Introduction to Data Analytics

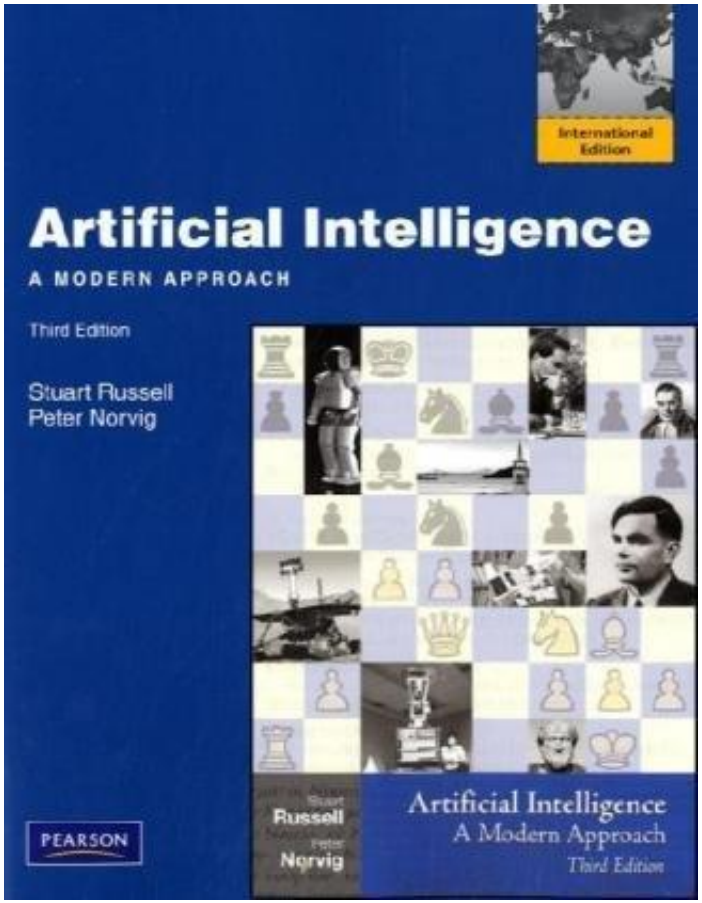
Xin Gao
Xin.gao@kaust.edu.sa
July 25, 2022
SDU

Course Content

- Basics and principles of artificial intelligence (AI), data mining (DM) and machine learning (ML)
- Goal
  - Understand these important research areas
  - Help your own research
  - Prepare for advanced courses

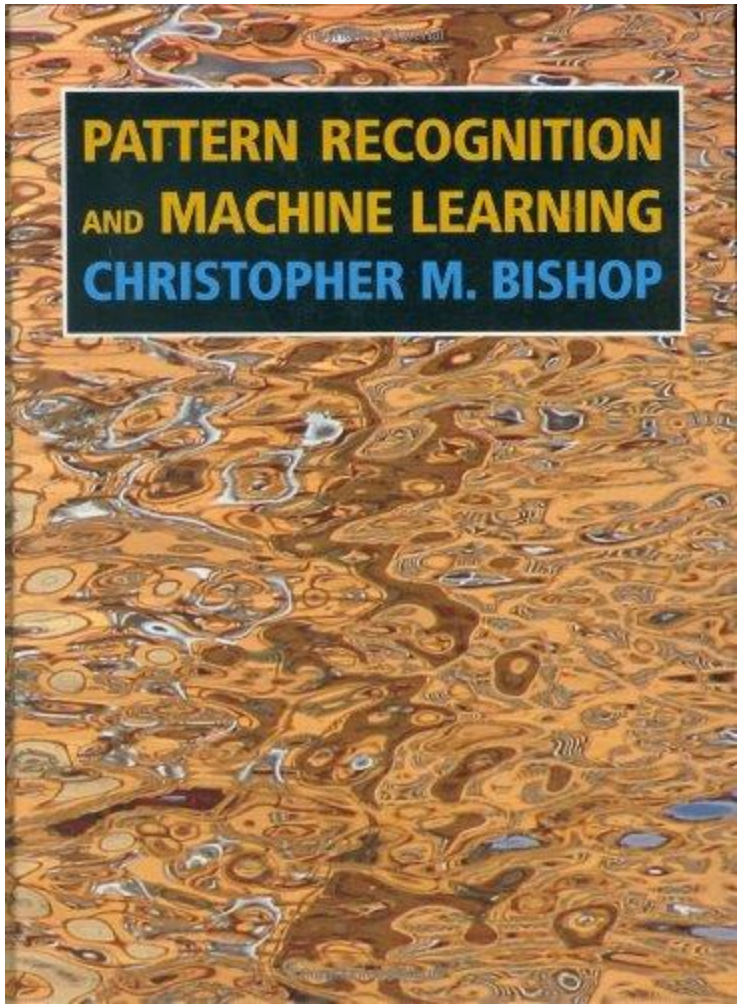
References (not obligatory)

- AI: Stuart Russel and Peter Norvig, “Artificial Intelligence: A Modern Approach (Second edition)”, Prentice Hall, 2002
- Third edition is also fine
- It is the dominant textbook in Artificial Intelligence
- Used in over 1100 universities in 102 countries (over 90% market share)



References (not obligatory)

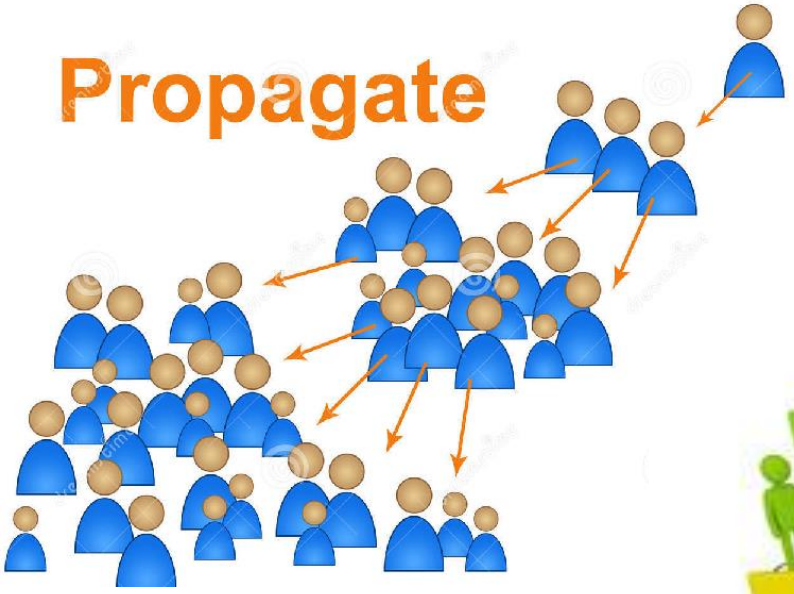
- Data mining & machine learning: Christopher Bishop, 2007
- All the course contents will be covered in class



Prerequisites

- Basic concepts in data structure and algorithms (tree, sorting, NP-hard, big O notations, etc) and probability theory (joint probability, conditional probability, etc).
- Ability to write code to implement algorithms and to test them. MATLAB and C/C++ are recommended, but any language is allowed

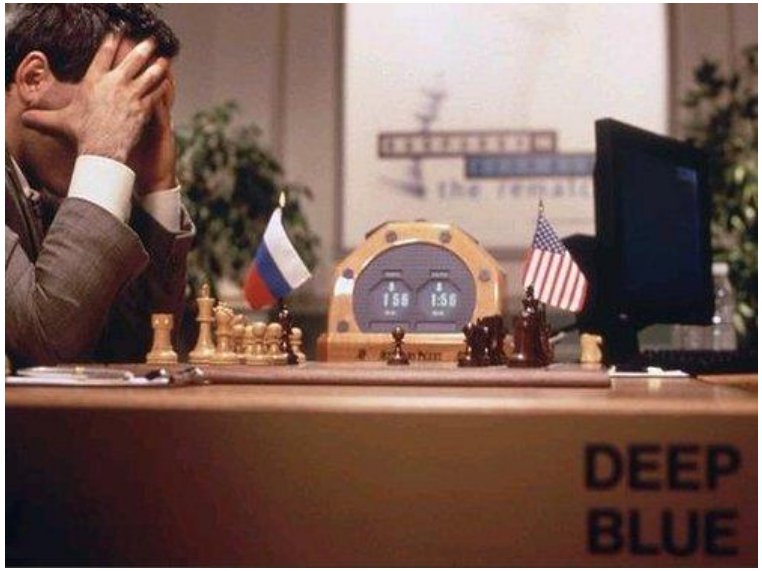
Artificial Intelligence
Example 1 – Information Propagation



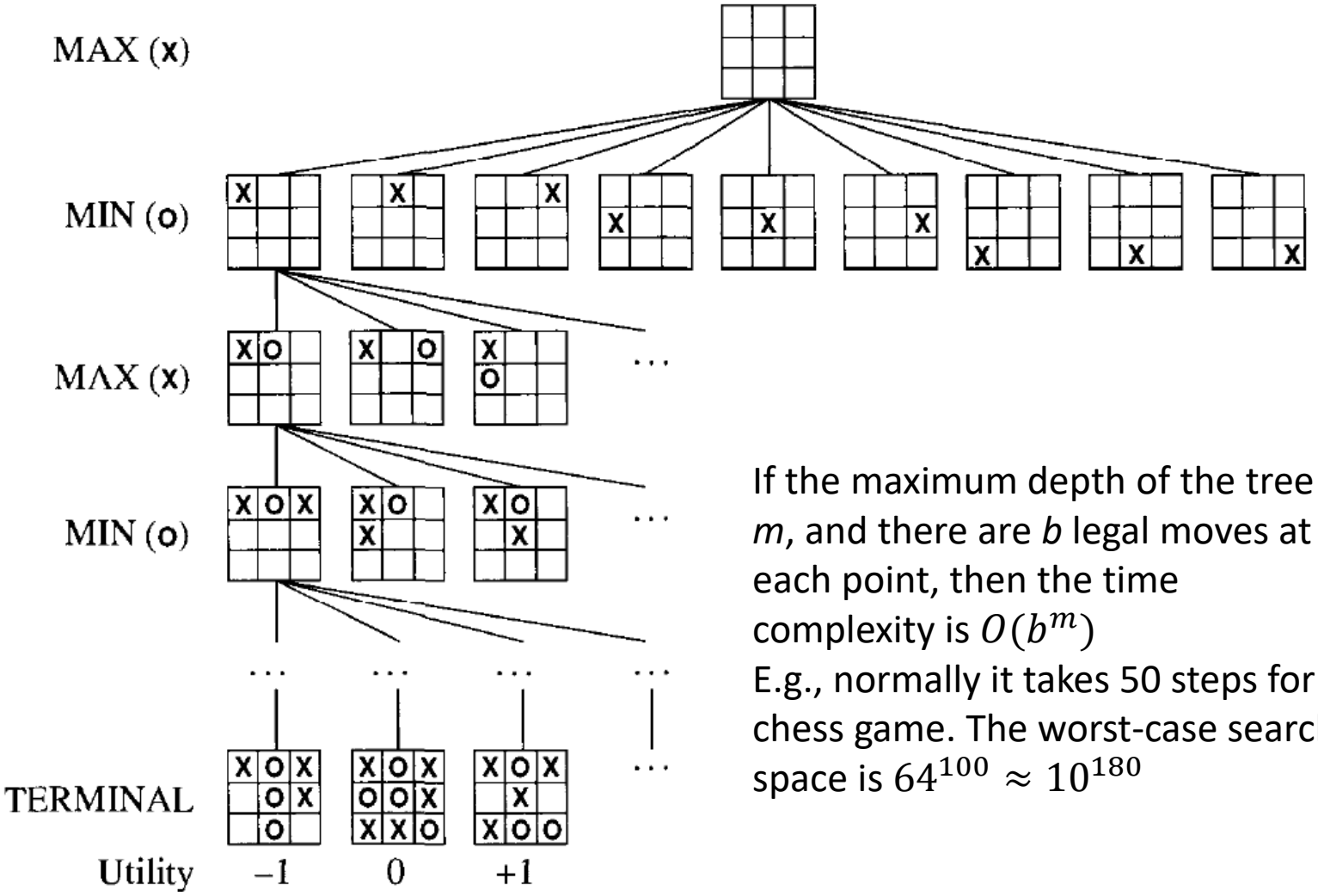
Artificial Intelligence
Example 2 – Games

- In 1996, IBM invented a supercomputer named "Deep Blue“. Able to compute more than 100 million chess positions per second, Deep Blue challenged the reigning world chess champion Garry Kasparov to a chess match. Kasparov won the match, with 3 wins, 2 ties, and 1 loss. This was the first time ever that a computer has beaten a reigning world chess champion.

Will computers soon surpass humans in chess playing and in other aspects of intelligence?



Searching – Adversarial Search



If the maximum depth of the tree is *m*, and there are *b* legal moves at each point, then the time complexity is $O(b^m)$

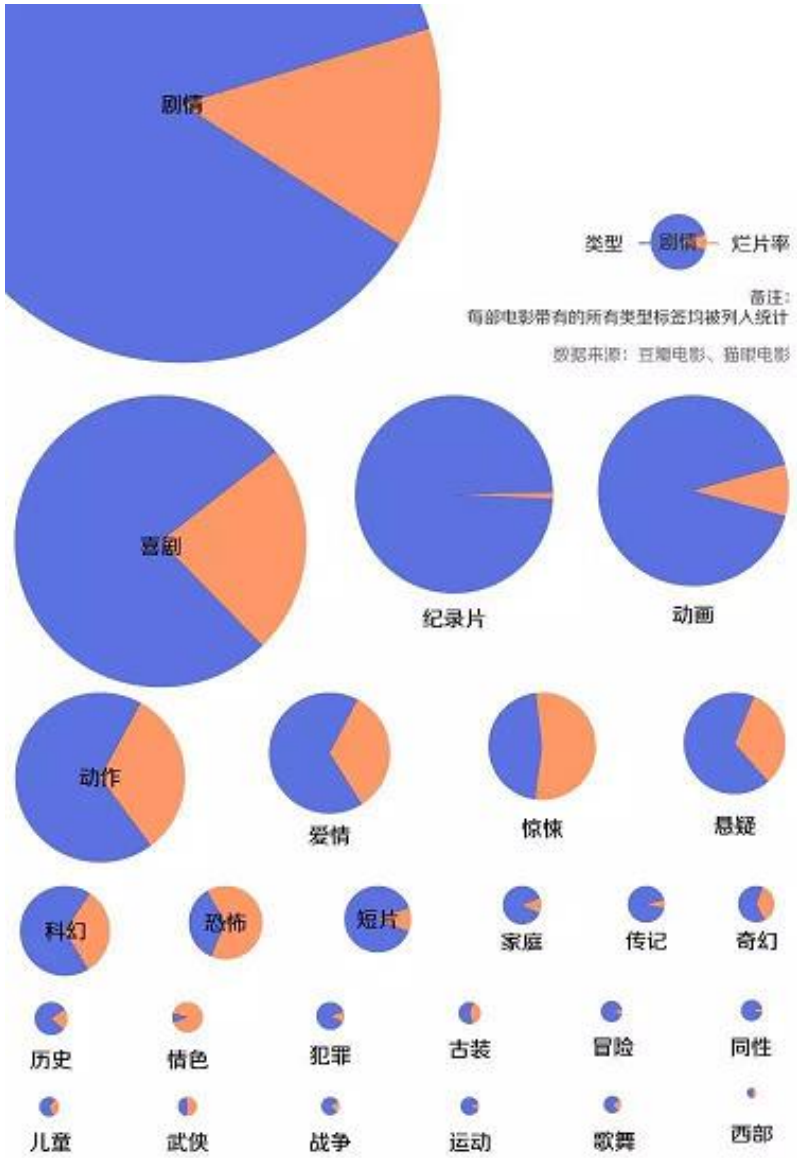
E.g., normally it takes 50 steps for a chess game. The worst-case search space is $64^{100} \approx 10^{180}$

AlphaGO

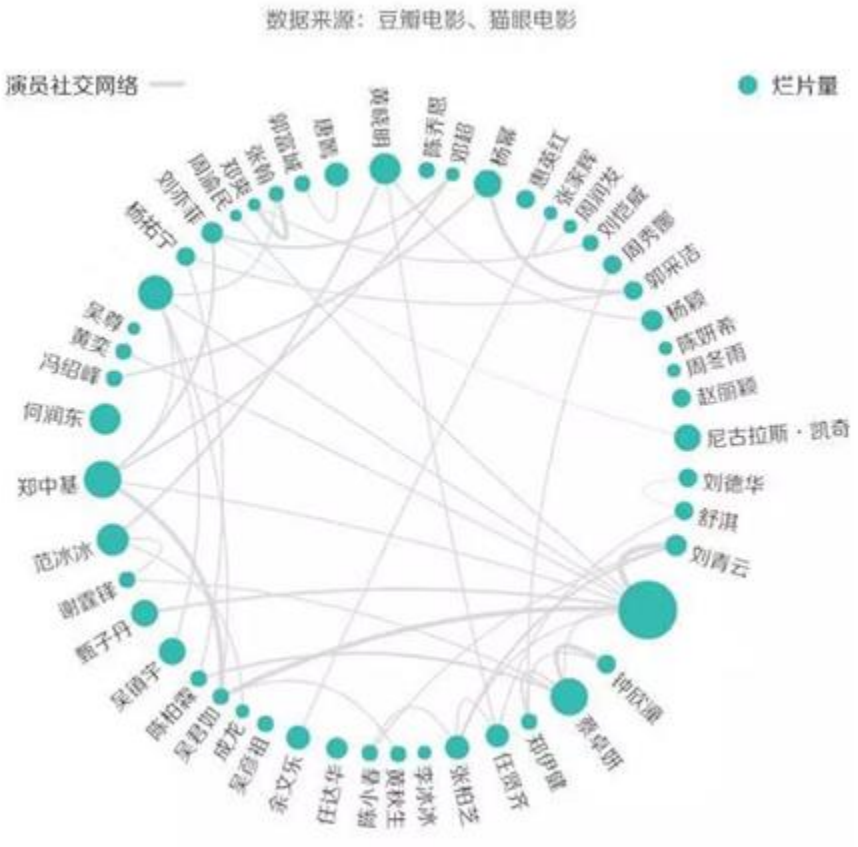

KBA
Google DeepMind
Challenge Match
8 - 15 March 2016
AlphaGo
ALPHAGO
01:58:31
Lee Sedol
AlphaGo
Google DeepMind
Challenge Match
8 - 15 March 2016

Data Mining
Example 1 – Movie Evaluation

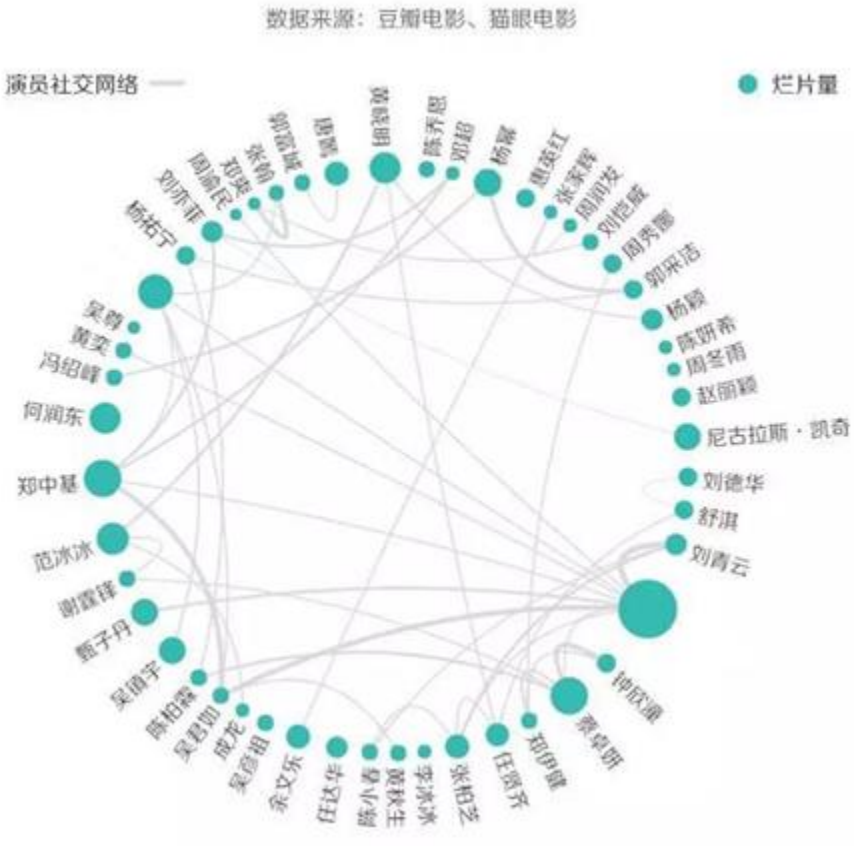
Data Mining
Example 1 – Movie Evaluation

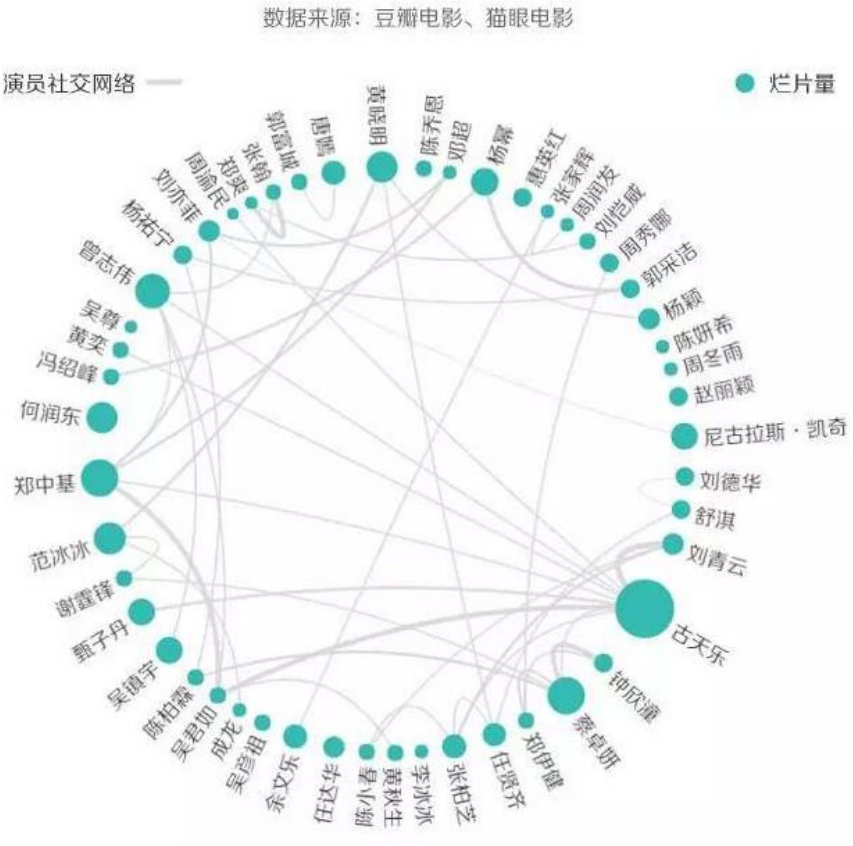


Data Mining
Example 1 – Movie Evaluation

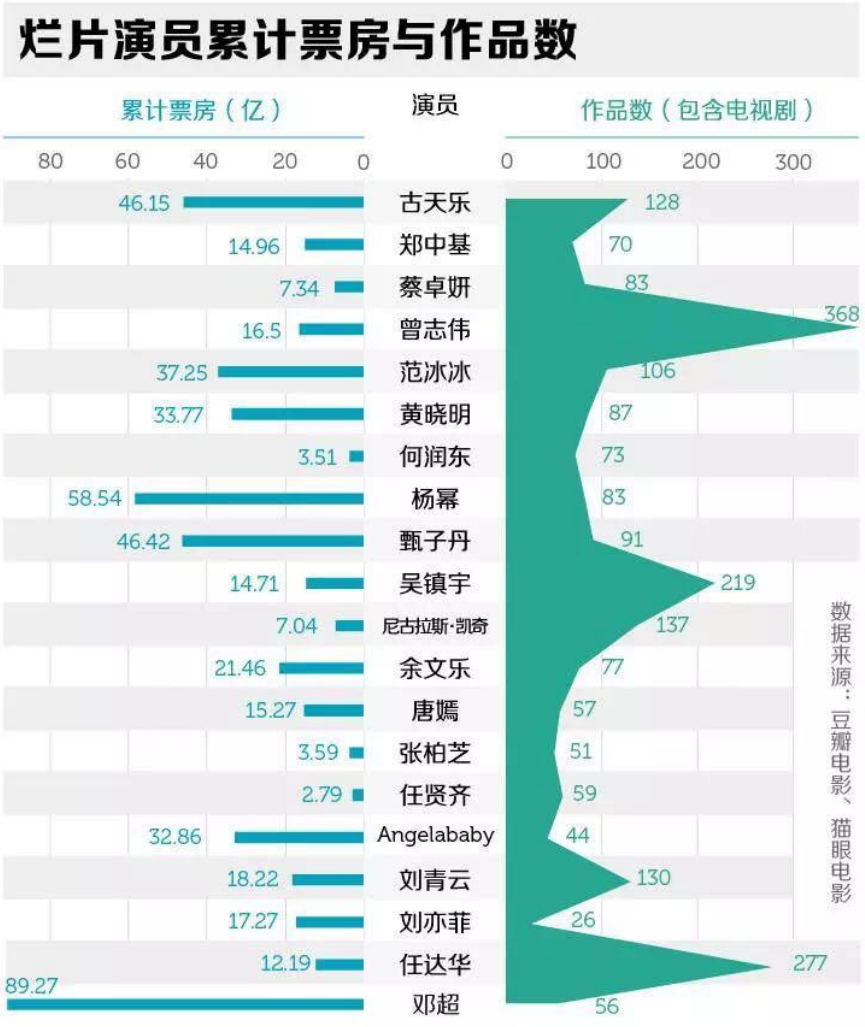


Data Mining
Example 1 – Movie Evaluation

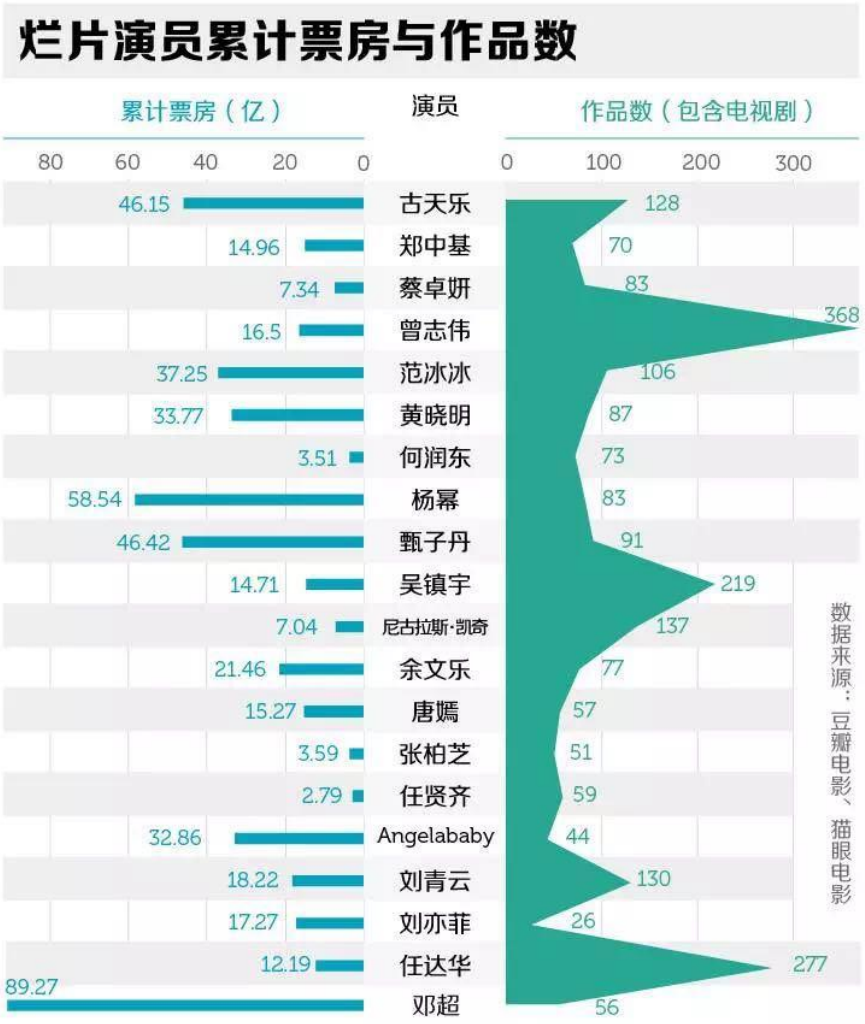


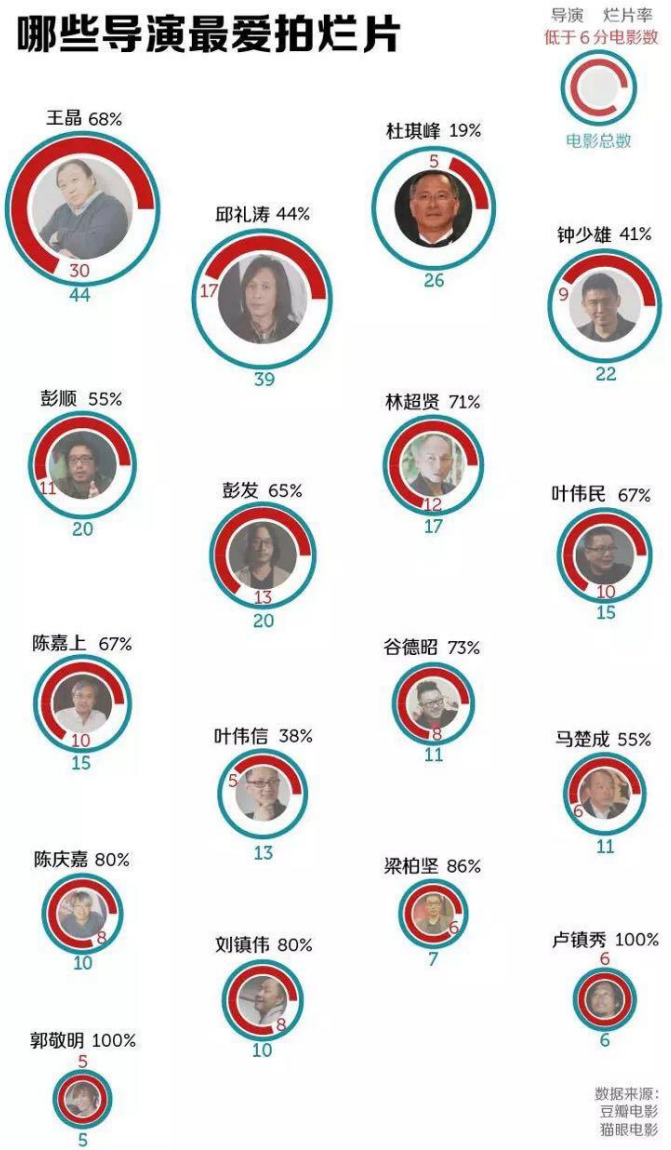


Data Mining
Example 1 – Movie Evaluation

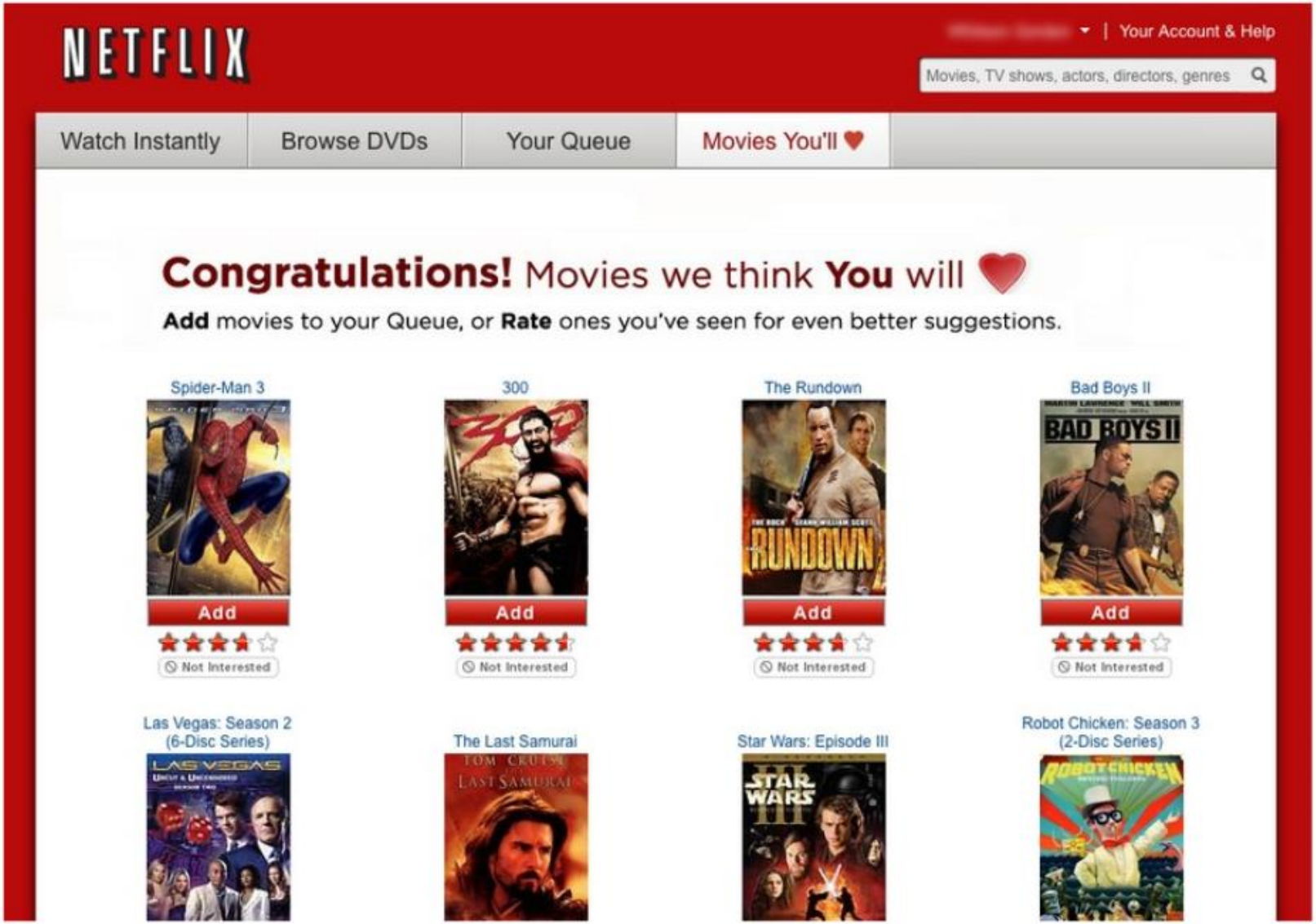


Data Mining
Example 1 – Movie Evaluation





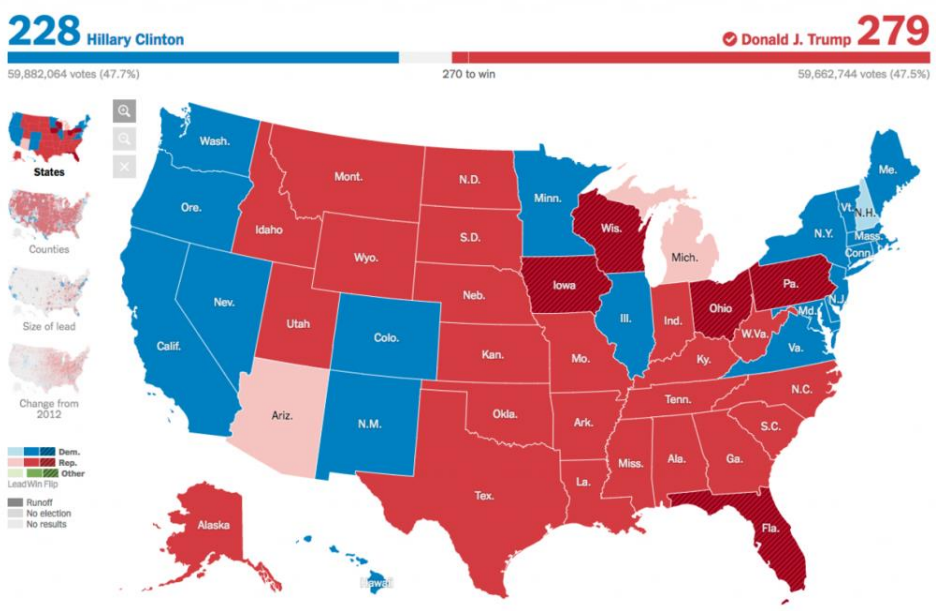
Data Mining
Example 2 – Netflix Recommender

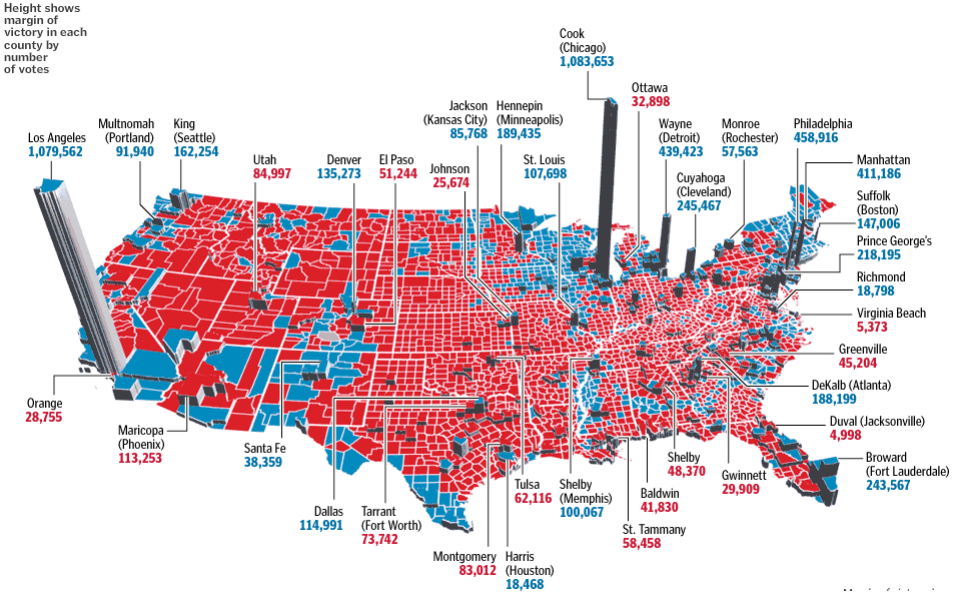


Feature Selection

- Feature to define distance:
  - Movie-movie distance
  - User-user distance
- Supervised feature selection
- Unsupervised feature selection

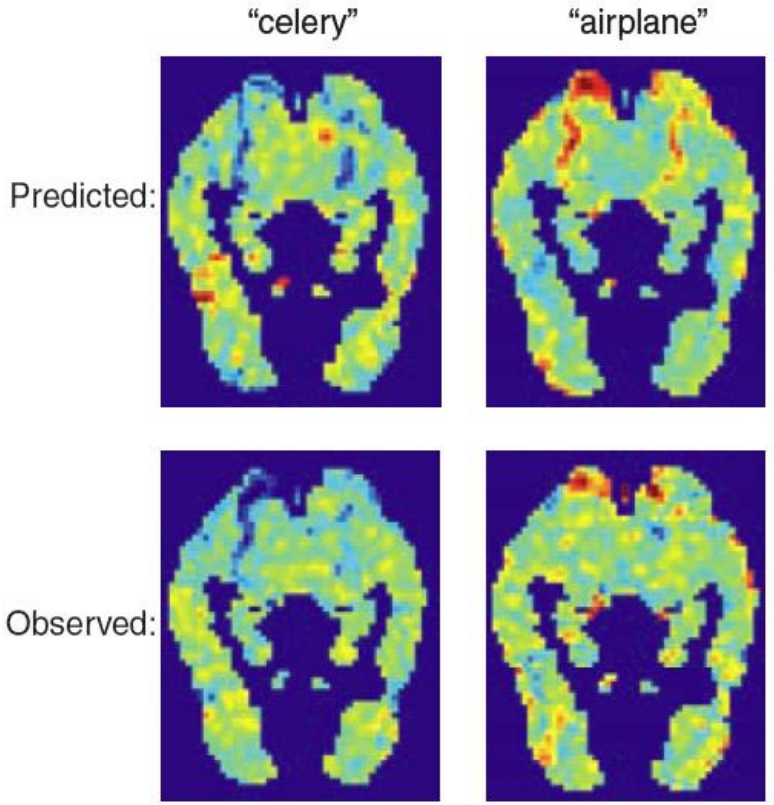
Machine Learning
Example 1 – Election Prediction





Machine Learning
Example 2 – fMRI Modeling

- Functional Magnetic Resonance Imaging (fMRI) is a type of specialized MRI scan. It measures the hemodynamic response (change in blood flow) related to neural activity in the brain or spinal cord of humans or other animals
- One of the most recently developed forms of neuroimaging

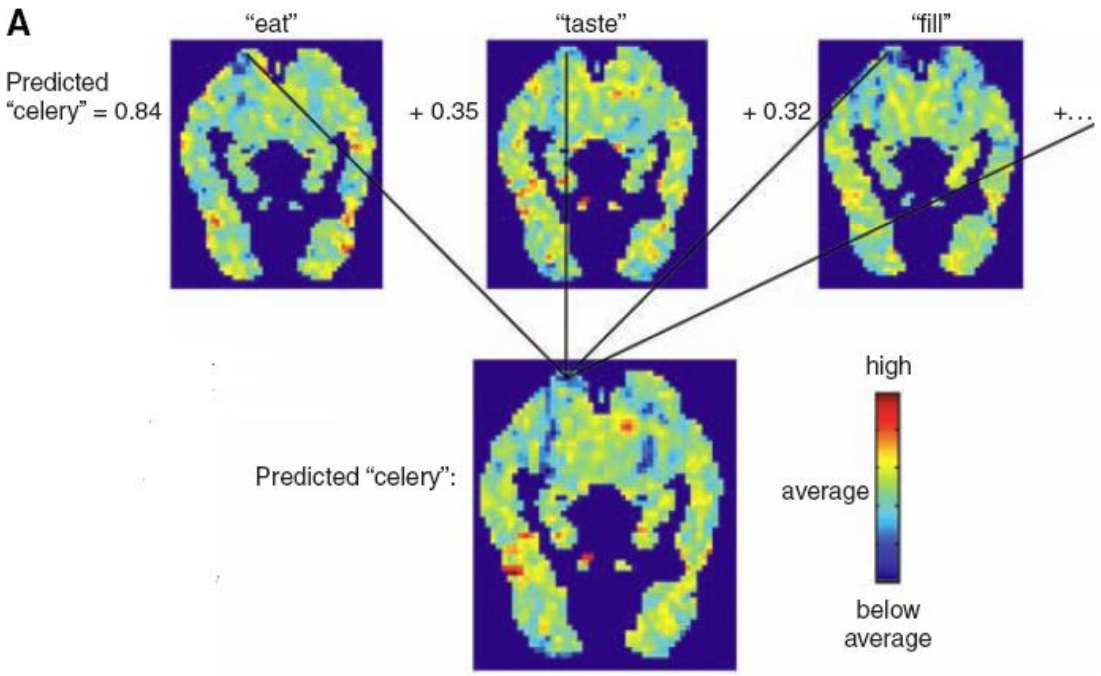


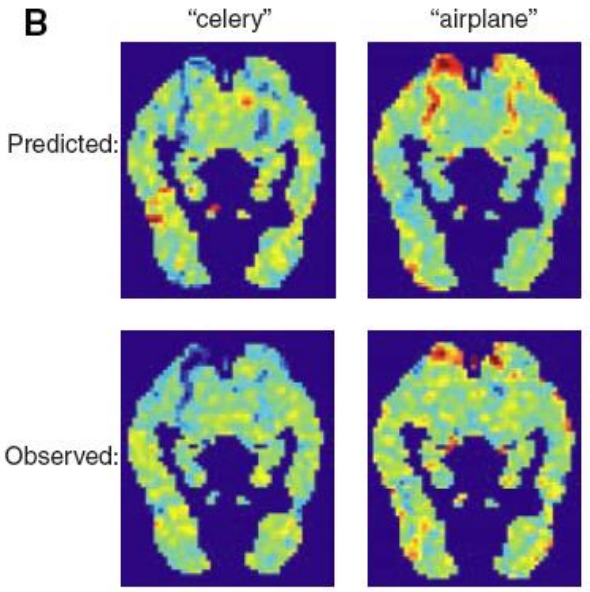
Tom Mitchell *et al*., Science 2008

Classification and Regression

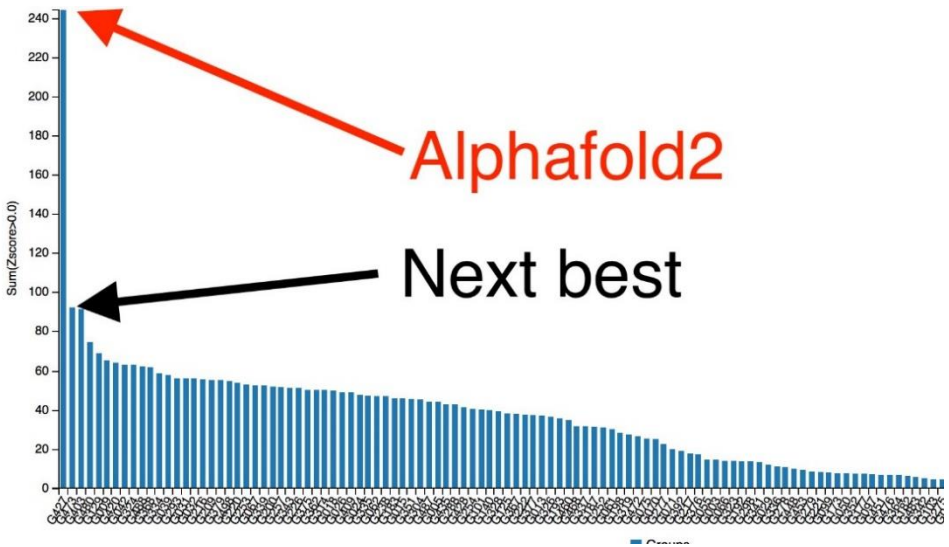
- Classification questions
  - Is he reading a sentence or viewing a picture?
  - Is he reading the word “Hammer” or “Apartment”?
  - Is he viewing a vertical or horizontal line?
  - Is he answering questions or getting confused?
- Regression questions
  - What should his fMRI be like if he’s reading “celery”?

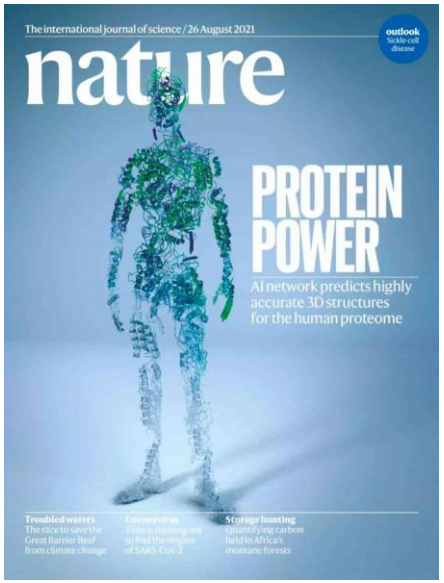
Understanding Brain Activities

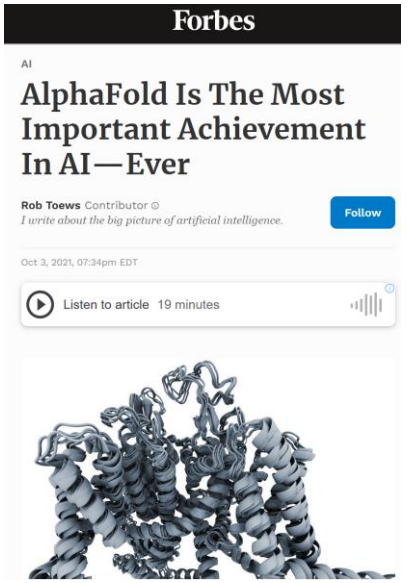


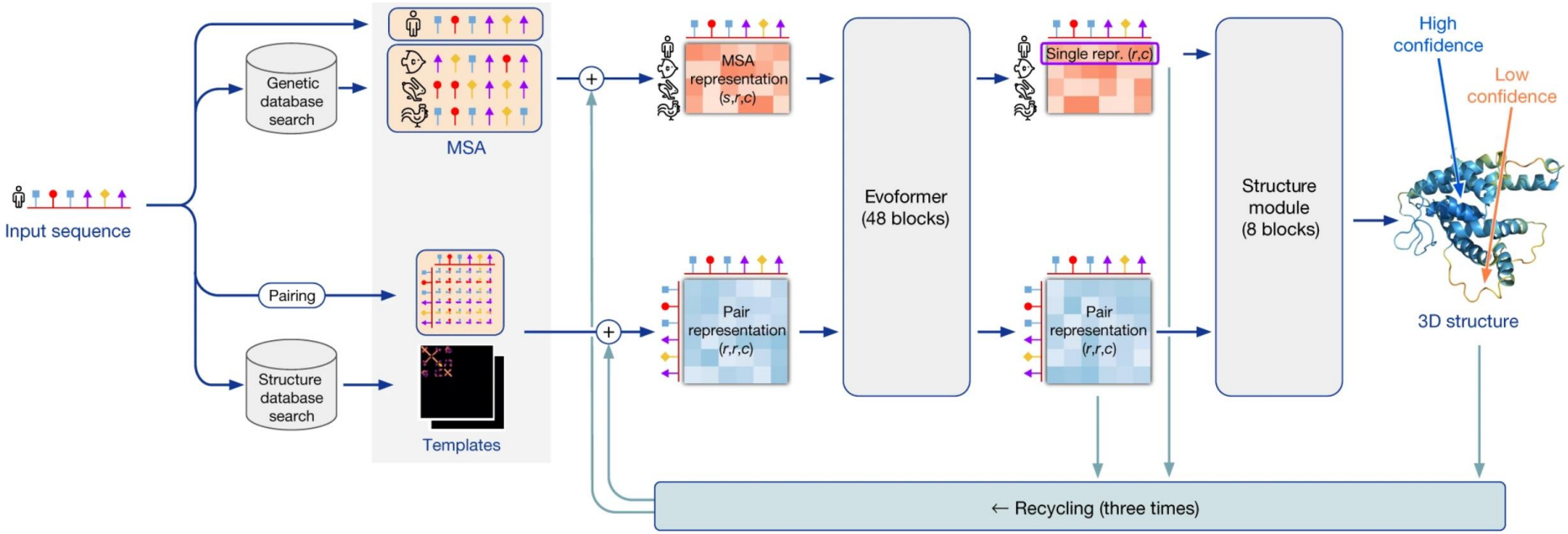


AlphaFold2 – The Largest Breakthrough in AI

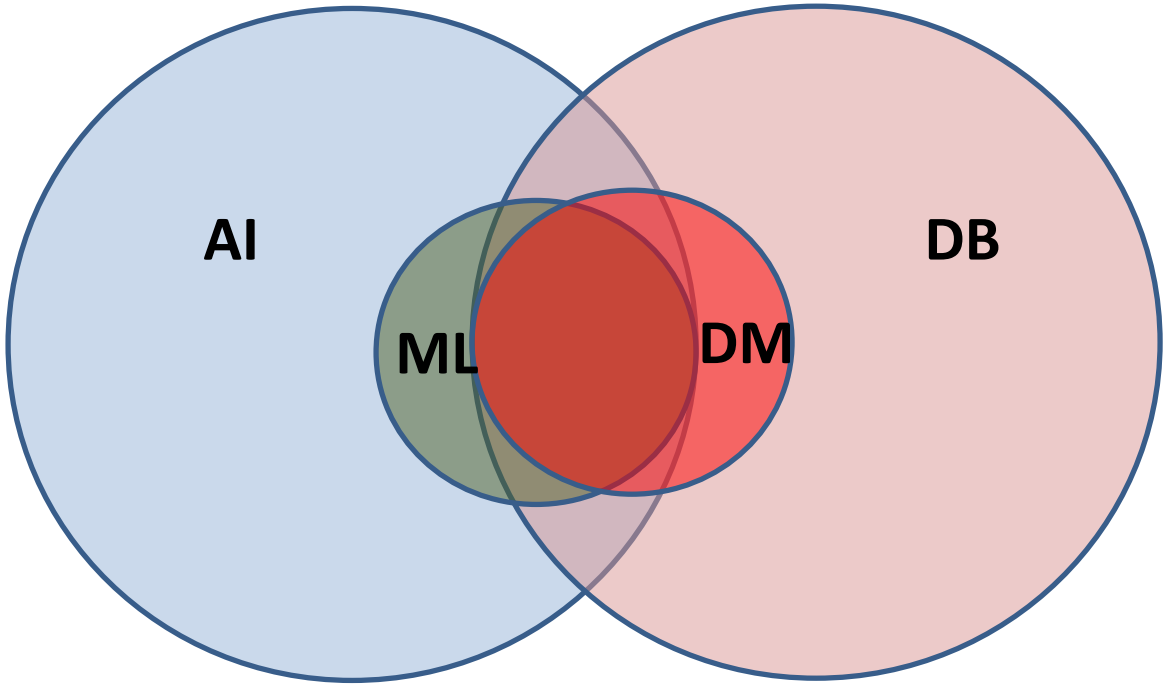




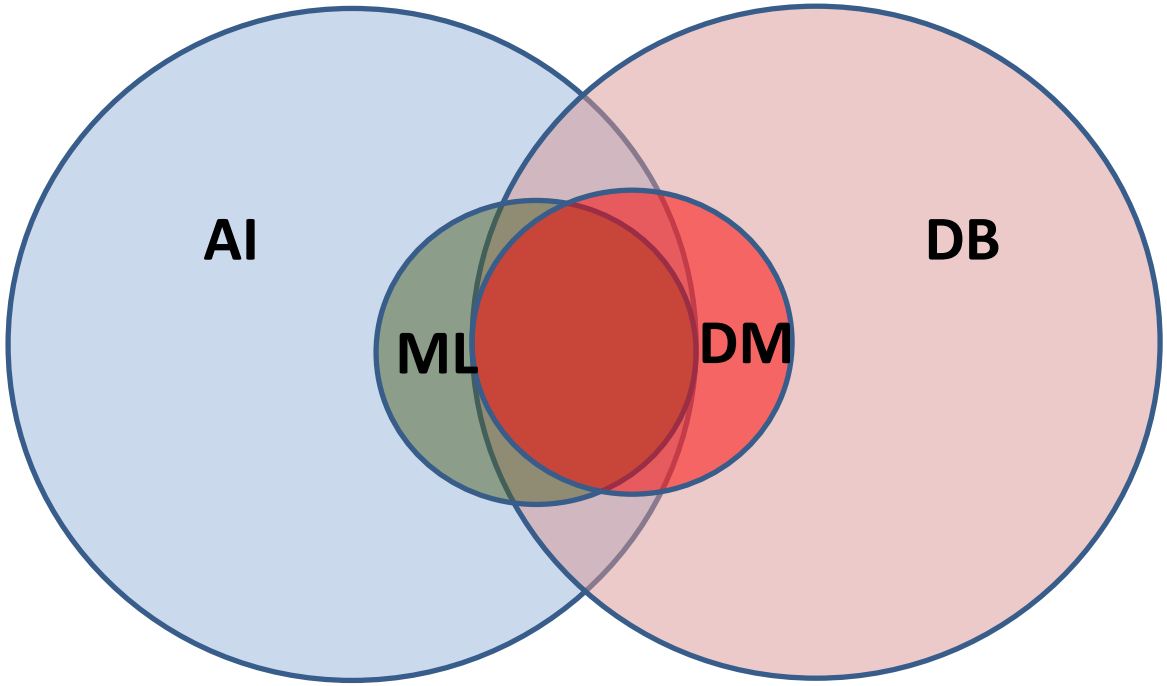




Relationship between AI, DM & ML



Relationship between AI, DM & ML



ML: focuses on prediction for unseen data based on properties from known data.
DM: focuses on discovery of unknown properties of data

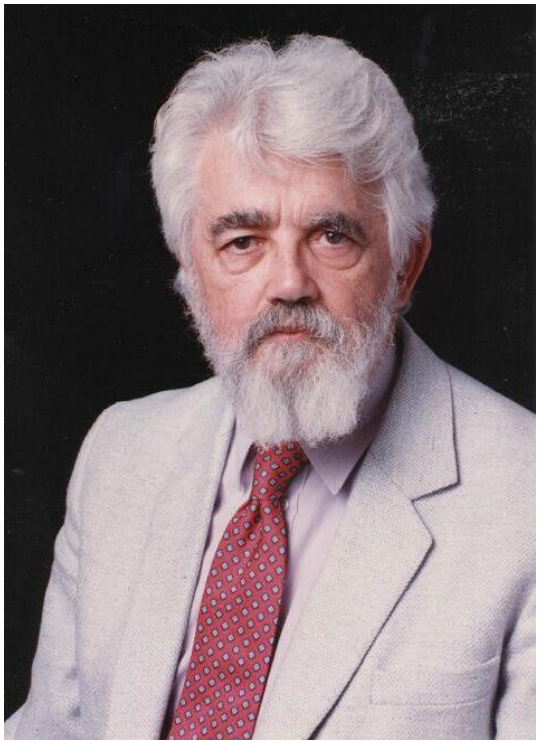
Topics (tentative)

- Task Environment & Performance Measure
- Concept of AI
- Data Mining
  - Data and patterns
  - Data exploration
  - Feature selection
- Machine Learning
  - Cross validation
  - Nearest neighbor
  - Naïve Bayes
  - Decision trees
  - Support vector machine
  - Deep learning

Artificial Intelligence

Where Does AI Come From?

- 1956: Happy birthday AI!
  - A two-month workshop at Dartmouth includes 10 major figures in AI: **John McCarthy** (logic), **Marvin Minsky** (neural network), **Claude Shannon** (information theory), **Nathaniel Rochester** (IBM 701), **Trenchard More** (array theory), **Arthur Samuel** (checkers program), **Ray Solomonoff** (algorithmic probability), **Oliver Selfridge** (machine perception), **Allen Newell** (information processing language), and **Herbert Simon** (cognitive science)
  - No breakthrough, but adopt McCarthy's new name to the field: artificial intelligence



Definitions

- What is intelligence?
  - Hint: what abilities do human have that are characteristic of intelligence?
    - Abstract concepts, mathematics, language, memory, planning, logical reasoning, emotions, morality, etc…
  - No clear definition, but ability to learn is an essential part of intelligence.

- What is artificial intelligence?

Definitions

The exciting new effort to make computers that think… machines with minds in the full and literal sense [Haugeland 85] [The automation of] activities that we associate with human thinking, such as decision making, problem solving, learning [Bellman 78]	The study of mental faculties through the use of computational models [Charniak & McDermott 85] The study of computations that make it possible to perceive, reason and act [Winston 92]
The art of creating machines that perform functions that require intelligence when performed by a human [Kurzweil 90] The study of how to make computers do things at which, at the moment, people are better [Rich & Knight 91]	A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes [Schalkoff 90] The branch of computer science that is concerned with the automation of intelligent behavior [Luger & Stubblefield 93]

Definitions

Systems that think like humans	**Systems that think rationally**
Systems that act like humans	**Systems that act rationally**

Definitions

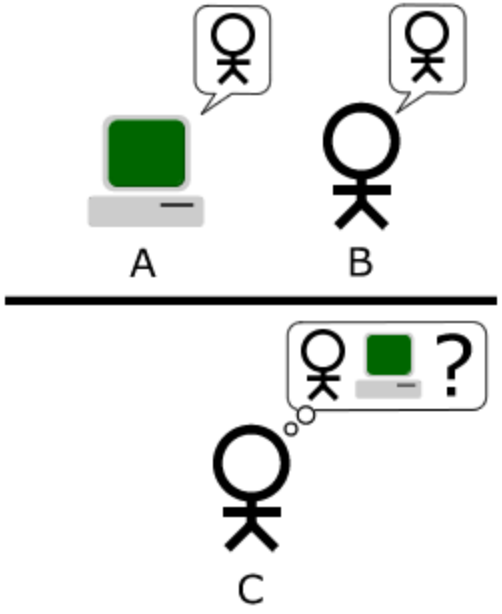
- Systems that think like humans
  - Need to know how humans think
  - Cognitive science: an interdisciplinary field that brings together computer models from AI and experimental techniques from psychology to try to construct precise and testable theories of the workings of human mind
  - Not covered in the course, but a fascinating area

Definitions

- Systems that think rationally
  - Aristotle (Greek philosopher): try to codify “right thinking”, i.e., irrefutable reasoning processes
  - Syllogism: argument structures that always yield correct conclusions when given correct premises. Initiated the field Logic

Definitions

- Systems that act like humans
  - The Turing Test:
    - A human interrogator, after posing some written questions, can not tell whether the written responses come from a person or not

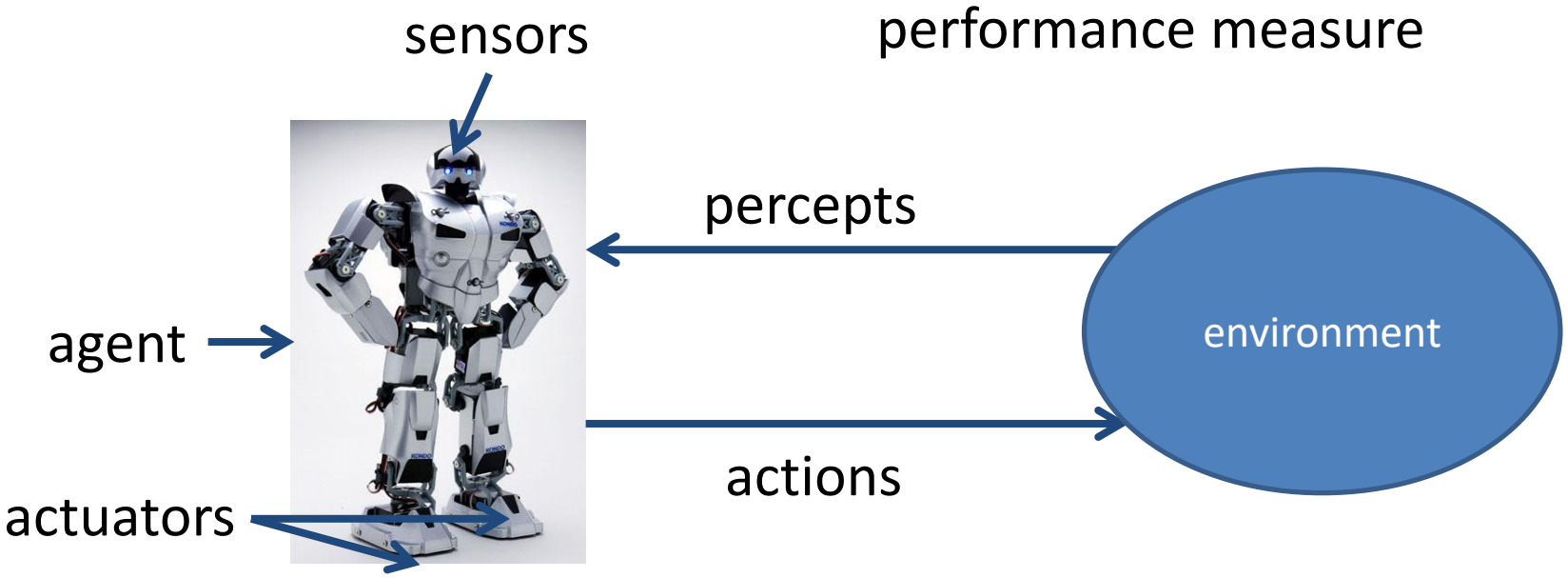


  - Suggested major components of AI: natural language processing, knowledge representation, automated reasoning, and machine learning

Definitions

- Systems that act rationally
  - Rational behavior: doing the right thing
  - Rational agent approach
    - Agent: something that perceives and acts
    - Rational agent: acts so to achieve the best outcome or the best expected outcome
- Better than thinking rationally
  - More general, because the correct inference is just one of several possible mechanisms for achieving rationality
  - More amendable, because the standard of rationality is clearly defined and completely general
- This is the approach we will cover in this course
  - General principles of rational agents
  - Components for constructing rational agents

Agents and Environment



- Agents include humans, robots, taxi, thermostats…
  - The agent function maps percepts to actions $f: P^* \rightarrow A$
  - The agent program runs on the physical architecture to produce f

Rational Agents

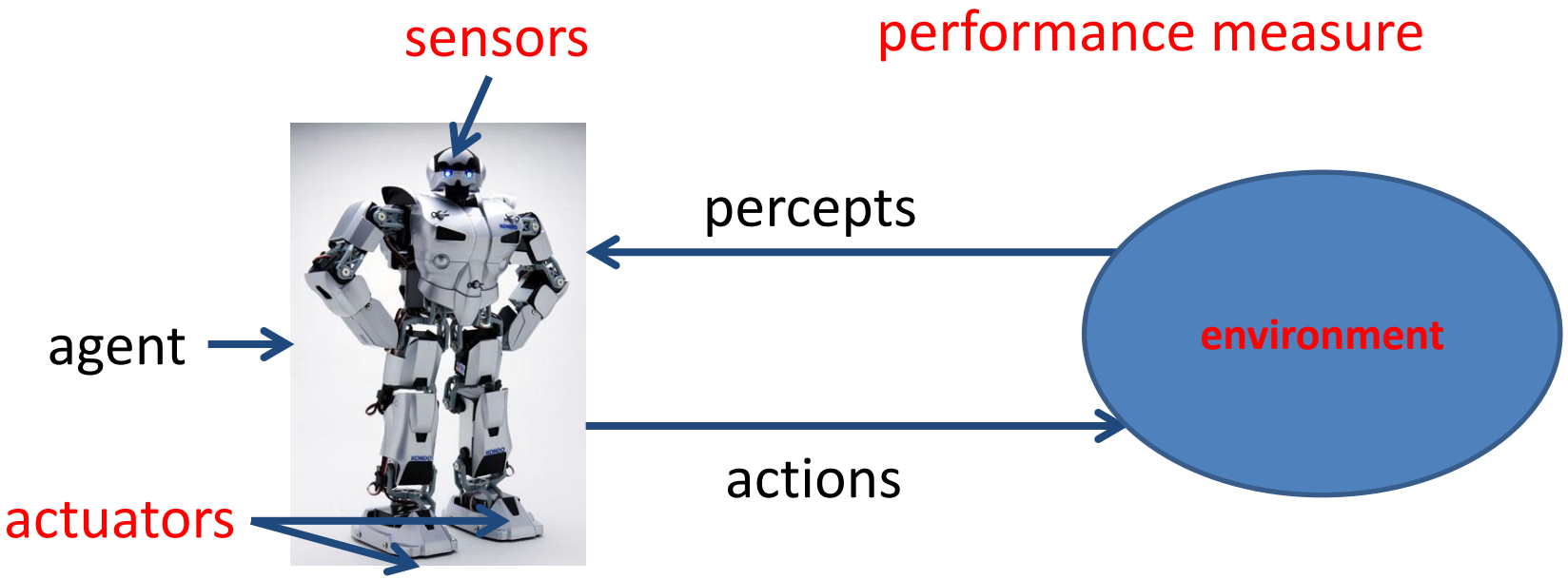
- A rational agent “does the right thing”
- Components for rationality
  - Performance measure that defines the criterion of success
  - Agent’s prior knowledge of the environment
  - The actions that the agent can perform
  - The agent’s percept sequence to date
- Definition: for each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given the evidence provided by the percept sequence and whatever build-in knowledge the agent has

Rationality

- Rationality ≠ Omniscience, Perfection, Success
- Rationality → Exploration, Learning, Autonomy

- Rationality maximizes the expected performance, while perfection maximizes the actual performance

Agents and Environment



PEAS

- Specify the task environment:
  - Performance measure, Environment, Actuators, Sensors

Example: Autonomous Taxi

Perf:

Envi:

Actu:

Sens:

Properties of Task Environment

- Fully observable vs. partially observable
  - Def: if sensors give access to the complete state of the environment at any time point, the environment is fully observable
  - Pros: know all the aspects related to the choice of my action, don't need to maintain an internal state to keep track of the world
  - Note: not available for most cases, because of the noise and inaccurate sensors

Properties of Task Environment

- Deterministic vs. stochastic
  - Def: if the next state of the environment is completely determined by the current state and the actions of the agent, the environment is deterministic
  - Pros: don't need to worry about uncertainty if fully observable and deterministic
  - Note: most of the multiagent environments are stochastic

Properties of Task Environment

- Episodic vs. sequential
  - Def: if the agent's experience can be divided into atomic episodes, and the next episode doesn't depend on the actions taken in previous episode, the environment is episodic
  - Pros: the action in each episode only depends on the episode itself
  - Note: in sequential environments, current actions could affect all future actions, i.e., short-term actions have long-term consequences

Properties of Task Environment

- Static vs. dynamic
  - Def: if the environment cannot change while the agent is thinking, the environment is static
  - Pros: don't need to keep looking at the world while thinking
  - Note: dynamic environments continuously asking the agent what to do, if the agent is still thinking, that counts as deciding "to do nothing"

Properties of Task Environment

- Discrete vs. continuous
  - Def: the distinction can be applied to the state, the way time is handled, and percepts and actions of the agent
  - Note: an environment can be mixed, e.g., taxi driving has continuous state, time, and actions, but discrete percepts, i.e., frames of digital cameras

Properties of Task Environment

- Single agent vs. multiagent
  - Def: seems trivial, but think again!
    - Should we treat any object in the environment as an agent, or a stochastically behaving objects, such as other people or side bar advertisements while online shopping
  - Note: depending on whether maximizing B's performance minimizes A's performance, there can be competitive multiagent environment and cooperative multiagent environment

Properties of Task Environment

- Hardest case: partially observable, stochastic, sequential, dynamic, continuous and multiagent, i.e., real-world!

Examples

5	3			7				
6			1	9	5			
	9	8					6	
8				6				3
4			8		3			1
7				2				6
	6					2	8	
			4	1	9			5
				8			7	9

Sudoku

Poker (Texas Hold'em)

Internet Shopping

Taxi Driving

Examples

Sudoku	Poker	Internet Shopping	Taxi Driving

Fully observable or Partially observable?

Examples

Sudoku	Poker	Internet Shopping	Taxi Driving
Fully Observable	Partially Observable	Partially Observable	Partially Observable

Deterministic or Stochastic?

Examples

Sudoku	Poker	Internet Shopping	Taxi Driving
Fully Observable	Partially Observable	Partially Observable	Partially Observable
Deterministic	Stochastic	Stochastic	Stochastic

Sequential or Episodic?

Examples

Sudoku	Poker	Internet Shopping	Taxi Driving
Fully Observable	Partially Observable	Partially Observable	Partially Observable
Deterministic	Stochastic	Stochastic	Stochastic
Sequential	Sequential	Episodic	Sequential

Static or Dynamic?

Examples

Sudoku	Poker	Internet Shopping	Taxi Driving
Fully Observable	Partially Observable	Partially Observable	Partially Observable
Deterministic	Stochastic	Stochastic	Stochastic
Sequential	Sequential	Episodic	Sequential
Static	Static	Dynamic	Dynamic

Discrete or Continuous?

Examples

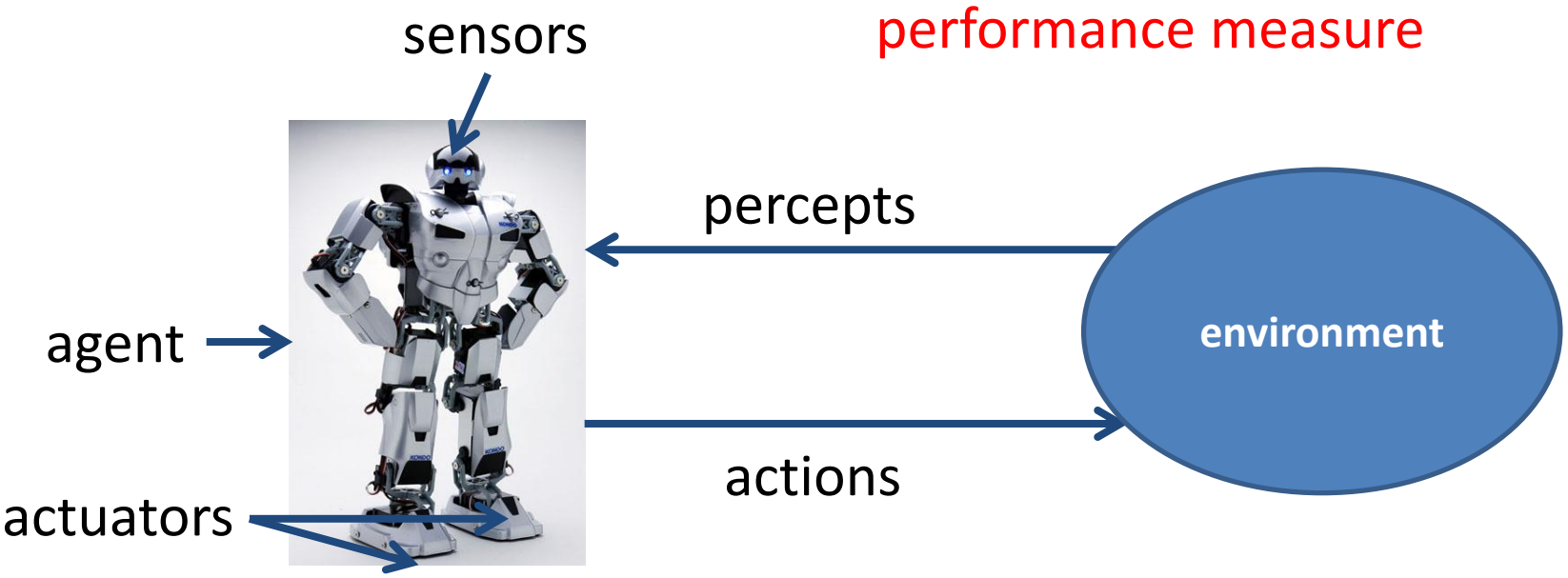
Sudoku	Poker	Internet Shopping	Taxi Driving
Fully Observable	Partially Observable	Partially Observable	Partially Observable
Deterministic	Stochastic	Stochastic	Stochastic
Sequential	Sequential	Episodic	Sequential
Static	Static	Dynamic	Dynamic
Discrete	Discrete	Discrete	Continuous

Single Agent or Multiagent?

Examples

Sudoku	Poker	Internet Shopping	Taxi Driving
Fully Observable	Partially Observable	Partially Observable	Partially Observable
Deterministic	Stochastic	Stochastic	Stochastic
Sequential	Sequential	Episodic	Sequential
Static	Static	Dynamic	Dynamic
Discrete	Discrete	Discrete	Continuous
Single Agent	Multiagent	Multiagent	Multiagent

Agents and Environment



Performance Measure

- Performance measure matters
- E.g., coin change problem
  - Find a way of using coins to pay for a value

2 Dollar Coin
"Toonie"

1 Dollar Coin
"Loonie"

25 Cent Coin
"Quarter"

10 Cent Coin
"Dime"

5 Cent Coin
"Nickel"

1 Cent Coin
"Penny"

Performance Measure

- Question 1: how to pay if I want the smallest number of coins, in Canadian/US coin system?

2 Dollar Coin "Toonie"

1 Dollar Coin "Loonie"

25 Cent Coin "Quarter"

10 Cent Coin "Dime"

5 Cent Coin "Nickel"

1 Cent Coin "Penny"

Performance Measure

- Question 1: how to pay if I want the smallest number of coins, in Canadian/US coin system?

2 Dollar Coin "Toonie" | 1 Dollar Coin "Loonie" | 25 Cent Coin "Quarter" | 10 Cent Coin "Dime" | 5 Cent Coin "Nickel" | 1 Cent Coin "Penny"

- Question 2: how to pay if I want the smallest number of coins, in an arbitrary coin system?

25 Cent Coin "Quarter" | 10 Cent Coin "Dime" | 1 Cent Coin "Penny"

Performance Measure

- Questions 3: how to pay if I want the minimum total weights of the coins?

- Question 4: how to pay if I want the maximum total weight with penalty for carrying too many pennies?