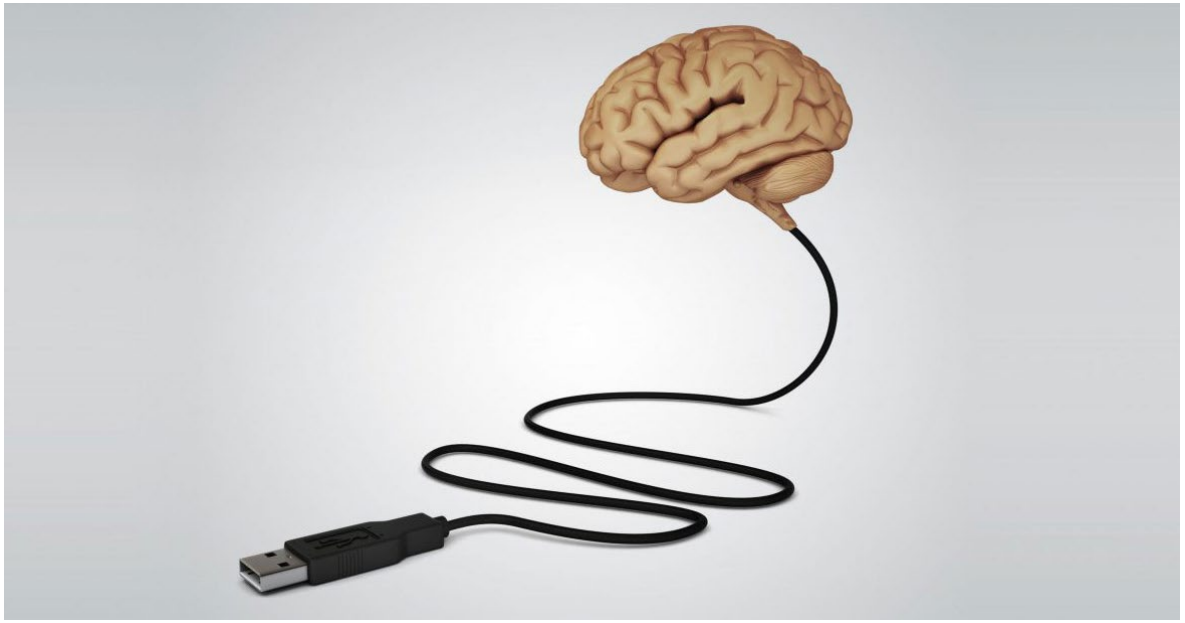


Can there be consciousness in a digital world? II

Chapman University. PHIL398. Lecture 16. 1/1/2021.
Kelvin J. McQueen

Reading schedule and 2nd assignment

- Schedule:
  - This week (week 8): ch15.

    Can there be *consciousness* in a digital world?
  - Next week (week 9): excursion into:

    Can there be *comprehension* in a digital world?

    Can there be *free will* in a digital world?

- 2nd short paper.
  - Due: Friday 11.59pm April 23.
  - Worth 15%.
  - Write a philosophical paper on a topic covered so far (distinct from 1st short paper). A bibliography is necessary and must include at least one reference (other than the Chalmers book). 600-800 words.

Discussion board (w8) posts due 4/3

- **Post 1**: Does the gradual uploading argument establish that there can be consciousness in a digital world? Explain your answer. 200-300 words.
- **Post 2**: Provide constructive feedback to a post on another student's thread. 150-250 words.

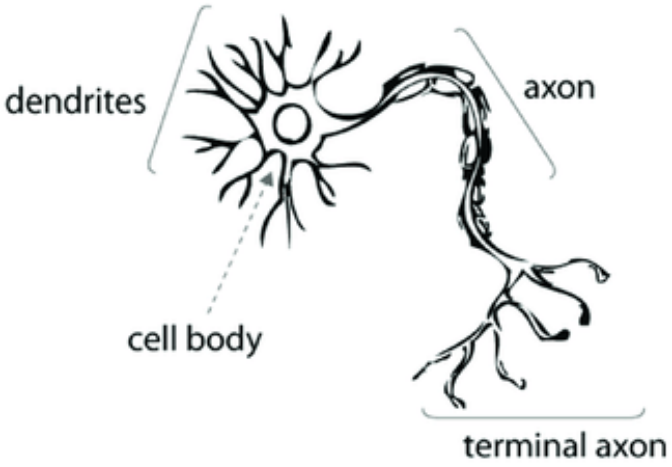
Assignment Rubric Details

Discussion Board Posts

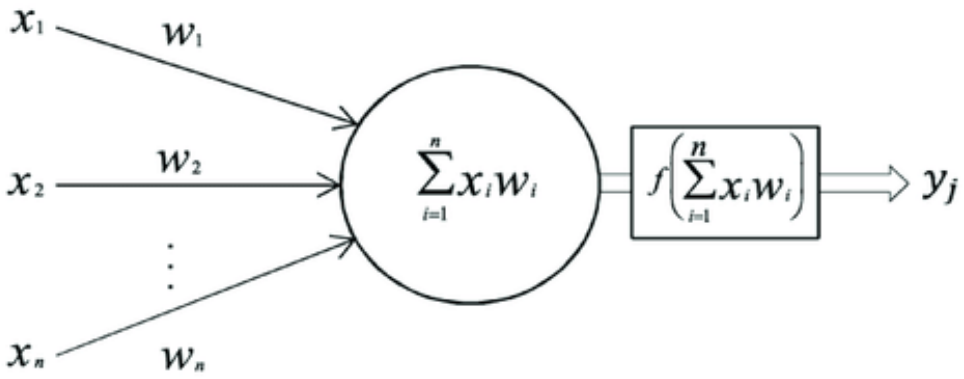
Criteria	Ratings			Pts
Post 1 Response to prompt	**5.0 pts** **Excellent** The post clearly answers the prompt, demonstrates understanding of the reading, and illustrates independent thinking.	**3.0 pts** **Adequate** The post attempts to answer the prompt, demonstrates partial understanding of the reading, but lacks independent thinking.	**1.0 pts** **Inadequate** Does not provide clear answer to the prompt and does not demonstrate understanding of the reading.	5.0 pts
Post 2 Constructive feedback	**5.0 pts** **Excellent** The post responds clearly to another student's post, offers constructive ideas, and is respectful.	**3.0 pts** **Adequate** The post attempts to respond to another student's post, but lacks either constructive ideas or respectful language.	**1.0 pts** **Inadequate** Does not respond to another student's post in way that demonstrates thoughtfulness.	5.0 pts

Total Points: 10.0

Biological versus artificial neurons



Biological (carbon-based) neuron



Artificial (silicon-based) neuron

Progress in building artificial brains

- The Blue Brain project, 2008.
  - Simulated 10,000 neurons connected by 10^8 synapses.
  - 4.76% of a rat brain.

- The K computer, 2013.
  - Simulation of 1.73 billion neurons connected by 10^{10} synapses.
  - 1% of a human brain.

Recap of main concepts

- Mind-uploading
  - The attempt to transfer our minds from biological brains to digital computers.
- Digital immortality
  - If the upload is me (*acts* like me, has a continued stream of *consciousness* with me), then I am immortal if the upload is immortal.
- Consciousness
  - A (mental) state is conscious iff there is something it is like to be in that state.
- Hard problem of consciousness
  - Structure and function alone do not entail consciousness.
- The problem of other minds
  - I cannot know one's consciousness from structure and function alone.
  - I only know it through *introspection* and then *behavioral/neural markers*.

The gradual uploading argument

- Example of gradual uploading:
  - Start with a brain you know to be conscious (e.g. your brain), call it B.
  - Consider a neuron-for-neuron artificial simulation of B, call it A.
  - To build A, we start by replacing *one* of B's neurons with an artificial replica.
  - Then, we do the same for another neuron, and so on.
  - Eventually, every biological neuron is replaced with an artificial neuron, and we will have built A.

- Argument:
  - There are three possibilities:
    1) *Magic-neuron hypothesis*: you lose consciousness all at once after the first replacement.
    2) *Fading consciousness hypothesis*: consciousness fades after a certain number of replacements.
    3) *Digital consciousness hypothesis*: consciousness remains intact!
  - Possibilities 1 and 2 are implausible:
    - Given functional equivalence of A and B, you cannot at any stage *notice* any change in consciousness, and you *report* that your consciousness remains intact.

Breakout rooms

Does the gradual uploading argument establish that there can be consciousness in a digital world?