Lecture #1: Introduction

Janardhan Rao (Jana) Doppa

School of EECS, Washington State University

Course Logistics

- CptS 570: Machine Learning
  - **Class Timings** – Tue and Thu 10:35 to 11:50 (Sloan 163)
  - **Instructor** – Jana Doppa (EME 133)
  - **Office Hours** – Mon, Fri 4-5?

Course Logistics

- CptS 570: Machine Learning
  - **Course website** – http://eecs.wsu.edu/~jana/classes/cs570-machine-learning-fall2016.html
  - **Course announcements and discussions** – Piazza
  - **Lecture Notes** – Slides and notes will be posted on Piazza

Grading Policy

- **6 Home works (35%)**
  - Best five will be counted towards your grade
- **2 Mid-term exams (40%)**
  - will decide on the timing via polling
- **1 Project (20%)**
  - Can be done in small groups (two or three students)
- **Class Participation (5%)**
  - Piazza and in-class

Late Policy

- All assignments, project proposal/report are due at the **start** of the class
- **Late Policy**
  - 0-24 hours late -- 80% of the final score
  - 24-48 hours late -- 50% of the final score
  - Beyond 48 hours -- 0%
- If you are late, please slip the assignment through my office door

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

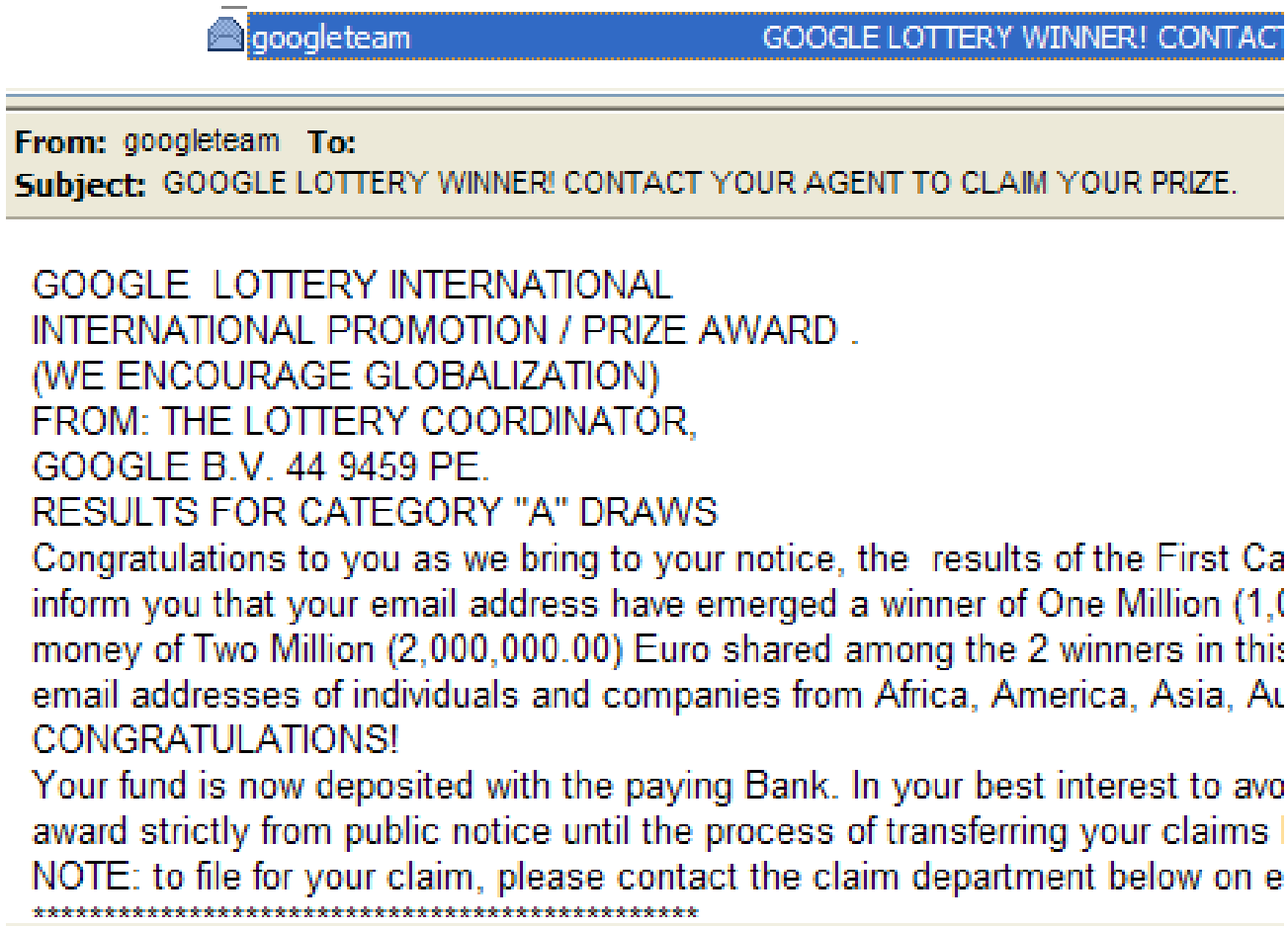
Machine Learning is Everywhere

- "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! CONTAC

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

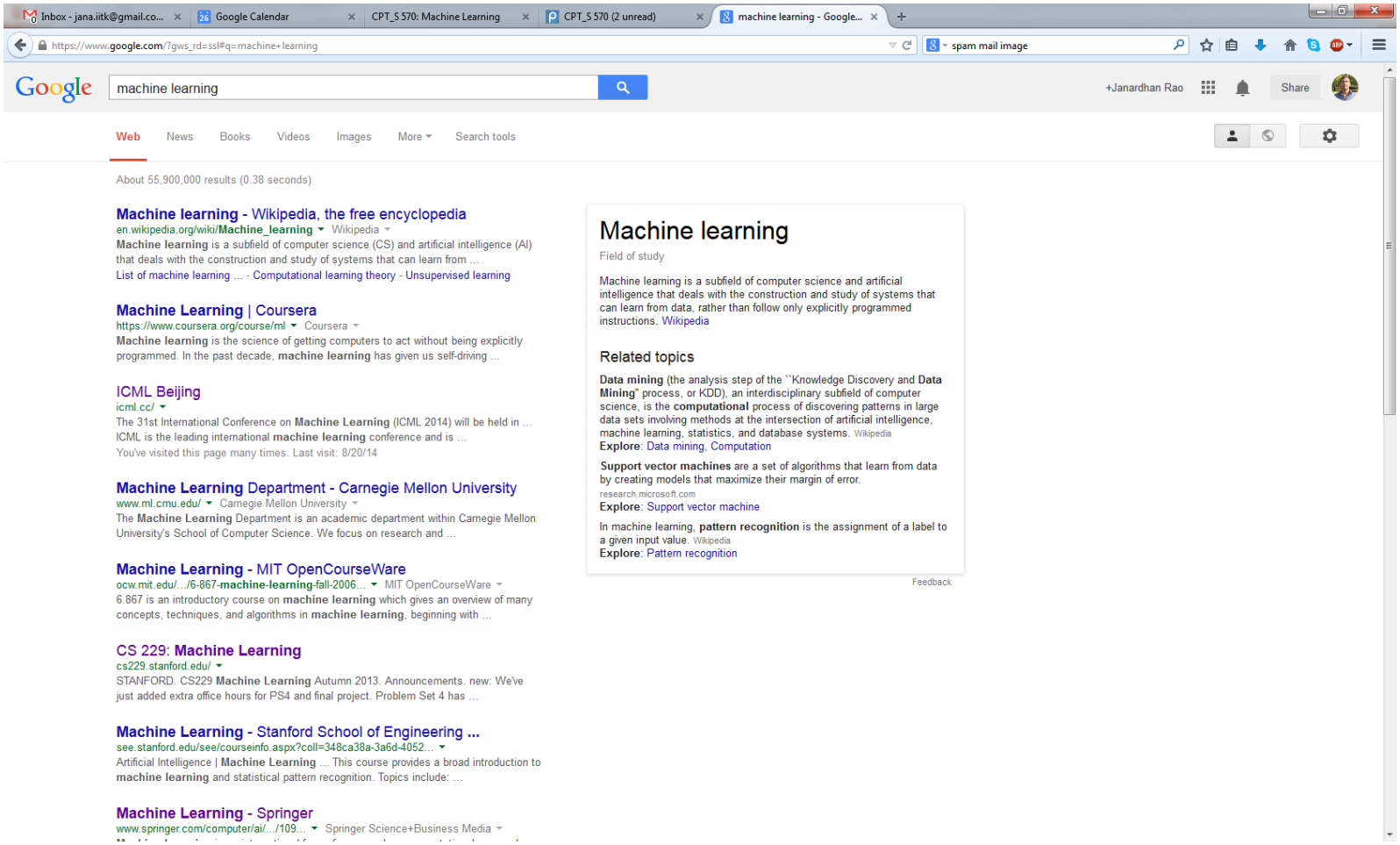
GOOGLE LOTTERY INTERNATIONAL
INTERNATIONAL PROMOTION / PRIZE AWARD .
(WE ENCOURAGE GLOBALIZATION)
FROM: THE LOTTERY COORDINATOR,
GOOGLE B.V. 44 9459 PE.
RESULTS FOR CATEGORY "A" DRAWS
Congratulations to you as we bring to your notice, the results of the First Ca
inform you that your email address have 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)

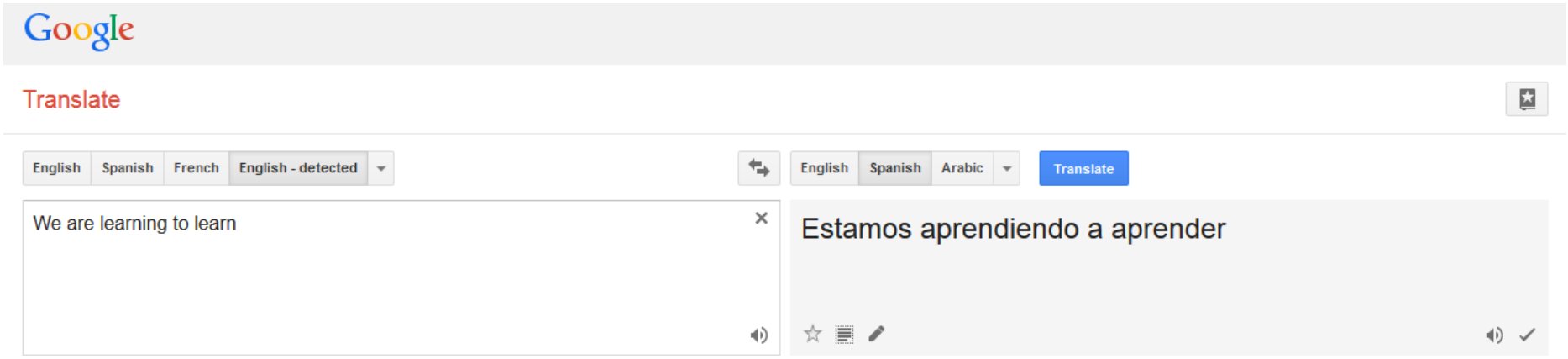
Machine Learning is Everywhere

- Search engines



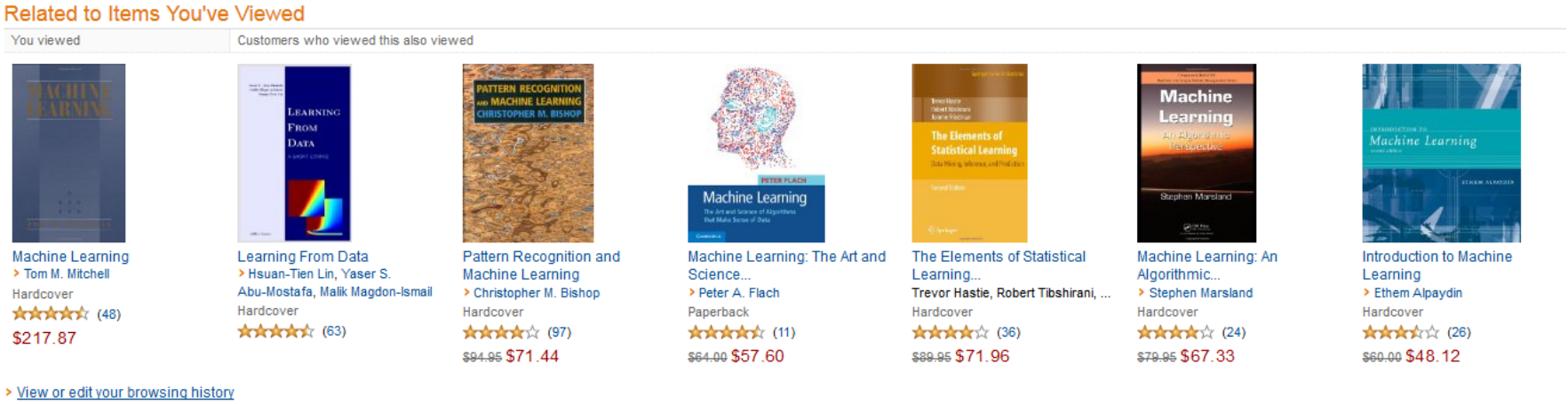
Machine Learning is Everywhere

- Automatic Translation



Machine Learning is Everywhere

- Recommendation Engines



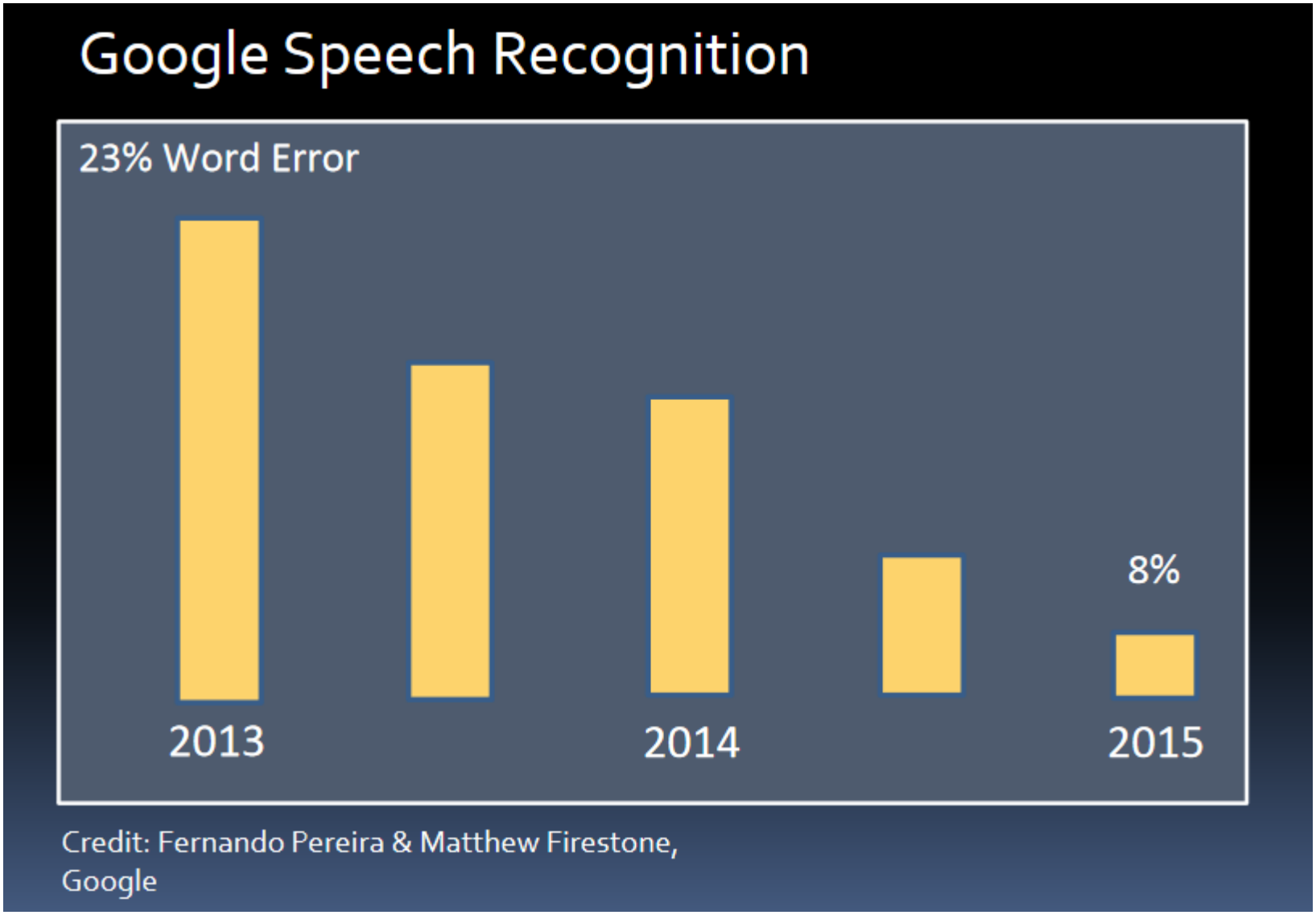
Machine Learning is Everywhere

- Self-driving cars

Google's Self Driving Car for Blind People

by EDITORS on Apr 6, 2012 • 4:07 pm

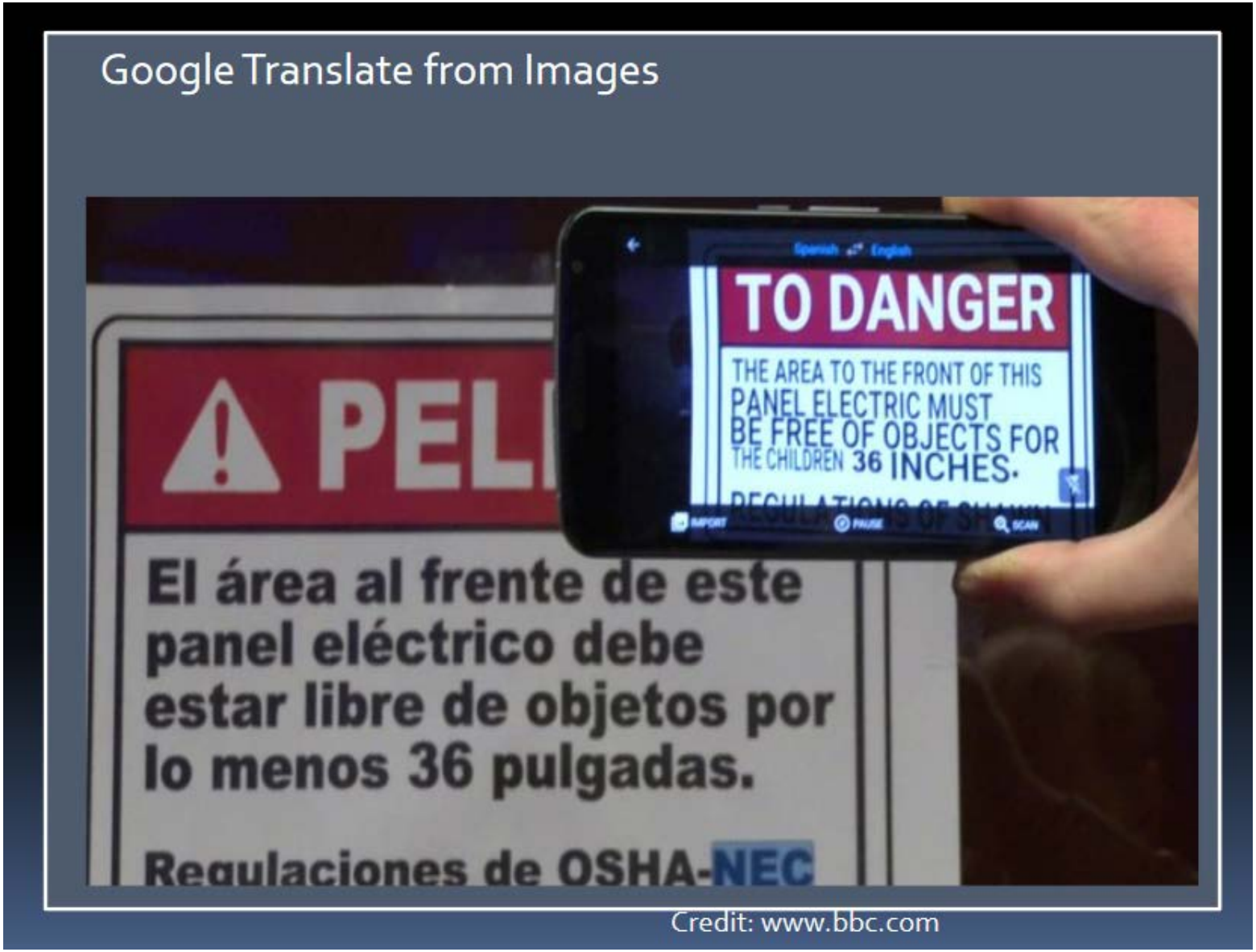
ML Successes: Perception



Credit: Tom Dietterich

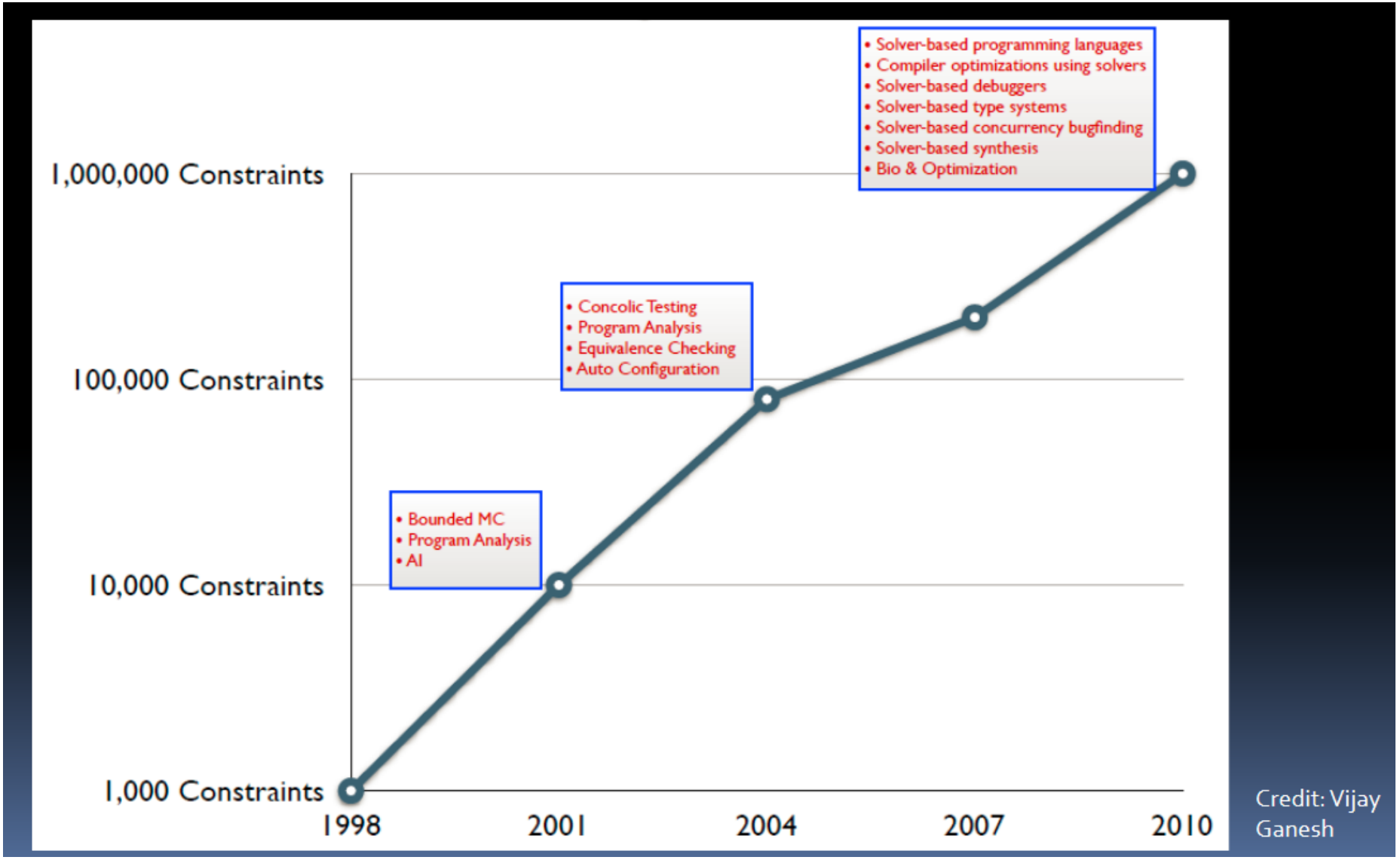
ML Successes: Image Captioning

ML Successes: Perception + Translation

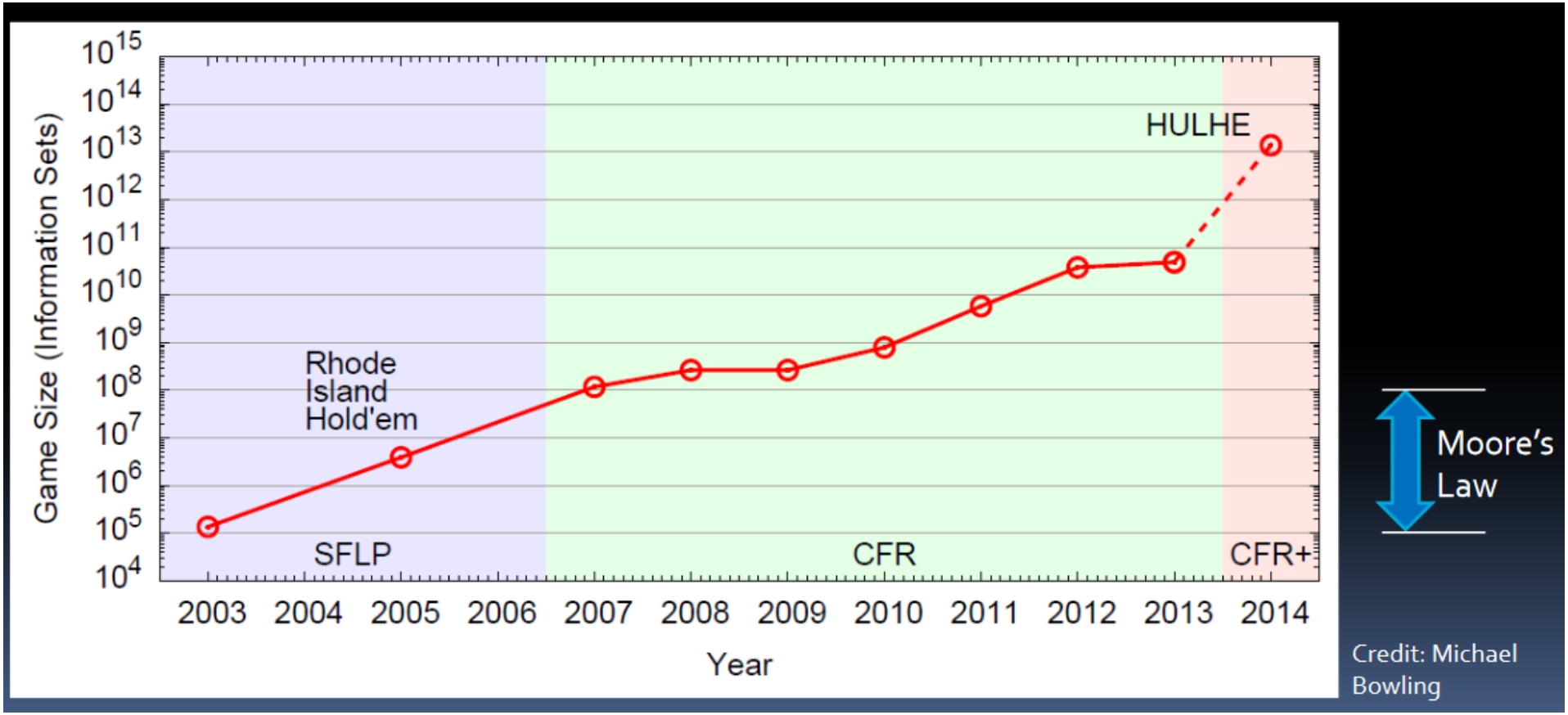


ML Successes: Skype Translator

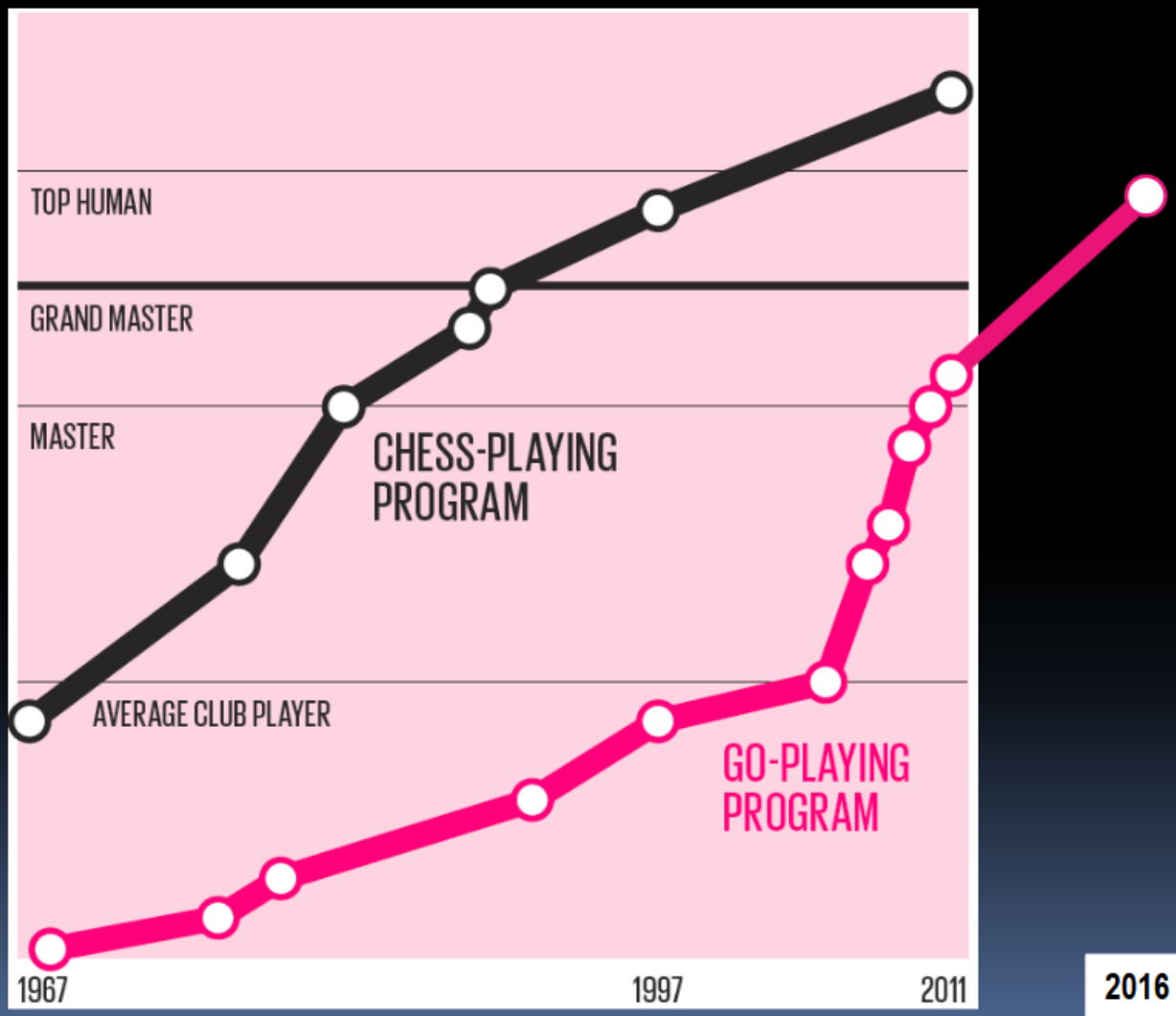
ML Successes: Reasoning (SAT)



ML Successes: Poker



ML Successes: Chess and Go

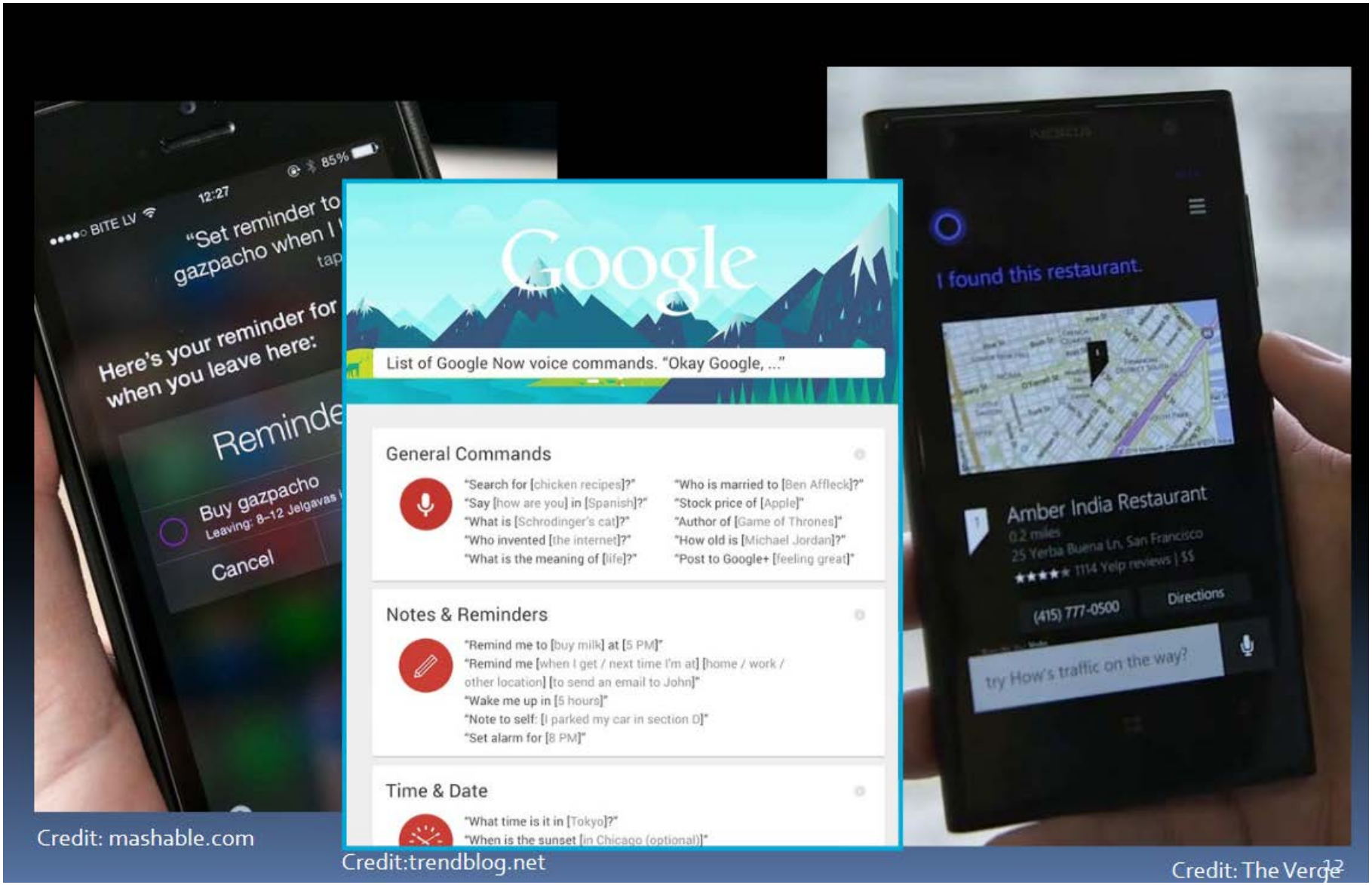


Silver, et al. (2016) *Nature*
Deep Learning +
Monte Carlo Tree Search

2016

Credit: Martin Mueller

ML Successes: Personal Assistants



Credit: mashable.com

Credit:trendblog.net

Credit: The Verge

High-Stakes Applications: Self-Driving Cars

Credit: The Verge

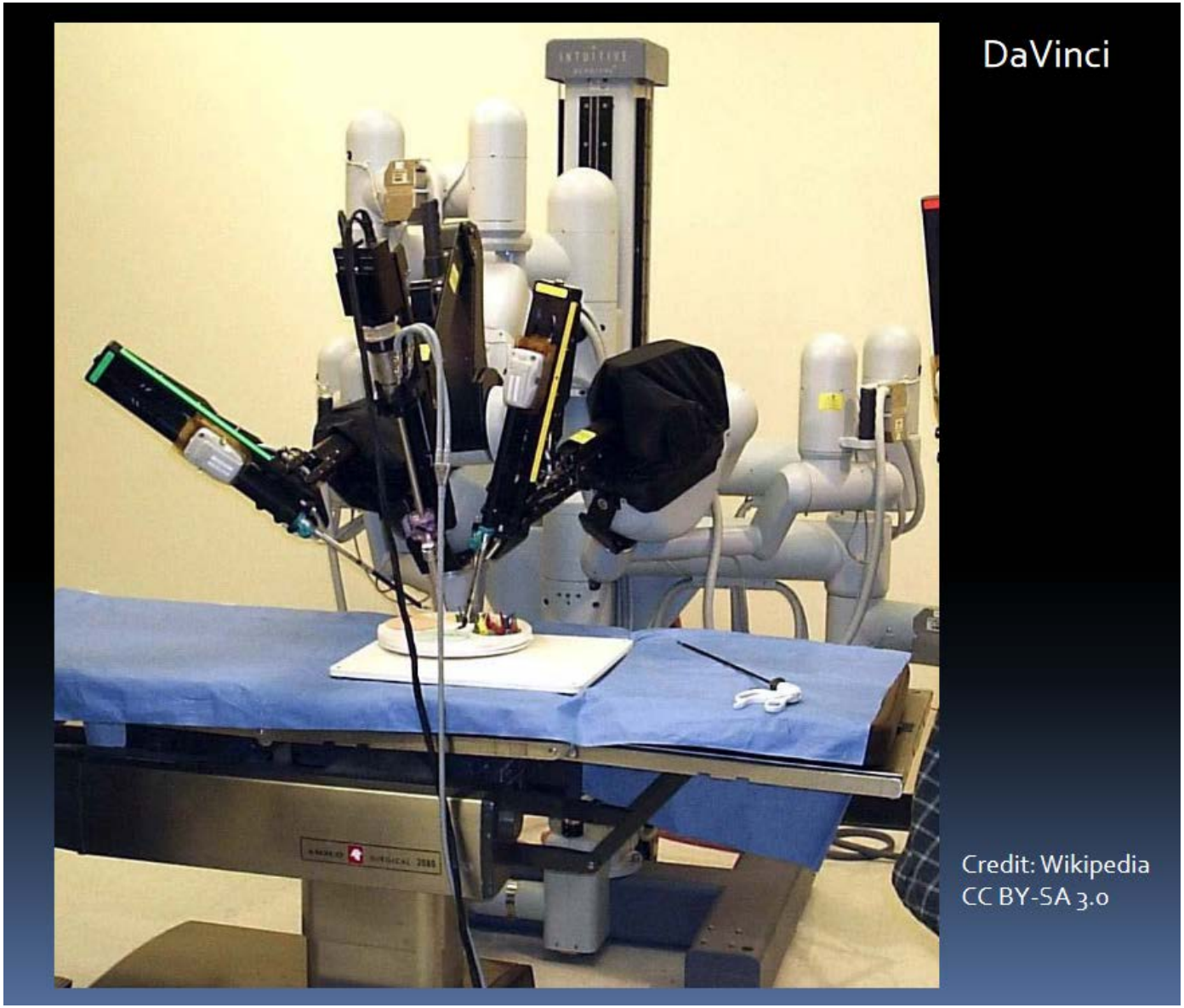
Credit: delphi.com

Tesla AutoSteer

Credit: Tesla Motors

14

High-Stakes Applications: Automated Surgical Assistants



High-Stakes Applications: AI Hedge Funds

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

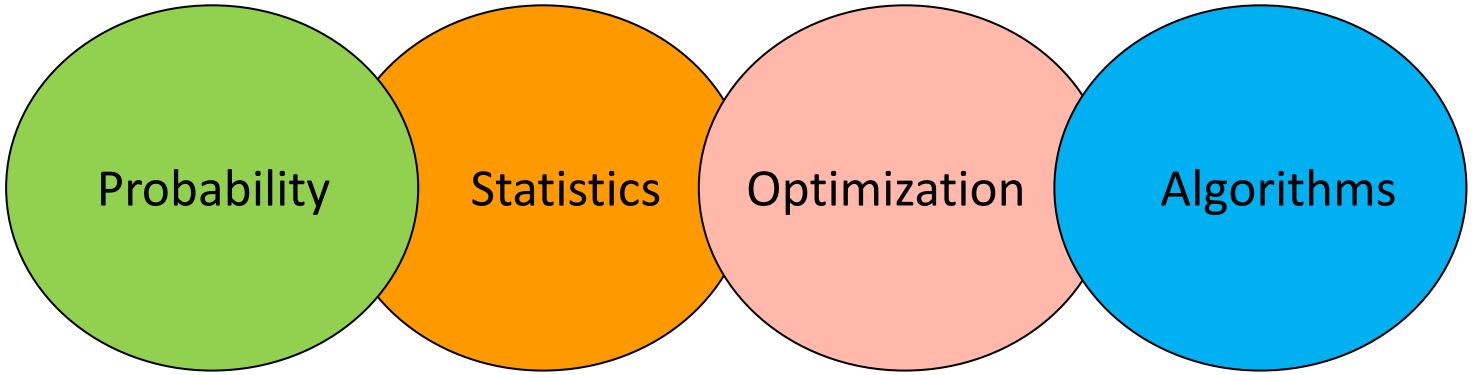
Credit: Duch.seb - Own work, CC BY-SA 3.0

Machine Learning is Everywhere

- **Analyzing and learning from social media data**
  - Facebook
  - LinkedIn
  - Twitter
  - …
- **Other Applications**
  - Health-care
  - Education
  - Assistive technologies

What is Machine Learning?

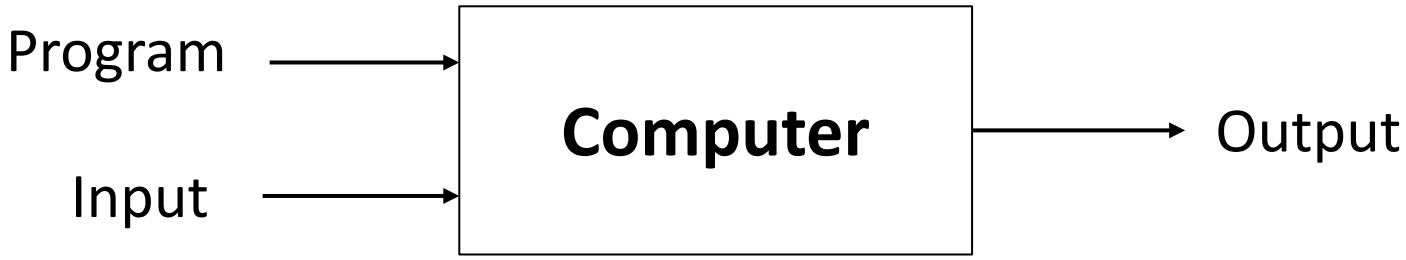
- **Machine learning is the branch of engineering that develops technology for automated inference**
  - It combines



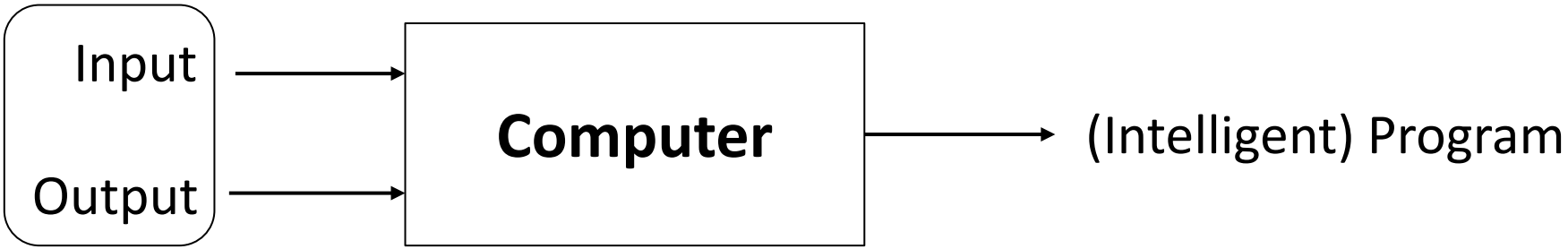
What is Machine Learning?

- **Machine learning = Automating Automation**

Traditional Programming



Machine Learning



Training data

Learning Paradigms

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

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

Supervised Learning

Learning a Classifier

( , male)

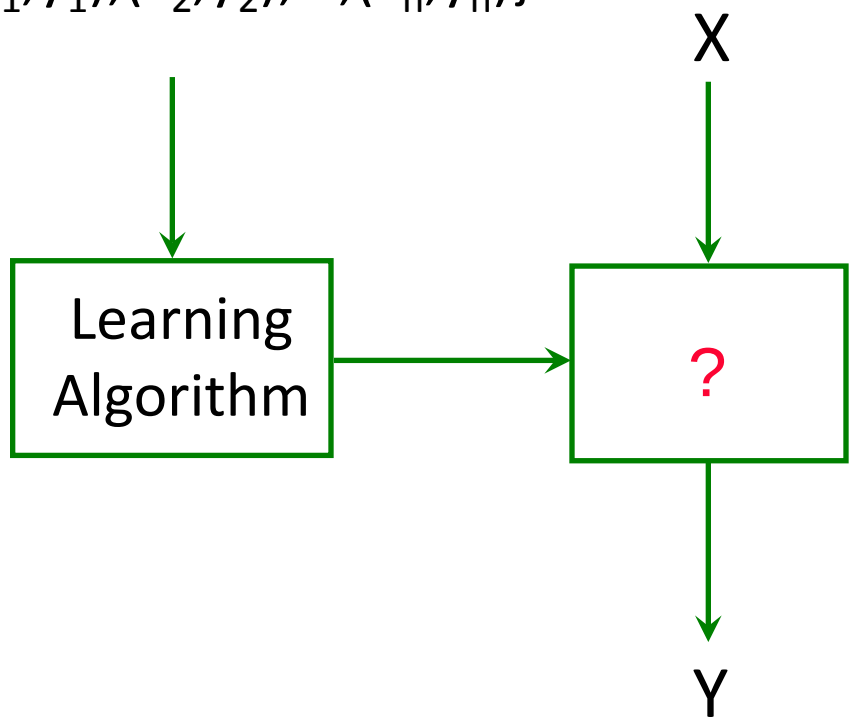
Example problem:

X - image of a face

Y ∈ {male, female}

Training Data

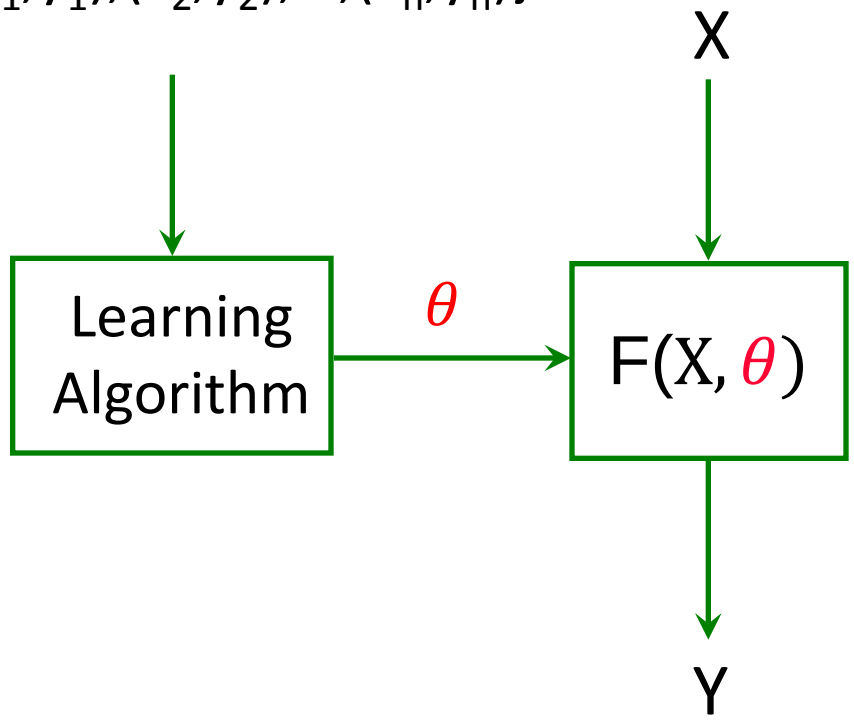
$\{(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)\}$



Learning a Classifier

Training Data

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

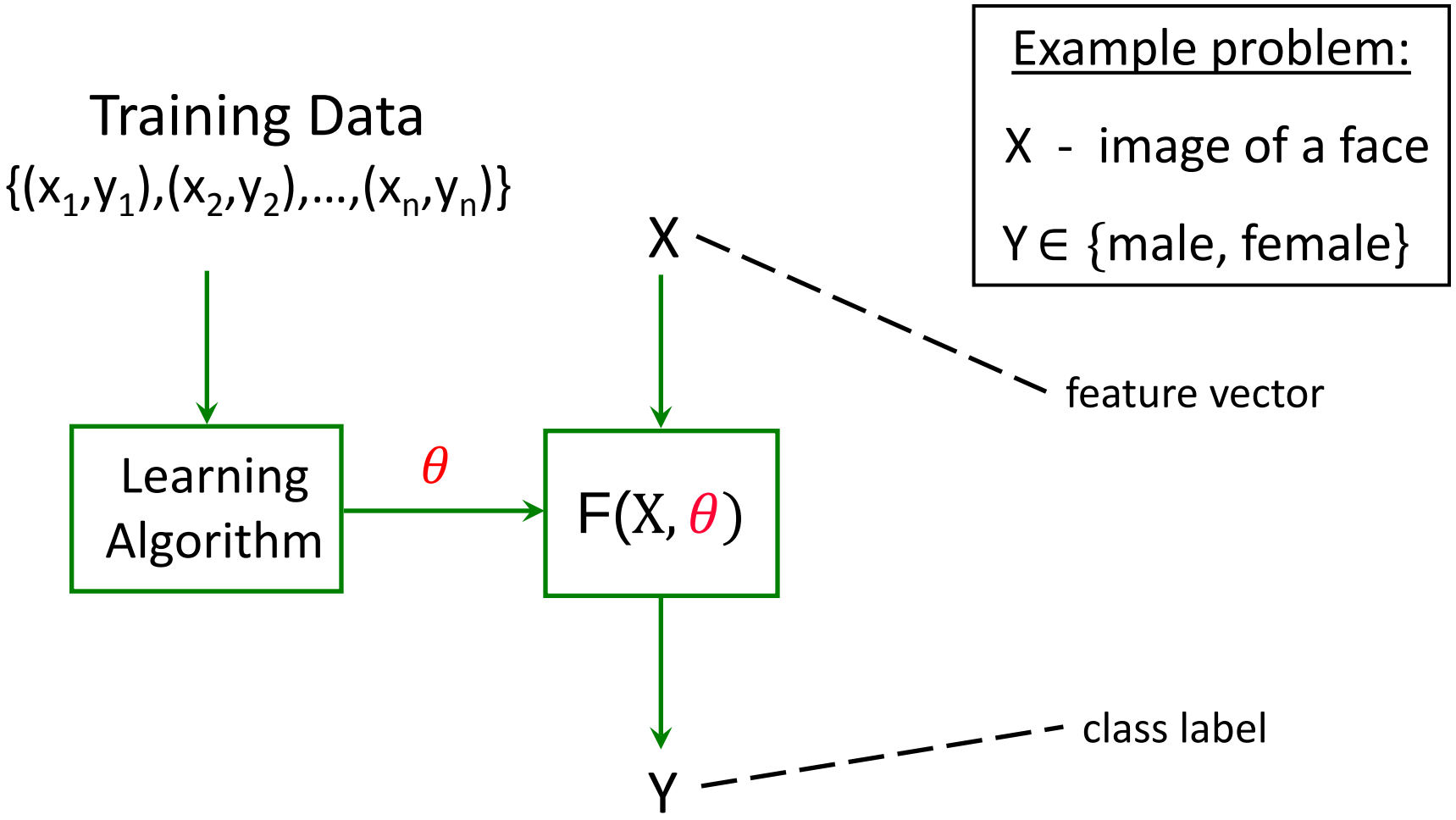


Example problem:

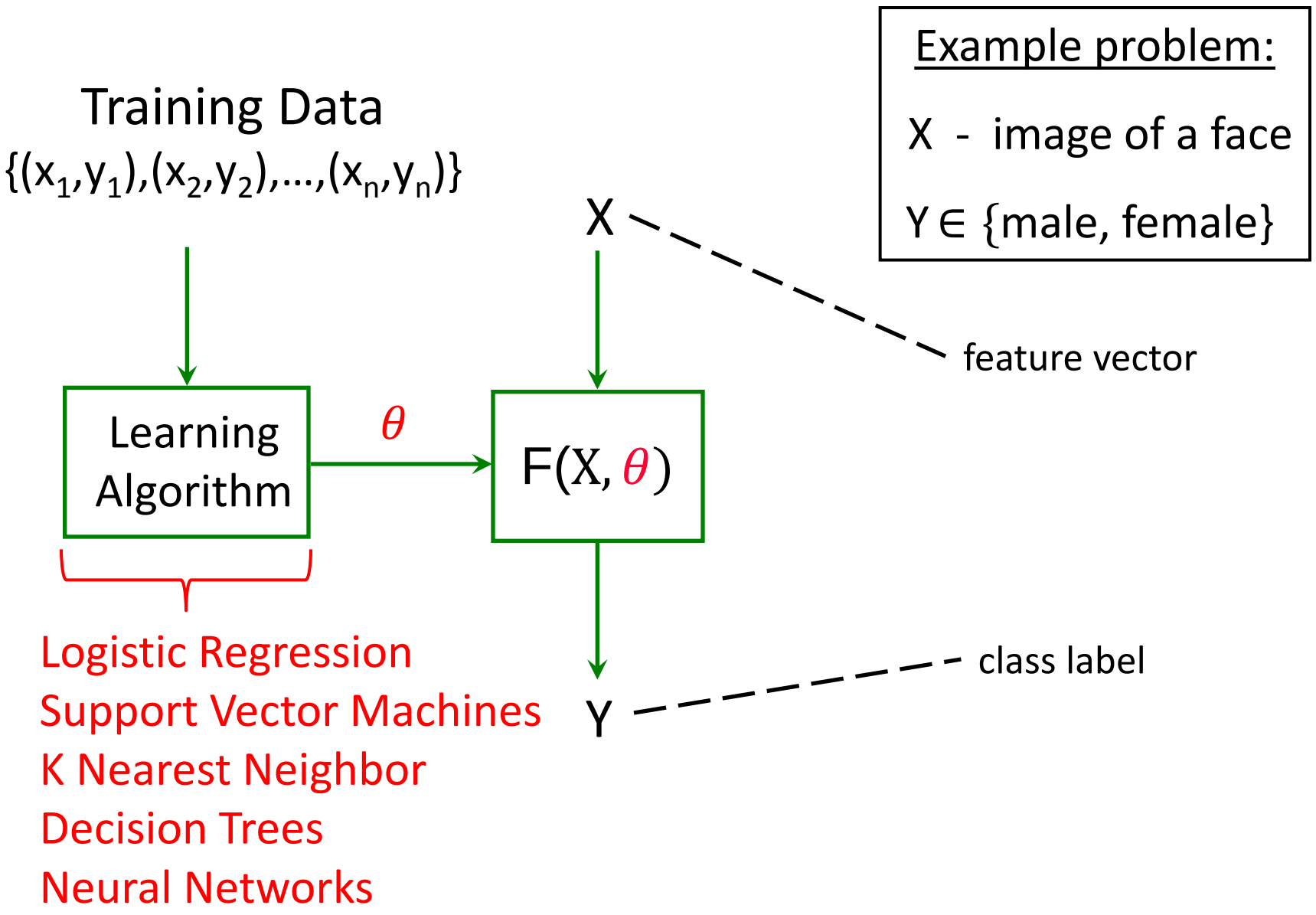
X - image of a face

Y ∈ {male, female}

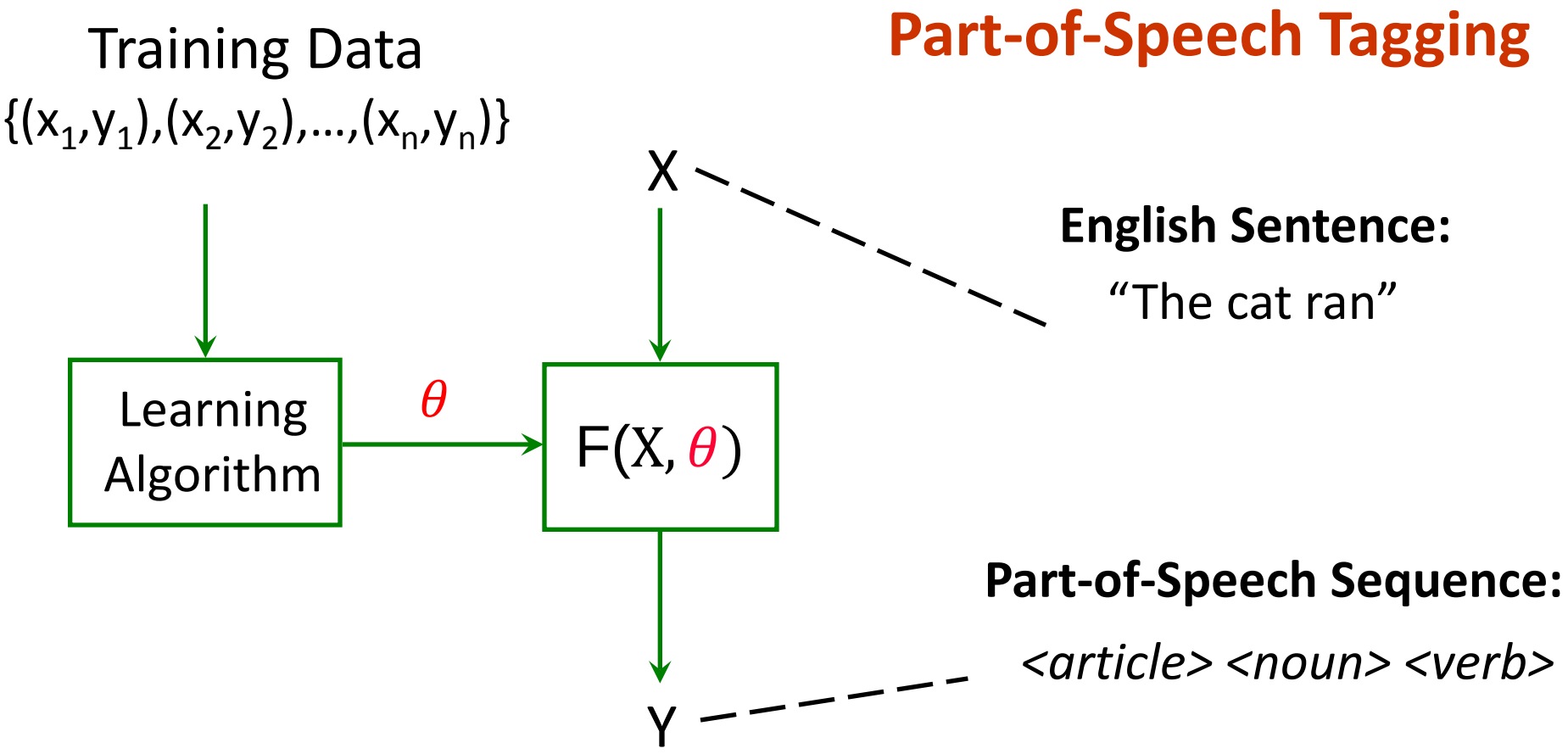
Learning for Simple Outputs



Learning for Simple Outputs



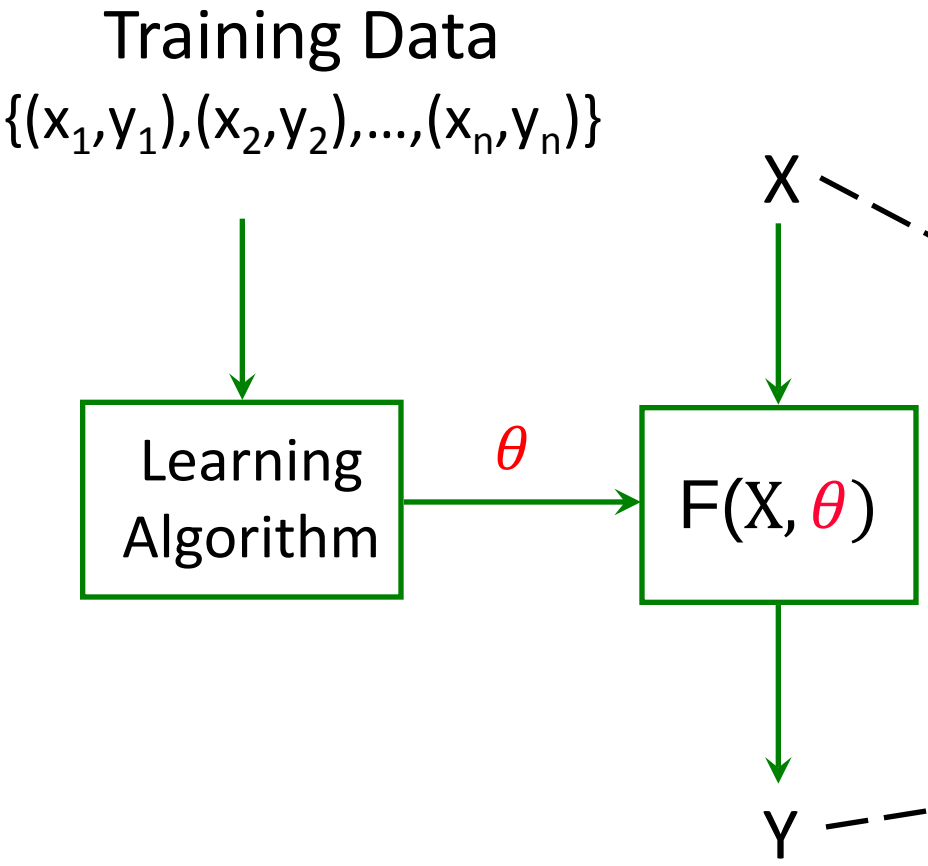
Learning for Structured Outputs



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

Exponential !!

Learning for Structured Outputs



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

Exponential !!

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:

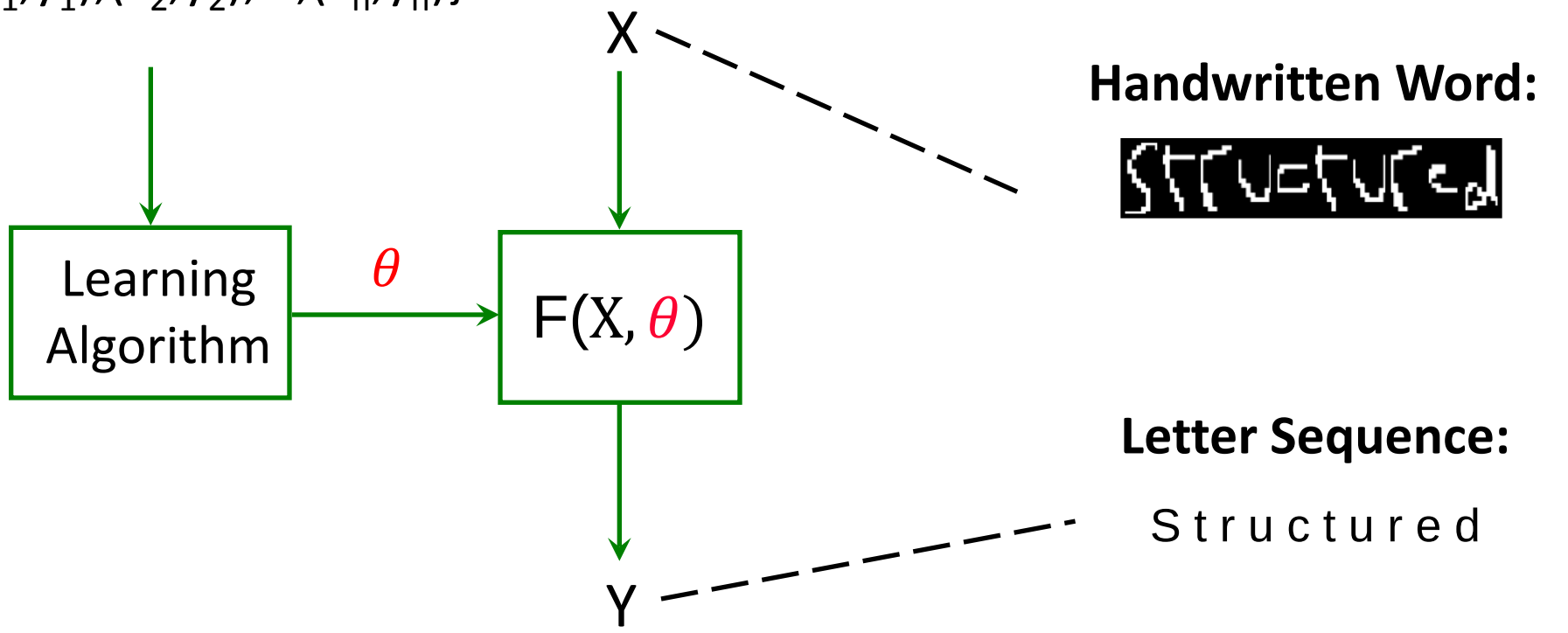
*"**Barack Obama** nominated **Hillary Clinton** as his **secretary of state** on Monday. **He** chose **her** because **she** had foreign affair experience as a former **First Lady**."*

Learning for Structured Outputs

Training Data

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

Handwriting Recognition



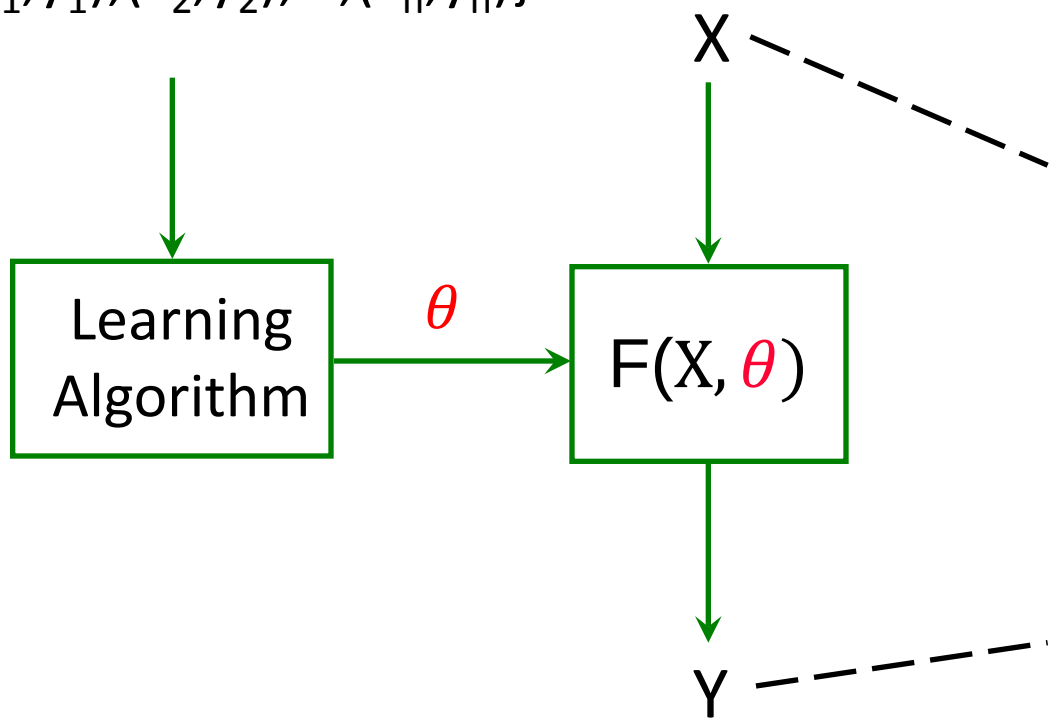
$\boldsymbol{Y}$ = set of all possible letter sequences

Exponential !!

Learning for Structured Outputs

Training Data

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



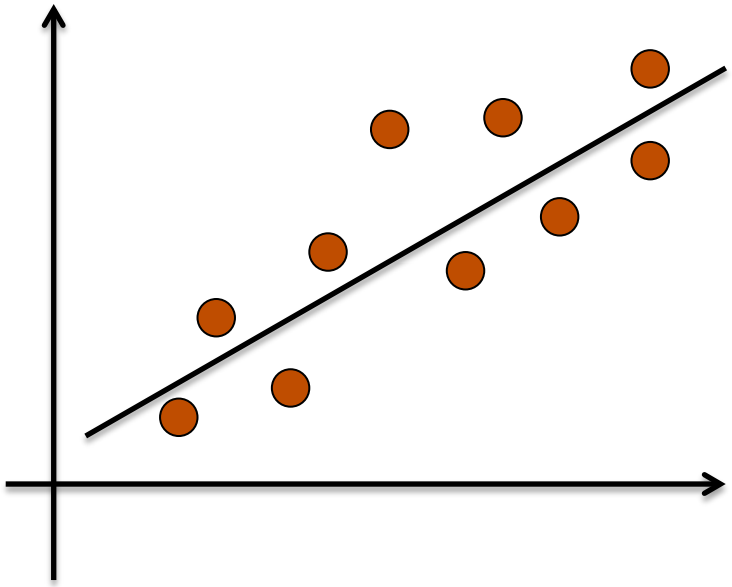
Y = set of all possible labelings

Exponential !!

Image Labeling

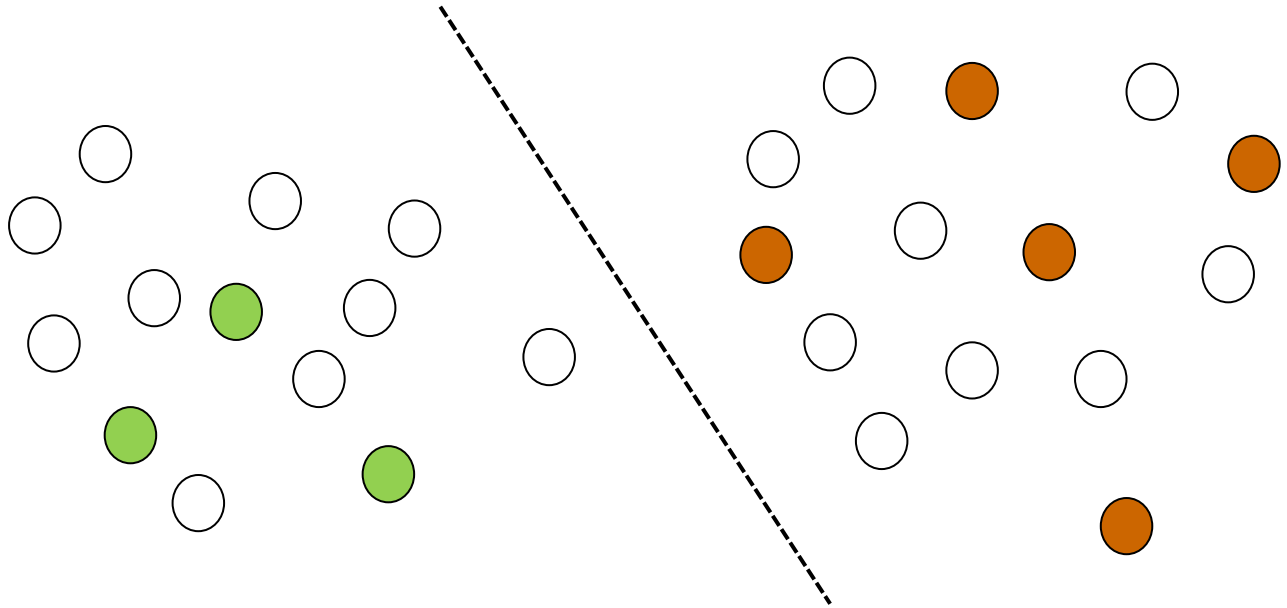
Regression

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



Semi-Supervised Learning

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



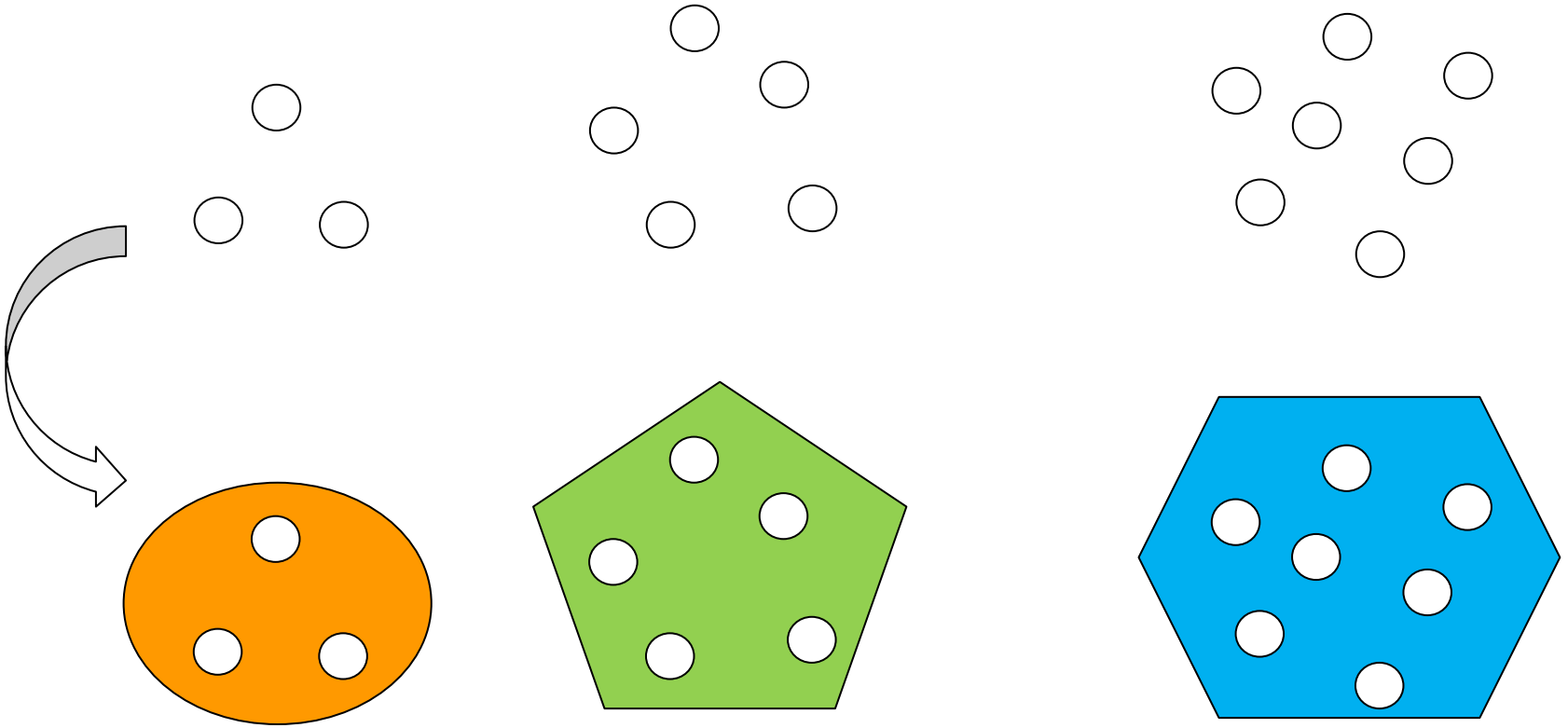
  - find a classifier that separates the labeled points and separates the unlabeled points "well"

Semi-Supervised Learning

- **Co-Training Style Algorithms**
  - Leverage diversity in the learners to learn from each other
  - Diversity comes from multiple (redundant) views of the input – In webpage classification, one view is the "words" on the page and another view is the "links" that point to that page
  - If only one view, employ learners with different hypothesis spaces to achieve diversity

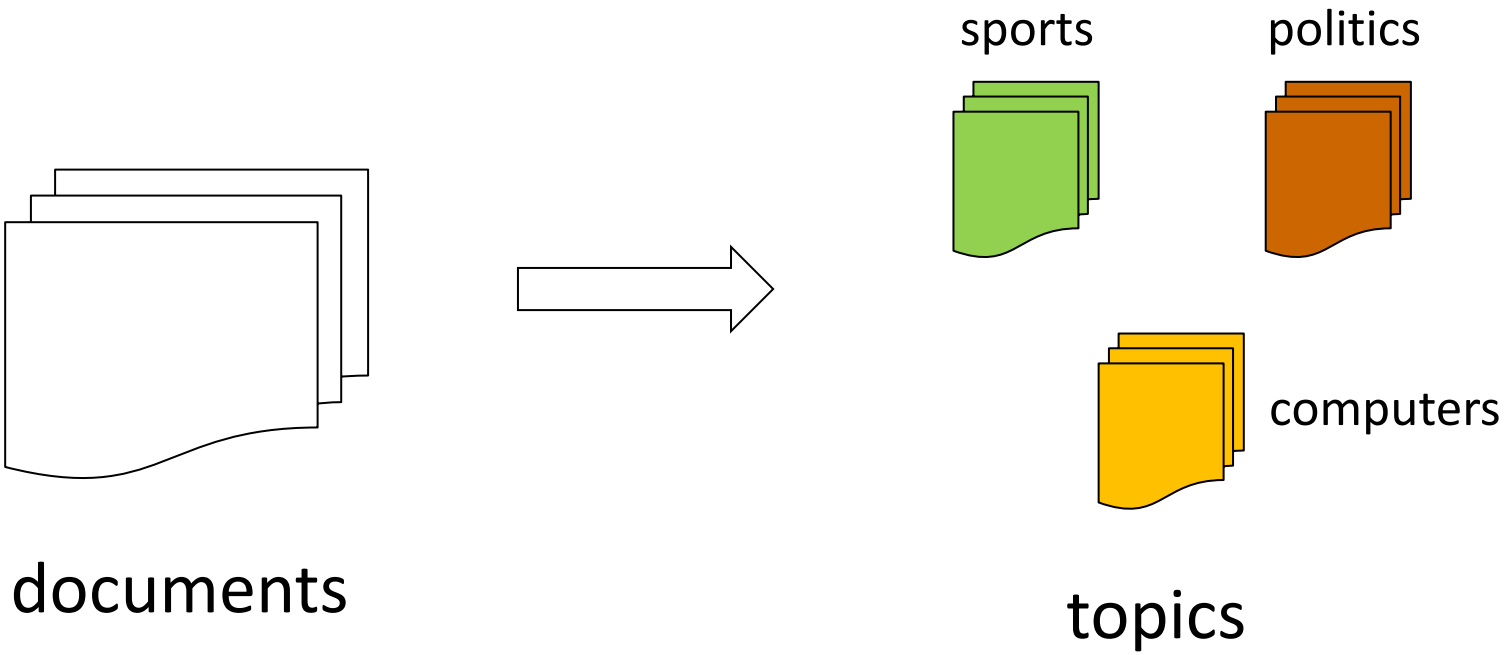
Unsupervised Learning

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



Unsupervised Learning

- **Text Clustering**

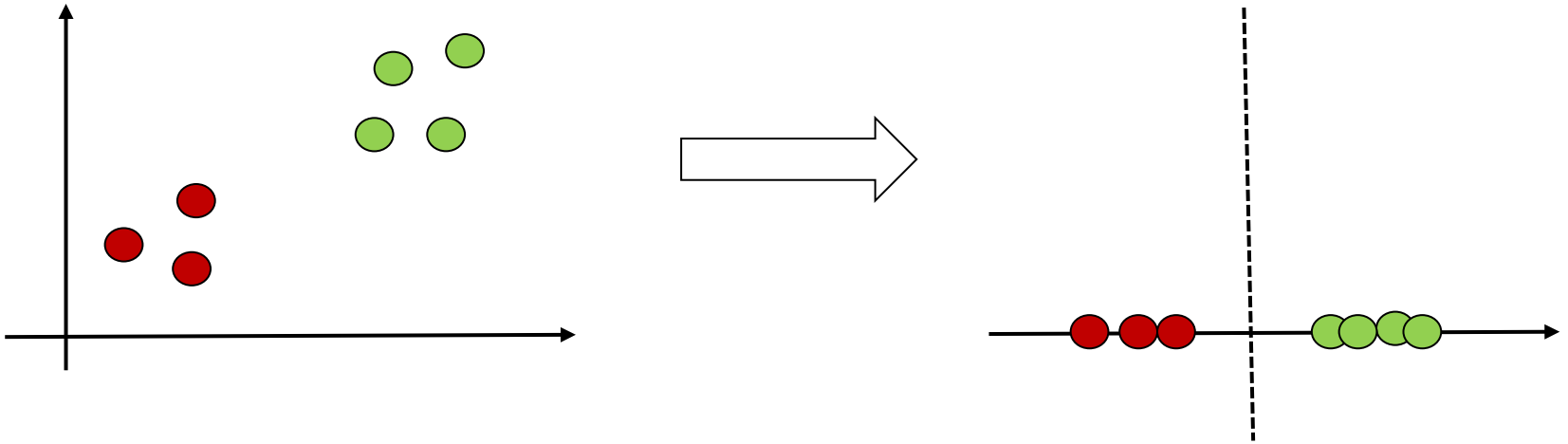


Unsupervised Learning

- **Image Segmentation**

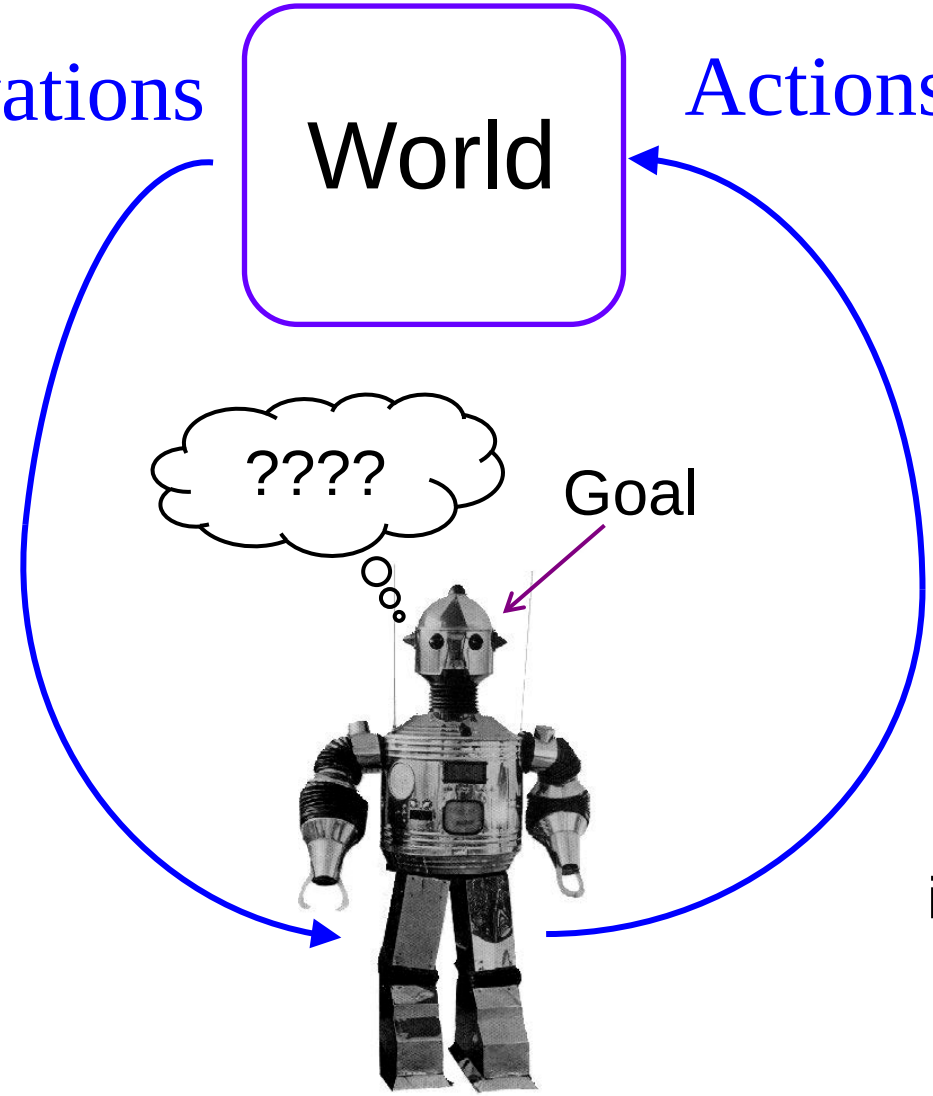
Unsupervised Learning

- **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

Reinforcement Learning



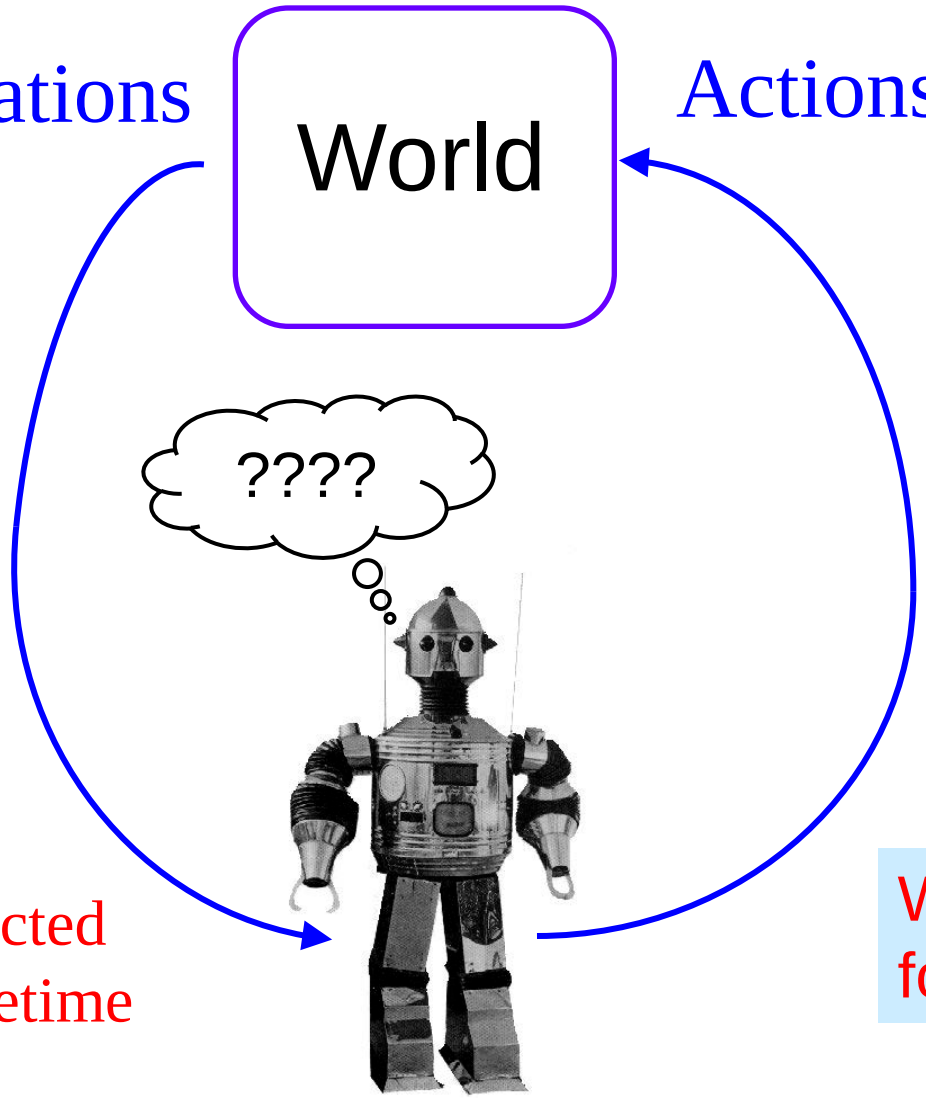
fully observable vs. partially observable

sole source of change vs. other sources

deterministic vs. stochastic

instantaneous vs. durative

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



fully observable

sole source of change

stochastic

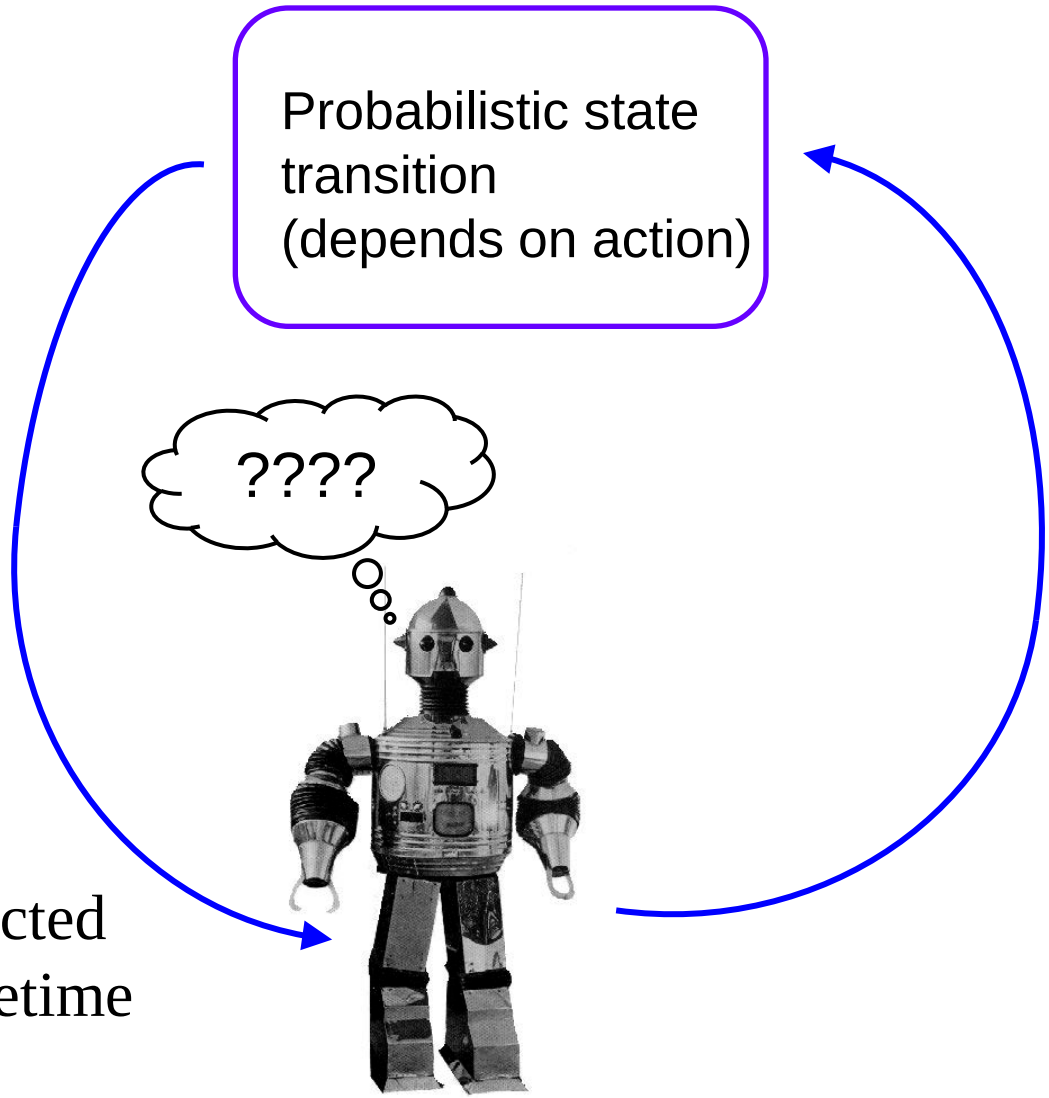
instantaneous

Goal

maximize expected reward over lifetime

We will primarily focus on MDPs

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

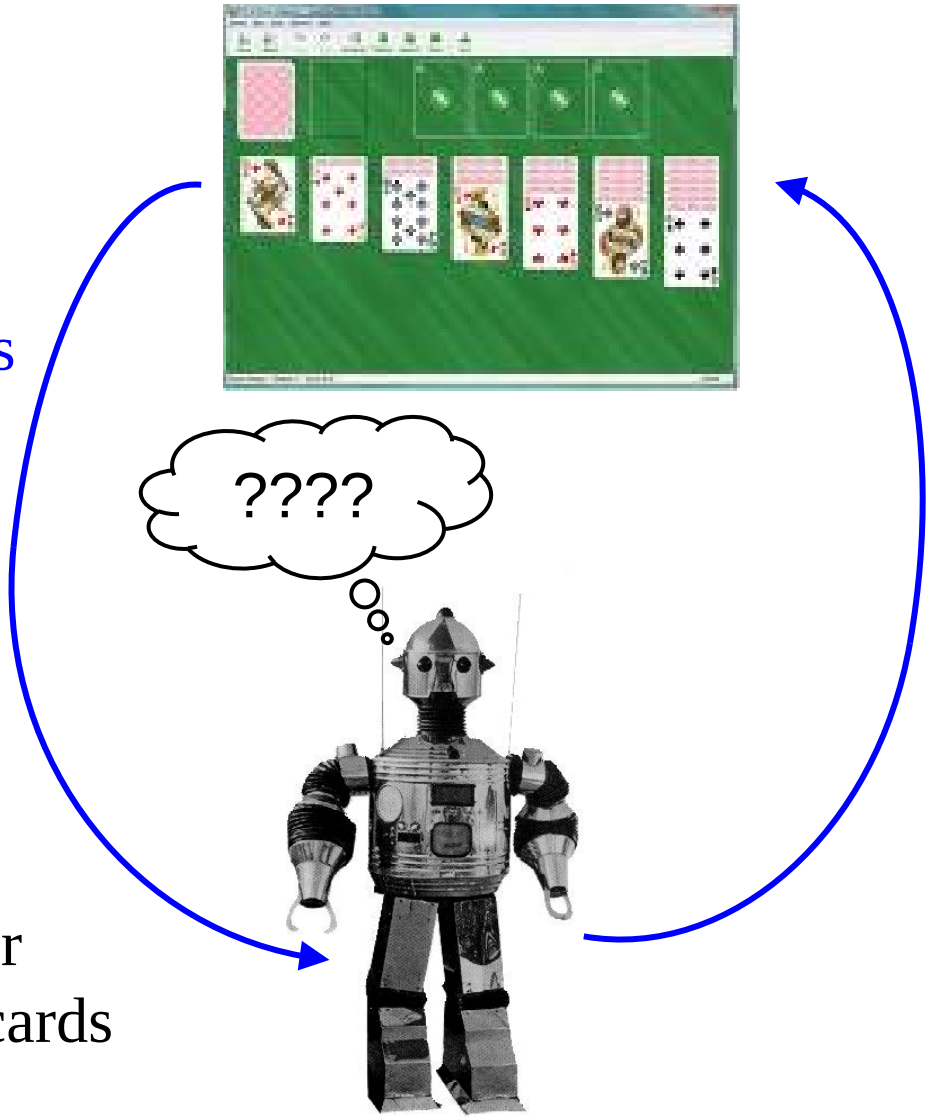


Goal

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

Questions ?