Philosophy and Psychology (UN3654): Computational Models of Mind

Instructors
Raphael Gerraty, rtg2116@columbia.edu
John Morrison, jmorrison@barnard.edu

Teaching Assistant
Andrew Richmond, ar3688@columbia.edu

Time and Location
Mondays and Wednesdays, 2:40pm-3:55pm
Location, Milbank 405

Office Hours
Raphael Gerraty, Monday, 4:00-5:00pm, and by appointment, Milbank 326b
John Morrison, Wednesday 4:00-5:00pm, and by appointment, Milbank 326b
Andrew Richmond, Tuesday, 2:00-4:00pm, Milbank 326b

Description
Computational models of perception, memory, decision-making, and motor control seem to give us important insights into the minds of human and non-human animals. According to many of these models, an animal perceives, remembers, decides, and moves as the result of computations involving representations of its environment. In this course, we will carefully consider the philosophical foundations of these models. In particular, we will focus on five questions:

- Which animals are capable of representing their environment?
- Which of their states are representations?
- What do those states represent?
- Why do they represent that?
- Which computations do they implement?

We will then discuss deep neural networks and probabilistic inference, two influential but highly controversial computational models.

Assessments

1. One five-page paper due by spring break (3/13) (20%)
2. One ten-page paper due by the end of the semester (5/4) (40%)
3. Final exam during finals week (20%)
4. Eight problem sets due throughout the semester (20%)

Problem sets will be distributed the day after the relevant lecture, and are due two weeks after they are distributed.

Schedule

Introduction

1/22 Course Overview
- Bison's "Alien/Nation"

1/27 Background
- Allen Newell and Herbert Simon's “Human Problem Solving”

1/29 Background

- David Marr, *Vision*, Chapter 1
- James Ward, *The Student's Guide to Cognitive Neuroscience* pp. 15-36
  - Problem set on Turing machines

<u>What do the states represent? And why?</u>

2/3 Causation: Simple causal accounts

- David Hubel and Torsten Weisel's "Receptive Fields of Single Neurons in the Cat's Striate Cortex"

2/5 Causation: Information Theory

- Randy Gallistel, *Memory and the Computational Brain*, Ch 1
  - Problem set on information theory

2/10 Causation: Sophisticated causal accounts

- Fred Dretske, *Knowledge and the Flow of Information*, Ch 3

2/12 Causation: Encoding and decoding accounts

- Nicole C. Rust's "Population-Based Representations: From Implicit to Explicit"
- J. Brendan Ritchie et al. "Decoding the Brain: Neural Representation and the Limits of Multivariate Pattern Analysis in Cognitive Neuroscience"

2/17 Selection: Selected Functions

- Philip Kitcher's "Function and Design"

2/19 Selection: Millikan

- Ruth Millikan's "Biosemantics"

2/24 Selection: Neander

- Karen Neander's "Towards an Informational Teleosemantics"

2/26 Explanation: Burge

- Tyler Burge's "Origins of Perception"
- Christopher Peacocke's "Perception, Biology, Action, and Knowledge"

3/2 Isomorphism: Cognitive Maps

- Edward Tolman's "Cognitive maps in rats and men"
- Edvard Moser, Emilio Kropff, and May-Britt Moser's "Place Cells, Grid Cells, and the Brain's Spatial Representation System"

3/4 Isomorphism: Cummins

- Gerard O'Brien and Jon Opie's "Notes Toward a Structuralist Theory of Mental Representation"

<u>Which systems represent?</u>

3/9 Learning

- Fred Dretske's "The Explanatory Role of Information"
- Yael Niv's "Reinforcement Learning in the Brain" (through Section 2.2)
- Problem set on learning

3/11 Means-Ends Reasoning

- Peter Carruthers's "On Being Simple Minded"

3/23 Dynamical Systems

- Shenoy et al's "Cortical Control of Arm Movements: A Dynamical Systems Perspective"
- Andy Clark, *Mindware*, Ch 7
  - Problem set on dynamical systems

3/25 Embodied Cognition

- Andy Clark, *Being There*, Introduction
- Ruth Millikan's "Pushmi-pullyu Representations"

3/30 Instrumentalism

- Daniel Dennett's "Intentional Systems"
- Lyn Rudder-Baker's "Instrumental intentionality"

Which inner states are representations?

4/1 Symbols and Languages

- Fodor's "Why There *Still* has to be a Language of Thought**"**

Which computations does it implement?

4/6 Implementation

- Gualtiero Piccinini, *Physical Computation*, Ch 2-3

Applications

4/8 Neural Networks

- Geoffrey Hinton, 10:00-25:42 of https://www.youtube.com/watch?v=VsnQf7exv5I
- Andy Clark, *Mindware*, Ch 4
  - Problem set on neural networks

4/13 Neural Networks

- Churchland and Sejnowski's "Neural Representation and Neural Computation "

4/15 Neural Networks

- Fodor and Pylyshyn's "Connectionism and Cognitive Architecture"

4/20 Neural Networks

- Daniel Yamins and James DiCarlo's "Using goal-driven deep learning models to understand sensory cortex"
- Hinton video, 25:42-34:22, https://www.youtube.com/watch?v=VsnQf7exv5I

4/22 Probabilistic Inference

- Chater, Tenenbaum, & Yuille "Probabilistic models of cognition: Conceptual foundations"
- McElreath's *Statistical Rethinking*, Ch 2
  - Problem set on probabilistic inference

4/27 Probabilistic Inference

- Weiji Ma, Konrad Kording, Daniel Goldreich's *Bayesian Modeling of Behavior*, Chapters 1-2
- Dobromir Rahnev's "The Bayesian brain: What is it and do humans have it?"
  - Problem set on Bayesian models

4/29 Neural Networks

- Brenden Lake, Tomer Ullman, Joshua Tenenbaum, and Samuel Gershman's "Building Machines That Learn and Think Like People," Sections 1-4

5/4 Probabilistic Inference

- Andy Clark's "Whatever next? Predictive brains, situated agents, and the future of cognitive science"
  - Problem set on predictive coding