Lecture #1: Introduction

Janardhan Rao (Jana) Doppa

School of EECS, Washington State University

Introductions

- **Jana Doppa, Huie-Rogers Chair Associate Professor**
- At WSU since Fall-2014
- PhD from Oregon State University (2014)
- Masters from IIT Kanpur (2006)
- Passionate about AI, machine learning, computing and data-driven solutions for real-world problems. Doing research for ~18 years.
- Like teaching courses on these topics

Introduce Yourself

- Your name
- Your discipline
- Motivation behind taking this class
- How are you dealing with pandemic and beyond?

Course Logistics

- CptS 570: Machine Learning
  - **Class Timings** – Tue and Thu 4:20 to 5:35pm (Spark 335)
  - **Instructor** – Jana Doppa (EME 133)
  - **Office Hours** – Mon 4-5pm (EME 133): From next week!
  - **Teaching Assistant** – Subhankar Ghosh
    - subhankar.ghosh@wsu.edu

Course Logistics

- CptS 570: Machine Learning
  - **Course announcements and discussions** – Piazza
  - **Lecture Notes** – Slides and notes will be posted on Piazza
  - Please use Piazza for communicating with course staff (Professor and TA). You will get faster response from me or TA or your classmates. Please avoid using email for communication.

Grading Policy

- **4 Homework assignments (36%)**
- **2 Exams: Mid-term (20%) and Final (20%)**
- **1 Course Project (20%)**
  - Can be done in small groups (one or two students)
- **Class Participation (4%)**
  - Piazza and in-class

Late Policy

- All assignments, project proposal/report are due at midnight
- **Late Policy**
  - 0-24 hours late -- 80% of the final score
  - 24-48 hours late -- 50% of the final score
  - Beyond 48 hours -- 0%
- All submissions will be handled through **Canvas**

Course Pre-requisites

- **Assume strong programming experience**
  - You are free to use any programming language
- **NO prior knowledge of Artificial Intelligence is needed**
  - This course stands on its own
- **Basic knowledge of the following is expected**
  - Probability and Statistics
  - Linear algebra and Multivariate calculus
  - Basic numerical optimization (e.g., gradient descent)
  - Algorithmic paradigms and Search algorithms

Course Materials

- **We will NOT follow a fixed textbook for this course**
  - Instructor will provide slides and lecture notes
  - Slides and notes will be posted on Piazza site
- **Optional Textbooks**
  - A Course in Machine Learning, by Hal Daume' III (free online book and easy to follow)
  - Machine Learning, by Kevin Murphy (Rich mathematical treatment)
  - Machine Learning, by Tom Mitchell
  - Pattern Recognition and Machine Learning, by Chris Bishop

Machine Learning is Everywhere

- "If you invent a breakthrough in artificial intelligence, so machines can learn," Mr. Gates responded, "that is worth 10 Microsofts."

(Quoted in NY Times, Monday March 3, 2004)

Machine Learning is Everywhere

- Spam filtering

googleteam GOOGLE LOTTERY WINNER! CONTACT

From: googleteam **To:**
Subject: GOOGLE LOTTERY WINNER! CONTACT YOUR AGENT TO CLAIM YOUR PRIZE.

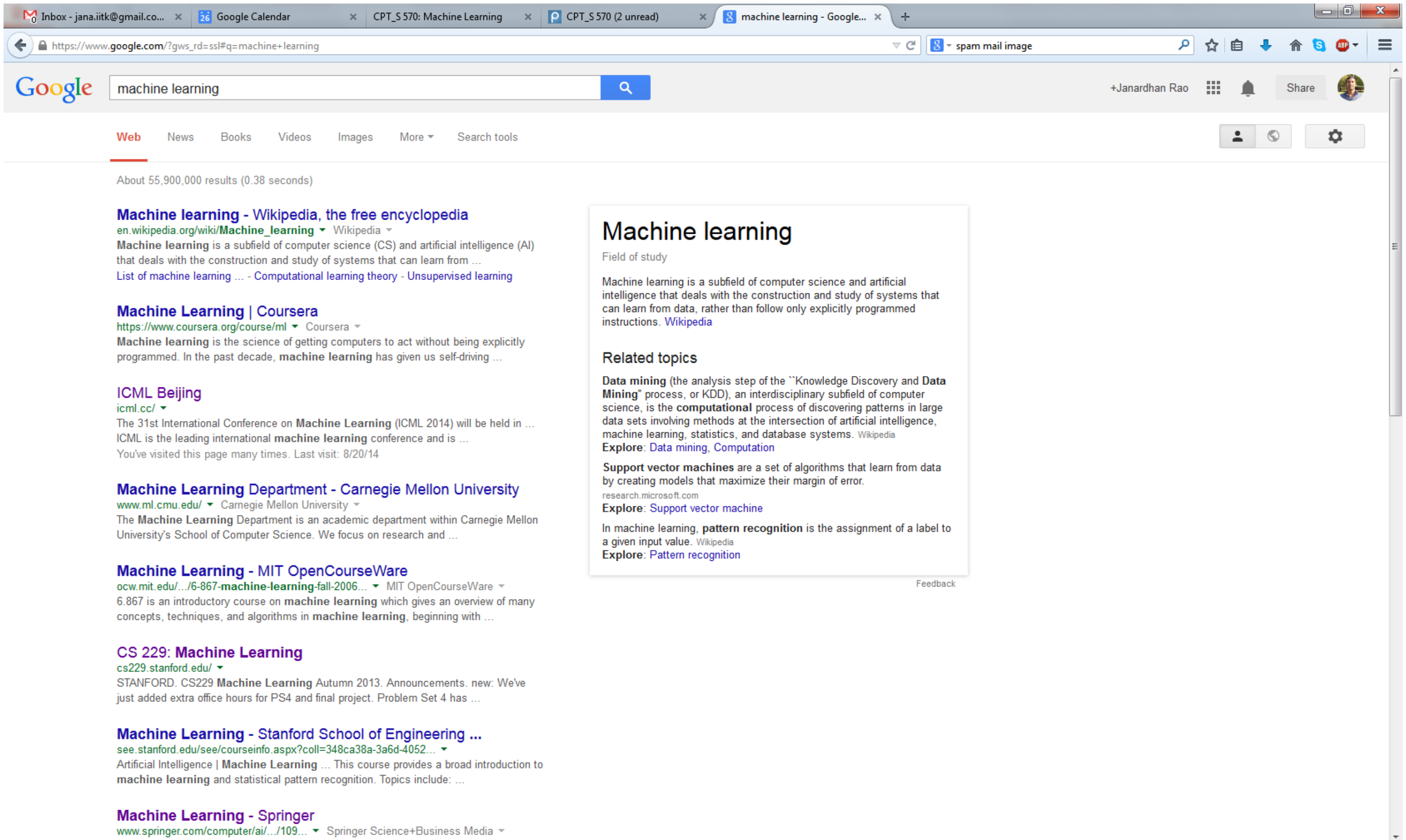
GOOGLE LOTTERY INTERNATIONAL
INTERNATIONAL PROMOTION / PRIZE AWARD .
(WE ENCOURAGE GLOBALIZATION)
FROM: THE LOTTERY COORDINATOR,
GOOGLE B.V. 44 9459 PE.
RESULTS FOR CATEGORY "A" DRAWS
Congratulations to you as we bring to your notice, the results of the First Ca
inform you that your email address have emerged a winner of One Million (1,(
money of Two Million (2,000,000.00) Euro shared among the 2 winners in this
email addresses of individuals and companies from Africa, America, Asia, Au
CONGRATULATIONS!
Your fund is now deposited with the paying Bank. In your best interest to avo
award strictly from public notice until the process of transferring your claims
NOTE: to file for your claim, please contact the claim department below on e
****

Machine Learning is Everywhere

- Optical Character Recognition (OCR)

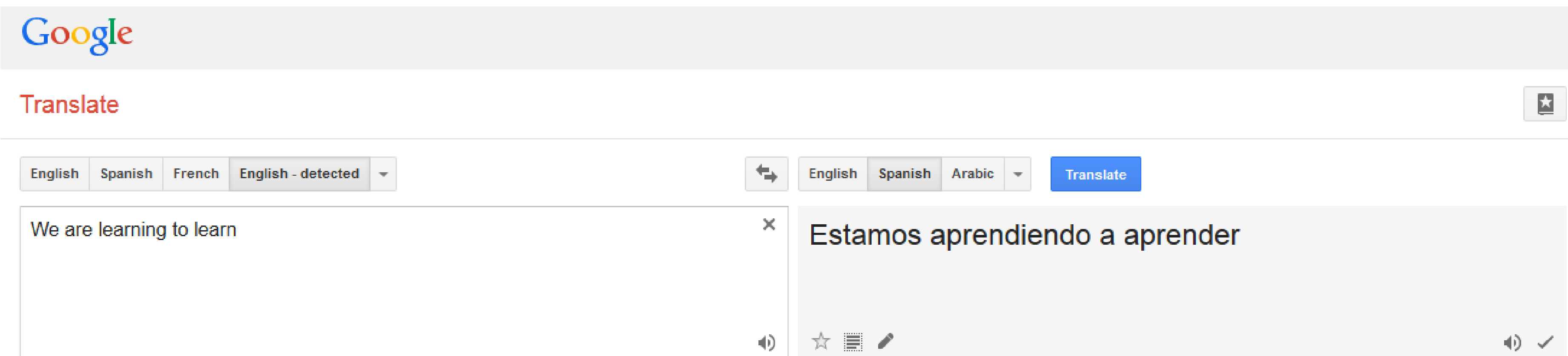
Machine Learning is Everywhere

- Search engines



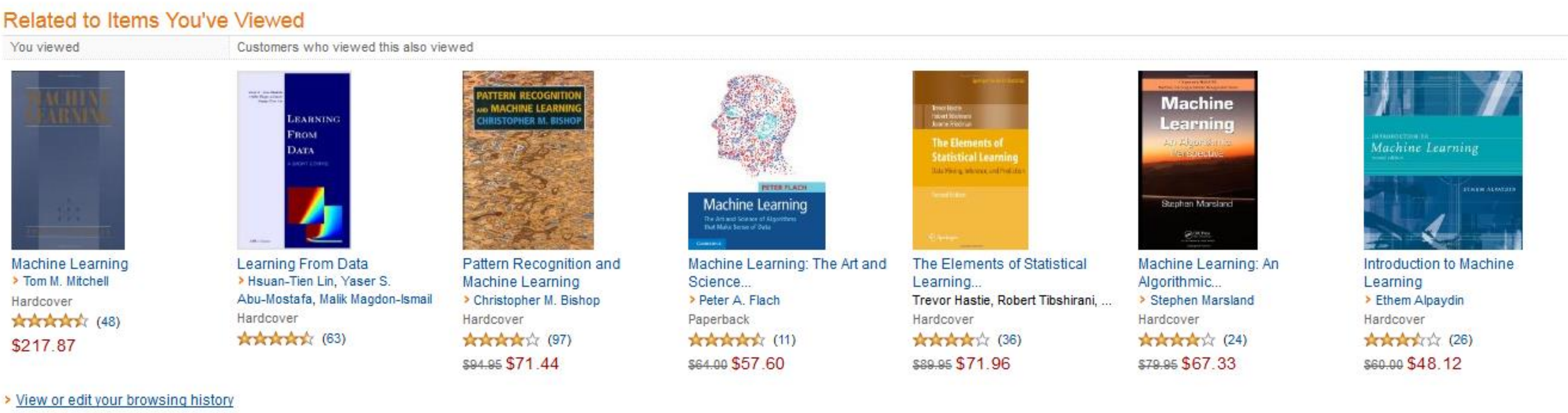
Machine Learning is Everywhere

- Automatic Translation



Machine Learning is Everywhere

- Recommendation Engines



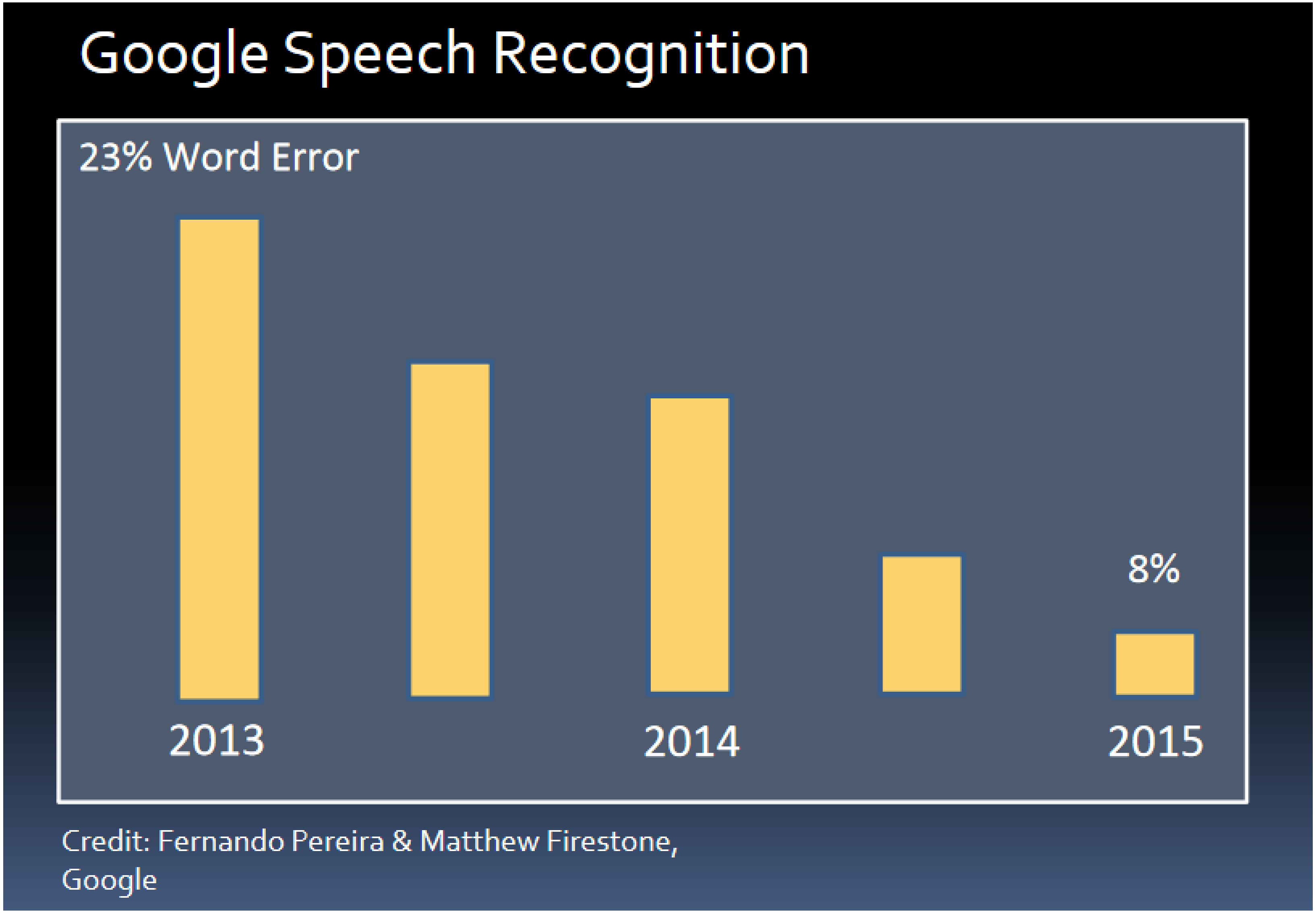
Machine Learning is Everywhere

- Self-driving cars

Google's Self Driving Car for Blind People

by EDITORS on Apr 6, 2012 · 4:07 pm

ML Successes: Perception



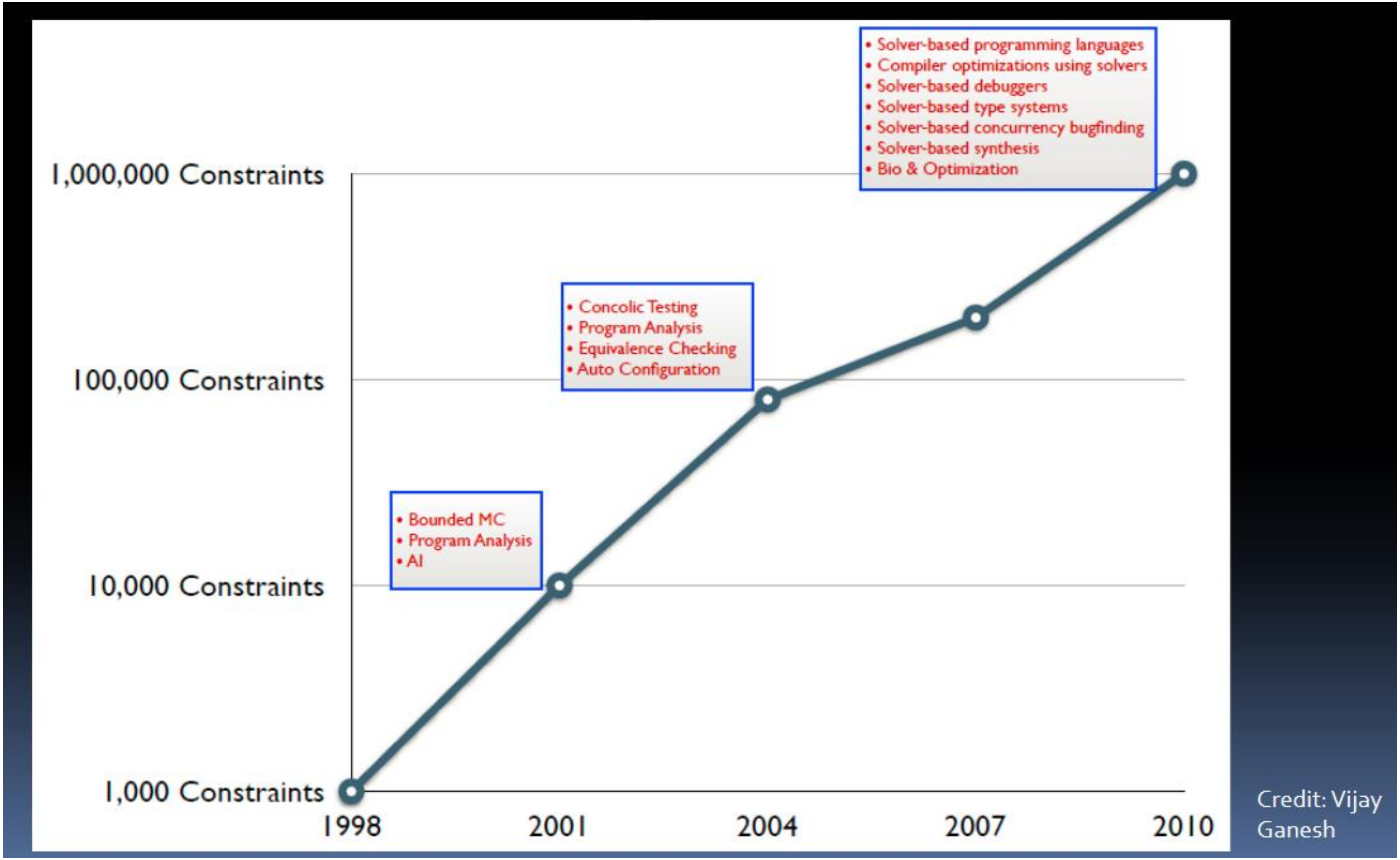
Credit: Tom Dietterich

ML Successes: Image Captioning

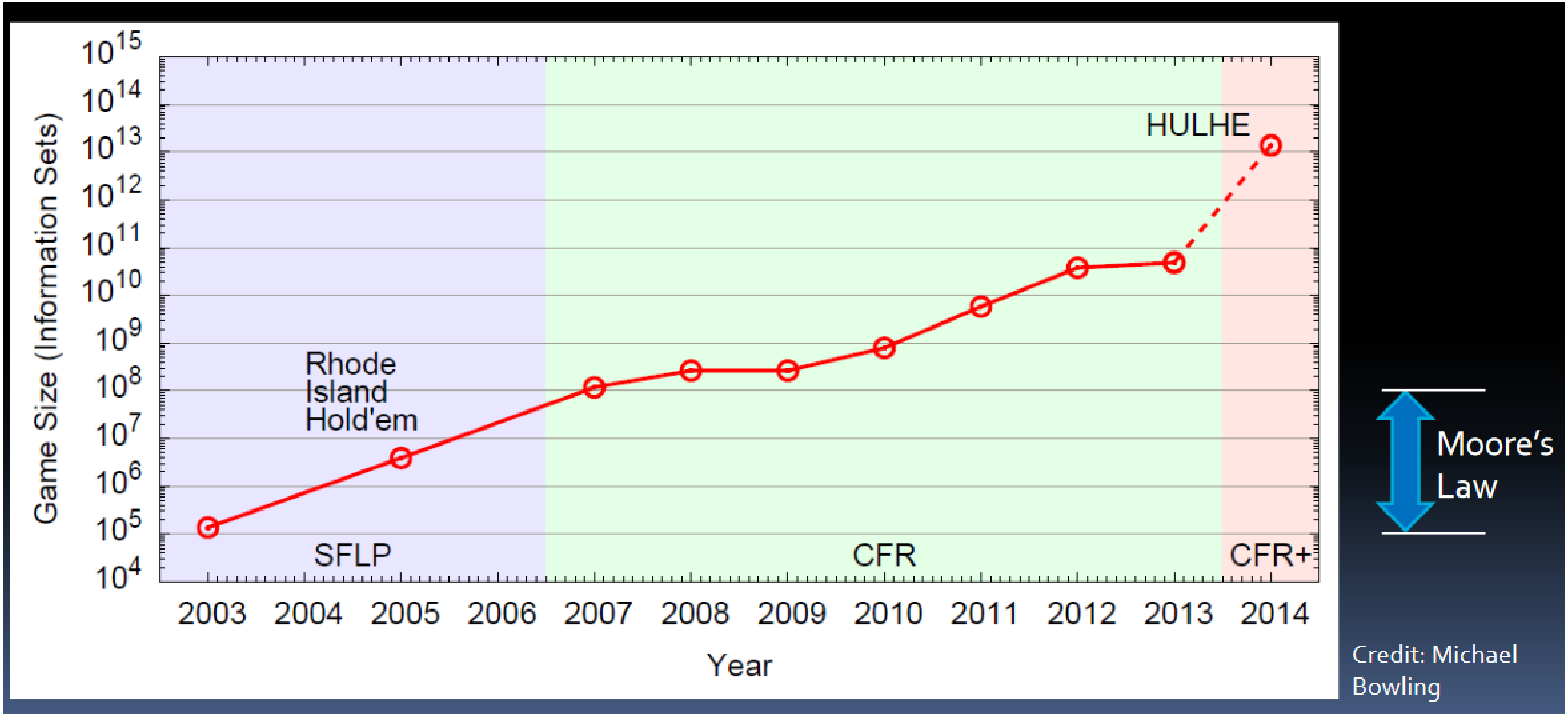
ML Successes: Perception + Translation

ML Successes: Skype Translator

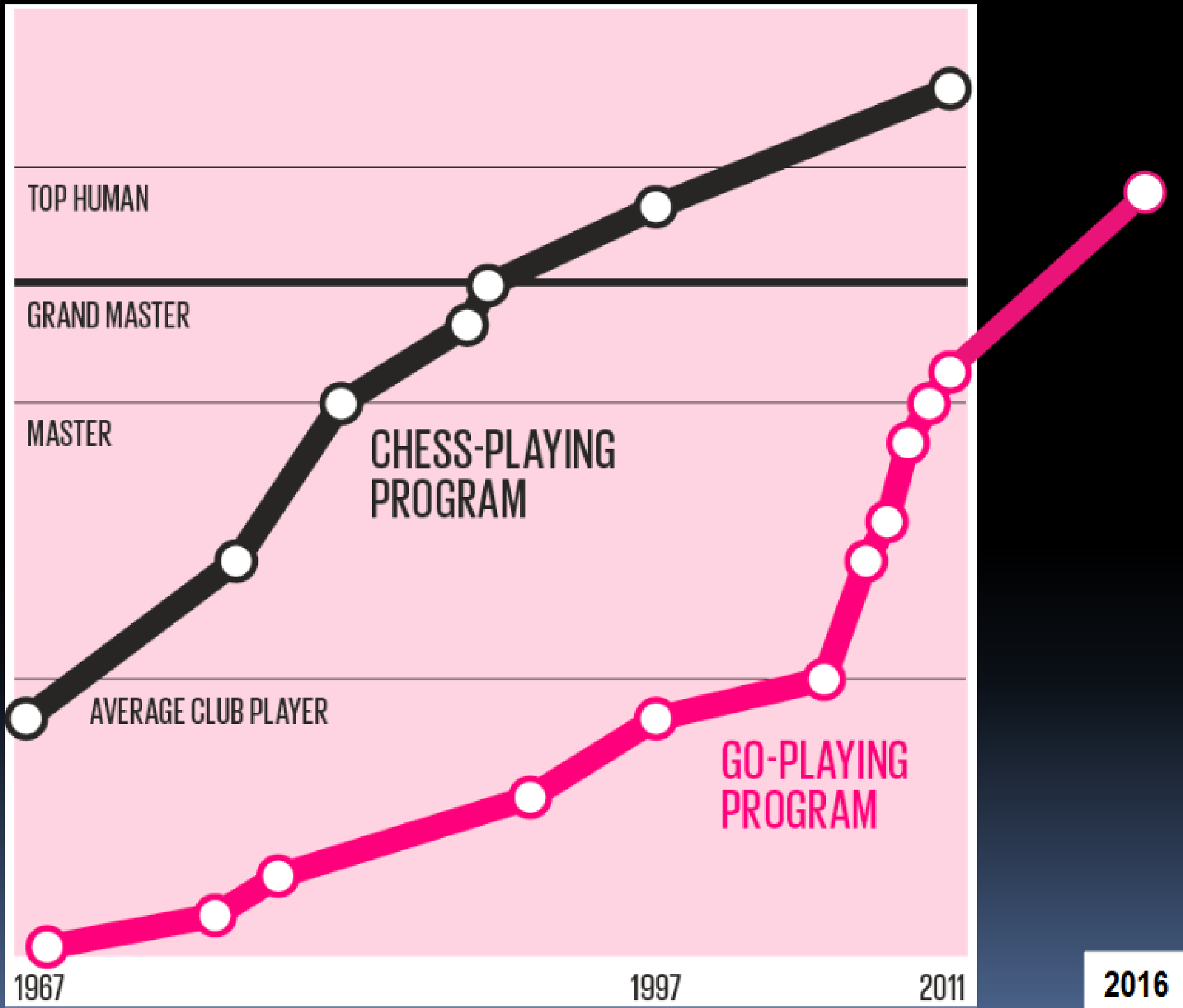
ML Successes: Reasoning (SAT)



ML Successes: Poker



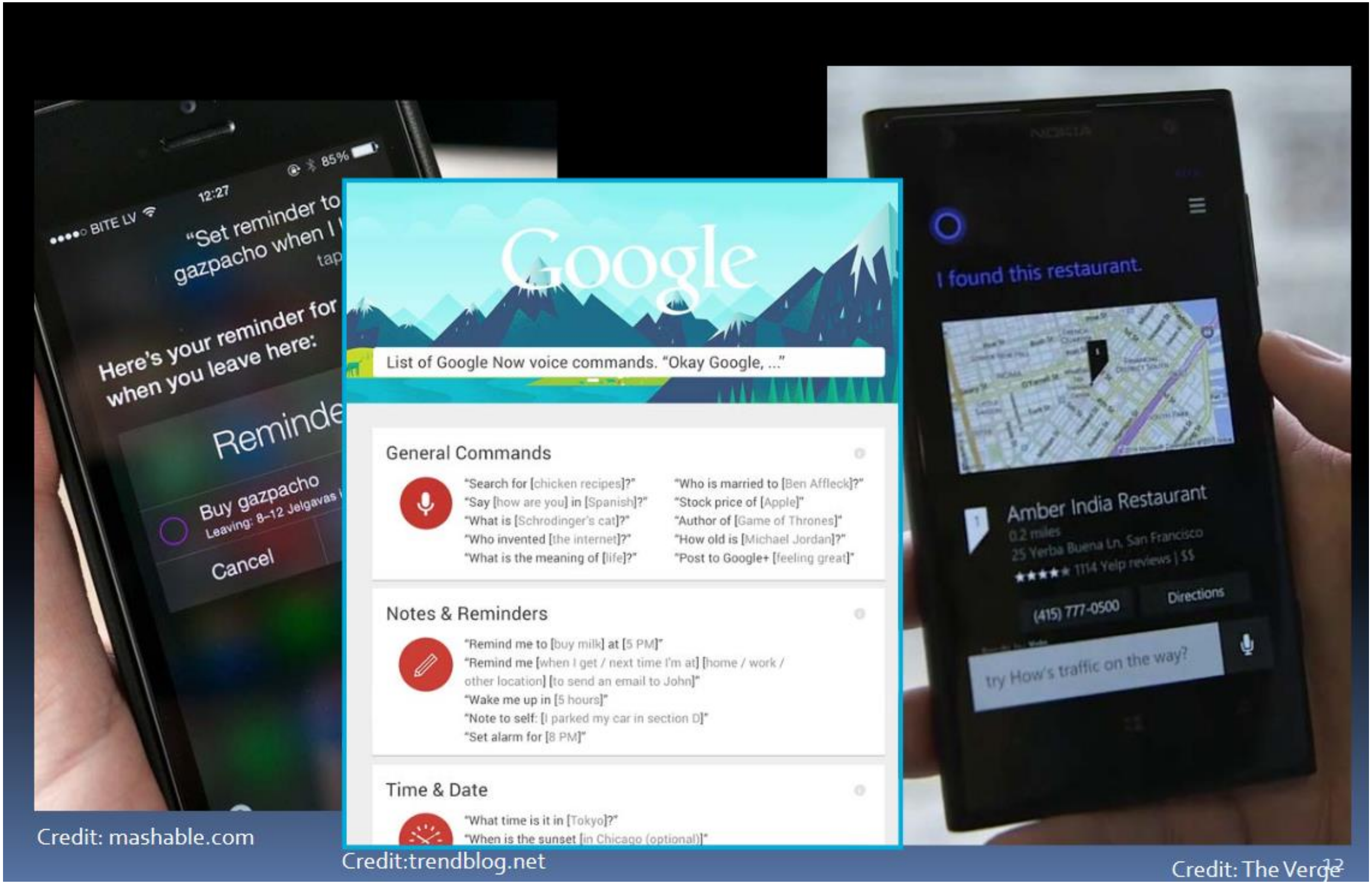
ML Successes: Chess and Go



Silver, et al. (2016) *Nature*
Deep Learning +
Monte Carlo Tree Search

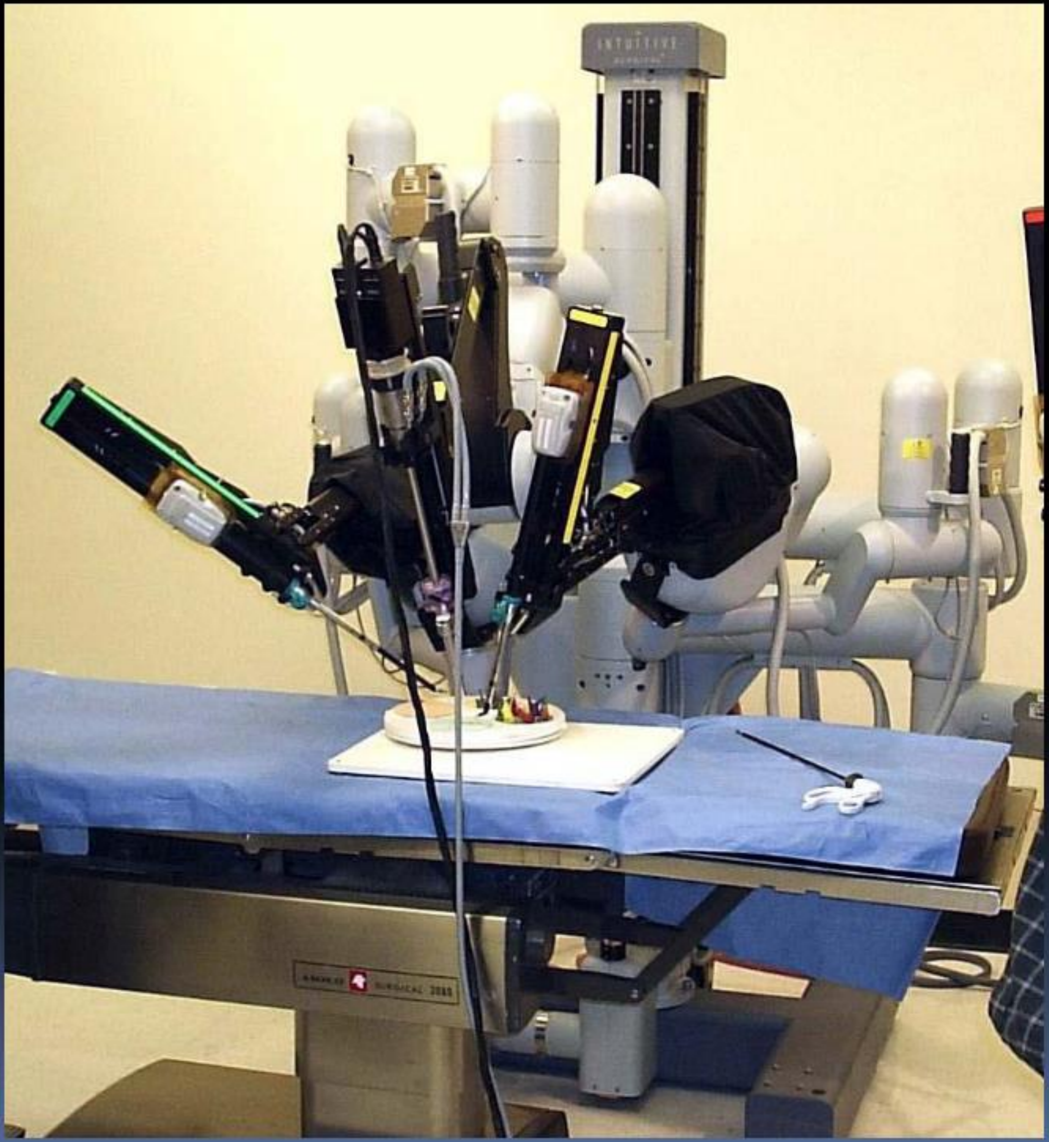
Credit: Martin Mueller

ML Successes: Personal Assistants



High-Stakes Applications: Self-Driving Cars

High-Stakes Applications: Automated Surgical Assistants



DaVinci

High-Stakes Applications: AI Hedge Funds

WIRED

CADE METZ BUSINESS 01.25.16 7:00 AM

THE RISE OF THE ARTIFICIALLY INTELLIGENT HEDGE FUND

High-Stakes Applications: Power Grid Control

CONTROLLING THE POWER GRID WITH ARTIFICIAL INTELLIGENCE

02.07.2015

Credit: EBM Netz AG

DARPA Exploring Ways to Protect Nation's Electrical Grid from Cyber Attack

Effort calls for creation of automated systems to restore power within seven days or less after attack

Credit: DARPA

High-Stakes Applications: Autonomous Weapons

Northroop Grumman X-47B

Credit: Wikipedia

Samsung SGR-1

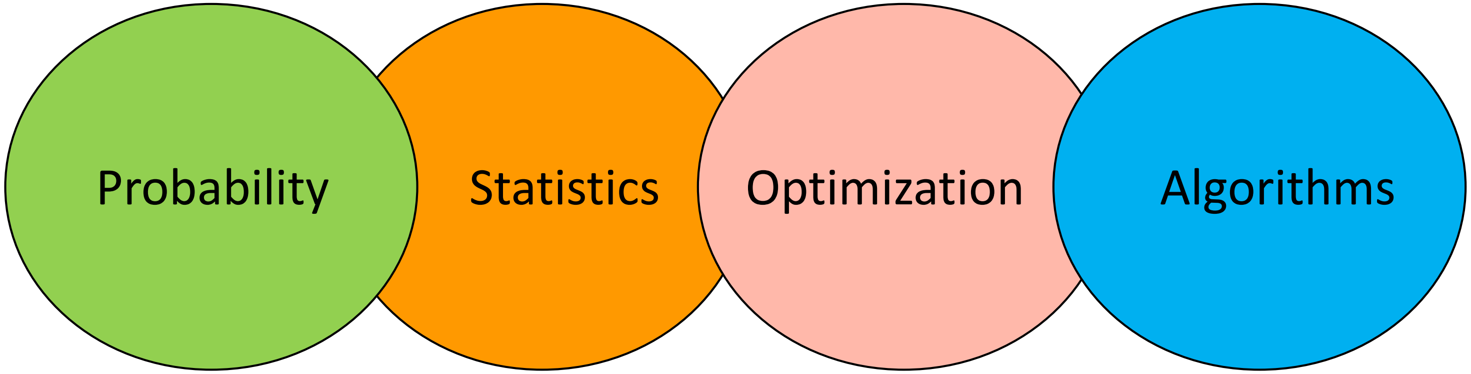
Credit: AFP/Getty Images

UK Brimstone Anti-Armor Weapon

Credit: Duch.seb - Own work, CC BY-SA 3.0

What is Machine Learning?

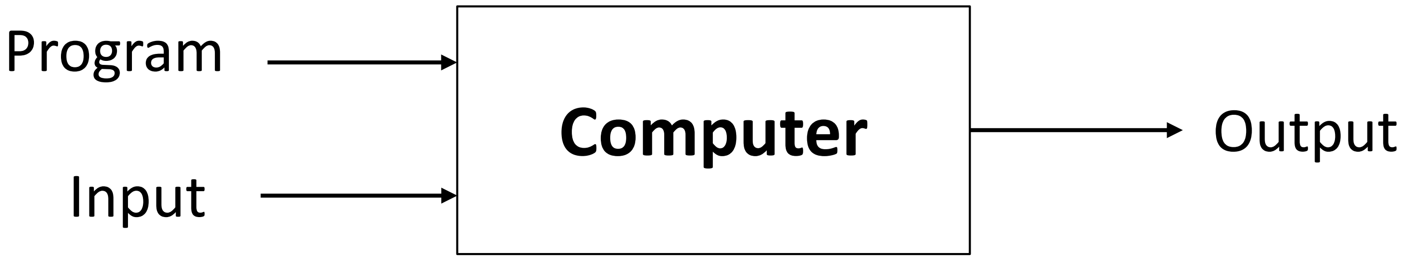
- **Machine learning is the branch of engineering that develops technology for automated inference**
  - It combines



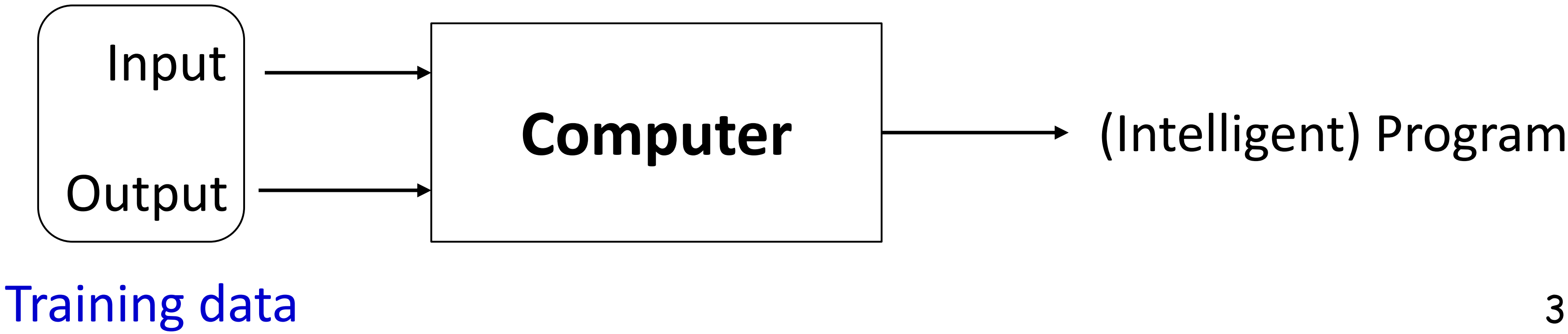
What is Machine Learning?

- **Machine learning = Automating Automation**

Traditional Programming



Machine Learning



Magic?

No, more like gardening

- **Seeds** = Algorithms
- **Nutrients** = Data
- **Gardener** = You
- **Plants** = Programs

- Credit: Liang Huang

Future of Software Engineering

- "See when AI comes, I'll be long gone (being replaced by autonomous cars) but the programmers in those companies will be too, by automatic program generators." --- an Uber driver to an ML prof

- Credit: Liang Huang

Learning Paradigms

- **Supervised Learning –** main focus of this course

- **Semi-Supervised Learning**
- **Unsupervised Learning**
- **Active Learning**
- **Reinforcement Learning**

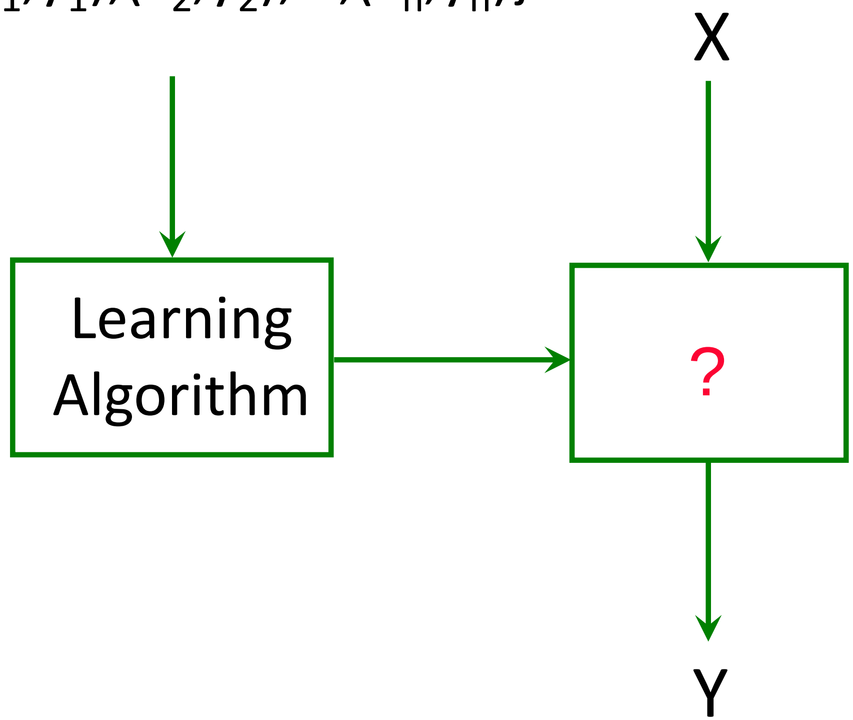
Supervised Learning

Learning a Classifier

( , male)

Training Data

$\{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\}$

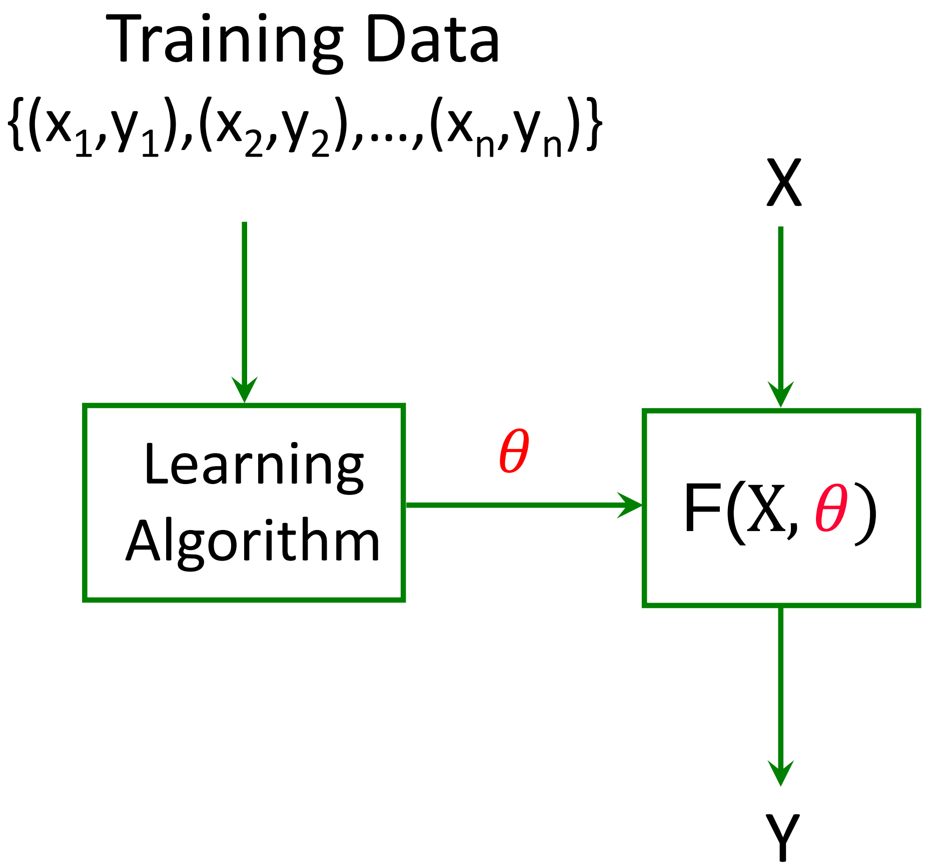


<u>Example problem:</u>

X - image of a face

$Y \in \{\text{male, female}\}$

Learning a Classifier

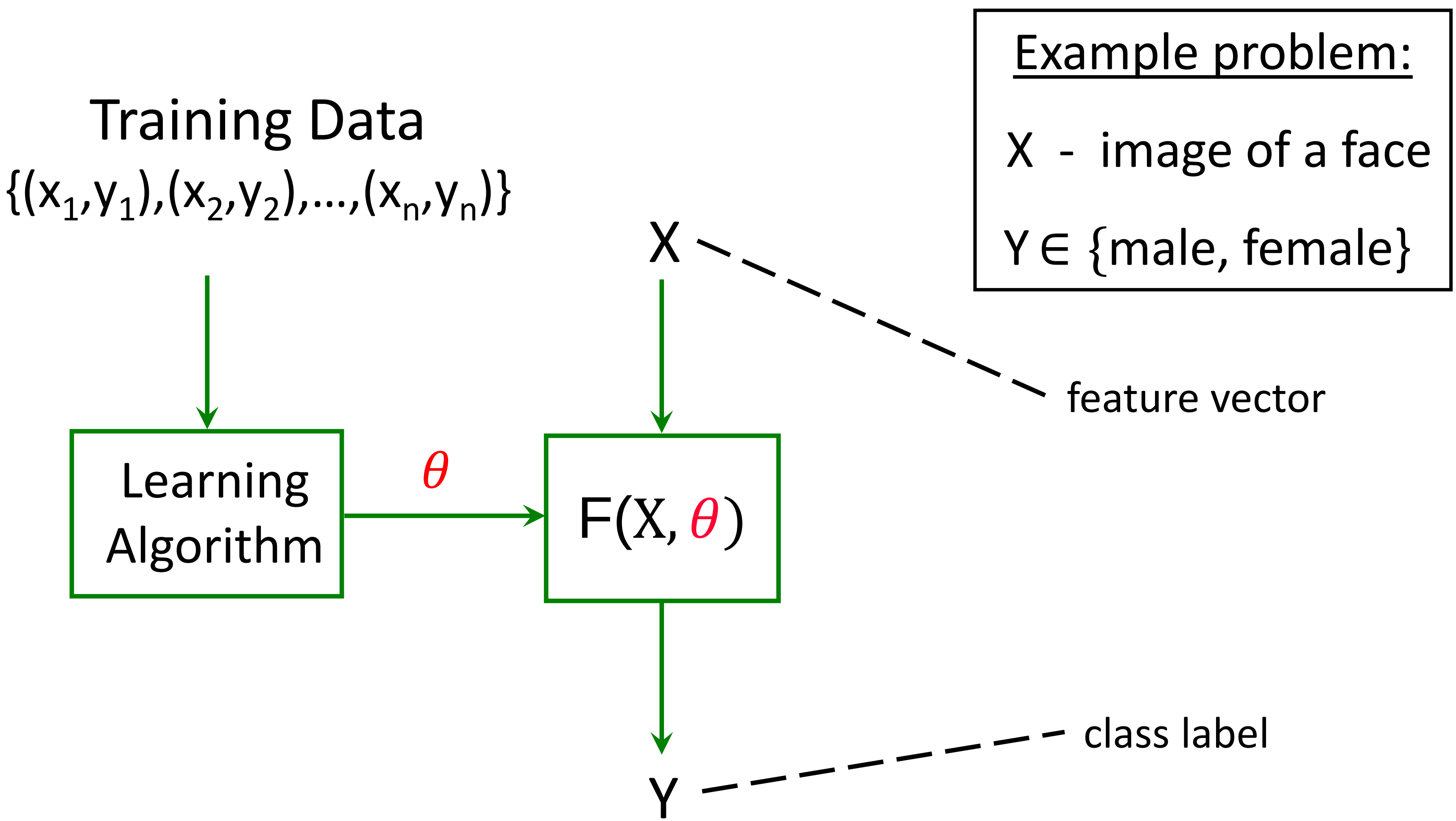


<u>Example problem:</u>

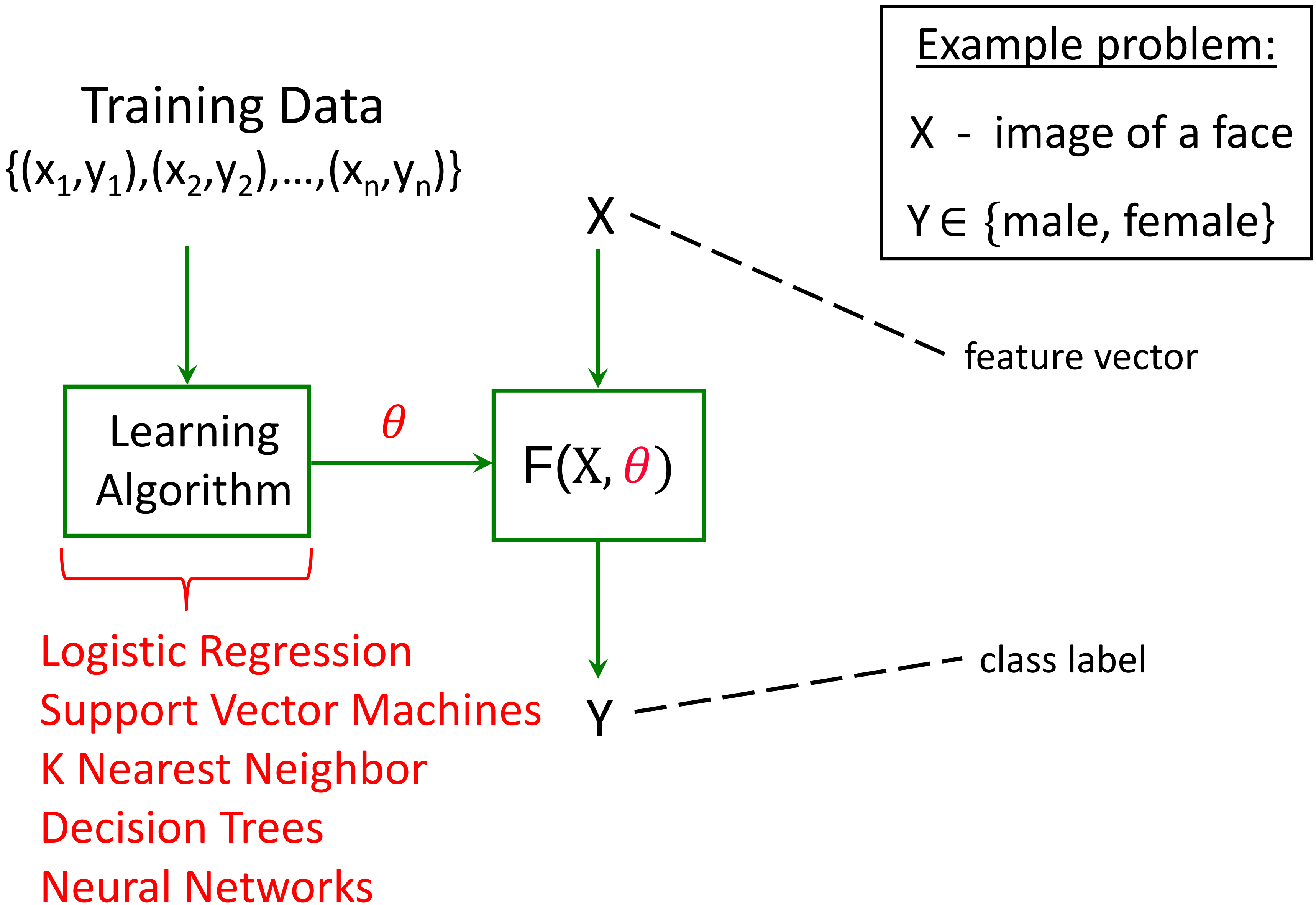
X - image of a face

$Y \in$ {male, female}

Learning for Simple Outputs



Learning for Simple Outputs

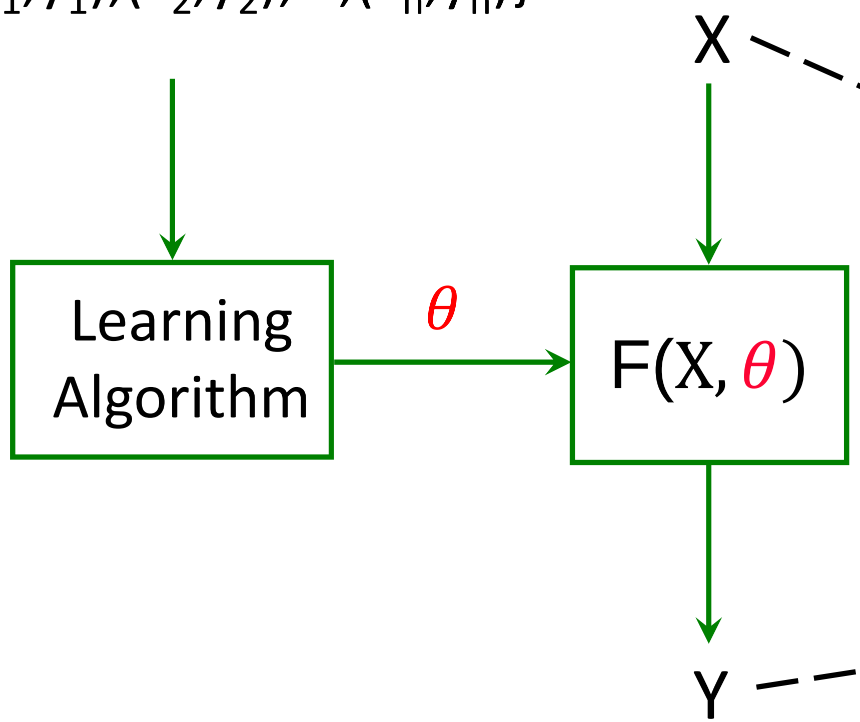


Learning for Structured Outputs

Part-of-Speech Tagging

Training Data

$\{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\}$



English Sentence:

"The cat ran"

Part-of-Speech Sequence:

<article> <noun> <verb>

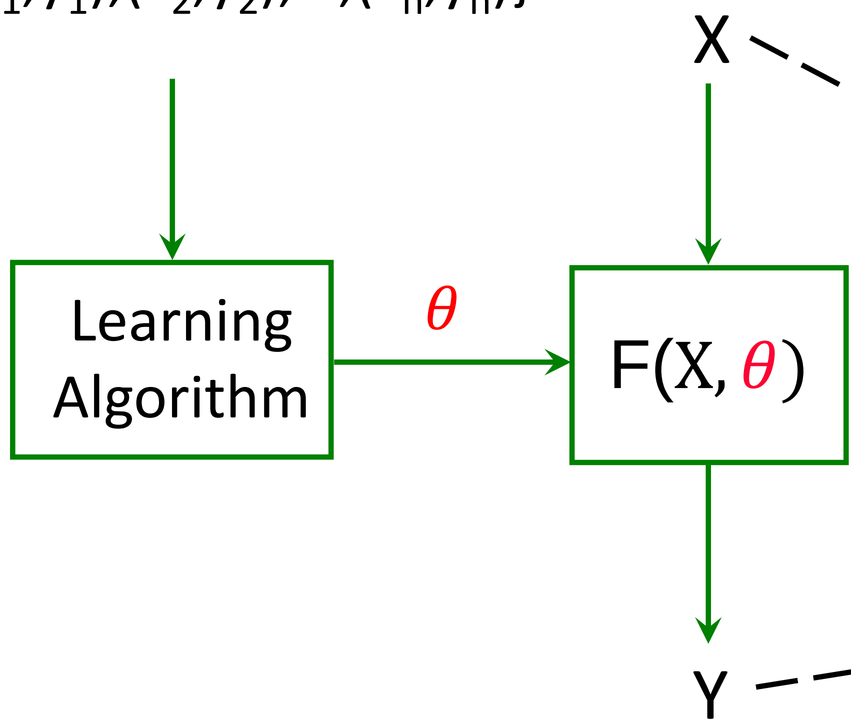
$\boldsymbol{Y} =$ set of all possible POS tag sequences

Exponential !!

Learning for Structured Outputs

Training Data

$\{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\}$



Co-reference Resolution

Text with input mentions:

*"**Barack Obama** nominated **Hillary Clinton** as his **secretary of state** on Monday. **He** chose **her** because **she** had foreign affair experience as a former **First Lady**."*

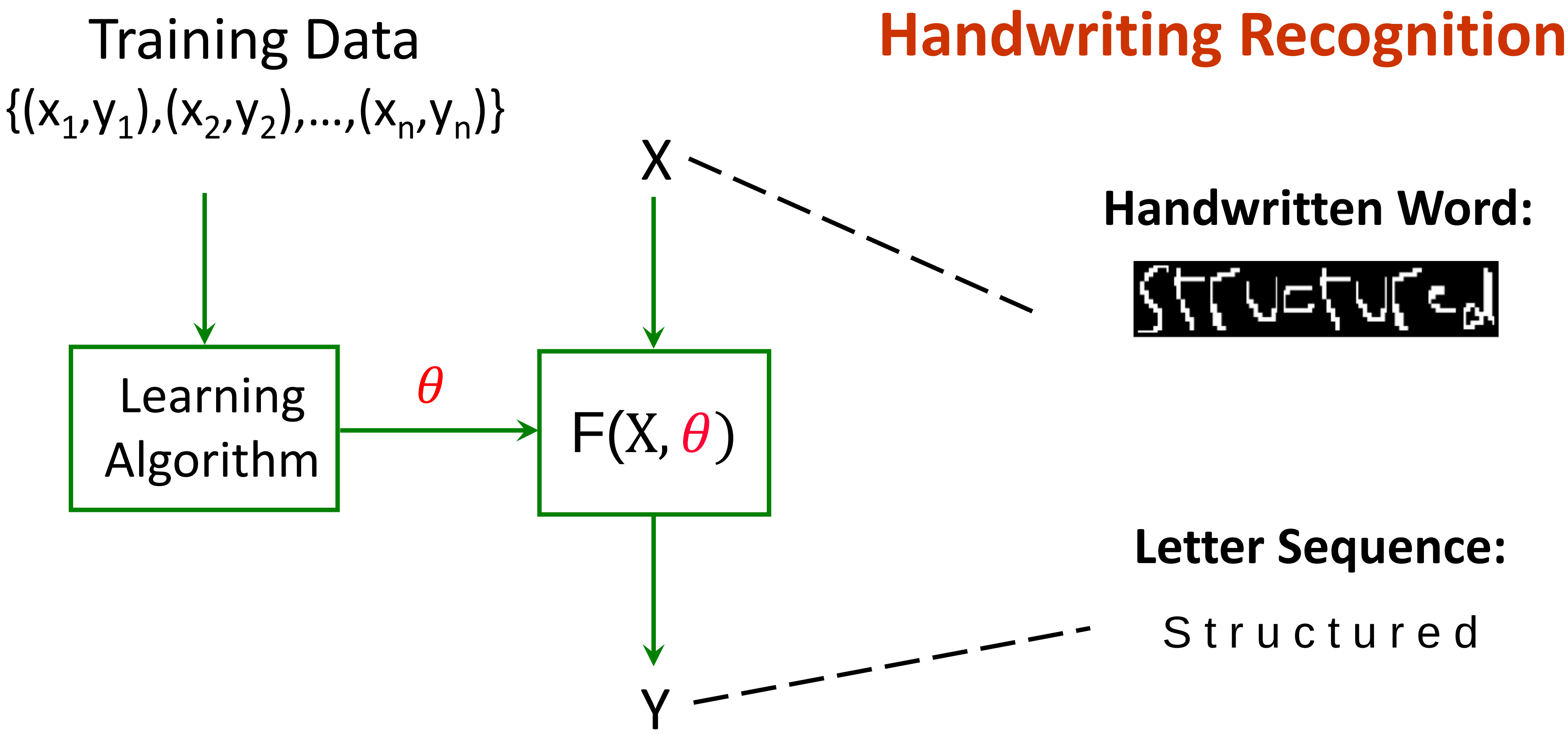
Co-reference Output:

*"**Barack Obama** nominated **Hillary Clinton** as his **secretary of state** on Monday. **He** chose **her** because **she** had foreign affair experience as a former **First Lady**."*

$\boldsymbol{Y} =$ set of all possible clusterings

Exponential !!

Learning for Structured Outputs



$\boldsymbol{Y} =$ set of all possible letter sequences

Exponential !!

Learning for Structured Outputs

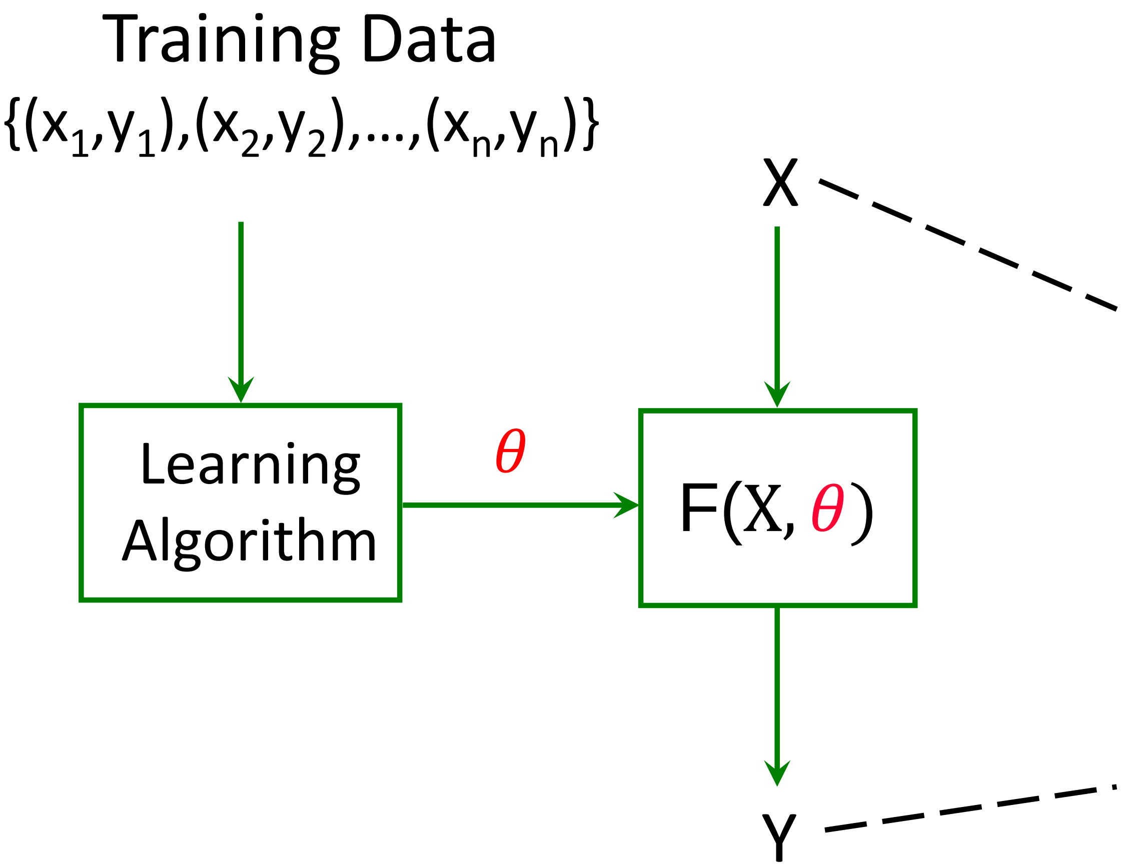
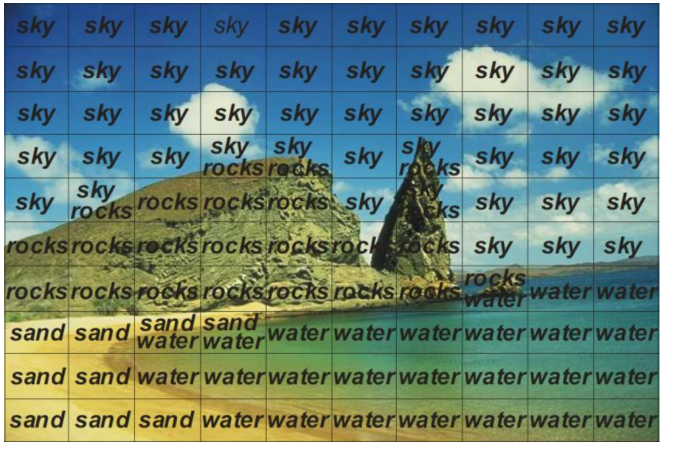


Image Labeling

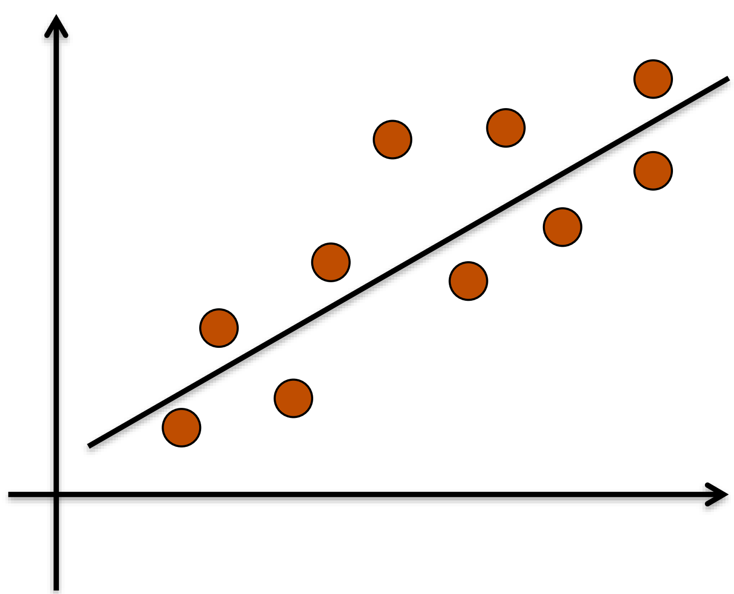


$\boldsymbol{Y}$ = set of all possible labelings

Exponential !!

Regression

- **Setting:** output y is a continuous value instead of a discrete value
  - Stock market price as a function of financial specs



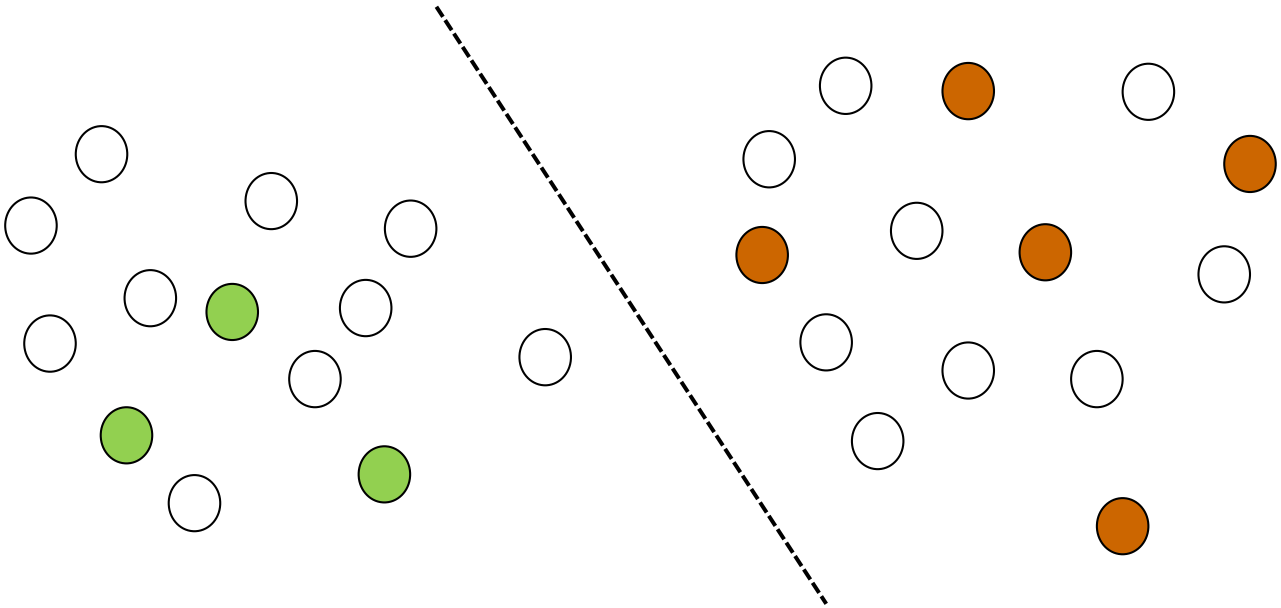
Learning Paradigms

- **Supervised Learning –** main focus of this course

- **Semi-Supervised Learning**
- **Unsupervised Learning**
- **Active Learning**
- **Reinforcement Learning**

Semi-Supervised Learning

- **Setting:** small amount of labeled data and large amount of unlabeled data



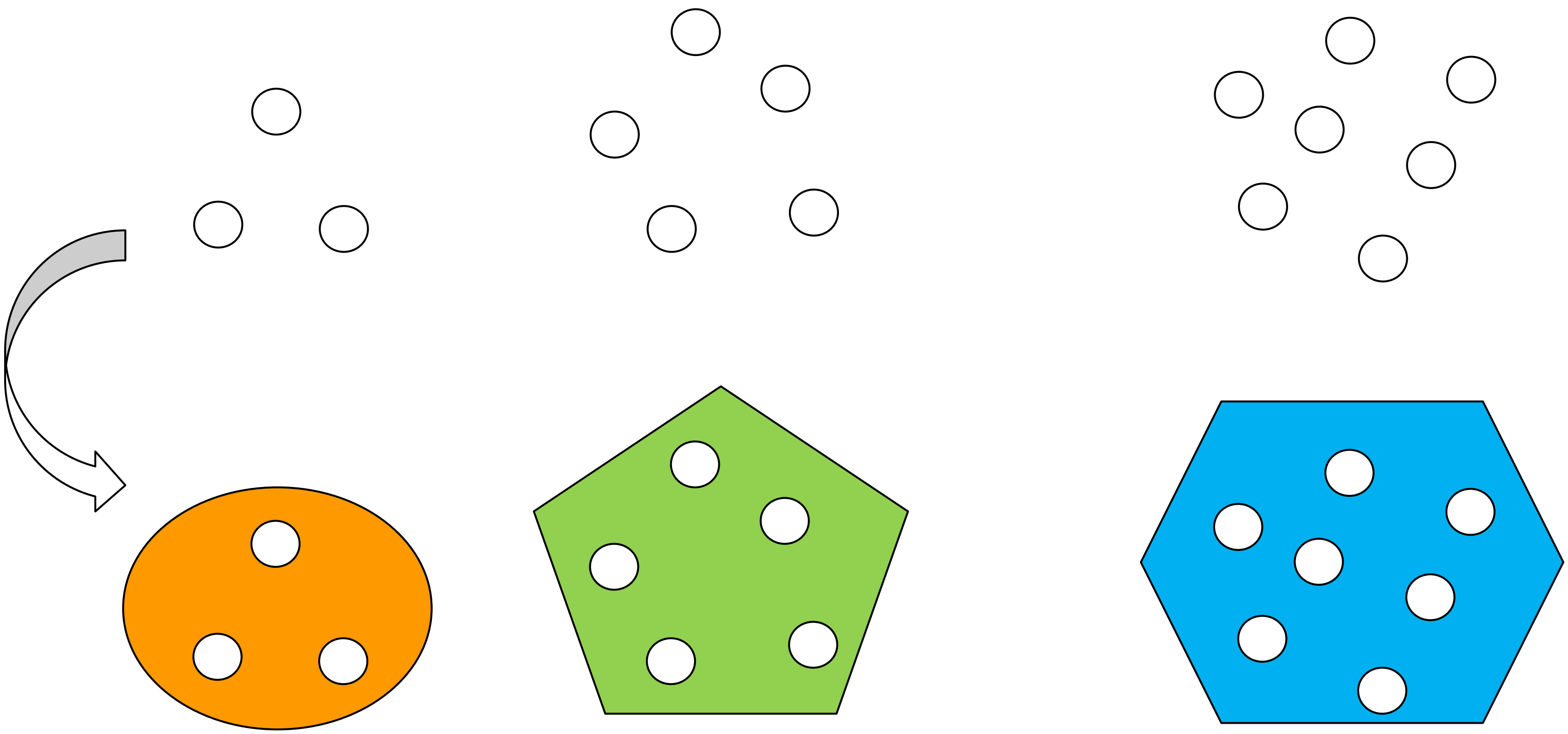
- find a classifier that separates the labeled points and separates the unlabeled points "well"

Learning Paradigms

- **Supervised Learning –** main focus of this course
- **Semi-Supervised Learning**
- **Unsupervised Learning**
- **Active Learning**
- **Reinforcement Learning**

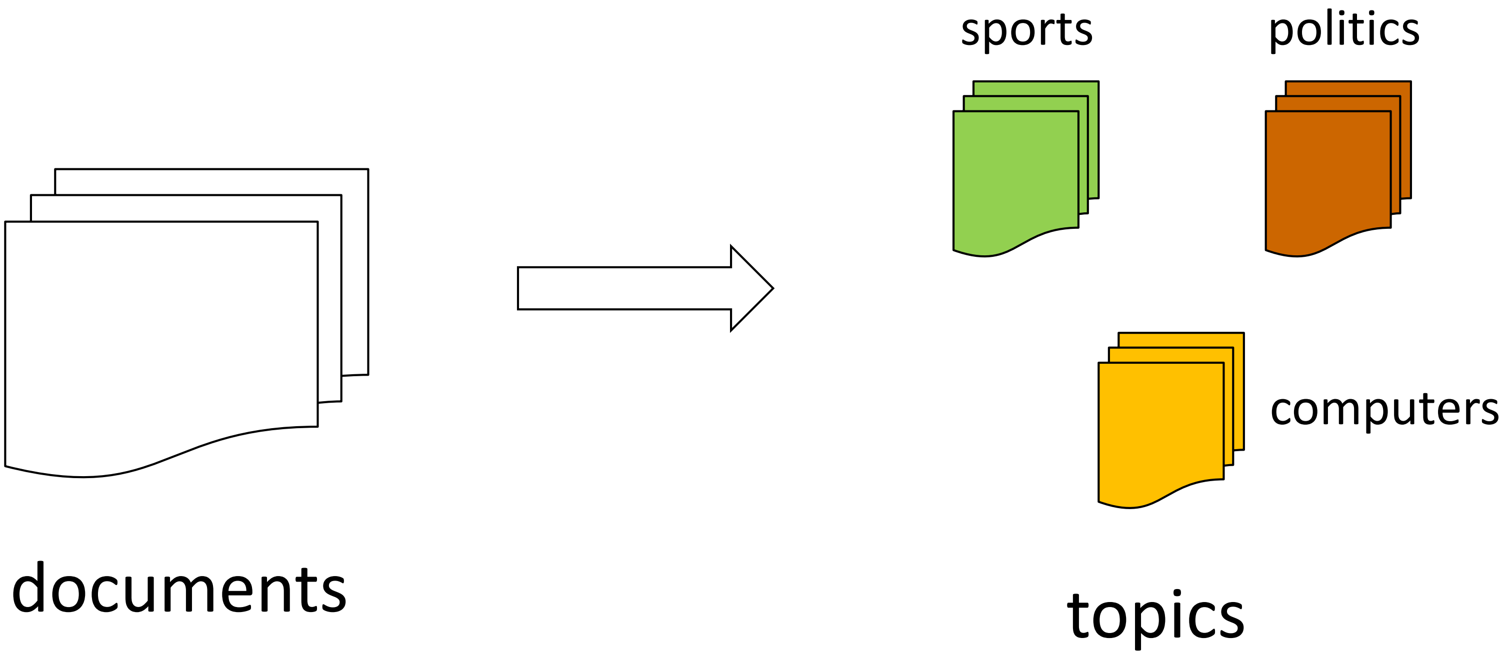
Unsupervised Learning

- **Clustering**
  - Given a collection of unlabeled examples (objects), discover self-similar groups in the data



Unsupervised Learning

- **Text Clustering**

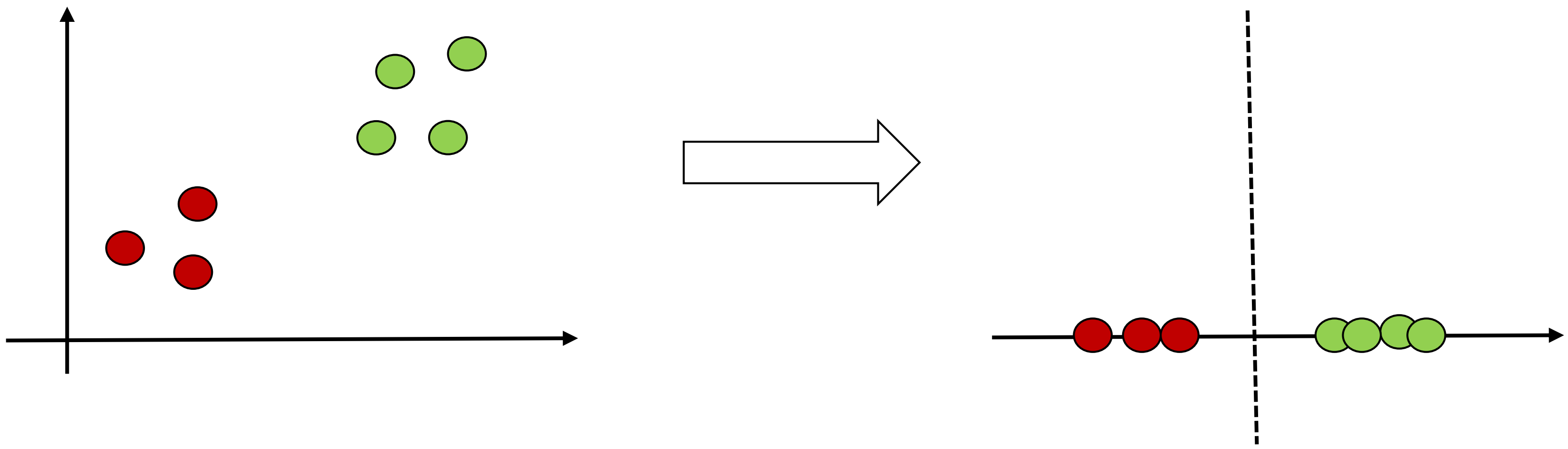


Unsupervised Learning

- **Image Segmentation**

Unsupervised Learning

- **Dimensionality Reduction (aka feature learning)**



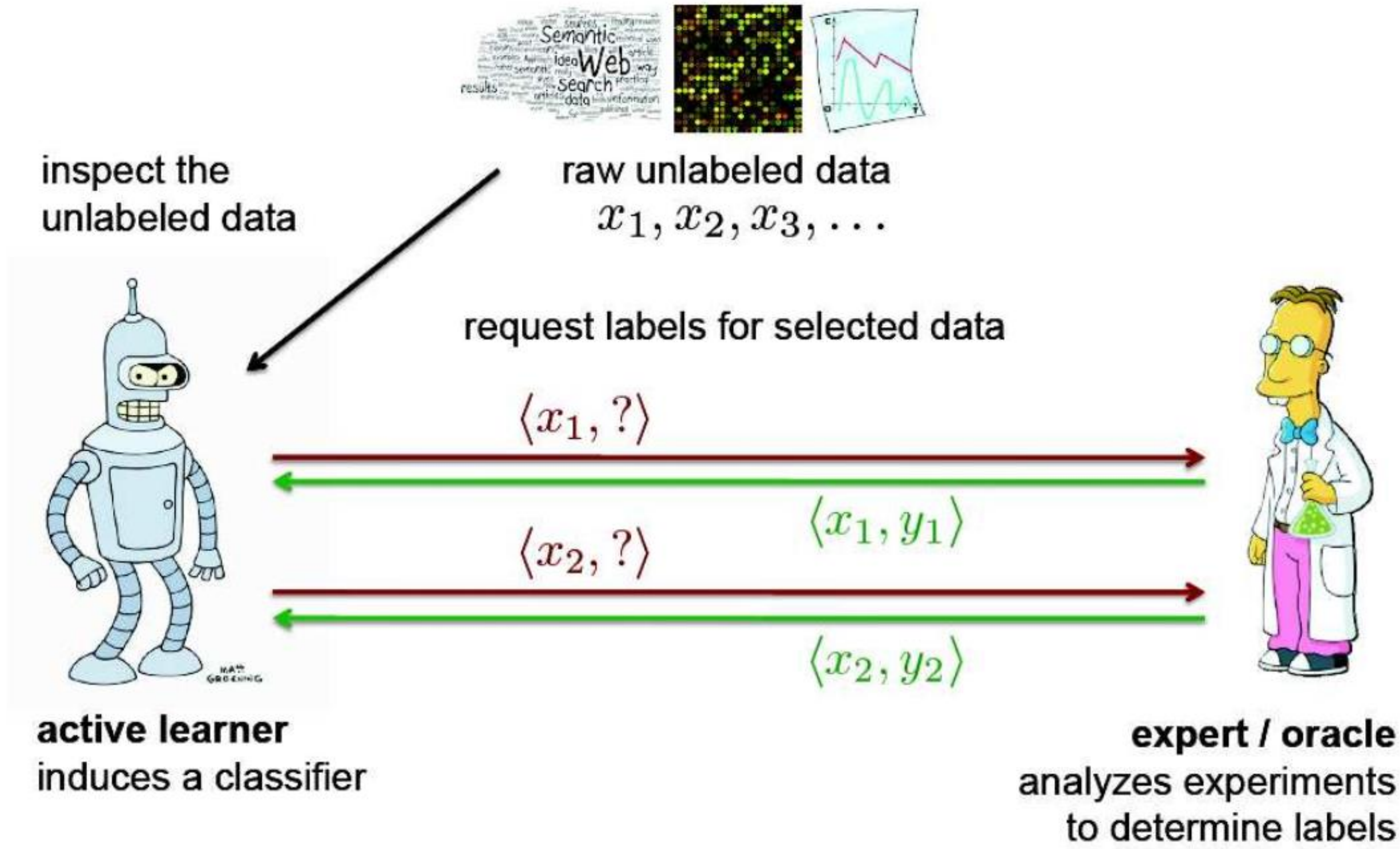
- find a mapping that preserves the “structure” of objects
- find relevant features (dimensions) for a task
- reduce dimensionality to manage the complexity of high-dimensional data

Learning Paradigms

- **Supervised Learning –** main focus of this course

- **Semi-Supervised Learning**
- **Unsupervised Learning**
- **Active Learning**
- **Reinforcement Learning**

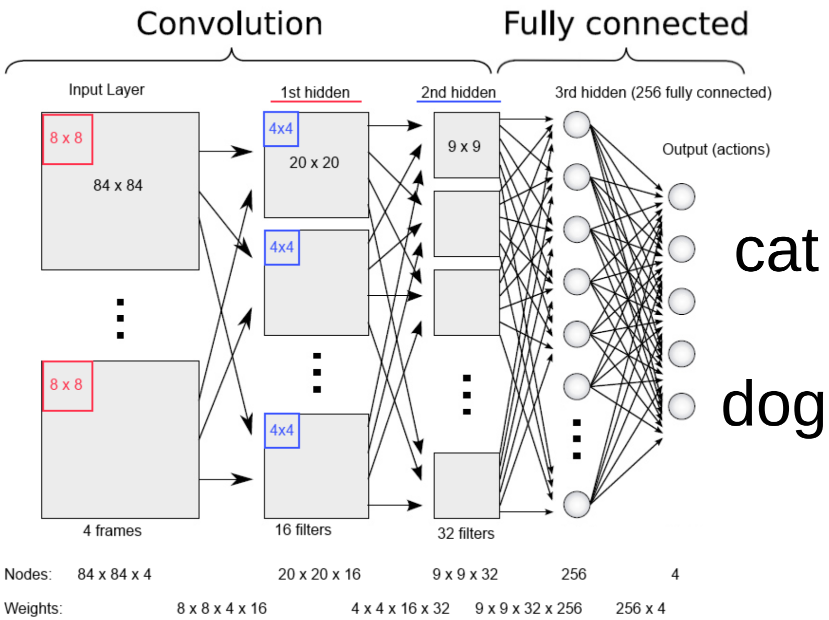
Active Learning



Learning Paradigms

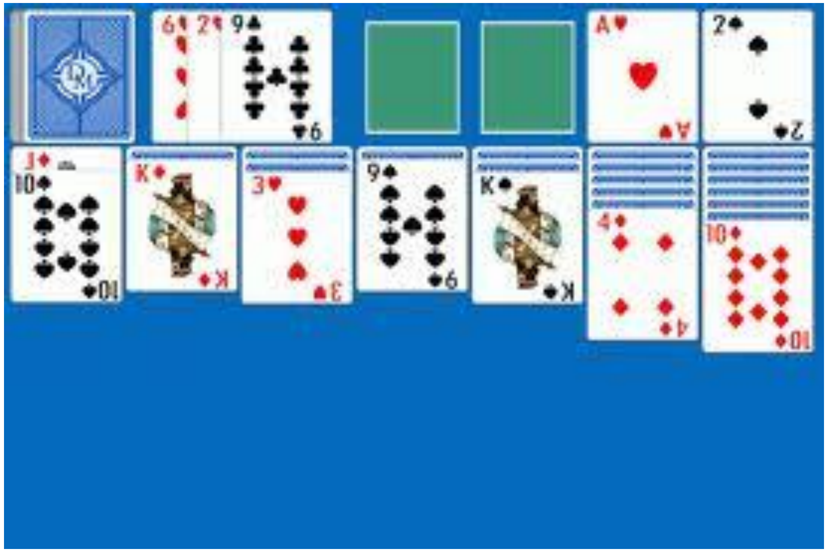
- **Supervised Learning –** main focus of this course
- **Semi-Supervised Learning**
- **Unsupervised Learning**
- **Active Learning**
- **Reinforcement Learning**

One Shot Decision Making



Each decision/classification can be made without considering future decisions making.

Sequential Decision Making

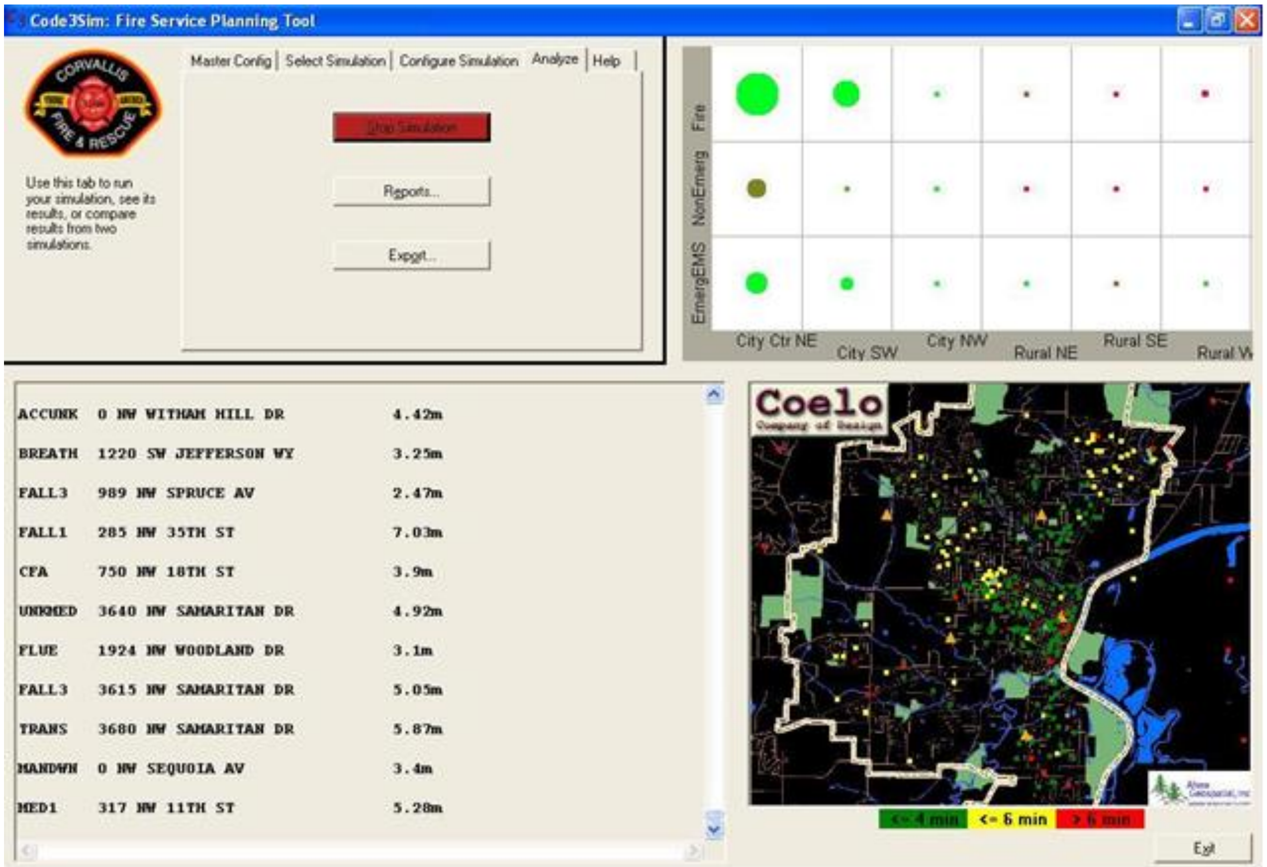


Klondike Solitaire

Real-Time Strategy Games

Sequential Decision Making

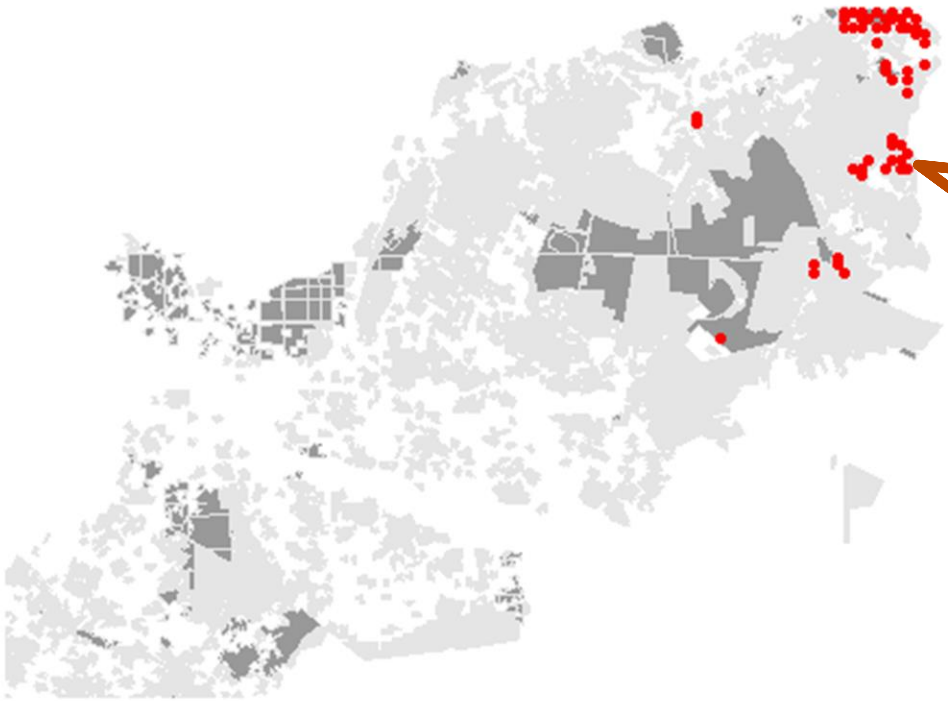
Optimizing Fire & Rescue Response Policies



Sequential Decision Making

Conservation Planning:

Recovery of Red-cockaded Woodpecker

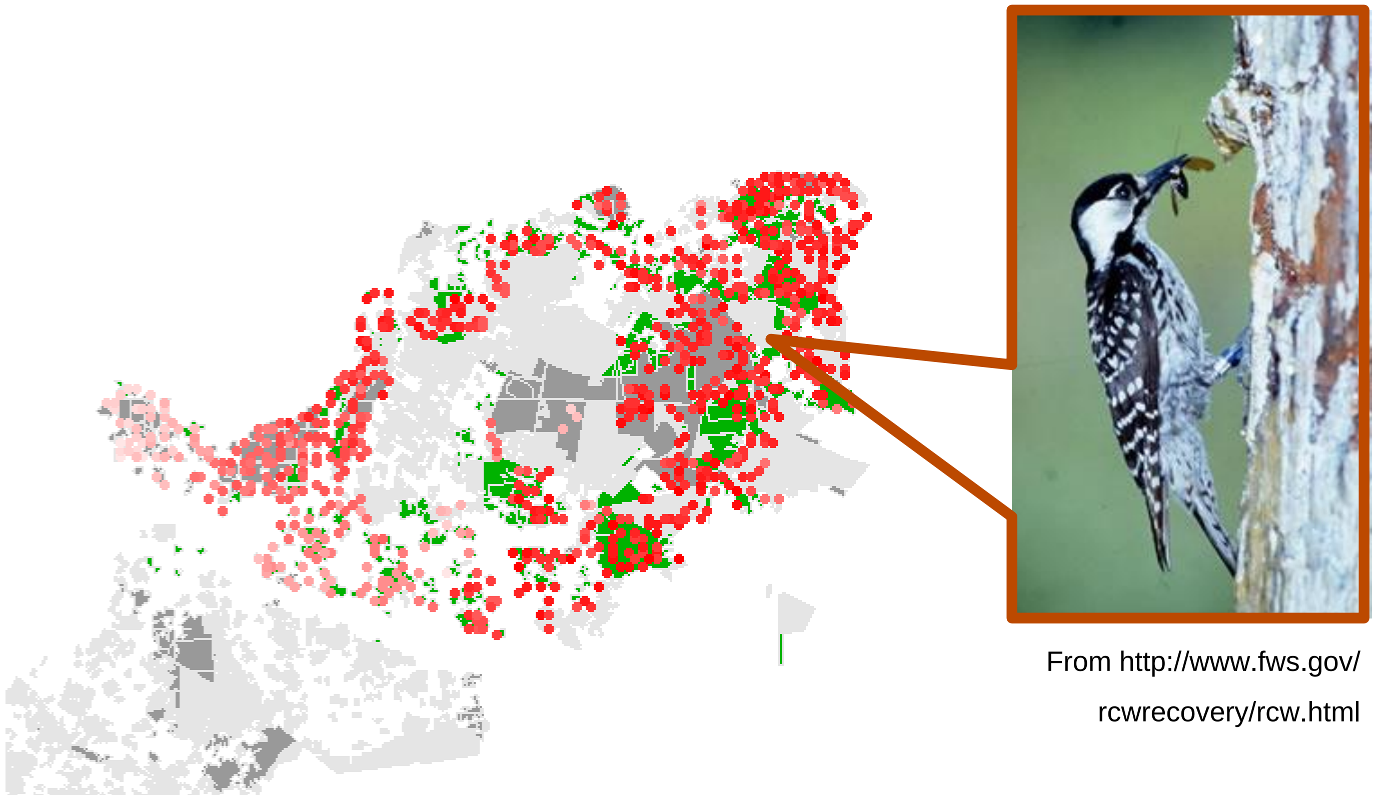


From http://www.fws.gov/
rcwrecovery/rcw.html

Sequential Decision Making

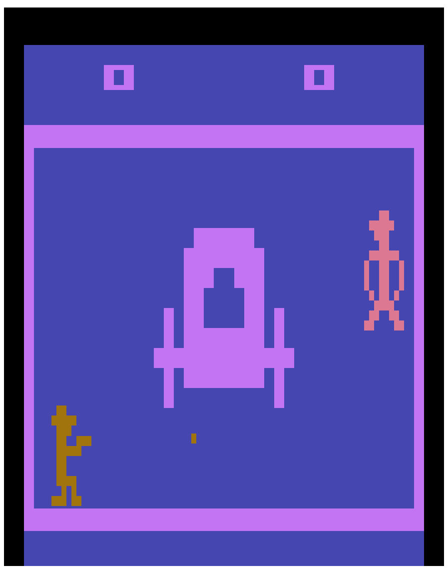
Conservation Planning:

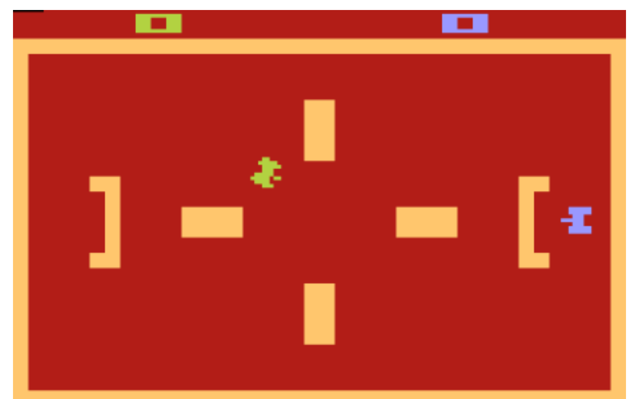
Recovery of Red-cockaded Woodpecker



From http://www.fws.gov/rcwrecovery/rcw.html

Sequential Decision Making

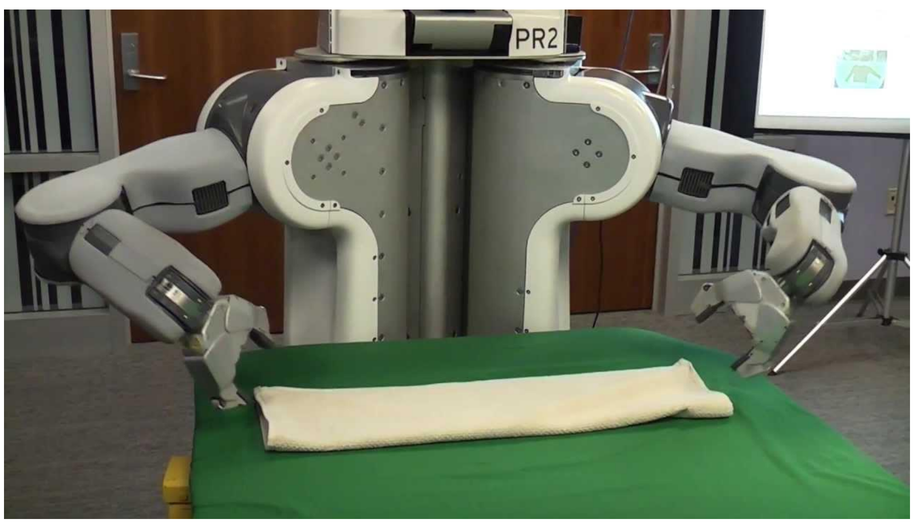




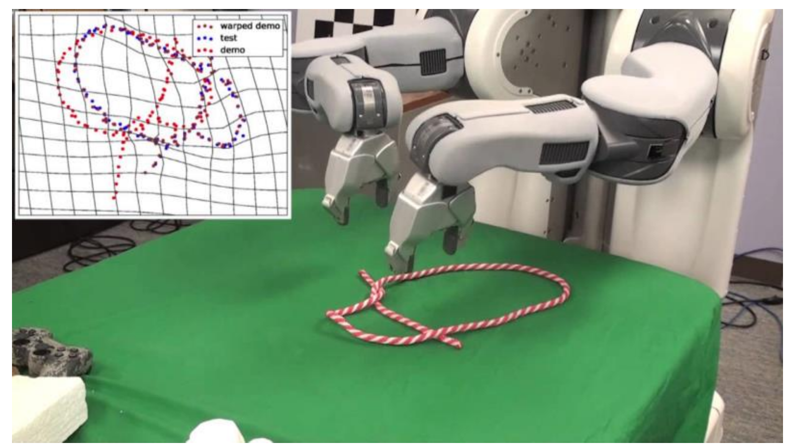
Sequential Decision Making

Helicopter Control

Legged Robot Control



Laundry



Knot Tying

AlphaGo

- Deep Learning + Monte Carlo Tree Search
- Learn from 30 million expert moves and self play
- Highly parallel search implementation
- 48 CPUs, 8 GPUs (scaling to 1,202 CPUs, 176 GPUs)

AlphaGo vs. Lee Sedol (9-dan pro w/ 18 world titles)

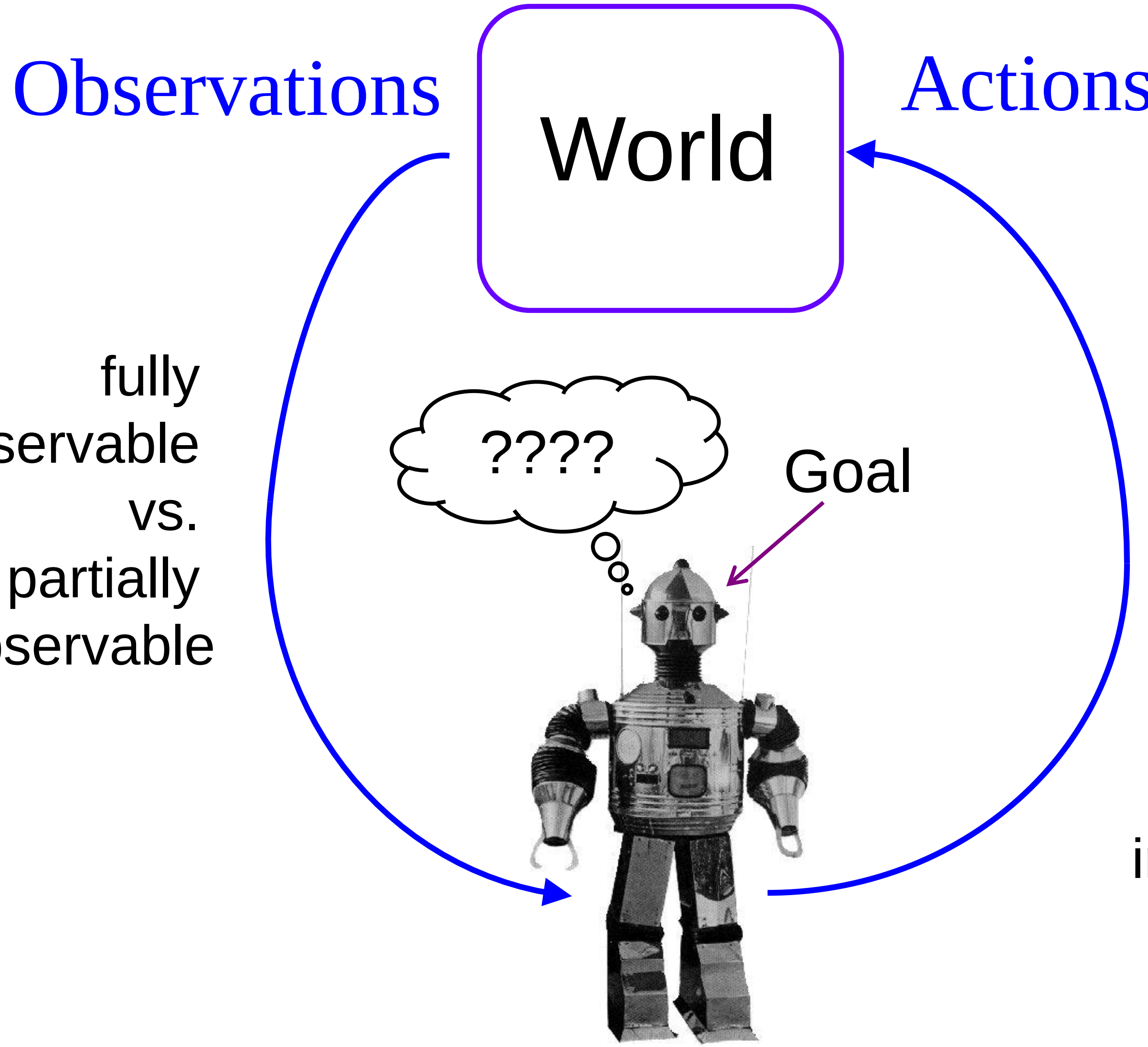
AlphaGo won 4 games to 1

https://deepmind.com/alpha-go.html

Common Elements

- We have a controllable system that can change state over time (in some predictable way)
  - The state describes essential information about system (the visible card information in Solitaire)
- We have an objective that specifies which states, or state sequences, are more/less preferred
- Can (partially) control the system state transitions by taking actions
- **Problem:** At each moment must select an action to optimize the overall objective
  - Produce most preferred state sequences

Reinforcement Learning



fully
observable
vs.
partially
observable

sole source
of change
vs.
other sources

deterministic
vs.
stochastic

instantaneous
vs.
durative

Stochastic/Probabilistic Planning: Markov Decision Process (MDP) Model

Observations

World

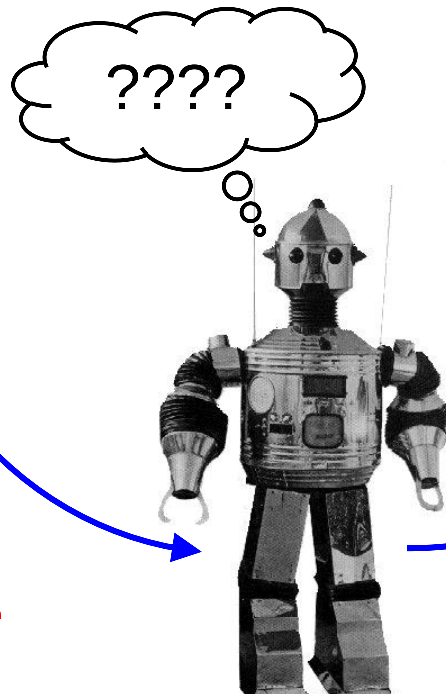
Actions

sole source
of change

fully
observable

stochastic

instantaneous

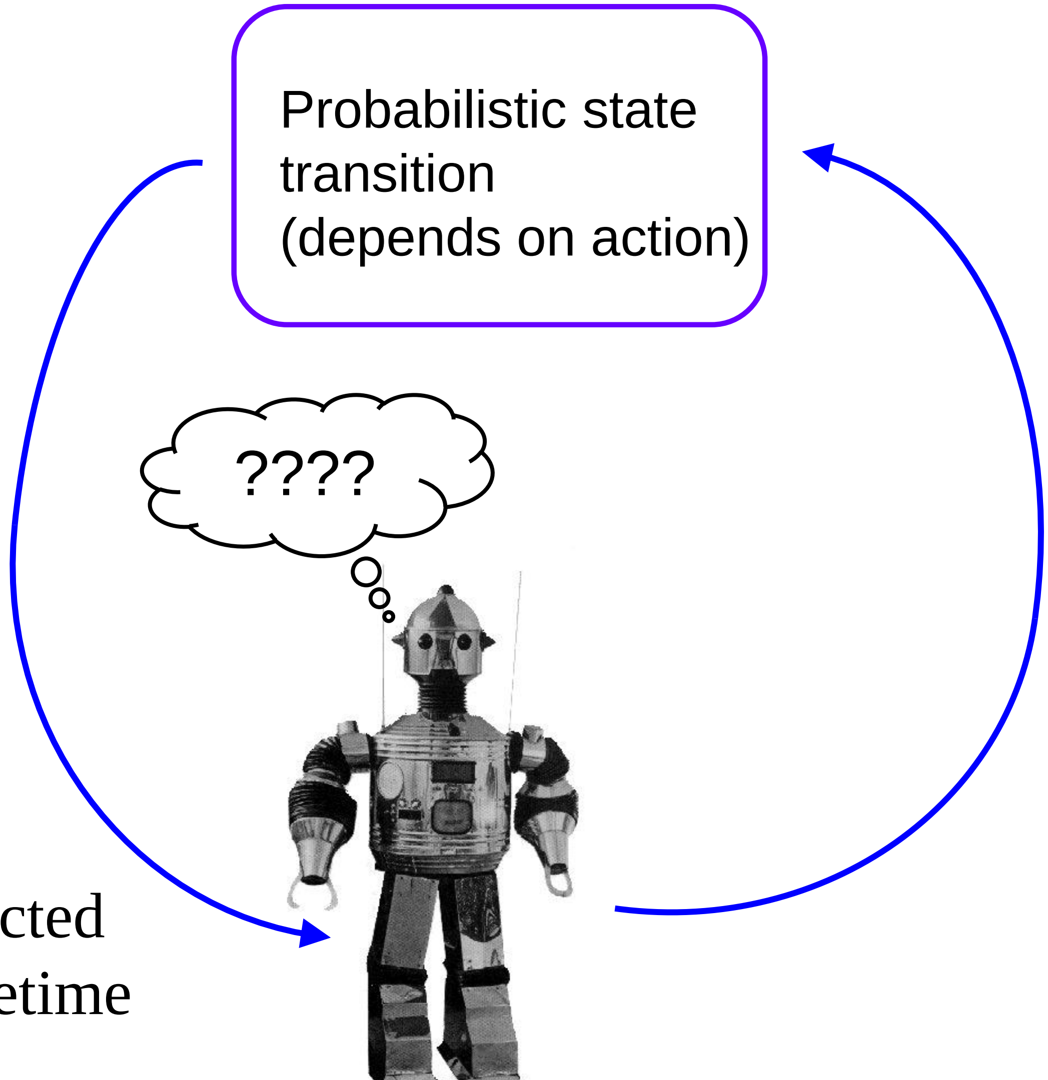


Goal

maximize expected
reward over lifetime

We will primarily
focus on MDPs

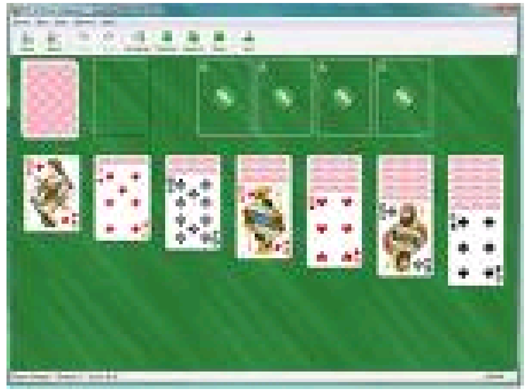
Stochastic/Probabilistic Planning: Markov Decision Process (MDP) Model



Goal

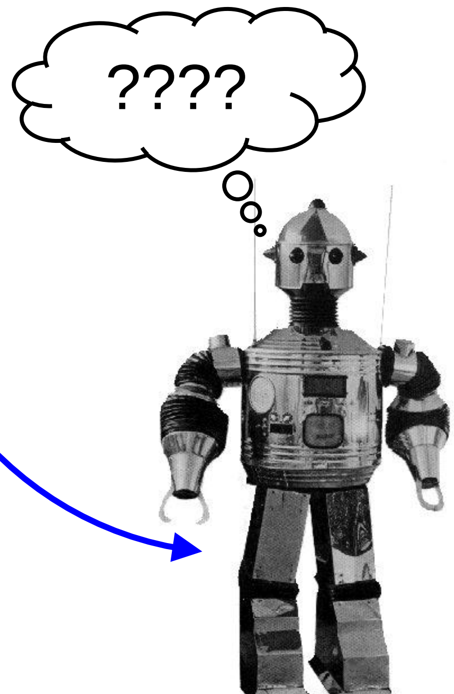
maximize expected reward over lifetime

Example MDP



State describes all visible info about cards

Action are the different legal card movements



Goal

win the game or play max # of cards