Chances, Counterfactuals, and Similarity

J.Robert G. Williams*
Revised Draft: Comments welcomed!
j.r.g.williams@leeds.ac.uk

(August 9, 2006)

Abstract

John Hawthorne in a recent paper takes issue with Lewisian accounts of counterfactuals, when relevant laws of nature are chancy. I respond to his arguments on behalf of the Lewisian, and conclude that while some can be rebutted, the case against the original Lewisian account is strong.

I develop a neo-Lewisian account of what makes for closeness of worlds. I argue that my revised version avoids Hawthorne's challenges. I argue that this is closer to the spirit of Lewis' first (non-chancy) proposal than is Lewis' own suggested modification.

1 Counterfactuals and Chance

The antecedents of some counterfactual statements render their consequent hugely probable, but not certainly true. That is, it is not impossible that a combination of unlikely coincidences could lead to a situation in which the antecedent is true and the consequent false. For example:

*Centre for Metaphysics and Mind, Department of Philosophy, University of Leeds. A version of this paper was presented at the Arché AHRC research centre modality seminar at the University of St Andrews. Thanks are due to all present for much valuable discussion. Thanks also to Ross Cameron, Antony Eagle, Adam Elga, Carrie Jenkins, Stephan Leuenberger, Andy McGonigal, Joe Melia, Daniel Nolan, Crispin Wright, and an anonymous referee for XX for enlightenment.

> (A) If I were to toss this fair coin 10,000,000 times, it would not come up heads every time.

David Lewis (1979) and John Hawthorne (2005) agree that such statements are counted as true in ordinary non-philosophical discourse. Nevertheless, were the antecedent satisfied, there is a calculable chance that the consequent would turn out false. Reflecting on this, we are inclined to endorse:

> (B) If I were to toss this fair coin 10,000,000 many times, it might be that it comes up heads every time.

And, in light of this, there is some pressure to withhold one's assent from (A), and endorse instead:

> (A*) If I were to toss this fair coin 10,000,000 many times, it would be extremely unlikely to come up heads every time.[1]

These examples give the flavour of the topic to be addressed here, but real bite is put into the issue by an acceptance of quantum mechanics, under an interpretation according to which the wave function for a physical system delivers objective probabilities of location. There is a real, albeit tiny, chance that I will spontaneously disappear from my present location while, simultaneously, an intrinsic duplicate of me appears on Mars. Analogous considerations effect almost every routine counterfactual. Hawthorne's example is:

> (C) If I had dropped the plate, it would have fallen to the floor.

[1] See, however, the appendices to Lewis (1979) for a treatment of the 'might' counterfactual which does not have this consequence.

As before, there is a small chance that the consequent fails to obtain, given the antecedent. Thus, the following is tempting:

(D) If I had dropped the plate, it might have flown off sideways.

This motivates the rejection of the following:

(E) If I had dropped the plate, it would not have flown off sideways.

But (C) and not-(E) are *prima facie* incompatible.[2] So, just as we retracted assent from (A) in favour of (A*), it seems that we should replace (C) with (C*).

(C*) If I had dropped the plate, it would very likely have fallen to the floor.

Hawthorne calls this the 'error theory' of ordinary counterfactual judgements. His task, and ours, is to examine ways of avoiding it, all of which are framed within the possible-worlds approach to the semantics of counterfactuals.[3]

The approach to be advocated here is a modification of Lewis' theory of counterfactuals. I will develop this view in the course of evaluating Hawthorne's critique of Lewis' own proposal. I first sketch Lewis' views, and then give Hawthorne's arguments against this account, giving what I take to be the best responses available to Lewis. I conclude, though, that Hawthorne's objections ultimately succeed in their aim of undermining the Lewis treatment. I then give my own version of a Lewisian view, drawing on the notion of 'typicality' introduced by Elga (2004). None of the original problems afflict my account.

[2]We assume that in all relevant worlds, p: (flying off sideways) and q: (falling to the floor) are incompatible. I.e. Every p world is a not-q world, and vice versa. Then the standard semantics for counterfactuals has it that (C) is true iff all the closest dropping-plate worlds are p worlds; and (E) is false iff there is some dropping-plate world which is a q world. So there must be some dropping plate world that is both a p and a q world. But this contradicts the above.

[3]It is noteworthy that other approaches to counterfactuals, e.g. Edgington (1995), are not obviously susceptible to the same concerns.

Lewis' approach

Lewis' account of counterfactuals consists of two elements. The first is a version of a now-standard analysis of counterfactuals:

> TRUTH CONDITIONS
> '$A \Box\!\!\rightarrow B$' is true iff B is true at all the A-worlds closest to the world of evaluation.[4]

Our previous concerns can be restated in this framework. Suppose a fair coin were to be flipped 10,000,000 times. Then there are $2^{10,000,000}$ equiprobable possible outcomes. Shouldn't each of these possible worlds be counted as equidistant from the actual world? But if so, then one of the 'closest' worlds will be one in which the coin lands heads each time—therefore the counterfactual (A) will be false.

Whether this argument is correct depends on whether the presumptions about closeness are vindicated; and indeed, the other element of Lewis' theory is an account of what makes one world closer than another. Initially, he proposed the following:

> SIMILARITY
> w_1 is more similar than w_2 to the world w_0 if the differences between w_1 and w_0 are of less weight than the differences between w_2 and w_1. The weighting of the differences is governed by the following principles:
>
> 1. It is of the first importance to avoid big, widespread, diverse violations of law.

[4]In fact, the gloss just given is only appropriate if we grant the "Limit assumption" (Lewis, 1973, p.19)—that there can never be an infinite series of *A*-worlds closer and closer to the world of evaluation. Lewis is not prepared to grant this assumption—in fact, he thinks it false—so he offers a more generally version of TRUTH CONDITIONS. First, let a non-empty set of worlds S be a *sphere around w* if there is no world as close to w as an element of S that is not already a member of S. Then we let '$A \Box\!\!\rightarrow B$' be true iff either (a) there are no A-worlds, or (b) there is some sphere S containing an A-world such that $\neg A \vee B$ holds throughout S (Lewis, 1973, p.16).

The additional generality gained from Lewis's official account does not alter any of the issues we will be considering, so I will continue to use the simpler version.

2. It is of the second importance to maximise the spatio-temporal region throughout which perfect match of particular fact prevails.
3. It is of the third importance to avoid even small, localised, simple violations of law
4. It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly

The justification for this understanding of 'similarity' is not that it matches our "explicit, snap judgements" about which worlds are similar to which others. If our snap judgements are typically sensitive to "imperfect match" over a whole region, the above account focuses rather on "perfect match" over a limited part of the region. Lewis, however, is keen to insist that this is still a legitimate notion of similarity, in the ordinary sense of the word.[5] This is the *analytic* ambition. Even if we were to deny this, however, the analysis above could be justified as a purely technical underpinning for Lewis' account of counterfactuals. It would then stand or fall to the extent that it tracks intuitions about the correctness of counterfactuals. This is the *instrumental* ambition.

The initial analysis goes wrong when worlds with chancy laws are introduced. In the actual world, I do not drop the plate. Now consider a counterfactual with antecedent 'I drop the plate'. The closeness of plate-dropping worlds is determined by the extent of spatio-temporal difference and law-violation that obtain. But we can minimise both, in a chancy world, by supposing that wave-functions collapse immediately after the dropping incident, returning the world to a state that exactly resembles the actual world at the same time. Hugely improbable, to be sure—but there is a definite positive probability that it will happen, and there are possible worlds involving no violation of physical law[6] where it occurs. Given SIMILARITY and TRUTH CONDITIONS, the effect is that, in the

[5] The above quotations are taken from Lewis (1979, p.54).

[6] Unless some violation is needed to make the world into a plate-dropping one; but this cannot be a factor in discriminating between worlds where the antecedent holds

majority of cases 'if it were that p then…' will be true just in case the consequent is true in the actual world, no matter what p is. Suppose I go walking along a cliff one day, and play football the next. The following would come out true: "If I had thrown myself over the cliff that day, I would have played football the next" would come out true—which is absurd. So the account of similarity needs amendment.

Lewis calls the unlikely 'convergence' events that occur in such worlds 'quasi-miracles'. He characterises 'quasi-miracle' thus

> What makes a quasi-miracle is not improbability *per se* but rather the remarkable way in which the chance outcomes seem to conspire to produce a pattern.
> (Lewis, 1979, p.60)

Whatever happens, in a chancy situation, it is exceedingly improbable that *exactly that* happened. *Any* particular ordered sequence of heads and tails is an equally improbable outcome of flipping a fair coin, notwithstanding that it is an entirely typical sequence. Not just any course of events should constitute a quasi-miracle, however, so Lewis adds the "remarkability" clause to his characterisation.

Lewis response to the convergence problem is as follows:

> What must be said, I think, is that a quasi-miracle …, though it is entirely lawful, nevertheless detracts from similarity … The quasi-miracle would be such a remarkable coincidence that it would be quite unlike the goings-on we take to be typical of our world. Like a big genuine miracle, it makes a tremendous difference from our world.
> (Lewis, 1979, p.60)

The suggestion is, I take it, that we incorporate a new condition into a revised version of SIMILARITY.

SIMILARITY*
w_1 is more similar than w_2 to the world w_0 if the differences between w_1 and w_0 are of less weight than the differences between w_2 and w_1. The weighting of the differences is governed by the following principles:

1. It is of the first importance to avoid big, widespread, diverse violations of law, or big, diverse, quasi-miracles.
2. It is of the second importance to maximise the spatio-temporal region throughout which perfect match of particular fact prevails.
3. It is of the third importance to avoid even small, localised, simple violations of law, (or small localised quasi-miracles?)
4. It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly

Given the formulation in terms of quasi-miracles, we are given answers to many of our earlier questions. There is a chance that, when dropped, the plate would fly sideways. But for it to do so would be a remarkable and low-probability event—i.e. a quasi-miracle. So worlds where this happens are *ipso facto* further away than typical plate-dropping worlds. So flying-sideways, or always-heads, worlds, are not among the closest where the respective antecedents holds. So (A) and (C) are vindicated, if SIMILARITY* can be sustained.

2 Hawthorne's criticisms, and Lewisian responses

John Hawthorne (2005) has put forward four objections to Lewis' revised analysis. In this section I describe each, and try to construct the best-possible Lewisian response.

Hawthorne frequently appeals to intuitive judgements about what events are, and which are not, remarkable. Since Lewis does not give anything like a substantial discussion of how this notion is to be taken, this is fair enough. It is also fair, however, that the opponent allow the Lewisian to precisify 'remarkableness' in whatever way will make her account strongest, so long as overall a coherent story about the notion emerges.

In the case of one of the puzzles that Hawthorne presents to the Lewisian, I think a direct rejoinder can be given. But in the other three cases, though the Lewisian may escape refutation, the puzzles succeed in identifying costs that the account must bear. Since these costs mount up, we are motivated to look for a revised account of remarkableness that will give a more satisfactory treatment of Hawthorne's cases.

The division problem

Quasi-miracles are remarkable, low-probability events. But there are remarkable events which do not have extremely low probability. Suppose that a monkey is at this moment so configured that, were it to start typing now, it would have a 20 per cent chance of producing a readable dissertation on anti-realism.

We do not want to endorse:

> (F) Were the monkey to start typing, it would produce junk, not a dissertation.

One can typically divide a reasonably probable event into low-probability subcases. It's probable that I will stand up sometime in the next few minutes; but highly unlikely that the way I will do so will fit atomistic description D_1. Every other atomistic description of a way of standing up is similarly improbable. These might be all the possible ways in which I could achieve standing up—so their exclusive disjunction is highly probable.

This goes equally for remarkable events which are not themselves of extremely low probability, such as that described above. There are many ways D_i in which the monkey could produce a dissertation on anti-realism. Each is highly improbable. And each is remarkable. By the above analysis, that makes each D_i world further away from actuality than every (improbable but) unremarkable world in which it produces junk. But this means we do get the counterfactual (F) “if the monkey were to start typing, it would produce junk, not a dissertation” coming out true, contradicting our earlier statement.

Response

One might be find the particular case that Hawthorne uses to illustrate the division problem problematic: a world with dissertation-writing monkeys might seem so removed from the world of our actual experience, that one might not feel comfortable in putting weight on one’s intuitions about what is or is not remarkable there.[7] I take it, however, that Hawthorne does not intend his example to be invoke a situation where monkeys in general have special literary abilities; rather, we are to suppose that some actual monkey, with actual monkey abilities, just happens to have its brain set up so that there’s a 20 per cent chance that the sequence of keys it strikes will produce a readable dissertation. It is legitimate, therefore, to apply actual-world standards of remarkableness.

One might not be quieted by this rejoinder, finding it hard to imagine what a monkey’s brain would have to be like to fit this description. But the division problem extends to other (perhaps more sober) examples where we can see exactly what is involved. I mention one in a footnote.[8] Setting aside these concerns, therefore, I will continue to

[7] I am grateful to an anonymous reviewer for bringing this issue to my attention.

[8] Our example is one mentioned (in another context) by Hawthrone: the event of getting a particularly remarkable hand dealt to one when playing Bridge: a hand containing all spades, for example. It does not seem at all absurd to suppose that someone sometime has actually been dealt such a hand. Even though the event of getting dealt all-spades strikes one as remarkable, things can be so set up that it is highly probable that it will happen. Suppose, for example, a deck of cards has been shuffled, and that every

discuss the division problem through Hawthorne's monkey example.

Hawthorne's objection rests solely on the claim that the event in question is *remarkable*. To see this, suppose that the chance that the monkey produces a dissertation on anti-realism is close to 1. Still, given that the event of the monkey's writing the dissertation is remarkable, we can argue such an event counts as quasi miraculous—and so the counterfactual (F) would still be false. For we can still partition the event of the monkey producing the dissertation into low-probability events (perhaps according to the precise timing of each keystroke). Each of these events will be remarkable and low-probability; hence quasi-miraculous.

Is the monkey's producing a dissertation, in the relevant sense, remarkable? The Lewisian should claim that it is not. If one flips a weighted coin, no wonder that it comes down heads most of the time. If one deals hands from a stacked deck where every fourth card is a spade, no wonder the fourth player gets an all-spades hand. And if one sets a 'weighted' monkey at a keyboard, it is no surprise if it produces a dissertation.

If an event is 'remarkable' just in case it is apt to be found surprising by agents, then the monkey's producing the dissertation is remarkable. But if 'remarkable' is read as 'apt to be found surprising by ideal agents in full knowledge of all relevant information', then not. The notion of remarkability that Lewis presents is fairly inchoate: the moral that the Lewisian should draw from Hawthorne's division problem is that the notion stands in need of precisification in the way indicated.

---

fourth card is a spade. Relative to this time, the objective chance that the fourth player getting an all-spades hand are high (the only way to avoid it is the dealer makes some mistake in dealing out the cards). The division argument says that still, this event will count as quasi-miraculous, since we can partition the event of being dealt that hand into low-probability subcases. We can suppose that in the actual world, the players were called away before the hand was dealt[9]: nevertheless, we want to say that *were* the hand to have been dealt, the fourth player would have had an all-spades hand. The division argument then steps in to say that each such outcome would be quasi-miraculous, by Lewis's lights, the closest worlds where the supposition is true will be ones where the player does not.

Though the response seems attractive, it brings a cost. Remarkableness itself is now being explicated in explicitly counterfactual terms: as *what would be found surprising by well informed people*. This explication is forced upon those who rely on 'quasi-miracles' to handle chancy counterfactuals, on pain of succumbing to Hawthorne's division problem. However, it appears circular. Counterfactuals are being explicated in terms of an appropriate notion of similarity; and similarity itself is now being explicated in terms of certain counterfactuals. This does not *trivialise* the analysis, for it enables one to explain a broad class of counterfactuals in terms of a small subset concerning the judgements of ideally informed agents. However, it is incompatible with the ambition the Lewisian originally held for the analysis of counterfactual similarity: to analyze counterfactuals in purely non-counterfactual terms. The strategic cost of this response to Hawthorne's division problem is thus significant.

The problem of the abundance of quasi-miracles

Take any remarkable fact about the world—an example Hawthorne suggests is the coincidence in apparent diameter of the Moon and the Sun. This is remarkable, and under some description, it is extremely improbable. So it counts as a quasi-miracle.

Hawthorne does not spell out exactly why the presence of quasi-miracles—even lots of them—in the actual world causes trouble for Lewis. I will consider three possible objections to Lewis' account based on the assumption that quasi-miracles are abundant in the actual world—I do not know which of these, if any, Hawthorne intended.

First, it might be argued that if quasi-miracles are abundant in the actual world, Lewis' characterisation of similarity will make a dull world (i.e. one just like the actual world but with the remarkable facts excised and replaced by unremarkable ones) more similar to the actual world than the actual world itself is. Perhaps the avoidance of 'large and diverse' quasi-miracles, will be more important than 'exact intrinsic match'. If so, then the dull world will be closer to actuality than the actual world itself.

Were this to be sustained, the Lewisian account would be in trouble. For if a world is closer to actuality than the actual world is, we will have failures of modus ponens. For consider a proposition q that is true at the dull world, but which is false in actuality; and take a proposition p that is true at both worlds (p might be that the Sun is roughly spherical, and q might be that the Sun and the Moon differ in apparent diameter.) Now $p \mathrel{\Box\!\!\rightarrow} q$ will be true, since it is true at the closest worlds where p obtains—i.e. at the dull worlds. So, at the actual world, we have p, $p \mathrel{\Box\!\!\rightarrow} q$ and $\neg q$—an inconsistent triad, given modus ponens.

The second potential problem is the following. Since the presence of quasi-miracles make for dissimilarity, worlds selected by counterfactuals will be as dull as possible *ceteris paribus*. Suppose that we consider what would happen if I were to shift my leg a little to the left. Now, since the avoidance of quasi-miracles must take priority over exact intrinsic match, the closest worlds where I shift my leg must be dull ones. So, the following counterfactual would be true: were I to shift my leg, the world would be dull. The point sounds really bad when we pick particular examples: were I to shift my leg, the Moon and Sun would not be of the same apparent diameter from the Earth.

Third, if quasi-miracles are abundant in actuality, it is not clear why the presence of quasi-miracles at a world should count as a respect in which the world is *dissimilar* from reality. If the absence of quasi-miracles is not a respect of similarity to the actual world, then by including this in his account of the ordering on worlds which fixes the truth conditions of counterfactuals, Lewis would have abandoned the claim to be analyzing counterfactuals *in terms of* the similarity of worlds.

Response

Hawthorne cites a variety of surprising facts and takes them to constitute quasi-miracles in the actual world. In doing so, he is presupposing (1) that they are remarkable (2) that they are improbable.

Consider, by way of illustration, the fact that the Sun and Moon have the same apparent diameter as viewed from the Earth. Is this remarkable? Certainly those ignorant of the difference of size of the two objects involved are unlikely to remark upon it. However, we have already seen that (because of the division problem) the Lewisian should not characterise 'remarkableness' of facts in terms of the knee jerk reactions of the folk; but rather, what well-informed opinion would count as surprising. And in this case, it is reasonably plausible that the fact under consideration will count as remarkable.

Granted this, they will indeed be quasi-miraculous if they are improbable. But to worry over whether or not such events are really improbable would be mere skirmishing if we accept the basic move deployed in the division argument earlier. Typically, we will be able to represent any remarkable event in the actual world as a disjunction of improbable events. By the reasoning of the division argument, given that the event itself is remarkable, so will be the various ways it can come about. Pick the one which is realised in the actual world. This particular way is uncontroversially low-probability, but *ex hypothesi* still remarkable.

There is, however, a proviso to be mentioned. The chances of events happening change over time. Before I toss a coin, there is a 50/50 chance of it coming down heads. After we have flipped it, and it has come down heads, the chance of it doing so (on that very occasion) is 1.[10] This is so no matter how unlikely the outcome originally was. So,

[10] For discussion, see Lewis (1980).

for example, the chance, at t, that the universe unfolds in the way it actually did up to t, is 1, even if the odds against that particular sequence of events happening were initially astronomical. Partitioning will not reinstitute the problem here.

The moral is that remarkable events are low-probability or high-probability, quasi-miraculous or not, only relative to a choice of time (presumably, a time fixed by the event figuring in the antecedent of the counterfactual).

It is false to say, without qualification, that we will find a quasi-miracle in actuality wherever we find some remarkable event, even one that was intuitively 'unlikely'. What is true is that, under any usual reading of 'remarkable' we will be able to find an abundance of quasi-miraculous events in the actual world in the time following t, relative to the chances at t. There looks little hope of resisting the thought that the actual world will contain an abundance of quasi-miracles (i.e. in the future, relative to present chances). Some damaging effects may be allayed by this observation: but the fundamental point—that the counterfactually nearest worlds will not contain quasi-miracles relative to any time, remain. So we can still expect to face versions of the criticisms sketched above.

Let us consider, then, the supposed consequences in turn. First, the threat that the actual world might not be the closest world to itself. The Lewisian might give short shrift to this worry. One of the formal features of orderings of worlds given in Lewis (1973) is "centering". This just *is* the requirement that any world w is closer than any other to w. So if centering is in place, there is no room for this complaint. [11]

The second problem, in its strongest form, threatened that under any minor coun-

[11] One might worry that this merely shifts the problem: we have now the worry that Lewis' account of similarity is inconsistent with the formal demands (such as centering) that he puts on the notion. But it is independently plausible that nothing can be more similar to a world than that world itself. Since this is the case, the Lewisian can simply add in the insistence that each world is closer than any other to itself, as part of the analysis of similarity between worlds, trumping all other concerns including the presence of quasi-miracles. Since such a clause sustains ordinary counterfactual judgements, the addition of a clause can be motivated both analytically and instrumentally.

terfactual hypothesis, we would have a world without the various remarkable, hence quasi-miraculous, features of our world. But the most damaging versions—the threat that "if I shift my leg, the Sun and Moon will not coincide in apparent diameter from the Earth" will come out true—can now be seen as flawed, since they ignore the time-relativity of quasi-miracles. Relative to the time at which I shift my leg, the chance that the Sun, Moon, and Earth are in their actual relative positions is 1; hence this remarkable feature of the world is not quasi-miraculous. Nevertheless, we still have odd results: for example, that under minor counterfactual assumptions, the future will be duller than the past. This seems wrong.

What of the third potential problem? This questions whether, if the actual world is full of quasi-miracles, we should class a world as less similar to our own in virtue of the quasi-miracles in it. Here, I think, Lewis must bite a bullet, and see the analysis of closeness of worlds in terms of quasi-miracles, not as an analysis of a pre-theoretically recognisable notion of similarity, but instead as instrumentally justified in getting a notion of closeness of worlds going that will deliver the right truth conditions for counterfactuals.

In sum, the Lewisian need not worry that Hawthorne's considerations will lead to a revisionary logic for counterfactuals. However, the Lewisian approach appears to have the following consequences: (1) under trivial counterfactual suppositions, it will be the case that the future is duller than the past; and (2) the ambition to analyze counterfactuals in terms of a *non-technical* notion of similarity would have to be given up. These are substantial costs for the Lewisian to accept.

The remarkable subpattern problem

If enough events happen, then it would be surprising if we didn't find remarkable patterns arising somewhere. If we had billions of duplicate fair coins, and we simultaneously and independently flipped each 10^6 times, then the probability of one of these coin-flipping sequences being "all-heads" tends to 1, as the number of coins (and hence, total number of flips) increases. [12]

Suppose the world contains N duplicate fair coins. Under the counterfactual assumption that each is flipped a million times, let $f_1, \ldots, f_N$ enumerate the sequences of heads and tails that respectively result. We can choose N large enough to make it likely that one of the f_i (sequence of one million coin flips) will turn up all-heads. Consider the following:

> (J) If N coins were each flipped 10^6 times, then none of the coins would come down heads every time.

This is clearly a false counterfactual—by construction, it would have been pretty likely that one of the f_i is all-heads. However, Hawthorne takes it that Lewis' account commits us to each of the following:

> (K_i) If N coins were each flipped 10^6 times, then the ith coin would not come down heads every time.

A sequence of a million coin-flips landing heads each time is, after all, just the kind of remarkable and unlikely event that constitutes a quasi-miracle. Since quasi-miraculous worlds are *ipso facto* further away than non-quasi-miraculous worlds, all

[12] The probability of none of these sequences is all-heads is $(\frac{2^{1000000}-1}{2^{1000000}})^n$. With n as 1, this figure is close to 1. As n increases, it tends to 0. Thus, the complement—the probability that one of the sequences of flips is all-heads—is, in the limit, 1.

the closest worlds where the N coins are each flipped 10^6 are ones where the results in a typical, random sequence of heads and tails—in particular, a sequence other than all-heads.

Now the principle of agglomeration, which Lewis' formal treatment of counterfactuals sustains, says the following:

> AGGLOMERATION
> If $A \mathrel{\Box\!\!\rightarrow} B$ and $A \mathrel{\Box\!\!\rightarrow} C$ then $A \mathrel{\Box\!\!\rightarrow} (B \wedge C)$.

Using this and the K_i, we get:

> (K) If N coins were each flipped 10^6 times, then none of the coins would come down heads every time.

But this is just J, which we already agreed was false. Lewis's account entails J; but J is false; so Lewis's account must be wrong.

Response

Let us symbolise the K_i as $(P_1 \wedge \ldots \wedge P_N) \mathrel{\Box\!\!\rightarrow} \neg H(f_i)$, where P_i stands for the hypothesis that the ith coin is flipped 10^6 times and $H(f_i)$) for the outcome of the series of flips of the ith coin being all-heads. The problem was that from these, and the principle AGGLOMERATION, we can derive $(P_1 \wedge \ldots \wedge P_N) \mathrel{\Box\!\!\rightarrow} (\neg H(f_1) \wedge \ldots \wedge \neg H(f_N))$. But this counterfactual is unacceptable.

I can see what the form of a response to this puzzle should be: one should reject each of the K_i, and explain away whatever intuitions there are in favour of them. I will sketch below what seems to me to be an attractive way of implementing this idea. It would be good for the Lewisian if her theory of counterfactual similarity of worlds would support the story I give—but whether or not Lewis' quasi-miracles analysis can do this is another matter, one I will discuss at the end of this section.

We should distinguish the K_i from the following:

> (L_i) If the ith coin were to be flipped 10^6 times, then it would not have come down heads every time.

In the notation above, the L_i can be formulated as $P_i \mathrel{\Box\!\!\rightarrow} \neg H(f_i)$. I take it that these are paradigmatic examples of conditionals that should come out true on a non error-theoretic treatment of counterfactuals and chance (at the least, I cannot see any way of developing a non error-theoretic account that does not render them true).

The truth of the L_i is quite compatible with the rejection of the K_i. To raise problems on the basis of L_i, we would need, in addition to AGGLOMERATION, the following principle:

> From $A \mathrel{\Box\!\!\rightarrow} B$ and $C \mathrel{\Box\!\!\rightarrow} D$, infer $A \wedge C \mathrel{\Box\!\!\rightarrow} B \wedge D$

But this is *not* supported by the standard logic of counterfactuals.[13]

The form that a response to the subpattern problem should take, I suggest, is the following: to declare the L_i are acceptable, but the K_i are not. But can Lewis's quasi-miracles support this classification?

Again, the reasoning of the division problem will make consideration of whether or not the pattern is low probability redundant. Any particular pattern will be extremely unlikely; so the particular one exemplified in a particular counterfactual world will be low-probability. The question of whether or not the patterns in question are quasi-miraculous

[13]Nor should it be. A counterexample would be: (1) If I were to disconnect the wiring, the bomb would be disarmed. (2) If I were to press the button, the bomb would go off. (1) and (2) could both be true, but we certainly should not infer (3) If I were to disconnect the wiring and press the button, the bomb would be disarmed and go off. The principle mentioned is of a kind with putative rules such as antecedent-strengthening (From $A \mathrel{\Box\!\!\rightarrow} C$, infer $A \wedge B \mathrel{\Box\!\!\rightarrow} C$) and transitivity (From $A \mathrel{\Box\!\!\rightarrow} B$ and $B \mathrel{\Box\!\!\rightarrow} C$ infer $A \mathrel{\Box\!\!\rightarrow} C$), either of which would allow us to infer (J) from the L_i—but which are notoriously invalid in conditional logics.

is thus a matter of whether they are remarkable. Surely, the situations depicted by the L_i are remarkable—these are worlds where a coin is flipped a million times, and comes up heads every time that it is tossed. But is the same true of the situations depicted by the K_i? Is a string of 10^6 heads, in the context of many trillions of coin-flips, remarkable?

We have already noted that remarkability of events needs to be understood as an information-relative notion, if Lewis' account is to be tenable at all. Given this, it is clearly *not* remarkable that there exists such a sequence somewhere within such an enormously long run of coin-flips: the well informed agent would see this as a statistical inevitability. So Lewis could reply to Hawthorne that *in the context of trillions of coin-flippings*, a sequence of a million heads is not remarkable, and not quasi-miraculous. If this is conceded, Lewis would be able to resist endorsing K_i. As sketched above, this gives rise to a stable position where, in particular, there is no commitment to the absurd (J).

Lewis' opponent will then distinguish the remarkableness, or otherwise, of *there being some sequence of a million heads* within a long coin-flipping sequence, from *this particular sequence of coin-flips turning up all-heads*. The former, he will concede, is not remarkable. However, the latter still seems remarkable: even if you know that *somewhere* coincidences will occur, it can still be remarkable that one occurs *right here*. Since it is particular *events* (such as segments of coin-flippings) that are classified as remarkable or not, rather than worlds as a whole, it seems that the Lewisian will have to regard each run of 1,000,000 heads as a quasi-miracle. The upshot is that the Lewisian is indeed committed to each K_i, and thus to the absurd J.

The problem of the exclusion of the more probable

Take some particular unremarkable sequence of coin-flips. Call it S. Consider the following:

(G) If you were to flip the coin 10,000,000 times, you would produce a sequence other than S

(H) If you were to flip the coin 10,000,000 times, you would produce a sequence other than all tails or all heads

(I) Producing a sequence of all tails or all heads is twice as likely as producing S.

On Lewis' account, (H) is true, since it concerns a remarkable outcome. But (G) is false, since the outcome it concerns is unremarkable—one of the closest worlds is an S-producing world. Hawthorne claims that these instantiate an uncomfortable pattern.

Response

I shall argue that resisting *this* problem does not require the Lewisian accept any new costs.

The alleged problem requires us to find the following claims in tension:

- $\neg[A \mathrel{\Box\!\!\rightarrow} B]$
- $A \mathrel{\Box\!\!\rightarrow} C$
- $A \mathrel{\Box\!\!\rightarrow}$ (B is at least as likely as C)

But what kind of tension is it that Hawthorne discerns here? Perhaps it is that it sounds odd to *assert* all three in the same context. Call this the *weak reading* of Hawthorne's criticism of Lewis. A much more damaging objection to Lewis would be the claim that H, I and $\neg G$ should be seen as *incompatible* propositions, whereas Lewis wrongly represents them as *compatible*. Call this the *strong* reading of Hawthorne's objection.

If Hawthorne's point is captured by the weak reading, it is not clear why the Lewisian should be worried. To begin with, notice that this criticism would not *directly* attack Lewis's approach, for Lewis's theory aims to tell us which counterfactuals are *true*, not which are *assertible*. To be sure, our total account of counterfactuals should, amongst other things, have something to say in explanation of the assertibility or non-assertibility of certain combinations of counterfactuals. However, for all we have so far said, the *overall* account of counterfactuals embedding the Lewisian analysis may include such an explanation. At this stage, all we can say is that *if* Lewis' account is right, the explanation of our reluctance to assert $H \wedge I \wedge \neg G$ will *not* be that the conjunction is false. If that is a cost of the Lewisian approach, it seems to me a minor one.

What certainly needs to be defused is Hawthorne's objection under the stronger reading. If can be maintained that H, I, and $\neg G$ not only fail to be jointly unassertable, but are positively *incompatible*, this would be a serious—even decisive—blow to Lewis's position. Here I respond directly: I think it is demonstrably the case that H, I and $\neg G$ are compatible.

We can reformulate the claim that G, H and $\neg I$ are incompatible as the claim that the following is a valid inference pattern:

EXCLUSION RULE

1. $A \mathrel{\Box\!\!\rightarrow} C$
2. $A \mathrel{\Box\!\!\rightarrow} (B$ is twice as likely as $C)$
3. Therefore: $A \mathrel{\Box\!\!\rightarrow} B$

The form of the objection is notable. It does not take the form of picking out a specific counterfactual, about which we can argue that Lewis' account delivers the wrong verdict. Here, we are rather asked to make judgements about the logical relations between counterfactual judgements in the abstract.

In cases where the relevant antecedents are actualised, we have straightforward counterexamples to the exclusion rule.[14] Suppose that we are about to toss a pair of dice, and in fact they will land on snake-eyes. Nevertheless, at the time at which they are tossed there is a 1/36 chance of them landing snake-eyes, and a 35/36 chance of them landing with some other combination of faces. By the centering assumption, we have:

- dice tossed $\square\!\!\rightarrow$ landing snake-eyes
- dice tossed $\square\!\!\rightarrow$ (dice landing in a combination other than snake-eyes is (more than) twice as likely as the dice landing snake-eyes)

But we do not have the result the exclusion rule would force upon us:

- dice tossed $\square\!\!\rightarrow$ dice will land in a combination other than snake-eyes

That this refutes the exclusion rule as stated is, I think, undeniable. But one might wonder whether some suitably refined version might still hold good. One might hold out hope for a version of the exclusion rule restricted to truly *counterfactual* cases—cases where the antecedent is false. One modification would simply be the following:

EXCLUSION RULE*

1. $\neg A$
2. $A \;\square\!\!\rightarrow C$
3. $A \;\square\!\!\rightarrow$ (B is twice as likely as C)
4. Therefore: $A \;\square\!\!\rightarrow B$

[14] I am extremely grateful to an anonymous referee for pointing to these.

If this is valid, it is still enough to underpin Hawthorne's criticism of Lewis, and so further investigation is called for.

Many invalid principles of conditional logic look good in abstract—transitivity and antecedent strengthening are classic examples. Most relevant to our discussion, the following rule looks plausible to many people when they first come across it:

CHANCE RULE[15]

1. $\Diamond A$
2. $A \,\Box\!\!\rightarrow$ non-zero chance that B
3. Therefore: $\neg(A \,\Box\!\!\rightarrow \neg B)$

However plausible this sounds in the abstract, it is, as Lewis shows, disastrous in action. Consider, for example, the counterfactual supposition that there is an unrealised chance that p, i.e. $q :=$ (non-zero chance that p) $\wedge \neg p$. Clearly, we should have the counterfactuals

$q \,\Box\!\!\rightarrow$ non-zero chance that p
$q \,\Box\!\!\rightarrow \neg p$

But in the presence of the chance rule these are contradictory.[16] One moral to be taken is that intuitions, in the abstract, about the validity of patterns of inference about coun-

[15]The first premise is needed because, on the standard Lewis-Stalnaker system, counterfactuals with impossible antecedents are vacuously true. *A fortiori*, for impossible A, $A \,\Box\!\!\rightarrow ch(B) > 0$ and $A \,\Box\!\!\rightarrow \neg B$ will both be vacuously true.

[16]For this argument, and a more complex one that does not involve counterfactuals with chancy vocabulary in their antecedents, see Lewis (1979, p.65).

Notice that the chance rule will also have centering-based counterexamples: suppose the dice will in fact land snake eyes. Then we have both: 'dice tossed $\Box\!\!\rightarrow$ landing snake-eyes' and 'dice tossed $\Box\!\!\rightarrow$ non-zero chance of landing in a combination other than snake eyes', yet the antecedent is clearly not impossible. As with the exclusion rule, therefore, we really need to restrict attention to instances of the chance-rule where the antecedent is *false*.

Notice that in Lewis's counterexample, the antecedent need not be actually true—so this is a counterexample to a weakened 'chance rule' where $\neg q$ is added as an extra premise.

terfactuals are unreliable. But we can go further: the exclusion rule itself falls to the same counterexamples as the chance rule.

Notice that the exclusion rule immediately gives us the following restricted version of the chance rule:[17]

RESTRICTED CHANCE RULE

1. $\Diamond A$
2. $A \mathrel{\Box\!\!\rightarrow} [ch(B) < 1/3]$
3. Therefore: $\neg(A \mathrel{\Box\!\!\rightarrow} B)$

The proof is as follows:[18]

1.	$\Diamond A$	Premise	1
2.	$A \mathrel{\Box\!\!\rightarrow} [ch(B) < 1/3]$	Premise	2
3.	$A \mathrel{\Box\!\!\rightarrow} [ch(\neg B) > 2 \times ch(B)]$	Probability theory, Logic, 2	2
4.	$A \mathrel{\Box\!\!\rightarrow} B$	Supposition for reductio	4
5.	$A \mathrel{\Box\!\!\rightarrow} \neg B$	3,4, Exclusion rule	2,4
6.	$\neg\Diamond A$	4,5 logic	2,4
7.	Contradiction	1,6	1,2,4
8.	$\neg(A \mathrel{\Box\!\!\rightarrow} B)$	Reductio on 4	1,2

A slight modification of Lewis's counterexample to the unrestricted chance rule rebuts the restricted chance rule, and hence the exclusion rule from which it follows.

[17] A weakened version of the exclusion rule adding the premise $\neg A$, will give rise to a weakened version of the restricted chance rule below, with added premise $\neg A$, as can be easily checked.

[18] The probability theory in step 3 just uses the fact that $ch(\neg p) = 1 - ch(p)$; thus if $ch(p) < 1/3$, $ch(\neg p) > 2/3$. Step 7 relies on some facts about the principle of conditional non-contradiction: stating that $p \mathrel{\Box\!\!\rightarrow} q$ and $p \mathrel{\Box\!\!\rightarrow} \neg q$ are incompatible. The only exceptions to conditional non-contradiction in Stalnaker-Lewis conditional logics occur where both conditionals are vacuously true: thus from a pair of "opposite" conditionals we can derive within these conditional logics that the antecedent is impossible.

Consider the proposition that some unlikely (chance 0.3) event will in fact occur (i.e. $ch(p) = 0.3 \wedge p$). This proposition is not impossible, and further we have both:

$$(p \wedge [ch(p) = 0.3]) \mathrel{\Box\!\!\rightarrow} p$$
$$(p \wedge [ch(p) = 0.3]) \mathrel{\Box\!\!\rightarrow} [ch(p) < 1/3]$$

But this is then a counterexample to the restricted chance rule. Notice again that the antecedent here can be supposed not to obtain at the actual world; so this is a counterexample that works against the reformulated EXCLUSION* rule as much as the original exclusion rule. [19]

In sum: Hawthorne's objection under the *weak* reading does not yet do enough to

[19]Indeed, if our logic contains the principle of *conditional excluded middle* (schematically, $A \mathrel{\Box\!\!\rightarrow} B \vee A \mathrel{\Box\!\!\rightarrow} \neg B$) then we can derive the original chance rule from the exclusion rule. The crucial principle, entailed by conditional excluded middle, is:

$$(A \mathrel{\Box\!\!\rightarrow} (B \vee C)) \leftrightarrow (A \mathrel{\Box\!\!\rightarrow} B) \vee (A \mathrel{\Box\!\!\rightarrow} C)$$

We need in addition a further assumption: that there are four mutually exclusive and exhaustive propositions p_1, p_2, p_3, p_4, such that $A \mathrel{\Box\!\!\rightarrow} [ch(p_i) = 0.25]$ (we can choose these to be propositions about when a certain radioactive atom will decay, for example. E.g. p_1 might be that a certain atom of Uranium 235 will decay before t; p_2 that it will decay between time t and time t', p_3 that it will decay between time t' and t'', and p_4 will decay after t''. With an appropriate choice of intervals, these propositions will meet the stated conditions.)

Note then that, by probability theory and logic we will have for each i, $A \mathrel{\Box\!\!\rightarrow} [ch(p_i) < 1/3]$ and hence $A \mathrel{\Box\!\!\rightarrow} [ch(B \wedge p_i) < 1/3]$. For possibly true A, we can apply the restricted chance rule to our previous result that $A \mathrel{\Box\!\!\rightarrow} [ch(B \wedge p_i) < 1/3]$, to derive $\neg(A \mathrel{\Box\!\!\rightarrow} (B \wedge p_i))$, for each i.

Moreover have: $\Box(B \leftrightarrow ((B \wedge p_1) \vee \ldots \vee (B \wedge p_4)))$, and from conditional excluded middle we have:

$$(A \mathrel{\Box\!\!\rightarrow} ((B \wedge p_1) \vee \ldots \vee (B \wedge p_4))) \leftrightarrow ((A \mathrel{\Box\!\!\rightarrow} (B \wedge p_1)) \vee \ldots \vee (A \mathrel{\Box\!\!\rightarrow} (B \wedge p_4)))$$

putting these together we can derive:

$$(A \mathrel{\Box\!\!\rightarrow} B) \leftrightarrow ((A \mathrel{\Box\!\!\rightarrow} (B \wedge p_1)) \vee \ldots \vee (A \mathrel{\Box\!\!\rightarrow} (B \wedge p_4)))$$

But we have already proved the negation of each disjunct on the right hand side; so we can conclude that the left hand side is false. That is, we can conclude $\neg(A \mathrel{\Box\!\!\rightarrow} B)$, as required.

In sum: the unrestricted chance rule can be derived from the restricted chance rule in the presence of conditional excluded middle. And we have already seen that the latter can be derived from the exclusion rule. So in the presence of conditional excluded middle, the exclusion rule enables one to derive the unrestricted chance rule. which we have already seen to be unacceptable.

Conditional excluded middle is a feature of Stalnaker's logic of conditionals, but Lewis rejects it.

embarrass the Lewisian story; but under the *strong* reading, the criticism cannot be sustained.

The state of play.

We have looked at Hawthorne's four problems in turn, and I have tried to give the best-possible Lewisian responses. Though resisting the exclusion problem brings at most minor costs to the Lewisian theory, in the other three cases, substantial costs are incurred. Resisting the division problem requires we introduce counterfactual elements into our general analysis of counterfactuals: vitiating the original Lewisian strategic ambitions for the theory. The abundance problem has at least two worrying consequences: it seems to vitiate Lewis' *analytic* ambition, to be giving an account of counterfactuals in terms of an intuitive notion of similarity; and further, under insignificant counterfactual assumptions, we have that no remarkable events would happen in the future, even though they are abundant in the past. The remarkable subpattern problems shows that the Lewisian is committed to a series of counterfactual judgements (the K_i), which collectively entail an absurd result.

I take it then, that Hawthorne's objections show that the Lewisian account of counterfactuals is in bad order. The response, however, should be to develop a better Lewisian analysis, more in keeping with the original non-chancy analysis, and not to abandon the account altogether. In the remainder of the paper, I outline one such approach.

3 Fit, Typicality and Randomness

Fit between worlds and laws

Within Lewis's overall system, the notion of possible worlds "fitting" with laws of nature comes into play twice. Firstly, as we have seen, a component determining the counterfactual similarity of worlds to actuality is that they maximise fit with the actual laws. Secondly, "fit" is given a central role in fixing the laws of nature that obtain in a given possible world: under Lewis's Humean "best systems analysis" of laws of nature, the laws of nature of world w are that axiomatic system that optimises simplicity, predicative power, and fit with the facts of w.[20]

In each case, there is the challenge to explicate what this notion of "fit" means when the laws of nature are chancy. The simple proposal that w fits with L when w contains no violations of the laws in L does not sufficiently constrain the relation for these theoretical purposes. A rule that says that the chance of a coin coming up "heads" is 0.5, and one that says that the chance is 0.6 are equally simple and have equal predicative power: but in a world where the relative frequency of heads in a billion coin-flips is 0.6, a Humean will wish the "best system" to embed the latter rather than the former as a law. Moveover, we saw at the beginning of this paper the difficulties that the simple "no violation" proposal causes for counterfactuals.

Lewis endorses different ways of patching his treatment of "fit" in the two cases. As we have seen, in the case of counterfactual similarity, he supplements his account by appeal to a new factor influencing whether or not worlds are similar to actuality: whether they contain remarkable improbable events. In the case of laws of nature, however, he tries a more direct patch, replacing the "no violation" analysis of fit with something else.

[20] Lewis' description of his overall system can be found in the introduction to Lewis (1986).

The new analysis of fit appeals to the probability that the world arises, given the laws of nature. The proposal is that w fits with L to the extent that L makes w probable. Take a world like that described above, consisting of a billion coin flips, with a relative frequency of heads to tails of 0.6. A law that assigns a chance 0.6 of an arbitrary flip resulting in "heads" assigns a far greater probability to this result, than a putative "law" that assigns chance 0.5 to the same event.

Lewis's patch in the case of counterfactual similarity is attacked—ultimately successfully, I have argued—by Hawthorne. But the patch in the case of laws of nature also has problems. The "zero-fit" problem (Elga, 2004) is that, when we deal with worlds that comprise infinitely many trials of the events in question, the probability of any given outcome is likely to be zero by the lights of the intuitively "correct" laws. For example, in a world which contains infinitely many flips of a fair coin, the probability assigned to the actual outcome will be zero.[21] The result is that the *probability of an infinite world by the lights of laws of nature* does not do the work we wanted it to do: it does not favour the law assigning chance 0.6 to a coin turning up "heads" over one assigning 0.5 to describe a world containing an infinite series of coin flips with limiting relative frequency of heads to tails of 0.6.

Hawthorne's attack on Lewis's use of "remarkableness", and the Elga zero-fit problem for Humean view of chancy laws, together demonstrate that the Lewisian must go back to the drawing board to get a decent theory of "fit" covering the chancy case. Clearly, the most attractive and economical route would be to find a single way of elucidating "fit" that would address the problems with both theories. I believe this can be done. In what follows, I will argue Elga's proposals for addressing the zero-fit prob-

[21]Keeping the laws of nature constant, all particular outcomes of a finite series of flippings will be equiprobable, which is the reason why Lewis's gloss of "fit" in the case of laws of nature doesn't look promising even initially as a gloss of "fit" for the case of counterfactual similarity.

lem can be used to formulate an analysis of counterfactual similarity that is immune to Hawthorne's attacks.

Typicality

Let us revisit the basic intuitions. Given the information that a coin is fair, how should you expect a long enough series of coin-flips to turn out? Well, you should not attempt to predict the particular outcome: the detail of how things go at each point is a matter of pure chance. But you can formulate some general expectations. You will expect, for example, that the limiting relative frequency of heads to tails will be 0.5. You will expect the sequence HTH to turn up just as often as THT. You will expect that looking at every other result will give you a sequence of coin flips just as "disordered" as the original series.[22]

What you can legitimately expect is that the outcome of the series of coin flips will have certain global properties. In short, you expect the outcome to fall within a broad class of outcomes: the typical ones.

I submit that requiring a world to be "typical" by the lights of its laws of nature is a legitimate explication of the constraint that the world "fits" those laws. Allowing this explication, we can address both the "zero-fit" concerns about the Humean theory of chancy laws, and (I shall argue) the Hawthorne puzzles over chancy counterfactuals. On the former point, Elga (2004) suggests the following. We admit that a variety of assignments of chances to coin flips are all (a) compatible with the actual outcome of infinitely many coin flippings; and (b) all the assignments assign the same (zero) chance to the particular outcome of the coin flippings. Nevertheless, the actual outcome is only typical according to some of these putative 'chancy laws'. For the outcome of an infinite

[22] Compare Elga (2004, p.72).

series of coin flippings to be intuitively typical, the limiting relative frequency of heads to tails must equal the single-case chance of the coin landing heads rather than tails. This immediately narrows down possible "best systems" to plausible candidates.

One can see also how such a notion will help us analyze counterfactual similarity. For a world to be optimally close to actuality, it will have to be *typical* by the lights of laws of nature: the hope will be that worlds where dropping a plate leads to it shooting off sideways will count as *atypical* by the lights of the chancy laws; and hence further away from actuality than worlds where such improbable coincidences do not arise.

Typicality as an objective feature of outcomes

No progress will have been made, however, if to invoke "typicality" is merely to invoke "remarkableness" in a new guise. And it might be suggested that a typical sequence, is *just* one which a well-informed agent would find *unremarkable*. I do not claim that such a 'projective' understanding of typicality is *inconsistent* with ordinary usage of that term. I maintain, however, that a non-projective property of typicality can be identified, and that this is the one of interest in the present context.

I contend that we can identify the required notion of typicality (relative to an assignment of chances) with the mathematical property of a set of outcomes being *random* (relative to an assignment of chances).[23] The identification of typicality with randomness should, on reflection, seem plausible. In the special case of the outcome of a series of coin flips, for example, randomness has all the characteristics we wanted typicality to have. It is concerned to pick out, from amongst a series of equiprobable outcomes, a class within which we can expect the outcome to fall. It is a holistic property of a whole series of outcomes—a random sequence will contain local patterning, so long as they

[23] In this I follow Elga (2004) and Gaifman and Snir (1982).

are swamped by the overall disorder. A sequence is not random if it exhibits biases—if, for example, in the long run the limiting relative frequency of heads to tails does not match the chances.

Randomness thus seems to have the *formal* features that we want typicality to have. And, I believe, intuitive judgements or how typical an outcome is by and large coincide with intuitive judgements of how random it is (by the lights of the governing probability function). That every typical sequence will be random seems to me beyond dispute; and that every random sequence will be typical seems highly plausible.[24]

For present purposes, we need not require some reductive analysis or mathematical definition of the notion of randomness: all we need do is convince ourselves that there is some such property that outcomes of chancy processes can objectively possess or fail to possess. We can then leave the question of whether this feature is reducible or must be taken as primitive for another occasion; our use of it will go through either way. However, it may help to look in a bit more detail at one option for giving a mathematical characterization of typicality/randomness: in particular, that found in Elga (2004) and Gaifman and Snir (1982).[25]

The basic idea common to the proposals of Elga (2004) and Gaifman and Snir (1982) is to look, not at the probability of a *particular* outcome arising, but at the probabilities

[24]It has been suggested to me that outcomes of a sequence of fair coin-flips which exhibit overall biases towards heads, say, may still count as random, so long as the outcomes are still intuitively "disordered". It certainly does not seem *obvious* to me that such sequences should count as random; and this is the kind of issue on which I would defer to whatever a best overall theory of randomness says on the issue (the von Mises-style mathematical characterization of randomness comes down *against* such sequences being random). It may be that the intuitive notion of 'randomness' is vague enough that there are a variety of tractable notions that would count as precisifications of that concept. If so, it is enough for my purposes if typicality turns out to be *one* (objective) precisification of randomness. As we shall see in the discussion below, one plausible candidate precisification is the target of the formal work of Gaifman and Snir (1982).

In connection with this, see Eagle (2005) for arguments that the usual formal characterizations do not match our ordinary concept of randomness (notice that for all that is argued in that paper, the formal characterizations may still succeed as a *direct* characterization of typicality).

[25]For a survey of some other attempts at analyzing randomness mathematically, see Eagle (2005).

of a suitable set of *properties* which that outcome instantiates.[26] When considering the outcome of flipping a fair coin, 'all heads' is a low-likelihood property (in the infinite case, it is probability 0 that the outcome has this property). 'Having as many heads as tails, in the long run' is a high-likelihood property: in the infinite case, it is probability 1 that the outcome has this property. The general theme is that an outcome is random to the extent that it possesses high-probability properties.

This gives us a general sense of what the objective randomness of an outcome (relative to a probability function) could consist in. It is the general idea behind the work of Gaifman and Snir (1982), who show how to shape this into a formal characterization: they define an outcome as typical iff it renders true all the probability-1 sentences within a certain class C. (By semantically ascending in this way—formulating the constraint in terms of *sentences* rather than properties or propositions—the definition becomes language-relative. For our purposes, it is best to remove this relativity by requiring the sentences to be formulated in a certain *canonical* language.[27])

Gaifman and Snir make a persuasive case that, for a range of choices of C, their definition gives an intuitively adequate characterization of typicality/randomness (so long as the outcomes are infinite in extent). In a sense, however, they give us an embarrassment of riches. What they deliver is not typicality *simpliciter*, but rather a whole hierarchy of relativized notions of C-typicality, corresponding to stricter and stricter

[26]To see why it is important that we restrict attention to a *suitable range* of properties, consider the following. Each outcome whatsoever of a sequence of coin-flippings possesses one 'exact distributional' property, specifying the result of each flip. But having this exact distributional property is low-probability (in the infinite case, it will be probability zero). Thus, if we required that a random outcome have nothing but high-probability properties, *no* outcome would be random. We need to find a way to select appropriately 'discriminating' or 'test' properties. More of this below.

[27]With Elga (2004), we may suppose that the canonical language will be one whose only predicates stand for 'perfectly natural' metaphysical or microphysical properties. The use of such restricted languages within the overall Lewisian project is familiar: it is for example, used within the Humean account of laws of nature, to make good a notion of theoretical simplicity. (Famously, Lewis argues that the invocation of perfectly natural properties is indispensable for a great range of projects. See Lewis (1983).)

versions of the intuitive notion of typicality.[28] The hierarchy starts by taking C to be quantifier-free sentences, and in ascending the hierarchy we add in sentences of greater quantificational complexity. Each choice of level in the hierarchy will give *a* candidate notion of typicality. But we need just one![29]

One challenge, therefore, is to get from the range of candidate notions of C-typicality that Gaifman and Snir provide, to a single unrelativized notion of worlds being *typical*, or *more typical than one another*. Three options are: (i) to pick out one level of complexity of sentences, and so identify typicality *simpliciter* with some particular C-typicality;[30]; (ii) to say that "typicality" simpliciter is a *vague* notion, with the various C-typicalities being the *precisifications* of this notion; (iii) to extract from the hierarchy of C-typicalities a principled *ordering* of worlds as more or less typical.

Elga (2004) takes approach (i). He identifies the Σ_2-sentences as the privileged class; consequently identifying typicality *simpliciter* as Gaifman/Snir's Σ_2-simplicity.[31] But he offers no reasons why this identification is particularly plausible. Approach (ii) appears less *ad hoc*. However, notice that for all that has been said so far, this notion may be strictly stronger than *any* of the Gaifman/Snir C-typicality properties, and for

[28] See Gaifman and Snir (1982, §5) for an account of how extant characterizations of randomness in the mathematical literature find their place within this hierarchy.

[29] As mentioned above, the Gaifman and Snir (1982) characterisation of randomness is not the only proposed mathematical characterisation of this notion, but it is uniquely interesting for our purposes. Unlike other proposed mathematical analyses of randomness their proposal not restricted to idealised coin-flipping scenarios, but can be applied to *arbitrary* possible worlds featuring the outcomes of probabilistic processes.

The Gaifman and Snir (1982) approach may be seen as a generalization of the von Mises-style characterisation there discussed, though the authors also connect it to the complexity theory approach deriving from Kolmogorov. Both the von Mises and the Kolmogorov approaches make essential appeal to properties of *linear sequences* of chance events, making them insufficiently general for our purposes.

[30] A variant on this approach would be to let the particular choice of C-typicality be a context-sensitive matter.

[31] The class of Σ_2 sentences are those that are equivalent to a sentence consisting of a block of existential quantifiers, followed by a block of universal quantifiers, followed by a quantifier-free formula.

that reason, we might pause before endorsing it.[32] It would be very nice, therefore, if we could find a workable version of option (iii). Leuenberger (2005, p.13) contains an elegant suggestion that can be adapted to the present situation.[33]

Thus, there are a number of options for focussing the 'embarrassment of riches' that Gaifman and Snir provide into something that can do the work we want it to. I shall take it, therefore, that this provides for a formally tight and intuitively adequate characterisation of objective typicality/randomness *in the case of infinitary outcomes*. There is a major qualification, however: the Gaifman/Snir characterizations do not give satisfactory results in finitary worlds. To be random in the Gaifman-Snir sense, a sequence need only to instantiate a suitable test set of probability-1 properties. In the finitary case, this doesn't adequately constrain matters: indeed, *every* sequence will have a finite chance of arising, and hence must instantiate all probability-1 properties. It follows that on the Gaifman-Snir definition, every finite sequence is random.

One response is to try to generalize the Gaifman/Snir characterization. Elga attempts to do so, by saying that we have to compare outcomes according to the probabilities assigned to the simple properties the outcomes respectively instantiate. So, for example, since the relative frequency of heads to tails in an 'all heads' finite outcome is *less prob-*

[32]To be determinately typical in this sense, a world would have to be C-typical for *every* choice of quantificational complexity C. Gaifman and Snir's work gives us no guarantee that this is not too strong.

[33]Let us first index the quantificational hierarchy by ordinals, so that Σ_i sentences are indexed by i, etc. (Leuenberger is not explicit about how this indexing is to proceed: there are probably a number of workable options.) Leuenberger then defines the *rank* of a world-theory pair:

$$\rho(T,w) := \sup\{\alpha : \alpha = ord(C) \wedge \text{w is C-typical relative to T}\}$$

We can now order a class of worlds W (by the lights of a probability function T) as follows: w is *less typical* than w' iff $\rho(w',T) < \rho(w,T)$.

Thus, worlds which are (intuitively) not typical at all, will not satisfy the probability-1 sentences at *any* non-zero level. Worlds which are intuitively typical will satisfy the probability-1 sentences of a certain bounded quantificational complexity. And within these, one world will be more typical than another if it continues to pass the tests 'further up' the hierarchy than the other.

able than that the relative frequency that occurs in an intuitively disordered outcome, on Elga's account the former is to that extent less random than the latter. However, Elga does not attempt a *formal* characterization of these 'fit comparisons'.

Perhaps Elga's characterization can be made precise, and will enable a general characterization of the typicality of sequences, finite and infinite. If so, we can appeal to it in the current case. But as noted above, we are not *obligated* to offer a reductive characterization of typicality/randomness in order for that notion to be available in characterizing counterfactual similarity. All we need is the concession that our intuitive judgements of the *randomness* or *typicality* of a finite sequence (relative to a set of chancy laws governing the generation of the sequence) are based, not just on projections of what we find remarkable, but on objective features of the sequences. We learn from the work of Gaifman and Snir that *infinite* sequences can be objectively random or non-random. What would it be to remain sceptical over whether there is an objective notion of randomness in the finitary case? The suggestion would have to be that our impressions of objective disorder, which in the infinite case track an objective feature of the outcomes, systematically deceive us in the finite case. To me, this suggestion seems grossly implausible. Whether or not we have a mathematical characterisation to hand, therefore, we have reason to think that there *is* some such feature for mathematical characterisations of randomness to aim at.

4 Hawthorne's worries reconsidered

We can generalise Lewis' original definition of counterfactual similarity to remove the reliance on a particular explication of 'fit' between laws and worlds that is appropriate only to the non-chancy case. We reach the following characterisation:

SIMILARITY**
w_1 is more similar than w_2 to the world w_0 if the differences between w_1 and w_0 are of less weight than the differences between w_2 and w_1. The weighting of the differences is governed by the following principles:

1. It is of the first importance to avoid big, widespread, diverse lack of fit with laws of nature
2. It is of the second importance to maximise the spatio-temporal region throughout which perfect match of particular fact prevails.
3. It is of the third importance to avoid even small, localised lack of fit with laws of nature.
4. It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly

We get the original formulation SIMILARITY when we fill in an analysis of 'fit' in terms of *no violation of the laws of* w_0. The suggestion to be investigated here is that where laws of nature are chancy, we instead fill in an analysis of 'fit' in terms of a world's being *maximally typical by the lights of the laws of* w_0. Filling this in, we arrive at:

CHANCY SIMILARITY
w_1 is more similar than w_2 to the world w_0 if the differences between w_1 and w_0 are of less weight than the differences between w_2 and w_1. The weighting of the differences is governed by the following principles:

1. It is of the first importance to minimize atypicality of the world as a whole, by the lights of the chancy laws of nature of w_0.
2. It is of the second importance to maximise the spatio-temporal region throughout which perfect match of particular fact prevails.
3. It is of the third importance to minimize even small, localised, atypicalities by the lights of the laws of w_0
4. It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly.

The first thing to be ascertained is that this gives rise to a non-error theoretic account of standard counterfactuals. Consider our paradigmatic counterfactual, (E):

If I had dropped the plate, it would not have flown off sideways.

On the typicality approach, quantum events conspiring to send the plate flying off sideways would constitute at the very least an atypical local space-time region, by the lights of the actual laws. Thus, due to clause (3) in the characterization of chancy similarity, such worlds will be more distant from actuality than those where (as expected) the plate falls to the floor and breaks.[34]

We can now reconsider Hawthorne's objections to the Lewisian approach. I shall not revisit the *exclusion problem* here, since I believe that an adequate response to Hawthorne is already available: the rule of inference on which the problem rests can and should be given up, on independent grounds. For the other three puzzles, I contend that in each case, the objections that undermine Lewis's analysis of chancy counterfactual similarity do not cause problems for the typicality account.

The Division problem avoided.

Recall the scenario: a monkey is currently so-configured that there is a relatively high probability that if it starts typing, it will produce a readable dissertation. The problem was that on the 'quasi-miracle' treatment of chancy counterfactuals, we could argue that 'if the monkey started typing, then it would not produce a dissertation' was true. To escape this, it looked like we would have to give up Lewis's ambition to explicate counterfactual similarity without appealing within the analysis (circularly?) to counterfactuals.

Lewis says that quasi-miracles are remarkable, low-probability outcomes. The division argument, in effect, aimed to show that the 'low-probability' aspect of this defini-

[34]Though see the 'lucky runs' concern below.

tion is redundant. Is there any corresponding worry for my favoured analysis in terms of typicality?

I suggest not. The whole point of typicality is that low-probability events (even probability 0 events) can be differentiated as typical or atypical, by looking at whether or not the overall pattern of events instantiate highly-probable *properties*.

In the case at hand, one would have to argue that any outcome having the property of *featuring a monkey producing a dissertation* is atypical. The way to do this is to argue that the property *featuring a monkey producing a dissertation* is the kind of low-probability property of outcomes which typical sequences should not instantiate. But the case at hand is precisely *not* one where this is the case: *ex hypothesi* the chances at the time relevant to assessing the counterfactual are set up so that there is a high probability that the outcome will have the property *featuring a monkey producing a dissertation*.

The remarkable subpattern problem defused

We are to consider a situation where we have an enormous number of fair coins, none of which are flipped. We are asking about counterfactual scenarios in which some or all of the coins are flipped a million times each. To avoid refutation, we must avoid commitment to:

> (K_i) If N coins were each to be flipped 10^6 times, then the ith coin would not have come down heads every time.

For these counterfactuals would entail the unacceptable result that *no* coin in the series of flips would have come down heads every time; whereas if we choose N large enough, we can make it almost inevitable that there would be *some* such 'all-heads' coin.

We distinguished the (K_i), above, from some related counterfactuals which are the kind of thing that our approach should declare true:

> (L_i) If the the ith coin were to be flipped 10^6 times, then it would not have come down heads every time.

I contend that the typicality account of counterfactual similarity makes the K_i false and the L_i true.

The point is a simple one. The counterfactual scenarios envisaged by the (L_i) are ones where only one coin is flipped. In such a world, the result that each flip lands heads amounts to a hugely atypical (non-random) sequence. The counterfactual scenario envisaged by each of the (K_i) involve a sequence of flippings of fair coins, with length $N \times 10^6$. N has been chosen long enough to make it almost inevitable that there is, somewhere in this sequence, a run of a billion heads. With that as antecedent, there is nothing atypical in any of the situations described by the consequent: there are optimally typical worlds wherein the ith coin lands heads every time. Each (K_i) is refuted by one such world.

The underlying difference between this approach and the Lewisian 'remarkableness' story, is that remarkableness is presented as a property of *local events*. Any consecutive string of a billion heads in flips of a fair coin is remarkable. We have no principled grounds for taking into consideration the wider setting, in which we see that the presence of such a string is a predictable feature of the world. By contrast, typicality (randomness by the lights of the chancy laws of nature) is a global property of the pattern of events that make up a possible world. In the context of the overall pattern of events, remarkable local facts can be required by to overall typicality.

The abundance problem defused

Quasi-miracles were remarkable, low-probability events. Hawthorne claims there are likely to be many of these in reality, and on this basis derives problems. On our revised approach, the potential abundance of *quasi-miracles* is of no relevance, since they play no role now in the analysis of counterfactuals. The analogous concern however, is this: what if the actual world is not as typical as it might be?

If we look at the revised version of SIMILARITY we see that the thing that is of the first importance in assessing whether a world is close to actuality, is whether overall it *fits* with actual laws of nature: and in the present context, this is the requirement that it be typical by the lights of those laws. Now, it seems indeed to be a possible that some non-actual world will be *more typical* by the lights of the actual laws of nature, than is actuality itself. Suppose that the actual (finite) outcome of a sequence of coin flips is a slight departure from the relative frequency predicted by theory. Then, by the lights of SIMILARITY**, a world which matches exactly the relative frequency predicted by theory is *ipso facto* closer than one that does not.

As in the previous case, three consequences threaten: (a) the actual world being less close to itself than some other world. (b) under slight counterfactual assumptions, the future will be typical even while the past remains atypical. (c) the thesis that counterfactuals are being analyzed in terms of a recognizable notion of similarity is threatened.

(a) is, I think, adequately dealt with in just the way that the original Lewisian response handled it: there is no need to think that (a) follows from the abundance result. The analogue of (b) was more discomforting: it seemed to commit us to the claim that if I had moved my leg slightly a second ago, the future would be 'duller' (less remarkable) than the past. The corresponding result in the present case, however, seems unproblematic. Suppose the past to have been atypical by the lights of the actual laws of nature. Then we are committed to the claim that if I had moved my leg slightly a second ago, then the future would be typical (and more typical than the past). But this is intuitively right: if the laws of nature are as we have assumed them to be, then the future would proceed in the way that those laws predict. It may be true that in this world simple-mindedly projecting past *trends* into the future would justify a *belief* that the future will exhibit similar trends, while the counterfactual truth is that such 'atypical' trends *would not continue*. However this is merely to note the fact that the sort of scenario under consideration is *counterinductive*, and gives no reason for thinking the relevant counterfactual false.

In respect of problem (c), the typicality account is again in a strong position. The problem, recall, was that if the *actual* world contained lots of remarkable, unlikely events, then in the intuitive sense, worlds similar to it should also contain remarkable, unlikely events. 'Dullness' of worlds makes them positively *dissimilar* to actuality in an intuitive sense: yet on Lewis's view dullness will place a world closer to actuality than it otherwise would have been.

Let us immediately note that the typicality account does not place *dull* worlds closer to actuality than others. It is only to be expected that in various places, surprising coincidences will occur. Indeed, a lack of surprising coincidences would be an *atypical* feature of a world like ours: hence, if anything, dullness of a world will increase the distance between that world and actuality, on the typicality account of the similarity metric.

But even if the original problem is avoided, does an analogous one emerge? The analogous problem would have to be this: if the actual world is atypical, then atypical worlds intuitively should be *ipso facto* more similar to reality than typical worlds.

This should be resisted. Recall the Lewisian concentrates *aspects* of similarity, rather than intuitions about similarity 'all things considered'. So *simply* pointing to a general feature of the actual world that is shares with another world (so that they are 'in some sense' similar) is not enough to cause problems for the approach.

In the non-chancy case we have a principled story about the two aspects of similarity that go into the similarity metric: one aspect is exact match to actuality on matters of particular fact; the other aspect is fit with the actual laws of nature. In the quasi-miracles case, we could not appeal to this story to explain the sense in which the 'similarity metric' captures genuine *similarity* of worlds, since quasi-miracles weren't invoked either as part of a general analysis of what it is to fail to 'fit with' the actual laws of nature, or as part of what it is to fail to match actual matters of particular fact. All Lewis tells us is that intuitively quasi-miracles 'detract from similarity with the actual world'. One can perfectly reasonably challenge this directly: on the grounds that *if* the actual world contains an abundance of quasi-miracles, then it is not at all clear why their presence should make for dissimilarity rather than similarity with the actual world. Note that it is the *departure* from the principled "intrinsic-match + fit-with-laws" story that makes this response available.

But in the revised case no such departure arises. The whole point is that typicality is proposed as an *explication* of what it is to fit chancy laws of nature. So it is perfectly reasonable to assess worlds as similar to actuality to the extent that they are typical *by the lights of the actual laws*.

The problem of nearby lucky runs

There is one problem, however, against which our previous moves will not help. In the discussion of the remarkable subpattern problem, we supposed that the world initially contained no coin-flippings, so that in relevant counterfactual scenarios the only relevant chance events were those described in the antecedent of the counterfactual. Our problem arises when we drop this assumption.

Let us now suppose the actual world contains infinitely many flips of a fair coin. Call any sequence where the coin comes down heads a million times in a row a *lucky run*. Since the world contains infinitely many flips of a fair coin, it is chance 1 that a sequence of a million heads occurs somewhere in the world (indeed, it is chance 1 that there will be infinitely many such sequences).[35] So it is chance 1 that the world contains lucky runs.

Now suppose in addition that I could have flipped the fair coin an additional million times, but decided not to. Characteristically, the non error-theoretic approach wants to say that the following counterfactual is true: "were I to have flipped the coin a million times, it would not have landed heads each time".

However, is a world where my coin flipping turns up heads each time more distant from actuality than one where the outcome is less surprising? Crucially, in the case at hand, the run of a million heads does not render the world as a whole less typical—such occasional lucky runs are a statistical inevitability. The difference between worlds

[35]To see that a sequence of a million heads has chance 1 of occurring, consider a die with 2^{10^6} sides. Roll it infinitely many sides. It is chance 1 that at some point, the very first side happens. But each side of this die can be conceived as a particular outcome to a sequence of one million coin flips, with the very first side corresponding to all-heads. And to flip a coin infinitely many times is exactly equivalent to rolling this die infinitely many times. So it is chance 1 that an all-heads sequence will happen at some point. The same argument shows that, given some n and ε we can choose an m such that the chance that a sequence of n heads occurs somewhere in a sequence of m flips of a fair coin is within ε of 1.

where my counterfactual coin flipping produces a typical outcome and one where it produces a sequence of all-heads is just in the location of the lucky run: in one case the atypicality occurs *when I flip the coin*, in other worlds the lucky run occurs somewhere else. But if these worlds differ only in *location* of lucky runs and not in how well the worlds match actual fact or fit with actual laws of nature, it looks like they will be equidistant from actuality. The result of this would be that "were I to have flipped the coin a million times, it would not have landed heads each time" would be *false* on the Lewisian analysis: for there would be an optimally close antecedent-world where the consequent fails.

This problem could be posed directly for the plate-dropping counterfactual (E). In a world where plate-dropping events occur infinitely many times, it is chance 1 that some plate-droppings will be followed by the plate flying off sideways. Again, the difference between counterfactual worlds where I drop my plate and it breaks, and counterfactual worlds where I drop my plate and it flies off sideways, is merely a matter of whether one of these rare but inevitable occurrences occurs *around here*.

Worlds where a lucky run occurs during *my* coin-flipping, and those worlds where a lucky run occurs *somewhere else*, are exactly on a par as regards (a) global typicality and (b) the fact they contain localised atypicalities. Similarly for worlds where the plate flies off sideways when *I* drop it in comparison to worlds where all plate-flying-sideways events happen to other people—again, CHANCY SIMILARITY seems to rate them equidistant from actuality. It looks, therefore, as if CHANCY SIMILARITY cannot help rescue the above counterfactuals.

The challenge is a serious one, but I think that a response is available. I suggest we make a small adjustment to CHANCY SIMILARITY, in effect relativizing the similarity of worlds to our current *perspective*.

We can take it that each counterfactual statement is made in a context where a particular spatio-temporal location is salient. E.g. when considering the counterfactual 'if I were to flip this coin one million times', the salient location is *here, now*. What we must do is hold that 'localized atypicality' are worse (make worlds more dissimilar) when they are *atypicalities over a region containing the salient location.* This ensures that the closest counterfactual worlds are ones where any 'lucky runs' are in spatio-temporal regions *distant* from the events invoked in the counterfactual. As anticipated, the similarity relation becomes context-sensitive; variations in which location is salient may change whether it is w_1 or w_2 that is closest to actuality.[36]

The final proposal for a neo-Lewisian analysis of chancy counterfactuals is thus the following:

> CHANCY SIMILARITY*
> w_1 is more similar than w_2 to the world w_0 if the differences between w_1 and w_0 are of less weight than the differences between w_2 and w_1. The weighting of the differences is governed by the following principles:
>
> 1. It is of the first importance to minimize atypicality of the world as a whole, by the lights of the chancy laws of nature of w_0.
> 2. It is of the second importance to maximise the spatio-temporal region throughout which perfect match of particular fact prevails.
> 3. It is of the third importance to minimize even small, localised, atypicalities by the lights of the laws of w_0 *especially when the atypical*

[36]In typical cases, the salient location will be the location of the events described by the antecedent of the counterfactual. But there are various cases where this would be difficult. Some antecedents will not describe located events at all (If there had never been any elephants...); some antecedents will describe multiple events located at different places (If Nixon had pushed the button and Brezhnev had invaded Western Europe...); sometimes the consequent will make salient a location that differs from that of the antecedent (If Adam had not eaten the apple, I would be living in Eden). There are several ways of handling such cases. Perhaps the simplest is to say that in these circumstances it will be *indeterminate* which location is salient; and consequently indeterminate what the relevant similarity ordering of worlds is. Given this, for a counterfactual to be unambiguously true it must be true on *every* way of resolving the indeterminacy. The apparatus of supervaluations (Lewis, 1970; Fine, 1975) can be deployed to this end. (Invoking supervaluations to resolve indeterminacies in the similarity ordering has precedent: see in particular (Stalnaker, 1984, pp.134-5).)

region is one that contains the salient location.

4. It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly.

I contend that this package avoids all of Hawthorne's worries and the concern just raised. Furthermore, it is an implementation of the original Lewisian idea of relating similarity to intrinsic match of particular fact and fit with laws of nature.

5 Conclusion

Hawthorne contends that the problems he adduces require either a contextualism about counterfactuals, or an abandonment of the Lewisian picture altogether. I think that the issues that Hawthorne raises do show that Lewis' rather *ad hoc* invocation of quasi-miracles has highly problematic consequences, and should be rejected. I have argued that this, the main thrust of his paper, is correct.

The diagnosis of the problem I have urged is that at the time that Lewis tried to extend his account of counterfactuals to the chancy case, he had not a worked out his Humean account of chancy laws of nature. Once this is in place, and understood in terms of typicality of a world by the lights of as actual laws, the analysis I have offered is a straightforward generalisation of the original account.

I contend, therefore, that the basic idea in Lewis' approach extends naturally to an account of chancy counterfactuals which does not fall to the objections that Hawthorne develops. Not only does this account avoid or ameliorate the specific points that Hawthorne makes, it is independently motivated and yields desirable conclusions.

Given that best empirical theory seems to tell us that the actual laws of nature are chancy, a credible theory of counterfactuals is of critical importance. Lewis's original analysis fails on this account; but the positive ‘typicality’ account of counterfactual similarity offered here remedies this deficiency.

References

Eagle, A. (2005). ‘Randomness is unpredictability’. *British Journal for Philosophy of Science*, **56**, 749–790.

Edgington, D. (1995). ‘On conditionals’. *Mind*, **104**, pp. 235–329.

Elga, A. (2004). ‘Infinitesimal chances and laws of nature’. *Australasian Journal of Philosophy*, **82**, 67–76.

Fine, K. (1975). ‘Vagueness, truth and logic’. *Synthese*, **30**, 265–300. Reprinted with corrections in Keefe and Smith (eds) *Vagueness: A reader* (MIT Press, Cambridge MA: 1997) pp.119-150.

Gaifman, H. and Snir, M. (1982). ‘Probabilities over rich languages, testing and randomness’. *The Journal of Symbolic Logic*, **47**(3), 495–548.

Hawthorne, J. (2005). ‘Chance and counterfactuals’. *Philosophical and Phenomenological Research*.

Leuenberger, S. (ms. April 2005). ‘The place of chance in a Humean metaphysics’.

Lewis, D. K. (1970). ‘General semantics’. *Synthese*, **22**, 18–67. Reprinted with postscript in Lewis, *Philosophical Papers I* (Oxford University Press, 1983) 189–229.

Lewis, D. K. (1973). *Counterfactuals*. Blackwell, Oxford.

Lewis, D. K. (1979). ‘Counterfactual dependence and time’s arrow’. *Noûs*, **13**, 455–76. Reprinted with postscript in Lewis, *Philosophical Papers II* (Oxford University Press, 1986) 32–51. Also reprinted in Jackson (ed) *Conditionals* (Oxford University Press, 1991) 46-76.

Lewis, D. K. (1980). ‘A subjectivist’s guide to objective chance’. In R. Jeffrey, editor, *Studies in Inductive Logic and Probability*, volume II, pages 263–93. University of California Press. Reprinted with postscript in Lewis, *Philosophical Papers II* (Oxford University Press, 1986) 83–113.

Lewis, D. K. (1983). ‘New work for a theory of universals’. *Australasian Journal of Philosophy*, **61**, 343–377. Reprinted in Lewis, *Papers on Metaphysics and Epistemology* (Cambridge University Press, 1999) 8–55.

Lewis, D. K. (1986). *Philosophical Papers*, volume II. Oxford University Press, Oxford, New York.

Lewis, D. K. (1999). *Papers on Metaphysics and Epistemology*. Cambridge University Press, Cambridge.

Stalnaker, R. (1984). *Inquiry*. MIT Press, Cambridge, MA.