Chancy-but-true Counterfactuals

Joshua Edward Pearson
pearsonj@mit.edu

WIP
10/3/24

I. A Classic Tension

(i) Were I to throw my pen, it would hit Kevin.

– *But wait!* By our best scientific theories, were I to throw my pen, there would be non-zero chance that the pen quantum tunnels through Kevin. 'Pessimists' worry this means (i) is false.[1]

– There's a debate to be had.[2] But: if we accept that our world is thoroughly chancy, and we're 'optimists' (i.e. anti-skeptical) about ordinary counterfactuals, we need some account of how a counterfactual $A \mathrel{\Box\!\!\rightarrow} C$ can be true despite $Ch(\neg C \mid A) > 0$.[3]

[1] And most other ordinary counterfactuals along with it. Hajek (2014) is the prominent pessimist.

[2] See e.g. Hawthorne (2005); Hajek (2014); Lewis (2016); Boylan (2023) for discussion

[3] You might be optimistic and deny this. On a Moorean approach, it's enough to rebut any purported argument for counterfactual skepticism without offering your own theory of counterfactuals. I'm trying to be more ambitious.

II. A Modern Classic Tension

An inconsistent triad:

(Optimism) Most ordinary counterfactuals are true.

(Ch-Tr) If $Ch(\neg C \mid A) \geq 0.5$, then $\neg(A \mathrel{\Box\!\!\rightarrow} C)$.

(Poss Pres) If $A \mathrel{\Box\!\!\rightarrow} C$ and $\neg(A \mathrel{\Box\!\!\rightarrow} \neg B)$ then $A\&B \mathrel{\Box\!\!\rightarrow} C$.

> **Leafy.**[4] Leafy the maple leaf shed today (October 3rd). Still, *had Leafy not shed today, it would have shed by spring*. **However!** It turns out that the physics of leaf-shedding works roughly as follows: for any particular morning on which a leaf has not yet shed, there is a 0.5 chance it still won't have shed the morning after.

[4] This is just a counterfactual variant of a now very famous case from Dorr et al. (2014). Leitgeb (2013) outlines a similar paradox, though I find his case less intuitive.

(1) (Leafy not shed on Oct 3) $\mathrel{\Box\!\!\rightarrow}$ (Leafy shed by Spring). (Optimism)

(2) not-[(Leafy not shed on Oct 3) $\mathrel{\Box\!\!\rightarrow}$ (Leafy shed on Oct 4)] (Stip + Ch-Tr)

(3) (Leafy shed on neither Oct 3 nor 4) $\mathrel{\Box\!\!\rightarrow}$ (Leafy shed by Spring) (1,2,Poss-Pres)

(4) not-[(Leafy shed on neither Oct 3 nor 4) $\mathrel{\Box\!\!\rightarrow}$ (Leafy shed on 5)] (Stip + Ch-Tr)

(5) (Leafy shed on neither Oct 3, 4 nor 5) $\mathrel{\Box\!\!\rightarrow}$ (Leafy shed by Spring) (3,4,Poss-Pres)

...Keep repeating until you get the absurd:

(N) (Leafy not shed by Spring) $\mathrel{\Box\!\!\rightarrow}$ (Leafy shed by Spring)

Fact: The majority of theories entail **Poss-Pres**.[5] So if they are optimistic, they deny **Ch-Tr**. But **Ch-Tr** looks very plausible!

[5] Both variably strict accounts like Lewis (1973) and Stalnaker (1968) and strict accounts like Von Fintl (2001), Lewis (2016) and Williamson (2020).

Mission: I'm an optimist. But I'm also an optimistic optimist. I want a non-skeptical theory of counterfactuals that doesn't so radically divorce counterfactual truth from chance! Why not deny **Poss Pres**?[6]

[6] Boylan and Schultheis (2021) deny **Poss Pres** for very different reasons. I need to think about how my approach here does/doesn't interact with theirs.

III. Counterfactual Lockeanism

(CL) $A \,\Box\!\!\rightarrow C$ iff $Ch(C \mid A) \geq T$. ($\frac{1}{2} < T < 1$)

I'm glossing over a lot of important details here, such as various relevant contextual parameters, and the matter of how the relevant chance function is to be picked out. On the latter, see Leitgeb (2012) and Leitgeb (2013).

- **Problem:** Beyond **Poss Pres**, the following principles *also* fail.[7]

  (Agglomeration) If $A \,\Box\!\!\rightarrow C_1$ and $A \,\Box\!\!\rightarrow C_2$ then $A \,\Box\!\!\rightarrow (C_1 \& C_2)$.[8]

  (Nec Pres) If $A \,\Box\!\!\rightarrow C$ and $A \,\Box\!\!\rightarrow B$ then $(A \& B) \,\Box\!\!\rightarrow C$.[9]

[7] Also faces the problems outlined in **V**.

[8] Hawthorne (2005) basically takes Agglomeration failure alone to be enough reason to reject **CL**.

[9] A point I thank to Goodman and Salow (fc).

IV. How to have your chancy cake and eat it too

How do we get a chance-friendly view that satisfies **Agglomeration**?

- For any set of propositions Γ, if we can identify a "strongest proposition" S in Γ — such that $P \in \Gamma$ iff S entails P — Γ is guaranteed to be "closed under conjunction".[10]
- Let $A^{\Box}$ (spoken "A squared") be the strongest proposition that would be true were A true: $A \,\Box\!\!\rightarrow C$ iff $A^{\Box}$ entails C.
- Now use facts about $Ch(\cdot \mid A)$ to determine $A^{\Box}$!

[10] Why's that?

- If $A \in \Gamma$, S entails A.
- So if $A, B \in \Gamma$, S entails both A and B
- So S entails $A \& B$.
- So $A \& B$ is in Γ.

There are many ways to implement this strategy — different ways of using facts about $Ch(\cdot \mid A)$ to determine $A^{\Box}$[11] — *page 5 offers a menu of different options with strengths and weaknesses.* Here I'll focus on a counterfactual analogue of Goodman and Salow (fc):

[11] One can see that last 12 years of probabilistic theories of belief as offering different ways to implement this strategy in a different context — Lin and Kelly (2012); Leitgeb (2014); Goldstein and Hawthorne (2021); Hong (fc); Goodman and Salow (fc); Pearson (ms)

(GaSC) $A^{\Box}$ is the smallest set S such that:

(i) For all $w \in S$, if $Ch(w' \mid A) \geq Ch(w \mid A)$, $w' \in S$.

(ii) $Ch(S \mid A) \geq T$; $\frac{1}{2} < T < 1$

- *Intuitively:* Add worlds into $A^{\Box}$, starting with the likeliest, then adding the second-most likely, and so on, until $A^{\Box}$ itself is at least T-likely.
- **GaSC** predicts **Ch-Tr**, **Agglomeration** *and* **Nec Pres** — at the price of **Poss Pres**. So far, so good![12]

[12] But see the menu for more details; one interesting problem is that **GaSC** violates "**Partition+**", which says: If $(A \& B) \,\Box\!\!\rightarrow C$ and $(A \& (\neg B)) \,\Box\!\!\rightarrow C$ then $A \,\Box\!\!\rightarrow C$.

V. Some Complications

V.I A Need for Partition-Sensitivity

The problems outlined in this section occur for all of the views on the menu!

- **GaSC** mentions the chances of various *possible worlds* — consistent and complete specifications of how the world might be. There's *a lot of these*; perhaps infinitely many. If so, arguably, they have a chance of at most 0. But then **GaSC** predicts $A^{\Box} = A$.

That's counterfactual skepticism!

- To solve this, understand the individuation of worlds in W as dependant on the conversation context: they can be pretty coarsely individuated — W only makes the distinctions needed for the conversational purposes.[13] That is, view W as a partition of all the possible worlds that is more-or-less fine grained.

[13] *Argument from authority:* that's what Agustin said last week!

- The move to partition-sensitivity has further motivation:

> **Hajek's Biased Coin.**[14] Alan has a coin that is extremely biased — 99.99% — towards heads. He's a busy guy: he didn't flip it 100 million times.

[14] From Hajek (2014); see also Boylan (2023).

(1H) Had Alan flipped it 100 million times, it would have landed heads on the first toss.

Shouldn't optimists like me like (1H)?

(nH) Had Alan flipped it 100 million times, it would have landed heads on the nth toss.

But then why not like (nH)?

($\forall$H) Had Alan flipped it 100 million times, it would have landed heads every time.

But then **Agglom** gives you ($\forall$H)!

($\exists$T) Had Alan flipped it 100 million times, it would have landed tails sooner or later.

And you'll have to deny ($\exists$T)!

- Partition-sensitivity gives us *a lot* of flexibility.[15] Whether the above claims are true depends very much on the relevant partition. This can be spun an advantage: I can get into two minds as to whether ($\exists$T) is true — now we can explain why!

[15] *Time for some context-sensitive gymnastics!*

V.II Not-so-counterfactual counterfactuals

- **The 'T+F=F' Problem.** Counterfactuals with true antecedents and false consequents sound false:

> **Missed Penalty.** You see what looks like Jobe Bellingham take a penalty for England and miss. You assert "Had his brother Jude taken that, it would have gone in." It turns out the penalty taker *was* Jude — Jobe and Jude are easily confused brothers. So didn't you say something false? After all, Jude *did* take the penalty yet the ball *did not* go in!

  - If the actual world is sufficiently unlikely relative to others, **GaSC** predicts failures of **T+F=F**.[16]
  - Fix: we can add a ad hoc condition that insures $A^{\Box}$ contains @ when $@ \in A$.[17] This feel ad hoc. *Too* ad hoc?...

[16] Surprisingly, Williamson (2020) denies **T+F=F**, but for reasons not relevant to here.

[17] e.g. add (iii) — if $@ \in A$, $@ \in S$ — as a third condition to **GaSC**

- **The 'T+T=T' Problem.** Counterfactuals with true antecedents and true consequents can sound true.

> **Theft.** I've stolen Hajek's biased coin. Kevin boldly asserts "Were to flip that, it would land tails." I disagree. Remarkably, I flip it and it lands tails. Can't Kevin say "See? I was right!"

  - Again, this can fail on **GaSC** if @ is very unlikely. No ad hoc fix this time — it's a genuine choice-point. Either give up **T+T=T** or divorce counterfactual truth from chance.[18]

[18] Thankfully I'm far from the only one to flirt with giving up **T+T=T**. A few examples are: Bennett (1974); Bennett (2003); Leitgeb (2013); Hajek (2014); Williamson (2020)

– Another move worth toying with. Give in to counterfactual skepticism, and propose theories like **GaSC** as an account of when counterfactuals are *assertable*, rather than true.

References

Bennett, J. (1974). Counterfactuals and possible worlds. *Canadian Journal of Philosophy*, 4(December):381–402.

Bennett, J. (2003). *Philosophical Guide to Conditionals*. Oxford University Press UK, Oxford, GB.

Boylan, D. (2023). Counterfactual skepticism is (just) skepticism. *Philosophy and Phenomenological Research*, 108(1):259–286.

Boylan, D. and Schultheis, G. (2021). How strong is a counterfactual? *Journal of Philosophy*, 118(7):373–404.

Dorr, C., Goodman, J., and Hawthorne, J. (2014). Knowing against the odds. *Philosophical Studies*, 170(2):277–287.

Goldstein, S. and Hawthorne, J. (2021). Knowledge from multiple experiences. *Philosophical Studies*, 179(4):1341–1372.

Goodman, J. and Salow, B. (fc). Belief revision normalized. *Journal of Philosophical Logic*.

Hajek, A. (2014). *Most Counterfactuals Are False*.

Hawthorne, J. (2005). Chance and counterfactuals. *Philosophy and Phenomenological Research*, 70(2):396–405.

Hong, F. (fc). Prefaces, knowledge, and questions. *Ergo*.

Leitgeb, H. (2012). A probabilistic semantics for counterfactuals. part a. *Review of Symbolic Logic*, 5(1):26–84.

Leitgeb, H. (2013). A lottery paradox for counterfactuals without agglomeration. *Philosophy and Phenomenological Research*, 89(3):605–636.

Leitgeb, H. (2014). The stability theory of belief. *Philosophical Review*, 123(2):131–171.

Lewis, D. (1973). *Counterfactuals*. Blackwell.

Lewis, K. S. (2016). Elusive counterfactuals. *Noûs*, 50(2):286–313.

Lin, H. and Kelly, K. (2012). Propositional reasoning that tracks probabilistic reasoning. *Journal of Philosophical Logic*, 41(6):957–981.

Pearson, J. E. (ms). Belief revision revised.

Stalnaker, R. C. (1968). A theory of conditionals. In Rescher, N., editor, *Studies in Logical Theory (American Philosophical Quarterly Monographs 2)*, pages 98–112. Oxford: Blackwell.

Williamson, T. (2020). *Suppose and Tell: The Semantics and Heuristics of Conditionals*. Oxford, England: Oxford University Press.

A Menu of Chance-friendly Theories of the Counterfactual

Counterfactual Lin and Kelly (CLaK)

Ingredients:

- $A^{\Box} = \{w_1 \in W : \forall w_2 \in W, \neg[Ch(w_2 \mid A) > c \times Ch(w_1 \mid A)]\}$. $(c \geq 1)$
- But if $@ \in A$, any world w such that $Ch(w \mid A) \geq Ch(@ \mid A)$ is also in $A^{\Box}$.

Description: *A daring dish in which $A^{\Box}$ contains all worlds from W that are not significantly less likely (given A) than some other world in W. Denies* ***Poss Pres*** *without giving up on any other plausible principle (e.g. neither* ***Nec Pres*** *nor* ***Partition+****). A complex combination of local flavors.*

1-star Google Review: `Horrendous. Advertised as sustaining a strong connection between counterfactual truth and chance, but when the meal arrived, I found a true counterfactual whose consequent had a greater than 50% chance of being false given the antecedent! (i.e.` **Ch-Tr** `fails) Needless to say, I sent it back immediately. 1 star.`

Counterfactual Goodman and Salow (GaSC)

Ingredients: $A^{\Box}$ is the smallest set S such that:

(i) For all $w \in S$, if $Ch(w' \mid A) \geq Ch(w \mid A)$, $w' \in S$.

(ii) $Ch(S \mid A) \geq T$; $\frac{1}{2} < T < 1$

(iii) If $@ \in A$, $@ \in S$.

Description: *A beautifully simple dish, building $A^{\Box}$ by gathering the likeliest worlds until the set is at least T-likely. A meal that satisfies a strong counterfactual-chance connection (***Ch-Tr***). Logical flavor with a probabilistic satisfying bite!*

1-star Google Review: `I saw this theory had great reviews, so I thought I'd check it out. But when I asked the waiter whether I could infer` C `from` $A \Box\!\!\rightarrow C$ `and` $(\neg A) \Box\!\!\rightarrow C$`, he said no!!! (i.e.` **Partition+** `fails.) ENDORSE THIS VIEW AT YOUR PERIL. 1 star.`

Counterfactual J.E. Pearson (CJEP)

Ingredients: $A^{\Box}$ is the smallest set S such that:

(i) If $Ch(w \mid A) \geq \tau$, $w \in S$ (for chosen $0 < \tau < 1$)

(ii) For all $w \in S$, if $Ch(w' \mid A) \geq Ch(w \mid A)$, $w' \in S$.

(iii) $Ch(S \mid A) \geq T$; $\frac{1}{2} < T < 1$

(iv) If $@ \in A$, $@ \in S$.

Description: *$A^{\Box}$ contains all the sufficiently (at least τ-likely) worlds, and then some, until $A^{\Box}$ becomes at least T-likely. A wonderfully flexible theory for modelling: whether it's complicated coin-flipping cases or composer-confusion, CJEP can handle it!*

1-star Google Review: `Super cool guy, but no respect for logicians like me. He said` B `would be true were` A`, and that` C `would be true were` A`, but then denied that` C `would be true were` $A\&B$`! Both` **Nec Pres** `*and*` **Partition+** `fail!! Also, what's with all the thresholds? 1 star.`

Relevant Principles:

(Poss Pres) If $A \Box\!\!\rightarrow C$ and $\neg(A \Box\!\!\rightarrow \neg B)$ then $A\&B \Box\!\!\rightarrow C$.

(Nec Pres) If $A \Box\!\!\rightarrow C$ and $A \Box\!\!\rightarrow B$ then $(A\&B) \Box\!\!\rightarrow C$.

(Partition+) If $(A\&B) \Box\!\!\rightarrow C$ and $(A\&(\neg B)) \Box\!\!\rightarrow C$ then $A \Box\!\!\rightarrow C$.

(Ch-Tr) If $Ch(\neg C \mid A) \geq 1 - T$ then $\neg(A \Box\!\!\rightarrow C)$.

For more information about the pros + cons of possible theories, see Goodman and Salow's "Belief Revision Normalized" (forthcoming in Journal of Philosophical Logic), making the required adjustments to fit the counterfactual setting + add some kind of ad hoc constraint for when $@ \in A$.