6-2016

The C3 Conditional: A Variably Strict Ordinary-Language Conditional

Monique L. Whitaker
Graduate Center, City University of New York

THE *C3* CONDITIONAL:

A VARIABLY STRICT ORDINARY-LANGUAGE CONDITIONAL

by

MONIQUE WHITAKER

A dissertation submitted to the Graduate Faculty in Philosophy

in partial fulfillment of the requirements for the degree of Doctor of Philosophy,

The City University of New York

2016

The *C3* Conditional: A Variably Strict Ordinary-Language Conditional

by

Monique Whitaker

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

_____ _____

Date Gary Ostertag: Chair of Examining Committee

_____ _____

Date Executive Officer

Supervisory Committee:

Graham Priest (Advisor)

Gary Ostertag (Committee Chair)

Melvin Fitting

Arnold Koslow

Stephen Neale

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

The *C3* Conditional: A Variably Strict Ordinary-Language Conditional

by

Monique Whitaker

Advisor: Graham Priest

In this dissertation I provide a novel logic of the ordinary-language conditional. First, however, I endeavor to make clearer and more precise just what the objects of the study of the conditional are, as a lack of clarity as to what counts as an instance of a given category of conditional has resulted in deep and significant confusions in subsequent analysis. I motivate for a factual/counterfactual distinction, though not at the level of particular instances of the conditional. Instead, I argue that each individual instance of the conditional may be interpreted either factually or counterfactually, rather than these instances dividing into distinct types. I examine the classic Oswald–Kennedy pair of sentences, typically taken to be the quintessential example of how conditionals must be split into two different categories, to show that they in fact do not demonstrate this. I then present my account of the logic underlying the ordinary-language conditional, the system *C3*, and a justification of the form it takes. This logic provides distinct interpretations of the conditional as it concerns what is factual or what is counterfactual, respectively. The factual interpretation is true when both antecedent and consequent are true at the actual world, false when the antecedent is true and consequent false, and not truth-apt otherwise. The counterfactual interpretation incorporates a *ceteris paribus* clause, ensuring that it is not falsified by extraordinary,

unforeseeable occurrences, and is true when, at all worlds at which the antecedent is *ceteris paribus* true, the consequent too is true. It is false when, at any of these worlds the consequent is false; and not truth-apt otherwise. I go on to examine alternative theories of the conditional—from the suppositionalist approach to evaluating the putative indicative conditional, to Stalnaker's combined indicative and counterfactual account, and Lewis's analysis of the counterfactual; among a number of others—and offer comparison to my own theory. Finally, I look at various challenges to my account of the conditional. Being a strict conditional (albeit variably so), that of *C3* is open to the objection that it fails to match ordinary speakers' intuitions as regards its truth-value assignments to, for instance, conditionals with necessarily true consequents. I address these so-called paradoxes of the strict conditional, and also discussion the question of truth-preservation in the *C3* system in the case of such inferences as those relying on transitivity and that of *modus ponens*. I maintain that these inferences are indeed truth-preserving under certain, specifiable, conditions, and I close by offering possible avenues for further research.

Contents

Acknowledgements

First and foremost, my very grateful thanks to my advisor, Graham Priest. Your incisive feedback has immeasurably improved this work. Any time that I had the feeling something was off with a argument or point I was making, you've invariably honed in on it in your comments and highlighted just what was going wrong, which has helped to make my writing clearer and more rigorous, though all remaining errors are my own. I'm also immensely grateful for the time and care you take with your students. The weekly meetings you organize for us to present work in progress are valuable as much for testing arguments and getting input, as for motivating continued progress and ensuring that we are in regular contact with you and our peers. It's a significant commitment of time and effort, but one that makes a great difference to us as your students. I also really appreciate the speed with which you unfailingly respond with feedback and assistance, despite the many demands on your time. I feel very fortunate to have an advisor who's genuinely there for his students; thank you.

Thank you, also, to my dissertation committee members, Stephen Neale and Gary Ostertag, for your input and direction on this project. I greatly appreciate it. And, Gary, thanks so much for all your support through the course of the graduate program. Your mentorship and encouragement has made such a difference to my academic development. Thank you for all your philosophical and practical guidance; it's meant an enormous amount to me.

My sincere thanks to everyone at the Kripke Center at the CUNY Graduate Center—to the late Meyer Kripke for his great generosity that provided for the Center and the Kripke Fellowship that made it possible for me to study at the Graduate Center; to Saul Kripke for your work, the *raison d'être* of the Center, and for your kindness to me; and to Gary Ostertag and Romina Padro for your exceptional scholarship and dedication to the Center.

To my fellow graduate students and dear friends, Tommy Kivatinos and Thomas Ferguson—I owe you a great deal. You've both been incredible sources of philosophical and personal support. Tommy, you've kept me grounded (no pun intended) through graduate school and truly inspired me with all that you've achieved. Thomas, you're as good a friend as you are a logician—your patient assistance with many logical difficulties has helped me enormously.

My family also deserve much gratitude—Dad, Mum, Jennie, Lynne, thank you for being there for me through all my years as student, and for your financial support during many of them. Thank you, too, to my siblings Joseph, Erin, Kai—and especially to Krista; having you just a phone call away in DC made graduate school that much easier. And thank to my dear godmother, Louise Becker. Your warmth

and care when I first arrived and throughout my time in New York meant a lot me.

I want to acknowledge, too, the great public school teachers I've had who helped me to love learning and fostered my curiosity, and who, along with my parents, taught me the skills of writing and thinking well, without which I would never have been able to pursue philosophy. So, thank you, Ms. Couzens at Penzance, Mr. Duell and Mr. Minnie at Grove, Mrs. Austen and Mr. Bisset at Westerford, and to Augustine Shutte at UCT for your inspiring introduction to philosophy that got me hooked in the first place.

And last, but certainly not least, to my friends who got me through graduate school: Thank you Rosemary Twomey, Sophia Bishop, and Romina Padro for your amazing friendship; Jennifer Corns for all your care and philosophical acumen; and Rebecca Traynor for the motivation and great writing sessions at the Hunter Digital Commons. Thank you, also, 3T Studio for keeping me sane and to my incredible fellow DSC officers for broadening my horizons and fighting the good fight. Finally, thank you so much Roscoe Sprong for all the love and cooking, making sure I still ate well when in the thick of it; and to Christina Nadler for everything—I couldn't have done it without you.

Introduction

Conditionals have not wanted for philosophical attention. As it is almost obligatory to observe at the start of any discussion of them—they've been puzzling us at least since Kallimachos remarked that, "Even the crows on the rooftops are cawing about the question of which conditionals are true."[1] Little has changed in the intervening two thousand-odd years, as the subtitle of Kratzer's (1979) introduction to her 'Conditional Necessity and Possibility' confirms: "Conditionals Are Important but Troublesome".

This ubiquitous yet opaque locution has generated an overwhelming body of literature. One so large that it might suggest a dissertation on the subject was not very urgently needed. But, despite the intense and prolonged scrutiny conditionals have received, there is hardly consensus as to which of them are true—and, indeed, now there is debate as to whether they even have truth conditions.

What has motivated the production of all this philosophical (and linguistic and

[1] Attributed by Sextus Empiricus, *Against the Mathematicians*: from B. Mates, *Stoic Logic*; quoted in Hájek (nd).

psychological) discussion of the conditional—and, more pertinently, why add to it? At it's heart, the appeal of understanding the conditional is its sheer ubiquity in our reasoning: Whenever we wonder what will happen given certain circumstances, speculate about how things could have been different, or predict how people will behave, we rely on conditionals; when we mull over the effects on the earth if climate-change predictions are accurate, how a friend would react if we told them the truth about their sartorial choices, whether we would handle our winnings responsibly if we won the lottery. But the appeal lies not just in how extraordinarily common and familiar conditionals are to us and the ease with which we use them in everyday life, but this in contrast with how remarkably resistant they've proven to clear analysis.

Moreover, providing a clear and accurate account of the conditional is of more than quotidian interest, given how integral conditionals are to numerous areas of philosophy. How I first came to conditionals is through an interest in the metaphysics of possibility and necessity, when I realized that the analysis of counterfactuals was central to one's position on alethic modality. Conditionals also play a key role in accounts of reasoning, and in many analyses of causality (notably Lewis's, 1973a) and dispositions.

The project of this dissertation is to argue for a particular logic of the conditional as it is used in ordinary discourse. But, perhaps more importantly, also to make clearer and more precise just what the objects of the study of the ordinary-language conditional are. In Chapter 1 I examine the nature of conditional, with particular regard to the classification of its types. A lack of clarity as to what counts as an instance of a given category of conditional has resulted in deep and significant

confusions in the subsequent analysis of these instances. Despite the great variety of forms they may take within and across natural languages, there doesn't appear to be any especial difficulty in or controversy over identifying a given locution as simply a conditional. Where the vagueness, contradiction, and general hand-waving come in is at the point of categorizing these instances into types.

Much of the trouble, I argue, arises from a good deal of the analysis of conditionals being confined just to English, which makes it hard to determine where language-specific features end and general characteristics of the conditional begin. Once we are clear on this distinction between the particularities of grammar and the nature of the conditional itself more generally, it becomes evident that we need to discard the popular notion of indicative and subjunctive conditionals in favor of the more perspicuous categories of factual and counterfactual; to do with how things actually are, and with how they might be, respectively. But, what also emerges, is that these classifications cannot be applied to instances of the conditional themselves—rather, these are ways in which a given conditional may be interpreted, depending on the context and the interlocutors' beliefs and interests therein. I examine the classic Oswald–Kennedy pair of sentences, typically taken to be the quintessential example of how conditionals must divide into two different types at the level of their instances, to show that they in fact demonstrate no such thing. Each of the pair turns out to be susceptible of both kinds of interpretation, though their grammatical form in English (and certain other languages) and widely-held beliefs as to the facts of Kennedy's assassination conspire to make salient one or other of these in each instance.

In Chapter 2 I present my account of the logic underlying the ordinary-language

conditional, the system $C3$, and a justification of the form it takes. This logic is based in significant part on Graham Priest's conditional logics—sharing with these the inclusion of a *ceteris paribus* clause as implicit in the conditional. $C3$, however, proposes a bifurcated evaluation, with distinct logical interpretations of the conditional as it concerns what is factual or what is counterfactual, respectively.

I maintain that the truth conditions of the conditional are such that it is true on a factual interpretation when both its antecedent and consequent are true at the actual world, and false when its antecedent is true but its consequent false. In those cases in which the conditional's antecedent is false at the actual world, though, I argue that it is not truth-apt on a factual interpretation, since there is little sense to be made of what does or does not follow from the antecedent's being the case (at the actual world), when it is not the case. Interpreted counterfactually, the conditional is true when, at all of those possible worlds at which the antecedent is, *ceteris paribus*, the case, the consequent is too; and false when, at any one or more of these antecedent-worlds, the consequent is false. The counterfactual interpretation is, therefore, one of a variably strict conditional—changing according to what is precluded by the *ceteris paribus* clause. I also give a combined interpretation, one on which the conditional can be termed simply true or false, without qualification. I take a conditional to be true on this joint interpretation when either both the factual and counterfactual interpretations are true, or at least one is true and the other non-truth-apt. Similarly, a conditional is false on this interpretation when it is false on both the factual and counterfactual interpretations, or false on at least one and non-truth-apt on the other.

For the counterfactual interpretation, much hinges on this *ceteris paribus* clause. Fully and determinately specifying what it includes and/or excludes is impossible—there is a surely infinite variety of exceptional, unlikely, and unpredictable, though still possible, things that may affect the consequent independently of the antecedent—yet it's evident that we manage very well in day-to-day life to make rough but sufficiently specific determinations of what constitutes 'all things being equal'. When someone commits to paying back a friend the money owed them on Wednesday, both take this to include an implicit *ceteris paribus* clause; it's not the case that the money will be paid back no matter what, even in the face of a catastrophic hurricane hitting their city or the severe injury of the borrower in a bus crash on Tuesday. We are quite able to distinguish between the sort of thing that would count as a legitimate reason to fail to deliver on this promise and the sort that wouldn't—being injured in a bus crash versus just not really feeling inclined to return the money.

Furthermore, we are able to adjust these assessments according to context. If, say, the money were a sizeable sum that the lender urgently needed repaid, and the borrower were a stuntperson who knew they stood a very high risk of being injured in bus stunt, which they had the option to schedule for either Tuesday or Thursday, then choosing to perform the stunt on the Tuesday and needlessly risking not being in a position to pay back the money the following day, would make the borrower's being injured in a bus crash (from the stunt gone wrong) reasonably part of all things being equal. This crash was a foreseeable incident, within the realm of ordinary life for a stuntperson, and thus not something that we would take the *ceteris paribus* to exclude, *in this particular context*.

Determining what is and what is not excluded by a given *ceteris paribus* clause is a complex and highly context-dependent feat, but one that permeates all aspects of our lives. I do not pretend to be able to detail or explain the mechanism by which we manage this, but I do argue that this problem is hardly unique to the analysis of the conditional alone.

In Chapters 3 and 4 I examine the various alternative accounts of the conditional against which my own is competing. Given the general division in the literature between, roughly, indicative and subjunctive or counterfactual conditionals, I divide the theories that I consider similarly—theories of the putative indicative conditional I consider in Chapter 3, and in Chapter 4 I turn to those that make use of a possible worlds semantics and that are concerned with the subjunctive or counterfactual conditional.

I begin, in Chapter 3, by looking at suppositionalist theories of the supposed indicative; that is, accounts on which the conditional is not susceptible of truth conditions, but is assessed in terms of its assertability or degree of believability. Theorists of this kind, most notably Edgington, do acknowledge the apparent dual interpretations that the conditional may be given, but, in contrast to my own approach, the conclusion they draw from this is that conditionals have no clear or determinate truth-value. Indeed, this seems to be the primary motivation for denying the conditional truth conditions. However, I argue that, once we recognize that any given instance of the conditional may be interpreted both factually and counterfactually, it's clear that there is no motivation to deny truth conditions of the conditional. There are, of course, other good reasons to think that conditionals do have truth-

values; not least of all because we treat them as if they do all the time in ordinary discourse.

I then turn to the pragmatic account of the indicative, which seeks to defend its having the logical form of the material conditional, though with constraints on its assertability provided by pragmatic concerns. On this view, the so-called paradoxes of the material conditional can be explained by there being a conversational implicature carried by the conditional that forms no part of its logical structure but nonetheless makes it unassertable when, roughly speaking, there is no reason to accept the consequent in light of the truth of the antecedent. Thus, conditionals such as, "If I am human, then apples are a fruit", are to be assessed as true (since both antecedent and consequent are true and the conditional's truth conditions are that of the material conditional), but not assertable, as the implicature that apples being a fruit in some way follows from my being human is violated.

Unfortunately for the pragmatic approach, there are notorious examples of cases where even such implicatures are unable to account for the intuitive truth-values of particular conditionals. Jackson's alternative pragmatic account attempts to address these concerns by adding the requirement of 'robustness' for a conditional to be assertable. A conditional is robust insofar as it is true, its consequent is highly probable given its antecedent, and it remains highly probable as a whole relative to its antecedent. But, while this further requirement does accommodate certain examples that are problematic for the Gricean view, it founders in other instances.

Finally in Chapter 3, I consider a restrictor analysis of the indicative of the sort that Kratzer gives. On this approach, the 'if'-clause of the conditional is taken to act

merely as a restriction on the domain of the associated adverb of quantification. As, of course, not all conditionals include an explicit adverb of quantification, Kratzer posits that those without one have an implicit modal, such as 'must'. However, there appear to be significant problems for the restrictor analysis, in that many other restrictive clauses cannot be accurately given in the form of a conditional without implying a connection between the antecedent and consequent that is absent in the original sentence. The difficulty appears to be that the 'if'-clause of the conditional is not merely a restrictor of the kind Kratzer argues for, but that the conditional also suggests some sort of relationship between antecedent and consequent that is not accounted for by this theory.

In Chapter 4 I focus on possible-worlds theories that offer joint accounts of both putative indicative and counterfactual conditionals, and accounts just of the counterfactual. The most prominent of these are, of course, those of Stalnaker and Lewis, respectively. I use Goodman's problematization of the analysis of counterfactuals as a starting point, to make clear the inherently vague nature of counterfactual conditionals and their consequent reliance on complex and somewhat indefinite intuitive judgments in order for them to be ascribed truth-values. Stalnaker's notion of closeness of worlds in terms of similarity, Lewis's of comparative overall similarity, mine of all things being equal—these are primitive notions that are grasped intuitively. Nothing more determinate can be forthcoming with regards to the counterfactual conditional (or, as I argue, the counterfactual interpretation of the conditional), because, absent any kind of body of data as to the exact nature of other possible worlds, we have only our experience of the actual world to go on in our assessments of them.

Stalnaker's theory requires that there is a unique closest possible world to the actual world at which the antecedent of any given instance of the conditional is true—this being the actual world itself in some cases. He argues for a selection function that picks out the closest (that is, the most similar) world at which the antecedent is true, relative to a given base world, with a conditional being true when its consequent too is true there and false when it is not. Similarity, for Stalnaker, is determined by the context set—that set of possible worlds not yet ruled out by what is presupposed by the participants in a given conversation. However, in particular instances, we wish to reach outside this context set in selecting a world at which the antecedent is true, as in the case of counterfactuals. One of the key difficulties for Stalnaker's theory is that there are cases in which there simply is not a single unique closest world, but two or more with equal claim of being the most similar to the actual world. Moreover, there are cases in which there would seem to be an infinite progression of ever-closer worlds, rather than one that is uniquely closest.

The $C3$ theory is immune to these sort of worries, seeing as it assesses the truth-value of the counterfactual interpretation of the conditional in terms of all of those worlds at which the antecedent is true (albeit a set of antecedent-worlds restricted by the *ceteris paribus* clause). But, of course, so too is Lewis's analysis of the counterfactual, which relies on spheres of similarity to a base world (in our case the actual world). These are able to accommodate both equally similar worlds as well as a series of infinitely more similar worlds, all of which may be contained within the same similarity sphere. A conditional's truth-value is then determined by whether the consequent is true or false at all of those most similar worlds at which the antecedent

is true.

However, though Lewis manages to avoid certain of the pitfalls of Stalnaker's account, his theory is unable to deal with others. Regardless of how it is qualified, similarity simply isn't always what we are most concerned with in the evaluation of conditionals. I conclude this chapter by looking at the problems with this notion as the key determinant of which antecedent-worlds to use in the assessment of counterfactuals, and argue that the greater flexibility of the *ceteris paribus* clause enables it to take significantly better account of the relevant context.

In the last chapter, Chapter 5, I address a variety of potential challenges to the $C3$ theory of the conditional. I consider the issue of embedded conditionals and compound antecedents and consequents, which I argue can be evaluated just as normal, as long as one takes adequate account of the relevant embeddings of *ceteris paribus* clauses in the case of counterfactual interpretations. I look, also, at the question of seemingly unrelated antecedents and consequents, in the case of conditionals that are evaluated as true on my approach but which appear intuitively false, and show that these may be explained as instances in which the implicature carried by the conditional (that there is some sense in which the consequent depends on or follows from the antecedent) is violated, even though their truth conditions are satisfied.

In addition, I examine the behavior of inferences making use of transitivity, *modus ponens*, and contraposition, and argue that none of these are formally valid on the $C3$ account, but that there are nonetheless clear conditions determining whether a given instance of reasoning is truth-preserving or not. I consider probability condi-

tionals, and counteridenticals and counteractuals, and make clear that these types of conditionals too can be accounted for on my view. And, finally, I look at conditionals with necessarily false antecedents, and how the logic of $C3$ could be modified to give meaningful truth-value assignments to conditionals of this kind.

Chapter 1: **The Ordinary-Language Conditional**

The logic that I propose posits that no given instance of the conditional, sans context, can be assigned to one category of conditional or another—that is, it rejects the traditional classification of particular everyday uses of the conditional according to grammatical/grammatical (indicative/subjunctive) or grammatical/metaphysical (indicative/counterfactual) type. My account does, however, distinguish between factual and counterfactual *interpretations* of specific instances of the conditional.[2] While grammatical categorization divides the conditional according to the grammatical mood of each given instance; metaphysical categorization divides it according to whether it concerns what is actual, what is the case at the actual world—or what is

[2]The terminology of "factual" and "counterfactual" interpretations of the conditional does accord somewhat with that of Nelson Goodman (1947), though he has the additional category of the "semifactual" conditional, and he takes the standard approach of assigning each instance of the conditional to one of these three categories exclusively, rather than seeing them as ways in which a conditional may be interpreted, as I propose.

counterfactual, what merely might have been or may be.

Instead of identifying particular conditionals as themselves either factual or counterfactual, though, I argue that no given instance of the conditional is itself inherently one or the other. Rather, each may be *interpreted* either factually or counterfactually, depending on the specific context in which the conditional is used. Typically, which of the two interpretations (if not both) is salient is evident to the relevant interlocutors, though there may be instances in which it is neither determinate which of the two interpretations is intended, nor is it communicatively relevant that it be so determinate.

On my account, a given conditional is true on a counterfactual interpretation (that is, interpreting the conditional as making a claim as to how things might be, rather than as to (or in addition to/regardless of) how they actually are in those cases where, at all those worlds at which the antecedent is *ceteris paribus* true, the antecedent is also. A conditional is factually true in those instances where both the antecedent and consequent are true at the actual world.

This theory provides a model of the logic implicit in our ordinary language usage of the conditional; it does not attempt to provide any kind of account of the psychological or neurological reality of how the truth-value of the conditional is in fact determined in actual instances of judgment. Moreover, my account of the conditional, as typified by the "If..., then..." construction, is not intended as specific to this or any other form in English, or to that in any other natural language. Rather, I aim to give an analysis of the underlying logic of the conditional, regardless of the form of its expression in any specific language—using the "If..., then..." construction

simply as a convenient shorthand for the conditional more generally.

Evidently, my analysis of the conditional itself makes use of any number of conditional constructions; there are instances throughout of the conditional being used rather than merely mentioned. In these cases, the conditional should be taken to have its common, intuitive and uncontroversial meaning—for my project is not to determine when the bulk of our ordinary uses of the conditional are true or false, but to elucidate the logical system underlying these truth-value judgments that we already quite comfortably make. The form of the logic implicit in these truth-value assignments is very much at issue here, but not the fact that competent natural language speakers (in this case of English) make these judgements as we do. These, in fact, are the very data which I seek here to explicate. Certainly, there are particular instances of the conditional where our intuitions are unclear, confused, or conflictory—but my exposition and argument do not depend on the *use* of such instances.

I propose a novel account of the ordinary-language conditional, whereby there are no distinct types of conditional per se, but rather two different interpretations available for each instance of the conditional; namely, a factual and counterfactual interpretation. Thus, the truth-value of the conditional is determined by the truth-value of the more salient of the two interpretations in a given context (or, in certain cases, by the combined truth-values of both where these are not in opposition to one another). The truth-value of the factual interpretation is true when, at the actual world, both antecedent and consequent are true, false when, at the actual world, the antecedent is true and the consequent false, and non-truth-apt otherwise; that is, when the antecedent is false at the actual world. The truth-value of the coun-

terfactual interpretation is given by C3, a modified version of the conditional logic C^+ (Priest, 2008). I argue that this latter system provides an accurate logic for the counterfactual interpretation of the conditional, providing both the best accommodation of our intuitions and also good reason to challenge or explain away those at odds with the answers it gives. This interpretation of the conditional is essentially that of a variably strict conditional, but incorporating a *ceteris paribus* clause, as seems well motivated by our intuitive use. Whatever the background assumptions or other shared presuppositions among speakers necessary for communication, they cannot all be directly built into the semantics of the conditional, and the *ceteris paribus* clause allows us to take account of this. This dual-interpretation account of the conditional provides the flexibility to handle both those conditionals for which we have a clear idea of their antecedents' and consequents' truth-values at the actual world, and those for which we do not.

Perfectly delimiting all and only instances of the conditional may be neither clear nor straightforward, but the general concept is evident enough. The ordinary-language conditional is the assertion of the consequent, given the truth of the antecedent, in everyday natural language usage. The use of this logical form—the expression of the conditionality of consequent on antecedent—appears to be universal to natural languages. As the linguist Wierzbicka has argued extensively, "IF is a universal human concept lexicalised in all languages". (1997: 25). She further observes that "Cross-linguistic evidence suggests that "an awareness of the conditional", an ability to say 'what if ...?' or 'if ... then...', is indeed a human universal." 1997: 53). Conditionals are argued by the linguist Hockett (quoted in von Fintel, 2011: 3) to be

"a primary exhibit for one of the 'design features' of human language: displacement". Thus legitimate explication of the ordinary-language conditional's underlying logical structure must be independent of linguistic variations of its expression within and across languages. Naturally, conditionals manifest in a variety of forms in different languages—some very close to those of English and others quite different—but we nonetheless have a good intuitive grasp of what is meant by a conditional statement, as used in everyday speech.

In English, of course, the conditional is typified by the "If..., then..." locution, though not every conditional necessarily takes that form. The following, though, is an example of the canonical "If..., then..." conditional locution,

(1) If they don't stop the walls from collapsing, then the roof will fall down.

That is, it is the case that the roof will fall down, conditional on their failing to stop the walls from collapsing.

There are other ways of expressing a conditional without using the "If..., then..." form. To give some indication of these, the following[3] each express a conditional (not necessarily with the same meaning as (1) or as one another) without making use of the typical English conditional locution, though I offer a suggested approximation of each example stated in canonical "If..., then..." form, to make evident how they are, in fact, conditional:

[3]Adapted from examples in Lewis (1973b: 4), Rawlins (2013: 112), Schwenter (1999: 95–96), and von Fintel (2011: 3).

(2) Had they stopped the walls collapsing, the roof wouldn't have fallen down.

If they had stopped the walls collapsing, the roof wouldn't have fallen down.

(3) Don't stop the walls from collapsing and the roof will fall down.

If you don't stop the walls from collapsing, then the roof will fall down.

(4) Stop the walls from collapsing or the roof will fall down!

If you don't stop the walls from collapsing, then the roof will fall down.

(5) Without the walls holding it up, the roof will fall down.

If the walls aren't holing it up, then the roof will fall down.

(6) Whether or not the walls collapse, the roof will fall down.

If the walls collapse or if they do not, (then) the roof will (still) fall down.

(7) No matter what happens, the roof will collapse.

If things happen however they will happen, (then) the roof will collapse.

(8) Given that the walls are going to fall down, the roof will collapse.

If the walls are going to fall down, then the roof will collapse. (And the

walls are going to fall down.)

(9) The *other* cement would have stopped the walls from collapsing and the roof falling down.

If the other cement had been used instead of the cement that was in fact used, then the walls wouldn't have collapsed nor the roof have fallen down.

(10) The walls can't have collapsed or the roof would have fallen down.

If the walls had collapsed, then the roof would have fallen down. (And the roof hasn't fallen down, so the walls haven't collapsed.

(11) No wall collapse, no roof collapse.

If the walls don't collapse, then the roof won't collapse.

Conversely, there are sentences that, in English, make use of the typical "If..., then'...' form but are not uncontroversially accepted as genuine conditionals—such as those variously termed biscuit, speech act, or relevance conditionals (DeRose and Grandy, 1999). For instance,

(12) If you want to stop your roof falling down, I've got a great book on effective building construction.

(13) Our repairs to stop the roof falling down were a great success, if I do say so myself.

Such cases could likely be accommodated without too much difficulty on my account, but I remain agnostic on their status as genuine conditionals and certainly do not take them as exemplars of the sorts of instances to be explained.

Taxonomizing the Conditional: Grammar, Metaphysics, and Tense

Though a sufficiently general notion of the conditional is easy enough to grasp, a failure to accurately taxonomize it has led to significantly confused categories of conditional types in even the most widely adopted explications. These problems of categorization, in turn, hamper the correct analysis of its underlying logical form. Thus, before the logical structure of the conditional can be effectively determined, we must first consider whether the conditional is to be treated as unified in form, or if it should be divided into different types, each with a distinct logical form. Clearly, there are grammatical characteristics—as well as metaphysical distinctions between what is actual and what simply possible—that encourage the division of the conditional into various types. It is important, though, to clarify whether any of these putative divisions reflect genuine logical distinctions in the ordinary-language conditional rather than merely superficial dissimilarities, particularly ones tied to the grammar of a specific language.

The taxonomy of the conditional is by no means a settled question. Various theorists have argued for a number of different pairs of conditional types—such as, indicative/subjunctive (Adams, 1970), indicative/counterfactual (Lewis, 1973b), prime/non-prime (Priest, 2009)—while Goodman (1983) favors a tripartite factual/semi-factual/counterfactual categorization, and Stalnaker (1968) argues for a single logical treatment of all conditionals, with minimal distinctions, though he does recognize certain differences between indicatives and counterfactuals.

The indicative/subjunctive split

The logical form of the conditional would obviously be rendered unjustifiably *ad hoc*, should it be made beholden to the particular vagaries of English grammar. Unsurprisingly, then, this is a significant difficulty in using the subjunctive mood to delineate a type of conditional, when this grammatical category has become significantly diminished in English over the past few centuries, with much of the full subjunctive form no longer in use. (Indeed, be you[4] in any doubt, consider that, as early as 1926, the renowned grammarian H.W. Fowler had already dismissed the English subjunctive as "moribund".) The thinking behind the original adoption of the indicative/subjunctive division of instances of the conditional was presumably that, insofar as the subjunctive mood in English (and a number of other European languages) expresses a lack of absolute definiteness—a degree of doubt, hypotheticality, or possibility rather than certainty—it would do well as a means of picking out those conditionals that concern things not as they are in reality, but as they might

[4]'Be you' being the now-archaic second-person present-tense subjunctive form.

have been or may be in the future.

Languages such as French may still maintain a robust subjunctive mood, but, along with the dubious health of the subjunctive in contemporary English usage, there is the further trouble that many others lack this grammatical form entirely. As linguist Bruce Donaldson confirms of Afrikaans, "[A]ll formal trace of the subjunctive has disappeared, mood being expressed periphrastically by means of *sou* 'would', should the need arise" (1994: 498). Those languages devoid of all verb inflections—among them Mandarin, Maybrat, Nahali, Sango, and Vietnamese—are, naturally, not possessed of grammatical moods, such as the indicative or subjunctive, though these may nonetheless be expressed by alternative linguistic means. (Bickel and Nichols, 2013) Additionally, many sign languages have grammatical indicators for conditionals—for example, raised eyebrows with signing of the antecedent, accompanied by a forward head thrust, among other non-manual markers, in American Sign Language and others (Liddell, 1986); and to an extent for distinguishing what are identified as counterfactual conditionals, typically, but not always, marked by an additional eye squint (Pfau and Quer, 2010)—but they do not have the sort of explicit grammatical moods, such as the subjunctive, found in certain spoken languages.

Linguists generally take indicators of counterfactuality to be the means by which languages without grammatical moods are able to convey the same sort of hypotheticality, possibility, or lack of certainty that the subjunctive offers in those languages that make use of it. Whether the subjunctive and the counterfactual conditional may be considered largely interchangeable is something I examine further below[5],

[5] *See p.33.*

but what is expressed by means of the subjunctive in English is nonetheless widely translated into languages without it by means of non-tense or -mood markers of counterfactuality.

Consider this example of a subjunctive conditional in English:

(14) If I had gone to the movies that evening, I could not have had dinner with my mom. (Yeh and Gentner, 2005: 2411)

As Yeh notes, counterfactuality is indicated by the use of the past perfect tense in the antecedent, and the subjunctive mood in the consequent. However, for the (transliterated) Chinese version,

(15) Ru guo wo na tien wan shan qu kan le dien ying, wo jiu bu neng gen wo ma qu chi wan fan. (Yeh and Gentner, 2005: 2411)

its direct translation into English is roughly, "If that night I go watch (past particle) a movie, I then cannot accompany my mom to go eat dinner." This is because there are no verb inflections in Chinese, and thus,

> [A] counterfactual is signaled by comparing the tense information—e.g., the past tense particle (le) after the verb—with contextual information as to

> whether the event actually occurred. The second clause is simply a consequence clause; it has no internal marker of counterfactuality. To detect counterfactuality, the hearer must compare the sentential assertion with context. (Yeh and Gentner, 2005)

This is echoed by South African Sign Language (SASL) practitioners, who stress the dependence on extensive context for the adequate communication of counterfactuality in and accurate translation of subjunctive English conditionals into SASL.[6] As Yong (2013) further argues, Chinese is possessed of various means of suggesting counterfactually, but these implicatures are still cancellable—and, it is also possible to express counterfactuality without making use of any such indications. What is perfectly evident, is that there is nothing like the subjunctive mood that could be used to divide instances of the conditional in any way resembling how these divisions might arguably be made in English or other subjunctive-possessing languages.

Similarly, in Japanese, there is often no clear distinction drawn between what may actually be the case and what is being considered purely hypothetically (Wierzbicka, 1997: 45):

(16) (Moshi) kare ga kanemochi nara/dat-tara/deare-ba, kocchi o erabu-daroo.

[6]Dr. Michiko Kaneko, head of the Department of South African Sign Language at the University of the Witwatersrand, and South African Sign Language interpreter Lindsey Rielly, in discussion with the author, July 6, 2015.

Wierzbicka (1997) gives the direct translation of this transliteration of the Japanese, as "([B]y any chance) he rich this choose will/would", which in turn—depending on the context in which (16) is used—may be translated into English as either of the following:

(17) If he is rich, he will choose this.

(18) If he were rich, he would choose this.

In cases such as this one, Japanese fails to make the grammatical distinction between indicative and subjunctive that the differing moods of the English sentences (17) and (18) do.

Some theorists have argued that there are languages that lack an explicit conditional construction, in which any form of subjunctive conditional would, very obviously, be absent. But do they genuinely fail to have a lexically encoded conditional? Various Australian languages are claimed to fall among these, and to "use paratactic means only" to establish conditionality, according to von Fintel (2011: 4), who gives Levinson's example from Guugu Yimithirr:

(19) Nyundu budhu dhada-a, nyundu minha maa-naa bira.

which may be translated verbatim as, "You maybe go, you meat get for sure"—equivalent to the English conditional, "If you go, you'll get meat." (von Fintel, 2011: 4)

However, Wierzbicka (1997), challenges the claim that examples such as (19) provide evidence for a lack of lexicalized conditional construction in the cited language, arguing that this is the result of failing to appreciate the degree of lexical polysemy in the relevant languages. She maintains that terms translated as 'perhaps' or 'maybe' may be used as either of the English words 'when' and 'if', but that these uses are linguistically distinguished, thus providing a conditional construction in such Australian languages. In fact, the Guugu Yimithir word 'budhu'—rendered by Levinson as "maybe"—is translated elsewhere as "if" (Haviland, 1979: 151–152). The very same sentence, (19) is given verbatim by Haviland as, "You if go, you meat get indeed',' or in grammatical English as either, "If you go, you'll get meat for sure," or "Should you go, you'll get meat for sure." (Haviland, 1979: 152) Here 'budhu' is explicitly translated as 'if'; as a word that "signals uncertainty, or questions the possibility of some outcome, sometimes very much like a subordinate conjunction, sometimes in a more modal sense." It is, thus, highly improbable that languages such as Guugu Yimithirr do not have any genuine conditional construction. I stress this because of its significance in establishing that the conditional is universal to natural language and that it is therefore that much more likely not to be a bound to the grammatical or other idiosyncratic features of any specific individual language or language family.

Irrespective of the lexical status of the conditional, though, it is entirely undis-

puted that nothing like the subjunctive mood exists in Guugu Yimithirr and its related languages, among a number of others—which presents an insurmountable obstacle for any theorist wishing to posit some sort of indicative/subjunctive division as the basis for a deeper logical divide of the conditional.

Another concern facing any proponent of dividing instances of the conditional according to the indicative/subjunctive split is that, even in languages such as English, there are instances in which there appear to be no substantive practical differences in meaning whether we use the indicative or the subjunctive mood[7]:

(20) *indicative:* It rained last night. So, if the burglars ran through the flowerbed, they left footprints. Let's go check...

(21) *subjunctive:* It rained last night. So, if the burglars had run through the flowerbed, they would have left footprints. Let's go check...

Even should there be examples in which some significance could be made of the differences in grammatical mood between (20) and (21), there are, notwithstanding, quite clear contexts in which no such substantive difference in meaning is evident—as were, say, a police officer at the scene of a house-breaking to ponder aloud the question of whether or not the burglars left footprints.

Further difficulty for the indicative/subjunctive division of the conditional lies in arguable distinctions even within the category of subjunctive conditional itself.

[7]Examples adapted from Abbott (2010).

This raises the possibility that an indicative/subjunctive split would not serve the purpose of distinguishing, even in English, between conditionals concerning reality and those concerning merely what's possible or at least potentially hypothetical. To take an example from Davis (1979: 546):

(22) If the light switch were flipped, the light would come on.

(23) If the light switch should be flipped, the light would come on.

Davis argues that there is a difference in what is implicated by (22) and (23), respectively; such that the former carries the implication that the switch will not be flipped, while the latter does not. As a native speaker of English, I do allow the possibility of a subtle distinction here, though I cannot say I'm confident I feel the force of this purported difference.[8] But, whether one is disposed to agree with Davis or not, his point nonetheless emphasises just how muddy the waters of the English subjunctive mood are.

Being altogether absent in a variety of languages, a conditional's mood is quite clearly a hopeless guide to its deeper logical structure. Moreover, there is some evidence to suggest that, even in languages where this distinction does hold, the split it

[8] Another interpretation, suggested to me by Gary Ostertag, is that (22) concerns a light switch that might be flipped at any time, whereas (23) concerns one that might be flipped in the future. This distinction, too, I have difficulty in feeling the intuitive force of, but this level of confusion and disagreement among native speakers aptly demonstrates the problematic state of the English subjunctive.

makes fails to demarcate an always logically significant divide, and, if Davis's point is accepted, these categories of grammatical mood make also insufficiently fine-grained distinctions.

Disentangling the subjunctive from the counterfactual

We have every justification, then—for the above multiple reasons—to reject the indicative and subjunctive moods as a basis for any kind of general, cross-linguistic, and logically significant distinction among instances of the ordinary-language conditional. But the category of subjunctive conditional is frequently used almost interchangeably with that of counterfactual conditional. Perhaps we might give the notion of the subjunctive the benefit of the doubt, and take it that what is really intended when it's used of a conditional, is to mark it as a counterfactual. This would, certainly, avoid the above-mentioned problems. However, even though the category of subjunctive conditional might perhaps be more usefully replaced with that of the counterfactual conditional, this latter runs quickly into difficulties of its own.

It is important, first, to clearly distinguish the subjunctive from the counterfactual. Though they are frequently treated as largely interchangeable in much of the philosophical literature, it is easy to see that these, respectively, grammatical and metaphysical categories are by no means so. The subjunctive mood is a feature of particular languages, and far from universal. Moreover, even in those languages where it is present, it fails to correspond always to the counterfactual.

Consider, for instance, the following conditionals. These all have as their focus counterfactual scenarios and not actuality, though they are not in the subjunctive

mood. As a result of the degraded state of the subjunctive mood in English, there are any number of conditionals that are of purely counterfactual interest, though grammatically they are in the indicative mood. This does not, of course, show that (24) to (27) could not apply to things as they are in reality, merely that they *may* concern what is counterfactual—and therefore that their grammatical mood in English cannot be determinative in this regard.

(24) If aliens invade Earth and kill us all, then it's not true that they will have come in peace. (I am presuming that aliens are not going to come to Earth, let alone invade and kill us all.)

(25) If I turn into a bear tomorrow, my family will be very shocked. (I am not going to turn into a bear at any time.)

(26) If you are Pegasus, then you are a magical winged horse. (You are not Pegasus.)

(27) If her plane was delayed, then she arrived too late to give her talk. (In fact, she arrived on time and gave her talk.)

None of these indicative conditionals describe any part of the world as it is in actuality, and yet all are in the indicative mood. Again, they could concern things

as they are in reality; the point here is purely that they need not necessarily do so. If we were to divide instances of the conditional by whether or not they were counterfactuals, in place of considering their grammatical mood, we clearly could not combine the category of counterfactuals with that of indicatives.

Just as there are seemingly counterfactual indicative conditionals, there are also subjunctive factual conditionals, ones primarily concerned what is in fact the case at the actual world (though, as above, this may be just contingently so). For example:

(28) *Where would they have hidden their keys? I wonder.* If they were as unimaginative as the average person, then they would just have put them under the doormat. *And, sure enough, here the keys are—under the doormat.*

(29) If it were lupus, we would expect the patient to be suffering from fever, fatigue, and joint pain. *And, indeed, the patient has a classic case of lupus, so it's not surprising he's experiencing all of these symptoms.*

(30) If she should have found the key you dropped earlier, then she would have been able to steal the plans before the alarm was reset! (She did find the key.)

(31) If you would put out that cigarette, I'd be very grateful. (You do put out the cigarette.)

It may be objected that the form of the subjunctive used in (30) is somewhat archaic, but, even without it, the remaining examples still demonstrate that there are instances of subjunctive conditionals that concern factual states of affairs; things as they are at the actual world.

Thus, while the categories of subjunctive and counterfactual do overlap significantly, the two terms are by no means interchangeable, even in English. This might, on the face of it, seem an advantage—for one, visual languages appear to distinguish counterfactuals from other conditionals (for instance, as discussed above with regards to American Sign Language, and as is the case in Israeli Sign Language, among others; Dachkovsky and Sandler, 2009), as do most, if not all, spoken languages. (Wierzbicka, 1997). However, it is in fact highly problematic, in light of the fact that the category of counterfactual conditional is typically contrasted with that of the indicative conditional.

The indicative/counterfactual conditional division

The division of the conditional into indicative and counterfactual instances (as in Lewis, 1973b) is an essentially confused one, given that one of its categories is grammatical and the other metaphysical—and that many conditionals fall squarely into both of these groups. The result is a pair of categories that are neither mutually exclusive nor exhaustive.

The categories of indicative conditional and counterfactual conditional, respectively, fail to be mutually exclusive as they overlap with one another. Examples such as those given above, in (24) to (27), clearly demonstrate that a conditional may easily enough be both indicative in mood and counterfactual in terms of the content it is intended to refer to. Consider one additional instance of an indicative counterfactual,

(32) If you throw that ball in the house, you'll break something. (You do not throw the ball in the house.)

At the same time, these two categories are also not exhaustive. Those conditionals that are neither in the indicative mood nor concern counterfactual states of affairs fail to be categorized at all. The examples (28) to (31), subjunctive in mood but concerning factual states of affairs, fall in neither of these categories; nor does the following instance,

(33) If she were to perform the surgery within the next hour, he would very likely live. (She does perform the surgery within this timeframe.)

As David Lewis himself admits in his formative work on counterfactuals, "Counterfactuals with true antecedents—*counterfactuals that are not counterfactual*—are not automatically false, nor do they lack truth value'. [...] You may justly complain, therefore, that my title 'Counterfactuals' is too narrow for my subject. I agree,

but I know no better." (Lewis, 1973b: 3, emphasis mine) Here Lewis is essentially conflating counterfactual with subjunctive conditionals.

In addition to these sorts of problems with placing certain conditionals in both groups and others in neither, taking the indicative mood to delimit one of the categories embroils the indicative/counterfactual taxonomization in all of the difficulties that come with identifying a type of conditional with the grammar of a particular language (or languages), as detailed above. An explication of the logical structure of the conditional, as it is actually used in natural language, can surely not be language-specific; otherwise, what sense can be made of the indicative category for classifying instances of the conditional in Chinese, in South African Sign Language, in Vietnamese, in Guugu Yimithirr? Unless we take an unwarrantedly parochial approach to taxonomizing types of the ordinary-language conditional, we must jettison English grammatical categories completely.

It is very evident, then, that the common division of conditionals into indicatives and counterfactuals is an incoherent and unjustified one. This cannot help but have deleterious effects on the project of determining the underlying logical structure of the conditional, when there are obvious instances that fall into both categories and others that fall into neither even in those languages that do have an indicative mood.

On the other hand, one might more charitably take the indicative/counterfactual split to be intended to be equivalent to a factual/counterfactual split—in other words, that the basic distinction being sort is really a metaphysical as opposed to grammatical or mixed one. Though not philosophically dispositive, it is nonetheless suggestive that the notion of a factual conditional is assumed in much of the linguistics liter-

ature on the conditional: "It is traditional to recognise, as special sub-categories of conditionals, the factuals and counterfactuals. In a factual conditional, the content of the *if*-clause is presumed to be the case, whilst in a counterfactual the content of the *if*-clause is taken to be contrary to fact." (Taylor, 1997)[9] And this is especially true of work on conditionals in those languages without verb inflections (*see* Fang, 2014; Pfau and Quer, 2010; Dachkovsky and Sandler, 2009).

However, while this categorization—in terms of conditionals whose antecedent concerns what is actually the case, and those concerning rather what might have been or may be—does avoid the myriad problems facing other divisions of conditional types so far discussed, the difficulty with the factual/counterfactual divide is that it fails to track any clear, corresponding distinction between instances of the conditional themselves. This is because there are any number of conditionals that, as used in everyday discourse, cannot be determined to be either factual or counterfactual at the time of their assessment.

Ordinary speakers have no trouble, for instance, determining the truth-value of the following,

(34) If you put that ice-cube in the fire, it will melt.

[9]However, there is also a third category presumed here; namely that of the hypothetical conditional, "in which the content of the *if*-clause is entertained as a possibility, neither in accordance with reality, nor necessarily inconsistent with it." (Taylor, 1997: 301–302)) This is a category that, I argue, can be dissolved on the factual/counterfactual division I propose, as detailed further below.

without having any idea whether the person being addressed will *in fact* put the ice-cube in the fire. That we, as ordinary language users, are able to, and frequently do, use such conditionals without difficulty, suggests that the logical form of the conditional in itself cannot differ between instances that are supposedly wholly factual and those supposedly wholly counterfactual. But, before I substantiate this criticism of even a factual/counterfactual classification of conditionals, I would like first to consider Graham Priest's alternative proposed categorization, which makes no direct use of either grammatical moods or metaphysical distinctions but also divides the conditional at the level of its various instances.

The prime/non-prime divide and subjunctive tense change

Priest (2009) argues for a distinction between prime and non-prime conditionals; namely, those whose antecedents can stand, grammatically, on their own as a sentence and with the same sense as they have embedded in the relevant conditional (prime conditionals), and those that cannot satisfy both of these criteria (non-prime). He argues that, "If there is a natural logical break in the genus of conditionals it is between prime and non-prime conditionals. Prime conditionals are always indicative. Non-prime conditionals can be indicative or subjunctive/counterfactual". What exactly, then, is the significance of this split? Consider Priest's example of a non-prime indicative conditional, his suggestion of its corresponding prime version, and the analysis he gives of this:

- If it rains, you will get wet.

The corresponding prime conditional is:

- If it is raining, you get wet.

The tense combination (antecedent, consequent) is ⟨*present*,*future*⟩; and what this gets us to do is to evaluate the prime conditional at a future (indefinitely specified) time, t. This is why, if we have 'It will rain' (i.e., 'It is raining at time t') we can apply *modus ponens*—at t, so to speak—to infer 'You get wet at time t' (i.e., 'You will get wet'.)" (Priest, 2009: 323–324)

However, I would argue that this fails to capture fully just what it is that the prime/non-prime distinction gestures toward—at least in English. The tense-shift in non-prime conditionals is taken to provide temporal information, telling us something about the time relative to which a particular conditional should be assessed, with the consequent giving us the time with which the content of the conditional is concerned and the antecedent the time at which it ought to evaluated. The antecedent of the non-prime conditional here is 'back-shifted' from that of the consequent—that is, the consequent, "You will get wet", is in the future tense, while the antecedent, "It rains", is in the present tense. According to Priest, this non-prime conditional must be evaluated by means of its prime equivalent at the time indicated by the tense of the original antecedent.

But, though tense is clearly significant here, the tense of the antecedent—'back-shifted' one degree to the present tense from the future tense of the consequent—is playing no sort of temporal role. Rather, this sort of tense shift is a marker of the subjunctive mood, even in apparently indicative conditionals and in the absence of any specifically subjunctive terms (such as 'were', 'would', and the like). (von Fintel, 2012; Molencki, 1996)

> In languages that have a past subjunctive (such as German), antecedents of conditionals of the second kind do indeed appear in the subjunctive mood but if a language does not have a past subjunctive, some other form is used. English uses an indicative antecedent with an additional layer of past tense morphology ("if Oswald didn't kill Kennedy" becomes "if Oswald hadn't killed Kennedy"). It is that additional layer of past tense morphology (which doesn't obviously contribute an actual past meaning) that quite reliably signals conditionals of the second kind [roughly, subjunctive conditionals]. (von Fintel, 2012)

In other words, in English where there is a back-shift from whichever tense gives the time relevant to what the conditional concerns, this shift is a marker of the subjunctive past. However, a stronger claim than von Fintel's is warranted—it is not merely that this additional layer of past tense fails to convey an obvious past meaning; it is in fact no kind of temporal marker at all. Instead, this layer of past morphology is an indicator that the conditional is concerned with what is merely possible and not actual—or, more accurately, with what at least *may not* be actual,

even should it turn out to be. And this is just what the subjunctive is, an indicator of a hope, wish, or possibility; something not certain—even in some instances in which the grammatical form is apparently indicative. Indeed, to return to Fowler (1926), he includes in the subjunctive that which is, "understood to be [grammatically] modally different from the indicative but is [...] indistinguishable from it in form"'. (For the remainder of this chapter only, I will subscript the term 'modal' with either a '$_G$' for grammatical, to mean modal in Fowler's sense; or with an '$_M$' for metaphysical, to mean modal in the sense of distinguishing between reality and possibility, what is actual and what simply counterfactually possible.)

This modal$_G$ role of tense back-shifting becomes more evident when we look at what I suggest is in fact the correct prime counterpart, (36), of the non-prime conditional Priest considers, repeated below in (35); namely,

(35) If it rains, you will get wet.

(36) If it's going to rain, you will get wet.

That the restatement, (36), has essentially the same sense as the non-prime original, (35), can be seen by applying Priest's *modus ponens* test: The antecedent of (36) can stand as the second premise of an argument that has the whole conditional of (35) as its first premise and the consequent of both as its conclusion. Thus,

1. If it rains, you will get wet.

2. It is going to rain.

3. You will get wet.

The antecedent of (36) clearly does, then, have the same sense as that of (35). And what this shows is that non-prime conditionals may be 'translated' into prime conditionals that indicate the correct time at which they should be assessed without any adjustment. Thus, a discrepancy between the tense of the antecedent and that of the consequent in a non-prime conditional does not tell us about the time with which the content of the conditional is concerned. Instead, the back-shifted tense of the antecedent in the above case is a modal$_G$ marker; that is, an indicator of the subjunctive mood and not a past time.

But, what of the many different tense pairings in various non-prime conditionals—over and above that of just ⟨*present*, *future*⟩—that Priest notes? For instance, we frequently find ⟨*pluperfect*, *pluperfect*⟩ combinations in conditionals concerning how things might have been had the past been different from the way it actually turned out. For instance,

(37) If Vusi hadn't forgotten his ticket, he would have caught the same train as us.

Here, the tense of the antecedent and that of the consequent match one another, but the time to which the content of the conditional as a whole relates, and relative to which it should be assessed, is best given by just the simple past—not an earlier time in relation to an already past time, as with a purely temporal use of the pluperfect.

This suggests that the category of non-prime conditionals may not fully capture what is really salient about these tense-shifts with regard to the accurate analysis of the conditional's truth-value. The characteristic that Priest identifies, of a clause being able to stand on its own with the same sense as it has embedded in the relevant conditional, is a useful indicator of when a tense is a modal$_G$ marker rather than a temporal one. But, in fact, it's applicable both to antecedent and consequent, not simply the antecedent alone. This is because in certain conditionals, where the tenses of antecedent and consequent match but both diverge from what would be the most accurate tense in purely temporal terms, neither is able to stand on its own with the same sense; the reason being that both make use of tense-shifts that are modal$_G$ in nature.

Linguist Andrew Nevins, following Iatridou (2000), argues for this contention that, in the case of such conditionals using the pluperfect (past-past tense) in the antecedent and consequent, this 'double past' form is really just the 'single past' in terms of tense, with the second 'layer' of past being a modal$_G$ and not a temporal indicator. (Nevins, 2002) As Iatridou (2000: 239) explains, regarding the use of tense to indicate the subjunctive mood in the following expressions,

> "To convey an unfulfilled desire about the present, a speaker uses past tense

morphology. We are dealing with a fake past since this past tense morphology does not receive a past interpretation. To convey an unfulfilled desire about the past, the speaker uses the pluperfect.

a.*I wish I have a car.

b. I wish I had a car (at present).

c. I wish I had had a car (back then).

To make this clearer, consider these further examples,

(38) If Catalina's train arrived in Coyoacán on time, then she presented her paper at the first conference session.

(39) If Catalina's train had arrived in Coyoacán on time, then she would have presented her paper at the first conference session.

Both of these conditionals concern the same past times, relative to the time, *t*, at which they are asserted—both antecedents, *t-2*, the time at which Catalina's train was scheduled to arrive, and both consequents, *t-1*, the time at which Catalina's paper was scheduled to be delivered by her. Thus it seems clear that (38) and (39) are temporally equivalent; the respect in which they differ being their modal_G import,

and not the specific times of the events with which they are concerned. The first sentence leaves open the possibility that the antecedent (and consequent) is true at the actual world, whereas the second is counterfactual, making clear that Catalina's train did not arrive on time at the actual world.

This analysis of past tense as indicator of grammatical mood gains further *prima facie* support from the fact that, when a tense's back-shifting has modal$_G$ import and an explicitly modal $_G$term is used (as in the consequent of (39)), this is often referred to as a form of modal perfect tense instead of simply an ordinary perfect tense. Obviously terms such as 'would', for instance, have modal$_G$ import, but the fact that this leads to the tense itself being described as modal$_G$ is suggestive. In addition, there are certain dialects of English in which speakers render conditionals like (39) as,

(40) If Catalina's train would have arrived in Coyoacán on time, then she would have presented her paper at the first conference session.

This inclination to add an explicitly modal$_G$ term, such as 'would', to the antecedent in these sorts of cases[10]—and the fact that the modal,$_G$ when used in this way, takes

[10]This usage is common in spoken language; to take just three examples of actual usage of this kind, from the Corpus of Contemporary American English:

i. "As president, if I would have signed that legislation, I would have had an action plan ready to go immediately," spoken, Fox News Channel, October 14, 2011. (Davies, 2008)

ii. "Jay McClain said, 'If she would have been healthy, she would have been our starting right fullback,"' quoted speech, *Denver Post*, June 18, 2006. (Davies, 2008)

the place of the second 'layer' of past tense (the pluperfect 'had')—gives further credence to the claim that one of the two 'layers' of past in the clause is to be taken as modal_G rather than temporal.

Distinguishing the Oswald–Kennedy pair

But, if the prime/non-prime distinction is not to be interpreted in the manner that Priest advocates—what then of that thorny Oswald/Kennedy pair, which he appears to explicate precisely by means of his analysis of this distinction?

(41) If Oswald didn't shoot Kennedy, then someone else did.

(42) If Oswald hadn't shot Kennedy, then someone else would have.

Of course, since there are languages in which grammatical moods are entirely absent, the putative indicative/subjunctive split cannot effectively distinguish between (41) and (42) cross-linguistically, at the level of logical structure. And, insofar as the common indicative/counterfactual division also incorporates the English grammatical category of the indicative mood, it too must be rejected here for the same reason.

iii. "'If you would have seen her working with the kids on the field, if you would have seen her in the classroom, if you would have seen her in life, she was a typical 25-year-old woman,' Chessmore said," quoted speech, *Associated Press*. (Davies, 2008)

However, the prime/non-prime classification of instances of the conditional might be thought able to avoid these concerns, and to provide an explanation of the inequivalence of the respective meanings of the Oswald–Kennedy pair. That is, one might contend that what the prime/non-prime split is in fact getting at is of significant logical interest and not necessarily dependent on surface grammatical form.

Priest argues that the prime/non-prime distinction here indicates the different times of evaluation that must be used for the assessment of (41) and (42) respectively. The claim is that, "a non-prime conditional with a certain tense structure has the same sense as the corresponding prime conditional, but evaluated at a different temporal location." (Priest, 2009: 12)

A prime conditional is to be evaluated relative simply to the time given by the tense of its antecedent—in the case of (41), this is a time in the past (relative to this present moment), after Kennedy had been assassinated. Thus, if it was true after the assassination that Oswald didn't shoot Kennedy, then the antecedent may be taken to be true. Since the consequent would then also be true, so too would be the conditional (41) as a whole. For a non-prime conditional, though, the time relative to which it must be assessed is provided by the tense of the antecedent, but the form of the conditional to be so assessed is given by the prime conditional corresponding to the non-prime one. In the case of (42) the tense of the antecedent is pluperfect (past past) which, Priest maintains, indicates that this conditionals's truth-value must be determined by assessing the corresponding prime conditional relative to a time two "degrees" into the past (that is, the past relative to the past); namely, before the assassination took place. The prime conditional he gives as corresponding to (42) is

(Priest, 2009: 12),

(43) If Oswald does not shoot Kennedy, someone else will.

In other words, the truth-value of (42) is given by the truth-value of (43), as assessed in the past past—prior to Kennedy's assassination. And, assuming that Oswald acted alone in killing Kennedy, (43) is false, thus making (42) false.

But, for the distinction to reveal something substantive for our evaluation of the conditional, it cannot be merely a quirk of English grammar, given the fatal difficulties faced by such forms of categorization discussed above. However, the prime/non-prime distinction does appear to rely on contingent features of this language, rather than tracking a deeper, logical difference.

To make this clearer, let's move away from English for the moment to get some perspective on its particular grammatical character. If we consider (41) and (42) translated into Afrikaans, it seems clear that we cannot rely on a distinction of the prime/non-prime sort to account for the difference in their meaning, as the distinctions evident in the English instances fail to match up with their Afrikaans translations:

(44) As Oswald nie vir Kennedy geskiet het nie, dan het iemand anders dit gedoen.

(45) As Oswald nie vir Kennedy geskiet het nie, dan sou iemand anders dit gedoen.

Neither of the Afrikaans sentences are prime, since the antecedent of (44) is, "Oswald nie vir Kennedy geskiet het nie,"' which cannot stand on its own grammatically. The auxillary verb "het" (roughly equivalent to the English "did") is exiled to the end of the dependent clause in a conditional construction, whereas, in order for this phrase to form a grammatical sentence on its own, "het" must come directly after the subject, "Oswald". Thus, rendered grammatical, the antecedent of (44) taken on its own would be, "Oswald het nie vir Kennedy geskiet nie."

Moreover, Afrikaans lacks a pluperfect tense[11], and so the subjunctive mood cannot be indicated in the antecedent alone by means of the mood-conveying (as opposed to temporal) back-shift made use of in English. The conditional (45) as a whole, though, is nonetheless in the subjunctive mood, in virtue of the use of "sou"—which translates as "would have"—in the consequent. However, it remains that no translation of (42), "If Oswald hadn't shot Kennedy, then someone else would have," will result in a prime conditional. The only conditionals that would be prime in Afrikaans are those with intransitive, adjectivally-unmodified verbs in the

[11]There are some pragmatic, mostly colloquial, ways in which to indicate in Afrikaans that the pluperfect rather than the simple past is intended, such as (to take (45) as an example) adding "nooit" ("never" in English) to the antecedent of (45), so that it reads, "As Oswald nooit vir Kennedy geskiet het" (Donaldson, 1994: 231–234). Also, the present tense is frequently used to describe past events where it is contextually quite clear that it is the past under discussion. Thus, the simple past may be employed in such instances as an ersatz pluperfect—the pragmatic effect, of using the past rather than the more-typical present tense, being to convey that the pluperfect is intended. (Raidt, 1995: 136)

present tense.[12] The introduction of any adverb or object, or the use of an auxiliary verb (which is required in almost every past-tense construction), adds an element that requires a change in placement of the main verb when this sentence is used as the antecedent of a conditional. For instance, (46) may form the antecedent of a conditional just as it is,

(46)	Die hond vreet.	As die hond vreet, sal sy siek word.
	(The dog eats.)	*(If the dog eats, she'll get sick.)*

and is, therefore, prime.

However, as soon as any adverb, object, or auxiliary verb is introduced to the sentence making up the antecedent, it cannot retain the same word order in a conditional construction and must be reordered to form the antecedent of a conditional. For instance,

(47)	Die hond vreet vinnig.	As die hond vinnig vreet, sal sy siek word.
	(The dog eats quickly.)	*(If the dog eats quickly, she'll get sick.)*

[12] In contrast, it seems very likely that almost every conditional in those various spoken and signed languages lacking verb inflections would turn out to be prime, also failing to separate out the members of the Oswald–Kennedy pair.

(48) Die hond vreet die brood. As die hond die brood vreet, sal sy siek word.
(The dog eats the bread.) *(If the dog eats the bread, she'll get sick.)*

(49) Die hond het gevreet. As die hond gevreet het, dan het sy siek word.
(The dog ate.) *(If the dog ate, then she got sick.)*

Indeed, this V2 (verb second) word ordering is not unique to nor even most pronounced in Afrikaans, but is a general feature of Germanic languages to varying degrees[13]. It is evident, for instance, in both Dutch and German, and occurs to a lesser extent in Swedish. For example, the word order of the Swedish statement (50) is altered when it is subject to a sentential modifier; for instance:

(50) De går på bio på måndag. Vantligtvis, går de på bio på måndag.
(They go to the movies on Monday.)
(Usually, they go to the movies on Monday.)

These considerations, then, tell against the prime/non-prime distinction, given its reliance on the contingencies of a given language's grammar. The fact that the

[13]English is in fact anomalous among Germanic languages in retaining only minimal vestiges of this phenomenon; for example, in the following cases, where V2 has not been entirely supplanted by Modern English VSO (verb, subject, object) order:

V2: Where is my pen? *VSO:* Where my pen is? / My pen is where?
V2: "Help!" cried the man. *VSO:* "Help!" the man cried.

distinction does not make the same discriminations across languages suggest that it cannot be an indicator of any deeper logical distinction in the ordinary speaker's evaluation of the conditional, which should one that cuts across natural languages.

But, even if the prime/non-prime division may not effectively pick out the difference between the Oswald–Kennedy conditionals using its own diagnostic tools, mightn't it nonetheless—albeit in an 'English-bound' manner—still be pointing toward a more fundamental distinction in types of conditional? Could the differences in the time relative to which each respective member of the pair should be evaluated not be the key to explaining the distinct truth-values ordinary speakers intuitively give them? Differences that could perhaps be picked out using a modified form of the prime/non-prime distinction. While the different between (41) and (42) may not be genuinely temporal, as we have seen, perhaps it could nonetheless indicate an inherent distinction in the correct time relative to which each of these conditionals should be assessed.

For (41), one could argue that it should be assessed at that point just after Kennedy has been been shot, when it is the case that someone has shot him—if not Oswald, then someone else. For (42), the argument is that the prime version of this conditional is (43), repeated here with its non-prime original:

(41) *non-prime:* If Oswald hadn't shot Kennedy, someone else would have.

(42) *prime version of (42):* If Oswald does not shoot Kennedy, someone else will.

The claim is that, (42)'s antecedent gives us the relevant time of evaluation, one *before* Kennedy is shot. And, if we evaluate (42)'s corresponding prime version, (43), then there is no reason to think at this time beforehand that someone else would commit the shooting should Oswald fail to—and, thus, this explains why the conditional (42) is false, while (41) is not. Even if this time of assessment cannot be given by the relevant conditional's antecedent in every language, could it not still be significant and perhaps derived in some other, more general, way?

However, if we strip (42) down to just its temporally-relevant tense, it becomes identical to (41), since these two conditionals are *temporally* equivalent. They concern the very same time on November 22, 1963; that of Kennedy's assassination, and it is hard to see how one might justify assessing each of these at different times. Moreover, it is not clear that (43) is the correct prime version of (42). If we take away the subjunctive markers, both obvious indicators of subjunctive mood and the subjunctive-indicating layer of past morphology, we are left with the following, which is simply (41):

> *revised prime version of (42):* If Oswald didn't shoot Kennedy, then someone else did.

For, without the extra, non-temporal layer of tense, the antecedent's pluperfect "had not shot" becomes just the simple-past "did not shoot", and the consequent's "would have" becomes "did". Remember, these are features of particular languages (in this

case English) and not linguistically universal, and so cannot be used to determine general distinctions in the conditional at the logical level.

To make this clearer, let's call the time immediately after Kennedy is shot *t*. For the prime version of (42) given by (43) it's claimed that the correct time of evaluation for (42) is time *t-1*, *before* Kennedy is shot. But, this requires that we take the second layer of past—which in fact forms part of this English sentence's subjunctive character—as genuinely temporal, when it is not. In any natural language without this sort of subjunctive-marking past morphology the truth-value ascriptions for the Oswald–Kennedy pair would not come out correctly on this approach.

What the revised prime version of (42) asserts is the case at time *t* is exactly what (41) does (obviously so, as they are identical); namely, that Kennedy has been shot and that either Oswald shot him or someone else did. Of course, the prime version of (42) that Priest gives is actually (43), rather than (41). But, as we have seen, there is no principled basis for taking the prime version of (42) to be (43), to be assessed at *t-1*, before Kennedy is shot. In fact, considered relative to this time, (41) would also have to be given as (43). The reason is that, since at *t-1* the shooting has not yet occurred, the claim (41) actually makes (taken relative to *t-1*) is that if Oswald was not that future shooter, then someone else was. Thus, from the perspective of *t-1*, (41) must say that either Oswald will shoot Kennedy or someone else will. This makes (41) and the prime version of (42), taken at *t-1*—at which they must both be rendered as (43)—identical to one another.

One might, though, deny that (41) really is equivalent to (43), relative to *t-1*, since this latter formulation appears to assume just what (41) does not; that Kennedy

will assuredly be shot, regardless of what Oswald does. In contrast, it seems that (42) does make this assumption, that Kennedy will be shot, independent of Oswald's participation or lack thereof. The trouble with this objection is that, if we consider what each of (41) and (42) implicitly assume, then there is no reason to evaluate these conditionals at any earlier time. If what constitutes the correct rendering of the conditional, *vis-à-vis* a given time, is determined by what is taken to be implicit in the original form of the conditional, then we are just as well off assessing (41) and (42) as is, in the present, in terms of precisely such determining factors.

Thus, there is nothing to be gained from assessing a conditional relative to one time rather than another; and, in fact, this cannot even be accurately done without first giving the sort of analysis of the original conditional that this move was intended to provide.

However, it remains that there is undeniably a difference between (41) and (42), though one not uniquely encoded by the English subjunctive or effectively captured by any of the permutations of conditional categorization I have so far surveyed. What distinguishes the truth-values of (41) and (42) is evidently a modal_M difference (to return to the metaphysical, as opposed to the grammatical, meaning of the term). But, this cannot be handled simply by splitting conditionals into factuals and counterfactuals, which one may suppose is what at least some of those using the indicative/counterfactual division are, in practice, doing. This is precluded by the fact that no instance of the conditional has any sort of *inherent* modality_M, in the sense of its being essentially factual or counterfactual in itself; rather, each conditional may be interpreted either factually or counterfactually.

On this account, what then separates the Oswald–Kennedy pair is the use of modal$_G$ terms and tense shifts in English; of non-manual grammar of eyebrow raises and forward head-thrusts in American Sign Language; of conversational context in Vietnamese—and any number of other grammatical, pragmatic, and contextual markers across the full spectrum of natural languages—which convey modal$_M$ distinctions, as captured in my analysis by the *ceteris paribus* clause implicit in the counterfactual interpretation of the conditional.

Chapter 2: The $C3$ Conditional

A Semantics for $C3$

Before defending the claim that the logical structure of the conditional, as abstracted from its ordinary use by English speakers, is best given by $C3$, let me first provide a semantics for this logic, based in significant part on Priest's (2008) $C, C^{+}, C1$, and $C2$ systems.

Let $\mathscr{F}$ be the set of all formulae of the language of $C3$, with formulae specified by the following recursive set of conditions:

For all atoms A, $A \in \mathscr{F}$.

For all $A \in \mathscr{F}$, $\neg A \in \mathscr{F}$.

For all $A, B \in \mathscr{F}$, $A \wedge B$, $A \vee B$, $A \rightharpoonup B$, $A \rightharpoondown B$, and $A \Rightarrow B$ are in $\mathscr{F}$.

A model for $C3$ takes the form of an ordered triple, $\langle W, \{R_A : A \in \mathscr{F}\},\ \nu\rangle$; W and ν each being non-empty sets, and $\{R_A : A \in \mathscr{F}\}$ a collection of binary relations

on W, R_A, one for each formula, A. ν is a function taking arguments from $W \times \mathscr{F}$ to a set of truth values $\{0, 1\}$.

W should be understood to be the set of all possible worlds, and, for all worlds w, $w \in W$. ν should be understood as the set of all functions such that each takes arguments from $W \times \mathscr{F}$ to a set of truth-values $\{0, 1\}$. We will abbreviate $\nu(w, A)$ as $\nu_w(A)$. wR_Aw' should be read as w''s being *ceteris paribus* the same as w, except that A, the antecedent of the conditional, is true at w'; in short, that w' is A-accessible from w.

Molecular formulae are constructed using the standard operators '$\neg$' (not), '$\wedge$' (and), '$\vee$' (or), and the conditional operators '$\rightharpoonup$' (factual interpretation), '$\rightharpoondown$' (counterfactual interpretation), '$\Rightarrow$' (combined interpretation); and assigned truth-values according to the following recursive set of conditions:

$\nu_w(\neg A) = 1$, *iff* $\nu_w(A) = 0$, and 0 otherwise.

$\nu_w(A \wedge B) = 1$, *iff* $\nu_w(A) = \nu_w(B) = 1$, and 0 otherwise.

$\nu_w(A \vee B) = 1$, *iff* $\nu_w(A) = 1$ or $\nu_w(B) = 1$, and 0 otherwise.

$$\nu_w(A \rightharpoonup B) = \begin{cases} 1 \textit{ iff } \nu_w(A) = 1 \text{ and } \nu_w(B) = 1, \\ 0 \textit{ iff } \nu_w(A) = 1 \text{ and } \nu_w(B) = 0, \\ \text{and not truth-apt otherwise.} \end{cases}$$

$$\nu_w(A \rightharpoondown B) = \begin{cases} 1 \textit{ iff} \text{ for all } w' \text{ such that } {}_wR_{Aw'},\ \nu_{w'}(B) = 1, \\ 0 \textit{ iff} \text{ for some } w' \text{ such that } {}_wR_{Aw'},\ \nu_{w'}(B) = 0, \\ \text{and not truth-apt otherwise.} \end{cases}$$

$$\nu_w(A \Rightarrow B) = \begin{cases} 1 \textit{ iff } \nu_w(A \rightharpoonup B) = 1 \text{ or } \nu_w(A \rightharpoondown B) = 1, \\ \text{and } \nu_w(A \rightharpoonup B) \neq 0 \text{ and } \nu_w(A \rightharpoondown B) \neq 0, \\ 0 \textit{ iff } \nu_w(A \rightharpoonup B) = 0 \text{ or } \nu_w(A \rightharpoondown B) = 0, \\ \text{and } \nu_w(A \rightharpoonup B) \neq 1 \text{ and } \nu_w(A \rightharpoondown B) \neq 1, \\ \text{and not truth-apt otherwise.} \end{cases}$$

One additional constraint is needed. Let $f_A(w)$ be the set of worlds that are A-accessible from w, under the relation R_A, and let $[A]$ be the class of all worlds at which A is true, $\{w : \nu_w = 1\}$. The constraint, then, is:

$(C3\ 1)\ f_A(w) \subseteq [A]$

This ensures that all worlds A-accessible from w—that is, worlds that are *ceteris paribus* the same as w, except that A is true there—are worlds at which A is in fact true. Of course, A may in fact be true at w, rather than differing in this respect from other A-accessible worlds, and thus itself be a member of $[A]$.

This constraint, ($C3$ 1), is the first of two Priest places on the conditional logic C (which results in its extension C^+). The second, however, is not part of $C3$. This second constraint is as follows, (Priest, 2008: 87)

Second C^+ constraint: If $w \in [A]$, then $w \in f_A(w)$

This constraint effectively gives us weak centering; namely, that the actual or base world is at least as close/similar (or the exemplar of whatever form of ordering one prefers) to itself as any other possible world. (Lewis, 1973b: 97; Kment, 2014: 51)

Now, on my account, there are two ways in which centering could be construed:

(a) As necessitating that the actual world is one of the set of antecedent-accessible worlds of the counterfactual interpretation, if just the explicit antecedent is true there.

(b) As necessitating that the actual world is one of the set of antecedent-accessible worlds of the counterfactual interpretation, if the explicit antecedent *and* its implicit *ceteris paribus* clause are both true there.

One could, as in construal (a), take centering to amount to the actual world necessarily being one of the set of A-accessible worlds, if the explicit antecedent is true

there. But, as will become clear below, there is reason to resist this claim, since there are certain exceptional cases in which the explicit antecedent is true at the actual world, and so too is the consequent, though this latter due only to extraordinary coincidence—cases in which, at those possible worlds where all relevant things are equal the consequent is not true; so that these cases are ones in which the actual world is not a member of the set of A-accessible worlds.

Consider, as an example, the following conditional:

(51) If you finish your paper today, there will be a storm tomorrow.

Now, it might turn out that you do finish your paper today and there is a storm tomorrow. But, this is pure happenstance. Nothing about your finishing your paper has the least connection with or effect on the fact of their being a storm the next day. In such a case, the *ceteris paribus* clause, as here construed, would exclude the actual world—at which antecedent and consequent are both true due simply to coincidence—and confine the A-accessible worlds just to those at which odd coincidences do not occur. And, evidently, this use of the *ceteris paribus* clause cannot be compatible with the definition of centering in (a).

The alternative conception of centering, as in construal (b), requires the actual world to be such that the antecedent's implicit *ceteris paribus* clause is not violated, in addition to just the explicit antecedent being true there, in order for the actual world to be included as one of the A-accessible worlds on the counterfactual interpretation. On this approach, centering is effectively respected on the $C3$ analysis without need

of any additional constraint, since, as we will see, the factual interpretation assesses the conditional with regard solely to the actual world, and whenever the *ceteris paribus* clause as well as the explicit antecedent is true at the actual world it will automatically be one of the A-accessible worlds (hereafter, simply A-worlds).

Whichever way centering is defined on my account, makes no substantive difference: It may be taken to lack a weak centering, understood as (a); or it may be taken to respect weak centering, understood as (b). Of course, $C3$'s lack, or its merely apparent lack, of weak centering (depending on how it is defined) has important implications for the validity of *modus ponens* in this system—something that I take up in detail in Chapter 5[14]; suffice it to say here, though, that *modus ponens* is truth-preserving for all reasoning using factual interpretations of the conditional.

C3 interpretations

The logic of $C3$ provides for two possible interpretations—three, if one counts the double harpoon as an additional interpretation—for each instance of the ordinary-language conditional. Here A represents a given conditional's antecedent and C its consequent:

[14] *See p.223.*

Factual interpretation: $A \rightharpoonup C$

Counterfactual interpretation: $A \rightharpoondown C$

Combined interpretation: $A \Rightharpoonup C$

The factual interpretation, the harpoon-up, is one which interprets the conditional as pertaining solely to how things are at the actual world—determining its truth in virtue of nothing but actuality. The counterfactual interpretation, the harpoon-down, interprets the conditional in terms of how things might have been or may be, rather than necessarily how they actually are—determining its truth in virtue of what is the case at relevant possible worlds. These interpretations may assign different respective truth-values, but the truth of one does not preclude that of the other (nor the falsity of one the falsity of the other).

On the logic of $C3$, the conditional is deemed true on the factual interpretation when, at the actual world, both the antecedent and the consequent are true; false when, at the actual world, the antecedent is true and the consequent false; and non-truth-apt otherwise; that is, it ascribes no truth-value in those instances in which the antecedent is false at the actual world. On the counterfactual interpretation, the conditional is deemed false when the consequent is false at any one or more of those possible worlds at which the antecedent is, *ceteris paribus*, true (those worlds A-accessible from w, or A-worlds). When the consequent is true at all such worlds,

the conditional as a whole is deemed true on the counterfactual interpretation. In those case where the antecedent is false at all possible worlds, the counterfactual interpretation does not ascribe a truth-value to the conditional.[15]

The double harpoon, or combined interpretation, provides an easy way to refer to the truth-value of a given conditional *simpliciter*, without having to specify which, of the factual or counterfactual, is the interpretation at issue. It is useful when their respective truth-values coincide (or at least fail to clash), particularly when it may be irrelevant or indeterminate which is the more salient in a particular context. However, it has no intuitive truth-value when the factual and counterfactual interpretations differ or neither returns a truth-value, since there is little sense in the idea of a single combined interpretation when the individual interpretations conflict or neither is truth apt. This combined interpretation is thus offered as a useful shorthand, but I am not especially concerned to defend it if it seems unintuitive.

At its heart, the claim I make about the ordinary-language conditional by arguing for $C3$ as an explication of its logical form, is one as to what we are effectively tracking when we, as ordinary speakers, make intuitive truth-value judgments about instances of the conditional. In logical terms, we are simply tracking whether the consequent is also the case when the antecedent is. And, so, the conditional carries a Gricean implicature[16] that the truth of the consequent in some way follows from that of the antecedent.

[15]Some philosophers maintain that conditionals with impossible antecedents, necessarily false at all possible worlds, are nonsensical, not truth-apt, or simply false—Stalnaker, for instance, renders all such conditionals equivalent in truth-value, as true, at the single absurd world, λ (Stalnaker, 1968: 103), while others wish to claim that they are just as susceptible of cogent analysis as any other conditional. This question I address in Chapter 5.

[16]*See p.68.*

The nature of our ordinary uses for language are such that, typically, we have reason to assert only conditionals in which the consequent at least *may* stand in this sort of dependence relation (whatever it in fact is) to the antecedent—as opposed, say, to conditionals with quite unrelated antecedent and consequent. This being the case, the respective truth-values of a conditional's factual and counterfactual interpretations then tend to coincide with one another and with that of its implicature (the implicature being that such a relationship between antecedent and consequent holds). But, since the logic of each interpretation, sans implicature, only indirectly captures this relation of antecedent to consequent, there are various possible cases in which the truth-values of the two interpretations and the implicature clash with one another[17].

As explained below[18], a given conditional's consequent may also be true at the actual world when, inter alia, the antecedent is due simply to coincidence—rather than to any kind of dependence of one on the other. The *ceteris paribus* clause of the counterfactual interpretation rules out such coincidences, but there is nothing insulating the factual interpretation from extraordinary coincidences since this interpretation is concerned simply with how things in fact are at the actual world. Nonetheless, even the *ceteris paribus* clause of the counterfactual interpretation does not insulate it from all implicature-violating instances, as I will discuss with reference to necessarily true consequents in Chapter 5[19].

The following example will help to illustrate the possibility of the truth-value of

[17] *See discussion of Table 1, p.85.*
[18] *See discussion of (tv2a), p.90.*
[19] *See p.209.*

the factual interpretation diverging from that of the conditional's attendant implicature. Should I buy three plants today and claim that,

(52) If I buy three plants today, it will be windy tomorrow.

and, by sheer happenstance, it does indeed turn out to be windy tomorrow, then, on $C3$'s factual interpretation (52) is deemed true. But, since there is no counterfactually reliable relationship between the antecedent and consequent, the conditional is judged false on the counterfactual interpretation. The latter interpretation seems to accord perfectly well with our everyday intuitions, but not the former. The ordinary speaker is inclined to deem (52) entirely false—how then to explain its truth on a factual reading?

The conditional's generalized conversational implicature

The factual truth of (52)—in virtue of both its antecedent and consequent happening to be true at the actual world—is at odds with the intuitive truth-value judgment of this conditional because, not only is it false on a counterfactual interpretation, but its assertion violates a generalized conversational implicature (Grice, 1975) carried by the conditional. As Levinson charcterizes it,

> Grice's notion of a generalized conversational implicature (GCI) is an essential explanatory notion. It amounts to the claim that there is a special species of

> pragmatic mechanism that yields inferences that are both defeasible and default in character. [... T]hese inferences, generated under the mutual assumption of tacit coordination through specific heuristics, have the force of strong presumptions. They belong to a broad third category or layer of meaning, midway between sentence-meaning and speaker-meaning (or utterance-token-meaning), namely utterance-type meaning, in which types of linguistic form come to have preferred, idiomatic readings. (2000: 73)

What is implicated by the conditional is that the consequent in some sense depends on or follows from the antecedent. This implicature is violated when a speaker fails to respect the conversational maxim of relation by asserting a consequent that lacks the relevant relation of following from or being dependent on the antecedent in a certain way. Grice's formulation of the maxim of relation was that "I expect a partner's contribution to be appropriate to immediate needs at each stage of the transaction; if I am mixing ingredients for a cake, I do not expect to be handed a good book, or even an oven cloth (though this might be an appropriate contribution at a later stage)." (1975: 47). To offer a consequent to which the truth or falsity of the antecedent is essentially irrelevant is to make a contribution (the consequent) that fails to be relevant to what immediately preceded it (the antecedent). Of the precise nature of the consequent's dependence on the antecedent, of how exactly the consequent may be said to 'follow from' the antecedent, I am not attempting to give any sort of explication. Whatever the specifics of this relationship may be, I argue that it is one that is implicated by the conditional and not part of the logical form

of either possible interpretation (the factual or the counterfactual).

This is much as is the case with a term such as "but", which has the logical form of simply a conjunction, but additionally carries an implicature of the second conjunct being, roughly, surprising in light of or somehow in contrast to the first. The notable distinction here, though, is that the implicature carried by the conditional is cancelable, unlike that of the conventional implicature carried by the term "but". For instance, one might say the following:

> If I buy three plants today, it will be windy tomorrow
>
> In fact, if I run a half-marathon or fall down a hill today, it will be windy tomorrow.
>
> Really, nothing I do or don't do will have any effect on the fact that it's going to be windy tomorrow...

Essentially, its being windy tomorrow is wholly independent of my actions, and in making this explicit I cancel the implicature carried by (52). A conditional such as this may, therefore, be factually true—should it be the case both that I happen to buy three plants today and that the weather service also accurately predicts tomorrow's wind. And, if (52) is used in the sort of context just given, the implicature's

cancellation would remove the conflict between what would otherwise be its truth-value (that is, false), and the conditional's truth on the factual interpretation.

Motivation for dual interpretations of the conditional

But what real motivation is there to accept that the conditional is susceptible of these two different interpretations? rather than holding with the more orthodox view that there are two distinct types of conditional; that any given instance of the conditional is in itself inherently factual or counterfactual, possessed of a single truth-value just as a conditional *simpliciter*. To begin with, consider our ordinary intuitions as to the truth-values of the two following conditionals, as accounted for by $C3$—when it happens, at the actual world, that it does in fact rain during Boniswa's commute, that she indeed has no umbrella, and that she fails to win the lottery:

(53) If it rains during Boniswa's commute and she has no umbrella, she'll get wet.

(54) If it rains during Boniswa's commute and she has no umbrella, she'll win the lottery.

In the case of (53)—whose antecedent and consequent we are presuming true, as assessed at the actual world—we intuitively judge this conditional to be true. And (54), with its true antecedent but false consequent at the actual world, we intuitively

deem false. But what of cases where the antecedent is false at the actual world? Let us assume, for instance, that it in fact does not rain during Boniswa's commute. Then, there is no sense to be made of the conditional as interpreted factually—it's neither true nor false that she either gets wet or wins the lottery at the actual world when it rains during her commute, since the antecedent is *not* true; it does not rain. We can, nonetheless, make good sense of what would have been the case *had* the antecedent been true, and so the counterfactual interpretation is the salient one here. However, as I discuss further below[20], it would not do simply to discard the factual interpretation in such cases, since it is still open to a confused or truculent interlocutor to maintain that the antecedent of any given conditional is factually true and on this basis to make a claim as to the truth-value on a factual interpretation of the conditional as a whole. One may still give a factual interpretation—albeit an inaccurate one. Such a claim remains perfectly intelligible and capable of being made, however poor the grounds for it may be.

With respect to the counterfactual interpretation, (53) is true on my account, since at every world at which, *ceteris paribus*, the antecedent is true it seems plain that the consequent also is true. However, in the case of (54) the conditional is false on the counterfactual interpretation, as it is obviously false that, at every single world at which the antecedent is *ceteris paribus* true, Boniswa also wins the lottery.[21]

[20] *See p.74.*

[21] I discuss the question of the content and/or exclusions of the *ceteris paribus* clause below (*see p.98 on*), but it seems evident that the relevant implicit *ceteris paribus* clauses of (53) and (54), respectively, are quite straightforward—effectively barring things such as, in the case of (53), an eccentric wealthy benefactor suddenly deciding to hire a flunky to follow Boniswa around and hold an umbrella over her; or, in the case of (54), a bizarrely benevolent criminal rigging the lottery so that it is won by whomever is first to walk past without an umbrella, and this person happening to be Boniswa.

Our intuitions appear to accord well with the dual interpretations, at least in these two instances. But is this true of conditionals across the board?

The Oswald–Kennedy pair revisited

To better apprehend the rationale behind the claim that any given conditional may be interpreted either factually or counterfactually, it's instructive to look at the classic example used to motivate for specific instances of the conditional being inherently factual[22] or counterfactual—the Oswald/Kennedy pair:

(55) If Oswald didn't shoot Kennedy, then someone else did.

(56) If Oswald hadn't shot Kennedy, then someone else would have.

The first of these two conditionals, (55), seems to take it for granted that Kennedy was undoubtedly shot; thus, if the shooter wasn't actually Oswald, it must have been someone else. This appears to concern how things are (or are not) at the actual world. In contrast, (56) seems to concern how things might have been. It might have been the case that Oswald didn't shoot Kennedy (though in fact he did), but from this it would not necessarily follow that there was any sort of larger conspiracy such that another person would have carried out the assassination had Oswald failed to (hence

[22]This pair (which originated with Adams, 1970: 90) is more commonly described as comprising an indicative and a counterfactual conditional, but I have made clear the problems of this sort grammatically-based categorization, above—not least of which is that it creates an *ad hoc* dependence of the logic of the conditional on the grammar of a particular language.

our intuitive judgment of this conditional as false). It really does appear that (56) is essentially counterfactual and (55) straightforwardly factual. Why, then, prefer my account?

Let's start with (56). One may reject the truth of this conditional for the usual reasons, just mentioned—which, on my analysis, is explained by the counterfactual interpretation, in terms of it not being the case that, at all the possible worlds where, *ceteris paribus*, Oswald did not shoot Kennedy, Kennedy was still assassinated. For surely there are A-worlds at which the former was home sick that day and the latter safely served out the rest of his term as President? giving the antecedent/consequent the respective truth-values of true/false and making the conditional as a whole false on a counterfactual interpretation.

However, there are other reasons for taking (56) to be false. That these may rely on implausible beliefs and a convoluted conspiracy theory, makes them no less intelligible—though it does help to explain why this alternative has been overlooked.

One might respond to (56) that it is false because, in actual fact, *Kennedy never was assassinated*[23]. It was all, as it turns out, an elaborate hoax wherein he faked his own death because he feared he was being targeted by the CIA, after which he fled the US and lived out his days in exile in Liberia. Thus, Oswald didn't shoot Kennedy, but nor did someone else (since *no one* in fact did). The conspiracy-loving contrarian believes the conditional's antecedent, (57), but rejects its consequent, (58):

[23]This is apparently not too implausible a position to be genuinely held. Conspiracy theorists "True Democracy Party" have obligingly claimed that JFK was indeed not assassinated. Spoiler: He reemerged in the guise of Jimmy Carter, which claim is supported with photos of the two presidents' apparently identically-shaped ears (http://truedemocracyparty.net/2013/04/jfk-assassination-faked-staged-event-fraud-just-like-the-sandy-hook-hoax-boston-marathon-staged-event-fake-batman-shooting/).

(57) Oswald didn't shoot Kennedy.

(58) Someone else (other than Oswald) shot Kennedy.

and, on the basis of their belief in the truth of (57) and the falsity of (58) at the actual world, the Kennedy-assassination doubter claims that (56) is false. They take (56) to be false because of (as they suppose) the truth of its antecedent, (57), and falsity of its consequent, (58), as evaluated *at the actual world*. The contrarian very evidently takes (56) to be making a claim that may be challenged by (putative) facts about the actual circumstances.

To make the contrarian's response to (56) clearer, consider the following exchange,

Ying: If Oswald hadn't shot Kennedy, then someone else would have.

Robin: Actually, no... Funny story—a friend of my grandfather's was in the CIA at the time and she told him what *really* went down that day. You see, Oswald was all set to shoot Kennedy (he'd been recruited by the CIA) and the CIA had a sharp-shooter of their own ready to shoot too, in case Oswald didn't. Turns out their back-up plan was a good idea, because Oswald was far too timid to be a killer and he couldn't pull the trigger.

But, the most ridiculous thing happened: Unbeknownst to anyone, there was a little ground instability, just in a very localize area there. This tiny sinkhole opened up—right there behind the bush the back-up CIA sharp-shooter was using to hide himself—and it collapsed under him just a second before he was going to shoot the president!

That's not all... Kennedy had got wind that an attempt was going to be made on his life around that time, and he'd devised a plan to stage his own assassination! It was that fake shooting that everyone saw, further along the route from where Oswald and the CIA shooter were.

And, because of that extraordinary coincidence, the back-up sharp-shooter missed the only time he could have shot Kennedy, and the president wasn't shot before he could go through with his fake assassination! So your claim isn't true, since neither Oswald nor anyone else did actually shoot Kennedy.

Ying: That's utterly ridiculous. And, even if your wild story is true, it doesn't matter. The point is that there were back-up shooters ready to kill Kennedy if Oswald missed, regardless of that totally bizarre coincidence. Whether some completely random stuff happened to prevent the CIA sharp-shooter from taking the shot or not just isn't relevant to my point—that Oswald wasn't the only one set to assassinate Kennedy that day.

Robin: Look, the facts are that Kennedy wasn't shot by Oswald or any one else,

no matter how that came about. So it's just not true that if Oswald hadn't shot Kennedy, someone else would have. Oswald didn't shoot Kennedy, and nobody else did either.

Ying: Well, as I said, I disagree. And anyway, I'm sorry, but your grandfather must have been pulling your leg. The reality is that Oswald did actually shoot Kennedy. Even you admitted that he was in on the plot. I've analyzed over 15,000 documents and film recordings, and I can conclusively say that Kennedy *was* shot, that Oswald did it, but that, if he'd missed or got cold feet, a highly trained CIA sharp-shooter would have shot the president.

Robin: I can understand your reluctance to believe me, but let me show you my extensive collection of annotated photographs of the "assassination"...

In the context of the above exchange, that Robin's remarks are felicitous (and they quite evidently are) is explicable only if (56) may be interpreted factually as well as counterfactually. Whether one countenances the patently ludicrous claims made about why the two shooters failed to actually shoot and how Kennedy faked his own assassination, is irrelevant to Robin's being a factual interpretation of this conditional. The claim that Kennedy was not assassinated is very much the stuff of half-baked conspiracy theory, but this makes the potential factual interpretation of (56) no less genuinely *an interpretation*—*qua* interpretation—that may quite legitimately be made and easily understood in ordinary discourse.

Nor can the disagreement between Ying and Robin be construed as failing to distinguish whether the conditional (56) is genuinely a factual or a counterfactual one—whether it is inherently one or the other—rather than as a dispute over the truth-values of the different possible interpretations. How they differ is in terms of which of the interpretations they are more interested in. Robin is concerned with what did actually happen, which clearly makes (56) factually false. Ying, on the other hand, is concerned with what would have happened had the normal course of things not been so bizarrely derailed, which makes the conditional counterfactually true.

If one is inclined simply to rule (56) false, *simpliciter*, and reject the claim that there is any counterfactual interpretation available here—provided that what Robin claims is true—consider why this conditional is typically rejected. It is usually deemed false, on the orthodox view, because Oswald was a lone shooter and no one else was in on the assassination plan. This, clearly, is a counterfactual interpretation. Thus, the counterfactual interpretation, given Oswald was actually the (lone) shooter, is taken to be false because there was no other assassin and, absent Oswald, Kennedy would not have been shot. Notice that what makes people typically reject this interpretation as false is precisely what concerns Ying here. That she believes it's true that Oswald wasn't alone and a CIA sharp-shooter was there to back him up, whereas we ordinarily take that to be false, is irrelevant. The point is that it's the truth-value of the claim that there were multiple shooters that matters to the truth-value of the counterfactual interpretation, irrespective of what that truth-value may be.

This becomes still clearer when one thinks of a conditional such as (56), but uttered before the assassination, say by someone who knows that Oswald is committed to shooting Kennedy:

(59) If Oswald were not to shoot Kennedy, someone else would.

To respond to this,

> "But what if a tiny sinkhole opens up under the back-up shooter just before he's about to take the shot?"

would be ludicrous. The obvious sort of reply an ordinary speaker would be inclined to make is,

> "And what if aliens blow up the earth tonight? Come on... Of course, if some freak occurrence like that happens then the back-up shooter couldn't take the shot—but that's totally irrelevant. The point is that Oswald won't be alone; there'll be another shooter backing him up."

What helps to hide the fact of there being two different interpretations is that, colloquially, speakers don't (need to, or in fact) make explicit (nor are they likely

even consciously aware) that there are alternative interpretations of the same conditional; rather, they typically speak as if the conditional is true or false *simpliciter*; exactly what leads to Ying and Robin's disagreement. Which interpretation the speaker is making use of at a given moment is, instead, tacitly indicated by factors such as context and explanation on the part of the speaker, as well as the way that the conditional is phrased. For instance, in the case of (56), the counterfactual interpretation is evidently taken to be salient by the speaker, as is the case for Ying when she asserts this conditional—this is suggested in the antecedent by the use of the subjunctive tense shift to the pluperfect, from the temporally-relevant simple past (which marks when the events that this conditional concerns took place); and by the subjunctive term "would" in the antecedent (not that this use of the subjunctive mood in English can simply be identified with counterfactuality, as we have clearly seen). However, notwithstanding this making-salient of the counterfactual interpretation, the factual interpretation is still available. This is what Robin avails himself of in denying Ying's assertion of (56).

One might object that what is really at issue in the disagreement between Ying and Robin is merely some sort of presupposition or conventional implicature generated by (56), so that what Robin is in fact denying is simply that Oswald did shoot Kennedy, as seems to be presupposed by the wording of this conditional. Certainly, one could object to (56) on this basis, but that cannot be all that Ying and Robin disagree on here. Ying's belief in the truth of (56) is not dependent on whether or not Oswald did actually shoot Kennedy; she accepts the conditional either way, whether the antecedent is factually true or not. In light of this, the truth-value of

the antecedent at the actual world is essentially irrelevant for Ying, and thus also whether any presupposition or implicature regarding Oswald having been the shooter is satisfied.

The not-so-counterfactual conditional

The essential idea underlying their being two possible interpretations of a given conditional is something that has been touched on elsewhere, though not examined or adequately addressed. David Lewis's analysis of counterfactuals is one of the most broadly accepted, but his reliance on a strong centering condition (1973b: 14–15) means that not all of his supposed "counterfactual" conditionals are genuinely counterfactual, as I mentioned in the previous chapter. I consider Lewis's account in detail in my next chapter, but the salient points here are that his view is based on the primitive notion of relative similarity between and among worlds, and that it makes use of strong centering.

Strong centering is the intuitively compelling claim that, for any given base world w, the world most similar to it is w itself. Thus, in those cases in which the antecedent of a given conditional is true at w, the most relevantly similar world to w at which the antecedent is true is w. And, roughly speaking, that set of worlds (or single world) most similar to a particular base world and at which the antecedent of the relevant conditional is true, is the set in terms of which this conditional's truth-value, relative to the base world, is to be determined (it being deemed true when all of these worlds are ones at which the consequent also is true, and false otherwise).

However, the upshot of Lewis's approach is that supposedly (inherently) coun-

terfactual conditionals may turn out to be factual. Lewis acknowledges this and, in fact, observes that, "counterfactuals with true antecedents reduce to material conditionals" (Lewis, 1973b: 26).

Take, for instance, the putative counterfactual conditional,

(60) If Anatoli were to win the lottery next Friday, he would donate half his winnings to Médecins Sans Frontières.

Lewis asserts that the use of a counterfactual with a true antecedent (true at the actual world, that is) is instance of an error: "[T]he counterfactual constructions of English do carry some sort of presupposition that the antecedent is false. It is some sort of mistake to use them unless the speaker does take the antecedent to be false" (Lewis, 1973b: 3) (though Lewis does not maintain that this makes such instances automatically false). However, there are clear cases, such as (60), in which the conditional is expressed using the subjunctive and might well turn out to be only counterfactually salient (most likely, Anatoli will not win the lottery), but could also happen to be true of the actual world—and there appears to be no sort of error in uttering such a conditional, even if one has reason to think the antecedent could turn out to be true at the actual world and no reason to take it to be definitively false. The same is true of the other instances of conditionals that may be given in the subjunctive mood in English but which the speaker knows to have a true antecedent, or at least does not know to have a false antecedent. Regardless, with conditionals like (60), it's impossible to say at the present moment which of these eventualities

will ultimately be the case. Thus, a theory that motivates for any sort of inherently counterfactual conditional simply cannot classify such an instance of the conditional either as a counterfactual or not. And, should Anatoli defy the odds and actually win the lottery (and, let's say, also donate half his winnings to Médecin Sans Frontières), it would then be the case that an apparently counterfactual conditional is in actuality a factual one.

That said, it really makes no matter to the plausibility of a dual interpretation what the truth-value of the antecedent does turn out to be. Still, one might argue that—even though a conditional such as (60) cannot be determinately classified as either a counterfactual or not until the time with which the antecedent is concerned has passed and the truth-value of the antecedent at the actual world is known—this is simply an epistemic limitation of ours, and certainly not definitive as regards the nature of the conditional itself.

However, this potential objection is essentially irrelevant; even if the antecedent does in the end happen to be actually false, this doesn't preclude someone from, say, believing that it really is the case and asserting it as factually true. Regardless of the *actual* truth-value of the factual interpretation of, for instance, (60), if it is merely possible for someone to take it to be factually true and evaluate the conditional as a whole on this assumption, then it is clearly susceptible to factual interpretation. Whether the truth-value ascribed to the conditional on the basis of this interpretation is correct or not is beside the point. My account requires merely that both a factual and counterfactual interpretation of each conditional be possible, not that both, or even either, be true. Given this, the position that conditionals are inherently either

factual or counterfactual cannot be sustained.

To return to the Oswald/Kennedy pair: As with (56), a supposedly quintessential example of the counterfactual conditional, even what is taken to be a classic instance of the indicative conditional, (55), may in fact be interpreted both factually and counterfactually.

Albert: If Oswald didn't shoot Kennedy, then someone else did.

Ying: Well, yes—but Oswald *did* actually kill Kennedy.

Absent a counterfactual interpretation of (55) there is no reason for Ying to grant any truth to what Albert says, given that she believes that, at the actual world, Oswald did in fact shoot Kennedy. Yet Ying's response is perfectly intelligible and felicitous. What leads her to accept the truth of (55) is that it is plausible to her on a counterfactual interpretation—namely, at all those worlds at which, *ceteris paribus*, Oswald did not shoot Kennedy, it is true also that someone else did. This is because it is quite evidently part of what is taken to be included in the *ceteris paribus* clause implicit in the counterfactual interpretation that Kennedy was indeed shot.[24]

To make clearer the necessity of recourse to a counterfactual interpretation of (55) in order to make sense of Albert and Ying's brief exchange above, consider the

[24]Not that this claim is unassailable, of course—despite both Ying and Albert here taking it to be implicit in the *ceteris paribus* clause. Its factuality may be freely challenged (as may the legitimacy of any tacit inclusion or exclusion of a given *ceteris paribus* clause). But it is, nonetheless, clear that it's here taken as read.

	***A* at *w* (@)**	***B* at *w* (@)**	$\rightharpoonup$	***A* at all *w′*, wR_Aw'**	***B* at all *w′*, wR_Aw'**	$\rightarrow$	$\Rightarrow$	**Example**
tv1	T	T	T	T	T	T	T	a. If you put paper in the fire, it'll catch alight. b. If it's measles, she'll be covered in spots. c. If bears hibernate, then seven is a prime number.
tv2	T	T	T	T	F	F	F	a. If I throw this ball into the air, an eagle will swoop down and grab it. b. If Mother Theresa was born in Skopje, Macedonia, then Obama is the current US president.
tv3	(T)	T	-	F	T	-	-	N/A
tv4	(T)	T	-	F	F	-	-	N/A
tv5	T	F	F	T	T	T	-	If it's sunny tomorrow, Jeremy will take his nieces to the beach.
tv6	T	F	F	T	F	F	F	If I snap my fingers three times, I'll turn into a horse.
tv7	(T)	F	-	F	T	-	-	N/A
tv8	(T)	F	-	F	F	-	-	N/A
tv9	F	T	-	T	T	T	T	If you drop your pencil, its lead will shatter.
tv10	F	T	-	T	F	F	F	If I'm wearing green shoes today, I'll forget my wallet on the subway.
tv11	F	T	-	F	T	-	-	If you square the circle, the sun will come up tomorrow.
tv12	F	T	-	F	F	-	-	If you square the circle, you'll find $20 in the street tomorrow.
tv13	F	F	-	T	T	T	T	a. If it's measles, she'll be covered in spots. b. If it were measles, she would be covered in spots.
tv14	F	F	-	T	F	F	F	a. If the Allies had lost WWII, Germany would had to have been renamed 'Pluto'. b. If you were Julius Caesar, you would live at Stonehenge.
tv15	F	F	-	F	T	-	-	If you square the circle, you'd be world famous for that.
tv16	F	F	-	F	F	-	-	If you square the circle, you'll win every national lottery ever held.

Table 1: *Possible truth-value combinations on contingent factual and counterfactual interpretations of the antecedent and consequent*

following conditional:

(61) If Oswald didn't shoot Kennedy, then someone else was playing the bassoon at the time.

Assuming that Oswald did shoot Kennedy, and no one was playing the bassoon then, (55) and (61) would have precisely the same antecedent and consequent truth-values at the actual world. Yet, were Ying's qualified agreement to (55) given in answer to the utterance of (61) rather, it would be impossible to account for her response as being explained by her finding the counterfactual interpretation of (61) plausible. This demonstrates that the assessment of (55) as true cannot be, say, simply the product of the antecedent's falsity at the actual world—or else precisely the same assessment would be made of (61), which has not just an antecedent with the very same truth-value but in fact the exact same antecedent as (55).

What would make the acceptance of (61)'s truth, on a counterfactual interpretation, inexplicable is that there is no counterfactually plausible connection between this conditional's antecedent and its consequent; that is, it appears quite intuitive that, of all those worlds where, *ceteris paribus*, Oswald did not shoot Kennedy, the great majority are such that no bassoon playing whatsoever was involved. At any rate, there is no reason to think that *none* of these worlds was without bassoon playing at the time of the assassination, which is what would be required for (61) to be true on a counterfactual interpretation. Thus, (61) would be deemed intuitively counterfactually false by Ying, and other ordinary speakers. And, since (55) and

(61) share their factual interpretations—that is, both fail to be factually truth-apt, given the falsity of their common antecedent at the actual world—it is only in terms of their counterfactual interpretations that they could differ. Hence, since they do indeed differ (given the facts as Ying believes them to be), (55) must have a possible counterfactual interpretation, despite its being the posterchild for the indicative (or factual) conditional.

None of the above is to say that there aren't specific conditionals (such as the very two comprising the Oswald/Kennedy pair) for which one or other of the possible interpretations, factual or counterfactual, is likely to be far more salient than the other in most, if not all, contexts of everyday discourse, nor that one or both of the interpretations may not be entirely irrelevant for all ordinary intents and purposes. It is no part of my argument that every possible interpretation of any given conditional is plausible, useful, or even one likely ever to be made in the entire history of everyday discourse—only that both of these interpretations are at least theoretically available in each case.

Truth-value combinations and their possible interpretations

I'd like to take a look at examples of each of the possible truth-value combinations of factual and counterfactual interpretations of the antecedent and consequent—both straightforward and seemingly problematic instances—to make clearer just what my account proposes[25].

To be absolutely clear, many of the conditionals given here as examples are ones

[25] *See Table 1, p.85.*

that *could equally well have different possible truth-values* on factual or counterfactual interpretations. The given truth-values are simply stipulated here, for the sake of explication, and may be wholly contingent—these conditionals are not necessarily all bound to have the specific truth-values assigned to them here. Rather, these are used as exemplars of particular instances of the conditional that at least could have such truth-values on these interpretations.

The first set of stipulated truth-values (tv1) for the respective antecedent and consequent of each of the given conditionals, provide factual and counterfactual interpretations on both of which the relevant conditional is deemed true. Instances (tv1a) and (tv1b) are essentially self-explanatory. The assumption as to the context is, in the first case ("If you put paper in the fire, it'll catch alight."), that it is one in which you do in actuality put the paper into the fire and in which it does indeed catch alight. Thus, on the factual interpretation the conditional (tv1a), given its true antecedent and true consequent, is true—and this accords with our ordinary usage. Similarly, the assumption in the case of (tv1b) ("If it's measles, she'll be covered in spots.") is that it is measles and she is, in fact, covered in spots. And, likewise, it seems perfectly uncontentious that this makes the conditional true on the factual interpretation; that is, true with respect to how things actually are.

What of the counterfactual interpretations of (tv1a) and (tv1b)? Both of these, too, are true on my account. At all of those possible worlds that are, *ceteris paribus*, just like the actual world and at which it is true that you put paper in the fire (and, in this case, the actual world is one of these worlds) it is intuitively true also that

the paper catches alight.[26] Similarly, since measles is a disease characterized by the sufferer developing spots all over their body, it seems clearly intuitive that all of those worlds at which, *ceteris paribus*, it is measles—all those worlds A-accessible from w (A-worlds), which in this case includes the actual world—are worlds at which she has measles. Hence, (tv1b), too, is true on the counterfactual interpretation.

These conditionals (tv1a) and (tv1b) are instances of the sort that are sufficiently straightforwardly true on both the factual and counterfactual interpretations that they may be used in contexts in which they're effectively asserted as being jointly factually and counterfactually true, or there may simply be no fact of the matter as to which of the interpretations a given speaker holds to be more salient, if either. It may be that the determination of which is salient is decided simply by whether the individual in question does in fact have spots or not. My provision of the double harpoon connective "$\Rightarrow$" is intended to usefully capture both interpretations in these sorts of instances, but may be quite irrelevant or infelicitous in other cases.

The third example, (tv1c): "If bears hibernate, then seven is a prime number", is rather less straightforward than the previous two. This conditional is deemed true on my account, on the factual interpretation, as both its antecedent and consequent are true at the actual world—bears do hibernate and seven is indeed a prime number. The obvious trouble is that there is no apparent connection between bears' hibernation and seven's being prime. I give a more detailed response to this concern in Chapter 5[27], in discussing challenges to my account of the conditional, but I'll

[26] Of course, the plausibility of the *ceteris paribus* clause at work here is crucial to strength of my account of the counterfactual interpretation. To that end, in the next section I offer an explanation and defense of my notion of the *ceteris paribus* clause (*see p.98*).

[27] *See p.209.*

provide a brief outline here.

It's unsurprising that, whatever the logical system implicit in our ordinary use of the conditional, its truth-value determinations should outrun our practical, everyday uses for them. Furthermore, since it is through our use of conditionals in such ordinary discourse that our intuitions as to their truth-value are forged, it is likewise unremarkable that ordinary speakers should have conflicting or unclear intuitions for the sorts of conditionals that are likely never to serve any practical purpose.

The problem here is that the logic of $C3$—which tracks the relevant conditional relationship between antecedent and consequent quite sufficiently for the typical extent of ordinary discourse—comes apart from the conditional's implicature in such contrived cases as (tv1c). Conditionals like (tv1c) are essentially useless in everyday discourse, but the logic itself is indifferent to whether a particular conditional happens to be oddly artificial or a mainstay of everyday conversation. And, where the relevant connection between antecedent and consequent fails to be captured by the logic, the attendant implicature explains this divergence between theory and intuition. Conditionals such as (tv1c) violate the implicature they carry that the consequent somehow depends on or follows from the antecedent—just (as mentioned above and explored more fully in Chapter 4) as the connective "but" has the inherent logic of nothing more than conjunction, of "and", yet also carries the implicature that the second conjunct in some way contrasts in an unexpected way with the first conjunct.

In the case of (tv2a), it's remarkable that both antecedent and consequent are true, since eagles are not prone to grabbing balls whenever they're thrown into the

air. Certainly, tennis would be a much more difficult game to play were this the case. The utterer of this sentence is "technically" entitled to boast that what they claimed in uttering (tv2a) was true: that, if they threw their ball into the air, an eagle would swoop down and grab it. However, since the truth of the consequent at the actual world is a remarkable fluke, it is open to an interlocutor to dispute the conditional's veracity, arguing that this was merely an extraordinarily lucky coincidence: that, in the vast majority of those worlds including both the utterer and the ball, the ball just falls back down without any sign of an interested eagle. The potential for this kind of dispute is in effect a result of the fact that the parties to the discussion disagree over which of the two possible interpretations is the more reasonable (though such disagreement ordinarily occurs without any explicit recognition of the fact of their being two alternative interpretations).

That there are, indeed, two possible interpretations of (tv2a) available here can be confirmed by what I'll call the,"Well, yes, but..." test:

Radha: See, I was right—I said if I threw up that ball an eagle would swoop down and grab it!

Shareen: Well, yes, but it was just an *incredible* fluke that it happened right after you said that. There's no way you could have known that an eagle would suddenly appear.

Radha: Ok, yeah, I'm not really psychic or some kind of eagle-summoner, but you have to admit that it was a really cool coincidence! And, what I said *was* technically true...

Shareen's comments seem entirely felicitous in this context, and her use of "but", and its implicature that there's a contrast between the conjuncts here, makes sense only if there is some distinction between her agreement with Radha's claim and her comment that the eagle's grabbing the ball is an extraordinary coincidence. And, what allows for the simultaneous truth of these two seemingly mutually exclusive views, is the fact of their being two different interpretations possible of the same, one instance of the conditional.[28]

The conditional (tv2b) has the same antecedent/consequent truth-values as (tv2a), and so also is true on a factual interpretation, the antecedent and consequent both being true at the actual world. In ordinary discourse, though, we would be more likely to characterize (tv2b) as false, which would appear to be problematic for a theory that purports to be one of just such ordinary use of the conditional.

But, the clear distinction between (tv2a) and (tv2b) is that the latter, but not the former, violates the generalized conversational implicature carried by the conditional. The utterance of (tv2b) obviously violates the maxim of relevance, since Obama's being president of the United States has nothing to do with the fact of Mother Theresa having been born in Skopje. These are two (contingent) wholly independent truths—and so their combination as the antecedent and consequent, respectively, of

[28]Of course, the 'Well, yes, but...' test makes evident the dual interpretations open for each of the Oswald–Kennedy pair, too—and for each instance of the conditional.

a conditional implicates a relation between the two that fails to obtain. It is this violation of the implicature that accounts for the divergence of the formal truth-value of the factual interpretation from the intuitive truth-value ascription of the ordinary speaker in a case such as (tv2b).

Moving further down the table of possible truth-value combinations, we come to those combinations that it is impossible even to construct, as they would require that the antecedent be both true at the actual world and simultaneously false at all possible worlds. Since the actual world is, of course, a possible world itself, this combination can never obtain. This rules out their being conditionals with the truth-value combinations given by (tv3), (tv4), (tv7), and (tv8).

The ruling out of these combinations should not be taken to suggest that any account of the implicit logic underlying usage of the conditional may be unprincipledly *ad hoc*, or arbitrarily adjusted where it might happen not to fit the evidence. Rather, whatever logical system is proposed as an explication of the ordinary-language conditional must be beholden to the metaphysical reality within which this real-world phenomenon is located. Simply because there is a line in the truth-table that the theory is able to generate does not mean that reality—let alone everyday language use—must (or even can) necessarily oblige. There is no reason to suppose—and every reason not to expect—that everyday use of the conditional would exhaust its underlying logic in terms of antecedent/consequent truth-value combinations. And, unlike the logic of a metaphysically untethered system, the nature and application of $C3$ must be constrained by the nature of reality and the limitations of human capabilities and interests.

A conditional such as (tv5) may have the truth-values given in *Table 1*[29]—antecedent/consequent truth-values at the actual world of true/false, and at all A-worlds of true/true. It is an instance in which the truth of the consequent is reliably counterfactually dependent on that of the antecedent, but also one in which the consequent is rendered false at the actual world due to some remarkable, fluke occurrence. An example of actual world circumstances that might make it true that it is sunny tomorrow, but simultaneously false that Jeremy takes his nieces to the beach, despite his fully intending to—are ones in which tomorrow does of course turn out to be sunny, but, say, all sufficiently nearby beaches are covered in toxic sludge from an overnight chemical spill. This is just the sort of extraordinary, unforeseen eventuality that is ruled out by the implicit *ceteris paribus* of the *counterfactual* interpretation.

On a factual interpretation, then, (tv5) is false, since although it is sunny Jeremy fails to take his nieces to the beach, making the antecedent true but the consequent false. And, indeed, one might lament that, though it's no fault of his, it's nonetheless a pity that it didn't in fact turn out to be true that he took them to the beach, even though the weather was actually sunny. However, we also presumably wish to account for the intuition that there was nothing dishonest in what he originally said, in asserting (tv5). Jeremy, let's assume, is a very honest and reliable individual, who had every intention of taking his nieces to the beach should the weather be fine. And, indeed, this is captured by the counterfactual interpretation of (tv5). At all the A-worlds, where the antecedent is *ceteris paribus* true, it seems intuitively clear

[29] *See p.85.*

that the consequent too is true. Though the bare antecedent is true at the actual world, in this case the actual world is not one of the A-worlds, as the *ceteris paribus* clause is made false by the exceptional circumstance of the surprise toxic spill that rendered the beach unusable. Thus, since the consequent is true at all A-worlds, the conditional as a whole is true on the counterfactual construal.

The conditional (tv6) is like (tv5) on a factual interpretation, but differs in the truth-value of the consequent on the counterfactual interpretation; that is, the consequent is not true at every A-world. As a result, this conditional is false both on the factual and the counterfactual interpretations, which accords very obviously with our ordinary intuitions. Needless to say, if I snap my fingers three times, it is true neither at the actual world—nor any other possible world at which the antecedent is *ceteris paribus* true—that I, or anyone else, turn into a horse.

Conditionals with the truth-value combinations of (tv11), (tv12), (tv15), (tv16) are ones with antecedents that are false not only at the actual world but also all possible worlds. On $C3$, these fail to be truth-apt on each respective interpretation. As I have argued above, this is because the way that the conditional is used in ordinary discourse is with regards to what does or does not follow from a given antecedent being the case. When that antecedent is not the case at the actual world, there is no sense to be made of claims as to what follows from its being true at the actual world. When the antecedent is false at all possible worlds, there is no sense to be made of claims as to what follows from the possibility of its being true; that is, true at at least one possible world. However, there is an argument to be made that conditionals with counter-logical antecedents may nevertheless be evaluated relative

to non-normal, impossible worlds; something I take up in Chapter 5. There are certainly instances in ordinary discourse in which we make use of conditionals with impossible antecedents.

The examples (tv9) and (tv10), and (tv13) and (tv14), are just the same as (tv1) and (tv2) and (tv5) and (tv6), respectively, as regards their counterfactual interpretations—the consequent of the first of each pair being true at all the respective A-worlds for each; and the consequent of the second of each pair failing to be true at one or more members of each respective set of A-worlds. Thus, at all of the A-worlds for (tv9), those at which *ceteris paribus* you drop your pencil, its lead shatters, and the conditional is therefore true on a counterfactual interpretation; whereas, it is not the case that I forget my wallet on the subway at every one of (tv10)'s A-worlds, those at which *ceteris paribus* I am wearing green shoes today, and the conditional is therefore false on a counterfactual interpretation. And, the same is the case, *mutatis mutandis*, for (tv13) and (tv14).

As for the factual interpretations of these conditionals—the stipulated truth-value combinations for (tv9), (tv10), (tv13) and (tv14) make the antecedent of each false; thus, none of these conditionals is truth-apt on a factual interpretation.

To help allay any doubt as to the non-truth-aptness of the conditional, on a factual interpretation, when its antecedent is false at the actual world, let's consider (tv14b) specifically in more detail,

(tv14b) If you were Julius Caesar, you would live at Stonehenge.

Now, interpreted factually, the claim being made by (tv14b) is that it follows from your being Julius Caesar that you live at Stonehenge. Putting aside the stipulated counterfactual truth-value of the consequent, I find myself hard-pressed to say just where I would live if I personally were Julius Caesar, though I can come up with various places, and accompanying explanations for them, for various different contexts in which (tv14b) might be put to me. But I have no conception of what the truth-value could be of the assertion that it follows from the fact that I *am*, at the actual world, Julius Caesar, that I live at Stonehenge. There is no sense in which it is either true or false that you have lived, currently live, or ever will live at Stonehenge because you are Julius Caesar. And, of course, that you certainly would, or wouldn't, live there at some possible world at which you are Julius Caesar, is irrelevant, since the factual interpretation concerns only what is the case at the actual world.

Frankly, it is hard to make sense of the idea that I even could possibly be me, myself, and simultaneously Julius Caesar. I can imagine being like Caesar in particular respects and I could likely work out if certain other things would also be true, or follow from this—but (tv14b) asks us to try to conceive of the bizarre case of one person also being another, which is, at the very least, patently factually false. And, so, this provides an extreme, and very clear, example of why there is no sense to be made from asking what follows from the truth of something that is not the case; that is, what the factual truth-value of this conditional is. It is simply not truth-apt.

However, one might think the antecedent of (tv14b) is so unintelligible as to be uninstructive with regard to ordinary cases. Let's consider a more prosaic instance, then:

(tv10) If I'm wearing green shoes today, I'll forget my wallet on the subway.

But I'm *not* wearing green shoes today. So what sense can I make of the claim that it follows from my wearing green shoes, at the actual world—when it is false at the actual world that I am wearing green shoes—that I will forget my wallet on the subway, also at the actual world? Such a case is very evidently not truth-apt when the antecedent is false.

The Ceteris Paribus Clause

That there is a *ceteris paribus* clause implicit in our everyday use of the conditional seems well motivated by our ordinary truth-value intuitions, which except conditionals from having to explicitly anticipate even the most unforeseeable and bizarre chance occurrences. For, whatever the background assumptions or other shared presuppositions among speakers necessary for communication, these cannot all be directly built into the semantics of the conditional, and the *ceteris paribus* clause allows us to take account of this. Clearly there is a *ceteris paribus* clause implicit in the counterfactual interpretation of the conditional, as I discuss just below—but this occasions the question of how such a clause is to be construed, and whether it can avoid being irredeemably *ad hoc*.

To better see that the counterfactual interpretation of the conditional must include an implicit *ceteris paribus* clause, consider again the example of Jeremy, who

utters the following to his nieces,

(62) If it's sunny tomorrow, I'll take you to the beach.

With an unqualified strict conditional, instances such as (62) are vulnerable to the challenge that it is not true in every possible case of the weather's being good tomorrow that we will go to the beach. Indeed, the weather may be perfectly fine, but if there's a toxic waste spill along the coast during the night, or we're tragically killed in a house fire that evening, then we will most definitely not be going to the beach tomorrow, regardless of how nice the weather is. This would, of course, make (62) false on the factual interpretation, since that is evaluated in terms of what is in fact the case at the actual world. Thus, one of the nieces might say: "If only what Jeremy said [that is, (62)] had been true; it's terrible that we're not going to the beach even though it's such a beautiful day. I can't believe the whole coast is covered in toxic waste!" It would be unreasonable, however, to suggest that—insofar as his original utterance constituted a commitment to take his nieces to the beach the next day should the weather be good—Jeremy could be said to be failing to honor this. Of course, it's not in any way his fault that the beach is contaminated with toxic waste, and so it seems odd to claim that Jeremy was speaking falsely when he uttered (62) or that he has broken his promise to his nieces. Yet, if we assess the counterfactual reading of (62) as purely a strict conditional, we cannot avoid coming to the same conclusion as on the factual interpretation, since the actual world would have to be

included as one of the A-worlds, and it would then be the case that the consequent fails to be true at every A-world.

However, while it is true that as a matter of fact about the actual world that (62) turned out to be false, it seems nonetheless quite correct to hold that, understood counterfactually, (62) is true. This is what licenses the natural inclination an ordinary speaker might have to admit that (62) was false "in a sense" (that is, on a factual interpretation), since it *was* sunny and Jeremy didn't actually take his nieces to the beach—but also the fact that this passes the earlier mentioned, "Well, yes, but..." test. In other words, it makes perfect sense to say, "Well, yes, but no one could have known that a freak accident would cover the beach in toxic waste." Why is that "but" appropriate here? Because the acknowledgment of the (factual) truth of (62) is in contrast to the fact that the toxic spill was a remarkable, unpredictable coincidence.[30]

One might further point out that the canceled outing to the beach was in no way Jeremy's fault, because he certainly *would have* taken his nieces to the beach, *had* the weather been good *and had there been no extraordinary, unforeseeable occurrences*

[30]Of course, one may use "Well, yes, but..." illegitimately—or, at least, in a way not indicative of another interpretation of the conditional. Imagine a scenario without any toxic spill in which Jeremy promises to take his nieces to the beach if it's sunny, but simply fails to do so. Still, someone might say, "Well, yes, but Jeremy really does mean well, you know." What distinguishes this from a case in which this is legitimately used and the test really is passed, is whether or not there is a disparity between the truth-value of the consequent at the actual world that the conjunct following the 'but' gives rise to, and what we would expect the truth-value of the consequent at the actual world to be. For instance, the freak toxic waste spill gives rise to the consequent of (62) being false at the actual world, which is inconsistent with the expectation (given Jeremy's reliable character) of its truth. However, if Jeremy is really just an unreliable person who didn't bother to keep his promise, then he could mean well without actually being at all dependable (so that the consequent is false), which is entirely consistent with the expectation, in this scenario, that he doesn't take his nieces to the beach (namely, that the consequent is false).

to prevent it.[31]

The implicit *ceteris paribus* clause ensures that, in such a case as (62), the claim being made *vis à vis* going to the beach tomorrow is that it is true—interpreted counterfactually—only in those circumstances in which the weather is good, *provided that all other things are equal.* Determining what exactly a given *ceteris paribus* clause is to be taken to encompass is by no means straightforward, but it seems uncontentious to assert that the unexpected spilling of toxic waste, or the fatal combustion of our house, should not be counted as being in the ordinary run of things. And, given that all of the special sciences rely on there being a robust answer to the question of just how to specify the *ceteris paribus* clause[32], its justification is by no means a problem unique to any theory of conditionals. Indeed, the notion would seem to be on at least as sound a footing as the special sciences themselves.

The problem of the specification of the *ceteris paribus* clause may, in fact, be taken to be just a special case of the frame problem more generally; that is, the problem confronted in the attempted development of artificial intelligence, that it appears one would have to provide an infinite list of what is and what is not relevant to any given judgment before an assessment can be made. Somehow, animals of all kind, ourselves included, ignore and discard irrelevant information and focus on what

[31]Should there be any lingering worry that a conditional such as (62) doesn't appear to have any possible counterfactual import (given it's categorically indicative form), one can just imagine a niece who complains to her parents that Uncle Jeremy broke his promise of (62), being told that she's being unfair, since, obviously, he would have taken her to the beach if he could have. Evidently, a counterfactual interpretation of (62) is perfectly possible.

[32]To take just two examples of its use in the special sciences, there is the 'law of definite proportions' in chemistry: "Any chemical compound consists of elements in unvarying proportions by mass, *ceteris paribus*." (Lange, 2002: 408); and from molecular biology, "[*C*]*eteris paribus*, the helices based on a strong alphabet [of RNA base pairs] are more stable than those on a weaker." (Kolchanov et al., 1996: 187)

is contextually relevant to their particular goals at a particular time and place, at least well enough that most individuals and groups successfully survive day-to-day life. Yet giving a systematic analysis of what exactly we do to achieve this, and how exactly we do it, has proven remarkably intractable. As Dennett observes apropos this difficulty of specifying just what an artificial intelligence should ignore and what they should pay attention to:

> The beauty of the a ceteris paribus clause in a bit of reasoning is that one does not have to say exactly what it means. "What do you mean 'other things being equal'? Exactly which arrangements of which other things count as being equal?" If one had to answer such a question, invoking the ceteris paribus clause would be pointless, for it is precisely in order to evade that task that one uses it. If one could answer that question, one wouldn't need to invoke the clause in the first place. One way of viewing the frame problem, then, is as the attempt to get a computer to avail itself of this distinctively human style of mental operation." (Dennett, 1984: 198)

Nonetheless, the manner in which the *ceteris paribus* clause inherent in the $C3$ conditional is cashed out is clearly central to my view. The *ceteris paribus* clause is what must do the heavy lifting in delimiting the set of antecedent-worlds at which a given conditional should be assessed—a role performed in other possible-world accounts by such means as orderings of worlds (for instance, in terms of similarity or

closeness)[33]. That the *ceteris paribus* clause is better able to successfully identify the relevant worlds, without the difficulties associated with other possible-world theories of the conditional, is a result of its responsiveness to the staggering variety of particular, and often unique, everyday contexts in which the ordinary-language conditional is actually used.

The type of *ceteris paribus* clause at work in the $C3$ conditional is one that serves to exclude such aberrances as would prevent all things, over and above the truth of the given conditional's antecedent, being equal; this exclusion taking the form of ruling certain possible worlds inadmissible for inclusion in the set of worlds according to which the conditional is assessed. Of course, this may just as well be characterized in terms of the *ceteris paribus* clause ruling which are to be included in the set of A-accessible worlds.

There are two ways one may go with a *ceteris paribus* clause—opting for one definite or indefinite; that is, one whose exclusions are (or could at least theoretically be) fully specified, or one whose exclusions remain open-ended. The main problem with the former option is that it is very difficult to nail down precisely, and with the latter that it risks being question begging or arbitrary. (Reutlinger et al., 2011) It seems highly unlikely that the *ceteris paribus* clause implicit in every respective instance of the conditional could be fully specified; at least for all practical intents and purposes, and surely not in all cases. And—even if it is in principle possible to detail the *ceteris paribus* clause's exclusions in full—as it is used implicitly in the conditionals that form part of our day-to-day language, it is effectively open-ended.

[33] *See p.165 on.*

At first blush this would seem a significant weakness of the *ceteris paribus* clause, and consequently of any view relying on it. It could be objected that the *ceteris paribus* clause of the $C3$ conditional is simply trivially true, amounting to nothing more than the assertion that any world that would make the relevant conditional false is to be excluded—or arbitrary, as one may seemingly make exclusions at will if there's no determinate way in which to fix the *ceteris paribus* clause's content.

Indeed, the *ceteris paribus* clause I wish to make use of is messy. And, I would argue, is inherently and ineliminably so. Its use might therefore seem like the introduction of a huge, illegitimate loophole, through which any and all recalcitrant instances may sneak off without giving proper account of themselves—a loophole that ensures the right truth-value may be assigned to the counterfactual interpretation of any given conditional, come what may; rendering my account of counterfactual interpretations vacuous.

As Earman observes, regarding difficulties generated for law-statements by their inclusion of a *ceteris paribus* clause,

> It seems that there could be no informative account of the truth-conditions of CP law-statements that did not render them vacuous. One way to see the problem is to note that we could specify the conditions under which such a statement is true if and only if we could specify the conditions under which it is false, but that is exactly what we cannot do with a CP law-statement. For such a statement will be violated exactly when the regularity contained in it is violated *and* "other things are equal", i.e. there is no "interference". But we cannot specify the conditions under which the second conjunct obtains;

> otherwise the CP clause is simply an eliminable abbreviation and what we have is not a genuine CP law-statement. (Earman et al., 2002: 292, emphasis in original)

While I'm characterizing not laws but ordinary-language instances of the conditional, this objection is nonetheless also relevant to my account of counterfactual interpretations. If the contents of the *ceteris paribus* clause used in the counterfactual interpretation of any given conditional cannot be systematically and determinately specified, then the concern is that the *ceteris paribus* clause can be made to exclude all and only those eventualities that would result in the falsity of the consequent; in other words, be used to select as A-accessible worlds only those at which the consequent is true. However, as I will show, what is taken to be included in the *ceteris paribus* clause is not simply stipulable by the utterer of a conditional, but something constrained by the norms and expectations of the language community in which it is used, and by the nature of reality itself.

It is important, for those sympathetic to Lewis's account of the (putative) counterfactual conditional, to note that the *ceteris paribus* clause is certainly no more mysterious nor obscure in its meaning than Lewis's notion of comparative similarity. The comparative similarity ordering of possible worlds (Lewis, 1973b, 1981) used to specify the set of worlds relative to which a given conditional should be assessed (like Stalnaker's (1968) concept of closeness of worlds), is one just as difficult to precisify as any *ceteris paribus* clause. As Lewis characterizes it, comparative similarity is "vague—very vague—in a well-understood way. Therefore it is just the sort of

primitive that we must use to give a correct analysis of something that is itself undeniably vague" (Lewis, 1973b: 91), that being what he termed the counterfactual conditional. Though I agree with Lewis's observation on the conditional's vagueness, of course I deny that comparative similarity is the concept in terms of which it may best be analyzed. The need for a great degree of flexibility of interpretation is evident in the problems encountered by analyses of the conditional that do not make use of a *ceteris paribus* clause. Mechanisms such as comparative similarity or closeness are indeed vague, but nonetheless too inflexible to effectively deal with the required degree of variation in the myriad different contexts of use of the conditional in everyday discourse.

However, the flexibility of the *ceteris paribus* clause need not render it vacuous. For, as I make clear below, what we may take to be allowed or excluded[34] by the *ceteris paribus* clause implicit in the counterfactual interpretation of any given conditional is very significantly constrained by what the utterer can reasonably take the hearer to understand to be ruled out by the clause, and, reciprocally, what the hearer may reasonably assume the speaker to have intended in using this conditional, and so on.

And, as it happens, we are highly adept at managing open-ended exclusions of the kind captured by the *ceteris paribus* clause, without lapsing into question begging or arbitrariness—however uncatalogable, and indeed perhaps infinite, the potential exclusions or inclusions may be. This is something we do almost every waking moment. For no-one walks around each day desperately trying, but failing,

[34]I characterize *ceteris paribus* clauses in terms of what they include or what they exclude, interchangeably. Nothing substantive hangs on which is used.

to calculate what they must hold fixed in order to negotiate daily life: that certain materials are solid enough to walk on; that particular substances may be safely eaten; that the sky does not need to be carefully watched less it fall on you unexpectedly; that staplers and paper tend to retain their shape when you move them about, but water and sand do not. And we are even sufficiently good at determining what others hold fixed to be able to interact largely successfully with them; typically, one does not wonder whether to hide the office stapler lest a perfectly phlegmatic colleague suddenly develop a fear that "all things being equal" fails to rule out the possibility they might be killed with it.

As Sven Ove Hansson notes,

> When discussing with my wife what table to buy for our living room, I said: 'A round table is better than a square one.' By this I did not mean that irrespectively of their other properties, any round table is better than any square-shaped table. Rather, I meant that any round table is better (for our living room) than any square table that does not differ significantly in its other characteristics, such as height, sort of wood, finishing, price, etc. This is preference *ceteris paribus* or "everything else being equal". *Most of the preferences that we express or act upon seem to be of this type.* (Hansson, 1996: 307, emphasis mine)

To take a fuller example; in August of 2011, New York experienced a minor earthquake. Being in an area not at all prone to noticeable seismic activity, this was

distinctly out of the normal run of things. Managers in various multistory New York buildings responded by instructing occupants to gather in their building's basement. Whether or not this was really the best available response, it was certainly taken to be at the time. But notice that it was also a response solely to the earthquake. Only one thing was adjusted—the fact of their having occurred an earthquake—but *all else was held fixed*. All things aside from the quake and its relevant consequences were considered "equal". Though there was a, quite possibly infinite, number of additional, unusual eventualities that could have been taken into account—all but a tiny fraction of these were completely ignored in determining how to respond. The difference from the ordinary run of things was taken to consist only in the quake and those particular things relevantly altered by it, as perceived by the various building managers. So, people were instructed to head to their building's basement, even though, were all other things not held to be fixed—were the possibility, say, of a coincidentally simultaneous flood considered—going to a basement that lay below sea level would be obviously ruled out. However, the expectation of the regular safe operation of these building's elevators was not held fixed; this was judged to be part of what counted as being altered by the quake and the elevators were shut down.

What was to be taken as changed—that there was currently earthquake activity, instead of the area's normal seismic stability; that the elevators could not run normally with an acceptable likelihood of safety; that heavy, unaffixed objects should be avoided; and so forth—was very likely not exactly and completely specifiable, nor even clearly determinate. Yet, in deciding on a response, numerous building managers were effectively able, on the fly, to pick out enough of what constituted

all things being equal in the highly unusual case of a light minor earthquake. It is precisely this skill that we use every moment, to determine what will change certain things about our environment but not (a potentially infinite and unspecifiable host of) others.

None of which is to say that we are any of us immune to error in our everyday determinations of what to take as being permitted or ruled out by a given counterfactually-interpreted conditional's implicit *ceteris paribus* clause. We are, in fact, constrained in our construal of any given *ceteris paribus* clause in everyday discourse by our own beliefs and intentions, those of others, and the brute metaphysical reality of the world. As Stephen Neale makes clear,

> Among the things that constrain the formation of A's saying-intentions are A's knowledge of the meanings of the words he is using and his (tacit) knowledge of the syntax of the language he is using. Thus A cannot (intend to) say that snow is white by uttering the sentence 'grass is green'. [...] More generally, he cannot (intend to) say that p by uttering X if he believes it is impossible for his audience B (or at least any rational, reasonably well-informed interpreter in B's shoes) to construe him as intending to say that p. (Neale, 2004: 77)

What Neale is pointing out is that we cannot—just as Humpty Dumpty instructively tried and failed to (Carroll, 1872: 123–124)—unilaterally stipulate meaning by intending to say anything we like with whatever words we prefer. No more can we intend to convey whatever we wish to (let alone actually do so), by means of

the implicit *ceteris paribus* clause of the counterfactual interpretation. It includes and excludes what the relevant community of language-users take it to include or exclude, in the given context. And, however it may be psychologically realized, ordinary speakers clearly do manage to effectively distinguish between what constitutes all things being equal and what is an unrelated, fluke occurrence. Certainly, effective language use is not a one-person show.

When reality and/or those around us disagree with our determinations and we fail to communicate as we wished, we make adjustments and test them on the world until we get as close to our aim as we're able. Similarly, when we use conditionals in ordinary discourse, others may dispute what we implicitly take to be excluded by or included in a particular *ceteris paribus* clause, so that we're forced to defend our assessment or concede the point. The following conditional,

(63) If Stalin had not joined forces with the Allies, the Axis powers would have won World War I.

might be taken to include a relatively straightforward *ceteris paribus* clause in everyday discussion, and so to be straightforwardly true on a counterfactual interpretation; but its *ceteris paribus* clause could be hotly contested by a group of World War I scholars, say, for whom the truth-value of (63) might be a highly debatable and complex question.

What we identify as aberrant and what we take to be constitutive of 'all things being equal', relative to the counterfactual interpretation of a specific conditional

used in a specific context, depends, as does all our behavior, on our specific knowledge, beliefs, and desires in that context. Consider the following:

(64) If the temperature of a body of gas increases, then its volume will increase proportionately.

What possible worlds someone takes to be ruled out by the *ceteris paribus* clause in assessing the truth-value of (64) may differ significantly depending on context; or, more specifically, one's knowledge, beliefs, and desires in this context. This is not in any way to say that all such variations are legitimate bases for determining the contents of the *ceteris paribus* clause. Instances of the conditional with identical explicit lexical form may have legitimately different truth-values, respectively, in different contexts—if, say, one context makes salient something that, as a result, must be included in the implicit *ceteris paribus* of that instance, while this fails to be the case in the other instance, which therefore would have a different *ceteris paribus* clause. On the other hand, it may be that a speaker or their audience is simply ignorant of some vital fact, and so misidentifies what the *ceteris paribus* clause should include, which may result in their being mistaken in the truth-value they ascribe to the conditional in question.

So, for instance, if Eva heard (64) uttered by a fellow high school student, with whom she'd just been discussing the role of pressure relative to the temperature and volume of a gas, respectively; there's every likelihood she would implicitly take the *ceteris paribus* clause to exclude those worlds at which the pressure of the body of

gas in question was varied. In this particular context, 'all things being equal' would include pressure invariance, and so (64) would be assessed as true on a counterfactual interpretation.

In contrast, we can imagine a situation in which students are discussing what the behavior of gasses would be like in a universe in which temperature and volume bore no relation to one another. It would then be perfectly legitimate for part of what is included in the *ceteris paribus* clause to be that a gas's volume does not vary in accordance with its temperature, and thus to rightly judge (64) to be false in the context of this discussion.

Consider another context, though—say one in which a student, Alfie, hears an utterance of (64) by a classmate, in a study group that hasn't yet properly covered the role of pressure in relation to gas volume and temperature, and whose members are currently under the impression that pressure can vary independently. Here, it's most likely that the implicit *ceteris paribus* clause employed by Alfie in determining the conditional's truth-value would *not* exclude those worlds at which the pressure of the gas in question was varied. In this discourse context, it seems that pressure variations wouldn't be taken to violate the requirement of all things being equal, so that those worlds at which these occurred would be included in the set of worlds at which (64) is assessed, making the conditional as a whole false on a counterfactual interpretation. Of course, Alfie wouldn't know that his construal of the implicit *ceteris paribus* clause he'd employed in counterfactually interpreting (64) was erroneously over-inclusive enough to mistakenly render (64) counterfactually false, but this is a purely epistemic limitation. On learning of the importance of pressure here, it would become evident

to him why his initial assessment of the conditional had been mistaken; that is (though he would very likely not think of it in these terms himself), that he had been mistaken in what he took the *ceteris paribus* clause to allow. Importantly, what distinguishes this case from that just above, in which context (64) is *rightly* deemed false (on the counterfactual interpretation), is that in the former context the truth-value of (64) is being assessed as regards a hypothetical universe in which gas temperature and volume vary independently, whereas in the latter (Alfie's) case this truth-value judgment is being made with reference to the actual laws of physics.

What this shows, then, is that the difficulty (if not the impossibility) of perfectly specifying the *ceteris paribus* clause implicit in any given conditional, is a product of precisely what allows for the conditional's great usefulness in ordinary life. Detailing how we are able to fill in *ceteris paribus* clauses is exactly the question of how we're able to assert of an entrance with a lock on it capable of thwarting opportunistic, would-be human burglars that,

(65) This entrance is secured against unwanted ingress.

though there's nothing to prevent a tiger clawing through the wood, a flamethrower incinerating the whole door, nor an earthquake from dislodging the frame from the surrounding wall. And it is this same aptitude for contextually sorting the pertinent from the irrelevant that allows us to determine that such an assertion as (65) would be true of the door to a modest apartment in a relatively safe town, but false of

the door to a collection of priceless paintings in a dangerous city riddled with highly organized crime syndicates.

All of which is not to provide any real explanation of our ability to discriminate in this way, but rather to show that we very evidently possess such a skill, allowing us to determine what counts as all things being equal—not infallibly, but reliably enough to serve in a given context, and sufficiently constrained by it to avoid arbitrariness. How exactly we perform this feat is far from clear. This is a substantial and complex problem—one to which I pretend to offer no solution, save to observe that to doubt that we, as ordinary natural-language speakers, are capable of adequately construing the inclusions and exclusions of the counterfactual interpretation's implicit *ceteris paribus* clause is to doubt that we are able to negotiate almost any aspect of everyday life.

Chapter 3: **Alternative indicative conditionals**

Now that we have a better understanding of the $C3$ analysis, we can examine what motivation there is to prefer this account over one (or more) of the extant theories of the conditional. I consider these alternative theories in two broad groupings: In this chapter, those that attempt to explicate the logical structure of the putative indicative conditional, and, in Chapter 4, those that concern themselves primarily with what they term the subjunctive or counterfactual conditional (or with a possible-worlds analysis of the conditional as a single type). Of course, in Chapter 1[35], I argued that these are neither accurate nor even coherent categories into which to divide the conditional—but, since this division is assumed by and basic to the great majority of these analyses, I follow it in my organization of the theories themselves. For ease of discussion I will not always refer to it as the "putative" or "supposed" in-

[35] *See p.24.*

dicative conditional (nor, likewise, with the subjunctive or counterfactual conditional in this and the following chapter), but, of course, my unqualified use of the term here is not to endorse this as a legitimate type of conditional in any substantive sense, and certainly not at the level of logical form. (Naturally, the indicative conditional is nonetheless a perfectly acceptable grammatical type, in English and a number of other languages—just not a categorisation that is relevant to the logical structure of the ordinary-language conditional as used across all of human natural language.)

The Suppositional Approach

The first of the alternative accounts of the conditional that I'll consider, confines itself to the indicative conditional.[36] Some theorists have rejected the idea that conditionals have truth conditions at all; that is, they deny that it is the truth-value of the antecedent and consequent respectively that, in whatever way, determine the truth-value of the conditional as a whole, and indeed that the conditional itself has a truth-value. In actual fact, though, the great bulk of the arguments given against truth-conditional analyses of the conditional are leveled against a truth-functional account, that of the material conditional, rather than one such as my own.

Ramsey proposed a reading of the conditional from which stemmed the suppositional view—one according to which one accepts or believes a given conditional, rather than judging it to be true or false, depending on the likelihood of the consequent given the antecedent.

[36]Though some attempts have been made to extend the view to counterfactuals (see Barnett 2010), the key focus of the approach is on the indicative.

The central claim Ramsey makes is that, "If two people are arguing 'If p, will q' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge, and arguing on that basis about q; ... they are fixing their degrees of belief in q given p" (Ramsey, 1990: 247). Subsequent theorists—notably, Adams (1965, 1966, 1975), Mackie (1973), Gärdenfors (1986, 1988), and Edgington (1995, 2008)—draw on this characterization in their respective accounts of the conditional.

The suppositional account has developed into the view, in essence, that the conditional cannot be assessed as if it is a single proposition whose truth-value is constituted by the combination of the truth-values of its parts. Rather, one should accept or believe a given conditional to the degree of probability one ascribes to the consequent, given one's supposition of the antecedent. That is, you suppose the antecedent to be the case, and judge how likely the consequent is on that supposition. David Lewis's (1976) proof of the triviality theorem shows that the probability of a particular consequent, supposing the relevant antecedent, is not equivalent to the probability of the truth of any one, single proposition. The upshot of this is that, if the suppositional approach is correct, the conditional as a whole cannot have truth conditions.

The Ramsey Test

What has come to be referred to as the Ramsey Test, applicable to so-called indicative conditionals, may be characterized as follows: "One should believe a conditional, 'if A then B' if one would come to believe B if one were to add A to one's stock of beliefs." (Read and Edgington, 1995) As Ramsey observes,

> [T]he belief on which the man acts is that if he eats the cake he will be ill, taken according to our above account as a material implication. We cannot contradict this proposition either before or after the event, for it is true provided the man doesn't eat the cake, and before the event we have no reason to think he will eat it, and after the event we know he hasn't. Since he thinks nothing false, why do we dispute with him or condemn him?[1] Before the event we do differ from him in a quite clear way: it is not that he believes p, we $\bar{p}$; but he has a different degree of belief in q given p from ours; and we can obviously try to convert him to our view. But after the event we both know that he did not eat the cake and that he was not ill; the difference between us is that he thinks that if he had eaten it he would have been ill, whereas we think he would not. But this is *prima facie* not a difference of degrees of belief in any proposition, for we both agree as to all the facts.
>
> [1]If two people are arguing 'If p, then q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q; so that in a sense 'If p, q' and 'If p, $\bar{q}$' are contradictories. We can say that they are fixing their degree of belief in q given p. If p turns out false, these degrees of belief are rendered *void*. If either party believes *not* p for certain, the question ceases to mean anything to him except as a question about what follows from certain laws or hypotheses. (Ramsey, 1990: 154–155, emphasis in original)

It is primarily from this footnote just quoted that various suppositional analyses of the conditional (as well as a certain truth-conditional accounts; for instance, Stalnaker, 1968) have been developed. The essential idea of the Ramsey Test is that one imagines one believes p (the antecedent), or assigns to it a probability of 1, and then, given one's hypothetical belief in p, one assesses the degree of credence one ascribes to q (the consequent), or the probability one assigns to it.

Do conditionals have truth-values?

Edgington is one of the primary proponents of the suppositional approach to indicative conditionals, arguing that conditionals do not have truth-values and are not "fact-stating" (Edgington, nd: 12). David Lewis's triviality proofs show that there is no single proposition, the probability of which is equal to the probability of the consequent of a given conditional, taking the antecedent to be true. On the basis of this Edgington argues that, "Believing that if A [the antecedent], C [the consequent] is not taking the attitude of belief to a proposition, $A * C$; it is believing that C, under the supposition that A." (nd: 13)

Part of the reason for Edgington's rejection of the idea that conditionals are susceptible of truth conditions appears to be pessimism as to the plausibility of the available accounts of this kind. She notes the "problematic and controversial" (Edgington, 2003: 385) state of conditionals' truth conditions, and from this concludes that we ought rather to focus on the analysis of conditionals that are not certainly either true or false. In asserting a conditional one is not, on the suppositional approach, making an assertion that some or other proposition is the case. You are

instead making the claim that the given conditional's consequent is the case, but only on the assumption that the antecedent is the case; that is, it is the assertion of the truth of the consequent, conditional on the truth of the antecedent, and not the assertion of the truth of the conditional itself.

I would argue that this denial of Edgington's, of the possibility of truth conditions for the conditional, arises from the failure to accurately taxonomize the conditional. There is evident confusion between the two possible interpretations available to a single instance of the conditional, as evinced by the following:

> Indeed, it is compatible with T3 [the suppositional approach] to say that 'If A, C? is true if A and C are both true, false if A is true and C is false, and neither true nor false if A is false. We need to add some explanation. On this assignment of truth values, belief that if A, C is not belief that it is true. For it is true only if A is true, and I may believe a conditional without believing its antecedent is true. For instance, I believe that if I touch the wire I will get a shock. But I don't believe that I will touch the wire. Symmetrically, to disbelieve a conditional is not necessarily to believe that it is false, for it is false only if A is true, and I might disbelieve 'If the Tories win they will nationalise the banks' without believing that the Tories will win. We say instead that to believe a conditional is to believe that it is true, on the supposition that it has a truth value—to believe that it is true, on the supposition that it is true or false. This is just our suppositional theory, T3, restated. (Edgington, 2003: 387)

What Edgington describes here is in part very close to my own account of the logical of the factual interpretation of the conditional. To use her terminology, on $C3$ a conditional 'If A, C' is deemed true on the factual interpretation if both A and C are true; false if A is true but C is false, and neither true nor false—namely, not truth-apt—if A is false. Where our views diverge is at the point of her conflation of the factual and counterfactual interpretations. Take the conditional she gives as an example,

(66) If I touch the wire, I will get a shock.

Edgington's objection to allowing that (66)) has truth conditions is that one cannot judge it to be true *simpliciter*—for she may believe it to be true that she will indeed get a shock if she touches the wire, but without believing that she *will* actually touch the wire. Thus, on her account this is an indicative conditional that can only be true *on condition* that the antecedent itself is true (being neither true nor false otherwise). My approach is in agreement that (66)) may be true even were I never to touch the wire, but on the $C3$ account this is analyzed as (66)) being not factually truth-apt but still true on a counterfactual interpretation. (And, this analysis allows for the possibility that I may in fact be foolish enough to actually touch the wire, in which case the conditional would be both factually true and counterfactually true.)

No longer confronted with a puzzling conflict between the conditional being potentially both not truth-apt and true, and able to attribute each judgment to its

relevant corresponding interpretation, we have no need to reject the possibility of truth conditions for the conditional.

In fact, Edgington herself observes the possibility of dual interpretations, though she does not put it in those terms (nor, of course, draw the same conclusion that I do as to how this is to be explicated):

> [W]e are arguing about whether, if you eat this apple, you will be ill. You throw it away in disgust. Our argument continues unabated—about whether you would have been ill if you had eaten it. We do not appear to have changed the topic of debate. Just before throwing it away you say, 'If I were to eat it, ...'; someone who left our company earlier says later on 'I'm convinced that if he ate the apple, he was ill.' The bipartite approach needs some explanation." (Edgington, 2008: 128)

Once clear on the different possible interpretations of the same conditional, we have just such an explanation of Edgington's apple example. The conditional, "If you eat this apple, you will be ill", is liable to interpretation both as concerning how things actually are, which would be a relevant interpretation were you in fact to eat it; but also as concerning how things might be, which is the more relevant interpretation after you throw away the apple and, presumably, preclude the possibility that the antecedent is true at the actual world. (Though, for the final speaker Edgington mentions in the above quotation, a factual interpretation is still a live possibility—as is, in effect, that the A-worlds to be considered on the counterfactual interpretation

may include the actual world; notwithstanding that, in actual fact but unbeknownst to them, this conditional has turned out not to be truth-apt on a factual interpretation.)

Similar concerns to Edgington's seem to motivate Ernest Adams's earlier suppositional account, which likewise rejects a truth-conditional analysis of the conditional. Adams asks whether the indicative conditional,

(67) If John does not arrive on the 10 o'clock plane, then he will arrive on the 11 o'clock plane. (1965: 169)

may be false when the antecedent is false—that is, when John does actually arrive on the 10 o'clock plane, and the negation of (67)'s antecedent is thereby true. In trying to answer this question we come to realize, according to Adams, that there is no clear answer to be had:

> [T]he term 'true' has no clear sense as applied to conditionals, particularly to those whose antecedents prove to be false [...]. This is not to say that conditional statements with false antecedents are not sometimes called 'true' and sometimes 'false', but that there are no clear criteria for applications of those terms in such cases. (1965: 169)

But, again, once we accurately distinguish between the two available (factual and counterfactual) interpretations of each conditional, the way is clear to accepting that a conditional may not be truth-apt on the factual interpretation when its antecedent is false at the actual world, but that the very same instance of the conditional may nonetheless have a truth-value on the counterfactual interpretation. And it's this counterfactual truth-value that accounts for the fact that we may still make use of a conditional such as (67) in ordinary discourse, even when the antecedent is in fact false. Realising that the statement of this conditional in the indicative mood in English does not preclude a counterfactual interpretation lets us handle such instances without abandoning truth conditions for the conditional.

To take Adams example, stated above—on my $C3$ account (67) would indeed fail to be truth-apt on a factual interpretation if it is true that, at the actual world, John does arrive on the 10 o'clock plane, which would make the antecedent of (67) false. However, a counterfactual interpretation—one taking (67) as being concerned with what would be the case were John not to have arrived on the 10 o'clock plane—is still perfectly truth-apt. It may be counterfactually true or false, depending on whether it is true at all of those A-worlds, ones at which John does not arrive on the 10 o'clock plane, that he does arrive on the 11 o'clock plane.

This analysis further explains the concern Adams raises that, on receipt of a telegram from John saying that, "I will arrive on the 10 o'clock plane. If I don't arrive on that plane, I will arrive on the 2 o'clock plane" (1975: 171), we are in a position to reject the antecedent of (67) and also the conditional as a whole—despite its apparent lack of a clear truth-value when the antecedent is false. This is

readily explicable on the $C3$ account; as just discussed above: The likely falsity of the antecedent at the actual world makes the conditional as a whole likely non-truth-apt on a factual reading, but John's assertion, that if he doesn't arrive on the 10 o'clock plane he will arrive on the 2 o'clock one, gives us good reason to deem (67) true on a counterfactual reading. It's possible, therefore, to effectively address the suppositionalist's worries about allowing the conditional truth conditions, without having to abandon the idea that conditionals do indeed have truth-values.

$C3$-independent reasons to think that conditionals have truth-values

Aside from the fact that the $C3$ account of the conditional provides a way for us to have our truth-value cake and eat the evident non-truth-apt examples as well, there are strong independent reasons not to give up truth conditions for conditionals. The suppositionalist's denial that we can properly ascribe truth-values to the conditional gives rise to serious objections. Most obviously, it clashes significantly with ordinary-language intuitions. Say one believes that,

(68) If Usain Bolt breaks one of his legs, he won't be able to run 100m in under 10 seconds while it is healing.

it seems intuitively obvious that it's perfectly acceptable to claim of the believer that they take this conditional to be true.

And, perhaps more significantly, the suppositional view fails to provide a means of assessing compound conditionals—where further conditionals are nested within the

main conditional—and compounds that include conditionals as a part. Consider, for instance, the following conditional,

(69) If Usain Bolt breaks one of his legs, then, if it's still healing, he won't be able to run 100m in under 10 seconds.

And these examples from the Corpus of Contemporary American English (Davies, 2008)),

(70) [I]f people delete comments then if other people are trying to look for that song that someone else suggested, they can't find it unless they ask someone else for it. ('New TVB Themesongs [Downloads and Goodies] — K for TVB', http://k-tvb.net/download2/)

(71) [I]f Song leaves arsenal then [if they] sign Sneijder [...] then we could play this formation[:] 4-2-3-1. ('Do you think Arsenal really need any more signings?', http://justarsenal.com/do-you-think-arsenal-really-need-any-more-signings/15265)

(72) If talks fail [i]f the summit fails, it would fall to the Government to make a fresh attempt to broker a deal when it takes over the six-month rotating presidency of the EU in January. ('Obama vows to press Burma for

further reforms during historic Asian visit', http://www.irishtimes.com/newspaper/world/2012/1119/1224326785043.html, November 19, 2012)

(73) If [P]auls think if obama taxes the rich more that they will receive this money [they're] sadly mistaken[.] ('Nearly Half of All Americans Don't Pay Income Taxes', http://blog.heritage.org/2012/02/19/chart-of-the-week-nearly-half-of-all-americans-dont-pay-income-taxes/, February 19, 2012)

The compound conditionals (69), (70), and (71) are of the form, "If A, then (if B, then C)"; the latter two, (72) and (73), of the form, "If (if A, then B), then C." Now, according to the suppositional approach, one should accept the conditional to the degree that, on the supposition that the antecedent holds, it is probable that the consequent is the case. But, here, that amounts to the probability, supposing A, that "if B, then C".

The clear problem is that what constitutes the consequent of the main conditional is itself a conditional. And, since Lewis's triviality theorem shows that the probability of a conditional is not equivalent to the probability of a particular consequent, supposing the relevant antecedent, the suppositional account is at a loss as to how to assess conditionals such as this. The suppositionalist might appeal to Jackson's suggestion that "If A, then if B, then C", may equally well be given as, "If A and B, then C" (which holds true also for conditionals of the form, "If A, if B, then C."). However, as Priest (2009) observes, this stratagem will not work for embedded

conditionals of the form, "If, if A then B, then C". As in the example,

(74) If, if you flip the switch, then the light comes on, then the power must have been reconnected.

Similar difficulties face this view when it comes to other compounds containing conditionals. And, as Lycan asks, why would just the indicative conditional,

> though 'assertible' under specifiable conditions, be barred from truth and falsity, especially when sentences of that type admittedly have probabilities, interact in truth-seeking discourse with uncontroversially truth-valued sentences, are agreed on or disagreed over, figure in valid argument, etc.?'

And, what is yet more strange, is to "talk, as Adams and Gibbard do, of a sentence's 'probability' when that term does not mean probability of *being true*." (Lycan, 2001: 74; original emphasis) In fairness, the suppositionalist may wish to respond that it is a probability of truth being considered here, but not that of the conditional sentence as a whole. Rather, the sentence is assertable when the probability (of the truth) of the *consequent*, given the antecedent, is sufficiently high.

However, this potential answer to his objection fails to deal with other of Lycan's (2001) concerns regarding the denial of truth-values to the indicative conditional.

For instance, this would seem to block such apparent equivalences as

(75) No incomplete forms will be processed by this office.

and,

(76) If a form is incomplete, it will not be processed by this office.

Of course, there are those who reject the commonly held logical equivalence of (75) and (76), but it would be difficult to deny that—certainly at least in many contexts that arise in ordinary discourse—the two may be used interchangeably. What the suppositionalist is then confronted with, is the difficulty of explaining how it is that (75) does have a truth-value, but the logically equivalent (76) does not.

As a practical example, imagine that the office in question handles purchase orders for a large university. Tired of having to follow up on forms with incomplete information, they put up a sign with (75) on it. And, let's say that they were highly annoyed at having to do that extra work and so stick ridgedly to this new policy. Even the suppositionalist would accept that the sentence on the sign had a truth-value—in this case, true. But now suppose that someone got a jacket button caught behind the sign as they brushed past and ripped part of it off the wall. They apologised and offered to replace it. The sign they print out and put up in its place has just (76) on it. Now, it seems wildly implausible that anyone could reasonably

object either that the the essential meaning of the sign had been changed (though one might, perhaps, think the new version sounded less forceful, say) or that the sign, which had previously had a perfectly clear and definite truth-value was now somehow truth-value-less.

Above all, the denial of truth conditions to the conditional flies in the face of the everyday use of ordinary-language conditionals. Consider just a few examples from the Corpus of Contemporary American English (Davies, 2008):

> The World Bank is far less defensible than the IMF. The bank was formed on the false assumption that *if rich countries lent money to poor countries, the poor countries would become rich and poverty would largely disappear.* In fact, rich countries became rich because they created the institutions and policies that allowed profit-seeking individuals to create goods and services for their fellow man. (*American Spectator*, 40(7): 20, September 2007; emphasis added)

The italicized conditional in this passage is quite evidently immediately preceded by the claim that it is *false*—"a false assumption". Moreover, the sentence immediately following the conditional, which is evidently intended to offer some explanation of the denial of the truth of the conditional, itself makes reference not just to the content of the consequent (should the suppositionalist wish to claim that it is essentially just the consequent, albeit given the antecedent, that is being deemed false here), but explicitly that of the antecedent as well—asserting that rich countries became wealthy due to particular institutions and policies, rather than by means of being

lent money, as is presented in the antecedent of the conditional as the path to the desired outcome of the consequent, that of the wealth and eradication of poverty in poor countries.

> COOPER: And true or false? *If a cough lasts longer than a week, go see a doctor?* GUPTA: You don't need to see your doctor, no matter how nice he might be, after a week. Three weeks, usually. COOPER: Really? After a week you don't? I was sure that was true. (spoken, Anderson Cooper, CNN news channel, January 10, 2011; emphasis added)

In this instance, Anderson Cooper asks whether it is true or false that one should see a doctor if a cough lasts longer than a week, and is told that this is unnecessary—namely, that this is *false*.

> But I mean, it is also true *if you kill this mission now, the James Webb, there won't be a future flagship mission that NASA will have.* (spoken, 'Funding for James Webb Space Telescope in Jeopardy', NPR Science, July 15, 2011; emphasis added)

Here, the claim being made is that it is *true* that, if the James Webb (Space Telescope) mission doesn't go ahead, then NASA will not have a future flagship mission.

> TRUE OR FALSE? *If an antiaging cream stings or tingles, it's normal.* TRUE. A little stinging or redness is totally okay and even pretty common! (*Cosmopolitan*, 245(3): 133, September 2008; emphasis added)

And, finally, *Cosmopolitan* asserts that it is true that, "If antiaging cream stings or tingles, it's normal."

What all of these examples serve to show is that, in ordinary speech (and writing), we quite obviously treat conditionals as having truth-values.

The Pragmatic Account

There is clear reason, then, to prefer an account of the logic of the conditional that does ascribe a truth-value to it. One such kind of approach is that of taking the indicative conditional to be identical with the material conditional of classical logic—which is true for all truth-value combinations of antecedent and consequent except that of a true antecedent and false consequent. The problems that then arise, where the material conditional appears to be strikingly at odds with our ordinary truth-value intuitions, are explained in terms of pragmatics.

The kind of difficulties that bedevil the material conditional, as applied to ordinary-language indicatives, are exemplified by the following instances, which the material conditional makes valid,

(77) It's cloudy today.

Therefore, if it's not cloudy today, then I'm a polar bear.

In any case, such as this one, in which the antecedent is false, the material conditional deems the conditional as a whole true—but it seems intuitively false that the conditional that here forms the argument's conclusion follows from the premise. It appears, at best, decidedly odd.

(78) There are no bananas left.

Therefore, if the earth revolves around the sun, then there are no bananas left, and if horses are made of rice, then there are no bananas left.

Similarly, in any instance of the conditional in which the consequent is true, the truth conditions of the material conditional ensure that the conditional as a whole must be true. But, again, the conditionals that form the conjunctive conclusion of (78) both seem intuitively false, since there is no evident connection between the earth's revolution around the sun and whether or not there are any bananas left, and it's even more puzzling how the fact of the dearth of bananas could possibly follow from the false claim that horses are made of rice.

The material conditional licenses further intuitively illegitimate inferences of the following sorts:

(79) If Roscoe is a dog, then he has a tail, or, if Roscoe is a human, then he has opposable thumbs.

Therefore, if Roscoe is a dog, then he has opposable thumbs, or, if Roscoe is a human, then he has a tail.

(80) If you strike this match, it will light.

Therefore, if you strike this match and you're under water at the time, it will light.

Moreover, a material conditional of the form, "If A, then $\neg A$", is true when A is false. Thus, if the truth conditions of the indicative are those of the material conditional, then this next instance must be true when A is false, despite its seeming intuitively false:

(81) If today's Saturday, then today's not Saturday.

Given these problems with the material conditional, what accounts for the various defenses of this analysis of the indicative?

One of the most appealing features of the material conditional is that it accords straightforwardly with such intuitively persuasive inferences as the following: Imagine that I borrowed a book from a friend, Abdul, and that he's arranged for one of his sisters—either Juwairiyyah or Shamila—to collect it from me. He's not sure

which of them will be coming to get the book, but he does know that one of them definitely will. Later, a woman greets me, and says that she's Abdul's sister and she's here to the collect his book. She's clearly in a hurry so I just give her the book and we don't get a chance to introduce ourselves. Then, when another friend who was with me asks which of Abdul's sisters that was, I am in a position to say that,

(82) That was either Juwairiyyah or Shamila.

And it seems intuitively that I could equally well respond that,

(83) If that wasn't Juwairiyyah, then it was Shamila.

The or-to-if inference appears clearly justified in this scenario. Since there are only two people it could have been, and I know it must be one of them, it can only be either Juwairiyyah or Shamila; and, so, if it wasn't Juwairiyyah, then it must be Shamila. The truth-tables for (82) and (83) are identical; both coming out true in all cases except that in which the person who collected the book from me was neither Juwairiyyah nor Shamila.

This is exactly what the material conditional predicts, and it is its intuitive fit with such cases that accounts for much of its appeal. But, of course, there are other cases in which things do not go nearly as smoothly for this analysis of the indicative.

Grice's defense of the material conditional as the indicative

Grice (1975, 1989) argues that pragmatics is able to account for seemingly odd or problematic cases generated by taking the indicative conditional to be simply the material conditional. The source of the trouble, according to Grice, is that we fail to distinguish between the truth-value of what is literally said and the assertability, in a given conversation context, of the conditional, given what is (in addition to its purely literal meaning) implicated by it. (Adler, 2008: 11–12)

The view that Grice defends is that, 'in standard cases to say 'if p then q' is to be conventionally committed to (to assert or imply in virtue of the meaning 'if') both the proposition that p ⊃ q and the Indirectness Condition"; this latter being "'[T]hat p would, in the circumstances, be a good reason for q,' 'that q is inferable from p,' 'that there are non-truth-functional grounds for accepting p ⊃ q"'. (Grice, 1989: 58) Grice takes the last of these to be the most compelling, though it would seem that, with regard to his examples of the cancelability of this as an implicature carried by the conditional, the other formulations would be more convincing.

In support of the claim that this Indirectness Condition is a conversational implicature, Grice offers a number of examples of its cancelability, a key feature of this type of implicature, the most persuasive of which are the following; firstly,

> To say "If Smith is in the library, he is working" would normally carry the implication of the Indirectness Condition; but I might say (opting out) "I know just where Smith is and what he is doing, but all I will tell you is that if he is in the library he is working." No one would be surprised if it turned

> out that my basis for saying this was that I had just looked in the library and found him working.

In other words, my reason for asserting the conditional just given is that I know Smith to be working the library, and therefore my utterance of this conditional does not implicate that the antecedent (Smith's being in the library) offers good reason for the consequent (Smith's working); rather, it is my direct knowledge of where he is and what he is doing that provides good reason for the conditional as a whole. (It will be evident, now, why the formulation of the Indirectness Conditional that Grice himself favors does not as readily support his own example here, since there may still be said to be non-truth-functional grounds for deeming the conditional true—namely, my knowing that he was in the library working.)

A further instance of cancelability that Grice provides is,

> There are now some very artificial bridge conventions. My system contains a bid of five no trumps, which is announced to one's opponents on inquiry as meaning "If I have a red king, I also have a black king." It seems clear to me that this conditional is unobjectionable and intelligible, carries no implicature of the Indirectness Condition, and is fact truth-functional.[37] (Grice, 1989: 60)

[37] As the casualty myself of a bridge partner's mistaking a purely conventional bid for a more natural indication of the contents of my hand—resulting in a disastrous championship defeat and followed by the utterance of belated reminders of the convention in question, in the form of conditionals of this kind—I can attest to plausibility of Grice's example.

Here, again, the implicature is absent, insofar as my having a black king does not follow from my having a red king, and nor is the latter good reason for the former.

The conversational implicature that Grice maintains is ordinarily carried by the conditional arises, roughly put, from the tacit, mutual conversational agreement that one should always assert the strongest claim that one is in a position to. More technically, it is violation of the conversational maxim of quantity—that one should be as informative as possible (though not more so than is necessary)—to assert a conditional when one is in a position to deny the truth of its antecedent or assert the truth of consequent, though other of Grice's maxims may also play a role in the relevant implicature generated by a given conditional, as is evident in the example below.

To clarify Grice's analysis of the indicative conditional, let's take an example (paraphrasing one of Lewis's, 1976): Imagine that we're walking through the woods, and that I point to some mushrooms growing next to a tree and say, "If you eat those mushrooms, you will die." Let's say that, not wishing to die, you unsurprisingly do not try the fungi. Taking this instance to be a material conditional—then, since the antecedent is false (you do not eat the mushrooms) the conditional as a whole is true. However, in the imagined scenario, I recognized that the mushrooms were a very rare and valuable, and quite edible, species and I in fact wanted to stop you from eating them so I could come back later to collect them. In this case, I have not strictly told a lie. Knowing that my assertion of this conditional would very likely lead you to avoid the mushrooms, what I said could be argued to be true. But, very obviously, it was also terribly misleading.

A benefit of Grice's account is that it explains the problematic nature of this sort of usage. His contention is that the truth-value of the indicative conditional is that of the material conditional—here, true—but in this instance I not only say something true, I also convey that the mushrooms are deadly poisonous. This is explained in terms of the pragmatic expectations we have of one another when we converse. There are, among others, the shared expectation that one will not assert things for which one lacks evidence or that one knows to be false. Here, I very clearly lack any evidence that the mushrooms I point out are poisonous; in fact, I know that they are not. Thus, in conveying (though it may not be part of the explicit meaning of my words), or implicating, that the mushrooms are poisonous, I am violating the pragmatic principles that we as ordinary speakers take to govern our conversations; namely, in this case, the supermaxim of quality (that one should try to make one's conversational contributions true), since I am implicating that the mushrooms are poisonous, which I know to be false.[38] So, we have an explanation of why, although this conditional is technically true, it is not appropriate to assert it in the scenario given.

At the heart of Grice's defense of the material conditional is the insight that there is an important distinction between what one may truthfully literally say, and what is legitimate to assert in a particular context, given what one literally says will

[38]The maxim of truthfulness (not contributing that which I know to be false or misleading) Grice (1989) took to apply to what was explicitly, literally said and not also to what was simply implicated, and so would not be violated in this case, since I needn't be saying anything *literally* false in uttering this conditional about the mushrooms. However, the implicature that the mushrooms are poisonous *is* false, and I am thus violating the supermaxim of quality (under which the maxim of truthfulness falls), since this applies to the whole of a speaker's conversational contribution, including what is implicated. (Wilson and Sperber, 2002)

implicitly implicate in such a context. (Edgington, 2014) It is this that allows Grice's theory to account for a number of otherwise troubling results of attributing the truth conditions of the material conditional to that of the ordinary-language indicative.

One of the difficulties with the material conditional is that it licenses the inference from the truth of Q to the truth of $P \supset Q$. So, for instance, if I know that Q,

Q: My friend's dog's name is Daisy.

I may infer the truth of any conditional having Q as its consequent, regardless of what P is. From the truth of Q it follows that, "If today is Friday, then my friend's dog's name is Daisy"; "If everyone on Earth is made of balloons, then my friend's dog's name is Daisy"; and so on, for any given P.

Grice's account offers us a way out of this, since, though any given conditional with Q as its consequent may be literally true, it is not the case that it is likewise assertable. The example, "If everyone on Earth is made of balloons, then my friend's dog's name is Daisy", is a case in point. Grice would argue that what it says is literally true, but that it is not assertable because it violates the maxim of quantity by not being as informative as is possible and necessary. I am in a position to assert the truth of the consequent on its own, thus in asserting it as part of a conditional instead I am not being as informative as I could be. Of course, I am also in the position to assert the falsity of the antecedent, which is a yet further way in which the relevant maxim is violated and this conditional confirmed in its unassertability.

Jackson's pragmatic alternative

Crucially, however, there are cases that a pragmatic approach is unable to adequately explicate (Jackson 1991, Bennett 2006). If the pragmatic account is correct that the logic of the indicative conditional is that of the material conditional, then it should also be able to explain logical inferences of the following kind. Consider this example from Jackson (1991):

(84) If it rains, it won't rain heavily.

Since the principle of contraposition[39] holds for the material conditional, it follows that,

(85) If it does rain heavily, it won't rain.

This seems very evidently intuitively false. And, worse, there is no pragmatic gloss that can explain why these two conditionals are not logically equivalent on Grice's account. Should it happen that it does indeed rain heavily, it's impossible for the consequent, too, to be true when the antecedent is. The conditional, (85), is then not equivalent in truth-value to its contrapositive, (84), and since this truth-value

[39]Contraposition of a given conditional gives us another conditional, logically equivalent to the first, in which the denial of the original consequent becomes the new antecedent and the denial of the original antecedent the new consequent. So, a conditional of the form, "If A, then B" becomes, "If not-B, then not-A". The status of contraposition on my own account I consider in Chapter 5 (*see p.232*).

clash is at the level of literal meaning it cannot be avoided by appealing to what is implicated in either case. Implicature can help to explain the apparent inconsistency of a theory's truth-value assignments with our intuitive judgements of a conditional's truth or falsity, but not clashes in the actual truth-values of what should be logically equivalent instances.

Frank Jackson offers an alternative pragmatic approach that's able to deal with examples such as that just given. On his conception of the indicative conditional, its truth-conditions are simply those of the material conditional, but as regards which conditionals we may actually use in a given context, both assertability and also what what he terms 'robustness' must be taken into account in addition to truth conditions:

> In the widest sense of 'meaning', $(P \rightarrow Q)$ and $(P \supset Q)$ do not mean the same. But their truth conditions are the same—they agree in sense or literal content. The extra element is that in using $(P \rightarrow Q)$, you explicitly signal the robustness of $(P \supset Q)$ with respect to P, and this element affects assertion conditions without affecting truth conditions. (Jackson, 1998: 13)

For a conditional to be robust, it must not only be true (according to the truth conditions for the material conditional) and assertable (having a consequent that's highly probable given the antecedent), the conditional itself must remain highly probable relative to P, the antecedent. So, to return to the case just considered above, the conditional, "If it does rain heavily, it won't rain," may be true (since Jackson's indicative has the truth conditions of the material conditional, it will be

true when the antecedent is)—but it is neither assertable nor robust. The probability of Q (the consequent), given P (the antecedent), is anything but high. Moreover, the probability of $(P \mid Q)$ given P is equally poor. Thus, Jackson's account is evidently able to avoid this sort of objection.

His analysis is also able to deal with other difficulties that the Gricean is not. As just discussed above, theorists who defend the claim that the indicative is equivalent to the material conditional, often explain the infelicity of certain instances by arguing that, while a conditional may be technically true, it is not assertable if the speaker is in a position to assert something stronger; that is, 'If A, then B' will be unassertable where it is the case, or it is highly probable, either that A is false or B is true. (Jackson, 1998: 4)

Take this example of Jackson's,

(86) If the sun goes out of existence in ten minutes' time, the Earth will be plunged into darkness in about eighteen minutes' time. (Jackson, 1998: 4–5)

As he observes, (86) is perfectly assertable, even though it is surely true (or at least extremely probable) that the antecedent is false; namely, that the sun will not go out of existence in ten minutes' time. One could certainly assert the negation of (86)'s antecedent, but this makes (86) itself no less assertable. Yet, a purely Gricean analysis is unable to account for this. Their admonition to 'assert the stronger' is no doubt useful in explaining the infelicity of some instances deemed true on

the material-conditional truth conditions they give the indicative, but it leads the Gricean into difficulties in cases such as this one.

Jackson's approach, however, can explain instances such as (86). Of course (86) is true, since the antecedent is (very likely) false. But, despite the very high probability that the antecedent on its own is false, this conditional itself is nonetheless assertable, since there is a very high probability that the consequent is true given that the antecedent is, and it is also robust—in other words, the probability that (86) is true remains high even if we take the antecedent to be true.

A theory of the indicative requiring both assertability and robustness, would seem to address many of the difficulties that arise on approaches arguing for the truth conditions of the material conditional—and so it does. However, there remain other challenges that it not able to overcome. Priest offers an objection to Jackson's view with the following example—it concerns Fred, the son of a couple who live in Autumn's Retreat, all of whose residents are elderly pensioners.

> We then wish to know whether Fred is under 10, and reason as follows:
>
> > If Fred is under 10, he lives with his parents. But, in that case, he lives in Autumn's Retreat, in which case he is an aged pensioner. Thus, *if Fred is under 10 he is an aged pensioner.* It follows that Fred is not under 10. (Priest, 2009: 319; emphasis in original)

As Priest points out, the italicized conditional is perfectly assertable, since it is a required step in determining whether or not Fred is under 10, which is what we wish to find out. However, Jackson is unable to account for this, as the consequent here is not in the least probable given the antecedent; in fact, given the antecedent, the consequent must be false. And, yet, despite failing to meet Jackson's requirements for assertability, this conditional clearly may be highly assertable in such a context. Furthermore, this conditional also fails to be robust in the manner specified by Jackson, since, given the antecedent, the consequent is not in the least probable, but it nonetheless remains quite assertable. It appears, then, Jackson's account is not an entirely satisfactory one, despite its not being susceptible to certain objections made to Grice's view.

What has gone wrong in Jackson's account? To make this clear, consider how Priest's example may be explained on the $C3$ approach. We know that Fred's parents live in Autumn's Retreat and that all its residents are pensioners. Now, let's assume that Fred is under 10 years old. Given our assumption, we may conclude that it is counterfactually true that,

(a) If Fred's under 10 he lives with his parents.

Of course—as we are concerned with what is the case when the conditionals necessary to make this argument are all true relative to the same given facts, assumptions, and relevant considerations—the counterfactual interpretations of these condition-

als must be made using the same *ceteris paribus* clause for each (this allows for transitivity to be preserved[40]).

With this in mind, since we have the counterfactual truth of (a) and the factual truth of its antecedent, it must be the case that the consequent of (a) also is factually true (that is, of course, assuming further that he hasn't been removed by the authorities or that he fails to live with his parents for some other exceptional reason): Fred does live with his parents. Similarly, we may conclude that, (b), if Fred lives with his parents he must live in Autumn's Retreat, is counterfactually true—and, following the same line of reasoning as with (a), that it is factually true as well, and so that it is factually true that Fred lives in Autumn's Retreat. And, given our two original facts, we known that, (c), if Fred lives in Autumn's Retreat he is a pensioner; so again, similarly, we can conclude that Fred is a pensioner at the actual world.

However, this cannot be the case, given our original assumption that he is under 10 years old at the actual world. Therefore, our assumption has led us to the contradictory conclusion that Fred both is and is not under 10 years old; hence, our starting assumption must have been false.

The problematic conditional for Jackson's analysis is,

(87) If Fred is under 10 he is an aged pensioner.

as it is seems intuitively quite assertable, but, according to Jackson's criteria, fails to be so. The difficulty is that the conditions for assertability and robustness are not

[40] *See Chapter 5, p.215, for more on transitivity.*

sensitive to the broader context. The reason that (87) fails to be assertable is that the consequent is not highly (or even slightly) probable, given the antecedent—and nor could it be robust, since the consequent's probability remains equally low taking the antecedent to be true (though, really, robustness is ruled out already by the failure of assertability). However, in the broader context of the argument being made, (87) is a perfectly felicitous conditional to assert. In order to assess these conditionals, on the $C3$ approach, this context has to be taken into account. The counterfactual interpretations cannot be given without first determining the appropriate *ceteris paribus* clause(s), given the relevant facts in evidence and assumptions being made. As it happens, since the conclusion is drawn by way of a *reductio*, some of the counterfactual (and factual) interpretations given were ultimately in fact false, but nonetheless quite legitimately made within the context of the argument at hand.

The Restrictor Analysis

The restrictor approach offers a somewhat different way of conceptualizing the structure of the conditional. On this analysis, the if-clause serves to limit the domain of the modal operator, either explicit or implicit, that has scope over the conditional.

The most prominent of the restrictor approaches is Kratzer's—developed from Lewis's (1975) claim concerning conditionals' if-clauses' behaving as restrictors of adverbs of quantification. On her account, 'If..., then...' is not a binary operator—rather, the if-clause acts as a restrictor of the domain of the associated adverb of quantification (such as 'most', 'many', 'sometimes', 'few'). In those instances in which there is no explicit adverb, Kratzer argues that there is an implicit modal;

namely, 'must' or 'necessarily'. She famously maintains that, "The history of the conditional is the history of a syntactic mistake. There is no two-place *if...then* connective in the forms for natural languages. *If*-clauses are devices for restricting the domains of various operators. Whenever there is no explicit operator, we have to posit one. [... E]pistemic modals are candidates for such hidden operators." (Kratzer, 1986: 11)

For instance, Kratzer (2012: 98) argues that the conditional,

(88) If the lights in his study are on, Roger is home.

has the logical form,

(89) (MUST: the lights in his study are on) (Roger is home)

In other words, (88) asserts that Roger is home, within the domain restricted to only those accessible worlds at which the lights in Roger's study are on.

To a certain extent, Kratzer's analysis of conditionals accords with my own, particularly with the counterfactual interpretation of $C3$. On this interpretation, the *if*-clause (or antecedent) determines the set of possible worlds at which the truth-value of the consequent is to be evaluated, which could be analyzed in terms of the restriction of the domain of evaluation of the truth-value of the consequent. My account could also be said to posit an implicit necessity operator, since the

counterfactual interpretation requires that the consequent be true at *every* A-world. However, there is a marked divergence between my use of an implicit *ceteris paribus* clause, on the counterfactual interpretation, and Kratzer's recourse to probability orderings of worlds. Need for a detailed discussion of this difference is preempted, though, by a more significant distinction between our theories of the conditional, as regards the taxonimization of the conditional on my account as opposed to restrictor approaches.

Barbara Abbott (2004: 6) attributes to Edgington a challenging objection to restrictor-type analyses of the conditional. Consider this pair of sentences, referring to a murder that has taken place (Stalnaker, 1975: 269; in Abbott, 2004: 6):

(90) Either the butler or the gardener did it.

(91) If the butler didn't do it, the gardener did.

Now, on a restrictor account, (91) is analyzed (using Kratzer's notation) as,

(92) (*MUST*: the butler didn't do it) (the gardener did it)

But (92) will be true only when the butler is not the murderer and the gardener is, while (90) is true when either the butler or the gardener committed the murder. Therefore, (92) fails to be equivalent to (90), despite the fact that it appears

intuitively clear that, at least in this sort of instance, the equivalence holds—and, indeed, that ordinary-language speakers use such conditionals interchangeably with the corresponding disjunction in many cases.

If my own account bears some similarity to the restrictor analyses of the conditional, then the question is whether this objection is similarly troubling for the $C3$ approach. In fact, though, the difference in taxonimization of the conditional (in terms of interpretations rather than types) ensures that this is not a challenge to my account.

Since the $C3$ analysis does not take the English indicative (or that of any other language) conditional as logically distinct type of conditional from the counterfactual, it is not as limited in its analysis. In the case of (91), provided that either the butler or the gardener is actually the murderer, the conditional is true on the counterfactual interpretation, regardless of which one out of the butler and the gardener did it. On the factual interpretation, the conditional is either non-truth-apt (if the butler is the murderer) or true (if the butler didn't commit the murder but the gardener did). And, in such cases, where each potential counterfactual interpretation is true and each factual interpretation is either true or non-truth-apt, the conditional is true also on the combined interpretation—and so we may say simply that (91) is true for every truth-value combination that (90) is.

Of course, should neither the butler nor the gardener be the murderer, (90) would be false, and (91) would be non-truth-apt on the factual interpretation and false on the counterfactual interpretation. And, this combination of factual and counterfactual truth-values would make the conditional as a whole false on the combined

interpretation. Thus, the $C3$ analysis confirms the intuitive equivalence of (90) and (91).

One could perhaps argue that the restrictor analysis has recourse to the claim that (90) is in fact equivalent not to (91) alone, but to the disjunction of (91) and,

(93) If the gardener didn't do it, the butler did.

which would allow for the equivalence on the restrictor approach as well as the $C3$ account.[41] While I prefer to reject this claim—because it doesn't seem to take account of how we actually use conditionals, as having (at least potentially) both factual and counterfactual import, which obviates the need to claim the disjunctive conditionals as the correct version of the simple disjunction—one needn't buy into my view to reject the restrictor analysis. There are further serious concerns that tell against this sort of account of the conditional.

Von Fintel and Iatridou frame the restrictor analyses of the conditional in terms of characterising conditionals as interchangeable with sentences in which the quantifier is restricted by a relative clause, derived from the conditional's if-clause. So, to take their example (von Fintel and Iatridou, 2002: 1), the following appear to be essentially equivalent:

[41] My account is essentially agnostic on whether these sorts of disjunctions are more accurately given in terms of conditionals as a single one—in this case, as (91)—or as two disjunctive conditionals—here, (91) or (93)—as both are equivalent in their truth-values.

(94) Every student will succeed if he studies hard. ≈ [This is approximately equivalent to:]

(95) Every student who studies hard will succeed.

In other words, on a restrictor account the conditional (94) could be given as,

(96) (*EVERY*: student studies hard) (student will succeed)

They consider various objections to the possibility of turning certain sorts of conditionals into sentences whose quantifiers are restricted by a relative clause, as well as potential responses to such objections. Ultimately, though, they look only at those conditionals that have clear and uncontentious quantifiers, since their point here is that even such paradigmatic instances can lead to problems—thereby undermining the restrictor analysis more generally.

To get a better idea of the apparent interchangeability of conditionals and sentences in which the quantifier is restricted by a relative clause, let's look at a few further examples:

(97) No giraffe, if it stretches, is unable to reach edible leaves.

(98) No giraffe who stretches is unable to reach edible leaves.

(99) If anyone has trouble with the project, they can phone you.

(100) Anyone who has trouble with the project can phone you.

(101) Most countries have a high trade deficit, if their currencies are too strong.

(102) Most countries whose currencies are too strong have a high trade deficit.

What von Fintel and Iatridou argue is that this supposed interchangeability runs into insurmountable difficulties in certain sorts of cases. Take the following sentence (von Fintel and Iatridou, 2002: 6):

(103) Every book that I needed for the seminar happened to be on the table.

If the if-clause of the conditional genuinely acts simply as a restrictor on the quantifier, then it should be possible to give (103) in conditional form without essentially altering its meaning,

(104) Every book happened to be on the table if I needed it for the seminar. (von Fintel and Iatridou, 2002: 6)

But (104) does not say the same thing as (103). The latter sounds distinctly odd, in that it intuitively seems to suggest that the happenstance of those books being on the table was in some way due to my needing them for the seminar. This is precisely the implicature that I have argued is carried by the conditional—namely, that the consequent in some way follows from or is dependent on the antecedent—but the restrictor view, in arguing that the antecedent is merely a restrictor on the sentence's (explicit or implicit) quantifier fails to account for this.

Finally, let's consider a further kind of problematic case, as exemplified by these two pairs of sentences (von Fintel and Iatridou, 2002: 7),

(105) Many/A few of the students will succeed if they work hard.

(106) Many/A few of the students who work hard will succeed.

(107) Nine of the students will succeed if they work hard.

(108) Nine of the students who work hard will succeed.

As they note, the second sentence of each pair does not quite capture the meaning of its corresponding conditional, as it lacks what they term a "weak existence presupposition". In the second of each pair, the relative clause appears intuitively to include

the presupposition that, in each respectively, there are students who work hard. But there is no such existence presupposition in the corresponding conditionals, which each appear to allow, respectively, for instances in which none of the students work hard. Von Fintel and Iatridou observe that (108) "is naturally read as presupposing that there are more than nine students who work hard. But [(107)] only presupposes that there are more than nine students simpliciter and does not presuppose anything about how many, if any, work hard." (von Fintel and Iatridou, 2002: 7)

This gives us additional reason to reject restrictor accounts as full and accurate analyses of the conditional—indicative or otherwise.

There are evidently many and diverse challenges to the various accounts of the indicative conditional discussed in this chapter. One of the key issues, though, is the problems that arise from the pervasive failure to accurately taxonimize the conditional. Taking particular instances to be, in themselves, indicative (or as concerning what is actually the case), consistently leads the theorist into difficulty. In those instances where the possible counterfactual interpretations are evident (albeit not *qua* counterfactual interpretations), they can make the truth-value of the conditional seem confused or contradictory, and even lead some philosophers to abandon truth conditions for the conditional all together. But, when we take adequate account of the potential alternative interpretations available, many of these difficulties are dissolved.

Chapter 4: **Possible-worlds and counterfactuals**

Accurately assessing counterfactual conditionals[42] is of broad philosophical importance. As Goodman reminds us,

> The analysis of counterfactual conditionals is no fussy little grammatical exercise. Indeed, if we lack the means for interpreting counterfactual conditionals, we can hardly claim to have any adequate philosophy of science. A [...] solution to the problem of counterfactuals would give us the answer to critical questions about law, confirmation, and the meaning of potentiality. (Goodman, 1947: 3)

[42] As I mentioned previously, I speak in chapters 3 and 4 of indicative and counterfactual conditionals—of conditionals as divided into inherently different types—but this is purely for ease of reference when discussing theories that assume such categorizations, rather than any endorsement of this.

Goodman's (1947) definition of the counterfactual is more determinate than that of a number of other accounts (notably Lewis's, 1973b): He specifies that he is concerned with only those conditionals that have irrevocably false antecedents and false consequents—that is, that they not concern the future or events of otherwise uncertain truth-value. The problem that Goodman identifies for the analysis of counterfactual conditionals is that they clearly cannot be treated as purely truth-functional, as having the truth conditions of the material conditional, since, given his definition of counterfactuals, they would all inevitably be true on this account in virtue of their false antecedents. Moreover, as he observes, we cannot look to empirical evidence to settle the truth-value of counterfactuals, since their antecedents are by definition false (at the actual world) on his account.

According to Goodman (1947), we need to address the fact that conditionals presuppose more than is given by the antecedent—further conditions must be met in order for the connection between antecedent and consequent to obtain. In addition, Goodman maintains that it is not usually a logical connection, between antecedent and consequent, that is suggested by a conditional. This, of course, is very much in line with my account of the counterfactual interpretation of the conditional, which posits an implicit *ceteris paribus* clause that must hold true at a world in order for it to be counted among the set of worlds at which the conditional should be assessed.

Edgington (1995: 248) characterizes the view that Goodman suggests along these lines:

> A counterfactual conditional "$A \rightarrow C$" is true if and only if there is

> a conjunction of truths T which include a law of nature [and satisfy condition X] such that $A\&T$ entails C.

X is a place-holder for the difficult bit.

That "difficult bit" being the determination of which things must be held constant and which taken to be different in order to accurately assess the conditional in question. To clarify this, consider an example Goodman (1947: 8) gives,

(109) If the match had been scratched, it would have lighted.

Now, this putative counterfactual conditional will be true only if certain things are taken to remain the same and certain others altered appropriately. For instance, in the former category would be things such as that the match is dry, that its head has a sufficient amount of the normal phosphorous-based flammable coating and has not been treated with any kind of fire-retardant, that the degree of oxygen in the air remains the same. In the latter category would be things such as that the match is picked up and scratched instead of just left lying there untouched. The law of nature referred to above is that (or those) required in order for this conjunction of truths (whichever may have been appropriately altered or held fixed) to result in the consequent being made true. In this case, presumably these would be laws about the flammability of particular substances and so forth.

The initial suggestion that Goodman makes, of how to deal with "condition X", is to provide an answer in terms of cotenability. S, in the following, is the set of true sentences that, together with a law of nature, constitutes T, mentioned above. His definition of cotenability (Goodman, 1947: 15) is that,

> A is cotenable with S, and the conjunction $A{\cdot}S$ self-cotenable, if it is not the case that S would not be true if A were.

The difficulty, then, is how to decide what satisfies this definition for a given antecedent, and what does not, without relying on knowledge of the truth-value of the counterfactual conditional containing that antecedent. Returning to the match example—how can cotenability give us the truth-value of (109)? Well, this counterfactual will be true, on Goodman's picture, in virtue of various laws of nature, in conjunction with a set of true sentences, S, cotenable with (109)'s antecedent, A. Consider the sentence,

(110) The match was dry.

This is cotenable with A,

(111) The match was scratched.

provided it's true that it's not the case that (110) would not be true if (112) were; namely, that

(112) It is not the case that the match was not dry if the match was scratched.

But, we can only determine the truth-value of (112) if we already know the truth-value of (109). We cannot look to empirical evidence at the actual world in order to settle counterfactual questions, so we must consider such things as whether the match would have lit if struck. If we take (109) to be true, then it must surely be the case that the match was dry, or it would not light if struck. The obvious trouble is that this is blatantly circular. We cannot settle the truth-value of (109) without already knowing the truth-value of sentences such as (110), but that in turn requires that we know whether (110) is cotenable with (111), the antecedent of (109); and that can only be determined if we already know the truth-value of (112), which is precisely what we were trying to establish in the first place.

This is the problem that Goodman leaves for any prospective analysis of the counterfactual conditional—the problem of how to provide an explication that doesn't simply rest on our intuitive understanding of counterfactuals' truth-values, but gives an account of why they have the truth-value they do.

I would argue that there is good reason to think that this problem has no solution—at least of the sort that Goodman sought. The best-known, and most

widely accepted, treatments of counterfactuals are those of Stalnaker (counterfactuals forming one part of his unified account of the conditional) and Lewis; and both of these make use of primitive terms that rely on our intuitive understanding of them—as does my notion of all things being equal. One might take this as sign that we should still be looking for a better theory, one that doesn't rely on terms of this kind; but that would be a fruitless search. The trouble lies in the very nature of counterfactuals, not (in this regard at least) in any lack in the analyses themselves.

Goodman himself diagnosed the problem; namely, that there is no independent body of empirical facts that can serve as the final arbiter of the counterfactual conditional's truth-value. Whether or not the particular term argued for by a given theory accurately captures just what constraints must be placed on the the worlds at which an explicit antecedent is true, in order to successfully assess the relevant counterfactual conditional—such notions as closeness, comparative similarity, or all things being equal are not merely *ad hoc*. It is by means of extrapolating, idealizing, and estimating from our myriad experiences of how the world is—of what generally follows from what and under what sort of circumstances, of how certain things typically affect others, etc.—that we come to a rough understanding of what sort of counterfactuals are true and what sort false. The messiness is inherent to counterfactuals; the question is, which is the best notion to capture this form of approximation by which we intuitively judge counterfactuals?

Stalnaker's Possible-Worlds Account of the Conditional

Unlike the accounts of the preceding chapter, which restrict themselves to consideration of just the putative indicative conditional, Stalnaker's conception of the conditional is a unified one.

Stalnaker remarks, of the claimed indicative/subjunctive division of conditionals, that,

> It is clear that the two kinds of conditionals have much in common, but also clear that there are semantic differences between them, since there are minimal pairs, differing only in that one is 'subjunctive' and the other 'indicative' that seem, intuitively to say quite different things. (2009: 228)

The Oswald–Kennedy pair are, of course, typically taken to be a paradigm example of this semantic difference, though we have seen that this claim does not stand up to scrutiny. In a footnote to the above quote, Stalnaker clarifies that,

> By a *semantic* difference, I mean here a difference in the assertive content that utterances of the contrasting conditional sentences would have in a similar situation. This is compatible with the hypothesis that the abstract semantics for the two conditionals is the same, but that the difference in content is explained by a difference in contextual determinants relative to which the contrasting kinds of conditionals are interpreted. (2009: 228; original emphasis)

This explanation goes some way toward my own account, in acknowledging that conditionals may change meaning between contexts, but does not address the possibility of their being more than one interpretation of a given conditional in a single context.

What, then, is Stalnaker's unified account of the conditional, and how does it handle the seeming differences between the two types he defends? "The difference between the two kinds of conditionals is explained in terms of different constraints imposed on the contexts relative to which the different forms of conditionals are interpreted." (Stalnaker, 1999: 64) "Consider a possible world in which A is true, and which otherwise differs minimally from the actual world. 'If A, then B' is true (false) just in case B is true (false) in that possible world." (Stalnaker, 1968: 102) And, as the world that differs least from the actual world may be the actual world itself, Stalnaker's approach encompasses both so-called indicative and counterfactual conditionals.

Let α be a possible world, A the antecedent of the conditional, B its consequent, and $f(A, \alpha)$ what Stalnaker terms the selection-function or s-function. This s-function picks out the closest possible world, relative to a particular base world α (for us, the actual world), at which A is true. Then, the conditional connective $>$ may be applied according to the following rules,

$A > B$ is true in α if B is true in $f(A, \alpha)$;

$A > B$ is false in α if B is false in $f(A, \alpha)$ (Stalnaker, 1968: 103)

What's required, then, is an explanation of how to determine the value of the s-function, $f(A, \alpha)$. Given that, we can ascertain whether B is true at this selected world and, if so, deem the conditional true (or, if not, false). Stalnaker defines λ as "the absurd world"—at which everything, including contradictions, is true—and makes use of it in the conditions he places on the s-function:

> [W]here $f(A, \alpha) = \beta$, A is the *antecedent*, α is the *base world*, and β is the *selected world*.
>
> (1) For all antecedents A and base worlds α, A must be true in $f(A, \alpha)$.
>
> (2) For all antecedents A and base worlds α, $f(A, \alpha) = \lambda$ only if there is no world possible with respect to α in which A is true. (Stalnaker, 1968: 104; emphasis in original)

Stalnaker's first condition here ensures that the relevant antecedent be true at the world given by the s-function; the second provides that, in cases where there is no possible world at which A is true—namely, that the antecedent is impossible—that the selection-function returns the absurd world, λ, at which any B whatsoever will be true. Further:

(3) For all base worlds α and all antecedents A, if A is true in α, then $f(A, \alpha) = \alpha$.

(4) For all base worlds α and all antecedents B and B', if B is true in $f(B', \alpha)$ and B' is true in $f(B, \alpha)$, then $f(B, \alpha) = f(B', \alpha)$. (Stalnaker, 1968: 104)

His third and fourth conditions seek to make clearer the ordering of possible worlds according to which the s-function selects the world closest to, or most closely resembling, the given base world relative to the antecedent concerned. The third condition provides, as mentioned above, that in those instances in which the antecedent is true at the base world itself, that the world selected is also this base world, since it is the world most similar to itself at which A is true. This is effectively a strong centering constraint. Stalnaker's fourth condition ensures that there is an ordering of worlds, such that if a world is given by the s-function as that most closely resembling the base world at which A is true, then no other distinct world may be determined to be closer to the base world. This is, in other words, a uniqueness constraint, which ensures that there is just one, unique closest or most similar world in terms of which the given conditional is assessed.

What exactly constitutes a particular world being the closest or most similar? Stalnaker says that the, "Relevant respects of similarity are determined by the context", and that, "when a speaker says 'If A', then everything he is presupposing to hold in the actual situation is presupposed to hold in the hypothetical situation in

which A is true." (1975: 144) That said, he acknowledges that there are contexts in which one wishes to give up certain presuppositions, but that some indication of this must be given; typically done, in English, by means of the subjunctive mood.

Differentiation, on Stalnaker's account, between the putative indicative and counterfactual conditionals, consists in whether the antecedent is true at worlds that are still live possibilities, given the common ground—the shared presuppositions among the conversational participants—in the context within which the conditional is used, or whether the antecedent can only be true at worlds outside this context set.

> To make an assertion is to reduce the context set in a particular way, provided that there are no objections from the other participants in the conversation. The particular way in which the context set is reduced is that all of the possible situations incompatible with what is said are eliminated. (Stalnaker, 1999: 86)

And, for indicative conditionals, it's a world from within this context set that makes true its antecedent; while, for a counterfactual conditional, one must go beyond the context set to find a world at which the antecedent is true. This is reflected in the further proviso he adds to the selection function, "[I]f the conditional is being evaluated at a world in the context set, then the world selected must, if possible, be within the context set." (Stalnaker, 1975: 275) This is motivated by Stalnaker's view that when we use conditionals in ordinary discourse we are typically concerned with how things are at the actual world, and thus we should first seek to evaluate

conditionals relative to worlds that have not yet been ruled out as possibly actual. Only then should we turn to worlds we know to be non-actual.

According to Stalnaker (2009), the propositions that are expressed by various instances of the conditional will typically differ in accordance with the grammatical features of the utterance, such as tense and mood, which in turn affect how the selection function applies to each conditional. So, though the same selection function is applied to both indicatives and counterfactuals, Stalnaker maintains that the proposition to which it's applied will usually differ in virtue of the conditional in question being either indicative or subjunctive in mood; and likewise for other grammatical features of English conditional constructions. These semantic distinctions between types of conditionals then result in differences in the pragmatics of how the selection function is constrained in each specific case.

Objections to Stalnaker's account

Of course, I believe there are good grounds on which to disagree with the claim that specific instances of the conditional are either indicative or subjunctive/counterfactual, for all the reasons that I have discussed in earlier chapters. To recap very briefly, the supposed difference Stalnaker posits in the propositions expressed by given conditionals, according to whether they are indicative or counterfactual, fails to account for cases in which either a factual or counterfactual interpretation is possible. Imagine, for example, you're on a building site and someone says to you,

(113) If you touch that wire, you'll get an electric shock.

As it happens, the wire in question actually has an insulating plastic coating. You may quite appropriately respond,

(114) No, I wouldn't get a shock. It's completely insulated by its plastic coating.

and then touch the wire, just to emphasize your point. Or, just as legitimately, you could answer with,

(115) No, look: See, I'm holding the wire and... no shock.

In both cases the context set is identical. Both include the actual world as a live possibility, yet the first response, (114), interprets (113) as concerning not just the actual world, and is made true by the consequent being false at all those worlds at which, *ceteris paribus*, the antecedent is true. The claim (114) takes the conditional to make is one broader than simply the fact that you weren't actually shocked, at the actual world, when you touched the wire. In the case of (115), the response concerns merely how things actually are—made true at the actual world by the fact that you touch the wire and don't get shocked. You might be an inordinately concrete-minded person, who has no interest in how things might be, concerned purely with how things actually are. This wire was actually insulated; therefore, you actually didn't get a shock. Whether you would or would not have got a shock in any circumstance like this one, may be a matter of complete indifference. The point is that both types

of response—and even one that combines or is indeterminate between the two—are available in a context like this, without any difference in the context set to account for it.

There is also the difficulty of Stalnaker's analysis being limited to English conditionals (or, at best, those of languages with both indicative and subjunctive moods), though the theory proposed would seem to be concerned with more than merely an explication of the particularities of the grammar of the conditional in a single language (or in a particular language-type). In fairness, though, one could argue that a language need not have a subjunctive mood in order for Stalnaker's approach to be applicable. Though use of the subjunctive may (however unreliably) signal instances of the conditional that must be assessed relative to worlds outside the context set in English, it seems reasonable to assume that what Stalnaker really is interested in here is counterfactual conditionals, with the subjunctive being simply a convenient general marker for this. And, so, whatever other means, grammatical or pragmatic, that a language may use to distinguish between uses of the conditional intended to concern purely what is the case and those intended to concern also what might be, one could substitute these for the subjunctive as an indicator of the difference. Of course, this does nothing to address the problems inherent in making such a distinction at the level of instances of the conditional themselves, rather than that of interpretation.

Another difficulty with Stalnaker's account of the conditional lies in the uniqueness requirement that he builds into the notion of closeness, or resemblance. This is problematic, because in the case of certain counterfactuals there's simply no clear

ordering such that there is a single, unique, closest world to the base world. David Lewis offers the following example,

> It is not the case that if Bizet and Verdi were compatriots, Bizet would be Italian; and it is not the case that if Bizet and Verdi were compatriots, Bizet would not be Italian; nevertheless, if Bizet and Verdi were compatriots, Bizet either would or would not be Italian. (Lewis, 1973b: 80)

The inability of Stalnaker's theory to deal with such examples is highly problematic for his view, as Lewis goes on to explain:

> However little there is to choose for closeness between worlds where Bizet and Verdi are compatriots by both being Italian and worlds where they are compatriots by both being French, the selection function still must choose. I do not think it *can* choose—not if it is based entirely on comparative similarity, anyhow. Comparative similarity permits ties, and Stalnaker's selection function does not. (Lewis, 1973b: 80; emphasis in original)

Stalnaker is aware of such instances and does address them—he maintains that there are truth-value gaps in cases such as that of a conditional like,

(116) If Bizet and Verdi were compatriots, then Bizet would be Italian.

since this sort of conditional fails to pick out a single particular situation. (Stalnaker and Thomason, 1970: 27–28) The answer is to provide a two-valued logic, as he does, as a useful idealization, for those cases in which we do have complete semantic information, but with a view to developing the s-function to allow for instances such as (116), which are semantically incomplete. (He gestures toward van Fraasen's 1966 as offering a means to achieve this.) The trouble, however, is that conditionals such as (116) do not appear to be too difficult to assign a truth-value. It seems intuitively false that Bizet would have to be Italian, were he and Verdi compatriots. Certainly, that is one way that the two could be compatriots, but by no means the only way. Yet, (116) presents Bizet's being Italian as being definitively the case, given that he and Verdi are compatriots. There are, of course, interpretations of the conditional on which specific instances are not truth-apt, but this does not appear to be an example of such an instance.

If, instead, we assess (116) on the counterfactual interpretation of $C3$, there is no such difficulty. This approach takes account of the fact that there are some worlds at which the antecedent is true, at which Bizet is Italian. But, as this interpretation is not limited to selecting a single world of evaluation, it considers also that there are A-worlds at which both Bizet and Verdi are French, and thus compatriots in this way. So, as it is not the case that Bizet would undeniably be Italian were the two compatriots, (116) is false (as Lewis, 1973b, evidently also takes it to be).

But what then of the pair of conditionals, (116) and

(117) If Bizet and Verdi were compatriots, then Bizet would not be Italian.

On my account both of these conditionals would be judged to be counterfactually false, yet it seems surely counterfactually true that,

(118) If Bizet and Verdi were compatriots, then Bizet would either be Italian or not be Italian.

And, indeed, on the $C3$ analysis (118) would be deemed counterfactually true (and true on the combined interpretation as well). The question, though, is how (118) can be true, when (116) and (117) are both counterfactually false, and the two latter conditionals appear to be disjunctively equivalent to the former. The answer is that—while the two respective consequents of (116) and (117) are lexically the same as the two disjuncts in the consequent of (118), respectively—$C3$ being a strict conditional (albeit variably strict) means that the disjunction of two separate conditionals' consequents is not equivalent to a single conditional with both of these combined into its single consequent. The straightforward reason for this is that, as a strict conditional, $C3$'s counterfactual interpretation requires that a given conditional's consequent be true at *every* A-world. While each of (116) and (117)'s consequents were true at *some* of these conditionals' respective A-worlds, neither was true at every single one of its respective set of A-worlds—though it is the case that at every A-world at which the consequent of (116) was false, (117) was true, and vice versa. (Naturally, I'm assuming that it's the same worlds forming the set of A-worlds in each of the three conditionals at issue here; namely, that the *ceteris paribus* clause of each is the same.)

Hence, when these two consequents from (116) and (117) are combined in the single disjunctive consequent of (118), since they're exhaustive of all the relevant *A*-worlds, this conditional is judged counterfactually true—which perfectly accords with our ordinary intuitive judgments of these conditionals.

Lewis and Strengthening the Antecedent

An invariably strict conditional, such as that introduced by C.I. Lewis (1918)—one on which the consequent must be true at every antecedent world in order to be deemed true—would help to deal with the sort of problems faced by Stalnaker's theory—in that we would not be forced to choose one unique world closest to the relevant base world, but would rather consider all of the worlds at which the antecedent is true. This would avoid the problem of the kinds of ties in closeness that David Lewis identifies.

However, David Lewis (1973) rejects a fixed or invariant strict conditional in favor of one that is variably strict (though, unlike Stalnaker's, his account is concerned with just the counterfactual conditional). The *C*3 conditional is in fact also a form of variably strict conditional, but one varying according to different criteria than does Lewis's (hereafter, references to 'Lewis' are to David Lewis, unless otherwise indicated). The question is, then, which account's criteria provide the better analysis of the counterfactual.

Lewis's theory of the counterfactual does not rely on the uniqueness and limit assumptions that Stalnaker's does. It will become clear why as I explain Lewis's approach, below. He gives the truth conditions of the counterfactual conditional,

$\Box\!\!\rightarrow$, as follows (Lewis, 1973b: 16),

> $\phi \Box\!\!\rightarrow \psi$ is true at a world i (according to a system of spheres $\mathfrak{S}$) if and only if either
>
> (1) no ϕ-world belongs to any sphere S in $\mathfrak{S}_i$, or
>
> (2) some sphere S in $\mathfrak{S}_i$ does contain at leat one ϕ-world, and $\phi \supset \psi$ holds at every world in S.

These spheres are demarcators of comparative similarity. Lewis says further that, "Corresponding to a variably strict conditional [...] there must be an assignment to each world i of a *set* $\mathfrak{S}_i$ of spheres of accessibility around i" (1973b: 13–14; emphasis in original), onto which he places these four conditions:

> Let $\mathfrak{S}$ be an assignment to each possible world i of a set of $\mathfrak{S}_i$ of sets of possible worlds. Then $\mathfrak{S}$ is called a (*centered*) *system of spheres*, and the members of each $\mathfrak{S}_i$ are called *spheres* around i, if and only if for each world i, the following conditions hold.
>
> (C) $\mathfrak{S}_i$ is *centered on* i; that is, the set $\{i\}$ having i as its only member belongs to $\mathfrak{S}_i$.
>
> [(a)] $\mathfrak{S}_i$ is *nested*; that is, whenever S and T belong to $\mathfrak{S}_i$, either S is included in T or T is included in S.

> [(b)] $\mathfrak{S}_i$ is *closed under unions*; that is, whenever $\mathcal{S}$ is a subset of $\mathfrak{S}_i$ and $\cup\mathcal{S}$ is the set of all worlds j such that j belongs to some member of $\mathcal{S}$, $\cup\mathcal{S}$ belongs to $\mathfrak{S}_i$.
>
> [(c)] $\mathfrak{S}_i$ is *closed under (nonempty) intersections*; that is, whenever $\mathcal{S}$ is a nonempty subset of $\mathfrak{S}_i$ and $\cap\mathcal{S}$ is the set of all worlds j such that j belongs to every member of $\mathcal{S}$, $\cap\mathcal{S}$ belongs to $\mathfrak{S}_i$. (Lewis, 1973b: 14; emphasis in original; footnote omitted)

These four conditions are given to ensure that the system of spheres around i, relative to which counterfactual conditionals are assessed, effectively represent the "comparative overall similarity" (Lewis, 1973b: 14) that the worlds within a given sphere bear to i.

The condition (C) provides for centering in Lewis's account. In Chapter 3 I discussed the question of centering as it relates to $C3$[43], which lacks a centering condition of the kind Lewis includes here, in order to accommodate those exceptional cases in which a conditional is made factual true due to some sort of bizarre coincidence at the actual world that must be ruled out on the counterfactual interpretation. (The implications for *modus ponens* I discuss in Chapter 5[44].) What centering amounts to here, for Lewis's view, is simply the presumption that any base world i is at least as similarly to itself as any other possible world may be, and so (C) ensures that i is included in every nonempty sphere around i.

[43] *See p.62.*
[44] *See p.223.*

Condition (a) provides for the nesting of spheres around i, which prevents any cases of a world j falling within sphere S, but not within sphere T, and a world k falling within sphere T but not sphere S. This is achieved by the requirement of this condition that any given sphere must either be nested within any other, or vice versa. As the closeness of spheres to i carries information as to the comparative similarity of the worlds in each sphere to i, a case such as that of S and T would allow for situations in which a world such as j could be both more similar to i than k and simultaneously less so. But, since the notion of similarity here is that of *overall* similarity, this would be incoherent and so must to be precluded.

Condition (b) simply ensures that—if the union $\cup\mathcal{S}$ of a set $\mathcal{S}$ of spheres around i contains a world j but not k, and that hence j falls within a sphere S in $\mathcal{S}$ but k does not; making $\cup\mathcal{S}$ a set all of whose member worlds are more similar to i than any outside it—such a set $\cup\mathcal{S}$ is a sphere around i.

In much the same way, condition (c) ensures that—if the intersection $\cap\mathcal{S}$ of a nonempty set $\mathcal{S}$ of spheres contains a world j but not k, and hence j falls within a sphere S in $\mathcal{S}$ but k does not; making $\cap\mathcal{S}$ a set all of whose member worlds are more similar to i than any outside it—such a set $\cap\mathcal{S}$ is a sphere around i. (Lewis, 1973b)

These conditions having been clarified, we can further explain the truth conditions, given above, that Lewis provides for the counterfactual conditional: "In brief: a counterfactual is vacuously true if there is no antecedent-permitting sphere, non-vacuously true if there is some antecedent-permitting sphere in which the consequent holds at every antecedent-world, and false otherwise." (1973b: 16) This, of course, makes any conditional with an impossible antecedent vacuously true, despite the

fact that it seems we can and do intuitively distinguish among the truth-values of different such conditionals. Consider, for instance, these examples (adapted slightly from Priest, 2008: 74),

(119) If Unathi squares the circle, I'll buy her a celebratory drink.

(120) If Unathi squares the circle, I'll give her my house.

Even though it's impossible that anyone can or will square the circle, (119) and (120) nonetheless seem quite intelligible, as well as being distinct in intuitive truth-value. If my friend Unathi squared the circle I would most definitely at least buy her a celebratory drink, but I'm equally certain that I would not give her my house. This cannot be accounted for on Lewis's theory, since both of these conditionals would be judged merely vacuously true, instead of (119) substantively true and (120) substantively false. Naturally, though, this is only a problem for Lewis if one feels the intuitive force of the claim that the truth-values of these conditionals differ from that of 'vacuously true', which his theory would assign them. My own theory, as it stands, is in much the same boat as Lewis's—but, in Chapter 5[45], I offer a potential modification of my approach that is able to handle impossible antecedents so as to legitimate these intuitive truth-value ascriptions.

To return to Lewis's account, discussion of the following examples (from Lewis 1973b: 10) will help to make clearer just how conditionals are to be assessed on his

[45] *See p.240.*

view:

(121) If Otto had come, it would have been a lively party.

(122) If Otto and Anna had come, it would have been a dreary party.

(123) If Otto and Anna and Waldo had come, it would have been a lively party.

The truth of (121) appears intuitively plausible, but would seem to be inconsistent with the equally plausible (122), although perhaps consistent with (123). If Lewis's were a fixed strict conditional, it would have to assess the truth-value of (122) and (123) at just the same set of worlds as it did (121), which would result in at least one of the three being assigned an intuitively incorrect truth-value, since they clearly could not all be true at every one of the very same worlds. This, roughly speaking, was the problem that motivated Lewis to opt instead for a variably strict conditional.

Unlike a simple strict conditional (like that of C.I. Lewis), David Lewis's is able to allow for the truth of each of (121), (122), and (123). The sphere of possible worlds most comparatively similar to the actual world at which the vivacious Otto did go the party, may well be such that all of the worlds within it are also ones at which the party was a lively one. But, since Lewis's conditional is variably strict, we are not forced to stick with this same sphere of evaluation when we assess (122). This conditional is assessed on Lewis's account by looking at all of those worlds in the

sphere closest to the base world—that is, the sphere containing those worlds most similar to the actual world at which both Otto and Anna went to the party. But, if, say, Otto and Anna dislike each other and tend both to be sulkily quiet when at the same function with one another (and no other guests are much fun without prompting), then it's perfectly plausible that the worlds most similar to the actual world at which the antecedent of (122) is true are ones at which it is also the case that the party is dreary.

What of (123)? We may now suppose that Waldo, a mutual friend of Otto and Anna's, is such a charmingly diplomatic person that whenever he too is there, Anna and Otto are persuaded to forget their feud and enjoy themselves in their infectiously jovial way. And, so, at those most similar worlds at which Otto, Anna, *and* Waldo go to the party, it may well be a lively one. In other words, Lewis's analysis allows us to account for the possibility that both (121), (122), and (123) are simultaneously true.

On my account, we can't say what the truth-values of (121), (122), and (123) would be on a factual interpretation, as we don't know whether or not their respective antecedents and consequents are true at the actual world; though it is possible to say, at least, that not all of them could be simultaneously factually true, since the party could not be both lively and dreary at the same time. But, it is obviously the counterfactual interpretation, rather than the factual, that is of primary salience and interest here.

The truth-value of (121) on a counterfactual interpretation is determined by whether the consequent, that the party at issue was a lively one, is true at every

$A_{(121)}$-world (I'm adding subscripts here to help avoid confusing each conditional's respective A-worlds). In other words, (121) is counterfactually true if and only if, at every world at which, *ceteris paribus*, Otto went to the party, it is true that the party was lively.

As the counterfactual interpretation of the $C3$ conditional is, like Lewis's, a variably strict one, it too is able to accommodate the simultaneous truth of all three of these conditionals. Thus, while all the $A_{(121)}$-worlds, if they made true (121) would render false (122)—this is essentially irrelevant. That's because, in $C3$ it is only $A_{(122)}$-worlds worlds that are used in evaluating the truth-value of (122), only $A_{(121)}$-worlds that are used in evaluating (121), and so on.

What could still be potentially problematic, is the claim that the antecedent of (122) does not obtain at any of the worlds $A_{(121)}$-worlds. In discussing the motivation for the *ceteris paribus* clause in the definition of the $C3$ conditional, I made the point that, in interpreting the conditional, "If I throw this ball into the air, an eagle will swoop down and grab it.", counterfactually, one obviously intends to exclude such improbable and exceptional eventualities as the beach being covered overnight in toxic sludge[46]; fine weather notwithstanding. But, turning to the examples in question here—whether the particular people named in (121), (122), and (123) do or don't attend the party concerned, none of the possible variations in their attendance seems wildly outside the ordinary course of things. Thus, one might want to object that there's no reason to think that, at all of the $A_{(121)}$-accessible worlds, neither Anna nor Waldo attended the party.

[46] *See p.94.*

There are, however, good grounds for maintaining that the *ceteris paribus* clause inherent in the counterfactual interpretation of my $C3$ account does rule out the attendance of Anna and Waldo at the party in the case of (121), since the matter of whether or not they attend has clear bearing on the truth or falsity of the consequent. So, should either or both of them go to the party, as well as Otto, all things would *not* be equal. The key point here is that violations of the *ceteris paribus* clause need not be remarkable in any broad or general sense—in the way that a plane crash or a horse born with only three legs would be. All that's required is that they be remarkable with regard to the facts of the conditional at issue.

In addition, whether or not Waldo's attendance is clearly excluded by the *ceteris paribus* clause in this instance, the attendance of Anna, Waldo, or both is, quite simply, not required in order to make the antecedent true, while—all other things being equal—everything else is as at the actual world.[47]

An Alternative to Similarity

The above considerations show that the $C3$ version of a variably strict conditional is as capable of handling these sorts of sequences of conditionals as Lewis's is. But

[47]Of course, the context of the conditional's use is crucial. I'm here assuming that there aren't some vital facts of which we're not apprised (though, of course, even if we don't know them, these could still affect the actual truth-values of the conditional's interpretations.) For instance, if I asserted (121) in a context in which Otto almost never went anywhere without Anna, and he was always the making of a party on his own, though he and Anna together invariably quashed any liveliness—in this sort of situation, (121) would surely be false, since it would be entirely plausible then to take the *ceteris paribus* clause not to rule out as A-worlds those at which Anna *did* also attend the party. (And, even if there were some possible worlds at which Otto came alone, despite his over-attachment to Anna, that would not serve to make (121) true, as $C3$ requires the consequent to be true at *all* A-worlds for the conditional to be true on a counterfactual interpretation.)

why, then, prefer my theory? It turns out that there are significant problems with the reliance of the Lewisian conditional on ordering the spheres of possible worlds around i in terms of comparative similarity. Though the *ceteris paribus* clause of the $C3$ account performs the same sort of function that the notion of similarity does for Lewis, there are crucial points of difference that make the former approach preferable.

Just what's wrong with similarity?

The primary issue with an ordering of worlds by means of similarity, is that there is good reason to doubt that this notion accurately picks out just those worlds at which the antecedent is true that are genuinely relevant to determining a given (for Lewis, counterfactual) conditional's truth-value. Fine's well-known objection in this regard is as follows:

> The counterfactual 'if Nixon had pressed the button there would have been a nuclear holocaust' is true or can be imagined to be so. Now suppose that there never will be a nuclear holocaust. Then that counterfactual is, on Lewis's analysis, very likely false. For given any world in which antecedent and consequent are both true it will be easy to imagine a closer world in which the antecedent is true and the consequent false. For we need only imagine a change that prevents the holocaust but that does not require such a great divergence from reality. (Fine, 1975: 452)

Clearly, the trouble here with the notion of similarity is that it seems to pick out the wrong world—one in which there is no nuclear holocaust, despite Nixon's pressing the button to launch the US's missiles—thereby falsifying this conditional, and contradicting not only our ordinary intuition that it is very likely true but also Lewis's own judgment as to the answer his account should give. Lewis (1979) does offer a response to Fine, arguing that similarity of the sort concerning counterfactual conditionals' truth-values is determined by the following similarity considerations, listed in order of declining priority:

1. It is of the first importance to avoid big, widespread, diverse violations of law.

2. It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails.

3. It is of the third importance to avoid even small, localized, simple violations of law.

4. It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly. (Lewis, 1979: 472)

Lewis contends that this overcomes Fine's objection, in that, were Nixon to press the button to launch the US's nuclear missiles, an extensive 'miracle'[48], or violation of actual-world physical law, would be required to prevent nuclear war and negate all of the other, undoubtedly numerous and widespread, effects of Nixon's decision to press the button, so as to return this possible world to a perfect match with the actual world from then onwards. And, since the avoidance of such widespread, diverse violations of law is prioritized over maximizing the spatio-temporal extent of perfect match of particular fact between worlds, the possible world that Fine considers would turn out to be less similar than one in which nuclear holocaust did indeed ensue from Nixon's pressing the button.

However, it's arguable whether Lewis's similarity considerations, and the relative priority he gives them, are independently justifiable, rather than an *ad hoc* counter to Fine's challenge. Regardless, it is also unclear that these guidelines for determining maximum similarity between or among worlds provide an adequate solution. Were Nixon to push the button to launch nuclear missiles while sleep walking, for instance, then it's possible he might remember nothing of the incident in the morning, so that it would have no wider repercussions, and only a second minor, localized 'miracle' (in addition to that which caused Nixon to push the button while sleep walking, making the antecedent true) would be required for the electrical signal sent by the button press to simply disappear and everything to continue at this possible world just as it has at the actual world—thus the third guideline is violated, but this is justified

[48]I put the term "miracle" in scare quotes, since this would be a miracle, or violation of physical law, only from the point of view of the actual world, but not in fact inconsistent with the laws at the possible world in question (Lewis, 1979).

in terms of Lewis's ordering of the similarity considerations in that it ensures the second, more significant guideline is respected.[49]

What makes the $C3$ account better able to handle such instances? Not least, it enables us to evaluate conditionals before the time with which they are concerned. Those who favor a division in types of conditional at the level of instances of the conditional, are really only able to guess, beforehand, at whether a particular conditional will turn out to be factual or counterfactual[50] according to their criteria for these distinct kinds of conditional. Take the example,

(124) If the US President were to press the button next year on March 15, then a nuclear holocaust will ensue.

Since this conditional relates to events that will or might take place in the future, on the standard sort of division of conditionals we have no way of saying for certain at this present moment whether this is a counterfactual or if it's in fact making a claim about how things will actually turn out. Unless all conditionals concerning the future are to be impossible to evaluate before the relevant time has passed, then any

[49]It perhaps sounds wildly far fetched that someone might almost start a nuclear war while asleep, but there are many documented cases of sleep walkers driving long distances, eating, having sex, and even killing other people. (See Schenck and Mahowald, 1995, and Bassetti, 2009, for relevant examples.)

[50]Of course, this is only the case for those accounts that do not tie their analysis of different types of conditional wholly to difference in grammatical form. But, though many do confuse indicative mood with a conditional's being concerned with actuality, and likewise for subjunctive mood and counterfactuality, they typically also make some metaphysical commitments regarding these categories—as, for instance, Lewis does in characterizing counterfactual conditionals largely in terms of those in the subjunctive moood, but excluding conditionals (even those subjunctive in mood) with antecedents true at the actual world.

account of the conditional must allow us to evaluate such cases without first being able to determine whether they concern the actual or simply the possible. Until March 15 of next year, Lewis can only guess whether he's correct in applying his theory of counterfactual conditionals to this specific conditional.

On my approach, though, there is no such split at the level of instances of the conditional themselves, and so we may evaluate a conditional such as (124) either factually or counterfactually. The factual interpretation we cannot settle until the time has passed and we know what the antecedent and consequent's truth-values are at the actual world, but we can rest assured that our counterfactual interpretation of (124) is perfectly legitimate even at this stage.

This highlights a further reason to prefer a counterfactual interpretation that does not mandate the inclusion of the actual world in the A-worlds at which the conditional is assessed, when merely the explicit antecedent but not the *ceteris paribus* clause is true, as this too would prevent any legitimate counterfactual interpretation from being made, in a great many instances, before the time with which a particular conditional is concerned. The reason for this is that it could happen that, at the actual world at this future time, the consequent's being false when the antecedent is true turns out to be the result of an exceptional and bizarre coincidence. Since the actual world would be ruled out by the *ceteris paribus* clause on the counterfactual interpretation made earlier—and let's say that the conditional was deemed counterfactually true then—forcing the actual world into the set of A-worlds simply on the basis of the truth of its explicit antecedent, would here change the truth-value of the conditional on a counterfactual interpretation to false.

The upshot would be that we could not make counterfactual assessments of conditionals before knowing the antecedent and consequent's truth-values at the actual world. For conditionals concerned with the very far future, or particular circumstances beyond our abilities to evaluate at the actual world, this would make them impossible to evaluate at all. It's presumably worries of this kind that motivate Goodman's limiting of his discussion of counterfactual conditionals solely to "those in which antecedent and consequent are inalterably false" (1947: 113), as I mentioned at the start of this chapter.

Lewis is one of those in the position of being unable to evaluate any conditional for whose antecedent we do not know the truth-value at the actual world (and, when the antecedent is true at the actual world, the truth-value of the consequent is required as well)—as are all those who endorse centering. Take a conditional such as,

(125) If humans still haven't established new communities on planets far from this galaxy 7.6 billion years from now, our species will be completely wiped out by the expansion of our sun into a red dwarf star.[51]

If C3 were to add a centering constraint[52]—of the sort that required the actual world to be included among the A-worlds when just the explicit antecedent of the relevant conditional is true there (weak centering), or even that the actual world constitutes the only A-world when the explicit antecedent is true there—it would then be unable

[51] Estimation of the time to the expansion of our sun taken from 'The Sun Will Eventually Engulf Earth—Maybe', *Scientific American*, (Appell, 2008).

[52] *See Chapter 2, p.62.*

to assess the truth-value of (125), as are Stalnaker and Lewis's theories, or any that categorize conditionals at the level of their instances[53].

If (125) is taken to be a factual conditional, it cannot be evaluated, since we have no way to know what actually will happen that far in future. If (125) is taken to be a counterfactual conditional, it cannot be evaluated on any theory with a centering condition, since this makes the actual world compulsorily one of the antecedent-worlds (or even the only antecedent-world, on strong centering), and we don't know whether the consequent of (125) will be true or not at the actual world.

A replacement for similarity, and further reasons to reject it

Returning to the issue of what the best criteria are for those determining the variability of a variably strict conditional—the $C3$ account, of course, does not need to appeal to the problematic notion of similarity, nor to any other form of closeness of worlds to the base world. What replaces considerations of this kind is the *ceteris paribus* clause.

How then would $C3$ handle Fine's example, for instance? The conditional at issue is the following,

(126) If Nixon had pressed the button, there would have been a nuclear holocaust.

[53] Again, other than those that tie these types of conditional wholly to grammatical categories.

Though (126) could be interpreted factually, say by someone who was convinced (however erroneously) that Nixon did press the button but, in fact, no nuclear holocaust ensued, for most of us what's really of interest here is the counterfactual interpretation. On the $C3$ approach, to determine the counterfactual truth-value of (126), all of the A-accessible worlds must first be identified. In this case, these are all of those worlds at which, *ceteris paribus*, Nixon pressed the button. The possible worlds considered must be ones at which the antecedent is true, with the caveat of all other things being equal, which would clearly seem to rule out the world Fine suggests, where the button is pressed but something then happens to prevent the missiles launching. So, too, would the world I imagine be inadmissible—all things would very obviously not be equal there, not only if Nixon were to push the button as a result of sleep walking, but also as concerns the very strange (indeed, physically impossible) outcome that the electrical signal from the button press is somehow 'lost' on its way to triggering the nuclear missiles. What the *ceteris paribus* clause does, but similarity and other related notions fail to do, is to effectively accommodate the crucially important context within which our ordinary discourse occurs. Our shared background assumptions as to how the world typically and relevantly is, apropos a given conditional, are an ineliminable part of our effective use and evaluation of conditionals.

Lycan cites another example, on which Lewis's (and also Stalnaker's) approach gives a straightforwardly wrong answer:

> Suppose that due to the peculiar effects of a rare and eccentric compound of kryptonite, Superman is rendered incapable of lifting any object weighing more than 100 pounds and less than 100.1 pounds. He can, however, still lift all objects he could previously lift which weigh 100 pounds or less, or more than 100.5 pounds. Since his exposure, Superman has not had occasion to lift anything weighing over 100.1 pounds, but he has attempted to lift several objects weighing between 100 and 100.1 pounds. After lifting an object weighing exactly 100 pounds, Superman tells Lois Lane, [(127)] 'If that [object] had been [...] heavier, I wouldn't have been able to lift it.' (Originally from George Schumm, cited as a personal communication in Nute, 1980*b*: 70. Not quoted entirely verbatim, since I have cleaned up several small infelicities.)[54] (Lycan, 2001: 51; parenthetical remarks his; bracketed changes and footnote mine)

The problematic conditional that Superman utters is,

(127) If that object had been heavier, I wouldn't have been able to lift it.

On Lewis's analysis, the sphere of worlds at which the object that Superman tries to lift was heavier than 100 pounds will be ones at which it is infinitesimally so, since

[54]It appears that Lycan has introduced an error of his own here, not present in Nute, 1980, with his reference to the weight "100.5" instead of "100.1", since it's unclear whether Superman would or would not be able to lift objects that weigh 100.5 pounds or less, but also 100.1 pounds or more. But, as it happens, this confusion doesn't affect the example.

these are the worlds most similar to the actual world, where the object is 100 pounds. As Lewis rejects the limit assumption—the assumption that there is a finitely closest world or worlds to the actual world—he is able to account for the possibility that there are infinite successively more similar worlds to the base world at which the object Superman tries to lift is over 100 pounds.[55]

The trouble with these worlds being the ones most similar to the base world, at which Superman does lift the 100-pound object, is that this makes (127) true on Lewis's theory—yet this conditional appears intuitively false. Superman says that he couldn't have lifted the object if it had been heavier, but if it had been any weight greater than 100.1 pounds (or 100.5, just to avoid the confusion in Lycan's version of the example), he could indeed have lifted it. The problem is, again, that similarity (even comparative overall similarity) isn't always what we're concerned with. Similarity, as a measure of which worlds to use in the assessment of a given conditional, isn't sufficiently flexible to account for all the different factors that may be relevant in a particular case. The most similar worlds at which the object is heavier than 100 pounds are ones at which Superman cannot lift it, but in this case what actually matters is whether the A-worlds are ones at which the object is simply

[55]This example turns out to be doubly problematic for Stalnaker's approach since his theory not only assigns an intuitively incorrect truth-value to (127), but, due to his acceptance of a limit assumption, is also faced with the fact that there is no uniquely closest world at which the object weighs more than 100 pounds. Whichever world it choses as supposedly closest or most similar, there will always be another between that and the base world at which the object is just that much fractionally lighter though still over 100 pounds—provided that there is no minimum quantum of weight, that is.

I say the *possibility* that there are infinite successively more similar worlds, to allow that there may in fact be an indivisible minimum quantum of weight, which would mean that there would, then, be a finitely closest world at which the object weighed over 100 pounds. Regardless, whether this is so or not doesn't alter the difficulty that this example poses for Lewis.

heavier (no matter by how much) and that there is no other unexpected circumstance interfering with Superman's lifting abilities or the like.

It should by now be clear how the $C3$ account can explain why (127) is intuitively false. This conditional is non-truth-apt on a factual interpretation (not least because, I'm afraid, Superman doesn't exist), and of no real relevance, and false on a counterfactual interpretation. Certainly, some of those worlds at which it is *ceteris paribus* the case that the the object is heavier than 100 pounds—namely, those at which it is also less than 100.1 pounds—are worlds at which the consequent is true: Superman cannot lift the object. But, all the other worlds at which the object is heavier than 100 pounds, and sufficiently so to avoid that weird quirk of the kryptonite, are worlds at which the consequent is false: Superman *can* lift the object. And since, on my theory, the consequent must be true at *all* A-worlds and not merely some in order for the counterfactual interpretation to be true, (127) is counterfactually false. Given that this conditional is also not factually truth-apt, it is false also on the combined interpretation, and so may be said to be simply false.

Finally, I would like to consider an objection to Lewis's use of comparative similarity as the means to order the closeness of spheres of worlds around the base world—that is, for our purposes in ordinary language, the actual world—from Edgington. She begins by quoting Lewis (Edgington, 1995: 254–255),

> "Though similarities and differences in laws have some tendency to outweigh difference or similarities in particular fact, I do not think they invariably do so", says Lewis (1973, p. 75). His reason is as follows. A tree blows over,

> destroying the roof of a house. [...] Consider
>
> > If the tree hadn't blown over, the roof would be intact.

If we were to prioritize similarity of laws absolutely over that of particular fact, we would have to suppose that (assuming the laws to be essentially deterministic, at least as regards macro-level objects) any world at which the tree didn't blow over must be radically different from the actual world. This is because the tree not blowing over would require whatever caused it to blow over to be different, which in turn would require its cause to be different, and so forth back to the start of the universe. And, of course, the differences required at the start of universe would affect not just the tree at this present time, but would also propagate outwards so as to change our world dramatically. Obviously, this would be little use to us in assessing the truth-value of conditionals, and Lewis must therefore opt here for a small, localized violation of law as preferable to a massive difference in matters of particular fact. (Edgington, 1995)

Hence, Lewis's further remark, "'Laws are very important, but great masses of particular fact count for something too ... . I therefore proceed on the assumption that the preeminence of laws ... is a matter of degree' (1973, p. 75)." (Edgington, 1995: 255; ellipses Edgington's) This move opens Lewis to counterexamples of the following kind:

> Suppose Hitler died in infancy. Then things would have been quite different in the 1930s and 1940s. But consider the world most similar to the actual world in which Hitler died in infancy. [...] That may be one in which some other child grew up to occupy a virtually identical Hitler-like role. Not that that would have happened mind you. Imagine two films in which Hitler died in infancy. One of them has a non-Hitler doing all the kinds of things Hitler did. It strikes you as remarkably like the actual world, almost indistinguishable from the newsreels. The other strikes you as a very plausible account of how the world would have been without Hitler—rather different. Judgements of similarity go one way, judgements about counterfactuals, the other. (Edgington, 1995: 255)

But it would seem that this is a worry already dispatched by Lewis's similarity considerations, discussed above. Last of the four is that, "It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly." (Lewis, 1979: 472)

Edgington's insight, though, is that approximate similarity of particular fact does in fact matter to many conditionals. The following is an example of just such a conditional (in italics), which I actually thought to myself recently while listening to a news item about the 1994 Rwandan genocide,

(128) I wish I'd realised at the time, back when I was a teenager, just how bad this event was. Then again, *if I'd realised back then how bad the situation was, it wouldn't have saved any lives.*

Here, it is precisely the preservation of approximate similarity of particular fact that makes this conditional true. The small 'miracle', the change required to make the antecedent true at other possible worlds, this makes the matters of particular fact at these worlds only approximately the same as those at the actual world rather than identical. But it is this very approximation at these worlds—that is, the genocide in question being just as bad as at the actual world—that secures the truth of the consequent, and thus of (128) as a whole.

The sort of case that prompts Lewis to disregard approximate similarity is that raised by Pavel Tichy (Edgington, 1995: 257–258):

> [W]hen Fred goes out, if the weather is bad, he always wears his hat; if the weather is fine, it's a random 50-50 whether he wears his hat. In fact the weather is bad, and he wears his hat. Consider "If the weather had been fine, he would have worn his hat." The fine-weather world in which he does so is more like the actual world than the fine-weather world in which he does not, but the counterfactual is not clearly true.

The diagnosis Edgington (1995) gives of the difference between the sort of case of which (128) is an example and the 'Fred's hat' type of case, is that in the former there is no causal relation involved; in fact, what's being implicated is a relation between the antecedent and consequent such that it follows from the antecedent that the consequent is unaffected by it—in effect, a denial of a causal relation between the two. But, as she points out, in the latter sort of instance the antecedent is

causally relevant to the consequent; whether the weather is fine or inclement directly affects the probability of Fred's wearing his hat (or not). Edgington observes that Lewis could try to incorporate considerations of causality in some way into his notion of comparative similarity, but that this is precluded by his using his theory of counterfactual conditionals to explicate causality.

How does my account deal with these disparate types of conditional? Let's take (128) first. The factual interpretation is irrelevant to me in this case, knowing as I do that the antecedent is false at the actual world and so that the conditional as a whole is not factually truth-apt. On the counterfactual interpretation, we must look at all the worlds at which the antecedent is *ceteris paribus* true; worlds at which I did properly realize in 1994 just how bad the situation in Rwanda was, without other relevant facts having changed. And, at all of these worlds, it seems clearly the case that, given the pertinent facts about my age, position, and personality then, nothing resulting from that realization at that time would have altered anything materially in Rwanda. Therefore, (128) is counterfactually true.

Turning to the example of Fred and his hat: The conditional in question is,

(129) If the weather had been fine, Fred would have worn his hat.

Not having any idea about the facts of Fred's life and doings, and taking from the subjunctive mood of this English rendition of the conditional that it is the case that the weather was not actually fine on the day in question, we can ignore the factual interpretation. The counterfactual interpretation asks us to consider all the worlds

at which, *ceteris paribus*, the weather was fine. Holding fixed everything other than what's necessary in order for the weather to be fine instead of rainy, we can then look at whether all of these A-worlds are ones at which the consequent is true. Since it's given that Fred only wears his hat half the time when it's sunny, there's no reason to think that he would certainly have worn it on this particular sunny day, and so (129) comes out false on a counterfactual interpretation.[56]

We, thus, have a strong basis for rejecting similarity as an adequate measure of what we are genuinely concerned with in assessing conditionals. As I've indicated, however, the $C3$ account is able to handle the problem cases discussed—giving us good reason to prefer this theory to those of Stalnaker and Lewis.

[56]The conditional could be altered to be counterfactually true, by changing the consequent to something like, "There's a 50% chance that Fred would have worn his hat", or "Fred would have been just as likely to wear his hat as not". (*See Chapter 5, p.236, for discussion of how probability conditionals are to be assessed.*)

Chapter 5: **Challenges for the $C3$ conditional**

There are various challenges for the factual and counterfactual interpretations of the $C3$ account of the conditional that need to be addressed in order for this theory to be a plausible one. In this chapter I respond to a number of these possible challenges to this approach, and indicate various ways in which my account could potentially be amended to accommodate additional sorts of conditionals.

Embedded Conditionals, and Compound Antecedents and Consequents

There are a number of different types of compound conditional constructions used in ordinary language. In this section I examine a series of examples to show how the $C3$ account is able to handle embedded conditionals, or those with a compound

antecedent or consequent.

Embedded conditionals

As we've seen in earlier chapters, embedded conditionals can cause significant problems for certain theories—particularly those that deny truth conditions to the conditional. On my approach, however, it's possible to deal with a range of different embeddings of conditionals. Take, for instance, a conditional of the form $(\varphi \Rrightarrow \lambda) \Rrightarrow \theta$ (where each instance of '$\Rrightarrow$' may, respectively, be replaced by either '$\rightharpoonup$' or '$\rightharpoondown$'):

(130) If, if Phumzile tutors him, Brian passes; then we still won't know whether he could have passed without extra lessons.

The factual interpretation of (130) is quite straightforward. This conditional is true if it is true at the actual world that Phumzile tutors Brian, Brian passes, and we don't know whether he could have passed without extra lessons. It is false if it is true at the actual world that Phumzile tutors Brian, Brian passes, but we do know whether he could have passed without extra lessons. And, it is non-truth-apt otherwise.

On a counterfactual interpretation, we need first to determine the A-worlds at which, *ceteris paribus*, Brian passes—let's called them A_B-worlds, to prevent any confusion. These worlds will be those at which Brian passes and it's not the case that he doesn't have extra lessons or anything else untoward occurs. Out of all of these A_B-worlds, we then have to identify the A_P-worlds, those at which, *ceteris paribus*, Phumzile tutors Brian. The conditional, (130), as whole will then be true

on a counterfactual interpretation if, at each of the A_P-worlds, it is not the case that we know whether Brian could have passed without extra lessons.

What about a mixed interpertation of (130); one, say, of the form $(\varphi \mathrel{\rightarrow\!\!\!\!\!\!\rightarrow} \lambda) \mathrel{\rightarrow\!\!\!\!\!\!\rightarrow} \theta$? The antecedent-embedded conditional's being true on a counterfactual interpretation is nonetheless the truth of such an interpretation at the actual world (much as the truth of its being possible that I might have grown up to be two inches shorter than I in fact am is the truth of this possibility at this, the actual, world). Thus, the truth, at the actual world, of this counterfactual interpretation of the antecedent-embedded conditional[57] constitutes the factual truth of the antecedent. Then, depending on whether we do or do not know the actual-world truth of whether Brian passed without extra lessons, we may determine the main conditional to be factually true or false, respectively.

All this said, this would be a decidedly odd reading of (130); but one we could encourage by phrasing the conditional thusly: "If, if Phumzile had tutored him, Brian would have passed; then we still don't know whether he passed without extra lessons." Of course, one could still reject that the most salient interpretation should be of the form $(\varphi \mathrel{\rightarrow\!\!\!\!\!\!\rightarrow} \lambda) \mathrel{\rightarrow\!\!\!\!\!\!\rightarrow} \theta$. And, indeed, it's very hard to see that such an interpretation could be relevant other than in the most extraordinary circumstances.

Similarly, it's difficult to see that we could have occasion to interpret (130) as taking the form $(\varphi \mathrel{\rightarrow\!\!\!\!\!\!\rightarrow} \lambda) \mathrel{\rightarrow\!\!\!\!\!\!\rightarrow} \theta$—in other words, interpreting the conditional in the manner suggested by the following wording, "If, if Phumzile tutored him,

[57]To avoid confusion, it should be borne in mind that, in general, references to the factual or counterfactual truth of a conditional are of course references to the factual or counterfactual truth of these conditionals *at the actual world*, since it is truth at this world that is of interest to us in the overwhelming majority of ordinary instances.

Brian passed; then we still wouldn't have known whether he could have passed without extra lessons." Again, notwithstanding this interpretation being likely quite unneeded in everyday life, it is available to us.

Let's turn now to a case in which the embedded conditional forms the consequent of the main conditional—that is, one of the form $\varphi \Rightarrow (\lambda \Rightarrow \theta)$ (where, again, each instance of '$\Rightarrow$' may, respectively, be replaced by either '$\rightharpoonup$' or '$\rightharpoondown$'):

(131) If the chair is strong enough to hold your weight, then, if you stand on it, you'll be able to change the lightbulb.

Firstly, let's consider this compound conditional, (131), on a factual interpretation. The consequent of the conditional whose antecedent is, "The chair is strong enough to hold your weight", has another, embedded, conditional as its consequent, so we need to tackle this first. The embedded conditional, "If you stand on it, you'll be able to change the lightbulb", is assessed as usual—it's factually true if both the antecedent and consequent are true at the actual world, false if the antecedent is true but the consequent false, and non-truth-apt otherwise. The main conditional is then, in turn, also assessed as normal, with the truth-value of the embedded conditional as a whole providing the truth-value of its consequent. Should the embedded conditional prove not to be truth-apt on a factual interpretation, then so too will the main conditional be, since it must be the case both that the antecedent is true at the actual world and that the consequent is either true or false at the actual world, in order for a

conditional to be ascribed a factual truth-value (as opposed to being judged non-truth-apt).

The counterfactual interpretation is a little more challenging, as the *ceteris paribus* clause of the embedded conditional is affected by its being embedded within another conditional. We have to determine the truth-value of the embedded conditional as a whole before we can assess that of the main conditional, but interpreting the former requires that we identify the A-worlds of the latter first, since the *ceteris paribus* clause of the embedded conditional is applicable to only these A-worlds.

What, then, are the A-worlds of the main conditional? These will be all those worlds at which the chair is strong enough to hold your weight and it's not the case that the legs have been half sawed through, someone's maliciously replaced the seat of the chair with a very convincing photographic reproduction printed on cardboard, or the like. Now, to assess the embedded conditional, we need to look at those of the A_m-worlds (the A-worlds of the main conditional) at which, *ceteris paribus*, you stand on the chair. The A_e-worlds (the A-worlds of the embedded conditional) will be those such that you stand on the chair but don't first move it into another room from the lightbulb or immediate faint, and so on. And if, at all of these A_e-worlds, you are able to change the lightbulb, then (131)'s embedded conditional is counterfactually true. Since all of the A_e-worlds are A_m-worlds, if the consequent of the embedded conditional is true at all of the former worlds, then the embedded conditional must be true at all of the latter worlds—making the main conditional, (131), as a whole, counterfactually true.

This particular example we can more readily make sense of in terms of a mixed

interpretation, say of the form $\varphi \ \rightarrow\!\!\!\!\!\sim \ (\lambda \ \rightarrow\!\!\!\!\!\sim \ \theta)$, than in the case of (130). What the assessment of such an interpretation would require is to determine whether the chair is in fact strong enough to hold your weight (by testing the actual load-bearing capacity of the chair in whatever way) and the truth-value of the counterfactual interpretation of the consequent-embedded conditional; namely, whether it is the case that at every A-world at which you stand on the chair you are able to change the lightbulb. The actual-world truth-value of this counterfactual conditional (which constitutes the truth-value of the consequent), in conjunction with that of the antecedent, would then allow us to assess the truth-value of the factual interpretation of the main conditional. Similarly, we would be able, *mutatis mutandi*, to give the truth-value of (131) were it to be of the form $\varphi \ \rightarrow\!\!\!\!\!\sim \ (\lambda \ \rightarrow\!\!\!\!\!\sim \ \theta)$.

Embedded conditionals may evidently be quite simply dealt with on a $C3$ account. Even more straightforward are those conditionals with conjunctive antecedents or consequents. In such cases, the conditional is interpreted as normal, either factually or counterfactually—just with the antecedent or consequent being made up of a conjunction, the truth-value of which is determined first and the conjunction then treated as a unit.

Disjunctive Antecedents

Disjunctive, rather than conjunctive, antecedents appear somewhat more complicated, however. Willian Star gives the following example of an intuitively valid inference, involving a conditional with a disjunctive antecedent. Despite its seeming intuitive validity, though, the inference this argument relies on is not endorsed

by certain well-known theories of the conditional (such as Stalnaker's, for instance). This inference is an instance of the simplification of disjunctive antecedents (Starr, 2014: 1048):

(132) If Bob had danced or Sarah had s[u]ng, Andy would have cried[.]

(133) So, if Bob had danced, Andy would have cried, and if Sarah had s[u]ng, Andy would have cried[.]

The intuitive validity of the inference from (132) to (133) is supported on the $C3$ account; at least in most ordinary contexts. More precisely, such inferences are not formally valid, but they are nonetheless truth-preserving in certain circumstances.

On a factual interpretation, if (132) is true, that can only be because both its antecedent and consequent are true. Then, in turn, given the truth of (132)'s antecedent and consequent, it must be the case that the antecedent of at least one of the conditionals conjoined in (133), and the consequents of both (since these are identical), all are true. But, though (133) could conceivably be true on a factual interpretation—should Bob's dancing and Sarah's singing both have reduced him to tears—this is by no means necessitated by the factual truth of (132).

However, the truth of (132) on a counterfactual interpretation would license the inference to (133). For the inference to hold, it cannot be simply that (132) is true on a counterfactual interpretation and, incidentally, (133) is also. The clear implication

here is that the truth of the conclusion, (133), is a result of or reliant on the truth of (132), not merely that they are coincidentally simultaneously true. What I take this to mean is that all of the A-worlds of (132) at which the explicit antecedent of the first conditional of (133) is true, must also be A-worlds of this latter conditional; and *mutatis mutandi* for the second conditional of (133).

This might seem an unwarranted claim—couldn't it be the case that Sarah is someone who cannot sing a note, and who anyway could never move Andy to either tears of joy or anguish with any attempt at song? Wouldn't this make the conditional in the second conjunct of (133) false on a counterfactual interpretation, since at all (or at least the vast majority) of those worlds at which Sarah did sing Andy would not be moved to cry? That, it seems, would make (133) counterfactually false, but note that this in turn would make (132) come out counterfactually false—for there must then be worlds among the A-worlds for (132) at which Sarah does sing, but Bob happens not to dance, and at these worlds the consequent of (132) would fail to be true; making the conditional as a whole counterfactually false.

But one may still be skeptical: Consider, for example, the worlds at which, though Bob does dance, Andy doesn't see Bob dancing and so does not cry in virtue of said dancing—and, at these same worlds, Sarah does sing, and Bob hears her and is moved to tears by her singing. Such worlds are ones at which both the antecedent and consequent of (132) are true, but they are not worlds at which (133) would be true, since they would render the antecedent in the first conjunction true, but its consequent false.

However, were such worlds not excluded by the *ceteris paribus* clause of (132),

then it surely must also include those worlds at which Bob dances unseen by Andy, who consequently does not cry at this dancing, and at which Sarah also does not sing—and the inclusion of these worlds in the set of (132)'s A-worlds would make the consequent of (132) false, and so (132) as a whole counterfactually false. More generally, in any contexts in which the *ceteris paribus* of (132) would allow worlds that would make the antecedent of the conditional in either conjunct of (133) true but its respective consequent false, it would have to make the consequent of (132) false too.

Finally, there may be some concern over an example like this, but set in a context in which it's widely know to be the case that Bob never misses a chance to dance, and reliably brings Andy to tears when he does; while Sarah hates to sing and has no talent for it. This would seem to be an instance where (132) would be counterfactually true, but (133) counterfactually false on my analysis. But, on reflection, this is quite unproblematic. Were one, in such a context, to hear someone utter (132), one would hardly take this to license (133), which would be intuitively false. There would be no question of why the intuitive truth of (132) failed to allow one to infer (133) from it, given the intuitive falsity of (133) in that context.

Thus, the inference from (132) to (133) is not a formally valid one, but arguments of this form are truth-preserving under certain conditions. Take the following argument schema, in which φ and λ stand for any disjuncts and θ any consequent:

(134) $(\varphi \vee \lambda) \rightharpoonup \theta$

(135) $(\varphi \rightharpoonup \theta) \wedge (\lambda \rightharpoonup \theta)$

Inferences of this form, from (134) to (135), are truth-preserving only when all of φ, λ, and θ are true at the actual world. In the case of the counterfactual interpretation,

(136) $(\varphi \vee \lambda) \rightarrow \theta$

(137) $(\varphi \rightarrow \theta) \wedge (\lambda \rightarrow \theta)$

the inference is truth-preserving only when the set of worlds at which φ is *ceteris paribus* true, the set of worlds at which λ is *ceteris paribus* true, and the set of words at which both φ and λ together are *ceteris paribus* true—when, at each member of all of these three sets of worlds, θ is true.

As Lycan (2001) points out, the simplification of disjunctive antecedents is an instance of antecedent strengthening, which is invalid on most analyses of the counterfactual. It is an instance of strengthening in that, in the original conditional (134), the truth of the consequent, θ, is claimed to follow from the truth of the disjunctive antecedent, $(\varphi \vee \lambda)$; whereas, in the simplification, the truth of that same consequent is claimed to follow both from one of those original disjuncts, on its own, *and* from the other, on its own—which is evidently a stronger claim. And so, the real question is very much that of what accounts for the simplification of disjunctive antecedents'

apparent truth-preservation in particular instances (in specific contexts), rather than what causes it to fail.

Let's consider another instance of conditional with a disjunctive consequent:

(138) If Lerato is practising the violin or Kerry put on that recording of her playing, that would explain the music I can hear in the background.

On a factual interpretation, this conditional is true if either Lerato is practising the violin or Kerry put on the recording of her playing (or both), and that explains the music I can hear in the background.

On a counterfactual interpretation, the relevant A-worlds will be those at which it is the case that, *ceteris paribus*, either or both of the disjuncts in the antecedent is true. The *ceteris paribus* clause presumably excludes those worlds at which either or both of those disjuncts are true but I can't hear either the playing or the recording, and so on. And, if at all of these A-worlds the music I can hear in the background is explained by either the playing or the recording (or both), then (138) is true on the counterfactual interpretation.

Let's turn now to an example of a conditional with a disjunctive consequent:

(139) If Bailey wins the 'Cutest Dog' contest, then Albie will have to get extra treats to make it up to him or he'll be jealous of all the attention Bailey gets.

This conditional is just as straightforwardly evaluable as (138). All that's required on the factual interpretation is that the antecedent and at least one of the disjuncts of the consequent are true; and on the counterfactual interpretation that either or both of the consequent's disjuncts is true at all of the relevant A-worlds.

Seemingly Unrelated Antecedents and Consequents

One seemingly difficulty for my account are those factual and counterfactual interpretations of the conditional that $C3$ judges to be true, but whose respective antecedents and consequents appear to be wholly unrelated to one another. The truth of such conditionals we ordinarily reject, because it seems, intuitively, that the truth-value of the consequent should rely in some way on that of the antecedent, which cannot be the case if the truth of one is completely independent of the truth of the other. This is, incidentally, one of the so-called paradoxes of the strict conditional (Priest, 2008) that still constitutes a challenge to the counterfactual interpretation of the $C3$ conditional, despite its being variably strict.

Take the claim that,

(140) If the sloth is a mammal, then seven is a prime number.

This conditional is true on both a factual and a counterfactual interpretation, according to $C3$, since it is necessarily the case that seven is a prime number, though

such a conclusion is hardly intuitively appealing. That seven is prime seems utterly unrelated to whether or not the sloth is a mammal.

How might this be accounted for? I would argue that cases such as this may be explained by appeal to pragmatic considerations that render (140) unassertable, rather than semantic ones that alter its truth-value[58]. This conditional is factually true, according to $C3$, because, at the actual world, it is true both that the sloth is a mammal and that seven is a prime number; and counterfactually true, because at every world A-accessible from the actual world (every A-world)—namely, all those at which sloths are mammals (whether or not this is necessarily or contingently so is immaterial to the example)—is also a world at which seven is prime (as, of course, this latter is the case at every possible world). The trouble is that, while I am claiming that (140) is quite true, this is certainly not our intuitive, pre-theoretical judgment as ordinary English speakers. What reason, then, is there to believe that ordinary intuitions about such conditionals may be explained by their unassertablity in everyday conversation, rather than to think that they are simply false, as seems to be more immediately plausible?

To answer this question it's necessary to identify what it is that gives rise to the offending intuition yet, following $C3$, fails to bring the conditional's truth-value into line with this intuition. The problem here seems to be that the way in which $C3$ determines truth-values does not require that there be any clear connection between

[58]Or, at least, it may be explained without appeal to semantic considerations that motivate a truth-value for the conditional at odds with that given by $C3$. I make this distinction because I would like to remain agnostic on certain debates over the location of relevant phenomena—whether they fall within the pragmatic or semantic arena. Which side the consensus comes down on is independent of the explanation I offer below.

the truth of the antecedent and the truth of the consequent. Provided that the latter is true at all worlds at which the former is, the conditional as a whole is deemed true. Whether or not there is any sort of dependence (or other) relation between the two—or at least the appearance, to ordinary language users, of such a relation—is no part of this determination. Whatever it may be about the nature of the universe that gives rise to the fact of seven being necessarily prime, it is certainly not evident to us how this is related to the matter of sloths being mammals, whether or not such a relation in fact exists.

That there be some kind of tie between antecedent and consequent, does however seem to be crucial to our intuitive truth-value judgments. But, I would argue that the notion of the consequent in some sense 'following from' the antecedent is a generalized conversational implicature of the 'If..., then...' locution (and other conditional constructions), as I discussed in Chapter 2[59]. If this seems implausible, consider terms such as 'but'. While it is uncontentious that the logic of 'but' is simply that of 'and' (namely, that both conjuncts must to be true for the statement as a whole to be true), it is still intuitively unpalatable that a statement such as,

(141) Disneyland is little more than a horribly over-commercialized attempt to empty one's wallet, but I always love going there.

is in fact equivalent in logical form simply to the conjunction of (142) and (143):

[59] *See p.68.*

(142) Disneyland is little more than a horribly over-commercialized attempt to empty one's wallet.

(143) I always love going to Disneyland.

that is,

(144) Disneyland is little more than a horribly over-commercialized attempt to empty one's wallet and I always love going there.

The strangeness, to an ordinary English speaker, of the mere conjunction of (142) and (143)—of (144)—being offered as the logical equivalent of (141), comes down to the conventional implicature carried by 'but'. The use of 'but' rather than 'and', implicates that the speaker takes the two conjuncts to appear incompatible, although they are both in fact true. (Grice, 1975, 1989) Similarly, the use of the conditional in ordinary discourse—as typified by, but not limited to, the 'If..., then...' construction—implicates that the consequent does, or at least could, follow from the antecedent in some sense evident to the speaker (and, presumably, that they take to be evident to their audience), though this is not actually part of the semantics of the expression.[60].

[60]Of course the category of conventional implicature is a controversial one, and it may be that what is supposedly implicated is in fact part of the semantics of the expression concerned. If the

In the interest of easing any lingering doubts, I'd like to consider the following even more striking example of intuitive unpalatableness in a statement whose logical structure is hardly controversial. First, take this very ordinary statement,

(145) I had an avocado sandwich for lunch, and I mowed the lawn for most of the afternoon.

I think it's safe to say that (145) is more dull than contentious. But, consider the following, which has precisely the same logical structure,

(146) I had an avocado sandwich for lunch, but I mowed the lawn for most of the afternoon.

Again, as in the Disney example (144) above, (146) seems odd and highly unlikely to be uttered in ordinary conversation, since our normal use of the conventional implicature of 'but' is not being respected. In (146), what causes this dissonance is that 'but' carries with it the conventional implication of some degree of contrast or clash between the two conjuncts—yet there is no discernible conflict of any kind between first eating an avocado sandwich and then mowing the lawn. Nonetheless,

---

explanation is a semantic one, though, it seems that expressions such as those containing 'but' or taking the form of a conditional (among others) must convey multiple propositions, but in a way that allows for the salience of one to obscure that of the other. (See Bach (1999), and Potts (2007), for such multiple-propositions accounts)

because 'and' and 'but' are logically equivalent, (146) shares its logical structure with (145).

Doubtless, one could probably come up with some sort particular context in which (146) was not infelicitous—let's imagine that I find avocado a great soporific, say, and typically have a long nap after eating it. In that sort of scenario I might find myself uttering something like (146) to assure any doubters that my lunch hadn't led me to shirk my mowing duty. But, precisely the same is true of the sort of conditionals at issue here.

Consider, (140): Even this conditional has a possible (however unlikely and convoluted) context in which it would not be infelicitous to utter it. Imagine a scenario in which children are being taught about prime numbers and helped to memorize the first five primes. Let's say, also, that their enthusiastic but not especially gifted teacher has developed an exercise of doubtful pedagogical efficacy—one in which each number from one to 12 is paired with a different animal. Since the children have just learned about mammals and reptiles, and the teacher wants them to do some work to figure which of the numbers are prime, he has paired each of the five primes with a picture of a mammal and each of the remaining numbers with the picture of a reptile. After informing the children of these pairings and tasking them with identifying the primes, he's asked by one of the children whether seven is prime; to which he responds with the, in this context entirely felicitous, conditional, "If the sloth is a mammal, then seven is a prime number."

The question, though, is not whether these sorts of conditionals, or sentences such as (146) *may* on occasion by intuitively acceptable, but whether they are ever

perceived as *not* being felicitous. In fact, a variety of instances of this kind (including conjunctions, conditionals, and others) may be rejected as unassertable and even be judged intuitively false.

As a final example, take this instance of a perfectly true statement,

(147) Shoes are designed to be worn on the feet, but the Empire State Building is in New York City.

Like (140), (147) is something that someone would never (or only in the oddest of circumstances) have occasion to utter in ordinary discourse. The ordinary speaker is very likely to consider this false, or be confused as to its truth-value or whether it is even truth-apt. Regardless, its intuitive unpalatability does nothing to undermine our confidence that it indeed has the logical form of any straightforward conjunction—however unassertable it may be. So, too, should we be similarly sanguine with regards to instances such as (140). They may not be paragons of ordinary discourse, but the blame for this cannot be laid at the door of their logical structure; more specifically, their intuitive oddness is no shortcoming of the $C3$ reading of the conditional.

Transitivity

Another puzzling feature of the conditional as it's used in ordinary discourse is that there are instances in which transitivity appears to hold, but others in which transi-

tivity very evidently fails. It seems highly implausible to suppose that the conditional follows some kind of variably applicable principle of transitivity that sometimes holds, sometimes not. How then can these discrepancies be explained? In the case of the $C3$ conditional, transitivity does hold for the factual interpretation of the conditional. Transitivity does not hold for the counterfactual interpretation of the conditional in every instance; however, there are nonetheless certain cases in which transitivity is truth-preserving.

First, let's consider a case where transitivity does apply, given the counterfactual interpretation of the following conditionals:

(148) If you were to raise enough money to build the proposed new children's shelter, then 100 more children in Cape Town would have safe accommodation.

(149) If 100 more children in Cape Town had safe accommodation, then the problem of homeless children would be somewhat lessened.

Therefore, by transitivity,

(150) If you were to raise enough money to build the proposed new children's shelter, then the problem of homeless children would be somewhat lessened.

Here, the inference from (148) and (149), by transitivity, to (150) appears to be perfectly valid. It is indeed true that if if the conditionals (148) and (149) are true, then so too must (150) be. Similarly, transitivity seems to hold for the following series of conditionals too,

(151) If skyscrapers were made of chocolate, they would be soft enough that one could break off pieces of them to eat.

(152) If skyscrapers were soft enough that one could break off pieces of them to eat, they would have grave structural problems.

Therefore, by transitivity,

(153) If skyscrapers were made of chocolate, they would have grave structural problems.

The inference from (151) and (152), by transitivity, to (153) seems just as acceptable as that immediately above it. So why, then, does transitivity appear to fail in cases such as that below, and what non-*ad hoc* way, if any, is there to distinguish these different sorts of instances from one another?

Consider Lewis's (1973b) example:

(154) If J. Edgar Hoover had been born in Russia, then he would have been a Communist.

(155) If J. Edgar Hoover had been a Communist, then he would have been a traitor.

Therefore, by transitivity,

(156) If J. Edgar Hoover had been born in Russia, then he would have been a traitor.

In the case of the two earlier series of conditionals, the ordinary, intuitive judgment is that these are truth-preserving applications of transitivity. If (148) and (149) are both counterfactually true, then it would certainly seem to hold that (150) must be true on a counterfactual interpretation; and similarly that the counterfactual truth of (153) follows from that of (151) and (152). However, in the case of the 'Hoover' example, just given, even if both (154) and (155) are true on a counterfactual interpretation, (156) does *not* intuitively follow. On the contrary, this last conditional seems obviously counterfactually false, and transitivity appears to have failed here.

What, then, accounts for the difference between this last example and the two preceding it, and how can the $C3$ theory of the conditional account for this fact that transitivity appears to hold in some cases involving a counterfactual interpretation of the conditionals concerned, yet not in others?

As I mentioned above, it's important to note that transitivity will always be truth-preserving on factual interpretations of the conditional. If it is the case that, *in actual fact*, both (154) and (155) are true, then (156) would have to be true on a factual interpretation. More generally, any argument of the following form will be truth-preserving—where φ, ψ, and λ stand for any antecedent and/or consequent; and (157) and (158) are the premises and (159) the conclusion,

(157) $\varphi \rightharpoonup \psi$

(158) $\psi \rightharpoonup \lambda$

(159) $\varphi \rightharpoonup \lambda$

Why the ordinary person doesn't interpret (154), (155), and (156) factually, is because it is common knowledge that none of their antecedents or consequents is true at the actual world. Moreover, there is good reason to think that, however mistaken we might turn out to be about what is in fact the case, it isn't possible for (154) and (155) both to be true at the same world. For these reasons, it makes far more sense

to ignore the potential factual interpretations and consider just the counterfactual ones.

In contrast, with the example of the conditionals (148), (149), and (150), it seems perfectly possible that both of (148) and (149) might be simultaneously true at the actual world, and hence that (150) be factually true also.

Now let's consider what would make (148) and (149), respectively, true on a counterfactual interpretation. The former, (148), may be deemed counterfactually true because every world at which it is *ceteris paribus* the case that you raise enough money to build the children's shelter, is a world at which it is also true that 100 more children will have safe accommodation. The latter, (149), is true because every world at which it is, *ceteris paribus*, the case that 100 more children have safe accommodation, is a world at which it is also true that the problem of homeless children will be somewhat lessened.

What is important here, is that the set of all worlds at which *ceteris paribus* the antecedent of the first premise, (148), is true (all of which are worlds at which its consequent too is true) is a sub-set of the set of all worlds at which the antecedent of (149) is true (and the consequent of this conditional is also true at each of these worlds). It is this that allows for the seeming success of transitivity here. In this case, the fact that the consequent of (148) and the antecedent are linguistically alike, and that they seem to share truth conditions, is neither necessary nor sufficient for the relevant inference to go through—for the reasons detailed below.

That transitivity holds in the second example can be explained in just the same way. The set of worlds at which the antecedent of (151) is, *ceteris paribus*, true (that

is, at which skyscrapers are made of chocolate), is a sub-set of the worlds at which the more general antecedent of (152), and consequent of (151), is true (that is, at which skyscrapers are soft enough to break pieces off to eat). And, since any consequent true at all of these antecedent-of-(152)-worlds must be true at any sub-set of such worlds, the truth of (151) and (152) ensures that the consequent of (152) is true at all antecedent-of-(151)-worlds—and the truth of (153) follows. This accounts for the preservation of truth by means of transitivity in this case.

In the third example, though, transitivity seems to fail. This is because that set of all worlds at which the antecedent of (154) is, *ceteris paribus*, true (let's call this set α)—at all of which the consequent of this conditional also is true—is *not* a sub-set of the set of all those at which the antecedent of (155) is *ceteris paribus* true (let's call this set β) and *its* consequent true. Now, this might at first seem counter-intuitive. But, despite any apparent plausibility of α's being a sub-set of β—and despite the linguistic similarity of the consequent of (154) and the antecedent of (155), and their perhaps seeming on the face of it to have the same truth conditions—this is not the case, and so the inference in this example fails to be truth-preserving.

And why is it the case that the set of worlds, at which the truth-value of the first conditional of this example is assessed, is not a sub-set of the set of worlds at which the truth-value of the second is assessed? The answer is that those worlds at which the antecedent of (155) is *ceteris paribus* true, are worlds at which the *ceteris paribus* clause that determines which antecedent worlds are relevant to the assessment of (154) is violated. Worlds at which, all else being equal, J. Edgar Hoover was a Communist are not among those worlds at which, *all else being equal*,

he was born in Russia—for the simple reason that this would require changing a crucial and far-reaching fact about his life, his nationality, and without anything in the statement of the antecedent to warrant this. And, indeed, any reasonably attempt to make good sense of an ordinary speaker's use of a conditional such as (155), as it would likely be used in everyday discourse, would seem to require that this possibility (Hoover's being born in Russia) be entirely excluded. Thus, (156) does not follow, notwithstanding the truth of (154) and (155).

The general principle underpinning the above example can be explained as follows: Take the following argument schema, in which φ, ψ, and λ stand for any antecedent and/or consequent; with (160) and (161) being the premises and (162) the conclusion,

(160) $\varphi \rightharpoonup \psi$

(161) $\psi \rightharpoonup \lambda$

(162) $\varphi \rightharpoonup \lambda$

Transitivity is truth-preserving, on my account of the conditional, in instances interpreted counterfactually just when the set of φ-worlds of (160) is a subset of the ψ-worlds of (161). Since, if this is the case and every world at which, *ceteris paribus*, ψ is true is one at which λ is, then it must be that every world at which φ is, *ceteris*

paribus, true (this latter set of worlds being a subset of the former) is one at which λ is true—which gives us (162).

Further, it should be evident from the preceding analysis that the $C3$ account would also allow for truth-preserving transitive inferences in cases where there is neither any linguistic identity (or even close similarity), nor shared truth conditions—between the consequent of one conditional premise and the antecedent of the other. Of course, this is because the factor crucial to transitivity here is, as just discussed, that the set of antecedent-worlds for the conditional in the first premise be a sub-set of, or identical to, the set of antecedent-worlds for the second (rather than simply sameness of truth conditions of the first conditional's consequent and the second's antecedent).

Modus Ponens

While there is, therefore, no reason to be concerned about transitivity on the $C3$ account, one might still be concerned with how exactly *modus ponens* fares on my analysis. There may indeed be cases in which we might expect *modus ponens* to hold but where it does not on this approach, though it, like transitivity, is always truth-preserving in the case of conditionals interpreted factually.

Let's consider an example to clarify the state of *modus ponens* in $C3$,

(163) If there's a quorum at today's faculty meeting, then they will vote unanimously to approve the curriculum changes.

As is the case with transitivity, whether truth is preserved by uses of *modus ponens* may vary between the respective factual and counterfactual interpretations of a given conditional involved in the inference, though, as just noted, it will always hold on the former interpretation. If (163) is factually true and its antecedent true at the actual world (as indeed it must be for the factual truth of the conditional to hold), then its consequent must also be factually true.

What of instances in which a conditional is interpreted counterfactually? Take (163), for instance, along with the following background assumptions: Almost none of the attendees ever vote in favor of or against any motions made in these faculty meetings, but instead consistently abstain; today, in a freak coincidence, an unscrupulous criminal—wanting their illegal development plans to be approved by the buildings committee holding their meeting next door to the faculty meeting—accidentally bribes and threatens all of the attending faculty (instead of the building committee members) into voting in favor of the first motion on the agenda (which is to approve the criminal's plans in the buildings meeting and to approve the curriculum changes in the faculty meeting).

As a result, someone not in the faculty meeting but familiar with its typical workings may, on hearing that there was a quorum, quite reasonably conclude that (163) is counterfactually false, since almost no one ever does anything but abstain. This serves to explain why the factual truth of (163) is so odd—dependent as it is on a bizarre and wildly unlikely mistake, which is ruled out by the *ceteris paribus* clause of the counterfactual interpretation. But, nonetheless, in the case of (163)—concerning as it does a single, actual event—what is more salient is what does in

fact happen at the actual world, and the counterfactual interpretation is thus not assertible as false (at least not without a degree of explanation that one is talking about what typically happens, rather than what actually did in this one instance) by those who believe that this conditional is factually true.[61]

With this in mind, let's use (163) as the first premise in an argument,

i. If there's a quorum at today's faculty meeting, then they will vote unanimously to approve the curriculum changes.

ii. There is a quorum at today's faculty meeting.

iii. They will vote unanimously to approve the curriculum changes.

The relevance of this to the preservation of truth by applications of *modus ponens* is that in a case such as that of this argument just stated, given the circumstances outlined, *modus ponens* will hold (that is, be truth-preserving) for the factual interpretation but not the counterfactual. This, however, is not typically the case—occurring

[61]If you are inclined to take this as an indication that there is no genuine counterfactual interpretation available in this instance, remember that one might, say, hear false rumors about what transpired at the faculty meeting and, on this evidence, conclude that (163) is both factually and counterfactually false, and there would then be nothing making the counterfactual interpretation unassertible. Certainly, you would be wrong about the factual interpretation not clashing with, and so rendering unassertible, the counterfactual reading, but that is quite irrelevant to whether or not these interpretations are genuine, assertible, interpretations. Essentially, to say that (163) is counterfactually false, is to make the claim that there is no counterfactual reliable connection between antecedent and consequent, which is not contradicted by their both being made true purely by happenstance.

only when the counterfactual connection between antecedent and consequent (here that of the conditional constituting the first premise) is absent on the factual interpretation due to some extraordinary circumstance. In general, though, *modus ponens* will go through or fail to do so on both the factual and the counterfactual interpretations of a given conditional. But one might still be concerned about the status of *modus ponens* in those cases where we must rely also on (apparent) transitivity.

Let's return to the series of conditionals, (154), (155), and (156), and further take it that (154) is counterfactually true,

(154) If J. Edgar Hoover had been born in Russia, then he would have been a Communist.

(155) If J. Edgar Hoover had been a Communist, then he would have been a traitor.

(156) If J. Edgar Hoover had been born in Russia, then he would have been a traitor.

If this is the case, then the consequent of (154) must be true—that is, at all A-worlds. Say that (155) is also counterfactually true, as seems quite plausible. Does the truth of the consequent of (154) and the truth of (155) as a whole then give us the truth of the consequent of (155)? No. And so, here, it appears that *modus ponens* fails.

But there is, in fact, no cause to accuse $C3$ of undermining *modus ponens*, since the truth of the consequent of (154) is actually only the truth of this consequent at the A-worlds of (154), not its truth at the actual world or all possible worlds. The implicit *ceteris paribus* clauses of (154) and (155), respectively, are different (as discussed above with regards to transitivity) and thus there is no real failure of *modus ponens*, only an apparent one when we fail to take into account this additional implicit element of the conditional.

If we give (154), (155), and (156) factual interpretations—assuming that all three are factually true, and thus that the antecedent of (154) is true at the actual world—then we could conclude, by *modus ponens*, that the consequent of (156) must also be true. This is because, as observed above, transitivity always holds for factual interpretations, since these have no implicit *ceteris paribus* clause to take into account. But the trouble with the Hoover examples, of course, is that it seems intuitive that (154) and (155) cannot both simultaneously be true at the actual world, and so no such inference would actually be able to go through.

What is the upshot of all this? Simply that *modus ponens*—though not formally valid—may still be truth-preserving on the $C3$ account. The only relevant thing to take into account is the presence of the implicit *ceteris paribus* clause when giving the counterfactual interpretation of any particular conditional, which may result in seeming failures of *modus ponens*. In general terms, then—taking φ and ψ as any antecedent and consequent, respectively,

(164) $\varphi \rightharpoonup \psi$

(165) φ

(166) ψ

modus ponens is not formally valid, but is nonetheless truth-preserving, just when, in addition to the truth of φ, there is nothing in the rest of the argument that violates the *ceteris paribus* clause implicit in (164). This requirement, which might be considered a nuisance, in fact has the striking benefit of enabling us to solve some of the apparent challenges to *modus ponens*.

Consider one of McGee's purported counterexamples to *modus ponens*:

> Having learned that gold and silver were both once mined in his region, Uncle Otto has dug a mine in his backyard. Unfortunately, it is virtually certain that he will find neither gold nor silver, and it is entirely certain that he will find nothing else of value. There is ample reason to believe
>
> > If Uncle Otto doesn't find gold, then if he strikes it rich,
> > it will be by finding silver.
> > Uncle Otto won't find gold.
>
> Since, however, his chances of finding gold, though slim, are no slimmer than his chances of finding silver, there is no reason to suppose that

If Uncle Otto strikes it rich, it will be by finding silver.

> These examples show that modus ponens is not an entirely reliable rule of inference. Sometimes the conclusion of an application of modus ponens is something we do not believe and should not believe, even though the premises are propositions we believe very properly. (McGee, 1985: 463)

The relevant conditionals, then, are,

(167) If Uncle Otto doesn't find gold, then, if he strikes it rich, it will be by finding silver.

(168) Uncle Otto won't find gold.

(169) If Uncle Otto strikes it rich, it will be by finding silver.

And the problem that McGee highlights is that, while we have perfectly good reason to take (167) and (168) to be true, we do not have reason to believe (169) to be true. Yet, if *modus ponens* is a truth-preserving form of inference, then the truth of (167) and (168) would seem to license the truth of (169). How to resolve this impasse?

We have no idea of the factual state of affairs, since Uncle Otto hasn't yet either found or conclusively failed to find anything in his backyard. So the factual interpretation cannot be settled with our current knowledge. But, on the counterfactual interpretation, (169) is only true if the specific *ceteris paribus* clause of the antecedent of the embedded conditional in (167) is not violated. If (169) is evaluated independently of (167), then it has a different *ceteris paribus* clause and is counterfactually false.

Let's consider this in more a bit more detail. What is the relevant *ceteris paribus* clause for the conditional embedded in (167)? Since it forms the consequent of the main conditional, and therefore falls within the scope of the main antecedent, the A-worlds of the embedded conditional—let's call them A_e-worlds, to prevent any confusion—must be selected solely from among the A-worlds of the main conditional—let's call these A_m-worlds. So, we need first to identity the A_m-worlds. These will be all those worlds at which, *ceteris paribus*, it's the case that Uncle Otto does not find gold. This set of worlds can then be narrowed down to those worlds at which, *ceteris paribus*, Uncle Otto strikes it rich. Another way to characterize what is essentially the same process, is to put it in terms purely of the *ceteris paribus* clause of the embedded conditional. Part of what counts as all things being equal in this specific context is that it's not the case that Otto finds gold, or gets rich by winning the lottery, say. Since McGee specifies in setting up the example that Uncle Otto will not find anything else of value, if he doesn't find gold, but does get rich, then it can only be in virtue of his finding silver. Hence, (167) is true on a counterfactual interpretation.

Given the truth of the main antecedent of (167), which the second premise, (168), provides, *modus ponens* licenses only the truth of the conditional embedded in (167)—not any other conditional, even one that happens to be lexically the same. One may take (169) to be independent of (167), and therefore to have a different *ceteris paribus* clause. But, if one does this, then the truth of the premises here, (167) and (168), cease to be relevant to the truth-value of (169). This is because, what the truth of the main antecedent of (167) gives us, is the truth of the embedded conditional *with a specific* ceteris paribus *clause*, and the independently evaluated (169) will have a different *ceteris paribus* clause and so have a distinct counterfactual interpretation from the conditional embedded in (167).

In contrast, though, if the counterfactual interpretations of (169) and that of the conditional embedded in (167) are taken to be the same—that is, their respective *ceteris paribus* clauses are taken to be generated within the very same context, which, as these conditionals are lexically the same, must result in their being identical—then this argument is truth-preserving and the counterfactual truth of (169) ensured by the premises.

As I noted above, we don't know enough about how things are at the actual world with respect to Uncle Otto and his treasure-hunting to work out the truth-value of these conditionals on a factual interpretation. However, we can assess what would be the case given various different potential actual-world scenarios. Most importantly, we can consider the question of why this argument is truth-preserving on a factual interpretation of each of the conditionals involved.

This argument is truth-preserving on factual interpretations of the conditionals

involved because it's not possible for the conclusion, (169), to be false when the premises are both true. For the conclusion to be factually true, it must be the case that Uncle Otto strikes it rich and that he does this by finding silver. But, if we assume the premises to be false, then, in the case of (167), it must be that Uncle Otto doesn't find gold, that he does strike it rich, but that this is not achieved by his finding silver. If that's so, however, then the conclusion must also be false. Thus, this argument is truth-preserving on a factual interpretation of the relevant conditionals. This is unproblematic, though, since the worry that, contra (169), if Uncle Otto strikes it rich it could be because he finds gold—his finding silver being no more likely than the former possibility—is absolutely ruled out from being the case at the actual world when the premises are true, as these explicitly deny that Uncle Otto finds gold.

Contraposition

Contraposition typically fails to preserve truth on non-material conditional accounts of the conditional, and $C3$ is no different, in that there are many cases in which contraposition is not truth-preserving. However, there are instances in which this inference, from a conditional to its contrapositive—the denial of the original conditional's consequent forming the contrapositive's antecedent, and the denial of the original antecedent the contrapositive consequent—is indeed truth-preserving.

The contrapositive of a given conditional is factually true in just those instances in which,

(F) both the antecedent and consequent of the original conditional are false at the actual world.

The contrapositive of a given conditional is counterfactually true in just those instances in which

(Cf1) the counterfactual interpretation of the original conditional is true.

(Cf2) the consequent of the counterfactual interpretation of the original conditional is not a necessary truth.

The condition for the factual truth of a contrapositive, (F), is required so as to ensure that both the antecedent and consequent of the contrapositive are factually true. Since the factual truth-values of the each of these latter will be the negation of those of the original conditional, respectively, the antecedent and consequent of the original conditional must both be factually false in order for those of the contraposition to be factually true.

The first condition for the counterfactual truth of a contrapositive, (Cf1), is required so as to make sure that the relevant counterfactual connection (whatever this may be) between the antecedent and consequent of the original conditional is present, which, in turn, must then hold too for the contrapositive. The second condition, (Cf2), is required because, if the original conditional has a necessary truth

as its consequent, this would result in its contrapositive having a necessarily false antecedent, which would rule out its being true on a counterfactual interpretation.[62]

Let's turn, now, to some examples of conditionals from the Corpus of Contemporary American English (Davies, 2008) in order to determine what explains the above contraposition conditions.

(170) If the radiation leak [at Fukushima] is not stopped completely, then it is likely that the total cumulative radiation could eventually exceed Chernobyl. (spoken, PBS NewsHour, April 12, 2011)

This conditional, (170), is an instance whose contrapositive is both factually and counterfactually true. As explained, above, the rationale for (F) is that only a conditional that's not truth-apt on a factual interpretation, in virtue of both its antecedent and consequent being false, could have a contrapositive that's factually true, since the respective truth-values of the contrapositive will be the opposite of those of the original conditional's antecedent and consequent. As the radiation leak at the Fukushima nuclear plant was stopped completely and, by all accounts, the total cumulative radiation leaked has and will not nearly exceed that of Chernobyl—both the antecedent and consequent of (170) are false at the actual world, satisfying (F).

[62] I consider conditionals with necessarily false antecedents later in this chapter (*see p.240*).

The contrapositive of this conditional, (170), happens also to be counterfactually true, since is satisfies both (Cf1) and (Cf2). But I will discuss the application of these conditions in more detail by means of another example.

This next instance, also from the Corpus of Contemporary American English (Davies, 2008), gives an example of a conditional that has a contrapositive that's true on a counterfactual interpretation, but not a factual one:

(171) If you were black and poor, then, really, nobody looked, especially the black and poor and Southern. ('Atlanta Child Murders', CNN News, January 15, 2011)

This conditional was uttered by an interviewee talking about serial killings going unnoticed in the US South in the past. The 'you' in the antecedent is referring to those who were murdered at the time, specifically identifying those who were black and poor (and especially those who are Southern as well)—and since there were people who fit this description at the relevant time, this antecedent is true at the actual world. And, given the lack of contemporary investigation (as described in the news segment from which this conditional is taken) into the murders of people who fell into these categories, it's certainly plausible that the consequent, too, is true at the actual world. Thus, both the antecedent and consequent are true at the actual world. The contrapositive of (171) is, therefore, not factually truth-apt.

However, (171) is true on a counterfactual interpretation, since, at every world at which it is, *ceteris paribus*, the case that one is a poor, black, Southern murder

victim (at the time with which this news segment was concerned), it is also true that no one would properly investigate one's death, which is the gist of the consequent here. The counterfactual truth of (171) then satisfies condition (Cf1) and licenses the truth of its contrapositive on a counterfactual interpretation—namely, the truth of the claim that, if one's murder was properly investigated, then one was not poor, black, and Southern. Further, (171) also satisfies (Cf2), in that its consequent is obviously not a necessary truth and, therefore, there are relevant A-worlds at which its negation—the antecedent of the corresponding contrapositive—will be true.

Thus, while contraposition may be not be formally valid in $C3$, there are nonetheless clear conditions under which the inference from a conditional to its contrapositive is truth-preserving, on factual and counterfactual interpretations.

Probability conditionals

I won't say a great deal about these here, other than to gesture toward a way in which it's possible to make sense of these kinds of conditionals on a $C3$ account.

Although rather glossed over above, a conditional such as (170),

(170) If the radiation leak [at Fukushima] is not stopped completely, then it is likely that the total cumulative radiation could eventually exceed Chernobyl. (spoken, PBS NewsHour, April 12, 2011)

is not making the claim that the part of the consequent after "likely" is true at every single one of the A-worlds, as in the counterfactual interpretation of an ordinary instance of the conditional such as we've looked at so far. Rather, the claim is that the consequent is *likely* true, given the truth of the antecedent; that is, that it is true at at least some of the A-worlds. Just how many of these worlds would need to be ones at which the consequent is true—in order to fulfill the requirement that it be *likely* that the total cumulative radiation eventually exceeds that of Chernobyl—will be determined by the definition of the term 'likely' and to an extent by pragmatic and contexual considerations.

The same sort of approach will allow for the explication of conditionals such as the following,

(172) If you have that server, there's only a 50 percent chance you'll get what you ordered.

(173) If his light's on, then he's probably home.

(174) If she putts carefully, she gets the ball in the hole in one shot almost every time.

(175) Even if you arrive very early, it's unlikely there'll still be any tickets available.

(176) If I study really hard for this exam, I might still be able to pass the course.

What of the factual interpretation of these sorts of conditionals? It's not clear just what sense to make of the notion that, at the actual world, you get what you ordered 50 percent of the time, when there is just one time under consideration. I think that these sorts of conditionals may either simply be assessed as not factually truth-apt, on the grounds that their consequents are not truth-apt as applied solely to the actual world; or, alternatively, a case could be made that, whether you get what you ordered or don't, both of these would make the consequent true at the actual world, since neither option is categorically ruled out, even though one may be claimed to be more likely that the other.

Regardless of which option is more appealing, it's not clear that it matters terribly much, since the nature of this sort of conditional is such that we're far more concerned with its counterfactual interpretation. What's salient here is whether the balance of possible worlds at which the part of the consequent aside from the modifier—("could", "50 percent chance", "probably", "almost every time", "unlikely", and "might", respectively), which denies the certainty of the consequent—is true, matches the claim made by that modifier.

So, for instance, let's look at (174): This conditional will be true on a counterfactual interpretation when, *ceteris paribus*, at almost every A-world—those at which she putts carefully—it's true that she gets the ball in the hole in one shot.

Determining the exact requirements of the term 'almost every' is something for an account of vagueness to settle.

To take another example: It will be counterfactually true that, "If his light's on, then he's probably home", when, *ceteris paribus*, at the majority of A-worlds—those at which his light is on—he also is home. Of course, there is a question, again, as to how precisely to define the modifier of the consequent. Just what size majority constitutes something's being probable? But this is beyond the scope of the work that a theory of the conditional needs to do.

I'd like to consider briefly, here, the issue of epistemic possibility. Probability conditionals are often considered, as by the suppositionalists among others, to be about epistemic rather than metaphysical possibility. The conditional (176), for instance; if this is taken to be wholly epistemic, then it may be paraphrased as, "If I study really hard for this exam, then, for all I know, I will pass." Of course, though 'might' conditionals are often singled out as being particularly epistemic in nature, every judgment of probability is heavily affected, and limited, by one's epistemic position. Trying to control for this, as reasonably possible, is effectively the role that the implicit *ceteris paribus* clause performs for us. If there is any sense at all to be made of metaphysical possibility, then the restriction of the A-worlds by the *ceteris paribus* clause leaves us with those worlds at which the antecedent is a real metaphysical possibility.

Naturally, there will be a complex interaction between the *ceteris paribus* clause and the sort of modifiers of the consequent discussed here. If the reading of 'might' conditionals as wholly epistemic is correct, then on a counterfactual interpretation

all of the A-worlds—worlds at which all I don't know and can't foresee is ruled out by the *ceteris paribus* clause—will be worlds at which the consequent is true, at which I pass. This, however, would make the "might" in the consequent wholly irrelevant, which cannot be correct. But, if these, and other probability conditionals, are such that they really concern metaphysical possibility, then the *ceteris paribus* clause will provide a set of A-worlds at which genuine metaphysical uncertainty is taken account of; that is, worlds at which metaphysical probability needs to be used to determine just how many of the A-worlds must be ones at which the consequent is true in order for the particular modifier in question to be satisfied.

In the case of (176), taking the "might" to be purely epistemic would absolutely guarantee that I pass, provided I study hard. But, taken to concern metaphysical possibility rather, it would still allow that I study hard but am crushed by a falling boulder on the way to my final exam—or that the professor takes an extreme dislike to me and doctors my grade—and that, consequently, I fail the class, despite my dedicated studying. Considering such conditionals to be solely epistemic, and so make its factual truth impervious to the kinds of unusual circumstances just considered, seems entirely unjustified.

Necessarily False Antecedents

A significant challenge to the $C3$ conditional is the problem of necessarily false antecedents. In any instance in which the antecedent is necessarily false, the conditional is not truth-apt, on either the factual or counterfactual interpretation. Thus, as it stands, there is no way to distinguish the truth-values of, for example, the following

conditionals:

(177) If the sum of 5 and 10 is 29, then the sum of 5 and 10 is a prime number.

(178) If the sum of 5 and 10 is 29, then the sum of 5 and 10 is 400.

Some argue, though, that there is an intuitive difference between these conditionals' respective truth-values; (177) being true, and (178) false. This is not an uncontroversial view—the standard line on necessarily false antecedents is generally taken to be that of Lewis and Stalnaker; namely, that all conditionals with such antecedents are simply vacuously true. Stalnaker (1968), for instance, makes use of the notion of a single absurd world, λ, at which all conditionals with necessarily false antecedents ('nf-conditionals', hereafter) may be assessed. And, Williamson (2010), among others, argues against the view that any nf-conditionals may be substantively true. He claims that, upon more careful examination, many instances of nf-conditionals that are supposedly either substantively true or false prove to break down, and that other examples rest on confused intuitions; though many have offered rebuttals of his contentions.

For instance, it's arguable that we're able to determine that (177) is true—since, taking the antecedent's claim that the sum of 5 and 10 is 29 to be true (and given that 29 *is* a prime number), the consequent correctly claims that the sum in question is prime—and that, in contrast, we can determine that (178) is false. If we take the

claim of the latter's antecedent (that 5 and 10 are 29) to be true, then (given that 29 is *not* equal to 400) the consequent cannot be correct in claiming that 5 and 10 are 400. Of course, an opponent of this view would respond that, since it's logically impossible for the sum of 5 and 10 to be 29, if we may take this impossibility to be true, it's not clear why or how we could rule out any other impossibility (such as 29 not being prime, or 29 being equal to 400).

It's my view that we do intuitively differentiate between true and false nf-conditionals in ordinary discourse, but those who disagree may simply take the $C3$ account as it is and ignore any efforts to accommodate such distinctions. There does, though, seem to be at least enough *prima facie* intuitive plausibility to the claim that we are able to meaningfully distinguish the truth-values of nf-conditionals such as (177) and (178) from one another, to warrant consideration of this question in a theory of the conditional.

Regardless of what sort of interpretation provides the best explanation of how it is we do in fact use nf-conditionals in ordinary language, I think that this may be accommodated by the $C3$ account. Should the evidence of ordinary usage show that we are genuinely making meaningful claims about what follows and what does not from a given impossibility, then this may be accounted for by adding impossible worlds to those at which conditionals are assessed. Conditionals with antecedents that aren't necessarily false would be unaffected, their truth-values determined just as before. But, a given nf-conditional would be assessed in terms of all those impossible worlds at which, *ceteris paribus*, its antecedent was true. The restriction provided by the *ceteris paribus* clause would ensure that the impossible worlds concerned were not

ones at which everything was true, but just those that differed logically, or otherwise, only as far as required for the antecedent to be true.

If the critics of such accounts are correct and there is convincing reason to hold that the relevant intuitions here are mistaken, the $C3$ conditional should remain just as it is. If not, it's relatively simple to augment the logic of this account to include impossible worlds of evaluation.

We would need to introduce non-normal worlds, NN, consisting of all those worlds $W - N$, where N is the set of all the normal worlds. Non-normal worlds are those at which logical truths may not hold (Priest, 1992, 2008), which allows us to accommodate conditionals with necessarily false antecedents in that these worlds permit the truth of such antecedents—which are, after all, the denial of logical truths. The connectives may be held the same as they are at normal worlds, except for that of the counterfactual interpretation of the conditional, which—if there were no A-accessible normal worlds—would need to consider those non-normal worlds at which A is *ceteris paribus* true, and evaluate the conditional in terms of these non-normal A-worlds.

Importantly, logical validity should still be defined in terms of truth preservation at normal worlds, since impossible worlds are precisely those at which the laws of logic fail (Priest, 1992, 2008; Berto, 2013). Then, aside from allowing for the truth of necessarily false antecedents, the non-normal A-worlds for a given conditional permit nothing else untoward, so that the truth-value of the consequent at such worlds is evaluated as normal.

Counteridenticals and Counteractuals

Counteridenticals

Goodman raises the issue of counteridentical conditionals, and considers this apparently problematic pair (Goodman, 1947: 115; updated slighty),

(179) If I were Julius Caesar, I wouldn't be alive in the twenty-first century.

(180) If Julius Caesar were I, he would be alive in the twenty-first century.

It's worth noting here that, if we both accept the necessity of identity (for instance, as articulated by Kripke, 1971) and take either or both of (179) or (180) to be meaningful, then we must accept that conditionals with necessarily false antecedents should be given substantive truth-values—since the necessity of identity would, of course, render the antecedents of both these conditionals necessarily false.

This issue aside, though, the evident trouble is that the two respective antecedents of (179) and (180) seem to be versions of just the same identity claim, that between Julius Caesar and myself. Further, the two respective consequents appear to exactly opposite claims about this Julius Caesar–I person; namely, that they would not be alive in the twentieth century and that they would be, respectively. That being the case, if these conditionals are asserted in the same context, how is it that both (179) and (180) are intuitively true?

The $C3$ is easily able to account for this difference. There isn't any difference between the meaning of the two explicit antecedents at the level of their semantics, but there is a very evident pragmatic difference in terms of the respective implicatures carried by each one. The Gricean maxim of manner explains why the antecedent of (179), "I was Julius Caesar", implicates that I went back to Julius Caesar's time—that, effectively, I was the person Julius Caesar and lived his life in some sense.[63] In contrast, the implicature of (180)'s antecedent, "Julius Caesar is I", is that Julius Caesar is, in some way, the person I am—living now, in the twenty-first century and doing the things I do, and so on.

The upshot of these pragmatic differences between the two semantically equivalent antecedents, is that their respective *ceteris paribus* clauses will differ in what they exclude and include. Clearly, the factual interpretation of these conditionals would be interesting only to those with delusions of being Julius Caesar, but the counterfactual interpretation is of more general interest. For this interpretation of (179), we must look to the worlds at which it is the case Julius Caesar and myself are one and the same person, but just those of them at which it is also the case that all things are equal. Given the pragmatics of the antecedent of (179), these will be those worlds at which I-who-am-Julius Caesar was born in 100 BCE and died in 44 BCE. Thus, all the A-worlds of (179)'s interpretation will be ones at which I-who-am-Julius

[63] I'm taking what's implicated by the antecedent of (179) to be of the kind that explicates that we (as ordinary speakers) take a sentence such as, "I put on my shoes and went for a walk", to implicate an ordering of the two conjunctions. The implicature is that I first put on my shoes, then went for the walk after this. There is debate, though, about the precise nature of this implicature (Davis, 1998), and likely no clear consensus on an implicature such as that of (179)'s antecedent—in terms of just what maxim gives rise to it and so on. But nothing hangs on these details; all that's needed is that there information additional to that conveyed by the antecedent's semantics that is given pragmatically here.

Caesar isn't alive in the twenty-first century, and so (179) is counterfactually true. On the other hand, in light of the different pragmatics of (180), the *ceteris paribus* clause of its counterfactual interpretation is one that restricts the A-worlds to those at which Julius Caesar-who-is-I was born in the twentieth century and has lived into the twenty-first. Therefore, Julius Caesar-who-is-I is alive in the twenty-first century at all of these A-worlds, and, as the consequent is true at all A-worlds, so too is (180) as a whole, on a counterfactual interpretation.

Counteractuals

So much for the problem of counteridenticals. But what of counteractuals on my approach? Take the following pair of conditionals (based on Yablo's examples, 2002: 113),

(181) If 'pigeon' meant table, then pigeons would (still) have two legs.

(182) If 'pigeon' actually meant table, then pigeons would have four legs.

The claim is that these two conditionals are both intuitively true. Assuming that the two above conditionals do differ substantively, what explains this? Or, put another way, how are we to evaluate a conditional such as (182), what Yablo (2002) terms a 'counteractual'?

What is it to consider things as they might *actually* have been? In fact, this is not fundamentally different to any other consideration of possibility. Think of what

a possible world is supposed to represent—a way the world might have been/be. Which world? The actual world. That's the world where we live and so the one we're concerned with. When we say that, "It's lucky that accident wasn't too bad. It could have been fatal.", we are (rightly or wrongly) making a claim about what's possible at the actual world. Hence the consideration of possibility in terms of the possible worlds accessible from the actual world, rather than any other. To be a world accessible from the actual world is to be taken as a representation of how things could possibly be at the actual world.

Why then does it seem that there is a distinction between conditionals such as (181) and (182)?

Consider the following pair of conditionals:

(183) If firecrackers healed burns, then setting off firecrackers would (still) put one at risk of getting burned.

(184) If firecrackers actually healed burns, then firefighters would carry firecrackers to treat burn victims.

What (183) and (184) show us is that what really needs explication is not the supposed 'counteractual' (184), but conditionals like (183). In the case of (184), the counterfactual interpretation appears more salient, and we get the intuitively correct answer if we evaluate this conditional just as normal, by taking all the worlds at

which it is *ceteris paribus* the case that firecrackers actually heal burns, which requires merely that each of these worlds takes itself to be actual—something no world, from its own perspective, can help but do. To take a world as actual means nothing more than to look at that world and see what's true there. And, at all those worlds at which (*ceteris paribus*), from the vantage-point of that world, firecrackers *actually* (that is, at this very world itself) heal burns, it is also the case that firefighters would carry firecrackers (which cause rather than treat burns back here at this world) as a cheap and effective way to treat burns. This is, again, no different from any ordinary counterfactual interpretation.

What's unusual is, in fact, the reading given of the first conditional of the pair, (183); unusual, because what it's taken to being conveying is a *lack* of connection between the antecedent and consequent, as opposed to the sort of relationship ordinarily implicated. Let's try to interpret (183) counterfactually: First, we need to identify all those worlds at which, *ceteris paribus*, firecrackers heal burns. Would these be useful at such worlds for treating burns?

According to those who argue for counteractuals, it's crucial, in assessing whether the consequent is true at these A-worlds, to determine the properties of firecrackers. Without the "actually" of (182) in its antecedent, the properties of the firecrackers, as referenced in the consequent of (183), are not what they would be were these objects considered at any one of the A-worlds, but what they are back at our world, the actual world. Here, firecrackers burn people rather than heal their burns. Thus, regardless of what properties they may have at any other worlds, they still have their same, familiar properties at the actual world. So, (183) is also counterfactually

true—because, at every A-world (indeed at every possible world there is), it is still the case that firecrackers *as they are at the actual world* cause burns. What this conditional, (183), is claimed to convey then, is that the properties of objects at the actual world are quite independent of what their properties might have been. So, even if we accept the reality of counteractuals, there's no need for any different sort of interpretation. The ordinary counterfactual interpretation of $C3$ will suffice.

However, as likely became clear in this analysis of (183), the claim that there is a distinct reading of counteractuals doesn't stand up to scrutiny. Essentially, the distinction drawn between (181) and (182) rests on the contingence of the meaning of the word 'pigeon'. It happens to mean a widespread and adaptable small bird, but it needn't have meant this. What's claimed is that, without "actually" in the antecedent, the use of the term 'pigeon' in the consequent refers to what the word picks out at the actual world, rather than at the A-worlds.

But this argument is equally applicable to the contingency of a particular type of object's properties. Firecrackers happen to have the property of causing burns, but it's arguable that this is not an essential property (unlike, say, the property of giving rise to a particular visual display after being lit, for instance—but, if one is a fan of intrinsic properties and causing burns seems to be one of them for firecrackers, you may simply substitute something like the property of having a red paper wrapper, say). It's hardly plausible, though, that (183) is counterfactually true; in other words, that, in ordinary discourse, anyone would take the reference to firecrackers in the consequent to be independent of that in the antecedent. The term "firecrackers", as used in the consequent, is surely making reference to what would be true of this

object if it had properties other than it actually does, as given by the antecedent. And, thus, it's more reasonable to give (181) this same reading, too, and so deem it counterfactually false.

That said, if one is deeply attached to the idea of distinct counteractuals, the $C3$ analysis is perfectly able to accommodate such readings.

Conclusion

I've presented, here, an analysis of the conditional as it's used in ordinary-language. The aim has been to provide the logical form of the condition, independent of any grammatical or other quirk of its use in a specific natural language. We are all, unavoidably, very rooted in our particular language communities, but it's still possible to abstract away from these to some degree. And, in so doing, it became quickly evident that the taxonomy of the ordinary-language conditional has been significantly confused by a failure to consider this locution cross-linguistically.

Independent of the logic, $C3$, I argue to be that underlying our intuitive day-to-day ascriptions of truth-values to the conditional, is my contention that there are not two distinct types of instances of the conditional but rather two distinct interpretations available for any given instance. Naturally, one of these may be more salient than the other in a given context; one may even require ridiculously far-fetched beliefs in order to be true (or false)—it's certainly not the case that each interpretation is relevant in every situation, even though it is available.

There are important advantages to the dual interpretation approach to the con-

ditional, not least that it's possible to get clear on the difference between factual and counterfactual, rather than having to rely on vague and poorly motivated categorizations of the latter any longer. Interpretations of the conditional as making a claim about purely what is actual, are factual. Interpretations of the conditional making a claim about what may be the case (though not excluding what is the case), are counterfactual. This dual view also undercuts motivation to deny truth conditions to the conditional, obviating the need for implausible accounts in terms of assertability or believability in order to deal with the apparent truth-value clashes and confusions generated by the failure to recognize both interpretations a conditional may take.

Of course, I do also argue for a particular logic of the interpretations I identify. My account claims that the conditional can be understood as asserting what does and does not follow given either the factual or counterfactual truth of the antecedent, but that there is little sense to be made of what follows from the truth of an antecedent that is false; hence the conditional not being truth-apt on either interpretation when the antecedent is either factually or counterfactually false, respectively. I made the case, too, for the inclusion of an implicit *ceteris paribus* clause in the counterfactual interpretation. It's clear that this interpretation must be a variably strict one, but, as the difficulties with Stalnaker's and Lewis's accounts show, similarity is not the correct notion for determining this variance. Our common, everyday experience supports the contention that the *ceteris paribus* clause is the best candidate for this job.

There are myriad challenges facing any theory of the conditional, but, in my final chapter, I offered some indication of how the $C3$ account is able to manage a variety

of potential concerns. My view doesn't provide for the formal validity of inferences of transitivity, *modus ponens*, or contraposition that the material conditional does—but it is able to explain why and provide clear conditions under which instances of these inferences are nonetheless truth-preserving.

Issues for further research abound—the vast sea of work on the conditional contains any number of additional puzzling instances and different forms of conditional for my account to address, such as conditional questions and commands, to take just two examples. There are also the projects of formally extending $C3$ to allow for non-normal worlds of evaluation for the assessment of conditionals with necessarily false antecedents, and of modifying the counterfactual interpretation to accommodate probability conditionals (and others).

In this dissertation, though, I hope to have made a plausible case for a ordinary-language conditional susceptible of dual interpretations, and one that genuinely respects our day-to-day usage.

Bibliography

Abbott, B. (2004). Some Remarks on Indicative Conditionals. *Proceedings from Semantics and Linguistic Theory (SALT) 14*, 1–19.

Abbott, B. (2010). *Contrasting Meanings in Languages of the East and West*, Chapter Conditionals in English and FOPL. Oxford, Bern, Berlin, Bruxelles, Frankfurt am Main, New York NY, Wien: Peter Lang.

Adams, E. W. (1965). A Logic of Conditionals. *Inquiry 8*, 166–97.

Adams, E. W. (1966). Probability and the Logic of Conditionals. In J. Hintikka and P. Suppes (Eds.), *Aspects of Inductive Logic*. Amsterdam: North Holland.

Adams, E. W. (1970, Feb.). Subjunctive and Indicative Conditionals. *Foundations of Language 6*(1), 89–94.

Adams, E. W. (1975). *The Logic of Conditionals*. Dordrecht: D. Reidel Publishing Company.

Adler, J. (2008). *Reasoning: Studies of Human Inference and its Foundations*, Chap-

ter Introduction: Philosophical Foundations. New York NY: Cambridge University Press.

Appell, D. (2008). The Sun Will Eventually Engulf Earth—Maybe. In *Scientific American*. September 1.

Bach, K. (1999). The Myth of Conventional Implicature. *Linguistics and Philosophy 22*, 327–366.

Barnett, D. (2010). Zif Would Have Been If: A suppositional view of counterfactuals. *Noûs 44*(2), 269–304.

Bassetti, C. L. (2009). *The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology*, Chapter Sleepwalking (Somnambulism): Dissociation Between 'Body Sleep' and 'Mind Sleep'. London, Burlington, San Diego: Academic Press.

Bennett, J. (2006). *A Philosophical Guide to Conditionals*. New York NY: Oxford University Press.

Berto, F. (2013). Impossible Worlds. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2013 ed.). Stanford CA: Stanford Center for the Study of Language and Information.

Bickel, B. and J. Nichols (2013). *Inflectional Synthesis of the Verb*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Carroll, L. (1872). *Through the Looking-Glass, and What Alice Found There*. London: Macmillan and Co.

Dachkovsky, S. and W. Sandler (2009). Visual intonation in the prosody of a sign language. *Language and Speech 52*(2–3), 287–314.

Davies, M. (2008). *The Corpus of Contemporary American English: 450 million words, 1990–present*.

Davis, W. A. (1979). Indicative and Subjunctive Conditionals. *The Philosophical Review 88*(4), 544–564.

Davis, W. A. (1998). *Implicature: Intention, Convention, and Principle in the Failure of Gricean Theory*. Cambridge: Cambridge University Press.

Dennett, D. C. (1984). Cognitive Wheels: The frame problem of AI. In C. Hookway (Ed.), *Minds, Machines and Evolution*. Cambridge: Cambridge University Press.

DeRose, K. and R. E. Grandy (1999). Conditional Assertions and "Biscuit" Conditionals. *Noûs 33*(3), 405–420.

Donaldson, B. (1994). *The Germanic Languages*, Chapter Afrikaans. Oxon, New York NY: Routledge.

Earman, J., J. T. Roberts, and S. Smith (2002). Ceteris Paribus Lost. *Erkenntnis 57*(3), 281–301.

Edgington, D. (1995). On Conditionals. *Mind 104*, 235–329.

Edgington, D. (2003, September). What If? Questions about conditionals. *Mind and Language 18*(4), 380–401.

Edgington, D. (2008). Conditionals. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Winter 2008 Edition)*. Stanford CA: Stanford Center for the Study of Language and Information.

Edgington, D. (2014). Indicative conditionals. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2014 ed.). Stanford CA: Stanford Center for the Study of Language and Information.

Edgington, D. (n.d.). Conditionals, truth and objectivity. Manuscript.

Fang, W. (2014). Factual Conditionals in Mandarin Chinese: Based on the analysis of the subcategories of conditionals. *Chinese Language Learning 2*.

Fowler, H. (1926). *A Dictionary of Modern English Usage*. Oxford: Oxford University Press.

Gärdenfors, P. (1986). Belief Revisions and the Ramsey Test for Conditionals. *Philosophical Review 95*, 81–93.

Gärdenfors, P. (1988). *Knowledge in Flux*. Cambridge MA: MIT Press.

Goodman, N. (1947, February). The Problem of Counterfactual Conditionals. *The Journal of Philosophy 44*(5), 113–128.

Goodman, N. (1983). *Fact, Fiction, and Forecast*. Cambridge MA, London: Harvard University Press.

Grice, H. (1975). Logic and Conversation. In P. Cole and J. Morgan (Eds.), *Syntax and Semantics*, Volume 3. New York NY: Academic Press.

Grice, H. (1989). *Studies in the Way of Words*. Cambridge MA: Harvard University Press.

Hájek, A. (n.d.). Most Counterfactuals are False. Manuscript.

Hansson, S. O. (1996). What is Ceteris Paribus Preference? *Journal of Philosophical Logic 25*(3), 307–332.

Haviland, J. (1979). *Handbook of Australian Languages*, Volume 1, Chapter Guugu Yimidhirr. Amsterdam: John Benjamins B.V.

Iatridou, S. (2000). The Grammatical Ingredients of Counterfactuality. *Linguistic Inquiry 32*(2), 231–270.

Jackson, F. (1991). *Conditionals*. Oxford: Oxford University Press.

Jackson, F. (1998). *Mind, Method and Conditionals*. London, New York NY: Routledge.

Kment, B. (2014). *Modality and Explanatory Reasoning*. Oxford: Oxford University Press.

Kolchanov, N., I. Titov, I. Vlassova, and V. Vlassov (1996). Chemical and Computer Probing of RNA Structure. *Progress in Nucleic Acid Research and Molecular Biology 53*, 131–196.

Kratzer, A. (1979). Conditional Necessity and Possibility. In A. v. S. Rainer Bäuerle, Urs Egli (Ed.), *Semantics from Different Points of View*. Springer-Verlag.

Kratzer, A. (1986). Conditionals. *Chicago Linguistics Society 22*(2), 1–15.

Kratzer, A. (2012). *Modals and Conditionals: New and Revised Perspectives*. Oxford, New York NY: Oxford University Press.

Kripke, S. (1971). Identity and Necessity. In M. K. Munitz (Ed.), *Identity and Individuation*. New York NY: New York University Press.

Lange, M. (2002, November). Who's Afraid of Ceteris-Paribus Laws? Or: How I Learned to Stop Worrying and Love Them. *Erkenntnis 57*(3), 407–423.

Levinson, S. C. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge MA: MIT Press.

Lewis, C. (1918). *A Survey of Symbolic Logic*. Berkeley CA: University of California Press.

Lewis, D. (1973a, October). Causality. *The Journal of Philosophy 70*(17), 556–567.

Lewis, D. (1973b). *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, D. (1975). *Formal Semantics of Natural Language*, Chapter Adverbs of Quantification. Cambridge: Cambridge University Press.

Lewis, D. (1976). Probabilities of Conditionals and Conditional Probabilities. *Philosophical Review 85*, 297–315.

Lewis, D. (1979, Nov). Counterfactual Dependence and Time's Arrow. *Noûs 13*(4), 455–476.

Lewis, D. (1981). Ordering Semantics and Premise Semantics for Counterfactuals. *Journal of Philosophical Logic 10*(2), pp. 217–234.

Liddell, S. K. (1986). Head Thrust in ASL Conditional Marking. *Sign Language Studies 52*(1), 244–262.

Lycan, W. G. (2001). *Real Conditionals*. Oxford: Clarendon Press.

Mackie, J. (1973). *Truth, Probability and Paradox*. Oxford: Clarendon Press.

McGee, V. (1985). A Counterexample to Modus Ponens. *Journal of Philosophy 82*(9), 462–471.

Molencki, R. (1996). *Advances in English Historical Linguistics*, Chapter Modals in Past Counterfactual Conditional Protases. Trends in Linguistics Studies and Monographs 112. Berlin, New York NY: Mouton de Gruyter.

Neale, S. (2004). This, that, and the other. In A. Bezuidenhout and M. Reimer (Eds.), *Descriptions and Beyond*, pp. 68–82. Oxford: Oxford University Press.

Nevins, A. (2002). Counterfactuality Without Past Tense. *Proceedings of the North East Linguistic Society 32*, 441–451.

Nute, D. (1980). *Topics in Conditional Logic*. Dordrecht, Boston MA, London: D. Reidel Publishing Company.

Pfau, R. and J. Quer (2010). *Sign Languages*, Chapter Nonmanuals: Their Grammatical and Prosodic Roles. Cambridge Language Surveys. Cambridge: Cambridge University Press.

Potts, C. (2007). Into the Conventional-Implicature Dimension. *Philosophy Compass 2*(4), 665–679.

Priest, G. (1992). What is a Non-Normal World? *Logique et Analyse 35*, 291–302.

Priest, G. (2008). *An Introduction to Non-Classical Logic: From If to Is*. Cambridge Introductions to Philosophy. Cambridge, New York NY: Cambridge University Press.

Priest, G. (2009). *Minds, Ethics, and Conditionals: Themes from the Philosophy of Frank Jackson*, Chapter Conditionals: A Debate with Jackson. Oxford, New York NY: Oxford University Press.

Raidt, E. H. (1995). *Language and Social History: Studies in South African Sociolinguistics*. Cape Town: David Philip Publishers.

Ramsey, F. (1929/1990). General Propositions and Causality. In D. H. Mellor (Ed.), *Philosophical Papers*, pp. 145–63. Cambridge: Cambridge University Press.

Rawlins, K. (2013, June). (Un)conditionals. *Natural Language Semantics 21*(2), 111–178.

Read, S. and D. Edgington (1995). Conditionals and the Ramsey Test. *Proceedings of the Aristotelian Society, Supplementary Volumes 69*, pp. 47–65, 67–86.

Reutlinger, A., G. Schurz, and A. Hüttemann (2011, March). Ceteris paribus laws. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2011 ed.). Stanford CA: Stanford Center for the Study of Language and Information.

Schenck, C. H. and M. W. Mahowald (1995). A Polysomnographically Documented Case of Adult Comnambulism with Long-Distance Automobile Driving and Frequent Nocturnal Violence. *Sleep 18*(9), 765–772.

Schwenter, S. (1999). *Pragmatics of Conditional Marking: Implicature, Scalarity, and Exclusivity*. New York NY: Garland.

Stalnaker, R. (1968). A Theory of Conditionals. In N. Rescher (Ed.), *Studies in Logical Theory*, pp. 98–112. Oxford: Basil Blackwell.

Stalnaker, R. (1975). Indicative Conditionals. *Philosophia 5*(3), 269–286.

Stalnaker, R. (1999). *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford, New York NY: Oxford University Press.

Stalnaker, R. (2009). Conditional Propositions and Conditional Assertions. In A. Egan and B. Weatherson (Eds.), *Epistemic Modality*. Oxford: Oxford University Press.

Stalnaker, R. and R. Thomason (1970). A Semantic Analysis of Conditional Logic. *Theoria 36*, 23–42.

Starr, W. B. (2014). A Uniform Theory of Conditionals. *Journal of Philosophical Logic 43*(6), 1019–1064.

Taylor, J. (1997). *On Conditionals Again*, Chapter Conditionals and Polarity. Amsterdam Studies in the Theory and History of Linguistic Science. Amsterdam: John Benjamins B.V.

van Fraassen, B. C. (1966). Singular Terms, Truth-Value Gaps, and Free Logic. *The Journal of Philosophy 63*(17), 481–495.

von Fintel (2012). *The Routledge Companion to Philosophy of Language*, Chapter Subjunctive Conditionals. London, New York NY: Routledge.

von Fintel, K. (2011). *Semantics: An International Handbook of Meaning*, Volume 2, Chapter Conditionals. Berlin, New York NY: Mouton de Gruyter.

von Fintel, K. and S. Iatridou (2002). If and When If-Clauses Can Restrict Quantifiers. Manuscript.

Wierzbicka, A. (1997). *On Conditionals Again*, Chapter Conditionals and Counterfactuals: Conceptual primitives and linguistic universals. Amsterdam Studies in the Theory and History of Linguistic Science. Amsterdam: John Benjamins B.V.

Williamson, T. (2010). Modal Logic within Counterfactual Logic. In A. Hoffman and B. Hale (Eds.), *Modality: Metaphysics, Epistemology, and Logic*, pp. 81–96. Oxford: Oxford University Press.

Wilson, D. and D. Sperber (2002). Truthfulness and Relevance. *Mind 111*(583-632).

Yablo, S. (2002). Coulda, Woulda, Shoulda. In T. S. Gendler and J. Hawthorne (Eds.), *Conceivability and Possibility*. Oxford: Oxford University Press.

Yeh, D. and D. Gentner (2005). Reasoning Counterfactually in Chinese: Picking up the pieces. In *Proceedings of the Twenty-seventh Annual Meeting of the Cognitive Science Society*, pp. 2410–2415.

Yong, Q. (2013). Typological Stage of Counterfactuals. In *Proceedings of the 27th Pacific Asia Conference on Language, Information and Computation*, pp. 329–338.