**As Time Goes By: Reichenbach on the Causal Direction of Time**

**Clark Glymour**

**Abstract**

We attempt to reconstruct Hans Reichenbach’s arguments for a macroscopic causal definition of the direction of time. Our analysis reveals that Reichenbach’s formulation of “screening off” is equivocal between the now common notion of conditional independence of two variables given others and a weaker notion that requires the conditional independence only for specific values of the variables. We also find that on the now common notion of screening off his own conditions for the “usual…conjunctive forks” are mathematically impossible for binary variables. Finally, we note that as a corollary to his familiar Principle of the Common Cause, Reichenbach’s argument embraces a No Fatalism principle that forbids explaining earlier probabilistic associations by values of later variables.

“The past never comes back.”

“We cannot change the past, but we can change the future.”

 -Hans Reichenbach, *The Direction of Time*

“The fundamental things apply

 As time goes by”

 -Herman Hupfeld

1. **Introduction**

The direction of time is many things and many issues, depending on the expositor. For Kant, it was the imposition of an a priori numerical order on the matter of experience. For others, it is a contingent, irreducible feature of the universe. For some, it is a question of how to obtain a direction of time from time-reversal invariant laws. For others, it is to specify a physical phenomenon—increasing entropy is the primary candidate—that corresponds to our cosmological theories and produces the time asymmetric phenomena of aging, memory and agency. For Hans Reichenbach, it was also how the direction of influence between macroscopic events can be derived from asymmetries of probability relations that we can observe or infer from events outside of ourselves. As in his much of his other work, he mixed metaphysics with conceptual analysis with epistemology, and so for that reason alone his story is sometimes difficult to untangle. And there is another feature of *The Direction of Time* that contributes to difficulty in unraveling the arguments. We do not know when Reichenbach began seriously thinking about the direction of time, or began the book manuscript which he did not complete before his death. The arguments in *The Direction of Time* are therefore unsurprisingly a pastiche, as though assembled from various notes, with ambiguities and extended discussions that seem to be digressions but were plausibly pieces of Reichenbach’s thinking at various times with attendant vagaries of notation. We will try to trace through some of Reichenbach’s reasoning, ignoring several digressions that are not essential to the main line of argument. Our conclusion is that Reichenbach’s attempt at a probabilistic, macroscopic theory of the direction of time based on the direction of causal relationships is unsound. That is no surprise. Perhaps more interesting is that Reichenbach’s notion of “screening off” is ambiguous, and with the now conventional understanding of that relation Reichenbach’s conditions for the “usual…conjunctive forks,” are impossible to satisfy. And, finally, that Reichenbach rests his argument on a kind of “No Fatalism” principle that complements his Principle of the Common Cause.

Reichenbach’s ambition to forge a plausible formal connection between probability relations and causal representations meets difficulties that were only remedied more than a generation after he wrote.

1. **The Principle of the Common Cause and the No Fatalism Principle**

Reichenbach supplemented his Principle of the Common Cause—of two probabilistically associated events, either one causes the other, or they have a common cause[[1]](#footnote-1)— with a No Fatalism Principle that associations of events cannot be explained by their effects; “fatalistic” explanation is not allowed. The No Fatalism Principle can be seen as a corollary of the Principle of the Common Cause to which Reichenbach appeals to save his attempted construction of the direction of time from difficulties arising from associations produced by conditioning on a common effect. Reichenbach was essentially wrestling with Berkson’s paradox (really, Sewell Wright’s paradox, 1918), which he seems to have rediscovered for himself.[[2]](#footnote-2)

**2. The “Lineal Ordering” of Events**

Reichenbach distinguishes an “order of time” from a “direction of time.”(32). By an “order of time” he intends a system of causal and temporal betweenness relations, which he calls a “lineal order.” The possible causal relations among three types of events, A, B, C are

* + - 1. A -> C -> B or
			2. A <- C <- B or
			3. A -> C <- B;
			4. A <- C -> B

The aim of this section of the book seems to be to establish that for deterministically related locally observable events, causal orders 1 and 2.—the causal chain cases—can be distinguished from 3 and 4 but not from each other, and 4, the common cause case, cannot be distinguished from 3, the common effect case. These results, Reichenbach, argues, are due to the time reversibility of classical physical dynamics.

﻿“…the laws of mechanics can very well inform us about temporal between-relations. They tell us that there is a causal line from A by way of B to C, or from C by way of B to A, and thus leave merely the direction of the line undetermined.”(33)[[3]](#footnote-3)

To illustrate why 4 and 3, the common cause and the common effect cases, are not distinguishable by local observation and classical physics, Reichenbach considers two balls observed at the same position A at some time, either moving to, or from, their respective position-moments E and C.

:

﻿

﻿Fig. 1. paths of two balls one passing from A to C and the other from A to E, in the respective directions of the solid arrows. The broken arrows indicate the motions in reversed time.

﻿“Both interpretations [of the direction of time] are compatible with the laws of mechanics. Thus we know that we have here two causal lines, which either both start at A, or both end at A. However, we cannot interpret one ball as moving in the direction of the broken arrow and the other ball as moving in the direction of the solid arrow. Therefore, though we do not know the direction of the lines, we do know a relation between the directions. We know that the two lines cannot be combined into one; they are counterdirected. The latter term expresses an order property, which holds for both directional interpretations. ” (32)

And he added:

“We will assume that strict coincidences, as well as approximate coincidences, are observable. This is a continuity assumption; it means that the approximate coincidences in the space-time neighborhood of a strict coincidence are defined observationally just as well as strict coincidence.” (33).

Which is intended, we think, to stipulate that the simultaneity, or not, of the balls at E and C can be observed.

“The determination of counterdirected lines leads to negative statements concerning the time order of events belonging to different lines. Thus, we conclude that the passage at A is not temporally between the passage at D and the passage at B. Either the passage at A is temporally before the two other passages or it is after them. We see that mechanical laws can establish and can deny the existence of temporal between-relations.” (36)

Why can’t we “interpret one ball as moving in the direction of the broken arrow and the other ball as moving in the direction of the solid arrow”? Reichenbach’s thought seems to have been that the simultaneity, or not, of balls at C and E can be observed, and it can be arranged that they are simultaneous, excluding cases 1 and 2.

Finally, we have this remark, which Reichenbach does not elaborate:

 “Applying similar methods to further events, we thus can construct a causal net which, as a whole, has a lineal order. This term means that, if a direction is assigned to one line, a direction is determined for each line.” (36)

The last sentence is not true. Given that A - > C; A – B; A – D, so that A is between C and B, and also between C and D and between B and D, the directions of B – A - D connections are not determined.

**3. Causal Betweenness from Probabilities**

Section IV of *The Direction of Time* has an extended discussion of the relation of macro and micro statistics, claims relating causal relations to entropy and to information theory and the first published statement of the Principle of the Common Cause, which Reichenbach had used without naming in *Experience and Prediction* as part of a causal discovery procedure.

To assign a direction to time based on causal asymmetries, Reichenbach appealed to his frequentist account of probability. Events are clustered as instances of the same type or class or sequence, the relative frequency of a class of events in a reference sequence is the estimate of its probability. Probabilities of events are probabilities of their respective classes or sequences. How the reference sequences or classes are chosen—what the relative frequencies are relative to--is not discussed; Reichenbach directed the reader to his *Theory of Probability.*

Some formal connection between causal order—which Reichenbach represented by directed graphs—and probability distributions is essential for inferences from probability relations to causal relations. That event A2 is causally between events A1 and A3 seems a place to start.

Reichenbach formalized the probabilistic criteria for causal betweenness this way. (His notation is ambiguous between using “Ai” to denote a positive (i.e., occurs) value of a binary random and denoting the random variable itself, i.e. specifying that the inequalities hold for all values of the random variables. In some contexts, as in inequalities (1) and (2) below his notation is most charitably interpreted in the former way, and in some contexts, as in inequality (3) below, in the latter way.)

Definition: Event A2 is causally between events A1 and A3 if the relations hold



Reichenbach’s notation for conditional probabilities--the variable or variable values on the left of the comma are conditioned on--is the reverse of modern notation, which we will use with a vertical bar in place of the comma. So Reichenbach’s P(A2, A1) becomes p(A1 | A2). Thus:

1 > p(A3 | A2) > p(A3| A1) > p(A3) > 0 1’

1 > p(A1 | A2) > p(A1 | A3) > p(A1) > 0 2’

 p(A3 | A2, A1) = p(A3 | A2) 3’

Reichenbach calls equation (3) “screening off.” Note that with the positive interpretation of the symbols (Ai = 1) Reichenbach’s \*(3) does not imply that

4’ p(A3 = 1 | A2 = 1, A1 = 1) = p(A3 = 1 | A2 = 1, A1 = 0).

So the value of A3 could depend on whether A1 does not occur. Later in the book (see equations (5) – (8) below), Reichenbach clearly takes his letters to denote positive values and introduces a complementation sign for negative values. Reichenbach’s notion of screening off is generally taken to require conditional independence, hence 4’, and we will understand it that way, admitting that Reichenbach’s intention is unclear to us.

Inequalities (1) and (2) entail that A1 and A3 are dependent. Reichenbach gave them a further purpose:

“The first inequality in (1) and the first inequality in (2) are required, because otherwise relations … (3) would be trivially satisfied. For instance, from p(A2, A1) = 1 we could immediately derive (3), because the addition of the factor “A1” in the first term of the probability expression “p(A2, A3)” does not change the value of p(A2, A3) if p(A2, A1) = 1” (191).

If that were the purpose of (1) and (2), it would have been sufficient to specify their respective inequalities with respect to 0 and 1. It seems Reichenbach’s real purpose with these conditions was to use (1) and (2) as probability measures of “causal distance.”

 “This statistical determination becomes weaker with growing causal distance, a fact which is expressed in relations (1)–(2).” (191-192)

That is of course not true if “causal distance” means the minimum number of edges in a directed path from one variable to another. For example, in Figure 2, even with each effect more probable conditional on any one of its causes than unconditionally, X4 can be more probable conditional on X1 than on either of X2 or X3 alone, although X2 and X3 are both closer to X4 than is X1:

 X1

 X2 X3

 X4

 Figure 2

Reichenbach takes note of this:

﻿“…it may happen that neither [event] screens off A1 from A3, whereas their conjunction does so. …We can therefore speak of a set …screening off A1 from A3.(203-204)

But he does not observe that this undoes his proposed account of causal distance for individual events, which would have to be made somewhat more complicated to take account of sets of intervening variables.

Reichenbach defines “causal relevance” as follows:

﻿DEFINITION 2. An event A1 is causally relevant to a later event A3 if and there exists no set of events . . . which are earlier than or simultaneous with A1 such that this set screens off A1 from A3. (204)

That, at least, is a tenable connection among causal relations, probabilities and time order if events have a discrete topology.

1. **Conjunctive Forks and Causal Markov Condition**

A central theme in Reichenbach’s account of macroscopic causation is the “conjunctive fork,” which he takes to describe either of two causal structures.

A B A B

 C E

Figure 3 (R's Figure 24) Figure 4 (R's Figure 25)

﻿“Fig. 24. Fork open toward the future, constituted by a common cause. Fig. 25 Fork open toward the past, constituted by a common effect.” (159)

A “double fork” (159) combines these:

 E

A B

C

﻿Figure 5 (R's Figure 23) Double fork, constituted by a common cause and a common effect.

Reichenbach then assumes that the causal structures in his Figure 24 satisfies these conditions (159):



Note that (5) and (6) assert that A and B are independent conditional on C as a binary variable). He says that these probability relations define a “conjunctive fork.” He claims that there are cases in which these conditions for a conjunctive fork are also satisfied by the common effect model of his Figure 25. So both structures, 24 and 25 depict causal relations that can be, and according to Reichenbach, usually are, “conjunctive forks.”[[4]](#footnote-4)

﻿

“Can we now use conditions (5)–(8) to determine the time direction of the fork, that is, to distinguish between figure 24 and figure 25? It is easily seen that this cannot be done directly. A common effect E, that is, the upper fork of figure 23, can very well satisfy the relations (5)–(8), with “E” in the place of “C”. In fact, this is usually true.” (160-161)

So Reichenbach says here that two independent causes of a common effect *can be* independent conditional on their common effect, and “this is usually true.” Further,

“Each of these [his figures 24 and 25] can satisfy relations (1) - (3). The common cause, as well as the common effect, can screen off the event A1 from A3.”(193)

No, it cannot if screening off requires conditional independence. With that understanding the passage just quoted adds a further constraint, the independence of the causes conditional on the effect in the common effect model of Figure 25.

In a change of notation, the following three conditions are necessary consequences of the “conjunctive fork” axioms “ with 4’ and with Reichenbach’s extra conditional independence constraint just noted:

1. X1 is not independent of X3 ; X2 is not independent of X3;
2. X1 is independent of X2;
3. X1 is independent of X2 conditional on X3

These conditions cannot be satisfied by any joint probability distribution on binary variables, X1, X2, X3. “Conjunctive forks” with common effect are impossible for binary variables.

By the law of total probability: P(X1, X2) = ΣX3 P(X1, X2, X3) = ΣX3 P(X1, X3)P(X2|X1, X3).

Then by condition 3:

P(X1, X2) = ΣX3 P(X1, X3)P(X2|X3) = ΣX3 P(X3|X1)P(X1)P(X2|X3) =

P(X1)ΣX3 P(X3|X1)P(X2|X3) (I)

By condition 2: P(X1, X2) = P(X1)P(X2) (II)

If P(X1) ≠ 0, from (I) and (II) we have: P(X2) = ΣX3 P(X3|X1)P(X2|X3) By the law of total probability: P(X2) = ΣX3 P(X3)P(X2|X3) Therefore, for condition 1 to be consistent with condition 2 and 3, the following (in)equations should be satisfied a. ΣX3 P(X3|X1)P(X2|X3) = P(X2) b. P(X3|X1) ≠ P(X3) When X3 is binary, there is no model satisfying both a and b because there is a unique solution of P(X3|X1) for a. When X3 is not binary, there are infinitely many solutions for P(X3|X1) that satisfy a, and we can find model satisfying the three conditions 1, 2 and 3.

A version of the argument applies as well for the case when X3 = 1.

1. **The Direction of Time**

Now, the punchlines:

﻿“…if there exists a conjunctive fork with respect to a common effect E, the simultaneous occurrence of A and B is more probable than a mere chance coincidence. Consequently, if there were no common cause C, the common effect would establish a statistical dependence between A and B; ﻿and explanation would be given in terms of a “final cause”. Referring to the discussion given in §18, we regard final causes as incompatible with the second law of thermodynamics and consider such forks impossible. In application to the present investigation this means: The principle of the common cause does not exclude, throughout, a statistical dependence with respect to a common effect; but it does exclude such dependence if ﻿there exists no common cause. In contrast, it is quite possible that there is no common effect although there is a common cause satisfying (5)–(8). This result can be used for a definition of time direction, as follows:

﻿“If a fork ACB is conjunctive, we say that the fork is closed at C. If there is no event E on the other side of the fork which satisfies (5)–(8), we say that the fork is open on that side. We now define:

﻿ ﻿DEFINITION. In a conjunctive fork ACB which is open on one side, C is earlier than A or B.” (162-163)”

So this is how Reichenbach would establish the macroscopic direction of time from macroscopic probability relations! To establish the direction of time from macroscopic probability relations among events, Reichenbach appeals to the inadmissibility of “final causes,” appeals opaquely to thermodynamics, and claims that, although “unusual,” a common effect of two variables could create an association between them without a common cause, such an occurrence would violate the Principle of the Common Cause because there would be an association of two variables even when they have no common cause and no causal influence of one on the other. In effect, Reichenbach specifies an a priori constraint on Nature: *no two events can be causes of a third event unless the causes themselves share a common cause.*  His metaphysics may be saved by the Big Bang, but as a practical principle for “local” scientific inference it is not tenable or useful.

**6. The Field Argument**

Harty Field has suggested a way that repeatable events could generate a direction of time from Reichenbach’s example using the Causal Markov Condition (CMC), an account that Reichenbach manifestly did not give. Field[[5]](#footnote-5) gives such an account explicitly; Arntzenius[[6]](#footnote-6) comes close. Field suggests this could be extended to order (almost) all events in time with the well-known PC algorithm.[[7]](#footnote-7)

1. **Conclusion**

It is fair, we think, to conclude that at the time of writing *The Direction of Time*, Reichenbach struggled to find a consistent connection between joint probability distributions and directed graphical representations of causal relationships. That problem was largely solved thirty years later. And he did not know quite what to do about probabilities conditional on common effects. In Reichenbach’s defense, at the time he was writing, and long after, many statisticians were confused about conditioning on common effects. Reichenbach’s critical puzzlement was that conditioning on a common effect of independent variables can result in their dependence. Although he thought wrongly that this is unusual, it challenges his own version of the Principle of the Common Cause, which we see he intended to apply to conditional as well as unconditional associations. The facts of the matter are almost the reverse of Reichenbach’s reasoning. It is the statistics of common effects, not the statistics of common causes, that can determine a causal ordering.

The most influential idea in *The Direction of Time* was the conjunctive fork, which, the considerable literature about it notwithstanding, does not exist as Reichenbach thought it most commonly applies. Typical discussions leave out Reichenbach’s requirement that the two causes be independent conditional on their common effect.

Nowadays, we know multiple circumstances in which statistical relations determine the direction of causal relations, even between just two variables.[[8]](#footnote-8) Reichenbach, of course, could not have known of these developments, but they essentially solve versions of the problem as he posed it. We are happy to think that were he living today Reichenbach would be fully engaged with developing methods to recover causal relationships from probabilities.

.

**References[[9]](#footnote-9)**

Arntzenius, F., & Dorr, C. (2014). *Space, time, and stuff*. Oxford University Press (UK).

Berkson, Joseph (1946). [Limitations of the Application of Fourfold Table Analysis to Hospital Data"](http://ije.oxfordjournals.org/content/43/2/511.full). *Biometrics Bulletin*. **2** (3): 47–53.

Field, H. (2003). Causation in a physical world. *Oxford handbook of metaphysics*, 435-460.

Reichenbach, H. *The Direction of Time*. Kindle Edition

Sober, E. (2011). Reichenbach’s cubical universe and the problem of the external world. *Synthese*, *181*(1), 3-21.

Sober, E. and Barrett, J. “Conjunctive Forks and Temporally Asymmetric Inference,”*Australasian Journal of Philosophy* Vol. 70, No. 1; March 1992

Spirtes, P., C. Glymour and R. Scheines, *Causation, Prediction and Search*, MIT, 2000.

1. We adopt Elliot Sober’s interpretation of what Reichenbach intended by the Principle of the Common Cause but did not write. Sober, E. (2011). Reichenbach’s cubical universe and the problem of the external world. *Synthese*, *181*(1), 3-21. [↑](#footnote-ref-1)
2. Berkson, Joseph (June 1946). [Limitations of the Application of Fourfold Table Analysis to Hospital Data"](http://ije.oxfordjournals.org/content/43/2/511.full). *Biometrics Bulletin*. **2** (3): 47–53. [↑](#footnote-ref-2)
3. All references to *The Direction of Time* are from the Kindle edition. [↑](#footnote-ref-3)
4. An exposition of why this is not “usually true” is given in Sober and Barrett, “Conjunctive Forks and Temporally Asymmetric Inference,”*Australasian Journal of Philosophy* Vol. 70, No. 1; March 1992 [↑](#footnote-ref-4)
5. Field, H. (2003). Causation in a physical world. *Oxford handbook of metaphysics*, 435-460. Some further assumptions beyond the CMC would be required and not all causal connections would be assigned directions by the PC algorithm. [↑](#footnote-ref-5)
6. Arntzenius, F., & Dorr, C. (2014). *Space, time, and stuff*. Oxford University Press (UK). [↑](#footnote-ref-6)
7. Spirtes, P., C. Glymour and R. Scheines, *Causation, Prediction and Search*, MIT, 2000. [↑](#footnote-ref-7)
8. A brief review and references is provided in Glymour, C., Zhang, K., & Spirtes, P. (2019). Review of causal discovery methods based on graphical models. *Frontiers in genetics*, *10*, 524. [↑](#footnote-ref-8)
9. Thank to Daniel Malinsky for many corrections. [↑](#footnote-ref-9)