CSE 40171:
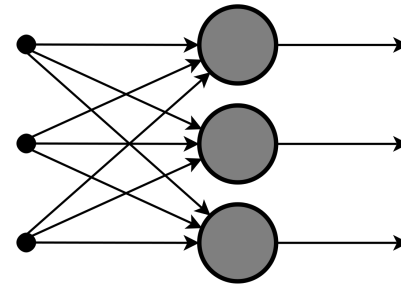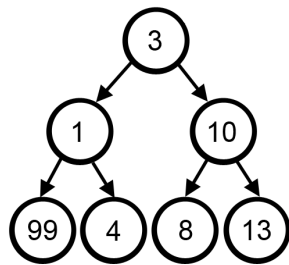# Artificial Intelligence

Introduction: Philosophical Perspectives

Fiona's Office Hours:

Mondays 12-1pm

Duncan Innovation Lounge

The First AI Winter (the second half of the 1970s)

What went wrong?

- Computing resources were meager

- Better understanding of intractability

- Hard to collect data pre-Internet

- Moravec's paradox

- Limits to logical inference

Lighthill Report on Artificial Intelligence (1973)

"Achievements… have to be judged in competition with… perfectly conventional methods of control engineering and data processing.

There is a consensus… that benefits from this work… have flowed primarily to the science of psychology…

[T]o judge the work on its own rules and its own aims [is] to conclude that the attempt to… [b]ridge [automation and cognitive science] is not succeeding."
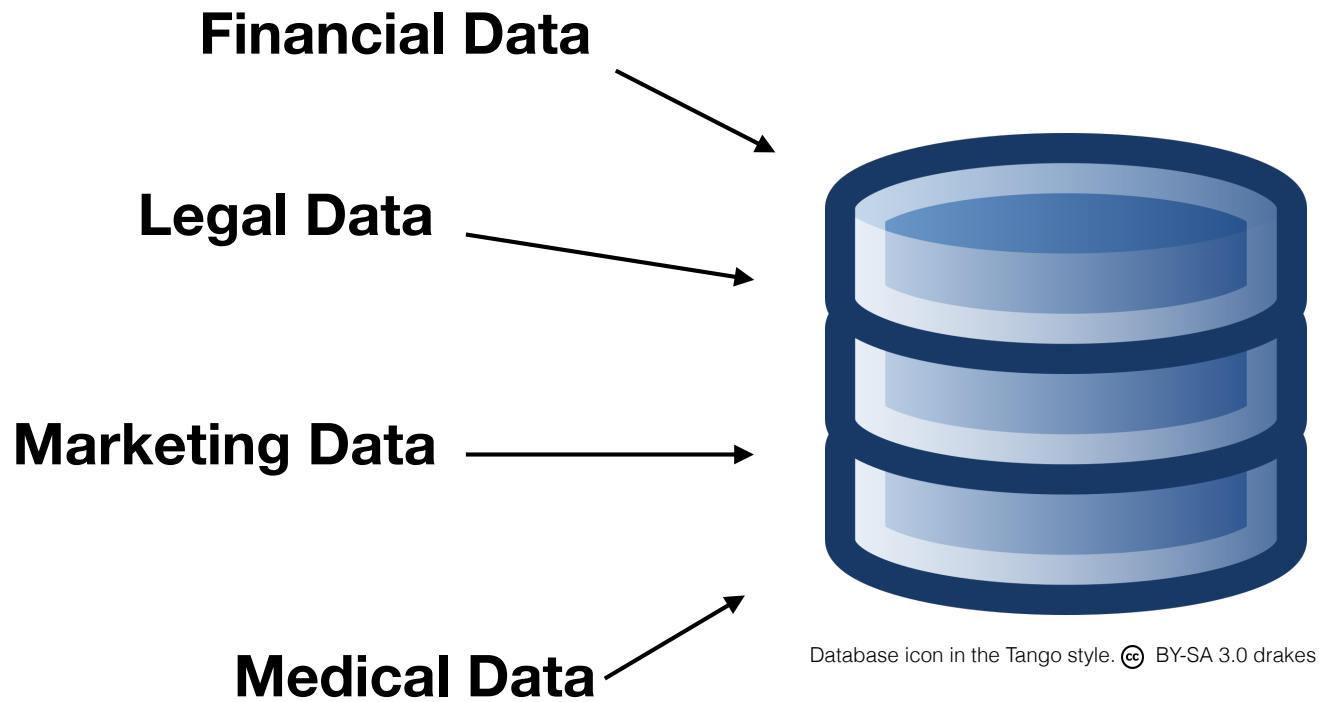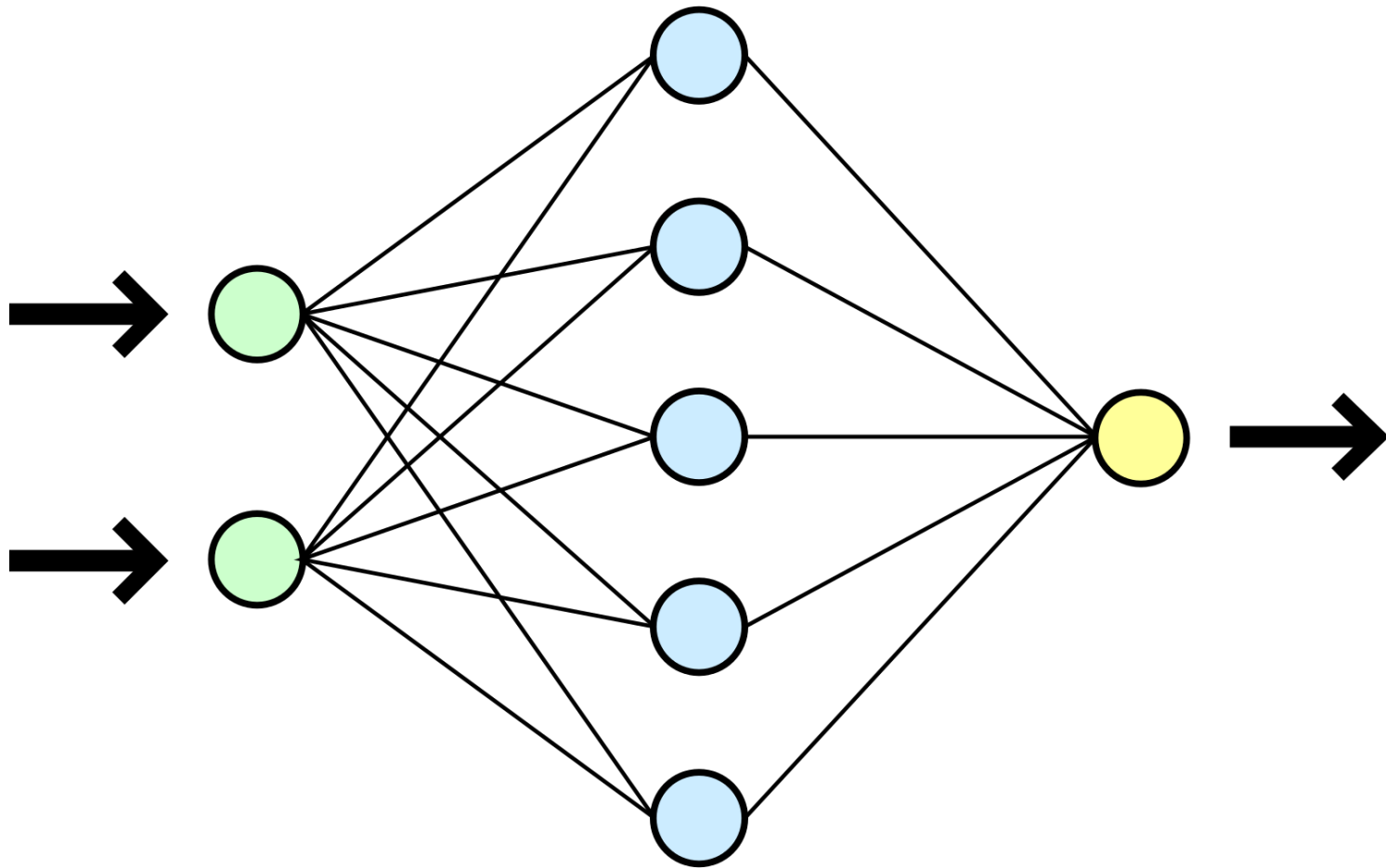
The 1980s:
The second AI boom

Expert Systems

A Symbolics Lisp Machine: an early platform for expert systems
BY-SA 3.0 Michael L. Umbricht and Carl R. Friend

Knowledge Bases

Financial Data

Legal Data

Marketing Data

Database icon in the Tango style. (cc) BY-SA 3.0 drakes

Medical Data

Artificial Neural Networks

A simplified view of an artificial neural network (cc) BY 1.0 Mysid

The Second AI Winter (arguably lasting until 2012)

- The field experienced an economic bubble

- Early expert systems were hard to update, could not learn, and were brittle in operation

- Pushback from the robotics community

The 90s…

Image Credit: "An example of an automated online assistant," Bemidji State University

Probabilistic Reasoning
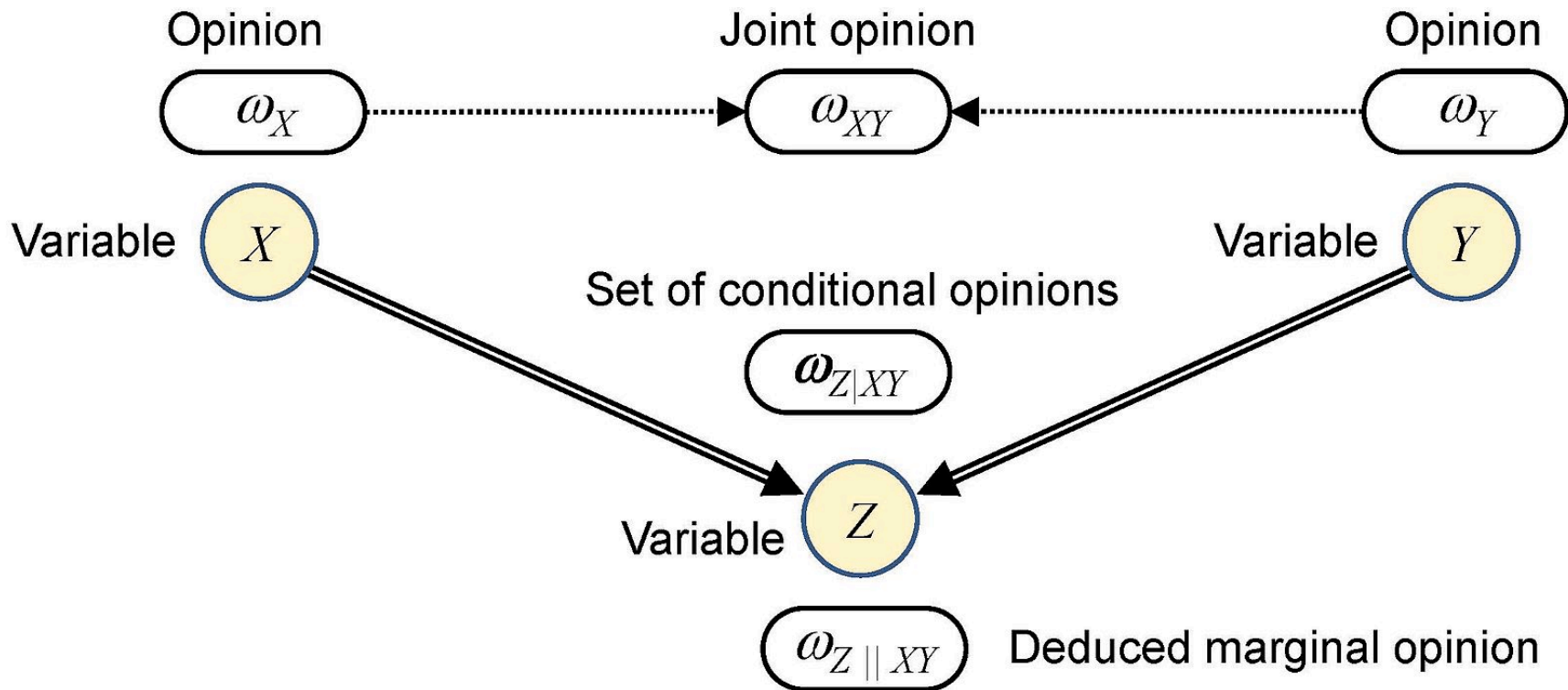
The Current AI Renaissance

Machine Learning
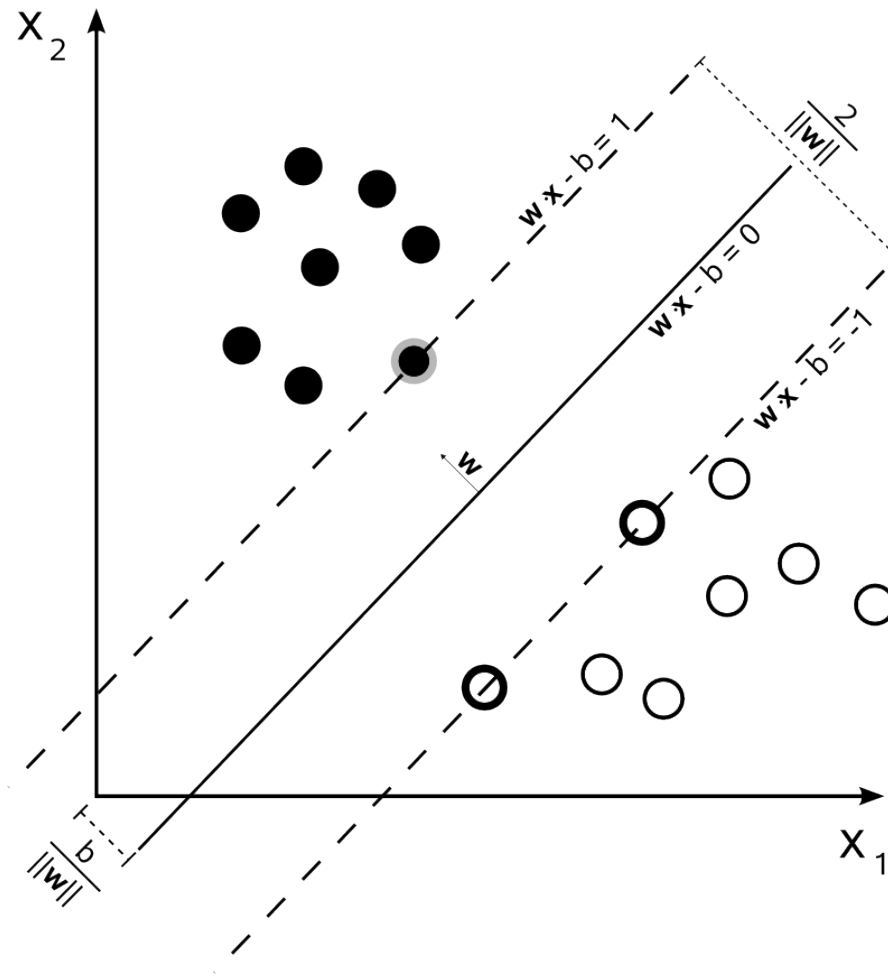
Enabling Factors

1. Web-scale Data

Image Credit: Andrej Karpathy, https://goo.gl/RehWdv

2. Cheap Parallel Hardware

AlexNet (2012)
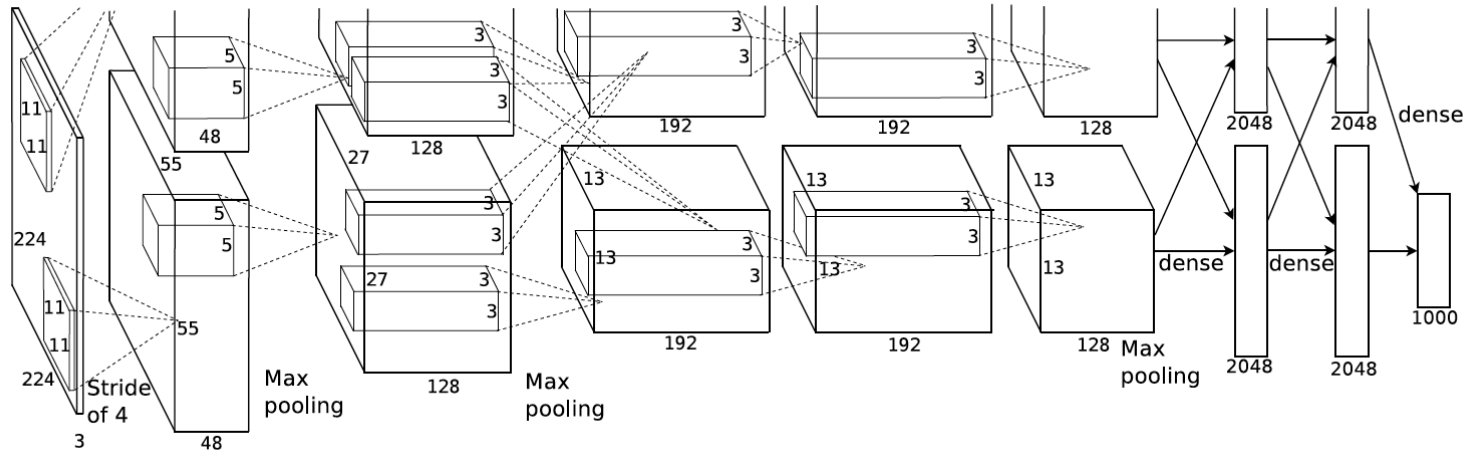
Model	Top-1 (val)	Top-5 (val)	Top-5 (test)
SIFT + FVs [7]	—	—	26.2%
1 CNN	40.7%	18.2%	—
5 CNNs	38.1%	16.4%	**16.4%**
1 CNN*	39.0%	16.6%	—
7 CNNs*	36.7%	15.4%	**15.3%**

ImageNet 2012
Object Classification Results

Krizhevsky et al. NIPS 2012

Macroanalysis: An Extension of NLP
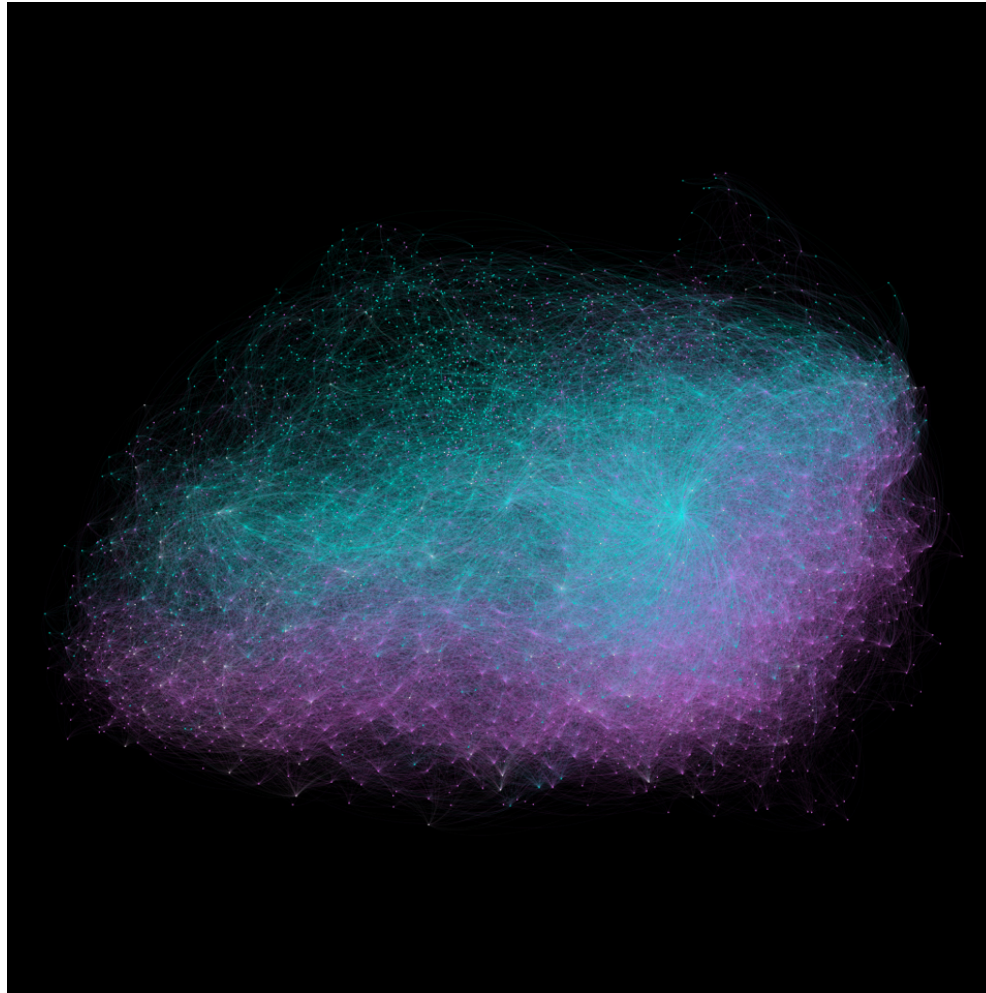
Texts organized by stylistic features;
colors indicate gender of writer

M. Jockers, Macroanalysis: Digital Methods and Literary History.

Neural Style Transfer

Original Image   **Target Style**   **Original w/ New Style**

Predictive Coding Networks
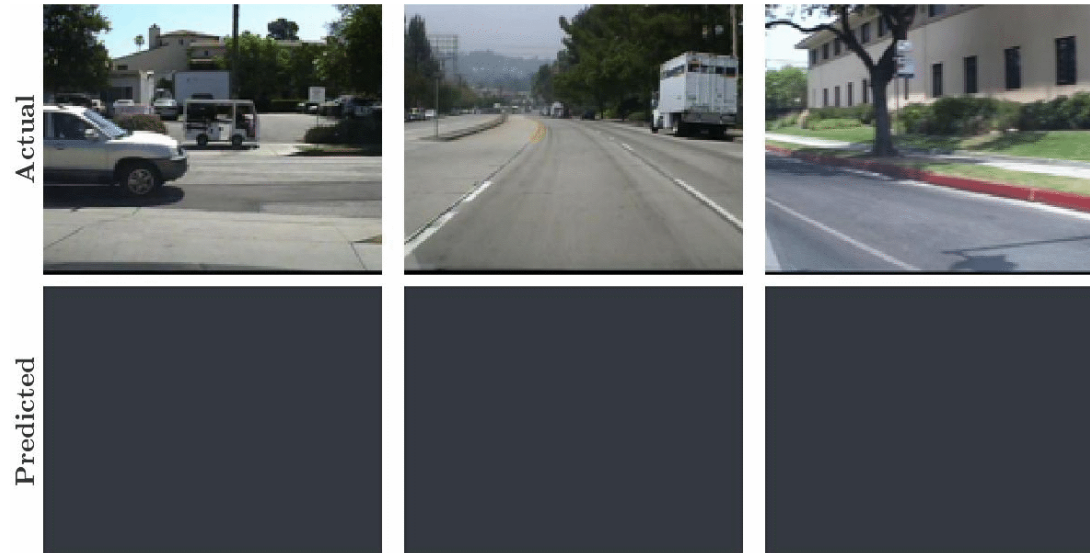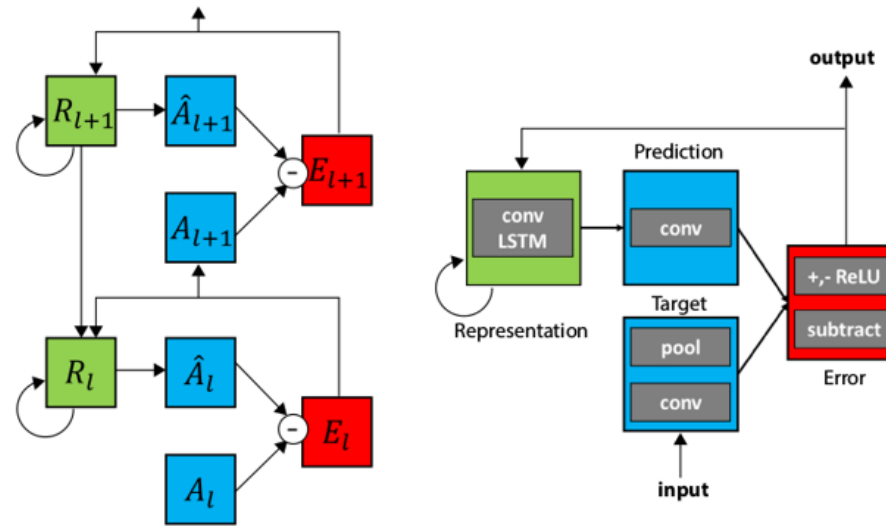
Models vs. Brains
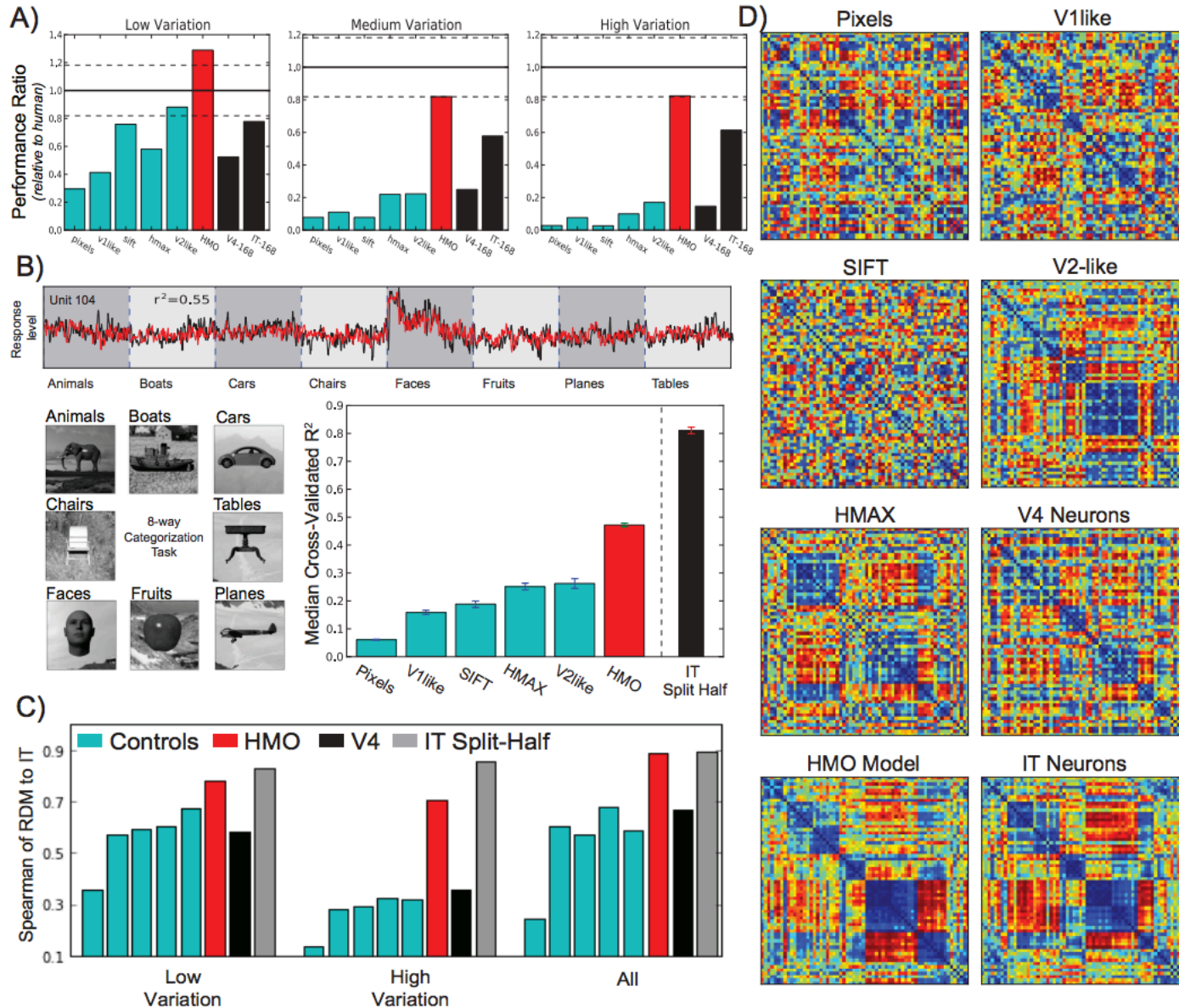
How did we get here?

Plato

"I think I am. I shouldn't like to take my oath on the whole story, but one thing I am ready to fight for as long as I can, in word and act—that is, that we shall be better, braver, and more active men if we believe it right to look for what we don't know…" (Meno 86b)
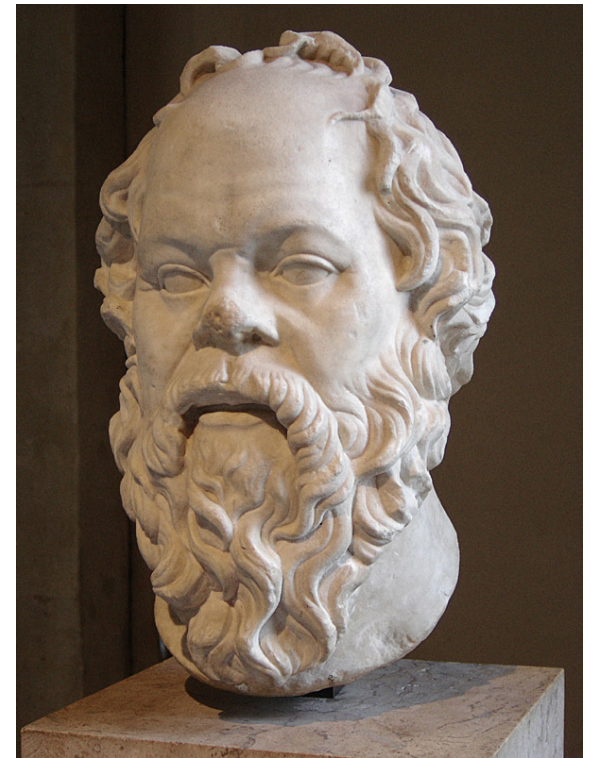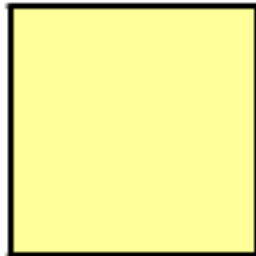
Portrait of Socrates (cc) BY-SA 2.5 Eric Gaba

Aristotle

"But how does it happen that thinking is sometimes accompanied by action and sometimes not…

What I need, I have to make; I need a cloak. I have to make a cloak. And the conclusion, the 'I have to make a cloak,' is an action." (*De Motu Animalium* 701 6-19)

Propositional Logic (3rd century BC)

$$Sentence \rightarrow AtomicSentence \mid Complex\ Sentence$$

$$AtomicSentence \rightarrow True \mid False \mid P \mid Q \mid R$$

$$ComplexSentence \rightarrow (\ Sentence\ ) \mid [\ Sentence\ ]$$

$$\mid \quad \neg\ Sentence$$

$$\mid \quad Sentence \wedge Sentence$$

$$\mid \quad Sentence \vee Sentence$$

$$\mid \quad Sentence \Rightarrow Sentence$$

$$\mid \quad Sentence \Longleftrightarrow Sentence$$

OPERATOR PRECEDENCE : $\neg, \wedge, \vee, \Rightarrow, \Longleftrightarrow$

Predicate Logic

$$Sentence \rightarrow AtomicSentence \mid Complex\ Sentence$$

$$AtomicSentence \rightarrow Predicate \mid Predicate\ (Term,...) \mid Term = Term$$

$$ComplexSentence \rightarrow (\ Sentence\ ) \mid [\ Sentence\ ]$$

$$\mid \quad \neg\ Sentence$$

$$\mid \quad Sentence \wedge Sentence$$

$$\mid \quad Sentence \vee Sentence$$

$$\mid \quad Sentence \Rightarrow Sentence$$

$$\mid \quad Sentence \Longleftrightarrow Sentence$$

$$\mid \quad Quantifier\ Variable,\ ...\ Sentence$$

$$Term \rightarrow Function\ (Term,\ ...)$$

$$\mid\ Constant$$

$$\mid\ Variable$$

$$Quantifier \rightarrow \forall \mid \exists$$

$$Constant \rightarrow A \mid X_1 \mid John \mid ...$$

$$Variable \rightarrow a \mid x \mid s \mid ...$$

$$Predicate \rightarrow True \mid False \mid After \mid Loves \mid Raining \mid ...$$

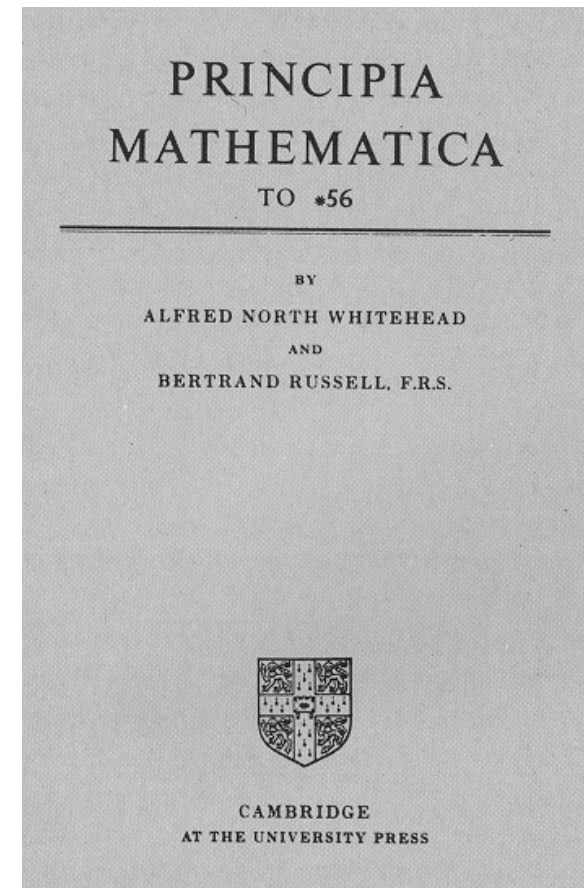$$Function \rightarrow Mother \mid LeftLeg \mid ...$$

OPERATOR PRECEDENCE : $\neg, =, \wedge, \vee, \Rightarrow, \Longleftrightarrow$

Principia Mathematica

Whitehead and Russell, 1910-1913

Goal: A set of axioms and inference rules in symbolic logic from which all mathematical truths could be proven

PRINCIPIA
MATHEMATICA
TO *56

BY

ALFRED NORTH WHITEHEAD
AND
BERTRAND RUSSELL, F.R.S.

CAMBRIDGE
AT THE UNIVERSITY PRESS

# Gödel's Incompleteness Theorems

Assume a consistent formal system F

First Incompleteness Theorem: there are statements of the language of F which can neither be proved nor disproved in F.

Second Incompleteness Theorem: an extension of the first, which shows that the system cannot demonstrate its own consistency.

Symbol Manipulation

- In spite of logical and conceptual problems, the idea of intelligence as direct symbol manipulation persisted

  - Newell and Simon: Logic Theorist (1955)

  - McCarthy: Advice Taker (1958)

- Critique from Dreyfus: human beings rarely invoke logic to solve problems

Dualism
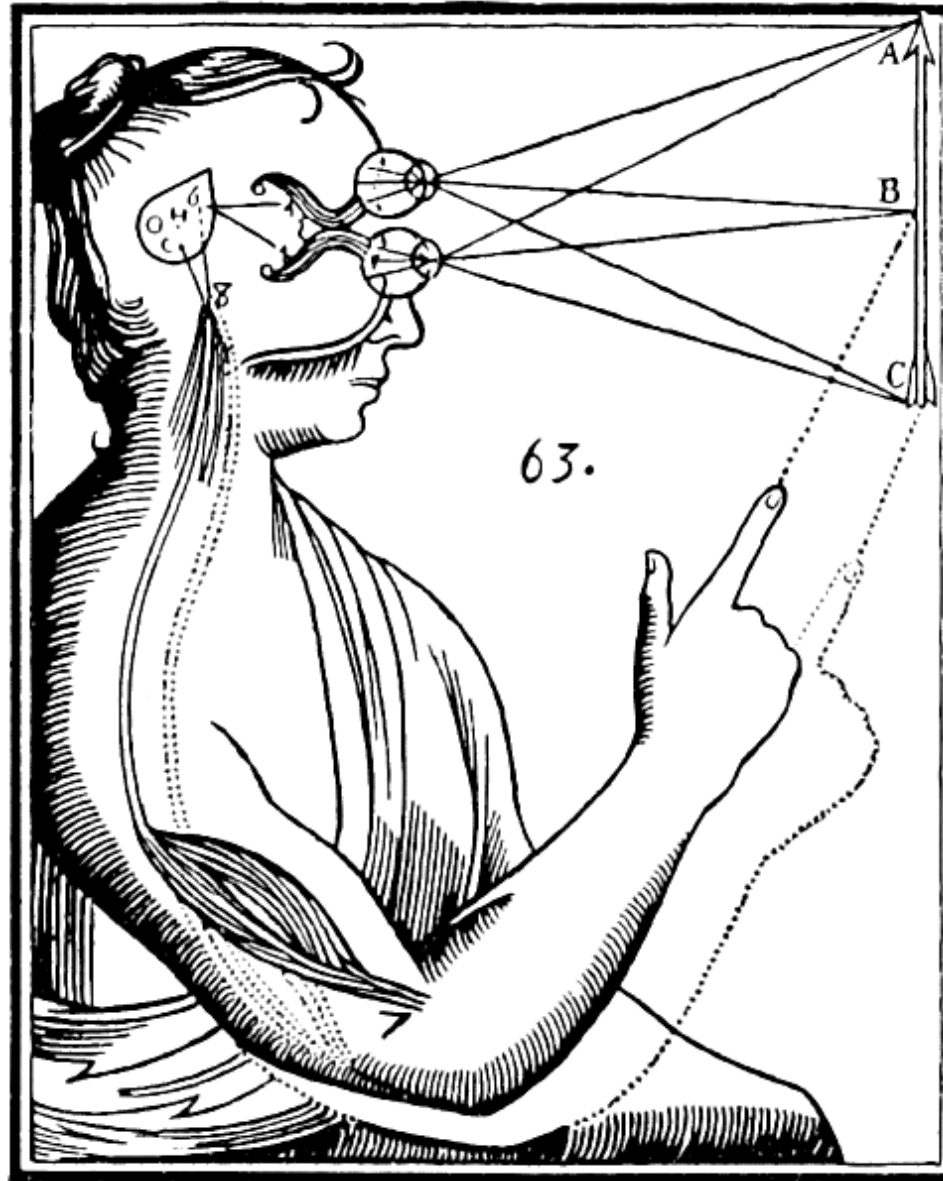
Rene Descartes

The mind and body are distinct, but closely joined

"Nature also teaches me, by the sensations of pain, hunger, thirst and so on, that I am not merely present in my body as a pilot in his ship, but that I am very closely joined and, as it were, intermingled with it, so that I and the body form a unit." (Med. 6)
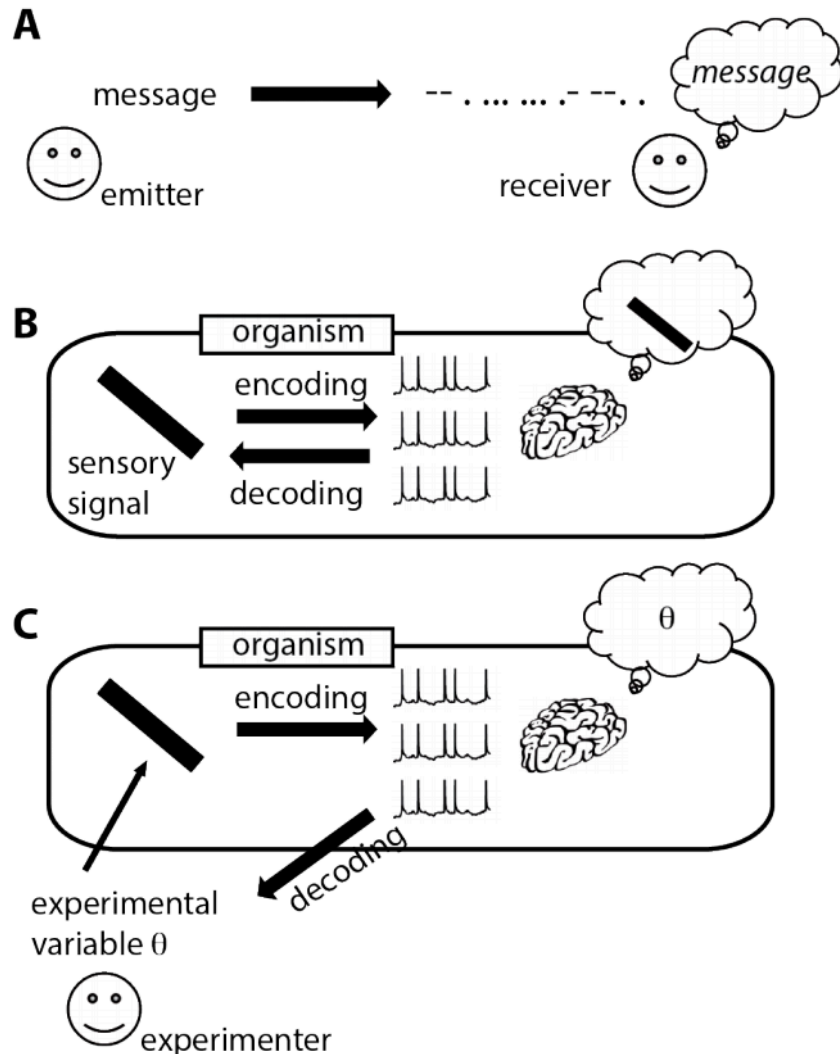
Scientific Consensus

- The brain is a computer

- The brain takes input from the sensory systems and processes it through the application of a series of non-linear functions

- It is possible to simulate neural computation on a digital computer

The problem: we can only measure second-order effects

Is coding a relevant metaphor for the brain?

• The neural coding metaphor is used broadly, in ways that are less obviously related to communication problems

• Is perception a communication problem?
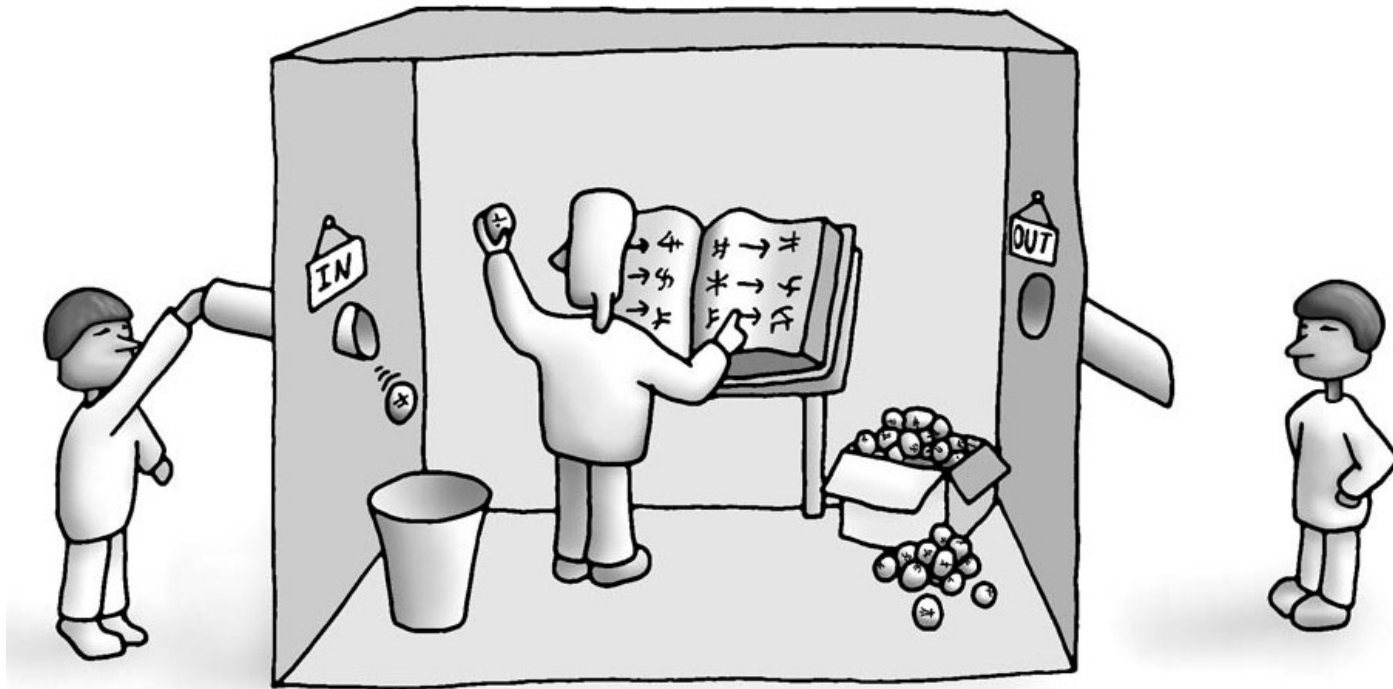
• And if so, with whom?

If we had a black box algorithm that claimed to be AI, how could we verify the claim?

Turing Test

The Chinese Room

Searle. Behavioral and Brain Sciences 1980
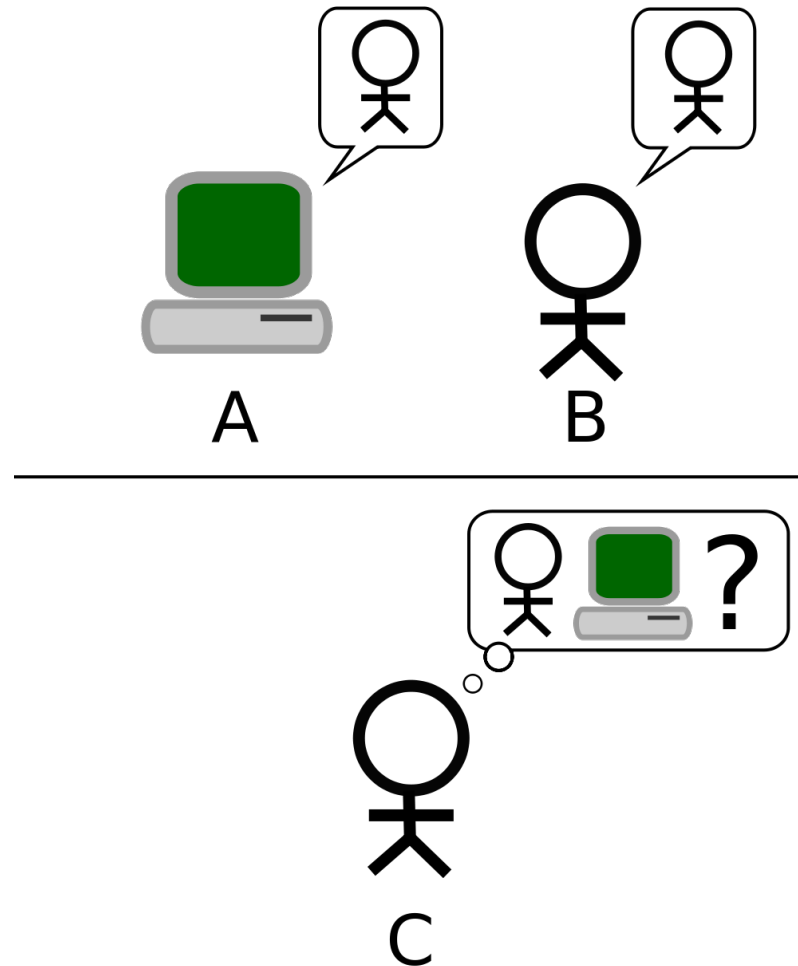
Key Question: Is the brain a Turing Machine?

Key Question: Can the brain even understand itself?