Review: Why Ain't Evidentialists Rich?

November 2022

I was engrossed by this paper, which is charming and incisive. In it, the author develops a "Why Ain'cha Rich?"-style objection to Evidential Decision Theory. The context is broadly game-theoretic. For example, Newcomb is represented like this:

	P1	P2
1	(m, a)	$(0, 0)$
2	$(m + t, 0)$	$(0 + t, b)$

where the payoffs for the predictor ("demon") are in the second coordinate; this is a simple way of operationalizing the idea that the demon is a good predictor because she is being rewarded for getting things right ($a, b > 0$). This was extremely natural to me and I found it easy to follow as well as convincing.

My questions concern how this interacts with at least two other bits of the "WAR" literature that are familiar to me: first, Lewis's own discussion of both the original objection (1981) and the relation of Newcomb to the Prisoners' Dilemma (1979), and second, the Wells-Ahmed exchange in *Mind* in 2019.

To the first: Lewis does claim, in "Why Ain'cha Rich?", that it is impossible to *tu quoque* the EDTer. The author appears to be aiming to do precisely this. So Lewis must have been wrong. It would be worth at least mentioning this. Lewis also claimed, in "Prisoners' Dilemma is a Newcomb Problem", that it was illuminating, when transposing between decision and game theory, to think of Newcomb as a PD. But the game in the table above is *not* a PD: it's not the case that there's a unique "defect-defect"-style equilibrium. So Lewis must have been wrong here too.

To the second: as far as I can tell, the author is aiming to do precisely what Wells aimed to do in his 2019 paper, "Equal Opportunity and Newcomb's Problem": to wit, to launch an *ain'cha rich* that doesn't lend itself to the obvious objection that the setup discriminates against the relevant decision theory from the start (an objection the original WAR argument is plausibly susceptible to!). Indeed, one of Wells's sections is called "why ya poor?". For what it's worth, I find the author's examples *much* easier to follow, on an intuitive level, than

Wells's. But I worry that the response, on behalf of the EDTer, will be quite similar to Ahmed's defense against Wells. In a nutshell, that defense is this. The argument should only work dialectically if the EDTer and CDTer can plausibly be modeled to be in the same decision problem. But this style of diachronic decision problem involves each theorist (EDTer and CDTer respectively) knowing what kind of theorist she is, in order to anticipate her future actions (such as, in the EDTer's case, paying to avoid information regarding what the Demon predicted). If that's so, then they can't know *the same things*, and therefore can't be in *the same problem*. While I'm not sure I find Ahmed convincing here, it does seem worth noting that he is *not* complaining that the EDTer is being discriminated against at the level of equal opportunity in Wells's sense (or, if I understand correctly, the sense the author would describe as "the EDTer['s being] harmed by Demon's malicious choices" (pg. 5).

This goes, by the way, to one of the sudden transitions in the manuscript as it stands: between the first example ("Split Newcomb") and the second two ("Coins and Signals" and "Coins and Newcomb"). We get this cheeky passage:

> I think [the above] responses work, but…they don't even fully convince me. And I'm normally more sympathetic to my arguments than others are. So let's look at a different example, one where Demon doesn't have these variable payouts.

I admit I chuckled aloud, but this does seem like exactly the sort of place some discussion of Ahmed's point would go.

Two Typos

- In the third column at the top of Figure 2.1, U (for "Up" I presume) should be uniformly replaced with H (for "Heads").
- On page 4, third full paragraph, the author has "But an EDTer will know that if they Play, they expect to get 4…". I think this should be "they expect to get 3…"; the author may want to double-check this at least.