ON STALNAKER'S *INDICATIVE CONDITIONALS*

Fabrizio Cariani

Stalnaker's *Indicative Conditionals* (1975, henceforth IC) is not *primarily* about conditionals. I would call this a "secret" if Stalnaker didn't go to some lengths to tell us himself, emphasizing that the essay's main goal is to showcase the power of a general-purpose theoretical framework and not to solve isolated puzzles in conditional semantics (IC, §VI). My aim is to introduce the central cogs, the main innovations and the general significance of that framework. Since IC does lead with a puzzle about indicative conditionals, that's what I'll do as well. §1-2 reconstruct the puzzle and Stalnaker's solution; §§3-5 develop three research themes that emerge out of IC. My focus throughout will be on unpacking Stalnaker's views and working through some of the unfinished agenda of IC.[1]

1 The core puzzle

1.1 Background

IC stages a competition between two analyses of conditional (*if*...*then*...) sentences. Some notation first: I use the two-place connective '▶' to symbolize our target—the conditional construction. I use 'A', 'B', 'C' as variables ranging over sentences.

According to the *material analysis*, A ▶ C is true at world w iff either A is false at w or C is true at w. (I use '$\supset$' to denote the material conditional).

According to the *selectional analysis* (Stalnaker, 1968; Stalnaker and Thomason, 1970), A ▶ C is true at world w iff C is true at the world v that

[1] For that reason, I will not survey the whole debate in which it fits. For broader surveys, see Gillies (2012, forthcoming).

is the closest A-world to w.[2] I call it 'selectional' because the analysis is often presented in terms of a "selection function" *sel* that inputs a proposition p and an evaluation-world w and outputs the selected world in p from w's perspective. ('>' will denote the selectional conditional.) Since Stalnaker grounds closeness in facts about similarity between worlds, he supplements the selectional analysis with restrictions on closeness, and so on selection functions.

The theoretical task is to incorporate one of these analyses in a theory that accurately predicts which inferences involving conditionals are ordinarily judged valid by speakers of a language like English.

1.2 The puzzle

Minimal assumptions about the logic of *if* make it equivalent to the material conditional. Collapse, as we call it, happens if $\mathsf{A} \blacktriangleright \mathsf{C}$ and the corresponding material conditional $\mathsf{A} \supset \mathsf{C}$ are co-entailing. Using '$\vdash$' to denote entailment, we can write this as follows:

COLLAPSE. $\mathsf{A} \blacktriangleright \mathsf{C} \dashv\vdash \mathsf{A} \supset \mathsf{C}$

The left-to-right direction of COLLAPSE follows from MODUS PONENS (and little else).[3]

MODUS PONENS. $\mathsf{A} \blacktriangleright \mathsf{C}, \mathsf{A} \vdash \mathsf{C}$

Stalnaker's focus is on the right-to-left entailment. All it takes to derive that is the plausible claim that sentences like *either Ada was guilty or Carmen was* entail sentences like *if Ada wasn't guilty, Carmen was*. In schematic form:[4]

OR-TO-IF. $\mathsf{A} \vee \mathsf{C} \vdash \neg\mathsf{A} \blacktriangleright \mathsf{C}$.[5]

[2]This demands the existence and uniquess of a closest world. An important debate thread in conditional semantics has focused on these demands. See Lewis (1973); Stalnaker's replies are in his (1981) and (1984, ch.7).

[3]The literature features many interesting attempts to invalidate MODUS PONENS (McGee, 1985*b*; Lycan, 1993; Kolodny and MacFarlane, 2010). These should not be dismissed but they aren't escape routes from collapse results because attempts to invalidate modus ponens do not usually extend to bare conditionals. That is, they do not extend to conditionals whose antecedents and consequences are not modal sentences, like *if Ada is innocent, Carmen is guilty*. A form of collapse that is restricted to bare conditionals is no better than general collapse.

[4]Stalnaker calls this "the direct argument"; the 'or-to-if' nomenclature comes, I believe from Bennett 2003.

[5]Adams (1975, p.15) gives an example that purports to show OR-TO-IF to be *probabilistically* invalid. What he means by this is that there are cases in which $\mathsf{A} \vee \mathsf{C}$ has high probability while $\neg\mathsf{A} \blacktriangleright \mathsf{C}$ has low probability.

To see this, suppose $\mathsf{A} \supset \mathsf{C}$; by truth-functional equivalence, we get $(\neg\mathsf{A} \vee \mathsf{C})$; but if the direct argument is valid, $\mathsf{A} \blacktriangleright \mathsf{C}$ follows.

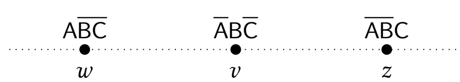
Maybe, the argument for COLLAPSE is an invitation to start loving the material analysis.[6] But, Stalnaker argues, that isn't so great either. If *if* were the material conditional, then negating *If Ada is not guilty, then Carmen is guilty* should entail *Ada is guilty*. That seems wrong: one may negate $\mathsf{A} \blacktriangleright \mathsf{C}$ by way of putting forward some alternative conditional $\mathsf{A} \blacktriangleright \mathsf{B}$, without settling the status of A. That's precisely what conditionals are for.[7] Formally, we want:

NEGATED CONDITIONALS. $\neg(\mathsf{A} \blacktriangleright \mathsf{C}) \not\models \mathsf{A}$

Unfortunately, the material analysis cannot fulfill that wish.

A subtler, more modern take (Gillies, 2004, 2009) accepts COLLAPSE but rejects the material analysis. According to this view, the conditional might well be co-entailing with the material conditional but co-entailment does not suffice for identity of meaning. In fact, it might not even suffice for logical equivalence—depending on how logical equivalence is understood. This view does not appear to be on IC's radar—and indeed, IC's theoretical development is a key step on the path that leads to it. For this reason, I mention it here out of duty to the fuller dialectic but I will refrain from developing it further.

The only other option is to reject the validity of OR-TO-IF. Of course, we didn't need the collapse result to know that OR-TO-IF fails for '$>$'. It is easy to sketch a counterexample. Suppose that in the actual world (w) Ada alone is guilty (A), but in the closest world in which she is innocent (v), Bo is guilty (B) and Carmen is not (C).



Then $\neg\mathsf{A} > \mathsf{C}$ is false at w even if $\mathsf{A} \vee \mathsf{C}$ is true there. The additional value of the collapse result is the implication that invalidating OR-TO-IF is required to avoid co-entailment with the material conditional.

[6]Possibly supplemented by implicature-based moves in the style of Grice 1989; Lewis 1976; Jackson 1987.

[7]See Edgington (1995) for a comprehensive array of nails in the material conditional's coffin.

Here, then, is IC's local puzzle: how might we invalidate OR-TO-IF while accounting for its plausibility?

2 Stalnaker's solution

Stalnaker solution is that OR-TO-IF, though invalid, is pragmatically acceptable. Spelling this out is a two-part task. The first step is to explain how an inference pattern might be acceptable even if invalid. To this end, Stalnaker develops the concept of "reasonable inference". Then, one must check that instances of OR-TO-IF count as reasonable inferences.

The explicit inspiration is Grice's pragmatic defense of the material analysis (Grice, 1989, ch. 4). Like Grice, Stalnaker proposes that much of the inferential profile of the conditional comes from features that are not determined by its truth-conditions. One difference is that, while Grice mostly applies the theory of implicatures to prune off some unwanted entailments of the material conditional, Stalnaker's central concern is that the selectional conditional undergenerates the acceptable inference patterns. Moreover, the patterns it misses are not rescued by the theory of implicatures.

There are important novel elements in Stalnaker's approach. I mention three. *One:* Stalnaker's system relies on constraints tied to the use of specific lexical items in context (see below for illustration). By contrast, the theory of conversational implicatures (at least, in vintage-Grice version) only leverages general principles about rationality and cooperation between conversational participants. *Two:* the acceptability of some reasonable inferences depends on relationships between the information conveyed and the common ground. *Three:* the account appeals to pragmatic principles that are themselves formalizable. Indeed, the appendix to IC goes to some lengths to state the pragmatic theory formally.

2.1 Introducing Reasonable Inference

It is worth quoting Stalnaker's characterization of reasonable inference in full. More recent appeals to IC's playbook often rely on simplified notions that have different structural properties from the original.[8]

[8] For example, reasonable inference is sometimes lumped together with the relation of "quasi-validity" (see e.g. Kolodny and MacFarlane, 2010, fn. 39). An argument is *quasi-valid* if, whenever the premises are known, the conclusion must also be known. Quasi-validity might function in some similar respect like reasonable inferenece (Nolan, 2003, p. 231), but it is significantly different in concept and mechanics. The closest ancestor to quasi-validity is the concept of "doxastic indefensibility" devel-

> [A]n inference from a sequence of assertions or suppositions (the premisses) to a conclusion is a reasonable inference just in case in every context in which the premises can be asserted or supposed, it is impossible for anyone to accept the premises without committing himself to the conclusion. (IC , p. 271)

Let us break down the main technical concepts in this definition with an eye towards coming up with an algorithm for testing reasonableness.

Context: A Stalnakerian context is a record of those propositions that are commonly accepted by the participants to a conversation. Stalnaker's thinking on the nature of context is complex and evolves over four decades (Stalnaker, 1974, 1978, 2002, 2016). One thing that stays constant is the idea that each context c is associated with a set of possible worlds—the *context set* of c. In particular, the context set of c is the set of worlds that verify each of the propositions that are common ground in c. Since the context set is the only feature of context we need to attend to, I will deliberately conflate contexts and their context sets.

Commitment and context acceptance: It is somewhat surprising to see Stalnaker talk about speakers' *commitments*. However, the technical appendix to IC reveals that talk of speaker's commitment is just shorthand for talk about the shared elements of context. The possible world framework Stalnaker adopts allows definitions that clarify the relevant cluster of concepts.

UNSTRUCTURED CONTENTS. $|\mathsf{A}| = \{w \mid \mathsf{A} \text{ is true at } w\}$

CONTEXT ACCEPTANCE. c accepts A iff $c \subseteq |\mathsf{A}|$.[9]

CONTEXT INCOMPATIBILITY. c is incompatible with A iff $c \cap |\mathsf{A}| = \emptyset$

A successful assertive utterance of A in c updates c to $c_{\mathsf{A}} = c \cap |\mathsf{A}|$. To say that a speaker who assertively utters A in c is committed to some other proposition p is just to say that $c_{\mathsf{A}} \subseteq p$.

oped by Hintikka (1962) to account for Moore's paradox. Other nearby notions that are often associated with reasonable inference are Veltman's (1996, p.224) *validity*$_2$ (*aka update-to-test*, in the terminology of van Benthem 1996 and Gillies 2004) and *validity*$_3$ (*aka test-to-test*, or also *informational consequence* in the terminology of Yalcin 2007). Finally, the distinction between preservation of truth and preservation of acceptance also shows up in the debate surrounding McGee's counterexamples to modus ponens (see for example, McGee, 1985*a*, p.90).

[9] In IC, context acceptance is a relation between a context and a set of worlds. For exposition purposes, I prefer to define it as a relation between contexts and sentences (but one that treats any two sentences that have the same unstructured content in the same way).

Assertibility. The definition of reasonable inference quantifies over contexts *in which the premises can be asserted (or supposed)*. To understand this restriction, note that in the background of IC is the idea that assertive utterances come with requirements that the context must meet—as we may call them, *presuppositions*. Some (though not all) of these requirements are specifically tied to the use of particular lexical items. One cannot assertively utter *Raz returned to Valencia* in a context that settles that Raz never was in Valencia. Stalnaker's account of OR-TO-IF essentially references the constraints imposed on context by conditionals and disjunctions.

The reasonable inference test. Let's put this all together. To check whether C is a reasonable inference from premises $A_1, \ldots, A_n$.

- consider an arbitrary context c in which all the premises are assertible.
- let c' be the context that results from asserting all of the premises in sequence.
- the inference is reasonable if and only if c' accepts C.

2.2 Stalnaker's explanation at a glance

It is useful to break down Stalnaker's account of OR-TO-IF in two modules—a general module and one for card-carrying Stalnakerians. The general module, which I introduce in this section, does not presuppose the possible-worlds framework. It only presupposes that there is some way of cashing out the relevant concepts of compatibility and acceptance, such that the following principles are satisfied:

D-CONSTRAINT. $A \vee C$ is assertible in c only if c is compatible with $\neg A \ \& \ C$.

D-UPDATE. Successful assertive utterance of $A \vee C$ in c updates c to a context that is:

- incompatible with $\neg A \ \& \ \neg C$.
- compatible with $\neg A \ \& \ C$ if c was.

C-ACCEPTANCE. If c is incompatible with $\neg A \ \& \ \neg C$ and c is compatible with $\neg A \ \& \ C$, then c accepts $\neg A \blacktriangleright C$.

D-CONSTRAINT is justified on familiar Gricean grounds. D-UPDATE is a direct consequence of the account of assertive update I sketched in §2.1 according to which $c' = c \cap |A \vee C|$. As for C-ACCEPTANCE, it might be defended in a variety of ways, depending on one's semantics for conditionals. I introduce Stalnaker's defense of C-ACCEPTANCE for §2.3.

For now, let us check that this is enough to classify OR-TO-IF as a reasonable inference.

1. Let c be an arbitrary context in which $\mathsf{A} \vee \mathsf{C}$ is assertible.

2. By D-CONSTRAINT, c must be compatible with $\neg\mathsf{A}\ \&\ \mathsf{C}$.

3. Let c' be the result of updating c with the assertion of $\mathsf{A} \vee \mathsf{C}$.

4. By D-UPDATE and 2., c' is incompatible with $\neg\mathsf{A}\ \&\ \neg\mathsf{C}$ but compatible with $\neg\mathsf{A}\ \&\ \mathsf{C}$.

5. By C-ACCEPTANCE and 4., c' must accept $\mathsf{A} \blacktriangleright \mathsf{C}$.

2.3 Inside Stalnaker's box

I postponed the defense of C-ACCEPTANCE. Here it is, in Stalnaker's words:

PRAGMATIC CONSTRAINT. If the conditional is being evaluated at a world in the context set, then the world selected must, if possible, be within the context set as well. [...] In other words, all worlds within the context set are closer to each other than any worlds outside it. (IC, pp. 275-276)

The formal upshot is that a selection function *sel* is only admissible if it satisfies the condition:[10]

$$\forall w \in c \text{ and } z \notin c,\ sel(w, A) = z \text{ only if } c \cap |A| = \emptyset$$

In turn, this constraint is justified as follows:

> Normally a speaker is concerned only with possible worlds within the context set, since this set is defined as the set of

[10]Stalnaker gives a different condition—namely: $\forall w \in c, sel(\mathsf{A}, w) \in c$. This condition does not match the informal presentation of the constraint. However, the condition does follow from a strengthening of the pragmatic constraint which I am about to discuss.

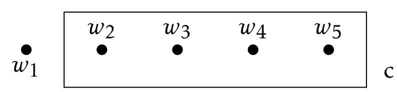
Another thing to consider is whether this stronger condition might be weakened by appealing to accessibility relations. While IC's formalism is a notch looser than Stalnaker (1968) and Stalnaker and Thomason (1970), both these earlier works involve an accessibility relation. In a recent paper, Mandelkern (2018) proposes a role for accessibility relation that bears directly on the formal analysis of the pragmatic constraint.

> possible worlds among which the speaker wishes to distinguish. So it is at least a normal expectation that the selection function should turn first to these worlds before considering counterfactual worlds — those presupposed to be non-actual. (IC, pp. 276)

Some comments in bullet-point format:

- C-ACCEPTANCE only holds by default, or in normal circumstances. This should draw our attention to those cases in which this default presumption is suspended. For instance, Stalnaker remarks that it is suspended by subjunctive conditionals, like *if I had a house, it would not have a front-yard.* (IC, p. 276-277).

- The pragmatic constraint breaks the visual analogy between spatial distance and modal similarity. Consider:

  In this depiction, w_2 is spatially closer to w_1 than it is to w_5. But because both belong to the context set c, w_2 and w_5 must be more similar to each other than w_2 is to w_1. It should be immediately apparent that this is no accident of the depiction: given any depiction of logical space as a convex set, worlds in c can be chosen that are arbitrarily close (spatially) to worlds outside c. Of course, there are ways to depict the structure of logical space that is imposed by the pragmatic constraint, especially if the depiction is allowed to include points that do not correspond to worlds.

  Pictorial musings aside, there is a broader theoretical point. If we try to supplement Stalnaker's proposal with a theory of context-dynamics, the pragmatic constraint will require that every (non-empty) informational update to the context must also rearrange the ordering of modal similarity.

Moving back to the heart of the dialectic, Stalnaker uses the pragmatic constraint to support:

ANTECEDENT COMPATIBILITY. A ► C presupposes that A is compatible with the context set (IC, p. 277).

This idea has been influential and it is sometimes simply quoted in place of the original constraint. However, it is clear that PRAGMATIC CONSTRAINT does not entail ANTECEDENT COMPATIBILITY. The former, but not the latter, is trivially satisfied if there are no antecedent worlds in the context set. IC's text strongly suggests that what fills the gap between PRAGMATIC CONSTRAINT and ANTECEDENT COMPATIBILITY is the availability of the subjunctive conditional as a tool to talk about worlds out of the context set (more on this in the next section). However, it is unclear exactly what principles are relied on in this derivation.

Be all that as it may, the pragmatic constraint is enough for the job at hand.

Fact 1 *Given the framework*, PRAGMATIC CONSTRAINT *entails* C-ACCEPTANCE

Proof: See appendix for proof of this and other facts.

That completes the official path to a pragmatic explanation of OR-TO-IF. There are, of course, other paths that leverage other theories of the conditional. As an illustration, let '$\rightarrow$' denote a strict conditional restricted to the context set. That is, say that $\mathsf{A} \rightarrow \mathsf{C}$ is true at w and context c iff c is incompatible with $\mathsf{A}\ \&\ \neg\mathsf{C}$. Then:

Fact 2 *If entailment is defined as preservation of truth at a world and a context,* $A \vee C \not\models \neg A \rightarrow C$.

Fact 3 *However,* $\rightarrow$ *satisfies* C-ACCEPTANCE.

Given Fact 3, OR-TO-IF for $\rightarrow$ can too be rescued as a reasonable inference. I don't mean to take this point very far, but it does show that the main strategy of IC is not exclusively available to defenders of the selectional analysis.

3 Interlude: indicative vs. subjunctive conditionals

It is a remarkable feature of *Indicative Conditionals* that one of its digressions opened its own research thread. The digression in question concerns the distinction between indicative and subjunctive conditionals. There are many important contrasts between indicatives and subjunctives. The central contrast in IC's dialectic is that indicatives and subjunctives pattern differently with respect to information that is settled by the common ground. Contrast:

(1) #Ada is guilty. If she isn't, Carmen is.

(2) Ada is guilty. If she weren't, Carmen would be.

There is an incongruence in (1), as if the speaker is changing their mind or hedging their assertion. No such incongruence affects (2).

One might account for contrasts like this by distinguishing two kinds of conditional *connectives*. According to this view, (1) uses "the indicative conditional" and the (2) uses "the subjunctive conditional".

Stalnaker would have us avoid this approach. According to him, we can hold on to a single meaning for *if*, while also explaining how indicative and subjunctive conditional *sentences* diverge in meaning. From a pragmatic point of view, Stalnaker thinks that subjunctives are not subject to either version of the pragmatic constraint. Interestingly, this pragmatic suggestion has a syntactic basis. The syntactic proposal is that subjunctive conditionals feature one ingredient that is not present in indicatives: subjunctive mood. Here is how a toy, IC-inspired syntax for (2) might go. Let S-MOOD be a sentential operator that captures the effect of subjunctive mood.

(3) S-MOOD(Ada is not guilty) > WOULD(Carmen is)

Many authors have since questioned the link between "subjunctive mood" and counterfactuality. The alternative view (Dudman, 1983, 1984; Lycan, 2001; Iatridou, 2000; Ippolito, 2013; Starr, 2014; Khoo, 2015) is that "subjunctivity" comes from an extra layer of past tense on the antecedent. What is special about this extra layer is that it does not appear (at least on the surface) to get a past interpretation.[11] As Iatridou notes, this is especially salient when "fake past tense" is restricted by future adverbials, as in:

(4) If Odysseus returned tomorrow, his dog would recognize him.

Despite this, the right perspective on the legacy of IC should be that, though it might have gotten the syntactic details wrong, it helped advance the powerful idea that subjunctive meaning ought to be constructed out of the syntactic ingredients of subjunctive conditionals.[12]

[11] There is a further debate, including among the authors I cited, as to whether the modal feel of subjunctive conditional might be recovered by giving that past tense an ordinary past interpretation.

[12] It should be mentioned that a critical step in this path was the application of Kratzer's (1991; 2012) framework to conditionals.

The other key component of IC's account of the contrast is the idea that there are pragmatic differences between indicatives and subjunctives. As I mentioned, the proposal is that subjunctive conditionals are not subject to the pragmatic constraint. But how is that supposed to work, given that we have already rejected the idea that *if* is ambiguous?

Stalnaker sketches an interesting proposal:

> I take it that the subjunctive mood in English and some other languages is a conventional device for indicating that presuppositions are being suspended, which means in the case of subjunctive conditional statements, that the selection function is one that may reach outside of the context set. (p. 276)

This is how I read this: the PRAGMATIC CONSTRAINT is a presupposition introduced by *if*. However, it can be *lifted* by other devices. Stalnaker's idea seems to be that, while presuppositions are requirements on context, some phrases may *suspend* those requirements (in other words: introduce permissions). I find this idea fascinating and deserving of careful development. As far as I know, it hasn't received one.

It is natural to see this lack of development as a significant lacuna. For instance, Starr writes:

> Why [...] did Stalnaker not propose a meaning for the subjunctive mood and construe the different propositions expressed by subjunctive and indicative conditionals as a function of this mood's contribution? (Starr, 2014, p. 1031)

I think that Starr is exactly right. It is not enough to point to an extra syntactic constituent, be it mood or tense, and claim that it *might* do the job. Much of the best contemporary work on the indicative/subjunctive contrast (including that of the authors I mentioned in this section) is precisely directed at developing a compositional account. Additionally, Starr (2014, p.1030) identifies some technical issues with Stalnaker's approach to the indicative/subjunctive contrast which are well worth working through—though space prevents me from doing that here.

4 On the intransitivity of reasonable inference

Arguably, the most important contribution of IC is the introduction of the concept of reasonable inference. Earlier, I have noted three innovations: (i) that reasonable inference involves formalizable principles (ii) that it is anchored by non-truth-conditional features of the meaning of particular lexical items and (iii) that it involves relations between utterances of conditional sentences and the common ground. In light of these innovations, it is natural to view the concept of reasonable inference as a critical step on the path to dynamic notions of consequence (á-la Veltman 1996).

But it would be hasty to assimilate Stalnaker's notion of reasonable inference and the dynamic concepts of consequence. Stalnaker himself has recently emphasized some differences between the dynamic semantics program and his preferred "dynamic pragmatics" (Stalnaker, 2016). Here, I want to emphasize some under-explored structural features of reasonable inference. In particular, I will focus on the little-known-but-obvious-once-you-see-it fact that reasonable inference is intransitive.[13] I argue that the intransitivity of reasonable inference might come with some explanatory benefits. However, I will also question whether Stalnaker is in a position to reap those benefits, given some of the other commitments he undertakes in IC. More than anything else, my aim is to stress that there are under-explored choice-points in developing the theory of reasonable-inference.

Let us use '$:\approx$' to symbolize the reasonable inference relation and restate its definition:

> $\mathsf{A}_1, \ldots, \mathsf{A}_n :\approx \mathsf{C}$ if and only if for all contexts c, if all of the A_i are assertible in c and c' is the result of asserting all of the A_i's in sequence, c' accepts C.

Here are some important facts concerning $:\approx$. First, every semantically valid argument is a reasonable inference.

Fact 4 *if* $\Sigma \models C$ *then* $\Sigma :\approx C$

Second, the inference from a sentence to its presuppositions is always reasonable.

Fact 5 *if* A *presupposes* C*, then* $A :\approx C$

Given these facts, it is easy to check that $:\approx$ is intransitive in the following sense.

Fact 6 (intransitivity) *Let* Σ *be a set of sentences.* $\Sigma :\approx B$ *and* $B :\approx C$ *do not guarantee* $\Sigma :\approx C$

[13] I learned this from Simon Goldstein and I have not found the issue discussed elsewhere.

In establishing this, it is convenient but not essential to add to the formal language a possibility operator $\Diamond$ that tracks what is compatible with the context set. So, $\Diamond$A is true at w in c iff $c \cap |A| \neq \emptyset$.

> By D-CONSTRAINT a disjunction $A \vee C$ is only assertible in those contexts that are compatible with $\neg A \ \& \ C$. In particular, $A \vee \neg A$ requires that c be compatible with $\neg A \ \& \ \neg A$. Though I have defined compatibility as a relation between contexts and sentences, the relation treats intensionally equivalent sentences in the same way. In other words, compatibility with $\neg A \ \& \ \neg A$ reduces to compatibility with $\neg A$. Using our new possibility operator, we say that $A \vee \neg A$ presupposes $\Diamond \neg A$. Thus, by Fact 5,
>
> $$A \vee \neg A :\approx \Diamond \neg A$$
>
> However, $A \vee \neg A$ is a logical truth; given Fact 4, it can be reasonably inferred from the empty set of premises (i.e. every context accepts it). By contrast, $\Diamond \neg A$ cannot be reasonably inferred from the empty set of premises. Putting it all together:
>
> $$:\approx A \vee \neg A$$
>
> $$A \vee \neg A :\approx \Diamond \neg A$$
>
> $$:\not\approx \Diamond \neg A$$
>
> $\square$

Though this fact is surprising, there is nothing intrinsically problematic about it. Reasonable inference is a technical concept that is invoked to play a certain role—to characterize the acceptable patterns of inference. There is no *a-priori* reason why that job should be best played by a transitive notion.

Still, we may ask whether intransitivity is indispensable to Stalnaker's explanation. A useful first experiment in this direction is to compare $:\approx$ with its transitive closure. Define $:\approx^+$ as the smallest transitive relation that extends $:\approx$. Since $:\approx^+$ extends $:\approx$, it classifies OR-TO-IF as acceptable. More generally, any empirical comparison between $:\approx$ and $:\approx^+$ must be based on inferences that are sanctioned as acceptable by $:\approx^+$ but not by $:\approx$.

This comparison reveals some unique advantages of $:\approx$. One of the advantages of Stalnaker's theory is that it doesn't validate the paradoxes of material implication.

FALSE ANTECEDENT. $\neg \mathsf{A} \not\models \mathsf{A} > \mathsf{C}$

TRUE CONSEQUENT. $\mathsf{C} \not\models \mathsf{A} > \mathsf{C}$

Not only are these patterns invalid, they are also not reasonable, in the sense of $:\approx$. However, both are sanctioned as acceptable by $:\approx^+$. We can see this by noting another violation of transitivity for $:\approx$:

Fact 7

(i) $\neg A :\approx \neg A \vee C$

(ii) $\neg A \vee C :\approx A > C$ *but*

(iii) $\neg A :\not\approx A > C$

(i) follows from Fact 4; (ii) is the direct argument; As for (iii), let c be a context in which ¬A's presuppositions are satisfied and also such that all the A-worlds are ¬C-worlds. Consider some arbitrary world $w \in c_{\neg \mathsf{A}}$. To evaluate $\mathsf{A} > \mathsf{C}$ at w we must reach out of $c_{\neg \mathsf{A}}$. We do know however that all the A-worlds outside of $c_{\neg \mathsf{A}}$ are ¬C-worlds, so $sel(\mathsf{A}, w) \notin |\mathsf{C}|$, which must mean that $\mathsf{A} > \mathsf{C}$ is false at w. Note that PRAGMATIC CONSTRAINT is vacuously satisfied in $c_{\neg \mathsf{A}}$, since there are no A-worlds to select. Furthermore, ANTECEDENT COMPATIBILITY is violated. However, neither of these is incompatible with (iii). □

By contrast, because $:\approx^+$ is transitive and extends $:\approx$, (i) and (ii) force $\mathsf{A} :\approx^+ \mathsf{A} > \mathsf{C}$.

Many theorists accept the validity of FALSE ANTECEDENT and TRUE CONSEQUENT as the cost of doing business. But the intransitive Stalnakerian package undoubtedly maps onto clear intuitions. Consider:

(5) Ada is not guilty

(6) Either Ada is not guilty or Carmen is guilty

(7) If Ada is guilty, Carmen is guilty

The intransitive package predicts that:

- the inference from (5) to (6) is reasonable
- the inference from (6) to (7) is also reasonable

- the inference from (5) to (7) is not reasonable.

These predictions seem just right and a system relying on intransitive notions is uniquely positioned to capture them.[14] None of this is a conclusive argument for Stalnaker's intransitive package. Any such argument would require a much more systematic analysis than what I can offer here. But it does mean that the intransitive package shows up to the jury of theory-choice criteria with a unique profile.

That is all good, but there is one more twist. The official story of IC is a bit more complicated than the story I just told. We can see this by attending to Stalnaker's discussion of the contraposition pattern.

CONTRAPOSITION. $\mathsf{A} > \mathsf{C} \vdash \neg\mathsf{C} > \neg\mathsf{A}$

Like OR-TO-IF, CONTRAPOSITION sounds plausible in the indicative case but not in the subjunctive case (Stalnaker 1968, p. 107; Lewis 1973, §1.8). Once again, Stalnaker proposes to invalidate the general pattern, while explaining the plausible cases *via* reasonable inference. Unfortunately, this strategy doesn't work.

Fact 8 *The reasonable-inference analogue of* CONTRAPOSITION *fails. (i.e.* $A > C \not\approx \neg C > \neg A$*).*

Stalnaker suggests a patch:

> Strictly, [CONTRAPOSITION] is reasonable only relative to the further assumption that the indicative conclusion is not inappropriate" (Stalnaker, 1975, note 15)

What could it mean that an inference pattern is reasonable relative to a further assumption?

One possibility is that there isn't one, monolithic notion of reasonable inference, but a family of them. In addition to $:\approx$, we ought to consider what I'll call *Strawsonian Reasonable Inference*.

> $\mathsf{A}_1,...,\mathsf{A}_n :\approx^S \mathsf{C}$ if and only if for all contexts c, if all of the A_i are assertible in c and c' is the result of asserting all of the A_i's, <u>if C is assertible in c'</u>, then c' accepts the proposition corresponding to C.

[14]For some related technical work involving an intransitive notion of entailment, see Cariani and Goldstein (forthcoming).

This is just like reasonable inference, except for the underlined bit. Moreover, just like reasonable inference $:\approx^S$ is intransitive (it is easy to check that the instance of intransitivity I sketched in establishing Fact 6 also works for $:\approx^S$). In essence $:\approx^S$ is a pragmaticized version of *Strawson Entailment*.[15] Since $:\approx^S$ extends $:\approx$, we can update the Stalnakerian proposal to the claim the acceptable arguments are exactly the ones that are captured by $:\approx^S$.

So, does this Strawsonized package catch CONTRAPOSITION, as Stalnaker claims? It depends.

Fact 9 *If the presupposition of indicatives is* ANTECEDENT COMPATIBILITY, *then* CONTRAPOSITION, TRUE CONSEQUENT *and* FALSE ANTECEDENT *are all Strawsonian Reasonable Inferences.*

Fact 10 *If the presupposition is just the* PRAGMATIC CONSTRAINT, *then none of* CONTRAPOSITION, TRUE CONSEQUENT *and* FALSE ANTECEDENT *count as Strawsonian Reasonable Inferences.*

Two things are interesting here. *One:* there is a tradeoff between invalidating the paradoxes of material conditional and classifying CONTRAPOSITION as acceptable. So it is not clear that Stalnaker can exploit the intransitivity of reasonable inference to avoid the paradoxes of material implication. *Two:* if we prioritize catching the CONTRAPOSITION pattern, we need the full force of ANTECEDENT COMPATIBILITY. This is unlike OR-TO-IF for which PRAGMATIC CONSTRAINT is enough.

5 On Fatalism

The word limit for this essay is nearing, and I must catch up with Stalnaker application of IC's apparatus to the "idle" argument for fatalism (see also Dummett, 1964). Extra motivation here is provided by the recent publication of the contemporary gold standard on the idle argument (Bledin, forthcoming), and by the fact that the idle argument is closely connected with the much-discussed miners paradox.[16] Consider:

[15] In a trivalent setting (e.g. in the context of a semantic account of presupposition), Strawson Entailment is equivalent to the claim that valid arguments are exactly those that cannot have true premises and false conclusions. (If this sounds like classical entailment, recall that in the trivalent setting the "standard" notions of consequence are preservation of truth and preservation of non-falsehood.) For discussion: Strawson (1952); Smiley (1967); von Fintel (1997); **?**.

[16] Kolodny and MacFarlane (2010); Charlow (2013); Cariani *et al.* (2013); Willer (2012); Cariani (2016).

(1) Either this essay will be mistake-ridden or it will not be.

(2) | Suppose it will be.
(3) | If I am careful, the essay will be mistake-ridden.
(4) | Being careful is ineffective.

(5) | Suppose it will not be.
(6) | If I am not careful, the essay will not be mistake-ridden.
(7) | Being careful is unnecessary.

(8) Being careful is either ineffective or unnecessary.

A satisfactory diagnosis of the idle argument ought to identify where it fails and explain why each step appears compelling. Here are the highlights of Stalnaker's account:

- The steps (2) $\Rightarrow$ (3) and (5) $\Rightarrow$ (6) are invalid, since they are instances of TRUE CONSEQUENT.
- The inferences corresponding to each of these steps might be "saved" pragmatically.
- The application of constructive dilemma in the last step in the derivation is not justified as a reasonable inference.
- The last step appears compelling because constructive dilemma is *valid* when the subarguments are valid.

The aggregated effect is that the argument fails both from the point of view of the entailment relation and from the point of view of reasonable inference. However, it appears plausible because each step is justifiable by at least one of these relations.

This dialectic is subtle and needs some unpacking. First, we need to clarify the claims about the constructive dilemma pattern. Contrast:

CD. if $A \models C$ and $B \models C$, $A \vee B \models C$,

RCD. if $A :\approx C$ and $B :\approx C$, then $A \vee B :\approx C$

The key observation is that while CD states a true claim about the logic, RCD does not. Although every entailment is a reasonable inference, it is not the case that, for every true claim about entailment, there is a corresponding true claim about reasonable inference. The observation that constructive dilemma behaves in this unusual way for some non-classical consequence relations that extend classical validity is important and it reappears in different guises in much of the contemporary discussion of information-sensitive modals.[17]

Another noteworthy item is that the dialectic I spelled out at the end of §4, resurfaces here. The steps (2) $\Rightarrow$ (3) and (5) $\Rightarrow$ (6) both have the form of TRUE CONSEQUENT. As we know, TRUE CONSEQUENT is not a reasonable inference. Stalnaker acknowledges this (fn. 17) and claims that the steps are reasonable in the Strawsonian sense (i.e. according to $:\approx^{S}$). As we know from Facts 9-10 this is only true if we stipulate ANTECEDENT COMPATIBILITY. In fact, IC's account of the idle argument *requires* this particular combination ($:\approx^{S}$ and ANTECEDENT COMPATIBILITY). Given the essential role of that combination in capturing CONTRAPOSITION we likely should characterize it as the "official" version of IC's pragmatics.

Acknowledgments

Special thanks to Simon Goldstein and Paolo Santorio for conversations and exchanges on the ins and outs of this dialectic. Thanks to Matt Mandelkern, and Zoltán Gendler Szabó for comments on the manuscript. Also thanks to David Boylan, Justin Khoo, John MacFarlane, Robert Van Rooij and Malte Willer for conversations and correspondences on these topics. Finally thanks to all three editors for the invitation and the volume's conception.

Appendix

Fact 1 *Given the framework,* PRAGMATIC CONSTRAINT *entails* C-ACCEPTANCE

> *Proof:* Suppose that c is incompatible with $\neg$A & $\neg$C and c is compatible with $\neg$A & C. Now consider any world w in c. We must show that the closest $\neg$A-world to w is a C-world. By the pragmatic constraint, the $\neg$A & C-world v that is in the context set must be closer to w than any $\neg$A & $\neg$C world. Now, because Stalnaker's selection functions are grounded by an ordering of similarity, the closest A-world is either the closest $\neg$A & C-world or the closest $\neg$A & $\neg$C-world. But it cannot be the latter, so it must be the closest $\neg$A & C-world. □

[17] Some key references for the discussion of constructive dilemma in, and around, dynamic semantics: Veltman (1996); Yalcin (2012); Willer (2012); Bledin (2014, forthcoming); Marra (2014); Moss (2015); Charlow (ms.); Goldstein (ms.).

Fact 2 *If entailment is defined as preservation of truth at a world and a context,* $A \vee C \not\models \neg A \rightarrow C$.

Fix c and w so that $\mathsf{A} \vee \mathsf{C}$ is true at w but also so that, in the closest $\neg\mathsf{A}$-world to w, C is false. □

Fact 3 *However,* $\rightarrow$ *satisfies* C-ACCEPTANCE.

Fix c and w. Suppose (i) $c \cap |\neg\mathsf{A}| \cap |\neg\mathsf{C}| = \emptyset$ and (ii) $c \cap |\neg\mathsf{A}| \cap |\mathsf{C}| \neq \emptyset$. Note that (i) alon is enough to guarantee $\neg\mathsf{A} \rightarrow \mathsf{C}$ at every world within c which is what needed to be shown. □

Fact 4 *if* $\Sigma \models C$ *then* $\Sigma :\approx C$

Fix a c in which every member of Σ is assertible. Let c_Σ be the context resulting from the sequential update of c with every premise in Σ. Let $|\Sigma|$ denote the set $\bigcap\{|\mathsf{A}| \mid \mathsf{A} \in \Sigma\}$. Because $c_\Sigma \subseteq |\Sigma|$ and $|\Sigma| \subseteq |\mathsf{C}|$, $c_\Sigma \subseteq |\mathsf{C}|$, which is what we need to show to establish that $\rightarrow$ satisfies the analogue of C-ACCEPTANCE. □

Fact 5 *if* A *presupposes* C*, then* $A :\approx C$

Suppose c is an arbitrary initial context in which A is assertible. Since the assertibility conditions for A include C, we must have $c \subseteq |\mathsf{C}|$. Since $c_\mathsf{A} \subseteq c$, we must have $c_\mathsf{A} \subseteq |\mathsf{C}|$. □

Fact 6 (intransitivity) *Let* Σ *be a set of sentences.* $\Sigma :\approx B$ *and* $B :\approx C$ *do not guarantee* $\Sigma :\approx C$.

[proven in the main text]

Fact 7

(i) $\neg A :\approx \neg A \vee C$

(ii) $\neg A \vee C :\approx A > C$ *but*

(iii) $\neg A :\not\approx A > C$

[proven in the main text]

Fact 8 *The reasonable-inference analogue of* CONTRAPOSITION *fails. (i.e.* $A > C :\not\approx \neg C > \neg A$*).*

This failure can be illustrated along the same lines as the failure of part (iii) of Fact 7. Let c be a context in which A > C's presuppositions are satisfied. Suppose that c only contains C-worlds, but also that there are some ¬C-worlds in the background model. Suppose however that all such ¬C-worlds are A-worlds. Let c' be the result of updating c with A > C. Now, consider the evaluation of ¬C > ¬A in c': since there are no ¬C-worlds in c', the selection function must reach outside of it to find some ¬C-worlds. However, such worlds are A-worlds, so c' does not accept ¬C > ¬A. □

Fact 9 *If the presupposition of conditionals is* ANTECEDENT COMPATIBILITY, *then* CONTRAPOSITION, TRUE CONSEQUENT *and* FALSE ANTECEDENT *are all Strawsonian Reasonable Inferences.*

Let c be a context in which A > C's presuppositions are satisfied. Let c' be the result of updating c with A > C. Because of Stalnaker's commitment to strong centering (if w makes A true, w is the closest A-world to itself), part of the effect of the update from c to c' is to rule out all the A & ¬C-worlds.

Now suppose c' satisfies the presuppositions of ¬C > ¬A. This must mean that c' contains a ¬C-world v. But we just argued that the A & ¬C-worlds have been excluded in the transition from c to c'. So v must be a ¬A & ¬C-world, and indeed this must be the case for any ¬C-world in v. Since the worlds in c' are to be considered closer than the worlds out of c', v (or some world relevantly like it) must be the closest ¬C-world. So the closest ¬C-world is a ¬A-world.

This same argument, however, will go through for TRUE CONSEQUENT and FALSE ANTECEDENT. □

Fact 10 *If the presupposition is just* PRAGMATIC CONSTRAINT, *then none of* CONTRAPOSITION, TRUE CONSEQUENT *and* FALSE ANTECEDENT *count as Strawsonian Reasonable Inferences.*

The countermodel we constructed in establishing Fact 8 vacuously satisfies PRAGMATIC CONSTRAINT in the updated context. (Similarly for analogous countermodels for TRUE CONSEQUENT and FALSE ANTECEDENT.) □

Bibliography

Adams, Ernest W (1975), *The Logic of Conditionals* (Springer).

Bennett, Jonathan (2003), *A Philosophical Guide to Conditionals* (Clarendon Press).

Bledin, Justin (2014), 'Logic informed', *Mind*, 123(490), 277–316.

——— (forthcoming), 'Fatalism and the Logic of Unconditionals', *Nous*.

Cariani, Fabrizio (2016), 'Deontic Modals and Probabilities: One Theory to Rule Them All?', in N. Charlow and M. Chrisman (eds.), *Deontic Modals*, 11–46 (Oxford University Press).

Cariani, Fabrizio and Goldstein, Simon (forthcoming), 'Conditional heresies', *Philosophy and Phenomenological Research*.

Cariani, Fabrizio, Kaufmann, Magdalena, and Kaufmann, Stefan (2013), 'Deliberative Modality under Epistemic Uncertainty', *Linguistics and Philosophy*, 36, 225–259.

Charlow, Nate (2013), 'What We Know and What To Do', *Synthese*, 190, 2291–2323.

——— (ms.), 'Modus Ponens and the Logic of Dominance Reasoning', University of Toronto.

Dudman, Victor H (1983), 'Tense and time in English verb clusters of the primary pattern', *Australian Journal of Linguistics*, 3(1), 25–44.

——— (1984), 'Conditional interpretations of if-sentences', *Australian journal of linguistics*, 4(2), 143–204.

Dummett, Michael (1964), 'Bringing about the past', *The Philosophical Review*, 73(3), 338–359.

Edgington, Dorothy (1995), 'On Conditionals', *Mind*, 104(414), 235–329.

von Fintel, Kai (1997), 'Bare Plurals, Bare Conditionals, and *Only*', *Journal of Semantics*, 14, 1–56.

Gillies, Anthony S. (2004), 'Epistemic Conditionals and Conditional Epistemics', *Noûs*, 38(4), 585–616.

——— (2009), 'On truth-conditions for *if* (but not quite only *if*)', *Philosophical Review*, 118(3), 325–349.

——— (2012), 'Indicative Conditionals', in Delia Graff and Gillian Russell (eds.), *Routledge Companion to Philosophy of Language* (Routledge).

——— (forthcoming), 'Conditionals', in Alex Miller Bob Hale and Crispin Wright (eds.), *Blackwell Companion to the Philosophy of Language, Blackwell* (Oxford University Press).

Goldstein, Simon (ms.), 'Free Choice and Homogeneity', Lingnan University.

Grice, Paul (1989), *Studies in the Way of Words* (Cambridge MA: Harvard University Press).

Hintikka, Jaako (1962), *Knowledge and Belief: an Introduction to the Logic of the Two Notions* (Cornell University Press).

Iatridou, Sabine (2000), 'The grammatical ingredients of counterfactuality', *Linguistic inquiry*, 31(2), 231–270.

Ippolito, Michela (2013), *Subjunctive Conditionals* (MIT Press).

Jackson, Frank (1987), *Frank Jackson* (Oxford, England: Blackwell).

Khoo, Justin (2015), 'On indicative and subjunctive conditionals', *Philosophers Imprint*.

Kolodny, Niko and MacFarlane, John (2010), 'Ifs and Oughts', *Journal of Philosophy*, 107(3), 115–143.

Kratzer, Angelika (1991), 'Conditionals', in A. von Stechow & D. Wunderlich (ed.), *Semantics: An International Handbook of Contemporary Research* (De Gruyter). From the *Semantics Archive*.

——— (2012), *Modals and Conditionals* (Oxford University Press).

Lewis, David (1973), *Counterfactuals* (Blackwell).

——— (1976), 'Probabilities of conditionals and conditional probabilities', *The Philosophical Review*, 85, 297–315.

Lycan, William G (1993), 'MPP, Rip', *Philosophical perspectives*, 7, 411–428.

Lycan, William H (2001), *Real Conditionals* (Oxford University Press).

Mandelkern, Matthew (2018), 'The case of the missing 'If': Accessibility relations in Stalnaker's theory of conditionals', *Semantics & Pragmatics*.

Marra, Alessandra (2014), 'For a Dynamic Semantics of Necessity Deontic Modals', in *International Conference on Deontic Logic in Computer Science*, 124–138 (Springer).

McGee, Vann (1985*a*), 'A Counterexample to *Modus Ponens*', *Journal of Philosophy*, 82, 462–471.

——— (1985*b*), 'A Counterexample to Modus Ponens', *The Journal of Philosophy*, 82(9), 462–471.

Moss, Sarah (2015), 'On the semantics and pragmatics of epistemic vocabulary', *Semantics and Pragmatics*, 8(5), 1–81.

Nolan, Daniel (2003), 'Defending a Possible-Worlds Account of Indicative Conditionals', *Philosophical Studies*, 116(3), 215–69.

Smiley, Timothy (1967), 'Mr. Strawson on the Traditional Logic', *Mind*, 76(301), 347–385.

Stalnaker, Robert (1968), 'A theory of conditionals', *Studies in Logical Theory, American Philosophical Quarterly*, Monograph Series 2, 98–112.

——— (1974), 'Pragmatic Presuppositions', in M.K. Munitz and P. Unger (eds.), *Semantics and Philosophy* (NYU Press).

——— (1975), 'Indicative Conditionals', *Philosophia*, 5, 269–86.

——— (1978), 'Assertion', in *Context and Content* (Oxford University Press).

——— (1981), 'A Defense of Conditional Excluded Middle', in William L. Harper, Robert Stalnaker, and Glenn Pearce (eds.), *IFS: Conditionals, Belief, Decision, Chance and Time*, volume 15 of *The University of Western Ontario Series in Philosophy of Science*, 87–104 (Springer Netherlands).

——— (1984), *Inquiry* (MIT Press (Bradford Books)).

——— (2002), 'Common ground', *Linguistics and philosophy*, 25(5-6), 701–721.

——— (2016), *Context* (Oxford University Press).

STALNAKER, Robert and THOMASON, Richmond H. (1970), 'A Semantic Analysis of Conditional Logic', *Theoria*, 36(1), 23–42.

STARR, William B. (2014), 'A Uniform Theory of Conditionals', *Journal of Philosophical Logic*, 43(6), 1019–1064.

STRAWSON, P. F. (1952), *Introduction to Logical Theory* (London: Methuen).

VAN BENTHEM, Johan (1996), *Exploring Logical Dynamics* (CSLI Publications).

VELTMAN, Frank (1996), 'Defaults in Update Semantics', *Journal of Philosophical Logic*, 25, 221–261.

WILLER, Malte (2012), 'A Note on *Iffy* Oughts', *Journal of Philosophy*, 109, 449–461.

YALCIN, Seth (2007), 'Epistemic Modals', *Mind*, 116(4), 983–1027.

——— (2012), 'A Counterexample to *Modus Tollens*', *Journal of Philosophical Logic*, 41, 1001–1024.