CHAPTER 12

The Arrows of Time

> Einen Weiser seh ich stehen
> Unverrückt vor meinem Blick;
> Eine Straße muß ich gehen,
> Die noch keiner ging zurück.
>
> (I see a signpost standing immovable before my gaze;
> I must go down a road by which no-one has come back.)
>
> Wilhelm Müller, 'Der Wegweiser'

The Hidden Signpost

Could time run backwards? When we try to visualize this, we imagine something like a film being shown in reverse. But to imagine this is to imagine being *outside* the process we are observing. As an observer, we are not ourselves in the film (even if we are watching a film of ourselves, for the person we are watching on the screen is not, whilst been filmed, watching the film). So there is no conflict in thinking of the order in which we view the film as the reverse of the order in which the events I am watching were originally filmed. But, in imagining seeing *time* running backwards, we cannot absent ourselves in this way. We are not outside the time we are imagining. So we are not at liberty to see things in a different order from the order in which they are happening. Visualization, then, is no help. We have to think in somewhat more abstract terms when we entertain the idea of time going backwards. The very idea invites paradox. For if 'time runs backwards' means 'events are occurring in the opposite order from the order in which they happen', it is simply contradictory. And 'later events happen before the earlier ones' is no better. Even the time-traveller is not putting time into reverse. It is just that certain events in the time-traveller's life that would normally proceed in one direction are proceeding, for a time, in the reverse direction. What, then, if *all* processes went into reverse? Consider the way in which natural processes normally proceed: drops of rain fall from clouds, soaking the earth, rivers run downhill, the spring of a clockwork mechanism unwinds, heat moves from a hotter body to a cooler. Now imagine all these processes going in reverse: water travelling upwards from the ground to the sky, rivers running uphill, springs winding themselves up, heat moving from cooler bodies. That, surely, is not an incoherent description. But would this be a case of time running backwards? The very idea of all things going into reverse implies that time (or at least its direction) is independent of the processes within it, and indeed carries on normally while everything *else* goes into reverse.

In any case, although we may detect no contradiction in the idea of all natural processes going backwards, we can still ask whether this is really possible. Is there not some hidden constraint, some underlying directedness, within time itself, that makes processes within time go one way rather than another? This idea is captured by the talk of the *direction,* or *arrow,* of time. That time must have a direction seems required by our experience of it: our youth, once past, never returns. Yet, despite the all-pervasive sense we have that time is passing, and passing in one direction only, one of the greatest mysteries of time is what explains this direction. Indeed, it is no straightforward task to give a precise account of what it is that we are trying to explain. What does it actually *mean* to say that time has a direction? Once again, as with the passage of time, our understanding may be assisted by examining the contrast with space. Clearly, we are, within certain limits, free to go where we want to in space: we can go backwards, forwards, up and down, sideways in either direction. And any constraints there may be on our movement have nothing (we suppose) to do with the nature of space itself, but rather with the behaviour and influence of objects within space—we cannot, without creating special circumstances, escape the force of gravity, for example. (Of course, if space has an edge, or is otherwise finite, then that represents a restriction on possible movements, and in general other features of the space, such as whether it is curved, may influence objects, as we saw in Chapter 4. But this does not undermine the contrast we are trying to draw here between space and time.) In short, there may be directions *in* space, but no direction *of* space. In contrast, we are not similarly free to go where we want in time. I cannot, time travel aside, revisit yesterday. So we might expect to see this difference show up when we compare temporal and spatial relations.

Consider, then, the relation 'earlier than'. This relation orders things in time, and does so, we might suppose, in a way that reflects the direction of time. How does it reflect this? Well, as a first attempt at explanation, we could point to the logical properties of the relation: in particular, it is asymmetric: necessarily, if A is earlier than B, B is not earlier than A. So perhaps the direction of time just consists in the asymmetry of the earlier than relation. But this cannot be right. For the relation *to the north of* is similarly asymmetric, yet we do not think that, because of this, space has a direction. So something more has to be said. Consider another asymmetric spatial relation, 'to the left of'. This requires a point of reference for its application. The desk may be to the left of the window from where I am sitting, but if I were to turn around, it would be to the right. So 'to the left of' really needs to be considered as short for 'is, from such-and-such a position, to the left of'. 'To the north of' and 'above' similarly presuppose a point of reference, but, unlike with 'to the left of', this is a fixed reference point. 'To the north of' refers implicitly to the North Pole, and so means, in effect, 'nearer the North Pole than'. Similarly, 'above', in its usual application, means 'nearer the surface of the Earth than'. In contrast, 'earlier than' does not require a reference point in time for its application. If it is true at 11 o'clock that boarding the train this morning was earlier than my arriving at work, then it is equally true at 4 o'clock. Whether one event is earlier than another does not change over time. We can express the contrast by saying that 'earlier than' is *intrinsically asymmetric*, intrinsic because its holding between two items does not depend on the existence of any other item. But, although this contrast is significant, it still does not quite capture the distinction we are trying to draw between time and space. Consider another asymmetric spatial relation: *inside*. Whether one object is inside another does not depend, it seems, on any spatial reference point. Moreover, the relation can be used to order a group of suitably related objects, such as the collection of increasingly diminutive figures inside a Russian doll. However, the 'inside' relation is not a *pervasive* one: it cannot be used to order all objects in space, for many objects are not inside any other object. 'Earlier than', in contrast, *is* pervasive: by means of it we can order all events in time. So being in time at all requires that one stands in the 'earlier than' relation to other events or times. Now there are spatial relations that are both intrinsically asymmetric and pervasive, for example 'larger than'. However, it is not a *positional* relation: it does not tell one anything about where in space each object is vis-à-vis the other. 'Earlier than', in contrast, is positional: it tells one where events are in time vis-à-vis other events.

In saying that the 'earlier than' relation is intrinsically asymmetric, pervasive *and* positional, we undoubtedly capture something that distinguishes time from space. But, still, direction appears to point to something else. A series can be ordered without being directed. Consider a series of integers, ordered by the relation 'greater than':

114 is greater than 113 is greater than 112 is greater than 111 . . .

We have here an ordered series. But we are not obliged to read the series in one direction rather than another. We can start with the smallest and work up to the largest, or vice versa. Nothing in the numbers themselves picks out a preferred direction. There is no hidden signpost, as it were. But time is both ordered *and* directed, or so we suppose. Certainly the passage of time indicates a preferred direction in which events occur: the present moves *from* earlier events *to* later events, never the other way around. It is an interesting and important question whether the idea of a direction of time has to be tied to the idea of the passage of time, a question we shall come back to later in the chapter.

Three Arrows, and Why Things Fall Apart

We have made some progress in trying to capture what is distinctive about time, and it is likely that this is intimately related to the idea of the direction of time, but we have not yet arrived at an analysis of this elusive direction. So let us try a different approach. What draws our attention to time's arrow is the behaviour of processes in time. Some processes, it seems, are *temporally asymmetric*: some stages in the process characteristically come before others. These different processes, in effect, each constitute a different arrow of time. Here are the most significant:

⇒ the *thermodynamic arrow:* the direction from order to disorder;
⇒ the *psychological arrow:* the direction from perceptions of events to the memories of those events;
⇒ the *causal arrow:* the direction from cause to effect.

In each of these cases, the direction of the process coincides with the direction from earlier to later. Thus disorder tends to increase (in a sense to be explained shortly), memories always come after, never before, the perceptions of which they are the memories, and causes always precede their effects (or do they? we take up this issue later). This raises some tricky questions:

- *Why does each arrow coincide with the direction from earlier to later?*
- *Why do all three arrows point in the same direction? Why, for example, should the direction from experiences to memories coincide with the direction from order to disorder?*
- *Is one arrow more fundamental than the others?*
- *Can we actually define the direction of time in terms of one of the other arrows?*

Answering some of these questions helps us answer others. For example, providing an answer to the first also answers the second. For another example, suppose, in answer to the fourth question, that we can indeed define the direction of time in terms of one of the other arrows. That would then answer the first question in relation to that arrow. For if the direction of time is to be explained in terms of, e.g. the psychological arrow, then it follows as a logical consequence that perceptions precede memories. It would also be grounds for asserting that arrow to be more fundamental than the other two, and a test of that assertion would be whether, by taking that arrow to be more fundamental, we can explain why the three arrows coincide with each other.

We will now consider each arrow in turn, bearing these questions in mind.

Take the thermodynamic arrow first. According to a familiar and informal statement of the Second Law of Thermodynamics, heat tends to move from a hotter body to a cooler one, thus warming it up. This means that there is a tendency for heat to become ever more widely and evenly distributed. What has this got to do with order? Well, consider a cup of freshly poured tea steaming away on the table. This is a relatively ordered situation, because the energy present in it is concentrated in a relatively small area. There is energy in the form of heat in the small area occupied by the tea. There is also energy in the forces holding the cup together. Finally, there is potential energy in the cup as a result of its being some distance above the ground. But now a careless hand accidentally sweeps the cup off the table. The cup descends, dashing itself to pieces on the floor, and spilling the tea around the room. The potential energy of the cup was transformed into kinetic energy as it fell, and then into sound and heat as it smashed on the floor. The heat from the tea is now rapidly dispersing itself into the air. Thus energy is more evenly distributed. This little vignette, often used to illustrate the Second Law, is but one instance of a wider phenomenon, also instanced by stones dropping into ponds, causing ripples to spread out to the edges, or by sunlight warming the bricks of a house, and by the tendency of things—buildings, books, and bicycles—to fall apart. In these cases, energy is becoming more widely dispersed, and this is what is meant, in this context, by increasing disorder, and it is referred to as an increase in *entropy*.

A word or two about the definition of entropy and some other relevant terms. Take a particular item or set of items one happens, as a scientist, to be interested in. It might be a volume of gas, or of liquid, or a solid lump of matter, or an array of objects in interaction with each other. Let us designate this item or set of items by the abstract name of 'system'. Now suppose a definite amount of heat to be introduced into the system: the gas is briefly warmed by the sun, or the liquid by the application of a flame, etc. Then the *change of entropy* of the system as a result of this introduction of heat is given by dividing the amount of heat introduced by the temperature of the system prior to its introduction. Where heat is lost by the system, the entropy change is negative. Change in the entropy of a system, however, does not require the introduction or loss of heat from the system; it can instead result from the transfer of heat from one part of the system to another. Suppose we place a very hot object next to a very cold object, and consider the two objects together as a single system. Heat flows from the hot body to the colder, so both objects experience a change in their entropy: the hot body experiences a negative change, the cold body a positive one.

The overall heat within the system is unchanged. However, the positive entropy change in the cold body is *greater* than the negative entropy change in the hot body (since in the former case one is dividing the amount of heat by a lower temperature), so the two entropy changes do not cancel each other out. The entropy of the system as a whole goes up. This gives us a more formal statement of the Second Law, as follows:

> *Where any change occurs in an isolated system, entropy either stays constant or increases.*

An 'isolated' system is one from which no energy can leave and into which no energy can be introduced. But what *is* entropy? One way of conceiving it is as follows: the entropy of a system is the degree to which the thermal energy (heat) in the system is so ordered that it is available for conversion into other forms of energy. The less available the energy is, the greater the entropy. So why is it that entropy tends to increase? Why do things fall apart?

One influential way of addressing this question is in terms of *probability*. Suppose you have, to continue with tea imagery, a large tea tray, and in one corner of this tray you have placed a number of sugar lumps. At the moment, they are neatly organized in two rows, one on top of the other, each row containing five lumps. You now begin to shake the tea tray vigorously and then set it down on the table. Assuming that the tray is flat and the table top not sloping, what are you likely to see? Not the same two neat rows of sugar lumps in one corner of the tray, that is for sure. Nor will they have gathered themselves neatly in the middle. Almost certainly, the lumps will have distributed themselves randomly around the tray. However many times you try the experiment, the result is likely to be the same. In other words, it is extremely *improbable* that the sugar should end up in an orderly arrangement, and concentrated in a small area. And this is because, whereas there are only a few such orderly arrangements, there are many more random or disorderly arrangements. So, given that the sugar is moving around randomly, the chances of its ending up in an orderly arrangement will be very small, reflecting the very small proportion of orderly arrangements amongst the total number of possible permutations. The reason energy tends to become more randomly distributed is the same: the permutations of random distributions outnumber the orderly permutations, so the probability is vastly in favour of a more random distribution.

But actually these probabilistic considerations do not really help us to see why disorder *increases* over time. A highly ordered state is indeed improbable, for the reasons given above, but it is equally improbable whether we are talking of an earlier time or a later time. Given an ordered state at a particular time, say the temperature distribution in my glass of gin and tonic (with ice and lemon) at 12 noon today, pure probability gives us no reason to suppose that the state of that system will be more disordered at *later* times than at *earlier* times. No more reason, in fact, than there is to suppose that things will be more disordered to the east of the gin and tonic than to the west. Of course, processes do proceed from earlier to later, and not vice versa, but we cannot just assume this without building in the very thing we are trying to explain, namely the direction of time.

Here is a quite different approach. Let us try to define the arrow of time in terms of the thermodynamic arrow. In other words, what makes it true to say that one state of the universe is before another state of the universe *just is* that the first is more ordered than the second. So, *by definition,* disorder tends to increase. This makes the thermodynamic arrow more fundamental than the others, because it, and it alone, defines the direction of time. More formally, what we might call the *thermodynamic analysis* of time order goes as follows:

> *The thermodynamic analysis of time order*: event A is earlier than event B if and only if the universe is, when B occurs, in a higher state of entropy than it is when A occurs.

We now have two explanations of why disorder increases on the table. The first of them, as we saw, is not a complete explanation, but we do not complete it by combining it, as we might be tempted to do, with the second explanation. We must *not* say 'Time order just is the direction from order to disorder, so where there is no thermodynamic arrow, there is no arrow of time either. And we know why disorder tends to increase: disorder is just a more probable state.' The first part of this statement renders the second redundant. If we have defined time order in terms of the thermodynamic arrow, then we do not need to hunt around for some *other* explanation of why entropy increases over time. It would be like wondering why everyone describing themselves as 'married' on a census form appears, on inspection, to have a spouse. Worse, if we say it is merely *probable* (even if overwhelmingly probable) that entropy increases, then we leave the door open for an occasional decrease in entropy. But in *defining* temporal direction in terms of entropy, we do not leave this door open at all: it becomes a necessary (indeed, trivial) truth that entropy increases. So probabilistic explanations of the thermodynamic arrow, one is tempted to say, diminish its philosophical significance, even if they increase its physical plausibility. Since we are concerned with the philosophy of time, we will confine ourselves in the remainder of this section to an exploration of the thermodynamic analysis of time order.

One consequence of the thermodynamic analysis is that a universe in which things were always as disordered as they could be would exhibit no direction of time at all, because there would be no (significant) changes in entropy in such a universe. This is only a problem, however, if we think that time must exhibit a direction in order to exist at all (we return to this issue in the final section).

The analysis implies that the thermodynamic arrow is the most fundamental of the three. But the test of this will be whether we can explain, by appeal to the fundamental nature of the thermodynamic arrow, why the three arrows point the same way. There is some hope that the direction of the psychological arrow can be explained, along the following lines. The thermodynamic arrow points to a *global,* rather than merely local, change. That is, it is the overall entropy in the universe as a whole that is increasing. Locally, there can be decreases in entropy, as when, for example, a gas is compressed to the extent that it liquifies, or a boulder is rolled to the top of a mountain. But this local decrease in entropy is always at the expense of a global *increase* in entropy, because energy is required to induce local decreases in disorder, and this energy is ultimately dissipated in the form of heat. Now, storing information in the form of a memory is one such local increase in order which must be made up for by a global increase in entropy. So the direction from experiences to memories of those experiences will coincide with the direction from order to disorder. (We assume that it is just a matter of definition rather than accidental fact that the formation of memories constitutes a local increase in order.)

But what of the causal arrow? Can this be explained in thermodynamic terms? Well, suppose increases in entropy are always a result of causal processes. Certainly all our examples of increasing entropy were causal processes: the falling of a cup onto the floor, causing it to smash; the stone thrown into a pond, causing ripples to spread out to the edges; the warming of the bricks by the sun; the formation of memories, wear and tear causing things to fall apart. Causing something to happen involves the transfer of energy, and this transfer results in energy being more widely or randomly distributed. It is causation that brings with it the increase in entropy, so the causal arrow must point in the same direction as the thermodynamic arrow. Is that a satisfactory explanation? Well, we might think of possible cases where causation actually resulted in a *decrease* in entropy. Imagine, for example, God intervening at some point to bring about a more ordered universe. Now, if the thermodynamic analysis is correct, the direction of entropy increase is also the direction of time, so the case we have just imagined would have to be a case of backwards causation: God intervening to bring about a more ordered universe at an *earlier* time. But that, of course, is not how we imagined it: we imagined God bringing about a later state of greater order. So if what we imagined is a genuine possibility, the thermodynamic analysis cannot be correct. Now, defenders of the thermodynamic analysis will simply say that what we imagined is not a genuine possibility: not even God can bring about a decrease in entropy. So, unless we are willing to countenance backwards causation, the thermodynamic and causal arrows will always point in the same direction.

This illustrates, however, the implausibility of the thermodynamic analysis. As a statement of probability, the Second Law of Thermodynamics may well be true: it is overwhelmingly likely that entropy will increase. This leaves the door open for occasional, isolated instances of entropy decrease. But the thermodynamic analysis firmly shuts this door, and makes it a *necessary truth* that entropy will only increase. And this just seems too strong. How can we just rule out even the possibility of a once-in-a-blue-moon decrease in entropy?

Let us move on to the psychological arrow.

The Mind's Past

When the seventeenth-century English philosopher John Locke was looking for the thread that linked our past selves with our present self, the unifying feature that explained our persistence as persons through time, he alighted on memory. There is some plausibility in this, for total loss of memory robs us of our sense of who we are. Equally, loss of our ability to form new memories drastically diminishes our status as agents. In *The Man Who Mistook His Wife for a Hat* the neurologist Oliver Sacks describes the case of a patient he calls 'Jimmie', whom he first met in 1975. Jimmie was suffering from advanced Korsakov's syndrome, unable to remember anything for more than a few seconds. Moreover, although he could recall in detail events up to 1945, the 30 years afterwards were almost a complete blank. As a result, he thought of himself as he was in 1945: a 19-year-old, wondering whether to stay in the American navy or go to college. Sacks comments:

'He is as it were,' I wrote in my notes, 'isolated in a single moment of being, with a moat or lacuna of forgetting all around him. . . . he is a man without a past (or future), stuck in a constantly changing, meaningless moment.'

. . . 'I may venture to affirm,' Hume wrote, 'that we are nothing but a bundle or collection of different sensations, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement.' In some sense [Jimmie] had been reduced to a 'Humean' being. (Sacks 1985, 28)

This apparently intimate connection between memory and the persisting self depends, however, on an even more intimate connection between memory and time. We remember only the past, never the future. And this is not just the trivial point that we would not *call* it 'memory' if we had experiences of the future; rather the past is revealed to us in a way in which the future is not. Why is this? Why does the psychological arrow point from earlier to later times? As with the thermodynamic arrow, let us see what happens if we try to define temporal precedence in terms of memory. Then we might get something like the following:

The psychological analysis of time order: A is earlier than B if and only if A is the content of a memory at the same time that B is the content of a perception.

That would undoubtedly explain why memories succeed perceptions: they do so by definition. What happened earlier *just is* what someone remembers. But it is a rather surprising definition, for it makes temporal order dependent on minds. If there were no one around to have experiences of things, and subsequently memories of those experiences, there would be no earlier and later. This, for some people, would be enough to dismiss the idea that the psychological arrow is somehow fundamental. 'Surely,' they would argue, 'it is a purely accidental matter whether or not an event is actually experienced by someone? Moreover, experienced events may be caused by events that are not experienced, and it would be absurd to suggest that an experienced event was earlier than the memory of the event, but the causes of that event were not.'

There are two quite distinct issues here. One is whether two events have to be experienced by someone in order for one event to be earlier than the other. The answer to that, surely, is 'no'. The other issue is whether, in order for one event to be earlier than another, *some* event or group of events (but not necessarily that same pair) has to be perceived. Is it possible that temporal precedence does depend somehow on minds? It is far less obvious that the answer to *that* question has to be 'no'.

So let us try to define time's arrow in terms of the psychological arrow in a slightly more subtle way. First, let us introduce the relation of betweenness. We can order a series, for example, a group of people standing in a line, from information about which member is between which two other members. So, on being told that Betty is between Frank and Walter, that Doris is between Harold and Enid, that Harold is between Walter and Doris, and that Walter is between Betty and Harold, we can construct the line as follows:

Frank—Betty—Walter—Harold—Doris—Enid

In the words of the first section, this is an ordered, but not yet directed series. We can give it a direction by specifying that Enid is first in the queue, but betweenness alone will not bestow directedness on a series. Now, by analogy, we can talk of the relation of temporal betweenness, and order times, or events in time, by means of this relation. Consider now the following two facts:

(a) The letter arrived between the kettle boiling and Alf setting off to work.

(b) The kettle boiled before the letter arrived, and the letter arrived before Alf set off to work.

Since (b) implies (a), we would naturally assume that (b) is the more fundamental fact, and in general that facts about precedence determine facts about betweenness. But let us for a moment entertain the idea that (a) is the more fundamental fact, by virtue of its holding quite independently of the state of any mind, whereas (b) depends upon two mental states: an experience and a memory of that experience.

Pursuing this thought, suppose we take it for granted that, if *e* is an experience of some event, and *m* is the memory of that experience, then *e* is, by virtue of that psychological relationship, earlier than *m*. Now let us consider some quite different, unperceived, event—call it *d*—and let us suppose that *e* occurs between *d* and *m*. Now, although *d* is neither experienced nor remembered, we can deduce that it is earlier than *e* and *m*. Finally, consider another unperceived event, *f*, and suppose that *m* occurs between *e* and *f*.Then we can deduce that *f* is later than *e* and *m*. We can also therefore deduce that *f* is later than *d*. So, provided that some events are experiences, and that they are subsequently remembered, we can discover, of two unexperienced events, whether one precedes the other, given information simply about the relevant betweenness relations. This allows us to define temporal precedence as follows:

> *The modified psychological analysis of time order:* A is earlier than B if and only if there exists an experience *e* and a memory of that experience *m* such that either
> (i) A is simultaneous with *e* and B is simultaneous with *m*, or
> (ii) A is simultaneous with *e* and B is between *e* and *m*, or
> (iii) B is simultaneous with *m* and A is between *e* and *m*, or
> (iv) A and B are between *e* and *m*, and A is between *e* and B, or
> (v) *e* is between A and B, and B is between *e* and *m*, or
> (vi) A is between *e* and B, and B is between *e* and m, or
> (vii) B is between A and *e*, but *m* is not between *e* and B, or
> (viii) A is between *m* and B, but *e* is not between A and *m*.

This could hardly be described as an elegant and economical analysis, but it shows that it is possible for the fact that A precedes B to be mind-dependent, without this implying either that A and B have to be perceived, or that judgements about precedence are somehow subjective, that is merely a matter of opinion. For the analysis to work, we have to recognize that relations of temporal betweenness are entirely mind-independent. It is the mind, however, that gives time its arrow.

But is this analysis viable? Here is a problem. Suppose Monica has an experience, *a*, and subsequently a memory of that experience, *b*. According to the modified psychological analysis, this makes it the case that *a* precedes *b*. Now take two other events (whose nature will remain secret for a few more moments), *c* and *d*, which are so situated in time that *c* is between *a* and *b*, and *b* is between *c* and *d*. According to the psychological analysis, *c* must precede *d*. Now let us reveal the nature of *c* and *d*: one of them is an experience had by Norman, and the other is Norman's memory of that experience. But which is the experience and which the memory? Since we already know that *c* precedes *d*, it must be *c* that is the experience and *d* the memory if the psychological analysis is correct. But why should this be so? What is there to prevent *d* from being the experience and *c* the memory? In other words, why shouldn't the order according to Monica be the reverse of the order according to Norman? Norman's mental states are not determined by Monica's mental states: they are quite independent. And if time order is mind-dependent, there is nothing other than minds that determines the order of their mental states. There is nothing, then, to prevent different minds from imposing different orders on the world. But, if we allow this, then we have to say that time order is not just mind-dependent, but actually relative to individual minds. So our modified psychological analysis would become:

> *The relative psychological analysis of time order*: A is earlier than B *for a given person* if and only if that person has an experience *e* and a memory of that experience *m* such that either
> (i) A is simultaneous with *e* and B is simultaneous with *m*, or *etc.* (as before).

But this is simply implausible. Moreover, unless the causal and thermodynamic arrows are similarly relative to individual minds (and surely it makes no sense to suppose that the second of these, at least, is so relative), then there is no longer any correspondence between the psychological arrow and the other arrows.

It is time, then, to consider the causal arrow's claim to be the most fundamental.

The Seeds of Time

In Act I of Shakespeare's *Macbeth* the King of Scotland's generals, Macbeth and Banquo, returning from battle, encounter three weird sisters upon a storm-blasted heath. The witches greet Macbeth with the news that he has been made Thane of Cawdor, and will at some future time be king. Banquo, seeing his companion 'rapt withal' at this extraordinary prophecy, addresses the sisters:

> If you can look into the seeds of time,
> And say which grain will grow and which will not,
> Speak then to me, who neither beg nor fear
> Your favours nor your hate.

Why does Banquo talk of 'the seeds of time' rather than simply 'what is yet to come'? Well, perhaps for Shakespeare and his contemporaries the idea of the future somehow being available for observation by those with the appropriate gifts would have seemed nonsensical, as the future simply does not exist in any sense. The best one can do is to infer what will happen from its present causes, and it is these that are the seeds of time, the generators of history. Let us see, then, whether we can generate time order from the causal relations between things in time.

Can we define time order in terms of causal order? The simplest version of the causal analysis of time order is as follows:

> *The basic causal analysis of time order*: A is before B if and only if A is among the causes of B.

It will be clear by now how the causal analysis could answer some at least of the other questions we posed about the arrows. The causal analysis makes the causal arrow the most fundamental of the three. Causes precede their effects because precedence has been defined in causal terms. And this explains why experiences precede memories of those experiences: the experiences are the causes of the memories and so, by the causal analysis of time order, must precede them. Can it explain why the causal and thermodynamic arrows (generally, but perhaps not invariably) point in the same direction? Well, the causal analysis cannot do this by itself, but if we grant that causation involves energy transfer, and such transfer tends to increase entropy, then both arrows will generally point in the same direction. And since the direction from earlier to later is defined as the direction from cause to effect, later states of the universe will tend to exhibit greater entropy than earlier states. So taking the causal arrow as the most fundamental would help to explain the temporal direction of the other two arrows. But it could do more than that. It also would explain why we cannot perceive the future. For perceiving something is a way of being causally affected by it. So, by the causal analysis, what one perceives must always occur earlier than one's perception of it. Perceiving the future would be an instance of backwards causation—an effect occurring earlier than its cause—and this the causal analysis rules out. The causal analysis of time order thus promises to be a very powerful tool. But it has some formidable objections to overcome.

Problems arise when we consider what to say of two events that are *not* causally connected. The causal analysis tells us that they cannot be related by the 'earlier than' relation. But this raises two issues. First, to put a rather obvious objection, it is surely possible for non-simultaneous events to be causally unconnected. Surely there could have been events in distant galaxies that have had no effect, and will have no effect, on us, and yet which lie in our past? Second, the basic causal analysis does not tell us whether events unconnected causally are simultaneous or completely unrelated temporally, but any adequate theory of time order would need to be able to distinguish between these quite different possibilities.

The first objection, that an event could be earlier than a causally unrelated event, could be met by replacing an actual causal connection with causal connectibility:

The modal causal analysis of time order: A is before B if and only if it is possible for A to be among the causes of B.

The term 'modal' here indicates that the notion of possibility is being invoked. What matters according to this analysis is not whether there is in fact a causal connection between two events, but whether there *could have been* a causal connection between them. If so, then the event that could have played the role of cause is the earlier. However, the shift from the basic to the modal causal analysis is fraught with danger. It is the appeal to possibility that is problematic, for there is more than one kind of possibility. The weakest kind is that of logical possibility, which we may define as anything that does not involve a contradiction. Thus it is logically possible that I could jump 20 feet in the air unaided, but not logically possible that I could prevent my own conception. More restricted notions of possibility would include reference to certain components of the actual situation. Given the law of gravitation and facts about my physical constitution, for example, it is not possible for me to jump 20 feet in the air. So what kind of possibility is appropriate for the causal analysis? If we choose the widest possible notion, namely logical possibility, then we face contradiction. Irrespective of how any two events, A and B, are related, it is always logically possible for A to be among the causes of B, and also logically possible for B to be among the causes of A (although both possibilities could not be realized together). It would therefore follow, from the modal causal analysis, that A is both earlier and later than B. But it is no good restricting the possibilities as follows: 'A is before B if and only if it is possible, *given their actual temporal relationship,* for A to be among the causes of B', for this would destroy the point of the causal analysis, in that it would make the temporal relation what determines the causal relation, and not the other way around. And if there was some more fundamental relation that determined whether A could be the cause of B, then one might as well refer directly to that relation, and leave out causation as redundant. So bringing in possibility achieves nothing.

Parallel Causes

Before introducing another, more promising, causal analysis, let us take a closer look at the second objection, that the basic causal analysis provides no means of distinguishing between simultaneous events and temporally unconnected events. Perhaps, in certain circumstances, we could use the causal account to distinguish between these two possibilities. Consider the two series of events in Figure 26. The arrows represent causal connections. So, in the first series, A causes B, which causes C, and so on. In the second series, α causes β, which causes γ, and so on. None of the members of the first series, however, is causally connected in any way to any member of the second series. According to the causal analysis, C precedes D, and γ precedes δ, but C neither precedes nor succeeds γ. But we are also able to infer, further, that C, for example, is not simultaneous with γ, for if it were, then B, being earlier than C, would be earlier than γ. But we have just said that B is causally unrelated to γ, and so cannot be earlier than it. According to the basic causal analysis, then, the two series are temporally unrelated.

However, although we have been able to use the basic causal theory to distinguish between simultaneity and temporal unrelatedness in this case, we have had to appeal to purely contingent features of the situation, namely that each series contains more than one member. What if there had been only two events, each causally

$$\mathrm{A} \rightarrow \mathrm{B} \rightarrow \mathrm{C} \rightarrow \mathrm{D} \rightarrow \mathrm{E} \rightarrow \mathrm{F}$$

$$\alpha \rightarrow \beta \rightarrow \gamma \rightarrow \delta \rightarrow \varepsilon \rightarrow \xi$$

Fig. 26. Parallel causal series

unrelated to the other? If it is to be an adequate account, the causal account should tell us whether or not the two events are simultaneous.

Consider a quite different example (Figure 27), one in which the universe contained converging causal series. Here F and ξ jointly cause Z. What is the relationship between F and ξ? It cannot be one of simultaneity, for the reasons given above. If F and ξ were simultaneous, then E would be earlier than ξ, and that is ruled out by the causal analysis, which requires an earlier event to be a cause. So the result is that the structure above is, according to the causal analysis, one in which time itself branches in the direction of the past: two unrelated temporal series converge to form a single series. But this is not a desirable result. Arguably, our universe contains many such branching causal series: series initially unrelated to each other converge at some point to bring about a single course of events, which then leads to diverging series of causal chains which then have nothing to do with each other. We surely do not want to say that this implies that time itself exhibits a branching structure.

All the problems considered so far could be dealt with very simply indeed, by allowing the causal theorist to make explicit appeal to the relation of simultaneity, as in the following variant of the causal analysis:

> *The augmented causal analysis of time order*: A is earlier than B if and only if A is simultaneous with one of the causes of B.

Since every event is simultaneous with itself, A's being a cause of B would satisfy the above requirement for being earlier than B. This

$$\begin{tikzcd}[row sep=small, column sep=small]
A \arrow[r] & B \arrow[r] & C \arrow[r] & D \arrow[r] & E \arrow[r] & F \arrow[dr] & \\
 & & & & & & Z \\
\alpha \arrow[r] & \beta \arrow[r] & \gamma \arrow[r] & \delta \arrow[r] & \varepsilon \arrow[r] & \xi \arrow[ur] &
\end{tikzcd}$$

Fig. 27. Converging causal series

new analysis allows (some) causally unrelated events to be temporally related. It also allows the causal theorist to specify under what conditions two events are temporally unrelated, namely when (i) they are not simultaneous with each other, *and* (ii) neither is the cause of the other. It therefore also allows us to distinguish between branching causal series and branching time. Of course, if 'simultaneous with' is defined as 'neither earlier nor later than', then we are simply smuggling in the very relationship we are trying to define into the analysis. But, as long as we allow for the possibility of multiple time-series or branching time, we will not want to identify simultaneity with the mere absence of the 'earlier than' relation.

Is Time Order Merely Local?

The introduction of simultaneity into the analysis raises a difficulty, however. We have already suggested, in our discussion of the psychological arrow, that temporal betweenness could be independent of temporal priority. In other words, it is possible for things to be temporally separated, without it being the case that some things occur earlier than others. Something else, not intrinsic to time, introduces the asymmetry. So, a proponent of the causal analysis could argue, a world where there were no causal relations could still be a world in which time existed, it is just that there would be no *arrow* of time. Only the existence of causal relations *somewhere* in the universe allows some times to be earlier than others. Now, the move from the basic to the augmented causal analysis was in part a response to the point that one event could be earlier than another even though no causal relation existed between them; all that is required is that they are appropriately related to items that are causally connected. But once this point is conceded, what is to rule out the possibility set out in Figure 28? Here, A is simultaneous with C, and B is simultaneous with D. C is a cause of B and D a cause of A. Since A is simultaneous with a cause of B it follows, from the

$$\begin{array}{ccc} A & \longleftarrow & D \\ C & \longrightarrow & B \end{array}$$

Fig. 28. Causal arrows running in opposite directions

causal analysis, that A is earlier than B. But B is also simultaneous with a cause of A, and so, again by the causal analysis, it also follows that *B is earlier than A*. This looks like a contradiction, but we cannot appeal to the causal analysis to rule out such a case, since it was the application of the causal analysis that enabled us to deduce the apparently inconsistent propositions that A is earlier than B and B is earlier than A. Nor can we say that the causal order of B and C is somehow constrained by the order of A and D, since the causal connection between A and D is quite independent of the connection between B and C. And we cannot appeal to any underlying asymmetry in time itself, forcing the causal arrows to point the same way, without undermining the causal analysis. Without causation, the analysis implies, there would *be* no earlier and later.

One quite tempting approach is simply to concede the result, but deny that it is at all contradictory. Only if we assume that temporal priority is a global relation is there a problem. Perhaps time order is simply *local* to causally connected series. We can allow, then, the existence of causally isolated series where the arrows point in opposite ways. In relation to *this* part of the universe, A is earlier than B; in relation to *that* part, it is B that is earlier than A. As long as they remain isolated, no anomalies will arise.

But what if these oppositely directed causal pairs do not remain isolated? Suppose, for example that the situation as in Figure 29 were to arise. Then, even granted that time order is simply local to a causal series, we still get a contradiction: A is both locally before and locally after B. We also get the contradiction that C is both earlier than, and simultaneous with, A. There are two possible ways out of this difficulty. The first, most radical, way is simply to *give up* *the notion of simultaneity altogether.* Without simultaneity, the problem cannot arise. We could, then, insist that without causation there is no time at all, not even undirected time. For there to be temporal relations of any kind between events, there must be causal connections. So, take a rather simple universe in which there are only a very limited number of events, and suppose that when we have mapped out all the causal connections we end up with the picture shown in Figure 30. Then we can say that both A and B are earlier than D and F, that D is earlier than F, and that C is earlier than E. Neither C nor E, however, stand in any temporal relationship to A, B, D,or F. What of the relationship between A and B? Well, although both are earlier than D, they are not themselves causally connected, and so neither is earlier than the other. In giving up simultaneity, we have moved back from the augmented to the basic causal analysis.

The less radical proposal involves appealing to certain facts about causation, and in particular the following:

> *The Betweenness Rule*: If B is causally between A and C (e.g. an effect of A and a cause of C), then B is also temporally between A and C.

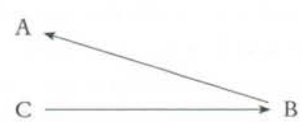

Fig. 29. A causal series changing direction

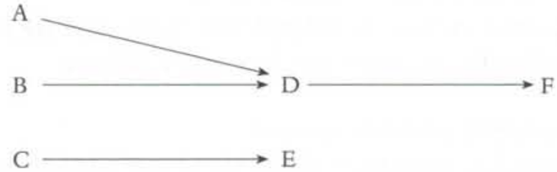

Fig. 30. A causal map

Appealing to this rule allows us to exclude cases where a series contains two causal arrows pointing in opposite directions. For example, the case we considered either, where A is simultaneous with C, C causes B and B causes A, violates the Betweenness Rule. This might at first sight seem a suspicious move. After all, surely we are trying to explain the properties of time in terms of the properties of causation, and not the other way around? But, as we have already indicated, there is no inconsistency in saying both that temporal betweenness is independent of causality, and that temporal priority is not. The Betweenness Rule does not, however, outlaw cases of different and isolated causal series where the causal arrows point in opposite directions. Unless we can find some means of doing so without undermining the causal analysis, it looks as if we have to accept that a consequence of that analysis is that time order is merely local.

Are Causes Simultaneous with Their Effects?

The causal analysis has not yet passed all the hurdles it needs to in order to be acceptable, however. It is a fundamental requirement of any causal analysis that causes precede their effects. But in some cases, cause and effect seem to be *simultaneous*: a crown resting on a cushion and causing an indentation in it, for example, or the passage of an electric current through a wire causing a magnetic field around the wire, or the engine of a train causing the carriages to move at the same time. If these genuinely are cases of simultaneous causation, then the causal analysis is doomed.

Perhaps these are not just isolated cases. There is an argument to the effect that *all* causation is simultaneous, as follows:

The simultaneity of causation argument

(1) Causes determine their effects, i.e. are such that, if the cause occurs it is impossible for the effect not to occur.

(2) If there were a temporal gap between cause and effect, then causes would not determine their effects, for something could intervene in the gap and prevent the effect from occurring.

Therefore:

(3) There is no gap between causes and effect.

Therefore:

(4) Causes are simultaneous with their effects.

The conclusion of the argument is worrying, for it seems to be incompatible with the notion of temporally extended causal chains, and hence with the idea that earlier events are causally relevant to later ones. However, despite the appearance of plausibility, the argument fails. We could attack the first premiss, and insist that causation is in fact indeterministic, but the major defect is the move from the first conclusion to the second. If the effect occurs *immediately* after the effect, there is no gap.

In at least some cases, the suggestion that cause and effect are simultaneous leads to contradiction. For example, a spoon placed in a cup of coffee will cool the coffee down, but in doing so must, by the Second Law, itself heat up at the same time. Suppose the coffee's cooling down at time t' is caused by the spoon's being placed in the coffee at t. An essential part of the cause is the temperature of the spoon at t. Suppose, then, that the spoon is 20 °C at t. It must therefore be more than 20 °C at t', because in cooling the coffee it will have warmed up. If $t = t'$, therefore (i.e. if causation is simultaneous in this case) we get the contradiction that at t the spoon is both 20 °C and more than 20 °C. Another example is provided by colliding billiard balls. A red ball, moving with some speed across the billiard table, crashes into the stationary black ball at t, causing it to move at t'. The momentum of the red ball at t we will designate by m. It is this momentum that is an essential part of the cause of the black ball's acquiring a non-zero momentum. But the red ball must necessarily lose momentum in causing the black ball to gain it. So the momentum of the red ball at t' is less than m. The consequence of identifying t and t', i.e. of assuming simultaneous causation, is th contradiction that the momentum of the red ball is both m and les than m at the same time. The general principle in play here is that th state of an object must necessarily change when that object causes a change in something else.

What of the plausible cases of apparently simultaneous causation we presented above? The causal theorist will, of course, resist the suggestion that these have to be viewed as cases of genuinely simultaneous causation. It cannot be denied that we are presented with two states, or processes, that are simultaneous. The crown's presence on the cushion is simultaneous with the depression in the cushion. The motion of the engine is simultaneous with the movement of the carriage. But simultaneous processes, even simultaneous processes between which there is a causal connection, do not imply simultaneous causation. These processes have distinct parts: the movement of the engine at 10 o'clock is a distinct state from its movement at 1 minute past 10 (or even a split second past 10). And the presence of the crown at breakfast time is a distinct state from its presence at teatime, even if it has not changed at any time. These distinct states have distinct effects: it is the movement of the engine at 10, *not* at 1 minute past, that takes the engine past the signal box. It is the presence of the crown at breakfast-time, *not* its presence at teatime, that causes me to see the crown at breakfast-time. So we can resist the temptation to describe these cases as genuine cases of simultaneous causation by reflecting on the following redescription of them: the movement of the engine at a time, t, is the cause of the movement of the carriage shortly *after* t; the presence of the crown at t is the cause of the depression in the cushion shortly after t. To show that this is indeed the correct account of what is going on, bring the engine to a sudden stop. Does the carriage stop moving *immediately*? No, inertia carries it on a short distance, slightly compressing the engine. Lift the crown off the cushion. Does the depression disappear immediately? No, the cushion only gradually starts to reinflate.

A Sense of Direction in a Directionless World

The causal analysis, it seems, has been able to resist some of the objections thrown at it. But there is a question we left hanging at the end of the first section of this chapter, and so far we have not answered it. An ordered series, we said, is not the same thing as a directed series. We can order the series of integers, but we are not obliged to think of that series as running in a certain direction—from smallest to largest, say, rather than from largest to smallest. Now the causal theory of time order is just that: a theory of *order.* It gives an account of what makes an event earlier than another. But, we are inclined to suppose, the direction of time is more than that: it is the direction *from* earlier *to* later. So what makes this the preferred direction? We will end this chapter by looking at two quite different responses to this question.

As we saw in Chapter 8, there is more than one way in which we can order events in time. Events ordered by the *earlier than* relation constitutes a B-series, recall. But events may also be ordered as an A-series, that is, in terms of their pastness, presentness or futurity. A-series positions, we observed, *change*, so that what was once future becomes present and then ever more past. And this is the basis of the passage of time. So perhaps the missing element in our discussion of time's arrow is the passage of time itself. It is the present moving from earlier to later events, we could argue, that gives time its preferred direction.

Setting aside for a moment the difficulties, discussed in Chapters 8 and 9, that the idea of the A-series raises, we should ask whether the flow of time really does help us understand time's arrow. The idea is intuitively appealing: we imagine the present moving along the dimension of time in a particular direction. But what makes it the case that the present is moving towards the *future*, rather than towards the *past*? Of course, there is a trivial answer to this question, and that is that the future is just defined as that towards which the present is travelling. So let us rephrase the question. Suppose occur. What entitles us to say that the present is moving from A to D, rather than from D to A? We can only answer this question in temporal terms: in order of time, A comes first. That is, A is, quite simply, *earlier* than D. But if this is the explanation, then it is hard to see what the A-series explains, as far as direction is concerned, that the B-series does not. It is true that, if the A-series exists, then events only stand in B-series relations by virtue of their A-series positions, so that A is earlier than D only by virtue of the fact that, e.g. A is past and D is present. But that dependence of the B-series on the A-series would not illuminate the issue of *direction*, because we cannot describe the direction in which time flows without appealing to facts about order, and order can be as well expressed in B-series terms as A-series terms.

Let me approach the issue in a different way. If events truly have A-series positions, then those positions change. Thus, to give the full picture of the passage of time, we have to describe, not just the A-series positions of events at any given moment, but also their A-series positions at other moments:

(i) A is present, B is future, C is future
(ii) A is past, B is present, C is future
(iii) A is past, B is past, C is present

So, the answer to the question why the present moves from A to C, and not vice versa, is that the present moves from (i*) to (iii*):

(i*) (i) is present, (ii) is future, (iii) is future
(ii*) (i) is past, (ii) is present, (iii) is future
(iii*) (i) is past, (ii) is past, (iii) is present

But now we have to ask why the present moves from (i*) to (iii*) and not vice versa. And so it goes on. At no stage do we get a full answer to our question. Clearly, it would be better not to embark on this fruitless quest in the first place, and say simply that A becomes present first because A is earlier than B and C.

What these reflections lead us to is the following radical thought: perhaps what we have been pursuing in our discussion of this difficult topic will turn out not to exist. We have been assuming that there is a difference between a merely ordered series and a *directed* one, and it is directedness that we have found so elusive. *But what if there is no such difference?* What if, in other words, the direction of time is nothing more than the asymmetry of the 'earlier than' relation? Well, the asymmetry of that relation does not in itself explain why causes occur before their effects, why disorder increases, and why memories only form themselves after the relevant experience. But there may be a more fundamental fact underlying all of these, namely the logical asymmetry of the *causal* relation: if *a* was the cause of *b*, *b* cannot have been the cause of *a*. If the 'earlier than' relation can be reduced to the more fundamental causal relation, then we not only explain why causes occur before their effects (they do so by definition), we also, it may plausibly be argued, explain why disorder increases and why memories follow experiences: by appeal to the fact that these processes are causal processes. So the direction of time is just the direction of causation. But, it will be objected, although we may explain why the direction from earlier to later is also the direction from cause to effect, we have not explained why the direction from earlier to later, and not from later to earlier, is *the* direction of time. To which it can be replied: we do not have to say that there is a preferred direction. To go on insisting that time must have a direction in this sense is rather like saying that all hills really go *up*hill, never *down*hill.

This spatial analogy will perhaps help us find the last piece of the jigsaw. Although hills themselves are not more uphill than downhill, our *experience* of them may be. Thus, we can experience a hill as going upwards rather than downwards, or vice versa: it just depends where we start from. So, perhaps, it is with time: we simply experience it as going one way. But, if there is no direction of time over and above time order, is our experience of apparent direction not precisely the problem that needs to be explained? We cannot rest content with 'There is no direction: we just experience time as if there were'; we have to explain our *sense* of direction in a directionless world. This is, essentially, the difficulty with which the hero of H. G. Wells's *The Time Machine* is grappling with in conversation with his incredulous friends:

'Clearly,' the Time Traveller proceeded, 'any real body must have extension in *four* directions: it must have Length, Breadth, Thickness, and—Duration. But through a natural infirmity of the flesh . . . we incline to overlook this fact. There are really four dimensions, three which we call the three planes of Space, and a fourth, Time. There is, however, a tendency to draw an unreal distinction between the former three dimensions and the latter, because it happens that our consciousness moves intermittently in one direction along the latter from the beginning to the end of our lives.'

'That', said a very young man, making spasmodic efforts to relight his cigar over the lamp; 'that . . . very clear indeed.'

'Now, it is very remarkable that this is so extensively over-looked,' continued the Time Traveller, which a slight accession of cheerfulness. 'Really this is what is meant by the Fourth Dimension, though some people who talk about the Fourth Dimension do not know they mean it. It is only another way of looking at Time. *There is no difference between Time and any of the three dimensions of Space except that our consciousness moves along it.*'

Evidently, our Time Traveller is a B-theorist: he has no use for the passage of time. (Although, to be fair to B-theorists, many of them would be quite unhappy to hear time described simply as a fourth spatial dimension. The differences between time and space do not all disappear with the demise of the A-series.) What then is his explanation of our sense of the passage and direction of time? It is that 'our consciousness moves along' (in one direction, note) the dimension of time. But if our consciousness really is moving along this dimension, and not just apparently so, then the very thing that the Time Traveller is attempting to banish has been reintroduced. For how are we to describe this movement of our consciousness except in temporal terms? Our consciousness *was* at that point of time, and *is now* at this point. If such movement is permitted, then there really is such a thing as the passage of time, and the difference between time and space, so passionately denied by the Time Traveller, is restored.

Perhaps, however, the talk of movement in this context is only meant metaphorically (although it is unlikely that the Time Traveller's friends would have so taken it). Perhaps all we have are different states of consciousness at different times. But how could this give rise to the impression of direction (and passage)? Think again of the experiential asymmetry between earlier and later times: we remember the past, perceive the very recent past, but have no experiences or other psychological traces of the future. We can explain this asymmetry by appeal to the asymmetry of causation. Since the direction from earlier to later just is the direction from cause to effect, it follows that backwards causation is impossible. But experiencing later events would be an instance of backwards causation, so such experience is impossible.

Finally, we may note one more useful feature of the causal analysis of time order: if it is true, it solves the mystery of why time has, indeed must have, only one dimension. The causal series is itself necessarily one-dimensional. There is only one way in which events can vary in their causal distance from other events. If time order just is causal order, then time too can have only one dimension.

Questions

Could time run backwards?

If time is all in the mind, why does it seem to have a direction?

Are there any cases where a cause is simultaneous with its effect?