solve our problems with partial compensation; for we still face a significant hurdle.

Settlement conditions can be said to be *finitely partition invariant* for some sentence φ if there is some finite partition $E_1, \ldots, E_n$ of worlds, such that for any agent, if (c, p) is a fair standard bet on φ for that agent, then $p(w) = p(w')$ whenever $w, w' \in E_i$ (for $1 \leq i \leq n$). Finite partition invariance for φ implies that there is an agent-independent fixed finite (though it can be arbitrarily large) set of factual possibilities that can be relevant when settling a bet on φ (the pay-off at each possibility can still be agent-dependent). A bet on a factual sentence trivially satisfies this (the partition consists of the set of worlds in which it is true and the set of worlds in which it is false); moreover, assuming the partial compensation condition, we get finite partition invariance for all sentences.

Observation 4 If the settlement conditions satisfy the minimal requirements, satisfy bet-additivity, yield rRT, do not constrain probabilities on factual sentences (more than requiring them to be coherent), and are such that any fair bet must give a pay-off of either \$0, \$1 or premium refund (no further 'partial compensations' allowed), then no sentence of the form $\neg A \wedge (A \to B)$ where $A \to B$ is logically contingent can have finitely partition invariant settlement conditions.

So even if we simply drop the weak cancellation condition there will be no straight-forward way of avoiding partial compensations. We can toy with the idea of rejecting finite partition invariance for a whole class of sentences (minimally: all sentences of the form $\neg A \wedge (A \to B)$), but it is hard to see how such an account would be developed, or what could count in its favour. Absent an implausibly fine-grained structure of factual considerations, one cannot account for rational probabilistic

attitudes towards complex conditional propositions by taking every factual possibility to be either irrelevant (premium refund) or as conclusive evidence for the truth or the falsity of the proposition. This is indirect evidence for the contention that conditionals do not express purely factual propositions. Importantly it would appear that—contrary to what one might expect—it is the combination of rRT and additivity that pushes us towards a commitment to partial compensations.

The most contentious implication of McGee's settlement conditions for bets—that pay-offs should allow for partial compensation—-comes from principles that many would otherwise accept. If one accepts them one also owes an explanation for the basis of the partial compensations.

Someone sceptical of McGee-measures but not of rRT or additivity, and who can stomach partial compensations, still has room for manoeuvre. For from Observation 2 we know that rRT and probabilistic additivity do not jointly imply GI; so there should be settlement conditions that allow for different pay-offs than those dictated by McGee's analysis. These can be had only by dropping or weakening the weak cancellation condition. For instance, one can replace it with the *very* weak cancellation condition:

VWC   A bet on φ is cancelled, if φ expresses a *simple* proposition that is strongly indeterminate.

This will still give us rRT without giving GI.

But merely weakening a condition is not sufficient. Given only the very weak cancellation condition, bet-additivity will not allow us to derive settlement conditions for bets on other indeterminate propositions. One must thus find some other settlement conditions for these, and, as shown, they must allow for partial compensations that presumably should be motivated in some way.

But why—if we can stomach the idea of partial compensations—would we think that we need to weaken the weak cancellation condition?

6 Counterexamples!?

6.1 The Counterexamples

McGee's analysis has repeatedly been held to ascribe unintuitive—indeed *wrong*—probabilities to complex sentences, which arguably has seriously affected the reception of the analysis. A number of counterexamples have been put forward to show this. The counterexamples have been aimed at GI and so strike equally at Stalnaker and Jeffrey's analysis.

Edgington [5] comes with the perhaps first documented counterexample; tersely delivered as follows:

Take an ordinary fair coin. "If it's first tossed at t_0, it will land heads, and if it's first tossed at t_1, it will land heads" should get 1/2, not 1/4. (p. 202, notation changed)

With the intuition that a coin can only be tossed for the first time once, and as the coin is fair this first toss—whenever it occurs—will have probability 1/2 of landing heads. But a McGee-measure will assign the conjunction probability 1/4 (which will decrease further if we add more conjuncts, like "If it's first tossed at t_2, it will land heads").

Lance [20] develops a counterexample in more detail. He colourfully sketches a scenario where a werewolf tossed a coin whether or not to stalk the neighbourhood where Jones lived. If the werewolf stalked the neighbourhood, everyone that was outside died. There is an even chance that Jones went out (an event that is probabilistically independent of whether the werewolf stalked the neighbourhood) and, given that he went out, an even chance that he went out the front door and an even chance that he went out the back door.

The situation can be represented as follows:

	J. went out front	J. went out back	J. stayed in
Werewolf	J. killed (.125)	J. killed (.125)	J. lives (.25)
¬Werewolf	J. lives (.125)	J. lives (.125)	J. lives (.25)

With these assumptions in place Lance asks us to consider the probability of the conjunction:

(3) If Jones went out the front door, he was killed by the werewolf and if Jones went out the back door he was killed by the werewolf.

Given rRT the probability of each conjunct should be .5. As all the factual probabilities are given, we can calculate the probability that a McGee-measure will assign to their conjunction: .25. However, Lance argues, this is the wrong answer; the right answer is .5.

> This is due to the fact that each conditional comes down to the question of where the werewolf went tonight. If she went to our neighbourhood she killed anyone outside. That is, if the werewolf is in our neighbourhood, then if Jones went out the front, it is certain that he was killed and if he went out the back, it is certain that he was killed. Similarly, if the werewolf is somewhere else, it is certain that Jones was not killed upon going out either door. So if the werewolf is in our neighbourhood, both conditionals are true (or highly assertible). If not, both are false. (p.271)

Bradley [2] has proposed a similarly structured counterexample directed at a special case of GI that doesn't require a conjunction of conditionals. He sketches the following scenario:

> Suppose that we have before us a coin that is known to be biased but that we do not know whether it is biased in favour of heads or in favour of tails (it is either a two-headed or two-tailed coin, say).

Bradley supposes, furthermore, that whether the coin is tossed is independent of whether it is biased heads or tails. Let bh = coin is biased heads, bt = coin is biased

tail, to = coin is tossed, H = coin lands heads. Bradley suggests that we should have:

$$\Pr(bh \wedge (to \rightarrow H)) = \Pr(bh).$$

This makes intuitive sense. For, Bradley argues, we should have:

$$\Pr_{bt}(to \rightarrow H) = 0,$$

because "given that the coin is biased toward tails, it is certain that the coin would not have landed heads had it been tossed" (p.553). So (as the ratio rule for conditioning is assumed) $\Pr(bt \wedge (to \rightarrow H)) = 0$, and so $\Pr(\neg bt \wedge (to \rightarrow H)) = \Pr(bh \wedge (to \rightarrow H)) = \Pr(to \rightarrow H)$. Due to independence, $\Pr(to \rightarrow H) = \Pr(bh)$. But then $\Pr(bh \wedge (to \rightarrow H)) = \Pr(bh)$.

Assuming that $\Pr(bh) = .5$ we thus have $\Pr(bh \wedge (to \rightarrow H)) = .5$. But a McGee-measure will give us $\Pr(bh \wedge (to \rightarrow H)) = .25 \neq \Pr(bh)$. Hence we have a counterexample to McGee's analysis.

McDermott [24] offers two counterexamples directed at McGee's account. Here is the first (p.26, numbering changed)), where he is considering the toss of a fair die:

> The following example is one about which everyone's intuitions are clear:
>
> (4) If it's odd it will be below three, and if it's even it will be above three.
>
> This is true if the result is a six, four, or one, false otherwise; so its assertability is 0.5.

McGee-measures will, by contrast, assign (4) probability 2/9.

The second counterexample that McDermott considers (p.27) also concerns the toss of a die:

(5) If it's even it will be above three, and if it's odd it will be above three.

McDermott suggests that "Most people think that (5) is equivalent to 'It will be above three'", an equivalence supported by his own semantic account, but not by McGee's. For a McGee-measure will never have

$$\Pr((A \rightarrow B) \wedge (\neg A \rightarrow B)) = \Pr(B)$$

when A and B have non-extreme (0 or 1) probabilities.

6.2 Proving more than one cares for

All the counterexamples were explicitly aimed at GI. The authors are all quite explicit that they are not challenging rRT. However:

Observation 5 Let $\mathcal{L}$ be a normal conditional logic (that is, a Tarskian logic satisfying MP$^{\supset}$, CL, CN, CK, and LLE) that in addition satisfies:

rMP $\quad A, A \rightarrow B \vdash_{\mathcal{L}} B$.
rCS $\quad A, B \vdash_{\mathcal{L}} A \rightarrow B$.
Sα $\quad A \rightarrow (B \wedge C) \vdash_{\mathcal{L}} (A \wedge B) \rightarrow C$.

Let Pr be an additive $\mathcal{L}$-based measure, and A, B and C be factual sentences.

1. If $\mathrm{Pr}((A \to B) \land (C \to B)) = \mathrm{Pr}(C \to B)$, $\mathrm{Pr}(A \land \neg B) > 0$, and
    $\mathrm{Pr}(\neg A \land C \land B) > 0$, then Pr does not satisfy rRT.
2. If $\mathrm{Pr}(C \land (A \to B)) = \mathrm{Pr}(C)$, $\mathrm{Pr}(A \land \neg B) > 0$ and $\mathrm{Pr}(\neg A \land C) > 0$, then Pr
    does not satisfy rRT.
3. If $\mathrm{Pr}((A \to B) \land (\neg A \to B)) = \mathrm{Pr}(B)$, $0 < \mathrm{Pr}(A) < 1$, and $0 < \mathrm{Pr}(B) < 1$,
    then Pr does not satisfy rRT.

The premises of (1) are satisfied in Edgington's and Lance's counterexample, the premises of (2) are satisfied in Bradley's, and the premises of (3) are satisfied in McDermott's second example (I will return to his first example). In no case can rRT be satisfied. So the counterexamples that were nominally levelled at GI somewhat unexpectedly (and, I take it, unintentionally) turn out to function as counterexamples to rRT.

The results of Observation 5 are unexpected, and what makes them unexpected, I would suggest, is that the counterexamples allow for a rather subtly misleading *partial* fulfilment of rRT. For instance, in Lance's example any pair of the factual propositions *a werewolf is about*, *Jones walked out the front door*, *Jones walked out the back door*, and *Jones was killed* can be combined into conditionals like *If a werewolf is about then Jones walked out the front door* that satisfy rRT while maintaining the intuitions of the example. But when these 'core' factual propositions enter into various boolean combinations (e.g. disjunctions), thereby widening the space of factual propositions, this no longer holds. For instance, what happens to a conditional like *If it is either the case that a werewolf is about or Jones walked out the front door, then Jones was killed*? From the proof of the above observation one can see that it is not possible for all such 'peripheral' conditionals to satisfy rRT while maintaining Lance's intuitions concerning the example. The same for Edgington's, McDermott's and Bradley's counterexamples. The violations of rRT are 'hidden' in these peripheral propositions and, correspondingly, difficult to detect.

The conclusion can of course be resisted by rejecting one of the logical principles that are presupposed, or the standard laws of probability. Rejecting the logical principles, however, is not so easy; for they are rather weak, at least given a commitment to rRT.

First, all the logical principles only concern base conditionals. Nothing is assumed about conditionals with non-factual consequents.

Second, all logical principles invoked are ICL-valid. So the logical principles make full room for McGee-measures, and these satisfy rRT. So the logical principles are not inherently hostile to rRT.

Third, from Observation 2 we can see that ICL+rRT does not imply GI. As GI is not assumed, the counterexamples will be in conflict with rRT even in the absence of GI. So the logical principles do not beg the question with regard to significant alternatives to McGee's analysis.

Fourth, the logical principles invoked are jointly strictly weaker than ICL. Importantly, no use is made of the most contentious principles of ICL: Import-Export and Conditional Excluded Middle.

Fifth, all the logical principles have independent appeal. Normality (CK + CN + LLE) is more or less a pre-condition for any semantically well-behaved conditional, and the only criticism of rMP that I am aware of is that it is *too weak* (by only being a *restricted* form of modus ponens). So, assuming that the Tarskian principles of classical logic and the standard laws of probability theory are beyond reproach in this context, only Sα and rCS can reasonably be held under suspicion.

Consider Sα first. It is valid in both Stalnaker's basic system C2 as well as in Lewis' system VC (following the terminology of [29]). Moreover, (and perhaps more importantly) given rRT it is 'probabilistically valid', in the sense that for any probability measure satisfying rRT we will have:

$$\Pr(A \rightarrow (B \wedge C)) \leq \Pr((A \wedge B) \rightarrow C).$$

The relationship between such probabilistic validity and logic is important. Quite generally we expect general constraints on a probability measure to be *neutral* with regard to the space of possibilities allowed by the semantics. That is, we expect that $\Pr(\varphi) \leq \Pr(\psi)$ will hold for all probability measures satisfying the constraints only if φ semantically entails ψ. For if φ does not semantically entail ψ, it must be possible that φ is true but that ψ is false, and if this possibility is sufficiently likely, the probability of φ will exceed that of ψ. To insist that it is possible that φ is true and that ψ is false, but—as a general probabilistic constraint—that the probability of this possibility is sufficiently low to ensure that $\Pr(\varphi) \leq \Pr(\psi)$, would be to add a probabilistic constraint that is not neutral with regard to the possibilities allowed by the semantics. So to reject Sα is to reject rRT as a semantically neutral constraint.[11]

So rCS remains. It too is valid in both Stalnaker's C2 and Lewis' VC. Moreover, it is probabilistically valid given rRT; for $\Pr(A \wedge B) \leq \Pr(A \rightarrow B)$. So, again, defending rRT from the counterexamples by weakening the logic will just strike back at rRT.

---

[11]The dialectical situation is somewhat nebulous as there is, to my knowledge, no one who actively endorses rRT who actively rejects Sα. However, the multi-dimensional framework developed by [2] (the same paper where he offers his counterexample to McGee) at least opens for the possibility. Bradley seeks a framework in which one can have rRT and additivity while satisfying the probabilities of the counterexamples and suggests that this can be done by weakening McGee's GI to a principle of restricted independence of antecedents rIA: $\Pr(\neg A \wedge (A \rightarrow B)) = \Pr(\neg A) \times \Pr(A \rightarrow B)$ (this is our principle IA restricted to factual consequents). As Bradley notes rIA implies rRT. Bradley's full semantic framework, outlined in his concluding section, is essentially a selection function semantics like McGee's, where selection functions are encoded as sequences of worlds. It features Centering, but Bradley does not state (or discuss) any further constraints, like the constraint that if $f(X) \in Y \cap Z$, then $f(X \cap Y) \in Z$ (central to Stalnaker-style treatments of conditionals and which validates Sα). Without further constraints all conditionals with non-equivalent antecedents become logically independent of each other, thus a logical principle like Sα will not be validated. Indeed, as Bradley suggests, it is plausible that weakening GI to rIA will allow the intuitions of the counterexamples to stand, but—and there is nothing to suggest that Bradley was aware of this—*only in the absence of Sα and similar properties*. Bradley offers no discussion of the problems associated with giving up Sα and related logical properties, including the problem that in the absence of Sα a principle like rRT will not be neutral to the space of possibilities. So Bradley's framework opens for the possibility that we have rRT without Sα, but he does not in any way argue for the plausibility of this combination. (Let me emphasize that these comments are in no way intended as a criticism of Bradley or his paper—it would be absurd to demand discussion of a result that did not exist at the time.)

In one way or other there is a hard logical price to pay for maintaining that the intuitions driving the counterexamples are logically consistent with rRT, given the standard laws of probability. I shall proceed under the assumption that they are not. We might then take the counterexamples to strike at rRT. But there remains the possibility that we cannot uphold the standard laws of probability in this context. After all, in the Dutch book argument, we either had to introduce bet-additivity (which gives us probabilistic additivity) as an *assumption* or derive it from the rather contentious partial compensation condition.

6.3 Dropping Additivity?

It is notable that the intuitions suggested by the counterexamples are all vindicated if we treat *partially* true propositions as epistemically *on par* with *determinately* true propositions. That is, if we stop treating partially true propositions as 'only partially' true, and instead treat them as *true*. Let us say that a sentence is *quasi-true* at w when it is either partially or determinately true at w, and, similarly, *quasi-false* at w when it is either partially or determinately false w. (Note that quasi-truth and quasi-falsity are 'factual' properties of sentences: for any world w, a sentence is either quasi-true at w, quasi-false at w or neither quasi-true nor quasi-false at w). So for instance both $A \rightarrow B$ and $(A \rightarrow B) \wedge (\neg A \rightarrow C)$ will be quasi-true when $A \wedge B$ is true (provided that $\neg A \rightarrow C$ is not a logical contradiction), and both $A \rightarrow \neg B$ and $(A \rightarrow \neg B) \vee (\neg A \rightarrow C)$ will be quasi-false when $A \wedge B$ is true (provided that $\neg A \rightarrow C$ is not a logical truth).

A *flat* measure (so-called as it flattens out the distinction between partial and determinate truth) is then given by the identity:

$$\Pr(\varphi) = \frac{\text{The probability that } \varphi \text{ is quasi-true.}}{\text{The probability that } \varphi \text{ is either quasi-true or quasi-false.}}$$

(When the probability that φ is either quasi-true or quasi-false is 0, $\Pr(\varphi)$ can be set to 1.) This corresponds to the following settlement conditions for a standard bet on φ:

1. Pay \$1 if φ is quasi-true.
2. Pay \$0 if φ is quasi-false.
3. Pay \$$\Pr(\varphi)$ (premium refund) if φ is neither quasi-true nor quasi false (i.e. if φ is strongly indeterminate or semantically neutral).

This will give us rRT and vindicates the intuitions of the counterexamples. Edgington's "If the coin was first tossed at t_0, it landed heads and if it was first tossed at t_1 it landed heads" will be quasi-true or quasi false if the coin was first tossed at either t_0 or t_1. It cannot be determinately true, but it is partially true iff it landed heads at that time. So, with a flat measure $\Pr((F_{t_0} \rightarrow H) \wedge (F_{t_1} \rightarrow H)) = \Pr((F_{t_0} \wedge H) \vee (F_{t_1} \wedge H))/\Pr(F_{t_0} \vee F_{t_1}) = 0.5$ (given that $F_{t_0} \vee F_{t_1}$ is probabilistically independent of H and $\Pr(H) = .5$), just as Edgington would have it. Lance's "If Jones left by the front door he was killed and if he left by the back door he was killed" is quasi-true iff Jones left by some door and was killed (probability .25), and is (quasi-)false iff he left by some door and wasn't killed (.25), and so will have probability .5. Bradley's "The coin is biased heads and if it was tossed it landed heads" is

partially true if the coin is biased heads and wasn't tossed and determinately true if the coin was biased heads and tossed, and so will be quasi-true iff the coin was biased heads, and so the conjunction will have probability .5. McDermott's example with a fair die "If it's odd it will be below three, and if it's even it will be above three" will be partially true if the die shows a one, four or six, and will be determinately false otherwise; so it will have probability .5 just as McDermott intuit, and his suggestion that $(A \rightarrow B) \wedge (\neg A \rightarrow B)$ is 'equivalent' to B is also vindicated: given that neither conjunct is a logical contradiction, the former is partially true iff the latter is true, and so they will have the same probability.

Treating partial truth as epistemically on par with determinate truth has some clear benefits. It brings together two truth-like concepts and offers an analysis of, well, *truth* (if we are willing to remove the modifier 'quasi'). Unlike McGee-measures flat measures capture the intuitions in the counterexamples. Of course, we know the theoretical cost for this: with flat measures we lose probabilistic additivity.[12] But we also know—recall Observation 5 of the previous section—that this is the price we must pay anyway for honouring the intuitions of the counterexamples (while maintaining rRT). So the intuitions seem to go together with a loss of additivity.

However, the apparently semantic grounds for the intuitions evoked by the counterexamples do not seem very stable.

Consider McDermott's suggestion that $(A \rightarrow B) \wedge (\neg A \rightarrow B)$ is equivalent to (and so should have the same probability as) B. This makes some intuitive sense. But the intuitive support fades quickly once A and B become probabilistically dependent. For instance, say that Lucy is a good runner and that it is highly likely that she won yesterday's race if she entered into it. It is also highly likely that she did enter into the race. So it is quite likely that she won yesterday's race. Of course, if she didn't enter into the race she didn't win. This much is certain. So "If she didn't enter into the race, she won" has probability 0 (or very close to 0). But—by Mcdermott's reckoning—"If Lucy entered into the race she won the race *and* if she didn't enter into the race, she won the race" has the same rather high probability as "Lucy won the race", even though the second conjunct has probability 0 (or close to 0). McDermott, on considering a similar example, bites the bullet, implying that we should be prepared to accept a conjunction even though we reject one of its conjuncts, but it seems a hard sell.

Likewise, the reasoning behind the suggestion that "If it's odd it will be below three, and if it's even it will be above three" should be assigned probability .5 will lead to strange consequences. For the first conjunct only has probability 1/3, which is lower than the probability of the conjunction. On the same reasoning, the following claim about a thousand-sided die would also be assigned probability .5 "If it's odd it will be a 1, and if it's even it will be above 2" as it is partially true in 500 cases and determinately false in 500 cases. But the first conjunct has probability .002. While

---

[12] The loss of additivity for flat measures is easily shown. By additivity $\Pr(A \rightarrow B) = \Pr(A \wedge (A \rightarrow B)) + \Pr(\neg A \wedge (A \rightarrow B)) = $ (by flatness) $\Pr(A \wedge B) + \Pr(\neg A)$. This will not in general be compatible with rRT: $\Pr(A \rightarrow B) = \Pr(A \wedge B)/\Pr(A)$.

one no doubt can feel the pull of McDermott's reasoning—partial truth has a 'truish' feel to it—it leads to conclusions that are very hard to accept.[13]

Lance and Bradley, by contrast, explicitly assume an element of probabilistic independence in their counterexamples. Lance assumes that the werewolf's whereabouts is probabilistically independent of whether Jones leaves the house or not. Bradley assumes that whether the coin is tossed or not is probabilistically independent of whether it is biased heads or tails. Neither makes it explicit why the independence assumption is important. But it does seem to be.

For consider the following modification of Lance's example. In one way or other we acquire more (perhaps not entirely trustworthy) information about what took place. In particular, say that we acquire good (though inconclusive) evidence that Jones is still alive (he seems to have survived the night!). This doesn't exclude that there was a werewolf about, but it strongly suggests that even if Jones went out, he wasn't killed, and so if he went out there was no werewolf about. There is now a perfectly legitimate—indeed perhaps *dominant*—interpretation of "If Jones went out he was eaten" whereby it acquires a lower, even considerably lower, probability than both "There was a werewolf about" and "If Jones *had* gone out, he *would have been* eaten".

In the quote given earlier Lance states "So, if the werewolf is in our neighbourhood, both conditionals are true (or highly assertible). If not, both are false". But this claim fails to establish that the conditionals have the truth-conditions that Lance's counterexample builds on. For if Jones is alive—and one has evidence for it—then "If Jones went out, he was eaten" doesn't strike one as true, even if one also knows that there was a werewolf about; so Lance's argument only works as long as one has no additional evidence of whether Jones is alive or not.[14]

Notably, when we switch from indicative conditionals to subjunctive conditionals Lance's argument makes sense irrespective of considerations of probabilistic independence, or of what we know. Even if we knew for sure that Jones is alive we would

---

[13] In his paper McDermott presents results from an informal survey on how people probabilistically assess compounds of conditionals. The results indicate that in the cases when compounds are only partially truth evaluable (on the current analysis), there is a strong tendency for the subjects to differ in their evaluation, yielding different probabilistic judgments. On being instructed to treat propositions that (on the current analysis) are actually *partially* true or *partially* false as either *true*, *false*, or *neither*, subjects gave differing responses. McDermott's hypothesized that the explanation for this inter-subject inconsistency is that the English 'and' and 'or' are ambiguous when they conjoin conditionals. A possible alternative explanation for the inconsistency is that speakers have different epistemic strategies for dealing with partial truth/falsity. One strategy is to equate partial truth with truth, another to equate it with undetermined, and McGee presents a third strategy by treating partial truths as true-ish (or, more exactly, the case when a proposition would be partially true evidentially speaks to its truth *to some degree*), giving them a more positive weight than undetermined, but not as positive as a fully fledged determinate truth. We can thus, in principle, explain the inconsistent responses in terms of the present semantic framework.

[14] One could, of course, hold that the semantic content of the conditional depends on what the speaker or assessor knows or has evidence for, leading into the debate on contextualist and relativist semantic theories. But it would seem that Lance's position in the paper is neither contextualist nor relativist; according to him it is the factual/causal (non-agent dependent) circumstances that make the conditional true. And, indeed, one can sense a reading according to which it *is* true in this more objective sense: the causal sense that is discussed below.

still accept "If Jones *had* gone out the front door he would have been eaten" (given that we know that there was a werewolf about).

Bradley's example follows the same pattern. Say that one learns that the person who was to flip the biased coin was very reluctant to flip it if the coin was biased tails. This makes one downgrade the probability of "If it was tossed it landed tails", but it doesn't make one downgrade the probability that it was biased tails. Now the intuition that "The coin was biased tails and if it was tossed it landed tails" says no more than— or at least has the same probability as—"The coin was biased tails" becomes seriously strained; far more so than the intuition that "The coin was biased tails and if it *had been* tossed it *would have* landed tails" has the same probability as "The coin was biased tails" (indeed, Bradley uses a subjunctive conditional to underscore his point "given that the coin is biased toward tails, it is certain that the coin would not have landed heads had it been tossed" (p.553)).

The reliance on probabilistic independence I think speaks against taking the intuitions as grounded in the semantics of an epistemic reading of the conditional. A possible alternative explanation comes from Kaufmann's [11, 12] observation that in certain circumstances indicative conditionals evoke a natural 'causal' reading that differs from its normal epistemic reading (in his terminology, indicatives have 'local' and 'global' readings).[15] In our case this is borne out by the fact that in both Lance's and Bradley's examples the intuitions become more stable and do not rely on probabilistic independence once we use subjunctive conditionals (the latter typically have a 'causal' reading). (Indeed [12], on independent grounds, analyses Lance's example in this way.) What kind of factors that would trigger such an alternative reading is disputed. [13] offers an account, and notes that one of the challenges is to explain why the phenomenon is relatively unusual. Perhaps in Lance's and Bradley's examples the triggering mechanism could be that probabilistic independence imposes a very small 'semantic cost' for adopting a causal reading—our judgments of indicative and subjunctive conditionals tend to coincide when probabilistic dependencies stem from our knowledge of causal relations rather than our knowledge of particular facts—and this is buttressed by an intuition that the sentences involved at least seem 'sort of' true in the stated conditions. Whatever the explanation would be, both Kaufmann and Khoo agree that even in cases where the 'causal' reading is dominant, an alternative 'evidential' reading persists.

If there is a distinctly causal reading, this at least undermines the argument that, due to the counterexamples, we need to abandon additivity. It would seem more appropriate to grab the other horn of the dilemma and reject rRT. After all, there is no reason to think that rRT would hold *in general* (e.g. also for more 'peripheral' propositions) for conditionals that are given a causal reading. If we allow that there is also an evidential reading that in these cases is not dominant (and so somewhat 'counter-intuitive'), rRT is not touched by the now merely apparent counterexamples. Nor—one should add—would the counterexamples work against McGee's analysis (or against the Stalnaker-Jeffrey analysis for that matter).

---

[15]One should add that for *future tensed* indicatives, a causal reading is often the most dominant. The discussion here concerns present and past tensed conditionals where the epistemic reading is usually the more dominant.

Edgington's example breaks the pattern as her indicative cannot be given a causal reading. But even while conceding the pull of her intuition in the example,[16] one can note that the four indicative conjunctions:

$$(F_{t_0} \rightarrow H) \wedge (F_{t_1} \rightarrow H) \qquad (F_{t_0} \rightarrow T) \wedge (F_{t_1} \rightarrow T)$$
$$(F_{t_0} \rightarrow H) \wedge (F_{t_1} \rightarrow T) \qquad (F_{t_0} \rightarrow T) \wedge (F_{t_1} \rightarrow H)$$

are (a) probabilistically incompatible, (b) jointly exhaustive and (c) equiprobable.[17] This on its own (without any particular theory of assigning probabilities to conjunctions of conditionals) suggests that each should have probability .25 (in accordance with McGee's analysis). But if we honour the intuition in Edgington's example they should each have probability .5, and this strikes me as highly counter-intuitive. Indeed, when one considers a long conjunctive string of mixed conditionals like $(F_{t_0} \rightarrow H) \wedge (F_{t_1} \rightarrow T) \wedge (F_{t_2} \rightarrow T) \wedge (F_{t_3} \rightarrow H)$ (which, by Edgington's account should have probability .5), implying multiple claims about the first flip of the coin, one starts sensing a friction in the original intuition: how can such an elaborate claim about varying possible outcomes at different points in time–that is indefinitely extendable to claims about other points in time–have the same probability as a claim about a single point in time? So, while Edgington's example cannot be explained away by giving the conditional a causal reading, what we can say in defence of additivity (and by extension McGee's-analysis) is that we are dealing with a clash of intuitions, and not with a clear-cut counterexample.

7 Concluding Remarks

McGee's semantic analysis of indicative conditionals manages to bring together several ideas that have been recurrent in the literature. It allows that conditionals with false antecedents can 'lack' a determinate truth value. It takes the antecedent of a conditional to restrict the space of possibilities relative to which the consequent is evaluated (a restrictor semantics), and it models this formally by a well-known tech-

---

[16] A possible origin of Edgington's intuition is our tendency to treat $(A \vee B) \rightarrow C$ as equivalent to $(A \rightarrow C) \wedge (B \rightarrow C)$. Indeed, semantically they are 'quasi-equivalent' as $(A \vee B) \rightarrow C$ is quasi-true iff $(A \rightarrow C) \wedge (B \rightarrow C)$ is quasi-true (as long as neither $A \rightarrow C$ nor $B \rightarrow C$ is logically false), which means that flat measures will assign them the same probability. By any account $\Pr((F_{t_0} \vee F_{t_1}) \rightarrow H) = .5$; so by the equivalence we would have $\Pr((F_{t_0} \rightarrow H) \wedge (F_{t_1} \rightarrow H)) = .5$, just as Edgington suggests. But there seem to be clear examples where the semantic 'quasi-equivalence' does not translate into a probabilistic identity. For instance, one may have high confidence in "If it was either the butler or the gardener that did it, then it was the gardener that did it", even though one has a low confidence in "If it was the butler that did it, then it was the gardener that did it, and if it was the gardener that did it, then it was the gardener that did it", as one has very low confidence in the first conjunct (this presupposes that one thinks it unlikely that *both* did it).

[17] It is presupposed here that sentences that are equivalent by ordinary boolean logic have the same probability. In addition claim (a) presupposes that $\Pr(F_x \rightarrow (H \wedge T)) = \Pr((F_x \rightarrow H) \wedge (F_x \rightarrow T)) = 0$ (for $x = t_0$ or t_1), and that $\Pr(\varphi \wedge \psi) = 0$ if $\Pr(\varphi) = 0$, which means that the conjunction of any two of the sentences will have probability 0. Claim (b) presupposes that $\Pr(F_x \rightarrow (T \vee H)) = \Pr((F_x \rightarrow H) \vee (F_x \rightarrow T)) = 1$ and that $\Pr(\varphi \wedge \psi) = 1$ if $\Pr(\varphi) = \Pr(\psi) = 1$, which means that the disjunction of all four sentences will have probability 1. (c) is given by the assumption that the coin is fair.

nical device: Stalnaker-style selection functions. The logic that McGee draws on (which matches the semantics) also brings together logical properties that have been recurrent in the literature (which, of course, McGee himself has been central in shaping): notably, the validity of Import-Export, and the failure of unrestricted modus ponens. The exact epistemic significance of the selection functions invoked in the formal machinery (e.g. what do they correspond to, pre-theoretically?) is still perhaps poorly understood, but by the construction invoked here, one can see that they can be given a purely epistemic interpretation.

McGee's rather opaque principle of independence (herein called GI) turns out to be equivalent to a more familiar independence principle that features as a recurrent theme in the literature: the antecedent of a conditional is probabilistically independent of the conditional (herein called IA). If we allow for a semantically sophisticated analysis of conditional probabilities, the principle IA turns out to be the unrestricted Ramsey Test.

Assuming that credences are otherwise well behaved (satisfy additivity) the underlying semantic analysis in combination with rRT force upon us a particularly contentious feature of McGee's analysis: the possibility that a proposition has indeterminate truth value can serve as either 'weak' positive or 'weak' negative evidence for the proposition, and the corresponding idea that bets—when used as measurements of credences—should provide partial compensations in the event that these possibilities materialise. The problem doesn't go away by dismissing talk of semantic indeterminacy. For as long as we accept the Ramsey Test and additivity, our conditions for settling bets will still require partial compensations, and this needs somehow to be explained.

The counterexamples that have been levelled specifically at McGee's probabilistic analysis all fail to achieve their intended purpose. For we cannot accommodate them merely by giving up the principle that they were aimed at (GI or IA). Instead— given fairly weak logical assumptions—the only way of accommodating them is by giving up either probabilistic additivity or the Ramsey Test, also in its restricted form. However, one can make a good case, at least for some of the counterexamples, that instead of accommodating the intuitions behind the counterexamples as features of an epistemic interpretation of conditionals, we should view them as additional evidence that indicative conditionals in certain contexts allow for a second, distinctly causal, sometimes dominant interpretation.

It is striking that the most contentious features of McGee's analysis stem from combining rRT and additivity and not from the much criticized GI. And, indeed, not all the intuitions behind the counterexamples can be explained away by appeal to a distinctly causal (non-epistemic) interpretation of the conditionals involved: some of the intuitions stand in an apparently irreconcilable conflict with intuitions supported by additivity. For an adherent of additivity it is perhaps a bit embarrassing that in this context one needs to resort either to general theoretical considerations or a contentious epistemology of indeterminacy to justify additivity. But many of its implications (e.g. a conjunction cannot be almost certainly true when one of its conjuncts is almost certainly false) make strong intuitive sense on independent grounds also when conditionals are involved, and these intuitions conflict with the intuitions driving the counterexamples. So, there is a genuine tension present which is not just

driven by mere theoretical considerations. McGee's *semantic* framework allows us to identify semantic mechanisms that plausibly explain the conflicting intuitions (a semantically robust but epistemically confounding notion of 'partial' truth and falsity). McGee's *epistemological* framework provides a plausible—but perhaps not compelling—means of resolving these conflicts.

In brief, the analysis that McGee presented in his 1989 paper stands on firmer grounds—both philosophically and in its ability to line up with both widely shared intuitions and the empirical facts—than is often realised.

Appendix: Proof of Observations and Theorems

Proof of Observation 2: The proof uses notions from Section 4. Consider a model where W contains four worlds, labelled 1, 2, 3, and 4. Let m be some probability mass on W that assigns non-zero probabilities to each world, and $\hat{m}$ be the McGee-mass generated by m. Consider some $\epsilon > 0$ such that $\hat{m}(s) > \epsilon$ and $\hat{m}(s) + \epsilon \leq 1$ for all sequences s. Let $\hat{m}'$ be a probability mass on sequences that is just like $\hat{m}$ except for the following four reassignments:

$$\hat{m}'(\langle 1, 3, 2, 4 \rangle) = \hat{m}(\langle 1, 3, 2, 4 \rangle) + \epsilon,$$
$$\hat{m}'(\langle 1, 3, 4, 2 \rangle) = \hat{m}(\langle 1, 3, 4, 2 \rangle) - \epsilon,$$
$$\hat{m}'(\langle 3, 1, 2, 4 \rangle) = \hat{m}(\langle 3, 1, 2, 4 \rangle) - \epsilon,$$
$$\hat{m}'(\langle 3, 1, 4, 2 \rangle) = \hat{m}(\langle 3, 1, 4, 2 \rangle) + \epsilon.$$

The net change is 0, so $\sum_{f \in \mathcal{F}} \hat{m}'(f) = 1$.

Assume a language where **1**, **2**, **3** and **4** serve as factual sentences true at (and only at) the corresponding world. A sequence like **134** will be understood as the *disjunction* **1** $\vee$ **3** $\vee$ **4**.

The first reassignment has the effect of adding probability ϵ to the following base conditionals (and any base conditional entailed by them):

$$\mathbf{1234} \to \mathbf{1} \qquad \mathbf{234} \to \mathbf{3} \qquad \mathbf{24} \to \mathbf{2}$$

Combining all three in a conjunction we get a sentence that is true only at the sequence $\langle 1, 3, 2, 4 \rangle$. The second reassignment has the effect of subtracting ϵ from the probability of the base conditionals:

$$\mathbf{1234} \to \mathbf{1} \qquad \mathbf{234} \to \mathbf{3} \qquad \mathbf{24} \to \mathbf{4}$$

So, the net effect of the two first reassignments leaves the probability of $\mathbf{1234} \to \mathbf{1}$ and $\mathbf{234} \to \mathbf{3}$ unchanged, while the probability of $\mathbf{24} \to \mathbf{2}$ has been raised by ϵ and the probability of $\mathbf{24} \to \mathbf{4}$ has been lowered by ϵ.

The third reassignment has the effect of lowering the following by ϵ:

$$\mathbf{1234} \to \mathbf{3} \qquad \mathbf{124} \to \mathbf{1} \qquad \mathbf{24} \to \mathbf{2}$$

The fourth reassignment has the effect of adding ϵ to the following:

$$\mathbf{1234} \to \mathbf{3} \qquad \mathbf{124} \to \mathbf{1} \qquad \mathbf{24} \to \mathbf{4}$$

So, the net effect of the third and fourth reassignments is that the probability of **24** → **2** has been lowered by ϵ while the probability of **24** → **4** has been raised by ϵ. In effect: undoing the changes (as regards base conditionals) of the first two reassignments.

So, taken together the four reassignments will not affect the probability of any base conditional. Moreover, the new measure is regular. As $\text{Pr}_{\hat{m}}$ satisfies rRT, $\text{Pr}_{m'}$ will satisfy rRT. But the probability of certain *compounds* of base conditionals have changed. For instance, the probability of

$$3 \wedge \mathbf{124} \rightarrow \mathbf{1} \wedge \mathbf{24} \rightarrow \mathbf{2}$$

has changed; for this conjunction is true only at the sequence $\langle 3, 1, 2, 4 \rangle$ to which we have added probability ϵ.

Given that GI uniquely determines the probabilities of all complex sentences on the basis of the factual probabilities, and that our original measure satisfied GI, we can thus conclude that the new measure does not. But one can also see this directly. For the probability of the conjunction **124** → **1** $\wedge$ **24** → **2** is unchanged by the reassignments (it is true at $\langle 1, 3, 2, 4 \rangle$ to which we have added ϵ and at $\langle 3, 1, 2, 4 \rangle$ from which we have subtracted ϵ). By GI

$$\text{Pr}_{\hat{m}}(\mathbf{3} \wedge \mathbf{124} \rightarrow \mathbf{1} \wedge \mathbf{24} \rightarrow \mathbf{2}) = \text{Pr}_{\hat{m}}(\mathbf{3}) \times \text{Pr}_{\hat{m}}(\mathbf{124} \rightarrow \mathbf{1} \wedge \mathbf{24} \rightarrow \mathbf{2}).$$

But $\text{Pr}_{\hat{m}}(\mathbf{124} \rightarrow \mathbf{1} \wedge \mathbf{24} \rightarrow \mathbf{2}) = \text{Pr}_{\hat{m}'}(\mathbf{124} \rightarrow \mathbf{1} \wedge \mathbf{24} \rightarrow \mathbf{2})$ and $\text{Pr}_{\hat{m}}(\mathbf{3}) = \text{Pr}_{\hat{m}'}(\mathbf{3})$. Still $\text{Pr}_{\hat{m}}(\mathbf{3} \wedge \mathbf{124} \rightarrow \mathbf{1} \wedge \mathbf{24} \rightarrow \mathbf{2}) \neq \text{Pr}_{\hat{m}'}(\mathbf{3} \wedge \mathbf{124} \rightarrow \mathbf{1} \wedge \mathbf{24} \rightarrow \mathbf{2})$. So:

$$\text{Pr}_{\hat{m}'}(\mathbf{3} \wedge \mathbf{124} \rightarrow \mathbf{1} \wedge \mathbf{24} \rightarrow \mathbf{2}) \neq \text{Pr}_{\hat{m}'}(\mathbf{3}) \times \text{Pr}_{\hat{m}'}(\mathbf{124} \rightarrow \mathbf{1} \wedge \mathbf{24} \rightarrow \mathbf{2}).$$

Thus GI is violated. Our measure $\text{Pr}_{\hat{m}'}$ satisfies rRT but not GI.   □

Proof of Theorem 2: Take any regular McGee-measure Pr. Consider the following consequence of GI:

(A)

$$\text{Pr}\left(\neg A \wedge \bigwedge_{1 \leq i \leq n} ((A \wedge B_i) \rightarrow a_i)\right) = \text{Pr}(\neg A) \times \text{Pr}\left(\bigwedge_{1 \leq i \leq n} ((A \wedge B_i) \rightarrow a_i)\right).$$

Here each a_i is an atomic sentence or a negated atomic sentence. Note that as Pr is regular we do not need to assume that $\text{Pr}(A \wedge B_i) > 0$. For if $\text{Pr}(A \wedge B_i) = 0$, then by regularity $\text{Pr}((A \wedge B_i) \rightarrow a_i) = 1$ and so the i-th term could without any effect be removed from the conjunction (if $\text{Pr}(A \wedge B_i) = 0$ for all i then both sides of the identity are equal to $\text{Pr}(\neg A)$).

From (A) we can establish:

(B)

$$\text{Pr}\left(\neg A \wedge \bigwedge_{1 \leq i \leq n} ((A \wedge A_i) \rightarrow a_i) \wedge \bigwedge_{1 \leq i \leq m} \neg((A \wedge C_i) \rightarrow \bot)\right) =$$

$$\text{Pr}(\neg A) \times \text{Pr}\left(\bigwedge_{1 \leq i \leq n} ((A \wedge A_i) \rightarrow a_i) \wedge \bigwedge_{1 \leq i \leq m} \neg((A \wedge C_i) \rightarrow \bot)\right).$$

For given regularity $\Pr(\neg((A \wedge C_i) \to \bot))$ is either 0 or 1. To see this assume that:

$$\Pr(\neg((A \wedge C_i) \to \bot)) > 0.$$

Then $1 - \Pr((A \wedge C_i) \to \bot) > 0$ and so $1 > \Pr((A \wedge C_i) \to \bot)$, By regularity it follows that $\Pr(A \wedge C_i) > 0$ and so by rRT $\Pr((A \wedge C_i) \to \bot) = 0$. So $\Pr(\neg((A \wedge C_i) \to \bot)) = 1$.

From (B) it follows by basic probabilistic principles (i.e. if $\Pr(\neg A \wedge \psi) = \Pr(\neg A) \times \Pr(\psi)$, then $\Pr(A \wedge \psi) = \Pr(A) \times \Pr(\psi)$) that:

(C)
$$\Pr\left(A \wedge \bigwedge_{1 \leq i \leq n} ((A \wedge A_i) \to a_i) \wedge \bigwedge_{1 \leq i \leq m} \neg((A \wedge C_i) \to \bot)\right) =$$
$$\Pr(A) \times \Pr\left(\bigwedge_{1 \leq i \leq n} ((A \wedge A_i) \to a_i) \wedge \bigwedge_{1 \leq i \leq m} \neg((A \wedge C_i) \to \bot)\right).$$

One can show (see [3]) that any sentence of the form $A \to \psi$ is logically equivalent to a disjunction $D(A \to \psi) = \delta_1 \vee \cdots \vee \delta_n$ where the disjuncts are mutually logically incompatible and each δ_j has the form

$$\bigwedge_{1 \leq i \leq k_j} ((A \wedge B_i^j) \to a_i^j) \wedge \bigwedge_{1 \leq i \leq m_j} \neg((A \wedge C_i^j) \to \bot).$$

Together with claim (C) this ensures that for any sentence $A \to \psi$ (where ψ need not be factual):

$$\Pr(A \wedge (A \to \psi))$$
$$= \Pr(A \wedge D(A \to \psi)) = \Pr((A \wedge \delta_1) \vee \cdots \vee (A \wedge \delta_n)) = \sum_{1 \leq j \leq n} \Pr(A \wedge \delta_j)$$
$$= \sum_{1 \leq j \leq n} \Pr\left(A \wedge \bigwedge_{1 \leq i \leq k_j} ((A \wedge B_i^j) \to a_i^j) \wedge \bigwedge_{1 \leq i \leq m_j} \neg((A \wedge C_i^j) \to \bot)\right)$$
$$= \Pr(A) \times \sum_{1 \leq j \leq n} \Pr\left(\bigwedge_{1 \leq i \leq k_j} ((A \wedge B_i^j) \to a_i^j) \wedge \bigwedge_{1 \leq i \leq m_j} \neg((A \wedge C_i^j) \to \bot)\right) \text{ (by C)}$$
$$= \Pr(A) \times \sum_{1 \leq j \leq n} \Pr(\delta_j) = \Pr(A) \times \Pr(D(A \to \psi)) = \Pr(A) \times \Pr(A \to \psi).$$

And this is IA.

For the other direction assume that Pr is a regular ICL-measure satisfying IA. By IA (omitting the subscript $1 \leq j \leq n$ for $\bigvee$, and $1 \leq i \leq n$ for $\bigwedge$):

$$\Pr\left(\bigvee A_j \wedge \left(\left(\bigvee A_j\right) \to \bigwedge(A_i \to \psi_i)\right)\right) =$$
$$\Pr\left(\bigvee A_j\right) \times \Pr\left(\left(\bigvee A_j\right) \to \bigwedge(A_i \to \psi_i)\right).$$

Note (letting $\Leftrightarrow$ denote logical equivalence):

$$\left(\bigvee A_j\right) \to \bigwedge(A_i \to \psi_i) \ \Leftrightarrow \ \bigwedge\left(\left(\bigvee A_j\right) \to (A_i \to \psi_i)\right) \quad \text{(Normality)}$$
$$\Leftrightarrow \bigwedge\left(\left(\left(\bigvee A_j\right) \wedge A_i\right) \to \psi_i\right) \quad \text{(Import-Export)}$$
$$\Leftrightarrow \bigwedge(A_i \to \psi_i). \quad \text{(LLE)}$$

So:

$$\Pr\left(\bigvee A_j \wedge \left(\bigwedge(A_i \to \psi_i)\right)\right) = \Pr\left(\bigvee A_j\right) \times \Pr\left(\bigwedge(A_i \to \psi_i)\right).$$

And so:

$$\Pr\left(\bigwedge \neg A_i \wedge \bigwedge(A_i \to \psi_i)\right) = \Pr\left(\bigwedge \neg A_i\right) \times \Pr\left(\bigwedge(A_i \to \psi_i)\right).$$

This implies GI.   □

Proof of Theorem 3: Let m be a probability mass on worlds. Let s be some non-empty sequence with domain $D(s)$ of cardinality k. Define an intermediary notion:

$$m^*(s) = \prod_{1 \leq i \leq k} \frac{m(s[i])}{\sum_{i \leq j \leq k} m(s[j])}.$$

Let $s - w$ be the sequence we get when we remove the world w from s. Trivially (when $D(s)$ contains more than one world):

$$m^*(s) = \frac{m(s[1])}{m(D(s))} \times m^*(s - s[1]).$$

As before where P is a set of sequences let $m^*(P) = \sum_{s \in P} m^*(s)$, and let $P - w$ be the set of sequences we get when we remove w from each sequence in P (if $P = \{\langle w \rangle\}$, then $P - w = \emptyset$). Let $P[w]$ be the set of sequences in P that have w as its first element.

When P is a set of sequences, let P° be the set of sequences in P that have domain $D(m)$. Our target construction $\hat{m}$ is then related to m^* by the identity:

$$\hat{m}(P) = m^*(P^\circ).$$

(m^* differs from $\hat{m}$ in that it will assign a non-zero weight also to a sequence that has a strict subset of $D(m)$ as its domain, and as a result $m^*(\mathcal{F})$ can be greater than 1).

When $s = \langle w_1, \ldots, w_k \rangle$ is a sequence such that $w \notin D(s)$ then the sequence $\langle w_1, \ldots, w_i, w, w_{i+1}, \ldots w_k \rangle$, for any i such that $0 \leq i \leq k+1$ is a w-*variant* of s (we get a w-variant by inserting w somewhere in the sequence). Let $I(s, w)$ denote the set of w-variants of s. We can now prove the core lemma.

Lemma 1 If $m(w) > 0$ and $w \notin D(s)$, then $m^*(I(s, w)) = m^*(s)$.

Proof Assume that $m(w) > 0$. When $m^*(s) = 0$ it trivially follows that $m^*(I(s, w)) = 0$. So assume that $m^*(s) > 0$. Proof by induction over the cardinality of $D(s)$. For the base case the cardinality is 1, so $D(s) = \{w'\}$ for some w'. There are two w-variants of s: $\langle w, w' \rangle$ and $\langle w', w \rangle$. Trivially $m^*(s) = 1$. Moreover $m^*(\{\langle w, w' \rangle, \langle w', w \rangle\}) = m(w)/(m(w) + m(w')) + m(w')/(m(w) + m(w')) = 1$.

For the induction hypothesis assume that the claim holds for any s with domain of cardinality k. Consider some sequence $s = \langle w_1, \ldots, w_{k+1} \rangle$ that has a domain of cardinality $k + 1$. Assume that $w \notin D(s)$. Let $s' = \langle w_2, \ldots, w_{k+1} \rangle$. $D(s')$ is of cardinality k so the claim holds for s': $m^*(I(s', w)) = m^*(s')$. We can get nearly all the elements of $I(s, w)$ by concatenating w_1 as an initial world to each element s'' of $I(s', w)$, (in symbols $\langle w_1; s'' \rangle$). The only element of $I(s, w)$ that we will miss in this way will be the sequence in which w is the initial element, that is, the sequence $\langle w; s \rangle$.

So

(A) $m^*(I(s, w)) = m^*(\{\langle w_1; s'' \rangle : s'' \in I(s', w)\}) + m^*(\langle w; s \rangle)$.

Focus first on the first term on the right-hand side of (A). We have:

$$m^*(\{\langle w_1; s'' \rangle : s'' \in I(s', w)\}) = \frac{1}{p} \times m(w_1) \times m^*(I(s', w)),$$

where $p = m(D(s)) + m(w) =$ the probability of the domain of the sequences involved. By the induction hypothesis:

$$\frac{1}{p} \times m(w_1) \times m^*(I(s', w)) = \frac{1}{p} \times m(w_1) \times m^*(s').$$

Note that

$$m^*(s) = \frac{1}{p - m(w)} \times m(w_1) \times m^*(s')$$

(w is not in the domain of s, thereby the term $p - m(w)$). So $m(w_1) \times m^*(s') = (p - m(w)) \times m^*(s)$. So:

$$\frac{1}{p} \times m(w_1) \times m^*(s') = \frac{1}{p} \times (p - m(w)) \times m^*(s).$$

The right-hand side will be our new first term for (A).

Now take the second term on the right-hand side of (A). We have:

$$m^*(\langle w; s \rangle) = \frac{1}{p} m(w) \times m^*(s).$$

So, combining the two new right-hand terms for (A):

$$m^*(I(s, w)) = \frac{1}{p} \times (p - m(w)) \times m^*(s) + \frac{1}{p} \times m(w) \times m^*(s)$$

$$= \left(\frac{1}{p} \times (p - m(w)) + \frac{1}{p} \times m(w) \right) \times m^*(s).$$

As $\frac{1}{p} \times (p - m(w)) + \frac{1}{p} \times m(w) = \frac{1}{p}(p - m(w) + m(w)) = 1$:

$$m^*(I(s, w)) = m^*(s).$$ □

A set of sequences P is *closed under w-variants* if for any $s \in P$ such that $s \neq \langle w \rangle$: $I(s - w, w) \subseteq P$. Intuitively, a proposition is closed under w-variants if the placement of w in an ordering is irrelevant for the truth of the proposition.

Corollary 2 If P is closed under w-variants and every $s \in P$ has a domain of at least two elements, then $m^*(P) = m^*(P - w)$.

Proof Assume that P is closed under w-variants. As each sequence has a domain of at least two elements $P - w$ will be a set of sequences. As P is closed under w-variants: for each $s \in P - w$, $I(s, w) \subseteq P$. Indeed $\{I(s, w) : s \in P - w\}$ makes up a partition of P. So $m^*(P) = \sum_{s \in P-w} m^*(I(s, w))$. By Lemma 1, $m^*(s) = m^*(I(s, w))$ for each $s \in P - w$. So $m^*(P - w) = \sum_{s \in P-w} m^*(I(s, w))$. So $m^*(P) = m^*(P - w)$. □

Corollary 3

1. $\hat{m}(\mathcal{F}) = 1$. ($\hat{m}$ is a probability mass on sequences)
2. $\hat{m}([A]) = m(F(A))$. ($\hat{m}$ extends m)
3. $\hat{m}([\neg A \wedge (A \to \varphi)]) = \hat{m}([\neg A]) \times \hat{m}([A \to \varphi])$ ($\hat{m}$ satisfies IA).

Proof (1)   First prove that if $X \subseteq D(m)$ and P is the set of all sequences with domain X, then $m^*(P) = 1$. By induction over the cardinality of X. The case when X is a singleton $\{w\}$ is trivial. So assume that the claim holds for cardinality k. Show that it holds for X of cardinality $k + 1$. Let P be the set of all sequences with domain X. Take any $w \in X$. Note that $P - w$ is the set of all sequences with domain $X - \{w\}$. By the induction hypothesis $m^*(P - w) = 1$. By Corollary 2, $m^*(P - w) = m^*(P)$. So $m^*(P) = 1$.

$\mathcal{F}$ is the set of all sequences and $\mathcal{F}^\circ$ is the set of all sequences with domain $D(m)$. So $m^*(\mathcal{F}^\circ) = \hat{m}(\mathcal{F}) = 1$.

(2)   Take any $w \in F(A)$. If $w \notin D(m)$ then $[A][w]^\circ$ is empty and $m^*([A][w]^\circ) = m(w) = 0$ (where, recall, $P[w]$ is the set of sequences of P that begin with w). So assume that $m(w) > 0$. Note that $[A][w]^\circ - w$ is the set of all sequences that have domain $D(m) - \{w\}$. By construction $m^*([A][w]^\circ) = (m(w)/m(D(m))) \times m^*([A][w] - w) = m(w) \times m^*([A][w]^\circ - w)$ (recall: $m(D(m)) = 1$). But by (1) of this Corollary, $m^*([A][w]^\circ - w) = 1$. So $m^*([A][w]^\circ) = m(w)$. This holds for all $w \in F(A)$ so $m^*([A]^\circ) = \hat{m}([A]) = m(F(A))$.

(3)   Note that $[A \to \varphi]$ is closed under w-variants for each $w \in F(\neg A)$. Consider $P = [A \to \varphi]^\circ$. P has domain $D(m)$ and will be closed under w-variants for each $w \in F(\neg A)$ such that $m(w) > 0$.

First show that $m^*(P[w]) = m(w) \times m^*(P)$ for each $w \in F(\neg A)$. Assume that $w \in F(\neg A)$. If $D(m) = \{w\}$ or if $m(w) = 0$ the claim follows immediately. So

assume that $D(m) \neq \{w\}$ and that $m(w) > 0$. By construction $m^*(P[w]) = m(w) \times m^*(P - w)$. By Corollary 2 $m^*(P) = m^*(P - w)$, so $m^*(P[w])) = m(w) \times m^*(P)$.

This holds for all $w \in F(\neg A)$ and so $m^*([\neg A] \cap P) = m(F(\neg A)) \times m^*(P)$.

By construction $m^*(P) = \hat{m}([A \to \varphi])$ and $m^*([\neg A] \cap P) = \hat{m}([\neg A \wedge (A \to \varphi)])$. Moreover, $m(F(\neg A)) = \hat{m}([\neg A])$. So:

$$\hat{m}([\neg A \wedge (A \to \varphi)]) = \hat{m}([\neg A]) \times \hat{m}([A \to \varphi]). \qquad \square$$

It follows immediately that $\Pr_{\hat{m}}$ is a regular McGee-measure. Uniqueness follows from McGee's results. $\qquad \square$

Proof of Theorem 4: Assume a measure m such that $m(X) > 0$. Assume a language where each set of worlds X has a corresponding atomic sentence $\mathbf{X}$ such that $F(\mathbf{X}) = X$. Given such a language one can for every sequence s construct a sentence $\mathbf{s}$ such that $[\mathbf{s}] = s$. Let $\Pr_X$ be the measure $\Pr_X(\varphi) = \Pr_{\hat{m}}(\mathbf{X} \to \varphi)$. As McGee demonstrated $\Pr_X$ thus defined is a McGee-measure. By rRT we know that for any factual A, $\Pr_X(A) = \Pr_{\hat{m}}(A \wedge \mathbf{X})/\Pr_{\hat{m}}(\mathbf{X})$. So the probability mass on worlds that would generate $\Pr_X$ is m_X. So $\Pr_X = \Pr_{\hat{m}_X}$ and so $\Pr_X(\varphi) = \hat{m}_X([\varphi])$.

Take any sequence s. We know that, for any s' such that $D(s') \cap X \neq \emptyset$, $s' \in [\mathbf{X} \to \mathbf{s}]$ iff $s'/X = s$. So

$$\hat{m}([\mathbf{X} \to \mathbf{s}]) = \sum_{s' \in [\mathbf{X} \to \mathbf{s}]} \hat{m}(s') = \sum_{s'/X=s} \hat{m}(s').$$

$\Pr_X(\mathbf{s}) = \hat{m}_X(s)$. Moreover, $\Pr_X(\mathbf{s}) = \hat{m}([\mathbf{X} \to \mathbf{s}])$, so $\hat{m}_X(s) = \sum_{s'/X=s} \hat{m}(s'). \qquad \square$

Proof of Theorem 5: (1) First a useful definition and an important property. Where f is a selection function and w is a world define f^w as follows (for any set of worlds X):

$$f^w(X) = \begin{cases} f(X), & \text{if } w \notin X \text{ and } D(f) \cap X \neq \emptyset, \\ w, & \text{if } w \in X, \\ \text{undefined, otherwise.} \end{cases}$$

f^w is 'just like' f except that it has w as its preferred world. Note that, trivially, if P is determinately true at w, then $f^w \in P$, and if P is determinately false at w then $f^w \notin P$.

Lemma 2 If P is a simple proposition that has an indeterminate truth value at w, then $f^w \in P$ iff $f \in P$.

Proof Let P be a simple proposition that has an indeterminate truth value at w. Take any f. Consider the case when P is positive, so $P = X > Y$ for some X and Y. If $w \in X$, then, contrary to assumption, $X > Y$ has a determinate truth value at w (it is determinately true at w if $w \in Y$ and determinately false at w if $w \notin Y$). So $w \notin X$. Note that as $D(f) - \{w\} = D(f^w) - \{w\}$, $D(f) \cap X = \emptyset$ iff $D(f^w) \cap X = \emptyset$ (as $w \notin X$). So if $D(f) \cap X = \emptyset$, then $f \in X > Y$ iff $f^w \in X > Y$. So assume that

$D(f) \cap X \neq \emptyset$. As $w \notin X$, by construction, $f^w(X) = f(X)$. So $f^w(X) \in Y$ iff $f(X) \in Y$, and so $f^w \in X > Y$ iff $f \in X > Y$.

Consider the case when P is negative, so $P = -(X > Y)$ for some X and Y. If $-(X > Y)$ has an indeterminate truth value at w, then so does $X > Y (= -P)$. By the above $f^w \in -P$ iff $f \in -P$, so $f^w \in P$ iff $f \in P$. □

Now, assume that $C(\mathcal{K}) = C(\mathcal{K}')$. Assume (i) that $f \in SIC_w(\mathcal{K})$ but for reductio (ii) that $f \notin SIC_w(\mathcal{K}')$.

From (ii) it follows that for every $K' \in \mathcal{K}'$ we have $f \notin SIC_w(K')$. Take any $K' \in \mathcal{K}'$. $\bigcap K'$ cannot be determinately true at w, for then $SIC_w(K') = \mathcal{F}$ and so would contain f contrary to assumption. So we have two alternatives. Either $\bigcap K'$ is determinately false in w, and then, trivially, $f^w \notin \bigcap K'$. Or $\bigcap K'$ has indeterminate truth value at w. Then there is some simple proposition $P \in K'$ such that P has indeterminate truth value at w and $f \notin P$. By Lemma 2, $f^w \notin P$. Accordingly, $f^w \notin \bigcap K'$. In either case: $f^w \notin \bigcap K'$. This holds for all $K' \in \mathcal{K}'$, so $f^w \notin C(\mathcal{K}')$.

We know from (i) that for some $K \in \mathcal{K}$, $f \in SIC_w(K)$. It follows that $\bigcap K$ is not determinately false at w and so that no $P \in K$ is determinately false in w. So take any $P \in K$. Either P has indeterminate truth value at w, in which case $f \in P$. By Lemma 2, $f^w \in P$. Or P is determinately true at w. Then, trivially, $f^w \in P$. In either case $f^w \in P$. This holds for all $P \in K$, so $f^w \in \bigcap K$. But then $f^w \in C(\mathcal{K})$, but then $C(\mathcal{K}) \neq C(\mathcal{K}')$ and we have a contradiction.

(2) Suppose that $C(\mathcal{K})$ has indeterminate truth value at w. So there are selection functions f and f' such that (i) $w_f = w_{f'} = w$, (ii) $f \in C(\mathcal{K})$ and (iii) $f' \notin C(\mathcal{K})$.

By (ii) there is some $K \in \mathcal{K}$ such that $f \in \bigcap K$. So $\bigcap K$ is not determinately false at w (and $\bigcap K$ can't be determinately true at w for then $C(\mathcal{K})$ would be determinately true at w). As $\bigcap K$ has indeterminate truth value at w and $f \in \bigcap K$, $f \in SIC_w(K)$ and so $f \in SIC_w(\mathcal{K})$. So $SIC_w(\mathcal{K})$ is not determinately false at w.

By (iii) for all $K \in \mathcal{K}$, $f' \notin \bigcap K$. Suppose, for reductio, that there is some $K \in \mathcal{K}$ such that $f' \in SIC_w(K)$. It follows that $\bigcap K$ is not determinately false at w. So $\bigcap K = Q \cap SIC_w(K)$, where Q is some proposition that is determinately true at w. But as $w_{f'} = w$, Q contains f'. So then $f' \in \bigcap K$, contrary to assumption. So $f' \notin SIC_w(K)$. This holds for all $K \in \mathcal{K}$ so $f' \notin SIC_w(\mathcal{K})$ and so $SIC_w(\mathcal{K})$ is not determinately true at w. □

Proof of Theorem 6: 2-to-1. Assume that Pr is coherent and that the class of fair bets for Pr satisfies the minimal requirements, bet-additivity and the weak cancellation condition. By the minimal requirements $\Pr(\varphi) = 1$, for any ICL-validity φ (as any ICL-valid sentence φ is always determinately true and a bet on φ will give pay-off \$1). Likewise, by bet-additivity $\Pr(\varphi \vee \psi) = \Pr(\varphi) + \Pr(\psi)$, when $\neg(\varphi \wedge \psi)$ is ICL-valid.

So Pr is an ICL-measure.

Consider a sentence $\varphi = A_1 \rightarrow B_1 \wedge \cdots \wedge A_n \rightarrow B_n$. Assume that no $A_i \rightarrow B_i$ is a logical truth or a logical contradiction (logical contradictions will have probability 0; if there are one or more logical truths in a conjunction they can be removed while

retaining the same probability). Consider a sentence C that is incompatible with each A_i. Take any world w in which C is true. Take a fair bet $b_1 = (\$\Pr(\varphi), p_1)$ on φ. We will have $p_1(w) = \$\Pr(\varphi)$ due to the weak cancellation condition. Take fair bets $b = (\$\Pr(C \wedge \varphi), p_2)$ and $b_3 = (\$\Pr(\neg C \wedge \varphi), p_3)$ on $C \wedge \varphi$ and $\neg C \wedge \varphi$ respectively. Due to bet additivity we have $p_1(w) = p_2(w) + p_3(w)$. That is, $\$\Pr(\varphi) = p_2(w) + p_3(w)$. But $\neg C \wedge \varphi$ is determinately false at w so $p_3(w) = 0$. So $p_2(w) = \$\Pr(\varphi)$. This will hold for all w where C is true. So the expectation value of b_2 is $\$\Pr(C) \times \Pr(\varphi)$. But the cost of b_2 is $\$\Pr(C \wedge \varphi)$. So $\Pr(C \wedge \varphi) = \Pr(C) \times \Pr(\varphi)$.

So we have GI.

(1-to-2) As McGee has provided a proof based on closely related assumptions only the finite case will be treated and only for regular measures. So assume, for simplicity, a finite model, and to avoid technicalities assume that the McGee-measure is defined on propositions, that propositions are sets of sequences, and that bets are made on propositions. Furthermore, let $D(\Pr)$ be the set of worlds w such that $\Pr(\mathcal{P}(w)) > 0$.

Where P is a proposition, let p_P be the pay-off function:

$$p_P(w) = \Pr(SIC_w(P)).$$

(This is the Partial Compensation condition!) Let $b_P = (\$\Pr(P), p_P)$. We need to show that this is a fair standard bet on P.

Take any $s = \langle w_1, \ldots, w_n \rangle$. We then have the unit proposition $\{s\}$. Let (recall: $>$ is our conditional operator on sets of worlds)

$$P_s = (\{w_2, \ldots, w_n\} > \{w_2\}) \cap \cdots \cap (\{w_{n-1}, w_n\} > \{w_{n-1}\}) \cap (\{w_n\} > \{w_n\}).$$
$$Q_s = \bigcap_{w \notin D(s)} (\{w\} > \emptyset).$$
$$R_s = -(\{w_2\} > \emptyset) \cap \cdots \cap -(\{w_n\} > \emptyset).$$

$\{s\} = \mathcal{P}(w_1) \cap P_s \cap Q_s \cap R_s$. Note that, due to Regularity, $\Pr(Q_s) = 0$ if there is some $w \in D(\Pr)$ such that $w \notin D(s)$. Meanwhile if there is some $w \in D(s)$ such that $w \notin D(\Pr)$, then $\Pr(R_s) = 0$. So if $D(\Pr) \neq D(s)$, then $\Pr(\{s\}) = 0$. Meanwhile, if $D(s) = D(\Pr)$, then $\Pr(Q_s) = \Pr(R_s) = 1$. So if $D(s) = D(\Pr)$, then $\Pr(\{s\}) = \Pr(\mathcal{P}(w_1) \cap P_s) = $ (by GI) $\Pr(\mathcal{P}(w_1)) \times \Pr(P_s)$. (which is also the cost of the bet $b_{\{s\}}$).

Note that $P_s \cap Q_s \cap R_s$ is the strongly indeterminate content of $\{s\}$ at w_1 (our $SIC_{w_1}(\{s\})$), while at all other worlds w, $SIC_w(\{s\}) = \emptyset$. So the expected pay-off for $b_{\{s\}}$ (when $\Pr(\{s\}) > 0$) is $\sum_{w \in W} \Pr(\mathcal{P}(w)) \times \Pr(SIC_w(\{s\}) = \Pr(\mathcal{P}(w_1)) \times \Pr(P_s) = \Pr(\{s\})$ which is also the cost of the bet and so $b_{\{s\}}$ is a fair standard bet.

This holds for unit propositions. We need to show that it holds for all propositions.

Lemma 3 For any propositional decompositions $\mathcal{K}$ and $\mathcal{K}'$, if $C(\mathcal{K}) \cap C(\mathcal{K}') = \emptyset$, then

1. $SIC_w(\mathcal{K}) \cap SIC_w(\mathcal{K}') = \emptyset$.
2. $SIC_w(\mathcal{K}) \cup SIC_w(\mathcal{K}') = SIC_w(\mathcal{K} \cup \mathcal{K})$.

Proof (1)   Assume that there is a $K \in \mathcal{K}$ and a $K' \in \mathcal{K}'$ and some f such that $f \in SIC_w(K) \cap SIC_w(K')$. It follows that no $P \in K \cup K'$ is determinately false in w. But then by Lemma 2 $f^w \in \bigcap K \cap \bigcap K'$. But then $f^w \in C(\mathcal{K}) \cap C(\mathcal{K}')$ which gives us a contradiction.

(2)   Follows trivially from the construction.   □

So, as the strongly indeterminate content is decomposition invariant, if $P \cap Q = \emptyset$, then $SIC_w(P) \cap SIC_w(Q) = \emptyset$ and $SIC_w(P \cup Q) = SIC_w(P) \cup SIC_w(Q)$. So $SIC_w(P) = \bigcup_{s \in P} SIC_w(\{s\})$. When $s \neq s'$, $SIC_w(\{s\}) \cap SIC_w(\{s'\}) = \emptyset$. So by the additivity of Pr:

$$\Pr(SIC_w(P)) = \sum_{s \in P} \Pr(SIC_w(\{s\})).$$

The pay-off for a bet on P in the event that w will be $p_P(w) = \Pr(SIC_w(P)) = \sum_{s \in P} \Pr(SIC_w(\{s\}))$, so the expected pay-off will be

$$\sum_{w \in W} \Pr(\mathcal{P}(w)) \sum_{s \in P} \Pr(SIC_w(\{s\}) = \sum_{s \in P} \sum_{w \in W} \Pr(\mathcal{P}(w)) \Pr(SIC_w(\{s\}))$$

$$= \sum_{s \in P} \Pr(\{s\}) = \Pr(P).$$

So b_P will be a fair standard bet for all P, and the set of fair standard bets will be additive. Moreover, the minimal requirements will hold (for if P is determinately true at w then $SIC_w(P) = \mathcal{F}$, and if P is determinately false at w then $SIC_w(P) = \emptyset$). Finally, the Weak Cancellation condition will hold trivially.   □

Proof of Theorem 7: The proof was—effectively—given in the proof of Theorem 6. For it was there shown that the class of fair standard bets on φ for a McGee-measure satisfies partial compensation and the latter entails the minimal requirements, bet-additivity and the partial compensation condition.   □

Proof of Observation 4: Assume a measure Pr that is coherent according to the settlement conditions. From bet-additivity: $\Pr(A \to B) = \Pr(A \wedge (A \to B)) + \Pr(\neg A \wedge (A \to B))$. Given rRT, $\Pr(A \to B) = \Pr(A \wedge B)/\Pr(A)$. Given the minimal requirements $\Pr(A \wedge A \to B) = \Pr(A \wedge B)$, so it follows that $\Pr(\neg A \wedge (A \to B)) = \Pr(\neg A) \times \Pr(A \to B)$.

Now assume factual sentences $E_F, E_1, \ldots, E_n$ that partition the worlds so that any fair bet (c, p) on $\neg A \wedge (A \to B)$ will satisfy: $p(w) = p(w')$ if $w, w' \in E_x$. So we can write $p(E_x)$ for the pay-off in partition E_x. Assume that E_F is the outcome that makes our sentence determinately false (i.e. E_F are the worlds in which A is true, and so worlds where $p(w) = \$0$). The remaining elements $E_1, \ldots, E_n$ partition possible ways in which $\neg A$ is true. Assume that $0 < \Pr(\neg A) < 1$. The expected value of a fair bet on $\Pr(\neg A \wedge A \to B)$ should be $\$\Pr(\neg A) \times \Pr(A \to B)$.

If $n = 1$ we get $\Pr(E_1) p(E_1) = \Pr(\neg A) \Pr(A \to B)$. But $\Pr(E_1) = \Pr(\neg A)$. So $p(E_1) = \$\Pr(A \to B)$. However, the pay-off $\$\Pr(A \to B)$ is not identical to either of the three allowed pay-offs $\$0$, $\$1$ or $\$\Pr(\neg A)\Pr(A \to B)$.

So $n > 1$. There are uncountably many ways of distributing the probability of $\neg A$ over the E_i's in a way that does not affect the probability of either $\neg A$ or $A \rightarrow B$ (given rRT). However, the *expected pay-off* on a bet on $\neg A \wedge A \rightarrow B$ for each distribution must be the same: for any given distribution Pr we need $\Pr(E_1)p(E_1) + \cdots + \Pr(E_n)p(E_n) = \Pr(\neg A)\Pr(A \rightarrow B) = (\Pr(E_1) + \cdots + \Pr(E_n))\Pr(A \rightarrow B)$. But given that each pay-off $p(E_i)$ can only take one of three values this cannot be achieved, not without imposing constraints on how the probabilities on the E_i's are distributed. $\qquad \square$

Proof of Observation 5: First a useful lemma.

Lemma 4 Assume that all logical and probabilistic properties stated in the Observation hold, and that rRT holds. Assume, moreover that $\Pr(\varphi \wedge (A \rightarrow B)) = \Pr(\varphi)$, and that $\Pr(A) > 0$. Then either $\Pr(A \wedge \neg B) = 0$ or $\Pr(\neg A \wedge C) = 0$ for every factual C that logically implies φ.

Proof Assume that $\Pr(\varphi \wedge (A \rightarrow B)) = \Pr(\varphi)$, that $\Pr(A) > 0$ and that C logically implies φ.

Assume for reductio that $\Pr(\varphi \wedge (A \rightarrow \neg B)) > 0$. As $\Pr(\varphi \wedge (A \rightarrow B)) = \Pr(\varphi)$, by ordinary probabilistic reasoning it then follows that $\Pr(\varphi \wedge (A \rightarrow B) \wedge (A \rightarrow \neg B)) > 0$. $\varphi \wedge (A \rightarrow B) \wedge (A \rightarrow \neg B)$ implies $(A \rightarrow B) \wedge (A \rightarrow \neg B)$ which by normality is equivalent to $A \rightarrow (B \wedge \neg B)$. As $\Pr(A) > 0$, we can apply rRT; but this will give $\Pr(A \rightarrow (B \wedge \neg B)) = 0$. So $\Pr((A \rightarrow B) \wedge (A \rightarrow \neg B)) = 0$ and so $\Pr(\varphi \wedge (A \rightarrow B) \wedge (A \rightarrow \neg B)) = 0$, and we have a contradiction.

So $\Pr(\varphi \wedge A \rightarrow \neg B) = 0$. As C implies φ, $\Pr(C \wedge (A \rightarrow \neg B)) = 0$.

Let
$$\psi = (A \vee \neg C) \rightarrow (A \wedge \neg B).$$

By Sα, ψ logically implies $((A \vee \neg C) \wedge A) \rightarrow \neg B$ and so by LLE ψ logically implies $A \rightarrow \neg B$. As $\Pr(C \wedge (A \rightarrow \neg B)) = 0$, $\Pr(C \wedge \psi) = 0$.

By rMP, $\neg A \wedge \neg C \wedge \psi$ and $A \wedge B \wedge \neg C \wedge \psi$ are logical contradictions and so will have probability 0. The four sentences C, $\neg A \wedge \neg C$, $A \wedge B \wedge \neg C$ and $A \wedge \neg B \wedge \neg C$ form a logical partition, and the first three get probability 0 when conjoined with ψ, so by additivity we have:
$$\Pr(\psi) = \Pr(A \wedge \neg B \wedge \neg C \wedge \psi).$$

However, by rCS, $A \wedge \neg B \wedge \neg C$ logically implies ψ so
$$\Pr(A \wedge \neg B \wedge \neg C \wedge \psi) = \Pr(A \wedge \neg B \wedge \neg C).$$

So:
$$\Pr(\psi) = \Pr(A \wedge \neg B \wedge \neg C).$$

By rRT:
$$\begin{aligned} \Pr(\psi) &= \Pr((A \vee \neg C) \wedge A \wedge \neg B)/\Pr(A \vee \neg C) \\ &= \Pr(A \wedge \neg B)/\Pr(A \vee \neg C). \end{aligned}$$

So:
$$\Pr(A \wedge \neg B)/\Pr(A \vee \neg C) = \Pr(A \wedge \neg B \wedge \neg C).$$

This can only hold if either $\Pr(A \wedge \neg B) = 0$, or $\Pr(A \vee \neg C) = 1$ (i.e. if $\Pr(\neg A \wedge C) = 0$). $\qquad\square$

Claim 1. Suppose that (i) $\Pr((A \to B) \wedge (C \to B)) = \Pr(C \to B)$, (ii) $\Pr(A \wedge \neg B) > 0$, (iii) $\Pr(\neg A \wedge C \wedge B) > 0$. Suppose, moreover, for *reductio* that Pr satisfies rRT. By rCS, $C \wedge B$ logically implies $(C \to B)$. By Lemma 4, either $\Pr(A \wedge \neg B) = 0$ or $\Pr(\neg A \wedge C \wedge B)) = 0$. The first alternative falls with premise (ii) and the second with premise (iii) and so we have a contradiction.

Claim 2. Suppose that (i) $\Pr(C \wedge (A \to B)) = \Pr(C)$, (ii) $\Pr(A \wedge \neg B) > 0$, and (iii) $\Pr(\neg A \wedge C) > 0$. Furthermore, by way of *reductio*, suppose that Pr satisfies rRT. C logically implies itself so from Lemma 4 we have either $\Pr(A \wedge \neg B) = 0$ or $\Pr(\neg A \wedge C) = 0$. The first alternative falls with premise (ii) and the second with premise (iii) and so we have a contradiction.

Claim (3). Suppose that (i) $\Pr((A \to B) \wedge (\neg A \to B)) = \Pr(B)$, (ii) $0 < \Pr(A) < 1$, (iii) $0 < \Pr(B) < 1$. Suppose, moreover, for *reductio* that Pr satisfies rRT.

By additivity $\Pr((A \to B) \wedge (\neg A \to B)) = \Pr(A \wedge (A \to B) \wedge (\neg A \to B)) + \Pr(\neg A \wedge (A \to B) \wedge (\neg A \to B))$. By rMP and rCS $A \wedge (A \to B)$ is logically equivalent to $A \wedge B$, and $\neg A \wedge (\neg A \to B)$ is logically equivalent to $\neg A \wedge B$. So $\Pr((A \to B) \wedge (\neg A \to B)) = \Pr(A \wedge B \wedge (\neg A \to B)) + \Pr(\neg A \wedge B \wedge (A \to B)) = \Pr(B)$ (by (i)). As $\Pr(B) = \Pr(A \wedge B) + \Pr(\neg A \wedge B)$ it follows that (1) $\Pr(A \wedge B \wedge (\neg A \to B)) = \Pr(A \wedge B)$ and (2) $\Pr(\neg A \wedge B \wedge (A \to B)) = \Pr(\neg A \wedge B)$.

Consider the first of these identities. $A \wedge B$ logically implies itself, so from Lemma 4 either (a) $\Pr(\neg A \wedge \neg B) = 0$ or (b) $\Pr(\neg \neg A \wedge A \wedge B) = \Pr(A \wedge B) = 0$. Consider the second identity. $\neg A \wedge B$ logically implies itself, so from Lemma 4 either (c) $\Pr(A \wedge \neg B) = 0$ or (d) $\Pr(\neg A \wedge \neg A \wedge B) = \Pr(\neg A \wedge B) = 0$. Four possibilities. If (a)+(c) hold then $\Pr(\neg B) = 0$ and this contradicts premise (iii). If (a)+(d) then $\Pr(\neg A) = 0$ and this contradicts premise (ii). If (b)+(c) then $\Pr(A) = 0$ and this contradicts premise (ii). If (b)+(d) then $\Pr(B) = 0$ and this contradicts premise (iii). Those are all the alternatives, so we have a contradiction. $\qquad\square$

References

1. Bradley, R. (2005). Radical Probabilism and Bayesian Conditioning. *Philosophy of Science*, *72*(2), 342–364.

2. Bradley, R. (2012). Multidimensional possible-world semantics for conditionals. *Philosophical Review*, *121*(4), 539–571.
3. Cantwell, J. (2021). An expressivist analysis of the indicative conditional with a restrictor semantics. *Review of Sumbolic Logic*, *14*(2), 487–530.
4. Douven, I., & Verbrugge, S. (2010). The Adams family. *Cognition*, *117*, 302–318.
5. Edgington, D. (1991). The mystery of the missing matter of fact. *Proceedings of the Aristotelian Society Supplementary Volume*, *65*, 185–209.
6. Evans, J., Over, D., & Handley, S. (2003). A theory of hypothetical thinking. In D. Hardman, & L. Maachi (Eds.) *Thinking: Psychological perspectives on reasoning and decision making* (pp. 3–22). Chichester: Wiley.
7. Gibbard, A. (1981). Two recent theories of conditionals. In Harper, Stalnaker, & Pearce (Eds.) *Ifs* (pp. 211–247). Dordrecht: Reidel.
8. Goldstein, S., & Santorio, P. (2021). Probability for epistemic modalities. Unpublished manuscript. url: https://philpapers.org/archive/GOLPFE-2.pdf.
9. Hájek, A., & Hall, N. (1994). The hypothesis of the conditional construal of conditional probability. In E. Eells, & B. Skyrms (Eds.) *Probability and Conditionals: Belief Revision and Rational Decision* (pp. 75–111). Cambridge: Cambridge University Press.
10. Jeffrey, R. (1983). *The Logic of Decision*, 2nd edn. Chicago: University of Chicago Press.
11. Kaufmann, S. (2004). Conditioning Against the Grain. *Journal of Philosophical Logic*, *33*, 583–606.
12. Kaufmann, S. (2009). Conditionals right and left: Probabilities for the whole family. *Journal of Philosophical Logic*, *38*, 1–53.
13. Khoo, J. (2016). Probabilities of conditionals in context. *Linguistics and Philosophy*, *39*, 1–43.
14. Khoo, J. (2022). *The Meaning of If*. Forthcoming. Oxford University Press.
15. Khoo, J., & Santorio, P. (2018). Lecture notes: Probability of conditionals in modal semantics. www.justinkhoo.org/teaching.
16. Kolodny, N., & MacFarlane, J. (2010). Ifs and oughts. *Journal of Philosophy*, *107*(3), 115–143.
17. Korzukhin, T. (2014). Contextualist theories of the Indicative conditional and Stalnaker's thesis. *Thought*, *3*, 177–183.
18. Kratzer, A. (1977). What 'must' and 'can' must and can mean. *Linguistics and Philosophy*, *1*(3), 337–356.
19. Kratzer, A. (2012). *Modals and Conditionals: New and Revised Perspectives*. Oxford: Oxford University Press.
20. Lance, M. (1991). Probabilistic dependence among Conditionals. *Philosophical Review*, *100*, 269–275.
21. Lewis, D. (1976). Probabilities of Conditionals and Conditional Probabilities. *Philosophical Review*, *85*, 297–315.
22. Lewis, D. (1986). Probabilities of Conditionals and Conditional Probabilities II. *Philosophical Review*, *95*, 581–589.
23. Mandelkern, M. (2019). Bounded modality. *The Philosophical Review*, *128*(1), 1–61.
24. McDermott, M. (1996). On the truth conditions of certain 'If'-sentences. *Philosophical Review*, *105*, 1–37.
25. McGee, V. (1985). A Counterexample to Modus Ponens. *Journal of Philosophy*, *82*, 462–471.
26. McGee, V. (1989). Conditional probabilities and compounds of conditionals. *Philosophical Review*, *98*, 485–542.
27. McGee, V., & McLaughlin, B. (1994). Distinctions without a difference. *The Southern Journal of Philosophy*, *33*, 203–251.
28. Milne, P. (1997). Bruno de Finetti and the logic of conditional events. *British Journal for the Philosophy of Science*, *48*(2), 195–232.
29. Nute, D., & Cross, C. B. (2001). Conditional logic. In D. Gabbay, & F. Guenthner (Eds.) *Handbook of Philosophical Logic*, (Vol. 4 pp. 1–98). Kluwer Academic Publishers.
30. Oberauer, K. (2006). Reasoning with conditionals: A test of formal models of four theories. *Cognitive Psychology*, *53*(3), 238–283.
31. Over, D., & Evans, J. (2003). The probability of conditionals: the psychological evidence. *Mind & Language*, *18*(4), 340–358.
32. Stalnaker, R. (2014). Context. Oxford University Press.
33. Stalnaker, R., & Jeffrey, R. (1994). Conditionals as random variables. In E. Eells, & B. Skyrms (Eds.) *Probability and Conditionals: Belief Revision and Rational Decision* (pp. 31–46). Cambridge: Cambridge University Press.

34. Thomason, R. (2012). The semantics of conditional modality. Unpublished manuscript. http://www.eecs.umich.edu/rthomaso/.

35. van Fraassen, B. (1976). Probabilities of Conditionals. In W. L. Harper, & C. A. Hooker (Eds.) *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science* (pp. 261–300). D. Reidel.

36. van Wijnbergen-Huitink, J., Elqayam, S., & Over, D.E. (2014). The probability of iterated conditionals. *Cognitive Science*, *39*(4), 788–803. https://doi.org/10.1111/cogs.12169

37. Vidal, M., & Baratgin, J. (2017). A psychological study of unconnected conditionals. *Journal of Cognitive Psychology*, *29*, 769–781.

38. Yalcin, S. (2012). A Counterexample to Modus Tollens. *Journal of Philosophical Logic*, *41*(6), 1001–1024.