Conditionals: between language and reasoning

Class 5: The problem of disjunctive antecedents

May 21, 2019

Part I:
The problem of disjunctive antecedents

(1) If Thorpe or Wilson were to win the next general election, Britain would prosper. (Fine 75)

(1) If Thorpe or Wilson were to win the next general election, Britain would prosper. (Fine 75)

$\rightsquigarrow$ If Thorpe were to win, Britain would prosper.
$\rightsquigarrow$ If Wilson were to win, Britain would prosper.

(1) If Thorpe or Wilson were to win the next general election, Britain would prosper. (Fine 75)

$\rightsquigarrow$ If Thorpe were to win, Britain would prosper.
$\rightsquigarrow$ If Wilson were to win, Britain would prosper.

(2) If you had pressed button A or button B, a light would have turned on.

(1) If Thorpe or Wilson were to win the next general election, Britain would prosper. (Fine 75)

$\rightsquigarrow$ If Thorpe were to win, Britain would prosper.
$\rightsquigarrow$ If Wilson were to win, Britain would prosper.

(2) If you had pressed button A or button B, a light would have turned on.

$\rightsquigarrow$ If you had pressed button A, a light would have turned on.
$\rightsquigarrow$ If you had pressed button B, a light would have turned on.

- In general, the following entailment patterns seems valid. They are known as simplification of disjunctive antecedents.

$$\frac{A \vee B > C}{A > C}\ (\mathrm{SDA}_1) \qquad \frac{A \vee B > C}{B > C}\ (\mathrm{SDA}_2)$$

- In fact, it seems that we have the following equivalence:

$$A \vee B > C \quad \equiv \quad (A > C) \wedge (B > C)$$

An argument for the validity of SDA: what has the speaker said?

(3) A: Last week, John went to London and to Paris. $L \wedge P$
B: That's not true: he didn't go to London.

B challenges A's claim $L \wedge P$ by denying L.
This makes good sense since $L \wedge P \models L$.

(4) A: Last week, John went to London or to Paris. $L \vee P$
B: #That's not true: he didn't go to London.

B challenges A's claim $L \vee P$ by denying L.
This seems to misunderstand what A said, since $L \vee P \not\models L$.

A: If Thorpe or Wilson were to win, Britain would prosper. $T \vee W > P$

B: That's not true! If Thorpe were to win, Britain would go bankrupt.

- B can sensibly challenges A's claim by denying that $T > P$.
- This suggests that A's claim implies that $T > P$:

$$T \vee W > P \quad \models \quad T > P$$

SDA is not valid in minimal change semantics

- If A and B are equally distant ($S_w^A = S_w^B$) then SDA goes through:

$$w \Vdash A \vee B > C \iff w \Vdash (A > C) \wedge (B > C)$$

- However, if B is more remote than A ($S_w^A \subset S_w^B$), then the closest $A \vee B$ worlds are just the closest A worlds

$$w \Vdash A \vee B > C \iff w \Vdash A > C$$

- Is this right?
- Perhaps we only considered cases where the disjuncts are equidistant. Let's try to consider a case where one disjunct is more remote.

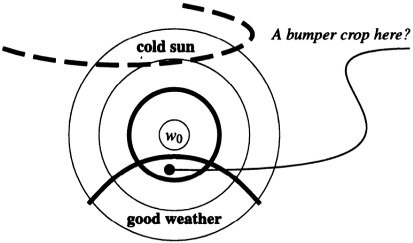
Consider the following scenario: the summer is over and you and I are visiting a farm. The owner of the farm is complaining about last summer's weather. To give us an example of its devastating effects, he points to the site where he used to grow huge pumpkins: there is a bunch of immature pumpkins and many ruined pumpkin plants. He then utters the counterfactual in (1):

(1) If we had had good weather this summer or the sun had grown cold, we would have had a bumper crop.
(A variation on an example in Nute 1975.)

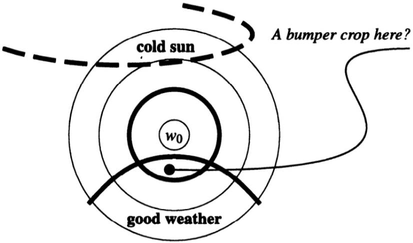
We conclude, right then, that there is something strange about this farmer. We have a strong intuition that the counterfactual in (1) is false: if we had had good weather this summer, he would have had a good crop, but we know for a fact—and we assume that the farmer does too—that if the sun had grown cold, the pumpkins, much as everything else, would have been ruined.

(Alonso-Ovalle 2009)

- Presumably, cold-sun worlds are more remote than the closest good-weather worlds.



- Presumably, cold-sun worlds are more remote than the closest good-weather worlds.

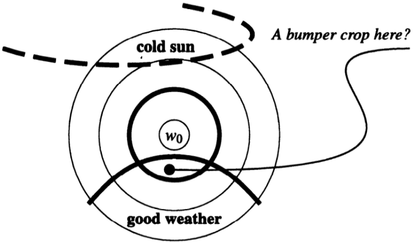


- Then the farmer's statement (5) is predicted to be equivalent to (6):

(5) If we had had good weather or the sun had grown cold, we would have had a bumper crop. $G \vee C > B$

(6) If we had had good weather, we would have had a bumper crop. $G > B$

- Presumably, cold-sun worlds are more remote than the closest good-weather worlds.



- Then the farmer's statement (5) is predicted to be equivalent to (6):

(5) If we had had good weather or the sun had grown cold, we would have had a bumper crop. $G \vee C > B$

(6) If we had had good weather, we would have had a bumper crop. $G > B$

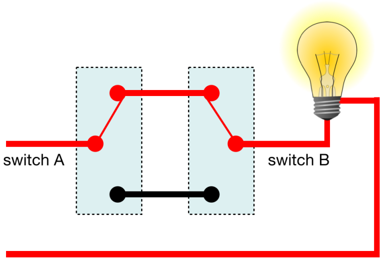
- This is wrong. In assessing (5), we do not disregard cold-sun worlds. We take them into account, and that's why we don't judge (5) as true.

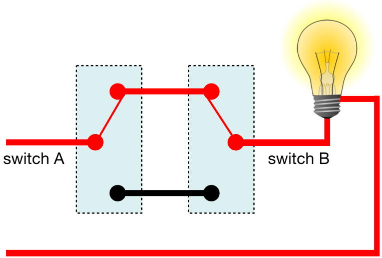
So far, it looks like SDA is an argument against minimal change semantics. One may hope that some other account of counterfactuals fares better here.

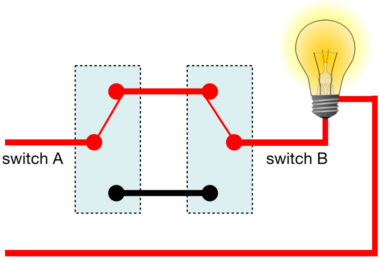
But the problem runs deeper: regardless of one's theory of counterfactuals, the facts about SDA call for a revision of our treatment of disjunction.

The context (Ciardelli, Zhang, and Champollion 2018)

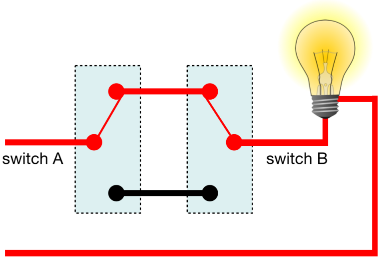
Imagine a long hallway with a light in the middle and with two switches, one at each end. One switch is called switch A and the other one is called switch B. As the following wiring diagram shows, the light is on whenever both switches are in the same position (both up or both down); otherwise, the light is off. Right now, switch A and switch B are both up, and the light is on. But things could be different...



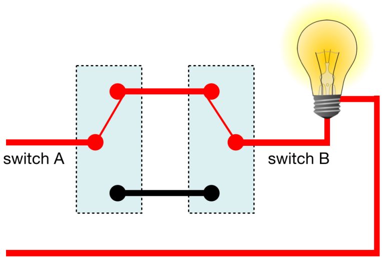

switch A
switch B



(7) If switch A was down, the light would be off.

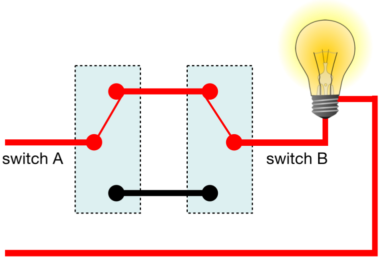


(7) If switch A was down, the light would be off. ✓



(7) If switch A was down, the light would be off. ✓

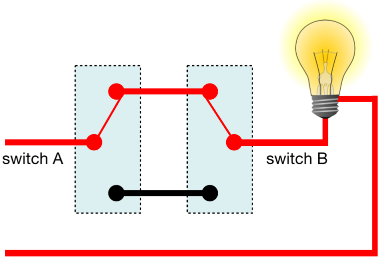
(8) If switch B was down, the light would be off. ✓



(7) If switch A was down, the light would be off. ✓

(8) If switch B was down, the light would be off. ✓

(9) If switch A or switch B was down, the light would be off. ✓

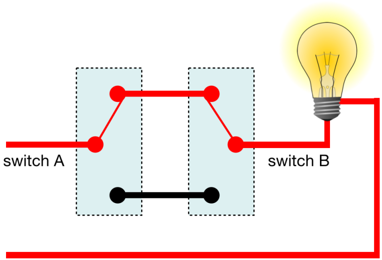


(7) If switch A was down, the light would be off. ✓

(8) If switch B was down, the light would be off. ✓

(9) If switch A or switch B was down, the light would be off. ✓

(10) If switch A and switch B were not both up, the light would be off.

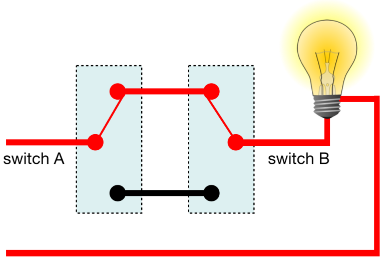


(7) If switch A was down, the light would be off. ✓

(8) If switch B was down, the light would be off. ✓

(9) If switch A or switch B was down, the light would be off. ✓

(10) If switch A and switch B were not both up, the light would be off. ✗



(1)	If switch A was down, . . . off.	$\overline{A} > Off$
(2)	If switch B was down, . . . off.	$\overline{B} > Off$
(3)	If switch A or switch B was down, . . . off.	$\overline{A} \vee \overline{B} > Off$
(4)	If switch A and switch B were not both up, . . . off.	$\neg(A \wedge B) > Off$

- Minimal change semantics fits within the framework of intensional semantics, where sentences are analyzed in terms of truth-conditions at possible worlds.
- The semantic value of a sentence φ is taken to be a set of possible worlds:

$$|\varphi| = \{w \in W \mid w \Vdash \varphi\}$$

- Consider the clauses corresponding to the antecedents of (3) and (4):

(5) Switch A or switch B is down. $\overline{A} \vee \overline{B}$

(6) Switch A and switch B are not both up. $\neg(A \wedge B)$

- They are truth-conditionally equivalent.
  Given down=not up, this is Morgan's law $\neg(A \wedge B) \equiv \neg A \vee \neg B$
- Thus, $|\overline{A} \vee \overline{B}| = |\neg(A \wedge B)|$

- Take an arbitrary account of counterfactuals which is compositional:

$$|\varphi > \psi| \ = \ \sigma(|\varphi|, |\psi|)$$

- Consider again (3) and (4):

(3)	If switch A or switch B was down, ...	$\overline{A} \vee \overline{B} > \mathit{Off}$
(4)	If switch A and switch B were not both up, ...	$\neg(A \wedge B) > \mathit{Off}$

- We have:

$$\begin{aligned} |\overline{A} \vee \overline{B} > \mathit{Off}| &= \sigma(|\overline{A} \vee \overline{B}|, |\mathit{Off}|) \\ &= \sigma(|\neg(A \wedge B)|, |\mathit{Off}|) \quad = \quad |\neg(A \wedge B) > \mathit{Off}| \end{aligned}$$

- So, (3) and (4) are predicted to have the same truth-conditions.
- Our observations contradict this: in our context, (3) is true but (4) is not.

- Assuming that the truth-conditions of a counterfactual depend only on the truth-conditions of antecedent and consequent leads to wrong predictions.

- This problem arises regardless of how counterfactuals work (as long as compositionality is assumed).

- Assuming that the truth-conditions of a counterfactual depend only on the truth-conditions of antecedent and consequent leads to wrong predictions.
- This problem arises regardless of how counterfactuals work (as long as compositionality is assumed).
- Counterfactuals require a representation of antecedents which is more fine-grained than provided by truth-conditional semantics.
- In particular, this representation should break de Morgan's law and assign a different semantic value to $\neg A \vee \neg B$ and $\neg(A \wedge B)$.

- Notice that we don't want to depart too radically from intensional semantics, saying that logically equivalent antecedents are generally not inter-substitutable.
- The relevant phenomenon seems to be tied specifically to disjunction.
- Normally, if two non-disjunctive antecedents are truth-conditionally equivalent, then the corresponding counterfactuals are equivalent:

(5) If it hadn't rained for a fortnight, the crop would have been ruined.

(6) If it hadn't rained for two weeks, the crop would have been ruined.

(7) If I had not passed that exam, I wouldn't have graduated on time.

(8) If I had failed that exam, I wouldn't have graduated on time.

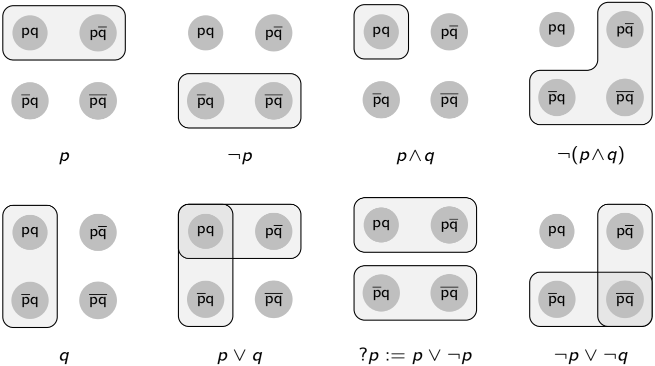
Part II
A solution in inquisitive semantics

Inquisitive semantics: a sketch

- The fundamental notion is not truth at a state of affairs, but rather support relative to a state of information ($s \models \varphi$).
- An information state is modeled as a set of possible worlds: those worlds which are compatible with the given information.
- The $\subseteq$-maximal information states supporting φ are called the alternatives. The set of alternatives is denoted by $\mathrm{Alt}(\varphi)$.
- So, a sentence determines a set of information states.
- The set of worlds where φ is true is $|\varphi| := \bigcup \mathrm{Alt}(\varphi)$.

Inquisitive semantics comes with a theory of propositional connectives, motivated by logical/algebraic considerations:

- $s \models p \iff s \subseteq V(p)$
- $s \models \varphi \wedge \psi \iff s \models \varphi$ and $s \models \psi$
- $s \models \varphi \vee \psi \iff s \models \varphi$ or $s \models \psi$
- $s \models \neg\varphi \iff$ there is no non-empty $t \subseteq s$ such that $t \models \varphi$



Inquisitiveness and counterfactuals

We adopt an idea due to Alonso-Ovalle (2009):

- Each alternative counts as a separate counterfactual assumption.
- The counterfactual is true if the consequent follows on each assumption.

Inquisitiveness and counterfactuals

We adopt an idea due to Alonso-Ovalle (2009):

- Each alternative counts as a separate counterfactual assumption.
- The counterfactual is true if the consequent follows on each assumption.

Implementation: inquisitive lifting (simplified)

Let $f(\cdot\,,\cdot)$ be a standard selection function for counterfactuals.

$$w \Vdash \varphi > \psi \iff \forall a \in \mathrm{Alt}(\varphi) : f(w, a) \subseteq |\psi|$$

Non-inquisitive antecedents

Consider a counterfactual $p > q$. Since $\mathsf{Alt}(p) = \{|p|\}$, we have:

$$\begin{aligned} w \Vdash p > q \quad &\iff \quad \forall a \in \mathsf{Alt}(p) : f(w,a) \subseteq |q| \\ &\iff \quad \forall a \in \{|p|\} \; : f(w,a) \subseteq |q| \\ &\iff \quad f(w,|p|) \subseteq |q| \end{aligned}$$

Thus we retrieve the standard truth-conditions.

Non-inquisitive antecedents

Consider a counterfactual $p > q$. Since $\mathsf{Alt}(p) = \{|p|\}$, we have:

$$\begin{aligned} w \Vdash p > q &\iff \forall a \in \mathsf{Alt}(p) : f(w, a) \subseteq |q| \\ &\iff \forall a \in \{|p|\} : f(w, a) \subseteq |q| \\ &\iff f(w, |p|) \subseteq |q| \end{aligned}$$

Thus we retrieve the standard truth-conditions.

The same holds whenever the antecedent α is non-inquisitive, i.e., $\mathsf{Alt}(\alpha) = \{|\alpha|\}$.

Non-inquisitive antecedents

Consider a counterfactual $p > q$. Since $\mathsf{Alt}(p) = \{|p|\}$, we have:

$$
\begin{aligned}
w \Vdash p > q \quad &\iff \quad \forall a \in \mathsf{Alt}(p) : f(w, a) \subseteq |q| \\
&\iff \quad \forall a \in \{|p|\} : f(w, a) \subseteq |q| \\
&\iff \quad f(w, |p|) \subseteq |q|
\end{aligned}
$$

Thus we retrieve the standard truth-conditions.

The same holds whenever the antecedent α is non-inquisitive, i.e., $\mathsf{Alt}(\alpha) = \{|\alpha|\}$.

In particular, if two antecedents are not inquisitive and have same truth-conditions, then the resulting counterfactuals are predicted to be equivalent:

(9) If it hadn't rained for a fortnight, the crop would be ruined.

(10) If it hadn't rained for two weeks, the crop would be ruined.

Disjunctive antecedents

Consider a counterfactual $p \vee q > r$. Since $\mathrm{Alt}(p \vee q) = \{|p|, |q|\}$ we have:

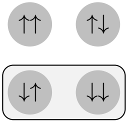
$$
\begin{aligned}
w \Vdash p \vee q > r \quad &\iff \quad \forall a \in \mathrm{Alt}(p \vee q) : f(w, a) \subseteq |r| \\
&\iff \quad \forall a \in \ \{|p|, |q|\} \ : f(w, a) \subseteq |r| \\
&\iff \quad f(w, |p|) \subseteq |r| \text{ and } f(w, |q|) \subseteq |r| \\
&\iff \quad w \Vdash p > r \text{ and } w \Vdash q > r
\end{aligned}
$$

Therefore, SDA is vindicated:

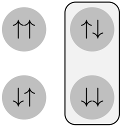
$$p \vee q > r \equiv (p > r) \wedge (q > r)$$

Back to our switches

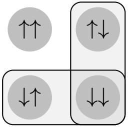
Here are the inquisitive semantics representations for our antecedents:



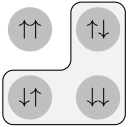
(i) $\overline{A}$



(j) $\overline{B}$



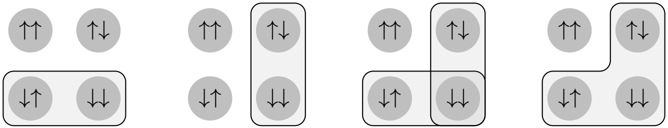
(k) $\overline{A} \vee \overline{B}$



(l) $\neg(A \wedge B)$

Back to our switches

Here are the inquisitive semantics representations for our antecedents:

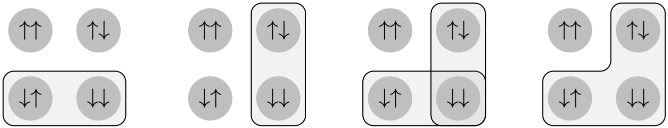


(m) $\overline{A}$ (n) $\overline{B}$ (o) $\overline{A} \vee \overline{B}$ (p) $\neg(A \wedge B)$

- This theory breaks de Morgan's law: $\overline{A} \vee \overline{B}$ and $\neg(A \wedge B)$ have the same truth-conditions, but different alternatives.

Back to our switches

Here are the inquisitive semantics representations for our antecedents:

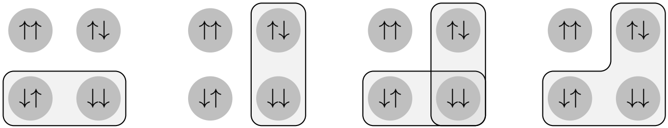


(q) $\overline{A}$ (r) $\overline{B}$ (s) $\overline{A} \vee \overline{B}$ (t) $\neg(A \wedge B)$

- This theory breaks de Morgan's law: $\overline{A} \vee \overline{B}$ and $\neg(A \wedge B)$ have the same truth-conditions, but different alternatives.
- The antecedents $\overline{A}$, $\overline{B}$, and $\neg(A \wedge B)$ are non-inquisitive: so the corresponding counterfactuals are interpreted standardly.

Back to our switches

Here are the inquisitive semantics representations for our antecedents:



(u) $\overline{A}$ (v) $\overline{B}$ (w) $\overline{A} \vee \overline{B}$ (x) $\neg(A \wedge B)$

- This theory breaks de Morgan's law: $\overline{A} \vee \overline{B}$ and $\neg(A \wedge B)$ have the same truth-conditions, but different alternatives.
- The antecedents $\overline{A}$, $\overline{B}$, and $\neg(A \wedge B)$ are non-inquisitive: so the corresponding counterfactuals are interpreted standardly.
- By contrast, $\overline{A} \vee \overline{B}$ is inquisitive: inquisitive lifting will ensure that $\overline{A} \vee \overline{B} > \mathit{Off}$ is interpreted as equivalent to $(\overline{A} > \mathit{Off}) \wedge (\overline{B} > \mathit{Off})$.

- Now suppose our selection function f is such that:
  - $f(w, |\overline{A}|) \subseteq |Off|$
  - $f(w, |\overline{B}|) \subseteq |Off|$
  - $f(w, |\neg(A \wedge B)|) \nsubseteq |Off|$

- Then the selection function approach delivers the right verdicts for (1), (2), (4):

(1)	$\overline{A} > Off$	✓
(2)	$\overline{B} > Off$	✓
(3)	$\overline{A} \vee \overline{B} > Off$	✓
(4)	$\neg(A \wedge B) > Off$	✗

- Now suppose our selection function f is such that:
  - $f(w, |\overline{A}|) \subseteq |\mathit{Off}|$
  - $f(w, |\overline{B}|) \subseteq |\mathit{Off}|$
  - $f(w, |\neg(A \wedge B)|) \not\subseteq |\mathit{Off}|$
- Then the selection function approach delivers the right verdicts for (1), (2), (4):

(1)	$\overline{A} > \mathit{Off}$	✓
(2)	$\overline{B} > \mathit{Off}$	✓
(3)	$\overline{A} \vee \overline{B} > \mathit{Off}$	✓
(4)	$\neg(A \wedge B) > \mathit{Off}$	✗

- The inquisitive account will make the same predictions for these sentences.
- But the original account would predict (3) to be false, since (3) $\equiv$ (4); instead, our inquisitive account will predict (3) to be true, since (3) $\equiv$ (1) $\wedge$ (2).

Part III:

SDA violations?

McKay and van Inwagen's example:

(5) If Spain had sided with the Axis or the Allies in WWII,
it would have sided with the Axis.

(6) If Spain had sided with the Allies, it would have sided with the Axis.

- Clearly, (5) does not imply (6), contrary to SDA.
- Plain minimal change semantics seems to make the right predictions: in the closest worlds where Spain takes sides, it sides with the Axis.

McKay and van Inwagen's example:

(5) If Spain had sided with the Axis or the Allies in WWII,
it would have sided with the Axis.

(6) If Spain had sided with the Allies, it would have sided with the Axis.

- Clearly, (5) does not imply (6), contrary to SDA.
- Plain minimal change semantics seems to make the right predictions: in the closest worlds where Spain takes sides, it sides with the Axis.
- However, there is a problem with this.

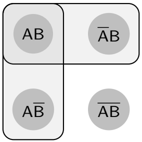
(7) If Spain had sided with the Axis or the Allies in WWII, Hitler would have been pleased.

- Intuitively, (7) is not true.
- But if the closest antecedent worlds are ones where Spain joins the axis, minimal change semantics would predict (7) to be true.
- It seems that failures of SDA occur only with very special consequents. . . why?

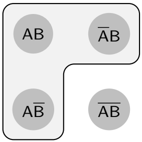
A proposal (essentially Alonso-Ovalle's)

- It is possible to merge the alternatives for the antecedent into one via an operator !:

$$\mathsf{Alt}(!\varphi) = \{ \bigcup \mathsf{Alt}(\varphi) \}$$



(y) $p \vee q$



(z) $!(p \vee q)$

- If ! is inserted, we retrieve the predictions of minimal change semantics.

- However, interpreting a disjunctive antecedent as $!(p \vee q)$ is possible only as a last resource, to avoid a contradictory interpretation.

(8) If Spain had sided with the Axis or the Allies in WWII,
it would have sided with the Axis.

- If Spain siding with the Allies is an entertainable supposition, and if it is impossible to side with both the Axis and the Allies, the SDA reading of (8) is contradictory.
- This is what motivates the insertion of '!'.
- On the other hand, the SDA reading of (9) is consistent, so ! is not inserted and we get an SDA reading.

(9) If Spain had sided with the Axis or the Allies in WWII,
Hitler would have been pleased.

- But there are other puzzling observations about such sentences.
- The following variants in (11) sound very odd.

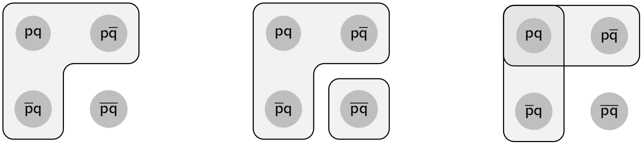
(10) a. If Spain had sided with the Axis or if it had sided with the Allies, it would have sided with the Axis.

b. If Spain had sided with the Axis, or had betrayed the Axis and sided with the Allies, it would have sided with the Axis.

- Making the disjunction syntactically prominent seems to force an SDA reading.
- This different is unexpected for both minimal change semantics and A-O: why would it matter whether the disjunction is at the VP level?

In inquisitive semantics, the availability of non-SDA reading is linked to the operator ! which cancels inquisitive content. This operator is also used to form polar questions.

(11)	a.	Mark went to London or to Paris.	$!(p \vee q)$
	b.	Did Mark go to London or Paris?	$?!(p \vee q)$
	c.	Did Mark go to London, or to Paris?	$p \vee q$



$!(p \vee q)$ $?!(p \vee q)$ $p \vee q$

We can observe a very interesting parallel with the SDA data:

(12)	Does Alice speak French or German?	✓ Polar
(13)	Does Alice speak French or does she speak German?	✗ Polar

Further readings of conditionals with disjunctive antecedents:

- Ciardelli 2016, Lifting conditionals to inquisitive semantics.
- Bledin 2017, Fatalism and the logic of unconditionals.
- Khoo 2018, Disjunctive antecedent conditionals.
- Cariani and Goldstein 2018, Conditional heresies.
- Santorio 2018, Alternatives and truth-makers in conditional semantics.

Appendix A

Experimental results

Testing the judgments

- We described the scenario by presenting the text and figure above.
- Each participant saw a target item and the filler, in random order, and was asked to judge these sentences as 'true', 'false', or 'indeterminate'.

Targets

(1) If switch A was down, the light would be off.

(2) If switch B was down, the light would be off.

(3) If switch A or switch B was down, the light would be off.

(4) If switch A and switch B were not both up, the light would be off.

(5) If switch A and switch B were not both up, the light would be on.

Filler

(6) If switch A and switch B were both down, the light would be off. False

Table 3: Results of the main experiment

Sentence	Number	True	(%)	False	(%)	Indet.	(%)
$\overline{A} >$ OFF	256	169	66.02%	6	2.34%	81	31.64%
$\overline{B} >$ OFF	235	153	65.11%	7	2.98%	75	31.91%
$\overline{A} \vee \overline{B} >$ OFF	362	251	69.33%	14	3.87%	97	26.80%
$\neg(A \wedge B) >$ OFF	372	82	22.04%	136	36.56%	154	41.40%
$\neg(A \wedge B) >$ ON	200	43	21.50%	63	31.50%	94	47.00%

Pre-test I

- Our aim was to test whether the antecedents of (3) and (4) are indeed perceived as truth-conditionally equivalent.

(7) Switch A or switch B is down.

(8) Switch A and switch B are not both up.

- More specifically, we wanted to check if (7) is interpreted exclusively.
- We presented participants with the picture below, and asked them to judge these sentences as 'true', 'false', or 'indeterminate'.

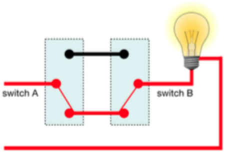
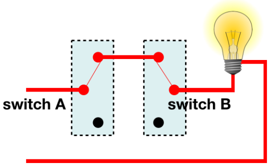


Table 1: Results of Pretest I

Sentence	Number	True	(%)	False	(%)	Indeterminate	(%)
$\overline{A} \vee \overline{B}$	145	118	81.38%	23	15.86%	4	2.76%
$\neg(A \wedge B)$	130	118	90.77%	11	8.46%	1	0.77%

Post-hoc test I

- We wanted to make sure that (4) are (5) are not rejected based on context-independent reasons (e.g., excessive processing load).
- For this, we tested our sentences in a different scenario: the light is on if and only if the switches are both up.



Sentence	Number	True	(%)	False	(%)	Indet.	(%)
$\overline{A}$ > OFF	52	41	78.85%	5	9.61%	6	11.54%
$\overline{B}$ > OFF	68	60	88.24%	5	7.35%	3	4.41%
$\overline{A}\vee\overline{B}$ > OFF	110	104	94.55%	1	0.91%	5	4.54%
$\neg(A\wedge B)$ > OFF	116	99	85.34%	9	7.76%	8	6.90%
$\neg(A\wedge B)$ > ON	103	19	18.45%	79	76.70%	5	4.85%

Post-hoc test II

- We wanted to make sure that our assumption of treating down as equivalent to not up was innocent.
- For this, we tested versions of our sentences where down is replaced by not up, in the setting of our main experiment.

Sentence	Number	True	(%)	False	(%)	Indet.	(%)
$\neg A >$ OFF	36	27	75.00%	1	2.78%	8	22.22%
$\neg B >$ OFF	43	28	65.12%	7	16.28%	8	18.60%
$\neg A \vee \neg B >$ OFF	80	48	60.00%	16	20.00%	16	20.00%
$\neg(A \wedge B) >$ OFF	372	82	22.04%	136	36.56%	154	41.40%
$\neg(A \wedge B) >$ ON	200	43	21.50%	63	31.50%	94	47.00%