



Smith ScholarWorks

Philosophy: Faculty Publications

Philosophy

6-3-2018

On Conditionals

Theresa Helke

Smith College, thelke@smith.edu

Follow this and additional works at: https://scholarworks.smith.edu/phi_facpubs



Part of the [Philosophy Commons](#)

Recommended Citation

Helke, Theresa, "On Conditionals" (2018). Philosophy: Faculty Publications, Smith College, Northampton, MA.

https://scholarworks.smith.edu/phi_facpubs/32

This Dissertation has been accepted for inclusion in Philosophy: Faculty Publications by an authorized administrator of Smith ScholarWorks. For more information, please contact scholarworks@smith.edu

ON CONDITIONALS

THERESA SOPHIE CAROLINE HELKE
(BA(Hons), Smith College)

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF PHILOSOPHY
NATIONAL UNIVERSITY OF SINGAPORE & YALE-NUS COLLEGE

2018

Supervisors:

Associate Professor Ben Blumson, Main Supervisor
Assistant Professor Malcolm Keating, Co-Supervisor, Yale-NUS College

Examiners:

Assistant Professor Robert Beddor
Assistant Professor Michael Erlewine
Emeritus Professor Frank Jackson, Australian National University

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.



Theresa Helke

Theresa Sophie Caroline Helke
3 June 2018

ACKNOWLEDGEMENTS

I'd like to thank

- supervisors Ben Blumson at the National University of Singapore (NUS) and Malcolm Keating at Yale-NUS College;
- original supervisor Jay Garfield, now back at Smith College;
- members of the 7 November 2017 reading group: Bob Beddor, Frank Jackson, Mike Pelczar, Abelard Podgorski and Tang Weng Hong;
- Linguistics professor mitcho Erlewine;
- audience members at the February 2018 Israeli Philosophical Association conference in Haifa, December 2017 New Zealand Association of Philosophy conference in Dunedin (especially Max Cresswell and Adriane Rini), November 2017 Australasian Postgraduate Philosophy Conference in Brisbane and March 2016 Kyoto-NUS-Chengchi conference in Kyoto;
- anonymous reviewers at *Synthese* and the *Australasian Journal of Philosophy*;
- friends and reviewers Josephine Sedgwick and David Zagoury; and
- Cambridge supervisor Steven Methven, now at the Other Place.

Without them, I wouldn't be submitting this thesis inasmuch as I wouldn't be submitting one at all, I'd be submitting one on a different topic, or I'd be submitting a version of this thesis with more problems.

TABLE OF CONTENTS

Introduction	1
Chapter 1: The Four Theories	5
Chapter 2: The Election Paradox	34
Chapter 3: The Barbershop Paradox	71
Chapter 4: On Premise Semantics and Apparent Counterexamples to <i>Modus Ponens</i> and <i>Modus Tollens</i>	105
Conclusion	147
References	150

SUMMARY

This thesis is about theories of the indicative conditional and apparent counterexamples to classically valid argument forms. Specifically, it applies the following four theories:

- material (inspired by Grice (1961, 1975 and 1989));
- possible-worlds (inspired by Stalnaker (1981); Lewis (1976); and Kratzer (2012)),
- suppositional (inspired by Adams (1975) and Edgington (1995 and 2014)); and
- hybrid (inspired by Jackson (1987))

to try and solve the following two counterexamples:

- Vann McGee's to *modus ponens* (1985); and
- Lewis Carroll's to *modus tollens* (1894).

I argue that none of the theories I consider can explain – without facing any problems – the individually plausible but jointly inconsistent theses that give rise to the apparent counterexamples. A theory can explain a thesis when it can account for why a naïve speaker might have the relevant intuition. In the case of McGee's Election Paradox, the theses are the following:

- McGee's argument is invalid;
- McGee's argument is an instance of *modus ponens*; and
- *modus ponens* is valid.

Similarly, in the case of Carroll's Barbershop Paradox, the theses are the following:

- Carroll's argument is invalid;

- Carroll's argument is an instance of *modus tollens*; and
- *modus tollens* is valid.

Despite consensus that the orthodox theories, those I examine, provide answers to questions about conditionals, the paradoxes (even some after over a century) persist: arguments we take obviously to be valid appear to have instances in which they aren't.

The thesis proceeds in four parts. In Chapter 1, I present the four theories and justify why they rather than others deserve our attention. Along the way, I disagree with Adams (1975). While he claims that on the suppositional theory, we can't evaluate an indicative conditional whose consequent is a conditional, I include a proof that enables us to.

In Chapter 2, I consider McGee's counterexample to *modus ponens* and how the theories might explain the theses of the relevant trilemma. I refute extant solutions: while the Dartmouth group (1986) thinks the material theory can solve the paradox and Edgington thinks the suppositional one can solve it and Jackson thinks the hybrid one can, I show that their solutions fall short of being comprehensive.

In Chapter 3, I do the same, this time considering Carroll's counterexample to *modus tollens*. I show that the paradox remains without a solution – and this despite the progress logicians made on conditionals in the centuries following the paradox's publication and the attempts logicians made at offering a solution.

Finally, in Chapter 4, I examine how premise semantics (Kratzer, 2012), a version of the possible-worlds theory, might explain McGee's and Carroll's counterexamples along with some others, these featuring not embedded

conditionals but embedded modals. I argue that, when it comes to counterexamples to *modus ponens* and *modus tollens* with (overt) modal verbs and adverbs, the theory gives us some results we want (e.g. invalidating Kolodny and MacFarlane's *Miners* examples (2010)) and others we don't (e.g. validating Cantwell's *Lottery* (2008)) – a point which the literature previously overlooked.

LIST OF TABLES

Table 1: Truth conditions of the indicative conditional according to the material theorist

Table 2: Possible combinations for who, between Carter and Anderson, might win if Reagan doesn't

Table 3: Possible combinations for who, among the barbers, might be in/out

Table 4: Breakdown of the blue/red/big/small marbles in the urn

LIST OF FIGURES

Figure 1: Venn diagram of the Election Paradox

Figure 2: Venn diagram of the Barbershop Paradox

LIST OF SYMBOLS

\supset	hook (material conditional)
\rightarrow	arrow (indicative conditional)
	given
\wedge	and
\vee	or
\neg	not
\therefore	therefore
+	plus
-	minus
\times	times
\div	divided by
=	equals
\geq	greater than or equal to
\leq	less than or equal to
\cap	intersection
\cup	union
\in	element of
$\{p\}$	the singleton set containing p
$[[p]]$	the proposition which p expresses
\emptyset	the empty set

INTRODUCTION

If you're reading this, then you're alive.

You're reading this.

Therefore, you're alive.

and

If you're Theresa Helke, then you wrote this example.

You didn't write this example.

Therefore, you aren't Theresa Helke.

are instances of *modus ponens* and *modus tollens* respectively. The arguments, variations perhaps on the *cogito ergo sum*, have the following forms:

If A , then B .

A .

Therefore, B .

in the case of the *modus ponens* and

If C , then D .

Not- D .

Therefore, not- C .

in the case of the *modus tollens*.

The first premise in each is a conditional. Here, in our *modus ponens*, it's 'If A , then B '; and in our *modus tollens*, 'If C , then D '. The second premise and conclusion in each do different things. In a *modus ponens*, the second premise confirms the antecedent of the first, here A , and the conclusion confirms the consequent of the first, here B . In a *modus tollens*, on the other hand, the second premise contradicts the consequent of the first, here D (i.e.

confirms ‘not-*D*’) and the conclusion contradicts the antecedent of the first premise, here *C* (i.e. confirms ‘not-*C*’).

You might think the two argument forms are valid. When both the premises are true, so too must be the conclusion, i.e. when it’s true both that if you’re reading this, then you’re alive *and* you’re reading this, then it must be true that you’re alive; likewise, when it’s true both that if you’re Theresa Helke, then you wrote the example *but* you didn’t write the example, then it must be true that you aren’t Theresa Helke.

Certainly, an introduction to Logic textbook would agree that *modus ponens* and *modus tollens* are valid. This thesis is about how they might not be: how arguments can have the forms we saw above and premises we’d accept but a conclusion we’d reject. (Maybe you aren’t alive after all! Maybe you’re me!) Indeed, the research project focuses on theories of the indicative conditional and apparent counterexamples to classically valid argument forms.

Specifically, it applies the material (inspired by Grice (1961, 1975 and 1989)), possible-worlds (inspired by Stalnaker (1981); Lewis (1976); and Kratzer (2012)), suppositional (inspired by Adams (1975) and Edgington (1995 and 2014)) and hybrid (inspired by Jackson (1987)) theories to try and solve Vann McGee’s counterexample to *modus ponens* (1985) and Lewis Carroll’s to *modus tollens* (1894).

Throughout, I understand an indicative conditional to be like the first – and not the second – of the Adams pair below:

- (i) If Oswald didn’t shoot Kennedy, then someone else did.
 - (ii) If Oswald hadn’t shot Kennedy, then someone else would have
- (from Khoo (2015, p. 1) who draws on Adams (1970)).

I agree with Justin Khoo (2015) that '(i) is about how the world was, given what we *now* know plus the supposition of the antecedent; (ii) is about how the world would have been (now that we know about it) were its antecedent to have held (pp. 1-2, emphasis in original).

There are both syntactic and semantic differences between (i) and (ii), which I understand to be a subjunctive conditional. The syntactic difference lies in the 'extra layer of past tense' (which in English we mark morphologically with the past perfect 'had') which (ii) carries and the presence in the consequent of (ii) of the modal auxiliary verb 'would' (Khoo, 2015, p. 1).

The semantic difference between (i) and (ii), on the other hand, lies in their truth conditions. Indeed, (i) is true while (ii) is false. We know that someone shot Kennedy and that there was no backup shooter (assuming we're Warrenites) (Khoo, 2015, p. 1).

If I'm focusing on indicative rather than subjunctive conditionals, it's because the conditionals in the apparent counterexamples to *modus ponens* and *modus tollens* – the paradoxes – I want to analyse are not subjunctive but indicative.¹

My argument is that none of the theories I consider can explain successfully the individually plausible but jointly inconsistent theses that give rise to the apparent counterexamples. Despite consensus that the orthodox theories, those I examine, provide answers to questions about conditionals, the

¹ Note that beyond limiting the entire thesis to indicative conditionals, I limit much of it to indicative conditionals where neither the antecedent nor the consequent contains any (overt) modal. Indeed, it's only in Chapter 4 that I consider conditionals whose consequent contains an overt modal.

paradoxes (even some after over a century) persist: arguments we take obviously to be valid appear to have instances in which they aren't.

In Chapter 1, I present the four theories and justify why they rather than others deserve our attention. In Chapter 2, I consider McGee's counterexample to *modus ponens* (the 'Election Paradox') and how the theories might explain the theses of the relevant trilemma. In Chapter 3, I do the same, this time considering Carroll's counterexample to *modus tollens* (the 'Barbershop Paradox'). And in Chapter 4, I examine how premise semantics, (Kratzer, 2012), might explain McGee and Carroll's counterexamples along with some others.²

² Note that throughout, I'll be evaluating the purported counterexamples as though a single speaker were uttering or considering the propositions. The results might be different if one evaluated each argument as – not a soliloquy but – a dialogue between two speakers. Thanks to mitcho Erlewine for pointing this out.

CHAPTER 1

THE FOUR THEORIES

Consider the following two propositions:

The Equivalence Thesis: The indicative conditional is the material conditional; and

Adams' Thesis: The assertibility of 'If Q , then R ' is equal to the conditional probability of R given Q .

For the time being, we understand assertibility of a sentence as the extent to which we are justified in asserting that sentence. We'll return to this concept in due course. Now, when it comes to accepting or rejecting these two propositions, there are four possibilities:

- (i) Accept the Equivalence Thesis but reject Adams' Thesis;
- (ii) Reject both the Equivalence Thesis and Adams' Thesis;
- (iii) Reject the Equivalence Thesis but accept Adams' Thesis; and
- (iv) Accept both the Equivalence Thesis and Adams' Thesis.

In this chapter, I consider in turn the four possibilities, each corresponding with a theory of the indicative conditional: (i) the material; (ii) the possible-worlds; (iii) the suppositional; and (iv) the hybrid theory. In sections 1 through 4, I'll outline their main tenets and note some merits and demerits – building a foundation for the chapters to come. In section 5, I conclude.

If these theories deserve our attention here, it's because they appear to be good candidates for solving the Election and Barbershop paradoxes we'll see later. The paradoxes challenge classical logic and the theories were built to explain a challenge to classical logic, i.e. that there appears to be some

difference between the material conditional we use to derive proofs and the indicative one we use in everyday speech. Of course, there are other theories: Cantwell (2008), Yalcin (2012), Bledin (2015) to mention only three. If I don't consider them here it's not because they don't deserve our attention. Rather, it's because considering them lies beyond the scope of this thesis.

1. The material theory

According to the material theorist, the Equivalence Thesis or the 'horseshoe analysis' holds: the indicative conditional (which I'll represent as the arrow \rightarrow) is the material conditional (which I'll represent as the hook \supset); an indicative conditional $Q \rightarrow R$ is true if and only if the corresponding material conditional $Q \supset R$ is true. In other words, 'If Q , then R ' is true if and only if Q is false or R is true. For example, let Q be 'Ben's on sabbatical' and R be 'he's in Australia'. The indicative conditional 'If Ben's on sabbatical, then he's in Australia' is true if and only if it's not the case that Ben's on sabbatical or it's the case that he's in Australia.

So, according to the material theorist, we can represent the truth conditions in the following truth table:

Q	R	$Q \rightarrow R$
True	True	True
True	False	False
False	True	True
False	False	True

Table 1: Truth conditions of the indicative conditional according to the material theorist

Looking at the table, one might think the truth conditions absurd: How can a conditional be true, say, when it has a false antecedent and true consequent? Or when both the antecedent and consequent are false? One way of making sense of the table is through the following story. Imagine one day I promise you ‘If you bring me durian, then I’ll pay you SGD 30’. The following day, when you bring me durian and I pay you SGD 30, I honoured the promise. When I said ‘If you bring me durian, I’ll pay you SGD 30’, I spoke the truth. This accounts for the true-true case. When you don’t bring me any fruit and I don’t pay you any money, I didn’t fail to honour the promise. In asserting the conditional, I spoke the truth. This accounts for the false-false case. Likewise, when you don’t bring me any fruit but (out of magnanimity?) I decide to give you SGD 30, I didn’t fail to honour the promise. In asserting the conditional, I spoke the truth. This accounts for the false-true case. But when you bring me durian and I don’t pay you any money, *then* I fail to honour the promise. This accounts for the true-false case.

So yes, according to the material theory, an indicative conditional is true if and only if it’s not the case that the antecedent is true and the consequent is false. In other words, an indicative conditional is true if and only if Q is false or R is true.

This point is important in our discussion in Chapters 2 through 4 of *modus ponens* and *modus tollens*. The validity of both argument forms relies on a conditional being false where it has a true antecedent and false consequent.

Next, according to the material theorist, it's not the case that Adams' Thesis holds. The material theorist concerns herself with the truth of the antecedent and consequent, rather than the conditional probability of the one given the other. Now, there are two kinds of material theorists. On the one hand, there's the earlier theorist. According to her, assertability goes with truth: a conditional is assertable if and only if it's true (Russell, 1905). On the other hand, there's the later theorist. According to her, assertability doesn't go with truth.

Logicians developed the later theory in light of objections to the earlier one. The earlier theory faced problems not when it came to conditionals such as the following two:

True-true If 2 is divisible by 2, then 2 is an even number; and

True-false If 2 is divisible by 2, then 2 is an odd number.

Consistent with the theory, we might be prepared to assert the first which, with a true antecedent and consequent, is true. Likewise consistent with the theory, we wouldn't be prepared to assert the second which, with a true antecedent and false consequent, is false.

No, the earlier theory faced problems when it came to conditionals such as the following two:

False-true If Angela Merkel is the Prime Minister of Singapore, then Angela Merkel is the Chancellor of Germany.

False-false If Vienna is the capital of England, then Vienna is the capital of Switzerland.

Both conditionals are true, according to the earlier (and later) theorist on the grounds that both have a false antecedent. Nonetheless, contra the earlier theorist, we mightn't be prepared to assert either. We would sound absurd!

Considering these objections, later theorists hold that it's not the case that the truth of a conditional is necessary and sufficient for our being prepared to assert it. Indeed, we mightn't be prepared to assert some truth conditionals such as *False-true* or *False-false*. Rather, and central to the later theory, we're prepared to assert an indicative conditional when, doing so, we'd be abiding by the Cooperative Principle.

Before we turn to the principle, note that henceforth when I write 'material theorist', I'm referring to the later kind.

1.1. The Cooperative Principle

According to the Cooperative Principle, you must 'Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged' (Grice, 1975, p. 45).

For example, suppose that you approach me in the street and ask me 'Where is the wet market?' To abide by the Cooperative Principle, I could respond: 'It is at the end of the street on the right', assuming that the market is indeed at the end of the street on the right. Here, I would be making a contribution such as is required (an answer); I would be doing so at the stage at which it occurs (right after you ask a question); by the accepted purpose or direction of the talk exchange in which we are engaged (addressing your specific question). Not to abide by the Cooperative Principle, I could respond

instead: ‘You could go down the street and turn *left*’. Here I might be contributing an answer right after you ask a question. However, I wouldn’t be responding to your specific question. Rather, I’d be responding to the question: ‘Where could I go?’

The material theorist is at pains to define the Cooperative Principle because of the existence of conversational implicature. *False-true* and *False-false* aside, when speaking, I might make a true statement while implicating a false one. For example, I could say ‘The banana seems yellow’. Assuming that the banana indeed seems yellow, I am speaking the truth. However, by using the verb ‘to seem’ rather than ‘to be’ (‘The banana *seems* yellow’ rather than ‘The banana *is* yellow’), I am suggesting that the banana mightn’t actually be yellow. I’m suggesting that the fruit might instead be a different colour.

When it comes to indicative conditionals, assuming that they’re identical to material ones, it’s easy to make true statements while implicating false ones. For example, I could say ‘If the market is not on the left, then it is on the right’. According to the truth table of the material conditional, this statement is true. It has a true consequent. However, truth table aside, I’m implicating a false one. I’m suggesting that the market might be on the left when I know that it is on the right.

Likewise, given a certain scenario, ‘You won’t eat those and live’ (Lewis, 1976, p. 306) is an example of a true but not assertable sentence. Imagine I utter the sentence while pointing at some non-toxic mushrooms and you, deferring to my apparent mycological knowledge, refrain from eating the mushrooms. I told no lie. Formally, the sentence is true. The sentence is a negated conjunction and one of the conjuncts is false (‘you eat those’). Indeed,

it's not the case that you eat the mushrooms. Nonetheless, the original sentence isn't assertable. It wouldn't be a cooperative thing to say.

1.1.1. Maxims

Now, within the Cooperative Principle, there are four categories of maxims by which we abide: Quantity, Quality, Relation and Manner. For the purposes of this dissertation, the following two maxims, one falling under the category of Quality and the other of Manner, are the most relevant:

- (i) Don't assert what you don't believe (Grice, 1975, p. 46); and
- (ii) Assert the stronger rather than the weaker (Jackson, 1979, p. 566).

Let's look at each in turn.

(i) Don't assert what you don't believe

Grice's words are 'Do not say what you believe to be false' but I think the stronger maxim 'Don't assert what you don't believe' is desirable. Grice's maxim allows that, so long as we don't believe a sentence to be false, we may assert it. This means that according to the maxim, if I don't believe my mother is in Geneva but I don't believe my mother's not in Geneva (maybe she changed her plans and is in Geneva?), I could still assert 'My mother is in Geneva'. According to the stronger maxim, however, we couldn't assert 'My mother is in Geneva'. We'd be asserting something we don't believe.

(ii) Assert the stronger rather than the weaker

A sentence Q is logically stronger than another R , where Q entails R but not vice versa. For example, let Q be S and R be $S \vee \neg S$. Here, S entails $S \vee \neg S$

but not vice versa. Suppose that S is ‘ x is divisible by 2’. From this, it follows that ‘ x is divisible by 2 or it is not divisible by 2’. However, from ‘ x is divisible by 2 or it is not divisible by 2’, it does not follow that ‘ x is divisible by 2’. x could be an odd number. According to (ii), we ought to assert S rather than $S \vee \neg S$. S is logically stronger than $S \vee \neg S$.

Of course, there are many possible reasons for wanting to assert R rather than Q . For example, Q might not be true or believable. ‘Hui Li is teaching’ might not be true while ‘Hui Li is teaching or Hui Li isn’t teaching’ is. Another reason for wanting to assert the weaker R rather than the stronger Q is that Q might be ‘unduly blunt’ (Jackson, 1979, p. 566). ‘I may or may not fail you’ is less unequivocal than ‘I will fail you’ and we might want to cultivate hope in our student. But abandoning the context of grading and focusing on epistemic and semantic considerations, there is, according to Jackson, no reason *not* to assert Q : ‘There is no significant loss of probability in asserting $[Q]$ and, by the transitivity of entailment, $[Q]$ must yield everything and more than $[R]$ does. Therefore, $[Q]$ is to be asserted rather than $[R]$, *ceteris paribus*’ (Jackson, 1979, p. 566).

1.2. Returning to False-true and False-false

The material theorist uses the idea of logically weak and strong statements to explain away counterexamples such as *False-true* and *False-false*. When it comes to conditionals $A \supset B$ where one of $\neg A$ or B is highly probable, the material theorist says that you should come right out and assert the logically stronger statement, namely ‘ $\neg A$ ’ or ‘ B ’ as the case may be (Jackson, 1979, pp. 566-7).

For example, we should assert not *False-true* but only its consequent. We believe *False-true* inasmuch as we believe that the consequent is true. The statement ‘Angela Merkel is the Chancellor of Germany’ is logically stronger than ‘If Angela Merkel is the Prime Minister of Singapore, then she is the Chancellor of Germany’. ‘She is the Chancellor of Germany’ entails ‘If Angela Merkel is the Prime Minister of Singapore, then she is the Chancellor of Germany’ but not vice versa. Granted, the conditional is true where the antecedent is false and the consequent true. On the equivalence thesis, however, we should assert only the true consequent.

Likewise, we should assert not *False-false* but only the negation of its antecedent. We believe *False-false* inasmuch as we believe that the antecedent is false. The statement ‘It’s *not* the case that Vienna is the capital of England’ is logically stronger than ‘If Vienna *is* the capital of England, then Vienna is the capital of Switzerland’. ‘It’s not the case that Vienna is the capital of England’ entails ‘If Vienna *is* the capital of England, then Vienna is the capital of Switzerland’ but not vice versa. The conditional is true where both the antecedent and consequent are false. On the equivalence thesis, however, we should assert only the negation of the antecedent.

Another way of dismissing *False-true* and *False-false* is by saying that we wouldn’t assert them because they are purposefully deceptive or confusing. *False-true* suggests that Angela Merkel is the Prime Minister of Singapore. *False-false* suggests that Vienna is the capital of England and Switzerland.

As we’ll see in the Chapters 2 through 4, this theory faces challenges. Moving forward, we might reject the Equivalence Thesis. Accepting it

required a principle to explain why we mightn't assert true conditionals. The possible-worlds theorist rejects the Equivalence Thesis. Let's turn to it next.

2. The possible-worlds theory

Like the material theorist and on the same grounds, the possible-worlds one rejects Adams' Thesis. She concerns herself with the truth of the antecedent and consequent, rather than the conditional probability of the one given the other. Unlike the material one, however, the possible-worlds theorist furthermore rejects the Equivalence Thesis.

According to the possible-worlds theorist, an indicative conditional 'If Q , then R ' is true in a possible world w if and only if R is true in all the Q -worlds which are most similar to w (and vacuously true when there's no Q -world), where Q -world is one in which the antecedent Q is true (Stalnaker, 1981, pp. 46-7; see also Lewis, 1973, though as the title suggests and unlike us here, Lewis is talking about counterfactual rather than indicative conditionals). For example, the conditional 'If Nicholas is at the office, then he's in London' is true in the actual world just in case Nicholas is in London in all most-similar worlds in which Nicholas is at the office.

And according to the possible-worlds theorist, it's not the case that we're prepared to assert an indicative conditional when there's a high probability in the consequent given the antecedent. Rather, we're prepared to assert a conditional if and only if it's true in every antecedent-world most-similar to the actual world.³

³ This isn't to say the possible-worlds theorist rejects the Equivalence Thesis so much as to say it's not applicable. The theorist doesn't use the Thesis to explain a difference between the material and indicative conditionals.

Now, possible-worlds theorists disagree on the number of most-similar antecedent-worlds to any given possible world. Some such as Stalnaker (1981) hold that there can be only one (p. 46). Let's call them the single-world theorists. Others such as Lewis (1973) hold that there can be more than one (pp. 97-8). Let's call them the multiple-world theorists and look at each in turn, considering their merits and demerits. (As Lewis's theory is about counterfactual conditionals and here we are looking at indicative ones, the multiple-world theorist is an imagined version of Lewis.)

2.1. Single-world theory

Predictably perhaps, the single-world theorist does a good job accounting for conditionals where there's a single most-similar antecedent-world. Consider the truth of the conditional $A \rightarrow B$ where, say, A is 'the train arrives late at 5:07pm' and B is 'I will miss the connecting 5:03pm train' and suppose that the train is due to arrive at 5:00pm. This conditional is true. We look at the most-similar A -worlds, those where it's true that the train arrives at 5:07pm. And in v , the A -world closest to w , the consequent is true. In v , everything is exactly the same as in w with the exception that the train arrives seven minutes late. The connecting train leaves on time at 5:03pm and I miss it. With a true antecedent and a true consequent in the antecedent-world closest to the actual one, the conditional is true.

Of course, there are other worlds in which the train arrives late, at 5:07pm, but I don't miss the connecting train. For example, in u , the connecting train might leave late, at 5:15pm, allowing me eight minutes to alight from the one train and board the other. In this world, the conditional is

false. It has a true antecedent but a false consequent. However, it doesn't make the conditional false in the actual world. u is not as close to w as v and we are concerned with the truth value of the conditional in the A -world closest to w . u is different not only in that the first train arrives seven minutes late but also in that the second train leaves 12 minutes late. Thus we need not worry about the truth value of the conditional in u , only in v .

The single-world theorist does less good a job accounting for conditionals where there's no most-similar antecedent-world. Such a scenario could arise for a couple of reasons: first, it may be that given any antecedent-world, there's another, more similar to the actual world; and second, it may be that two or more worlds are equally similar to the actual world and more similar to it than any other worlds (Lewis, 1973, p. 80).

Here are a couple of examples where there's no most-similar antecedent-world for each of those reasons.

(i) *Taller* (inspired by Edgington, 1995, p. 252)

If I'm taller than 2.5m, then I must pay to ride the bus.

In the case of the conditional above, for any given antecedent-world, there's another, more similar one. The more my height tends (from above) toward 2.5m in a possible world, the more similar that world is to the actual world, where I'm shorter than 2.5m. However, there'll be no most-similar world. For every antecedent-world where I am 2.55m tall, there's another, more similar world where I'm 2.525m tall. And for every antecedent-world where I'm 2.525m tall, there's another, even more similar one where I'm 2.5125m tall, etc. Without a most-similar possible world, the theory can't determine the

truth value of the conditional and without a truth value, the theory can't determine whether we'd be prepared to assert the conditional.

(Of course, short of being able to determine the truth value, we might think this is a case where the conditional is vacuously true (no Q -world?) – but that isn't desirable either. As Edgington points out, it would follow that 'If were taller than I am, no one would know the difference' would come out as true (1995, p. 252). While this isn't an indicative conditional, it's relevant nonetheless as one with potentially no most-similar antecedent world.)

(ii) *Les compatriotes* and *I compatrioti*

If François and Italo are compatriots, then they're both French; and

If François and Italo are compatriots, then they're both Italian.⁴

In the case of each conditional in the pair above, there are two most-similar antecedent-worlds: the world in which Italo is French and the world in which François is Italian; and in all respects except the nationality of the relevant man, the two worlds are exactly the same as the actual world. With no single most-similar antecedent-world, the single-world theorist can't determine the truth value of the conditional and, thereby, can't determine whether we'd be prepared to assert it.

That said, the single-world theorist would take the disjunction of the two (i.e. where *Les compatriotes* and *I compatrioti* each form a disjunct) to be true. Indeed and unlike the multiple-world one, the single-world theorist can derive the principle of Conditional Excluded Middle, according to which $(Q \rightarrow R) \vee (Q \rightarrow \neg R)$ is a tautology. This will prove valuable in explaining a

⁴ See Quine, 1952, p. 15 for the original pair of counterexamples with Bizet and Verdi; or Lewis, 1973, p. 80 for more recent a discussion.

thesis (#2) in Chapter 3 (subsection 2.2.). We'll save her proof of the principle until then.

2.2. The multiple-world theory

In contrast but still predictably perhaps, the multiple-world theorist does a good job accounting for the two cases above. According to her, both conditionals come out as false because for *each* there are two most-similar antecedent-worlds *each* assigning a different truth value to the consequent – and the theory considers such conditionals false. That said, the multiple-world theorist doesn't systematically do better than the single-world one as we'll see in Chapters 2 and 3. Indeed, her rejection of the principle of Conditional Excluded Middle prevents the multiple-world theorist from explaining a thesis (#2) which the single-world one can in Chapter 3.

Finally and briefly, there's a third kind of possible-worlds theorist: we'll call her the premise semanticist. I'll wait until Chapter 4 to give a full exegesis. The theory offers an account of statements with embedded modals and in that chapter we'll consider some.

3. The suppositional theory

The suppositional theorist rejects the Equivalence Thesis but accepts Adams' Thesis. According to her, conditionals are things we evaluate in terms of probabilities (Adams (1975) and Edgington (1995 and 2014)). Indeed, according to her, we're prepared to assert an indicative conditional $Q \rightarrow R$ when there's a high conditional probability of the consequent R , on the

supposition that the antecedent Q is true (so long as the probability of Q isn't equal to zero) – and the Equation holds true.

3.1. The Equation

According to the Equation, the probability of an indicative conditional $Q \rightarrow R$ is equal to the conditional probability of the consequent R , on the supposition that the antecedent Q is true so long as the probability of Q isn't equal to zero. Formally, writing ' P ' for probability and ' $|$ ' for given, we get the following:

$$P(Q \rightarrow R) = P(R | Q) \text{ provided that } P(Q) \neq 0.$$

So, according to the Equation, the probability of the conditional 'If Josie's at home, then she's in New York' is equal to the conditional probability of 'Josie's in New York' given 'she's at home'.

Two things to note at this point:

- (i) logicians disagree on the range of Q and R in the Equation. Some take Q and R to range over only propositions. Adams (1975) and McGee (1989) are cases in point. Others take Q and R to range over propositions and non-propositions (i.e. sentences which don't necessarily have a truth value). (McGee allows right- but not left-nested conditionals (Hájek, 2012, pp. 150-1) and I'll side with him.) This disagreement is an important matter, as we'll see when deriving the Triviality Result below and when defining *modus ponens* and *modus tollens* in Chapters 2 and 3; and
- (ii) the Equation is different from Adams' Thesis: while the Equation equates probability of a conditional with conditional probability, Adams' Thesis equates assertibility with conditional probability. In Chapter 2 (subsection 4.1), we'll see the difference prevent the hybrid theorist from adopting as it is

a proof by the suppositional theorist. The proof relies on the Equation which, unlike the suppositional one, the hybrid theorist doesn't accept. Indeed, the difference will require the hybrid theorist to *adapt* the proof.

Now, to calculate the conditional probability of the consequent, given the antecedent, we rely on the Ratio Formula.

3.2. *The Ratio Formula*

According to the Ratio Formula, for any proposition Q and R , the conditional probability of R given Q equals the probability of the conjunction of Q and R divided by the probability of Q so long as the probability of Q isn't equal to zero (Hájek 273). Formally, we get

$$P(R | Q) = P(Q \wedge R) / P(Q) \text{ provided that } P(Q) \neq 0.$$

So, the conditional probability of 'Josie's in New York' given 'she's at home' is equal to the probability of 'Josie's at home and she's in New York' divided by the probability of 'Josie's in New York'.

3.3. *Returning to False-true & co and evaluating the theory*

The suppositional theorist can account for the conditionals we've been considering. According to her, we aren't prepared to assert *False-true* because, assuming that Angela Merkel is the Prime Minister of Singapore, there won't be a high probability that she's the Chancellor of Germany. Among the facts of the world is the following one: Singapore and Germany are two different countries and no person is simultaneously the head of state in both.

We aren't prepared to assert *False-false* because assuming that Vienna is the capital of England, there won't be a high probability that Vienna is the

capital of Switzerland. Also among the facts of the world is the following one: England and Switzerland are two countries and no city is simultaneously the capital of both.

The suppositional theory can account for *Taller* the same way in which it accounts for other conditionals. We'd evaluate the consequent of *Taller* on the *supposition* of the antecedent.

Finally, consistent with our intuition, we wouldn't be prepared to assert *Les compatriotes* or *I compatrioti*. Assuming François and Italo are compatriots, there wouldn't be a high probability that they were both French, or in the other case, that they were both Italian.

The suppositional theorist does a good job explaining why we would or wouldn't assert or could evaluate *False-true* and the other conditionals we've considered so far. However, the theory makes a serious departure from the notions of truth and validity of classical logic – which some might view as a demerit – and without that departure, the Triviality Result follows (Lewis, 1976).

3.4. *The Triviality Result*

The Equation in part implies that a conditional is non-propositional. Indeed, the Equation plus the assumption that a conditional is propositional would lead to absurdity: they imply that no change in any proposition's probability affects another's. For example, an increase in the probability that Roger Federer will win the Wimbledon final against Rafael Nadal (say, Roger's serving match point, leading six games to four in the third set and two sets to zero) doesn't

affect the original probability that Rafael *will* win. Logicians call such a scenario, where no change affects any change, trivial.

Specifically, the scenario arises from the fact that conditionals could embed everywhere and the probability of any conditional would be the probability of its consequent. For any sentence ‘Roger will win’, we could create a conditional ‘If x , then Roger will win’ and no matter what we choose for x , the probability of the conditional will be that of the consequent ‘Roger will win’. x could be ‘Roger will lose’ and that wouldn’t affect the probability of the conditional. Absurdly, it would remain that of ‘Roger will win’.

Now, to derive the Triviality Result, apart from relying on the Ratio Formula, the Equation and Conditionalisation (which we’ll define), we rely on the following axioms and theorems:

Axioms

A1. $P(A) \geq 0$

A2. If T is a tautology, then $P(T) = 1$

A3. $P(A \vee B) = P(A) + P(B)$ if A and B are inconsistent

Theorems

T1. If C is a contradiction, then $P(C) = 0$

T2. $P(A) = P(B)$ if A and B are equivalent

T3. $P(A \wedge B) = P(A | B) \times P(B)$ provided that $P(B) \neq 0$

We begin by proving a lemma – which will serve us not only in this proof but later.

3.4.1. Import-Export

We call Import-Export the rule of inference according to which we can derive $Q \rightarrow (R \rightarrow S)$ from $(Q \wedge R) \rightarrow S$ and vice versa. We'll also call Import-Export the probabilistic equivalence between the probabilities of those two sentences: $P(Q \rightarrow (R \rightarrow S)) = P((Q \wedge R) \rightarrow S)$. Before we derive this equivalence, we note that Import-Export will prove important in this thesis.

In Chapter 2, for example, the possible-worlds theorist will appeal to her rejection of Import-Export to reject in turn the proposition according to which McGee's argument is an instance of *modus ponens*. While McGee formalises the first premise as having the form $Q \rightarrow (R \rightarrow S)$, the possible-worlds theorist would formalise it as having the form $(Q \wedge R) \rightarrow S$ and, citing the invalidity of Import-Export according to her theory (i.e. rejecting that one can derive the one relevant conditional from the other), she would deny that McGee's argument is counterexample to *modus ponens*.

Also in that chapter, as in Chapter 3, the suppositional theorist will appeal to her acceptance of Import-Export to calculate the probability of the first premise in McGee and Carroll's respective arguments. Contra Adams (1975, pp. 30-3), I show that she can evaluate the probability of a conditional with a conditional as consequent: by evaluating the probability of the equivalent conditional with a conjunction as antecedent.

McGee notes that his argument suggests it's not the case *modus ponens* and Import-Export can both be valid where the conditional connective is stronger than \supset , as we'd expect \rightarrow to be (McGee, 1985, pp. 465-6). While we find both valid where the connective is the material conditional, '[w]e have

explicit examples to show that the indicative conditional does not satisfy modus ponens', McGee writes (1985, p. 466). I'll dedicate Chapter 2 to examining one of the examples he offers.

Here, first, to prove the probabilistic equivalence of $Q \rightarrow (R \rightarrow S)$ and $(Q \wedge R) \rightarrow S$, we rely on not only the Equation but also Conditionalisation.

3.4.1.1. Conditionalisation

According to Conditionalisation, for any formulae Q and R , the probability of R after we find out that Q is equal to the conditional probability of R given Q . Formally, we get the following equation, where $P_Q(R)$ is the posterior probability of R after we discover Q (Lewis, 1976, p. p. 299):

$$P_Q(R) = P(R | Q)$$

So, according to Conditionalisation, the probability of 'Josie's in New York' after we find out that 'she's at home' is the conditional probability of 'she's in New York' given 'Josie's at home'.

Now, assuming $P(Q \wedge R) \neq 0$ and using the Equation and Conditionalisation, we can prove that $Q \rightarrow (R \rightarrow S)$ and $(Q \wedge R) \rightarrow S$ are probabilistically equivalent (see Alan Hájek's proof in Bennett, 2003, p. 62):

1. $P(Q \rightarrow (R \rightarrow S)) = P((R \rightarrow S) | Q)$ by the Equation
2. $= P_Q(R \rightarrow S)$ from 1 by Conditionalisation
3. $= P_Q(S | R)$ from 2 by the Equation
4. $= P_Q(S \wedge R) \div P_Q(R)$ from 3 by the Ratio Formula
5. $= P((S \wedge R) | Q) \div P(R | Q)$ from 4 by Conditionalisation
6. $= [P(S \wedge R \wedge Q) \div P(Q)] \div [P(R \wedge Q) \div P(Q)]$
from 5 by the Ratio Formula

$$7. = P(S \wedge R \wedge Q) \div P(R \wedge Q) \quad \text{from 6 by cancellation of } P(Q)$$

$$8. = P(S \mid (Q \wedge R)) \quad \text{from 7 by the Ratio Formula}$$

$$9. = P((Q \wedge R) \rightarrow S) \quad \text{from 8 by the Equation}$$

Taking it from top to bottom and still assuming $P(Q \wedge R) \neq 0$,

$$10. P(Q \rightarrow (R \rightarrow S)) = P((Q \wedge R) \rightarrow S)$$

from 1 and 9 by transitivity of
identity

which is Import-Export.

The suppositional theorist welcomes this result. The equivalence allows her to evaluate the probability of a conditional whose consequent is itself a conditional. Without the result, she couldn't as she can evaluate only a conditional whose antecedent and consequent are propositional, and a conditional isn't propositional.

Here, assuming the suppositional theorist can derive the proof, I disagree with Adams (1975). He claims that on the suppositional theory we can't evaluate conditionals whose consequents are conditionals (Adams, 1975, pp. 30-3). This proof gives us a means to do so.

Next, we turn to prove that for any conditional $Q \rightarrow R$, the probability of that conditional is equivalent to the probability of the consequent: the Triviality Result proper.

3.4.2. Triviality

Proof (from Bennett, 2003, p. 63. Bennet himself takes the proof from Blackburn, 1986, pp. 218-20):

$$11. P(R \mid Q) = P(Q \rightarrow R) \quad \text{by the Equation}$$

$$12. = P((R \wedge (Q \rightarrow R)) \vee (\neg R \wedge (Q \rightarrow R)))$$

from 11 by T2

$$13. = P(R \wedge (Q \rightarrow R)) + P(\neg R \wedge (Q \rightarrow R))$$

from 12 by A3

$$14. = [P((Q \rightarrow R) | R) \times P(R)] + [P((Q \rightarrow R) | \neg R) \times P(\neg R)]$$

from 13 by T3

$$15. = [P(R \rightarrow (Q \rightarrow R)) \times P(R)] + [P(\neg R \rightarrow (Q \rightarrow R)) \times P(\neg R)]$$

from 14 by the Equation

$$16. = [P((R \wedge Q) \rightarrow R) \times P(R)] + [P((\neg R \wedge Q) \rightarrow R) \times P(\neg R)]$$

from 15 by Import-Export

$$17. = [P(R | (R \wedge Q)) \times P(R)] + [P(R | (\neg R \wedge Q)) \times P(\neg R)]$$

from 16 by the Equation

$$18. = [(P(R \wedge R \wedge Q) \div P(R \wedge Q)) \times P(R)] + [(P(R \wedge \neg R \wedge Q) \div P(\neg R \wedge Q)) \times P(\neg R)]$$

from 17 by the Ratio Formula

$$19. = [1 \times P(R)] + [0 \times P(\neg R)]$$

from 18 by algebra, T1 and T2

$$20. = P(R)$$

from 19 by algebra

3.4.3. Avoiding the result

Now, to avoid the Triviality Result while still holding onto the Equation, the suppositional theorist has various options, some more attractive than others. One option is to reject 1. She might say the Equation doesn't hold where Q and R range over non-propositions and here there's nothing preventing Q or R from being indicative conditionals.

Another option is to reject the move from 11 to 12. This step seems to assume that a conditional is propositional and the suppositional theorist rejects that assumption. The step assumes that the probability theorem T2 applies but, for the theorem to apply, the conditionals would have to be logically equivalent and in 11 they're merely probabilistically so, meaning the theorem doesn't apply.

A third option to avoid the Triviality Result is to reject the move from 12 to 13. This step also seems to assume that a conditional is propositional. The step assumes that the probability axiom A3 applies but, for the axiom to apply, the conditionals would have to be logically inconsistent and here they're merely probabilistically so, meaning the axiom doesn't apply. To be logically inconsistent, the conditionals would have to have truth values and, according to the suppositional theorist, they don't.

Here, the second and third options are more attractive than the first. While rejecting 1 prevents the suppositional theorist from deriving the proof of Import-Export, rejecting the move from 11 to 12 or 12 to 13 doesn't – and Import-Export will serve the suppositional theorist later, as we'll see in Chapters 2 and 3, when she tries to explain our intuitions in the case of apparent counterexamples to *modus ponens* and *modus tollens*.

Either way, according to the suppositional theorist, two propositions don't combine into a single proposition we judge as probably true when we judge the second to be probably true on the supposition of the first (Edgington, 1995, p. 305). So, according to the suppositional theory it's not the case that 'If Josie's at home, then she's in New York' can be true or false.

Likewise, according to the suppositional theorist, it's not the case that classical validity is relevant when we're talking about arguments with indicative conditionals. Classical validity concerns itself with truth-preservation and here we're taking indicative conditionals to be non-propositional. Rather, when we're talking about arguments with indicative conditionals, probabilistic validity is relevant.

To define probabilistic validity, we define uncertainty. We define the uncertainty of a formula Q as one minus the probability of Q (Adams, 1975, p. 3). Formally, writing ' U ' for uncertainty, we get

$$U(Q) = 1 - P(Q)$$

Using this definition of uncertainty, we define an argument as probabilistically valid when the uncertainty of the conclusion can't exceed the sum of the uncertainties of the premises (Bennett, 2003, p. 129; Adams, 1975, pp. 1-2). In contrast to classical validity which bars falsity from entering 'along the way from the [true premises] to the conclusion', probabilistic validity bars improbability from entering along the way (Bennett, 2003, p. 129).

The departure from notions of truth and validity we know in classical logic will prove helpful for the theory when explaining the intuitions we might have when it comes to the Election and Barbershop paradoxes, as we'll see in Chapters 2 and 3. Still, some might view the departure as a demerit for the theory, given the consequences we've seen above.

4. The hybrid theory

According to the hybrid theorist, the Equivalence Thesis and Adams' Thesis both hold true. The hybrid theory combines elements from the material theory

and the suppositional theory. On the one hand, like the material theory, the hybrid one concerns itself with truth. On the other hand, like the suppositional one, the hybrid theory concerns itself with probability. For example, an indicative conditional is true where the corresponding material one is; and it's assertible where the conditional probability of the consequent given the antecedent is high. The hybrid theorist speaks of an assertible conditional as being robust: it's such that we wouldn't abandon belief in it upon learning that its antecedent is true (Jackson, 1979, pp. 569-70).

4.1. Assertibility

Note that the material theory's concept of assertability with an 'a' and the hybrid theory's concept of assertibility with an 'i' aren't the same. The difference in vowels reflects a difference in meaning. The assert-a-bility of a sentence depends on 'local' factors: 'how important and relevant is the information to present concerns, is the information already widely known, and so on and so forth?' (Jackson, 1987, p. 11). In contrast, the assert-i-bility of a sentence depends on factors 'governing when it is justified or warranted – in the epistemological sense, not in a purely pragmatic one – to assert it, or, as this comes in degrees, to what extent it is justified to assert it under the circumstances' (Jackson, 1987, p. 8).⁵

While assertibility and probability of truth go hand in hand for many sentences, they don't for many others, notably indicative conditionals. Let's

⁵ For those familiar with the Gricean contrast between conversational implicature and conventional implicature, I'll add that assertability is related to conversational implicature while assertibility is related to conventional implicature. So here in the hybrid theory subsection, I'm talking about conventional implicature.

take a simpler example – one whose main connective is not a conditional but a conjunction – first. Consider the following sentence:

Ming Bin is a good student even though he's intelligent.

The sentence has a low assertibility but a high probability: it has a high probability if and only if the conjunction 'Ming Bin is a good student and is intelligent' has a high probability – and arguably, it does. The sentence has a low assertibility, however, given the presence of 'even though'. The phrase places an 'additional [assertibility] requirement' on the sentence, over and above the probability requirement – and the additional requirement isn't satisfied here (Jackson, 1987, p. 60).

As in the example, in the case of indicative conditionals, assertibility doesn't go hand in hand with probability. The assertibility of an indicative conditional depends on the *conditional* probability of its consequent given its antecedent (Jackson, 1987, p. 11). Formally, we get Adams' Thesis, which we saw at the very beginning and according to which the assertibility of $Q \rightarrow R$ is equal to the conditional probability of R given Q so long as the probability of Q isn't equal to zero.

If this weren't the case (i.e. if the assertibility of indicative conditionals *did* go hand in hand with probability), then the assertibility of 'If it's not the case that it's raining, then it's raining' would be whatever the probability of rain were, since the probability of the conditional would be the probability of the disjunction 'It's raining or it's raining' which in turn would be the probability of the disjunct 'it's raining'.

As it is, the assertibility of the conditional 'If it's not the case that it's raining, then it's raining' is zero. The conditional probability of 'it's raining'

given ‘it’s not the case that it’s raining’ is zero, since according to the Ratio Formula, the probability of the consequent given the antecedent is equal to the probability of ‘it’s not the case that it’s raining and it’s raining’ divided by ‘it’s raining’. Here, since the numerator is the probability of a contradiction, it has a probability of zero and, regardless of the denominator, the quotient (i.e. the conditional probability of ‘it’s raining’ given ‘it’s not the case that it’s raining’) will be zero.

Note further that, in certain contexts, a conditional might be assert-ible but not assert-a-ble. In a silent reading room, ‘If my birth certificate is correct, Jill is my mother’ is a case in point. Granted, generally, it might make sense for me to say it because there’s a high probability that Jill is my mother, given that my birth certificate is correct. However, it wouldn’t make sense for me to say it in a room where I’m not permitted to speak (Jackson, 1987, pp. 10-1).

4.2. Returning to False-true & co and evaluating the theory

According to the hybrid theory, we can dismiss both *False-true* and *False-false* on the grounds that they aren’t robust. We would abandon belief in *False-true* upon learning that its antecedent is true. Supposing it true that Angela Merkel is the Prime Minister of Singapore, we would not believe that she is also the Chancellor of Germany. Singapore and Germany are two different countries and one person is not simultaneously the head of state in both. Likewise, we’d abandon belief in *False-false* upon learning that its antecedent is true. Supposing it true that Vienna is the capital of England, we would not believe that Vienna is also the capital of Switzerland. England and

Switzerland are two different countries and one city is not simultaneously the capital of both.

As to *Taller*, *Les compatriotes* and *I compatrioti*, the hybrid theorist would evaluate them in the same way as a suppositional one.

Among the merits of the theory are the following two: first, of course, it can explain the sentences; and second, it aligns with the truth-conditional theory of meaning. Indeed, it doesn't face the problem the suppositional theory did when it comes to aligning with the truth-conditional theory of meaning. According to the hybrid theory, an indicative conditional can be true – and so, on the truth conditional-theory of meaning, it *is* possible for us to know its meaning. We need only know the conditions under which the conditional is true.

As to the demerits, there's the following one: it requires the seemingly ad hoc adoption of rules to account for Import-Export preserving assertibility (i.e. the idea that the assertibility of $Q \rightarrow (R \rightarrow S)$ is equivalent to the assertibility of $(Q \wedge R) \rightarrow S$). I discuss this as the matter arises in Chapters 2 and 3.

5. Concluding

In this chapter, I outlined the topic of the thesis: theories of the indicative conditional and apparent counterexamples to classically valid argument forms; and I considered in turn four possible positions when it comes to accepting or rejecting the Equivalence Thesis or Adams' Thesis, each position corresponding with a theory of the indicative conditional: the material; the possible-worlds; the suppositional; and the hybrid theory.

Along the way, I disagreed with Adams (1975). While he claims that on the suppositional theory, we can't evaluate an indicative conditional whose consequent is a conditional, I included a proof that enables us to.

Next, we turn to Chapter 2.

CHAPTER 2

THE ELECTION PARADOX

So far, we've seen various theories of the indicative conditional coming under four headings: the material, possible-worlds, suppositional and hybrid theories. Next, we turn to apply those theories to an alleged counterexample to an apparently valid argument form. In this chapter, I proceed largely from first principles in analysing Vann McGee's alleged counterexample to *modus ponens* (1985). I focus on responses the theories would offer to the trilemma McGee's argument presents.

The counterexample takes as context the 1980 US presidential election. Shortly before voting day, opinion polls reveal five things:

- (i) the Republican Ronald Reagan is in first place;
- (ii) the Democrat Jimmy Carter is in second;
- (iii) another Republican John Anderson is in third;
- (iv) Reagan is substantially ahead of Carter; and
- (v) Carter is substantially ahead of Anderson.

Intuitively, the following two premises seem true:

If a Republican wins the election, then if it's not Reagan who wins it will be Anderson; and

A Republican will win the election.

The first seems true inasmuch as there are two Republicans in the race; and, logically, if a Republican wins and the one Republican loses, the other one will win. Moreover, the second premise seems true inasmuch as the polls

show the Republican Reagan in first place and far ahead of Democrat Carter in second.

Also intuitively, however, the conclusion seems false even though it follows by *modus ponens* from the two premises:

If it's not Reagan who wins, it will be Anderson (McGee, 1985, p. 462).

The conclusion seems false inasmuch as the polls show Carter in second position – not Anderson. If Reagan loses, then presumably Carter, in second position, would win.

This appears to be a counterexample to *modus ponens*. Writing *A* for 'a Republican wins', *B* for 'Reagan doesn't win', *C* for 'Anderson wins', ' \rightarrow ' for the indicative conditional connective and ' \therefore ' for 'therefore', we get the following argument⁶:

$$A \rightarrow (B \rightarrow C)$$
$$A$$
$$\therefore B \rightarrow C$$

The premises seem true and the second premise *A* is the antecedent of the first premise $A \rightarrow (B \rightarrow C)$. The conclusion $B \rightarrow C$, however, seems false even though it's the consequent of the first premise. Indeed, there are two possible combinations if Reagan doesn't win and between those two, it's not necessarily the case that it will be Anderson:

⁶ Here, I'm departing from McGee's original wording. I'm saying e.g. the consequent of the conclusion is 'Anderson wins' rather than 'it will be Anderson'. I *don't* depart from the wording everywhere henceforth and, even if I did, I don't think the departure has any significant implication except, perhaps for the reader, ease of understanding.

	Reagan	Carter	Anderson
1	Loses	Wins	Loses
2	Loses	Loses	Wins

Table 2: Possible combinations for who, between Carter and Anderson, might win if Reagan doesn't

It's possible – and, given the results of the poll, likely – that if it's not Reagan who wins, it will be Carter.

I'm not alone in offering a discussion of McGee. Walter Sinnott-Armstrong, James Moor and Robert Fogelin (1986) (later 'the Dartmouth group'), E.J. Lowe (1987), Christian Piller (1996), Bernard D. Katz (1999), Joseph S. Fulda (2010) and most recently Justin Bledin (2015) do so too. I'll cite or discuss some where relevant. That said, I won't discuss Bledin's positive view, for example, as his informational theory of logic doesn't fit within the theories of conditionals I'm considering.

In this chapter, I refute extant solutions: while the Dartmouth group thinks the material theory can solve the paradox and Edgington thinks the suppositional one can solve it and Jackson thinks the hybrid one can, I'll show that their solutions fall short of being comprehensive.

I proceed by considering McGee's argument in light of four theories of the indicative conditional: the material (inspired by Grice 1961, 1975 and 1989), possible-worlds (inspired by Stalnaker (1981) and Lewis (1976)), suppositional (inspired by Adams (1975) and Edgington (1995 and 2014)) and hybrid (inspired by Jackson (1987)). Specifically, I examine how the theories would solve an inconsistent triad. The three individually plausible but jointly inconsistent theses are the following:

#1 the premises are true and the conclusion is false;

#2 the argument is an instance of *modus ponens*, i.e. is one with the form $Q \rightarrow R; Q; \therefore R$, where one has replaced the letters with propositions; and

#3 *modus ponens* is valid, i.e. in any argument with the form $Q \rightarrow R; Q; \therefore R$, the truth of the premises guarantees the truth of the conclusion.

As we'll see, not all theories can explain the plausibility of all three theses at once – and to solve the trilemma they must. Moreover, the theories which can explain all three theses face problems of their own.

In section 1, I present the material theory and show that it rejects #1 but can't explain the plausibility of the thesis without undermining itself. In section 2, I show that the possible-worlds theory rejects #1 too but can't explain its plausibility. In section 3, I present the suppositional theory and show that it rejects all three theses but can explain at most the plausibility of two. In section 4, I present the hybrid theory and show that while it rejects #1, it can explain the plausibility of the thesis and while the theory accepts #2 and #3, it fails in its original mission. In section 5, I conclude.

1. The material theory

Recall that according to the material theorist, an indicative conditional $Q \rightarrow R$ is true if and only if the corresponding material conditional $Q \supset R$ is true. In other words, 'If Q , then R ' is true if and only if Q is false or R is true. For example, let Q be 'Ben's on sabbatical' and R be 'he's in Australia'. The indicative conditional 'If Ben's on sabbatical, then he's in Australia' is true if

and only if it's not the case that Ben's on sabbatical or it's the case that he's in Australia.

In this section, I show how the material theory accepts Theses #2 as well as #3 but rejects #1 and can explain its plausibility. I show also that in explaining the plausibility of #1, the material theory deals itself a blow: it finds unassertable a sentence we might assert.

1.1. Accepting #2 and #3 but rejecting #1

According to the material theorist, the following holds true:

- the argument is indeed an instance of *modus ponens*. We can reduce it to the form $Q \rightarrow R; Q; \therefore R$;
- *modus ponens* is valid. It's impossible for the premises to be true and the conclusion false; and
- both premises are true – and so is the conclusion. Given the results of the poll, the second premise is true. So is the conclusion inasmuch as the antecedent of the conditional is false. And the first premise is true inasmuch as the consequent of the conditional is true: it's the conclusion and, as we've just seen, the conclusion is true.

The material theorist can explain nonetheless the plausibility of Thesis #1 by appealing to the concept of assertability. Recall from Chapter 1 that according to the Cooperative Principle, some sentences are true but not assertable. Given a certain scenario, 'You won't eat those and live' is a case in point (Lewis, 1976, p. 306). Imagine I say the sentence while pointing at some non-toxic mushrooms and you, deferring to my apparent mycological

knowledge, refrain from eating the mushrooms. I told no lie. Formally, the sentence is true. The sentence is a negated conjunction and one of the conjuncts (viz. 'you eat those') is false. Indeed, it's not the case that you eat the mushrooms. Nonetheless, the original sentence isn't assertable. It wouldn't be a cooperative thing to say.

Looking at McGee's argument, a material theorist would argue we take the conclusion to be false because it's unassertable. Again, it wouldn't be a cooperative thing to say. According to a maxim of cooperative communication, one must 'Assert the stronger rather than the weaker' and here I'd be asserting a weaker sentence when I could be asserting a stronger one (Grice, 1961, p. 132). If I'm asserting the conditional because I know it has a false antecedent, I might as well assert the negation of the antecedent. Here, 'It's Reagan who wins' is a stronger statement. 'It's Reagan who wins' implies the conditional while 'If it's not Reagan who wins, it will be Anderson' doesn't imply the negation of the antecedent.

Moreover, the material theorist's solution to a paradox of material implication commits her to saying that the conclusion is unassertable. The paradox arises from the fact that, according to the material theory, the following argument is valid: $\neg Q; \therefore Q \rightarrow R$ – yet this contradicts our intuition. For example, where Q is 'I live until 120' and R is 'I die at 36', we get

It's not the case that I live until 120.

Therefore, if I live until 120, then I die at 36.

Here, we might accept the premise but reject the conclusion, even though it follows by a valid rule of inference. The material theorist solves the paradox

by saying the conclusion isn't assertable. When it comes to the truth values of the antecedent and consequent, the conclusion corresponds to the *False-true* conditional we saw in Chapter 1:

If Angela Merkel is the Prime Minister of Singapore, then Angela Merkel is the Chancellor of Germany.

While *False-true* and the conclusion – ‘If I live until 120, then I die at 36’ – are true, we'd only assert them because we knew that they had a false antecedent. In each case, we might as well assert the negation of the antecedent. So too in the case of the conclusion of McGee's argument.

The Dartmouth group notes that the conclusion might be true while unassertable but fails to note the following: unfortunately for the material theorist, her solution to another paradox of material implication commits her to saying the first premise isn't assertable either (Sinnott-Armstrong et al., 1986, p. 296). This paradox arises from the fact that, according to the material theory, the following argument is valid: $R; \therefore Q \rightarrow R$ – yet this contradicts our intuition. For example, keeping Q as ‘I live until 120’ and R as ‘I die at 36’, we get the following argument:

I die at 36.

Therefore, if I live until 120, then I die at 36.

Here, we might accept the premise but reject the conclusion, even though it follows by a valid rule of inference. The material theorist solves the paradox by saying the conclusion isn't assertable. Asserting it would mean flouting the ‘Assert the stronger’ maxim (as we saw in Chapter 1, the maxim's wording comes from Jackson, 1979, p. 566). While the conditional is true, we'd assert

it only because we knew it had a true consequent. In this case, we must assert just the consequent.

Similarly, in asserting the first premise of McGee's argument, we'd be flouting that maxim. The whole conditional is true but the consequent is stronger than the whole conditional. 'If it's not Reagan who wins, it will be Anderson' implies 'If a Republican wins the election, then if it's not Reagan who wins it will be Anderson'. 'If it's not Reagan who wins it will be Anderson' is true and thus implies any conditional which takes it as consequent.

Moreover, the negation of the antecedent of the embedded conditional is stronger than the whole embedded conditional. 'It's not the case that it's not Reagan who wins' or, when we eliminate the double negation, 'It's Reagan who wins' is true and thus its negation, the original 'It's not Reagan who wins', implies any conditional which takes it as antecedent. So, in the case of the first premise and to abide by the maxim of cooperative communication, we'd have to assert the negation of the antecedent of the embedded conditional. Asserting it would be asserting something stronger than not only the embedded conditional but the whole conditional.

(Note that the other maxim we saw in Chapter 1 isn't relevant here. In asserting the premises, we wouldn't be flouting the 'Don't assert what you don't believe' maxim. Given the results of the poll, we might believe that 'If a Republican wins the election, then if it's not Reagan who wins it will be Anderson' and 'If it's not Reagan who wins, it will be Anderson'.)

We see that the material theorist rejects the thesis according to which the conclusion is false while explaining its plausibility. Yet, in introducing the

concept of assertability to explain the plausibility, the theory faces new problems: namely, it finds unassertable a sentence we might assert.

2. *The possible-worlds theory*

Recall that according to the possible-worlds theorist, an indicative conditional ‘If Q , then R ’ is true in a possible world w if and only if R is true in all Q -worlds which ‘differ minimally’ from w (and vacuously true when there’s no Q -world) (Stalnaker, 1981, pp. 46-7; see also Lewis, 1973).

A Q -world is one in which the antecedent Q is true. For example, the conditional ‘If Nicholas is at the office, then he’s in London’ is true in the actual world just in case Nicholas is in London in all most-similar worlds in which Nicholas is at the office.

In this section, I show how the possible-worlds theory accepts Theses #2 and #3 and rejects #1 but can’t explain its plausibility.

2.1. *Accepting Thesis #3: modus ponens is valid*

According to the possible-worlds theorist, *modus ponens* is valid. It’s impossible for the premises to be true while the conclusion is false. Suppose that the premises are true and the conclusion is false in the actual world. This means that in all most-similar worlds in which the antecedent of the first premise is true, the consequent of the first premise is also true. Here, since the second premise is true and the second premise is the antecedent of the first premise, the most-similar world in which the antecedent of the first premise is true is the actual world.

This leads us, however, to a contradiction. Since in the actual world the consequent of the first premise is the conclusion and the conclusion is false, the consequent of the first premise is also false. So, the first premise is both true and false. We see that, according to the possible-worlds theory, in the case of *modus ponens*, it's impossible for the premises to be true while the conclusion is false.

2.2. Accepting Thesis #2: the argument is an instance of *modus ponens*

Here, the possible-worlds theorist agrees with the material theorist. Yes, the argument as we've formalised it is an instance of *modus ponens*. Having said that, the possible-worlds theorist might argue that, in fact, the argument has another form, namely

$$(A \wedge B) \rightarrow C$$

$$A$$

$$\therefore B \rightarrow C$$

We can't reduce this to an instance of *modus ponens*. According to the possible-worlds theorist, Import-Export isn't valid. Indeed, from $(A \wedge B) \rightarrow C$, we can't necessarily derive $A \rightarrow (B \rightarrow C)$ – and to reduce the new argument form to an instance of *modus ponens*, we'd need to.

Let A be true in the actual world. Let A and C be false in all most-similar B -worlds. And let C be true in all most-similar $(A \wedge B)$ -worlds. $(A \wedge B) \rightarrow C$ would come out as true in the actual world. Since in all most-similar $(A \wedge B)$ -worlds C is true, the conditional $(A \wedge B) \rightarrow C$ is true. $A \rightarrow (B \rightarrow C)$, however, would come out as false in the actual world. Since a most-similar A -

world is the actual world and, in all most-similar B -worlds C is false, the whole conditional $A \rightarrow (B \rightarrow C)$ has a true antecedent and a false consequent.

Moreover, this new argument with a conjunction embedded in the antecedent of the first premise is invalid according to the possible-worlds theory. We can build a counterexample. We want in the actual world the premises to be true but its conclusion to be false. For the first premise to be true in the actual world, it must be the case that, in all most similar $(A \wedge B)$ -worlds, C is true. For the second premise to be true, it must be the case that, in the actual world, A is true. And for the conclusion to be false, it must be the case that, in the most-similar B -world, C is false.

Using the same truth-values as those we used above to show that $(A \wedge B) \rightarrow C$ and $A \rightarrow (B \rightarrow C)$ weren't logically equivalent, we get a counterexample. The first premise is true. In all most-similar $(A \wedge B)$ -worlds, C is true. The second premise is true. A is true in the actual world. The conclusion, however, is false. In the most-similar B -world, C is false.

If the possible-worlds theorist can successfully argue the first premise in McGee's argument has an embedded conjunction in the antecedent, she can reject #1 yet explain its plausibility – it appears to be a *modus ponens* because it appears to have an embedded conditional in the consequent. But short of providing a positive reason for reformulating the premise, she can't explain the plausibility of #1.

2.3. Rejecting Thesis #1: the premises are true and the conclusion is false

According to the possible-worlds theorist, the conclusion is false. The conditional 'If it's not Reagan who wins, it will be Anderson' is true in the

actual world just in case it'll be Anderson who wins in all most-similar worlds in which it's not Reagan who does. In all most-similar worlds in which it's not Reagan who wins, however, it's Carter who does. He was in second place. With a false consequent in a most-similar antecedent-world, the conclusion is false.

According to the possible-worlds theory, however, it's not the case the premises are true. While the second premise (viz. 'A Republican will win the election') is true, the first isn't. The conditional 'If a Republican wins the election, then if it's not Reagan who wins it will be Anderson' is true in the actual world just in case if it's not Reagan who wins it'll be Anderson in all most-similar worlds in which a Republican wins the election.

Now, the most-similar world in which a Republican wins the election is the actual world itself. However, in that world, it's not the case that if it's not Reagan who wins it'll be Anderson. In a most-similar antecedent-world to the actual world in which it's not Reagan who wins, it'll be Carter who does. So, with a false consequent, the first premise is false. It's not the case that both premises are true (McGee, 1985, pp. 466-7).

As Thesis #1 stands, the possible-worlds theory can't explain its plausibility. The theory can't, for example, appeal to the concept of assertability as the material theory did. That concept explained the situations where sentences were true but not assertable. Here, in contrast, we have a sentence which the theory finds to be false yet nonetheless we'd assert.

3. The suppositional theory

Recall that according to the suppositional theorist such as Adams and Edgington, conditionals are things we evaluate in terms of probabilities. Indeed, according to her, the probability of a conditional $Q \rightarrow R$ is equal to the conditional probability of the consequent R , on the supposition that the antecedent Q is true so long as the probability of Q isn't equal to zero. In Chapter 1, we called this the Equation.

Since the suppositional theorist takes the indicative conditional to be non-propositional, it's clear that she'll reject Theses #1, #2 and #3. They imply that conditionals are propositional. She can, however, explain the plausibility of #1 using the definition of Import-Export.

In this section, I show how the suppositional theorist does so and why she faces a dilemma between explaining #2 and #3.

3.1. Rejecting Thesis #1: the premises are true and the conclusion is false

The suppositional theorist would explain the plausibility of #1 by saying the uncertainty of the conclusion exceeds the sum of the uncertainties of the premises.

To know the uncertainty of the premises, we need to know their probabilities. To know the probability of the first premise, we calculate the probability of 'If a Republican wins *and* it's not Reagan who does it will be Anderson'. As we saw in Chapter 1, this formula is probabilistically equivalent to the first premise, where two formulae are probabilistically equivalent if and only if necessarily they have the same probability. Formally, we get

$$P(A \rightarrow (B \rightarrow C)) = P((A \wedge B) \rightarrow C)$$

This equivalence, which in Chapter 1 we called Import-Export, enables us to calculate the probability of a conditional whose consequent is also a conditional. Using it to calculate the probability of the first premise, we can see that McGee's argument isn't probabilistically valid. Let's look at a possible probability assignment, calculate the probability of the first premise and conclusion (we need not calculate that of the second premise as its probability will be whatever figure we assign A) before calculating the uncertainty of the premises and conclusion.

We know there's a high probability that a Republican will win the election so let $P(A) = 0.80$. We know there's a low probability that it's not Reagan who wins so let $P(B) = 0.24$. We know there's also a low probability that it'll be Anderson who wins. Moreover, we know this probability is even lower than the probability that it's not Reagan who wins because if it's not Reagan who wins it'll be – not Anderson but – Carter so let $P(C) = 0.04$. We know the probability that a Republican wins *and* it's not Reagan is low and is close to the probability that a Republican wins and it's not Reagan but Anderson. The probability of a Republican other than Reagan winning is close to zero so let $P(A \wedge B) = 0.04$ and $P(A \wedge B \wedge C) = 0.04$ as well. Finally, we know the probability of it being not Reagan who wins but Anderson is also close to zero so let $P(B \wedge C) = 0.04$.

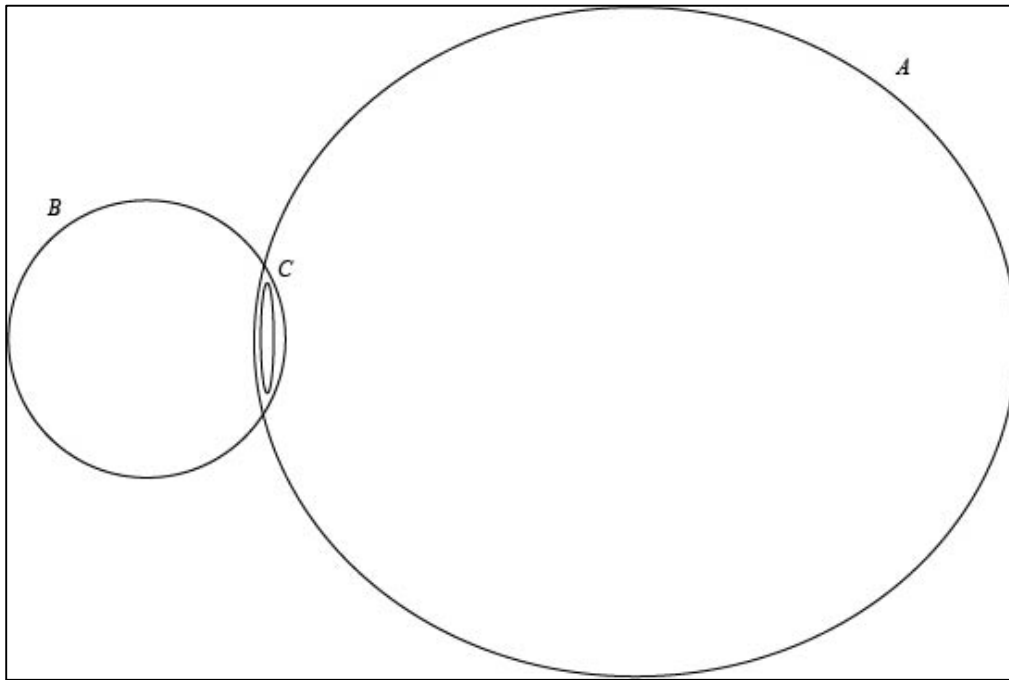


Figure 1: Venn diagram of the Election Paradox

If I'm letting $P(C)$, $P(A \wedge B)$, $P(A \wedge B \wedge C)$ $P(B \wedge C)$ each to be equal to 0.04, it's because I'm taking the probability of a Republican other than Reagan or Anderson winning to be zero – and so I'm taking 'Anderson wins', 'A Republican wins and it's not Reagan', 'A Republican wins and it's not Reagan but it's Anderson' and 'It's not Reagan who wins but Anderson wins' to be the same.

Now, turning to the first premise: from Import-Export and the Equation, we know that $P(A \rightarrow (B \rightarrow C))$ is equal to $P(C \mid (A \wedge B))$. From this and the Ratio Formula, we know that $P(C \mid (A \wedge B))$ is equal to $P(A \wedge B \wedge C) \div P(A \wedge B)$. Introducing our numbers, we find that the probability of the first premise is 1. This is consistent with our intuition that the first premise has a high probability.

Next, turning to the conclusion: from the Equation, we know that $P(B \rightarrow C)$ is equal to $P(C | B)$. From this and the Ratio Formula, we know that $P(C | B)$ is equal to $P(B \wedge C) \div P(B)$. Introducing our numbers, we find that the probability of the conclusion is 0.17. This is consistent with our intuition that the conclusion has a low probability.

Finally, we move to calculating the uncertainty of the premises and conclusion. From our definition of uncertainty, the uncertainty of the first premise is $1 - P(A \rightarrow (B \rightarrow C)) = 0$, that of the second is $1 - P(A) = 0.20$ and that of the conclusion is $1 - P(B \rightarrow C) = 0.83$. We see that the argument isn't probabilistically valid. The uncertainty of the conclusion ($= 0.83$) exceeds the sum of the uncertainties of the premises ($= 0 + 0.2 = 0.2$). This being so, the suppositional theorist can explain the plausibility of #1. We might take the premises to be true and the conclusion false because the uncertainty of the conclusion can exceed the sum of the uncertainties of the premises.

3.2. Rejecting Thesis #3: *modus ponens* is valid

Depending on how she defines *modus ponens*, the argument form is or isn't *probabilistically* valid according to the suppositional theorist. If she defines it as requiring necessarily that we replace each of the letters in $Q \rightarrow R; Q; \therefore R$ with a proposition (i.e. a formula with a truth-value/not containing an indicative conditional), then it's probabilistically valid. Let's call the argument meeting the definition 'simple *modus ponens*'. To prove the validity of simple *modus ponens*, we derive the following lemma:

Lemma: Conditional Contradiction

Assuming $P(Q) \neq 0$, the probability of $\neg R$ given Q is equal to one minus the probability of R given Q :

$$1. P(Q) \div P(Q) = 1 \quad \text{by algebra}$$

$$2. [P((Q \wedge R) \vee (Q \wedge \neg R))] \div [P(Q)] = 1$$

from 1 by T2 (which we saw in Chapter 1 subsection 3.4. and according to which $P(A) = P(B)$ if and only if A and B are equivalent)

$$3. [P(Q \wedge R) + P(Q \wedge \neg R)] \div [P(Q)] = 1$$

from 2 by Finite Additivity

$$4. [P(Q \wedge R) \div P(Q)] + [P(Q \wedge \neg R) \div P(Q)] = 1$$

from 3 by algebra

$$5. P(R | Q) + P(\neg R | Q) = 1 \quad \text{from 4 by the Ratio Formula}$$

$$6. P(\neg R | Q) = 1 - P(R | Q) \quad \text{from 5 by algebra}$$

Using this lemma and still assuming that $P(Q)$ isn't zero, we prove the probabilistic validity of *modus ponens* where the conclusion is a proposition (i.e. doesn't contain an indicative conditional). The first step is to prove the uncertainty of an indicative conditional is equal to the quotient of the uncertainty of the corresponding material conditional divided by the probability of the antecedent. To do so, we appeal to not only Conditional Contradiction but also Uncertainty, the Ratio Formula and the equivalence of $Q \supset R$ and $\neg(Q \wedge \neg R)$:

$$1. U(Q \rightarrow R) = 1 - P(Q \rightarrow R) \quad \text{by Uncertainty}$$

$$2. = 1 - P(R | Q) \quad \text{from 1 by the Equation}$$

3. $= P(\neg R | Q)$ from 2 by Conditional Contradiction
4. $= [P(Q \wedge \neg R)] \div [P(Q)]$ from 3 by Ratio Formula
5. $= [1 - P(\neg(Q \wedge \neg R))] \div [P(Q)]$ from 4 by Conditional Contradiction
6. $= [U(\neg(Q \wedge \neg R))] \div [P(Q)]$ from 5 by Uncertainty
7. $= [U(Q \supset R)] \div [P(Q)]$ from 6 by T2

Taking it from top to bottom,

8. $U(Q \rightarrow R) = [U(Q \supset R)] \div [P(Q)]$ from 1 and 7 by transitivity of identity (this corresponds to Bennett's 'fourth result' (2003, p. 134)).

So far, we've proved the uncertainty of an indicative conditional is equal to the quotient of the uncertainty of the corresponding material conditional divided by the probability of the antecedent. Building on the first step, the second is to prove that an indicative conditional probabilistically entails a material one. In other words, the uncertainty of a material conditional $Q \supset R$ can't exceed the uncertainty of the corresponding indicative one $Q \rightarrow R$. To prove this, we appeal to algebra:

9. $[U(Q \rightarrow R)] \times [P(Q)] = U(Q \supset R)$ from 8 by algebra
10. $U(Q \rightarrow R) \geq U(Q \supset R)$

from 9 by algebra (because $P(Q)$
is less than or equal to 1)

By now, we've proved that an indicative conditional probabilistically entails a material one. Building on this previous step, the final one is to prove that an indicative conditional and its antecedent probabilistically entail the consequent. In other words, the uncertainty of the consequent can't exceed the sum of the uncertainties of the conditional and the antecedent. To prove this, we appeal to algebra again:

11. $U(Q) + U(Q \supset R) \geq U(R)$ because Q and $Q \supset R$ entail R
12. $U(Q \supset R) \geq U(R) - U(Q)$ from 11 by algebra
13. $U(Q \rightarrow R) \geq U(R) - U(Q)$ from 10 and 12 by algebra
14. $U(Q \rightarrow R) + U(Q) \geq U(R)$

from 13 by algebra (this verifies
Bennett's 'Security Thesis'
(2003, p. 141))

And now we've proved what we wanted to prove. Line 14 says *modus ponens* is probabilistically valid for the indicative conditional according to the suppositional theorist at least when the conclusion is propositional, i.e. simple *modus ponens*. It tells us it's not the case the uncertainty of the conclusion in a *modus ponens* can be greater than the sum of the uncertainties of the premises.

If, however, the suppositional theorist defines *modus ponens* as not requiring necessarily a propositional conclusion (i.e. allowing that we substitute a conditional in place of R in the argument $Q \rightarrow R; Q; \therefore R$), then it isn't probabilistically valid. Let's call the argument which this definition includes but that of simple *modus ponens* doesn't 'complex *modus ponens*'.

As we saw, McGee’s argument – a complex *modus ponens* with its conclusion being a conditional – isn’t probabilistically valid. It’s possible for the uncertainty of the conclusion to exceed the sum of the uncertainties of the premises. McGee’s argument would be a counterexample to the probabilistic validity of *modus ponens*.

While we might think we could adapt the proof we’ve just seen to prove the probabilistic validity of an argument where the conclusion is an indicative conditional, we’d be wrong. The non-propositional nature of indicative conditionals bars us from doing so. Specifically, it bars us from adapting Conditional Contradiction on which the proof relies. Consider the first three lines of the lemma where we’ve replaced R with $R \rightarrow S$:

$$1. P(Q) \div P(Q) = 1 \quad \text{by algebra}$$

$$2. [P((Q \wedge (R \rightarrow S)) \vee (Q \wedge \neg (R \rightarrow S)))] \div [P(Q)] = 1$$

from 1 by the logical equivalence
theorem T2

$$3. [P(Q \wedge (R \rightarrow S)) + P(Q \wedge \neg (R \rightarrow S))] \div [P(Q)] = 1$$

from 2 by Finite Additivity

2 no longer follows from 1 by T2. Q is a proposition but $(Q \wedge (R \rightarrow S)) \vee (Q \wedge \neg (R \rightarrow S))$ isn’t because it contains indicative conditionals and indicative conditionals aren’t propositional. This being so, Q and $(Q \wedge (R \rightarrow S)) \vee (Q \wedge \neg (R \rightarrow S))$ can’t be logically equivalent and we can’t derive $(Q \wedge (R \rightarrow S)) \vee (Q \wedge \neg (R \rightarrow S))$ from Q by T2.

Similarly, 3 no longer follows from 2 by Finite Additivity. $Q \wedge (R \rightarrow S)$ and $Q \wedge \neg (R \rightarrow S)$ contain conditionals and conditionals aren’t propositional. Strictly speaking, then, they can’t be logically inconsistent and

so we can't derive $[P(Q \wedge (R \rightarrow S)) + P(Q \wedge \neg (R \rightarrow S))]$ from $[P((Q \wedge (R \rightarrow S)) \vee (Q \wedge \neg (R \rightarrow S)))]$ by Finite Additivity because that rule applies to logically inconsistent statements.

So here, we see a significant divergence between the classical logician's conception of *modus ponens* and the suppositional theorist's. For the classical logician, *modus ponens* is valid, whether the conclusion is itself a conditional or not. For the suppositional theorist, on the other hand, whether *modus ponens* is valid depends on whether she limits her definition to simple or complex instances.

If she defines *modus ponens* as allowing that the conclusion be a conditional, i.e. allowing a non-propositional conclusion, then she'll find *modus ponens* probabilistically invalid. If she defines it as barring that the conclusion be a conditional, i.e. barring a non-propositional conclusion, then she'll find *modus ponens* probabilistically *valid*.

3.3. Rejecting Thesis #2: the argument is an instance of *modus ponens*

The suppositional theorist rejects #2. It's not that she denies the second premise is the antecedent of the first or the conclusion is the consequent of the first. Rather, it's that #2 implies that the conclusion of McGee's argument – one with the form $Q \rightarrow R; Q; \therefore R$ – as being a proposition and here, being a conditional, the conclusion isn't a proposition.

So, according to the suppositional theorist, McGee's argument is no more an instance of *modus ponens* than the following argument, in which the conclusion is an imperative:

If it's sunny, then go outside!

It's sunny.

Therefore, go outside!

Here, again, the conclusion – ‘Go outside!’ – is non-propositional and so the whole argument can't be an instance of *modus ponens* according to the definition which #2 offers.

To be clear, the suppositional theorist here and that in Piller (1996) each reject #2 for a different reason. The suppositional theorist here rejects #2 on the grounds that it requires the conclusion in a *modus ponens* to be propositional and that in McGee's argument isn't. The suppositional theorist in Piller (viz. she who subscribes to the ‘Adams-Appiah theory’), however, rejects #2 on the grounds that the second premise isn't the antecedent of the first and thus McGee can't apply *modus ponens* ‘in the way he intended to’ (Piller, 1996, p. 40).

Indeed, according to her, the second premise is A and the antecedent of the first is $A \wedge B$. Piller's suppositional theorist claims that ‘what seems to be an embedded conditional is, in fact, no embedded conditional. McGee's premise seems to be of the form “If A , then if B , then C ” but is in fact of the form “If A and B , then C ” (Piller, 1996, p. 40 [I've changed the letters]). (Edgington agrees (1995, p. 284 and 2014).) This is because there is no formula to calculate the probability of an embedded conditional and this with good reason: ‘Conditionals cannot be embedded in other conditionals. The parts of a conditional, its antecedent and its consequent, must have truth-values. But because conditionals do not have truth-values, conditionals cannot be part of conditionals’ (Piller, 1996, p. 40).

Since Piller (1996) published his paper, there have been developments. We know now how to calculate the probability of an embedded conditional. As we've seen above, the probability of $A \rightarrow (B \rightarrow C)$ is equal to the probability of $(A \wedge B) \rightarrow C$ – and there's a formula for calculating the probability of the latter. So here, rather than reject #2 on the grounds that the conditionals can't embed and the second premise isn't the antecedent of the first, *our* suppositional theorist rejects it on the grounds that the conclusion of a *modus ponens* must be propositional and the conclusion in McGee's argument isn't.

Now, depending on how she defines *modus ponens*, our suppositional theorist here will find that the argument is or isn't an instance of *modus ponens*. If she defines *modus ponens* as having the form $Q \rightarrow R; Q; \therefore R$ and allowing the embedding of conditionals, then she'll find that McGee's argument *is* equivalent to *modus ponens*. Indeed, the argument would meet the definition.

If, however, she defines *modus ponens* as having the form $Q \rightarrow R; Q; \therefore R$ and barring the embedding of conditionals, then she'll find McGee's argument *isn't* equivalent to *modus ponens*. Indeed, in his argument, the R is a conditional.

And now she faces a dilemma. If she chooses the more-inclusive definition, then she can explain the plausibility of #2: it mightn't be the case that we're replacing $Q \rightarrow R; Q; \therefore R$ with propositions but we do, according to her theory, have a *modus ponens*. In choosing the more-inclusive definition, however, she bars herself from explaining the plausibility of #3. As we saw, *modus ponens* with embedded conditionals is probabilistically invalid.

If, conversely, she chooses the less-inclusive definition, then she can explain the plausibility of #3. She has a proof that *modus ponens* without embedded conditionals is valid. Similarly, however, in choosing the less-inclusive definition, she bars herself from explaining the plausibility of #2.

We see that the suppositional theorist rejects #1, #2 and #3 and while she can explain #1, she must choose between explaining #2 and #3.

4. The hybrid theory

Recall that the hybrid theory combines elements from the material theory and the suppositional theory. On the one hand, like the material theory, the hybrid one concerns itself with truth. On the other hand, like the suppositional theory, the hybrid one concerns itself with probability. For example, an indicative conditional is true where the corresponding material one is; and it's assertible where the conditional probability of the consequent given the antecedent is high. Also recall that assertibility is different from assertability (see Chapter 1).

Now, in this section, I show that the hybrid theorist accepts Theses #2 as well as #3 but rejects #1 while explaining its plausibility.

4.1. Rejecting Thesis #1: the premises are true and the conclusion is false

The hybrid theorist takes the first conjunct in #1 to be true. Indeed, according to her, the premises *are* true – and this for the same reason as the material theorist takes them to be true. Moreover, just as the premises are probable according to the suppositional theory, they are assertible according to the

hybrid one. The assertibility of $Q \rightarrow (R \rightarrow S)$ is exactly the same as that of the corresponding sentence of the form $(Q \wedge R) \rightarrow S$ (Jackson, 1987, p. 130). According to the hybrid theory, Import-Export preserves not only truth but also assertibility.

The adoption of Import-Export might seem ad hoc. Treating conditionals with the form $Q \rightarrow (R \rightarrow S)$ requires the introduction of extra rules. As Jackson admits,

There is no obvious, inevitable answer as to what supplemented equivalence theorists [who believe that \rightarrow and \supset are not identical in meaning inasmuch as \rightarrow carries a ‘conventional signal’ which \supset doesn’t] should say about ordinary language occurrences of indicative conditionals within the scope of indicative conditionals. The matter needs further investigation which, if all goes well, will lead to a further, *detachable*, bit of theory designed to handle the more complex constructions. (1987, p. 128, emphasis in original)

In contrast, the suppositional theory could more easily account for Import-Export thanks to Conditional Contradiction.

The hybrid theorist can’t just rely on truth-functional equivalences and say that since \supset is equivalent to \rightarrow and $Q \supset (R \supset S)$ is equivalent to $(Q \wedge R) \supset S$, the assertibility of $Q \rightarrow (R \rightarrow S)$ is equivalent to the assertibility of $(Q \wedge R) \rightarrow S$. The theorist denies the possibility of relying on them in the case of simple conditionals. ‘If Q , then R ’ and ‘Not- Q or R ’ might have the same truth table but there’s a difference in practice between the two. In practice, we’ll assert ‘If Q , then R ’ when the probability of R given Q is high. The same isn’t true about ‘Not- Q or R ’.

Likewise, the hybrid theorist can't just claim the embedded conditional is material. Jackson acknowledges this: 'Sometimes there is no theory to be had for an "if" in the scope of an "if"; other times there is a theory, and one consonant with the supplemented equivalence theory, but it is not obtained by simply treating the " \rightarrow " of smallest scope as " \supset "' (Jackson, 1987, p. 129). We'll return to this point when discussing #2 and a suggestion Lowe (1987) makes.

Now, Jackson reaches the conclusion that Import-Export preserves assertibility based on the fact that he can't find a counterexample. He sets out to look for one in 'the place to look for [them]', namely, 'the cases where $(Q \wedge R) \rightarrow S$ is highly assertible, Q is highly assertible, but $(R \rightarrow S)$ is highly unassertible' in vain (Jackson, 1987, pp. 130-1; see also Bennett, 2003, pp. 99-100).

That said, the hybrid theorist might be able to adapt the suppositional theorist's proof of Import-Export. To do so, the hybrid theorist must edit lines 1, 2, 3 and 9 such that the proof relies on not the Equation (which the hybrid theory doesn't assume) but Adams' Thesis (which it does):

Line 1 will equate the assertibility of $A \rightarrow (B \rightarrow C)$ with the probability of $B \rightarrow C$ given A , citing Adams' Thesis as justification;

Line 2 will derive the posterior assertibility of $B \rightarrow C$ after we discover A from 1 by Conditionalisation;

Line 3 will derive the posterior probability of C given B after we discover A from 2 by Adams' Thesis; and

Line 9 will derive the assertibility of $(A \wedge B) \rightarrow C$ from 8 by Adams' Thesis.

Of course, the proof assumes the current assertibility of $A \rightarrow (B \rightarrow C)$ is equal to the future assertibility of $B \rightarrow C$ after we find out that A . In other words, the proof assumes we can rely on Conditionalisation when talking about not probability but assertibility: we are saying the assertibility of $B \rightarrow C$ after we find out that A is equal to the conditional probability of $B \rightarrow C$ given A . Whether the hybrid theorist would assume this is unclear. But it seems plausible that she might, because assertion is the outward expression of belief – and if she does assume this, she can show that the premises are assertible.

Next, the hybrid theorist takes the second conjunct in #1 to be false. Indeed, according to her, it's not the case that the conclusion is false. Like the material theorist, the hybrid one takes the conclusion to be true, because it has a false antecedent. Nonetheless, the hybrid theorist can explain the plausibility of the thesis. While the conclusion might be true, it isn't assertible. As the suppositional theorist showed, the consequent has a low probability given the antecedent. From this and Adams' Thesis, we know that the conclusion has a low assertibility (Piller, 1996, p. 43).

4.2. *Accepting Thesis #3: modus ponens is valid*

The hybrid theorist agrees that *modus ponens* preserves truth. She would have to concede, however, that it doesn't preserve assertibility (Piller, 1996, p. 44). While the suppositional theorist choosing the less-inclusive definition could say that McGee's argument, although probabilistically invalid, wasn't a true counterexample to *modus ponens*, the hybrid theory can't. Strictly speaking (and we'll return to this in subsection 4.3), McGee's argument is an instance of *modus ponens*. The conclusion is a proposition and true. Thus the hybrid

theorist must say that McGee's argument, despite not showing that *modus ponens* fails to preserve truth, shows that *modus ponens* fails to preserve assertibility.

The hybrid theorist can't adapt the suppositional one's proof that *modus ponens* preserves probability such that it proves *modus ponens* preserves assertibility. Dashing any hope of this is the fact that we can't equate unassertibility with uncertainty. Granted, rather than define uncertainty, we can define unassertibility as follows:

Assuming that one is the maximum degree of assertibility, for any formula Q , the unassertibility of Q is equal to the remainder of one minus the assertibility of Q .

Moreover, we can edit the lemma such that it's about unassertibility and line 7 equates the assertibility of $Q \rightarrow \neg R$ with the remainder of one minus the assertibility of $Q \rightarrow R$, citing Adams' Thesis. And we can edit the proof such that:

Line 1 equates the unassertibility of $Q \rightarrow \neg R$ with the remainder of one minus the assertibility of $Q \rightarrow R$, citing the definition of unassertibility;

Line 2 derives the remainder of one minus the assertibility of R given Q by Adams' Thesis rather than the Equation;

Line 3 derives the assertibility of $Q \rightarrow \neg R$, from 1 by Unassertibility;

Line 4 cites again Adams' Thesis rather than the Equation;

Line 9 equates the unassertibility of $Q \rightarrow R$ with the quotient of the uncertainty of $Q \supset R$ divided by the probability of Q ;

Line 10 equates the product of the unassertibility of $Q \rightarrow R$ multiplied by the probability of Q with the uncertainty of $Q \supset R$; and

Line 11 finds the unassertibility of $Q \rightarrow R$ to be greater than or equal to the uncertainty of $Q \supset R$.

The problem is that we've proved something other than what we want. We want line 11 to find the unassertibility of $Q \rightarrow R$ to be greater than or equal to the unassertibility of $Q \supset R$. Instead, we've proved the unassertibility of $Q \rightarrow R$ to be greater than or equal to the *uncertainty* of $Q \supset R$.

The hybrid theorist might try to patch up her theory by saying that the uncertainty of $Q \supset R$ is equal to the unassertibility of $Q \supset R$. But this would be a false start. We know uncertainty and unassertibility can come apart. Just as probability of truth and assertibility don't go hand in hand for indicative conditionals, neither do their counterparts uncertainty and unassertibility. So in line 11, because R itself is an indicative conditional (as the conclusion is in McGee's argument), the unassertibility of $Q \supset R$ might well be greater than the uncertainty of $Q \supset R$.

4.3. *Accepting Thesis #2: the argument is an instance of modus ponens*

According to the hybrid theorist, the argument is an instance of *modus ponens* – and this, again, for the same reason as the material theorist takes it to be an instance of *modus ponens*.

Granted, like the suppositional theory, the hybrid one concerns itself with conditional probability, taking a conditional to be assertible where the conditional probability of the consequent given the antecedent is high. Unlike the suppositional theory, however, the hybrid one concerns itself also with

truth, taking an indicative conditional to have the truth value of the corresponding material conditional. And because of this, McGee's argument with its conditional conclusion meets the definition of *modus ponens* in #2.

We can see the tenet according to which conditionals are propositional as endowing the hybrid theory with an advantage and a disadvantage. On the one hand, it allows the theory to be consistent with the truth-conditional theory of meaning. Unlike the suppositional theory, it can say that to know the meaning of a conditional is to know what it takes for that conditional to be true. On the other, it prevents the theory from dismissing McGee's argument as a counterexample to *modus ponens*. McGee's argument constitutes an instance of *modus ponens* and fails to preserve assertibility.

The hybrid theorist can't escape the problem by saying:

The best reply, in my view, to this argument points out that we sometimes need to do a certain amount of massaging of surface linguistic structure in order to display logical form. ... Now it is plausible that the second premise of the putative counter-example should strictly be written as 'If a Republican wins and Reagan does not win, then Anderson will.' The sentence whose surface form is $A \rightarrow (B \rightarrow C)$ has logical form $(A \wedge B) \rightarrow C$. Hence, the alleged counter-example is not really an instance of *modus ponens*. (Jackson, 2006, p. 216, [I've edited the symbols for consistency])

The claim that a sentence whose surface form is $A \rightarrow (B \rightarrow C)$ has the logical form $(A \wedge B) \rightarrow C$ is ad hoc. Nonetheless, the hybrid theorist takes the two propositions to be logically equivalent – and, in doing so, doesn't help the hybrid theory. If the two *are* logically equivalent, then it remains for the

theorist to explain McGee's argument as a counterexample to *modus ponens*. Strictly speaking, an argument whose first premise is $(A \wedge B) \rightarrow C$, second is A and conclusion is $B \rightarrow C$ is logically equivalent to a *modus ponens*: relying on the logical equivalent of $(A \wedge B) \rightarrow C$ and $A \rightarrow (B \rightarrow C)$, we can derive an argument whose second premise is the antecedent of the first and whose conclusion is the consequent of the first.

This is a particularly hard blow for the theory because one of its *raison d'être* was to ensure that *modus ponens* preserved assertibility (Jackson, 1987, p. 29). Yet here we see that *modus ponens* doesn't. Efforts to save the theory by saying that the first premise wasn't assertible in the first place would be in vain. The probability of the consequent given the antecedent is high. Moreover, we take the premise to be assertible. (Here, I go further than Piller who wrote before Jackson published his self-described 'best reply' to McGee (Jackson, 2006, p. 216). Piller says Jackson can solve the paradox, assuming his theory is correct. I say Jackson can't without abandoning one of the most important motivations for devising his theory.)

Now, one might think that the hybrid theorist might have a way out. The response above, of course, assumes that, according to the hybrid theory, we can embed indicative conditionals in indicative conditionals. If we reject the idea, however, we would reject Thesis #2. Instead of seeing the conclusion as the consequent of the first premise, we might, as Lowe does, see the conclusion as an indicative conditional and the consequent of the first as a material one:

$$A \rightarrow (B \supset C)$$

A

$\therefore B \rightarrow C$

Note that, in what follows, I'll be talking about a fictional Lowe. While the real one makes no mention of probability, I'm casting the fictional one as a hybrid theorist who will, as we'll see in a counterexample, appeal to Adams' Thesis. Think 'A New Lowe', if you will.

Now, one reason, perhaps, for rejecting the embedding of indicative conditionals is that they carry implicatures and implicatures don't survive all embeddings. For example, if we say 'He's good at maths but not chess', we're committed to the implicature according to which there's a contrast between being good at maths and being good at chess. However, if we say 'He believes he's good at maths but not chess', we're not committed to that implicature. Similarly, then, when we say 'If B , then C ', we might be committed to a set of implicatures while when we say 'If A , then if B , then C ' we are not committed to the same set.⁷

That said, implicatures do survive some embeddings. For example, if we say 'She knows she's good at maths but not chess', we're committed to the implicature to which we weren't committed in 'He believes he's good at maths but not chess'. Namely, we're committed to the implicature according to which there's a contrast between being good at maths and being good at chess.

Now, coming back to Lowe where we left him, i.e. before I tried to reconstruct his motivation for rejecting the embedding of indicative

⁷ Note that some theorists – such as e.g. Christopher Potts – classify 'but' as a conventional implicature, and think that one distinguishing feature of conventional implicatures is that they do project out of pretty much all embeddings. I don't think this affects the larger points here but the interested reader can turn to Potts (2004).

conditionals – he does not argue that the first premise has the form $A \rightarrow (B \supset C)$. Instead, he claims, the burden of proof is on McGee's shoulders to show that the first premise does *not* have this form. For Lowe,

it is enough that $A \rightarrow (B \supset C)$ is not a patently implausible interpretation. ... The burden of proof lies rather with McGee to show that his first premise genuinely *is* of the form $A \rightarrow (B \rightarrow C)$, since it is he who is relying on this assumption in order to challenge a deeply rooted principle of deductive inference. (1987, p. 46 [I've changed the letters])

As far as Lowe sees,

a reasonable degree of logical conservatism entitles us to see in McGee's example not a breakdown of *modus ponens* but rather a demonstration that the English indicative conditional is sometimes interpretable as a material conditional and sometimes not. (1987, p. 46)

There's a problem with this response. The burden of proof *does* lie with Lowe and he doesn't discharge it. To reject #2 and prove that McGee's argument isn't a *modus ponens*, it is not enough to provide an interpretation of the first premise which is not patently implausible. Just because it's not patently implausible that I'm currently at the library doesn't undermine the fact that I'm not. To begin to defend *modus ponens* as an argument which preserves assertibility, Lowe must provide some positive reason for believing that the first premise has the form $A \rightarrow (B \supset C)$. Not having done so, Lowe fails to respond to McGee.

If he *could* provide a reason, then one might think his theory an attractive one. He could explain the plausibility of the other two Theses. He'd

accept #1 according to which the premises are true and the conclusion false. He finds the first premise assertable inasmuch as, on his formulation, it corresponds to ‘If a Republican wins the election, then either it will be Reagan who wins or it will be Anderson’ (Lowe, 1987, p. 45). He gives no reason to find the second premise unassertible but does give one to find the conclusion so:

no conversational point is normally served by *asserting* something of the form $B \supset C$ where B and C are reasonably believed to be false, any more than it is by asserting something of the equivalent form $\neg B \vee C$. This is indeed why, although the circumstances of McGee’s example it would be *reasonable to believe* the disjunction

Either it will be Reagan who wins the election or it will be Anderson,

it would not be conversationally appropriate to *assert* this. (Lowe, 1987, p. 46, emphases in original)

Considering probabilities, we can bear out Lowe’s position – finding the premises assertible but the conclusion not so, where the embedded connective in the first premise is a material conditional but the main connective in the conclusion is an indicative one. Adopting where necessary the numbers we used in the suppositional theory section (subsection 3.1), we can calculate the following probabilities for the premises and conclusion.

Turning to the first premise: from Adams’ Thesis, we know that the assertibility of $A \rightarrow (B \supset C)$ is equal to $P((B \supset C) | A)$. From material implication, we know that $P((B \supset C) | A)$ is equal to $P((\neg B \vee C) | A)$. After all, $(B \supset C)$ is logically equivalent to $(\neg B \vee C)$. Next, applying the Ratio

Formula, we know that $P((\neg B \vee C) | A)$ is equal to $P((\neg B \vee C) \wedge A) \div P(A)$. From the stipulation that $\neg B \vee C$ is equal to A , we know that $P((\neg B \vee C) \wedge A) \div P(A)$ is equal to $P(A \wedge A) \div P(A)$. This stipulation that the probability that ‘Reagan wins or Anderson does’ is equivalent to the probability that ‘A Republican wins’ isn’t controversial: there are two Republicans running in our scenario, Reagan and Anderson. And applying algebra, we calculate that $P(A \wedge A) \div P(A)$ is equal to $P(A) \div P(A)$.

Taking it from top to bottom, by arithmetic, the assertibility of $A \rightarrow (B \supset C)$ is equal to $P(A) \div P(A)$. So, whatever $P(A)$ (so long as it isn’t zero), the probability of the first premise is 1.

Turning to the second premise: its probability is 0.8. The second premise is A to which, in subsection 3.1. of this chapter, we’d assigned this high probability.

And turning to the conclusion: its probability is 0.17. The conclusion is $B \rightarrow C$ which, as we calculated in subsection 3.1. of this chapter also, has a low probability.

Lowe would accept #3 according to which *modus ponens* is valid. As he writes, ‘it is more reasonable to appeal to the validity of *modus ponens* to show that McGee has misinterpreted the form of one of the sentences he invokes in his example’ (Lowe, 1987, p. 46). Still, Lowe would have to provide a positive reason for believing that the first premise has the form $A \rightarrow (B \supset C)$.

But Lowe’s theory is no more attractive than that of his hybrid theorist colleague who took the argument to have the form $A \rightarrow (B \rightarrow C); A; \therefore B \rightarrow C$. Lowe can’t say *modus ponens* preserves assertibility either. If McGee’s

argument were a *modus ponens* (i.e. the connective in the conclusion were not \rightarrow but \supset and thereby the same connective as that in the consequent of the first premise), then the assertibility of the conclusion could be low still while those of the premises were high. Line 1 of the counterexample would say that the assertibility of $(B \supset C)$ is equal to the probability of C given B and from here the calculation would be the same as that for the other hybrid theorist. (Even the non-fictional version of Lowe, as we saw earlier in this subsection, finds that it wouldn't be appropriate to assert the disjunction version of the conclusion.)

5. Concluding

In this chapter, I sought to analyse largely from first principles Vann McGee's alleged counterexample to *modus ponens*. I focused on responses the material, possible-worlds, suppositional and hybrid theories of the indicative conditional would offer to the trilemma McGee's argument presents. The theses of the trilemma were:

#1 the premises are true and the conclusion is false;

#2 the argument is an instance of *modus ponens*; and

#3 *modus ponens* is valid.

We saw that the material theorist respond by rejecting #1. According to her, the conclusion is true inasmuch as it had a false antecedent. In trying to explain the plausibility of #1 by defending a surrogate thesis, however, the theorist introduced new problems.

Similarly, the possible-worlds theorist responded by rejecting #1. According to her, the first premise isn't true. In trying to explain the

plausibility of #1, however, the theorist offered a new formulation of the first premise but gave no positive reason for our accepting that new formulation.

The suppositional theorist responded to the trilemma by rejecting all theses. As she takes indicative conditionals to be non-propositional, she finds it's not the case that the premises are true or the conclusion false, that the argument is an instance of *modus ponens*, or that *modus ponens* preserves truth. While she could explain the plausibility of #1, she had to choose between explaining #2 and #3. Prevent her from explaining all three was the fact that however she defined *modus ponens*, her definition barred her from finding either that *modus ponens* is probabilistically valid or McGee's argument is an instance of *modus ponens*.

And the hybrid theorist responded by rejecting #1. For the same reason as the material theorist, the hybrid one finds the conclusion to be true: it has a false antecedent. The theory's tenet according to which conditionals are propositional proved an advantage and a disadvantage. On the one hand, it allowed the theory to remain consistent with the truth-functional account of meaning. On the other hand, however, it implied an unwelcome result: according to the theory, McGee's argument was an instance of *modus ponens* and probabilistically invalid – and this, when one of the theory's *raison d'être* was to guarantee the probabilistic validity of *modus ponens*.

Along the way, I disagreed with Piller (1996) and a fictional Lowe (1987). While Piller's suppositional theorist rejects #2 on the grounds that the second premise in McGee's argument (viz. A) isn't the antecedent of the first (viz. $A \wedge B$), mine does so – if at all – on the grounds that the conclusion and

consequent of the second premise aren't propositional. And while Lowe takes the hybrid theory to solve the paradox, I showed a flaw in his argument.

CHAPTER 3
THE BARBERSHOP PARADOX

“‘What, *nothing* to do?’ said Uncle Jim. ‘Then come along with me down to Allen’s. And you can just take a turn while I get myself shaved’” (Carroll, 1894, p. 436). So begins Lewis Carroll’s ‘ornamental presentment’ of an apparent counterexample to the validity of *modus tollens*.

Despite the progress logicians made on conditionals in the centuries following its publication and the attempts logicians made at offering a solution, the paradox remained without one. In this chapter, as I did in the case of McGee’s in the previous one, I proceed largely from first principles in analysing Carroll’s argument. I focus on the responses four theories of the indicative conditional would offer to the trilemma Carroll’s argument presents.

Allen’s is a shop where three barbers work: the eponymous Allen, Brown and Carr. The narrator, the young ‘Cub’, heads there with his two uncles, Jim and Joe. On the way, Uncle Joe makes a claim: ‘Carr’s *certain* to be in’ and proceeds to prove it (Carroll, 1894, p. 436).

First, he submits that ‘If Carr is out, it follows that if Allen is out Brown *must* be in’ (Carroll, 1894, p. 436, emphasis in original). Uncle Jim grants this conditional: ‘Of course he must ... or there’d be nobody to mind the shop’ (Carroll, 1894, p. 436). Next, Uncle Joe submits that ‘the Hypothetical “if Allen is out Brown is out” is *always* in force’ (Carroll, 1894, p. 437, emphasis in original). Again, Uncle Jim grants the conditional: ‘ever

since he had that fever [Allen]’s been so nervous about going out alone, he always takes Brown with him’ (Carroll, 1894, p. 437).

From these two conditionals, ostensibly by *modus tollens*, Uncle Joe concludes ‘Therefore Carr *cannot* be out’ (Carroll, 1894, p. 437, emphasis in original). This leaves Uncle Jim looking ‘thoroughly puzzled’ (Carroll, 1894, p. 437).

Now, if we formalise the argument without modals and let C be ‘Carr is out’, A be ‘Allen is out’, B be ‘Brown is *in*’, we get:

$$C \rightarrow (A \rightarrow B)$$

$$(A \rightarrow \neg B)$$

$$\therefore \neg C$$

Of course, according to classical logic, this isn’t an instance of *modus tollens*. While the conclusion might be the contradictory of the antecedent of the first premise, the second premise isn’t the contradictory of the consequent of the first. To transform the argument into a *modus tollens* while maintaining the formalisation of the first premise, we’d rewrite the second as

$$\neg(A \rightarrow B)$$

If I’m not formalising the second premise in this way, i.e. if I’m formalising the negation as taking not wide but narrow scope, it’s for three reasons:

(i) the negation in Carroll’s original argument takes narrow scope.

Uncle Joe says ‘If Allen is out Brown is out’ or in other words ‘If Allen is out Brown is not in’ (narrow). He doesn’t say ‘It’s not the case that if Allen is out Brown is in’ (wide). And I want to remain true to the scope Carroll used in his original argument;

(ii) if we assume the negation takes wide scope, we might not look further into Carroll's argument. With the negation taking wide scope, Carroll's argument would be similar to McGee's, which we've already examined: Carroll's second premise would have the form of the negation of McGee's conclusion and Carroll's conclusion the form of the negation of McGee's second premise. If we didn't look further into Carroll's argument, we'd miss out on the opportunity to make some interesting findings about how different theories of the indicative conditional grapple with negation – and I want not to miss out on that opportunity; and

(iii) for some of the theories we'll consider, namely the single-world and suppositional ones, $A \rightarrow \neg B$ and $\neg (A \rightarrow B)$ are equivalent, assuming A is possible or probable. So if we go ahead and formalise the negation with wide scope, we'd be foreclosing the possibility of considering a question under debate, namely 'Is Carroll's argument equivalent to a *modus tollens*?' – and I want not to foreclose that possibility.

Carroll isn't alone in offering a counterexample to *modus tollens*. James Williams Forrester (1984), Frank Veltman (1985), John Cantwell (2008), Niko Kolodny and John MacFarlane (2010), and Seth Yalcin (2012) do so too. I won't discuss Veltman explicitly as the argument he presents doesn't raise any question I don't answer in this chapter or the previous one. Veltman's argument is similar to Carroll's inasmuch as one premise contains an embedded conditional; and it's similar to McGee's inasmuch as there's no question about the scope of negation (Veltman's negation takes wide scope

(1985)). I'll discuss Forrester (1984), Cantwell (2008), Kolodny and MacFarlane (2010) and Yalcin (2012) in Chapter 4, as the arguments they present contain – not embedded conditionals but – embedded modals.

Likewise, I'm not alone in offering a discussion of Carroll. Alfred Sidgwick (1894), William Johnson (1894), Bertrand Russell (1905), Pavel Florensky (1914/1997), Arthur Burks and Irving Copi (1950) and, more recently, Michael Rhodes (2012) do so too. None of the initial respondents had modern theories of indicative conditionals at their disposal; and Rhodes focuses on comparing two of the earlier twentieth-century solutions rather than discussing modern theories.

In this chapter, I consider Carroll's argument in light of the material (inspired by Grice (1961, 1975 and 1989)), possible-worlds (inspired by Stalnaker (1981) and Lewis (1976)), suppositional (inspired by Adams (1975) and Edgington (1995 and 2014)) and hybrid (inspired by Jackson (1987)) theories. Specifically, I examine how the theories respond to an inconsistent triad of the following three individually plausible but jointly inconsistent theses:

#1 the argument is invalid, i.e. it's possible that the premises are true, while the conclusion is false;

#2 the argument is equivalent to an instance of *modus tollens*, i.e. the second premise of the argument is the contradictory of the consequent of the first, and the conclusion is the contradictory of the antecedent or since it's plausible that $A \rightarrow \neg B$ is equivalent to $\neg(A \rightarrow B)$; and

#3 *modus tollens* is valid, i.e. any argument in which the second premise is the contradictory of the consequent of the first and the

conclusion is the contradictory of the antecedent of the first premise is valid.

Of course and again, according to classical logic, #2 isn't true. But part of the challenge for the theories is to explain why it might be plausible. As we'll see, not all theories can explain the plausibility of all three theses at once – and to solve the trilemma they must. Moreover, the theories which can explain the plausibility of all three theses at once face problems of their own.

In section 1, I present the material theory and show that it rejects Thesis #2 and can explain at most the plausibility of two theses. In section 2, I present the possible-worlds theory and show that it rejects Thesis #1 but can't explain its plausibility. In section 3, I present the suppositional theory and show that it rejects all theses and, like the material theorist when it comes to – not the theses but – the total, can explain at most two theses. In section 4, I present the hybrid theory and show that it rejects Thesis #2 and can explain its plausibility. In section 5, I conclude.

1. The material theory

Recall that according to the material theorist, an indicative conditional $Q \rightarrow R$ is true if and only if the corresponding material conditional $Q \supset R$ is true. In other words, 'If Q , then R ' is true if and only if Q is false or R is true. For example, let Q be 'Ben's on sabbatical' and R be 'he's in Australia'. The indicative conditional 'If Ben's on sabbatical, then he's in Australia' is true if and only if it's not the case that Ben's on sabbatical or it is the case that he's in Australia.

Now, considering Carroll's story, the material theorist might see the following fact as the source of Uncle Jim's puzzlement: among the viable combinations of who might be in/out, it's not necessarily the case that Carr is in. Indeed, there are the following eight *possible* combinations = 2^3 :

	<i>A</i>	<i>B</i>	<i>C</i>
1	In	In	In
2	In	Out	In
3	Out	In	In
4	Out	Out	In
5	In	In	Out
6	In	Out	Out
7	Out	In	Out
8	Out	Out	Out

Table 3: Possible combinations for who, among the barbers, might be in/out

And among those eight, only five are viable in Carroll's scenario. The second premise rules out possible combinations #3 and #7. Each suggests that Allen is out and Brown is in. Similarly, the first premise rules out possible combination #8. It suggests that all barbers are out. And among the five viable combinations, #5 and #6 have Carr out.

In this section, I show how the material theory can explain Thesis #3. I also show how the theory must choose between explaining Theses #1 and #2.

1.1. Explaining Thesis #3: modus tollens is valid

According to the material theory, Carroll's argument doesn't falsify Thesis #3. There's no question whether *modus tollens* is valid, i.e. where the premises of an argument are those of a *modus tollens* and true, the conclusion will necessarily also be true.

Suppose that both premises of a *modus tollens* argument are true. Then the consequent of the first must be false. The second premise is true and the contradictory of the first. But then, according to the material theory, the antecedent of the first premise must be false. To be true, a conditional with a false consequent must have a false antecedent. It follows that the conclusion must be true, since the antecedent of the first premise is false and the contradictory of the conclusion.

Carroll's argument isn't a counterexample to this. Its premises aren't those of a *modus tollens* (Russell, 1905). According to the material theory, the premises have the form:

$$C \supset (A \supset B)$$

$$A \supset \neg B$$

On the material account, the second premise $A \supset \neg B$ isn't the contradictory of the consequent of the first $A \supset B$. As we have seen, according to the account, an indicative conditional $Q \rightarrow R$ is true if and only if the corresponding material conditional $Q \supset R$ is true. And it isn't the case that $A \supset B$ is true if and only if $A \supset \neg B$ is false. For example, where A is false, both conditionals are true. With the second premise not being the contradictory of the consequent of the first, Carroll's argument doesn't justify worries about *modus tollens* not being truth-preserving, since it isn't an instance of *modus tollens*.

Moreover, it doesn't justify worries about valid argument forms not being truth-preserving either, since it's not classically valid. The argument can have true premises and a false conclusion. Just let A be false and C be true. Here, the truth value of B doesn't matter. Where C is true and A is false, the first premise will be true. The main connective is a conditional which has a true antecedent and true consequent. Where A is false, the second premise is also true. The main connective is a conditional which has a false antecedent. And where C is true, the conclusion is false. The main connective is a negation whose argument is true.

Even if we avoid the scenario where the second premise seems to be the contradictory of the consequent of the first, though logically it isn't, the material theorist faces problems. One way to avoid the scenario is to reformulate the second premise such that it's the logical contradictory of the consequent of the first:

$$\neg(A \supset B)$$

Now, the new premise reads 'It's not the case that if Allen is out, then Brown is in' and $\neg(A \supset B)$ is much stronger an assertion than the original $A \supset \neg B$ 'If Allen is out, then Brown is out'. According to the material theorist, the new formulation is equivalent to $A \wedge \neg B$ 'Allen is out and Brown is out' – a conjunction rather than a conditional and one we wouldn't accept. We don't know whether either conjunct is true, whether Allen or Brown is out (Veltman, 1985, p. 26).

1.2. Distinguishing between truth and assertability

So far, the material theory has a virtue: it explains the thesis according to which the argument is invalid. The theory shows that the conclusion doesn't follow from the premises. The theory has a failing, however: it doesn't explain Thesis #2 according to which the second premise is the contradictory of the consequent of the first. We take 'If Allen is out, then Brown is out' to deny 'If Allen is out, then it's not the case that Brown is out' yet the theory maintains that the one doesn't contradict the other (as we saw in passing in subsection 1.1.). Both conditionals can be true when it's not the case that Allen is out.

Drawing on the concept of assertability we saw in Chapter 1, a material theorist would argue that we take the second premise of Carroll's argument to be the contradictory of the consequent of the first not because they can't both be true – as we've seen, they can be – but rather because the one is assertable if and only if we can deny the other. In normal circumstances, again, saying both 'If Allen is out, then Brown is out', and also 'If Allen is out, then it's not the case that Brown is out' wouldn't be reasonable. However, as we'll see, in introducing the concept of assertability, the material theorist faces a dilemma.

1.3. Framing the dilemma

Having introduced the concept of assertability, the material theorist now faces a dilemma: show that $A \supset \neg B$ denies $A \supset B$ or not. Either way, she can't explain Thesis #1 or Thesis #2.

Reformulating them in the relevant terms for the material theorist:

Thesis #1: the argument doesn't preserve assertability, i.e. while the premises are such that we might assert them, the conclusion is such that we wouldn't.

Thesis #2: the second premise in Carroll's argument denies the consequent of the first. In other words, 'If Allen is out, then Brown is out' denies 'If Allen is out, then Brown is in'.

If the theorist can show that $A \supset \neg B$ denies $A \supset B$, then she can explain Thesis #2 but not #1. And if she can't show that $A \supset \neg B$ denies $A \supset B$, then she can explain #1 but not #2. Let's consider each horn in turn, taking the second first.

Note, before we do, that the dilemma which the material theorist faces here is different from that which the suppositional theorist faced in Chapter 2. The suppositional theorist had to choose between explaining the plausibility of – not #1 and #2, i.e. that *the argument under consideration is invalid and it has the form of a classically valid argument form* but – #2 and #3, i.e. *the argument under consideration has the form of a classically valid argument form and the classically valid argument is valid*.

1.4. Trying to explain reformulated Thesis #2: the second premise denies the consequent of the first

To explain Thesis #2, the material theorist would deny

(i) $A \supset \neg B$ is true if and only if $A \supset B$ isn't

and argue

(ii) we can assert $A \supset \neg B$ if and only if we can deny $A \supset B$

claiming that the thesis supports (ii) and not (i). But there's a difference between the thesis the theorist wants to prove (ii) and, as we'll see, the thesis the theorist can prove (iii):

(iii) we can assert $A \supset \neg B$ only if we cannot assert $A \supset B$

The difference is twofold: she wants to prove a biconditional (if and only if, i.e. that the one side is a necessary and sufficient condition for the other) but can prove only a conditional (only if, i.e. that the one side is a necessary – but not sufficient – condition for the other); and she wants to prove that if we can assert $A \supset \neg B$, then we can deny $A \supset B$ but can prove only that we cannot assert $A \supset B$.

To *try* to show that the thesis supports (ii) and not (i), the material theorist could seek to prove that $A \supset \neg B$ and $A \supset B$ are not co-assertable. The proof relies on two maxims of conversation:

Assert the stronger rather than the weaker (Grice, 1961, p. 132); and

Don't assert what you don't believe (Grice, 1975, p. 46).

(See Chapter 1 (subsections 1.1.1. and 1.2.) for discussion of these maxims.)

Suppose $A \supset \neg B$ and $A \supset B$ are co-assertable. This implies that we believe both conditionals to be true. If we didn't believe $A \supset \neg B$ and $A \supset B$ to be true, they wouldn't be assertable. In saying them, we would be saying something we didn't believe to be true. This in turn implies that we believe A to be false. If A were true, then according to the truth table of the material conditional, one conditional would be false. But if we believe A to be false, then neither conditional is assertable. In asserting the conditionals, we would be asserting weaker sentences when we could be asserting a stronger one. We

could assert $\neg A$. It entails $A \supset \neg B$ and $A \supset B$ but not vice versa. So we must abandon our initial supposition that $A \supset \neg B$ and $A \supset B$ can be co-assertable.

This proof doesn't help the theory. The proof shows that the second premise isn't assertable when the consequent of the first one is and vice versa (iii). It doesn't show, however, that the second premise denies the consequent of the first and vice versa – and to explain Thesis #2 it must. From the fact that, when we assert the second premise, we can't assert the consequent of the first it doesn't follow that, when we assert the second premise, we can deny the consequent of the first. (Here, I understand 'denying p ' as being equal to 'asserting not- p '.) Just because, not knowing where Mike is, I can't assert 'Mike is in the USA' doesn't mean I can deny 'Mike is in the USA'. He might very well be there, I don't know!

Here, to explain Thesis #2, the material theorist must be able to show that when we can assert the second premise, we can deny the consequent of the first. Showing that the two aren't co-assertable is not sufficient. In other words, the proof shows something weaker than we want here. It shows (iii) rather than (ii).

While showing (iii) if not (ii) gives the material theorist a partial solution, one needn't think she's in better a position than other theorists. While the material theorist might receive a pass here, she'd receive a fail elsewhere. Recall that, in the previous chapter, the material theorist found unassertable a sentence we might assert.

1.5. Trying to explain reformulated Thesis #1: the argument doesn't preserve assertability

Moreover, a successful explanation of Thesis #2 would bar the material theorist from explaining Thesis #1, according to which the conclusion is such that we wouldn't assert it. Given that the material theorist has explained Thesis #2, and shown that the second premise denies the consequent of the first, the following two principles of assertability would commit the material theorist to saying that the conclusion must also be assertable:

Principle #1

If we can assert a conditional and deny its consequent, then we must be able to deny its antecedent.

Principle #2

We can deny a statement if and only if its negation is assertable.

Briefly on Principle #1 – assuming the material theorist accepts *modus tollens* and wants to have a surrogate for Thesis #3, she will accept Principle #1. It's just a reformulation of *modus tollens* in terms of assertability rather than validity. Granted, the material theorist wouldn't say that every argument which is valid preserves assertability. One of the paradoxes of material implication – Q ; Therefore, if R , then Q – is a counterexample. But it doesn't follow that the material theorist would say that *modus tollens* doesn't preserve assertability. *Modus tollens* is different from the paradoxical argument. The paradox is (materially) valid but seems invalid. In contrast, *modus tollens* is (materially) valid and seems valid.

Now if, as we have seen, we can assert $C \supset (A \supset B)$ and can deny $A \supset B$, then, by Principle #1, we will be able to deny C . And if we can deny C , then, by Principle #2, we will be able to assert $\neg C$. $\neg C$ will be assertable – despite the thesis that it isn't.

Note that the principles aren't controversial. For example, if we can assert 'If Elena isn't in the Grad Room, then she's at the library' and can deny that she's at the library, then we can deny that she isn't in the Grad Room (Principle #1). Furthermore, if we can deny that she isn't in the Grad Room, then we can assert 'Elena's in the Grad Room' (Principle #2).

Now, the material theorist can explain the thesis that the conclusion isn't assertable and avoid undermining the main tenet by abandoning the claim that $A \supset \neg B$ denies $A \supset B$. By doing so, she allows a scenario where we can assert $A \supset \neg B$ and $C \supset (A \supset B)$ without having to deny C . While we can't assert $A \supset B$ in this scenario, it doesn't follow from the fact that we can't assert $A \supset B$ that we can deny it, and so Principle #1 need not apply.

This doesn't, however, help the theory either. Granted, it enables the material theorist to explain Thesis #1. However, obviously, in abandoning the explanation that $A \supset \neg B$ denies $A \supset B$, the theorist can no longer even begin to explain Thesis #2.

1.6. Making Carroll's argument classically valid

Even if we avoid the scenario where the second premise fails logically to contradict the consequent of the first, the material theorist faces problems. One way to avoid the scenario is to reformulate the second premise such that it's the logical contradictory of the consequent of the first:

$$\neg(A \supset B)$$

Now, the new premise reads 'It's not the case that if Allen is out, then Brown is in' and $\neg(A \supset B)$ is much stronger an assertion than the original $A \supset \neg B$ 'If Allen is out, then Brown is out'. According to the material theorist, the new

formulation is equivalent to $A \wedge \neg B$ ‘Allen is out and Brown is out’ – a conjunction rather than a conditional and one we wouldn’t accept. We don’t know whether either conjunct is true, whether Allen or Brown is out.

Here, the material theorist might claim that it’s not always clear what a speaker who says ‘It’s not the case that if Q , then R ’ is committing himself to or intends to convey (Grice, 1989, pp. 80-1). She could be saying that it’s not the case that ‘If Q , then R ’ is true, or that it’s not the case that it’s assertable. According to Grice, there are three kinds of cases but, as we will see, only two are relevant to Carroll’s example:

‘(1) Cases in which the unnegated conditional [‘If Q , then R ’] has (would have) no [meaning beyond its literal meaning], and in which the total signification of its utterance is representable by the content of the material conditional’ (Grice, 1989, p. 80). In other words, the denial of ‘If Q , then R ’ has the meaning of denying the literal content of ‘If Q , then R ’, in other words, of asserting ‘ Q and not- R ’.

This interpretation isn’t relevant. According to it, the second premise is equivalent to ‘Allen is out and Brown is out’ which we have already rejected as too strong; next

(2) cases where the denial of ‘If Q , then R ’ ‘is naturally taken as a way of propounding a counterconditional, the consequent of which is the negation of the consequent of the original conditional’: ‘If Q , then not- R ’ (Grice, 1989, p. 80).

This interpretation *is* relevant. According to it, $\neg(A \supset B)$ has the meaning of the counterconditional $A \supset \neg B$, which is the original formulation.

However, this just brings us back to the original problem of explaining Thesis #2: we would be trying to explain why $A \supset \neg B$ denies $A \supset B$ despite the fact that, according to their truth tables, the two conditionals can be true when A is false; and finally

(3) cases where the denial of ‘If Q , then R ’ ‘has the effect of a refusal to assert the conditional in question, characteristically because the denier does not think that there are adequate non-truth-functional grounds for such an assertion’ (Grice, 1989, p. 81).

Let’s consider this more closely. A truth-functional ground for asserting ‘If Q , then R ’ would be that you know R or you know not- Q . So if you know that Amelia is in the USA, you have a truth-functional ground for asserting ‘If she’s not in Morocco, then she’s in the USA’. Likewise, if you know that Amelia *is* in Morocco, you have a truth-functional ground for asserting ‘If she’s not in Morocco, then she’s in the USA’.

A non-truth functional ground for asserting ‘If Q , then R ’, then, would be that you know ‘Either R or not- Q ’ but you don’t know R and you don’t know not- Q either. So if you know that either Amelia is in the USA or she’s in Morocco but you don’t know that currently she’s in the USA and you don’t know that currently she’s in Morocco, you have a non-truth functional ground for asserting ‘If she’s not in Morocco, then she’s in the USA’.

Now, a person might *not* have adequate non-truth functional grounds to assert ‘If Q , then R ’ for a couple reasons: she has a reason but it’s a truth-functional one; or she has no reason whatsoever to assert the conditional in the first place. So if you know only that Amelia is in the USA, you do not have an adequate non-truth functional ground to assert ‘If she’s not in Morocco, then

she's in the USA'. Same if you know only that Amelia is in Morocco. In those two cases, you know only a truth-functional ground for asserting the conditional: you know the consequent is true; or you know the antecedent is false.

And you also lack a non-truth functional ground to assert the conditional should Amelia be standing in front of you in Singapore. In that case, you have no reason to assert the conditional in the first place. You have no reason to think that Amelia is either in Morocco or the USA.

Like the second, this third interpretation is relevant. According to it, $\neg(A \supset B)$ is a refusal to assert $A \supset B$. Here, we would refuse to assert 'If Allen is out, then Brown is in' not because we have a truth-functional ground to assert it. It's not that we know that it's not the case that Allen is out or that we know that Brown is in. On our way to the barbershop, we have no idea who we'll find there. Rather, it's that we have no reason to assert the conditional in the first place.

Note that here I'm not saying 'There's no reason to assert anything'. That would be false because we *do* have a reason to assert 'If A , then not- B '. Indeed, through the story, we know Allen never goes out alone and always takes Brown. But the fact that we *have* this reason is irrelevant. It's relevant only when we're talking about asserting 'If A , then not- B ' and above we're talking about asserting 'If A , then B ' – a conditional with the same antecedent A but a negation before the consequent B . No, instead of saying 'There's no reason to assert anything', I'm saying 'There's no reason to assert "If A , then B "' and that's true.

Alas, the material theorist is now back in the position of being unable to explain Thesis #2. Her problem was $A \supset \neg B$ didn't deny $A \supset B$. One solution was to change $A \supset \neg B$ such that it explicitly denied $A \supset B$, in other words, change $A \supset \neg B$ to $\neg(A \supset B)$. But since it turns out this explicit denial is really just a refusal to assert $A \supset \neg B$, we are back at square one, in the position of having $A \supset \neg B$ not denying $A \supset B$ – and being unable to explain Thesis #2.

2. The possible-worlds theory

Recall that according to the possible-worlds theorist, an indicative conditional 'If Q , then R ' is true in a possible world w if and only if R is true in all the Q -worlds which are most similar to w (and vacuously true when there's no Q -world) (Stalnaker, 1981, pp. 46-7; see also Lewis, 1973).

A Q -world is one in which the antecedent Q is true. For example, the conditional 'If Nicholas is at the office, then he's in London' is true in the actual world just in case Nicholas is in London in all most-similar worlds in which Nicholas is at the office.

As we saw in Chapter 1, possible-worlds theorists disagree on the number of most-similar antecedent-worlds to any given possible world. Some such as Stalnaker (1981) hold that there can be only one (p. 46). We're calling them the single-world theorists. Others such as Lewis (1973) hold that there can be more than one (pp. 97-8). We're calling them the multiple-world theorists. (Remember that since Lewis's theory is about counterfactual conditionals and here we are looking at indicative ones, the multiple-world theorist is an imagined version of Lewis.)

In this section, I show how both possible-worlds theories can, like the material theory, explain Thesis #3. I also show how, while the single-world theorist can explain #2 but not #1, the multiple-world theorist can explain neither.

2.1. *Accepting Thesis #3: modus tollens is valid*

Carroll's argument doesn't falsify Thesis #3. There's no question whether *modus tollens* is truth-preserving, i.e. where the premises of an argument are those of a *modus tollens* and true, the conclusion will also be true.

Suppose that both premises of a *modus tollens* argument are true. This means that in the most-similar worlds in which the antecedent of the first premise is true, the consequent of the first premise is also true. Likewise, in the most-similar antecedent worlds in which the antecedent of the second premise is true, the consequent of the second premise is also true. But since the second premise is the contradictory of the consequent of the first, the antecedent of the first premise can't be true in the actual world. We assumed that the first premise is true in the actual world and we see that it has a false consequent there. It follows that the conclusion is true. The conclusion is the contradictory of the antecedent of the first premise and the antecedent of the first premise is false.

2.2. *Facing trouble with the single-world theory*

The single-world theory can explain Thesis #2 but not Thesis #1. On this theory, as long as A is possible (i.e. there exists a world in which A is true), $A \rightarrow \neg B$ is the contradictory of $A \rightarrow B$; in other words, $A \rightarrow \neg B$ and $\neg(A \rightarrow B)$

are equivalent; in yet other words, the principle of Conditional Excluded Middle, which we saw in Chapter 1 (subsection 2.1.) and according to which $(A \rightarrow B) \vee (A \rightarrow \neg B)$ is a tautology, is true. The proof is as follows:

First, assuming A is possible, $A \rightarrow \neg B$ implies $\neg(A \rightarrow B)$. For $A \rightarrow \neg B$ to be true, it must be the case that, in the nearest antecedent-world (and there will only be one), B is false. It follows that $\neg(A \rightarrow B)$ is true, since for $\neg(A \rightarrow B)$ to be true, it must be the case that $A \rightarrow B$ is false, and for $A \rightarrow B$ to be false, it must be the case that, in the nearest antecedent-world, B is false – and, as we have seen, B is indeed false in that world.

And second, $\neg(A \rightarrow B)$ implies $A \rightarrow \neg B$. For $\neg(A \rightarrow B)$ to be true, it must be the case that $A \rightarrow B$ is false. For $A \rightarrow B$ to be false, it must be the case that, in the nearest antecedent-world, B is false, and so $\neg B$ is true. It follows that $A \rightarrow \neg B$ is true, since in the nearest antecedent-world, A is, of course, true, and, as we have seen, in that world, $\neg B$ is true.

So the theory can account for Thesis #2. It can't do the same for Thesis #1. According to the theory, the argument is valid. For suppose the argument is invalid (i.e. suppose the premises are true and the conclusion false in the actual world). Then in the actual world, C is true. Since the actual world is a most-similar antecedent-world to itself in which C is true, it follows that $A \rightarrow B$ is true there. Now, suppose v is a most-similar antecedent-world to the actual world in which A is true (so long as A is possible, we can make this assumption). Since $A \rightarrow \neg B$ is true in the actual world, $\neg B$ is true in v . But since $A \rightarrow B$ is also true in the actual world, B is true in v . So we have both B and $\neg B$ in v , which is a contradiction. On this theory, the argument must be

valid and can't account for the thesis that we might assert the premises but we wouldn't assert the conclusion.

2.3. Facing trouble with the multiple-world theory

While holding that there can be only one most-similar antecedent-world yields bad results, holding that there can be more than one such world yields even worse ones. The multiple-world theory can't explain either Thesis #1 or Thesis #2. With the principle that there can be more than one most-similar antecedent-world, the second premise isn't the contradictory of the consequent of the first. Indeed, the principle of Conditional Excluded Middle isn't true. $A \rightarrow B$ and $A \rightarrow \neg B$ could both be false at the same time. When trying to evaluate the truth value of $A \rightarrow B$ in the actual world, we look for the most-similar A -worlds. Among them, we might find two equally-similar worlds where B is true in one and false in the other. If that is the case, then we have a scenario where $A \rightarrow \neg B$ and $A \rightarrow B$ are both false.

And here too, the argument is valid, so the multiple-world theory can't explain Thesis #1 either. The proof of validity we saw in the context of the single-world theory is applicable here. No step in it relied on there being no more than one antecedent-world. Moreover, in rejecting one premise – or even accepting the conclusion – the possible-worlds theorist would be abandoning any attempt at explaining Thesis #1.

We see that neither kind of possible-worlds theories can explain all three theses. One way forward might be to abandon the concept of possible worlds and truth. The suppositional theory does both. Let's turn to it now.

3. *The suppositional theory*

Recall that according to suppositional theorists such as Adams (1975) and Edgington (1995 and 2014) the Equation holds true: writing ‘ P ’ for probability,

$$P(Q \rightarrow R) = P(R | Q) \text{ provided that } P(Q) \neq 0.$$

In prose, the probability of a conditional $Q \rightarrow R$ is equal to the conditional probability of the consequent R , on the supposition that the antecedent Q is true so long as the probability of Q isn’t equal to zero.

So, according to the Equation, the probability of the conditional ‘If Josie’s at home, then she’s in New York’ is equal to the conditional probability of ‘Josie’s in New York’ given ‘she’s at home’.

The Equation implies that a conditional is non-propositional. If a conditional were propositional, then for any conditional there would be a proposition whose probability were equal to the conditional probability of its consequent given its antecedent. But there isn’t (Lewis, 1976). The two propositions don’t combine into a single proposition we judge as probably true when we judge the second to be probably true on the supposition of the first (Edgington, 1995, p. 305).

So, according to the suppositional theory it’s not the case that ‘If Josie’s at home, then she’s in New York’ can be true or false. As we’ll see, this framework allows the suppositional theorist to explain the plausibility of all three theses.

3.1. Rejecting Thesis #2: the argument is equivalent to an instance of modus tollens

The suppositional theorist rejects #2 on the grounds that the thesis requires that conditionals be propositional. As we've seen, she takes conditionals to be non-propositional. For the second premise and consequent of the first to be logical contradictories, when one is true, the other must be false – and conditionals aren't true or false. This being so, according to her, Carroll's argument is no closer to being equivalent to an instance of *modus tollens* than the following one in which the second premise is an imperative:

If it's sunny, then do go outside!

Don't go outside!

Therefore, it's not the case that it's sunny.

Here, as in Carroll's argument, the second premise – 'Don't go outside!' – is non-propositional and can't be the logical contradictory of the consequent of the first premise. Since, to accept #2, the suppositional theorist must find the second premise in Carroll's argument propositional and she doesn't, the theorist doesn't accept #2.

Now, according to her account, the second premise $A \rightarrow \neg B$ and the consequent of the first $A \rightarrow B$ are – if not logically, then – *probabilistically* contradictory. We define two formulae as probabilistically contradictory where necessarily the sum of their probabilities is equal to one and the probability of the conjunction is zero. Formally, for any formulae Q and R , Q and R are probabilistically inconsistent if and only if necessarily $P(Q) = 1 - P(R)$ and $P(Q \wedge R) = 0$. In Carroll's argument 'If Allen is out, then Brown is out' and 'If Allen is out, then it's not the case that Brown is out' are probabilistically contradictory, since $P(A \rightarrow \neg B) = 1 - P(A \rightarrow B)$ and $P((A \rightarrow \neg B) \wedge (A \rightarrow B)) = 0$ provided that $P(A)$ isn't zero. Because of this, we can't

resolve the trilemma by pointing out that $A \rightarrow \neg B$ and $A \rightarrow B$ are *not* contradictory – probabilistically, they are (Adams, 1975, p. 42).

To derive a proof, the suppositional theorist would rely on the Conditional Contradiction and the Equation. Recall the lemma we proved in Chapter 2 (subsection 3.2.) showing that, assuming the probability of Q isn't zero, the probability of not- R given Q is one minus the probability of R given Q . Now, here's how the suppositional theorist would prove that the second premise and the consequent of the first are probabilistically contradictory, provided that $P(A)$ isn't zero:

1. $P(B | A) = 1 - P(\neg B | A)$ by Conditional Contradiction
2. $P(A \rightarrow B) = 1 - P(A \rightarrow \neg B)$ from 1 by the Equation

That said, whether the suppositional theorist can say that – beyond the second premise and consequent of the first being probabilistically contradictory – the argument is equivalent to an instance of *modus tollens* depends on her definition of *modus tollens*. She can define it in two ways: one more-inclusive, another less-inclusive.

According to each definition, of course, the argument will have two premises and a conclusion; the second premise and antecedent of the first will be contradictories; likewise the conclusion and the antecedent of the first premise. Differentiating the more- and less-inclusive definitions, however, is whether the second premise and consequent of the first (and, albeit less relevant here, the conclusion and antecedent of the first premise) must be propositional. According to the more-inclusive definition, it's not the case that they must.

If she adopts the more-inclusive definition, then she can say that Carroll's argument is *probabilistically* equivalent to a *modus tollens*. Indeed, this definition includes instances of complex *modus tollens*, arguments such as Carroll's with an embedded conditional. If she adopts the less-inclusive definition, then she can't say that Carroll's argument is probabilistically equivalent to a *modus tollens*. Indeed, this definition includes only instances of simple *modus tollens*, arguments without an embedded conditional or imperative.

Granted, considering #2, the one definition might appear more attractive to us than the other inasmuch as the more-inclusive definition allows the suppositional theorist to explain the plausibility of the thesis and the less-inclusive one doesn't. However, as we'll see when considering #3, the opposite will be true: the less-inclusive definition might appear more attractive than the more-inclusive one.

3.2. *Rejecting Thesis #1: the argument is invalid*

The suppositional theorist would deny that the argument is invalid – i.e. deny that it's possible for the premises of the argument to be true, while the conclusion is false – on the grounds that it's impossible for the premises to be true at all. Instead, she'd argue that the argument is *probabilistically* invalid – i.e. it's possible for the premises to be probable (say, have a probability over 0.5), while the conclusion is improbable (say, have a probability under 0.5) (Adams, 1975, p. 1)). She would defend the surrogate thesis over the original one, claiming that conditionals don't have truth values but probabilities.

To prove that the argument is probabilistically invalid, we show that it's not the case that the uncertainty of the conclusion can't exceed the sum of the uncertainties of the premises. And to calculate the uncertainties, first we calculate the probabilities of the premises and the conclusion.

Suppose (i) $P(C)$ is very high; (ii) $P(\neg B \wedge C)$ is zero; and (iii) $P(\neg B | A)$ is very high (as illustrated in Figure 1). From (iii) and the Equation, it follows that the probability of the second premise is high. (The area representing the probability of $\neg B$ takes up a large proportion of the area representing the probability of A). From (ii), it follows that $P(C \wedge A \wedge \neg B) \div P(C \wedge A)$ is zero. From this and the Ratio Formula, it follows that $P(\neg B | (C \wedge A))$ is zero. From this and the Equation it follows that $P((C \wedge A) \rightarrow \neg B)$ is zero. And from this and the result in the last subsection about probabilistically contradictory conditionals, it follows that the probability of the first premise is one. (The intersection of the areas representing the probabilities of A and C is entirely covered by the area representing the probability of B .) Finally, from (i) it follows that $P(\neg C)$ is very low.

Translating probability into uncertainty, we see that the uncertainty of the conclusion can exceed the sum of the uncertainties of the premises. The sum of the uncertainties of the premises is close to zero: as the probability of the first premise is one, the uncertainty of the first premise is zero; and, as the probability of the second premise is close to one, the uncertainty of the second premise is close to zero. However, as the probability of the conclusion is itself very close to zero, the uncertainty of the conclusion is very close to one – exceeding the sum of the uncertainties of the premises.

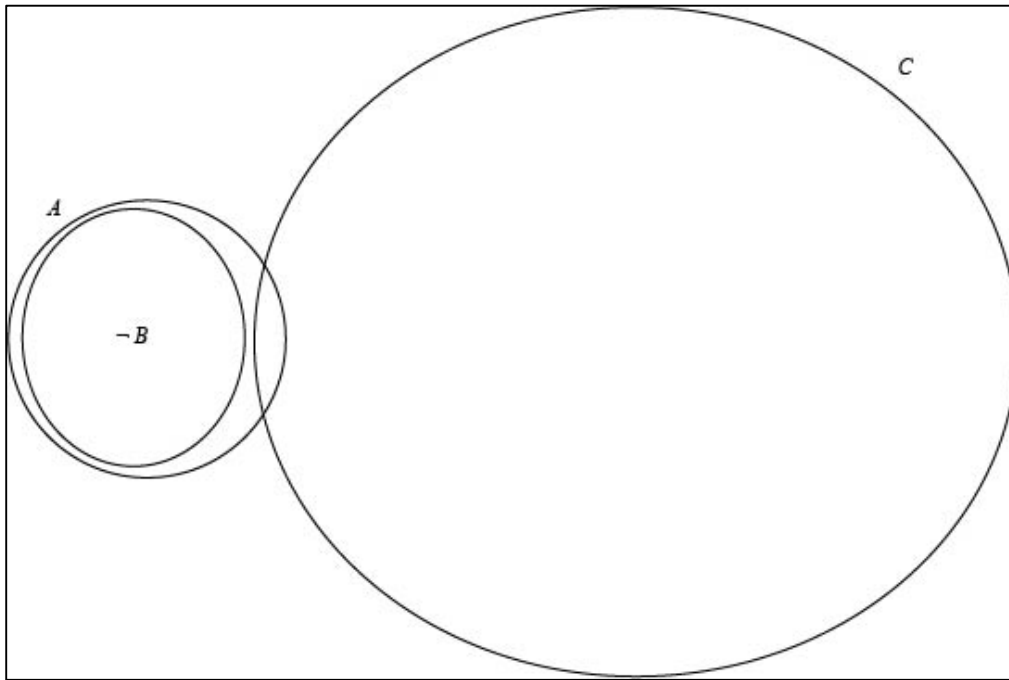


Figure 2: Venn diagram of the Barbershop Paradox

As we see, the argument is probabilistically invalid. In other words, we can be justified in accepting that ‘If Carr is out, then if Allen is out, then it’s not the case that Brown is out’ and ‘If Allen is out, then Brown is out’ while rejecting ‘It’s not the case that Carr is out’.

3.3. *Rejecting Thesis #3: modus tollens is valid*

Depending on how she defines *modus tollens*, the argument form is or isn’t *probabilistically* valid according to the suppositional theorist. If she defines it as being equivalent to simple *modus tollens*, then it’s probabilistically valid. If she defines it as being equivalent to complex *modus tollens*, then it isn’t.

To prove the probabilistic invalidity of complex *modus tollens*, the theorist could refer to the counterexample we saw in the previous subsection (3.2).

To prove the probabilistic validity of simple *modus tollens*, she'd need to prove that for any argument with the form $Q \rightarrow R; \neg R; \therefore \neg Q$ where we replace Q and R with propositions, it's not the case the uncertainty of the conclusion can be greater than the sum of the uncertainties of the premises. And she can do that by drawing on the proof she derived in Chapter 2 (subsection 3.2) showing that simple *modus ponens* was probabilistically valid:

$$1. U(\neg R) + U(Q \supset R) \geq U(\neg Q)$$

because $\neg R$ and $Q \supset R$ entail $\neg Q$

$$2. U(Q \supset R) \geq U(\neg Q) - U(\neg R)$$

from 1 by algebra

$$3. U(Q \rightarrow R) \geq U(\neg Q) - U(\neg R)$$

from 2 and the fact that an indicative conditional probabilistically entails a material one (see line 10 of proof in Chapter 2 (subsection 3.2.)) by algebra

$$4. U(Q \rightarrow R) + U(\neg R) \geq U(\neg Q)$$

from 3 by algebra

And she would have proved what she wanted. Line 4 says that simple *modus tollens* is probabilistically valid for the indicative conditional according to the suppositional theorist.

So just as she did while responding to the trilemma giving rise to the Election Paradox, the suppositional theorist faces a dilemma: explain the thesis according to which the argument is equivalent to an instance of a classically valid argument form; or explain the thesis according to which that

classically valid argument form is valid. Just as before in the case of *modus ponens*, the definition of *modus tollens* she adopts will allow her to explain either #2 or #3 – but not both. Adopting the more-inclusive definition, she can explain #2 but not #3. Adopting the less-inclusive definition, she can explain #3 but not #2.

In the following section, we turn to a theory which agrees with the suppositional one on our being prepared to assert a conditional where we find it has a high probability but disagrees with it on conditionals being non-propositional. In so doing, the hybrid theory succeeds in explaining the theses without facing the problems the suppositional theory does when it comes to the truth-conditional theory of meaning.

4. The hybrid theory

The hybrid theory combines elements from the material theory and the suppositional theory. On the one hand, like the material theory, the hybrid one concerns itself with truth. On the other hand, like the suppositional theory, the hybrid one concerns itself with probability. For example, an indicative conditional is true where the corresponding material one is; it's assertible where the conditional probability of the consequent given the antecedent is high. Recall, the material theory's concept of assertability with an 'a' and the hybrid theory's concept of assertability with an 'i' aren't the same.

The hybrid theorist can explain all three theses.

4.1. Accepting Thesis #1: the argument is invalid

The hybrid theorist accepts that valid arguments might not preserve assertibility: ‘I hope I have already accustomed you to the idea that a valid inference may lead from the highly assertible to the highly unassertible’ (Jackson, 1987, p. 133). To explain Thesis #1, however, the hybrid theorist must show that an invalid argument isn’t assertibility-preserving.

Some invalid arguments *are* assertibility-preserving. The following is a case in point:

Alley is good at mathematics but not at chess.

Therefore, being good at mathematics contrasts with not being good at chess.

Here, the argument preserves assertibility but not truth. If we are prepared to assert the premise, then we’ll be prepared to assert the conclusion. The ‘but’ in the premise conversationally implies the contrast which the conclusion mentions. However, if the premise is true, it’s not necessarily the case that the conclusion will also be true.

Now, the theory *can* explain Thesis #1. As we’ve already seen, just as the argument doesn’t preserve probability, according to the suppositional theory, it doesn’t preserve assertibility, according to the hybrid one. As we saw in the previous chapter, the assertibility of $Q \rightarrow (R \rightarrow S)$ is exactly the same as that of the corresponding sentence of the form $(Q \wedge R) \rightarrow S$ (Jackson 1987, p. 130). According to the hybrid theory, Import-Export preserves not only truth but also assertibility. Moreover, it’s possible for the premises $(C \wedge A) \rightarrow B$ and $A \rightarrow \neg B$ to be highly assertible while the conclusion $\neg C$ isn’t.

Applying Adams' Thesis to the suppositional theory's proof of probabilistic invalidity shows this.

4.2. Rejecting Thesis #2: the argument is equivalent to an instance of modus tollens

The theory can also explain the plausibility of Thesis #2. Granted, according to the hybrid theory, they aren't logically contradictory. It's not the case that $A \rightarrow \neg B$ is true if and only if $A \rightarrow B$ is false. The conditionals can both be true when A is false. However, just as the suppositional theorist could show that the probability of the first was 1 minus the probability of the second, the hybrid theorist can show that the assertibility of the first is 1 minus the assertibility of the other (and vice versa).

Just as the probability of $A \rightarrow B$ was 1 minus the probability of $A \rightarrow \neg B$, the assertibility of $A \rightarrow B$ is 1 minus the assertibility of $A \rightarrow \neg B$. The hybrid theorist's proof proceeds the same way as the suppositional theorist's up to line 6 before diverging. Line 7 follows from line 6 by – not the Equation but – Adams' Thesis. According to Adams' Thesis, the assertibility of 'If Q , then R ' is the conditional probability of R given Q . So, the assertibility of the conditional 'If it's raining, then it's cloudy' is equal to the conditional probability of 'it's cloudy' given 'it's raining'.

This proof is sufficient to explain the thesis. Indeed, it goes beyond the material theory in saying not only that $A \rightarrow \neg B$ and $A \rightarrow B$ aren't co-assertible. On the material theory, if we can assert the one, then we can't assert the other – but this doesn't mean that we can go so far as to deny the other. In contrast, on the hybrid theory, if we can assert $A \rightarrow \neg B$, then we can deny A

$\rightarrow B$ and vice versa. As Jackson writes, the two conditionals ‘have a kind of “see-saw” relationship. As the assertibility of one goes up, the assertibility of the other goes down’ (Jackson, 1987, pp. 12-3).

4.3. *Accepting Thesis #3: modus tollens is valid*

Moreover, the theory can explain Thesis #3, according to which *modus tollens* is valid. Carroll’s argument doesn’t falsify the claim that *modus tollens* preserves truth. Like the material theorist, the hybrid one would point out that the argument doesn’t strictly have the form of a *modus tollens*. The second premise isn’t the logical contradictory of the consequent of the first.

Note that the hybrid theorist’s explanation is different from the suppositional theorist’s. The hybrid theorist isn’t saying that the second premise isn’t the logical contradictory of the consequent of the first because the second premise is non-propositional. Rather, the hybrid theorist is saying that the second premise and the consequent of the first can both be true when A is false.

That said, Carroll’s argument falsifies the claim that *modus tollens* preserves assertibility. When we reformulate the premises and conclusion such that they have the relevant logical form, namely

$$C \rightarrow (A \rightarrow B)$$

$$\neg (A \rightarrow B)$$

$$\therefore \neg C$$

the first premise is assertible, so too is the second – ‘It’s not the case that if Allen is out, then Brown is in’ – but still the conclusion isn’t. Despite it being assertible, we might nonetheless reject the second premise in its new

formulation. $\neg(A \rightarrow B)$ is logically equivalent to $A \wedge \neg B$ – ‘Allen is out and it’s not the case that Brown is in’ – and on our way to the barbershop, we are agnostic about the current whereabouts of Allen and Brown.

5. Concluding

In this chapter, I proceeded largely from first principles in analysing Lewis Carroll’s alleged counterexample to *modus tollens*. I focused on the responses four theories of the indicative conditional would offer to the trilemma Carroll’s argument presents. According to the material theory, (i) Carroll’s argument is logically invalid and doesn’t preserve assertability (with an ‘a’); (ii) the second premise isn’t the logical contradictory of the consequent of the first and, while the theory falls short of showing that if you can assert the one, then you can deny the other, the theory can show that the two are not co-assertable; and (iii) *modus tollens* is logically valid and preserves assertability.

According to the single-world theory, (i) Carroll’s argument is logically valid; (ii) the second premise *is* the logical contradictory of the consequent of the first; and (iii) *modus tollens* is logically valid. The multi-world theorist agrees when it comes to (i) and (iii) but disagrees when it comes to (ii). According to her, the second premise isn’t the logical contradictory of the consequent of the first.

According to the suppositional theory, (i) Carroll’s argument is probabilistically invalid; (ii) while the second premise and consequent of the first aren’t logical contradictories, they are probabilistic ones; and (iii) whether *modus tollens* is probabilistically valid or Carroll’s argument is an instance of *modus tollens* depends on her definition of the argument form: if it extends to

include Carroll's argument, then *modus tollens* isn't probabilistically valid while if it doesn't extend to include his argument, then *modus tollens* is probabilistically valid.

Finally, according to the hybrid theory, (i) Carroll's argument is logically invalid and doesn't preserve assertibility (with an 'i'); (ii) the second premise and consequent of the first aren't logical contradictories but the assertibility of the one is high if and only if the assertibility of the other is low; and (iii) *modus tollens* preserves truth if not assertibility.

Carroll ends his 'ornamental presentment' with the young Cub saying: 'How long this argument might have lasted, I haven't the least idea. I believe either of [Uncle Jim or Uncle Joe] could argue for six hours at a stretch. But, just at this moment we arrived at the barber's shop; and, on going inside, we found – ' (Carroll, 1894, p. 437).

CHAPTER 4

ON PREMISE SEMANTICS AND APPARENT COUNTEREXAMPLES TO *MODUS PONENS* AND *MODUS TOLLENS*

This chapter is about premise semantics for the indicative conditional and apparent counterexamples to *modus ponens* and *modus tollens*. Specifically, it's about how a theory designed to account for conditionals with embedded modals deals with instances of the two classically valid argument forms where a premise is a conditional with an embedded modal.

The chapter proceeds in three parts. In section 1, I offer an exegesis of premise semantics (drawing on Kratzer (2012)). In section 2, I argue that, when it comes to counterexamples to *modus ponens* and *modus tollens* with (overt) modal verbs and adverbs, the theory gives us some results we want (e.g. invalidating Kolodny and MacFarlane's *Miners* examples (2010)) and others we don't (e.g. validating Cantwell's *Lottery* (2008)) – a point which the literature previously overlooked. In section 3, I analyse the Election and Barbershop paradoxes through the lens of premise semantics. And in section 4, I conclude.

1. Exegesis

According to this theory, conditionals are modal restrictors, so we're going to talk now about modal semantics. In the possible-worlds semantics we are building here, we identify propositions with sets of possible worlds. Indeed, a proposition p is true in a possible world w in the set of possible worlds W if and only if w is an element of p . Formally,

A proposition p is true in $w \in W$ if and only if $w \in p$.⁸

A conversational background is a function from worlds to premise sets – or sets of propositions (Kratzer, 2012, p. 20). A conversational background can be realistic, totally realistic or empty.

Realistic

A realistic conversational background is a function f such that for any world $w \in W$, $w \in \cap f(w)$. That is, f assigns to every possible world a set of propositions that are true in it (Kratzer, 2012, p. 32).

Totally realistic

A totally realistic conversational background is a function f such that for any $w \in W$, $w \in \cap f(w) = \{w\}$. That is, f assigns to any world a set of propositions that characterises it uniquely. For each world, there are many ways of characterizing it uniquely (Kratzer, 2012, p. 33).

Empty

The empty conversational background is a function f such that for any $w \in W$, $f(w) = \emptyset$. Since $\cap f(w) = W$ if $f(w) = \emptyset$, empty conversational backgrounds are also realistic (Kratzer, 2012, p. 33).

We can think of conversational backgrounds in terms of not only premise sets

⁸ In this chapter, I'm using lowercase letters (such as p) to denote sets of possible worlds, which we call propositions. In previous chapters, I was using uppercase letters (such as A , B , C) to denote truth-apt sentences. (The exception being where the suppositional theorist takes e.g. R to denote an indicative conditional. Of course, conditionals aren't truth-apt according to her.) So the use here of lowercase letters following the use before of uppercase ones represents a change in – not notation but – subject.

but also the kinds of witnesses in a criminal trial. For example, a conversational background containing only premises we know is realistic: if we know the premises, then they're true. The witness corresponding to the realistic background is one who tells the truth (albeit not the whole truth: there are premises we don't know).

Next, a conversational background containing enough premises to single out the actual world is totally realistic. The witness corresponding to the totally realistic background is one who tells the whole truth and nothing but the truth.

And finally, a conversational background containing as premises the laws of an anarchic society is empty: an anarchic society has *no* laws. The witness corresponding with the empty background is one who remains silent.

And, when analysing conditionals, we consider two kinds of conversational backgrounds: a modal base and an ordering source. Let's define each in turn.

Modal base

Phrases such as 'in view of what we know' identify a modal base (Von Stechow and Heim, 2011, 41). Our knowledge differs from one possible world to another. In one possible world, I might know that Mongolia has a capital city but not know its name. In another possible world, I might know that Mongolia has a capital and that the name of the capital city of Mongolia is Ulaanbaatar.

In each world, what we know consists of a set of propositions. So in our example, in the first possible world, the set includes the proposition (1) where

(1) Mongolia has a capital city.

In the second, the set includes the propositions (1) and (2) where

(2) The name of the capital city of Mongolia is Ulaanbaatar.

The modal base then is that function which assigns to every possible world the set of propositions we know in that world.

Returning to our definition of truth in a world, we identify (2) with the set of possible worlds where (2) is true. So ‘The name of the capital city of Mongolia is Ulaanbaatar’ is false in the first world but true in the second. The first world isn’t an element of the proposition. The second one is.

Having identified propositions with sets of possible worlds, we can get the set of worlds in which all the propositions in a set are true. For example, we can identify the worlds which are compatible with everything we know. To define compatibility formally,

A proposition p is compatible with a set of propositions A if and only if $A \cup \{p\}$ is consistent (Kratzer, 2012, p. 10).

Returning to our example once more and assuming that (1) is the only thing we know in the first world, what we know in the first world is compatible with the second world. (1) is compatible with the set containing (1) and (2).

The modal base might not always be explicit. Consider

(3) (In view of what we know,) Mongolia has a capital city.

While the ‘in view of’-phrase explicitly signals the intended modal base, we can still infer the epistemic base which the speaker assumes, should she omit the phrase. Other clues in the discourse might help us here (Von Stechow and Heim, 2011, p. 41).

Ordering source (from Von Fintel and Heim, 2011, pp. 59-61)

An ordering source is a kind of conversational background – and as such, it’s a function from possible worlds to premise sets. This kind of conversational background ranks ‘worlds according to how close they come to the normal course of events in the world of evaluation, given a suitable normalcy standard’ (Kratzer, 2012, p. 39). In other words, a set of propositions A can induce an ordering \leq_A on a set of possible worlds \mathcal{W} such that a possible world w is at least as close as a possible world z to the ideal A determines if and only if all propositions of A that are true in z are true in w as well (Kratzer, 2012, pp. 39-40).

One reason we need a second conversational background is to provide for deontic cases. Consider the proposition

(4) Amarbold must pay a fine.

The truth of such a proposition depends both on facts about the law and facts about Amarbold’s actions. For example, we’ll judge (4) as true if (i) the law states that bringing durian on a bus is fined; and (ii) Amarbold has brought durian on a bus. Conversely, we’ll judge (4) as false if it’s not the case that bringing durian on a bus is fined; or Amarbold didn’t bring durian on a bus.

Here, using one conversational background isn’t sufficient. Unless people had not previously infringed the law, the conversational background would be empty, (4) would be vacuously true and this result would contradict our intuition that (4) is true if (i) and (ii) are true. By using two conversational backgrounds, we can make explicit the difference between the ways in which facts about Amarbold’s actions and facts about the law impact the truth conditions of propositions such as (4).

The modal base f assigns to any evaluation world a set of propositions describing the relevant circumstances, say, Amarbold's actions. Since in our evaluation world Amarbold brought durian on a bus, the modal base will assign the proposition that 'Amarbold brought durian on a bus' to this world. And the ordering source g will assign to any evaluation world a set of propositions P whose truth the law demands – and which we can use to order the worlds in the modal base.

For any pair of worlds u and v we say that u comes closer than v to the ideal $g(w)$ establishes if and only if the set of propositions from $g(w)$ that are true in v is a proper subset of the set of propositions from P that are true in u .

For our example, suppose that P contains the following two propositions: (i) nobody brings durian on a bus; and (ii) anybody who brings durian on a bus pays a fine. Any world in the modal base in which Amarbold pays a fine will count as better than an otherwise similar world where he doesn't. In other words, (4) claims that in the best worlds (among those where Amarbold brings durian on a bus), he pays a fine.

Using the definitions we've seen so far, we can prove that if g is a realistic conversational background, then for all worlds w and v , w is at least as close as v to an ideal which the ordering source determines. We'll rely on this lemma later.

Lemma 1: If g is a realistic conversational background, then for all worlds w and v , w is at least as close as v to an ideal which the ordering source $g(w)$ determines.

Proof: Suppose g is a realistic conversational background, so from the definition of a realistic conversational background, w is compatible with the

conversational background (i.e. every proposition in the conversational background $g(w)$ is true in w). But suppose p is in the conversational background $g(w)$ and p is true in v . Since p is in the conversational background $g(w)$ and w is compatible with that background (i.e. every proposition in that background is true in w), p is true in w . So all the propositions in the conversational background that are true in v are true in w . So w is at least as close as v to an ideal which the ordering source $g(w)$ determines.

So much for defining the concepts of world w , modal base f and ordering source g , and proving a lemma. We turn to defining additional modal relations (such as necessity and possibility) and the truth conditions of the indicative conditional.

Necessity and possibility

‘Simplifying slightly, a proposition is a necessity just in case it is true in all accessible worlds that come closest to the ideal determined by the ordering source’ (Kratzer, 2012, p. 40). So ‘Amarbold must pay a fine’ would be a necessity if in all worlds that came closest to the ideal – where (i) nobody brings durian on a bus; and (ii) anybody who brings durian on a bus pays a fine – it were true that Amarbold pays a fine. Simplifying less, a proposition p is a necessity in w with respect to f and g if, for all worlds u in the set of worlds where all propositions in the modal base are true, there’s a world v in that set of all worlds such that:

- (i) v is at least as close as u to an ideal which the ordering source $g(w)$ determines

and

(ii) for all worlds z in the set of worlds where all propositions in the modal base $f(w)$ are true: if z is at least as close as v to the ideal which the ordering source $g(w)$ determines, then p is true in z .

Again, using this definition, we can prove that for any world w and realistic backgrounds f and g and proposition p , if ‘necessarily p ’ is true in w with respect to f and g , then p is true in w with respect to f and g . We’ll rely on this lemma later as well.

Lemma 2: For any world w and realistic backgrounds f and g , if ‘necessarily p ’ is true in w with respect to f and g , then p is true in w with respect to f and g .

Proof: Suppose ‘necessarily p ’ is true in w with respect to f and g . Then from the definition of a realistic conversational background, all propositions in the modal base $f(w)$ are true in w . So from the definition of ordering source, there’s a world v in the set of worlds where all propositions in the modal base are true in it and:

(i) v is at least as close as w to the ideal which the ordering source $g(w)$ determines

and

(ii) for all worlds z in the set of worlds where all propositions in the modal base $f(w)$ are true: if z is at least as close as v to the ideal which the ordering source $g(w)$ determines, then p is true in z .

From (ii) and the fact that all propositions in the modal base $f(w)$ are true in w it follows that if v is at least as close as w to the ideal which the ordering source $g(w)$ determines, p is true in w . But from Lemma 1, we know

v is at least as close as w to the ideal which the ordering source $g(w)$ determines. So p is true in w . QED.

Next, turning to possibility: ‘A proposition is a *possibility* in w with respect to f and g if and only if its negation (that is, its complement) is not a necessity in w with respect to f and g ’ (Kratzer, 2012, p. 40). So ‘Amarbold must pay a fine’ is possible if in at least one world that came closest to the ideal, Amarbold pays a fine.

Finally, we turn to conditionals.

Indicative conditional

According to premise semantics, for any propositions p and q , ‘If p then (modal) q ’ is true in w with respect to f and g if and only if ‘(modal) q ’ is true in w with respect to f^* and g , where for all worlds u , $f^*(u) = f(u) \cup \{p\}$, or the original modal base with the addition of p (Kratzer, 2012, p. 94).⁹ Note that where the conditional is a bare conditional ‘If p , then q ’ (i.e. where there is no modal), we add a necessity modal such that it reads ‘If p , then necessarily q ’.

Beyond indicative conditionals, different kinds of conditionals depend on the settings for f and g . For example:

Material implication

⁹ Later, the ‘+’ in f_+ will serve the same purpose as the * in f^* , viz. to represent the original modal base \cup the relevant conditional’s antecedent, i.e. the new modal base. I’ll be using + rather than * so that later, when I prove Import-Export (in Theorem 3 in subsection 3.1.2.), I can speak of two new modal bases separately before showing that they are one and the same.

Characterising the material conditional are a totally realistic modal base f (again, where f assigns to any world a set of propositions that characterises it uniquely) and an empty ordering source g (Kratzer, 2012, pp. 65-6).

Proof for conditional ‘If p , then q ’ (Kratzer offers a ‘Sketch of proof’ and, in what follows, I fill out her sketch):

Case 1: Suppose p is true in w . Then $\cap f_+(w) = \{w\}$. In other words, the set of all worlds where all the propositions in the new modal base are true contains only w . The set of all worlds where all the propositions in the original modal base are true contained only w (because we’re taking f to be a totally realistic modal base); the set of all worlds where all the propositions in the new modal base are true is equal to the set of all worlds where all the propositions in the original modal base *and* p are true; and all the propositions in the original modal base are true in w and, with the supposition, so too is p . But then we get the following biconditional:

q is a necessity in w with respect to f_+ and g if and only if q is true in w .

To prove the left-to-right proposition in the biconditional:

Suppose q is true in w to prove that q is a necessity in w with respect to f_+ and g . According to the definition of necessity, for q to be a necessity in w with respect to f_+ and g , it must be the case that for all worlds u (in the set of worlds where all propositions in the modal base are true), there’s a v such that

(i) v is at least as close as u to an ideal which the ordering source $g(w)$ determines; and

(ii) for all worlds z in the set of worlds where all propositions in the modal base $f(w)$ are true: if z is at least as close as v to the ideal which the ordering source $g(w)$ determines, then q is true in z .

And that *is* the case for all worlds u . There's only one world in the set of worlds where all propositions in the modal base are true, namely w ; and for it, there does exist a world such that (i) and (ii), namely w again. When it comes to (i), w is at least as close as itself to an ideal which the empty ordering determines; and when it comes to (ii), again, there's only one world in the set of worlds where all the propositions in the modal base are true, namely w , it's at least as close as itself to the ideal which the ordering source determines and, given our assumption, q is true in w . So q is a necessity in w with respect to f^+ and g .

Next, to prove the right-to-left proposition in the biconditional ' q is a necessity in w with respect to f^+ and g if and only if q is true in w ':

Suppose q is a necessity in w to prove that q is true in w . Again, according to Lemma 2, for q to be a necessity in w with respect to f^+ and g , q must be true in w . So, given our assumption, q is true in w .

Case 2: Suppose p is false in w . Then $\cap f^+(w) = \emptyset$. In other words, the set of all worlds where all the propositions in the modal base are true is empty. But then q is trivially a necessity in w with respect to f^+ and g . According to the definition of necessity, q is a necessity in w if, all worlds u (in the set of worlds where all propositions in the modal base are true), there's a v such that

(i) v is at least as close as u to an ideal which the ordering source $g(w)$ determines

and

(ii) for all worlds z in the set of worlds where all propositions in the modal base $f(w)$ are true: if z is at least as close as v to the ideal which the ordering source $g(w)$ determines, then q is true in z . And here, there are no worlds u .

*Multiple-world-theorist conditionals*¹⁰

Characterising a conditional as the multiple-world theorist understands it is an empty modal base f and a totally realistic ordering source g (Kratzer, 2012, p. 66). To spell this out, a conditional ‘If α , then (necessarily) β ’ is true in w with respect to an empty modal base f and a totally realistic ordering source g if and only if, for all worlds u in the set of worlds where all propositions in the modal base *and* α are true, there’s a world v in that set of all worlds such that:

(i) v is at least as close as u to an ideal which the ordering source $g(w)$ determines

and

(ii) for all worlds z in the set of worlds where all propositions in the modal base $f(w)$ *and* α are true: if z is at least as close as v to the ideal which the ordering source $g(w)$ determines, then β is true in z .

Here, we can interpret the ‘at least as close as ... to the ideal which the ordering source determines’ relation (the symbol for which we saw earlier being $\leq_{g(w)}$) in premise semantics as denoting the similarity relation in the multiple-world theory.

¹⁰ Kratzer (2012) calls them counterfactuals, following Lewis (1973). But we can call them multiple-world-theorist conditionals, following our imaginary Lewis (which I introduced in Chapter 1 section 2). Kratzer’s definition leaves open the possibility of the antecedent being true, like in an indicative.

For the premise semanticist's relation to correspond with the multiple-world theorist's it would need to possess the following three properties: (i) reflexivity; (ii) transitivity; and (iii) that if a possible world w is closer than another v to the ideal which the ordering source determines, then v isn't equal to w .

And the premise semanticist's relation possesses (i), (ii) and (iii). We get (i) and (ii) from the definition of a totally realistic ordering source. As we've seen, according to the definition, a set of propositions A can induce an ordering \leq_A on a set of possible worlds W such that a possible world w is at least as close as a possible world z to the ideal A determines if and only if all propositions of A that are true in z are true in w as well. Here, the 'all' in the second half of the biconditional endows the relation with the property of reflexivity: for all modal bases f , worlds w and v , $v \leq_{f(w)} v$. The 'all' also endows the relation with the property of transitivity: for all f , w , u , v , t , if $u \leq_{f(w)} v$ and $v \leq_{f(w)} t$, then $u \leq_{f(w)} t$.

And we get (iii) from the definition of totally realistic conversational background as we apply it to the ordering source. A totally realistic ordering source is a function g such that for any w , $w \in \cap g(w) = \{w\}$. Here, a totally realistic g assigns to any world a set of propositions that characterises it uniquely and so it's not possible for two worlds to hold the same place in the ordering – or be equally similar – unless they are one and the same.

Note that (iii) is stronger than Lemma 1. Recall, according to Lemma 1, if g is a realistic conversational background, then for all worlds w and v , w is at least as close as v to an ideal which the ordering source $g(w)$ determines.

This allows for a relation of ‘less than or equal to’. (iii) in contrast doesn’t allow equality. (iii) ensures that the relation is one of ‘*strictly* less than’.

Now, using the definition of indicative conditional, we can prove that *modus ponens* (with realistic backgrounds f and g) is valid. We’ll rely on this lemma to prove that *modus tollens* (with realistic backgrounds f and g) is valid and later as well.

Theorem 1: For any world w and realistic conversational backgrounds f and g , if ‘If p , then (necessarily) q ’ is true in world w with respect to f and g and p is true in w , then q is true in w .

Proof: Suppose ‘If p , then (necessarily) q ’ is true in world w with respect to f and g and p is true in w . Then from the definition of an indicative conditional, ‘(necessarily) q ’ is true in w with respect to f^* and g where the new modal base $f^*(w)$ is the old one plus p .

Since f is realistic, w is compatible with the old modal base $f(w)$. But since p is true in w , it follows that all the propositions in $f^*(w)$, the old modal base $f(w)$ plus p , are true in w . So w is compatible with the new modal base $f^*(w)$.

So from the definition of necessity, there’s another world v compatible with the new modal base $f^*(w)$ such that v is at least as close as w to the ideal the ordering source $g(w)$ determines, and for any world z compatible with the new modal base $f^*(w)$, if z is as close as v to the ideal with respect to $g(w)$, then q is true in z .

But since g is realistic, it follows from Lemma 1 that w is as close as v to the ideal with respect to $g(w)$. So it follows that q is true in w . QED.

Using this lemma, we can prove that *modus tollens* (with realistic backgrounds f and g) is valid.

Theorem 2: For any world w and realistic conversational backgrounds f and g , if ‘If p , then (necessarily) q ’ is true in world w with respect to f and g and $\neg q$ is true in w , then $\neg p$ is true in w .

Proof: Let ‘If p , then (necessarily) q ’ and $\neg q$ be true in world w with respect to f and g . Now suppose for *reductio* that $\neg p$ is false in w . Then p is true in w . So by *modus ponens*, q is also true in w . But this contradicts the initial assumption that $\neg q$ is true in w . So $\neg p$ is true in w . QED.

2. *The good and the bad*

So far, we’ve seen premise semantics’ truth conditions for an indicative conditional, necessity and possibility. Next, we turn to the theory’s response to apparent counterexamples to *modus ponens* and *modus tollens*. When it comes to such arguments, premise semantics give us some results we want and others we don’t. Let’s consider responses to counterexamples which fall into three categories: (i) having a realistic modal base but non-realistic ordering source; (ii) having a non-realistic modal base but a realistic ordering source; and (iii) having realistic modal base and ordering source.

2.1. *Realistic f and non-realistic g*

Premise semantics can explain Miners, an apparent counterexample to *modus tollens*.

Miners (Kolodny & MacFarlane, 2010, pp. 115 and 128)

The argument takes the following context:

Ten miners are trapped either in shaft *A* or in shaft *B*, but we do not know which. Flood waters threaten to flood the shafts. We have enough sandbags to block one shaft, but not both. If we block one shaft, all the water will go into the other shaft, killing any miners inside it. If we block neither shaft, both shafts will fill halfway with water, and just one miner, the lowest in the shaft, will be killed (Kolodny & MacFarlane, 2010, p. 115).

Given this scenario, Kolodny and MacFarlane build a paradox with the following two premises and conclusion:

- (5) It's not the case we ought to block shaft *A*.
- (6) If the miners are in shaft *A*, then we ought to block shaft *A*.
- (7) Therefore, it's not the case the miners are in shaft *A*.

They see this as a paradox inasmuch as we might accept the premises while rejecting the conclusion even though the conclusion follows by *modus tollens*. We might take (5) as the 'obvious' outcome of our deliberation. We don't know in which shaft the miners are but we know that, given our limited knowledge, blocking neither shaft is the option we ought to take: we ought to save the greatest number of miners and, by blocking neither shaft, we could guarantee that all but one miner survived; by not doing so, i.e. blocking shaft *A* or blocking shaft *B*, we could risk killing all the miners. (5) follows from the proposition 'We ought to block neither shaft'. We might accept (6) on the grounds that again, we ought to save the greatest number of miners. However, we reject (7) on the grounds that we remain agnostic about the exact location of the miners (Kolodny & MacFarlane, 2010, p. 128).

Premise semantics can account for this paradox. The theory takes arguments with the following form to be invalid:

Not-Ought q .

If p , then Ought q .

Therefore, not- p .

Where we are dealing with ‘oughts’, the ordering source isn’t realistic because it’s not the case that the actual world is most-similar to itself. If it were (i.e. if the ordering source with ‘oughts’ were realistic), then it would follow by Lemma 2 that whatever we ought to do, we do. Lemma 2 shows that for any world w and realistic backgrounds f and g , if ‘necessarily p ’ is true in w with respect to f and g , then p is true in w with respect to f and g . So with a realistic ordering source, if I ought to do my homework, I do my homework – and since that would be a bad result, we say the ordering source isn’t realistic.

Now, *modus tollens* with a non-realistic ordering source is invalid. Here, with the actual world containing fewer true/more false propositions than other worlds relative to the ideal world, we have a non-realistic ordering source and thereby an invalid *modus tollens*.

Consider the following counterexample:

Counterexample 1

Let w be the actual world, t be a possible world, f be a realistic modal base and g be a non-realistic ordering source. Furthermore, let ‘The miners are in shaft A ’ and ‘It’s not the case that we block shaft A ’ be true in w , and ‘The miners are in shaft A ’ and ‘We block shaft A ’ be true in t . Finally, let the ordering

source $g(w)$ contain ‘The miners are in shaft A and we block shaft A ’, and let the modal base $f(w)$ be empty.

Then *modus tollens* is invalid. The premises can be true and the conclusion false. The second premise ‘If the miners are in shaft A , then we ought to block shaft A ’ is true in w with respect to f and g if and only if ‘We ought to block shaft A ’ is true in w with respect to the updated modal base $f^*(w)$, and the ordering source $g(w)$. And ‘We ought to block shaft A ’ is true in w with respect to the updated modal base $f^*(w)$, and the ordering source g if and only if for all possible worlds u in the set of worlds where all propositions in the updated modal base $f^*(w)$ are true, there’s another possible world v in that set such that:

(i) v is at least as close as u to an ideal which $g(w)$ determines; and

(ii) for all worlds z in the set of worlds where all the propositions in the updated modal base $f^*(w)$ are true, if z is at least as close as v to the ideal which $g(w)$ determines, then ‘We block shaft A ’ is true in z .

And that is the case. Since ‘The miners are in shaft A ’ is true in both w and t , w and t are in the set of all worlds where all the propositions in the updated modal base $f^*(w)$ are true, and for both w and t , (i) and (ii) are satisfied. Where u is w , there’s a v , namely t . (i) For t to be at least as close relative to $g(w)$ as w , at least as many propositions in $g(w)$ must be true in t as in w and that’s the case. In t , one proposition in $g(w)$ is true (namely ‘The miners are in shaft A and we block shaft A ’) while in w , no proposition in $g(w)$ is true.

And (ii), for all worlds z in the set of worlds where all the propositions in the updated modal base $f^*(w)$ are true, if z is at least as close as v to the

ideal which g determines, then ‘We block shaft A ’ is true in z . Here, the only relevant worlds z can be is w and t . The conditional in (ii) holds true for w and t : since ‘We block shaft A ’ is indeed true in w and t , the conditional has a true consequent.

Next, where u is t , there’s also a v , etc. namely, t itself. (i) is satisfied because t is at least as close as itself. Likewise, (ii) is satisfied for the same reason as above: the consequent of the conditional is true. So the second premise is true. So too is the first.

‘It’s not the case we ought to block shaft A ’ is true in w with respect to the modal base f and the ordering source g if and only if ‘We ought to block shaft A ’ is false in that context. And it is false in that context if and only if there exists no possible world u (in the set of worlds where all propositions in the modal base $f(w)$ are true) for which there is a possible world v in that set such that:

(i) v is at least as close as u to an ideal which $g(w)$ determines; and

(ii) for all worlds z in the set of worlds where all the propositions in the modal base are true, if z is at least as close as v to the ideal which $g(w)$ determines, then ‘We block shaft A ’ is true in z .

And this is the case. Where u is t , there are two possible v -worlds – t itself and w – and in neither case are (i) and (ii) both satisfied. Where the v -world is t itself, (i) is satisfied (t is as close as t) but (ii) isn’t: let the z -world be w . w is in the set of worlds where all propositions in the modal base $f(w)$ are true but it’s not the case that ‘We block shaft A ’ is true in w . Moreover, where the v -world is w , (i) isn’t satisfied: w isn’t as close as t . While one proposition in $g(w)$ is true in t , none is true in w .

Next, where u is w , again there are two possible v -worlds – w itself and t – and in neither case are (i) and (ii) both satisfied. Where the v -world is t , (i) is satisfied (t is at least as close as w ; one proposition in $g(w)$ is true in t , none is true in w) but (ii) isn't: let the z -world be w . w is in the set of worlds where all propositions in the modal base $f(w)$ are true but it's not the case that 'We block shaft A ' is true in w . Moreover, where the v -world is w itself, (i) is satisfied (w is as close as w) but (ii) isn't: again, w is in the set of worlds where all propositions in the modal base $f(w)$ are true but it's not the case that 'We block shaft A ' is true in w .

So 'We ought to block shaft A ' is false so 'It's not the case we ought to block shaft A ' is true in w with respect to the modal base f and the ordering source g , and so the first premise is true.

The theory can also explain *Miners (bis)*, an apparent counterexample to *modus ponens*. Using the same scenario as they did in the original *Miners*, Kolodny and MacFarlane build a paradox with the following four premises, and conclusion:

Miners (bis) (Kolodny & MacFarlane, 2010, pp. 115-6)

- (8) We ought to block neither shaft.
- (9) If the miners are in shaft A , we ought to block shaft A .
- (10) If the miners are in shaft B , we ought to block shaft B .
- (11) Either the miners are in shaft A or they are in shaft B .
- (12) Therefore, either we ought to block shaft A or we ought to block shaft B .

They see this as a paradox inasmuch as we might accept the premises while rejecting the conclusion. As in their first paradox, we might take (8) as the ‘obvious’ outcome of our deliberation. If we don’t know in which shaft the miners are, blocking the shaft in which they aren’t would result in the death of all miners and blocking neither shaft would guarantee that all but one miner survived, we ought to take the option that guarantees we will save the greatest number of miners. We accept (9) and (10) on the grounds that again, we ought to save the greatest number of miners. We accept (11) given the specific details of the scenario. While (12) appears to follow from (8) through (11), we might not accept it. It’s incompatible with (8) (Kolodny & MacFarlane, 2010, pp. 115-6).¹¹

Premise semantics can also account for this counterexample. We are effectively drawing the conclusion using two *modus ponens* with the following form and the theory takes arguments with such a form to be invalid:

p.

If *p*, then Ought *q*.

Therefore, Ought *q*.

As we saw above, where we are dealing with ‘oughts’, the ordering source isn’t realistic and *modus ponens* with a non-realistic ordering source is invalid.

¹¹ Note that I’m taking the modal to be a subjective ‘ought’ rather than an objective ‘ought’. Following Wedgwood (2016), I’m taking ‘ought’ to express more objective concepts when what an agent ought to do at a given time may be determined by facts that neither agent nor friends or advisers either knows or is even in a position to know. And I’m taking ‘ought’ to express more subjective concepts when what an agent ought to do is in some way more sensitive to the informational state that the agent (or his advisors or the like) find themselves in at the conversationally salient time.

Here, given our informational state, it’s true that we ‘ought’ (subjective) to block neither shaft *A* nor shaft *B*.

Here, with the actual world containing fewer true/more false propositions than other worlds relative to the ideal world, we have a non-realistic ordering source and thereby an invalid *modus ponens*.

Consider the following counterexample:

Counterexample 2

Assume the same as in *Counterexample 1*, the counterexample to *modus tollens* with a non-realistic ordering source. Then *modus ponens* is invalid. The premises can be true and the conclusion false. ‘If the miners are in shaft *A*, then we block shaft *A*’ is true in *w* with respect to *f* and *g* for the same reason in *Counterexample 1*. So the second premise is true. So too is the first given our assumption that ‘The miners are in shaft *A*’ is true in *w*. And the conclusion is false. ‘We ought to block shaft *A*’ is true in *w* with respect to the modal base *f* and the ordering source *g* if and only if ‘It’s not the case we ought to block shaft *A*’ is false in that context. And, as we saw above, that isn’t the case. Indeed, ‘It’s not the case we ought to block shaft *A*’ is true in *w* with respect to the modal base *f* and the ordering source *g* and so ‘We ought to block shaft *A*’ – the conclusion – is false in that context.

2.2. Non-realistic f and realistic g

Premise semantics can also explain *Marble*, an apparent counterexample to *modus tollens*.

Marble (Yalcin, 2012, pp. 1001-2)

This argument involves the following context:

We have an urn containing 100 marbles, a mix of blue and red, big and small.

The breakdown is the following:

Marbles	Blue	Red
Big	10	30
Small	50	10

Table 4: Breakdown of the blue/red/big/small marbles in the urn

We select at random a marble and place it under a cup.

Given this context, we can derive the following *modus tollens*

(13) If the marble is big, then it's likely red.

(14) The marble is not likely red.

(15) Therefore, the marble is not big.

In this argument, we're inclined to believe the premises but disinclined to believe the conclusion – even if it appears to follow by *modus tollens*.

Premise semantics' diagnosis would be the following: the argument is invalid; it's a *modus tollens* with a realistic ordering source but a non-realistic modal base – and such arguments are invalid.

Consider the following counterexample:

Counterexample 3

Let w be the actual world and f be a non-realistic modal base. Furthermore, let 'The marble is big' and 'The marble is not red' be true in w . Finally, let 'It's not the case the marble is big' be a proposition in the modal base. That means the set of worlds where all the propositions in the updated modal base are true

is empty. The updated modal base contains the contradictory propositions ‘It’s not the case the marble is big’ and ‘The marble is big’.

Then *modus tollens* is invalid. The premises are true and the conclusion false. The first premise is true in w with respect to the updated modal base, and the ordering source because for all possible worlds u (in the set of worlds where all the propositions in the updated modal base are true) in the definition of necessity, the definition is vacuously satisfied.

The second premise is also true. ‘The marble is not likely red’ is true in w with respect to the modal base f and the ordering source g if and only if ‘The marble is likely red’ is false in that context. And it’s false in that context if and only if there exists no possible world u (in the set of worlds where all propositions in the modal base $f(w)$ are true) for which there is a possible world v in that set such that:

(i) v is at least as close as u to an ideal which $g(w)$ determines; and

(ii) for all worlds z in the set of worlds where all the propositions in the modal base are true, if z is at least as close as v to the ideal which $g(w)$ determines, then ‘The marble is red’ is true in z .

And there is no such possible world u . As there’s no possible world compatible with the modal base simpliciter, it’s vacuously true that there’s no possible world compatible with the modal base for which there’s a possible world v in that set such that (i) and (ii). We get this for free.

So ‘The marble is likely red’ is false so ‘The marble is not likely red’ is true in w with respect to the modal base f and the ordering source g , and so

the second premise is true. And the conclusion is false given our assumption that ‘The marble is big’.¹²

Of course, for this counterexample to apply to *Marble*, the premise semanticist must argue that rather than signaling probability, ‘likely’ signals a non-realistic modal base. And this she can do. In *Marble*, the modal base is non-realistic because the line of inference from ‘Likely p ’ to ‘Therefore, p ’ is invalid. For example, just because it was likely that Britain would vote to remain in the EU, it didn’t follow that Britain would vote to remain in the EU. And for the line of inference to be invalid, the modal base or ordering source must be non-realistic.

Here, the former is. Supposing ‘Likely p ’ is a member of the modal base, we can imagine another proposition u – which follows from ‘(Likely) p ’ but which is false – will also be a member. In our Brexit example, let p be ‘Britain would vote to leave the EU’, we can imagine u being ‘The pound would not weaken on 24 June 2016’. u follows from ‘(Likely) p ’ but is false.¹³

¹² While one might think I’m treating ‘likely’ as a necessity modal identical to ‘must’, I’m not. I’m taking ‘likely’ to be dissimilar to ‘must’ inasmuch as ‘likely’ doesn’t have a realistic modal base (as we’ll see below) while ‘must’ does. (If ‘likely’ *did* have a realistic modal base (and a realistic ordering source), ‘likely p ’ would entail p – a result we’d reject. One might imagine a weak ‘must’ which is closer to ‘likely’ – e.g. ‘You *must* be joking!’ – but that is different from the strong (epistemic) ‘must’ with which I’m contrasting ‘likely’ above.

¹³ Here, granted but still consistent with the statement in the previous footnote, I’m taking ‘likely’ to be similar to ‘must’ inasmuch as it’s a quantifier over possible-worlds compatible with the modal base. Again, one might imagine a weak ‘likely’ which is closer to ‘most’. Then, perhaps, one might argue that ‘likely’ takes a realistic modal base while still explaining why the inference from ‘likely p ’ to p is invalid. If I’m not arguing this, it’s because the argument would require a different theory, one facing the daunting task of explaining the semantics of ‘most’ – and not one falling within the remit of this thesis.

Turning back to *Marble*, we see that ‘likely’ signals a non-realistic modal base and, as the counterexample demonstrates, *modus tollens* with such a modal base is invalid. Just as it did with *Miners*, premise semantics produce a desired result with *Marble*.¹⁴

Premise semantics give us, however, the undesired result that a *modus ponens* version of *Marble* is invalid – when we might not take it to be so.

Consider the following argument, assuming the same context as that of the original *Marble*:

Marble (bis)

- (13) If the marble is big, then it’s likely red.
- (16) The marble is big.
- (17) Therefore, the marble is likely red.

On premise semantics, we can construct a counterexample to such a *modus ponens*, i.e. one where the ordering source is realistic but the modal base isn’t.

Counterexample 4

Assume the same as above. The first premise is true for the same reason as above. The second is so too given our assumption that ‘The marble is big’ is

¹⁴ One might wonder whether one can analyse ‘likely’ in terms of quantification over a set of worlds which an ordering source orders, or whether one must appeal to probability functions (in the mathematical sense) to model its semantics. If I don’t address this matter here, it’s because I don’t want to get into a level of detail such as that. One reason being the particular semantics of ‘likely’ don’t seem to be crucial to the argument. Indeed, one could change ‘likely’ to ‘normally’ and construct another counterexample. Nevertheless, I point the interested reader to Yalcin (2010). It gives some arguments against the premise semanticist’s treatment of ‘likely’ and one in favour of putting to work probability functions in the semantics.

true in w . And the conclusion is false. As we saw above, it's not the case 'The marble is likely red' is true in w .

2.3. Realistic f and non-realistic g (bis)

Moreover, premise semantics can't explain *Lottery*, another apparent counterexample to *modus tollens*.

Lottery (Cantwell, 2008, p. 331)

(18) If you don't buy a lottery ticket, you can't win.

(19) You can win.

(20) Therefore, you do buy a lottery ticket.

Here, we have an argument that appears to have realistic both modal base and ordering source – and, according to premise semantics, such an argument is valid. The proof relies on the validity of *modus ponens* with a realistic modal base and a totally realistic ordering source.

Proof: Let p be 'You buy a lottery ticket' and q be 'You win'. For any world w and a realistic modal base f and a totally realistic ordering source g , if 'If not- p , then necessarily not- q ' and 'possibly q ' are true in world w with respect to f and g , then p is true in w .

Suppose for *reductio* that 'If not- p , then necessarily not- q ', 'not- p ' and 'possibly q ' are true in world w with respect to f and g . Then from the definition of an indicative conditional, 'necessarily not- q ' is true in w with respect to f^* and g where the new modal base $f^*(w)$ is the old one plus not- p .

Since f is realistic, w is compatible with the old modal base $f(w)$. But since $\text{not-}p$ is true in w , it follows that all the propositions in $f^*(w)$ (the old modal base $f(w)$ plus $\text{not-}p$) are true in w . So w is compatible with the new modal base $f^*(w)$.

So from the definition of necessity, there's another world t compatible with the new modal base $f^*(w)$ such that t is at least as close as w to the ideal the ordering source $g(w)$ determines, and for any world y compatible with the new modal base $f^*(w)$, if y is as close as t to the ideal with respect to $g(w)$, then $\text{not-}q$ is true in y . And since g is realistic, it follows from Lemma 1 that w is as close as t to the ideal with respect to $g(w)$. So it follows that $\text{not-}q$ is true in w .

Moreover, it follows that 'necessarily $\text{not-}q$ ' is true in w with respect to f and g , since according to the definition of necessity a proposition p is a necessity in w with respect to f and g if, for all worlds u in the set of worlds where all propositions in the modal base are true, there's a world v in that set of all worlds such that:

(i) v is at least as close as u to an ideal which the ordering source $g(w)$ determines

and

(ii) for all worlds z in the set of worlds where all propositions in the modal base $f(w)$ are true: if z is at least as close as v to the ideal which the ordering source $g(w)$ determines, then p is true in z .

Here, for all worlds u , there's a v , namely w . w satisfies both (i) and (ii): w is at least as close as u to an ideal which the ordering source $g(w)$ determines (so we saw following from Lemma 1); and p is true in w . The

relevant conditional reads ‘if z is at least as close as w to the ideal which the ordering source determines, then not- q is true in w ’ and not- q is indeed true in z – and, since we’re assuming a totally realistic ordering source, the only world z could be is w .

But then ‘necessarily not- q ’ and ‘possibly q ’ would both be true in w . Or in other words, ‘not possibly q ’ and ‘possibly q ’ would both be true in w . So not- p must be false in w . In other words, ‘not not- p ’ must be true there. In yet other words, p must be true in w . QED.

It’s important to specify that here we’re assuming a totally realistic ordering source g (or ‘centering’ as Lewis would call it (Lewis, 1998, p. 82)). According to this specification, there can be no world u as close to the world w as w itself (where u and w aren’t identical). This assumption is necessary for us to find w is the only possible world z could be. If we didn’t assume a totally realistic ordering source g , then we could derive a proof that *Lottery* is *invalid*. We could show that it’s possible for the premises to be true and the conclusion false, as follows.

Suppose p and q are false in the actual world w ; p and q are true in world u ; and the modal base is realistic: $f(w)$ is equal to $\{w, u\}$. In this scenario, ‘If not- p , then (necessarily) not- q ’ is true in w with respect to f . The antecedent restricts the modal base to contain only not- p -worlds, in our scenario w is the only not- p world ($f^*(w) = \{w\}$), and not- q is true in w . So the first premise in *Lottery* would be true. So too would the second because there exists a world in the unrestricted modal base in which q is true, namely u . But the conclusion would be false because p is false in w .

We see where the ordering source isn't totally realistic, *Lottery* is invalid. This is consistent with the results so far. For realistic *f* and partly realistic *g*, *modus tollens* falls – as does *modus ponens*. The symmetry when it comes to the validity/invalidity of the two argument forms persists.

Still, maintaining the assumption, we see that premise semantics don't offer us a result we want: namely that *Lottery* is invalid. Moreover, they offer us a result we don't want: namely that *Lottery* is valid. This is just the beginning. Premise semantics can't explain *Break-in*, yet another counterexample to *modus tollens* – and in the following section, we'll see that the theory offers us more results we don't want.

Break-in (Yalcin, 2012, p. 1003)

(21) If there is a break-in, the alarm always sounds.

(22) It is not the case that the alarm always sounds.

(23) Therefore, there is no break-in.

Here, just as 'must' (can't win=necessarily you do not win=must not win) did, 'always' takes realistic conversational backgrounds and so premise semantics find the argument valid.

Of course, in an attempt to save her theory, the premise semanticist could argue that must/always need not take such backgrounds. In doing so, however, she'd be giving up the non-straightforward version of not only *modus tollens* but also *modus ponens* (and this would mean giving up the *modus ponens* version of *Lottery*, which we think is valid. It reads 'If you don't buy a lottery ticket, you can't win' and 'You don't buy a lottery ticket',

‘Therefore, you can’t win’). Indeed, as we’ll see in the following section, for her, *modus ponens* and *modus tollens* stand or fall together.

3. Premise semantics on the Election and Barbershop paradoxes

So far, we’ve seen premise semantics offers proofs and counterexamples which explain some of our intuitions but contradict others. We noted that for the premise semanticist, *modus ponens* and *modus tollens* stand or fall together. Where we can derive proofs that one argument form (with realistic or non-realistic modal base or ordering source) is valid, we can derive proofs that the other (with the equivalent modal base or ordering source) is also valid. And where we can build counterexamples for the one (with realistic or non-realistic modal base or ordering source), we can build counterexamples for the other (with the equivalent base or source). Whether we are looking at *modus ponens* or *modus tollens*, the first premise will be true for the same reason. The second premise will be true given our assumptions. And the conclusion will be true or false given our assumptions too.

Next, we turn to examine how the premise semanticist would respond specifically to the paradoxes we saw in Chapters 2 and 3, i.e. the Election and Barbershop paradoxes.

3.1. The Election Paradox

Recall the three individually plausible but jointly inconsistent theses:

- #1 the argument’s premises are true and conclusion, false;
- #2 the argument is an instance of *modus ponens*; and
- #3 *modus ponens* is valid.

3.1.1. On #2

The premise semanticist could reject #2, arguing that the argument isn't an instance of *modus ponens*. According to the premise semanticist, an instance of *modus ponens* is an argument with the form 'If p , then (necessarily) q ' and p therefore q where we replace p and q with propositions – and the Election paradox doesn't meet this definition. Indeed, the covert necessity modal doesn't appear in the right place. While the definition places the '(necessarily)' before the consequent of the main connective conditional, the Election paradox places it before the consequent of the embedded conditional. Indeed, the first premise in the Election paradox has the form 'If p , then if q , then (necessarily) r '.

Of course, strictly speaking, the premise semanticist's *modus ponens* is not the same as the classical logician's. The premise semanticist's *modus ponens* looks like the following:

(i)

If p , then (necessarily) q .

p .

Therefore, q .

While the classical logician's looks like the following:

(ii)

If p , then q .

p .

Therefore, q .

In the argument meeting the premise semanticist's definition, the proposition q falls within the scope of a covert modal in the first premise but not in the conclusion. In the argument meeting the classical logician's, that's not the case.

That said, the premise semanticist's Election paradox is consistent with the classical logician's *modus ponens*. Aside from seeing the form of the second premise and the antecedent of the first as being identical, the premise semanticist sees the consequent of the first premise and the conclusion as so too: in each conditional, the proposition q doesn't fall within the scope of a covert modal while the proposition r does. And we can reduce an argument with the form 'If p , then if q , then (necessarily) r ' and p , therefore 'if q , then (necessarily) r ' to a simple classical logic *modus ponens* ((ii) above).

So, while rejecting #2, the premise semanticist would acknowledge that the argument has the form of the semantic counterpart to *modus ponens*. The second premise and antecedent of the first appear to have the same meaning; likewise the conclusion and the consequent of the first premise.

3.1.2. On #3

The premise semanticist would accept #3. Assuming her definition of *modus ponens*, she'd argue the argument form is valid (Khoo, 2013, p. 161). Her definition extends only to arguments with the form I mention above as (i) and will call simple *modus ponens*': 'If p , then (necessarily) q ' and p therefore q – where we replace p and q with propositions. It doesn't extend to arguments with the different form I'll call 'complex "*modus ponens*": 'If p , then if q , then (necessarily) r ' and p therefore 'If q , then (necessarily) r '.

Indeed, while we can derive a proof that premise semantics validates simple *modus ponens*, we can't derive one that validates complex *modus ponens* too. Rather, we can build a counterexample showing that it's possible for 'If p , then if q , then (necessarily) r ' to be true in world w with respect to f and g , p to be true in w and 'If q , then (necessarily) r ' to be false in w .

Let p be true in worlds w and v , q be true in v and u , r be true in v . Let the modal base f be empty. And let the ordering source g contain not- q and not- r . As q isn't true in w , then the modal base for 'If q , then (necessarily) r ' won't be realistic and, despite the premises being true, the conclusion will be false (this draws on the counterexample in Khoo (2013, p. 161, footnote 11)).

This result is consistent with the proofs we've derived above. They prove the validity/invalidity of simple *modus ponens*.

That said, the counterexample suggests that the premise semanticist is saying two things which are in contradiction: on the one hand, premise semantics (at least with a realistic f and totally realistic g) is equivalent to the multiple-world theory; and on the other hand, Import-Export is valid and complex *modus ponens* (at least with realistic f and g) is invalid – when, contrariwise, the multiple-world theorist finds Import-Export invalid and complex *modus ponens* valid.

Now, the counterexample offers a second reason for thinking that premise semantics isn't a generalisation of the possible-worlds theory. The first is that, as we've seen, the possible-worlds theory invalidates Import-Export and, as we'll see, premise semantics validates it.

Theorem 3: For any world w and realistic conversational backgrounds f and g , if 'If p , then if q , then (necessarily) r ' is true in world w with respect to

f and g if and only if ‘If p and q , then (necessarily) r ’ is true in world w with respect to f and g .

Proof: Suppose ‘If p , then if q , then (necessarily) r ’ is true in world w with respect to f and g . By the definition of the indicative conditional, it follows that ‘If q , then (necessarily) r ’ is true in world w with respect to f^* and g , where $f^*(w)$ is equal to the old modal base $f(w) \cup \{p\}$. And by once more the definition indicative conditional, it follows that ‘(necessarily) r ’ is true in world w with respect to f^{**} and g , where $f^{**}(w)$ is equal to the old modal base $f(w) \cup \{p\} \cup \{q\}$.

Next, suppose ‘If p and q , then (necessarily) r ’. By the definition of the indicative conditional, it follows that ‘(necessarily) r ’ is true in world w with respect to f_+ where $f_+(w)$ is equal to the old modal base $f_+(w) \cup \{p \text{ and } q\}$.

Here, $f^{**}(w)$ and $f_+(w)$ are equivalent.

Proof:

$$1. \cap f^{**}(w) = \cap (f(w) \cup \{[[A]]^{f,g}\} \cup \{[[B]]^{f,g}\})$$

by definition

$$2. \cap f(w) \cap \{[[A]]^{f,g}\} \cap \{[[B]]^{f,g}\} \quad \text{from 1 by distributivity}$$

$$3. \cap f(w) \cap [[A \wedge B]]^{f,g}$$

from 2 by the fact that the set of worlds at which the conjunction of two propositions is true is the intersection of the worlds at which the conjuncts are true

$$4. \cap (f(w) \cup \{[[A \wedge B]]^{f,g}\}) \quad \text{from 3 by distributivity}$$

$$5. \cap f_+(w) \quad \text{from 4 by stipulation}$$

Here, we start with the set of all things compatible with $f^{**}(w)$. And by stipulation, distributivity and the identity between intersection and conjunction, we can derive the set of all things compatible with $f^+(w)$.

Now, there appear to be two possible premise semanticists: let's call them the single- and multiple-box theorists.

Single-box theorist

The single-box theorist, on the other hand, formulates the first premise in the following way:

‘If A , then if B , then (necessarily) C ’

According to her, Import-Export is valid and she would cite the proof we saw above.

Multiple-box theorist

The multiple-box theorist formulates the first premise in the Election paradox in the following way:

‘If A , then (necessarily) if B , then (necessarily) C ’

This theorist corresponds to the multiple-world theorist. Like the multiple-world theorist, the multiple-box theorist would find Import-Export invalid and cite the proof we saw in Chapter 2 (subsection 3.2.).

The premise semanticist would do well in choosing to be the single-box theorist. This theorist can solve the Election Paradox in a way better than the multiple-box/multiple-world theorist.

The premise semanticist might try to have her cake and eat it too: she could say that sometimes a conditional contains a single box and other times it contains multiple boxes. This would allow her to say that *modus ponens* or *modus tollens* are valid when the semantic structure of the premises and conclusion have the relevant form, and invalid when they don't.

Of course, if she did adopt this position, it would be incumbent on the premise semanticist to do some more explaining about invisible boxes and why, at any given time, conditionals contain the number of boxes she says they contain. (Kratzer (2012) and Khoo (2013) suggest there should always be a single box but don't explain why there couldn't be two beyond saying that we most naturally analyse two if-clauses which appear in a row as successively restricting one and the same operator (Kratzer, 2012, p. 105)). Or that, since set intersection models domain restriction, conditionals with two if-clauses (i.e. 'If q , then if r , then s '-type conditionals) and conditionals with one if-clause and a conjunction as antecedent (i.e. 'If q and r , then s '-type conditionals) 'amount to the same thing' (Khoo, 2013, p. 157). Indeed, each conditional contains one modal: while in the case of the two-if clauses 'if q ' and then 'if r ' successively restrict a single modal, in the case of the one if-clause 'if q and r ' restrict a single modal as well.

3.1.3. On #1

The single-box theorist would accept #1. She would find the premises true and the conclusion false. The scenario is exactly that in the counterexample above. q ‘it’s not Reagan who wins’ is false in the actual world and so while the premise semanticist would find ‘If a Republican wins, then if it’s not Reagan who wins, then Anderson will’ and ‘a Republican wins’ come out as true, the conclusion comes out as false.

The multiple-box premise semanticist, on the other hand, would say what the multiple-world theorist says.

3.2. *The Barbershop Paradox*

Recall the three individually plausible but jointly inconsistent theses:

#1 the argument is invalid;

#2 the argument is equivalent to an instance of *modus tollens* (i.e. the second premise and the consequent of the first are contradictories; and the conclusion and the antecedent of the first premise are so too); and

#3 *modus tollens* is valid.

3.2.1. *On #1*

The single-box theorist would accept #1. It’s possible for the premises to be true and the conclusion false. Adopting Khoo’s counterexample to complex ‘*modus ponens*’ (2013, p. 161), we can derive a counterexample showing that, for any world w and realistic conversational backgrounds f and g , it’s possible for ‘If p , then if q , then (necessarily) r ’ and ‘If q , then (necessarily) not- r ’ to be true in world w with respect to f and g , and not- p to be false in w . Let p be true in w and v , q be true in v and u , and r be true in v . Then the first premise is

true: p , q and r are true in v . So is the second. As r isn't true in u , so too is the second premise. And as p is true in w , the conclusion is false.

3.2.2. On #2

With no extra constraint, the premise semanticist would reject that the second premise and the consequent of the first are contradictories. According to her, it's possible for neither 'If A , then B ' nor 'If A , then not B ' to be true. Consider the following scenario: let B be true in world u but false in world v and u and v be equidistant to the actual world w . Furthermore, let $g(w) = \text{'Not } A\text{'}$. This does the trick: the intersection of each u and v and $g(w)$ is empty. (And note that g is realistic because 'Not A ' is true in w .) Here, we have two sentences which are contrary (i.e. can't both be true) so long as A is compatible with the modal base but, since they're not subcontraries (here, it's possible that they both be false), the two aren't contradictories.

That said, adding the constraints which characterise the single-world theory, the premise semanticist would accept that the second premise and the consequent of the first are contradictories. The constraints are the following two: (i) no world w can be more similar to itself than w itself; and (ii) no two worlds can be equally similar to a third.

Writing the constraints in the language of the premise semanticist, for the second premise and consequent of the first to be contradictories, we need a function f such that, for any possible worlds w and v , and proposition p , if $v \in f^*(w)$ (where $f^*(w) = \cap f(w) \cup [[p]]$), then $f^*(w) = \{v\}$. This function guarantees, as (i) and (ii) do, that if a possible world w is closer than another v to the ideal which the ordering source determines, then v isn't equal to w .

With (i) and (ii), premise semantics telescope into the single-world theory and according to which, as we saw in Chapter 3, the second premise and the consequent of the first are contradictories.

3.2.3. On #3

The premise semanticist would accept #3 assuming her definition. It extends only to arguments with the form of a simple *modus tollens* viz. ‘If p , then (necessarily) q ’ and ‘not q ’ therefore ‘not p ’ where we replace p and q with propositions. It doesn’t extend to arguments with the form of a complex ‘*modus tollens*’, viz. ‘If p , then if q , then (necessarily) r ’ and ‘Not-(If q , then (necessarily) r)’ therefore not- p .

The same counterexample as we saw in the case of complex ‘*modus ponens*’ works here to show that complex ‘*modus tollens*’ isn’t valid. It shows that it’s possible that, for any world w and realistic conversational backgrounds f and g , it’s possible for ‘If p , then if q , then (necessarily) r ’ to be true in world w with respect to f and g , ‘not (if q , then (necessarily) r)’ to be true in w and Not- p to be false in w . If q isn’t true in w , then the modal base for ‘If q , then (necessarily) r ’ won’t be realistic and, despite the premises being true, the conclusion will be false (again, see Khoo (2013 p. 161, footnote 11) for a complete counterexample).

4. Conclusion

This chapter was about premise semantics for indicative conditionals and apparent counterexamples to *modus ponens* and *modus tollens*. I offered an exegesis of premise semantics, defining truth conditions for the indicative

conditional, necessity and possibility. Moreover, I argued that, when it comes to counterexamples to *modus ponens* and *modus tollens* with (overt) modal verbs and adverbs, the theory gives us some results which are consistent with our intuitions, others which aren't. Finally, I imagined the response she might give to the Election and Barbershop paradoxes.

CONCLUSION

In this thesis, I considered theories of the indicative conditional and apparent counterexamples to classically valid argument forms. Specifically, I applied the following four theories which I called:

- material
- possible-worlds
- suppositional; and
- hybrid

to try and solve the following two counterexamples:

- McGee's to *modus ponens*, which I called the Election paradox;
and
- Carroll's to *modus tollens*, which I called Barbershop paradox.

I argued that none of the theories I considered could explain – without facing any problems – the individually plausible but jointly inconsistent thesis that gave rise to the apparent counterexamples. In the case of Election, the theses were the following:

- #1 the argument is invalid;
- #2 the argument is an instance of *modus ponens*; and
- #3 *modus ponens* is valid.

Similarly, in the case of Barbershop, the theses were the following:

- #1 the argument is invalid;
- #2 the argument is an instance of *modus tollens*; and
- #3 *modus tollens* is valid.

The material theory couldn't explain the plausibility of Thesis #1 in Election without undermining itself. By introducing the concept of assertability, the theory found unassertable a sentence we might assert.

In Election, the possible-worlds theory couldn't explain the plausibility of #1 at all. It found the first premise – 'If a Republican wins the election, then if it's not Reagan who wins it will be Anderson' – false yet we might assert it.

As to the suppositional theory, it could only ever explain the plausibility at most of two theses. In Barbershop, for example, the suppositional theorist had to choose between explaining Thesis #2 and explaining Thesis #3. Whether she chose to define *modus tollens* as allowing embedded conditionals ('complex') or not ('simple'), she couldn't explain both that the argument was an instance of *modus tollens* and that *modus tollens* was valid.

The hybrid theory failed in its *raison d'être* of making *modus ponens* assertibility-preserving. And premise semantics couldn't explain why we might take the following argument – a *modus ponens* version of *Marble* – to be valid:

If the marble is big, then it's likely red.

The marble is big.

Therefore, the marble is likely red.

Indeed, according to premise semantics, we can construct a counterexample where the ordering source is realistic but the modal base isn't. And it couldn't explain either why we might take the following argument – *Lottery* – to be invalid:

If you don't buy a lottery ticket, you can't win.

You can win.

Therefore, you do buy a lottery ticket.

Indeed, according to premise semantics, we can derive a proof that the argument (with realistic both modal base and ordering source) *is* valid.

While the literature already acknowledged the surprising implication that bare conditionals contain a covert necessity modal operator (Kratzer, 2012, pp. 97-8), it had yet to acknowledge the implication that the theory invalidates arguments we might take to be valid and validates arguments we might take to be invalid.

Looking forward, we might turn toward the theories I don't consider here on the grounds that they lie beyond the scope of the thesis: dynamic or expressivist approaches to conditionals (Cantwell, 2008, Yalcin, 2012 and Bledin, 2015).

REFERENCES

- Adams, E. (1970). Subjunctive and indicative conditionals. *Foundations of Language*, 6(1), 89-94.
- Adams, E. (1975). *The logic of conditionals*. Dordrecht: Reidel.
- Bennett, J. (2003) *A philosophical guide to conditionals*. Oxford: Oxford University Press.
- Blackburn, S. (1986). How can we tell whether a commitment has a truth condition?, in C. Travis (Ed.), *Meaning and interpretation* (pp. 201-32). Oxford: Basil Blackwell.
- Bledin, J. (2015). Modus ponens defended. *The Journal of Philosophy*, 112(2), 57-83.
- Burks, A. & Copi, I. (1950). Lewis Carroll's barber shop paradox. *Mind*, 59(234), 219-22.
- Cantwell, J. (2008). Changing the modal context. *Theoria*, 74(4), 331-51.
- Carroll, L. (1894). A logical paradox. *Mind*, 3(11), 436-38.
- Edgington, D. (1995). On conditionals. *Mind*, 104(414), 235-329.
- Edgington, D. (2014). Indicative conditionals. *Stanford Encyclopedia of Philosophy*. Retrieved from <http://plato.stanford.edu/entries/conditionals/>
- Florensky, P. (1997) *The pillar and ground of truth*. (B. Joakim, Trans.). Princeton: Princeton University Press. (Original work published in 1914)
- Forrester, J. (1984). Gentle murder, or the adverbial samaritan. *The Journal of Philosophy*, 81(4), 193-7.

- Fulda, J. S. (2010). Vann McGee's counterexample to modus ponens: An enthymeme.' *Journal of Pragmatics*, 42(1), 271-3.
- Grice, H. P. (1961). The causal theory of perception. *Proceedings of the Aristotelian Society Supplementary Volume 35*, 121-52.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Speech acts (Vol. 3)* (pp. 41-58). New York: Academic Press.
- Grice, H. P. (1989). *Studies in the Way of Words*. Massachusetts: Harvard University Press.
- Hájek, A. (2012) The fall of 'Adams' Thesis'? *The Journal of Logic, Language and Information*, 21, 145-61.
- Jackson, F. (1979). On assertion and indicative conditionals. *The Philosophical Review*, 88(4), 565-89.
- Jackson, F. (1987). *Conditionals*. Oxford: Basil Blackwell.
- Jackson, F. (2006). Conditionals. In M. Devitt & R. Hanley (Eds.), *The Blackwell Guide to the Philosophy of Language* (pp. 212-24). Oxford: Wiley-Blackwell.
- Johnson, W. E. (1894). A logical paradox. *Mind*, 3(12), 583.
- Katz, B. D. (1999). On a supposed counterexample to modus ponens. *The Journal of Philosophy*, 96(8), 404-15.
- Khoo, J. (2013). A note on Gibbard's proof.' *Philosophical Studies*, 166, supp. 1, 153-64.
- Khoo, J. (2015). On indicative and subjunctive conditionals. *Philosophers' Imprint*, 15(32), 1-40.
- Kolodny, N. & MacFarlane, J. (2010). 'Ifs and oughts.' *The Journal of Philosophy*, 107(3), 115-43.

- Kratzer, A. (2012) *Modals and Conditionals*. Oxford: Oxford University Press.
- Lewis, D. (1973). *Counterfactuals*. Oxford: Basil Blackwell.
- Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *The Philosophical Review*, 85(3), 297-315.
- Lewis, D. (1998). *Papers in philosophical logic*. Cambridge: Cambridge University Press.
- Lowe, E. J. (1987). Not a counterexample to modus ponens. *Analysis*, 47(1), 44-7.
- McGee, V. (1985). A counterexample to modus ponens. *The Journal of Philosophy*, 82(9), 462-71.
- McGee, V. (1989). Conditional probabilities and compounds of conditionals. *Philosophical Review*, 98(4), 485-541.
- Piller, C. (1996). Vann McGee's counterexample to modus ponens. *Philosophical Studies*, 82(1), 27-54.
- Potts, C. (2004) *The logic of conventional implicatures*. Oxford: Oxford University Press.
- Quine, W. V. O. (1952). *Methods of Logic*. London: Routledge and Kegan Paul.
- Rhodes, M. (2012). Note on Florensky's solution to Carroll's "barbershop" paradox: Reverse implication for Russell?. *Philosophia*, 40(3), 607-16.
- Russell, B. (1905). The existential import of propositions. *Mind*, 14(55), 398-402.
- Sidgwick, A. (1894). A logical paradox. *Mind*, 3(12), 582.

- Sinnott-Armstrong, W., Moor, J. & Fogelin, R. (1986). A defense of modus ponens. *The Journal of Philosophy*, 83(5), 296-300.
- Stalnaker, R. (1981). A theory of conditionals. In W. L. Harper, R. Stalnaker & G. Pearce (Eds.), *Ifs: Conditionals, belief, decision, change, and time* (pp. 41-55). Dordrecht: Reidel.
- Stalnaker, R. (1999). *Context and content: Essays on intentionality in speech and thought*. Oxford: Oxford University Press.
- Veltman, F. (1985). *Logics for conditionals* (Doctoral dissertation). University of Amsterdam, Amsterdam.
- Von Fintel, K., & Heim, I. (2011). Intensional semantics [PDF document]. 24.973 *Advanced Semantics*, Massachusetts Institute of Technology, Spring. Retrieved from <http://web.mit.edu/fintel/fintel-heim-intensional.pdf>
- Wedgwood, R. (2016) Objective and subjective ‘ought’. In N. Charlow and M. Chrisman (Eds.), *Deontic modality*. Oxford: Oxford University Press.
- Yalcin, S. (2010) Probability operators. *Philosophy Compass*, 5(11), 916-37.
- Yalcin, S. (2012). A counterexample to modus tollens. *The Journal of Philosophical Logic*, 41(6), 1001-24.