PHI 374. Philosophy of Mind: Artificial Intelligence

Instructor: Gui Sanches de Oliveira
Course level: upper level

Course description:

"Artificial Intelligence" (AI) means different things to different people. In some circles, AI refers to the project of understanding the mind by creating virtual or concrete machines that literally instantiate human intelligence in artificial media. In other circles, AI refers more simply to the project of designing and building mechanisms that imitate some aspects of human behavior with the goal of supporting humans in their everyday activities (e.g., finding relevant information or products online) without the ambition of artificially recreating the human mind. This course examines philosophical questions arising from both kinds of projects in AI, emphasizing, on the one hand, metaphysical assumptions about the nature of the mind and its relation to embodiment, and, on the other, moral and political implications of AI and the tension between assumptions of neutrality and issues relating to gender, race and class. [Required text: Margaret Boden (2018) *AI: Its Nature and Future*. OUP.]

Student Evaluation:

- Discussion Questions, one set per reading (25%)
- Concept Maps, one per unit (25%)
- Midterm and final papers (20+30%)

Schedule overview:

Weeks 1-2: Minds and Machines, an Overview
- selections from Margaret Boden (2018) *AI: Its Nature and Future*
- Gerd Gigerenzer & Daniel Goldstein (1996) "Mind as Computer: Birth of a Metaphor"

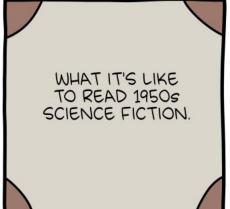Weeks 3-6: Classical Philosophical Critiques of Classical AI
- selections from Hubert Dreyfus (1972) *What Computers Can't Do*
- John Searle (1980) "Minds, brains, and programs"
- Daniel Dennett (1984) "Cognitive wheels: The frame problem of AI"
- Paul Churchland & Patricia Churchland (1990) "Could a Machine Think?"

Weeks 7-10: Distributed Computation, Artificial Life and Evolutionary Robotics
- Edwin Hutchins (2014) "The Cultural Ecosystem of Human Cognition"
- J. Bongard & M. Levin (2021) "Living Things Are Not (20th Century) Machines: Updating Mechanism Metaphors in Light of the Modern Science of Machine Behavior"

Week 11 to end: Artificial Intelligence and Real People
- Amy Kind (2021) "Computing Machinery and Sexual Difference: The Sexed Presuppositions Underlying the Turing Test"
- Abeba Birhane (2021) "The Impossibility of Automating Ambiguity"
- selection from Ruha Benjamin (2020) *Race After Technology: Abolitionist Tools for the New Jim Code*