

Our Big Questions

Course Description

This course will be an introduction to philosophy in the analytic tradition, by focusing on a few representative issues:

1. How can we tell whether animals and future computers have “minds” — that is, their own thoughts, experiences, ambitions, self-awareness, and so on — or whether they’re instead just mindless automata?
2. Relations between minds, brains, and machines: Are your mind and body made of different stuffs? If a machine duplicates the neural structure of your brain, would it have the same thoughts and other mental states that you have?
3. What does it take to have free will? Is this incompatible with one’s choices being programmed or physically determined?

A lot of this class will be focused around notion of **having a mind** (sometimes people say "souls," "selves," "consciousness"… we'll introduce ways these words can be understood differently later) — that is:

- having opinions, being able to think things through, solve problems, and learn
- having feelings, experiences (conscious), emotions
- having ambitions, preferences, plans
- aware of itself as "me" (conscious), reflect on its own states (revise them), and its own limits (compensate for them)

1. What kinds of mental lives — *how much* of these things — can animals and/or future computer systems have? *How can we tell* whether/when they do? or whether they're just mindless automata, like Roombas or (perhaps) bees?
2. How are minds connected to organic brains and bodies? are they made of different stuffs: does having a mental life require having a non-physical/immaterial soul?
3. What about to other physical bodies, such as machines — AIs of the near or far future? Can they ever have mental lives too?

Not the same as: would their mental lives be the same as ours? (different cultures / genders / generations)

Or does some or all of our mental life essentially depend on something the machine will lack (our biology? a non-physical soul?)

a) Can a machine have "its own" thoughts, or only ones its designers preselect for it?

b) Can it feel frustration or distress? Can it suffer?

c) *What if we "uploaded" the structure of a living organic brain into a computer: what would the program running that structure be missing out on? Would it be having the same thoughts and feelings the living brain did? Would it have the same awareness/consciousness of itself and its own existence?*

d) *How could we tell* whether/when a machine acquired any of these capacities? What would persuade you it had happened? What if your best friend (or everyone else but you) turned out to have a machine brain?

Part of our class will focus on a special part of our mental lives: **making choices**.

4. If machines ever can have their own preferences and plans, would they be able to *make their own choices*? Would they have any (even shared) *responsibility* for the choices they make? Or would their actions just be results of their hardware, their programming, and the inputs they receive, and so not really "up to" the machine/under its control?

5. What about us with organic brains and bodies? Are our choices subject to the same physical laws that govern our bodies? Does that mean we're predetermined and unfree too?

Or is it possible to make free choices, and be responsible for them, compatibly with being programmed or physically determined to do so?