CS 730/830: Intro AI

Philosophy and AI

1 handout: slides

Philosophy and AI

Arguments Against Weak AI

Can machines simulate intelligence? (Turing Test)

disability   be kind, resourceful, have initiative, sense of humor, judge right from wrong, make mistakes, fall in love, enjoy strawberries, learn, use words properly, do something new

incompleteness

informality   hard to formalize (similar to qualification problem): no finite set of rules

Arguments Against Strong AI

Can machines *really* think?

intentionality   are representations in machine about something real?

if simulated storms are not wet, why are simulated thoughts actual mental processes?

consciousness   can machine be aware of its own mental states and actions?

Turing: 'polite convention'

phenomenology   can a machine have subjective experience

The Mind-Body Problem

dualism Descartes (1641): separate realms that intersect. How?

gap: explaining brain function does not explain subjective experience

physicalism mental states are physical states. But is brain-in-a-vat experience a different mental state?

functionalism mental state is causal connection between input and output. But in neuron replacement experiment, does experience stay the same?

1. if output is same, artificial brain is conscious. then is a lookup table conscious.

2. if causation not crucial, artificial brain is not conscious. consciousness is epiphenomenal (not causally related to behavior).

3. experiment is impossible

Consciousness

- **naturalism:** consciousness caused by biological details in brain, not pure 'syntax'. eg, Chinese room.
- hard to measure!
- subjective experience (qualia). eg, inverted spectrum experiment.
- does it matter for making intelligent systems?

Ethics

- should AI be pursued? jobs, evil uses, danger
    autonomous weapons
    value alignment (eg, uncertainty in objective)
    transhumanism
- should robots be moral agents?
    Bryson: no

EOLQs

- What question didn't you get to ask today?
- What's still confusing?
- What would you like to hear more about?

Please write down your most pressing question about AI and put it in the box on your way out.

Thanks!