Lecture 1: Computation and the Mind

Todd M. Gureckis
Department of Psychology
New York University
What makes a mind?

Debated for thousands of years. If you don’t have an immediate answer, don’t feel bad. Various proposals have been thrown around from by Plato, Buddha, Aristotle, Zoroaster, ancient Greek, Indian, and Islamic philosophers, and even a few folks at NYU.
What makes a mind?

What do they do?

Minds encompass our thoughts, which are the mental processes which allow us to deal with the world. These include not only explicit wishes, desires or intentions but unconscious processes as well.
What makes a mind?

Does MIND = BRAIN?

We know that we can’t have a mind or thoughts without a brain, but does that mean that minds and brain are synonymous?
In Ch. 1 Edelman uses the “slippery slope” argument to try to convince us that minds are not literally brains, but encompass anything that is organized as a set of represented mind states that accurately reflect aspect of the world.
Figure 1.1 — The BRAIN/MIND RIDDLE. What is common to the minds of various sentient creatures that look at the scene in the center of this picture and see three objects? This question can be elaborated (by asking it about two cylinders and a sphere rather than “three objects”), or extended to other cognitive processes such as thought or discourse that need not involve vision or any other particular perceptual modality.
Edelman’s arguments

- What is common to all sober observers viewing the same scene and who are in agreement about what is viewed?

- Can’t literally be neurons. My neurons are my own, and you can’t borrow them to solve your own problems.

- Well maybe is the the literal organization of the human nervous system (up to the limit of correspondence). However, we know (or at least believe) that cats have a very similar visual system and view the world much like we do. Is it the mammalian visual system? What about other animals?

- What about artificial systems formed of computers and video cameras that can accurately recognize the scene as well?

- The key to minds may be not the physical substrate in which they are embodied but the relations that various states of the system have to one another and to the environment/world.
The “organizational” view of the mind

- Minds aren’t human neurons or cat neurons or robot parts, but the organization of dynamic, continually evolving systems that relate ongoing internal (i.e., mind) states and external (i.e., world) states.

- Correspondences can be made between the evolution of two systems to describe what they are doing independent of the exact things they operate on.

- Such correspondences are particularly well described in the language of computation, simply because the THEORY OF COMPUTATION offers use formal insight into how ostensibly dissimilar systems can be formally identical.

- Everything that can be expressed in one system can be expressed in a different, but functionally identical system.
The “organizational” view of the mind

Minds are what brains do

Brains perform computations

Computation is the manipulation of representations through various processes
Representation and Process

- What is a representation?
  - Edelman uses the example of a calculator... there are operations you can perform (addition, multiplication) but someplace are are electronic representation of the operands (the numbers you are adding or multiplying)
  - These are **physical symbol systems** because a physical state (be it electronic or neural) represents some other entity
  - Whole branches of philosophy devoted to the idea of representation (semiotics)
Representation and Process

Very nice description comes from Markman’s (1999) “Knowledge Representation” textbook

- Representations have four components

1. A represented world: what the representations are about

2. A representing world: the domain that contains the representations

3. Representing rules: the things that relate the two above things (basically a map that draws correspondences between the represented and representing systems)

4. A process that uses the representations: The first three just make the potential for representation. Representations are inert unless some process makes use of them
Representations can be entirely abstract (in Edelman’s terms) because they don’t have to resemble the represented world in any particular way.

Even more impressively, they don’t have to represent the past or present, but can extend to predict the future!!

---

FIG. 1.2. Various ways of representing temperature. The top row depicts water that is frozen, at room temperature, and boiling. The next two rows depict possible analog representations. The two following rows show numerical temperature notations. Finally, the last row depicts temperature with the darkness of the square.
Kinds of Representations

Analog
The representing world has a structure and the form of the isomorphism is not arbitrary.

Symbolic
The rules relating numerical forms to temperature is arbitrary. There is not direct, physical correspondence.

FIG. 1.2. Various ways of representing temperature. The top row depicts water that is frozen, at room temperature, and boiling. The next two rows depict possible analog representations. The two following rows show numerical temperature notations. Finally, the last row depicts temperature with the darkness of the square.
FIG. 1.2. Various ways of representing temperature. The top row depicts water that is frozen, at room temperature, and boiling. The next two rows depict possible analog representations. The two following rows show numerical temperature notations. Finally, the last row depicts temperature with the darkness of the square.
2010

Arabic numerals

MMX

Roman numerals
explicit

Directly signifies or stands for something else (e.g., the temperature examples)

implicit

A representation that only makes sense through another representation (e.g., motor program to control arm only meaningful in relation to another system)
Where does meaning come in?

Most famously argued by Searle, if minds are just pushing around and manipulating symbols/representations where does the content/meaning come in?
Where does meaning come in?

Most reasonable answer is that meaning can come from the relation between elements within the same system. The mind is not literally a chinese room, but representations at one level or point introduces dependencies in other domains.
Who programs the mind’s computer?

- Computers today are clearly symbol manipulating machines.

- All representations ultimately are recoded into a set of numbers (binary numbers), but a variety of rich types of structured representations can exist on top of that (this presentation file, a webpage, a movie, a photo, etc...).

- The possibility of using numbers to represent basically anything is a first piece of the puzzle.
Who programs the mind’s computer?

• Next, we need a way to talk about what “computation” really is. I like Edelman’s phrase “a dogged adherence to a sequence of instructions each of which is explicit, simple, and unequivocal.”

• All of math can be reduced to a strict set of rules or algorithms that can be carried out without insight, problem solving, or really anything beside a pencil and a piece of paper.

• You should sense a hint of something interesting there just using your computer (which gains new capabilities in software all the time without radically changing the underlying computations).... *Computers 10 years ago were basically the same just slower!*
In computer science/mathematics, a key advance was made by Alan Turing and Alonzo Church (known as the Church-Turing thesis).

Simplest computational device that can, in principal, execute any program or procedure one can devise (in particular the “computable set” this excludes a couple very tricky “undecidable programs” who’s solution can never be guaranteed to end in finite time)
Turing Machines

- The ingredients of the Turing machine are quite simple

- A **tape** which is cut up into lots of little parts. Each cell can have a symbol (blank or something like a number) written to it. The tape infinite both left and right of the “head”

- A **head** which can read and write symbols from the tape and move left and right one cell at a time.

- A **finite table** of instructions that tell the head what to do (there are various specifications of the table but...):
  - Either erase of write a symbol and then, move the head to the left or right, and then either assume the same state or new state

- A **state register** that just stores the current internal state
So what?

- The Turing-Church hypothesis proves that the Turing machine has “universal computational” principals in that it can mimic any other Turing machine and can model any program that any other computer can.

- In Turing’s words “It is possible to invent a single machine which can be used to compute any computable sequence. If this machine U is supplied with a table on the beginning of which is written a [description] of some computer machine M, then U will computer the same sequence as M.”

- This makes more concrete some of our intuitive arguments for the relationship between minds across physical substrates: ASSUMING minds are computational devices, the implementation details are not strictly important... might as well just be coded as a program in a Turing machine!!! If you like rigor in your philosophy, this might be the thing to convince you.
To Summarize

• We can talk about minds without reference to any particular brain or biological substrate... this seems a logical necessity of any world-view that doesn’t hold human cognition as a singularity in the universe

• One of the important functions of the mind can be understood as providing representations of the world and the relationship between successive mental “states”

• We talked about some of the important properties of representations

• We dismissed simple arguments holding “meaning” or “intentionality” as refutation of the power of symbol-manipulating machinery

• We discovered why Turing is so famous and gathered more support for point 1.
Where to next?

- The BRAIN!
- So far the idea of computation has been quite abstract. We might be able to envision some idea or metaphor of our computer doing some kind of computation, but remember,

\[
\begin{align*}
\text{Minds are what brains do} \\
\text{Brains perform computations} \\
\text{Computation is the manipulation of representations through various processes}
\end{align*}
\]