General Principles of Human and Machine Learning

Lecture 1: Introduction

Dr. Charley Wu
Dr. Charline Tessereau

https://hmc-lab.com/GPHML.html

Overview

- Organization
  - Contact information and office hours
  - Introductions
  - Course organization
  - Grading
  - Schedule
- What is learning?

Course & Contact Info


Charley


Charline

Instructors
Dr. Charley Wu (charley.wu@uni-tuebingen.de)
Dr. Charline Tessereau (charline.tessereau@internationalbrainlab.org)

Teaching Assistants
Hanqi Zhou (hanqi.zhou@uni-tuebingen.de)
Turan Orujlu (turan.orujlu@tuebingen.mpg.de)
Mani Hamidi (manihamidi@gmail.com)


Hanqi


Turan


Mani

General information
Location**: Ground floor seminar room, AI building, Maria-von-Linden-Str. 6, D-72076 Tübingen**
Lecture time: **10:15** - 12:00 Thursdays
Tutorial time: **16:15** - **17:30** Fridays
Office Hours: Charley directly after lectures; Charline by appointment (email)
Course website: https://hmc-lab.com/GPHML.html

Introductions

- What is your name?
- What do you study?
- What do you hope to learn from this course?
- [Bonus] Name each of the people prior to you

Course organization

Lectures

- Read assigned paper
- Show up to class and participate in discussion

Tutorials

- Combination of hands-on exercises, (paper) discussions, programing challenges, and pop-quizzes (see Grading on next slide)
- Student responsibilities:
  - Keep up with material (complete assigned readings, re-visit lecture slides, visit office hours, ask TAs)
  - Show up and participate

Grading

- [**30% of grade**] Best 3 out of 4 pop-quizzes
  - They are designed to make sure you are following the material and are relatively easy marks
  - If you are unable to attend any tutorials, please email both instructors 24 hrs in advance (or as early as possible)
  - If you have well-documented absences, we may consider make-up quizzes or alternative solutions
- [**70% of grade**] Final exam
  - Tentative dates: July 27th and October 12th
  - Questions will be a combination of multiple choice and short answer questions

Course overview

What are the guiding principles of human and machine learning?

How have these two fields informed one another?

Which mechanisms of learning are shared across fields?

Where have we seen convergence?

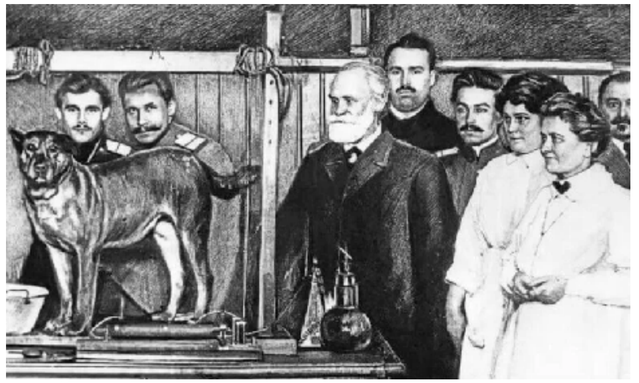
Syllabus

Date	Lecturer	Topic
Week 1: April 20,21	Wu/Tessereau	Introduction
Week 2: April 27, 28	Wu	Origins of biological and artificial learning: Behaviorism and Connectionism
Week 3: May 4, 5	Wu	Cognitive maps: Learning a representation of the environment
Week 4: May 11,12	Tessereau	Introduction to RL: Model-free vs. model-based learning and biological realism
Week 5: No classes		*Christihimmelfahrt: No lecture, no tutorial*
Week 6: May 25, 26	Tessereau	Advances in RL: Modern implementations and recent breakthroughs
Week 7: No classes		*Pfinstpause: No lecture, no tutorial*
Week 8: No classes		*Fronleichnam: No lecture, no tutorial*
Week 9: June 15,16	Wu	Concepts and Categories: Representations Learning in Humans
Week 10: June 22,23	Wu	Supervised and Unsupervised learning: Representation Learning in AI
Week 11: June 29,30	Wu	Generalization: Predicting in Novel Situations
Week 12: July 6,7	Tessereau	Common tools for understanding brains and neural networks
Week 13: July 13,14	Wu	Language and semantics
Week 14: July 20,21	Wu/Tessereau	General Principles (overview)

Origins of Biological and Artificial Learning

Behavioralism

- Understanding intelligence through behavior
- Trial and error learning
- Classical and operant conditioning
- Rescorla-Wagner model as proto-RL



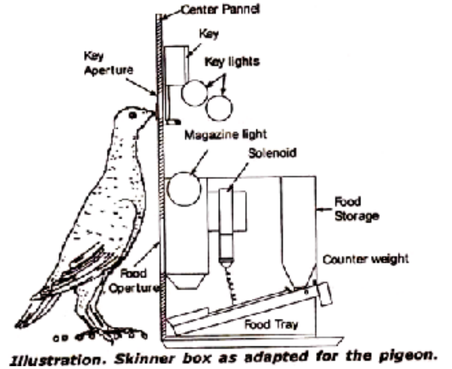
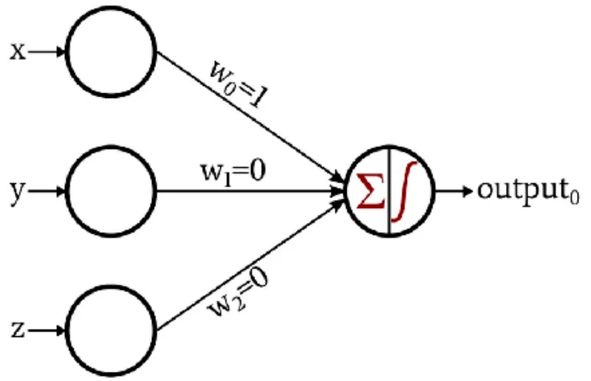


Illustration. Skinner box as adapted for the pigeon.

Connectionism

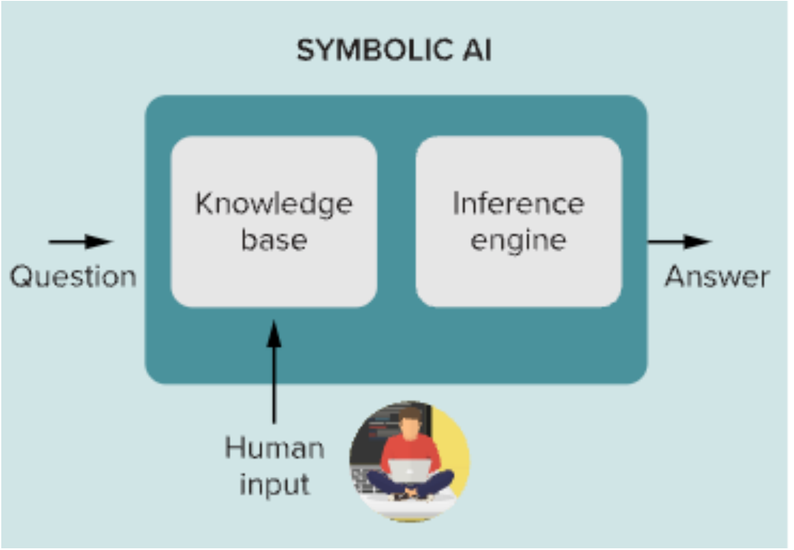
- Understanding intelligence through artificial neural networks
- Perceptrons, logical operators, gradient descent, and backpropagation



Symbolic AI and Cognitive Maps

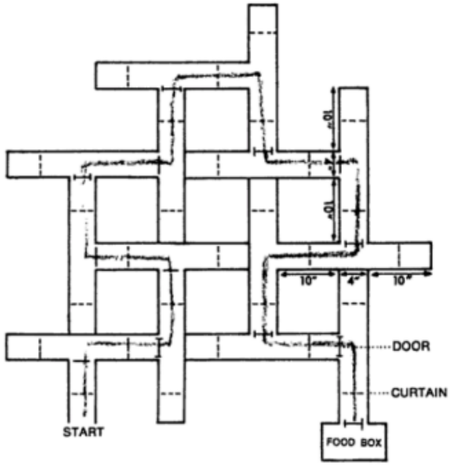
Symbolic AI

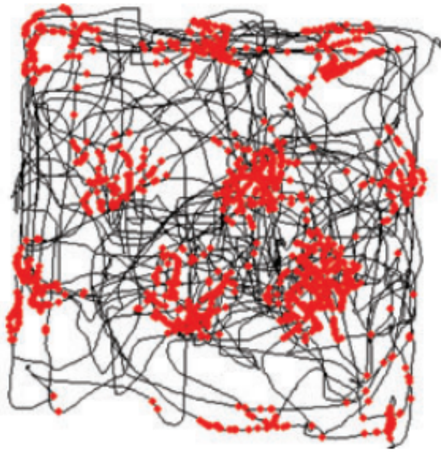
- What happened during the AI winter?
- Intelligence as manipulating symbols through rules and logical operations
- Learning as search



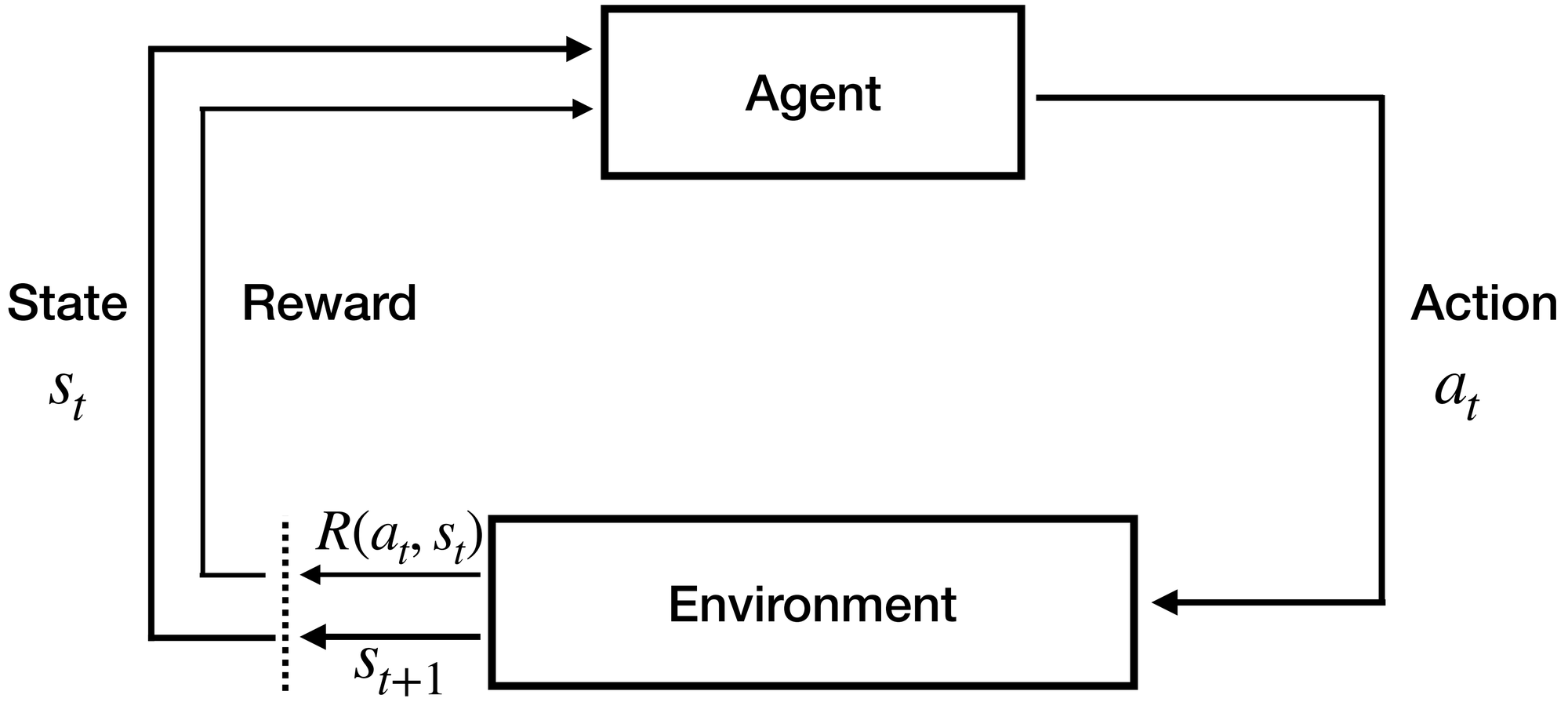
Cognitive Maps

- From Stimulus-Response learning to Stimulus-Stimulus learning
- Constructing a mental representation of the environment
- Neurological evidence for cognitive maps in the brain





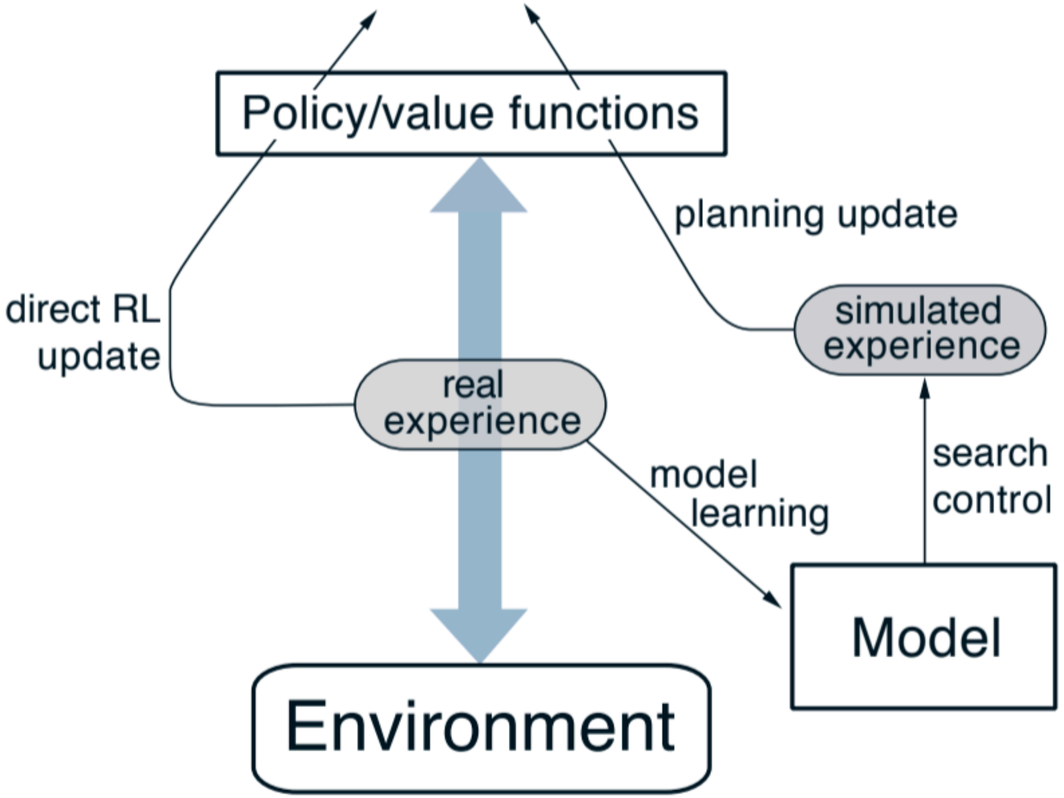
Introduction to RL



Sutton & Barto (1998)

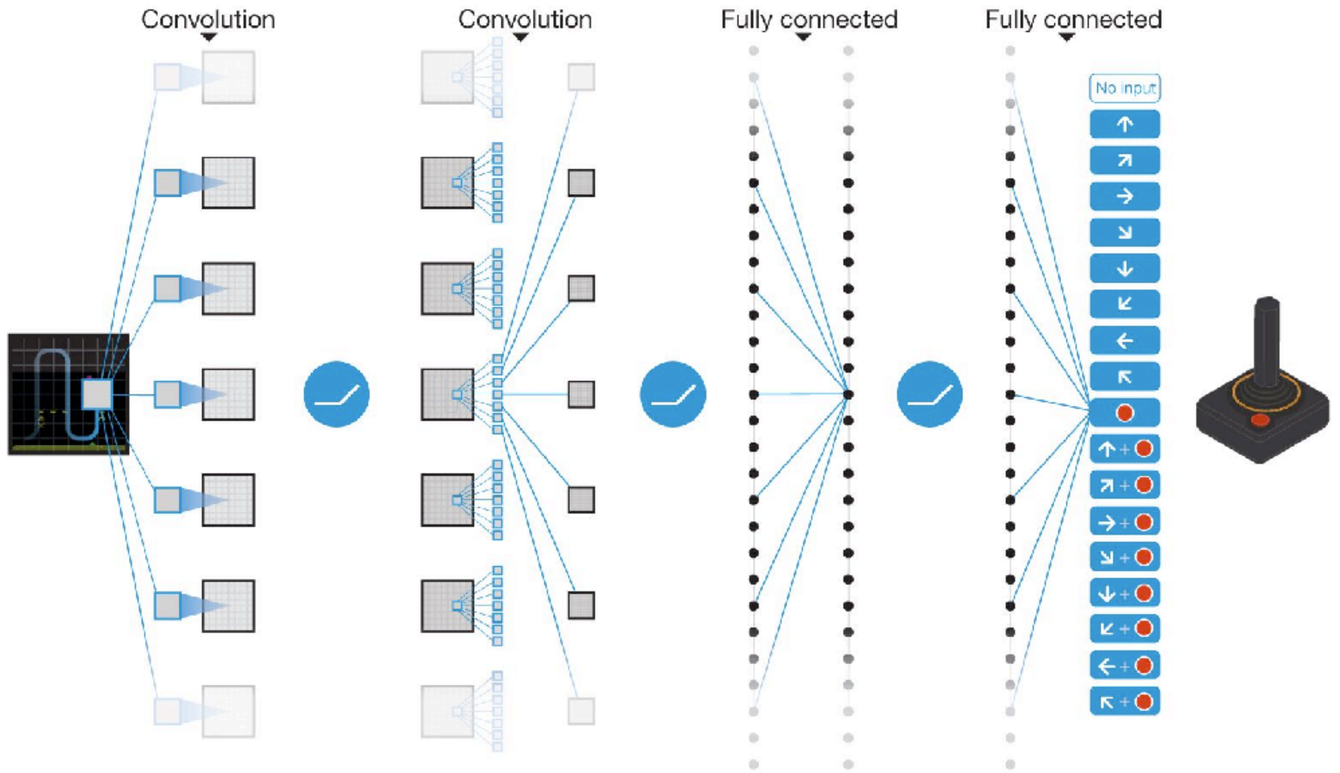
Advances in RL

Dyna



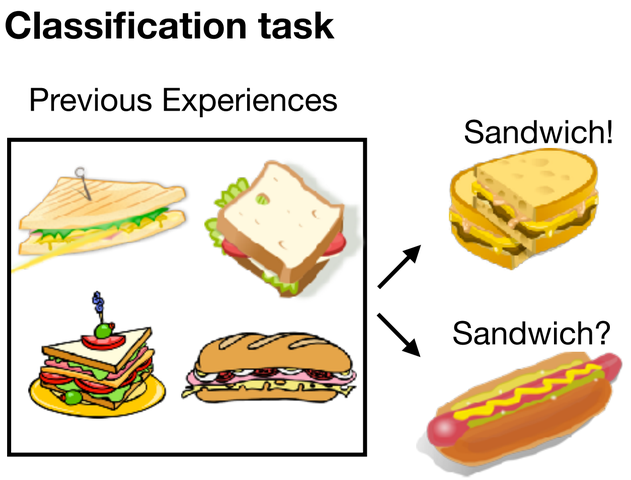
Sutton (1991)

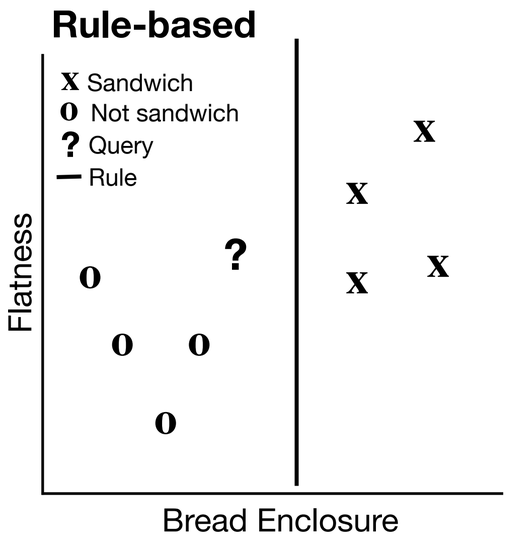
Deep RL

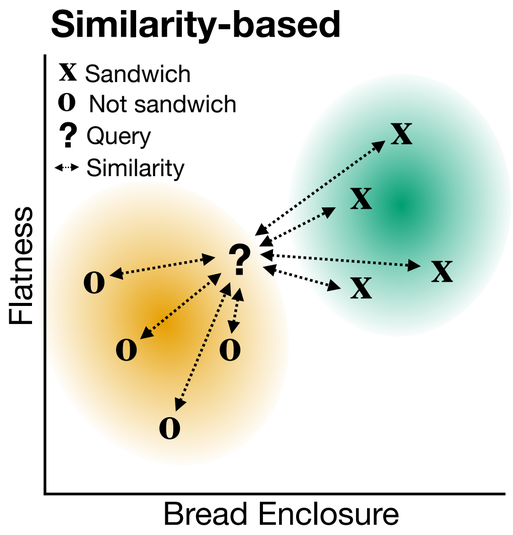


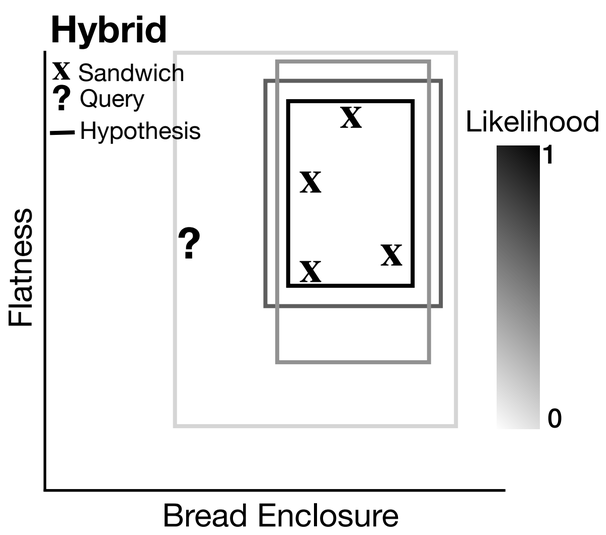
Mnih et al., (2015)

Concepts and Categories



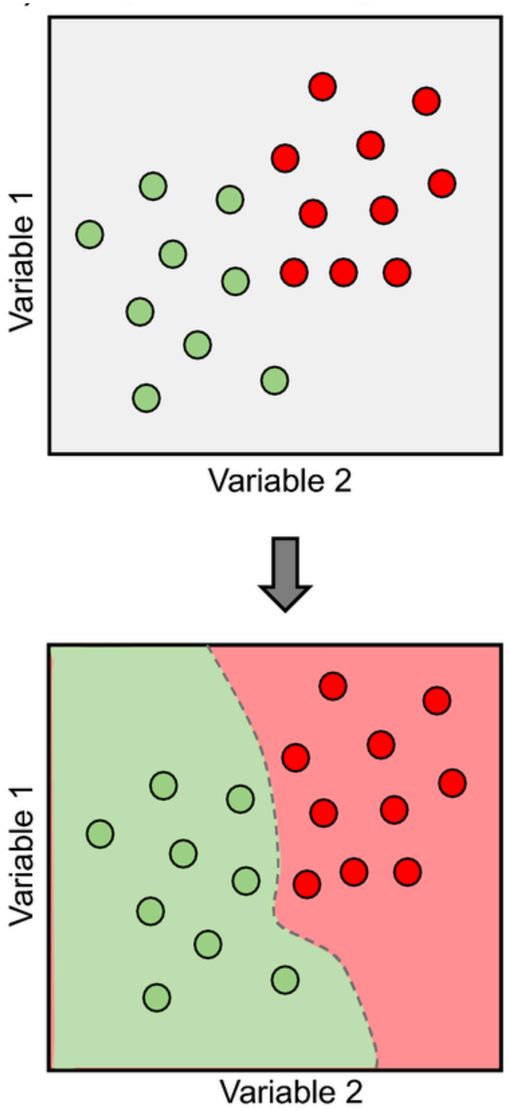


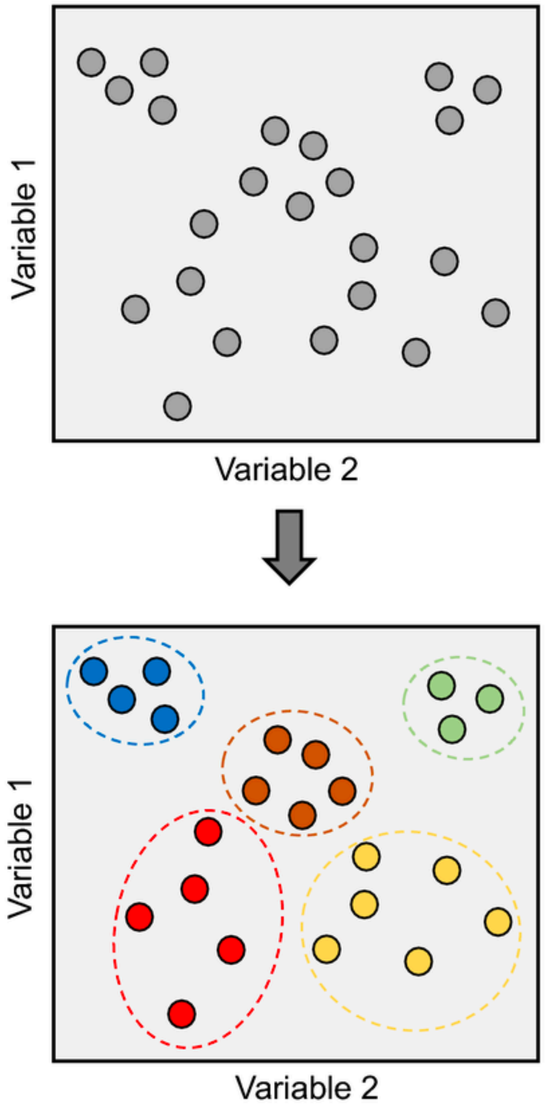




Wu et al., (in prep)

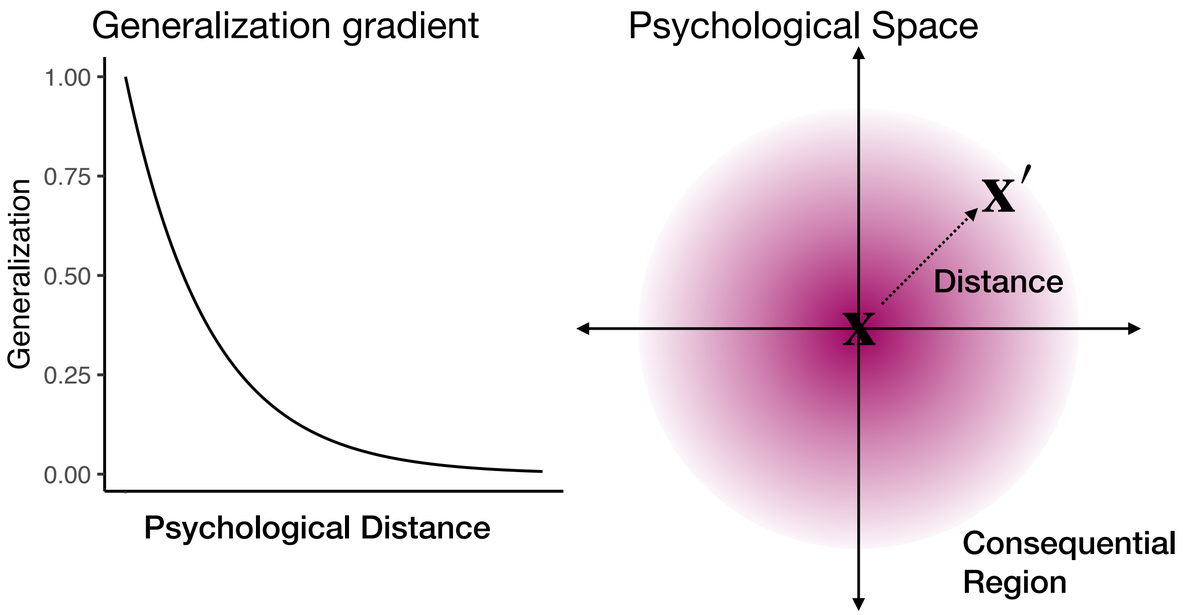
Supervised and Unsupervised Learning



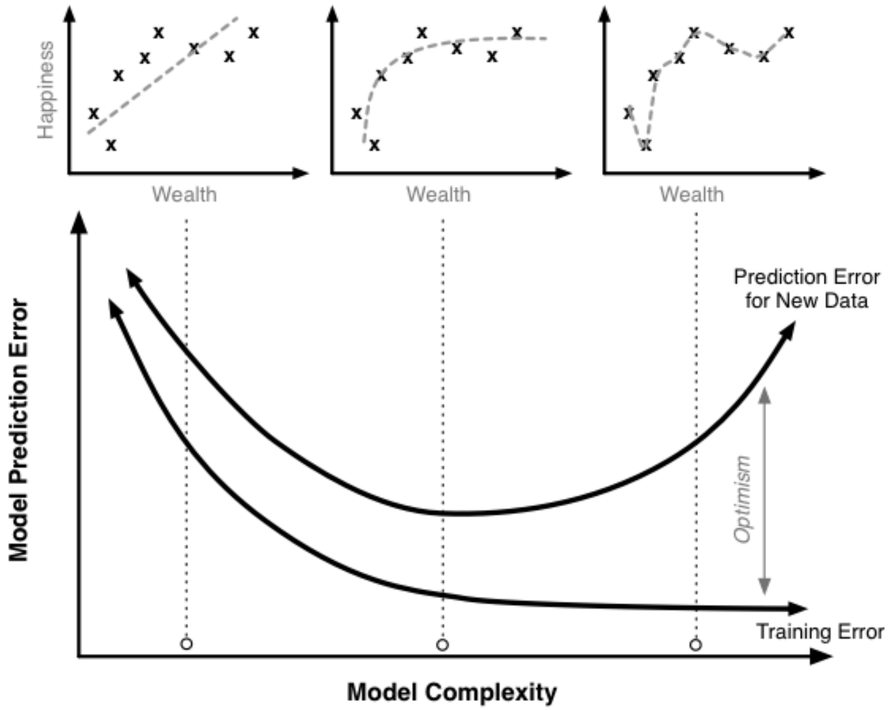


Generalization

Shepard's (1987) Law of Generalization



Bias-Variance trade-off and how deep learning breaks it



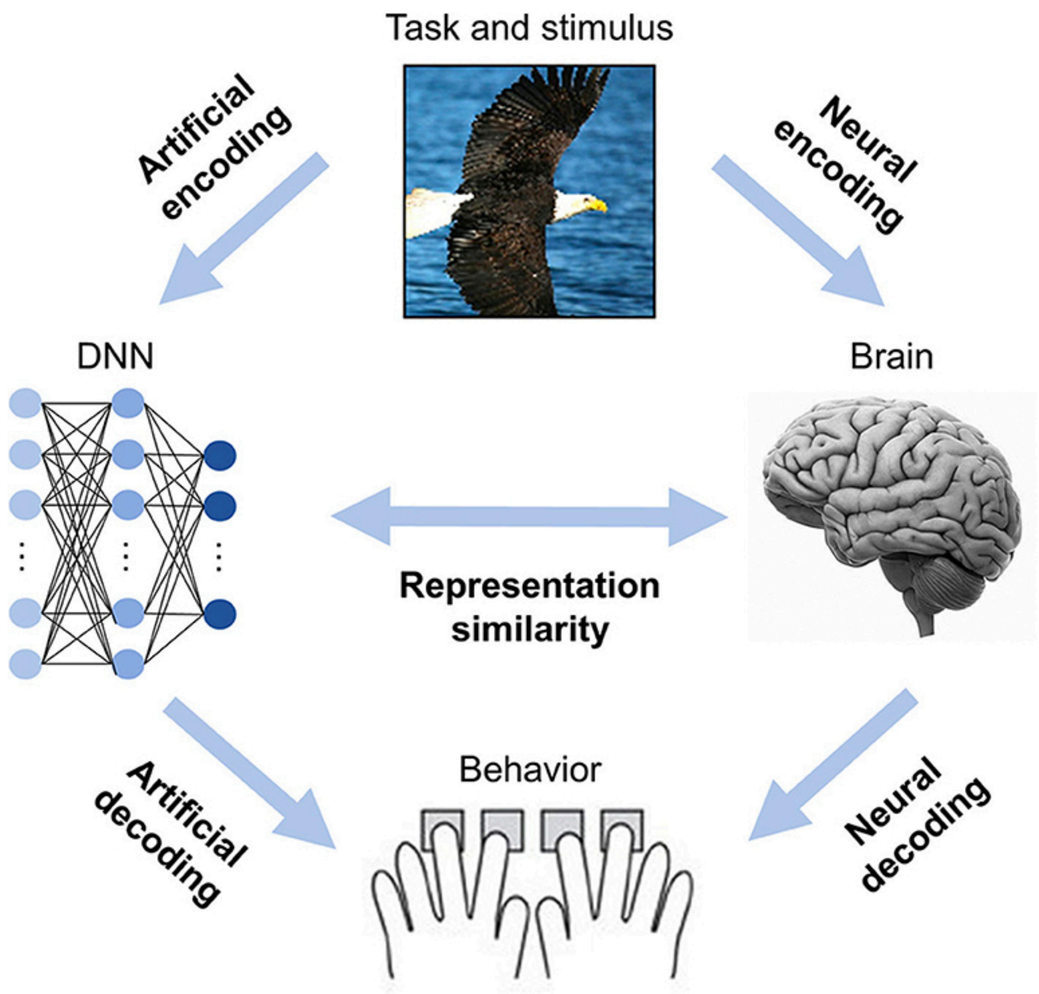
Common tools for understanding brains and neural networks

Overview of empirical methods used for understanding both brains and neural networks

- Manifold Analysis
- Representational Similarity Analysis

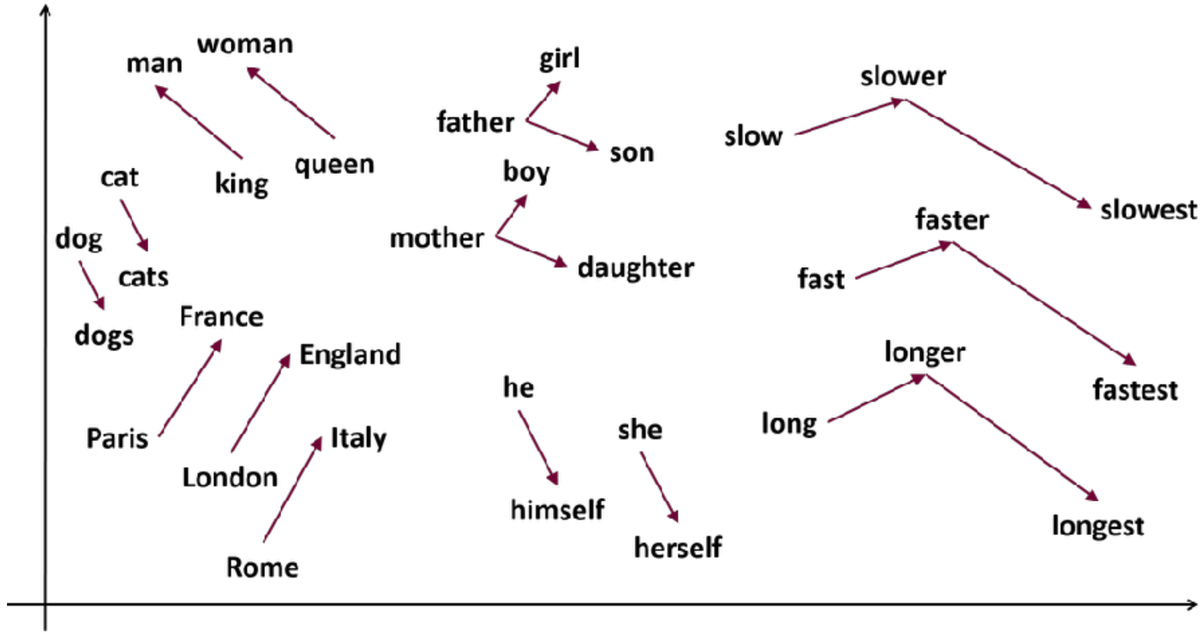
When things go wrong…

- Link to computational psychiatry

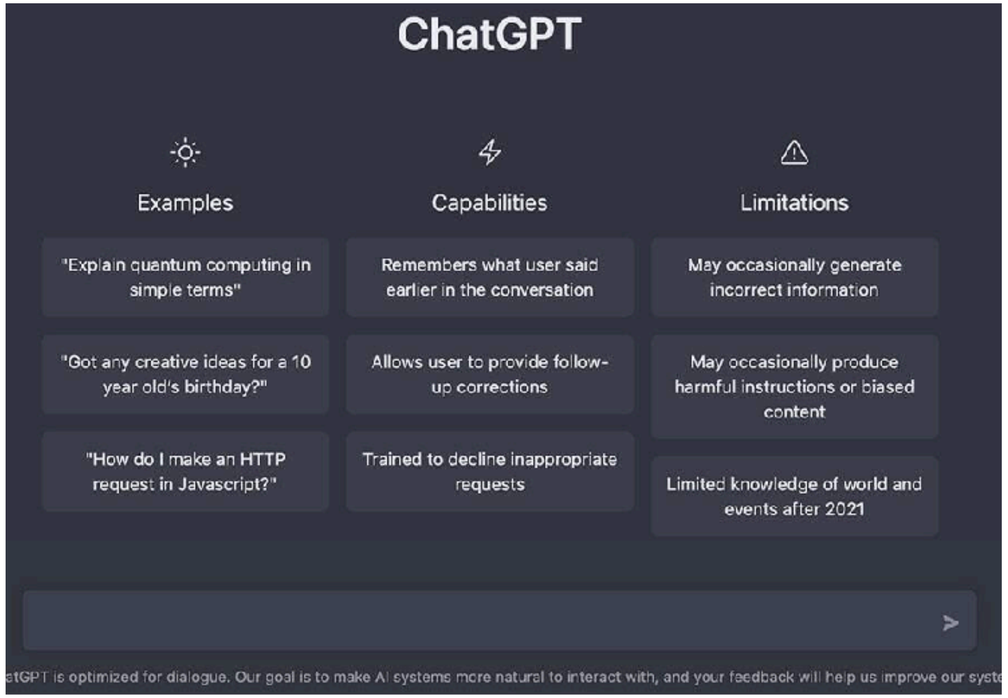


Language and Semantics

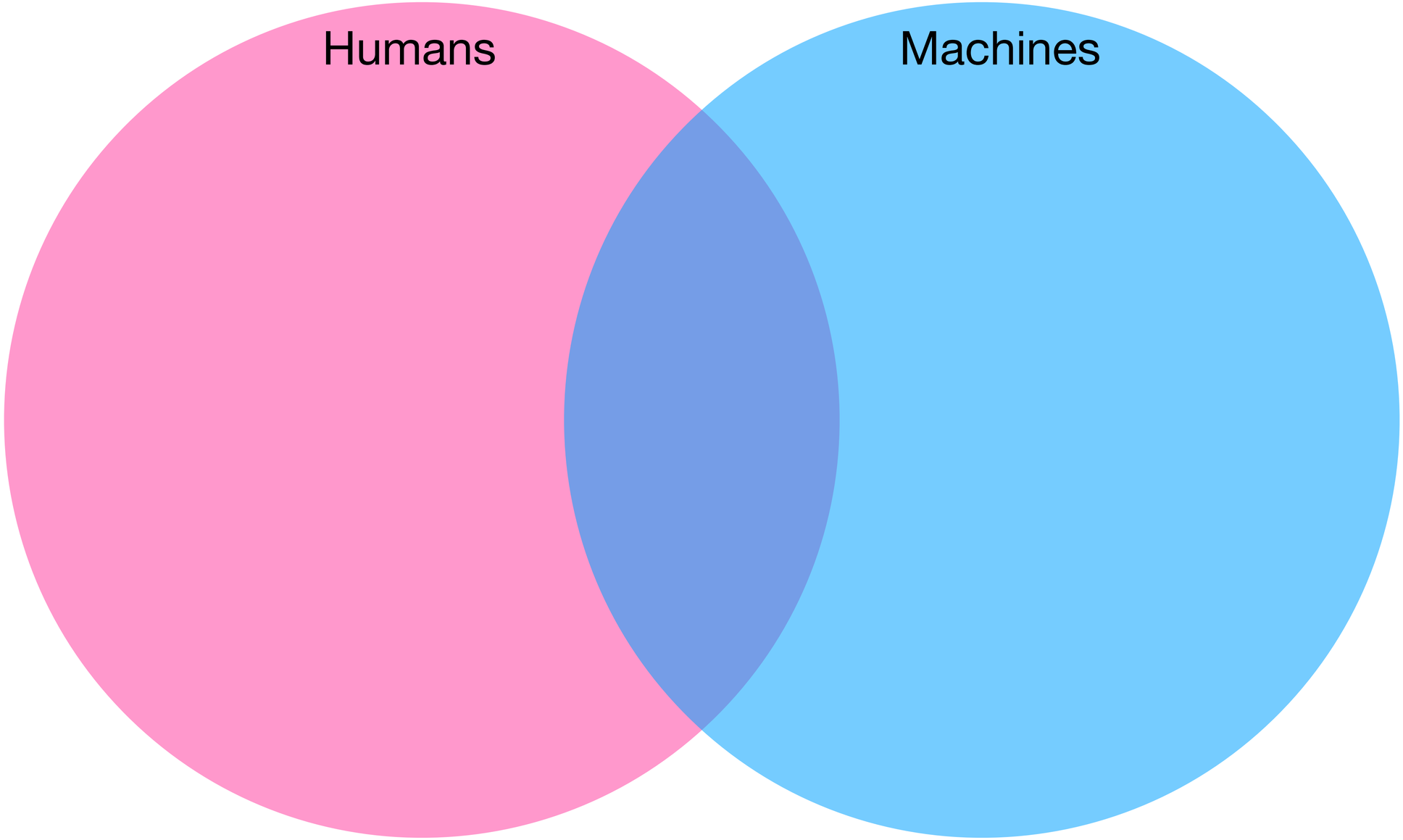
Vector Space Semantics



Large Language Models

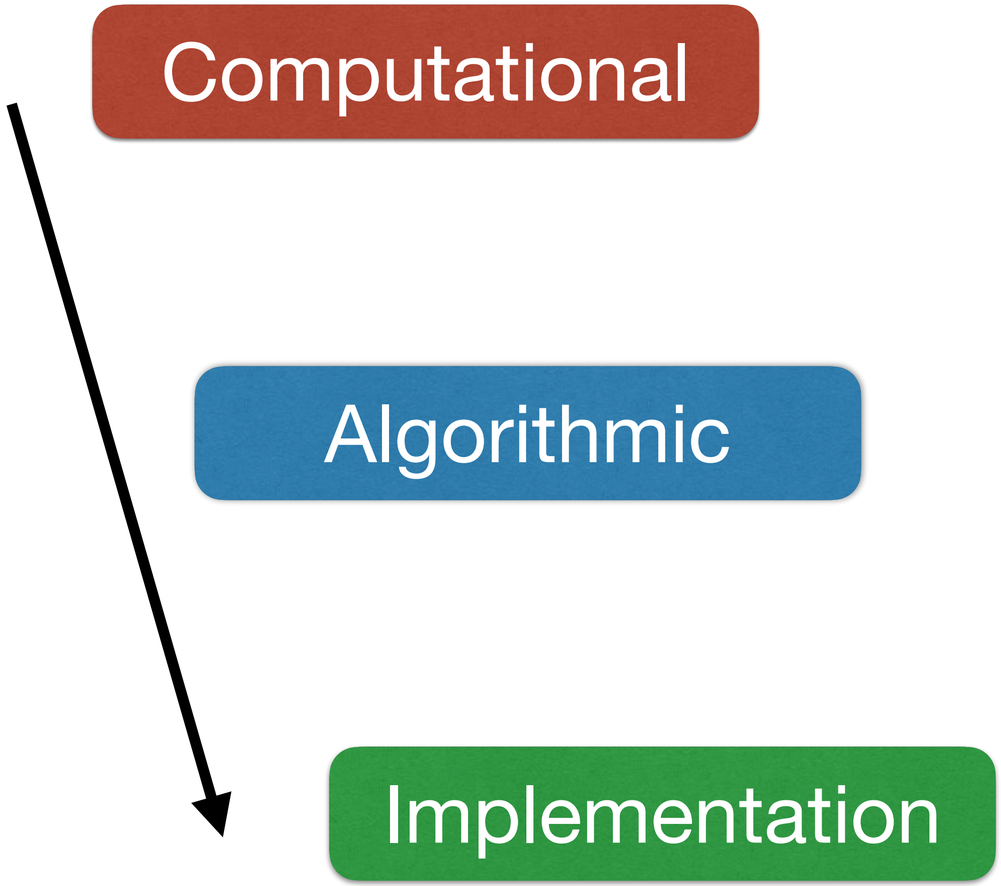


General Principles

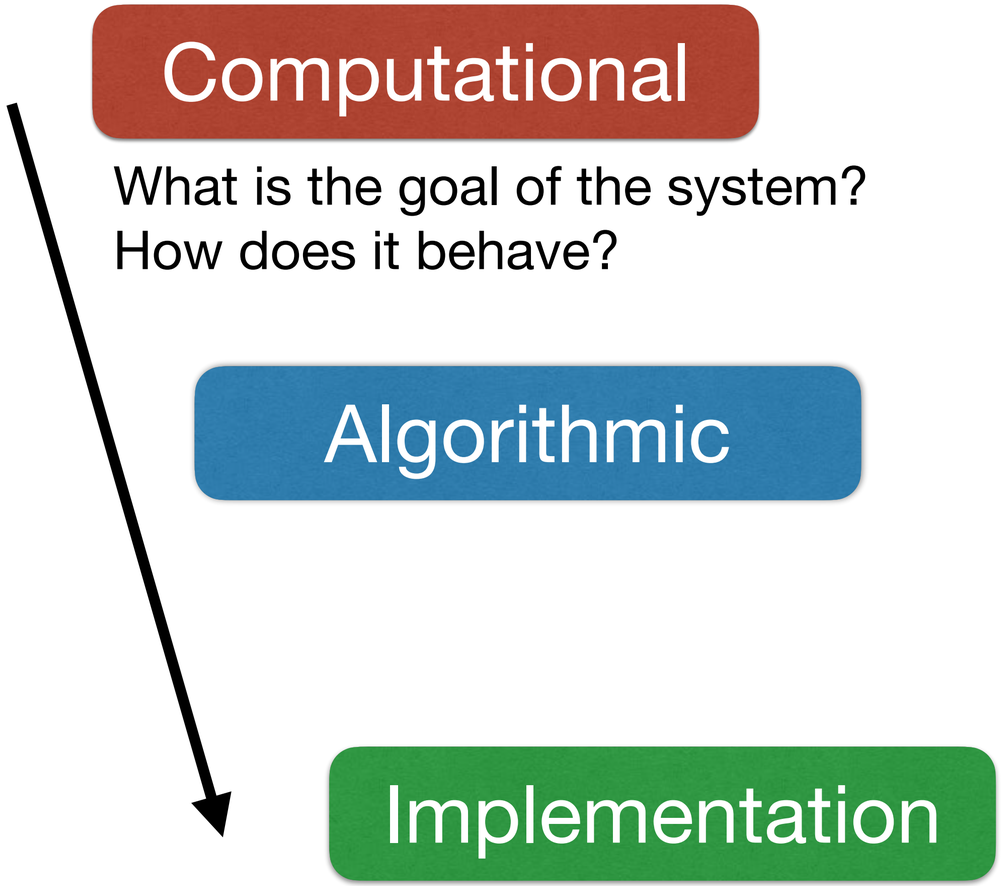


What is learning?

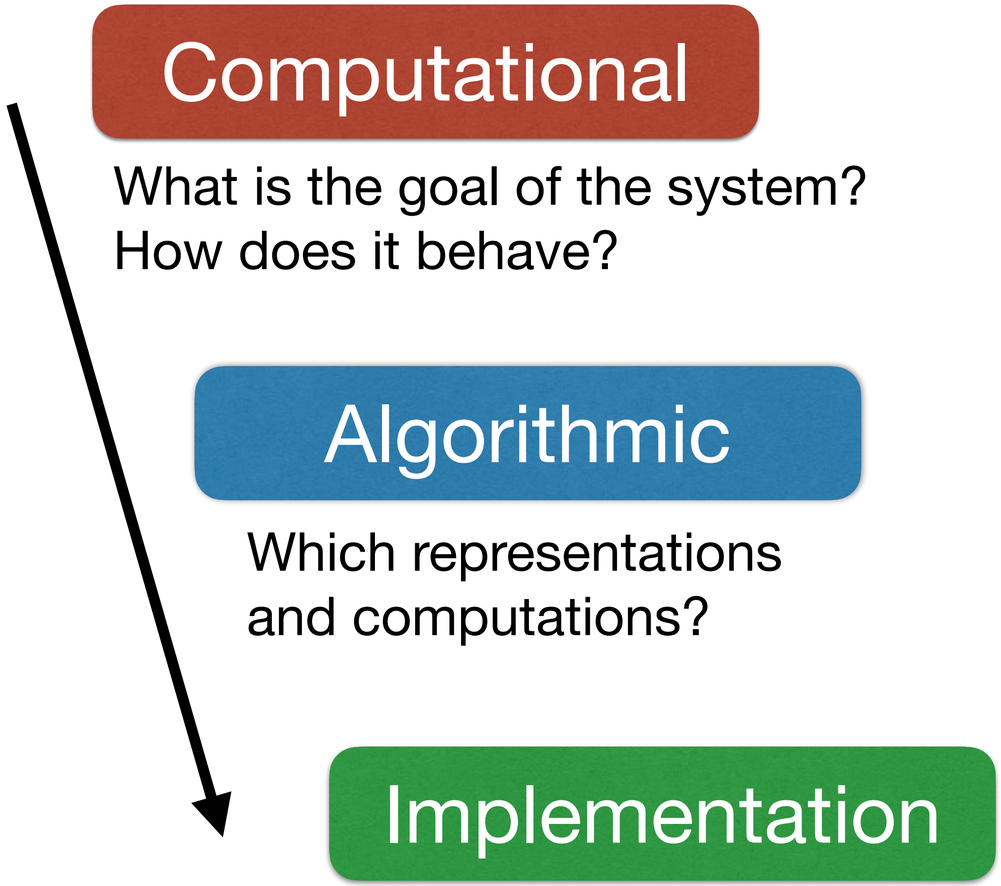
Marr’s Levels of Analysis (1982)



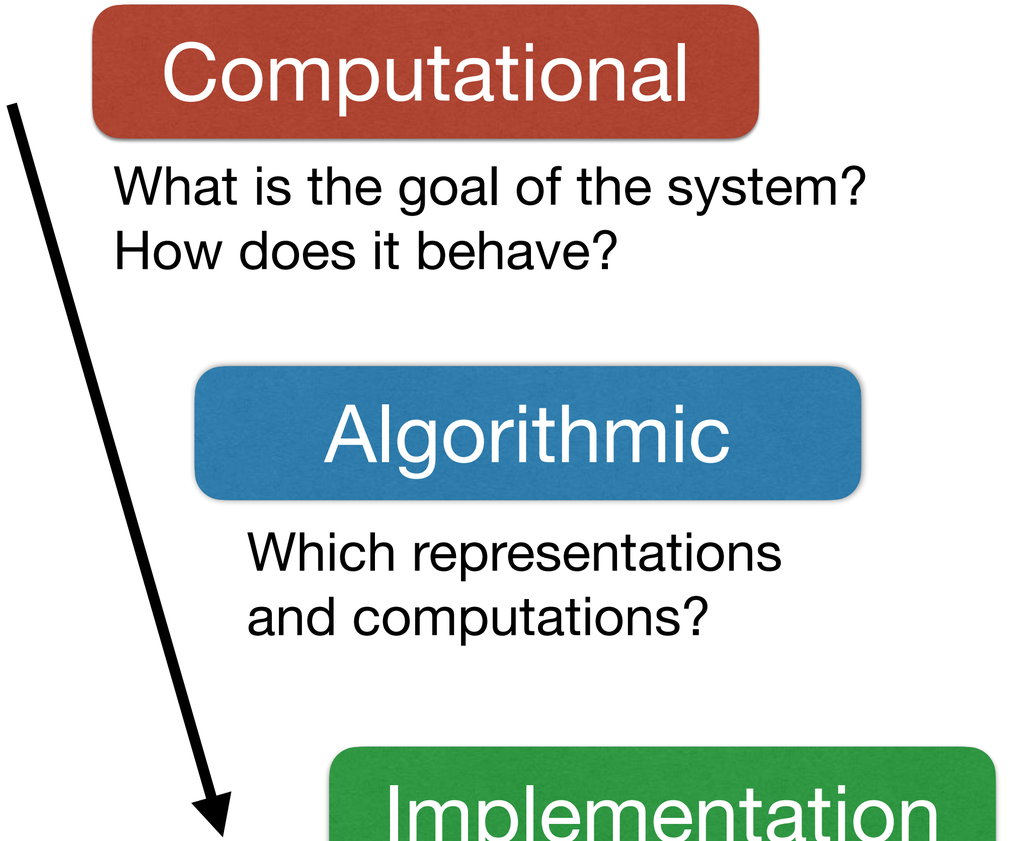
Marr’s Levels of Analysis (1982)



Marr's Levels of Analysis (1982)



Marr's Levels of Analysis (1982)



Marr's Levels of Analysis (1982)

Flight

Computational

What is the goal of the system?
How does it behave?

Flapping

Algorithmic

Which representations
and computations?

Feathers

Implementation

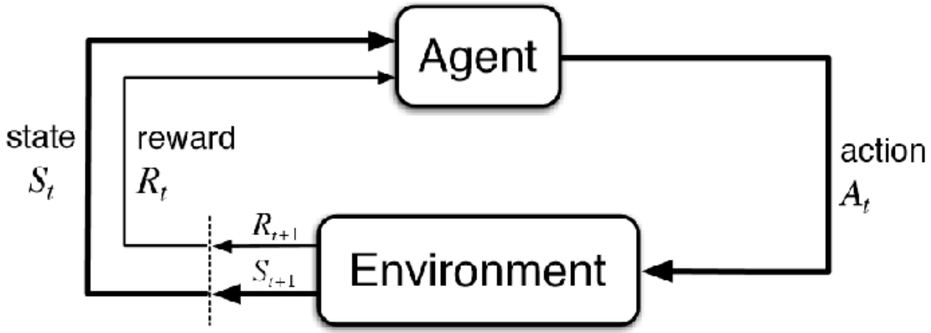
How is the system realized?

Marr's Levels of Analysis (1982)

Flight

Computational

What is the goal of the system?
How does it behave?



Flapping

Algorithmic

Which representations
and computations?

Feathers

Implementation

How is the system realized?

Marr's Levels of Analysis (1982)

Flight

Computational

What is the goal of the system?
How does it behave?

Flapping

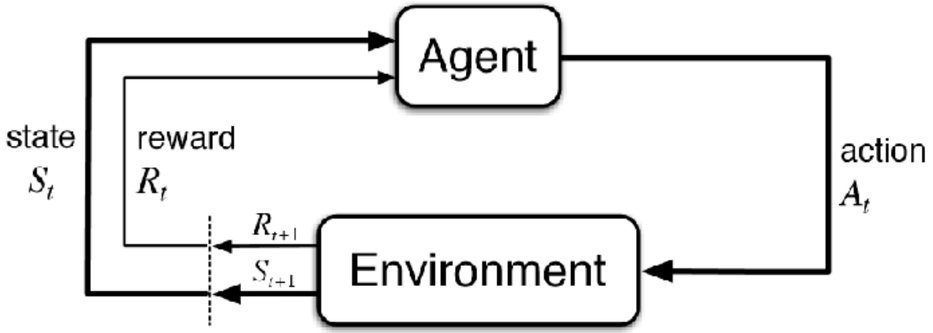
Algorithmic

Which representations
and computations?

Feathers

Implementation

How is the system realized?



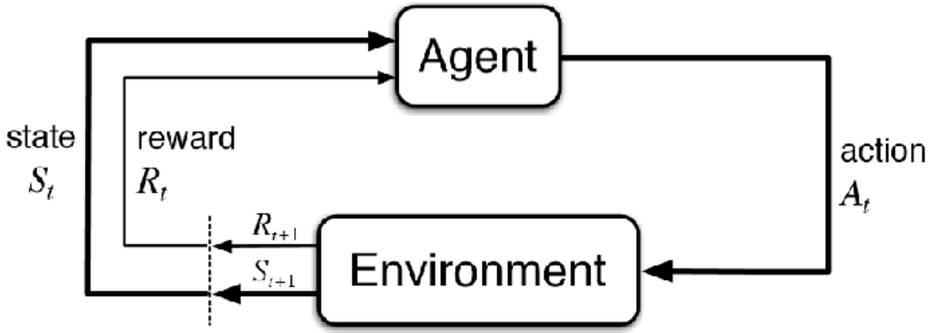
Initialize $Q(s,a)$ arbitrarily
Repeat (for each episode):
  Initialize s
  Repeat (for each step of episode):
    Choose a from s using policy derived from Q (e.g., ε-greedy)
    Take action a, observe r, s'
    $Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right]$
    $s \leftarrow s'$;
  until s is terminal

Marr's Levels of Analysis (1982)

Flight

Computational

What is the goal of the system?
How does it behave?

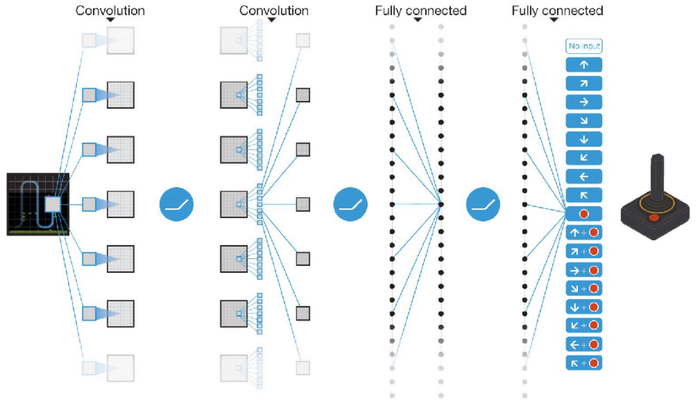


Flapping

Algorithmic

Which representations
and computations?

Initialize $Q(s,a)$ arbitrarily
Repeat (for each episode):
  Initialize s
  Repeat (for each step of episode):
    Choose a from s using policy derived from Q (e.g., ε-greedy)
    Take action a, observe r, s'
    $Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right]$
    $s \leftarrow s'$;
  until s is terminal



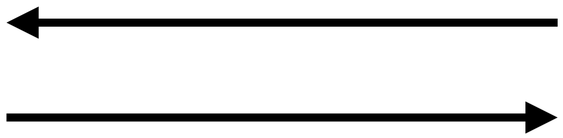
Feathers

Implementation

How is the system realized?

Categorize each definition of “learning” using Marr’s levels

How can machines inform our understanding of human learning?



How does human learning inform the development of machine learning?

See you next week

- Don’t forget to finish your assigned reading before the tutorial tomorrow
  - Spicer & Sanborn (2019)
- Next week, we look at the the origins of research on biological and artificial learning