Sly Pete in dynamic semantics

Goldstein, Simon David

SLY PETE IN DYNAMIC SEMANTICS

Simon Goldstein

Keywords: Sly pete, conditionals, dynamic semantics

Abstract

In 'Sly Pete' or 'standoff' cases, reasonable speakers accept incompatible conditionals, and communicate them successfully to a trusting hearer. This paper uses the framework of dynamic semantics to offer a new model of the conversational dynamics at play in standoffs, and to articulate several puzzles posed by such cases. The paper resolves these puzzles by embracing a dynamic semantics for conditionals, according to which indicative conditionals require that their antecedents are possible in their local context, and update this body of information by eliminating the possibilities where the antecedent is true and the consequent is false. In this way, the dynamic analysis draws on insights from the material conditional and contextualist analyses, while explaining how standoffs are genuine disagreements.

1 Introduction

This paper explains 'Sly Pete' or 'standoff' cases, like the following:

> Suppose three people know that a coin may or may not be flipped at noon. Alfred is told ahead of time that if it lands head he will be told that within a minute of it being flipped. Bob is told ahead of time that if it lands tails he will be told that within a minute of it being flipped. Catherine gets no information, but both Alfred and Bob are allowed to pass a note to Catherine. Suppose neither Alfred nor Bob has heard anything about the coin flip at 12:01. Alfred passes a note to Catherine that says, 'If it was flipped, it landed [heads].' Bob passes a note to Catherine that says, 'If it was flipped, it didn't land [heads]'. Catherine concludes that it wasn't flipped.[1]

In this example, Catherine receives the following conditionals by testimony:

(1) a. If it was flipped, it landed heads.
    b. If it was flipped, it didn't land heads.

Catherine seems to learn from both pieces of testimony, because she concludes that the coin wasn't flipped. But this is puzzling. The two conditionals she has learned seem incompatible. How can you learn something from two incompatible

[1] Rothschild 2015. See Bennett 2003 and Gibbard 1981 for similar cases involving someone named 'Sly Pete'.

claims? In this way, standoffs seem to violate the principle of 'Conditional Non-Contradiction', which says that two conditionals with the same antecedent and incompatible consequents are themselves incompatible, whenever the antecedent is possible.

This paper uses the framework of dynamic semantics to offer a new model of the conversational dynamics at play in standoffs, and to articulate some general puzzles posed by such cases. The paper goes on to resolve the puzzles by embracing a dynamic semantics for conditionals, according to which indicative conditionals require that their antecedents are possible in their local context, and then update this body of information by eliminating the possibilities where the antecedent is true and the consequent is false. In this way, the dynamic analysis draws on insights from the material conditional and contextualist analyses, while explaining how standoffs are genuine disagreements.

2 Modelling preliminaries

My first task is to offer a model of the conversational dynamics involved in Sly Pete cases. To do so, I first introduce some formal preliminaries. I model a simple propositional language with the indicative conditional $\rightarrow$, and a possibility operator $\Diamond$.

I work within a possible worlds framework. Worlds are functions from atoms to truth values. An agent's state of information s is modelled with a set of possible worlds. To give a theory of communication, I introduce two notions. First, say that a sentence is assertible by an agent with information s just in case s accepts A ($s \models \mathsf{A}$). Second, I model the result of learning A with an update function, which assigns each state s and sentence A a posterior state $s[\mathsf{A}]$, which is the result of learning A in s.

Definition 1.

1. W is the set of worlds w, functions from atoms to truth values.

2. An information state s is a set of worlds.

3. $\models$ is a relation between states and sentences.

4. $[\cdot]$ is a function from sentences to functions from states to states.

The two crucial notions in this system are acceptance and update, which induce corresponding norms on speakers and hearers. The first rule governs assertion.

(2) **Assertion Rule.** Suppose that S's information state is s. Then S ought to: assert A only if $s \models \mathsf{A}$.

The second rule governs updating:

(3) **Update Rule.** Suppose that S's information state is s, S' asserts A to S and S trusts S''s assertion. Then S's new information state is $s[\mathsf{A}]$.

Neither the Assertion Rule nor the Update Rule specifies what in particular an agent's information state is. One natural thought is that a trusting hearer should update her beliefs with [A]. If all goes well, she may also come to know A.

3 Sly Pete in dynamic semantics

I now characterize more precisely what is puzzling about standoffs. I'll show that in standoffs: (i) agents can accept a conditional without accepting a proposition; and (ii) agents can update on conditionals without thereby accepting them.

I begin with a model of standoffs. Let A be the claim that the coin was flipped and B be the claim that it landed heads. Standoffs involve information exchange between three agents. Let s_1 and s_2 denote the information states of the two speakers. Let h_1 and h_2 denote the information state of the hearer before and after updating on their testimony. I model each information state as in Figure 1.

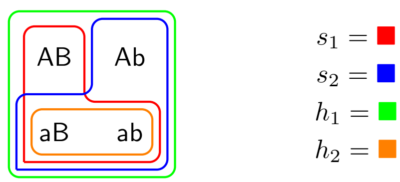

Figure 1: A standoff

Puzzles result from the following features of this example. First, s_1 and s_2 are compatible: they agree about what happens when A is false. Second, s_1 and s_2 accept incompatible conditionals: $\mathsf{A} \rightarrow \mathsf{B}$ and $\mathsf{A} \rightarrow \neg\mathsf{B}$. Third, the hearer h_1 can learn from both s_1 and s_2, by updating on $\mathsf{A} \rightarrow \mathsf{B}$ and $\mathsf{A} \rightarrow \neg\mathsf{B}$. When she learns these claims, she enters a new, consistent state h_2. Finally, in this new state the hearer goes on to accept $\neg\mathsf{A}$ without accepting either $\mathsf{A} \rightarrow \mathsf{B}$ nor $\mathsf{A} \rightarrow \neg\mathsf{B}$. Summarizing:

(4) a. $s_1 \cap s_2 \neq \emptyset$
b. $s_1 \models \mathsf{A} \rightarrow \mathsf{B}$
c. $s_2 \models \mathsf{A} \rightarrow \neg\mathsf{B}$
d. $h_1[\mathsf{A} \rightarrow \mathsf{B}][\mathsf{A} \rightarrow \neg\mathsf{B}] = h_2 \neq \emptyset$
e. $h_2 \models \neg\mathsf{A}$
f. $h_2 \not\models \mathsf{A} \rightarrow \mathsf{B}$
g. $h_2 \not\models \mathsf{A} \rightarrow \neg\mathsf{B}$

In addition to this setup, I hold fixed an analogue of Conditional Non-Contradiction for the state of acceptance: when A is consistent, no consistent state can accept both $\mathsf{A} \rightarrow \mathsf{B}$ and $\mathsf{A} \rightarrow \neg\mathsf{B}$. I'll call this specific version of the principle 'CNC' throughout.[2]

[2] Williams 2008b develops a theory of standoffs which rejects CNC. On that theory, $\mathsf{A} \rightarrow \mathsf{B}$

(5) **CNC.** If $\mathsf{A} \not\models \bot$, then there is no s besides $\emptyset$ where $s \models \mathsf{A} \to \mathsf{B}$ and $s \models \mathsf{A} \to \neg\mathsf{B}$

The assumptions so far create several puzzles about assertion and updating.

First, consider the following constraint on acceptance: believing a conditional amounts to believing a proposition. In the framework above, we can characterize propositional belief precisely. Where p is a proposition, or set of worlds, accepting A is a matter of believing p just in case any state s accepts A iff s is included in p.

(6) **Propositionality.** There is some p where for every s: $s \models \mathsf{A}$ iff $s \subseteq p$

Standoffs are counterexamples to Propositionality. To see why, it will help to formulate some general notions of disagreement which are exemplified by standoffs, and which rule out Propositionality. First, say that two agents disagree when they accept claims that no single agent could accept.[3]

(7) **Disagreement.** s and t disagree regarding A and B iff:

a. $s \models \mathsf{A}$
b. $t \models \mathsf{B}$
c. There is no u besides $\emptyset$ where $u \models \mathsf{A}$ and $u \models \mathsf{B}$

Propositionality implies that whenever two agents disagree, only one of them can be right about which world is actual: their two states cannot overlap. To raise trouble for Propositionality, hold fixed CNC. Given Propositionality, this implies that when A is consistent, the propositions associated with $\mathsf{A} \to \mathsf{B}$ and $\mathsf{A} \to \neg\mathsf{B}$ do not overlap. This means that if there are two states s and s' where s accepts $\mathsf{A} \to \mathsf{B}$ and s' accepts $\mathsf{A} \to \neg\mathsf{B}$, then these states have no worlds in common. For suppose that $\mathsf{A} \to \mathsf{B}$ expresses the proposition p and $\mathsf{A} \to \neg\mathsf{B}$ expresses p'. Then s is a subset of p, s' is a subset of p', and $p \cap p' = \emptyset$; so $s \cap s' = \emptyset$ also.

But now return to the standoff. The two speakers accept $\mathsf{A} \to \mathsf{B}$ and $\mathsf{A} \to \neg\mathsf{B}$ respectively. But their two information states are compatible: they leave open common possibilities where the coin wasn't flipped. Given CNC, Propositionality rules this out: in accepting $\mathsf{A} \to \mathsf{B}$, s_1 locates herself in a region of logical space disjoint from the one in which s_2 inhabits by accepting $\mathsf{A} \to \neg\mathsf{B}$.

I understand this idea in terms of a special kind of disagreement. Two agents disagree 'compatibly' regarding two claims when they accept claims that no single agent could accept, but their overall information states are nonetheless

and $\mathsf{A} \to \neg\mathsf{B}$ are vacuously accepted by any state which lacks any A worlds. For this reason, Williams 2008b rejects (4-f) and (4-g).

[3] The semantic significance of disagreement is carefully explored in recent work on semantic relativism. For an overview of relevant questions and a taxonomy of different notions of disagreement, see MacFarlane 2014. For an application of this debate about disagreement to standoffs, see Weatherson 2009, who expresses reservations about whether the agents in a standoff count as genuinely disagreeing. For my purposes, it is not necessary that the agents in a standoff intuitively count as disagreeing; the crucial question is whether they satisfy the structural properties in terms of which I define disagreement, rather than whether the term 'disagreement' correctly applies to them.

compatible.[4]

(8) **Compatible disagreement.** s and t disagree compatibly regarding A and B iff:

a. s and t disagree regarding A and B

b. $s \cap t \neq \emptyset$

Compatible disagreements are ruled out by Propositionality. If s and t disagree compatibly about A and B, then they disagree about A and B. Since they disagree about A and B, no consistent state u can accept both A and B. This means that the propositions expressed by A and B are incompatible. So s and t are incompatible. So $s \cap t = \emptyset$.[5] Summarizing:

Observation 1. Compatible disagreements violate Propositionality.

So far, I've shown that standoffs are incompatible with Propositionality. That is, standoffs impose significant constraints on our theory of acceptance. Standoffs also constrain the theory of updating. First, consider the constraint that after a trusting hearer has updated on A, she is now in a position to herself assert A.

(9) **Idempotence.** $s[\mathsf{A}] \models \mathsf{A}$

Language can be used to pass information down a chain of agents. When one agent shares information with another, the new agent is herself in a position to share that information with others.[6]

Standoffs conflict immediately with Idempotence. I imagined an agent h_1 who updates on both of $\mathsf{A} \to \mathsf{B}$ and $\mathsf{A} \to \neg\mathsf{B}$, and enters a new state h_2. Crucially, this new state did not accept either $\mathsf{A} \to \mathsf{B}$ or $\mathsf{A} \to \neg\mathsf{B}$. But this generates a failure of Idempotence: the state $h_1[\mathsf{A} \to \mathsf{B}]$ is updated with $\mathsf{A} \to \neg\mathsf{B}$, but does not thereby accept $\mathsf{A} \to \neg\mathsf{B}$.

Now consider one more structural constraint: that when an agent believes something, she continues to believe it after updating consistently on new information.

(10) **Persistence.** If $s \models \mathsf{A}$, then $s[\mathsf{B}] \models \mathsf{A}$

Standoffs produce violations of Persistence. s_1 accepts $\mathsf{A} \to \mathsf{B}$. But if s_1 is updated with $\mathsf{A} \to \neg\mathsf{B}$, it becomes state h_2, which does not accept $\mathsf{A} \to \mathsf{B}$. On the theory I later develop, this is because s_1 is consistent with A, but h_2 is not.

[4] An anonymous referee helpfully observes that one might distinguish two notions of compatibility: categorical and complete. Two information states are categorically compatible when they have worlds in common, but may not yet be completely compatible if they accept incompatible conditionals. In this way, one might not think compatible disagreements are so strange, since the conditional facts might vary freely from the categorical facts. I'll show in a moment, however, that standoffs give rise to further problems (involving idempotence, persistence, and the possibility of productive disagreement) that cannot be resolved in this way.

[5] In using standoffs to reject Propositionality, I generalize the argument in Gibbard 1981.

[6] For recent defense of idempotence, see Klinedinst and Rothschild 2015 and Yalcin 2015.

The failures of Idempotence and Persistence are instances of a more general kind of disagreement exemplified by standoffs. CNC implies that s_1 and s_2 disagree regarding $\mathsf{A} \rightarrow \mathsf{B}$ and $\mathsf{A} \rightarrow \neg\mathsf{B}$, since no state can accept both claims. Nonetheless, h_1 can consistently update on both of their testimony. Standoffs are productive disagreements, where two agents who disagree with one another can still provide helpful information to a third party who trusts both of them.[7]

(11) **Productive disagreement.** s and t disagree productively regarding A and B iff:

a. s and t disagree regarding A and B

b. There is some h where $h[\mathsf{A}][\mathsf{B}] \neq \emptyset$

Productive disagreements are ruled out by Idempotence and Persistence. If s and t disagree productively regarding A and B, then no consistent state u can accept both A and B. But Idempotence and Persistence imply that $h[\mathsf{A}][\mathsf{B}]$, which is consistent, does accept both A and B. Standoffs are cases of productive disagreement; so standoffs must be failures of Idempotence or Persistence.

Observation 2. Productive disagreements violate either Idempotence or Persistence.

Productive disagreements are interesting in another way. Suppose that s_1 and s_2 share their information with one another, and suppose they are trusting. Once s_1 updates on $\mathsf{A} \rightarrow \neg\mathsf{B}$, she enters state h_2. Similarly, when s_2 updates on $\mathsf{A} \rightarrow \mathsf{B}$, she enters state h_2. While our two agents initially disagree, they end up sharing the same belief state. This means that productive disagreements can be resolved through communication.

To make sense of the compatible and productive disagreements involved in standoffs, requires a theory of assertion and updating on which Idempotence, Persistence, and Propositionality can fail. I'll first show that existing theories do not satisfy the desiderata above, and then develop a dynamic theory that does.

4 Possible worlds semantics

First, let's see what propositional theories of meaning make of the puzzles above. On these theories, the meaning of an atom is the set of worlds where it is true. The meaning of $\neg\mathsf{A}$ is the set of worlds where A is false. The meaning of $\mathsf{A} \wedge \mathsf{B}$ is the set of worlds where both A and B are true. The meaning of the indicative conditional $\mathsf{A} \rightarrow \mathsf{B}$ is some function $>$ of $[\![\mathsf{A}]\!]$ and $[\![\mathsf{B}]\!]$. For example, $>$ generates the material conditional when it maps any $[\![\mathsf{A}]\!]$ and $[\![\mathsf{B}]\!]$ to $(W - [\![\mathsf{A}]\!]) \cup [\![\mathsf{B}]\!]$. Variably strict and strict analyses of the conditional are generated by corresponding choices of $>$.

Definition 2.

[7] Again, the theory in Williams 2008b would deny that standoffs involve a disagreement in this sense, because $\mathsf{A} \rightarrow \mathsf{B}$ and $\mathsf{A} \rightarrow \neg\mathsf{B}$ would be jointly accepted by anyone who rejects A.

1. $[\![\mathsf{p}]\!] = \{w \mid w(\mathsf{p}) = 1\}$

2. $[\![\neg\mathsf{A}]\!] = W - [\![\mathsf{A}]\!]$

3. $[\![\mathsf{A} \wedge \mathsf{B}]\!] = [\![\mathsf{A}]\!] \cap [\![\mathsf{B}]\!]$

4. $[\![\mathsf{A} \rightarrow \mathsf{B}]\!] = [\![\mathsf{A}]\!] > [\![\mathsf{B}]\!]$

Updating and acceptance are defined in terms of truth. Updating s with A narrows down s to the worlds where A is true. Then a state accepts A when updating with A has no effect.

Definition 3.

1. $s[\mathsf{A}] = s \cap [\![\mathsf{A}]\!]$

2. $s \models \mathsf{A}$ iff $s[\mathsf{A}] = s$ iff $s \subseteq [\![\mathsf{A}]\!]$

Regardless of the particular operation assigned as the meaning of an indicative conditional, this proposal cannot resolve the puzzles described above:

Observation 3. Definitions 2 and 3 validate Propositionality, Idempotence, and Persistence.

Regardless of the particular choice of $>$ in the above, this simple framework validates Propositionality. After all, Definition 3 immediately implies that $s[\mathsf{A} \rightarrow \mathsf{B}] = s \cap [\![\mathsf{A} \rightarrow \mathsf{B}]\!]$, and $[\![\mathsf{A} \rightarrow \mathsf{B}]\!]$ is some proposition p fixed independently of s. For similar reasons, the theory validates Idempotence and Persistence.[8]

Since the theory above accepts Propositionality, it also rules out compatible disagreements. Since s_1 and s_2 are compatible (sharing some worlds in common), the theory must either reject CNC, or must deny that s_1 accepts $\mathsf{A} \rightarrow \mathsf{B}$ and s_2 accepts $\mathsf{A} \rightarrow \neg\mathsf{B}$.[9]

Since the theory validates Idempotence and Persistence, it also rules out productive disagreements. $h_1[\mathsf{A} \rightarrow \mathsf{B}][\mathsf{A} \rightarrow \neg\mathsf{B}]$ accepts both $\mathsf{A} \rightarrow \mathsf{B}$ and $\mathsf{A} \rightarrow \neg\mathsf{B}$, since $h_1[\mathsf{A} \rightarrow \mathsf{B}][\mathsf{A} \rightarrow \neg\mathsf{B}]$ is the result of intersecting h_1 with the worlds where $\mathsf{A} \rightarrow \mathsf{B}$ are true and the worlds where $\mathsf{A} \rightarrow \neg\mathsf{B}$ are true.

5 Test semantics

Now I'll evaluate the test semantics for conditionals defended in Dekker 1993 and Gillies 2004.[10] While this theory rejects Propositionality and Idempotence, it continues to treat conditionals as Idempotent.

[8]For take any state s. In the framework above, $s[\mathsf{A}] = s \cap p$ for some fixed choice of p. So $s[\mathsf{A}][\mathsf{A}] = (s \cap p) \cap p = s \cap p = s[\mathsf{A}]$. So $s[\mathsf{A}] \models \mathsf{A}$.

[9]The particular premise rejected depends on the choice of conditional operator $>$. The material analysis rejects CNC: when s is incompatible with consistent $[\![\mathsf{A}]\!]$, it accepts both $\mathsf{A} \rightarrow \mathsf{B}$ and $\mathsf{A} \rightarrow \neg\mathsf{B}$. By contrast, suppose that $>$ is the variably strict conditional. Then either s_1 will not accept $\mathsf{A} \rightarrow \mathsf{B}$, or s_2 will not accept $\mathsf{A} \rightarrow \neg\mathsf{B}$. For consider the worlds in $s_1 \cap s_2$. At most one of $\mathsf{A} \rightarrow \mathsf{B}$ and $\mathsf{A} \rightarrow \neg\mathsf{B}$ can hold there. So either $s_1 \not\models \mathsf{A} \rightarrow \mathsf{B}$, or $s_2 \not\models \mathsf{A} \rightarrow \neg\mathsf{B}$, or both.

[10]See Heim 1983 and Veltman 1996 for the general theoretical framework.

In this style of dynamic semantics (sometimes called 'update semantics'), updating is taken as primitive rather than defined in terms of truth conditions or acceptance (see Heim 1983 and Veltman 1996). Updating a state s with an atom simply narrows down the state to the worlds where the atom is true. Updating s with $\neg\mathsf{A}$ requires removing from s any worlds that survive updating with A. Updating s with $\mathsf{A} \wedge \mathsf{B}$ is a matter of updating sequentially, first with A and then with B. Indicative conditionals are tests. On this proposal, updating any state with the indicative either produces the original state or the absurd state $\emptyset$. The test for $\mathsf{A} \rightarrow \mathsf{B}$ is passed by s just in case $s[\mathsf{A}]$ accepts B.

Definition 4.

1. $s[\mathsf{p}] = \{w \in s \mid w(\mathsf{p}) = 1\}$
2. $s[\neg\mathsf{A}] = s - s[\mathsf{A}]$
3. $s[\mathsf{A} \wedge \mathsf{B}] = s[\mathsf{A}][\mathsf{B}]$
4. $s[\mathsf{A} \rightarrow \mathsf{B}] = \{w \in s \mid s[\mathsf{A}] \models \mathsf{B}\}$

A state accepts a sentence when the sentence has no effect on the state.

Definition 5.

$s \models \mathsf{A}$ iff $s[\mathsf{A}] = s$

Consider how this theory responds to the puzzles about standoffs. First, both Idempotence and Propositionality are satisfied for bare conditionals.

Observation 4. Definitions 4 and 5 imply that $\mathsf{A} \rightarrow \mathsf{B}$ is idempotent, persistent, and propositional for every consistent A and B.

Idempotence holds because $s[\mathsf{A} \rightarrow \mathsf{B}]$ is either s or $\emptyset$. In the former case, s passes the test imposed by $\mathsf{A} \rightarrow \mathsf{B}$, and so $s \models \mathsf{A} \rightarrow \mathsf{B}$. The same holds trivially of $\emptyset$. To see why Propositionality holds, note that in the above theory a state accepts $\mathsf{A} \rightarrow \mathsf{B}$ just in case it accepts the corresponding material conditional, itself true at a fixed set of worlds.[11]

Since Idempotence and Persistence are satisfied, standoffs are not productive disagreements. A trusting listener can't faithfully update on both $\mathsf{A} \rightarrow \mathsf{B}$ and $\mathsf{A} \rightarrow \neg\mathsf{B}$ without being able to assert either one themselves. This manifests itself in a few respects. First, the theory above does not predict that $h_2 = h_1[\mathsf{A} \rightarrow \mathsf{B}][\mathsf{A} \rightarrow \neg\mathsf{B}]$. Rather, $h_1[\mathsf{A} \rightarrow \mathsf{B}][\mathsf{A} \rightarrow \neg\mathsf{B}]$ is $\emptyset$. This follows from a more general feature of the semantics above: if modal sentences are tests, returning either the

[11] At first glance, this observation is somewhat surprising. One of the hallmarks of dynamic semantics is the rejection of exactly such properties (See Rothschild and Yalcin 2015). Here, two clarifications are in order. First, Propositionality, Idempotence, and Persistence fail in the system above for more complex constructions. Idempotence fails for example in the case of $\neg(\top \rightarrow \mathsf{A}) \wedge \mathsf{A}$. For let $s = \{w, v\}$, where A is true at w and false at v. $s[\neg(\top \rightarrow \mathsf{A}) \wedge \mathsf{A}] = \{w\}$. But $\{w\}[\neg(\top \rightarrow \mathsf{A}) \wedge \mathsf{A}] = \emptyset$. Similarly, negated conditionals of the form $\neg(\mathsf{A} \rightarrow \mathsf{B})$ are not propositional or persistent. Second, an analogue of Propositionality for updating (that $\exists p : s[\mathsf{A} \rightarrow \mathsf{B}] = s \cap p$) fails in this system.

initial state or the absurd state, then the update function $[\cdot]$ cannot model the proper response of a hearer to modal testimony. If hearers performed the test update in response to conditional testimony, then their beliefs would either stay the same or become absurd. For this reason, the theory trivially validates:

(12) a. $h_1[\mathsf{A} \to \mathsf{B}][\mathsf{A} \to \neg\mathsf{B}] \models \mathsf{A} \to \mathsf{B}$
b. $h_1[\mathsf{A} \to \mathsf{B}][\mathsf{A} \to \neg\mathsf{B}] \models \mathsf{A} \to \neg\mathsf{B}$

Likewise, the theory above accepts Propositionality for bare conditionals. The semantics doesn't classify standoffs as compatible disagreements, because s_1 and s_2 don't count as disagreeing about $\mathsf{A} \to \mathsf{B}$ and $\mathsf{A} \to \neg\mathsf{B}$. The theory gives up CNC. When $s \models \neg\mathsf{A}$, the theory above predicts that $s \models \mathsf{A} \to \mathsf{B}$ and $s \models \mathsf{A} \to \neg\mathsf{B}$.

6 Antecedent possibility

I'll now show that the standoff puzzles can be resolved within a dynamic semantics, by enriching the material analysis with the familiar idea that indicative conditionals require that their antecedent is possible.

A long tradition of theorizing about indicative conditionals claims that $\mathsf{A} \to \mathsf{B}$ comes with a requirement that A is possible.[12] To model this idea, I introduce a dynamic possibility operator, $\Diamond$, that tests a body of information or 'local context' to see if that information is consistent with its complement. When the test is passed, the information stays the same; otherwise, it is destroyed (see Veltman 1996).

Definition 6. $s[\Diamond\mathsf{A}] = \{w \in s \mid s[\mathsf{A}] \neq \emptyset\}$

I then model the conditional as a conjunction of $\Diamond\mathsf{A}$ and the material conditional.[13]

Definition 7. $s[\mathsf{A} \to \mathsf{B}] = s[\Diamond\mathsf{A} \wedge \neg(\mathsf{A} \wedge \neg\mathsf{B})] = \begin{cases} (s - s[\mathsf{A}]) \cup s[\mathsf{B}] & \text{if } s \models \Diamond\mathsf{A} \\ \emptyset & \text{otherwise} \end{cases}$

Here, I depart from existing work that relies on the possibility of the antecedent. Much of that work relies on the test semantics rather than the material conditional. But in order to model standoffs, the conditional must non-trivially update the hearer's information. Otherwise, how would the hearer who learns both of $\mathsf{A} \to \mathsf{B}$ and $\mathsf{A} \to \neg\mathsf{B}$ enter a new state in which she accepts $\neg\mathsf{A}$? On the other hand, if I used a non-test conditional stronger than the material conditional, it

[12] Stalnaker 1975 appeals to this requirement to explain why the Direct Argument, that $\neg\mathsf{A} \vee \mathsf{B} \models \mathsf{A} \to \mathsf{B}$, is a reasonable inference. Gillies 2009 appeals to it to explain why antecedent strengthening and contraposition seem to be invalid inferences. Starr 2014 appeals to the same requirement as part of a uniform account of indicative and subjunctive conditionals. Willer 2017 applies the requirement to puzzles about indicative Sobel sequences. And Starr 2021 uses it to explain judgments about the probability of conditionals.

[13] Most theorists cited above treat $\Diamond\mathsf{A}$ as a presupposition rather than a part of the at-issue meaning. But I suppress this complexity, since standoffs only involve unembedded conditionals.

would be difficult to explain how the hearer can come to learn something from both of the conditionals together, since the two claims would conflict.

My theory predicts that conditionals are neither idempotent, persistent, nor propositional.

Observation 5. Definition $\mathsf{A} \to \mathsf{B}$ is neither idempotent, persistent, nor propositional for every consistent A and B.

I'll illustrate this observation in the setting of standoffs, by returning to the earlier model.

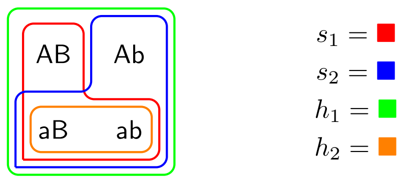

Figure 2: A standoff

First, the semantics makes the correct predictions about the setup:

(13) a. $s_1 \cap s_2 \neq \emptyset$
b. $s_1 \models \mathsf{A} \to \mathsf{B}$
c. $s_2 \models \mathsf{A} \to \neg\mathsf{B}$
d. $h_1[\mathsf{A} \to \mathsf{B}][\mathsf{A} \to \neg\mathsf{B}] = h_2 \neq \emptyset$
e. $h_2 \models \neg\mathsf{A}$
f. $h_2 \not\models \mathsf{A} \to \mathsf{B}$
g. $h_2 \not\models \mathsf{A} \to \neg\mathsf{B}$

s_1 accepts $\mathsf{A} \to \mathsf{B}$, and s_2 accepts $\mathsf{A} \to \neg\mathsf{B}$, because accepting an indicative is equivalent to accepting the material conditional whenever the antecedent is consistent with the state.

After updating on each of $\mathsf{A} \to \mathsf{B}$ and $\mathsf{A} \to \neg\mathsf{B}$, the hearer learns $\neg\mathsf{A}$. Since A is possible in h_1, $[\mathsf{A} \to \mathsf{B}]$ simply narrows down h_1 to the worlds where A is false or B is true. Since A is still possible in $h_1[\mathsf{A} \to \mathsf{B}]$, $h_1[\mathsf{A} \to \mathsf{B}][\mathsf{A} \to \neg\mathsf{B}]$ shrinks further to the worlds where A is false or B is false. This results in h_2, which is simply the $\neg\mathsf{A}$ worlds in h_1.

The analysis then departs from the material analysis regarding what the hearer accepts in the posterior state h_2. After updating on each conditional, the hearer has eliminated the antecedent A. For this reason, she does not accept either conditional in her posterior state.

This creates failures of Idempotence and Persistence: the hearer updates h_1 with each of $\mathsf{A} \to \mathsf{B}$ and $\mathsf{A} \to \neg\mathsf{B}$, entering the new state h_2. But in this new state, they accept neither conditional.

(14) a. $h_2 \not\models \mathsf{A} \to \mathsf{B}$
b. $h_2 \not\models \mathsf{A} \to \neg\mathsf{B}$

At the same time, the theory departs from the material analysis by validating CNC. When A is consistent, $\mathsf{A} \to \mathsf{B}$ and $\mathsf{A} \to \neg\mathsf{B}$ cannot both be accepted. Accepting the former requires being in a state with an A world, and where every A world is a B world. Accepting the latter requires leaving open an $\mathsf{A} \wedge \neg\mathsf{B}$ world. These requirements are incompatible.

Standoffs are productive disagreements. s_1 accepts $\mathsf{A} \to \mathsf{B}$ and s_2 accepts $\mathsf{A} \to \neg\mathsf{B}$. But CNC predicts that the two conditionals cannot be accepted by any single agent. So s_1 and s_2 disagree productively about $\mathsf{A} \to \mathsf{B}$ and $\mathsf{A} \to \neg\mathsf{B}$.

Propositionality fails, because accepting a conditional is not a matter of accepting a proposition. When an information state s accepts $\Diamond\mathsf{A}$ and accepts $\neg(\mathsf{A} \wedge \neg\mathsf{B})$, it accepts $\mathsf{A} \to \mathsf{B}$. But once s is updated with $\neg\mathsf{A}$, $\mathsf{A} \to \mathsf{B}$ is no longer accepted. This rules out Propositionality, which implies that if s accepts $\mathsf{A} \to \mathsf{B}$ then $s[\neg\mathsf{A}]$ does too. This is why standoffs are compatible disagreements. s_1 accepts $\mathsf{A} \to \mathsf{B}$ while s_2 accepts $\mathsf{A} \to \neg\mathsf{B}$, despite the fact that s_1 and s_2 overlap.

7 Comparisons

Before concluding, I briefly compare my analysis of conditionals to other proposals.

7.1 The material conditional

The analysis avoids two paradoxes of material implication:

(15) **True Consequent.** London is in England. ??So: if Paddington Station is in France, London is in England.

(16) **False Antecedent.** The moon is not made of cheese.??So: if the moon is made of cheese, it is made of ketchup.

In each inference, the conclusion is unlicensed because its antecedent is inconsistent with the speaker's information.

On the other hand, in the presence of the usual dynamic meaning for negation, the theory shares with the material conditional the usual bad predictions about conditionals under negation:

(17) **Negation.** It's not true that if Alex is at the party, Billy is at the party. ??So: Alex is at the party and Billy isn't.

True Consequent and False Antecedent involve unembedded conditionals, while Negation involves an embedded conditional. My story about standoffs essentially appeals to the update and acceptance properties of unembedded conditionals. But nothing in that story is immediately committed to a prediction about how the conditional behaves under the scope of negation.[14]

[14] One solution follows Willer 2018 and uses a bilateral dynamic system, which defines positive and negative updates for each expression. A positive update with $\neg(\mathsf{A} \to \mathsf{B})$ is the negative

7.2 Contextualism

Kratzer 1986, Kremer 1987, Weatherson 2001, Nolan 2003, and Rothschild 2013 all defend contextualist theories where the conditional expresses a proposition, but this proposition is dependent on the relevant information in context. For example, imagine that $\mathsf{A} \to \mathsf{B}$ is true in a context just in case the speaker's knowledge implies the corresponding material conditional $\neg(\mathsf{A} \wedge \neg\mathsf{B})$. The resulting theory predicts that each speaker's conditionals is true in their context of utterance, and in this way the theory uses context sensitivity to resolve the prima facie incompatibility of their respective utterances.

In some ways, my theory is broadly sympathetic to this strategy. Where a contextualist lets a conditional express a different proposition in different contexts, I let the conditional's acceptance in a state depend on structural properties of that state (in particular, whether the antecedent is consistent with it). But the difference between my theory and the contextualist's is that I make immediate predictions about how the hearer should update. Gibbard 1981 objects to a simple contextualism about conditionals on exactly this basis.[15] Gibbard 1981 observes that the hearer in standoffs can be totally ignorant of the identity of the speaker, and hence of the proposition expressed by the speaker's utterance, given the contextualist semantics. In that case, which proposition should the hearer add to her beliefs? My theory solves the problem, making clear predictions about how the hearer should respond to conditional assertions. When the hearer updates their beliefs, the relevant body of information for assessing the possibility of the antecedent is their own, not that of a contextually supplied knower. In addition, my proposal captures the subtle ways in which speakers disagree (compatibly and productively) regarding conditional information. If each speaker expressed a different proposition with their respective conditionals, it is hard to see why their two assertions seem to conflict, and in particular why no single hearer can accept both of the propositions at once.

7.3 Expressivism

Because I deny Propositionality, my theory is consonant with the expressivism defended by Gibbard 1981 and Edgington 1995 in response to standoffs. My theory differs from these by supplying clear predictions for how the hearer should change their beliefs in response to conditional assertions. One might formalize those theories, for example, by letting the meaning of a sentence be a set of probability functions, and letting the meaning of the conditional $\mathsf{A} \to \mathsf{B}$ be the set of states where the probability of B given A is sufficiently high. But such a theory provides only one half of a theory of communication. It leaves unsettled

update with $\mathsf{A} \to \mathsf{B}$, which could be the positive update with $\mathsf{A} \to \neg\mathsf{B}$. Another solution departs further from the dynamic framework, and does not define updating compositionally. Instead, one can define acceptance and rejection compositionally, and define updating in terms of acceptance, so that $s[\mathsf{A}]$ is the largest subset of s that accepts s. Then one could say that s accepts $\mathsf{A} \to \mathsf{B}$ when $s[\mathsf{A}]$ accepts B and s accepts $\Diamond\mathsf{A}$, and that s accepts $\neg(\mathsf{A} \to \mathsf{B})$ when s accepts $\mathsf{A} \to \neg\mathsf{B}$ and s accepts $\Diamond\mathsf{A}$.

[15] See Stalnaker 1984 also.

how a hearer should respond to a speaker who asserts a claim when her credence function accepts it.

8 Conclusion

My theory predicts that when someone has ruled out an antecedent, they can no longer update on or accept indicative conditionals with that antecedent. This prediction is controversial.[16] Dorst Forthcoming argues against this idea, using data like the following:

(18) a. I know that Oswald shot Kennedy. But I also know that if he didn't, then someone else did.
b. I know Steph stole some cookies from the jar. But I don't know that if she didn't, then someone else did.

Dorst Forthcoming explains the contrast through the principle that believing a conditional requires having conditional belief. Conditional belief can in turn be understood in terms of a disposition to believe the consequent if one were to learn the antecedent. In (18), the agent is disposed to believe the consequent upon learning the antecedent. In (18-b), she isn't.

There are two ways one might respond. The first is to reject the direct semantic significance of the data. For one thing, the data seems sensitive to order:

(19) ??I know that if Oswald didn't shoot Kennedy, someone else did. But I also know that Oswald did shoot Kennedy.

For another, the data is one instance of a general pattern that cries out for pragmatic explanation. von Fintel and Gillies 2021 worry about analogous challenges by Lassiter 2016 to the strength of epistemic necessity claims:

(20) The car must have travelled 138,000. But I don't know for sure.

von Fintel and Gillies 2021 argue that this is a pragmatic phenomenon where 'speakers make some strong claims and then back off them a bit or... hesitate in the presence of a salient chance of error from making those claims.' (p. 20).

Another response takes the challenge more seriously, and agrees that an agent who rules out A can still accept $\mathsf{A} \to \mathsf{B}$ as long as she conditionally believes B, given A. Here, one natural next step would draw on recent work by Ciardelli 2020, who offers an account of indicative conditionals according to which asserting $\mathsf{A} \to \mathsf{B}$ narrows down our information to incorporate a requirement that when we revise that information with A, it accepts B. The question here would be how

[16] Compare Morton 1997: 'the Gibbard phenomenon is not that when we know that the antecedent of an indicative conditional is false we know there is no correct opinion as to its truth. For in the standard example we *know* that Oswald shot Kennedy, and also know that if he did not someone else' (p. 102).

to integrate these insights into an account of standoffs that retains the structural properties from earlier.

The reflections above do little to affect the basic desiderata on a theory of standoffs. Each speaker in a standoff considers the antecedent possible, as does the hearer up until her final update. So the treatment of conditionals with epistemically impossible antecedents should not affect what one says here. Rather, this contingency is only relevant for the hearer after they have updated on each conditional, and ruled out the possibility of the coin flipping. Here, though, the proposal which takes conditional belief seriously will agree with this paper's verdicts about what the hearer accepts. In particular, such a proposal will agree that the hearer does not accept either conditional after hearing from both speakers. This is because in standoffs the hearer does not have the relevant conditional beliefs after updating on conditional testimony. Rather, her epistemic position regarding heads versus tails is symmetric. Even when we imagine her revising her beliefs to include the coin being flipped, we cannot thereby imagine her counting as believing that the coin landed heads rather than tails.[17]

[17] I leave for future research whether the account can extend to analogues involving subjunctive conditionals. There is some controversy in the literature about whether such cases exist. Edgington 1995, Edgington 1997, and Swanson 2013 argue that there are such cases, while Morton 1997 and Williams 2008a among others express skepticism.

References

Kent Bach. Applying pragmatics to epistemology. *Philosophical Issues*, 18:68–88, 2008.

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

Ivano Ciardelli. Indicative conditionals and graded information. *Journal of Philosophical Logic*, 49:509–549, 2020.

Paul Dekker. *Transsentential Meditations: Ups and Downs in Dynamic Semantics*. PhD thesis, Universiteit van Amsterdam, 1993.

Kevin Dorst. Abominable kk failures. *Mind*, Forthcoming.

Dorothy Edgington. On conditionals. *Mind*, 104(414):235–329, 1995.

Dorothy Edgington. Truth, objectivity, counterfactuals, and gibbard. *Mind*, 106 (421):107–116, 1997.

Allan Gibbard. Two recent theories of conditionals. In William L. Harper, Robert Stalnaker, and Glenn Pearce, editors, *IFS: Conditionals, Belief, Decision, Chance and Time*, volume 15 of *The University of Western Ontario Series in Philosophy of Science*, pages 211–247. Springer Netherlands, 1981.

Anthony S. Gillies. Epistemic conditionals and conditional epistemics. *Nous*, 38 (4):585–616, 2004.

Anthony S. Gillies. On truth-conditions for *if* (but not quite only *if*). *Philosophical Review*, 118(3):325–349, 2009.

Paul Grice. Logic and conversation. In *Studies in the Way of Words*. Harvard University Press, Cambridge MA, 1975/1989a.

Irene Heim. On the projection problem for presuppositions. *WCCFL*, 2:114–125, 1983.

Nathan Klinedinst and Daniel Rothschild. Quantified epistemic modality. Manuscript, 2015.

Angelika Kratzer. Conditionals. *Chicago Linguistics Society*, 22:1–15, 1986.

Michael Kremer. ‘if’ is unambiguous. *Noûs*, 21(2):199–217, 1987.

Jonathan Kvanvig. Assertions, knowledge, and lotteries. In Patrick Greenough and Duncan Pritchard, editors, *Williamson on Knowledge*. Oxford University Press, 2009.

Jonathan Kvanvig. Norms of assertion. In Jessica Brown and Herman Cappelen, editors, *Assertion: New Philosophical Essays*. Oxford University Press, 2011.

Daniel Lassiter. *Must*, knowledge, and (in)directness. *Natural Language Semantics*, 24(2):117–163, June 2016.

John MacFarlane. *Assessment Sensitivity: Relative Truth and its Applications*. Oxford University Press, 2014.

Adam Morton. Can edgington gibbard counterfactuals? *Mind*, 106(421):101–105, 1997.

Daniel Nolan. Defending a possible-worlds account of indicative conditionals. *Philosophical Studies*, 116(3):215–69, 2003.

Daniel Rothschild. Do indicative conditionals express propositions? *Noûs*, 47(1), 2013.

Daniel Rothschild. Conditionals and propositions in semantics. *Journal of Philosophical Logic*, 44(6):781–791, 2015.

Daniel Rothschild and Seth Yalcin. On the dynamics of conversation. *Noûs*, 2015.

Robert Stalnaker. Indicative conditionals. *Philosophia*, 5:269–86, 1975.

Robert Stalnaker. *Inquiry*. Cambridge University Press, 1984.

William B. Starr. A uniform theory of conditionals. *Journal of Philosophical Logic*, 43(6):1019–1064, 2014.

William B. Starr. Indicative conditionals, strictly. 2021.

Eric Swanson. Subjunctive biscuit and stand-off conditionals. *Philosophical Studies*, 163:637–648, 2013.

Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25(3):221–261, 1996.

Kai von Fintel and Thony Gillies. Still going strong. *Natural Language Semantics*, 29(1):91–113, 2021.

Brian Weatherson. Indicative and subjunctive conditionals. *Philosophical Quarterly*, 51:200–16, 2001.

Brian Weatherson. Conditionals and indexical relativism. *Synthese*, 166:333–357, 2009.

Malte Willer. Lessons from sobel sequences. *Semantics and Pragmatics*, 10(4): 1–57, 2017.

Malte Willer. Simplifying with free choice. *Topoi*, 37(3):379–392, 2018.

J. Robert G. Williams. Chances, counterfactuals and similarity. *Philosophy and Phenomenological Research*, 77(2), 2008a.

J Robert G Williams. Conversation and conditionals. *Philosophical Studies*, 138 (2):211–223, 2008b.

Timothy Williamson. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000.

Seth Yalcin. Epistemic modality *De Re*. *Ergo*, 2(19), 2015.