

What You (Want to)* Want

paulgraham.com · Paul Graham · 2022-11 · [source](#)

|

November 2022

Since I was about 9 I've been puzzled by the apparent contradiction between being made of matter that behaves in a predictable way, and the feeling that I could choose to do whatever I wanted. At the time I had a self-interested motive for exploring the question. At that age (like most succeeding ages) I was always in trouble with the authorities, and it seemed to me that there might possibly be some way to get out of trouble by arguing that I wasn't responsible for my actions. I gradually lost hope of that, but the puzzle remained: How do you reconcile being a machine made of matter with the feeling that you're free to choose what you do?

[1]

The best way to explain the answer may be to start with a slightly wrong version, and then fix it. The wrong version is: You can do what you want, but you can't want what you want. Yes, you can control what you do, but you'll do what you want, and you can't control that.

The reason this is mistaken is that people do sometimes change what they want. People who don't want to want something — drug addicts, for example — can sometimes make themselves stop wanting it. And people who want to want something — who want to like classical music, or broccoli — sometimes succeed.

So we modify our initial statement: You can do what you want, but you can't want to want what you want.

That's still not quite true. It's possible to change what you want to want. I can imagine someone saying "I decided to stop wanting to like classical music." But we're getting closer to the truth. It's rare for people to change what they want to want, and the more

"want to"s we add, the rarer it gets.

We can get arbitrarily close to a true statement by adding more "want to"s in much the same way we can get arbitrarily close to 1 by adding more 9s to a string of 9s following a decimal point. In practice three or four "want to"s must surely be enough. It's hard even to envision what it would mean to change what you want to want to want to want, let alone actually do it.

So one way to express the correct answer is to use a regular expression. You can do what you want, but there's some statement of the form "you can't (want to)* want what you want" that's true. Ultimately you get back to a want that you don't control.

[2]

Notes

[1]

I didn't know when I was 9 that matter might behave randomly, but I don't think it affects the problem much. Randomness destroys the ghost in the machine as effectively as determinism.

[2]

If you don't like using an expression, you can make the same point using higher-order desires: There is some n such that you don't control your n th-order desires.

Thanks to Trevor Blackwell,

Jessica Livingston, Robert Morris, and
Michael Nielsen for reading drafts of this.

|