

PROCEEDINGS OF THE ARISTOTELIAN SOCIETY

ISSUE 1 | VOLUME CXIII | 2012 - 2013

'Actual Instead'

SARAH BROADIE
UNIVERSITY OF ST. ANDREWS

DRAFT
PAPER

1 8 8 8 | C E L E B R A T I N G 1 2 5 Y E A R S | 2 0 1 3

PROCEEDINGS OF THE ARISTOTELIAN SOCIETY
134TH SESSION

ISSUE NO. I
VOLUME CXIII
2012 - 2013

PRESIDENTIAL ADDRESS:
'ACTUAL INSTEAD'

SARAH BROADIE
UNIVERSITY OF ST. ANDREWS

MONDAY, 8 OCTOBER 2012

17.30 - 19.15

THE CHANCELLOR'S HALL
SENATE HOUSE
UNIVERSITY OF LONDON
MALET STREET
LONDON WC1E 7HU
UNITED KINGDOM

*Followed by a wine reception
in the Grand Lobby of Chancellor's Hall*

CONTACT

mail@aristoteliansociety.org.uk
www.aristoteliansociety.org.uk

BIOGRAPHY

Sarah Broadie is in the Moral Philosophy Department at the University of St Andrews. She has previously worked in philosophy departments at Princeton, Rutgers, Yale, the University of Texas at Austin, and Edinburgh University. Her publications include *Nature, Change, and Agency in Aristotle's Physics* (1982); *Passage and Possibility: A Study of Aristotle's Modal Concepts* (1982); *Ethics with Aristotle* (1991); *Aristotle, Nicomachean Ethics: Commentary*, with translation by Christopher Rowe (2002); *Aristotle and Beyond, Essays on Metaphysics and Ethics* (2007); *Nature and Divinity in Plato's Timaeus* (2011); *Philoponus on Aristotle, Physics 4. 10-14 (on Time)*, translation and notes (2012). She is a Fellow of the British Academy, the Royal Society of Edinburgh, the American Academy of Arts and Sciences, and the Academia Europaea.

This year's Presidential Address marks the inauguration of Sarah Broadie as the 105th President of the Aristotelian Society.

EDITORIAL NOTE

The following paper is a draft version that can only be cited with the author's permission. The final paper will be published in *Proceedings of the Aristotelian Society*, Issue No. 1, Volume CXIII (2013). Please visit the Society's website for subscription information: www.aristoteliansociety.org.uk.

It is argued that acceptance of determinism sits badly with the way we use counterfactual conditionals when considering gains and losses in light of how things would have been if such and such had or had not happened; it is further suggested that one type of indeterminism runs into the same difficulty; also that the difficulty may escape notice through failure to distinguish different uses of counterfactuals.

I. PRELIMINARIES

The idea of determinism continues to exert a grip on our philosophical imagination; or so one would think, given the unremitting stream of publications on compatibilism. Compatibilism, according to current usage, is the thesis that moral responsibility and free will, or one or other of these if they do not necessarily hang together, are not excluded by determinism. Defenders of compatibilism and their opponents would presumably not debate the question unless they assumed that determinism, anyway at the outset of discussion, deserves to be taken somewhat seriously. Obviously this applies to whoever enters the lists already convinced, or inclined to be convinced, of determinism, just as it does to whoever takes up cudgels against determinism on the ground that it rules out responsibility or free will. But it equally applies to interested agnostics who see it as important to show that responsibility and free will are not at risk even if determinism is true.

This paper will oppose a kind of compatibilism, but not the well-known kind. The target (at least to begin with) is the assumption that determinism poses no threat to a common and quite possibly indispensable use of counterfactual conditionals. *This* compatibilist assumption is logically more basic than the debate about determinism *vis à vis* free will and moral responsibility. Freedom and responsibility relate to actions, and action or agency is normally a matter of bringing it about that things are different from how they would have been if the agent had not so acted.¹ Thus our understanding of action rests on counterfactual conditionals. The point holds even if one maintains, for whatever reason, that the human agent’s direct input is confined to choosing or willing or trying to do so and so, this being the cause of those overt events in which we normally locate the action. For we still have causation of a sequence that differs in specific ways from what would have come about otherwise, i.e. if the agent had not so chosen.

I shall argue that an ordinary practice of ours with counterfactuals assumes in effect that determinism is not a real epistemic option. The thesis is not that this practice entails the falsehood of determinism, but, more weakly, that it excludes treating it as a sensible or reasonable position, one that we have to consider as a serious contender. More fully, the proposal is that so far as we regard ourselves as reasonable in forming certain counterfactual judgments, to that extent we treat determinism as out of the

¹ For the present purpose I set aside questions of over-determination, pre-emption, masking, and finkishness; I treat relevant omissions as actions; and I also assume that choices etc. are generally efficacious as such, regardless of whether they ultimately are or consist in neurological events.

question or not up for rational consideration.²

This topic falls within territory where deep and highly technical expertise holds sway: the study of conditionals, and of how counterfactuals relate to causal statements. While I have learned from some of this work, I remain an amateur in that direction. Yes, it does take a fool to rush into precincts where angels fear to tread; but for any proverb there is usually another that preaches the opposite message, and this case is no exception. I can only hope that what follows will land somewhere on the border between temerity and timidity.

The kind of determinist theory with which I shall be concerned is that which figures in most arguments that try to relate it to free will and responsibility. In the first place, the focus is on macro-events and –states of affairs; thus it is assumed legitimate to ignore the question of indeterminism in quantum mechanics. Secondly, the issue is universal macro-determinism, i.e. determinism not of specific systems within the universe, but of every state of the world. Thirdly, the determinism in question runs from earlier to later, even though there is a sense in which, at least for some processes, the later may be said to determine the earlier. Fourthly, it is assumed that laws of nature or general causal patterns constrain particular events rather than the reverse: in other words, it is not the case that laws and patterns supervene on particulars.³

The counterfactual conditionals at issue here are ones that assume causation running from one particular actual event to another, and hence a causal nexus between the counterfactual antecedent and the counterfactual consequent. Furthermore, the cases I have in mind are not simply subjunctive conditionals, but ones whose meaning implies that what the subordinate clauses represent is *not* the case. Thus they are literally counterfactual. This is by contrast with subjunctive conditionals that leave this open while perhaps suggesting that what is represented is only a remote possibility. In addition, the counterfactuals in question refer to the past: they say what would have happened or obtained at a past time if something which actually happened or obtained then or a bit earlier had not - although there may also be a focus on the difference this makes for the present state of things. Further, they have to do with cases where something practical is at stake: where the difference between the actual and the counterfactual course of events amounts to a difference of harm and benefit, loss and gain. Obviously this includes cases where we focus on an *intended* outcome, contemplating how things would have been better or worse if someone had or had not acted in a certain way. But not all the relevant cases are like this. There are also ones that focus on mindless or unpurposed events such as storms and traffic-jams and lottery-wins whose occurrences and non-occurrences are noticed for leaving people better or worse off than they would have been otherwise. And there are cases where the kind of event envisaged in the antecedent might or might not have been intentionally brought about by someone: e.g. a missed train, being late for an interview, encountering a friend in the street. The common thread is that these are all cases where we not only talk or think about how things would have been if something that happened hadn't, or if something that didn't happen had, but we also, and thereby, find it very natural to compare actual and counterfactual outcomes in terms of loss or gain, thinking how someone would have fared better or worse if what happened hadn't

² This is not itself a ground for taking the universe to be indeterministic. It is not written obviously in stone that philosophers ought to regard ordinary practice with counterfactuals as reasonable; and even if we cannot help so regarding it, moving from that to conclusions about the world raises huge and difficult questions.

³ Determinism as just described is the focus of the much discussed 'consequence' argument.

happened, i.e. if some contrary outcome had been *actual instead*.⁴

Here are some scenarios for illustration.

A (Party): We say: ‘If it hadn’t been for the storm we’d have held the children’s party out of doors, and it would have been a lot more fun for them than it actually was: what a pity about the storm.’

B (House Fire): At 2 am a neighbour smells smoke next door and immediately phones the fire brigade, which comes at once; the story ends with a rescue just in time. Later we say: ‘If he hadn’t noticed and acted at once, the people sleeping inside would have died that night one way or another: they were lucky the neighbour was awake and alert.’

C (Shareholder): After carefully studying the situation, I move a large investment out of entity JKL and into entity WVU. A few days later things have turned out as I thought they very well might, and I’m saying to myself or others: ‘What a good thing I followed my judgment and switched when I did; otherwise I’d now be holding a bunch of shares worth a quarter of what I paid for them.’

In ordinary life we often judge that counterfactuals like these are true or well founded and correct,⁵ and I am going to assume that we are often right about this.

II. FIRST STEP

It is a familiar point that accepting a counterfactual conditional depends on ‘holding fixed’ certain features of the actual situation. That is, we take it that the situation envisaged in the counterfactual antecedent has background features in common with the actual situation. It is equally familiar that our willingness to treat the conditional as true or well founded, rather than the reverse, depends on which background features we ‘hold fixed’. The conditional incorporates or is qualified by a *ceteris paribus* clause, and it all depends on what we include under the ‘other things’ that are said to be ‘equal’. (So, e.g., in B [House Fire] we hold it fixed that the rooms next door had no working fire alarms; it’s not at all clear that with working alarms in place the neighbour’s action would have made a crucial difference.) Now the first step is that in the examples above, we don’t just *choose to* hold it fixed that the background of the counterfactual situation was the same as in the actual one, as if this were up to us and we could hold fixed something different if we wanted to; rather, it is as if we are factually confronted with the same background in both scenarios. Instead of its being merely our hypothesis or supposition or optional assumption that the actual background features also belong to the counterfactual situation, we non-optionally presuppose, or by implication assert, that they *are* or *were* part of it. The phrase ‘*ceteris paribus*’ makes life difficult by being ambiguous: it can be meant hypothetically, amounting to ‘let’s suppose that the other things were the same’, or it can be meant categorically, amounting to: ‘it’s a (non-optional) given that the other things *were* the same’. The first step of the paper consists in underlining that in our examples the categorical sense is operative.

⁴ This locution incorporates the commonsense assumption that, whether or not there are other possible worlds, this world alone is actual.

⁵ The wording aims to leave open whether what is assessed is the consequent in light of the antecedent or the conditional as a whole.

The examples present conditionals, whose antecedents and consequents are of course not categorical. But is it so clear that there is in fact something categorical in the vicinity? It may seem that there isn't, even when we add more detail about the background. Thus we might fill out House Fire by saying (α): 'Last night, if the neighbour hadn't noticed etc., the people in the house would have died, given that they would have been sleeping heavily - just as they were in the actual situation', and here the background condition of their sleeping heavily is phrased in what appears to be the subjunctive: it figures as what 'would have' obtained. We can, however, re-phrase this in the past indicative as follows (β): 'Last night, given that the people were all sleeping heavily, they'd have died if the neighbour hadn't noticed etc.' As far as I can see this change imports no addition or loss of meaning. It also seems perfectly legitimate to say 'since' instead of 'given that' here, which yields (γ): 'Last night, *since* the people were all sleeping heavily, they'd have died if etc.' With 'since' the indicative and categorical status of the clause about sleeping is even more obvious. The sleeping and other circumstances were the causal basis of a causal connection running from the purely suppositional (because counterfactual) non-intervention of the neighbour to the purely suppositional (because counterfactual) death of the people in the house: but it is the actual, not merely suppositional, sleeping etc. that was this causal basis. Similarly, the actual dryness of this dry match is (with other factors) the causal basis of the match's igniting-were-it-to-be-struck, even if it is never struck and never ignites. We rightly say, in view of this match's actual state, that if it were struck it would light; and this is different from saying that if it were struck, then on the supposition of its being dry as well as being struck it would light.

The apparent subjunctivity, in (α) above, of the clause 'given that they would have been sleeping heavily' may suggest that we are dealing with a supposition rather than something categorical. There are, however, ways of explaining away the appearance of subjunctivity. One is to think of the verb as grammatically attracted into the mood of the undeniably subjunctive clause 'if the neighbour hadn't noticed' and its counterfactual consequent. Another is to identify 'they would have been sleeping' here as the past indicative of 'they will have been sleeping', with prospective use of the latter legitimating retrospective use of the former. That is to say: the counterfactual conditionals we are concerned with can be viewed as past-tense or retrospective updates of previously relevant future-oriented conditionals. Thus from a supposed earlier perspective on the House Fire we observe the people heavily sleeping, we note the half extinguished cigarette under the bed as well as other relevant conditions, and we ask: 'Will an outsider notice? Will help come in time? If Yes, they will escape death; if No, they will perish.' We take it that whether the answer is 'Yes' or 'No', the factors mentioned – sleeping human beings, smouldering cigarette, flammable materials lying about, the laws of nature – *are* the same and *are* going to continue the same either way - until the situation develops into disaster or into rescue. From the earlier perspective we already predict that one or other of these outcomes will be actual, providing material for a later causal report using past-tense categoricals, e.g. 'Someone did notice and because of that the people were saved', while the other outcome provides material for a past-oriented counterfactual conditional. Prospectively we take it for granted that the background conditions are and will have been the same in both scenarios.⁶ Whichever turns out to be the non-actual one, it is still a scenario that rests on the actual background situation considered as actual. We don't in advance downgrade this background to the status of a mere supposition in

⁶ '... will have been the same': i.e. up to or overlapping with the time that is retrospectively viewed as crucial: i.e., in this example, the time when help arrived or the time when, absent help, disaster became unavoidable.

relation to whichever train of salient events (we don't yet know which train it will be) will turn out not to take place. We don't in advance restrict actuality of the background so that it is only in place for the as yet unidentified set of salient happenings that's going to be the actual one. Correspondingly, from the retrospective standpoint we are, I submit, likewise even-handed. Retrospectively there is, of course, the major difference of our knowing now which sequence did take place and which did not; but we still view the counterfactual absence of neighbourly intervention and the counterfactually resulting disaster as occurring under the same background conditions - the same not merely in description but also in actuality - as the actual intervention and its actual result.⁷

And arguably the even-handedness applies to suppositions as well. That is to say: if for some reason we confine ourselves to merely *supposing* that the people would have been heavily asleep in the counterfactual event of no neighbourly intervention, then (arguably) we should likewise only *suppose* them to have been heavily sleeping in the actual event of the neighbour's noticing smoke and getting help. That certainly seems strange given that we are in an epistemic position to assert that they were in fact sleeping. What is more, if we only *suppose* them to have been sleeping when the neighbour acted, we can't any longer say that the neighbour *saved* them and that they *were* thereby extremely lucky. We can at best say that he *would have saved* them, and they *would have been* thereby lucky, *supposing* that they had been heavily asleep. They might have been awake and put the fire out themselves in good time, or never have started it. Again, we say that the storm *spoilt* the party on the ground e.g. that the children arrived bursting with energy and only had outdoor toys (it doesn't seem to have been a very well planned party), and were going to continue in this state storm or no storm. If we grant that the storm did happen but regard the outdoor toys/high energy combination as only suppositional, then it can't be said of the actual storm that it *spoilt* the party, only that it *would have spoilt it if* that particular combination were to have been in place along with the storm. This, obviously, is consistent with a highly successful party of pre-exhausted children only too glad of an excuse to curl up in front of the television while the storm raged outside. And similarly if it's a matter of saying that the investor by her actual intelligent move *avoided* a major loss, as distinct from saying that she made the move and thereby *would have avoided* a big loss *if* the share price was about to plummet.

III. BACKTRACKING

The idea of the background has been prominent, but a few more words are needed on what it involves. Relevant background includes laws of nature, general causal patterns, and particular concurrent or antecedent states of things, events, processes, that empirically ground e.g. the positive claim that the neighbour's intervention saved the people's lives along with the associated counterfactual that had he not intervened they would have perished. (In our examples, the background also includes features that do not affect the bare bones of what occurs and its counterfactual alternative, but

⁷ The prospective conditionals should be hedged by a distinct *ceteris paribus* clause to the effect that between now and the crucial moment nothing unexpected will occur to disrupt the background conditions that narrow the outcomes to 'they will escape' and 'they will perish'. This 'absent disruption' *ceteris paribus* clause is suppositional. If and when we make the corresponding retrospective judgments, it is because we now know that no disruption did occur, and so we discard the proviso.

which make the outcome welcome or unwelcome to people concerned.⁸) The presence of relevant background is causally independent of the occurrence or non-occurrence of the intervention-event and its counterfactual alternative. For example, suppose Ann is furious that but for a train delay she wouldn't have missed a vital exam: she stops this fuming if she learns that the power-cut which caused the delay also knocked out the exam venue, so that the exam has been postponed: i.e. the background conditions had developed in such a way that there *was* no exam for her to miss through lateness on the day in question. In general, our assessment of what would or would not have resulted, and what the gains or losses would have been, if some event *E* had happened or not, is only as good as our assessment of what the background conditions were and were going to be whether *E* happened or not.

The problem posed by determinism is an old chestnut. Determinism says that, given the laws of nature plus the antecedent state of the world, only one subsequent state is possible at every stage. It follows that if any given stage had been different in some particular there must have been some prior difference in respect of some particular. But at whatever point we start this backward looking counterfactual reasoning, we imply prior counterfactual differences right back to the beginning of the universe. It becomes apparent early on in the argument that for each prior stage we can at best postulate the counterfactuality of a huge conjunction of factors many of which we are totally unable to identify. It must be that one or another of them would have been absent: but the disjunction covering this is indefinitely long, many or most of its disjuncts are unknown to us, and in addition the question of *which* disjunct at any given stage would have been different is undecidable in principle. It also becomes apparent that, the world being what it is, we are committed to a vast number of unknown counterfactual effects or outcomes resulting from postulated earlier differences and ramifying (as is the way with effects of causes in our world) ever further in space and onward in time. It follows that we don't have the slightest reason to suppose that, for example, if a storm had not blown up when and where that children's party was in fact held, the very conditions would have obtained that made it the case that a party was held just there that was a flop instead of a success. We don't know whether there would have been these or similar children, whether there would in the vast womb of time have developed such things as parties or even the human race itself. We do not know whether there would have been a physical world like this at all. We cannot begin to guess what would have been actual instead of what is. But in general when people judge that if it hadn't been for the storm the party – in that same place, with those same children, etc. - would have gone a whole lot better, they do so with what would generally be regarded as reasonable confidence based on ordinary experience of the world. Yet ordinary experience cannot possibly justify confidence that even if the universe had been different right back to the beginning in unguessable ways, there would still have come into being these children, this party, these capacities for pleasure and disappointment, etc. Thus if we judge 'The storm spoils the party' with what we regard and perhaps cannot help regarding as reasonable confidence, in effect we are simply setting aside the notion that the hypothetical non-occurrence of the storm entails hypothetical differences at every stage between now and the beginning of the universe. But in that case we are setting aside determinism. We are not or should not be judging it false, for it is logically possible both that determinism is true and that the storm spoils this particular party; i.e. both that determinism is true and that on the

⁸ Obviously the same outcome can be welcome and unwelcome to different persons given their different conditions and circumstances. This complexity does not affect the present argument.

occasion in question the background obtained that made it the case that with the storm the children were bored and cross and without it they would have had a good time instead. But this logical possibility does not alter the fact that we set determinism aside as lying beyond the horizon of what can be reasonably entertained, given that we see ourselves as making a reasonable judgment that the storm spoils the party.

Someone may object that it is downright illegitimate ever to let that unlimited backtracking get started: for if we allow it, we render ourselves unable to evaluate the truth or correctness of a great many everyday empirically founded counterfactual conditionals! I agree: but the present contention is that putting a curb on backtracking comes with a price, namely refusal to treat determinism as having real epistemic possibility.

IV. WIDENING THE TARGET

The unlimited backtracking unleashed by a determinist picture of actual history faces counterfactual reasoners with the impossible task of deciding which entire course of history – if any - would have been actual instead of the one we know we are in, the one that led e.g. to the children's party and the storm. I think that a parallel argument can be mounted against a certain kind of *indeterminist* picture of actual history. Indeterminism means that it is not the case that for each stage of the universe, however stages are discriminated, only one next stage is nomologically possible – perhaps because some or all laws of nature and causal patterns are probabilistic rather than deterministic. Some philosophers view possible worlds, including this actual world of ours, in terms of world-books, i.e. maximal conjunctions of propositions such that each conjunction details an entire possible world – a world that is actual if the conjunction is true.⁹ Now let us consider the world book of an indeterministic world, and let us assume that the book represents every detail of the world in question throughout its entire history. Thus whether or not the world in question is actual (or actualized), it is fully determinate throughout or as a whole. That is: the contents of any given moment of its history are determinate at every moment. (The contrast is with a world whose contents, or some of them, only become determinate as time goes on.) Now let us suppose that our actual world is fully determinate in the way just indicated, and let us return to our counterfactuals. It seems to me that even if this fully determinate world is indeterministic, we have much the same problem as before in deciding how things would have been if the neighbour had not intervened or if there hadn't been a storm. If the world in its spatio-temporal spread is actual or actualized as a fully determinate whole, then the hypothesis that the neighbour didn't intervene amounts to the hypothesis that this whole world was not actual or actualized: that its world-book contains at least one false conjunct and so is false as a whole. But now if we ask in the way we do when framing ordinary counterfactual conditionals 'How in that case would things have been instead? – *what* difference did it make that he intervened, etc.?' we are faced with essentially the same situation as before: if this whole world – the world as defined by its complete world-book - had not been actual, no doubt some world or other would have been actual instead, but there is not the slightest reason for believing that what would have been actual instead is a world with just the features necessary for it to be true that there was

⁹ One such conjunction describes the completely empty world if such is metaphysically possible. I leave aside books for impossible worlds.

that neighbour and that house containing those sleepers who would have died, and so on. Such a world logically, and no doubt metaphysically, *could* have been actual instead; it's just that there is no reason at all to believe that it *would* have been. In short, these counterfactual conditionals are a problem for any view whereby world-history is actual or actualized - or else not - as a whole, since we cannot know what, if anything, would have been actual or actualized instead of this whole. A deterministic history is holistic because nothing could have been different at any stage without incalculable differences at every previous and subsequent stage; and an indeterministic history is, if anything, even more 'all-or-nothing' if one assumes that its contents at every moment are, at every moment, fully determinate.

So according to what I am arguing here, our practice with counterfactual conditionals such as A – C above looks as if it presupposes a world-view in which a measure of indeterminism in the sequence of events combines with a measure of indeterminacy about the future at each stage. This yields a version of the familiar (although in some quarters philosophically unloved) picture of world-history becoming determinate bit by bit as the future becomes present. *Incompleteness* in respect of determinacy at any given time is one feature of a world that fits our practice with counterfactuals. The other feature is an indeterministic *looseness*, such that not all world-stages occur as the sole possible successors to what preceded them.¹⁰ I leave aside for now any discussion of precise ways in which these features of indeterminacy and indeterminism must and must not apply to the ingredients of our examples, namely the salient sequence of cause and effect, and the background conditions.¹¹

V. THE 'LOCAL MIRACLE' ACCOUNT

David Lewis proposed a determinism-compatible account of counterfactuals in terms of 'closest worlds' in which the counterfactual antecedent is true. The content of those relevantly closest worlds includes a minor breach of natural law (as judged by the nomological standard of this world). Specifically, this 'local miracle' is such that the antecedent event emerges from a past which almost up to the moment of emergence

¹⁰ Indeterminacy in the sense envisaged here is compatible with universal determinism; see e.g. Tooley, 2000, 22-24. For illustration, suppose that a god or spirit of history brings about (so as to render determinate) each stage of the universe just in the light of what he/it perceives as the implemented concrete detail of the previous stage (I leave aside the problem of how he individuates stages). Suppose further that he/it necessarily discerns at each juncture exactly what *the most appropriate* next stage would be (whatever the standard for 'appropriate'), and necessarily proceeds to implement this. Moreover, there cannot be ties for 'most appropriate next stage'. The key idea is that the supernatural agent, like many a human artist, has in some sense to experience or be confronted by the completed actuality of each stage before knowing exactly what to make happen next. Thus the agent produces history bit by bit even though, given the assumptions, each next stage is the only possible way forward from the previous one.

¹¹ Why exactly would a 'loose and incomplete' world resist pernicious (unlimited) backtracking with counterfactuals? The most I can presently say is that, given such a world, the past moment when one of several previously contingent trains of events *became determinately actual*, and the others *became determinately excluded* from actual history, constitutes a logically natural firewall against further backtracking. Given (for any *t*) the difference in ontological status between what at *t* becomes actual and what at *t* is already actualized, it is not legitimate to think in terms of 'this whole world' while its history is still going on. It may be a different matter at the end of the world's history, if that notion makes sense. But for anyone in the midst of history the only alternatives to some stretch *S* of actual history are ones that share an actual past (which would have been their past, had they been actual) with *S*.

exactly matches the actual past in terms of particulars.¹² The overall account is designed to uphold our ordinary discriminations between acceptable and unacceptable counterfactual conditions, and in particular is tailored so that it blocks unlimited backtracking. On a very rough first impression one might complain that it simply asks us to pretend (in a rather elaborate way) that determinism is not true even though, in the context of defending compatibilism, we are assuming for the sake of argument that it *is*! But that would be much too crude. It is, however, fair to note that the account does carry a cost. It postulates different systems of natural law for the actual world and the relevantly closest other worlds. Since the laws of a world would seem to be overwhelming candidates for inclusion among background conditions of any salient particulars of that world, the account entails abandoning one or another natural assumption: either (1) the prospective-standpoint assumption that background conditions are and are going to be the same either way whether or not e.g. the neighbour calls the fire brigade, or (2) the assumption that retrospective counterfactuals inherit that prospective feature unchanged except for tense (see Section II). Since the device of a closest-worlds ‘local miracle’ incurs this cost and is not (as far as I can see) an independently attractive postulate, it’s fair to ask whether it is more than an *ad hoc* measure to save determinism from triggering the backtracking that would destroy the rational basis of many everyday counterfactual judgments. And if this is a not unfair suspicion – the suspicion that the ‘miracle’ is *ad hoc* – then the debate has surely reached a point where anyone who entertains any sympathy for universal macro-determinism, whether or not they finally endorse it, ought to step forward to show why this theory is so worthwhile that one should seriously consider buying it today even at significant cost. Perhaps the thought is that empirical science may be going to show that universal macro-determinism is our best theory, so we had better prepare ourselves by constructing analyses that uphold its compatibility with everything we hold dear. However, it is not at all clear whether empirical science could ever show any such thing.¹³ If, moreover, the task of upholding compatibilism were allowed to recede in importance, we could channel the energies saved into fuller explorations of the theoretical potential of an unapologetic – and far from ‘panicky’ – macro-indeterminism.¹⁴

Lewis thought of possible worlds, whether deterministic or not, as fully determinate in the sense discussed in the previous section. This entails, I think, that a closest-worlds ‘local miracle’ account is vulnerable – except on an assumption to be mentioned in a moment – to the type of objection raised there. Although the past in the closest worlds (once we get back beyond the ‘miracle’) exactly matches the actual past, we still have no reason to hold that if (e.g.) it hadn’t been the case that the neighbour smelt smoke etc., *one of those closest worlds would have been actual instead of our actual history*. As before, we have absolutely no idea what sort of entirely determinate world would have been actual instead of this one. Thus we are not in a position to say that but for the neighbour’s alertness the people in the house would have perished *instead* of surviving: nor, therefore, to say that his alertness *saved* them. However, Lewis himself is not open to this objection because he would have rejected the whole idea that one

¹² Lewis, 1979.

¹³ Cf. Steward, 2011 and 2012 on universal determinism as a metaphysical rather than empirical theory.

¹⁴ Cf. Strawson 1962, *ad fin.*

world could be actual *instead* of another, or *rather than* another. This follows from his extreme realism about possible worlds, according to which the phrase ‘the actual world’ is purely indexical, so that each possible world is actual from the point of view of itself. There is no world that is, absolutely, the uniquely actual one, any more than, given the plurality of selves or places, there is only one Me or only one Here. Hence it is nonsense to speak of one world being actual *instead* of some other. Lewis’s extreme modal realism is the assumption that blocks the objection. But most philosophers reject extreme modal realism. In general, the pros and cons of that debate lie well beyond the topic of this paper: but it is interesting to note that one reason for embracing extreme modal realism might be that it enables philosophers to remain on compatibilist terms with determinism! Yet, as before, we may now feel that it’s time someone explained exactly what determinism these days has going for it that would justify such a steep philosophical price. It is not enough to repeat what we all know: that determinism has a hoary history and has been endorsed by some very great earlier philosophers.

VI. WAYS WITH COUNTERFACTUALS

Throughout I have spoken of a loss-or-gain-oriented *use* of counterfactual conditionals about particulars. This is because there is surely a distinct use of such conditionals, one which I’ll simply call ‘theoretical’, where the attention is not on how the counterfactual outcome would or might have affected anyone for good or ill. The focus is purely on the nexus of antecedent to consequent regardless of whether the result or non-result would have been to anyone’s advantage or detriment. One can note that if that match had been struck it would have ignited, with emphasis purely on the point that if something else, e.g. a metal pin, had been scratched on the matchbox there would have been no ignition. One can reckon that if the oil tank had been filled last month it wouldn’t be running low today, with emphasis purely on the specific rate of consumption. The statement that if John had lifted this dumb-bell he would have injured his back may simply be meant to underline a difference between the strength of someone with John’s physique and that of a stronger or fitter person, or between the weight of this dumb-bell and another. In the theoretical use we are not primarily thinking of ways the counterfactual events, had they occurred, would have contributed to the on-going history of the world. We are not thinking about them as events whose occurrence, had they occurred, would have been embedded in a history stretching forward into the future and back into the past. We are thinking of them as indicative of the natures of the objects or processes involved or of the causal connections between them. Thus, in saying or thinking ‘If that match had been struck it would have ignited’ we may be attending to striking that match as opposed to striking a pin, or to striking it as opposed to breaking it into two pieces, or to its igniting rather than, say, melting. We are not focusing on what difference it would have made to the surrounding situation if that particular match had ignited at time *t*, e.g.

on whether it would have got a pleasant hearth-fire going or have caused an explosion because gasoline vapour was present. So although this counterfactual conditional is about concrete particulars, in theoretical use one abstracts from the context in order to make a point about the objects, processes, and relations represented within the conditional. Obviously the conditional is true or correct in virtue of certain concurrent background conditions, but the abstract nature of its theoretical use is such that we systematically ignore earlier antecedents just as we ignore wider and subsequent consequences. Thus

we screen out questions about how it came about that this match wasn't struck at or just before time *t*. As far as we are concerned the world complete with match etc. might as well have been created a few seconds before *t*. In other words, questions that trigger backtracking don't arise. Nor do questions about wider and subsequent differences made by the match's *not* igniting, hence by its having remain *unstruck*. As for the fact itself of the match's staying unignited rather than igniting, we have (I submit) very little interest in this (we take it for granted) by comparison with our interest in the fact that what it signally didn't do was ignite rather than melt. Contrast this with the House Fire case where, in practical mode, we are all hung up on the difference between the people's surviving versus their perishing, and therefore on the difference between the neighbour's having taken action and his not having done so. This is because the practical concern is bound up with seeing the events or non-events as historically situated and as belonging in the biographical narrative of anyone believed to be affected:¹⁵ which is also why here we are not free to refuse to backtrack if it seems to us that the counterfactual antecedent is such that, had it been actually true, the past must have been different.

I also suspect that in the theoretical use it makes little difference if we merely *suppose* the background conditions to have obtained in both actual and counterfactual scenarios (cf. Section II). In suppositional mode we get 'If the conditions were the same, then if the match had been struck it would have ignited'; i.e. the conditional that connects match struck and match igniting is a conditional conditional as distinct from a plain categorical one. Now, it's not at all clear that this difference affects the information conveyed, or our interest in it, if what concerns us is match struck versus pin struck, or match struck versus broken in two, or match igniting versus melting. Contrast the practical use, where the riveting thought is that actual conditions were such that if the neighbour hadn't stepped in the people *would have died* period. How starkly different from the thought that *if* such conditions obtained, they would have died if the neighbour hadn't stepped in! In one case - if we wish the people well - we, like them, *are* thankful for the difference the neighbour *made*. In the other we form, at most, a 'cool and indolent judgment of the understanding' that we *would have been* thankful for the difference he *would have* made. But the main point is that backtracking, hence the unlimited backtracking precipitated by determinism, only targets actualities regarded categorically - as having obtained *period*. Counterfactuals in theoretical use are therefore out of danger if this use can dispense with categoricity.

If philosophers overlook our different ways with counterfactuals the topic becomes more slippery even than nature made it. Focusing just on the theoretical use gives easy assurance that determinism introduces no problem. Within limits this is true, since that use emerges as immune to pernicious backtracking. Presumably, however, the practical use is what controls examples that come to mind when philosophers worry over whether they can rationally believe in determinism along with believing in moral agency and free will. Here, the old chestnut hasn't been cracked.

¹⁵ Suppose the story goes the other way - the neighbour failed to act, the sleepers died - and consider the naturalness of comments like: 'To think he [one of the deceased] had just celebrated his twenty first birthday/returned fit and well from fighting in Afghanistan/been awarded a prestigious scholarship/heard that his wife was expecting their first child...', even when these circumstances had no causal bearing on what happened.

*Departments of Philosophy
University of St. Andrews
Edgecliffe
The Scores
St Andrews
Fife KY16 9AR*

REFERENCES

Lewis, David 1979: ‘Counterfactual Dependence and Time’s Arrow’, *Nous* 13, 455-476, reprinted with postscripts in D. Lewis, *Philosophical Papers* vol. II, Oxford 1986, 32-66.

Steward, Helen 2011: ‘Moral responsibility and the concept of agency’, *Free Will and Modern Science*, ed. Richard Swinburne, Oxford, 141-157.

Steward, Helen 2012: *A Metaphysics for Freedom*, Oxford.

Strawson, P. F. 1962: ‘Freedom and Resentment’, *Proceedings of the British Academy* 48, 1–25.

Tooley, Michael 2000: *Time, Tense, and Causation*, Oxford.



THE ARISTOTELIAN SOCIETY

PRESIDENT: Sarah Broadie (St. Andrews)

PRESIDENT-ELECT: E.J. Lowe (Durham)

HONORARY DIRECTOR: Lucy O'Brien (UCL)

EDITOR: Matthew Soteriou (Warwick)

LINES OF THOUGHT SERIES EDITOR: Scott Sturgeon (Oxford)

EXECUTIVE COMMITTEE: Ben Colburn (Glasgow) / Alison Hills (Oxford) / Rosanna Keefe (Sheffield)
Marie McGinn (UEA) / Samir Okasha (Bristol) / Ian Rumfitt (Birkbeck) / Robert Stern (Sheffield)

EXECUTIVE ADMINISTRATOR: Mark Cortes Favis

ASSISTANT EDITOR: David Harris

EDITORIAL ASSISTANT: Lea Salje