Morgenbesser cases and closet determinism

Ian Phillips

Sidney Morgenbesser brought to attention cases of the following form:

(MC1) Chump tosses an indeterministic coin and, whilst it is in mid-air, calls heads. The coin lands tails, and Chump loses. His betting was causally independent of the coin's fall. Chump seems right to say: 'If I had bet tails, I would have won.'[1]

(MC2) Champ narrowly misses a plane which subsequently crashes due to unexpected indeterministic events. His missing the plane was causally independent of those events. Champ's relief seems well expressed by the conditional: 'If I had caught that plane, I would probably be dead.'[2]

Such cases highlight counterfactuals that seem intuitively correct and yet whose correctness can only be established on the basis of hindsight. This feature has been taken to have significant implications when it comes to theorizing about counterfactuals. For example, Schaffer (2004: 303) offers (MC1) as a counter-example to Lewis's standard semantics for counterfactuals as complemented by his rules for determining world-similarity. (See Lewis 1973a, 1973b, 1979 and 1986.) And Bennett (2003: §142) cites (MC2) as a decisive reason against extending a suppositional account of the indicative conditional to the subjunctive.

In §1, I set out the way that Morgenbesser cases are employed by Schaffer and Bennett. In §2, I dispute the standard understanding of the cases that such employment relies upon. In order to correctly evaluate such phenomena we must divest ourselves of our deeply entrenched deterministic mindset. Once liberated, it becomes clear that Morgenbesser cases do not pose any obstacle to suppositional accounts of the subjunctive conditional and, though they arguably do still pose an obstacle to Lewis's semantics, it is not quite the one Schaffer suggests. Finally, in §3, I respond to an objection based on the use of counterfactuals in empirical reasoning that Edgington raises against the account of Morgenbesser cases that I offer.

---

[1] This is Morgenbesser's original case as cited in Slote 1978: 27.

[2] This example is taken from Edgington 2003.

1. *The standard use of Morgenbesser cases*

As Schaffer (2004) makes clear, Lewis's semantics for counterfactuals supplemented by his rules for determining world-similarity evaluate the counterfactual in (MC1) as false. Chump merely *might* have lost: the world in which he bet tails and won lies at the same distance from actuality as the world in which he bet tails and lost. Thus, if the counterfactual, 'If I had bet tails, I would have won', *is* correct as Schaffer contends, then Lewis's account must be rejected.

Morgenbesser cases also pose an apparent problem for a natural extension of a suppositional account of indicative conditionals to subjunctives.[3] The basic thought behind such an extension involves claiming that

> confidence in [a] counterfactual expresses the judgment that it was probable that B given A, at a time when A had non-zero probability, even if it no longer does; and even if you do not now have a high degree of belief in B given A. (Edgington 1995: 265)[4]

Thus, crucially, the extension of the suppositional view to subjunctives relies on the following general claim:

> (CT) For any A and B, if A $\square\rightarrow$ B is the right thing to think at a certain time, then at some earlier time A $\rightarrow$ B was the right thing to think. (Bennett 2003: 366)

Bennett rejects (CT) on account of Morgenbesser cases. The reason is that such cases precisely appear to show the existence of counterfactuals that we intuitively regard as correct but for which there seems to be no previously right indicative: 'would haves' without any previously acceptable 'will'. For whilst the coin is in mid-air or the plane on the ground it would be unacceptable to say, 'If I bet tails, I will win,' let alone, 'If I catch the plane, I will die.' Given indeterminism, such claims would be quite groundless.

---

[3] The suppositional account of indicatives holds that a 'conditional judgement involves two propositions, which play different roles. One is the content of a supposition. The other is the content of a judgement made under that supposition. They do not combine to yield a single proposition which is judged to be likely to be true just when the second is judged likely to be true on the supposition of the first' (Edgington 2001, §3.1). Thus, a conditional like, 'If the today is Wednesday, then tomorrow is Thursday', is a conditional assertion of the consequent. If, in fact, it is Wednesday, then this remark is equivalent in force to an assertion of, 'Tomorrow is Thursday.' If today is not Wednesday, no *proposition* is asserted. Conditionals are not true or false. They are evaluable in terms of their conditional probabilities. For further elucidation and defence see Edgington 1995 and 2001. Edgington also defends the extension to subjunctives in her 2003.

[4] See also Adams 1975, Ellis 1984 and Skyrms 1981, 1994.

Edgington (2003: 22) insists that there *is* (in fact) a previously correct indicative even though it would then have been irrational to endorse it. She thinks that with 'the benefit of hindsight' we would judge even a crazed doomsayer *right* if they had said: 'If Champ boards the plane, he will not live.' After all, on hearing about the plane crash Champ might exclaim: 'My God, the doomsayer was right!'

Bennett questions Edgington's ability to make this move. Witness how careful Edgington is to say that the doomsayer 'was right' or 'vindicated' rather than that they spoke the truth. This is because conditionals are not truth-apt on her account. Indicative conditionals instead have objective conditional probabilities. Given this, can we say that the doomsayer was right in so far as Champ's death did (in fact) have an objectively high probability conditional on his catching the plane? No: even in hindsight the plane crash *was then* objectively extremely unlikely. Thus, Bennett complains:

> It is not clear to me what the probabilities are in the light of which the indicatives are judged to be 'right' in hindsight…. [F]or the fortune-teller's conditional to be 'vindicated', room must … be found in the story for a nearly 100 per cent probability of the plane's crashing given that [Champ] was on it. I cannot find … anything allowing us to say that the predictor's conditional probability for the plane's crashing given [Champ's] being on it was, though not 'justified at the time', correct, right, vindicated. (2003: 367)

Edgington's response to this problem is just to say:

> Lucky guesses are sometimes right … The value to be assigned to the hindsightful counterfactual trumps the most rational value to be assigned to the forward-looking indicative. The chance that C given A, beforehand provides the best available opinion on whether C if A, but it can be overturned by subsequent events, not predictable in advance. (2003: 23)

The suggestion seems to be that the objective chance of the crash given Champ's being aboard really was near 100% *despite there being nothing in the world that could have made it so and everything to have made it less than 1%*. Thus, hindsight does not merely reveal what was once the most rational assessment to be incorrect. Hindsight *determines* what was not determined before. As Bennett nicely puts it,

> This is not merely hind-*sight* but hind-*rightness-making* … In cases like this, the idea of the forward indicative's being 'right' *depends on* the idea of subjunctive's being right – the explanatory direction runs from subjunctive to indicative, not the other way. (368)

There are two related worries here. Firstly, according to the suppositional theory, indicative conditionals are to be understood in terms of conditional probabilities. We are to proportion our credence in 'If A, then B' to the subjective conditional probability we give B given A; and the rightness or wrongness of the conditional is to consist in the objective chance of B given A (see Edgington 1995, §8). But, if we take on Edgington's new conception of rightness, we cannot maintain this. The new conception of rightness 'floats free', as Bennett puts it, of the suppositional theory and that cuts us off from our old theoretical explanation of indicative conditionals.

Secondly, the suppositional theory hoped to explain counterfactuals in terms of indicatives. But the above explanation has reversed the order of priorities. It is our independent grasp on subjunctives that seems to determine the rightness or wrongness of the indicatives. Thus, whether or not we hold on to (CT), we remain in need of an independent theory of subjunctives.

Edgington's move cannot save her attractive monistic account of conditionals. However, a re-examination of Morgenbesser cases shows she should never have had to make it.

2. Closet determinism

In his original footnote, Slote made the following comment about (MC1):

> I know of no theory of counterfactuals which can adequately explain why [the counterfactual] statement seems natural and correct. But perhaps it simply *isn't* correct, and the correct retort is 'no, you're wrong: if I had bet (heads), the coin might have come up differently. …' (1978: 27)

Slote's retort has largely been dismissed. Schaffer summarizes the standard reaction:

> [One] might bite the bullet here, as Slote did … [One] might dismiss our intuitions as vestiges of a deterministic mindset. I would agree with Dorothy Edgington that this would be 'wishful thinking'. (2004: 303)[5]

In turn, Edgington supports her claim by considering some of the things we would be forced to say if Slote's response were correct:

> Consider: you are watching [an indeterministic] lottery draw on television and to your dismay your arch … rival wins a prize.… If Slote's 'retort' were correct, so would this be: … 'If I had scratched

<hr />

[5] See also Kvart (1986) and Barker (1999).

my nose a minute ago, he very probably would have lost. What a pity I didn't scratch my nose!' (2003: 17)

I agree: if Slote is right and our intuitions *are* 'vestiges of a deterministic mindset', then this last remark must be correct (even if conversationally infelicitous). I now argue that this is a perfectly palatable bullet to bite.

Two ingredients are essential to run a Morgenbesser case: the causal independence of antecedent and consequent and indeterminism. As Edgington notes:

> [Y]ou have to countenance the possibility of indeterminism for these examples to be a problem.… It seems to me that a decent theory of counterfactuals should cater for that possibility. But for someone who thinks, on something like a priori grounds, that determinism must be true, the standard view is not in trouble: sufficient causes of the plane crash were there back before the fork. (2003: 17–18)

Indeterministic thinking does not come easily to us and our initial judgments often fail us. This is particularly true when our thinking is prejudiced by choice of example. Coin flips and plane crashes are usually perfectly amendable to macroscopic, deterministic treatment and merely saying that they are indeterministic does not remove the bias. To remedy this, deterministic and indeterministic cases must be considered alongside each other.

Consider a three player (Texas Hold'em) poker game using a standard deck of cards. Doyle is dealt the three and five of clubs. Unsurprisingly, he folds. The other players stay in to see the next cards to be dealt from the top of the pack ('the flop'). They are the ace, two and four of clubs. Doyle thinks, surely rightly, 'Damn, if I had stayed in I would have won.'[6]

Now consider the same situation but played out with a random card generator determining the cards. This generator:

(a) Generates cards based on a truly indeterministic physical process. (I assume there are such processes for the purpose of the argument.)

(b) Does not generate a 'shuffled deck' at the beginning of each game. It only generates cards at a fixed time after betting begins such that the time of generation is independent of whether one bets or not.

Once again, Doyle folds pre-flop. The flop comes down with the ace, two and four of clubs. He thinks: 'Damn, if I had stayed in I would have won.'

---

[6] In Texas Hold'em you form the best five card hand using your two personal cards (here Doyle's three and five of clubs) and five shared cards the first three of which are the 'flop'. In the case in point, the flop gives Doyle an unbeatable straight flush.

Is he right to think this in the second situation? Edgington thinks so. Here is one tempting line of thought that supports her judgment:

> Of course it would be right. Nothing causally relevant would have been different if he had bet. And as long as you don't mess with causally relevant factors, then the world will turn out just the same.

Tempting but wrong. For *to assume the causally relevant factors are enough is just to assume determinism*. Fixing all causally relevant factors is – by definition in an indeterministic world – insufficient to fix outcome. Thus, when considering a situation distinct from the actual one – a situation in which Doyle called or I scratched my nose – we cannot assume that in that situation the outcome will be the same *just as long as the differences are causally independent of that outcome*. Fixing the causally relevant factors only justifies the claim that the probabilities will be the same. And that is not enough to support the counterfactual.

A different line of thought is this:

> It wouldn't be right for Doyle to think, 'If I'd stayed in, I would have won.' To say this would be to think of the outcome as revealing some hidden fact about the earlier situation – a fact which would have been the same even if he'd bet. If the process were truly indeterministic, then the outcome reveals no such thing. If the process were truly indeterministic, then, in a different situation – say one in which he had bet – the outcome might well be different.

This piece of reasoning is not only tempting but correct. We can represent the causal possibilities of indeterministic worlds in terms of a branching structure. In the poker case there are four relevant branches: <Bet, Flop = Ac, 2c, 4c>, <Bet, Flop ≠ Ac, 2c, 4c>, <Fold, Flop = Ac, 2c, 4c>, <Fold, Flop ≠ Ac, 2c, 4c>. Just because we *actually* end up on the third branch, that doesn't mean we can claim that the first or third branches *were* inevitable. Put this way, the counterfactual seems quite wrong.[7]

Certainly, Doyle can still think, 'If I'd stayed in *and* those cards had fallen, then I would have won.' Indeed, the fact that in the ordinary deterministic situation these two thoughts are equivalent misleads us into thinking they are in the indeterministic one too. Also misleading is the fact that even in an indeterministic world we would have no *use* for counterfactuals like, 'If I'd scratched my nose, then things would most likely have been different.' But uselessness is no barrier to correctness.

---

[7] Derek Parfit asked me whether a commitment to eternalism would alter this verdict. The interactions between accounts of time and determinism are too complex to discuss here. However, I am strongly inclined to think that our intuitive assessments of counterfactuals are robust with respect to such issues in the metaphysics of time. A commitment to eternalism is not a commitment to fatalism.

If this is right, the Morgenbesser problem disappears. The suppositional theorist who wishes to extend her account to the indicative can simply deny that there is any need to employ hindsight and so avoid any talk of hind-making-right. (CT) is saved without cost.

Lewis may also avail himself of this defence. If the Morgenbesser counterfactual is not correct, then it does not amount to a counter-example to his semantics. The defence is short-lived, however. As Schaffer (*op. cit.*) notes and, indeed, as Lewis (1979) suggests, one can plausibly rig the story in such a way that differences in imperfect future match force his semantics to assess the original counterfactual as true in conflict with the above treatment.[8]

3. *Hindsight and empirical reasoning*

At the end of her 2003 paper, Edgington suggests that we 'can do better than just appeal to intuitions' (17) in rejecting Slote's retort by considering the important role counterfactuals play in empirical reasoning.

> [W]e need, for the empirical inferences we make, not only judgments to the effect that such-and-such *is* (more, or less) likely; but judgments that such-and-such *was* (more, or less) likely, or likely given some-thing else … But if this is what we do, in explaining and drawing inferences from what we see and hear, of course we will use hindsight. (25)

She illustrates this with an example of Jackson's. Millennia ago a volcano erupted and began oozing lava. Valley *A* looked in danger of being submerged. Then an unexpected earthquake occurred and diverted lava flow into valley *B*. A modern day geologist aware of the earthquake reasons: 'There must have been an eruption, because Valley *B* is sub-merged and that's what would have happened if there had been an erup-tion given what I know about the earthquake.'

Edgington suggests that the geologist relies here on a hindsightful coun-terfactual, namely: 'If the volcano had erupted, valley *B* would have been submerged.' She concludes that our empirical employment of counterfac-tuals is hindsightful. I fail to see this. Clearly we *can* evaluate the hind-sightful conditional probabilities relevant to the geologist. They are the conditional probabilities associated with the outcome conditional on an eruption *and an earthquake*. But we can also consider the counterfactual without considering the earthquake: what would have happened condi-tional on an eruption. On that supposition, the mostly likely thing that would have happened (in an indeterministic world) *was* a non-earthquake

---

[8] Schaffer 2004 also contains several distinct counter-examples to Lewis's account which remain problematic.

future. Empirical reasoning, then, is not impugned by ignoring hindsight in indeterministic cases. Of course, human geologists will often employ hindsight to tell them about past sufficient causes – but that is because they assume (for very good reason) a macroscopic determinism.[9]

All Souls College
Oxford OX1 4AL, UK
ian.phillips@all-souls.ox.ac.uk

References

Adams, E. 1975. *The Logic of Conditionals*. Dordrecht: Reidel.

Barker, S. 1999. Counterfactuals, probabilistic counterfactuals and causation. *Mind* 108: 427–69.

Bennett, J. 2003. *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press.

Edgington, D. 1995. On conditionals. *Mind* 104: 235–329.

Edgington, D. 2001. Conditionals. In *The Stanford Encyclopedia of Philosophy* (Fall 2001 Edition), ed. E. Zalta.

Edgington, D. 2003. Counterfactuals and the benefit of hindsight. In *Causation and Counterfactuals*, ed. P. Dowe and P. Noordhof, 12–27. London: Routledge.

Ellis, B. 1984. Two theories of indicative conditionals. *Australasian Journal of Philosophy* 62: 50–66.

Kvart, I. 1986. *A Theory of Counterfactuals*. Indianapolis: Hackett Publishing.

Lewis, D. 1973a. *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, D. 1973b. Counterfactuals and comparative possibility. *Journal of Philosophical Logic* 2: 418–46.

Lewis, D. 1979. Counterfactual dependence and time's arrow. *Noûs* 13: 455–76.

Lewis, D. 1986. Postscripts to 'Counterfactual dependence and time's arrow'. In his Philosophical Papers: Volume II, 52–66. Oxford: Oxford University Press.

Schaffer, J. 2004. Counterfactuals, causal independence and conceptual circularity. *Analysis* 64: 299–309.

Slote, M. 1978. Time in counterfactuals. *Philosophical Review* 87: 3–27.

Skyrms, B. 1981. The prior propensity account of subjunctive conditionals. In *Ifs*, ed. W. Harper, R. Stalnaker and G. Pearce, 259–65. Dordrecht: Reidel.

Skyrms, B. 1994. Adams' conditionals. In *Probability and Conditionals*, eds. E. Eells and B. Skyrms, 13–26. Cambridge: Cambridge University Press.