Exhaustification
in the semantics of cause and because*

Dean McHugh
ILLC, University of Amsterdam

MLC Seminar
Friday, 5 November 2021

In this talk we analyze the semantics of sentences such as (1).

(1) a. The light turned on because Alice flicked the switch.
b. Flicking the switch caused the light to turn on.

Figure 1: An everyday example of causality.

Research on the semantics of causal claims raises two questions:

(2) The modeling question: What kind of information do we use when we judge that a causal relation holds?

(3) The meaning question: Once we have a causal model, what is the semantics of causal claims, given in terms of that model?

*For fruitful discussions on the present material, I am grateful to Lorenzo Pinton, Morwenna Hoeks, Katrin Schulz, and Zhuoye Zhao. This research was funded by the NWO PhDs in the Humanities Grant (PGW.18.028).
1 Three properties of the meaning of ‘cause’ and ‘because’

Property 1: The comparative nature of cause and because.

- Lewis (1973, 536). An event \( e \) causally depends on an event \( c \) just in case the following two counterfactuals are true: if \( c \) had occurred, \( e \) would have occurred, and if \( c \) had not occurred, \( e \) would not have occurred.

- Wright (1985, 2011). NESS (Necessary Element of a Sufficient Set) test for causation: something is a cause just in case it is an element of a set of conditions that are jointly sufficient for the effect, but are not sufficient when the cause is removed from the set.

- Mackie’s INUS condition states that a cause is “an insufficient but non-redundant part of a condition which is itself unnecessary but sufficient for the result” (Mackie 1974, 64).

- Beckers. An actual causal claim is true just in case the cause produced the effect, and if the cause had not occurred, the absence of the cause would not have also produced the effect.

While these analyses differ in their details, they have the same overall shape. They consist of two conditions: one that considers what would happen in the presence of the cause (what we will call a ‘positive condition’) and one that considers what would happen in the absence of the cause (what we will call a ‘negative condition’).

Property 2: Asymmetry in strength between positive and negative conditions.

The second property of the semantics of cause and because we consider is an asymmetry in strength between the positive condition and the negative condition.

(4) a. Positive condition: in all scenarios where the cause occurs the relevant condition is met
b. Negative condition: in some scenario where the cause does not occur the relevant condition is met.

What exactly this ‘relevant condition’ is depends on the analysis in question; for example, in the NESS and INUS conditions it is the effect occurring, according to Beckers it is that the cause produces the effect.
**Property 3:** The positive and negative conditions have the same background.


(5)  
a. The light turned on because Alice flicked the switch.
b. Alice flicking the switch caused the light to turn on.

- These background facts are involved in checking the positive condition.
- E.g. in the NESS and INUS tests above one checks whether the presence of the cause is sufficient for the effect given some background facts: the ‘set’ in the words of the NESS test; the ‘condition’ in the words of the INUS test.
- There is also a background involved when evaluating the negative condition: the facts from the actual world that are held fixed when evaluating what would happen if the cause had not occurred.
- We will see evidence that these two backgrounds must be the same.

## 2 Preliminaries

- Let \( \Box_{f,g}(p)(q) \) be a universal counterfactual modal with modal base \( f \), ordering source \( g \), restrictor \( p \) and nuclear scope \( q \).
- We roughly paraphrase \( \Box_{f,g}(p)(q) \) as *Given the circumstances (f), the truth of \( p \) guaranteed the truth of \( q \).*
- Exhaustification is defined as follows, and is akin to a silent *only*.\(^1\)

(6) Exhaustification operator
\[
O_{\text{ALT}} \varphi \ := \ \varphi \land \forall \psi \in \text{ALT}((\varphi \text{ does not entail } \psi) \rightarrow \neg \psi).
\]

- We plug in the sufficiency condition, \( \Box_{f,g}(p)(q) \), for the prejacent.
- The alternatives result from replacing \( p \) with its polar alternative, \( \neg p \).
  \[
  \text{ALT} = \{ \Box_{f,g}(p)(q), \Box_{f,g}(\neg p)(q) \}
  \]
- For brevity we write this as \( \{ p, \neg p \} \).

\(^1\)For simplicity’s sake we use Krifka’s (1993) entry for *only*. The result in (7) also follows from Fox’s (2007) exhaustivity operator, based on the notion of innocent exclusion. For an overview and comparison of exhaustivity operators see Spector (2016).
3 The meaning of ‘cause’ and ‘because’

3.1 The simplified semantics

\[ O_{(p,\neg p)} f_g(p)(q) = \square f_g(p)(q) \land \neg \square f_g(\neg p)(q) \]

- The positive condition states that if the cause had occurred, the effect is guaranteed to occur (given the circumstances \( f \)), i.e. the cause was sufficient for the effect given the circumstances.
- The negative condition states that if the cause had not occurred, the effect might not have occurred.
- This is essentially the NESS test above, formalized in terms of counterfactual modality\(^2,3\)

\( O_{(p,\neg p)} f_g(p)(q) \)

\[ \Leftrightarrow p \land f_g(p)(q) \land \neg (p \land f_g(\neg p)(q)) \]

\[ \Leftrightarrow p \land f_g(p)(q) \land (p \lor \neg f_g(\neg p)(q)) \]

\[ \Leftrightarrow p \land f_g(p)(q) \]

\(^2\)We place the condition that the cause occurred (\( p \)) outside the scope of exhaustification because otherwise exhaustification would be vacuous, as we see in the following chain of equivalences.

\[ O_{(p,\neg p)} (p \land f_g(p)(q)) \]

\[ \Leftrightarrow p \land f_g(p)(q) \land \neg (p \land f_g(\neg p)(q)) \]

\[ \Leftrightarrow p \land f_g(p)(q) \land (p \lor \neg f_g(\neg p)(q)) \]

\[ \Leftrightarrow p \land f_g(p)(q) \]

\(^3\)Note that we do not need to add \( q \) as a conjunct to (8) since \( p \land f_g(p)(q) \) entails \( q \): if \( p \) is true and is sufficient for the truth of \( q \), then \( q \) is also true.

Our entry for because in (8) assigns the same status to the condition that the cause occurred (\( p \)) as we do to the other conjunct \( O_{(p,\neg p)} f_g(p)(q) \). Both are entailments. Alternatively, one might propose that \( p \) is encoded as a presupposition in the lexical semantics of because. Such a stipulation does not account for why some inferences rather than others are selected as presuppositions in the first place (see Abrusán 2011, 2016, for discussion). Moreover, the inferences from cause and because that their arguments are true is a soft presupposition in the sense of (Abusch 2002, 2010), as they are easily suspendable, as shown in (i).

(i) a. The outcry which followed Morgan was not because the House of Lords had changed the law but because the public mistakenly thought it had done so. (Source: Temkin 2002)

b. No, the coronavirus did not cause the death rate to drop in Chicago... Overall, deaths don’t appear to be declining. (Source: Politifact.com, 3 April 2020)

Romoli (2012, 2015) explains that the projection properties of because are in fact due to a scalar implicature. An utterance of \( \neg (q \text{ because } p) \) triggers the alternatives \( \neg p \) and \( \neg q \). Since \( \neg (q \text{ because } p) \) – whose meaning according to (8) is given in (iia) – entails neither alternative, we derive the implicatures in (iib).
(8) Semantics of *because* (simplified).

\[ \text{[because]} = \lambda p_{(s,t)} \lambda q_{(s,t)} : p \land O_{(p,\neg p)} \Box_{f,g}(p)(q). \]

3.2 Properties 1, 2, and 3 via exhaustification

Properties 1, 2 and 3 fall out immediately from exhaustification.

Accounting for Property 1.

- The comparative character of *because* (Property 1) results from the comparative nature of exhaustification, which compares the prejacent with its alternatives.

- We stipulate that in the semantics of *because*, the alternatives are the cause’s polar alternatives.

Accounting for Property 2.

- The asymmetry in strength between the positive and negative conditions (Property 2) results from the fact that exhaustification negates alternatives.

- Given the duality between universal and existential quantification, the negation contributed by exhaustification turns a necessity modal into a possibility modal, generating an asymmetry in strength.

- This parallels the behaviour of *only* when it composes with a universal quantifier.

(9) The effect is guaranteed to occur only if the cause occurs.

a. \( \Rightarrow \) The effect is not guaranteed to occur if the cause does not occur.

b. \( \nRightarrow \) The effect is guaranteed to not occur if the cause does not occur.

(10) Only\(\{p,\neg p\}\) \(\Box_{f,g}(p)(q)\)

a. \( \Rightarrow \neg \Box_{f,g}(\neg p)(q) \)

b. \( \nRightarrow \Box_{f,g}(\neg p)(\neg q) \)

\[(ii)\]

a. \( \neg(q \text{ because } p) \Leftrightarrow \neg p \lor \neg \Box_{f,g}(p)(q) \lor \Box_{f,g}(\neg p)(q) \)

b. \( O_{\text{ALT}} \neg(q \text{ because } p) \Leftrightarrow \neg(q \text{ because } p) \land p \land q, \) where

\( \text{ALT} = \{\neg(q \text{ because } p), \neg p, \neg q\}. \)

Given Romoli’s account, we can capture the projection properties of *because* without needing to assign a special status to \( p \) in the lexical semantics of *because*.
• In (9) we assume that broad focus on the if clause triggers its polar alternative.  

**Accounting for Property 3.**

• The fact that the positive and negative conditions have the same background (Property 3) falls out from the fact that exhaustification simply copies the modal’s parameters – the modal base ($f$) and ordering source ($g$) – without altering them.

### 3.3 The full semantics

• We replace $q$ in the simplified semantics with $p$ produce $q$.

• The idea, informally, is that $p$ produces $q$ just in case there is a chain of dependence from $p$ to $q$.

$$(11) \quad O_{\{p, \neg p\}} \Box_{f,g}(p \text{ produce } q) = \Box_{f,g}(p \text{ produce } q) \land \neg \Box_{f,g}(\neg p)(\neg p \text{ produce } q)$$

$$(12) \quad \text{Semantics of because (full).}$$

$$[\text{because}] = \lambda p \langle s, t \rangle \lambda q \langle s, t \rangle \cdot p \land O_{\{p, \neg p\}} \Box_{f,g}(p \text{ produce } q).$$

### 4 Sufficiency

(13)   a. The robot taking First Street caused it to take Road B.
   b. The robot took Road B because it took First Street.

(13) are intuitively false in this context.  

**Sufficiency inferences**

(14)   a. Serena Williams won the tournament because she won the semi-final.
   b. Serena Williams winning the semi-final caused her to win the tournament.

---

1For more on how polar alternatives are generated by interaction with focus, see e.g. Biezma and Rawlins 2012, Kamali and Krifka 2020.

2This observation is evidence against analyses of causation proposing that counterfactual dependence is enough for the corresponding causal claim to be true, such as Lewis (1973), Pearl (2000), Hitchcock (2001), Woodward (2003), Hall (2004, 2007), Halpern and Pearl (2005), Weslake (2015), Halpern (2016), Beckers (2016) and Nadathur and Lauer (2020).
(15) a. Casey got the job because they applied for it.
    b. Casey applying for the job caused them to get it.

5 Difference-making

An engineer is standing by a switch in the railroad tracks. A train approaches in the distance. She flips the switch, so that the train travels down the right-hand track, instead of the left. Since the tracks reconverge up ahead, the train arrives at its destination all the same.

(Hall 2000, p. 205)

(16) a. The train reached the station because the engineer flipped the switch.
b. The engineer flipping the switch caused the train to reach the station.

5.1 Overdetermination

Suzy and Billy, expert rock-throwers, are engaged in a competition to see who can shatter a target bottle first. They both pick up rocks and throw them at the bottle, but Suzy throws hers before Billy. Consequently Suzy’s rock gets there first, shattering the bottle. Since both throws are perfectly accurate, Billy’s would have shattered the bottle if Suzy’s had not occurred, so the shattering is overdetermined. (Hall 2004, 235)

(17) a. The bottle broke because Suzy threw her rock at it.
   b. Suzy throwing her rock at the bottle caused it to break.

(18) a. The bottle broke because Billy threw his rock at it.
   b. Billy throwing his rock at the bottle caused it to break.

Intuitively, the sentences in (17) are true and the sentences in (18) are false.

• Simplified semantics
  – Suzy ⃝
    + Suzy’s throw guaranteed the bottle would break ✓
    – If Suzy hadn’t thrown, the bottle might not have broken ⃝
  – Billy ⃝
    + Suzy’s throw guarantees the bottle will break ✓
    – If Suzy hadn’t thrown, the bottle might not have broken ⃝

• Full semantics
  – Suzy ✓
    + Suzy’s throw guaranteed that her throw would produce the bottle to break ✓
    – If Suzy hadn’t thrown, her throw might not have produced the bottle to break ✓
  – Billy ⃝
    + Billy’s throw guaranteed that his throw would produce the bottle to break ⃝
    – If Billy hadn’t thrown, his throw might not have produced the bottle to break ✓
6 Property 2: Asymmetry in strength between the positive and negative conditions

(19)  
| a. He has an American passport because he was born in Boston.  
| b. Reyna was born at Royal Bolton Hospital but received a Danish passport because her mother was born in Copenhagen.  
| c. Naama Issachar ... could spend up to seven-and-a-half years in a Russian prison because 9.5 grams of cannabis were found in her possession during a routine security check.  
| d. A 90-day study in 8 adults found that supplementing a standard diet with 1.3 cups (100 grams) of fresh coconut daily caused significant weight loss.

(McHugh 2020, example 6)

Intuitively, (61d) does not say that 100 grams is the minimal amount required to cause significant weight loss.

(20)  
| a. Computers do an awful lot of deliberation, and yet their every decision is wholly caused by the state of the universe plus the laws of nature.  
| b. If anything is happening at this moment in time, it is completely dependent on, or caused by, the state of the universe, as the most complete description, at the previous moment.  
| c. If you keep asking “why” questions about what happens in the universe, you ultimately reach the answer “because of the state of the universe and the laws of nature.”

(McHugh 2020, example 7)

(21)  
Reyna was born at Royal Bolton Hospital but received a Danish passport because her mother was born in Copenhagen.

- If Reyna’s mother hadn’t been born in Copenhagen, she could have been born outside Denmark, or somewhere else in Denmark.

---

6https://rupaulsdragrace.fandom.com/wiki/Charlie_Hides  
8https://www.healthline.com/nutrition/coconut-meat  
9http://commonsenseatheism.com/?p=899  
11https://www.edge.org/response-detail/10164
• If she had been born somewhere else in Denmark, Reyna would still have received a Danish passport.

• On the current analysis, (63) implies that Reyna’s mother being born in Copenhagen produced Reyna to have a Danish passport.

• Using strong dependence, this implies there is a chain of strong dependence beginning with Reyna’s mother being born in Copenhagen.

(22) a. $\Box_{f,g}(\neg \text{born in Copenhagen})(\neg \text{Danish passport}).$ Strong: X
b. $\neg \Box_{f,g}(\neg \text{born in Copenhagen})(\text{Danish passport}).$ Weak: ✓
(23) a. $\Box_{f,g}(\neg \text{Copenhagen})(\neg (\neg \text{Copenhagen}) \text{ produce Danish}).$ Strong: X
b. $\neg \Box_{f,g}(\neg \text{Copenhagen})(\neg (\neg \text{Copenhagen}) \text{ produce Danish}).$ Weak: ✓

6.1 Further evidence for weak negative condition: implicatures

(24) A: Does John have a French passport because he was born in France?
B: He has a French passport because he was born in [Paris]$_F$.

(25) A: Did John get invited to the summer school because he passed syntax?
B: He got invited to the summer school because he passed syntax [and semantics]$_F$.

(26) Upward entailing environments: scope of every; conditional consequents
a. (i) A: Was every student born in France?
   (ii) B: Every student was born in [Paris]$_F$.
      $\not\rightarrow \neg$(Every student was born to France).

b. (i) A: If John works hard will he pass syntax?
   (ii) B: If he works hard he will pass syntax [and semantics]$_F$.
      $\not\rightarrow \neg$(If John works hard he will pass syntax).

(27) Non-monotonic environments: between $n$ and $m$; exactly $n$

a. (i) A: Were between 10 and 20 students born in France?
   (ii) B: Between 10 and 20 students were born in [Paris]$_F$.
      $\not\rightarrow \neg$(Between 10 and 20 students were born in France)

b. (i) A: Did exactly 10 students enrol in syntax?
   (ii) B: Exactly 10 students enrolled in syntax [and semantics]$_F$.
      $\not\rightarrow \neg$(Exactly 10 students enrolled in syntax)

The acceptability of the following dialogues helps illustrate that B’s answers in (27) do not trigger the stated implicatures.
A: Did between 10 and 20 students go to France?
B: Between 10 and 20 students went to [Paris]$_F$.
A: Ok, but how many went anywhere in France?
B: In total, 18 students went to France.

A: Did exactly 10 students enrol in syntax?
B: Exactly 10 students enrolled in syntax [and semantics]$_F$.
A: Ok, but how many enrolled in syntax?
B: 10 students enrolled in syntax. The students have to take syntax and semantics at the same time.

- We can use the implicatures triggered by because’s cause argument to empirically distinguish the strong and weak negative conditions.

\[(30)\]

\begin{align*}
\text{a. } & q \text{ because}_{\text{weak}} p = p \land \Box f,g (p) (q) \land \neg \Box f,g (\neg p)(q). \\
\text{b. } & q \text{ because}_{\text{strong}} p = p \land \Box f,g (p) (q) \land \Box f,g (\neg p)(\neg q).
\end{align*}

- The weak and strong negative conditions predict different the monotonicity properties for because.

- The weak condition predicts that because is downward entailing (DE) in its cause argument.

- The strong condition predicts that because is non-monotone in its cause argument

\[(31)\]

\begin{align*}
\text{a. } & (p^+ \land (q \text{ because}_{\text{weak}} p)) \Rightarrow q \text{ because}_{\text{weak}} p^+ \\
\text{b. } & (p^+ \land (q \text{ because}_{\text{strong}} p)) \not\Rightarrow q \text{ because}_{\text{strong}} p^+
\end{align*}

- We assume that when it comes to implicature calculation, universal modals act as if they are DE in their restrictor (see von Fintel 1999, 2001)

A: Would John have gotten a French passport if he had been born in France?
B: John would have gotten a French passport if he had been born in [Paris]$_F$.
\[
\Downarrow \leadsto \neg(\text{John would have gotten a French passport if he had been born in France}).
\]

A: Would John have gotten invited to the summer school if he had passed syntax?
B: John would have been invited to the summer school if he had passed syntax [and semantics].

\[ \neg (\text{John would have been invited to the summer school if he had passed syntax}) \]

\begin{align*}
\square_f(p)(q) & \Rightarrow \square_f(p^+)(q) \\
\neg \square_f(p^+)(q) & \Rightarrow \neg \square_f(p)(q) \\
\neg \square_f(\neg p)(q) & \Rightarrow \neg \square_f(\neg p^+)(q) \\
\square_f(\neg p^+)(\neg q) & \Rightarrow \square_f(\neg p)(\neg q) \\
\square_f(\neg p)(\neg q) & \Rightarrow \square_f(\neg p)(\neg q)
\end{align*}

(Contraposition)

\begin{align*}
(p^+ & \Rightarrow p, \text{ so } \neg p \Rightarrow \neg p^+) \\
(p^+ & \Rightarrow p, \text{ so } \neg p \Rightarrow \neg p^+)
\end{align*}

Positive condition    Weak negative condition

\begin{align*}
p^+ & \quad \square_f(p)(q) \\
\downarrow & \quad \downarrow \\
p^+ & \quad \square_f(p^+)(q) \\
& \quad \neg \square_f(\neg p^+)(q)
\end{align*}

Strong negative condition

\begin{align*}
p^+ & \quad \square_f(p)(q) \\
\downarrow & \quad \downarrow \\
p^+ & \quad \square_f(p^+)(q) \\
& \quad \square_f(\neg p^+)(\neg q)
\end{align*}

Figure 4: Entailment relations between stronger and weaker causes on the simplified semantics.

7 The circumstances

(34) a. Given how things were on Monday, the robot took Road B because it is programmed to prefer tree-lined roads.

b. Given how things were on Wednesday, the robot took Road B because it is programmed to prefer tree-lined roads.
a. Given how things were on Monday, the robot’s preference for tree-lined roads caused it to take Road B.

b. Given how things were on Wednesday, the robot’s preference for tree-lined roads caused it to take Road B.

There is a contrast between the (a) sentences and the (b) sentences. Intuitively, the (a) sentences are false and the (b) sentences are true.

7.1 The circumstances as modal base

(36) He might have won the game.  

(37) a. He might have (already) won the game (# but he didn’t).

    Epistemic reading: present perspective, past orientation

b. At that point he might (still) have won the game but he didn’t in the end.

    Counterfactual reading: past perspective, future orientation

(38) a. The robot took Road B because it is programmed to prefer tree-lined roads.

b. The robot’s preference for tree-lined roads caused it to take Road B.

Both readings are available

(39) a. The robot took Road B because it is programmed to prefer tree-lined roads. For, its programming made it take Road B rather than Road A.

b. The robot didn’t take Road B because it is programmed to prefer tree-lined roads. For It could have turned right on Tuesday, in which case it would have taken Road C or D, not road B.
7.2 Property 3: the positive and negative conditions have the same modal base

Consider (38) in this context.

(38) a. The robot took Road B because it is programmed to prefer tree-lined roads.
   b. The robot’s preference for tree-lined roads caused it to take Road B.

Intuitively, (38) are false in this context.

These results are summarized in Table 1.

<table>
<thead>
<tr>
<th>Temporal perspective</th>
<th>Positive condition</th>
<th>Negative condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Wednesday</td>
<td>✓</td>
<td>✗</td>
</tr>
</tbody>
</table>

Table 1
7.3 Generating each modal independently?

(40) a. \( p \land \Box_{f,g}(p)(q) \land \neg \Box_{f,g'}(\neg p)(q) \)
   
   b. \( p \land \Box_{f,g}(p(p \text{ produce } q)) \land \neg \Box_{f,g'}(\neg p)(\neg \text{produce } q) \)

We could, of course, add a constraint that the modals’ parameters must be identical as a stipulation.

(41) a. \( p \land \Box_{f,g}(p)(q) \land \neg \Box_{f,g'}(\neg p)(q) \land f = f' \land g = g' \)
   
   b. \( p \land \Box_{f,g}(p(p \text{ produce } q)) \land \neg \Box_{f,g'}(\neg p)(\neg \text{produce } q) \land f = f' \land g = g' \)

An alternative proposal:

- the modals’ parameters in the positive and negative conditions are determined by conversational context, which is of course the same for both the positive and negative conditions.

- A modal’s parameters can shift within the same conversation—and even within the same sentence. Consider, for example, (42).

(42) I want to go to the gym and I don’t want to go to the gym.

8 Economy

The previous sections provided evidence that the semantics of cause and because satisfies properties 1, 2 and 3. These three properties all point to the presence of an exhaustification operator in the lexical semantics of cause and because. Now, one may wonder about the status of this operator. It is always present in the semantics, or subject to licensing conditions?

To answer this question, a key test case is how cause and because behave under negation. It is commonly assumed that exhaustification is subject to an economy condition that prevents it from appearing when it produces an overall weaker meaning (Chierchia 2013, Fox and Spector 2018). If the exhaustification operator in the semantics of cause and because is subject to this constraint, we would expect the following parses to be ruled out by Exhaustification Economy.

(43) a. \( \neg (p \land O_{(p,\neg p)} \Box_{f,g}(p)(q)) \).
   
   b. \( \neg (p \land O_{(p,\neg p)} \Box_{f,g}(p(p \text{ produce } q))) \).

These are equivalent, respectively, to the following.

(44) a. \( \neg p \lor \neg \Box_{f,g}(p)(q) \lor \Box_{f,g}(\neg p)(q) \)
b. \( \neg p \lor \neg \Box f_g(p \text{ produce } q) \lor \Box f_g(\neg p \text{ produce } q) \)

These meanings are weaker than those without exhaustification.

(45) a. \( \neg (p \land \Box f_g(p)(q)) \Leftrightarrow \neg p \lor \neg \Box f_g(p)(q) \).

b. \( \neg (p \land \Box f_g(p \text{ produce } q)) \Leftrightarrow \neg p \lor \neg \Box f_g(p \text{ produce } q) \).

However, it turns out that the only parse of not ... because and not ... cause that correctly accounts for the data is one that violates Exhaustification Economy, as we see now.

8.1 *Because and economy: data*

In de Saint-Exupéry’s *The Little Prince*, the protagonist visits a king who claims to be able to command the sun to set. Suppose the king commands the sun to set, and sure enough, some time later it sets. Unfortunately for the king’s ego, the following sentences are false.

(46) a. The sun set because the king commanded it.

b. The king’s command caused the sun to set.

The simplified and full semantics explain the falsity of (46) in different ways. On the simplified semantics (46) are false because the sun would have set even if the King hadn’t commanded it.

(47) a. \( \Box f_g(\text{King command})(\text{sun set}) \)

b. \( \Box f_g(\neg\text{King command})(\text{sun set}) \)

While on the full semantics (46) are false because the King’s command did not produce the sun to set.

(48) a. \( \neg \Box f_g(\text{King command})(\text{King command produce sun set}) \)

b. \( \neg \Box f_g(\neg\text{King command})(\neg\text{King command produce sun set}) \)

According to the simplified semantics, (46) are false for the same reason that (16), repeated below, are false in the train track scenario: the train would have reached the station anyway.

(16) a. The train reached the station because the engineer flipped the switch.

b. The engineer flipping the switch caused the train to reach the station.

(49) a. \( \Box f_g(\text{pull lever})(\text{train reach station}) \)

b. \( \Box f_g(\neg\text{pull lever})(\text{train reach reach}) \)
While according to the full semantics, (16) are false for a different reason. Pulling the lever produced the train to reach the station (because there is a chain of events beginning with the engineer pulling the lever, through the train taking the side track, to the train reaching the station). But symmetrically, not pulling the lever would have also produced the train to reach the station, so the full semantics predicts (16) to be false.

(50) a. \( \Box_{f,g}(\text{pull lever})(\text{pull lever produce train reach station}) \)

b. \( \Box_{f,g}(\neg\text{pull lever})(\text{pull lever produce train reach}) \)

8.2 *Because and economy: analysis*

For the simplified semantics, these data are compatible with two parses. The first, \( \neg(p \land O_{\{p,-p\}} \Box_{f,g}(p)(q)) \), violates Economy. The second, which Fox and Spector (2018, ex. 70) discuss, features a higher exhaustification operator whose alternative is the prejacent without exhaustification. This parse does not violate Economy.

(51) \( O_{\text{ALT}}[\neg O_{\{p,-p\}}[\Box(p)(q)]] \) where \( \text{ALT} = \{\neg O_{\{p,-p\}}[\Box(p)(q)], \neg \Box(p)(q)\} \).

Table 2 shows that both parses of the simplified semantics are compatible with the data in (16) and (46).

<table>
<thead>
<tr>
<th>Parse of not ... because</th>
<th>Simplified meaning of parse</th>
<th>(46)</th>
<th>(16)</th>
<th>Economy</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \neg \Box(p)(q) )</td>
<td>( \neg \Box(p)(q) )</td>
<td>F X</td>
<td>T ✓</td>
<td>✓</td>
</tr>
<tr>
<td>( O_{{p,-p}}[\neg \Box(p)(q)] )</td>
<td>( \neg \Box(p)(q) \land \Box(\neg p)(q) )</td>
<td>F X</td>
<td>F X</td>
<td>✓</td>
</tr>
<tr>
<td>( \neg O_{{p,-p}}[\Box(p)(q)] )</td>
<td>( \neg \Box(p)(q) \lor \Box(\neg p)(q) )</td>
<td>T ✓</td>
<td>T ✓</td>
<td>X</td>
</tr>
<tr>
<td>( O_{\text{ALT}}[\neg O_{{p,-p}}[\Box(p)(q)]] )</td>
<td>( \Box(p)(q) \land \Box(\neg p)(q) )</td>
<td>T ✓</td>
<td>T ✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 2: Possible parses of not ... because on the simplified semantics.

However, as Table 3 shows, only one parse of the full semantics is compatible with the data above. This is also the only parse that violates Economy. In the Table we use \( p \leftrightarrow q \) as shorthand for \( p \text{ produce } q \). Similar to (51) above, \( \text{ALT} = \{\neg O_{\{p,-p\}}[\Box(p)(p \leftrightarrow q)], \neg \Box(p)(p \leftrightarrow q)\} \).

9 *Production*

(52) **The idea.** \( p \text{ produces } q \) iff there is a chain of dependence from \( p \) to \( q \).

Questions:
Parse of *not ... because* | Full meaning of parse | (46) | (16) | Economy
--- | --- | --- | --- | ---
\(\neg \Box (p) (p \leftrightarrow q)\) | \(\neg \Box (p) (p \leftrightarrow q)\) | T ✓ | F ✗ | ✓
\(O_{(p, \neg p)} [\neg \Box (p) (p \leftrightarrow q)]\) | \(\neg \Box (p) (p \leftrightarrow q) \land \Box (\neg p) (p \leftrightarrow q)\) | F X | F X | ✓
\(\neg O_{(p, \neg p)} [\Box (p) (p \leftrightarrow q)]\) | \(\neg \Box (p) (p \leftrightarrow q) \lor \Box (\neg p) (p \leftrightarrow q)\) | T ✓ | T ✓ | X
\(O_{\text{ALT}} [\neg O_{(p, \neg p)} [\Box (p) (p \leftrightarrow q)]]\) | \(\Box (p) (p \leftrightarrow q) \land \Box (\neg p) (p \leftrightarrow q)\) | F X | T ✓ | ✓

Table 3: Possible parses of *not ... because* on the full semantics.

1. What is the chain made up of?
2. What does “dependence” mean?
3. What do “from \(p\)” and “to \(q\)” mean?

### 9.1 The chain

“\(e\) depends causally on \(c\) iff \(c\) occurs, \(e\) occurs, and if \(c\) had not occurred, then \(e\) would not have occurred at all, or would have occurred later than the time that it actually did occur”.

(Paul 1998, 193)

“replace the words ‘or would have occurred later than the time that it actually did occur’ by the words ‘or would have occurred at a time different from the time that it actually did occur’.”

(Lewis 2000, 187)

(53) a. Socrates died because he drank poison.
   b. Socrates drinking poison caused him to die.

(54) a. The bottle broke because Suzy threw a rock at it.
   b. Suzy throwing her rock caused the bottle to break.

**Definition 1** (Temporally-indexed proposition). *Where \(p\) is a proposition and \(t\) is a point in time, let us call \((p, t)\) (also denoted \(p_t\)) a temporally-indexed proposition and define that \(p_t\) is true at a world \(w\) just in case \(p\) is true at \(w\) at time \(t\).*

**Why temporally-indexed propositions?** Evidence from overdetermination.

(55) a. Socrates died because he drank poison.
   b. Socrates drinking poison caused him to die.

(56) a. The bottle broke because Suzy threw a rock at it.
b. Suzy throwing a rock at the bottle caused it to break.

Compositionality suggests that *Socrates died* has the same meaning under *because* has in other environments (similarly for *the bottle broke*). For example,

\[(57) \quad \lbrack \textit{Socrates died} \rbrack = \exists e (dying(e) \land \text{agent}(e) = \text{Socrates}).\]

We can even have coreference between an event that is asserted to be inevitable and one that is asserted to have a cause.

\[(58) \quad \text{a. Socrates drinking poison caused an event, his death, which was bound to happen eventually.}\]

\[\text{b. Socrates’ death was inevitable. According to the \textit{Phaedo}, it was caused by him drinking poison.}\]

Idea (Paul 1998, Lewis 2000, and many more): production involves dependence, not in whether the proposition *Socrates died* is true or false in the actual world, but in something about how the proposition is true.

9.2 What is a chain?

**Definition 2** (Chain). *A chain is a set \{p_t : p \in P, t \in T\} where \(P\) is a set of propositions and \(T\) is a set of time intervals\(^{12}\) such that the intervals in \(T\) do not overlap (i.e. \(t\) and \(t'\) are disjoint for all \(t, t' \in T\)).*

![Figure 8: A chain](image)

**Definition 3.** *A chain \{p_t : p \in P, t \in T\} is dense just in case \(\bigcup T\) is an interval.*

![Figure 9: A dense chain](image)

\(^{12}\)An interval \(I\) is a set of points in time such that for any points in time \(x, y, z\), if \(x, y \in I\) and \(x \leq z \leq y\) then \(z \in I\).
9.3 Chains of dependence

**Definition 4.** A chain \( C = \{p_t : p \in P, t \in T\} \) is a chain of strong dependence just in case for all \( p_t, q_{t'} \in C \), if \( t \leq t' \) then \( q_{t'} \) counterfactually depends on \( p_t \), i.e. \( \Box (\neg p_t) (\neg q_{t'}) \).

**Why does the chain need to be dense?** Evidence from overdetermination.

(55) a. Socrates died because he drank poison.
   b. Socrates drinking poison caused him to die.

(56) a. The bottle broke because Suzy threw a rock at it.
   b. Suzy throwing a rock at the bottle caused it to break.

On the proposed analysis, these require:

(59) a. Not: if Socrates had not drunk poison, him not drinking poison would have produced his death.
   b. Not: if Suzy had not thrown her rock, her not throwing her rock would not have produced the bottle to break.

Socrates drank poison at time \( t \) and died at time \( t' \). If he had not drunk poison, he would have died at time at some other time, \( t^* \). Without density, we could build the chain out of just the cause and effect. The problem is this: Socrates dying at time \( t^* \) counterfactually depends on him not drinking poison at time \( t \). For if he had drunk poison at time \( t \), he would not have died at time \( t^* \). So without density, we wrongly predict (59a) to be false: if Socrates had not drunk poison, him not drinking poison would have produced his death. Thus we wrongly predict (55) to be false.

Similarly, without density we predict: if Suzy had not thrown her rock at the bottle, then her not throwing her rock would have produced the bottle to break. Consider the non-dense chain: Suzy doesn’t throw her rock at time \( t \), the bottle breaks at time \( t' \). If Suzy had thrown her rock at \( t \), the bottle would not have broken at time \( t' \). So without density we wrongly predict (59b) to be false, and so wrongly predict (56) to be false.

With density, we need the chain to be sustained throughout time. There is no dense chain from Socrates not drinking poison to his death. And there is no dense chain from Suzy not throwing her rock to the bottle breaking.

9.4 Chain widening

**Why do we need dependence to be strong?**
(60)  
a. Strong dependence: $\Box(\neg p)(\neg q)$
b. Weak dependence: $\neg\Box(\neg p)(q)$

Evidence from chain widening.

Figure 10: How production is supposed to work.

Figure 11

Production under weak counterfactual dependence.

- If the cause had not occurred at $t$, the world would have been in a different state at $t'$ (where $t < t'$).
- If the world had been in a different state at $t'$, there are countless ways in which it could have been different.
- Weak dependence requires that in some of these cases, the next eventuality on the chain does not occur.
If this domino had not fallen, the dominoes would be in a different state.

Figure 12: Chain widening.

If the dominoes were in a different state, there are many other states they could have been in.

Figure 13: A problem with defining production using weak counterfactual dependence.
• Strong dependence requires that in all of these cases, the next eventuality on the chain does not occur.

• For any proposition $p$ whatsoever that is true at $t'$, if the world had been different at $t'$, the difference could be that $p$ is false.

• Weak dependence allows chain widening to trivialise production, while strong dependence does not.

9.5 ‘From’ the cause, ‘to’ the effect

(61) a. He has an American passport because he was born in Boston.\textsuperscript{13}

b. Reyna was born at Royal Bolton Hospital but received a Danish passport because her mother was born in Copenhagen.

c. Naama Issachar ... could spend up to seven-and-a-half years in a Russian prison because 9.5 grams of cannabis were found in her possession during a routine security check.\textsuperscript{14}

d. A 90-day study in 8 adults found that supplementing a standard diet with 1.3 cups (100 grams) of fresh coconut daily caused significant weight loss.\textsuperscript{15}

(McHugh 2020, example 6)

Intuitively, (61d) does not say that 100 grams is the minimal amount required to cause significant weight loss.

(62) a. Computers do an awful lot of deliberation, and yet their every decision is wholly caused by the state of the universe plus the laws of nature.\textsuperscript{16}

b. If anything is happening at this moment in time, it is completely dependent on, or caused by, the state of the universe, as the most complete description, at the previous moment.\textsuperscript{17}

c. If you keep asking “why” questions about what happens in the universe, you ultimately reach the answer “because of the state of the

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{13} https://rupaulsdragrace.fandom.com/wiki/Charlie_Hides
\item \textsuperscript{15} https://www.healthline.com/nutrition/coconut-meat
\item \textsuperscript{16} http://commonsenseatheism.com/?p=899
\end{itemize}
\end{footnotesize}
universe and the laws of nature.”

(McHugh 2020, example 7)

(63) Reyna was born at Royal Bolton Hospital but received a Danish passport because her mother was born in Copenhagen.

• If Reyna’s mother hadn’t been born in Copenhagen, she could have been born outside Denmark, or somewhere else in Denmark.

• If she had been born somewhere else in Denmark, Reyna would still have received a Danish passport.

• On the current analysis, (63) implies that Reyna’s mother being born in Copenhagen produced Reyna to have a Danish passport.

• Using strong counterfactual dependence, this implies there is a chain of strong counterfactual dependence beginning with Reyna’s mother being born in Copenhagen.

(52) The idea. \( p \) produces \( q \) iff there is a chain of dependence from \( p \) to \( q \).

What do “from \( p \)” and “to \( q \)” mean? Two ideas: exact and inexact.\(^\text{19}\)

(64) Let \( C \) be a chain.

a. \( C \) is exactly from \( p \) iff \( p \) is the first proposition on \( C \).

b. \( C \) is inexactly from \( p \) iff \( p \) entails the first proposition on \( C \).

c. \( C \) is exactly to \( q \) iff \( q \) is the last proposition on \( C \).

d. \( C \) is inexactly to \( q \) iff \( q \) entails the last proposition on \( C \).

Problem: strong dependence is too strong.

Proposed solution: production is interpreted inexactly.

(65) Initial idea (to be revised). \( p \) produce \( q \) just in case there is a dense chain of strong dependence inexactly from \( p \) to \( q \).

\(^{\text{18}}\)https://www.edge.org/response-detail/10164

\(^{\text{19}}\)These conditions are formalised as follows.

(i) Let \( C = \{ r_t : r \in P, t \in T \} \) be a chain.

a. \( C \) is exactly from \( p \) just in case \( p_t \in C \) for some \( t \in T \) such that \( t \leq t' \) for all \( t' \in T \).

b. \( C \) is inexactly from \( p \) just in case \( p \) entails \( p' \) for some \( p'_t \in C \) such that \( t \leq t' \) for all \( t' \in T \).

c. \( C \) is exactly to \( q \) just in case \( q_t \in C \) for some \( t \in T \) such that \( t' \leq t \) for all \( t' \in T \).

d. \( C \) is inexactly to \( q \) just in case \( q \) entails \( q' \) for some \( q'_t \in C \) such that \( t' \leq t \) for all \( t' \in T \).
• Reyna’s mother being born in Copenhagen produced Reyna to have a Danish passport because
  – being born in Copenhagen entails being born in **Denmark**, and
  – there is a chain of strong dependence from Reyna’s mother being born in **Denmark** to Reyna receiving a Danish passport.

So interpreting production inexactly allows producers to be stronger than strictly required.

For the effect, it turns out that *exactly to* \( q \) is equivalent to *inexactly to* \( q \). To see this, note that strong dependence is DE in its scope.

(66) If \( q^+ \) entails \( q \) then \( \Box(\neg p)(\neg q) \) entails \( \Box(\neg p)(\neg q^+) \).

This is because universal quantifiers are UE in their scope and negation is DE. Together they result in a DE environment.

\[
q^+ \Rightarrow q \\
\neg q \Rightarrow \neg q^+ \\
\Box(\neg p)(\neg q) \Rightarrow \Box(\neg p)(\neg q^+)
\]

It follows that “exactly to” and “inexactly to” are equivalent, as we show now.

Suppose there is chain of strong dependence *inexactly to* \( q \). Then \( q \) entails \( q' \), where \( q' \) is the last proposition on the chain, and \( q' \) strongly depends on every event on the chain. Since \( q \) entails \( q' \) and strong dependence is DE in its effect, \( q \) also strongly depends on every event on the chain. So there is a chain of strong dependence *exactly to* \( q \). So “inexactly to” entails “exactly to”.

Conversely, if there is a chain of strong dependence *exactly to* \( q \), then as \( q \) entails itself, there is a chain of strong dependence *inexactly to* \( q \). So “exactly to” entails “inexactly to”.

On the other hand, strong dependence is UE in its restrictor.

(67) If \( p^+ \) entails \( p \) then \( \Box(\neg p^+)(\neg q) \) entails \( \Box(\neg p)(\neg q) \).

This is because restrictors of universal quantifiers are DE and negation is DE. These together create a UE environment.

\[
p^+ \Rightarrow p \\
\neg p \Rightarrow \neg p^+ \\
\Box(\neg p^+)(\neg q) \Rightarrow \Box(\neg p)(\neg q)
\]
As we saw, the exact and inexact interpretations of “to p” are not equivalent. The exact interpretation wrongly predicts (63) to be false, while the inexact interpretation predicts (63) to be true.

9.6 Production and monotonicity

The inexact interpretation of “to p” makes production DE in its producer argument.

- p produce q just in case there is a dense chain of strong dependence inexactly from p to q
- i.e there is a dense chain of strong dependence such that p entails the first proposition on the chain and q entails the last proposition on the chain.
- Then for any p⁺ that entails p, if p⁺ is true then there is a dense chain of strong dependence such that p⁺ entails the first proposition on the chain and q entails the last proposition on the chain.

9.7 A problem: trivial producers

Example. Alice flicks a switch, causing the light to turn on. At the same time she flicks the switch, a leaf falls from a tree.

- Alice flicking the switch produced the light to turn on.
- (Alice flicking the switch or the leaf falling) produced the light to turn on.
  - If Alice hadn’t flicked the switch, there wouldn’t have been electricity in the wire, ...
  - ⇒ If Alice hadn’t flicked the switch and the leaf hadn’t fallen, there wouldn’t have been electricity in the wire, ...
- “The leaf fell” entails “Alice flicked the switch or the leaf fell”.

- p produces q
- ⇒ (p ∨ r) produces q
- r entails p ∨ r
- Inexact interpretation of “to” implies: r produces q, for any r whatsoever

Problem: interpreting “from” as “inexactly from” in the definition of implies that the leaf falling produced the light to turn on.

Toward a solution. Notice that if the leaf hadn’t fallen from the tree, we could have made exactly the same argument:
• $p$ produces $q$
• $\Rightarrow (p \lor \neg r)$ produces $q$
• $\neg r$ entails $p \lor \neg r$
• Inexact interpretation of “to” implies: $\neg r$ produces $q$, for any $r$ whatsoever

The propositions expressed by the leaf fell and the leaf did not fall are symmetric with respect to producing the light to turn on.

However, this does not hold for born in Copenhagen and born in Denmark.

(68)  
   a. $r$ entails $p \lor r$, and $\neg r$ entails $p \lor \neg r$
   b. Copenhagen entails Denmark, but $\neg$Copenhagen does not entail Denmark

Proposed solution: exhaustification with respect to the cause’s polar alternative.

(69)  
Definition of production (final). $p$ produces $q$ just in case
   a. If $p$ had occurred,  
      there is a chain of strong dependence inexactly from $p$ to $q$, and  
   b. If $p$ had not occurred,  
      there is no chain of strong dependence inexactly from $\neg p$ to $q$.

Remarkably, this fix keeps production DE in its producer argument.\textsuperscript{20}

(70)  
If $p^+$ entails $p$ then $p^+ \land (p$ produce $q)$ entails $p^+$ produce $q$.

References


\textsuperscript{20}This is due to the following entailments.

- There is a chain of strong dependence inexactly from $p$ to $q$
  $\Rightarrow$ There is a chain of strong dependence inexactly from $p^+$ to $q$

- There is a chain of strong dependence inexactly from $\neg p^+$ to $q$
  $\Rightarrow$ There is a chain of strong dependence inexactly from $\neg p$ to $q$

- There is no chain of strong dependence inexactly from $\neg p$ to $q$
  $\Rightarrow$ There is no chain of strong dependence inexactly from $\neg p^+$ to $q$


