This is a pre-copyedited, author-produced PDF of an article accepted for publication in *Analysis* following peer review. The definitive publisher-authenticated version (A. Ahmed, 'Out of the Closet', *Analysis* 71: 77-85) is available online at: http://intl-analysis.oxfordjournals.org/content/71/1/77.extract

Out of the Closet

ARIF AHMED

In 'Morgenbesser cases' a non-actual event *A* could have occurred before a causally unrelated and chancy but actual event *B*. The standard example: whilst a fair indeterministic coin is in the air you bet on its landing heads; it lands tails. Here *A* is the possible but non-actual event of your betting on tails instead; *B* is the actual, subsequent and causally unrelated event of its *landing* tails.

Morgenbesser cases matter because they appear to refute an attractively simple theory of counterfactuals: the *suppositional theory*. According to it, the correct credence to give the counterfactual 'If *A* had occurred then *B* would have occurred', where *A* and *B* describe possible and dated events – '*A* > *B*' from now on – is the chance, at some earlier time when *A*'s chance was non-zero, of *B* conditional upon *A*. In Edgington's words:

> 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)

Typically the time in question is taken to be just prior to the supposed time at which *A* begins.[1]

For example: there is a small but non-zero chance at 2pm one day that it will soon rain (*A*); and the chance at 2pm of the match's being cancelled (*B*), conditional upon rain just after 2pm, is 2/3. In fact it doesn't rain that afternoon. But that evening somebody says: 'If it *had* rained just after 2pm, the match would have been cancelled' ($A > B$). The suppositional theory tells us that our confidence in that counterfactual at the time of its utterance ought to be the chance of *B* conditional on *A* at 2pm: in this case, 2/3.

Morgenbesser cases appear to refute the suppositional theory. Consider the standard coin-tossing example described above; let *A* be the event of your betting tails and *B* be the event of the coin's landing tails. The Suppositional Theory implies that our *ex post* credence in:

(1) If you had called tails instead the coin would have landed tails.

should be no more than 50%. For at no point before your actual bet did the coin have a greater than 50% chance of *landing* tails, conditional upon your *betting* tails. But intuition reckons (1) definitely true (or at least very likely), as follows from its taking the same view of:

(2) If you had called tails instead you would have won.

[1] More precisely, it is the time of what Edgington calls the 'fork' i.e. the first supposed deviation from actuality (Edgington 2004: 13). But we can safely ignore this complication.

– which the suppositional theory assigns a probability of no more than 50% for the same reason as (1). This clash between intuition and the suppositional theory looks like a good reason to reject the suppositional theory.

Not so, says Phillips in an intriguing and extremely instructive paper (2007[2]). He claims: *first*, that the problematic intuitions are wrong: what generates them is an illegitimate and tacit assumption of determinism. *Second*, that rejecting intuition, and so counterfactuals like (1) and (2), does not disrupt our purposes in using counterfactuals. This paper will contest both claims.

1. Closet Determinism

Phillips compares two scenarios. The second is a Morgenbesser case.

Scenario A: Doyle is playing Texas Hold'em. He receives the three and five of clubs and folds. The next three cards to be dealt (the 'flop') are the ace, two and four of clubs. Doyle thinks (ruefully):

(3) If he had not folded he would have won.

– since in that case he would have had a straight flush.

Scenario B. As in *A* except that a random card-generator produces the flop on a truly *indeterministic* basis (not just a deterministic but unknown one, as in *A*). And it does so after Doyle has folded. Seeing the ace, two and four of clubs, Doyle thinks (3).

[2] Except where stated all citations are to this paper.

Everyone agrees that in scenario *A* – where genuine chance plays no role – Doyle was right to think (3). Now intuition endorses (3) in scenario *B* as well, because it endorses:

(4) If he had not folded the flop would have been what it actually was.

But Phillips thinks that the intuition in favour of (3) and (4) is a 'vestige of a deterministic mindset' (42, quoting from Schaffer 2004); and of course *that* basis for them is absent from scenario *B*.

But *how* does the intuition behind (3) and (4) rely upon determinism? Phillips suggests just one line of thought connecting determinism with them, so this must be what he has in mind (47). Here it is, written out in a form that distinguishes premisses from conclusion.

(5) Nothing causally relevant to the flop would have been different if he had bet (premiss).

(6) 'As long as you don't mess with causally relevant factors, then the world will turn out just the same' (47) (premiss).

(7) Therefore (4) is true.

It is clear how the assumption of determinism supports premiss (6). 'For to assume the causally relevant factors are enough is just to assume determinism' (47). But in scenario B both determinism and the premiss that rests upon it are false: 'Fixing all causally relevant factors is – by definition in an indeterministic world – insufficient to fix outcome' (47).

But it seems to me that the intuition behind (4) rests upon quite another thought, one that does *not* assume determinism and which therefore applies to scenario *B* just as well as to scenario *A*. It runs like this. 'The only actual events that would have occurred differently (or not at all) if Doyle had not folded are those to which Doyle's folding *makes* a difference. And it *makes* a difference only to those events to which it is causally relevant.[3] But his folding is not causally relevant to the flop. Hence the flop would have occurred just as it actually did if Doyle had bet; hence (4) is true – and so is (3).'

We can see more clearly that *this* argument does not rely upon determinism (and hence also that it applies to scenario *B*) if we write it out as follows:

(8) If *C* makes no difference to an actual event *E* then *E* would still have occurred even if *C* had not (premiss).

(9) *C* makes no difference to any actual events to which it is causally irrelevant (premiss).

(10) Doyle's folding is causally irrelevant to the flop (premiss).

(11) Therefore (4) is true.

Each of its premisses, I claim, is both intuitive and independent of determinism. So contra Phillips, there *is* an intuitive route to (3) and (4) that doesn't presuppose determinism.

[3] It does no harm in this case to understand causal relevance in terms of conditional chance i.e. *A* is causally relevant to *B* iff $\mathrm{Ch}\,(B \mid A) \neq \mathrm{Ch}\,(B)$ at the time of *A*'s occurrence. However this won't work in general: if *A* has an effect upon *B* via e.g. two distinct and *antagonistic* causal paths, it may turn out that $\mathrm{Ch}\,(B \mid A) = \mathrm{Ch}\,(B)$ at the time of *A*, though intuitively *A* is of causal relevance to *B*. But even if we take causal relevance to be primitive the argument in the text still stands: all that *it* requires is that we can understand *causal* relevance independently of *counterfactual* dependence.

First objection: ‘This argument begs the question. Nobody would accept premiss (10) unless he *already* granted that if Doyle had folded then the flop would have been what it actually was, and hence also tacitly granted (11) all along.’

That would be a mistake. It *is* possible to settle questions of *causal* relevance between events – as in (10) – without first settling their *counterfactual* dependence – as in (11). For instance, one can know that the assassin’s shot was *causally* relevant to the victim’s death without knowing whether the assassin had a backup (and so without knowing whether the death was *counterfactually* dependent upon the shot).

Second objection: ‘It’s true that none of the premisses(8)-(10) *individually* presupposes determinism. But taken together (8) and (9) *do* presuppose – that is, entail – it. For taken together they entail:

> (12) If *C* is causally irrelevant to an actual event *E* then *E* would have occurred even if *C* had not.

And (12) is equivalent to saying that “as long as you don’t mess with causally relevant factors” the outcome will be the same i.e. to determinism.’

But (12) is *not* equivalent to determinism. Phillips’s loose formulation of that thesis – ‘as long as you don’t mess with causally relevant factors the world will turn out the same’ – is ambiguous between two claims, depending on whether we render it as a strict or as a subjunctive conditional. One (‘strict’) reading of it is:

> (13) For any post-*C* event *E*, *every* world that agrees with actuality over (i) the laws (ii) the causal factors for *E*, will agree with actuality over *E*.

The other ('subjunctive') reading is:

> (14) For any post-*C* event *E*, the *closest* ~*C*-worlds that agree with actuality over (i) the laws (ii) the causal factors for *E*, will agree with actuality over *E*.

(12) entails (14) – given the uncontentious premiss that all the closest ~*C*-worlds are legal after the time of *C* – but *not* (13), which is stronger than (14). But it is (13) and not (14) that captures determinism – at any rate the consequence of determinism that we are assuming to be violated in Morgenbesser cases. So I think that (12) does *not* entail determinism, since it does not entail (13). Nor therefore do (8) and (9) presuppose it. Nor therefore does the route that I have sketched, from our intuitions about Morgenbesser cases to the application of such conditionals as (3) and (4) thereto, and thence to the denial of the suppositional theory.

2. The Benefit of Hindsight

But it is not only intuition that prefers (3) and (4) to their denials. Edgington has argued (2004) that if they really did conform to the sort of theory that denies (3) and (4) – as we are supposing that the suppositional theory does –, counterfactuals could not fill the role in empirical reasoning about the actual past that constitutes much of their point.

Into what sorts of empirical reasoning about the actual past do counterfactuals enter? There are two. In this illustration of the first I emphasize the operative counterfactual:

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 submerged and *that's what would have happened if there had been an eruption* given what I know about the earthquake.' (48)

I call this sort of reasoning *counterfactual abduction*. An argument is a counterfactual abduction if and only if it instantiates the following schema:

(15) Q, $P > Q$, therefore P

– where '>' is the counterfactual conditional. Clearly not everything that instantiates this schema is valid; still such an argument's premisses will invariably transmit *some* support to its conclusion.[4] The volcano example instantiates (15) with P = 'There was an eruption' and Q = 'Valley *B* is submerged'.

Here is an illustration of the second sort of reasoning (emphasis as before):

The grass was dry this morning. *If it had rained last night then the grass would have been wet this morning.* Therefore, it didn't rain last night.

This is called *counterfactual* modus tollens (Edgington 2004: 23). An argument is a counterfactual modus tollens if and only if it instantiates the following schema:

[4] In general the *degree* of confirmation that they give depends—in a manner that Edgington (2004: 24) makes clear—upon the relative likelihood of ($P > Q$) and ($\sim P > Q$)

(16) Q, $P > {\sim}Q$, therefore ${\sim}P$

Any instance of (16) *is* valid: if its premisses are true then its conclusion is true; moreover the conclusion is at least as likely as the conjunction of the premisses[5]. The grass/rain example instantiates (16) with P = 'It rained last night' and Q = 'The grass was dry this morning.'

Edgington argues that counterfactuals such as (4) are empirically useful for reasoning about Morgenbesser cases because of their role in the *first* sort of reasoning i.e. counterfactual abduction. Go back to the volcano/earthquake argument. The counterfactual premiss in that argument was

(17) If there had been an eruption then valley *B* would have been submerged

– and, just as (2) relies upon (1) and (3) upon (4), so too (17) relies upon:

(18) If there had been an eruption then the earthquake would (still) have occurred.

[5] This follows from the principle that $P > {\sim}Q$ entails the corresponding material conditional $P \supset {\sim}Q$, which is itself derivable via modus ponens for >. Edgington nowhere disputes that principle, but she *does* say that an argument of the form of (16) may be probabilistically *in*valid; and hence that the conjunctions of its premises may be more likely than its conclusion. This will happen, she says, if ${\sim}P > {\sim}Q$ is sufficiently probable (Edgington 2004: 24).

I think that is false. Even assuming (on *any* interpretation of probability) that the probability of Q is fixed and high, if the probability of ${\sim}P > {\sim}Q$ increases the probability of $P > {\sim}Q$ will *fall*; at no point will the probability of $(P > {\sim}Q)$ & Q exceed that of ${\sim}P$, on pain of incoherently assigning a non-zero probability to the inconsistent conjunction of P, Q and $P > {\sim}Q$.

For instance, if I am certain that the grass would have been wet if it *hadn't* rained (because there is a sprinkler that runs automatically in dry conditions), and if I am certain that the grass is dry, I will suppose that something must have stopped the rain from wetting the grass (say, an artificial cover). Hence my confidence *that if it had rained the grass would have been wet* will fall to zero. Therefore it will not exceed my confidence that it didn't rain last night. Contra Edgington, a high probability for some ${\sim}P > {\sim}Q$ does not defeat the probabilistic *validity* of the corresponding instance of (16) but only reduces the joint probability of its *premisses*.

Suppose now that the earthquake was a causally undetermined actual event whose chance of occurrence at any time *prior* to its occurrence was very low, say 1% (and let us now reserve the term 'chancy' for just such events). If (18) is false on that supposition then so is (17). But if (17) and (18) *are* false then we cannot use any counterfactual abduction that relies upon them, which defeats Edgington's geologist. But that is wrong: the geologist's empirical reasoning was both sensible and useful as long as the earthquake was *actual*, whether or not it was also *inevitable*.

Phillips replies that dropping (17) and (18) in the indeterministic case poses no problems for the geologist. Instead of asking what would have happened if there had been an eruption, she can ask what would have happened if there had been an eruption *and* an earthquake. That is, the counterfactual premiss of her abduction builds all her relevant knowledge into its antecedent. So that premiss is not (17) but:

> (19) If there had been an eruption *and the earthquake had (still) occurred*, valley B would have been submerged. (48) [6]

For the suppositional theory allows that (19) is true even if the earthquake is chancy, whereas it reckons both (17) and (18) false upon that hypothesis. Hence even a geologist whose theory of counterfactuals rules out (1)-(4) (and (17) and (18)) *can* still reason by counterfactual abduction about Morgenbesser cases. In the present case her knowledge, that (19) is true and valley *B* is submerged, transmits some support to

[6] This is the third to last sentence on that page. Two points of interpretation: first, he actually writes of the *conditional probability* of the submergence upon the eruption and earthquake. But that he equates conditional probabilities—whose introduction here is otherwise irrelevant—with probabilities of the corresponding conditionals is implicit in his next sentence. Second: he concludes the paper by saying that actual geologists often 'assume… a macroscopic determinism' (49). Maybe they do make that assumption *sometimes*, but it is quite unnecessary *here*. If the application of (19) to the present case is meant to *contrast* with the non-application of (17)—and if not then what is its point?—then Phillips must intend that (19) be available to the geologist *whether or not* the earthquake was causally determined.

the conclusion that there was an eruption *and* an earthquake, because that argument instantiates (15). Hence it supports the desired corollary that there was an earthquake.

That meets the objection that *one* kind of empirical reasoning requires such counterfactuals as (1)-(4), (17) and (18). But I mentioned *two* kinds of empirical reasoning using counterfactuals: counterfactual abduction and counterfactual modus tollens. Phillips only considers counterfactual abduction. But counterfactual modus tollens creates the greater difficulty for the suppositional theory and its handling of Morgenbesser cases.

Consider that if E is chancy and C is a possible but non-actual event that would have been both prior and causally irrelevant to E, then the suppositional theory, which gives a low probability to $C > E$, will give a high one to its opposite $C > \sim E$. For instance, at no time prior to its occurrence did a chancy earthquake have a chance of more than 1% of occurring. So according to the suppositional theory our confidence in

(20) If the eruption had occurred the earthquake would *not* have occurred.

should be at least 99%.[7]

Suppose also that we know that the earthquake *did* occur. Then the following also deserves a confidence of 99%:

[7] We must sharply distinguish (i) the claim that we should *now* assign 99% *credence* to (20) and (ii) the claim that if the eruption had occurred then the *chance* of no-earthquake *then* would have been at least 99%. (ii) is plausible—at least none of my arguments have any bearing upon it—but unlike (i) it has nothing to do with the theory that Phillips sought to defend. That theory is not about what probability-assigning counterfactuals we should believe; it is about what probabilities we should assign to 'straight' counterfactuals (see e.g. Edgington 2004: 15-16).

(21) The earthquake occurred, and if the eruption had occurred the earthquake would *not* have occurred.

But from (21) and the validity of (16) it follows that *there was no eruption*. So the proposition that there was no eruption must inherit via (16) at least the probability of (21) i.e. at least 99%. This is obviously wrong: the whole point of the example was to describe fairly good evidence *for* an eruption! So on the suppositional theory, counterfactual modus tollens is empirically *misleading*. Indeed we could use modus tollens to argue against the actual occurrence of *any* possible event that (would have) preceded *any* unlikely and causally unrelated actual one.

Even worse, the theory presumably gives a 99% probability not only to (21) but also to this:

(22) The earthquake occurred, and if the eruption had *not* occurred the earthquake would not have occurred.

(since the earthquake is causally independent of the eruption). And by the same reasoning our confidence that there *was* an eruption should *also* be at least 99%.

This absurdity is a *reductio* of counterfactual modus tollens in indeterministic settings. At least it is if you want to insist upon the suppositional theory. But if you find counterfactual modus tollens useful in empirical reasoning about such settings then you had better drop that theory.

Nothing like Phillips's reply to the objection from counterfactual abduction is available to him here. Suppose we insist that counterfactual modus tollens arguments should always proceed from conditionals that build our knowledge of actuality into their antecedents (on the model of (19)). In that case the disastrous consequences of (21) and (22) are blocked; instead we get the anodyne consequences of:

> (23) The earthquake occurred, and if the eruption had occurred *and the earthquake had (still) occurred* then the earthquake would have occurred.

But the conditional conjunct of (23) is tautological. So (23) tells us no more than what is inferrable from its non-conditional conjunct.

More generally, counterfactuals that build our knowledge of actuality into their antecedents are useless for any counterfactual modus tollens whose non-conditional premiss is known. From my knowledge that the grass was dry this morning and of what would have happened if it had rained last night, I can infer something about last night’s weather. But from my knowledge that the grass was dry this morning and of what would have happened if it had rained last night *and the grass had been dry this morning*, I can infer nothing at all about last night’s weather.

So the only alternative to dropping the suppositional theory is the abandonment of (16) in indeterministic settings. But that would certainly make counterfactual expressions less useful than they presently are; for the fact is that we *do* use instances of (16) with confidence, and without risk, in deterministic *and* indeterministic settings. So setting aside the desire—which I do not share—that our use of ‘>’ meet the tidy constraints of the suppositional theory, it is hard to see what could possibly motivate it.[8]

University of Cambridge

[8] Thanks to the Leverhulme Trust for its support during the writing of this paper.

Sidgwick Avenue

Cambridge CB3 9DA

ama24@cam.ac.uk

References

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

———. 2004. Counterfactuals and the benefit of hindsight. In *Causation and Counterfactuals*, ed. P. Dowe and P. Noordhof, 12-27. London: Routledge.

Phillips, I. 2007. Morgenbesser cases and closet determinism. *Analysis* 67: 42-49.

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