Alternatives and Truthmakers in Conditional Semantics

Paolo Santorio
UNIVERSITY OF LEEDS

Draft of September 7, 2015

1 Introduction: a trilemma

My starting point is a classical puzzle about counterfactuals in natural language. Counterfactuals seem to satisfy three logical constraints that are individually plausible, but jointly inconsistent.

Constraint #1. Counterfactuals invalidate Antecedent Strengthening. I.e., one cannot replace the clause appearing in a counterfactual antecedent with a stronger one and preserve truth value.

**Failure of Antecedent Strengthening**

\[ \varphi \rightarrow \psi \not\vdash \varphi^+ \rightarrow \psi \]

(with \( \varphi^+ \models \varphi \), \( \varphi \not\models \varphi^+ \))

The argument for this constraint (Stalnaker 1968, Lewis 1973a, 1973b) is that discourses that exemplify violations of Antecedent Strengthening—so-called Sobel sequences—can be heard as consistent.

(1) If the US threw its weapons into the sea, there would be war.
If the US and all other nuclear powers threw their weapons into the sea, there would not be war.

Constraint #2. Counterfactuals validate Simplification of Disjunctive Antecedents. I.e., a counterfactual with a disjunctive antecedent entails the counterfactuals whose antecedents are the individual disjuncts.

**Simplification**

\[ (\varphi \lor \psi) \rightarrow \chi \models \varphi \rightarrow \chi, \psi \rightarrow \chi \]

The argument for this constraint is that this pattern seems systematically validated by all counterfactuals with disjunctive antecedents. For example:
(2) If Alice or Bob went to the party, the party would be fun.
   a. ~⇒ If Alice went to the party, the party would be fun.
   b. ~⇒ If Bob went to the party, the party would be fun.

**Constraint #3.** Counterfactuals validate Substitution of Logical Equivalents (SLE) in antecedent position. I.e., replacing an antecedent with a logically equivalent antecedent preserves truth value.

**Substitution**
\[ \varphi \rightarrow \psi \models \varphi' \rightarrow \psi' \]
(with \( \varphi \) and \( \varphi' \) logically equivalent)

The argument for this constraint is theoretical, rather than empirical. Possible worlds semantics provides an elegant account of counterfactuals, which fits well into a general account of linguistic modality (Kratzer 1981a, 1981b, 1986, 1991, 2012). But this semantics is intensional, i.e. validates the replacement of necessarily equivalent clauses in all positions. A fortiori, it validates Substitution.

Unfortunately, if we hold on to a Boolean semantics for disjunction (i.e., if we take ‘or’ to mean ‘\( \lor \)’), Simplification and Substitution immediately entail Antecedent Strengthening. Hence the three constraints are inconsistent and at least one of them must go. The standard solution consists in retaining Failure of Antecedent Strengthening and Substitution, and jettisoning Simplification. Counterfactuals with disjunctive antecedents like (2) are an acknowledged problem, but it is assumed that they can be accommodated via a local fix.

This paper argues that no local fix will do, and that the trilemma should push us to reconsider some features of the semantics of conditionals. In particular, the correct semantics for conditionals is hyperintensional and hence invalidates Substitution. Similarly to several nonstandard accounts, this semantics makes use of a notion of a truthmaker. But, differently from other truthmaker accounts, this notion of a truthmaker is cognitive rather than metaphysical and is defined exclusively via linguistic means.

The key idea behind the account is that conditionals are alternative-sensitive. It is widely agreed that natural language includes expressions and mechanisms

---

1 Proof:
\[ i \quad \varphi \rightarrow \psi \]
\[ ii \quad (\varphi \lor \varphi^+) \rightarrow \psi \quad \text{(from i, by SLE)} \]
\[ iii \quad \varphi^+ \rightarrow \psi \quad \text{(from ii, by SDA)} \]

As will become clear in the paper, the right semantics for disjunction is not what is at stake, so I won’t question the Boolean account of or here.
that manipulate alternatives to the linguistic material that is pronounced. One uncontroversial example is only. Consider:

(3) Only Alice came to the party.

On standard analyses, (3) presupposes that Alice came to the party, and asserts that none of a set of alternative individuals did. This is captured by letting only manipulate a set of alternatives. Roughly, alternatives are clauses that are generated from the material that is pronounced by replacing parts of it: for the case of (3), relevant alternatives may be Bob came to the party, Cynthia came to the party, etc. Similar mechanisms relying on alternatives inform contemporary accounts of focus, scalar implicature, and so-called free choice effects.

I argue that, similarly, conditional antecedents are alternative-sensitive. In particular, I use alternative-sensitive mechanisms to define a set of propositions—ways for the antecedent to be true, or truthmakers—that are denoted by conditional antecedents. The resulting truth conditions are, on a rough pass:

\[ \varphi \square \rightarrow \psi \equiv true \text{ iff: the propositions } p_1, p_2, \ldots, p_n \text{ that are ways for } \varphi \text{ to be true are such that the closest } p_1^-, p_2^-, \ldots, p_n^- \text{-worlds make } \psi \text{ true.} \]

This semantics makes conditionals hyperintensional, since intensionally equivalent sentences can have different alternatives (via their different syntactic structure). But this kind of hyperintensionality is very different from that postulated by existing truthmaker accounts. On the one hand, it is well-understood and independently needed. On the other, it can be combined with standard tools from possible worlds semantics for conditionals.

My semantics drops Substitution but doesn’t quite vindicate Simplification (though it gets close). This is also a welcome feature, since it allows my semantics to accommodate examples that are vexing for other truthmaker accounts.

A word about the scope of the paper: issues concerning Simplification have been discussed mostly in the literature on counterfactuals. But the data naturally generalize to all kinds of conditionals. Here I assume that all conditionals have a structurally uniform semantics, and that hence my claims apply to conditionals across the board. If you disagree, you may take all that I say as applying specifically to counterfactuals.

After setting up some background in section 2, I lay out a number of problems, all related to disjunctive antecedents, in section 3. Section 4 and 5 develop the positive account, and section 6 closes the paper with a brief theoretical discussion.
2 Background: comparative closeness semantics

In this section, I introduce a basic possible worlds semantics for conditionals. I follow closely classical ordering semantics for counterfactuals as formulated by Stalnaker (Stalnaker 1968) and Lewis (1973a, 1973b). Contemporary semantics for conditionals often diverge from ordering semantics in several ways, but the differences are irrelevant for my purposes.2

The key element of Stalnaker/Lewis semantics is a relation of comparative closeness \( \preceq_w \). \( \preceq_w \) compares worlds with respect to their closeness to a benchmark world \( w \): \( w' \preceq_w w'' \) says that \( w' \) is closer to \( w \) than \( w'' \). Both Stalnaker and Lewis take \( \preceq_w \) to be a total preordering: \( \preceq_w \) is transitive, reflexive, and total (in the sense that it is defined over all pairs of worlds). The basic function of \( \preceq_w \) is singling out a set of worlds that verify the antecedent and that at the same time are ‘maximal’, i.e. are such that no other world is more similar to \( w \) then they are. Conditionals quantify over the maximal set of worlds so individuated. Using, as is standard, ‘[ ]’ and ‘ ]’ for the interpretation function, here are schematic truth conditions:3

\[
\left[ \varphi \square \rightarrow \psi \right] \preceq_w \text{true if and only if for all } w' \in \max_{\preceq_w} \{ w' : \left[ \varphi \right] \preceq_{w'} \text{true} \}, \left[ \psi \right] \preceq_{w'} \text{true (where } \max_{\preceq_w} \{ w' : \left[ \varphi \right] \preceq_{w'} \text{true} \} \text{ is the set of closest } \varphi\text{-worlds).}
\]

This says: all the maximally close \( \preceq_w \) \( \varphi \)-worlds are \( \psi \)-worlds. For shorthand, we can say that each counterfactual antecedent selects, on the basis of \( \preceq_w \), a set of worlds it quantifies over.

This semantics produces an elegant account of Sobel sequences. Consider:

(1) If the US threw its weapons into the sea, there would be war.
    If the US and all other nuclear powers threw their weapons into the sea, there would not be war.

The set of maximally close US-throwing-weapons worlds need not overlap

---

2 Some versions of comparative closeness semantics (e.g. Veltman 1976, or Kratzer 1981a, Kratzer 1981b, 1986, 1991, 2012) exploit so-called premise sets rather than a closeness ordering. Others (for example, von Fintel 2001 and Gillies 2007) remove some crucial elements from the semantics proper and place them in a dynamic account of contextual information. These differences won’t matter for my purposes. Premise semantics for counterfactuals are translatable into ordering semantics (precisely, a subtype of ordering semantics is equivalent to premise semantics—see Lewis 1981), and dynamic accounts suffer from the same problems I raise.

3 This is an approximation to both of Stalnaker and Lewis’s accounts. For Stalnaker, the ordering singles out, for each world \( w \), a unique world \( w' \) that is closest to it. Lewis rejects the so-called limit assumption, i.e. the assumption that there is a \( \preceq_w \)-maximal set of antecedent worlds.
with the set of maximally close US-and-other-nuclear-powers-throwing-weapons worlds. In particular, the latter might be farther off than the former:

![Diagram](image)

This situation allows for both conditionals in (1) to be true, since the two antecedents select distinct domains of quantification. Hence the discourse in (1) is consistent.

Throughout the paper, I also make assumptions about the syntactic structure of conditionals. With Kratzer (1981a, 1981b, 1986, 1991, 2012), I assume that all conditionals are modalized statements. The *if*-clause is used to restrict the background domain of quantification of the modal, which is usually called *modal base*. For example, the structure of the first conditional in (1) is:

(4) \[ \text{if the US threw its weapons into the sea} \] \[ \text{would there be war?} \]

The modal *would* has two propositional arguments: one is the proposition expressed by the *if*-clause, the other the proposition expressed by the consequent clause (usually called *prejacent*).

There is a large literature on how closeness should be interpreted for various kinds of conditionals. These questions are orthogonal to all my main points in this paper, so I ignore them throughout.

---

4 Some classical papers for the case of counterfactuals are Fine 1975, Jackson 1977, Lewis 1979; see Bennett 2003 for a useful overview. For some recent literature on the topic, see Schaffer 2004, Williams 2008.
3 Antecedent Distribution

3.1 Disjunctive antecedents

Recall from the introduction: a conditional with disjunctive antecedents seems to entail the conditionals with the two individual disjuncts as antecedents:

(2) If Alice or Bob went to the party, the party would be fun.
   a. —if Alice went to the party, the party would be fun.
   b. —if Bob went to the party, the party would be fun.

I use the label ‘Antecedent Distribution’ for the phenomenon exemplified by (2)–(2-b) (after the intuition that the conditional is ‘distributed’ over the disjuncts). The intuitions supporting Antecedent Distribution can be sharpened by considering the corresponding Sobel sequences:

(5)  #If Alice or Bob went to the party, the party would be fun.
     If Bob went, the party would be dreary.

The infelicity of (5) is unexpected on closeness semantics. Just assume that the closest Anna-going-to-the-party worlds are all closer than the closest Bob-going-to-the-party-worlds. Then the set selected by the two counterfactuals in (5) are disjoint, hence (5) is predicted to be consistent.

There is broad consensus that closeness semantics is correct and that these data should be accommodated via a local fix. In this section, I consider and discard the two main options for a solution of this sort.

3.2 Against pragmatic accounts

A first option for a local fix is a pragmatic account. The natural suggestion is that (somehow or other) pragmatic reasoning generates the following assumption when a conditional with disjunctive antecedents is uttered:

---

5 The problem was noticed independently by Fine 1975 and Nute 1975, and has been discussed extensively. For some treatments of disjunctive antecedents within a classical or minimally modified comparative closeness framework, see Lewis 1977, Nute 1980, Klinedinst 2007 and 2009, Alonso-Ovalle 2009.

6 Early discussions of the problem (for example, Lewis 1977 and McKay & Van Inwagen 1977), often endorse a different solution, i.e. that we just regiment conditionals with disjunctive antecedents, à la Quine, as conjunctions of conditionals. Since then, the goal of a theory of counterfactuals has shifted from regimentation to a genuine compositional semantics. Of course, in the latter context this solution is a non-starter.
Diversity Condition: The worlds that count as closest for the purposes of evaluating a conditional of the form $\Box \varphi \lor \psi \rightarrow \chi^{-}$ include both $\varphi$- and $\psi$-worlds.

(Deriving the Diversity Condition in a principled way is not trivial, and may actually require some changes to conditional semantics. But I won't worry about this here.) If the Diversity Condition is satisfied, standard conditional semantics ensures that (2) entails (2-a) and (2-b) (more on this in a few pages).

One argument for a pragmatic account is that Antecedent Distribution is not without exception. Consider:

(6) If Spain had fought with the Axis or the Allies, she would have fought with the Axis.  
(McKay & Van Inwagen 1977)

It seems absurd to claim that (6) entails:

(7) If Spain had fought with the Allies, she would have fought with the Axis.

Despite this, I am going to insist that Antecedent Distribution is a semantic, and not a pragmatic phenomenon. I give three arguments below. I get back to examples like (6) in section 5.

3.2.1 Argument #1: downward entailing environments

My first argument concerns so-called downward entailing (henceforth, DE) environments. DE environments are linguistic environments that reverse the direction of entailment. Classical examples are negation and verbs like doubt.

(8) a. Jane runs $\Rightarrow$ Jane moves  
b. Jane doesn't move $\Rightarrow$ Jane doesn't run  
c. I doubt that Jane moves $\Rightarrow$ I doubt that Jane runs

DE environments provide a test for distinguishing semantic and pragmatic phenomena. In particular, pragmatic effects (understood, broadly, as optional strengthening effects that go beyond the basic meaning of a sentence) usually

---

7 The state-of-the-art account in this strand is Klinedinst’s (2007, 2009), who switches to a semantic where modals quantify over pluralities of worlds. My account is very different, but I am indebted to Klinedinst for pointing out the relevance of the semantics of plurals to conditionals.

8 Much recent literature treats phenomena that are standardly thought to be within the purview of the pragmatics as generated by compositional processes. One salient example is scalar implicature (see, among many, Chierchia 2004, 2013, Fox 2007, Chierchia et al. 2008). For current purposes, I count approaches of this sort as pragmatic, since they involve strengthening operations that go beyond basic meaning.
disappear in DE environments.

Scalar implicatures (Grice 1975, Gazdar 1979, Sauerland 2004) are a case in point. Scalar implicature is a strengthening effects that involves conjoining a sentence with the negation of one or more stronger alternatives (I say more about what alternatives are in play in section 4). A typical example is the exclusivity implicature one gets for disjunctive sentences: \( p \lor q \) get strengthened with the negation of \( p \land q \). For example, (9-a) is normally read as having the stronger meaning in (9-b).

\[
\begin{align*}
(9) & \quad \quad a. \quad \text{Jane talked to Mary or Sue.} \\
& \quad \quad b. \quad \approx \text{Jane talked to exactly one of Mary and Sue.}
\end{align*}
\]

But the stronger meaning disappears under a DE operator.

\[
\begin{align*}
(10) & \quad \quad a. \quad \text{It's not the case that Jane talked to Mary or Sue.} \\
& \quad \quad b. \quad \not\approx \text{It's not the case that Jane talked to exactly one of Mary and Sue.}
\end{align*}
\]

\[
\begin{align*}
(11) & \quad \quad a. \quad \text{I doubt that Jane talked to Mary or Sue.} \\
& \quad \quad b. \quad \not\approx \text{I doubt that Jane talked to exactly one of Mary and Sue.}
\end{align*}
\]

(10-a) is not read as having the meaning in (10-b). If it did, then it could be used to say that John either talked to none or both of John and Mary—which is obviously not the case. Similarly, mutatis mutandis, for (11-a).

From a theoretical perspective, the disappearance of scalar implicatures in DE environments is fully expected. Scalar implicatures are optional effects that aim at increasing the information carried by a sentence. But, in environments where the direction of entailment is reversed, they would produce an overall weakening. Hence speakers have a preference for not competing implicatures in these cases.

So-called free choice effects also disappear under DE operators. Free choice consists in the ‘distributive’ interpretation of certain linguistic phrases in the scope of existential quantifiers.\(^9\) A classical illustration involves, again, disjunction:

\[
\begin{align*}
(12) & \quad \quad \text{Mary may go to Paris or Berlin.} \\
& \quad \quad a. \quad \rightarrow \text{Mary may go to Paris.} \\
& \quad \quad b. \quad \rightarrow \text{Mary may go to Berlin.}
\end{align*}
\]

Alternatives and Truthmakers in Conditional Semantics

There is no agreement on the proper account of free choice, but most theorists converge on the idea that it is a kind of implicature. The main argument (Kratzer & Shimoyama 2002, Alonso-Ovalle 2005, Fox 2007) is just that it disappears in DE environments:

(13) a. It's not the case that Mary may go to Paris or Berlin.
    b. \( \not\approx \) It's not the case that: Mary may go to Paris and she may go to Berlin.

(14) a. I doubt that Mary may go to Paris or Berlin.
    b. \( \not\approx \) I doubt that: Mary may go to Paris and she may go to Berlin.

Now, my first argument is simple: differently from implicatures and free choice effects, Antecedent Distribution does not disappear in DE environments.

(15) It's not the case that, if Alice or Bob went, the party would be fun.
    a. \( \not\rightarrow \) It's not the case that, if Alice went, the party would be fun.
    b. \( \not\rightarrow \) It's not the case that, if Bob went, the party would be fun.

(16) I doubt that, if Alice or Bob went, the party would be fun.
    a. I doubt that, if Alice went, the party would be fun.
    b. I doubt that, if Bob went, the party would be fun.

This is surprising on a pragmatic view, and expected on a semantic view, of Antecedent Distribution. Incidentally, let me notice that the distributive mechanism exemplified by (15) and (16) is slightly different from what we observed in unembedded conditionals. If the conditional in (say) (16) was 'distributed' over the disjuncts, as it happens for (2), it would mean:

(17) I doubt the following: it is both the case that, if Alice went to the party, the party would be fun, and that, if Bob went to the party, the party would be fun.

which is not what (16) says. I get back to this in section 5.

3.2.2 Argument #2: probability operators

Antecedent Distribution happens also in probably-conditionals:

(18) If Alice or Bob went to the party, probably Mary went too.
    a. \( \not\rightarrow \) If Alice went to the party, probably Mary went too.
    b. \( \not\rightarrow \) If Bob went to the party, probably Mary went too.
Let me make some assumptions about probably-conditionals. Following Kratzer (1981b, 1991, 2012), I assume that if-clauses work as restrictors of the domain of quantification (modal base) of modals. Moreover, following recent work (Yalcin 2010, Lassiter 2011, Holliday & Icard 2013), I assume that probably has a probabilistic semantics (or at least a semantics that yields an equivalent logic). Roughly, $\Gamma$ probably $\varphi \rightarrow \Gamma$ says that the probability of $\varphi$ is higher than .5.

These assumptions result in intuitive truth conditions for probably-conditionals. $\Gamma$If $\varphi$, probably $\psi \rightarrow \Gamma$ simply says that the conditional probability of $\psi$, given $\varphi$, is higher than .5.10

Given this setup, Antecedent Distribution cannot be derived on standard accounts, even with the aid of pragmatic machinery. Suppose that, thanks to some kind of pragmatic reasoning, we are able to derive the counterpart of the Diversity Condition for probably-conditionals:

**Diversity Condition**: The worlds quantified over by a conditional of the form $\Gamma$ if $\varphi$ or $\psi$, probably $\chi \rightarrow \Gamma$ include both $\varphi$- and $\psi$-worlds.

For the case of most conditionals, including counterfactuals, the Diversity Condition is sufficient to derive Antecedent Distribution. This derivation is guaranteed by a special property of closeness orderings: if the closest $\varphi$-or-$\psi$-worlds include a set of $\varphi$-worlds, then the latter also count as the set of closest $\varphi$-worlds (ditto for $\psi$-worlds). Call this property ‘Stability’. Given the Diversity Condition, Stability guarantees that, if all closest $\varphi$-or-$\psi$-worlds make $p$ true, the closest $\varphi$-worlds also make $p$ true.11

But no version of Stability holds for probably-conditionals. It might be that, within the set of $\varphi$-or-$\psi$-worlds, the $p$-worlds are assigned a greater amount of probability than the non-$p$-worlds and that, at the same time, within the set of $\varphi$-worlds, the non-$p$-worlds have greater probability. Here is an example:

It’s .7 likely that Alice goes to the party and .3 likely that Bob goes; moreover, the intersection of these propositions, i.e. the

---

10 This is compatible with the well-known triviality results about the link of probabilities of conditionals and conditional probabilities. (See Hájek & Hall 1994 for an overview.) These results concern the probabilities that we, qua theorists, should assign to bare conditionals. My assumptions concern the truth conditions of object language conditionals involving probably.

11 A formal definition of Stability, in the context of Lewis semantics with limit assumption:

**Stability.** For any sets of worlds $S$, $S'$ s.t. $S' \subseteq S$: if $S' \cap \max_{\preceq_w}(S) \neq \emptyset$, $\max_{\preceq_w}(S') = \max_{\preceq_w}(S) \cap S'$

As Schlenker 2004 points out, Stability generalizes a property entailed by condition 4 of Stalnaker’s 1968 semantics. I borrow the label ‘Stability’ from Charlow 2013, who uses it for an analogous property of deontic selection functions.
proposition that they both go, is .1 likely. Alice is fun, so it’s certain (probability 1) that the party will be a success given that she goes. Bob is extremely awkward and it’s certain (probability 1) that he will make everyone uncomfortable and the party will be a disaster, unless Alice is there to save the day.

On these assumptions, we get the following conditional probabilities:

\[
\begin{align*}
Pr(\text{Fun} | A \lor B) &= .78 \\
Pr(\text{Fun} | B) &= .33
\end{align*}
\]

Given the semantics I’m assuming, (18) is predicted to be true; moreover, the Diversity Condition\(^*\) is satisfied. But (18-b) is false.

Of course, this by itself doesn’t establish that the pragmatic account is wrong for other conditionals. But it shows that this account is not general enough to cover examples that seem obvious instances of the same phenomenon.

3.2.3 Argument #3: nonclosest worlds

My third argument exploits conditional logic. Consider again:

(2) If Alice or Bob came, the party would be fun.

We can prove that, in certain contexts, the set of worlds selected by the antecedent of (2) is the union of two discontinuous segments of the ordering (i.e. two segments such that all worlds in the first segment are strictly closer than all worlds in the second):

This situation is straightforwardly incompatible with closeness semantics, if we hold on to the assumption that conditionals with disjunctive antecedents denote a unique proposition. Hence this is a counterexample to any account that sticks to standard semantics.
The argument requires some setup. First, I borrow a scenario and some judgments from Lewis (1973a, p. 33):

Otto is Waldo's successful rival for Anna’s affections. Waldo still tags around after Anna, but never runs the risk of meeting Otto. Otto was locked up at the time of the party, so that his going to it is a far-fetched supposition; but Anna almost did go.

(19) If Anna had gone to the party, Waldo would have gone. ✓
(20) If Otto had gone to the party, Anna would have gone. ✓
(21) If Otto had gone to the party, Waldo would have gone. ✗

(19)–(21) is one of the triads Lewis uses to show that transitivity (below) is invalid for counterfactuals.

\[
\begin{align*}
\text{Transitivity} & : \varphi \rightarrow \psi, \psi \rightarrow \chi \models \varphi \rightarrow \chi
\end{align*}
\]

Second, I observe that Antecedent Distribution persists also in the backdrop of this scenario. For example, (22) is still infelicitous.

(22) #If Otto or Anna had gone, it would have been a lovely party.
     If Otto had gone, it would have been a dreary party.

Finally, I assume that the Diversity Condition (repeated below) holds for (22) in the relevant context.

Diversity Condition: The worlds that count as closest for the purposes of evaluating a conditional of the form \( \Gamma \varphi \lor \psi \rightarrow \chi \neg \) include both \( \varphi \)- and \( \psi \)-worlds.

Given the infelicity of (22), this seems a natural assumption, and one that the defender of the pragmatic account will be willing to go along with.

The judgments about (19)–(21), together with the Diversity Condition, yield a surprising conclusion. Configurations like (19)–(21) force certain facts about the ordering. In particular, Anna-worlds must be strictly closer than Otto-worlds. This descends from a general fact about transitivity-violating configurations.

Fact. Consider any triplet of counterfactuals of the form:

\[
\begin{align*}
\text{(a)} & : \varphi \rightarrow \psi \\
\text{(b)} & : \psi \rightarrow \chi \\
\text{(c)} & : \varphi \rightarrow \chi
\end{align*}
\]
Alternatives and Truthmakers in Conditional Semantics

if (a), (b) are true and (c) false, then the closest $\psi$-worlds must be strictly closer than the closest $\varphi$-worlds.\footnote{12}

Hence the judgments on (19)–(21) require that the closest Anna-worlds must be closer than the closest Otto-worlds. Now consider again:

$$
\text{(23) If Otto or Anna had come, it would have been a lovely party.}
$$

By the Diversity Condition, (23) quantifies over both Otto and Anna worlds. Hence (23) quantifies over discontinuous segments of the ordering:

\begin{proof}
I consider a Lewis semantics without limit assumption; the result follows immediately for stronger semantics, including a Lewis-style semantics with limit assumption and Stalnaker semantics. The truth conditions of (a)–(c) are, respectively:

\begin{enumerate}
\item[(a)] $\exists v \in \varphi : \forall u ((u \preceq_w v) \supset u \in (\varphi \supset \psi))$
\item[(b)] $\exists v \in \psi : \forall u ((u \preceq_w v) \supset u \in (\psi \supset \chi))$
\item[(c)] $\exists v \in \varphi : \forall u ((u \preceq_w v) \supset u \in (\varphi \supset \chi))$
\end{enumerate}

Now, consider (b): it says that there is a $\psi$-world (call it '$w_\psi$') such that all worlds at least as close as it make true the material conditional $\Box \psi \supset \chi$. In the Lewis framework, the claim stated in \textbf{FACT} translates as follows:

For all $\varphi$-worlds $w'$: $w_\psi \precsim_w w'$

Suppose, for \textit{reductio}, that this is not the case. Then there is a $\varphi$-world, call it '$w^*$', such that $w^* \preceq_w w_\psi$. Now, consider (a): it says that there is a $\varphi$-world such that all worlds at least as close as it validate the material conditional $\Box \varphi \supset \psi$. There are two cases: either (i) $w^*$ is such a world; or (ii) there is some $w' \preceq_w w^*$ that is such a world. In either case, we have that there is a $w_\varphi$ that works as a witness of (b), and such that $w_\varphi \preceq_w w^*$. By the transitivity of $\preceq_w$, we have that $w_\varphi \preceq_w w_\psi$. Now, by (a), we know that all worlds at least as close to $w_\varphi$ validate $\Box \varphi \supset \psi$. Since $w_\varphi \preceq_w w_\psi$, we know that those worlds also validate $\Box \psi \supset \chi$. But then, by the transitivity of $\supset$, we have that all worlds at least as close to $w_\varphi$ validate $\Box \varphi \supset \chi$; i.e., $\forall u ((u \preceq_w w_\varphi) \supset u \in (\varphi \supset \chi))$. But now, by existential generalization on $w_\varphi$, we get exactly (c), i.e. the truth condition of $\Box \varphi \supset \psi$. Hence $\Box \varphi \supset \psi$ is true after all. \textbf{Contradiction}.

I should note that the result does not follow for a Kratzer-style semantics with partial orderings. But, in that case, the point can be made by replacing (21) with:

\begin{enumerate}
\item[(i)] If Otto had gone to the party, Waldo would not have gone.
\end{enumerate}

(i) sounds true in Lewis’s party scenario. In Kratzer’s semantics, this gives us strictly more information than the falsity of (21) and allows us to prove that (23) quantifies over separate segments of the ordering. The proof is left as an exercise to the reader.
\end{proof}
Call this the problem of nonclosest worlds. Standard accounts, which are just designed around the idea that conditionals select the closest antecedent worlds, have no way of accommodating nonclosest worlds.

### 3.3 Antecedent Distribution, generalized

A second local fix consists in switching to a non-Boolean semantics for disjunction. In this section I show that this move would be pointless, since Antecedent Distribution extends to a large variety of other lexical items.

#### 3.3.1 Generalizing the problem

Antecedent Distribution happens with all kinds of lexical items. Here is a partial list, with some representative examples of infelicitous Sobel sequences.

**Entailing determiner phrases/pairs of a determiner phrase and a name**

(24) #If a former president confessed to adultery, there would be an uproar.
    If Bill Clinton confessed to adultery, there would not be an uproar.

(25) #If one of my students came to the reading group, they would be lost.
    If my student John came, he would be on top of the material.

(26) #If exactly three students came, the reading group would be interesting.
    If students Elizabeth, Jennifer, and Sarah came, the reading group would not be interesting.

---

13 Proposals in this spirit have been floated in the literature on disjunctive antecedents, in particular, by Alonso-Ovalle 2005, 2009 and van Rooij 2006; see also Aloni 2007 for a similar proposal concerning imperatives.

14 Building on the linguistics literature, I assume that it makes sense to talk about entailment in a generalized way, so that we can cover also lexical items. For a characterization of this notion, see von Fintel 1999.
Alternatives and Truthmakers in Conditional Semantics

(27) #If most enrolled students came, we would have a lively discussion. If exactly seven [out of ten] enrolled students came, we would have a terrible discussion.

Entailing names

(28) ??If John had gone to France, he would have had a great holiday. If John had gone to Avignon, he would have had a terrible holiday.

(29) ??If John lived in New York City, his apartment would be within 10 minutes of an upscale coffee shop. If John lived in the Bronx, his apartment would not be within 10 minutes of an upscale coffee shop.

Entailing nouns

(30) ??If John had a pet, he would be happy. If he had a dog, he would be unhappy.

(31) ??If I had a sibling, I would be much happier. If I had a sister, I would be much less happy.

Entailing adjectives

(32) ??If the applicant was European, we could not offer her a scholarship. If the applicant was German, we could offer her a scholarship.

(33) ??If your bag was red, it would not match your jacket. If your bag was crimson, it would match your jacket.

The quality of the data is variable, as indicated. But, for most speakers at least, none of (24)–(33) is fully felicitous. This is a substantial difference with respect to sequences like (1), which are perfectly acceptable without need of contextual setup (more on the quality of the data below). At the same time, closeness semantics predicts that there are scenarios that verify all of (24)–(33).

Call the lexical items that distinguish the two counterfactuals in a Sobel sequence contrast items. The contrast items in classical examples of Sobel sequences are invariably individual conjuncts and a conjunction, or pairs of determiners like some and all. I'm suggesting that we run into problematic data as soon as we consider other contrast items. A precise characterization of what contrast items generate problematic data must wait for the next section. For the moment, let me point to the intuition that contrast items give rise to bad Sobel sequences in cases where one is more specific than the other.
3.3.2 Strengthening the data: the Klecha test

Let me add two clarifications.

First, I am claiming that there is a systematic phenomenon behind the infelicity of Sobel sequences in (24)–(33). Whenever the contrast items in a Sobel sequence are (say) a pair of a more specific and a less specific determiner, or a pair of entailing adjectives, etc., the Sobel sequence is infelicitous (module context shift—more below). The phenomenon depends on syntactic form, and not on specific features of the scenario, or world knowledge that we use to evaluate the counterfactuals. Hence the infelicity of (24)–(33) is due to very different factors from the infelicity of, e.g., the following:

(34) ??If you stirred your coffee, it would taste better.
    If you stirred your coffee with a spoon, it would taste worse.

(34) sounds infelicitous because of world knowledge: using a spoon is the most common way to stir coffee, hence the second conditional sounds surprising after hearing the first. (24)–(33) are infelicitous because of their syntactic form. Their infelicity doesn't depend on what we assume about the background situation.

The contrast can be better appreciated via a test recently devised by Peter Klecha (2014). Kelt points out, following Edgington 1995, that genuine Sobel sequences (as opposed to other sequences of conditionals that may look superficially analogous) are intended to be “single pointful pieces of discourse”. Uttering the second conditional in a Sobel sequence does not call for a retraction of the first, or trigger a shift to a context where the second conditional is no longer taken to hold. This can be seen clearly if we put a Sobel sequence in dialog form:

(35) a. A: If Alice had been there, the party would have been fun.
    b. B: Yes, and if Alice and Bob had been there, the party would have been terrible.
    c. A: Indeed.

The dialog is fully consistent. A and B agree on everything that has been said without the context changing in any relevant way. Now, consider:

(36) a. A: If John stirred his coffee, it would taste better.
    b. B: Yes, and if John stirred his coffee with a spoon, it would taste worse.
    c. A: Indeed.

15 Thanks to [information omitted for blind review] for suggesting this example.
This is also a consistent discourse. To be sure, upon hearing it you may think that A and B have strange beliefs about coffee-stirring. But that’s beside the point. Like (35), (36) is fully consistent; speakers agree on all that has been said without any shifts in context. While their attitudes about stirring coffee may be off, their conversation is perfectly smooth. And now, consider:

(37)  
   a. A: If exactly three students came, the reading group would be interesting.  
   b. B: # Yes, and if students Elizabeth, Jennifer, and Sarah came, the reading group would not be interesting.  
   c. A: # Indeed.

(37) cannot work as a “single pointful piece of discourse”. Upon overhearing this conversation, you would not think that the speakers have bizarre beliefs. You would be thinking that they’re lunatics, or at least not competent with conditionals. The dialog in (37) cannot be felicitous, at least not without modifying in a way that a shift of context is introduced between A’s and B’s utterances.

Second, and relatedly, I am happy to grant that revised versions of (24)–(33) can be uttered felicitously. In particular, (24) and (33) are made felicitous by adding to the second conditional locutions like *but of course*.

For example:

(38)  
   Context: John is an unusually talented student.  
      If one of my students came to the reading group, they would be lost.  
      But of course, if my student John came, he would be on top of the material.

This does not threaten my empirical claims. Locutions like *but of course* are standard devices of context shift.\(^\text{16}\) In (38), the speaker is signaling that they are switching contexts: the modal base of the first conditional didn’t include any John-worlds, which are now being quantified over.

Let me close by sidestepping some residual skepticism about the data. A minority of speakers still find part of the data (i.e. (28)–(33)) unconvincing. Even if this skepticism was grounded, this would not threaten my main claims in the paper, including my positive account. All I need is that Antecedent

---

\(^\text{16}\) For example:

(i) I must either destroy the evidence or else claim that I did it to stop communism. *Though, of course*, there is one other possibility—I could put the public interest first for once. (adapted from Lewis 1978)
Distribution concerns a wide variety of conditionals, and not just conditionals with disjunctive antecedents. The data about quantifiers in (24)–(27) is sufficient to make this point. The exact extent of the phenomenon is secondary, and can be accommodated by modifying moving parts of my positive theory.  

3.4 Taking stock

I have raised three problems for pragmatic accounts of Antecedent Distribution. First, Antecedent Distribution persist where pragmatic effects normally disappear; Second, Antecedent Distribution happens also in probably-conditionals, where pragmatic accounts fail to predict it; third, some conditionals with disjunctive antecedents provably quantify over discontinuous segments of the ordering, contrary to any version of closeness semantics. Finally, I have observed that Antecedent Distribution generalizes far beyond disjunction, showing that it can’t be accounted for by adopting a nonstandard semantics for or.

4 Alternative-sensitivity

My basic proposal is that conditional antecedents denote sets of propositions, each of which specifies a way for the antecedent to be true. These propositions may combine ‘pointwise’ with the main modal in a conditional, giving rise to Antecedent Distribution. Roughly the resulting truth conditions are:

\[ \gamma \varphi \rightarrow \psi \gamma = \text{true iff: the propositions } p_1, p_2, \ldots, p_n \text{ that are ways for } \varphi \text{ to be true are such that the closest } p_1\gamma, p_2\gamma, \ldots, p_n\gamma\text{-worlds make } \psi \text{ true.} \]

This proposal immediately solves all the problems raised in the previous section. First, the distribution effect is semantic, hence it is expected to persist in all linguistic environments. Second, since it is tied to the semantics of if-clauses, we expect the effect to take place in all conditionals, including probably-conditionals. Finally, if the closest worlds verifying each of the two disjuncts are at different distances from the actual world, the proposal allows for quantification over nonclosest worlds.

17 In particular, it can be accommodated by using a different approach to alternatives than the one I use. As I go on to say, this part of the theory is something I take entirely off the shelf. My main contribution is specifying an algorithm that generates truthmakers, starting from a specification of alternatives.
While the basic idea is intuitive, it's unclear how to define, in a systematic and principled way, the notion of a way for a sentence to be true. This section shows how to do this merely on the basis of syntactic alternatives.

### 4.1 Alternatives

I start by giving a precise specification of alternatives. For ease of exposition, I use scalar implicature as my benchmark example of an alternative-sensitive phenomenon. I assume (following the literature) that the same alternatives are in place in alternative-sensitive mechanisms throughout language.

Given a sentence $S$, speakers systematically take some sentences, and not others, to work as alternatives for $S$. For example, a conjunctive sentence like (39)-b normally works as an alternative for a disjunctive sentence like (39)-a.

\begin{align*}
(39) & \quad (a. \text{ Johanna talked to Mary or Sue}) \\
       & \quad (b. \text{ Johanna talked to Mary and Sue})
\end{align*}

This is showed by facts about implicature. Scalar implicature are generated by negating more informative (i.e. stronger) alternatives to a sentence. The scalar implicature normally generated by (39)-a shows that (39)-b works as an alternative to it (and hence is what is negated in order to produce the implicature.)

\begin{align*}
(40) & \quad \text{Johanna talked to Mary or Sue} \\
     & \quad \rightarrow \text{Johanna did not talk to both Mary and Sue}
\end{align*}

Conversely, (39)-c, even though it is a relevant and more informative variant of (39)-a, does not work as an alternative to it.

\begin{align*}
(39) & \quad (c. \text{ Johanna talked to exactly one between Mary and Sue})
\end{align*}

If it did, then an utterance of (39)-a would implicate (via negation of (39)-c) that Johanna talked to both Mary and Sue—which is obviously wrong.\(^{18}\)

The problem of specifying alternatives is the problem of specifying, in a principled way, which sentences work as alternatives of which others. Traditional accounts of alternatives merely stipulate that alternatives are part of the meaning of lexical items. For example, it is part of the lexical meaning of or that or is on a 'lexical scale' that also includes and. Recently, a nonstipulative option has

\(^{18}\) The problem of explaining why (39)-b, but not (39)-c, is an alternative to (39)-a, is called 'symmetry problem'. The symmetry problem was first noticed by Kroch 1972 (standing to the historical information in Fox 2007), and earned its name in class notes by Kai von Fintel and Irene Heim at MIT. For a contemporary statement of the problem, see Sauerland 2004.
emerged, thanks to Katzir 2007. In what follows, I’m going to assume Katzir’s theory of alternatives, though nothing in my account depends on this.

Katzir’s account is built around two principles: relevance and complexity. Roughly, alternatives to a sentence $S$ are all and only those sentences that are relevant in the context and no more complex than $S$. The notion of complexity here is technical. Roughly, $S'$ counts as at most as complex as $S$ just in case we can derive $S'$ from $S$ by either deleting syntactic constituents from $S$, or replacing them with syntactic items that are part of a given substitution source. The substitution source is defined as the union of the whole lexicon with items that have been pronounced in the context. I relegate the precise formulation of Katzir’s algorithm to a footnote.

It’s useful to show how Katzir’s theory deals with the examples in (39). Katzir predicts, correctly, that (39)-b is an alternative to (39)-a, since it can be derived from the latter just by replacing $or$ with $and$. He also predicts that (39)-c is not an alternative to (39)-a. The reason is that (39)-c is more complex than (39)-a—it cannot be derived from (39)-a by replacing $or$ with an item from the lexicon. This, again, is a welcome prediction.

### 4.2 Specificity

This section is the heart of the paper. It describes the algorithm for generating ‘ways for a sentence to be true’—or, for short, truthmakers of a sentence. While the starting point—i.e. what alternatives are in play—is shared with existing literature, the algorithm is entirely new.

I allow myself to be sloppy in two ways. First, while alternatives are syntactic items, sometimes I treat them as propositions. Second, I suppress all reference

---

19 For approaches based on lexical scales (commonly called Horn scales), see Horn 1972, Gazdar 1979, Horn 1989. For a refinement of Katzir’s approach, see Fox & Katzir 2011.

20 First, we define the notion of a structure being at most as complex as another in context $c$ (represented as ‘$\preceq_c$ ’):

\[
S' \preceq_c S \text{ iff } S' \text{ can be derived from } S \text{ by successive replacements of syntactic sub-constituents of } S \text{ with elements of the substitution source for } S \text{ in } c, SS(S, c)
\]

(The notion of a sub-constituent is commonplace in syntax; see e.g. Carnie 2013.) Then we define the notion of a substitution source for a sentence $S$ in a context $c$, as follows:

The substitution source for sentence $S$ in context $c$ is the union of:

i. The lexicon;

ii. the subconstituents of $S$;

iii. the set of salient constituents in $C$. 

Alternatives and Truthmakers in Conditional Semantics

to context. In both cases, the sloppiness is harmless.

Here is the basic proposal. Let $\text{ALT}_S$ be the set of alternatives to $S$. The ways for $S$ to be true are propositions that are (a) stronger than that expressed by $S$, and (b) individuated by the subsets of $\text{ALT}_S$ that are \textit{specific} and \textit{minimal}. Specificity is the key new notion. I say that a subset of $\text{ALT}_S$ is specific iff it is \textit{consistent with the negation of every alternative that is not a member of it}. The intuition is that a set of alternatives is specific just in case it contains enough information to stand alone—even if all other alternatives are false, it’s still consistent to suppose that all sentences in the set are true.

I state formal definitions in a few paragraphs, but let me walk you through an example first:

(41) Otto or Anna went to the party

For the time being, I simply assume that the alternatives to (41) are:

(42) \[
\begin{cases}
\text{Otto or Anna went to the party} & O \lor A \\
\text{Otto went to the party} & O \\
\text{Anna went to the party} & A \\
\text{Otto and Anna went to the party} & O \land A
\end{cases}
\]

(Some extra alternatives may be present—for example, if a third individual, John, is salient in the context, \textit{John went to the party} will be an alternative. I show below that this is irrelevant.) Crucially, the alternatives in (42) can be ordered by logical strength (stronger alternatives are above weaker ones):

\[
\begin{array}{c}
O \land A \\
O \\
O \lor A \\
A
\end{array}
\]

We proceed by checking what the specific subsets of $\text{ALT}_{(41)}$ are. It’s easy to see that these subsets are $\{O \lor A,O\}$, $\{O \lor A,A\}$, and $\text{ALT}_{(41)}$ itself. For a comparison, consider $\{O \lor A\}$: this set is not specific, since it’s inconsistent when supplemented with the negation of all the other alternatives.

Of all the specific subsets of $\text{ALT}_{(41)}$, we take only the minimal ones—i.e. the ones that are not proper supersets of any other specific subset of $\text{ALT}_{(41)}$.\footnote{21 Why minimality? One might think that \textit{all} the specific alternative sets to the antecedent should...}
are left with \( \{O \lor A, O\} \) and \( \{O \lor A, A\} \):

\[
\begin{array}{c}
O \land A \\
O \lor A \\
O \\
A
\end{array}
\]

The last step is using this machinery to capture the notion of a way for a sentence \( S \) to be true—what I call a *truthmaker* of \( S \). This step is simple. First, we use sets of alternatives to individuate propositions. In particular, we take the propositions denoted by the conjunction of all sentences in each set of alternatives. In our example, we obtain the proposition that Otto went to the party—call this ‘\( o \)’—and the proposition that Anna went to the party—call this ‘\( a \)’. Then, of the propositions obtained in this way, we keep only those that entail \( S \) itself. Since it’s useful to have a term, I call the propositions that pass this test the *truthmakers* of \( S \). In our example, both \( o \) and \( a \) entail the proposition expressed by (41), hence they’re both truthmakers of (41).

The entailment condition—i.e., the condition that a truthmaker must entail the proposition expressed by the antecedent—also screens off irrelevant alternatives. I discuss a case in detail in a footnote.22

---

22 Consider again (41). Suppose that a third individual—call him ‘John’—is contextually relevant, so that he affects what alternatives are computed. The alternatives to (41) now are:

---

be considered. But I have learned of decisive examples from [name omitted for blind review] (p.c.). Here is a variant on his examples:

Scenario. The three passengers in a small plane, contrary to the pilot’s recommendations, clustered on the left-hand side of the plane because they enjoyed sitting together. As a result, the plane was unbalanced and crashed.

(i) If some passengers had sat on the right-hand side, the plane would not have crashed.

(i) seems true. This suggests that the relevant ‘ways for the antecedent to be true’ won’t include the way characterized by the strongest alternatives in the set. If the relevant ‘ways’ included the proposition *All passengers sat on the right-hand side*, the conditional would not be true (since then the plane would still have crashed).

22 Consider again (41). Suppose that a third individual—call him ‘John’—is contextually relevant, so that he affects what alternatives are computed. The alternatives to (41) now are:
Technicalities

Following common usage, I use \( \sigma^- \) to denote the set of negations of the sentences in \( \sigma \); and I use the notion of a set of sentences being consistent with another set of sentences as a natural extensions of the notion of consistency for propositions.

I say that a subset of \( \text{ALT}_S \) is specific with respect to \( \text{ALT}_S \) iff it is consistent with the negation of every alternative that is not a member of it.

\[
\sigma \subseteq \text{ALT}_S \text{ is specific (with respect to } \text{ALT}_S \text{) iff } \sigma \cup (\text{ALT}_S - \sigma)^- \not\models \bot
\]

This is combined with minimality in the obvious way.

\[
\sigma \subseteq \text{ALT}_S \text{ is minimal specific (with respect to } \text{ALT}_S \text{) iff }
\begin{align*}
(i) & \quad \sigma \text{ is specific with respect to } \text{ALT}_S, \text{ and } \\
(ii) & \quad \neg \exists \sigma' : \sigma' \text{ is specific and } \sigma' \subset \sigma.
\end{align*}
\]

Notice that, on this definition: (a) specificity immediately entails consistency; (b) the specific subsets of \( \text{ALT}_S \) are closed under weaker alternatives.

The notion of minimal specificity is related in interesting ways to other notions appearing in the literature on alternatives and implicature. I discuss these connections in a footnote.\(^{23}\)

| Otto or Anna went to the party | \( O \lor A \) |
| Otto went to the party | \( O \) |
| Anna went to the party | \( A \) |
| Otto and Anna went to the party | \( O \land A \) |
| John went to the party | \( J \) |
| Otto or John went to the party | \( O \lor J \) |
| John or Anna went to the party | \( J \lor A \) |
| Otto and John went to the party | \( O \land J \) |
| John and Anna went to the party | \( J \land A \) |

The minimal specific subsets of the alternative set are: \{\( O \lor A, A \), \{\( O \lor A, O \), \{\( J \lor A, J \), \{\( J \lor A, A \). By conjoining the clauses in each set, we get three propositions: \( \text{Otto went to the party, Anna went to the party, John went to the party} \). But only the former two entail the prejacent and hence qualify as truthmakers.

23 First, minimal specificity is something like the converse of the notion that Danny Fox (2007) dubs ‘maximal exclusion’. A maximal exclusions of \( S \) is a maximal subset of alternatives \( \sigma \subseteq \text{ALT}_S \) such that the negation of all the alternatives in \( \sigma \) is consistent with \( S \). For example, the maximal exclusions of \( \text{ALT}_S(43) \) are \{\( O, O \land A \) and \{\( A, O \land A \). Fox uses maximal exclusions to characterize the algorithm that generates scalar implicature. Second, minimal specific sets of alternatives seems to correspond to the alternatives that Chierchia (2013) dubs ‘domain alternatives’. Chierchia gives a semi-formal characterization of domain alternatives for a disjunction or an existential quantifier as ‘all the subdomains of the domain of disjunction/existential quantification’ (2013,
Finally, here is the definition of a truthmaker.

\( p \) is a truthmaker of \( S \) iff

(i) for some \( \sigma \subseteq ALT_S \) such that \( \sigma \) is maximal specific with respect to \( ALT_S \), \( \llbracket \bigwedge \sigma \rrbracket = p \); and

(ii) \( p \models \llbracket S \rrbracket \).

The truthmakers of \( S \) are the propositions denoted by the conjunctive closure of minimal specific sets of alternatives to \( S \), and that are at least as strong as \( S \).

The denotation of \( \text{if} \)-clauses on the new semantics is just the set of truthmakers of the \( \text{if} \)-clause.

\[ \llbracket \text{if } \varphi \rrbracket^{w} = \{ p : p \text{ is a truthmaker of } \varphi \} \]

5 Conditionals as descriptions

I have explained how truthmakers are derived. But I have not explained what role they play in an overall semantics for conditionals. A first, natural suggestion is that conditionals quantify universally over truthmakers.

\( \neg \varphi \Rightarrow \psi \uparrow = \text{true iff: for all truthmakers } p' \text{ of } \varphi, \text{ the closest } p' \text{-worlds make } \psi \text{ true.} \)

This is the route that existing version of truthmaker semantics (e.g. Van Fraassen 1969, Fine 2012b, 2012a) take, and that leads to a semantic vindication of Simplification. This route is natural, but it is wrong. In this section, I propose a better alternative: conditionals refer to, rather than quantify over, sets of propositions. The difference is structurally analogous to that between universally quantified determiner phrases (like all boys) and plural definite descriptions (like the boys). The resulting semantics is quite similar to one using existential quantification, but handles much better a number of problem cases.

p. 116; Chierchia relies on the idea that disjunction can be understood as an existential quantifier over the disjuncts). It’s unclear how this formulation can be generalized to cases where the relevant lexical material doesn’t involve a quantifier domain argument. Minimal specificity improves on Chierchia’s characterization, both because it’s more precise and because it’s more general.
5.1 Homogeneity

My argument against the validity of Simplification connects to the data about distribution effects in DE environments.

(15) It's not the case that, if Alice or Bob went, the party would be fun.
   a. $\rightarrow$ It's not the case that, if Alice went, the party would be fun.
   b. $\rightarrow$ It's not the case that, if Bob went, the party would be fun.

(16) I doubt that, if Alice or Bob went, the party would be fun.
   a. $\rightarrow$ I doubt that, if Alice went, the party would be fun.
   b. $\rightarrow$ I doubt that, if Bob went, the party would be fun.

(15) and (16) show that Antecedent Distribution persists, in some form, in DE environments. But they also show that it doesn't work the way it should if Simplification was valid. In that case, (16) (say) would mean:

(44) I doubt the following: it is both the case that, if Alice went to the party, the party would be fun, and that, if Bob went to the party, the party would be fun.

The problem extends systematically to all occurrences of conditionals in DE environments. This is a strong argument against the semantic validity of Simplification.

The pattern displayed by (15) and (16) is well-known and naturally suggests an account. Here I illustrate it via a parallel with plural definite descriptions, though several items in language exhibit it. In their unembedded occurrences, plural descriptions are interpreted universally:

(45) The boys went swimming.
    $\approx$ Each boy went swimming.

But this interpretation disappears in DE environments. There the universal quantifier appears to take scope outside the DE element.

(46) I doubt that the boys went swimming.
    $\not\approx$ I doubt that each boy went swimming.
    $\approx$ For each of the boys, I doubt that he went swimming.

Schematically (‘$op_{DE}$’ stands for ‘DE operator’):

24 An incomplete list includes: generic statements, bare plurals, embedded interrogatives, and statements about the future involving will. See Gajewski 2005, as well as other references in footnote 25, for a comprehensive discussion.
The $Fs$ are $G \Rightarrow$ For each individual $x$ that is $F$, $x$ is $G$

$op_{DE} [The \ Fs \ are \ G] \Rightarrow$ For each individual $x$ that is $F$, $op_{DE} [x \ is \ G]$

This puzzling effect is known as homogeneity effect\(^{25}\) (the intuition is that the set of boys is homogenous with respect to the property denoted by the predicate: either all boys possess it, or none does). My observation is that conditionals generate an analogous homogeneity effect with respect to the truthmakers of their antecedents. Schematically:\(^{26}\)

$\varphi \square \Rightarrow \psi \Rightarrow$ For each truthmaker $\rho$ of $\varphi$, $\Gamma \rho \square \Rightarrow \psi^\uparrow$

$op_{DE} [\varphi \square \Rightarrow \psi] \Rightarrow$ For each truthmaker $\rho$ of $\varphi$, $\Gamma op_{DE} [\rho \square \Rightarrow \psi]^\uparrow$

Given this parallel, I take the semantics of plural descriptions as a guide to the semantics of conditional antecedents. The analogy between descriptions and conditionals is not new, hence my account connects naturally to other theories currently on the market (Bittner 2001, Schein 2003, Schlenker 2004). The key difference is that these theories analyze conditionals as descriptions of worlds, while I analyze them as descriptions of truthmakers.

5.2 The analogy with plural descriptions

Most approaches to plurals take plural expressions to denote pluralities, or plural individuals. For current purposes, we can take pluralities to be sets of atomic individuals.\(^{27}\) Plural terms denote sets of individuals: e.g., $Alph \ and \ Bob$ denotes the set $\{a, b\}$ containing the atomic individuals Alph and Bob. Plural predicates denote sets of sets: e.g., $boys$ denotes the set of all sets of boys (equivalently, the powerset of the denotation of the singular $boy$).

Plural descriptions are treated as referring to the largest plurality\(^{28}\) of indi-

---


26 To keep things simple, I am sloppy with use and mention here.

27 Pluralities are sometimes understood as sets, and sometimes as mereological sums of individuals. The locus classicus for the introduction of the sum approach is Link 2002; see also Landman 1989. Set-type approaches are presented by Hoeksema 1983 and Schwarzchild 1996, among others (though Schwarzchild uses a kind of nonstandard set theory). The sum approach seems dominant nowadays; I use the set approach for ease of exposition. For some useful overviews about the semantics of plurals, see Nouwen 2014a, Nouwen 2014b.

28 More precisely, they are treated as referring to the maximal plurality of individuals satisfying the predicate, and normally ‘maximal’ is understood as ‘largest’. This is the classical theory derived from the work of Sharvy 1980 and Link 2002. Recently, von Fintel et al. 2014 have given a convincing argument to the effect that the relevant measure of maximality is based on informativity. So far as I can see, shifting to von Fintel, Fox, and Iatridou’s proposal makes no
 Alternatives and Truthmakers in Conditional Semantics

individuals that satisfies the predicate appearing in the description:

\[ \text{⟦The } \varphi \text{⟧ = the (plural) individual } i \text{ s.t. } i \text{ is the unique maximal individual of which } \varphi(i) \text{ is true.} \]

Hence, if Alph, Bob, and Chad are all and only the boys within the domain of quantification, we have:

\[ (47) \quad \text{⟦The boys⟧ = \{a, b, c\}} \]

This basic account is supplemented with some extra features, two of which are relevant here.

First, plural descriptions admit of both distributive and collective readings. These are illustrated by, respectively, (48) and (49):

\[ (48) \quad \text{The boys carried a backpack} \]
\[ \approx \quad \text{For each of the boys, he carried a backpack} \]

\[ (49) \quad \text{The boys carried a piano together} \]
\[ \approx \quad \text{All of the boys jointly carried a piano} \]

Compositionally, it is usually assumed that the distributive reading involves an optional distributivity operator, \( \text{DIST} \), that is adjoined to the predicate. Roughly, the distributivity operator takes as argument a property and a plurality of individuals, and says that the property is true of each individual that is a part of the plurality. Schematically:

\[ \text{⟦[The Fs] DIST[Gs]⟧ = true iff } \forall x : x \text{ is atomic and } \text{⟦Fs⟧(x) = 1, } \text{⟦Gs⟧(x) = 1} \]

Second, distributivity operators carry a semantic presupposition that either all the individuals picked out by the description satisfy the predicate, or they all don’t.\(^{29}\) This all-or-nothing presupposition serves to capture homogeneity effects on the distributive reading (collective readings immediately validate homogeneity).\(^{30}\) By negating \textit{The boys went swimming} we get (via the distributivity

\[ \text{Here is the entry:} \]
\[ \text{⟦DIST⟧ = } \lambda P. \lambda x x : \]
\[ \forall y y \leq x x . (\text{Atom}(y y) \rightarrow P(y y)) \lor \forall y y \leq x x . (\text{Atom}(y y) \rightarrow \neg P(y y)). \]
\[ \forall y y \leq x x . (\text{Atom}(y y) \rightarrow P(y y)) \]

\[ \text{(Following a widespread convention, I use double variables like ‘xx’ to range over pluralities.)} \]
\[ \text{I should note that this is only one of the ways to capture homogeneity (for which see Von Fintel 1997, Gajewski 2005), and there is no agreement that it is the correct one. Nothing in my} \]

\[ \text{difference to my account.} \]

\(^{29}\) This or-nothing presupposition

\(^{30}\)
operator) that not all boys went swimming. Combined with the all-or-nothing presupposition, this gets us the perceived truth conditions, i.e. that none of the boys went swimming.

My account consists simply in importing these features to conditionals. In a slogan: conditionals are descriptions of truthmakers. My account and its predictions fall out immediately by switching truthmakers for individuals in the semantics I just sketched.

To start with, if-clauses denote sets of propositions. In addition, I assume the existence of an optional distributivity operator Dist$_\pi$, analogous to Dist but operating over propositions. Dist$_\pi$ takes as arguments a function from propositions to truth values (i.e. the denotation of the consequent of a conditional) and a set of propositions, and says that the function maps each proposition in the set to true. Schematically:

$$\llbracket \text{if } \varphi \text{ } \text{Dist}_\pi \text{[} \psi \text{]} \rrbracket = \text{true iff } \forall p: p \in \llbracket \text{if } \varphi \rrbracket, \llbracket \psi \rrbracket(p) = \text{true}$$

Dist$_\pi$ is also the bearer of an all-or-nothing presupposition, again analogous to the one in use on the distributivity operator for individuals.

Dist$_\pi$ is what generates Antecedent Distribution. Take our running example (2), and assume that it has the distributive LF, represented below:

(50) $\llbracket \text{If Alice or Bob went to the party } \text{Dist}_\pi \text{[} \text{would } \text{the party be fun} \rrbracket \rrbracket$

The if-clause in (50) denotes the set of the propositions expressed by the two disjuncts. Dist$_\pi$ ensures that these propositions are combined individually with the rest of the clause. The resulting truth conditions are:

(51) $\forall p \in \{\text{Alice went}, \text{Bob went}\}, \max_{\preceq_w}(p) \in \llbracket \text{party be fun} \rrbracket$

Hence (2), on the parsing in (50), entails

(2) a. If Alice went to the party, the party would be fun.
   b. If Bob went to the party, the party would be fun.

Since I assume that the distributivity operator is optional, again in analogy with the semantics of plurals, I am committed to readings on which Antecedent Distribution fails. In fact, this prediction is borne out. The counterpart of collective readings for conditionals are just examples like:

(6) If Spain had fought with the Axis or the Allies, she would have fought with the Axis.

account of conditionals depends on accounting for homogeneity via this route.

28
Hence, differently from other truthmaker theories, I do not vindicate Simplification. But this failure is expected and descends from independently motivated features of the theory.

**Technicalities**

The main innovation is letting if-clauses denote sets of propositions. To do this, I treat if as a set-forming operator: if takes as argument a clause and a set of alternatives, and generates a set of truthmakers for that clause.

\[
\left[ \text{if} \right] \equiv_{w} = \lambda_{p_{s,t}}. \lambda_{ALT}. \{q_{s,t} : q \text{ is a truthmaker}_{ALT} \text{ of } p\}
\]

The distributivity operator \( \text{DIST}_{\pi} \) takes as argument a set of propositions and quantifies over singletons of propositions within that set. (The quantification is over singleton rather than over the propositions themselves for type-theoretic reasons—this way, the input argument to the modal is of the same type whether \( \text{DIST}_{\pi} \) is present or not.) In addition, exactly like the distributivity operator for individuals, it carries the homogeneity presupposition. This is the lexical entry (I use ‘p’ as a type for sets of propositions):

\[
\left[ \text{DIST}_{\pi} \right] \equiv_{w} = \lambda_{\Phi_{s,t}}. \lambda_{S_{p}} : \forall p \in S. \Phi\{\{p\}\} = 1 \lor \forall p \in S. \Phi\{\{p\}\} = 0.
\]

Finally, modals appearing in conditionals work in a standard way, aside from two tweaks. First, the outermost argument of a modal is a set of propositions, rather a proposition. Second, modals ‘extract’ from their set-type argument a proposition in the following way: they take the proposition generated by the disjunctive closure of the propositions in the set. [EXPLAIN THIS!] When the set is a singleton, the result is just the unique proposition in the set. When the set is not a singleton, this gets back the proposition expressed by the antecedent. The latter case is the one that generates violations of Simplification.

For an example, here is the lexical meaning of would:

\[
\left[ \text{would} \right] \equiv_{w} = \lambda_{S_{p}}. \lambda_{p}. \forall w' \in \max_{\equiv_{w}} (\forall S) = 1, p(w') = 1
\]

### 6 Hyperintensionality and syntactic complexity

I started from an inconsistent triad of apparent *desiderata* about conditional logic: Failure of Antecedent Strengthening, Simplification of Disjunctive Antecedent, and Substitution of Logical Equivalents. Standard semantics for condition-
als drops Simplification and preserves the other two. My account drops both Simplification and Substitution, but it is very much in the spirit of Simplification-vindicating accounts (and it does allow conditionals to vindicate Simplification, provided that they are parsed with a distributivity operator).

The main theoretical move of my account is dropping Substitution. This makes conditionals hyperintensional: we can’t preserve truth value under substitutions of necessarily equivalent antecedents. Pairs of conditionals exemplifying this failure are easy to find. For example:

(52) If Anna came to the party, the party would be fun.
(53) If Anna, or Otto and Anna, came to the party, the party would be fun.

The antecedents of (52) and (53) are logically equivalent, yet (on the distributive parsing of (53)) the two conditionals have different truth conditions.

Dropping Substitution goes against a long tradition of work on which the semantics for conditionals is intensional. But my account also diverges from existing hyperintensional semantics, which generally rely on a metaphysics of fine-grained entities and drop all talk of worlds and comparative closeness. All the main tools I use (possible worlds, closeness, alternatives) are already available in the literature. Moreover, my notion of a truthmaker is not metaphysical, but cognitive: truthmakers are just standard propositions; which propositions count as truthmakers is determined on the basis of grammatical mechanisms tied to alternatives.

The kind of hyperintensionality resulting from my view is pretty mundane. Its source is the fact that logically equivalent sentences may have different alternatives. This kind of hyperintensionality is also widespread through language. Consider:

(54) a. John only read *Anna Karenina*.
   ⇒ John did not read *War and Peace*.

   b. John only read either *Anna Karenina*, or *War and Peace* and *Anna Karenina*.
   ⇔ John did not read *War and Peace*.

This kind of hyperintensionality has its origins in grammar, rather than in metaphysics. The grammatical processes that generate it are independently needed, and independently understood. Incorporating it into a theory of conditionals brings in extra complexity, but no ontological or explanatory costs.

31 For a counterfactual semantics employing impossible worlds, see, among many, Nolan 1997. For a semantics using states, see Fine 2012a and 2012b, as well as the version of Fine’s semantics in Briggs 2012.
Alternatives and Truthmakers in Conditional Semantics

References


Alternatives and Truthmakers in Conditional Semantics


Magri, Giorgio. 2014. An account for the homogeneity effects triggered by plural definites and conjunction based on double strengthening.


Alternatives and Truthmakers in Conditional Semantics