This is the accepted version of the following article: ‘Walters on Conjunction Conditionalization’, Proceedings of the Aristotelian Society 111: 115-122 (2011), which has been published in final form at:

<http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9264.2011.00301.x/abstract;jsessionid=8F680847C1303D210A9F42FE75695231.f01t02>

***Walters on Conjunction Conditionalization***

*Abstract*

This discussion note examines a recent argument for the principle that any counterfactual with true components is itself true. That argument rests upon two widely accepted principles of counterfactual logic to which the paper presents counterexamples. The conclusion speculates briefly upon the wider lessons that philosophers should draw from these examples for the semantics of counterfactuals.

1: Introduction

The present paper concerns an argument for Conjunction Conditionalization i.e. the schema:

1. (X ∧ Y) ⊃ (X > Y)

—where ‘>’ denotes the subjunctive or counterfactual conditional and legitimate replacements for ‘X’ and ‘Y’ report the occurrence of actual dated events.

In a very impressive paper Lee Walters presents an argument for Conjunction Conditionalization as applied to the so-called ‘Morgenbesser’, a fictitious scenario of which the following four features are characteristic:

(Toss) I toss a coin

(~ Bet) You decline to bet on its landing heads

(Heads) It lands heads; so

(~Win) You do not win the offered bet.

Two more things are clearly true of the Morgenbesser: first, that the coin was tossed *and* it landed heads; and second, that if you *had* bet on its landing heads then you *would* have won that bet.[[1]](#footnote-1) So in the Morgenbesser we have:

1. Toss ∧ Heads
2. Bet > Win

And Walters’ argument for Conjunction Conditionalization in the Morgenbesser—his ‘Argument 1’ (Walters 2009: 370)—proceeds from (3) to:

1. Toss > Heads

The argument thus shows that the Morgenbesser is not (as some have claimed) a counterexample to Conjunction Conditionalization: at least not to the instance of it that we get by substituting ‘Toss’ and ‘Heads’ in (1) for ‘X’ and ‘Y’ respectively.[[2]](#footnote-2)

This doesn’t by itself achieve Walters’ broader aim of thoroughly vindicating Conjunction Conditionalization, that is, of establishing the truth of every instance of (1), or at least of every one that could possibly matter. To this further end Walters offers ‘Argument 3’ on p. 377, as follows.

Let ‘A’ and ‘C’ report or denote *any* two actual dated events. Let ‘B’ report or denote some non-actual event whose occurrence is counterfactually irrelevant to A and C, so that B > A and B > C. (I shall leave it to context and the reader’s charity to settle when they are reporting and when they are denoting.) Then Walters argues as follows:

*Argument 3*

1. B > C premise
2. (B > C) ⊃ ((B ∧ A) > C) premise
3. ((B ∧ A) > C) ⊃ (A > C) premise
4. A > C from (5)-(7)

Of course the Morgenbesser itself illustrates Argument 3 with A = Toss, B = Bet and C = Heads: if you had bet then the coin would [still] have landed heads; if this is so then if you had bet and the coin had been tossed, it would still have landed heads; if *that* is so then if the coin had been tossed it would have landed heads; hence if the coin had been tossed it would have landed heads.

But Walters’ point is that for *any* actual A and C, as long as we can find a B that is counterfactually irrelevant to both, we can be sure that the corresponding instance of Conjunction Conditionalization holds (i.e. that in which ‘A’ replaces ‘X’ and ‘C’ replaces ‘Y’ in schema (1)). Since it is plausible that for any A and C that anybody might reasonably care about we *can* find just such a B,[[3]](#footnote-3) it is also plausible that every instance of Conjunction Conditionalization that anyone might care about is true.

But there are reasons to resist Argument 3.

2: The case for (6)

They all turn on premise (6), which says that if C would have occurred if B had, then C would have occurred if B *and* A had. Walters gives two arguments for this principle.

The first argument is as follows:

1. B > (B ∧ A) premise
2. ((B > (B ∧ A)) ∧ ((B ∧ A) > B) ∧ (B > C)) ⊃ ((B ∧ A) > C)

premise

1. (B > C) ⊃ ((B ∧ A) > C) from (9), (10)

The conclusion (6) follows from the premises (9) and (10) given the trivial (B ∧ A) > B. Premise (9) is a straightforward consequence of B > B and B > A. Premise (10) is an instance of *Substitution* i.e. the schema:

1. ((X > Y) ∧ (Y > X) ∧ (X > Z)) ⊃ (Y > Z)

Every instance of (11) is a theorem of three standard counterfactual logics: Lewis’s VC, Pollock’s SS and Stalnaker’s C2.[[4]](#footnote-4)

The second argument is as follows:

1. (B > A) premise
2. ((B > A) ∧ (B > C)) ⊃ ((B ∧ A) > C) premise
3. (B > C) ⊃ (B ∧ A) > C from (12), (13)

The conclusion (6) follows from (12) and (13). Premise (12) is a straightforward consequence of B’s definition. Premise (13) is an instance of *Very Limited Antecedent Strengthening* or *VLAS* i.e. the schema:

1. ((X > Y) ∧ (X > Z)) ⊃ ((X ∧ Y) > Z)

Every instance of (14) is a theorem of VC, SS and C2.

3: The case against (6)

Walters’ two arguments for (6) therefore rely upon Substitution and Very Limited Antecedent Strengthening, that is upon (11) and (14), respectively. The following story furnishes counterexamples to both.

A and B are assassins; C, D and E are prominent statesmen of the northern Republic (formerly Kingdom) of Zembla. A has orders to shoot C and D. She is an excellent shot and very determined. So unless the police catch her before she reaches C nothing will stop her from shooting C; once she has shot C nothing will stop her from shooting D unless somebody else has got to him first. B has orders to shoot D and E. He too is an excellent shot and very determined. So unless the police catch him before he reaches D nothing will stop him from shooting D; once he has shot D nothing will stop him from shooting E. As it happens neither A nor B gets to shoot anyone because independent policemen catch them both before either gets an opportunity. A is caught trying to enter Zembla from the East and B is caught trying to enter Zembla from the West. The authorities imprison both. There is no other threat to the lives of C, D or E.

*3.1 Substitution*. Clearly if somebody had shot C then it would have been A: A was the only person who *could* have shot him. So if somebody had shot C then A would have gone on to shoot D. So if somebody had shot C had then somebody would have shot C and somebody would have shot D. Let us write ‘C’, ‘D’ and ‘E’ for the propositions that somebody shoots C, that somebody shoots D and that somebody shoots E respectively. (So ‘C’, ‘D’ and ‘E’ abbreviate sentences as well as denoting imaginary persons. Once again, context should make their meaning quite obvious in any particular case.) Then we have:

1. C > (C ∧ D)

On the other hand it is not true that if *somebody* had shot C and *somebody* had shot D then it would have been A who shot *both*. It might have been A who shot C and *B* who shot D. In that case, B would have gone on to shoot E. So it is not true that if somebody had shot C and somebody had shot D then E *would* not have been shot, for E would only have survived if A had been the one to shoot *both* C and D. So we have:

1. ~ ((C ∧ D) > ~E)

Perhaps you will object to (16) that nearby worlds at which A shoots both of C and D involve just one causally unprecedented departure from actuality, namely A’s evasion of capture; whereas nearby worlds at which A shoots C and B shoots D involve *two* such departures, namely A’s evasion of capture *and* B’s evasion of capture.[[5]](#footnote-5) And when we are trying to establish what would have happened if somebody had shot C and somebody had shot D, we ought only to consider worlds at which those events come about through a *minimal* departure from actuality. At all such worlds only A evades capture [the objector continues], because close worlds at which only A evades capture involve *fewer* causally unprecedented departures from actuality than any at which B also evades capture. Hence at all such worlds A shoots both C and D and B remains in prison; it follows that (16) is false after all.

But the principle, that when asking what would have occurred if some non-actual event F occurs, we should only consider F-worlds that depart minimally from actuality—that principle is false if ‘minimal’ is so defined that worlds involving *fewer* causally unprecedented departures from actuality *ipso facto* count as involving a more minimal departure from actuality than cases involving *more* departures. We can see this from the following example.

A slightly unorthodox lottery pays a prize of £10*n* if you match the *lowest* *n* out of the 6 numbers on your ticket with those that come out in the draw, *n* > 0. (So e.g. if the numbers drawn are 4, 12, 21, 32, 48, 59 then you win $1000 if you have 4, 12, 21, 34, 44, 52 on your ticket; you win nothing if you have 5, 12, 21, 32, 48, 59 on your ticket). One purchases a ticket by paying £1 for a slip with numbers from 1-60 and marking 6 numbers on the slip; the *m*th number must exceed 10(*m* – 1) but not 10*m*. You buy a ticket every week and you choose each number independently of any other. On this occasion you don’t match *any* of the numbers and so win nothing. You console yourself with the reflection that:

1. If I had won anything it would only have been £10.

Plainly (17) is false, because it is also false that if I had matched *at least* one number then I’d have matched *just* one number.

But if worlds involving *fewer* causally unprecedented departures from actuality *ipso facto* count as involving a more *minimal* departure from actuality than worlds involving *more* departures, then the principle, that when asking what would have occurred if some non-actual event F occurs, we should only consider F-worlds that depart minimally from actuality—that principle makes (17) *true*. For worlds at which you win a prize because you have matched *just* one number involve *fewer* causally unprecedented departures from actuality (in your chosen numbers) than worlds at which you win a prize because you have matched *more* than one number.[[6]](#footnote-6) So the objection to (16) fails because the principle behind it does. And (16) stands.

Now it is trivial that if somebody had shot C and somebody had shot D then somebody would have shot C:

1. (C ∧ D) > C

Finally, recall that B was in prison at the only time that A could have shot C, and nobody else *could* have shot C. It follows that if somebody had shot C then nobody would have shot E, that is:

1. C > ~E

For only B *could* have shot E, and we have just seen that even if somebody had shot C, B would (still) have been in prison and so not able to shoot anyone.

But now if we let X = C, Y = C ∧ D and Z = ~E then we have a counterexample to the Substitution schema (11). For (15), (18) and (19) conjoin to give the antecedent of an instance of (11) whilst (16) negates that instance’s consequent. Hence that instance of the Substitution schema is false; so we should reject it (and also Lewis’s VC, Pollock’s SS, and Stalnaker’s C2). This undercuts Walters’ first argument for (6).

*3.2 Very Limited Antecedent Strengthening*. If somebody had shot C then it would have been A: A was the only person who *could* have shot him. So if somebody had shot C then A would have gone on to shoot D. So if somebody had shot C then somebody would have shot D:

1. C > D

But now if we let X = C, Y = D and Z = ~E then (19) and (20) conjoin to give the antecedent of an instance of the VLAS schema (14) whilst (16) negates that instance’s consequent. Hence that instance of the VLAS schema is false; so we should reject it (and also Lewis’s VC, Pollock’s SS and Stalnaker’s C2). This undercuts Walters’ second argument for (6).

4: Conclusion

The challenge for Walters is therefore to supply counterfactual principles that are severally more obvious than (6) and which jointly entail it. I am not at all confident that this cannot be done. I am more confident of something else that the counterexample shows: namely, that what I called the ‘standard’ counterfactual logics VC, SS and C2 are unfaithful to the intuitions of intelligent and philosophically lay speakers of English. They must go.

And if they do go, then it is hard to see how semantics for counterfactuals could be based upon an absolute ordering of possible worlds. I think that we should prefer some such antecedent-relative ordering as that implicit in Edgington 2004.[[7]](#footnote-7) But that is a matter for another paper.[[8]](#footnote-8)

***References***

Ahmed, A. Out of the Closet. *Analysis* forthcoming.

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

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

Phillips, I. 2007. Morgenbesser cases and closet determinism. *Analysis* 67.1: 42-9.

Pollock, J. 1976. *Subjunctive Reasoning*. Dordrecht: D. Reidel.

Stalnaker, R. 1981. A theory of conditionals. In Harper, W., R. Stalnaker and G. Pearce (eds), *Ifs: Conditionals, Beliefs, Decision, Chance and Time*. Dordrecht: D. Reidel: 41-56.

Walters, L. 2009. Morgenbesser’s coin and counterfactuals with true components. *Proceedings of the Aristotelian Society* cix: 365-79.

1. Ian Phillips has in fact objected to the second claim in his 2007. I defend it in my forthcoming. [↑](#footnote-ref-1)
2. Walters’ basic argument for (4) takes the form of the argument that I criticize below. But I am not going to contest (4). The premises that support *it* are themselves derivable from assumptions (such as VRSDA: Walters 2009: 372) that lie beyond the scope of my arguments in this paper. [↑](#footnote-ref-2)
3. For a counterexample that nobody cares about: let C hold only at the actual world. [↑](#footnote-ref-3)
4. Lewis 1973: 132, Pollock 1976: 42-3, Stalnaker 1981: 48. [↑](#footnote-ref-4)
5. I am calling an event *e* a ‘causally unprecedented departure from actuality’ at a world *w* if (i) *e* occurs at *w* but *not* at the actual world; (ii) all of *e*’s causal antecedents at *w* occur at the actual world. (So if actuality is deterministic then any causally unprecedented departure from it is miraculous.) [↑](#footnote-ref-5)
6. This example also works against the principle that if every causally unprecedented departure from actuality at world *w*1 is also a causally unprecedented departure from actuality at *w*2 but not vice versa, then *w*1 involves a more minimal departure from actuality than does *w*2. For every causally unprecedented departure from actuality at any nearby world at which you win a prize of just £10 is a causally unprecedented departure from actuality at every nearby world which you win a prize of *more* than £10, but not vice versa. So if ‘minimal departure’ is going to mean anything relevant to the evaluation of counterfactuals then that principle entails (17) too; but (17) is false. [↑](#footnote-ref-6)
7. See especially Edgington 2004: 17. Those who are familiar with that paper will note that (16) does *not* play the role in empirical confirmation and disconfirmation that Edgington considers essential to counterfactuals (2004: 24): ultimately this is because the information that C and D have both been shot does not confirm the hypothesis that B has escaped any more than the information that C has been shot. In my view this fact does not tell against *either* (16) *or* Edgington’s own semantics for counterfactuals. What it *does* tell against is the idea that a counterfactual conforming to *that* semantics is thereby well fitted for the epistemic role that Edgington emphasizes. But that is a matter for another paper. [↑](#footnote-ref-7)
8. I am grateful to Lee Walters and to a referee for the *Proceedings* for their very helpful comments. [↑](#footnote-ref-8)