

ELLIOTT SOBER

TEMPORALLY ORIENTED LAWS¹

ABSTRACT. A system whose expected state changes with time cannot have both a forward-directed translationally invariant probabilistic law and a backward-directed translationally invariant law. When faced with this choice, science seems to favor the former. An asymmetry between cause and effect may help to explain why temporally oriented laws are usually forward-directed.

1. THE PROBLEM

Much has been written about laws that say that some quantity increases with time in the evolution of systems of a certain type. The second law of thermodynamics is the most famous example. Such laws are said to embed an asymmetry between earlier and later. In this paper, I want to discuss a rather different property of laws. Laws can allow us to calculate the future from the past or the past from the future.² Laws that do one of these but not the other, I will call *temporally oriented*.

The strict second law is not temporally oriented. If I observe the present entropy of a closed system, I can infer that the entropy will be no less in the future *and* that it was no greater in the past. A law that posits a monotonic increase (or nondecrease) in a quantity permits inference in both directions.

Non-probabilistic laws that describe a conditional relationship between earlier and later are not temporally oriented. A law of the form 'If the system is in state E at an earlier time, then it will be in state L at a later time' supports inferences in both directions. A glimmer of the asymmetry we seek is to be found in the idea of conditional probability. A law that assigns a value to a probability of the form 'Pr (the system later is in state L/the system earlier is in state E)' permits forward-, but not backward-, directed inferences.³ And, of course, precisely the opposite will be true if conditioning and conditioned propositions are reversed.⁴

In saying that 'Pr (the system later is in state L/the system earlier is in state E)' permits a forward, but not a backward, inference, I set to one side the possibility of using this law in a Bayesian format wherein

prior probabilities plus the forward-directed law allow one to compute \Pr (the system earlier is in state E/the system later is in state L). The point is that the forward-directed law, *by itself*, permits no such inference. Bayesians insist that priors are always available; their critics disagree. However, for the purposes of investigating the concept of a temporally oriented law, this controversy may be bracketed.

Are scientific laws often temporally oriented? Among laws that are temporally oriented, are more of them oriented to the future than to the past? When a law is temporally oriented, is this an artifact of our interests or does it reflect some objective feature of the world? These are the sorts of questions I wish to investigate.

2. A PROOF

Consider any three times t_1 , t_2 , and t_3 , which occur in that order, and which are equally spaced. Let ' $\Pr(t_i = x)$ ' denote the probability that the system at time t_i is in state x ($i = 1, 2, 3$; $x = 0, 1, 2, \dots$). The laws⁵ governing the evolution of the system will be given by conditional probabilities of the form $\Pr(t_i = x/t_j = y)$. I will say that the law is *forward-directed* if $i > j$ and *backward-directed* if $i < j$.⁶

A forward-directed law is *invariant under time translation* if $\Pr(t_2 = x/t_1 = y) = \Pr(t_3 = x/t_2 = y)$, for any equally spaced t_1 , t_2 , t_3 and for any x and y . That is, the 'date' at which a temporal interval begins or ends is irrelevant to calculating the probability of the system's state after some fixed amount of time has elapsed; all that matters is the system's (undated) initial state (specified in the conditioning proposition) and the amount of time that elapses between it and the end state.⁷ Likewise, a backward-directed law is time translationally invariant if $\Pr(t_1 = x/t_2 = y) = \Pr(t_2 = x/t_3 = y)$, for any equally spaced t_1 , t_2 , t_3 and for any x and y .

It is worth noticing two features of this definition of time translational invariance. The definition requires that two conditional probabilities be equal, for any *fixed amount of spacing* between the three times and for any *states* that the system may occupy at a time. To see what these two constraints amount to, let us consider models in population genetics that relate the gene frequencies at one time in a population to some probability distribution of gene frequencies in the population some time later. An example is provided by models of neutral evolution, in which gene frequencies evolve by random walk (Crow and Kimura 1970;

Kimura 1983). These models are time translationally invariant in the strong sense defined. First, the fixed amount of spacing between the three times may be any number of generations you please. Second, the model applies to all possible initial gene frequencies. That is, the theory of neutral evolution is time translationally invariant in the following sense:

For any starting time t , for any number of generations n , and for any pair of gene frequency distributions i and j , $\Pr(\text{Population is in state } j \text{ at time } t + n / \text{Population is in state } i \text{ at } t) = \Pr(\text{Population is in state } j \text{ at time } t + 2n / \text{Population is in state } i \text{ at } t + n)$.

This strong concept of time translational invariance will be assumed in what follows, although weaker notions of invariance are certainly possible. A law might be invariant for some amounts of spacing between the three events but not others; and it might be invariant for some possible states of the system but not others. A consequence of weakening the definition will be considered in due course.

I now will show that a system whose expected state⁸ changes with time cannot have both a forward-directed time translationally invariant law and a backward-directed time translationally invariant law. If we assume that laws must be translationally invariant,⁹ then this simplifies to: if a system's expected state changes, then it cannot have both a forward-directed probabilistic law and a backward-directed probabilistic law.

I'll begin by assuming that the system must be in one of two states (0 or 1) at any time; I'll dispense with this assumption shortly.

If the system possessed a backward-directed translationally invariant law, the following equality would have to hold:

$$(B) \quad \Pr(t_1 = 1/t_2 = 0) / \Pr(t_1 = 0/t_2 = 0) = \Pr(t_2 = 1/t_3 = 0) / \Pr(t_2 = 0/t_3 = 0).$$

Bayes' theorem allows (B) to be expanded into the following:

$$\Pr(t_2 = 0/t_1 = 1) \Pr(t_1 = 1) / \Pr(t_2 = 0/t_1 = 0) \Pr(t_1 = 0) = \Pr(t_3 = 0/t_2 = 1) \Pr(t_2 = 1) / \Pr(t_3 = 0/t_2 = 0) \Pr(t_2 = 0).$$

If the system possessed a forward-directed time invariant law, it would be true that:

$$(F) \quad \frac{\Pr(t_2 = 0/t_1 = 1)}{\Pr(t_2 = 0/t_1 = 0)} = \frac{\Pr(t_3 = 0/t_2 = 1)}{\Pr(t_3 = 0/t_2 = 0)}.$$

If we assume (F) and that the ratios mentioned in it are well defined but not zero, then the Bayesian expansion of (B) simplifies to:

$$(C) \quad \Pr(t_1 = 1)/\Pr(t_1 = 0) = \Pr(t_2 = 1)/\Pr(t_2 = 0).$$

This last statement means that the expected state of the system does not change with time.

The proof easily generalizes to any finite number of states. Let 'x' and 'y' be variables that take states (0, 1, 2 . . .) as values. Parallel reasoning then entails that for any x and y:

$$\Pr(t_1 = x)/\Pr(t_1 = y) = \Pr(t_2 = x)/\Pr(t_2 = y).$$

If each pairwise ratio of this form must remain constant, then the expected state of the system cannot change.¹⁰

To say that the expected state does not change with time differs from saying that the state can be expected not to change. A particle doing a random walk on an open line has the same expected position throughout its history; on average it goes nowhere.¹¹ But this does not mean that we expect the particle to stay at the same location. Proposition (C) implies a lack of *directionality* in the system's probability trajectory.¹²

This result does not rule out evolution in the expected state of a system when the system has a forward-directed law *for one set of properties* and a backward-directed law *for some other set of properties*. Note that the argument considered forward- and backward-directed laws with respect to the same set of properties ($t_i = 0, 1, 2 . . .$). The point is that for each set of properties, a choice must be made between a forward-directed law and a backward-directed law, if laws are to be time translationally invariant and the system's expected state is to evolve.

The above argument cannot be carried through for systems all of whose transition probabilities are either 0 or 1; for in such cases the ratios in (F) either will equal zero or will not be well defined. The asymmetry we have established does not apply to *deterministic laws* (in at least one sense of that term).¹³

In the process here described, the effect of what happens at t_1 is itself the cause of what happens at t_3 . But the idea that one and the same event is both effect and cause is not essential to the proof just

offered. Consider, for example, the causal relationship between smoking (S) and dying of lung cancer (D). Smoking at one time causally contributes to dying sometime later; but dying at one time does not contribute to subsequent smoking. Even so, it is easy to show that our proof applies. Two applications of Bayes' theorem yield the following equation:

$$\frac{P(S/D)}{P(-S/D)} = \left[\frac{P(D/S)}{P(D/-S)} \right] \left[\frac{P(S)}{P(-S)} \right]$$

The left side of the equation describes a ratio of backward-directed conditional probabilities. The first product term on the right describes a ratio of forward-directed conditional probabilities. Suppose this forward ratio is time translationally invariant. Then the ratio on the left side of the equation is time translationally invariant only if $P(S)/P(-S)$ is as well. I conclude that forward-directed and backward-directed conditional probabilities cannot both be time translationally invariant, if the probability of the cause (smoking, in this case) changes.

I mentioned earlier that my argument makes use of a rather strong concept of time translational invariance. It is worth considering the effect of weakening it in a certain way. Instead of considering time translational invariance as covering all amounts of spacing between the three times, let us consider the weaker idea of time translational invariance with respect to the temporal interval k (where k is a constant). If the forward- and backward-directed laws are both time translationally invariant with respect to k , then we can deduce that for any states x and y

$$\Pr(t_1 = x)/\Pr(t_1 = y) = \Pr(t_2 = x)/\Pr(t_2 = y),$$

where t_1 and t_2 are any two times separated by a temporal interval that is k units in length. In such a case, we cannot conclude that the system's expected state never changes, but only that it has the same value every k units of time. That is, with this weakened notion of invariance, we may conclude that the system's expected state is constant *or that it cycles with period k* . Not surprisingly, a weaker result is obtained for a weaker notion of invariance.

3. BREAKING THE SYMMETRY

It is fairly clear that science finds it natural to think in terms of forward-directed probabilities. We talk of the half-life of uranium. This is the amount of time it takes for the radioactivity of a sample to drop to half its initial value; it reflects a forward-directed probability that is assumed to be translationally invariant. In population genetics, Mendelism is a process by which parental genotypes confer probabilities on offspring genotypes; this, too, reflects a forward-directed probability assumed to be translationally invariant.

This temporal asymmetry in scientific concepts is also present in the concepts we use in everyday life. Roads are *dangerous*; skies are *threatening*; people are *friendly*. These familiar dispositional concepts, once they are understood probabilistically,¹⁴ can be seen to describe forward-directed probabilities.

The proof of the previous section shows that we must choose between forward-directed and backward-directed laws. It appears that science and common sense have opted for the former. Is there some general feature of the world that explains why this should be so?

The probability of getting cancer if you smoke has changed through time, because other causal factors (for example, the amount of asbestos in the environment) have also changed. But once these other causal factors are held fixed, the conditional probability of getting cancer relative to them seems like it should be translationally invariant.¹⁵

The probability of having been a smoker, given that you have lung cancer, has also changed with time. But if we hold fixed the other effects of smoking (emphysema, for example), there is no reason to expect that the backward-directed conditional probability is temporally invariant.

One (rough) formulation of (forward-directed) *determinism* is the thesis that the state of the universe at time t_i is incompletely specified by a law, if the law assigns an intermediate value to a probability of the form $\Pr(t_j = x/t_i = y)$ (where $i < j$). The idea that completely specified causes generate time translationally invariant forward-directed laws makes no such deterministic assumption; the idea is that such probabilities should be stable under time translation, not that they should have values of 0 or 1.

Imagine an experimenter who runs a series of trials on an experimental setup one day, then waits a week and runs a second series of

trials. Suppose that the two series of trials show markedly different frequency distributions. The experimenter reasons that although it is possible that the two series were generated by the same underlying probabilities, it strains one's credulity to think so. So the experimenter infers that the probabilities of outcomes have changed from one week to the next.

The natural inference to make here is *not* to attribute the change in probabilities to 'the passage of time' and let things go at that. Rather, what the experimenter will look for is a physical change in the experimental setup. Each series of trials is to be specified by its own conditional probability; the relevant conditioning propositions must differ. The problem is to find out how. We see here a practical consequence of accepting the idea that completely specified forward-directed probabilistic laws must be time translationally invariant.

The slogan 'same cause, same effect' is a common expression of the idea of determinism. I suggest that a vestige of this principle survives the dismantling of determinism; it is the idea of 'same cause, same probability of effect'. The slogan's converse – 'same effect, same cause' – was not much cited by those who accepted a deterministic world picture, though perhaps it should have been, since Newtonian physics is time symmetric. In any event, its probabilistic analog – 'same effect, same probability of cause' – is evidently not a plausible constraint on stochastic model building in the sciences.

So the preference for forward-directed laws reflects the assumption that 'same cause, same probability of effect' is true, but 'same effect, same probability of cause' is not. In each case, sameness of cause or of effect must be understood in terms of a complete specification; without this rider, neither principle is plausible.

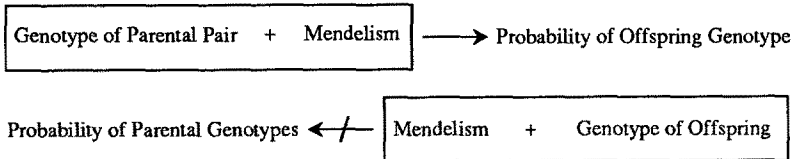
I do not claim that the choice of forward-directed over backward-directed laws is explained by the acceptance of one of these principles but not the other. Rather, our preference for forward-directed laws and our opting for 'same cause, same probability of effect' are two expressions of the same underlying idea. Is there some basic feature of the world that explains both sorts of choices?

Let us look carefully at the example of Mendelism and describe with more care how probabilities are defined there. Consider a one-locus, two-allele example. There are three genotypes (AA, Aa, and aa) possible in each sex, so there are, in principle, nine possible parental pairs; if we ignore the sex of each parent and consider only the genotypes

they contribute, then there are six pairs. For each pair we can say how probable it is that an offspring should have a particular genotype. The rows and columns denote the parental genotypes and the cells provide the probabilities of offspring genotypes, conditional on their coming from a particular cross:¹⁶

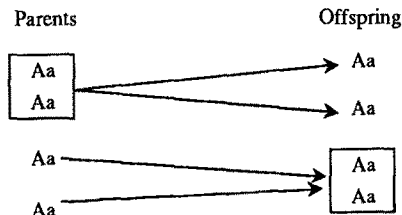
			Female	
		AA	Aa	aa
Male	AA	1(AA)	1/2(AA); 1/2(Aa)	1(Aa)
	Aa	-	1/4(AA); 1/2(Aa); 1/4(aa)	1/2(Aa); 1/2(aa)
	aa	-	-	1(aa)

Notice that the Mendelian process does *not* allow you to compute probabilities for parental genotypes from the genotypes of offspring. If the sons and daughters are heterozygotes, how probable is it that the parents were also heterozygotes? This question is unanswerable, until prior probabilities are given for the genotypes of parental pairs.¹⁷ The asymmetry under consideration can be illustrated as follows:



As noted before, an assumption of completeness is needed here. If there is unequal mortality among offspring genotypes and if the offspring genotypes are censused sometime after the egg stage, then the predicted Mendelian ratios for the offspring will be violated. The assumption, therefore, is that no such interferences impinge.

Let us now look at this asymmetry from a different angle. Suppose that two heterozygote parents produce two heterozygote offspring. We can think of the parental pair as the common cause of two effects, or we can think of the offspring pair as the common effect of two causes:



The genotype of the parental pair not only confers a probability on the offspring genotypes; in addition, the genotype of the parental pair renders the genotypes of the two offspring (O_1 and O_2) independent of each other:

$$\Pr(O_1 \text{ is Aa} \ \& \ O_2 \text{ is Aa} / \text{parents are Aa}) = \Pr(O_1 \text{ is Aa} / \text{parents are Aa}) \Pr(O_2 \text{ is Aa} / \text{parents are Aa}).$$

That is, the common cause forms a *conjunctive fork* with its two effects (Reichenbach 1956).

In contrast, the two offspring genotypes do *not* render the genotypes of the two parents (P_1 and P_2) statistically independent of each other:

$$\Pr(P_1 \text{ is Aa} \ \& \ P_2 \text{ is Aa} / \text{offspring are Aa}) \neq \Pr(P_1 \text{ is Aa} / \text{offspring are Aa}) \Pr(P_2 \text{ is Aa} / \text{offspring are Aa}).$$

I believe that this Mendelian asymmetry is typical of a larger pattern: *a common cause often forms a conjunctive fork with its joint effects, but a common effect rarely forms a conjunctive fork with its joint causes* (Sober and Barrett 1992). In the next section, I'll defend this pair of claims. For the present, I want to show how this asymmetry, if true, explains why temporally oriented laws are usually forward-directed.

If I want to describe the probability of an offspring's genotype, given the genotype of its parents, *I do not have to take into account what else is true simultaneously with that offspring*. I do not need to know what that offspring's sibs are like; I also do not need to know the prior probability of offspring genotypes. This is a consequence of the fact that the first fork displayed above is conjunctive.

In contrast, if I wish to describe the probability of a parent's genotype, given the genotypes of its offspring, I must take into account facts that pertain to the parental generation. I need to know the pattern of mating (random or assortative) and the prior probability of parental genotypes. This is a consequence of the fact that the second fork is not conjunctive; the genotypes of the offspring do not render statistically irrelevant all other facts simultaneous with the parent's having the genotype it does.

Because the second fork is not conjunctive, the probability of a cause, given its effects, cannot be purely a matter of law. The happenstance of what else is true simultaneously with the cause must also be taken into account. In contrast, the probability of an effect, given its causes, may be a matter of law, since the happenstance of what else is true

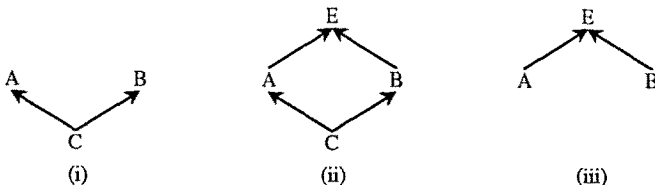
simultaneous with the effect may be ignored, if the fork is a conjunctive one. My suggestion is that laws are temporally oriented towards the future because of this asymmetry between forks.

In the example just discussed, I have treated the rules of Mendelism as providing a law governing the reproductive behavior of the organisms in a population. There is room to quarrel with the idea that organisms obey Mendelian rules purely as a matter of law. After all, the population is the result of evolution and it is perfectly reasonable to see its conformity to Mendelian patterns as a consequence of that evolution. This makes Mendelism a matter of happenstance, not a matter of law (Beatty 1981).

Although this raises interesting questions about the role of laws in evolutionary theory, it does not undermine the main point I wish to argue for here. In a Mendelian system, one can describe a forward-directed conditional probability that does not depend for its correctness on the prior distributions of genotypes at either time or on the system of mating; this cannot be done for the backward-directed conditional probability. If the prior distributions (or the system of mating) are not matters of law, then forward-directed conditional probabilities may be laws, but backward-directed conditional probabilities cannot be.

4. FORKS

Reichenbach (1956) advanced two theses that together entail an asymmetry between future and past. First, he maintained that if A and B are simultaneous and correlated events, they must have a common cause that renders them conditionally independent of each other. Reichenbach's second claim is that if A and B have a common effect, this common effect will 'usually' form a conjunctive fork with A and B. What is not allowed is that A and B lack a common cause but form a conjunctive fork with a common effect. That is, Reichenbach maintained that of the three possible patterns that conjunctive forks might display, the first two are common, whereas the third is impossible:



As Reichenbach put it, (conjunctive) forks open to the future are possible, whereas forks open to the past are not.

Reichenbach's first claim, which he called the *principle of the common cause*, is incompatible with results stemming from quantum mechanics (van Fraassen 1982). In addition, it conflicts with a perfectly 'classical' probabilistic phenomenon: two causally independent processes may be correlated because each shows a monotonic increase in some quantity (Sober 1988). Yet consistent with these corrections, a rather Reichenbachian thesis can be advanced concerning common causes and their joint effects: quantum mechanical phenomena aside, a complete specification of the causes must render the effects conditionally independent of each other.

Reichenbach's second claim is also off the mark. Conjunctive forks open to the past are not impossible. Even so, joint causes (whether they share a common cause) *only rarely* form conjunctive forks with their common effect (Sober and Barrett 1992).

To see why, let us schematize the problem as follows. A and B are joint causes of E. First, I define the four conditional probabilities of the form $\Pr(E/\pm A \& \pm B)$ as follows:

		Pr(E/-)	
		B	-B
A	w	x	
-A	y	z	

I'll call these probabilities the *efficacies of the causes*. They are all forward-directed (assuming that cause precedes effect); no assumption will be made as to whether they are invariant over time. The four probabilities of the form $\Pr(\pm A \& \pm B)$ I'll call the probabilities of *combinations of causes*.

AEB is a conjunctive fork if and only if the following equality is true:

$$\Pr(A \& B/E) = \Pr(A/E)\Pr(B/E).^{18}$$

Bayes' theorem allows these three conditional probabilities to be expressed as follows:

$$\begin{aligned}
\Pr(A\&B/E) &= \Pr(E/A\&B)\Pr(A\&B)/\Pr(E) \\
&= w\Pr(A\&B)/\Pr(E) \\
\Pr(A/E) &= \Pr(E/A)\Pr(A)/\Pr(E) \\
&= [\Pr(E/A\&B)\Pr(B/A) \\
&\quad + \Pr(E/A\&-B)\Pr(-B/A)]\Pr(A)/\Pr(E) \\
&= [w\Pr(B/A) + x\Pr(-B/A)]\Pr(A)/\Pr(E) \\
\Pr(B/E) &= \Pr(E/B)\Pr(B)/\Pr(E) \\
&= [\Pr(E/A\&B)\Pr(A/B) \\
&\quad + \Pr(E/-A\&B)\Pr(-A/B)]\Pr(B)/\Pr(E) \\
&= [w\Pr(A/B) + y\Pr(-A/B)]\Pr(B)/\Pr(E)
\end{aligned}$$

So the fork is conjunctive precisely when:

$$\begin{aligned}
w\Pr(A\&B) &= [w\Pr(A\&B) + x\Pr(A\&-B)][w\Pr(A\&B) \\
&\quad + y\Pr(-A\&B)]/\Pr(E).
\end{aligned}$$

This simplifies to:

$$(*) \quad wz/xy = \Pr(A\&-B)\Pr(-A\&B)/\Pr(A\&B)\Pr(-A\&-B).$$

Note that the left-hand side of (*) describes a relationship among the *efficacies* of the causes, while the right-hand side describes a relationship among the *frequencies* of the causes.

The frequencies of causes often undergo change; so do the efficacies of causes (at least when the causes are incompletely specified). My claim is that when the frequencies and/or efficacies of causes change, the changes are almost never coordinated so as to keep (*) true. If AEB is a conjunctive fork at a given time, this is a mathematical accident that is soon canceled by changes in the system at hand.

So even though Reichenbach's conjectured fork asymmetry is not correct, another can be substituted in its stead (Sober and Barrett 1992). Reichenbach claimed that (i) and (ii) are common patterns displayed by conjunctive forks, whereas (iii) is impossible. I have argued that (ii) and (iii) are both possible – but are very rare – and that (i) requires hedging only because of quantum mechanical phenomena.

I now want to apply this general schema to the Mendelian case discussed before. Reichenbach's exposition and the result just described assume that the events in the fork are dichotomous. But in the simplest of Mendelian cases, this isn't so. The parental pair has six possible states and each offspring has three. Yet it is clear enough how to generalize Reichenbach's idea to cover this situation.

I'll begin with the forward fork formed by a common cause and its two effects. The two parents are heterozygotes and we wish to describe the probabilistic relationship between them and their two offspring, whom we'll assume are heterozygotes as well. The parental pair is the cause and the two offspring are the effects.

The conjunctive fork idea requires that the different possible states that the parental pair might occupy each render the offspring's genotypes independent of each other. Focusing on the case in which both parents are heterozygotes, this means that:

$$\text{(FOR)} \quad \Pr(\text{Off}_1 = \text{Aa} \ \& \ \text{Off}_2 = \text{Aa} / \text{Par} = \text{Aa} \& \text{Aa}) = \Pr(\text{Off}_1 = \text{Aa} / \text{Par} = \text{Aa} \& \text{Aa}) \Pr(\text{Off}_2 = \text{Aa} / \text{Par} = \text{Aa} \& \text{Aa}).$$

It is quite clear that (FOR) is a standard assumption in models of the Mendelian process.

Let us turn now to the backward fork formed by a common effect and its two causes. We wish to determine under what circumstances the genotypes of the offspring form a conjunctive fork with the genotypes of the two parents (P_1 and P_2). For simplicity (but without loss of generality), I'll focus on the case of a single heterozygous offspring whose two parents are heterozygotes as well.

If the fork is to be conjunctive, the different possible states of the offspring must each render the parental genotypes independent of each other. That is, the fork will be conjunctive only if:

$$\text{(BACK)} \quad \Pr(P_1 = \text{Aa} \ \& \ P_2 = \text{Aa} / \text{Off} = \text{Aa}) = \Pr(P_1 = \text{Aa} / \text{Off} = \text{Aa}) \Pr(P_2 = \text{Aa} / \text{Off} = \text{Aa}).$$

I prove in the Appendix that (BACK) is true if and only if:

$$(**) \quad \Pr(P_2 = \text{Aa} / P_1 = \text{Aa}) / \Pr(P_2 = \text{Aa}) = 1 / [2\Pr(\text{Off} = \text{Aa})].$$

The left side of (**) describes the degree of assortative mating that occurs among parents; it tells you how much the choice of mates departs from randomness. The right side describes the frequency of heterozygotes among the offspring. Notice that for any degree of assortative mating, there is exactly one point-value for the offspring frequency of heterozygotes that satisfies the requirements for a conjunctive fork.

If like always mates with like, then the assortative mating is said to be perfect and heterozygosity declines to zero. If the positive assortative mating falls short of this absolute degree, then heterozygosity declines

until it reaches an equilibrium value (Crow and Kimura 1970, p. 144). And of course other evolutionary forces besides pattern of mating can modify the frequency of heterozygotes.

If the assortative mating is perfect, (**) reduces to

$$2 = \Pr(P_2 = Aa)/\Pr(\text{Off} = Aa)$$

and indeed it is a property of this system of mating that heterozygosity is halved in every generation. So the requirement provided by (**) is satisfied in this case. However, the conjunctive fork idea imposes other demands. In particular, the other possible offspring genotypes must render the parental genotypes independent of each other:

$$\Pr(P_1 = Aa \ \& \ P_2 = Aa/\text{Off} = AA) = \Pr(P_1 = Aa/\text{Off} = AA)\Pr(P_2 = Aa/\text{Off} = AA).$$

This will not be true when there is perfect positive assortative mating.

It is not impossible for a population experiencing positive (though imperfect) assortative mating to satisfy (**) over the long haul. What is required is that the intensity of positive association between mates should adjust itself so as to satisfy a criterion specified in terms of the offspring frequency of heterozygotes. The rarer heterozygotes become, the more monomaniacally must like mate with like (all this in accordance with an exact *quantitative* formula). Rube Goldberg devices for achieving this coordination are not ruled out *a priori*, but it is not surprising that no living system happens to possess one.¹⁹

The above argument about the backward fork focuses on the case in which two parents have a single offspring. What is the effect of increasing the number of offspring? If the parental pair has an infinite number of offspring, the likelihood terms (the probability of offspring genotypes conditional on the genotypes of the two parents) approach 1 or 0, and the fork becomes conjunctive degeneratively. But short of this limiting case, the conclusion that conjunctive forks are hard to come by remains in place.

Correlation between sibs is the inevitable outcome of Mendelian reproduction (in a segregating population). Correlation between parents is not inevitable, but is the familiar pattern that is called assortative mating. When sib genotypes are correlated, it is easy to explain why this is so by invoking a model that says that each parental pair forms a conjunctive fork with the children it produces. A model that says that a child forms a conjunctive fork with its two parents allows one to

deduce the pattern of association found between the parents. One reason for claiming that this model is not an explanation is to insist that later events don't explain events that happen earlier. But, in addition, it is worth noting that the purely probabilistic relationships invoked in such a model are rarely if ever satisfied in nature.

5. CONCLUDING REMARKS

The concept of *law* naturally carries with it the idea of time translational invariance. This requirement, by itself, induces no temporal orientation in the laws we posit. The same can be said of the idea of *probability*. That a law should be stochastic, and that it should say of certain systems that their expected state changes with time, also induces no temporal asymmetry. But these two ideas, each of them innocent when taken singly, require that temporal symmetry be broken when they are taken together.

It is a separate question which way the symmetry is broken. If probabilistic laws were backward-directed about as often as they are forward-directed, we might not expect there to be a single underlying explanation for why symmetry is broken in the way it is. But if most (or all) temporally oriented laws are forward-directed, the suspicion arises that there is a single explanation of this fact. I have floated the idea that the source of the bias in favor of forward-directed laws is to be found in an asymmetry concerning causality. We can talk of the probability of an effect, given a cause, and also of the probability of a cause, given an effect. With incomplete descriptions of both cause and effect, neither of these conditional probabilities can be expected to be time translationally invariant. However, when the causal facts are completely circumscribed, it is plausible to maintain that the forward-directed conditional probability is time translationally invariant. The same cannot be said for the backward-directed probability when the description of the effects is rendered complete. The reason for this, I suggest, is that (completely specified) common causes often form conjunctive forks with their joint effects, but (completely specified) common effects rarely form conjunctive forks with their joint causes. This may go some way to explaining why temporally oriented laws are forward-directed.

At the beginning of this paper, I raised the question of whether the temporal asymmetry described here is an artifact of our interests or an objective feature of the world. Different aspects of my argument answer

this question in different ways. The proof with which I began owes nothing to our interests; it is a mathematical fact that systems of the kind specified cannot have both forward-oriented and backward-oriented laws, if laws must be time translationally invariant. I then argued that science opts for forward-directed laws because conjunctive forks open to the future are common while forks open to the past are rare. This difference between forks derives, in part, from the fact that the efficacies of causes are not tightly coupled with their frequencies, and again I take this relative independence to be an objective feature of the world. But what of our choice of *descriptors*? What entitles us to carve up the world into efficacies and frequencies that are not tightly coupled? Could not a 'gruification' of familiar terminology produce a vocabulary in which efficacies and frequencies are bound together, thus ensuring that conjunctive forks open to the past are as common as dirt? Having pursued the question of objectivity back this far, I will not attempt to pursue it farther. For now my conclusion will be a conditional one: *given* the descriptors we use of the probabilities and efficacies of causes, it is an objective feature of the world that temporally oriented laws are oriented towards the future.

APPENDIX

I want to prove that

$$\text{(BACK)} \quad \Pr(P_1 = Aa \ \& \ P_2 = Aa / \text{Off} = Aa) = \Pr(P_1 = Aa / \text{Off} = Aa) \Pr(P_2 = Aa / \text{Off} = Aa)$$

is true if and only if

$$\text{(**)} \quad \Pr(P_2 = Aa / P_1 = Aa) / \Pr(P_2 = Aa) = 1 / [2\Pr(\text{Off} = Aa)].$$

Bayes' theorem allows (BACK) to be rewritten as:

$$1/2[\Pr(P_1 = Aa \ \& \ P_2 = Aa)] / \Pr(\text{Off} = Aa) = [\Pr(\text{Off} = Aa / P_1 = Aa) \Pr(P_1 = Aa) / \Pr(\text{Off} = Aa)]^2.$$

Note that:

$$\begin{aligned} & \Pr(\text{Off} = Aa / P_1 = Aa) \\ &= \Pr(\text{Off} = Aa / P_2 = AA \ \& \ P_1 = Aa) \\ & \quad \times \Pr(P_2 = AA / P_1 = Aa) \\ & \quad + \Pr(\text{Off} = Aa / P_2 = Aa \ \& \ P_1 = Aa) \end{aligned}$$

$$\begin{aligned} & \times \Pr(P_2 = Aa/P_1 = Aa) \\ & + \Pr(\text{Off} = Aa/P_2 = aa \ \& \ P_1 = Aa) \\ & \times \Pr(P_2 = aa/P_1 = Aa) = 1/2. \end{aligned}$$

This allows (BACK) to be further simplified to

$$\frac{1/2[\Pr(P_1 = Aa \ \& \ P_2 = Aa)\Pr(\text{Off} = Aa)]}{[\Pr(P_1 = Aa)\Pr(P_1 = Aa)]} = 1/4,$$

from which (**) follows.

NOTES

¹ My thanks to Martin Barrett, James Crow, Carter Denniston, Ellery Eells, Malcolm Forster, Richard Lewontin, and the anonymous referees of this journal for useful comments.

² Or neither. But I'll ignore laws of simultaneous compossibility.

³ One can infer the later state of the system using this conditional probability if the conditional probability is high or low, once one has observed that the conditioning proposition is true. And even if the probability is middling, the conditional probability plus the observation allow one to assign a probability to the conditioned proposition.

⁴ The idea of a law's subserving one sort of inference but not the other will be clarified presently.

⁵ Nothing very demanding is implied by talk of 'laws' here. Issues about the modal status of laws don't matter, for example.

⁶ A forward-directed law, thus defined, can permit a backward inference based on likelihood, not probability. Suppose one observes that the system is presently in state *s*; this supports the hypothesis that it was earlier in state *i* better than it supports the hypothesis that it was earlier in state *j*, if $\Pr(\text{system is now in state } s/\text{system was earlier in state } i) > \Pr(\text{system is now in state } s/\text{system was earlier in state } j)$.

⁷ In the theory of Markov processes, such laws are said to be *time stationary*.

⁸ By 'expected state', I mean the state's *mathematical expectation*. In general, the mathematical expectation of a quantity is not the value we would expect the quantity to have. Consider a population in which half the parents have one child and half have two. The expected number of offspring of a parent drawn at random from this population is 1.5, although we would not expect any parent to have that number of children.

⁹ Earman (1986) describes a number of writers who hold to this principle. He does not dissent from it. If laws describe causal variables, then the principle embodies the idea that the date of an event is not a causal variable. I do not claim that this idea is an *a priori* constraint on our concept of law, plausible though it is in scientific practice.

¹⁰ Kemeny and Snell (1960, pp. 26, 105) point out that if a forward process is a Markov chain, the backward process will not generally be a Markov chain. They note that symmetry is restored if the process begins at equilibrium.

¹¹ See Berg (1983) for details. The same point holds for random genetic drift. The initial gene frequency is the expected value throughout the population's evolution; yet, the

probability of the gene's going to an absorbing state (0% or 100%) increases with time. See Crow and Kimura (1970).

¹² To see what is involved in this claim, consider a random walk on an open line, with the marker beginning at position 0. After a unit of time elapses, there is a chance of moving one unit to the left and an equal chance of moving one unit to the right. The process may continue for as long as you please. What is the expected location of the marker on the line from $+\infty$ to $-\infty$ after n units of time elapse? The expected location is at location 0. The longer the process continues, the less probable it is that the marker is at location 0. Yet this remains its expected position.

Suppose the first step of this process, during one of its realizations, has the marker move one step to the left. Its expected position thereafter, conditional on this fact, is +1. But earlier, its expected position was 0. Isn't this a case in which the expected position has changed?

No – not in the sense intended. Conditional on the position of the marker at any time (for example, the beginning of the process), we can ask what the expected position is after n moves. The point is that the answer to this question does not depend on n . It is in this sense that the expected position is the same throughout the process.

¹³ In the language of the theory of Markov processes, the above result applies to systems that are *ergodic*: for each pair of states i and j that the system might occupy, there is a nonzero probability that a system beginning in state i will remain there and a nonzero probability that it will go into state j . I believe that the result also applies to Markov chains that are nonergodic, but *decomposable*; this is the case in which the state space can be divided into regions, between which there is no chance of passing, but within which each state is accessible to every other.

¹⁴ I suggest that dispositional concepts like *solubility* should be understood probabilistically. If indeterminism were to imply that a lump of sugar emersed in water under normal conditions has a tiny but nonzero probability of *not* dissolving, I don't think we would conclude that sugar is not water soluble (Sober 1984).

¹⁵ See Eells (1991) for discussion of this point.

¹⁶ I am here ignoring the phenomenon of meiotic drive; that would complicate the example, but would not affect the main point.

¹⁷ The probability of a parental pair is itself a result of two factors: the probabilities of genotypes within each sex and the system of mating.

¹⁸ The conjunctive fork idea also requires that $\Pr(A\&B/-E) = \Pr(A/-E)\Pr(B/-E)$, but the argument would not be affected by considering this condition separately.

¹⁹ Although there is dispute among population geneticists about how assortative mating ought to be represented mathematically, one standard model postulates that a fraction r of each genotype pairs up with like individuals and the remainder $(1 - r)$ mates at random (Crow and Kimura 1970, p. 144); r is here said to be the coefficient of assortative mating. Under this arrangement

$$\Pr(P_1 = Aa \ \& \ P_2 = Aa) = r\Pr(P_1 = Aa) + (1 - r)\Pr(P_1 = Aa)^2.$$

If we let $p = \Pr(P_1 = Aa)$ be the population frequency of Aa in the current generation and p' be the frequency in the next generation, then the condition for a conjunctive fork (***) becomes:

$$r + (1 - r)p = p/2p'$$

The frequency of heterozygosity in the next generation can be expressed as a function of the heterozygosity in the previous generation

$$p' = rp/2 + (1 - r)p_0,$$

where p_0 is the frequency of heterozygosity at the beginning of the process. This allows (***) to be restated as:

$$r + (1 - r)p = p/2p' = p/[rp + 2(1 - r)p_0].$$

As in the simpler treatment provided in the text, it is extremely improbable that an evolving population should satisfy this requirement.

REFERENCES

- Beatty, J.: 1981, 'What's Wrong with the Received View of Evolutionary Theory?', *PSA 1980*, Vol. 2, Philosophy of Science Association, East Lansing, MI, pp. 397-426.
- Berg, H.: 1983, *Random Walks in Biology*, Princeton University Press, Princeton.
- Crow, J. and M. Kimura: 1970, *Introduction to Population Genetics Theory*, Burgess Publishing Co., Minneapolis.
- Earman, J.: 1986, *A Primer on Determinism*, Reidel, Dordrecht.
- Eells, E.: 1991, *Probabilistic Causality*, Cambridge University Press, Cambridge.
- Kemeny, J. and J. Snell: 1960, *Finite Markov Processes*, Van Nostrand, Princeton.
- Kimura, M.: 1983, *The Neutral Theory of Molecular Evolution*, Cambridge University Press, Cambridge.
- Reichenbach, H.: 1956, *The Direction of Time*, University of California Press, Berkeley.
- Sober, E.: 1984, *The Nature of Selection: Evolutionary Theory in Philosophical Focus*, MIT Press, Cambridge.
- Sober, E.: 1988, *Reconstructing the Past: Parsimony, Evolution, and Inference*, MIT Press, Cambridge.
- Sober, E. and M. Barrett: 1992, 'Conjunctive Forks and Temporally Asymmetric Inference', *Australasian Journal of Philosophy* **70**, 1-23.
- Van Fraassen, B.: 1982, 'The Charybdis of Realism: Epistemological Implications of Bell's Inequality', *Synthese* **52**, 25-38.

Department of Philosophy
 University of Wisconsin
 5185 Helen C. White Hall
 Madison, WI 53706
 U.S.A.