Consciousness & Integrated Information Theory
What is consciousness?
Agenda

- History
- Who is conscious?
- Consciousness and neuroscience
- Integrated Information Theory
Locke (1690). “Essay Concerning Human Understanding”:

“the perception of what passes in a man’s own mind”

Dictionary today:

“... awareness or perception of an inward psychological or spiritual fact: intuitively perceived knowledge of something in one’s inner self ...”
Who is conscious?

- I am (I think therefore I am)
- Other people
- Animals
- Plants?
- Computers?

Cogito ergo sum
- Descartes
Panpsychism and Artifact Consciousness

- One of the oldest philosophical theories: everything is conscious
- Ancient theme in mythology

Can machines think? ← Bad question! Let’s change it

Instead, let’s play a game!
The Imitation Game

Original game:

3 players: A (man), B (woman), C (either)

C tries to determine which is which, A tries to trick C while B tries to help C make the right choice

New game:

3 players: A (machine), B (human), C (human)

The Turing Test!

But it avoids the philosophical problem!
The Chinese Room Argument

Thought experiment formulated by John Searle in 1980

Differentiate between understanding and simulating (strong AI vs. weak AI)

- Closed room
- Book of rules
- Pen and paper
Consciousness and neuroscience

- Connections in the brain
- Meditation
- Psychedelic substances
- Dream states
The mind body problem

- Western thought
  - Mind/Soul
  - Abrahamic religions

- Eastern thought
  - Hinduism and Buddhism
  - Mind and body are one
  - Creeping into western world
Integrated Information Theory (IIT)

-An attempt to characterize consciousness mathematically, both in quantity and quality
Axioms
(self-evident truths) → Postulates
(unproven assumptions) → Theory of consciousness
5 axioms

- **Existence**
  *Consciousness exists.*

- **Composition**
  *Each experience consists of multiple aspects in various combinations.*

- **Information**
  *Each experience differs in its particular way from other possible experiences.*

- **Integration**
  *Each experience is strongly irreducible.*

- **Exclusion**
  *Each experience excludes all other experiences.*
A minimally conscious system

Experience of light ≥2
Integrated Information

Information that is generated by a mechanism above and beyond the information generated by its minimal parts.
On if majority of inputs = 1

On if even number of inputs = 1

On if all inputs = 1

On if majority of inputs = 1
On if majority of inputs = 1

On if even number of inputs = 1

On if all inputs = 1
Implications of IIT

- Framework for measuring consciousness
- Consciousness is universal
- Rethink our internal models of consciousness
- Is the internet conscious?
- More conscious when using computer?
Limitations of IIT

- Very hard to calculate
- Possible to test?
- Why consciousness exists?
- How did consciousness emerge in the first place?
Discussion points

- One of the oldest philosophical theories is "everything is conscious", is there any truth to this statement? How do you intuitively define consciousness?

- From an evolutionary perspective, what benefit for the survival of an organism flows from consciousness?

- IIT attempts to rigorously define consciousness. Do you think it is possible to define a limit (intelligence, behavior) of when an organism exhibits consciousness?

- As humans we consider ourselves conscious, but are there levels of consciousness even for us? Are there situations where you, according to your level of consciousness, are a vegetable?