Causal and Evidential Conditionals

Mario Günther[1,2] [ID]

Abstract
We put forth an account for when to believe causal and evidential conditionals. The basic idea is to embed a causal model in an agent's belief state. For the evaluation of conditionals seems to be relative to beliefs about both particular facts and causal relations. Unlike other attempts using causal models, we show that ours can account rather well not only for various causal but also evidential conditionals.

Keywords Conditionals · Causal Models · Evidential vs. Causal

1 A Problem for Causal Model Semantics

When should an intelligent agent believe a conditional? To answer this and related questions, Pearl (2000, Chap. 7.1) proposes a semantics for conditionals based on causal models.[1] Since then many semantics for causal conditionals have been proposed, for instance by Hitchcock (2001), Woodward (2003), Hiddleston (2005), Schulz (2011), Halpern (2016), Ciardelli et al. (2018), and Santorio (2019). While these different causal model semantics account well for our intuitions about causal conditionals, they do not account for what we call evidential conditionals at all.

What is an evidential conditional? Kratzer(1989, p. 640) provides an example similar to the following:

Example 1 King Ludwig of Bavaria likes to spend his weekends at Leoni Castle. Whenever the King is in the Castle, the Royal Bavarian flag is up and the lights are on. You see that, at the moment, the lights are on, but the flag is down. You say:

---

[1] Pearl does not answer the question how an intelligent agent learns conditionals, but see Günther (2017, 2018).

✉ Mario Günther
Mario.Guenther@lmu.de

1 LMU Munich, Munich, Germany

2 Australian National University, Canberra, Australia

(1)   If the flag had been up, the King would have been in the castle.

    Kratzer says that this conditional is acceptable. To see why the conditional is evidential, let us distinguish two readings of (1).

(2)   If the flag had been up, this *would have caused that* the King is in the castle.
(3)   If the flag had been up, this *would have been evidence that* the King is in the castle.

(3) but not (2) seems to be an appropriate paraphrase of (1). In general, the antecedent of an evidential conditional provides, or would provide, evidence for the consequent. Now, the above mentioned causal model semantics capture the unacceptability of (2), but fail to account for (3) and thus (1). This is simply because the flag being up has no causal influence on whether or not the King is in the castle.[2] The flag merely constitutes evidence for the presence of the King.

    Here we propose an account for when an agent should believe causal and evidential conditionals. We assume that an agent has beliefs about both particular facts and causal relations. We embed a causal model in an agent's belief state to represent beliefs about causal relations. Then we illustrate how beliefs about causal relations and particular facts influence the evaluation of conditionals by walking through several examples.

2 An Epistemic Account

We represent an agent's belief state by $\langle M, B \rangle$, where M is a causal model $\langle V, E \rangle$ and $B \subseteq W$ a set of value assignments to the variables in V. W denotes the set of all relevant possible worlds. The agent believes what is true in all value assignments, or equivalently worlds, $w \in B$ that are most plausible. The plausibility of worlds is lexicographically determined: (i) worlds that assign to more variables the values about which the agent is certain are more plausible than worlds which assign those values to fewer variables; and whenever there are ties in (i), (ii) the worlds that satisfy more of the structural equations in E are more plausible than worlds which satisfy fewer.[3] An agent's certainty about particular facts trumps the believed causal relations.

---

[2] Note that most of the cited causal model semantics are meant for counterfactual conditionals only. They should thus solve the Kratzer-style example. As we will see later, our causal/evidential distinction cross-cuts the subjunctive/indicative distinction, and likewise the more traditional ontic/epistemic distinction.

[3] The plausibility order induced by (i) and (ii) is a strict partial order on the set of worlds. The "more" in both clauses is to be understood in a numerical sense: 2 variable assignments about which the agent is certain is more than 1 such variable assignment; likewise, a world that satisfies 3 of the structural equations satisfies more than a world that satisfies only 2. Since the plausibility order is determined by the particular facts that are believed for certain and the set of structural equations, we may leave it implicit in what follows.

A causal model is a tuple $\langle V, E \rangle$, where V is a set of variables and E a set of structural equations. The set E of structural equations represents the causal relations the agent believes to be true. For simplicity, we restrict ourselves to binary variables representing that a fact obtains or else does not. The variables thus come with a range of two values, we call true 1 and false 0. For instance, let K be the variable that represents whether or not Kennedy has been shot. K taking the value 1, $K = 1$, says that Kennedy has been shot, while $K = 0$ says that he has not been shot. Whenever the binary variables behave like propositional variables, we abbreviate $K = 1$ by k and $K = 0$ by $\neg k$.

Each structural equation specifies a relation of direct causal dependence between some of the variables in V. For each variable X in V, there is at most one structural equation of the form:

$$X = f_X(Pa_X), \tag{*}$$

where the set Pa_X of X's parent variables is a subset of $V \setminus \{X\}$. A structural equation says that the value of the variable on the left-hand side is causally determined by the values of the variables on the right-hand side. The value of X is determined by the values of the variables in Pa_X, in the way specified by f_X. In what follows, we will only use a subclass of the functions applicable to binary variables, the Boolean operators $\neg, \wedge, \vee$.

Sides matter for structural equations. In (*), the value of X is determined by $f_X(Pa_X)$, but the value of $f_X(Pa_X)$ is not determined by X. A structural equation expresses how the variable on the left-hand side depends on the variables on the right-hand side. Hence, (*) encodes a set of conditionals that have the following form:

If the variables in Pa_X were to take on these or those values, X would take on this or that value.

The structural equation $K = S \vee O$, for instance, encodes four conditionals:

(i)   if s and o were the case, k would be the case.
(ii)   if s and $\neg o$ were the case, k would be the case.
(iii)   if $\neg s$ and o were the case, k would be the case.
(iv)   if $\neg s$ and $\neg o$ were the case, $\neg k$ would be the case.

Causal models provide a semantics for causal or interventionist conditionals.[4] The causal conditional $a >_c c$ says that c would result if a were set by intervention. An intervention on an arbitrary variable A can be represented as follows: replace the equation f_A by an equation $A = Y$, where Y is the value to which the variable is set. The equations for the other variables remain untouched. The result is the causal model $M_{A=Y}$ obtained from M by setting A to the value Y.

---

[4] Causal models give also rise to a proper logic of causal conditionals. This logic has first been worked out by Galles and Pearl (1998) and Halpern (2000). It has been refined by Briggs (2012) and Zhang (2013).

As soon as you assign values to all the "parentless" variables of a causal model M, the values of the remaining variables can be computed. This solution of M is unique if the structural equations can be ordered such that no variable occurs on the left-hand side of an equation after having occurred as a parent on the right-hand side. In what follows, we consider only such *acyclic* causal models.

The result M_a of intervening on M by a propagates the effects of the intervention causally downstream according to the structural equations. The "parentless" variables include now A. The result can be thought of as the causal possible world, where A takes the value 1, and which is otherwise most similar to the actual world. If the variable C takes the value 1 in M_a, the causal conditional $a >_c c$ is true relative to M and a value assignment to the (parent) variables in V.

An agent believes a conditional $a > c$ to be true iff $a > c$ is true at each $\langle M, w \rangle$ where $w \in B$. In words, the agent believes those conditionals which are true at each most plausible world. It remains to state the definition for when a causal conditional is true at what we will call a causal world $\langle M, w \rangle$:

Definition 1 *Causal Conditionals*
   A causal conditional $a >_c c$ is true at $\langle M, w \rangle$ iff c is true at $\langle M_a, w \rangle$.

As above, M_a represents the causal model obtained from M where the structural equation f_A has been replaced by $A = 1$. Let us turn to evidential conditionals. We define when an evidential conditional is true at a causal world as follows:

Definition 2 *Evidential Conditionals*
   An evidential conditional $a >_e c$ is true at $\langle M, w \rangle$ iff c is true at all $\langle M, w' \rangle$, where w' satisfies a and is otherwise most plausible.[5]

Recall that a world w is more plausible than another w' if w, as compared to w', assigns to more variables the values about which the agent is certain. When there are ties on the first comparison, a world w is more plausible than another w' if w satisfies more of the structural equations in E than w. If an agent believes some particular facts for certain, the corresponding worlds are more plausible. By contrast to believing for certain, an agent merely believes causal relations. The worlds corresponding to the agent's certain beliefs are more plausible than the worlds corresponding to the agent's causal beliefs. In a slogan, the particular facts about which the agent is certain trump the believed causal relations. In what follows, we apply our account to several examples.

---

[5] The truth of $a >_e c$ at $\langle M, w \rangle$ does not depend on w. We might thus say as well: $a >_e c$ is true at M iff c is true at all $\langle M, w' \rangle$, where w' satisfies a and is otherwise most plausible. We have chosen the tuple formulation so that Definition 2 parallels Definition 1.

3 King Ludwig of Bavaria

Recall *Example* 1. You observe that the flag is down ¬**f** and that the lights are on **l**, so you believe these for sure. You also believe that the presence of the king causes the flag to be pulled up $(F = K)$ and the lights to be on $(L = K)$. The facts you believe for certain, ¬**f** and **l**, trump the causal relations you believe. No worlds are eliminated, some are just more plausible than others. On the left is a graphical representation of your causal model, where the set E of structural equations is $\{F = K, L = K\}$.[6] On the right is a table containing the set W of all possible worlds for the variables K, F, L. The most plausible worlds are w_3 and w_4.

	K	F	L
w_1	1	1	1
w_2	0	1	1
w_3	1	0	1
w_4	0	0	1
w_5	1	1	0
w_6	0	1	0
w_7	1	0	0
w_8	0	0	0
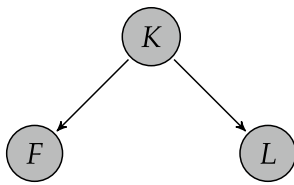
$$F = K$$
$$L = K$$
K: Is the king in the castle?
¬**f**: The flag is down.
l: The lights are on.

3.1 Causal Conditional

You believe the causal conditional $f >_c k$ to be true iff $f >_c k$ is true at $\langle M, w_3 \rangle$ and $\langle M, w_4 \rangle$. But $f >_c k$ is false at $\langle M, w_4 \rangle$ because k is false at $\langle M_f, w_4 \rangle$. The intervention by f replaces the structural equation $F = K$ by $F = 1$. K is unaffected by the intervention. For K is on the right-hand side of the structural equation, and so is not determined by an intervention on the left-hand side. More colloquially, causation flows only downstream and K is causally upstream from F. Hence, k is not true at $\langle M_f, w_4 \rangle$. If the flag had been up, this would not have caused the King to be in the castle. The causal reading (2) of conditional (1) is thus inappropriate. And the extant causal model semantics cited in the Introduction have no problem to account for this.

---

[6] This causal model is, of course, incomplete. There are other causal factors influencing whether or not the flag is up and the lights are on. However, we take it to be an advantage of our account that "small" causal models are sufficient for arriving at intuitive results.

3.2 Evidential Conditional

You believe the evidential conditional $f >_e k$ to be true iff $f >_e k$ is true at $\langle M, w_x \rangle$ for $x \in \{3, 4\}$. The evidential conditional $f >_e k$ is true at $\langle M, w_x \rangle$ because k is true at all $\langle M, w' \rangle$, where w' satisfies f and is otherwise most plausible. The antecedent f overwrites your certain belief in ¬**f**. The most plausible worlds that satisfy f are those that satisfy **l** as well, and the structural equations. The most plausible world is thus w_1. Hence, we need only to check whether k is true at $\langle M, w_1 \rangle$. This is the case. So you believe $f >_e k$ to be true. If the flag had been up, this would be evidence that the king is in the castle. The evidential reading (3) of conditional (1) is thus appropriate according to our account. The extant causal model semantics cannot account for this.

Kratzer (1989, p. 640) agrees that (1) is acceptable but asks: "why wouldn't the lights be out and the King still be away?" Well, on an evidential reading, you first keep fixed your certain beliefs backed by the available evidence as much as the supposition of the antecedent allows; in a second step, you identify the worlds which satisfy more of the structural equations than any other worlds; and these are your most plausible worlds. Unlike w_1, w_6 does not keep fixed your certain belief in **l**; moreover, unlike w_1, w_6 does not satisfy the structural equations. The answer is thus that Kratzer's alternative is not compatible with your certain beliefs backed by your evidence and the causal relations you believe to be true. Or so says our account.

4 Oswald–Kennedy Example

Adams (1970, p. 90) puts forth an example similar to the following.

Example 2 You believe that Oswald shot Kennedy (o) and that he acted alone: no one else shot Kennedy (¬s). However, you are not quite certain about this. By contrast you believe for sure that Kennedy has been shot (**k**) either by Oswald or by someone else. So you accept the indicative conditional:

(4)   If Oswald didn't shoot Kennedy, someone else did.

And yet you do not believe the corresponding subjunctive conditional:

(5)   If Oswald hadn't shot Kennedy, someone else would have.

On the left, you see a graphical representation of your causal model, where $E = \{K = S \vee O\}$. On the right, you see the set W of all possible worlds for the variables S, O, K. You believe for sure that Kennedy has been shot **k**. Moreover, you believe that it was either Oswald or someone else. w_1, w_2, w_3 are thus the most plausible worlds.

	S	O	K
w_1	1	1	1
w_2	0	1	1
w_3	1	0	1
w_4	0	0	1
w_5	1	1	0
w_6	0	1	0
w_7	1	0	0
w_8	0	0	0
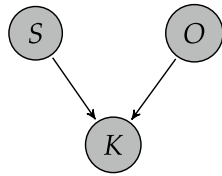
$K = S \lor O$

o: Oswald shot Kennedy.

$\neg s$: No one else shot Kennedy.

k: Kennedy has been shot.

4.1 Causal Conditionals

You believe the causal conditional $\neg o >_c k$ to be true iff $\neg o >_c k$ is true at $\langle M, w_x \rangle$ for $x \in \{1, 2, 3\}$. But $\neg o >_c k$ is false at $\langle M, w_2 \rangle$ because k is false at $\langle M_{\neg o}, w_2 \rangle$. If $\neg o$ is set by intervention in w_2, the structural equation determines $\neg k$. Hence, you do not believe $\neg o >_c k$. You do not believe "If Oswald had not shot Kennedy, Kennedy would have been shot".[7] The fact **k** you believe for certain is not kept fixed when evaluating the causal conditional. Rather the value for the variable K is overwritten by the downstream effects of intervening by $\neg o$ in world w_2.

Let us modify the example slightly. Suppose you believe for sure that someone else shot Kennedy. That is, you believe **s**. Hence, you believe w_1 or w_3 to be the actual world. Now, you believe $\neg o >_c k$ to be true. Given that someone else shot Kennedy: if Oswald had not, Kennedy would still have been shot.

4.2 Evidential Conditionals

You believe the evidential conditional $\neg o >_e s$ to be true iff $\neg o >_e s$ is true at $\langle M, w_x \rangle$ for $x \in \{1, 2, 3\}$. The evidential conditional $\neg o >_e s$ is true at each $\langle M, w_x \rangle$ because s is true at all $\langle M, w' \rangle$, where w' satisfies $\neg o$ and is otherwise most plausible. The only such w' is w_3. So you believe $\neg o >_e s$ to be true. Given you believe for sure that Kennedy has been shot, if Oswald did not shoot Kennedy, (this is evidence that) someone else did.[8]

---

[7] This conditional can be paraphrased as "If Oswald had not shot Kennedy, this would have caused that Kennedy was shot."

[8] The evidential conditional could also be paraphrased by "if Oswald did not shoot Kennedy, it must be that someone else did". The optional "it must be that" expresses that there is no other possibility for you under the assumption of the antecedent. This is good evidence indeed.

To evaluate an evidential conditional, the variable assignments that are believed for sure must be true in the worlds at which the consequent is evaluated, unless this conflicts with the antecedent. In the present example, **k** is believed for sure and thus must be true at the evaluation worlds. By contrast, suppose our agent were not sure whether Kennedy was shot. Then she would not believe $\neg o >_e s$ to be true. For the evidential conditional is not true at w_8.

Adams (1970) uses his example to argue that indicative and subjunctive conditionals differ. We have already seen that the indicative/subjunctive distinction cross-cuts the causal/evidential one. The subjunctive conditional (1) is evidential while the subjunctive (5) is causal, and the corresponding indicative (4) is evidential again.[9]

5 Hamburger Example

Rott (1999, Sec. 3) discusses an example due to Hansson (1989).

Example 3 Suppose that one Sunday night you approach a small town of which you know that it has exactly two snackbars. Just before entering the town you meet a man eating a hamburger (**h**). You have good reason to accept the following indicative conditional:

(6)   If snackbar *A* is closed, then snackbar *B* is open.

Suppose now that after entering the town, you see that *A* is in fact open (**a**). The question is whether we would accept the corresponding subjunctive conditional

(7)   If snackbar *A* were closed, then snackbar *B* would be open.

It seems clear to me that it is not justified to accept this conditional.

According to Rott, the example illustrates that subjunctive conditionals are understood in an ontic way while indicative conditionals are understood in an epistemic way. The indicative (6) tells you how you would revise your beliefs upon learning that snackbar *A* is closed. By contrast, the subjunctive (7) tells you what the world would be like if snackbar *A* were closed. And *A* being closed does not cause *B* to be open.[10] Rott's reason for rejecting the subjunctive conditional (7) is thus that you think there is no causal relation between the antecedent and the consequent.

On the left, you see a graphical representation of your causal model, where $E = \{H = A \vee B\}$. On the right, you see the set W of all possible worlds for the

---

[9] Moreover, we have shown that counterfactuals—subjunctives whose antecedent is (believed to be) false—can have an epistemic reading. (1) can be understood as "If the flag had been up, I would have *believed* that the King is in the castle."

[10] For details on the distinction between epistemic and ontic conditionals, see Lindström and Rabinowicz (1995).

variables H, A, B. You believe for certain that snackbar A is open **a** and the man eats a hamburger **h**. Moreover, you believe that the man can only eat the hamburger if snackbar A or snackbar B is open. w_1 and w_3 are thus the most plausible worlds.

	A	B	H
w_1	1	1	1
w_2	0	1	1
w_3	1	0	1
w_4	0	0	1
w_5	1	1	0
w_6	0	1	0
w_7	1	0	0
w_8	0	0	0
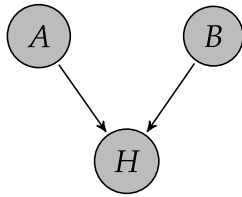
$$H = A \vee B$$

a: Snackbar A is open.
h: Man eats a hamburger.
B: Is snackbar B open?

5.1 Causal Conditional

You believe the causal conditional $\neg a >_c b$ to be true iff $\neg a >_c b$ is true at $\langle M, w_x \rangle$ for $x \in \{1, 3\}$. But $\neg a >_c b$ is false at $\langle M, w_3 \rangle$ because b is false at $\langle M_{\neg a}, w_3 \rangle$. Even if $\neg a$ is set by intervention in w_3, b is false there. Hence, you do not believe $\neg a >_c b$. You do not believe "If snackbar A were closed, this would cause snackbar B to be open"; for all we know, snackbar B might be closed as well if A were.

5.2 Evidential Conditional

Before seeing that snackbar A is in fact open, you believe **h** but you have no beliefs about a and b (except the structural equation). That is, you believe the worlds w_1, w_2, and w_3 to be most plausible. Why? Because only these worlds satisfy **h** and the structural equation.

You believe the evidential conditional $\neg a >_e b$ to be true iff $\neg a >_e b$ is true at $\langle M, w_x \rangle$ for $x \in \{1, 2, 3\}$. The evidential conditional $\neg a >_e b$ is true at $\langle M, w_x \rangle$ because b is true at all $\langle M, w' \rangle$, where w' satisfies $\neg a$ and is otherwise most plausible. The only world w' that satisfies $\neg a$, **h**, and the structural equation is w_2, and b is true there. So you believe $\neg a >_e b$ to be true. Given you believe for sure that the man eats a hamburger, if snackbar A is closed, (this is evidence that) snackbar B is open.[11]

Our account delivers the desired results. And it seems that the evidential/causal distinction is similar to the epistemic/ontic distinction. In light of the King Ludwig

---

[11] Again, the evidential conditional could be paraphrased by "If snackbar A is closed, it must be that snackbar B is open." See footnote 8.

example, however, we refrain from associating these distinctions with the mood of a conditional.
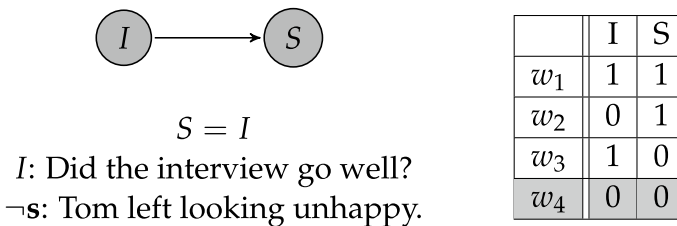
6 Backtracking

A backtracking conditional traces some causes from an effect: if this effect had not occurred, (it must have been that) some of its causes would have been absent.[12] To illustrate such conditionals, consider the following scenario inspired by Veltman(2005, p. 179).

Example 4 Tom and Marianne wait in the lobby to be interviewed for a great job. Tom goes in first, Marianne continues to wait outside the interview room. When Tom comes out he looks rather unhappy. Marianne thinks:

(8)   If Tom had left the interview smiling, the interview would have gone well.

(8) is a backtracking conditional. Tom had an interview and comes out looking rather unhappy. Marianne believes the causal hypothesis that the interview influences whether or not Tom looks happy. In particular, she believes that the interview going bad ($\neg i$) caused him to look rather unhappy ($\neg s$). Assuming that the effect—looking unhappy—is not the case, the cause—the interview not going well—must have been different.

On the left, you see a graphical representation of your causal model, where $E = \{S = I\}$. On the right, you see the set W of all possible worlds for the variables S, I. You believe for certain that Tom looks rather unhappy $\neg s$. Moreover, you believe that whether or not the interview went well affects whether or not Tom is smiling. So w_4 is the most plausible world.

$$S = I$$

I: Did the interview go well?
$\neg s$: Tom left looking unhappy.

	I	S
w_1	1	1
w_2	0	1
w_3	1	0
w_4	0	0

---

[12] For details, see Lewis (1979).

6.1 Causal Conditional

Like all backtracking conditionals, you do not believe the backtracking counterfactual (8) under the causal reading. An intervention on a variable effects changes only causally downstream, and so does not change its causal past. You believe the causal conditional $s >_c i$ to be true iff $s >_c i$ is true at $\langle M, w_4 \rangle$. But $s >_c i$ is false at $\langle M, w_4 \rangle$ because i is false at $\langle M_s, w_4 \rangle$. Even if s is set by intervention in w_4, i is false there. Hence, you do not believe $s >_c i$. You do not believe "If Tom had left the interview smiling, this would have caused the interview to go well".[13]

6.2 Evidential Conditional

You believe the evidential conditional $s >_e i$ to be true iff $s >_e i$ is true at $\langle M, w_4 \rangle$. The evidential conditional $s >_e i$ is true at $\langle M, w_4 \rangle$ because i is true at all $\langle M, w' \rangle$, where w' satisfies s and is otherwise most plausible. The only world w' that satisfies s and the structural equation is w_1, and i is true there. So you believe $s >_e i$ to be true. If Tom had left the interview smiling, this would have been evidence that the interview went well. Our evidential conditional captures backtracking.

7 A Generalization

So far, our account relies on a simple distinction between believing a particular fact for certain and not doing so. This simple distinction factors into the plausibility order over worlds. We have said that worlds that assign more variables the values about which the agent is *certain* are more plausible than worlds which assign those values to fewer variables. But the underlying distinction is too simple, as we will show now. Then we will generalize our account in response.

Consider a modification of the Oswald-Kennedy example.[14]

Example 5 You believe that Oswald shot Kennedy (o) and that he acted alone: no one else shot Kennedy ($\neg s$). You also believe that Kennedy has been shot (k) either by Oswald or by someone else. You are, however, not certain about any of your beliefs. Still, you are much more certain that Kennedy has been shot than that it was Oswald. So you accept the conditional:

(4)   If Oswald didn't shoot Kennedy, someone else did.

---

[13] By contrast, you *do* believe "If the interview goes well, Tom will leave it smiling". This causal conditional in the indicative mood completes our claim that our causal/evidential distinction cross-cuts the subjunctive/indicative distinction.

[14] The modification has been suggested by an anonymous reviewer to whom we are very grateful.

On our present account, however, you do not believe (4). The reason is that you have no certain beliefs. The most plausible worlds are thus all the worlds that satisfy the structural equation $K = S \vee O$. These worlds are w_1, w_2, w_3, and w_8 of the original Oswald-Kennedy example. You believe $\neg o >_e s$ to be true iff $\neg o >_e s$ is true at $\langle M, w_x \rangle$ for $x \in \{1, 2, 3, 8\}$. But $\neg o >_e s$ is false at $\langle M, w_8 \rangle$ because s is false at w_8. Hence, you do not believe $\neg o >_e s$.

The result would be acceptable if you were just as sure that Kennedy was killed than that it was Oswald (cf. Sect. 4.2). As long as you are more certain that Kennedy was killed, however, the result is inacceptable. The problem seems to be that the present account is only sensitive to certain beliefs about particular facts.[15] In the modified Oswald-Kennedy example, by contrast, it seems to matter that you are much more certain that Kennedy has been shot than that it was Oswald. But the present account is blind for this relative certainty.

We modify our account. The idea is that it is not certainty but relative certainty that actually matters. You believe that a particular fact is more certain than another. And this makes—in your view—some worlds more plausible than others. We therefore replace (i) in our plausibility order as follows: worlds that assign to more variables the values about which the agent is *most certain* are more plausible than worlds which assign those values to fewer variables. An agent is most certain about a variable value if (a) she is at least quite certain about the variable value, and (b) there is no variable value of which she is more certain.

The modification accounts for *Example* 6. To see this, recall that you are quite confident that Kennedy has been shot and more certain of this than that it was Oswald. Moreover, you are more certain that Oswald did it than somebody else. There are no other variable values involved. Hence, you are most certain that Kennedy has been shot—even though you are not certain of it. There is no other variable value of which you are most certain. In the presence of the structural equation, the most plausible worlds are thus w_1, w_2, w_3. And so you believe (4). Our final account treats the modified Oswald-Kennedy example like our preliminary account treats the original one.

Our final account is a proper generalization of the preliminary account. If you believe a variable value for certain, you also believe it for most certain. Believing a particular fact for *certain* is—so to speak—the highest degree of being *most certain*. It does therefore not come as a surprise that our final and preliminary account agree on all the examples of the previous sections. What we have learned in this section is this: in general it is relative certainty that matters for the evaluation of evidential conditionals.

---

[15] On the formal level, the problem seems to be that *all* the worlds that satisfy the structural equation are equally plausible. But, intuitively, w_1, w_2, w_3 should not be treated on a par with w_8.

8 Conclusion

We have put forth a causal model account for the evaluation of causal and evidential conditionals. An agent comes equipped with beliefs about causal relations and beliefs about particular facts. Our account, like many other causal model accounts, captures causal conditionals rather well. Unlike other causal model accounts, ours captures as well evidential conditionals. Embedding a causal model in an agent's belief state enables the agent to evaluate evidential conditionals better than without the information encoded in the causal model. Unsurprisingly, the believed causal relations help determine which facts are to be kept fixed when evaluating a conditional. Notably, however, neither the sophisticated causal model account of conditionals due to Pearl (2011) nor the one due to Halpern (2016) provide the desired results for evidential conditionals like the one in the King Ludwig example. We have thus provided the first extension of a causal model account to what we call evidential conditionals.

Deng and Lee (2021) proposed a causal model semantics that has structural similarities to our account.[16] Their semantics is designed to account for conditionals in the indicative and subjunctive mood, respectively. And indeed, they capture the Oswald-Kennedy conditionals and other minimal pairs in a way that parallels our account. However, like the other causal model semantics, theirs faces troubles with the King Ludwig conditional. For Kratzer's If the flag had been up, the King would have been in the castle is a subjunctive conditional which comes out false under their account of subjunctive conditionals. Our account, by contrast, can explain why this evidential subjunctive is acceptable.

We have observed that the distinction between evidential and causal conditionals cross-cuts the distinction between indicative and subjunctive conditionals. Our account relies on beliefs about causal relations and can therefore explain why certain evidential subjunctives are acceptable. Furthermore, our distinction between evidential and causal seems to refine the distinction between epistemic and ontic conditionals drawn by Lindström and Rabinowicz (1995). In the Hamburger example, both the evidential/epistemic and the causal/ontic conditional are epistemic in the sense that they are evaluated relative to an agent's beliefs.

Of course, we do not claim that our account captures all readings of all conditionals. It has, in particular, no resources to model uncertainty about causal relations. A generalization to remedy this situation must await another occasion. Still our account is quite widely applicable given its simplicity. So we may hope that our account is a first step towards a more sophisticated account based on which we can capture our intuitions about causal and evidential conditionals. A causal model account of the type proposed here may one day even instruct an intelligent machine to evaluate conditionals just like we do.

---

[16] We would like to thank an anonymous reviewer for pointing this out.

References

Adams, E. W. (1970). Subjunctive and indicative conditionals. *Foundations of Language, 6,* 89–94.

Briggs, R. (2012). Interventionist counterfactuals. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 160*(1), 139–166.

Ciardelli, I., Zhang, L., & Champollion, L. (2018). Two switches in the theory of counterfactuals. *Linguistics and Philosophy, 41*(6), 577–621.

Deng, D.-M., & Lee, K. Y. (2021). Indicative and counterfactual conditionals: A causal-modeling semantics. *Synthese, 199*(1), 3993–4014.

Galles, D., & Pearl, J. (1998). An axiomatic characterization of causal counterfactuals. *Foundations of Science, 3*(1), 151–182.

Günther, M. (2017). Learning conditional and causal information by Jeffrey imaging on Stalnaker conditionals. *Organon F, 24*(4), 456–486.

Günther, M. (2018). Learning conditional information by Jeffrey imaging on Stalnaker conditionals. *Journal of Philosophical Logic, 47*(5), 851–876.

Halpern, J. Y. (2000). Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research, 12*(1), 317–337.

Halpern, J. Y. (2016). *Actual causality*. The MIT Press.

Hansson, S. O. (1989). New operators for theory change. *Theoria, 55*(2), 114–132.

Hiddleston, E. (2005). A causal theory of counterfactuals. *Noûs, 39*(4), 632–657.

Hitchcock, C. (2001). The intransitivity of causation revealed in equations and graphs. *Journal of Philosophy, 98*(6), 273–299.

Kratzer, A. (1989). An investigation of the lumps of thought. *Linguistics and Philosophy, 12*(5), 607–653.

Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs, 13*(4), 455–476.

Lindström, S., & Rabinowicz, W. (1995). The Ramsey test revisited. In G. Crocco, L. F. Nasdel-Cerro, & A. Herzig (Eds.), *Conditionals: From philosophy to computer science* (pp. 131–182). Oxford University Press.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.

Pearl, J. (2011). The algorithmization of counterfactuals. *Annals of Mathematics and Artificial Intelligence, 61*(1), 29.

Rott, H. (1999). Moody conditionals: Hamburgers, switches, and the tragic death of an American President. In *JFAK. Essays dedicated to Johan van Benthem on the occasion of his 50th birthday* (pp. 98–112).

Santorio, P. (2019). Interventions in premise semantics. *Philosophers' Imprint, 19*(1), 1–27.

Schulz, K. (2011). If you'd wiggled A, then B would've changed. *Synthese, 179*(2), 239–251.

Veltman, F. (2005). Making counterfactual assumptions. *Journal of Semantics, 22*(2), 159–180.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press.

Zhang, J. (2013). A Lewisian logic of causal counterfactuals. *Minds and Machines, 23*(1), 77–93.