

Dennett's Compatibilism

Determinism just isn't relevant

Dennett rejects PAP

Instead of rising to the defense of any of the earlier analyses—many of which are quite defensible so far as I can see—I will go on the offensive. I will argue that *whatever* “could have done otherwise” actually means, **it is not what we are interested in** when we care about whether some act was freely and responsibly performed.

PAP = “the **P**inciple of **A**lternate **P**ossibilities”
= the requirement that a free agent “could have done otherwise.”

- N.B. there are two basic stances that a compatibilist can take towards PAP:
 - Accept PAP, but interpret it in a way that is compatible with determinism (e.g. Stace, Hume)
 - Reject PAP (Dennett)

Frankfurt cases

Jones hates Smith and decides, in full possession of his faculties, to murder him. Meanwhile Black, the nefarious neurosurgeon (remember him?), who also wants Smith dead, has implanted something in Jones' brain so that *just in case Jones changes his mind* (and chickens out), Black, by pushing his special button, can put Jones back on his murderous track. In the event Black doesn't have to intervene; Jones does the deed all on his own.

Dennett agrees, but ...

... his counterexamples are rather special and unlikely cases, and they invite the defender of the principle to try for a patch: modify the principle slightly to take care of Frankfurt's troublesome cases. Exotic circumstances do little or nothing to dispel the illusion that in the normal run of things, where such overdetermination is lacking, the regnant principle is indeed that if a person could not have refrained (could not have done otherwise), he would not be held responsible.

First assault on PAP

“Here I stand,” Luther said. “I can do no other.”
Luther claimed that he could do no other, that his conscience made it *impossible* for him to recant. ...
...Whatever Luther was doing, he was not trying to duck responsibility.

(Similarly with Dennett being offered \$1000 to torture an innocent person.)

What do you think?

1. Kane's response

“My response to Dennett is to grant that Luther could have been responsible for this act ... though he could not have done otherwise then and there and even if his act was determined. But this would be so to the extent that **he was responsible for his present motives and character** by virtue of many earlier struggles and self-forming choices (SFAs) that brought him to this point where he could do no other.”

- Robert Kane, “Free Will: New Directions for an Ancient Problem”, 2002.

2. My response: What kind of modality?

- A libertarian might respond that this argument equivocates on the modal terms, like 'cannot', that come in various flavours.
- E.g. there is not just **nomic** possibility, but also **epistemic** possibility and **deontic** possibility. Was Luther saying that all alternative actions in this situation were *nominally* impossible?

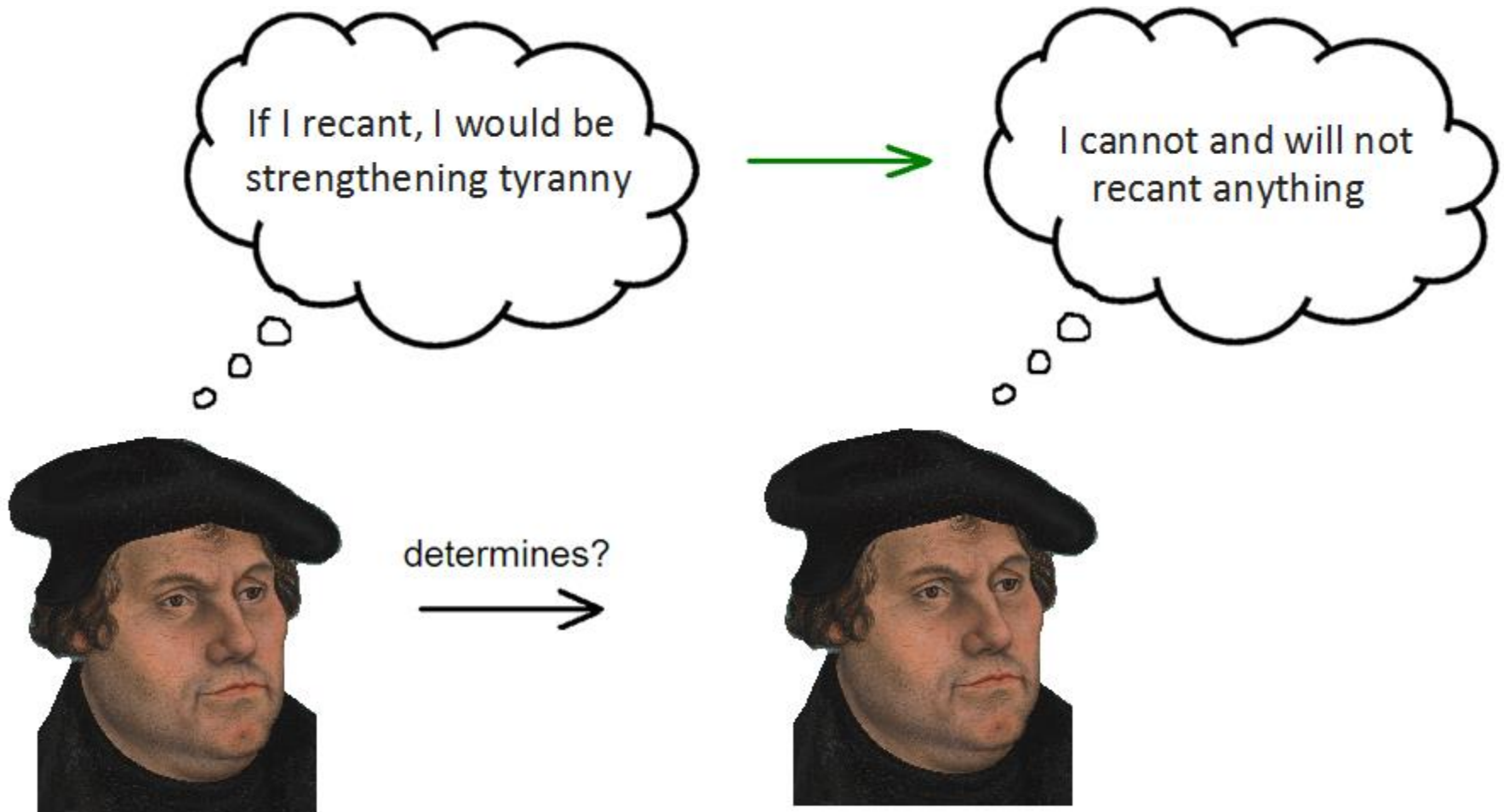
(Some kinds of modality)

Nomically possible:	Consistent with the laws of physics and the actual past
Epistemically possible:	Consistent with my knowledge
Deontically possible:	permissible, or consistent with my moral obligations

It seems to me that Luther is talking about *deontic* possibility here. His conscience is clear, so he is sure the present action is morally required, and anything else morally impermissible.

- Dennett actually says something similar:
“But in other cases, like Luther’s, when I say I cannot do otherwise I mean that **I cannot because I see so clearly what the situation is** and because my rational control faculty is not impaired. **It is too obvious what to do; reason dictates it;** I would have to be mad to do otherwise, and since I happen not to be mad, I cannot do otherwise.” Dennett, p. 4
- But notice how Dennett is committed to the thesis that when someone like Luther is very sure that he’s morally obliged to do X (and is not mad) he’s also nomically determined to do X.
- Is this thesis true?
- N.B. According to ‘opaque thought’ libertarianism, if Luther’s action is truly caused by his intentional states, it actually *cannot* be deterministic.

How do these levels relate?



Martin Luther at the Diet of Worms, 1521

How do these levels relate?

- If Luther is subject to deterministic laws, then they would seem to be *psychological* laws – perhaps laws of folk psychology.
- Folk psychology is a pretty useful and reliable theory, on the whole, but we also know that **it fails to capture the full depths of the mind**. (We can't even predict our own behavior, much of the time.) People suffer from akrasia (= 'weakness of will'), laziness, cognitive errors, etc.
- So folk psychology gives no evidence that the best physical theory describing Luthor's microphysical states is deterministic.

Perhaps this is what some people think: they think that if I were right when I claimed I could not do otherwise in such cases, I would be some sort of zombie, “programmed” always to refuse thousand-dollar bribes. A genuinely free agent, they think, must be more volatile somehow. If I am to be able to listen to reason, if I am to be flexible in the right way, they think, I mustn’t be too dogmatic. ...

... That would be fallacious reasoning. Seeing both sides of the question does not require that one not be overwhelmingly persuaded, in the end, by one side.

(Dennett, p. 5)

Agree? (Note that an ‘opaque thought’ libertarian disagrees with this.)

A good person *takes away* their own CDO

- Dennett considers the case of a werewolf who always arranges to have himself locked up during a full moon. And similar, more realistic, cases, where one tries to make oneself *the kind of person* who would not do evil (e.g. commit rape).

“Doesn’t a considerable part of being a responsible person consist in making oneself *unable to do* the things one would be blamed for doing if one did them?”

- Is this right?

- Thinking in terms of control, this seems like the younger self exerting some control over the older self. (Full control in the case of the werewolf.)
- Robert Kane, as we have seen, says something very similar. A good person uses SFAs (self-forming actions) to cause himself to act rightly in the future.
- An “opaque thought” libertarian thinks this is generally correct, but a bit extreme. SFAs exist, and one responsibility we have is to make ourselves virtuous. (But this doesn’t entirely take away your later CDO.)

Second assault on PAP:

‘a most peculiar problem of ignorance’

“If our responsibility really did hinge, as this major philosophical tradition insists, on the question of whether we ever could do otherwise than we in fact do *in exactly those circumstances*, we would be faced with a most peculiar problem of ignorance: it would be unlikely in the extreme, given what now seems to be the case in physics, that *anyone would ever know* whether anyone has ever been responsible.” (p. 7)

1. *Assume* PAP (FW \Rightarrow CDO)
2. Given our present understanding of quantum mechanics, and the complexity of our brains, **it's unclear** whether or not our decisions are causally pre-determined in a given exact case.

- \therefore 3. We will never have *any* good reason to believe, of any particular act, that it was or was not responsible.
- \therefore 4. There is something badly wrong with PAP (since it leads to 3.)

“The critical difference [between free and unfree actions] would be **utterly inscrutable** from every macroscopic vantage point, and practically inscrutable from the most sophisticated microphysical vantage point imaginable. ...

... To say the very least it is hard to take seriously the idea that something that could matter so much could be so magnificently beyond our ken.” (p. 8)

Parallel case: deterministic chances?

1. There are objective probabilities, or ‘chances’, in the natural world. (E.g. the probabilities in quantum mechanics.)
2. On *some* theories of chance, there would be no objective probabilities (other than 0 and 1) if determinism were true.
3. However, right now we don’t have any proof of indeterminism.

∴ Any such theory of chance is wrong

- E.g. Carl Hoefer (2007, p. 557): “Any view of chance that implies that there may or may not be such a thing after all—it depends on what the laws of nature turn out to be—must be mistaken”

Another parallel case

- Imagine an advocate of continental drift arguing (in 1940 say) that colliding continental plates are a necessary condition for fold mountains to exist.
- A critic might have objected that, since the alleged drift is too small to measure, it cannot be known to exist. Then the colliding-continent theory of mountain formation entails that fold mountains might not exist.
 - “Any view of mountains that implies that there may or may not be such things after all—it depends on whether or not the continents move—must be mistaken”.

- Surely in this case one can argue that:
 - (i) We know mountains to exist, by direct observation.
 - (ii) Drift provides the best explanation for mountainsHence, (iii) The existence of mountains **gives indirect evidence of drift.**

Similarly, the existence of objective chances might provide indirect evidence for indeterminism, if the best theory of chance requires indeterminism. (?)

Similarly, if we take the existence of moral responsibility as a fixed premise, as Dennett does, then any argument for incompatibilism will be indirect evidence for CDO?

Introspection is reliable?

- Dennett seems to assume that the only way we can judge whether a given act is deterministic is **scientifically**, using brain scanners, physical theories, and the like.
- But perhaps our “feeling of freedom” is like a sense organ, or like introspection? Our sense organs and introspection often give us information that couldn’t be obtained scientifically (or not till recently, in some cases).

Introspection is reliable?

- E.g. we can just tell someone's emotions by looking at their face, or body language. Can quantum mechanics tell us this?
- We know our own mental states, e.g. beliefs, pains, by introspection. Can a brain scan tell us this?
- The feeling that one has the power to do A, and the power to do B, could be the output of a reliable cognitive mechanism, and hence a case of knowledge (according to one analysis of knowledge).

Part 2

What do we care about?

3rd Assault on PAP: “What we care about”

- Why do we ask: “could Jones have done otherwise?”
- Dennett says that we want to know *general* information about Jones, about his character, his *dispositions* to act in certain ways, in certain situations.
- In other words, we want to know whether this was ‘in character’ for Jones, something he would always or usually do in similar circumstances, or whether it was just a freak occurrence (like a speck of dust getting into an unlikely spot).

- Dennett discusses the case of a deterministic robot that does something undesirable in a given situation.

“What concerns the engineers when they encounter misperformance in their robot is whether or not the misperformance is a telling one: does it reveal something **about a pattern of systematic weakness**, likely to recur, or an inappropriate and inauspicious linking between sorts of circumstances and sorts of reactions? Is this *sort* of thing apt to happen again, or was it due to the coincidental convergence of fundamentally independent factors, highly unlikely to recur?”

- Why do we ask “could he have done otherwise?” We ask it because something has happened that we wish to interpret. An act has been performed, and we wish to understand how the act came about, why it came about, and what meaning we should attach to it. That is, we want to know what conclusions to draw from it *about the future*.
- Responses?
- In my opinion, we *do* wonder if an unexpected act reveals a character flaw, with implications for the future. (E.g. we often say, “If he did it once, he’ll probably do it again.”)
- Is that one reason why we ask whether the person could have done otherwise?
- If so, is it the *only* reason?

Be a self controller

“Knowing that I will always be somewhat at the mercy of the considerations that merely happen to occur to me as time rushes on, knowing that I cannot entirely control this process of deliberation, I may take steps to bias the likelihood of certain sorts of considerations routinely “coming to mind” in certain critical situations.” (p. 17)

- Dennett says that the value in regretting past mistakes, and wishing one had done otherwise, is to *alter one's dispositions*, and so avoid making similar mistakes in the future. (Do you agree?)
 - (I think I agree with that. People can be blameworthy for being careless, as well as for doing deliberate harm.)

Does determinism say that you never underperform?

- van Inwagen (1975, pp. 49-50) says, “To deny that men have free will is to assert that **what a man does do and what he can do coincide.**” In a deterministic world what sense could we make of the exhortation to do the best we can? It does seem to us that sometimes people do less well than they are able to do. How can we make sense of this?”
- N.B. The consequence argument tries to show that determinism erases the distinction between the things that are under our control, and the things that aren't. This is a similar claim by van Inwagen.

Dennett's response

- Even if the world is deterministic, Dennett says, there seems to be a distinction between what *happens*, and what *must* happen.
- E.g. a coin *did* land heads, but it wasn't *necessary*. It *could* have landed tails. (It could *not* have zoomed off to Alpha Centauri, however.)
- In a deterministic universe, if a particular oxygen atom never bonds to any hydrogen atoms, then it would be physically impossible for it to do so. But surely that's a crazy statement?

- Assuming determinism, in what sense of ‘could have’ is it true that the coin could have landed heads, but could not have zoomed off? Ideas?
 - Dennett suggests *epistemic* possibility. I.e. the coin “could have” landed tails, in the sense that this is consistent with our knowledge of coins. (So it’s an illusion?)
 - (An alternative view: Perhaps it’s instead a matter of consistency with the laws *only*, vs. consistency with the laws *plus the initial state*? (Still just an illusion?))

‘epistemic’ randomness is enough

- In science, Dennett argues, it seems to make no difference whether ‘random’ phenomena such as genetic mutations are truly random, or rather just patternless, unpredictable, independent of other things, etc.
- Evolution (or ‘Mother Nature’) is also unable to predict chaotic events, and so has to give us general capacities to handle what comes at us. She gives us some ‘avoidance machinery’.

Reflections

- Dennett has a rather ambitious goal in this chapter, to argue that:
It doesn't matter at all whether or not, in a given case, we could have done otherwise
- It's not relevant, it's not something we care about, or are interested in.
- When we *seem* to be interested in CDO (in the strict metaphysical sense) we're actually interested in something else.

Girl: I want a pony!

Dennett: I guess you want to get around more easily.
Here's a scooter.

Girl: That's nothing like a pony!

Dennett: Ok. How about this? It's what you really want.



CDO and “originative value”

Nozick (1981, p. 313) claims that we all want “originative value,” but the only conditions under which we would have this are (on his analysis) conditions that apparently demand the metaphysical reading of “could have done otherwise”: “We want it to be true that in that very same situation we could have done (significantly) otherwise, so that our actions will have originative value.”

Once again, is it plausible at all that something we care so much about (if Nozick is right) is something we could never know to be the case? Put another way, **if originative value requires this, why would anyone care about having originative value?** (Dennett, p. 10, note 2)

Freedom and creativity

- If you look at the history of technology, a lot of new stuff has been invented since the Stone Age. (Is an iPhone just a slight variation on a flint axe?)
- Same with the history of music, art, politics, commerce, etc.
- We say that (at least some) humans have the power of *creativity*, the ability to invent things that are really new, not just 'derivative'.

Freedom and creativity

- Human creativity is probably not well understood.
- But in engineering, it surely requires *understanding* of the problem, and of the materials and technologies available for a solution. ('Intentionality' is rather mysterious.)
- (Creative engineering also seems to need flashes of insight, but those are even more mysterious.)

Freedom and creativity

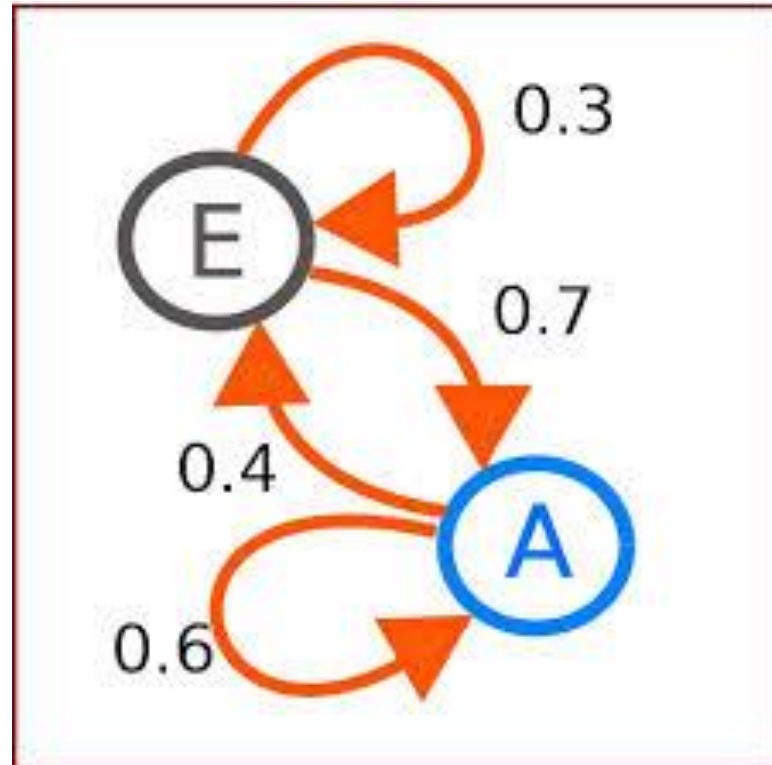
- Can *deterministic* systems be creative in this way?
- Does *free will* have anything to do with creativity (novelty, originality, etc.)?
- In a deterministic system, the whole history is a logical consequence of the initial state and laws.
 - The history contains no new information ('yawn')
- In a stochastic system (with transition probabilities rather than deterministic laws) the **stationary probabilities are fixed**.

- A simple Markov chain, with just two states (A and E). The “transition probabilities” represent the dynamical laws of physics.

Stationary probabilities:

E – 0.36

A – 0.64



Stationary probabilities

- The transition probabilities determine (under typical conditions) a “stationary probability” for each state.
- This measures the extent to which the dynamics favours that state. (State A is favoured on the previous slide – stationary probability 0.64)
- These stationary probabilities are fixed by the dynamics, and never change. The system keeps doing the same *kinds* of thing.

Creativity?

- In other words, a physical system does what it is “programmed” to, and this also applies to stochastic systems.
- **The laws of physics and initial state determine what is probable and what isn’t**, and there’s no changing these probabilities significantly in the absence of very unlikely events.
- We tend to think that creativity involves “breaking out of” these limits. E.g. if the process of evolution is creative, then the process *itself* creates high stationary probabilities for functional objects.
- But such creativity is apparently impossible in a physical system.

“the evolution of life marks the end of a physics world view of law entailed dynamics. Our considerations depend upon discussing the variability of the very “contexts of life”: the interactions between organisms, biological niches and ecosystems. These are ever changing, intrinsically indeterminate and even unprestatable: we do not know ahead of time the “niches” which constitute the boundary conditions on selection. More generally, by the mathematical unprestatability of the “phase space” (space of possibilities), **no laws of motion can be formulated for evolution**. We call this radical emergence, from life to life.”

“No entailing laws, but enablement in the evolution of the biosphere”,
Giuseppe Longo, Maël Montévil, Stuart Kauffman (arXiv, 2012)

Robert Nozick seems to agree:

- “The probability of a strike at bowling is not altered by anything the ball does; that probability is fixed by other events” (p. 311)
- “A being with originaive value, one whose acts have originaive value, can make a difference” (p. 312)