

Tossing Morgenbesser's Coin

Zachary Goodsell

December 11, 2021

1 Morgenbesser's Coin

An indeterministic coin is tossed and while it is in the air you decline to bet that it will land heads. Then it lands heads. Morgenbesser observed that the following counterfactual seems obviously true: if you had accepted the bet, then you would have won. This observation of Morgenbesser's was reported by Slote (1978), whose theory requires that the apparently true counterfactual be false. Undaunted, Slote disputes our judgement about the counterfactual:

I know of no theory of counterfactuals that can adequately explain why such a statement seems natural and correct. But perhaps it simply isn't correct and the correct retort to it is: "no, you're wrong; if I had bet (heads), the coin might have come up differently and (so) I might have lost – assuming the coin was random."
(1978: 27)

This would be a fair response if Morgenbesser's case was an outlier, disconnected from the more central cases of counterfactual reasoning. But it is not. As Edgington (2004) points out, Morgenbesser's intuition seems to flow from a pervasive disposition to infer counterfactual independence from causal independence. Here are some cases of the sort that Edgington uses to illustrate the point:

Two Coins

Two indeterministic and causally isolated coins are tossed, a nickel and a dime. The nickel lands heads and the dime lands tails. You would have won a prize if both had landed heads.

Lottery

Your rival enters an indeterministic lottery with a one in a million chance of winning, and he wins. You could have scratched your nose moments earlier, but by chance you did not.

Football

A goalkeeper has to dive left or right before she knows where the striker will aim (the ball flies too fast). She dives right, and luckily prevents a goal.

In the Two Coins case, the isolation of the coins seems to ensure that how the nickel landed wouldn't have been different if the dime had landed heads. That's why it makes sense to wish that the dime had landed heads: if it did, you would have won the prize. In the lottery case, your rival would still have won even if you had scratched your nose earlier. But as Edgington points out, Slote seems committed to the claim that your scratching would have almost certainly resulted in your rival's not winning. What a pity you didn't! And in the Football case, the striker should think: 'If only I had aimed the other way, I would have scored.' The goalkeeper should likewise think: 'Lucky I dived right, otherwise she would have scored.' But if Slote is right then both of these regrets are incorrect.

Intuitions seem to tell in favour of Morgenbesser. I argue here, however, that Slote's wishful thinking pays off. Counterintuitively, Morgenbesser was simply wrong about the coin. We cannot say whether it would have landed heads if you had bet. The argument relies on the following case:

Infinitely Many Coins

Infinitely many indeterministic and causally isolated coins are tossed. By coincidence, they all land heads.

Take an arbitrary one of these coins c . The Morgenbesser intuition suggests that c would still have landed heads no matter what pattern the other coins might have landed in. In particular, say some infinite set I of coins *lands in a normal pattern* if infinitely many of the members of I land heads, and infinitely many do not. And let C^- be the set of the coins besides c . We are inclined to judge:

1. If C^- had landed in a normal pattern, then c would still have landed heads.

By contrast, let C be the set of all the coins including c (i.e., $C = C^- \cup \{c\}$). Since all the members of C might be on a par, as far as their disposition to land heads or tails, we are not at all inclined to accept:

2. If C had landed in a normal pattern, then c would still have landed heads.

Rather, (2) seems to have about a fifty-fifty chance of being true, as opposed to (1)'s certainty.

But there is a problem: it is necessary that C^- lands in a normal pattern if and only if C does. To see why, suppose that C^- lands in a normal pattern. Then, infinitely many of the coins besides c land heads and infinitely many do not. So infinitely many of the coins in C land heads and infinitely many do not, so C lands in a normal pattern. On the other hand, suppose that C lands in a normal pattern. The addition of c to a set cannot take it from finite to infinite, so infinitely many of the coins in C^- land heads and infinitely many don't. So C^- lands in a normal pattern.

Therefore, the intuitive judgements about (1) and (2) are incompatible with the following orthodox principle:

(Substitution) Necessarily equivalent propositions have the same counterfactual consequences.

$$\Box(p \leftrightarrow q) \rightarrow ((p \Box\rightarrow r) \leftrightarrow (q \Box\rightarrow r))$$

Substitution is widely – but not unanimously – accepted. It is built into the two dominant theories of counterfactuals, Stalnaker's (1968) and Lewis' (1973). The principle has also been explicitly defended in the literature, for example by Williamson (2020: 214–21) and Bacon (forthcoming).

I also accept Substitution, but making the case here would take us too far afield. In what follows, Substitution will be assumed. Thus we have a puzzle: counterintuitively, we cannot accept (1) without also accepting (2).

The argument schematically goes as follows: given Substitution, we must concede either that (2) is true or that (1) might be false. In Section 2 I argue that the former option is untenable, so we must withhold on (1). We must withhold, that is, on whether c would still have landed heads if the other coins had landed in a normal pattern. Section 3 shows that the judgement

that (1) is true is, modulo some irrelevant differences, the Morgenbesser intuition. So the Morgenbesser intuition should be relinquished (if Substitution is correct).¹

2 Against Independence

2.1 A Dilemma

Given Substitution, the options are to cling to (1) and say that (2) is therefore certainly true, or to concede that (1) is not certainly true because (2) is either uncertain or false. This section considers the first option. That is, we accept

1. If C^- had landed in a normal pattern, then c would still have landed heads.

And we use Substitution to conclude that,

2. If C had landed in a normal pattern, then c would still have landed heads.

Those who accept (2) face a dilemma. Granting that c would still have landed heads if C had landed in a normal pattern, we can ask: do the other coins also have this property? Either all the coins have the property, or not all do.

2.2 Horn I: Homogeneity

On the first option we say that for each coin x in C , x would still have landed heads if the coins in C besides x landed in a normal pattern. Thus, each such x would still have landed heads if the coins in C including x landed in a normal pattern. But this seems inconsistent. If every coin would have landed heads, if the coins in C landed in a normal pattern, where could we find those infinitely many tails that a normal pattern requires?

1. The puzzle is similar to one used by Fine (2012) to argue against Substitution, and the Infinitely Many Coins case is a variant of a case Builes (2020) uses to argue against an analogue of Substitution for probability operators. Thus, it is not out of the question to use the present puzzle as a *reductio* of Substitution. Still, all can agree that the *reductio* can only be performed with a clear picture of what Substitution forces us to say about the puzzle, which is what this paper provides. (See also Bacon, forthcoming for a response to Fine, and Dorr, Hawthorne, and Isaacs 2021 for a response to Builes.)

For this reason, very few authors who have considered puzzles of this sort can take Horn I of the present dilemma. For example, Fine (2012), as well as Pollock (1976), Herzberger (1979), Caie (2018) and Bacon (forthcoming), all endorse a principle of the following sort:

(Infinitary Closure) If each of some propositions is a counterfactual consequence of something, then so is their conjunction.

On Horn I, each of the coins' landing heads is a counterfactual consequence of C 's landing in a normal pattern. So by Infinitary Closure it is a counterfactual consequence of C 's landing in a normal pattern that each of the coins in C lands heads. And this is absurd: a set of coins cannot land in a normal pattern while landing all heads. Since the aforementioned authors have made a convincing case for Infinitary Closure, there is not much else to be said for Horn I.

However, it is worth pointing out that Horn I is problematic even without the assumption of Infinitary Closure. The problem can be exacerbated by supposing that exactly one of the coins – c' , say – is not fair, but is instead overwhelmingly predisposed to land tails. Suppose again that against all odds, all coins land heads. We are still inclined to judge that c' would still have landed heads if the *other* coins in C had landed in a normal pattern. After all, the coins are all isolated from each other. It may be miraculous that c' landed heads, but Morgenbesser's reasoning suggests that how the other coins landed would not have taken that miracle away.

Now, it seems especially odd to think that if the coins in C had landed in a normal pattern, c' would still have landed heads. This time the coins are not on a par. c' is especially likely to land tails. So if there were infinitely many coins tossed tails, it is surely at least reasonably likely that c' would be among them. That is, we should not think that c' would still land heads if the coins in C landed in a normal pattern. So taking Horn I is not a viable solution to the puzzle.

2.3 Horn II: Heterogeneity

On the second horn, we say that not every coin would have landed heads if C had landed in a normal pattern, and thus that not every coin would have landed heads if the coins in C besides that one had landed in a normal pattern.

Those who choose this horn must explain why we have ended up talking about c , for whom we know the conditionals (1) and (2) are true, even when there are many coins c^* for which analogous conditionals (1*) and (2*) concerning c^* are false. The only somewhat plausible explanation I see is a linguistic one. Somehow or other, in considering c , we become more likely to consider one notion of counterfactual consequence over others. And the notions we consider have a bias in favour of c 's landing how it actually did, but not c^* 's (if we considered c^* to begin with instead of c then things would presumably be reversed).

To investigate this suggestion, we can suppose that there are a range of counterfactual-conditional-like binary operators that one can express by the counterfactual conditional in English. These are, let us suppose, the binary operators $\Box \rightarrow_0$, $\Box \rightarrow_1$, and so on. Supposing that Substitution is to be respected and Horn I of the present dilemma avoided, we will say that for each such $\Box \rightarrow_i$, that the set of coins x for which (the coins in C besides x land in a normal pattern $\Box \rightarrow_i x$ still lands heads) excludes infinitely many members of C . But we can maintain the acceptability of (1) by supposing that the meaning of 'if ... would' in an utterance of (1) in context will always be chosen to be some $\Box \rightarrow_i$ such that (the coins in C land in a normal pattern $\Box \rightarrow_i c$ lands heads) is true.

Although it seems likely that the counterfactual conditional in English is context-sensitive, this strategy only becomes plausible if a story can be told as to why we choose to express conditionals that give our chosen coin c special status. And this is where the strategy starts to face problems. The point of uttering (1) is to express that the way c landed is independent, in some interesting way, from the other coins' landing in a normal pattern. It is not to state that c has a property that was specifically chosen so as to make c have it. There would be no point in predicating such a property of c , and utterances of (1) would not be worth making.

Consider an analogous case: we might wonder which of us is a *self-sufficient* person: someone who would do ok even without other people around. Suppose we determine that Anne is self-sufficient by reflecting on her can-do attitude, and we proclaim:

Anne, unlike many other people, would do ok even without other people around.

It would undermine the point of this whole exercise if, in declaring Anne self-sufficient, we chose to employ a conditional that is somehow biased in

favour of her doing ok were she by herself. If we were in the business of doing that then we would not need to reflect on her can-do attitude. She could be identical for all practical purposes to everyone else and still the proclamation would be true. But if that were so, then the property expressed by ‘would do ok even without other people around’ in the proclamation would be a completely uninteresting one, and it would be hard to justify ever bothering to express it.

The contextualist strategy for pursuing the second horn of the dilemma tells us that the proclamation of (1) is unfairly biased in favour of coin c ’s landing heads on the salient reading of the counterfactual. But since there is no important difference between c and the other coins, the contextualist is left with the awkward task of explaining why we the speakers would bother pick out such a proposition, rather than one which is not biased, or one that favours some coins besides c instead.

There is another important problem. (1) seems plausible, but we could just as easily have argued directly for the truth of

3. Every coin is such that it would have landed heads if the coins besides it had landed in a normal pattern.

(3) can be motivated in exactly the same way (1) was: by reflecting on the independence of each coin from how the others land. If you deny (3), you suggest that you think that the results of some of the coins could be affected by the others. But by Substitution, (3) is equivalent to

4. Every coin is such that it would have landed heads, if the coins in C landed in a normal pattern.

And to accept (4) would be to take Horn I of the dilemma.

This time the appeal to context-sensitivity does not work. For whatever relation $\Box \rightarrow_i$ is expressed by the counterfactual conditional in a given context, if $\Box \rightarrow_i$ satisfies Substitution, then the sentences (3) and (4) are equivalent in that context. Since the arguments against Horn I are not sensitive to context, (4) will be false in that context, so (3) will be false in that context as well.

Since the context was arbitrary, the sentence (3) must be denied in the present context, and once we are in the business of doing so, it doesn’t seem any worse to withhold assent from (1). Taking Horn II, then, is also not a viable solution.

3 Tossing Morgenbesser's Coin

Since (1) and (2) are equivalent and (2) cannot be affirmed, we must also withhold on (1). Despite appearances, we are in no position to accept

1. If C^- had landed in normal pattern, then c would still have landed heads.

This is so even though we know perfectly well that c is causally isolated from all the other coin tosses, never in history has there been any interaction between c and the other coins, that there is no spooky action at a distance, and so on.

This result is surprising enough. Our judgement that (1) is the case is initially quite definite. One might hope, though, that the pathological behaviour of (1) is isolated from the core of the Morgenbesser intuition, that we seem to rely on so heavily in our everyday counterfactual thought. Unfortunately, there are compelling reasons to think that this hope is false.

The problem is that the Morgenbesser intuition is supposed to generalise beyond Morgenbesser's particular case, and once we appreciate two dimensions along which it is supposed to generalise, we find that our initial judgements about (1) and (2) in the Infinitely Many Coins case are in fact central instances of the Morgenbesser intuition.

The first way in which the Morgenbesser intuition generalises is in the exact mechanism of the two events under consideration. Suppose I toss a coin on earth and a radioactive particle flies past Jupiter. The particle decays, and my coin lands heads. The Morgenbesser intuition compels us to say:

5. If I had tossed tails, the particle would still have decayed.
6. If the particle had survived, I would still have tossed heads.

Or suppose that instead of a coin toss or a particle, we are concerned with whether or not an indeterministic computer residing in the Alpha Centauri system, causally isolated from us and the particle, would have printed a YES or a NO. If it did in fact print a NO, then we think:

7. If I had tossed tails (or if the particle had survived), NO would still have been printed.

8. If the computer had printed **YES**, I would still have tossed heads (and the particle would still have decayed).

The second dimension of generalisation concerns the probabilities of the two events. Coins are often presupposed to be fair, but the Morgenbesser intuition does not hang on this. If the coin was really unlikely to land heads, then you would still think ‘woops, I would have won if I bet’ upon seeing it land heads. And the judgements (5) through (8) clearly maintain their plausibility even supposing that the particle was really unlikely to decay, the coin was really unlikely to land heads, or that the computer was really unlikely to print **NO**.

Let me tell you a bit more about the computer. It is indeterministic because it contains an infinite bank of indeterministically-tossable and causally independent coins. When you press ‘print’, it tosses the coins and checks to see whether they landed in a normal pattern. It prints **YES** if they land in a normal pattern and **NO** otherwise. Let C^- be the set of coins in the computer, and let c be the coin I happen to toss light-years away on Earth. Inside the computer and on Earth, all coins land heads by coincidence. So the computer prints **NO** and I toss heads. The Morgenbesser intuition has just commanded us to accept (8): if the computer had printed **YES**, c would still have landed heads. But we now know that the computer would print **YES** if and only if the coins in the computer landed in a normal pattern. So where C^- is the set of coins in the computer, we must conclude:

1. If C^- had landed in a normal pattern, c would still have landed heads.

(1) is therefore an unexceptional case of the Morgenbesser intuition. We who accept Substitution cannot accept (1), so should agree with Slote that Morgenbesser was wrong in the original case.²

References

Bacon, Andrew. Forthcoming. “Counterfactuals, Infinity, and Paradox.” In *Outstanding Contributions to Logic: Kit Fine*, edited by Federico L. G. Faroldi and Frederik van De Putte.

2. For helpful discussion and comments, thanks go to Andrew Bacon, John Hawthorne, Samuel Lee, Weng Kin San, Juhani Yli-Vakkuri, a referee for *Analysis*, as well as participants in Zoë Johnson King’s Dissertation Seminar at USC (Spring 2021) and USC’s Speculative Society.

- Builes, David. 2020. "A Paradox of Evidential Equivalence." *Mind* 129 (513): 113–127.
- Caie, Michael. 2018. "Benardete's Paradox and the Logic of Counterfactuals." *Analysis* 78 (1): 22–34.
- Dorr, Cian, John Hawthorne, and Yoaav Isaacs. 2021. "Solving a Paradox of Evidential Equivalence." *Mind* 130 (520): 1159–82.
- Edgington, Dorothy. 2004. "Counterfactuals and the Benefit of Hindsight." In *Cause and Chance: Causation in an Indeterministic World*, edited by Phil Dowe and Paul Noordhof. Routledge.
- Fine, Kit. 2012. "Counterfactuals Without Possible Worlds." *Journal of Philosophy* 109 (3): 221–246.
- Herzberger, Hans G. 1979. "Counterfactuals and Consistency." *Journal of Philosophy* 76 (2): 83–88.
- Lewis, David. 1973. *Counterfactuals*. Cambridge, MA, USA: Blackwell.
- Pollock, John L. 1976. "The 'Possible Worlds' Analysis of Counterfactuals." *Philosophical Studies* 29 (6): 469–476.
- Slote, Michael A. 1978. "Time in Counterfactuals." *Philosophical Review* 87 (1): 3–27.
- Stalnaker, Robert C. 1968. "A Theory of Conditionals." In *Studies in Logical Theory (American Philosophical Quarterly Monographs 2)*, edited by Nicholas Rescher, 98–112. Oxford: Blackwell.
- Williamson, Timothy. 2020. *Suppose and Tell: The Semantics and Heuristics of Conditionals*. Oxford, England: Oxford University Press.