This is a preprint of an article whose final version appears in *The Journal of Philosophy*, accepted 2024.

The Branchpoint Proposal and the Role of Counterfactuals

Alison Fernandes

I introduce a novel method for evaluating counterfactuals. According to the branchpoint proposal, counterfactuals are evaluated by 'rewinding' the universe to a time at which the antecedent had a reasonable probability of coming about and considering the probability for the consequent, given the antecedent. This method avoids surprising dynamics, allows the time of the branchpoint to be determined by the system's dynamics (rather than by context) and uses scientific posits to specify the relevant probabilities. I then show how the branchpoint proposal can be justified by considering an evidential role for counterfactuals: counterfactuals help us reason about the probabilistic relations that hold in a hypothetical scenario at which the antecedent is maximally unsettled. A result is that we should distinguish the use of counterfactuals in contexts of control from their use for reasoning evidentially. Standard Lewisian accounts run into trouble precisely by expecting a single relation to play both roles.

Acknowledgments: For helpful feedback, comments and suggestions, my warm thanks to Heather Demarest, Erica Shumener, Christian Löw, Isaac Wilhelm, Michael Hicks, David Albert and Barry Loewer, as well as two anonymous reviewers at this journal. This paper grew out of comments presented at the Society for the Metaphysics of Science Annual Conference (2018) responding to a paper by Chris Dorst and a paper, 'Evidence-based Science' presented at the online pre-read workshop Digital Metaphysics organised by Kadri Vihvelin (2020).

What would be the case, were Hamish to drink the wine? The correct answer depends on what 'counterfactuals' are true: what would be the case, were something in the world to be a certain way.[1] Counterfactuals seem to be reasonably objective. They are the kind of thing we might be wrong about and have empirical evidence concerning. Yet, regarding counterfactuals with false antecedents, they don't seem to be directly settled by events in the actual world. They are, at least, not settled merely by what happens when the antecedent is true. As with many modal relations, which are *prima facie* about the possible rather than the actual, two puzzles then arise. Why do we care about counterfactuals and why should we evaluate them in any given way?

The standard approach to justifying methods for evaluating counterfactuals is to argue that the counterfactuals that result match our intuitive judgements or the use of the English subjunctive conditional (Lewis 1973, 1979; Bennett 2003). Some also argue that the method should be

[1] Counterfactuals, as I use the term here, can have true or false antecedents.

temporally neutral, making no explicit reference to the past or future (Lewis 1979; Loewer 2007, 2012; Albert 2000, 2015). If a method is temporally neutral, it seems any asymmetries in what counterfactual are true will be due to asymmetries in the world, rather than the method, so that the method can be used to trace temporal asymmetries in counterfactuals, and in derivative relations like causation, back to empirical features of the actual world.

While recovering our intuitive judgements might be a starting point, this standard approach fails. First, it doesn't explain *why* we care about counterfactuals or *why* we should evaluate them in a given way. Second, it leaves open the possibility that a method has simply been reverse-engineered to deliver intuitive results. Reverse-engineering risks delivering incorrect results whenever our intuitions are unreliable. It also provides no deeper understanding of why the results delivered are appropriate and so why counterfactuals matter in contexts such as decision-making (Lewis 1986c) or the analysis of causation (Lewis 1986b). For related concerns, see Bennett (1984), Horwich (1987, p. 172), Woodward (2003, p. 137) and Price and Weslake (2009). Third, the standard approach doesn't ensure that a method is appropriate for explaining temporal asymmetries. There are many temporal asymmetries in the world. Latching onto one of these using a temporally neutral method is not enough to ensure that we have latched onto the one that explains the relevant temporal asymmetry. To ensure we latch onto the right asymmetry, a method requires justification.

To justify a method of evaluating counterfactuals, we need to show that it delivers counterfactuals that answer to some specific role—not merely *that* counterfactuals matter to decision-making, for example, but *how* they matter. The aim is not to deliver a conceptual analysis,truth conditions of natural language counterfactuals or our intuitive judgements. Instead, the aim is to identify a precise role for counterfactuals and determine what method is appropriate, given this role.

Justifying a method by appeal to the role of counterfactuals is to adopt a broadly 'functionalist' approach to the metaphysics of science, one that uses the role of scientific relations to give accounts of them. This methodology is naturalistic, in the sense of Quine's (1995) methodological naturalism, as interpreted by Verhaegh (2018): one adopts science when doing metaphysics. Given our modal discourse is in order, we would, on naturalistic grounds, expect our modal discourse to serve some function. So, there are naturalistic grounds for expecting modal relations to serve specific functions, which can then be used in giving accounts of them. Sometimes a focus on the function of modal terms has been associated with anti-realism and a reluctance to provide accounts (Price 2013; Ismael 2017). But a term or relation having a function does not imply that the relation is less than fully real or that one can't provide an account. For any relation that we are realist about, we can ask why we reason about the world using that relation. The fact that a relation exists can only ever be part of the answer—we also need to identify its role.

There is reason to think this functionalist methodology will deliver specific methods for evaluating counterfactuals. Distinguish between 'physically fundamental' modal relations (those that feature in how fundamental physical theories are formulated) and 'higher-level' modal relations (those that only feature elsewhere in science).[2] If one is convinced by Russell's arguments (1912−13) that causal relations don't feature in fundamental physics and Cartwright's arguments (1979) that causal relations are needed in other sciences then one should think that causal relations are higher-level relations. Similar arguments are plausible in the case of counterfactuals. Counterfactuals don't feature in how candidate fundamental physical theories are formulated. Yet they are used in explanations in higher-level sciences (Strevens 2012). If one takes fundamental physics to be the science that can explain the success of higher-level sciences, there must be some way of specifying what fundamental physical conditions obtain in the actual world when a given counterfactual is true—which is all that giving a method of evaluating counterfactuals requires.

In the following, I identify an evidential role for counterfactuals. Counterfactuals direct evidential reasoning and evidence gathering. More precisely, counterfactuals indicate the probability of the consequent, conditional on the antecedent and further states in a hypothetical scenario where the antecedent is maximally unsettled (Section 2). Provided certain conditions are met, counterfactuals that play an evidential role are also relevant in contexts of control. But not always. We need to distinguish the relations used in contexts of control from those used for reasoning evidentially. We should not expect a single analysis of counterfactuals to play both roles (Section 4).

I use the evidential role of counterfactuals to justify the following method of evaluating them. This method can be employed whether or not a given counterfactual involves backtracking—where the

[2] I take a complete fundamental physical theory to be one that is universal in scope and can explain the success of all other scientific theories. We don't yet have a candidate fundamental physical theory, but we can consider what form it will take.

counterfactual world differs from the actual world prior to the antecedent. While I justify this method by appeal to the role of counterfactuals, the method is not subjective and counterfactuals have a healthy degree of objectivity. According to the 'branchpoint proposal', we 'rewind' the universe to a time at which the antecedent had a reasonable probability of coming about, conditional on certain states up to and at that time. Counterfactuals are then evaluated by considering the probability of the consequent, given the antecedent and those states. While accounts using 'forks' of various kinds have been defended before (Lewis 1979; Edgington 1995, 2004; Bennett 2003; Leitgeb 2012), none makes use of the probabilistic machinery I employ here to identify branchpoints and none adequately identify a distinct role for counterfactuals (Section 2). Nor do previous accounts explain the temporal asymmetry of the method (Section 3).

There are several advantages to the branchpoint proposal. First, unlike Lewisian accounts, the branchpoint proposal avoids surprising dynamics at both the micro and macrolevel—there are no Lewisian miracles or surprising macroscopic dynamics. Second, unlike altered-states recipe approaches (Maudlin 2007a, Ch. 1; Paul and Hall 2013, pp. 47–53; Dorst 2022), the branchpoint proposal establishes the time of the branchpoint using the system's dynamics, rather than by using context—implying a greater degree of objectivity to counterfactuals. Third, the proposal uses scientific posits to specify the relevant probabilities, offering better prospects for scientifically explaining temporal asymmetries and providing a unified treatment of 'might' and 'would' counterfactuals. All three features help ensure the scientific credentials of counterfactuals and their fit within a naturalistic metaphysics of science.

Before I begin, some disclaimers. First, absent a finished fundamental physical theory, I will work with higher-level candidate approximations, including classical Boltzmannian statistical mechanics and the GRW version of quantum mechanics. Second, talk of the past, present and future is to be treated indexically (typically relative to the time of the antecedent), rather than as referring to metaphysically distinct regions of time. Third, talk of past and future *directions* in time is to be understood as referring to the directions in which various temporally asymmetric phenomena lie, rather than intrinsic directions of time. Finally, I will only discuss counterfactual antecedents that are events or states of affairs regarding particular times and places. Considering other antecedents will require dealing with further difficulties (Edgington 1995, pp. 322−3). I don't offer the branchpoint proposal as a method of evaluating *all* counterfactuals.

The paper proceeds as follows. In Section 1 I outline the branchpoint proposal for evaluating counterfactuals and consider some of its advantages. In Section 2, I justify the proposal by considering an evidential role for counterfactuals. In Section 3, I argue that the temporal asymmetry in the method reflects a temporal asymmetry in the universe's probabilistic structure. In Section 4, I distinguish evidential and control roles for counterfactuals and argue that it is a mistake to look for a single relation to play both roles.

1. The Branchpoint Proposal

Consider the following counterfactual. 'If Hamish were to drink the wine, he would get a headache'. Say Hamish doesn't drink the wine. Barring an unusual setup, the truth of the counterfactual seems to depend on facts like the constitution of Hamish and the wine and lawlike regularities concerning their relation. It doesn't depend on how Hamish or the wine came to be there. Many methods can get 'forwards-looking' counterfactuals like this right. Provided a method alters the present state in a minimal way to satisfy the antecedent and evolves it forwards in a lawlike way, the method will correctly determine the truth of the counterfactual.

But here is a more difficult case—one where what happens counterfactually depends on how the antecedent came about. These kinds of counterfactuals have played less of a role in recent literature, because many accounts aim to minimise backtracking on the way to delivering causal relations (Section 4). But it is worth considering whether we can give a systematic account of them.

Consider what would be the case, were your friend Clare to fail to meet you at the café, given that she meets you in the actual world. Plausibly, her not being there would have implications for future states of affairs—perhaps you would wait before leaving. Clare failing to be there would also, plausibly, have implications for *past* states of affairs and, via these, further implications for the future. Perhaps if Clare weren't there, this would be due to some previous mishap—a past state that would lead to future states such as her messaging you. After all, the following counterfactual sounds plausible: 'If Claire had failed to meet you, you would have received a text from her shortly thereafter, explaining why.' A way to put this thought is in terms of inference to the best explanation: when evaluating counterfactuals concerning Clare, we sometimes reason to a past state that would have explained her absence, in the sense of implying that her absence was reasonably probable and more probable than it would otherwise be. Plausibly, we also expect the past state we reason to, to be reasonably probable. We don't reason that Clare's absence would be due to her inexplicably swerving past the café or uncharacteristically snubbing you—even though those states would also lead to her absence.

Evaluating counterfactuals in this way countenances backtracking. Backtracking is required to keep the counterfactual world evidentially 'well-behaved', with the same lawlike regularities we observe in the actual world. There are plausibly limits to how far such backtracking will go. One is unlikely to reason that, if Clare had been absent, she would not have been your friend or that cafés would not have existed. Instead, we look to states nearer in time that had some reasonable probability of occurring and that would have led, with sufficient probability, to her absence.

Here is a method of evaluating counterfactuals that attempts to capture this informal reasoning, while still getting cases like Hamish's right—thus offering a unified treatment of forwards-looking and backtracking counterfactuals. Call this the 'branchpoint proposal'. Consider a time t, the 'branchpoint', that is simultaneous with or as close as possible previous in time to the time of the antecedent (t_A), such that the antecedent occurring (A) and the antecedent failing to occur (¬A) are both 'reasonably probable' (see Figure 1). These probabilities are probabilities conditional on *certain* states up to or at t—'Prior'. In other words, both P(A|Prior) and P(¬A|Prior) must be reasonably probable. I'll discuss Prior and the interpretation of these probabilities below. For now, note the probabilities are objective conditional probabilities ('chances') that are not relativized to times. In Clare's case, the branchpoint is the time at which her being at the café (or not) are both reasonably probable (given Prior)—plausibly just before a potential mishap might befall her. What counts as 'reasonably probable' will depend on the system's dynamics. For many ordinary cases, 'reasonably probable' is a probability of around 0.5. More precisely, one 'rewinds' to a branchpoint such that the probability of A and ¬A are both as close to 0.5 as possible, provided there are no later times at which the probability of A or ¬A is as close to 0.5. In cases where states never have a probability as high as 0.5, lesser probabilities will therefore suffice, right down to a limiting case where A or ¬A is very improbable.

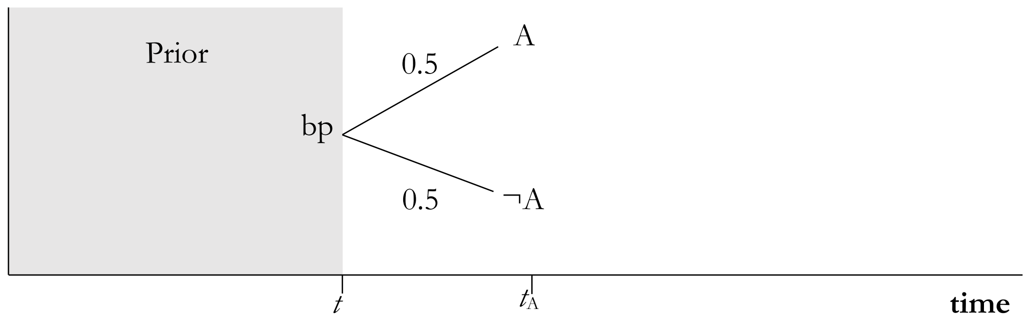

Figure 1: To evaluate counterfactuals using the branchpoint proposal, consider a time *t* (the branchpoint, 'bp') at which A and not-A are both reasonably probable, where the probabilities are conditional on certain states up to and including *t* (Prior).

According to the branchpoint proposal, the following counterfactuals are true: 'If A were to be the case, C would have a probability of P(C|A.Prior)'. 'If A were *not* to be the case, C would have a probability P(C|¬A.Prior)'. *Mutatis mutandi* for ¬C. Counterfactuals always have probabilistic consequents. Counterfactuals with non-probabilistic consequents are approximations. The counterfactual 'If A were to be the case, C would be the case' can be treated as roughly true just in case P(C|A.Prior) is very high—in that case, A 'probabilistically settles' C.[3] See Appendix 1 for some brief notes on the logic of these counterfactuals. What counts as 'very high' and 'sufficiently high', here and below, will depend on the system's dynamics and contextual features. I won't attempt to settle the standards for these approximations. For ease, I will typically use 'consequent' to refer to the state (C) featuring in the consequent. Antecedents and consequents may concern macro or microstates.

Often, we will also be interested in whether A counterfactually 'makes a difference' to C—what would standardly be captured by counterfactual dependence. The precise relations that the branchpoint proposal delivers are: a) A is *positively* counterfactually relevant to C just in case Prob(C|A.Prior) > Prob(C|¬A.Prior); b) A is counterfactually *relevant* to C just in case Prob(C|A.Prior) ≠ Prob(C|¬A.Prior). Because A and ¬A form a partition and P(A|Prior) and

---

[3] Leitgeb defends a similar condition using 'approximate truth' (2012, p. 56). The main difference between our accounts is that I use the dynamics of the system to determine the time of the branchpoint, rather than using a temporal reference point (ibid., pp. 63–9). The approximation conditions also play less of a role on my account, since I don't justify counterfactual methods by their ability to deliver ordinary English counterfactuals.

P(¬A|Prior) are both greater than 0, a) and b) also hold for Prob(C|Prior) on the right-hand side. In other words, whenever A is (positively) counterfactually relevant to C, A changes (or increases) the probability of C compared to not conditionalising on A. There are also approximations: 'C counterfactually *depends* on A' can be treated as roughly true just in case P(C|A.Prior) and P(¬C|¬A .Prior) are both 'sufficiently high'.[4] While the precise counterfactuals the branchpoint proposal delivers are probabilistic, I will sometimes talk of 'dependence' for grammatical ease.[5]

Applied to Hamish's case, whether Hamish would get a headache (H), were he to drink the wine (W), depends on whether the probability of his getting a headache given he drinks the wine is sufficiently high. This probability is evaluated conditional on states up to and at the time when his drinking the wine is maximally unsettled—plausibly when he's deciding whether to drink. Let's say that P(H|W.Prior) is sufficiently high. Then drinking the wine probabilistically settles that Hamish will get a headache. If we're looking for useful ways to reason about the world, we're also likely to be interested in whether drinking the wine *changes* the probability of his headache (compared to not drinking the wine). If W doesn't change the probability of H, then we needn't look for clues as to W to figure out H—W is irrelevant, and H is settled (or not) by other states. If P(H|W.Prior) > P(H|¬W.Prior), Hamish's drinking the wine is *positively* counterfactually relevant to his getting a headache.

Applied to Clare's case, we can recover the informal reasoning above. For details, see Appendix 2. To evaluate what would happen, were Clare to fail to arrive (F), we first identify a time (*t*) at which Clare's failing to arrive (F) or not (¬F) are both reasonably probable (given Prior). We then consider what other states would be probable, given her failure (F). Let us say that, given Prior, Clare's failing to arrive would imply, with high probability, that some mishap (M) had befallen her. Since F is reasonably probable (given Prior) and since M is highly probable given F (and Prior), M must itself

[4] Standardly, 'If A were to be the case C would be the case' is true if A and C are both true in the actual world (Lewis 1979). Because I take counterfactuals to concern conditional probabilities, I don't adopt that assumption here. C can only counterfactually *depend* on A when A is positively relevant to C. The standard assumption leads to odd consequences in cases of chancy outcomes: the outcome counterfactually depends on any prior event (no matter how unrelated) (Bennett 2003, p. 240; Edgington 2004, p. 17). Loewer (2007) rejects the assumption, as does Lewis (1986b, p. 176; 1986c, p. 334) in the case of indeterministic laws.
[5] See Edgington (1995, pp. 249−50; 2004), Bennett (2003, pp. 252−4) and Leitgeb (2012, p. 56 ff) for semantics using probabilities to indicate approximate truth or degree of acceptability, sometimes by building in a contextualist parameter. See Hájek (2014) for concerns with Leitgeb's account and Hájek (2021) for more general concerns with contextualist accounts.

be at least approximately reasonably probable (given Prior). In other words, the branchpoint will be a time at which the mishap *may well* befall Clare. Given the above conditions, the relevant mishap will imply that her failure is at least reasonably probable (given Prior) (Appendix 2, Equation 4). Altogether, we reason to past states that are reasonably probable and that imply that the antecedent was reasonably probable. This recovers part of the informal reasoning above.

However, the informal reasoning also included that the past state implies that the antecedent was more probable than it would otherwise be. If this informal reasoning is to be applicable in Clare's case, we need a further assumption: that Clare's failure is positively counterfactually relevant to an earlier mishap. Let's assume from now that this assumption holds. The mishap then implies that Clare's failure to arrive was more probable than it would otherwise be (Appendix 2, Equations 5−8). This recovers the final part of the informal reasoning above. Whether a given past state satisfies the above conditions will depend on the dynamics of the system. But the method vindicates the informal reasoning, for cases that do satisfy these conditions.

The branchpoint proposal uses probabilities, even if the laws are deterministic. I take the probabilities concerned to be objective worldly probabilities, often known as 'chances'. Chances are objective probabilities that apply in the single case and that are 'worldly': though they guide belief, they are not mere recommendations for belief. They are features of reality, derivable from the posits of physical theories. Roughly, chances are as 'worldly' as the fundamental dynamical laws.

These probabilities are defined as follows. Take the dynamical laws of fundamental physics, an appropriate probability distribution (in the case of deterministic laws and some indeterministic laws), and any other information about contingent states (B). These define conditional *objective probabilities*: probabilities for states (A) conditional on B: $P(A|B)$. In the case of deterministic laws, there are no restrictions on what is included in B: probabilities are well-defined towards both past and future. I'll call such probabilities 'deterministic chances'. In the case of classical Boltzmannian statistical mechanics, these probabilities derive from Newtonian dynamical laws and a flat probability distribution given by the Lebesgue measure: the 'Lebesgue postulate'. The Lebesgue postulate specifies a probability measure over microstates at a single time, although it has implications for states at other times via the dynamical laws.[6] The Lebesgue postulate does similar work to Albert's (2000) 'statistical postulate', which is applied to the initial macrostate. An advantage to the Lebesgue postulate is that it can be applied at *any* time and will deliver the same conditional chances. It can be applied to the entire 'phase space' of the universe (phase space is a continuous space representing possible states of a system) or to the states one conditionalises on at any given time. In either case, the Lebesgue postulate plus the dynamical laws deliver the same conditional probabilities.

Unlike standard accounts of chance (Lewis 1986a; Edgington 1995; Leitgeb 2012) or Albert (2000) and Loewer's (2007) 'statistical-mechanical chances' there is no relativization of chances to times or other states that one must conditionalise on. Because there are no further restrictions on what is conditionalised on, chances can play a direct credence-guiding role: one is guided by known chances that are conditional on known states. Even though the chances *relevant* to a reasoner will typically depend on what they know, the *values* of the conditional chances do not depend on what they know. So, despite their direct-credence guiding role, chances remain objective.[7]

For indeterministic laws, the situation is more complex. In general, chances are given by whatever posits are needed to derive macroscopic probabilistic behaviour. The indeterministic laws we're most familiar with, including those of GRW, specify 'transition probabilities'—probabilities for later microstates given earlier microstates. These deliver 'future-directed chances'—chances for later states given earlier states. In the case of GRW, Albert (2007, Ch. 7) argues that *any* probability distribution will lead to (roughly) the same probabilities for later macrostates after a short time. If so, the probability measure may not be part of the content of science and future-directed chances may derive from dynamical laws alone. However, if that condition is not satisfied, then a particular probability distribution is required. In this case, the Lebesgue postulate is used, applied to the earliest state conditionalised on.

---

[6] See Fernandes (2022a) for defence of this postulate. There is debate over whether one should employ a particular probability distribution (Albert 2000, Ch. 3; 2015, Ch. 1; Loewer 2007) or simply require that it satisfy certain minimal conditions, such as being continuous with the Lebesgue measure (Maudlin 2007b)—perhaps features so minimal that the probability distribution is not part of the content of science. My stance is that the probability distribution is a matter for science. However, it will not matter if one takes the probabilities to be less objective than the laws, provided one accepts counterfactuals that are less objective than the laws.

[7] Hall (2004) also denies additional restrictions.

I'll now consider what states are included in 'Prior'. The most general case, and the one I will commit to for deterministic laws, is for Prior to include the full *macrostate* of the universe *up to and at t*. By including the full macrostate up to and at *t*, macrostates at or prior to the branchpoint cannot counterfactually depend on antecedents. Why specify Prior in this way? I'll consider its justification further in Section 2. Roughly, the branchpoint proposal takes counterfactuals to model certain hypothetical cases of evidential reasoning. The probabilities we reason with typically concern macrostates, and, in the case of deterministic laws, a restriction to something larger than microstates is necessary to apply probabilistic reasoning. The above way of specifying Prior takes macroscopic information at and prior to the time of the branchpoint to be 'accessible' in a way that information about later states or microstates is not. These idealisations are approximations. Regarding the limitation to macrostates, we may sometimes have relevant knowledge of microstates. In Section 4, I'll consider why Prior is limited to earlier states.

For some indeterministic laws, another specification of Prior is available. If future-directed chances derive from dynamical laws alone, Prior can include the full *microhistory* of the universe up to and including the branchpoint. This specification provides a uniform treatment of macro and microstates—no states at or prior to the branchpoint can counterfactually depend on antecedents. For this reason, I will commit to this specification when it is available—that is, whenever a probability distribution is not required to define chances.

Both specifications of Prior will often deliver similar results for forwards-looking counterfactuals, even for those concerning microstates. For example, take an isolated box of gas at equilibrium containing some dust particles. Consider the counterfactual 'the dust particle would have been deflected to the left (L), if it had been hit on the right by an air molecule (R)'. If Prior is the *microhistory* of the universe, the branchpoint is a time just prior to R—the latest time at which R is maximally unsettled, given the microhistory. The counterfactual can be treated as roughly true just in case the probability for L given R and the microhistory is sufficiently high. If Prior is the *macrohistory* of the universe, the branchpoint will likely be *the time of R*—this is the latest time at which R is maximally unsettled, give then *macrohistory*, since the locations and velocities of air molecules are not part of the macrostate. The counterfactual can be treated as roughly true just in case the probability for L given R and the macrohistory is sufficiently high. In neither case is L certain given R. In the microhistory case, there are trajectories the molecules and particles *could* take that are compatible with R and the microhistory but not L—such as the particle being hit simultaneously on the left side. In the macrohistory case, there are *microarrangements* of molecules and particles at the time of the branchpoint whose trajectories are compatible with R and the macrohistory but not with L—such as arrangements which lead to the particle being hit simultaneously on the left side.

There are some differences that result from the specification of Prior. If Prior is the *macrohistory* and the antecedent is *less than macroscopic*, the time of the branchpoint *may* be the time of the antecedent. If this is so, and the consequent is *macroscopic*, there is no transition period and no backwards counterfactual dependence. For example, say that Hamish's *decision* to drink the wine is maximally unsettled by the macrohistory up to and including the time of his decision (*t*). If so, no macrostates at or prior to *t* will counterfactually depend on Hamish's decision.[8] Under a microscopic specification of Prior, there may be some backwards counterfactual dependence in such cases, but it is likely to be minimal—the branchpoint will be just before Hamish's decision and there won't be macrostates between then and the decision that counterfactually depend on the decision.

On the other hand, if the consequent concerns *microstates*, there is sometimes more scope for backwards counterfactual dependence under a macroscopic specification of Prior—because there can be backwards counterfactual dependence *before* the branchpoint. For example, say in the box of gas example that the particle is *not* hit on the right. If it had been hit on the right (R), the world would *not* have started out in the microstate than it did—this result is probabilistically ruled out. If it hadn't been hit on the right, the world *may* have started out in the microstate that it did—this result is not probabilistically ruled out. Under a suitably low threshold for the approximation conditions, the initial microstate counterfactually depends on R. Such dependencies are ruled out by a microscopic specification of Prior.

Altogether, differences in results between the specifications of Prior show up mostly clearly when microscopic consequents and antecedents are at issue, but, for macroscopic antecedents and consequents and for forwards-looking counterfactuals, the specifications tend to produce the same results. This is what we should expect. Something about our reasoning, especially concerning past

[8] Loewer (2007, p. 317) claims that an agent's decisions will generally satisfy this condition with respect to the macrostate of the universe at *t*. For dissention, see Fernandes (2022b, 2023).

states and microstates, is sensitive to the form of fundamental dynamics. But much of our reasoning is not. One should also keep in mind that the specifications of Prior are only idealisations.

Other specifications of Prior are available. A method closer to Albert (2000, Ch. 6) and Loewer's (2007, 2012) is to take Prior to include only the full state of the universe *at t* (for indeterministic laws like GRW) or the full *macrostate* of the universe *at t* (for deterministic laws), and perhaps further relevant known past states, such as the 'Past Hypothesis'. [9] This alternative might seem preferable: it would allow (macroscopic) consequents prior to the branchpoint to have non-trivial probabilities. It is also closer to a temporally neutral account—remaining asymmetries are only that one conditionalises on the Past Hypothesis and that *t* is stipulated to be prior to the antecedent. However, (macroscopic) antecedents won't generally be counterfactually relevant to states prior to the branchpoint, since the state at *t* largely screens off any such probabilistic dependencies.[10] So it will do no harm to help ourselves to an idealisation that holds the (macroscopic) past of the branchpoint 'fixed'.

One may worry, from the other direction, that including the full macrohistory in Prior will restrict future microtrajectories too much in the deterministic case, such that probabilities for futures states conditional on Prior will approach extreme values of 1 or 0. But this won't be so if the macrodynamics of our worlds is 'Markovian', meaning that the present macrostate screens off the probabilistic relevance of past macrostates. In the causal modelling tradition, the closely-related 'casual Markov condition' is a standard assumption for many systems (Hitchcock and Rédei 2020).

A more temporally neutral alternative, suggested by David Albert, is to take the branchpoint to be the *closest* time at which the antecedent and its negation are both reasonably probable, given the macrostate at that time. This alternative allows the branchpoint to be before or after the antecedent. However, for reasons mentioned above, I don't take temporal neutrality absent justification to be any kind of advantage. As I will argue (Section 3), the correct way to justify a proposal that includes a branchpoint requires the branchpoint to be prior. It is worth keeping in mind that holding the

[9] A remaining difference is that Albert and Loewer's accounts specify the state of the universe at the time of the antecedent rather than the time of the branchpoint and introduce further restrictions (see Section 3). Loewer has since defended an account (currently unpublished) that is much closer to the branchpoint proposal.
[10] See Loewer (20007) for details. There may only be probabilistic dependence when a macroscopic event in the past is correlated with a *microscopic* event at *t*—the structure underlying the so-called 'Atlantis' case, attributed to Adam Elga.

present fixed *or* holding the past and present fixed are both idealisations, given we never know the full state at *t*. More generally, the vagueness of what is included in 'Prior' reflects the fact that counterfactual reasoning contains an idealisation—it is designed to capture something that holds by and large of our evidential reasoning, but nothing more precise than that.

In other work (2020, 2021), I use cases of time travel to motivate more subjective alternatives, which take Prior to include only information *known* or *directly available*. These subjective alternatives do better in contexts involving backwards causation, time travel and causal loops. So, they may provide a more general account of counterfactuals, of which the branchpoint proposal is a special case. The justification for the branchpoint proposal is tied to an evidential asymmetry that holds in the actual world—that the forwards evolution of the universe is reasonably probable, while its backwards evolution is not (see below and Section 3). In cases where that asymmetry fails, the specifications of Prior and the method may need to change. But the specifications of Prior above remain defensible as idealisations that holds of the actual world that deliver counterfactuals that are sufficiently objective to guide our evidential reasoning. We may disagree about what counterfactuals are true, given that we have different evidence. But what counterfactuals are true does not depend on what evidence we have.

The branchpoint proposal has several advantages. First, it can be employed regardless of whether the fundamental dynamical laws are deterministic or indeterministic. I take this to be an advantage, since most of our higher-level counterfactual reasoning should be insensitive to this feature of fundamental physical laws. In the case of deterministic chances, probabilities are well-defined independently of when the states conditionalised on are located, so the probabilities required by the branchpoint proposal can be calculated directly. In the case of future-directed chances, the probabilities can be derived. The probabilities we need are P(C|A.Prior). If C is after both the branchpoint and A, future-directed chances are well-defined. If C is before the branchpoint, in many cases, C or its failure will be included in Prior—if so, the relevant probability is 1 or 0.[11] If C is after the branchpoint but before A, the required probability can be calculated indirectly, using the definition of conditional probability:

[11] Cases in which C or its failure are not included in Prior are those in which C is more fine-grained than Prior. If so, the relevant probabilities are calculated conditional on a complete past micro or macrostate (PS), prior to C.

1) $P(C|Prior.A) = P(A.C|Prior) / P(A|Prior)$.

Second, unlike standard Lewisian accounts, the branchpoint proposal avoids surprising dynamics at both the macro and micro level and ensures that the antecedent had a reasonable probability of coming about. Lewis (1979), by contrast, allows violations of the fundamental dynamical laws called 'miracles'—spatiotemporal areas in which our laws don't hold—in order to rule out backwards counterfactual dependence prior to the miracle. For concerns, see Dorr (2016). Statistical mechanical accounts by Albert (2000, 2015) and Loewer (2007, 2012) allow for improbable macroscopic behaviour, in order to minimise macroscopic differences between the actual world and the counterfactual world at the time of the antecedent. The branchpoint proposal keeps regular micro and macroscopic behaviour unchanged in counterfactual worlds, so that counterfactuals can facilitate accurate evidential reasoning. While it is a positive feature of statistical mechanical accounts that they avoid violations of the fundamental laws, if counterfactuals are to be justified based on their evidential role, violations to macroscopic behaviour should also be avoided. The use of miracles or surprising macroscopic dynamics in these other accounts is primarily motivated by the need to avoid backtracking counterfactuals—a motivation I'll suggest we give up (Section 4).

A third advantage is that the branchpoint proposal allows the time of the 'fork' from the actual world to be determined by the system's dynamics, rather than being (at best) determined by context, as with alternative accounts (Edgington 1995, 2004; Maudlin 2007a, Ch. 1; Leitgeb 2012; Paul and Hall 2013, pp. 47–53; Dorst 2022). On the branchpoint proposal, while contextual features may be used when determining what state of affairs is referred to by a particular (linguistic) antecedent, context is not used to determine the time of the fork or what is held 'fixed' in the counterfactual world.[12] Altered-states recipe approaches (Maudlin 2007a, Ch.1; Paul and Hall 2013, pp. 47–53; Dorst 2022), Edgington's (1995) and Leitgeb's (2012) accounts are similar to the branchpoint proposal, in that they involve rewinding the universe and using the laws (or chances) to determine whether the consequent occurs. But they are more limited in what they achieve. When evaluating

[12] Since context partly determines what state of affairs is picked out by a linguistic antecedent, there can be variation in whether a counterfactual is true, dependent on context. Context may even sometimes require that the putative antecedent is brought about in a particular way. So, there is scope for the branchpoint proposal to use contextual features to respond to counterexamples. (My thanks to a referee for this observation.) The important point is that context plays no role in determining what probabilistic counterfactuals are true, once the antecedent and consequent have been fully specified.

counterfactuals concerning future states, these accounts can sometimes reasonably take the time of the fork to be the time of evaluation of the counterfactual (Leitgeb 2012, p. 67). But this contextualist alternative is not available when evaluating counterfactuals concerning past states, where a previous fork cannot be the time of evaluation. More generally, it is difficult to see how present contextual features alone can determine the time of the fork when evaluating counterfactuals concerning past states. It is perhaps for this reason that altered-states recipe approaches don't allow for the evaluation of counterfactuals where the consequent is prior to the antecedent. While some allow for changes to the time of the antecedent outside the area of the antecedent, these are all specified by 'context' (Maudlin 2007a, p. 24).[13]

A fourth advantage to the branchpoint proposal is that it employs probabilities and a precise means of specifying them, derived from scientific posits. The use of probabilities means the account has some answer to the complaint that most counterfactuals are false (Edgington 1995, 2004; Hájek 2021). On Lewis' account, if the antecedent is imprecisely specified, many ordinary counterfactuals will come out false, since the consequent won't be true at all the nearest antecedent-satisfying worlds. By using counterfactuals with probabilistic consequents, one can employ a semantics that allow consequents with high probability to come out as acceptable or approximately true (Edgington 1995; Leitgeb 2012). By using probabilities, the proposal also provides a unified treatment of so-called 'might' and 'would' counterfactuals—'would' counterfactuals being those with high probability consequents and 'might' counterfactuals being those with lesser probability consequents. Finally, the use of scientifically-derived probabilities provides a closer link between metaphysics and science and better prospects for explaining temporal asymmetries in the method rather than treating them as primitive (Section 3). Many alternative accounts either fail to justify asymmetries in how counterfactuals are evaluated (Paul and Hall 2013) or fail to adequately explain the source of these asymmetries (Lewis 1979)—see Elga (2001).

2. The Evidential Role of Counterfactuals

Having laid out the branchpoint proposal and some of its advantages, I'll now justify it by considering the evidential role of counterfactuals. I'll justify the method's temporal asymmetry in Section 3.

[13] Maudlin (2007a, p. 32) suggests we may evaluate counterfactuals with earlier consequents in a non-trivial way but doesn't square this with his earlier claim.

As the informal reasoning in Section 1 suggested, one important role of counterfactuals is to help us reason evidentially. Counterfactuals point us towards probabilistic structure that helps us reason about what is the case, given evidence of other states. More precisely, counterfactuals aim to capture how one *should* reason about a counterfactual consequent (C), given an antecedent (A), in situations that are relevantly similar to a hypothetical scenario where A is maximally unsettled by information one could reasonably be expected to have access to. Counterfactuals capture whether A probabilistically settles C and whether A changes the probability of C in these hypothetical scenarios, which indicate general evidential relations that holds between A and C that are useful for reasoning in other settings.

In Clare's case, counterfactuals indicate whether A (failure to arrive) settles C (mishap) and raises its probability in a scenario where A is unsettled. Knowing these general evidential relations, we can reason *from* A to C in similar scenarios where only A is known. For example, we can reason that Georg's failure to arrive (A) settles an earlier mishap (C), in cases where we only know that he has failed to arrive. We can also use counterfactuals to reason from C to A. For example, we can reason that Darcy's earlier mishap implies that her failure to arrive is at least reasonably probable and that her mishap raises the probability of her failing to arrive. This result makes use of the informal reasoning above (see Appendix 2). In cases where *neither* A nor C is known, counterfactuals deliver hypothetical information—what one would learn on learning A or C—which can direct further information gathering. For example, if we're worried that Evelyn may have had a mishap, we can check whether she's arrived—knowing that failure to arrive (A) will indicate an earlier mishap (C)—both in the sense of settling an earlier mishap and raising its probability.

If counterfactuals capture how one should reason about C, given A, in the hypothetical scenario, there is reason for evaluating them using the branchpoint proposal. The branchpoint proposal requires 'rewinding' the universe to a time when A and not-A are both reasonably probable, given Prior. This is the condition of the hypothetical scenario, where A is maximally unsettled, given states one could reasonably expect to have access to. While we never have access to Prior in its entirety, conditionalising on Prior ensures that evidential relations are robust as new information accessible up to that time is conditionalised on—similar to the use of 'admissibility' (Lewis 1986a) and 'resiliency' (Skyrms 1980) in accounts of chance. In Clare's case, for example, the evidential relations between her failure and her earlier mishap do not change, given further information about her upbringing—as the relevance of her upbringing is included in Prior. Given the specification of the branchpoint, we evaluate counterfactuals by considering the chance of C conditional on A and Prior. Because conditional chances are reliable evidential guides when known (Section 1), these chances indicate whether A settles C and whether A raises the probability of C in the hypothetical scenario and in relevantly similar scenarios.

What are relevantly similar scenarios? Broadly speaking, they are those that share enough probabilistic structure with the hypothetical scenario, such that the counterfactual relevance of A for C is not undermined. For example, A can be relevant for C, under various conditions of A and C being known or having occurred, provided the scenarios are structurally like the hypothetical scenario—the cases of Georg, Darcy and Evelyn above. But A may no longer be relevant for C in the case of Benedict, who is habitually careless about his appointments. A may also no longer be a good indicator of C, by the time A arrives, if surprising events happen between the branchpoint and A that change the counterfactual relevance of A for C. Our use of counterfactuals relies on a) different systems (or temporal repeats of the same system) sharing enough probabilistic structure that we can reason from one to another and b) chances often enough being robust as states between the branchpoint and the antecedent are conditionalised on. Regarding b), we expect many chances to be robust in this sense—the chance of the coin landing heads is insensitive to most interventions after the coin is tossed. Regarding a), we often enough expect chances, conditional on reasonably complete states of the universe at a time (chances for singular *token* events) to indicate chances for other singular token events. This expectation underlies our use of type-level chance reasoning.[14]

This evidential justification of counterfactuals contrasts with those given by Edgington (2004) and Dorst (2022), endorsed by Loewer (2024, Ch. 3). Edgington and Dorst argue that counterfactuals are required when we acquire unanticipated evidence concerning C and need to reason about A. Say I observe tyre tracks (C) and am trying to ascertain a car's earlier speed (A) (Dorst 2022). According to Edgington and Dorst, we need to reason counterfactually—we need to consider some earlier time at which the car's speed could have been different and consider what differences in the car's speeds

[14] The physical conditions required for a) and b) are similar to those that allow us to divide the world into subsystems that behave similarly, even though the behaviour of any subsystems is sensitive to the global state of the universe at that time. For a general outline of what physical conditions are required to underwrite type-based reasoning about local subsystems, see Elga (2007).

would imply about differences in the length of the tyre tracks. The reasoning involves inference to the best explanation, where one consider the probability of a later C (tyre tracks) given an earlier A (car's speed): Cr(C|A.K) or P(C|A.K). Dorst (2002, p. 553) and Edgington (1995 p. 265; 2004, p. 25) seem to think as follows. We can't reason using epistemic probabilities because we already *know* the length of the tyre tracks. Conditional on what we know, we can't reason about what the lengths of the tyre tracks would have been given differences in the car's speed. We can't reason using chances because the chance now of the tyre tracks being as they are is 1. Chances concerning past events are 'trivial'—they take the value of 0 or 1 depending on whether the event has occurred. So, there aren't well-defined non-trivial chances about how the tracks would have differed given differences in the car's speed. If we had anticipated acquiring the later evidence, we could have calculated the required epistemic probabilities or chances. But we didn't. So, we don't now have probabilities suitable for ascertaining the relevance of C for A and hence we need to make counterfactual judgements about the relevance of A for C.

But this argument fails. Regarding epistemic probabilities, because A is unknown, we can reason directly about A, given known information including C, using Cr(A|C.K)—where Cr are credences and K is everything known. Even if C comes after A, epistemic probabilities are well-defined and non-trivial. If we don't know these epistemic probabilities, they can be derived using Bayes' theorem from epistemic probabilities concerning the length of the tyre tracks conditional on the car's earlier speed. Using Bayes' theorem, one must identify prior probabilities for the car's speed and so *some* earlier time at which these priors are calculated—but nothing like a particular branchpoint. One can also often reason using credences simply by removing the unanticipated evidence (C) when conditionalising on known information (Bennett 2003, p. 336): Cr(C|A.(K−C)). If the fact that I have observed the tyre tracks is what prevents me reasoning about how the tyre tracks may have differed, I can remove the tyre tracks from my stock of beliefs, as well as other information that needs to be explained.

Regarding chances, in the case where the laws are deterministic, chances are well-defined and non-trivial, independently of when the states conditionalised on occur. We can reason about the relevance of C for A using P(A|C.K). If the underlying laws are indeterministic, then not all chances are well-defined. But, provided one is willing to conditionalise on sufficiently complete past states (PS), chances for A given C will also be well-defined and non-trivial, even if C comes after A. We can reason using P(A|C.PS). For example, on Loewer's statistical mechanical account of indeterministic chances (2024, Ch. 8), chances are always conditional on the initial macrostate of the universe, the 'Past Hypothesis' (PH). If so, the chance of the earlier car speed, given the later tyre tracks, P(A|C.PH), is well-defined and non-trivial. Later or more local states will also suffice. Even if chances for past events are always 1 or 0, chances that obtained at some earlier time can still guide reasoners. One can reason that the chance of the tyre tracks being thus and so *was* much higher than otherwise given the car's earlier speed. Reasoning in this way doesn't require a specification of the branchpoint. It just requires one to consider the chances at *a* time when the relevant chances were well-defined and non-trivial. In the case of both epistemic probabilities and chances, because we don't need to specify a branchpoint, we don't get a full specification of what counterfactuals are true.

None of this is to say that counterfactuals aren't useful in cases where C is known and A is unknown. But they're not essential. We can always reason about A using conditional probabilities, without specifying the branchpoint. Counterintuitively, it is actually when both A and C are known that we need to specify the branchpoint (and Prior) and thus employ a method of evaluating counterfactuals.[15] These counterfactuals direct us to evidential relations that are then useful (but not required) for reasoning in other settings. Recall, in cases where both A and C are known, the aim is not to figure out A or C but to ascertain more general evidential relations: whether A is counterfactually relevant for or settles C. Figuring out these general relations requires figuring out how A is evidentially related to C in a similar hypothetical scenario where A is maximally unsettled—which requires us to make assumptions about what information we'd have access to in that hypothetical scenario. Because counterfactuals pick up on only certain kinds of general probabilistic relations, idealisations (the branchpoint and Prior) feature in counterfactual reasoning that are absent from probabilistic reasoning.

With the above evidential justification in mind, we can address a remaining detail: whether future states should ever be included in Prior. Some methods of evaluating counterfactuals adopt the feature 'hindsight': chancy events in the future of an antecedent are held fixed when evaluating

---

[15] Recall, I dropped the standard assumption that 'If A were to be the case C would be the case' is true if A and C are both true of the actual world (see Footnote 4) and did not assume counterfactuals have false antecedents (Footnote 1). My concern is with relations that can be used even if the truth or falsity of the antecedent is unknown.

counterfactuals, provided they're not causally downstream of the antecedent (Edgington 2004; Bennett 2003, Ch. 15; Maudlin 2007a, Ch. 1; Dorst 2022). These accounts motivate hindsight as follows. Say you don't bet on Heads (¬B), a chancy coin is tossed (T) and it lands Heads (H). Your bet isn't causally relevant to the outcome of the coin. It seems the following counterfactual is true: 'If you had bet on heads, you would have won'. So, it seems chancy events such as the coin landing heads should be held 'fixed' in counterfactual scenarios.

One could add hindsight to the branchpoint proposal. Since I don't use counterfactuals to analyse causation, there is no threat of circularity. But I prefer not to hold events in the future of the branchpoint fixed, because events in the future of the branchpoint aren't included in the hypothetical scenario. Consider the coin toss case, where the antecedent is your bet on heads. The branchpoint is likely the time of your deliberation, t_D,—your bet is unsettled at that point. At t_D there aren't relevant states that settle the coin landing heads. If there were, it would not be a chancy coin toss. So, insofar as counterfactuals are guide your evidential reasoning at t_D, H should not be held fixed. If one is interested in choosing B or not, or interested in figuring out what one can reason to concerning the outcome of similar coin tosses, one can do no better than assign equal probability to H and ¬H. We may have intuitions to the contrary that lead us to regret our failed bets. I suspect this is because, in non-chancy cases, such outcomes are settled by prior states and therefore can in principle be known in advance. We adopt this expectation in the chancy case.

There will still be cases where we hold chancy future events like H fixed when reasoning. To take a case from Dorst (2022), we might hold the chancy event of a car's anti-lock braking system failing to engage fixed when working out what later tyre tracks should lead us to infer about the car's earlier speed. But, *contra* Dorst, such cases don't justify hindsight. As discussed above, epistemic probabilities and chances allow us to reason in such cases: we simply use our knowledge about what happened at any time to constrain our reasoning about the unknown. We consider the probability of the car's earlier speed being thus, conditional on the car's anti-lock braking system failing to engage, the tyre tracks, the weather, etc. We do not need to hold other parts of the unknown 'fixed'. So, we don't need to employ counterfactuals and such cases don't motivate hindsight. Even in cases where we consider a time prior to the antecedent, either because the laws are indeterministic or the epistemic probabilities are inaccessible, there is no need to hold future states 'fixed'. One simply considers the probabilities for all relevant possible future states conditional on the antecedent and states up to a chosen time, and then consider which antecedent 'best explains' the future states observed, using Bayes' theorem. If our evidence settles that the car's anti-lock braking system fails to engage, it will be part of the set of future states that the antecedent needs to explain—and so is relevant for reasoning about other states.

3. Temporal Asymmetry

I've argued that key features of the branchpoint proposal can be justified by considering an evidential role for counterfactuals. One remaining feature to consider is its temporal asymmetry. The branchpoint proposal is temporally asymmetric: the branchpoint is prior to the antecedent and therefore so are the states conditionalised on (Prior). It might seem that the method illicitly builds in a temporal asymmetry 'by hand' and that it would be preferable to employ a more temporally neutral method. One possibility is to take the branchpoint to be the *nearest* time to the time of the antecedent at which the antecedent is maximally unsettled.[16] A temporally neutral method might be used to derive temporal asymmetries of counterfactuals and causation.

There are three responses for why the branchpoint proposal doesn't illicitly build in a temporal asymmetry 'by hand'. First, because I don't use counterfactuals to reduce causal relations, the asymmetry does not prevent one providing a substantial account of why causation is temporally asymmetric. For such an account, see Fernandes (2017).

Second, as I argued above (Section 2), a method of evaluating counterfactuals requires justification. If we consider the evidential role of counterfactuals, there is reason to evaluate them temporally asymmetrically. The asymmetry is not put in 'by hand'. I know of no such justification for the nearby temporally neutral alternative.

The reason the branchpoint proposal is temporally asymmetric is that it reflects a general probabilistic asymmetry that holds in our universe: the universe's forward evolution is reasonably probable, in the sense that, conditional on micro or macrostates up to any point, subsequent macrostates are reasonably probable. This feature does not hold in reverse: the backwards evolution of the universe is typically highly improbable. Conditional on future states, individual past

[16] Barry Loewer defends such a method (unpublished).

macrostates that obtained are typically highly improbable. The probabilistic asymmetry is pervasive and underwrites asymmetries in how we reason about the past and future.[17] So it should be maintained in counterfactual scenarios.

The temporal asymmetry of the branchpoint proposal ensures that the universe's future evolution is reasonably probable in counterfactual scenarios. Placing the branchpoint prior to the antecedent ensures that, conditional on relevant micro or macrostates up to the branchpoint, the antecedent is reasonably probable. Because Prior contains at least the full macrohistory, the evolution of the universe prior to the branchpoint is reasonably probable.[18] The use of the Lebesgue postulate and the dynamical laws in deriving states after the branchpoint makes it highly probable that the universe's subsequent evolution will be reasonably probable. Putting together these three pieces, the evolution of the universe forwards in time in counterfactual scenarios is (with high probability) reasonably probable. But, because the probabilistic asymmetry does not hold in reverse, it makes no sense to place a branchpoint after the antecedent and require the universe's backwards evolution to be reasonably probable. So, there is no justification to employ a method in which *past* states are reasonably probable.

There is a second way that the universe's probabilistic asymmetry helps underwrite the usefulness of counterfactuals. The hypothetical scenario that counterfactuals model is one in which the antecedent, A, is maximally unsettled by states that are reasonably epistemically accessible at the time of the branchpoint. Counterfactuals indicate what one would learn on learning A. Such knowledge is particularly useful if there is the possibility of 'later' learning A, where 'later' indicates the direction of time in which we acquire and retain information. This was the case of Evelyn above, where neither A nor C is currently known, but A may become known. Counterfactuals ascertainable at early times can then indicate evidential relations useable at later times. Counterfactuals are only useful in this fashion if the branchpoint is prior to the antecedent. There are *future* states we could later come to know of that aren't settled by states accessible now and that have further implications. But there aren't generally *past* states we could later come to know of that aren't settled by states accessible now. While there are past states we don't know, we expect them to be either recoverable

---

[17] For discussion, see Reichenbach (1956),Albert (2000) and Fernandes (2022a). From the probabilistic asymmetry one can derive that the entropy of the universe increases towards the future and not the past.

[18] More precisely, it ensures this to the degree that our universe's evolution is probable. We should expect minor deviations where some subsequent states are improbable, as determined by probability theory.

from present states or not recoverable at all. We don't expect there to be past states we'll 'later' learn of that have any implications, beyond what is derivable from the present. Once again, it is the probabilistic asymmetry that explains these asymmetries, including the fact that we retain information in one temporal direction and not the other (Reichenbach 1956; Albert 2000, Ch. 6; Fernandes 2022a). For this reason, branchpoints prior to antecedents are useful to us in ways that branchpoints after antecedents would not be.

Here is another reason to adopt the temporally asymmetric branchpoint proposal over its temporally neutral alternative. There would be some advantage to a temporally neutral proposal if it could rule out backwards counterfactual dependence and so backwards causation. But, even in cases where the branchpoint is prior to the antecedent, the method allows for significant backwards counterfactual dependence. Whenever the consequent is *before* the antecedent but *after* the branchpoint (see Figure 1), the consequent may counterfactually depend on the antecedent, implying backwards counterfactual dependence. This is also a third reason that the branchpoint proposal does not put in an asymmetry by hand—it does not rule out backwards counterfactual dependence.

In fact, the branchpoint proposal allows more scope for backwards counterfactual dependence than other accounts. Typically accounts of counterfactuals aim to minimise the 'transition period'—the time between when the state of the counterfactual world diverges from that of the actual world and the antecedent—to minimise the scope for backwards counterfactual dependence and backwards causation. Remaining problem cases are dealt with by arguing that counterfactual dependence during the transition period is insufficiently 'definite and detailed' (Lewis 1979, p. 463) or otherwise robust (Albert 2015, Ch. 2; Loewer 2012). For example, Lewis (1979) minimizes the transition period by maximizing the spatiotemporal area of perfect match between the actual world and the counterfactual world. Statistical-mechanical accounts minimise the transition period by requiring the macrostate of the counterfactual world to *match* that of the actual world at the time of the antecedent, outside the area of the antecedent (Albert 2000; Loewer 2007, 2012) or by minimizing such changes.[19] These stipulations minimise the transition period by allowing either miracles (areas of counterfactual worlds that violate our fundamental laws) (Lewis 1979) or highly improbable forwards evolutions (Albert 2000; Loewer 2007, 2012). These stipulations rule out macroscopic

[19] See Kutach (2002) and Fernandes (2023) for the latter interpretation of Albert (2000).

correlations we would otherwise expect between the antecedent and other events at the time of the antecedent.

The branchpoint proposal avoids such stipulations and so doesn't limit the scope for backwards counterfactual dependence. Rather than having the branchpoint set by context or by such stipulations, the method allows for longer transition periods, by requiring the forwards evolution of the universe to remain reasonably probable—in keeping with regular macroscopic behaviour. For an antecedent such as Clare failing to arrive at the café to come about in a way that is reasonably probable, one may have to countenance backtracking to hours beforehand, to a time when she had a reasonable probability of slipping in the bathtub. One does not limit such backtracking by allowing the antecedent to come about in a highly improbable way, such as by Clare inexplicably swerving past the café at the last moment.

Finally, it is worth keeping in mind that competing methods that look temporally neutral can sometimes smuggle in temporal asymmetries. Loewer's (2007) statistical-mechanical account evaluates counterfactuals using probabilities derived from the statistical postulate, the dynamical laws, the Past Hypothesis and the full macrostate of the universe. It might seem that Loewer's account does better at explaining temporal asymmetries, since Prior only includes the present state. But appearances are misleading. Loewer restricts the antecedents to decisions and stipulates that these are less than macroscopic and have a 'reasonable' probability of coming or failing to come about conditional on the macrostate *at or prior to* the time of the antecedent. The assumption that antecedents are probabilistically independent of past states builds in an unexplained asymmetry. Moreover, requiring antecedents to be less than macroscopic and probabilistically independent of states in the present restricts the scope of Loewer's account.[20] The branchpoint proposal is more general and its temporal asymmetries are explained rather than presumed.

4. Control

Counterfactuals, evaluated as the branchpoint proposal suggests, are useful for evidential reasoning. But they are also relevant in contexts of control. As we deliberate, our decisions and subsequent actions are often unsettled by present and previous states. Our deliberations often take place at

[20] For further concerns, see Fernandes (2022b).

branchpoints with respect to our decisions and actions. So, it is useful for us to know about the evidential relations that hold in the hypothetical scenario—they often indicate what our decisions and actions will probabilistically settle about the world, which can guide decision-making. Hamish's case is of this type. Hamish can use knowledge of what would happen, were he to drink the wine, in choosing whether to drink the wine.

Why do our deliberations typically take place at branchpoints? This arises partly from the complexity of decision-making: we respond to a variety of reasons in nuanced ways, sometimes by revising the decision-process itself (Price 2012; Ismael 2012). Norms on deliberation also suggest that deliberation takes place at branchpoints. Deliberation may not make sense if we are already certain of its result. While sometimes there may be evidence which we don't have or ignore, typically our decisions and actions won't be settled by states accessible at the time of deliberation.

It is not, however, *necessary* that our deliberations take place at branchpoints. We might sometimes be reliable responders to the world, such that our decisions are probabilistically settled by previous states (Fernandes 2023). In such cases, the relevant branchpoints are earlier than the states we reliably respond to and there is backwards counterfactual dependence of previous states on our decision. For example, you might be a reliable fly-swatter such that your swatting and your decision to swat are settled by the presence of a fly in the previous moment (and you don't decide to swat otherwise). Say there is a fly before you in the actual world and you therefore decide to swat. The branchpoint at which you decision is probabilistically unsettled by Prior is a time at which the fly may or may not have flown before you—not the time of your decision or deliberation. Given this Prior, your not deciding to swat probabilistically settles the fly not being before you in the previous moment. So, if you hadn't decided to swat, there wouldn't have been a fly there. The fly's location in the past counterfactually depends on your decision. Because of the possibility of such cases, the branchpoint proposal does not imply a temporal asymmetry of counterfactual dependence, even when antecedents are decisions. Nor do branchpoint counterfactuals reliably indicate what we control.

There are other approaches to explaining temporal asymmetries of control. My preference is to use norms on what counts as 'reasonable' deliberation and to argue that agents cannot use their decisions now to raise the probability of past states (conditional on their evidence) during reasonable deliberation (Fernandes 2017). Even if agents don't know the full state of the universe while deliberating, they do have access to states of their own psychology, including their desires and beliefs. They have access to states that mediate whatever correlations there are between previous states and their decisions. Following Price (1991), this evidence is enough to screen off correlations between decisions or actions and previous states. If we define control asprobability raising through decisions in reasonable deliberation, agents won't control the past. This kind of 'agent-based' account may also be extended to explain the temporal asymmetry of causation. Causal relations can be defined as evidential relations of a type that agents could use in reasonable deliberation to raise the probability of outcomes they seek. If so, there is no backwards causation. These explanations of the temporal asymmetry of control and causation rely on norms of deliberation which are absent from the branchpoint proposal. They also employ a more local kind of screening off than the branchpoint proposal—one that does not involve conditionalising on full states of the universe.

I've argued that the probabilities used to evaluate relations of control do not always overlap with those used in the hypothetical scenario. There is a more general lesson to be had: it is a mistake to expect a single analysis of counterfactuals to satisfy two competing roles: 1) indicate relations of control (control role) and 2) indicate general evidential relations of use to evidential reasoners (evidential role). Even if one uses counterfactuals (or probabilities) to evaluate both causal relations and evidential relations, one needs to employ different methods in each case.

To satisfy the control role (1), the method of evaluating counterfactuals must rule out backtracking counterfactuals, given that there is no backwards causation or control in the actual world. To satisfy the evidential role (2), the method of evaluating counterfactuals must keep the laws and macroscopic regularities of the actual world unchanged in counterfactual scenarios. These requirements come into conflict in the transition period. If backtracking is ruled out, the transition period must be zero—no events prior to the antecedent can depend on the antecedent. Yet, a zero-length transition period implies violations of the laws and macroscopic regularities whenever the antecedent does not already have a reasonable probability of occurring. A transition period that is evidentially reasonable must allow for backtracking. But such a transition period will not rule out backwards causation and so will not satisfy the control role.

Lewisian accounts run into trouble, precisely in attempting to satisfy these competing roles. Lewis (1979), Albert (2000; 2015) and Loewer (2007; 2012) attempt to minimise the transition period to minimise the scope for backwards counterfactual dependence. They therefore must accept either miracles (Lewis) or surprising macroscopic dynamics (Albert, Loewer). These counterfactuals are not reliable evidential guides. But, as the transition period cannot be eliminated entirely, without producing evidentially *very* strange counterfactuals, they must still accept backwards counterfactual dependence during the transition period (Bennett 2003, pp. 288−91; Dorr 2016, pp. 262−5). So Lewisian counterfactuals, without further restriction, don't satisfy the control role.

The solution, I suggest, is to sharply distinguish the relations that satisfy the evidential role and the control role. My preferred choice is to adopt an account of counterfactuals that satisfies the evidential role and a probabilistic account of causation that satisfies the control role. One might use counterfactuals (or probabilities) to satisfy both roles. The important point is that the methods used in their evaluation must be distinct.

5. Conclusion

The branchpoint proposal for evaluating counterfactuals has significant advantages: it avoids surprising micro or macroscopic dynamics, it allows the time of the branchpoint to be set by the system's dynamics rather than by context and it uses scientific posits to specify the relevant probabilities. While the proposal has a temporal asymmetry, this feature reflects a general temporal asymmetry in the universe's probabilistic structure, rather than being put in by hand. The branchpoint proposal is justified by considering an evidential role for counterfactuals. Counterfactuals indicate general evidential relations that hold in a hypothetical scenario at which the antecedent is maximally unsettled. Knowing these counterfactuals is useful for reasoning evidentially. While branchpoint counterfactuals are often relevant in contexts of control, they aren't always so. Crucially, we should not expect a single account of counterfactuals to deliver relations that satisfy both roles.

Appendix 1

Here are some results concerning the logic of counterfactuals with non-probabilistic consequents that can be treated as roughly true. In each case, I'll consider some (non-exhaustive) ways in which inference patterns are invalidated. First, as with similarity accounts (Lewis 1973), the branchpoint proposal invalidates inference patterns associated with antecedent monotonicity. For example:

<u>Antecedent Strengthening</u>

A □→ C ⊮ (A & B) □→ C

P(C|A.B.Prior) may be significantly less than P(C|A.Prior). It's probable I'm happy, given I dance at the cèilidh. But it's not probable I'm happy, given I dance at the cèilidh on a full stomach. The branchpoint may also be different for antecedents A and A.B.

<u>Transitivity</u>

A □→ B, B □→ C ⊮ A □→ C

P(B|A.Prior) and P(C|B.Prior) may both be high without P(C|A.Prior) being high. Say I can either stay home, go east to the post office or west to the grocers. It may be highly probable that I leave the house, given I pick up a letter. It may be highly probable that I go to the grocers, given I leave the house. Yet it will not be probable that I go to the grocers, given I pick up a letter.

<u>Contraposition</u>

A □→ B ⊮ ¬B □→ ¬A

P(¬A|¬B.Prior) may be significantly less than P(B|A.Prior). For example, P(B|A.Prior) may be high because B is contained in Prior. Yet Prior can change, depending on the antecedent. It's probable I went to Rome, given I get a suntan (the branchpoint being when I'm considering going for a walk in Rome). It's not probable that I would be pale, given I don't go to Rome. Given I don't go to Rome, it's highly probable I went to the Algarve and got a suntan there.

There are also inference patterns that similarity accounts maintain, which the branchpoint proposal invalidates.

<u>Modus Ponens</u>

A □→ B, A ⊮ B

P(B|A.Prior) being very high and A are logically compatible with not-B. It's probable I'm happy, given I dance at the cèilidh. I dance at the cèilidh. But I'm not happy—someone trod on my toes.

Agglomeration

A □→ B, A □→ C ⊮ A □→ B & C

P(B|A.Prior) and P(C|A.Prior) may both exceed the threshold of being 'very high', while P(B.C|A.Prior) does not. For example, if B and C are uncorrelated (given A.Prior):

P(B.C|A.Prior) = P(B|A.Prior) . P(C|A.Prior)

For concerns, see Hájek (2014, p. 248). There are no violations if B and C are perfectly correlated. The likelihood of violation increases to the degree B and C are anti-correlated. It's probable I dance, given I go to the cèilidh. It's probable I drink some wine, given I go the cèilidh. But it's significantly less probable that I dance and drink some wine, given I go the cèilidh—I have a tendency to avoid wine when dancing.

Some of these violations could be avoided, on a case-by-case basis, by restricting Prior, restricting the antecedents and consequents or adjusting the 'very high' threshold. But I suspect we should learn to live with them. Counterfactuals with *probabilistic* consequents are approximations, derived using idealisations. Counterfactuals with non-probabilistic consequents are additionally approximate, derived using a threshold—we shouldn't expect their logic to be neat. An analogy might be personal identity. If personal identity is a higher-level approximation, one shouldn't be surprised at violations of the logic of identity, such as violations of transitivity in cases of fission. One can keep transitivity, but only by devices such as accepting co-located persons. Whether we want to use devices in particular cases depends on what we use the approximate relations for—something I have not explored here, beyond the general role for counterfactuals.

Appendix 2

Here is how the branchpoint proposal vindicates the informal reasoning in Clare's case (Section 1). To evaluate counterfactuals about what would happen, were Clare to fail to arrive (F), identify a time (*t*) at which Clare's failing to arrive (F) or arriving (¬F) are both reasonably probable (given Prior):

1) $P(F|Prior) \approx P(\neg F|Prior) \approx 0.5$

Then consider what other states would be probable, given her failure to arrive (F). Say Clare's failing to arrive would imply, with high probability, that some mishap (M) had befallen her.

2) $P(M|F.Prior) \gg 0.5$

Since F is reasonably probable (given Prior) and since conditionalising on F implies that M is highly probable (given Prior), M must itself be at least approximately reasonably probable (given Prior).

3)
$$P(M|Prior) \geq P(F|Prior) . P(M|F.Prior)$$
$$\approx 0.5$$

Given the above conditions, the relevant mishap implies that her failure is at least reasonably probable (given Prior).

4)
$$P(F|M.Prior) = P(M|F.Prior) . P(F|Prior) / P(M|Prior)$$
$$P(F|M.Prior) \quad \text{Maximum} \quad = 1$$
$$\text{Minimum} \quad \approx 0.5$$

If F is positively counterfactually relevant to M,

5) $P(M|F.Prior) > P(M|\neg F.Prior)$

Using the law of total probability,

6) $P(M|F.Prior) > P(M|Prior)$

Using the definition of conditional probability,

6) $P(F|M.Prior).P(M|Prior) / P(F|Prior) > P(M|Prior)$

Implies

7) $P(F|M.Prior) > P(F|Prior)$

8) $P(F|M.Prior) > P(F|\neg M.Prior)$

We reason to past states that are at least approximately reasonably probable (equation 3) and that imply that the antecedent is at least approximately reasonably probable (equation 4). If the antecedent is positively relevant to the past state, the past state implies that the antecedent is more probable than it would otherwise be (equation 8).

References

Albert, David. *Time and Chance*. Cambridge, Mass.: Harvard University Press, 2000.

———. *After Physics*. Cambridge, Mass.: Harvard University Press, 2015.

Bennett, Jonathan. "Counterfactuals and Temporal Direction." *The Philosophical Review* 93, no. 1 (1984): 57–91. https://doi.org/10.2307/2184413.

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

Cartwright, Nancy. "Causal Laws and Effective Strategies." *Noûs* 13, no. 4 (1979): 419–37. https://doi.org/10.2307/2215337.

Dorr, Cian. "Against Counterfactual Miracles." *Philosophical Review* 125, no. 2 (2016): 241–86. https://doi.org/10.1215/00318108-3453187.

Dorst, Chris. "Why Do the Laws Support Counterfactuals?" *Erkenntnis* 87, no. 2 (2022): 545–66. https://doi.org/10.1007/s10670-019-00207-1.

Edgington, Dorothy. "On Conditionals." *Mind* 104, no. 414 (1995): 235–329. https://doi.org/10.1093/mind/104.414.235.

———. "Counterfactuals and the Benefit of Hindsight." In *Cause and Chance: Causation in an Indeterministic World*, edited by Phil Dowe and Paul Noordhof, 12–27. New York: Routledge, 2004.

Elga, Adam. "Statistical Mechanics and the Asymmetry of Counterfactual Dependence." *Philosophy of Science* 68, no. 3 (2001): S313–24. https://doi.org/10.1086/392918.

———. "Isolation and Folk Physics." In *Causation, Physics, and the Constitution of Reality*, edited by Huw Price and Richard Corry, 106–19. Oxford: Oxford University Press, 2007.

Fernandes, Alison. "A Deliberative Approach to Causation." *Philosophy and Phenomenological Research* 95, no. 3 (2017): 686–708. https://doi.org/10.1111/phpr.12443.

———. "Freedom, Self-Prediction, and the Possibility of Time Travel." *Philosophical Studies* 177, no. 1 (2020): 89–108. https://doi.org/10.1007/s11098-018-1181-9.

———. "Time Travel and Counterfactual Asymmetry." *Synthese* 198, no. 3 (2021): 1983–2001. https://doi.org/10.1007/s11229-019-02186-w.

———. "How to Explain the Direction of Time." *Synthese* 200, no. 5 (2022a): 1–30. https://doi.org/10.1007/s11229-022-03818-4.

———. "Back to the Present: How Not to Use Counterfactuals to Explain Causal Asymmetry." *Philosophies* 7, no. 2 (2022b): 43. https://doi.org/10.3390/philosophies7020043.

———. "Time, Flies, and Why We Can't Control the Past." In *The Probability Map of the Universe: Essays on David Albert's Time and Chance*, edited by Barry Loewer, Brad Weslake, and Eric Winsberg, 312–34. Cambridge, Mass.: Harvard University Press, 2023.

Hall, Ned. "Two Mistakes About Credence and Chance." *Australasian Journal of Philosophy* 82, no. 1 (2004): 93–111. https://doi.org/10.1080/713659806.

Hájek, Alan. "Probabilities of Counterfactuals and Counterfactual Probabilities." *Journal of Applied Logic* 12, no. 3 (2014): 235–51. https://doi.org/10.1016/j.jal.2013.11.001.

———. "Counterfactual Scepticism and Antecedent-Contextualism." *Synthese* 199 (2021): 637–59. https://doi.org/10.1007/s11229-020-02686-0.

Hitchcock, Christopher, and Miklós Rédei. "Reichenbach's Common Cause Principle." In *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/sum2021/entries/physics-Rpcc/.

Horwich, Paul. *Asymmetries in Time: Problems in the Philosophy of Science*. Cambridge, Mass.: MIT Press, 1987.

Huggett, Nick. "Reading the Past in the Present." In *The Probability Map of the Universe: Essays on David Albert's Time and Chance*, edited by Barry Loewer, Brad Weslake, and Eric Winsberg, 271–93. Cambridge, Mass.: Harvard University Press, 2023.

Ismael, Jenann. "Decision and the Open Future." In *The Future of the Philosophy of Time*, edited by A. Bardon, 141–58. London: Routledge, 2012.

———. "An Empiricist's Guide to Objective Modality." In *Metaphysics and the Philosophy of Science: New Essays*, edited by Matthew Slater and Zanja Yudell, 109–25. New York: Oxford University Press, 2017.

Kutach, Douglas. "The Entropy Theory of Counterfactuals." *Philosophy of Science* 69, no. 1 (2002): 82–104. https://doi.org/10.1086/338942.

Leitgeb, Hannes. "A Probabilistic Semantics for Counterfactuals." *The Review of Symbolic Logic* 5, no. 1 (2012): 26–121. https://doi.org/10.1017/s1755020311000153.

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

———. "Counterfactual Dependence and Time's Arrow." *Noûs* 13 (1979): 455–76. https://doi.org/10.2307/2215339.

———. "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61 (1983): 343–77. https://doi.org/10.1080/00048408312341131.

———. "A Subjectivist's Guide to Objective Chance." In *Philosophical Papers: Volume II*, 83–132. Oxford: Oxford University Press, 1986a.

———. "Causation." In *Philosophical Papers: Volume II*, 159–213. Oxford: Oxford University Press, 1986b.

———. "Causal Decision Theory." In *Philosophical Papers: Volume II*, 305–56. Oxford: Oxford University Press, 1986c.

———. "Humean Supervenience Debugged." *Mind* 103 (1994): 473–90. https://doi.org/10.1093/mind/103.412.473.

Loewer, Barry. "Determinism and Chance." *Studies in History and Philosophy of Modern Physics* 32 (2001): 609–29. https://doi.org/10.1016/s1355-2198(01)00028-4.

———. "Counterfactuals and the Second Law." In *Causation, Physics, and the Constitution of Reality*, edited by Huw Price and Richard Corry, 293–326. Oxford: Oxford University Press, 2007.

———. "Two Accounts of Laws and Time." *Philosophical Studies* 60 (2012): 115–37. https://doi.org/10.1007/s11098-012-9911-x.

———. *Laws of Nature and Chances: What Breathes Fire into the Equations*. Oxford: Oxford University Press, Forthcoming 2024.

Maudlin, Tim. *The Metaphysics within Physics*. New York: Oxford University Press, 2007a.

———. "What Could Be Objective about Probabilities?" *Studies in History and Philosophy of Modern Physics* 38 (2007b): 275–91. https://doi.org/10.1016/j.shpsb.2006.04.006.

Price, Huw. "Agency and Probabilistic Causality." *British Journal for the Philosophy of Science* 42 (1991): 157–76. https://doi.org/10.1093/bjps/42.2.157.

———. "Causation, Chance, and the Rational Significance of Supernatural Evidence." *Philosophical Review* 121, no. 4 (2012): 483–538. https://doi.org/10.1215/00318108-1630912.

———. "Where Would We Be Without Counterfactuals?" *MS*. http://philsci-archive.pitt.edu/10040/ (accessed June 17, 2020).

Price, Huw, and Brad Weslake. "The Time-Asymmetry of Causation." In *The Oxford Handbook of Causation*, edited by H. Beebee, C. Hitchcock, and P. Menzies, 414–43. Oxford: Oxford University Press, 2009.

Quine, Willard van Orman. *From Stimulus to Science*. Cambridge, Mass.: Harvard University Press, 1995.

Reichenbach, Hans. *The Direction of Time*. Edited by M. Reichenbach. Berkeley-Los Angeles: University of California Press, 1956.

Russell, Bertrand. "On the Notion of Cause." *Proceedings of the Aristotelian Society, New Series* 13 (1912–13): 1–26.

Skyrms, Brian. *Causal Necessity*. New Haven, Conn.: Yale University Press, 1980.

Strevens, Michael. "The Explanatory Role of Irreducible Properties." *Noûs* 46, no. 4 (2012): 754–80. https://doi.org/10.1111/j.1468-0068.2010.00802.x.

Verhaegh, Sander. *The Nature and Development of Quine's Naturalism*. Oxford: Oxford University Press, 2018.

Woodward, James. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press, 2003.