Chapter 5: Freedom and Determinism

Let me state at the outset a basic point that will reappear again below with its justification. The title of this chapter (and many other discussions too) make it appear as if the opposite of determinism is freedom. As we shall see, the opposite or denial of determinism is indeterminism. Freedom is a separate matter.

Taylor begins with an argument for (what he calls) the metaphysical thesis of determinism:

1. At each time \( t \) the world is perfectly determinate in all detail.

[Let us grant this for the sake of argument. Relativity raises some questions about the meaningfulness of the expression “each time \( t \)” Quantum mechanics provides strong reason to doubt “perfect determinacy”, but at the start we can ignore these points.

One might think of this premise in terms of language. It says that any sentence describing the world at any time \( t \) is either true or false,
depending on the way the described part of the world is.]

2. That is, at each time $t$ the world is in a definite state.

[We could view this as a definition of the term ‘state’. Let me give it a bit more content by quoting from the beginning of Earman’s “Defining Determinism”: “At each instant, the state of the world is fully characterized by specifying the values of the relevant physical magnitudes--instantaneous values of the positions and velocities of particles, instantaneous values of electric and magnetic field vectors, and the like.” (5) This is the way one would characterize a state of the world, or indeed a state of any system. One has to specify all the values of all the relevant magnitudes describing the system at some given time.]

3. The reason that the world in some given state at $t$ is that it was in some other state $t^*$ just before $t$.

4. And so on, forward (and perhaps also backward) in time.

5. The past antecedent conditions of our present state cannot be changed, since one cannot change the past.
:: (D)“Things [now and in the future] could not be other than they are.” That is, the present state of the world cannot be other than it is, and a similar argument applies to any other state at any other time as well.

Taylor says that this conclusion is “an exact statement of the metaphysical thesis of determinism.” (36)

He then adds: “More loosely, it [the metaphysical thesis of determinism] says that everything, including every cause, is the effect of some cause or causes; or that everything is not only determinate but causally determined.” (36)

But in the rest of the section called “Determinism” Taylor says that “the principle of determinism” is the principle that all events have causes.

If one looks carefully at the third premise Taylor’s argument, one might see a problem. Given that the world is in a particular state at t* and then in another particular state at t, in what way does the state at t depend on the earlier state at t*? Put another way, given that the world is in state S* at
time \( t^* \), why is it in the particular state \( S \) at \( t \)? Or put yet another way, what is it that governs the transition from state \( S^* \) to state \( S \)? If any state \( S \) could follow \( S^* \), we would have something quite unlike determinism, for the world could then be any way at all at \( t \). So for determinism to follow from these premises, the evolution of the world (or whatever system one is considering) from state \( S^* \) to \( S \) must be governed by some rule (by, in fact, some law or set of laws) that requires a system in state \( S^* \) to evolve into a system in one particular state \( S \) at \( t \).

So something crucial has been left out of Taylor’s argument, the law(s) governing the transition mentioned in premise (3).

In general, the description of the evolution of any system requires two elements. (1) A complete description of the system at some given time \( t \) (a complete description of its state), and (2) a law of set of laws that govern the transition from one state at one time to other states at other times.

Suppose, for example, that our “system” is a ball just as it is let fall at time 0 from the top of a tower. If its further states are given by the distance that it has fallen, then its state at time \( t \) (until it hits the earth) is \( S = \frac{1}{2} gt^2 \). (‘\( g \)’ is a
certain constant) This formula is the law (of falling bodies near the surface of the earth) that tells one how the states of the falling body change.

A long as nothing interferes with the ball (that is, as long as our system is *closed*), the law of falling bodies requires the ball to be in precisely one state at any later given time $t$ (before it hits the ground). That is, *it is the law that is deterministic*, and the crucial concept for the success of Taylor’s argument is one that is suppressed in the statement of the premises.

It is worth noting that if one changes the initial state of a system at a time (say one drops the ball from a tower of different height), then its future evolution will in general be different, even if governed by the same law or laws. Its past states will have to have been different as well.

If a law should fail to be deterministic, what does (or could) this mean? It clearly should mean that, if we have a system in state $S^*$ at time $t^*$, then it is compatible with that law that at some later time $t$ the system could be in any one of some number of distinct states $S_1, \ldots, S_n$. 
Let me try to give more intuitive content to this idea, starting from a classic discussion by William James in “The Dilemma of Determinism”. First he characterizes determinism:

**Determinism** “professes that those parts of the universe already laid down absolutely appoint and decree what the other parts shall be. The future has no ambiguous possibilities hidden in its womb: the part we call the present is compatible with only one totality. Any other complement than the one fixed from eternity is impossible. The whole is in each and every part, and welds it with the rest into an absolute unity, and iron block, in which there can be no equivocation or shadow of turning.” (569-70)

By way of contrast,

**Indeterminism** “says that the parts have a certain amount of loose play on one another, so that the laying down of one of them does not necessarily determine what the others shall be. It admits that possibilities may be in excess of actualities.... Of two alternative futures which we conceive, both may now be really possible; and the one become impossible only at the very moment when the other excludes it by becoming real itself.... [Indeterminism] says there is a certain
ultimate pluralism in [the world]; and, so saying, it corroborates our ordinary unsophisticated view of things.” (570)

[Note: The denial or opposite of determinism is indeterminism, as I said at the beginning of these notes.]

Indeterminism involves **chance**, “and chance [James writes] is something a notion of which no sane mind can tolerate in the world.” (572)

But James argues that chance is a useful, negative concept. “All you mean by calling [something] chance is that this is not guaranteed, that it may also fall out otherwise. For the system of other things has no positive hold on the chance-thing. Its origin is in a certain fashion negative: it escapes, and says, Hands off! coming, when it comes, as a free gift, or not at all.” (572-73)

At this point I take as having been established that the denial of determinism is indeterminism and that these two notions have to do with the transitions (one-to-one, or one-to-many) between states of systems. It is more controversial, but I think true, that one can speak of causes and effects even in indeterministic systems. Consider the following example.
Imagine that I rig up a box with a genuinely randomizing device with 100 equiprobable states. To this device I attach a revolver which will fire iff the device goes into state 17 (which has probability .01). In front of the revolver, in honor of Erwin Schrödinger, I put a cat. Suppose that I turn on the device, that it goes into state 17, and that the cat is killed.¹

After the humane society has me arrested for animal cruelty because I killed the cat, I argue in response that I did not cause the death of the cat (and so could not have killed the cat). Since the notion of cause, I argue, can not be applied to a genuinely randomizing device, my turning it on did not (indeed, could not) cause the cat’s death. I would not expect any judge or jury to accept this argument. Would you?

If you don’t accept it, then you can see just how far Taylor’s argument for determinism fall short of establishing its conclusion. Even if every event has a cause, if causes can be indeterministic or, as is usually said, probabilistic, then no conclusions about unique evolution can be derived from the thesis, even if true, that every event has a cause.

¹ Example taken from Fred Dretske’s “Causal Irregularity” in Phil. of Science (1972).
Even if the state $S^*$ at $T^*$ causes the state $S$ and $t$, if that cause is probabilistic, then the system could have been in other states at $t$ (and we may even be able to assign exact probabilities to those other states).

That is, according to Taylor the thesis of determinism is

(D) “Things [now and in the future] could not be other than they are.”

But we have just shown that, (D) does not follow from the premises of Taylor’s argument, even if one adds the additional premise that every event has a cause.

This is a long way round to agree with Earman’s crisply-stated view in §2 of his chapter that the concept of determinism is not clarified by invoking the concept of causation, a concept which is itself more obscure and difficult to pin down than the one it is supposed to clarify.

One instead has to look at the character of the laws governing the transitions between states of systems. This character cannot be determined by a priori inspection, even if we assume that we know what the basic laws of nature are. (For example,
Newton’s laws of motion are typically expressed as second-order differential equations. The question of determinism becomes this: under what conditions do these equations have unique solutions? The answer is by no means obvious, and similar questions with regard to more sophisticated putative laws are even more difficult.

Therefore I am arguing neither for determinism nor indeterminism. I am arguing that one or the other of these two “isms” must be true, since they are logically exclusive but mutually exhaustive concepts, if one assumes that the evolution of the world and any of its closed subsystems is governed by natural laws. We have pretty good grounds in experience for this latter assumption.

Now back to Taylor...

He thinks of determinism as a (conscious or unconscious) datum. We always act as if this thesis is true. We suppose that events or changes have causes, and we often seek to know the causes.

Since, as Taylor says, I am part of the world (37), his thesis of determinism applies to me and to my behavior. Since at the end of chapter 4 Taylor concluded that persons are living animal organisms
(34) and since it’s reasonable to suppose that living animal organisms are part of the world, then it follows that I, a person, am indeed part of the world. (If one wishes to deny that persons are “part of the world”, then one has to deny some step in the argument above, and we have seen in some detail in chapters 2 to 4 that it is not easy to defend such a denial.)

If determinism as characterized by Taylor is true, then it is supposed to follow that

(i) I cannot in any circumstances or at any time act in any way other than the way I actually do, did, or will act.
(ii) Since one is held morally responsible only for actions that one has voluntary control over, for actions that could have been otherwise, [“Ought implies can.”], we cannot be held morally responsible for any of our actions (since none of our actions could have been otherwise).

Taylor starts his investigation from the following two pieces of data. [Remember, that data are starting points and may eventually be given up in a more considered final view.]

1. I sometimes deliberate about what to do.
2. It is sometimes up to me what I do.

I can deliberate about my own future behavior as long as (1) I do not already know what I am going to do in the given situation and (2) what I do in the given situation is “up to me”.

An action is “up to me” if it is within my direct control, like moving my finger in one direction or another. I can (normally) move my finger to the left, says Taylor, and I can move it to the right. These actions are possible, but not merely in the sense of logically possible. They are, as Taylor says in another place, “within my power”. (43) It’s logically possible that I move my finger to the left, even when it’s in a cast so that (in the relevant sense) I can’t move it at all. It is not within my power to do so.

What this shows is that logical necessity (or possibility) is not the only kind of necessity (or possibility), and not the kind of modality that’s most important for discussions of determinism. There is at least one alternative sort of modality, one that Taylor calls causal necessity (or possibility). Others may call this nomic necessity, since the Greek word for law is nomos.
Indeed, one might think of the modal term ‘can’ as polysemous (which is a more accurate term than ‘ambiguous’). One can think of a particular shade of meaning of ‘can’ as depending on the range of assumptions that one takes to be compossible in the given situation. Alternatively, one can think that the term varies with set of worlds it deems possible. A logically possible world is one constrained only by logic, by consistency. A nomically possible world is one constrained by logic and the laws of nature, whatever they may be. So there are fewer nomically possible worlds than there are logically possible worlds. This topic will be further discussed in later lectures. I expect to have more to say about the notion of power there.

But perhaps it is helpful to say at this point that a more exact formulation of the thesis of determinism would go something like this: (D’) If one has a system in state S* at T*, then there is only one state S that is nomically possible at t. As I have said, I neither endorse nor dissent from (D’). I want only to (try to) clarify what is actually involved in the thesis.

To say that an action is “up to me” is to say that I am free to do it or to refrain from doing it. This, in turn, presupposes that there is no obstacle that prevents me from doing it and no constraint that
forces me to do it. I am not compelled, constrained or coerced to do it, if I am free.

If this is all there is to freedom, then freedom is compatible with determinism (= causally determined), because having a cause and being compelled, coerced or constrained are entirely distinct concepts that need not coincide (or be co-extensive) in their application.

In Taylor’s view, philosophers are often led to make too much of this compatibility and to espouse a view that he calls (following James) soft determinism. **Soft determinism** involves the following three theses:

1. The thesis of determinism is true.... All human behavior is caused and determined.
2. Voluntary behavior is free to the extent that it is not externally constrained (forced, coerced, compelled) or impeded.
3. Such voluntary behavior is caused by certain states of the agent—volitions (or acts of will), choices, decisions, desires, and so on.

But now, argues Taylor, ask a soft determinist this question: in a given situation, could I have acted
otherwise? Yes, they will say, if my volitions or desires, etc. had been otherwise.

Well, could my volitions or desires have been otherwise? By (1), determinism, they were caused. So they were determined by antecedent conditions to be what they were and could have been different only if those conditions had been different... And so on.

But, given determinism, none of these conditions could actually have been otherwise. So none of my actions (or the volitions, desires, decisions, etc. that were causally sufficient for them) could have been different (actually, rather than conditionally). So how can I be held responsible for them?

This argument is buttressed by the example of the “ingenious physiologist” who can implant volitions. The subject of these implantings fits the soft determinism definition of a free agent yet is in reality a mere puppet of the physiologist. (That is, the conditions for free actions were supposed to be sufficient conditions. If so, they would remain free even if more details about them and their origin were added. Yet one is supposed to deny that the actions are really free when the ingenious physiologist intervenes.)
Taylor believes he has shown, then, that both hard and soft determinism are views that have consequences that are very difficult to accept. The only alternative left seems to be what he calls **simple indeterminism** but he finds that unacceptable too. Simple indeterminism is the view that some of our actions (or some of the mental acts that lead to our actions, like our volitions or decisions) are uncaused.

But that is to say that they are a matter of chance (cf., James) or are random. Consider, urges Taylor, the picture one gets of one of my arms or legs just spontaneously (because uncaused) moving in some direction or other from time to time. Is this a picture of a free person? No, says Taylor. “[T]he conception that now emerges is not that of a free person, but of an erratic and jerking phantom, without any rhyme or reason at all.” (48)

So simple indeterminism also seems incompatible with our normal sense of self and with attributions of responsibility. Suppose, for instance, that certain movements of my arm are uncaused or chancey.

...[N]either I nor anyone else can tell what this arm will be doing next. It might seize a club and lay it on the head of the
nearest bystander, no less to my astonishment than his. There will never be any point in asking why these motions occur, or in seeking any explanation for them, for under the conditions assumed there is no explanation. They just happen, from no causes at all. (48)

How, then, could I (or anyone) be held responsible for these actions over which I had no control (or, which were not “up to me”)?

[As far as I know, the argument that indeterminism or chance is incompatible with freedom and moral responsibility originated in Hume’s Treatise. Hobart argues along the same lines that moral responsibility requires “determination”.]

At this point, Taylor briefly recapitulates the problems with the three views discussed so far. I’d like to highlight one remark of his: “But if determinism is true, then there are always conditions existing antecedently to everything I do, sufficient for my doing just that, and such as to render it inevitable.” (49) Notice that, officially speaking, to say that some condition C is sufficient for (say) an action A is to assert the material conditional: C ⊃ A.
So we have the following argument, which is valid, since it is a simple application of the inference rule *modus ponens*: $C, C \supset A \therefore A$.

The premises seem warranted by the thesis of determinism, but the conclusion is just that $A$ occurs or takes place, *not* that it is inevitable. To say that $A$ is inevitable, that in some sense $A$ must occur, is to invoke a modal notion. It is to say something like: $\Box A$ or $\sim \Diamond \sim A$, which is of course equivalent. How is the introduction of a modal notion supposed to be justified? That is, the argument $C, C \supset A \therefore \Box A$ is not valid, so where does inevitability (represented by a formula with a modal operator) come from?

The following *is* true: $\Box [(C & C \supset A) \supset A]$. That is, here is a conditional sentence that is a necessary truth, because it is a logical truth. But it does not follow from the truth of this sentence that $A$ is “inevitable”. If one could assert as a premise $\Box (C & C \supset A)$, then one could conclude $\Box A$, but nothing so far that has been said justifies that premise, if ‘$\Box$’ is logical necessity. Taylor explicitly denies that the causal relation is logically necessary, but $\Box (C & C \supset A)$ entails both $\Box C$ and
\(\square (C \supset A)\). We may return to this line of argument eventually, thinking instead in terms of nomic rather than logical necessity (or possibility).

There is one more possible view regarding determinism (Taylor thinks), the theory of agency.

“The only conception of action that accords with our data is one according to which people—and perhaps other things too—are sometimes, but of course not always, self-determining beings; that is, beings that are sometimes the causes of their own behavior.” (51)

This view presupposes a couple of remarkable metaphysical theses:

(i) I am a self or person—a self-moving being.
(ii) A substance (e.g. a self) can cause (originate, initiate) an event, without anything (else) causing it to do so.

[It also presupposes the denial of Laplacean determinism as defined by Earman, since Taylor says: “In the case of an action that is free, it must not only be such that it is caused by the agent who
performs it, but such also that no antecedent conditions were sufficient for his performing just that action.” (51)]

According to Taylor,

- The theory of agency avoids the absurdities of simple indeterminism by conceding that human behavior is caused.
- The theory of agency avoids the difficulties of determinism by denying that every chain of causes and effects is infinite.

Nevertheless, Taylor wonders whether the theory of agency is, when fully considered, so odd that one might not wish to go back and question the data he began with. It is an odd view, and R. E. Hobart, I believe, got to the bottom of what is odd about it, as we shall see when we discuss his paper.

For the moment, however, consider the following problem. Consider some action A that I perform at some time t. I, my self, typically have existed for a long time before and after the action A at t. So the existence of my self (or Self) is not sufficient for the occurrence of A, though it may be necessary. What, then, completes the set of standing necessary conditions at (or just before) t so that A
occurs at \( t \)? That *triggering condition* is, presumably, something dated, something that happens at a time—that is, a change or event. *But it is exactly this event that is left out of the theory of agent causation even though typically a triggering event in a set of standing conditions is singled out as “the cause” of what follows, as striking a match is considered the cause of its lighting.*

One might think that this line of argument begs the question, since Taylor clearly affirms that prior to the action’s occurring (or being willed?) no sufficient antecedent conditions for the performance of just that action existed. So by hypothesis there is no answer to the question: why did this particular action occur. But it seems to me that one can legitimately ask the different question: why did that action occur *at the particular time that it did*, as opposed to any other time when my Self existed? How can a continuing substance (the Self) cause an action *at a particular time*?

[One can also ask why a particular radioactive atom decays at the precise time that it does decay, and there is no answer to this question, though there is a general probabilistic law of radioactive decay. The theory of agency is committed, I think,
to the view that the Self is constrained by no laws of any kind, whether deterministic or indeterministic.]

In fact, Taylor clearly thinks of the theory of agency as a kind of deterministic theory (52), but one that avoids the difficulties of determinism by denying that causal chains stretch indefinitely back into the past. They come to a stop at selves as causes, or initiators, of actions. Some causal chains have a beginning.

But how does this differ from indeterminism as I have characterized it? The theory is certainly not deterministic in Earman’s sense (of future Laplacean determinism).

The theory of agency entails that the Self cannot be an object of scientific study. The Self, according to that theory, “rises above” its past and present circumstances to initiate an action or a chain of causes leading to an action. It is precisely by being able to escape the past history and present constraints, the theory says, that the Self acts freely. But laws connect state transitions to earlier states or conditions, so when the Self acts freely it cannot be governed by law (whether deterministic or indeterministic).
Here we may be at a basic divide. Some see this consequence, that Selves can somehow rise above the natural order, as absurd. If we are “living animal organisms”, then our behaviour must be governed by laws and amenable to scientific study, just like other living animal organisms. Others think that science simply reaches a limit or a barrier that it cannot cross when it studies the behaviour of human beings and it is in this un-explorable territory that our humanity resides.

I can offer no argument for the latter view, but the steady advance of the science of psychology lends plausibility to the former view.