Intro to Deep Learning

ECE408 / CS483 / CSE408

Carl Pearson

pearson@illinois.edu

October 26, 2017

Outline

Introduction

Machine Learning Basics

- How to solve problems you can’t formally describe
- How to determine what features of data are relevant

The story so far

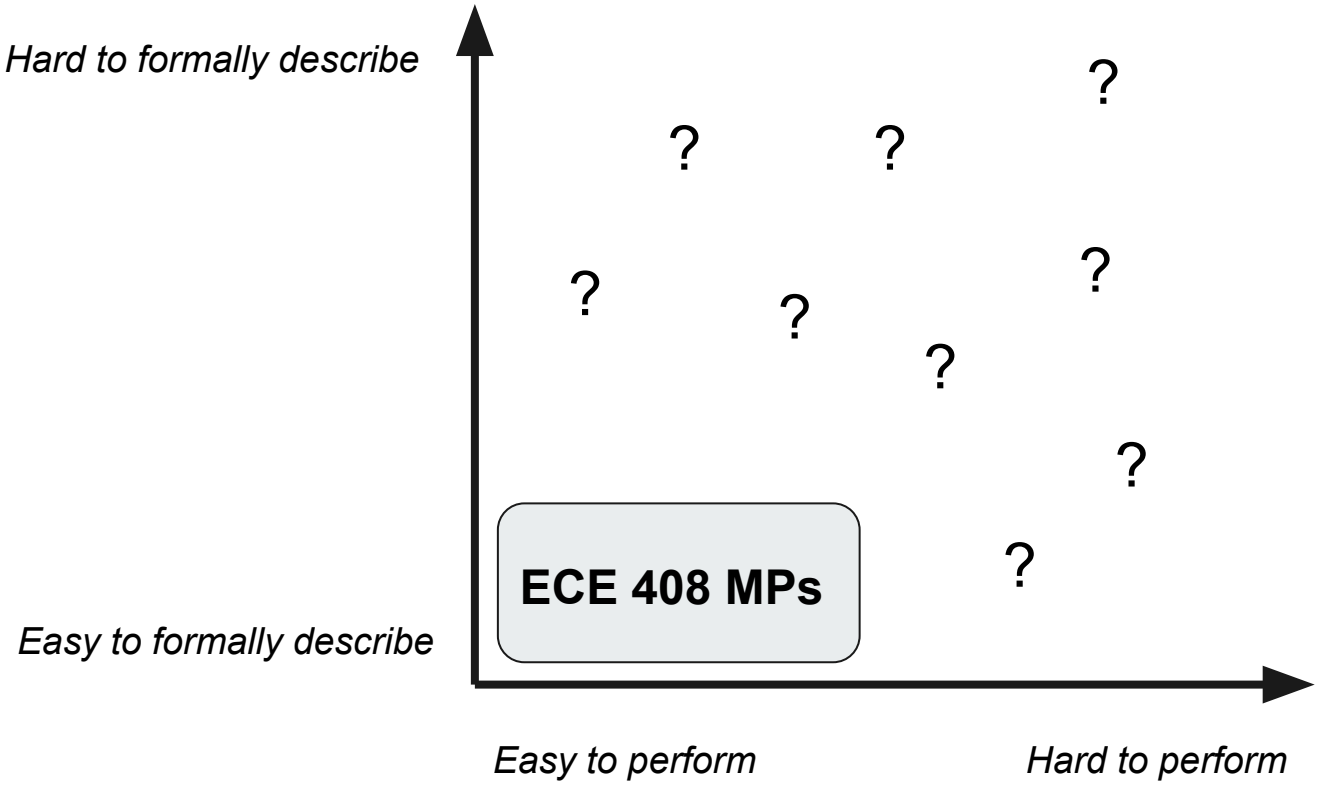
The problem: easily defined

The algorithm: easily specified

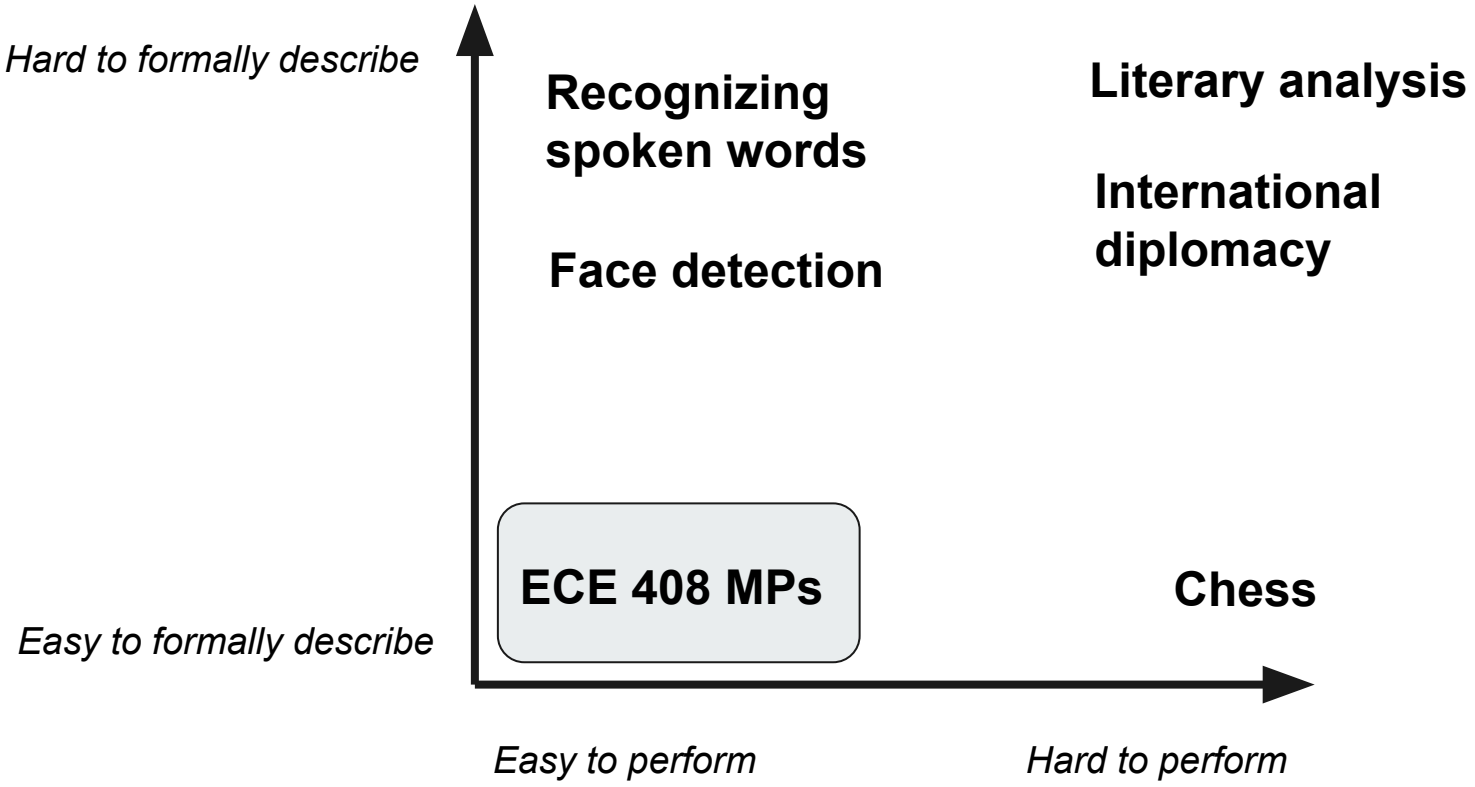
The output: doesn't take too long to produce

Figure adapted from deeplearningbook.org

Types of problems

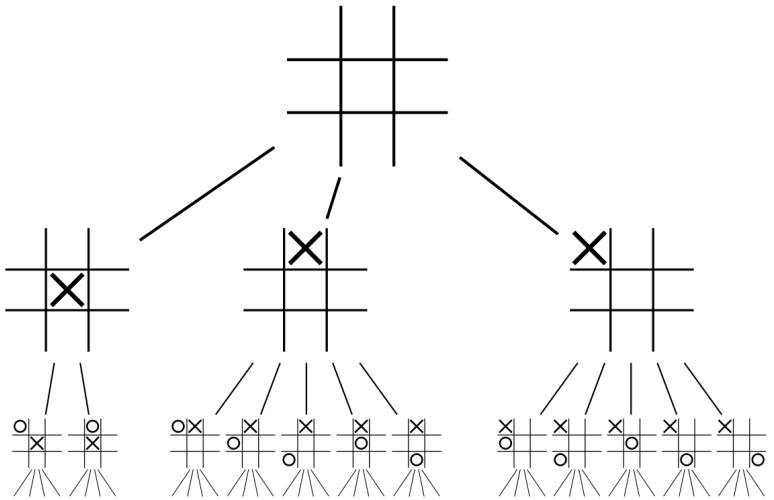


Types of problems



Chess as an AI Success (1)

- Easy to formalize
  - 64 locations, 32 pieces
  - Few allowable moves
- Score each leaf in a tree of possible moves
- Proceed down path with best score



2-ply game tree for tic-tac-toe

Chess as an AI Success (2)

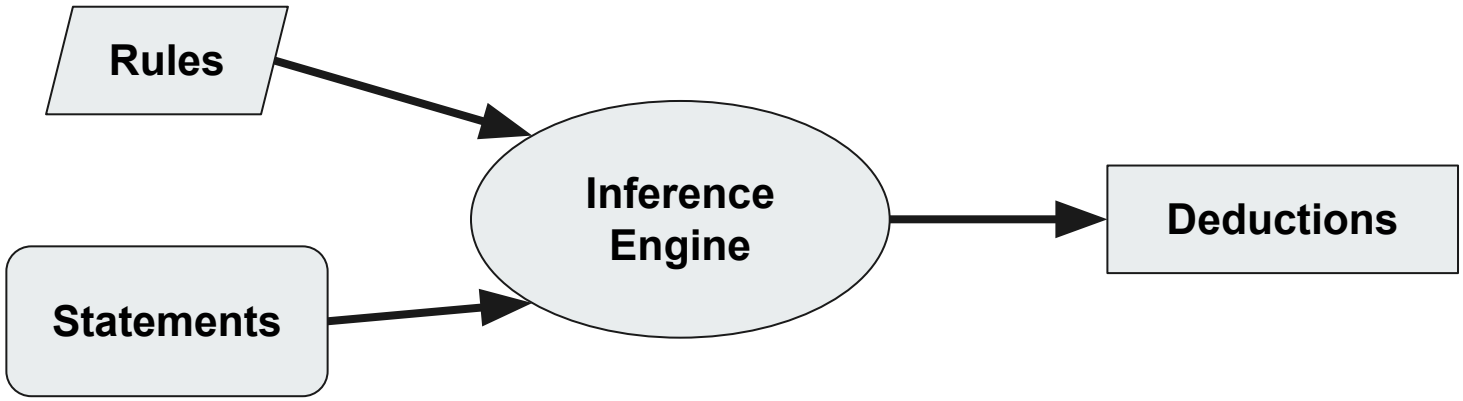
Deep Blue defeated Gary Kasparov in 1997

- Hard to perform
  - ~30 legal moves per position
  - 10^{15} moves for 10-ply lookahead
  - 30 years at 1M positions/sec
- Heuristics, pruning, parallel search, and fast computers

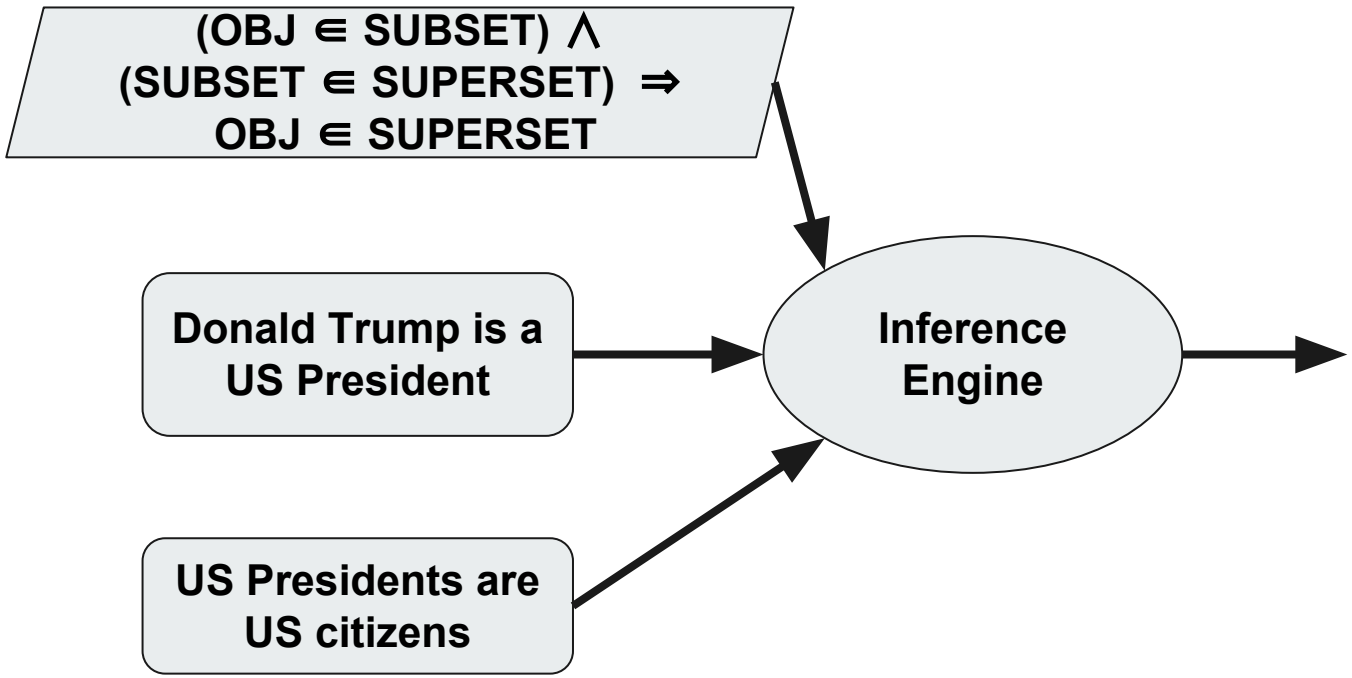
Cyc: Extending Rule-based Systems to the Real World

Comprehensive ontology and knowledge base of common sense

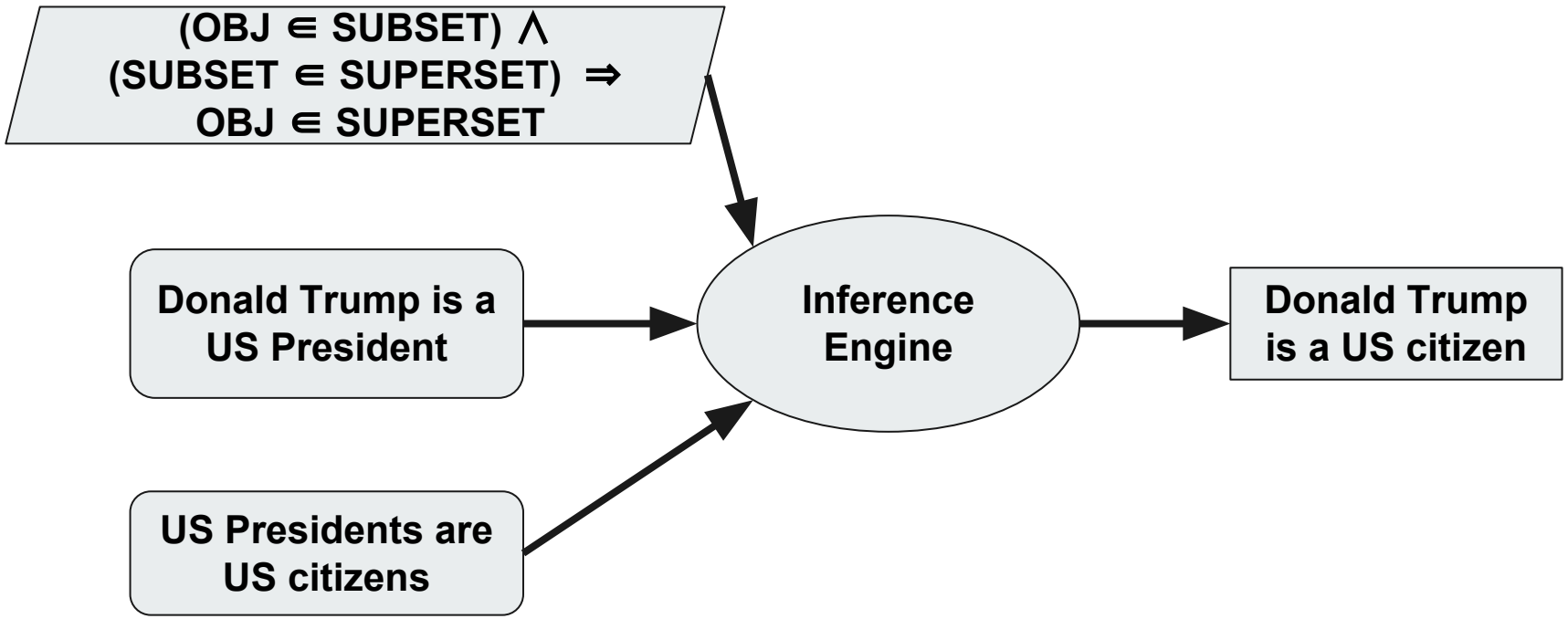
Cyc reasons about formal statements about the world



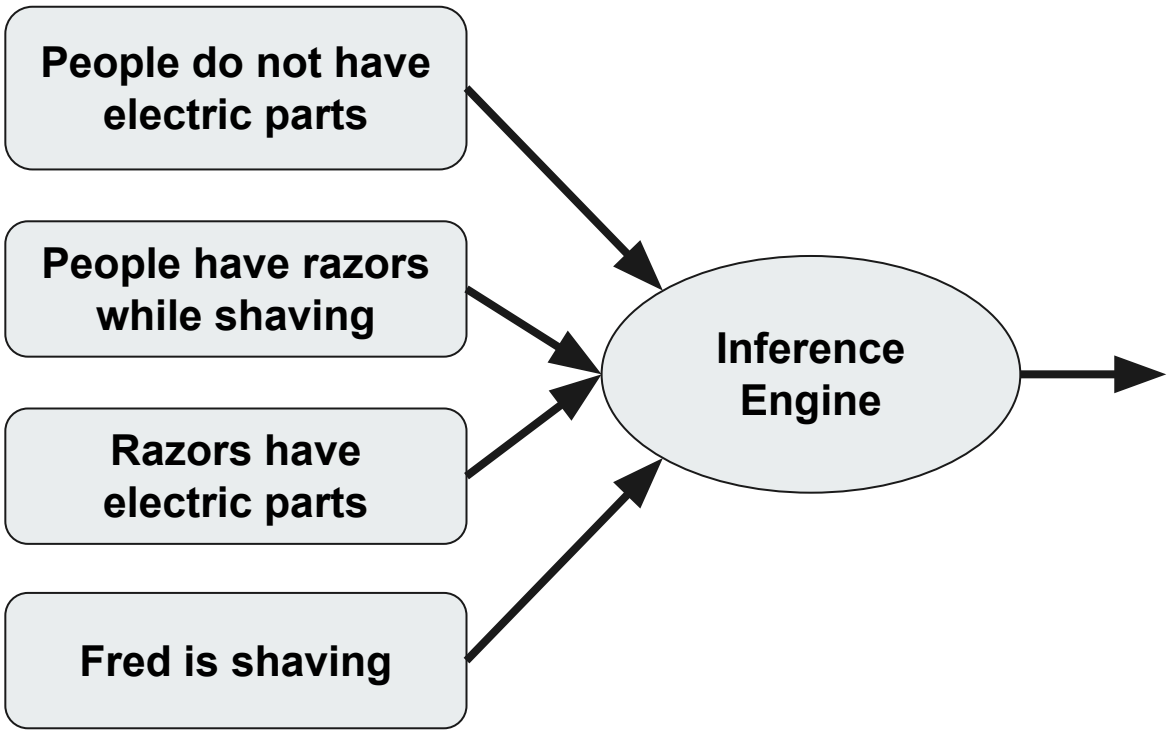
Cyc: a simple example



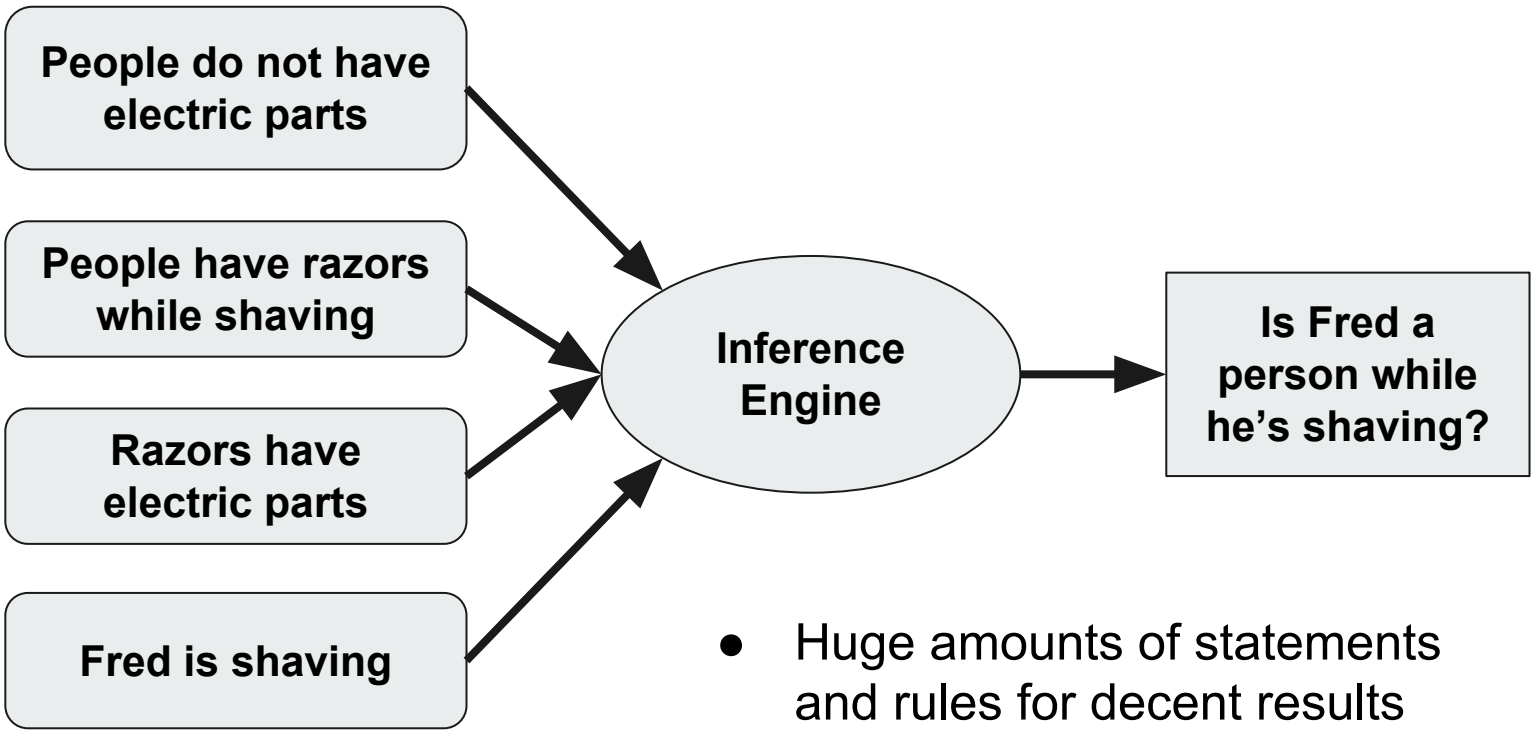
Cyc: a simple example



Cyc: FredWhileShaving

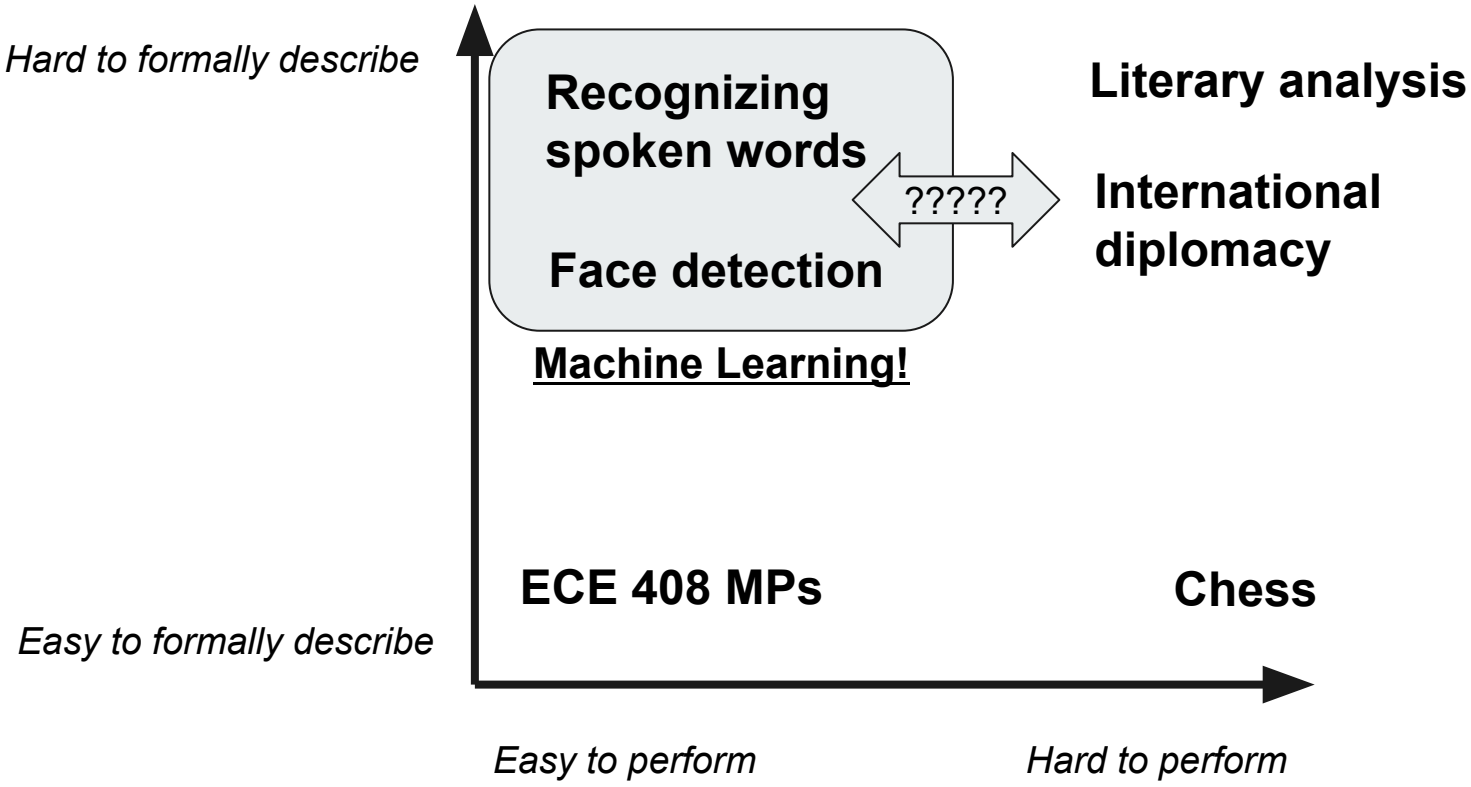


Cyc: FredWhileShaving



- Huge amounts of statements and rules for decent results
- Cannot learn new rules or statements on its own

Types of problems



The “Machine Learning” Approach

Challenge	Solution
Hard to formalize the problem	Don’t formalize the problem Let the machine learn from experience

Classic Machine Learning

Learn how features are associated with outputs

Does not help choose features

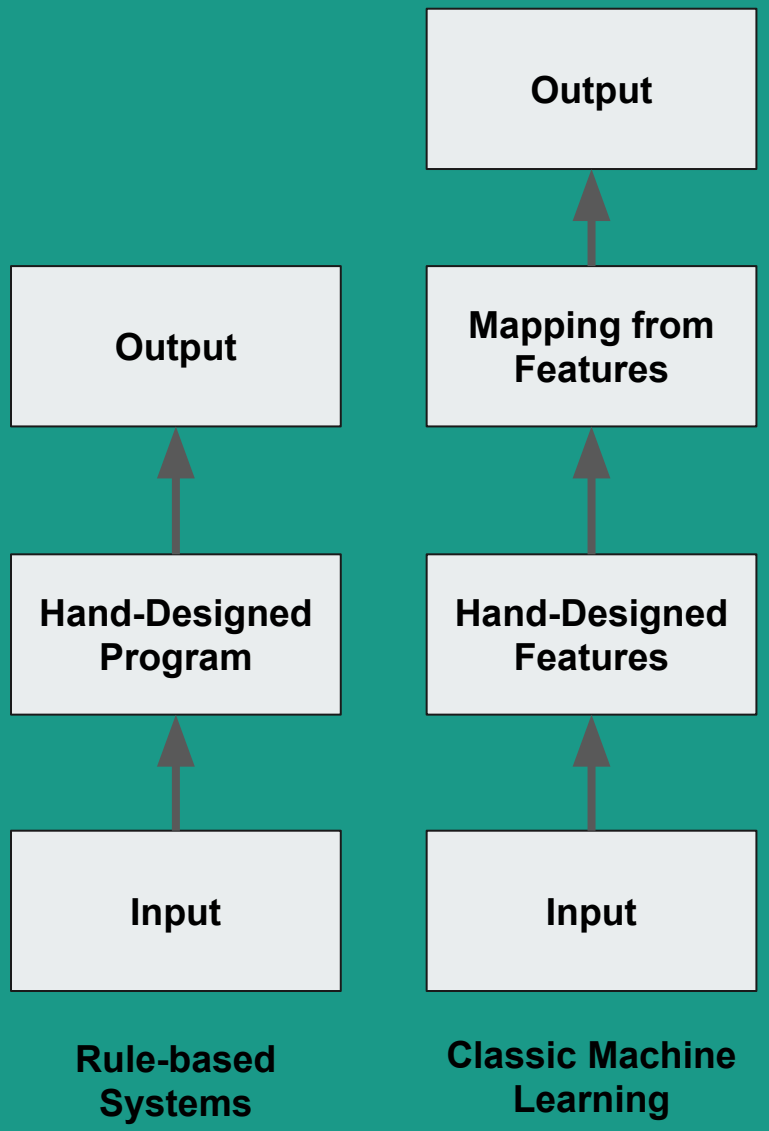


Figure adapted from deeplearningbook.org

You may have heard of...

Naive Bayes

Features as independent contributors to output

Logistic Regression

Learn how to weight each feature to output, usually through gradient descent*

* more on this in a later lecture...

MCMLIV

– MDCCCVI

MCMLIV

− MDCCCVI

CXXXIX

1945

\- 1806

$$\begin{array}{r} 1\,9\,\cancel{4}^{3}\,5^{1} \\ -\;\;1\,8\,0\,6 \\ \hline \\ \end{array}$$

Importance of Data Representation

Arithmetic

MCMLIV - MDCCCVI

1945 - 1806

Importance of Data Representation

Arithmetic

MCMLIV - MDCCCVI

1945 - 1806

Searching

Binary Tree vs Linked list

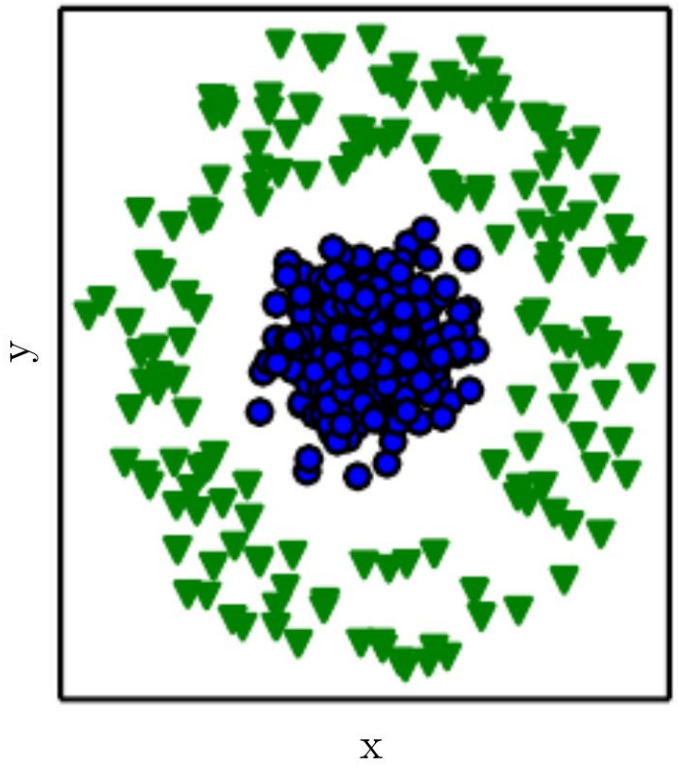


Figure adapted from deeplearningbook.org

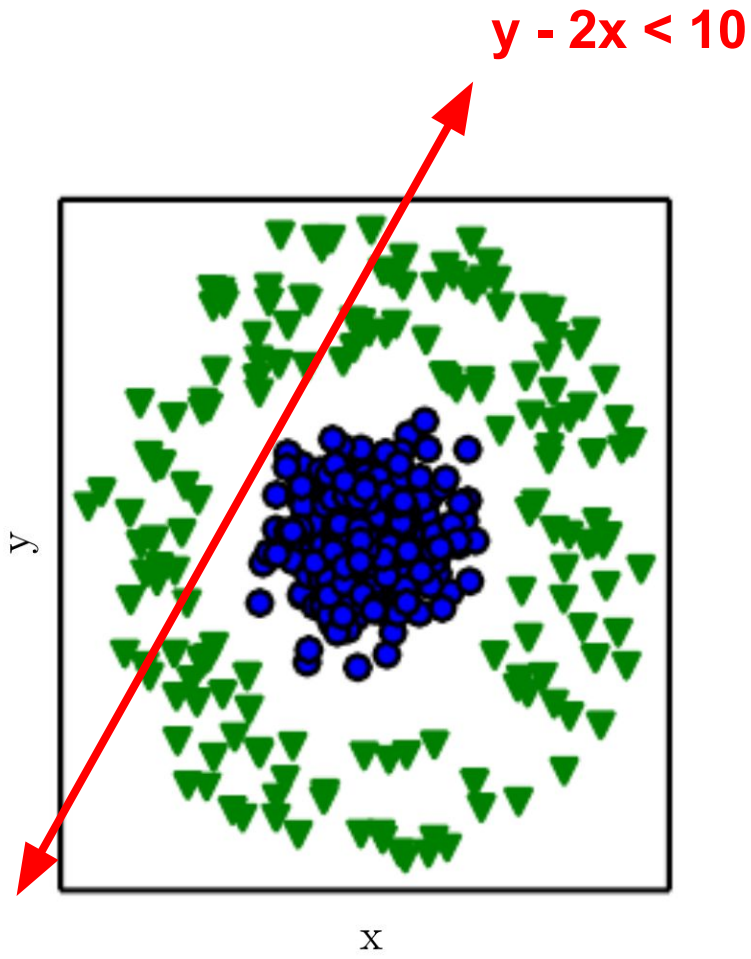


Figure adapted from deeplearningbook.org

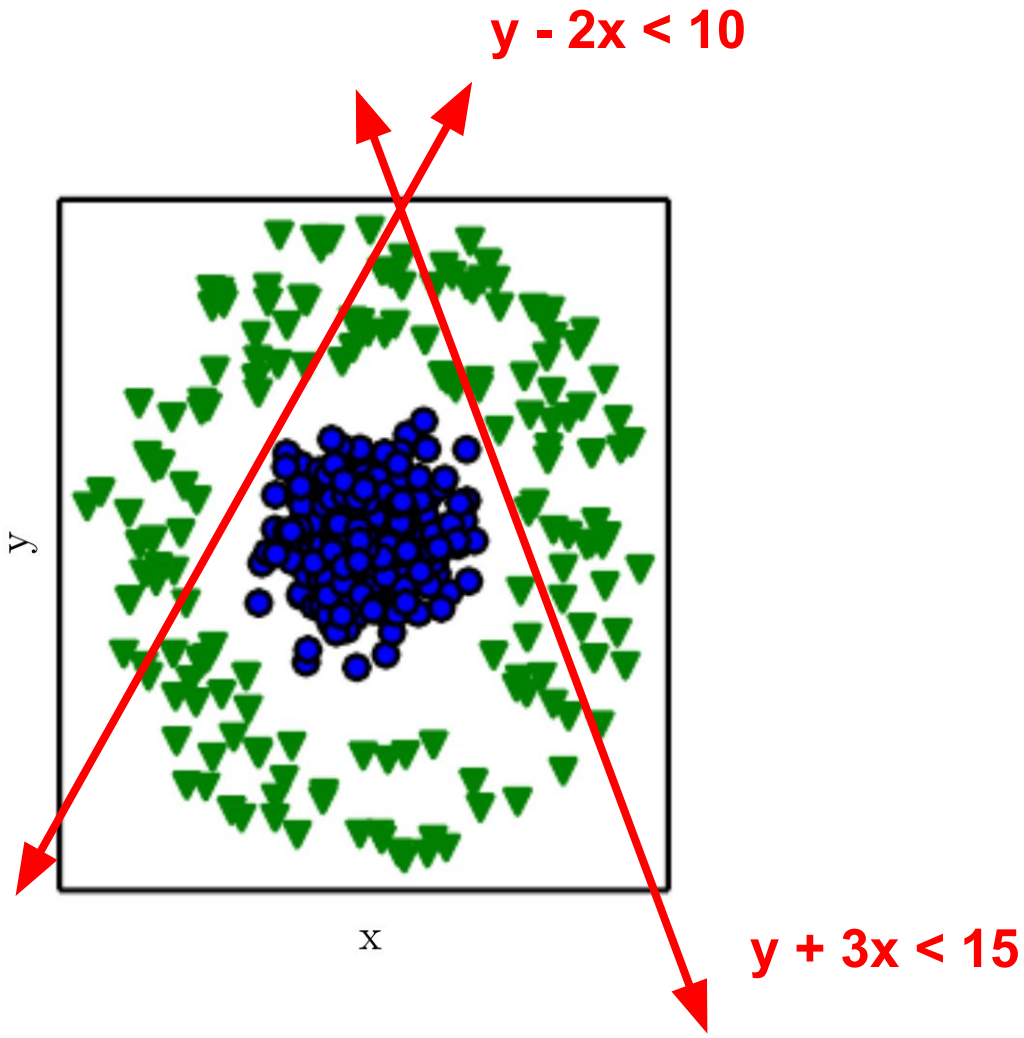


Figure adapted from deeplearningbook.org

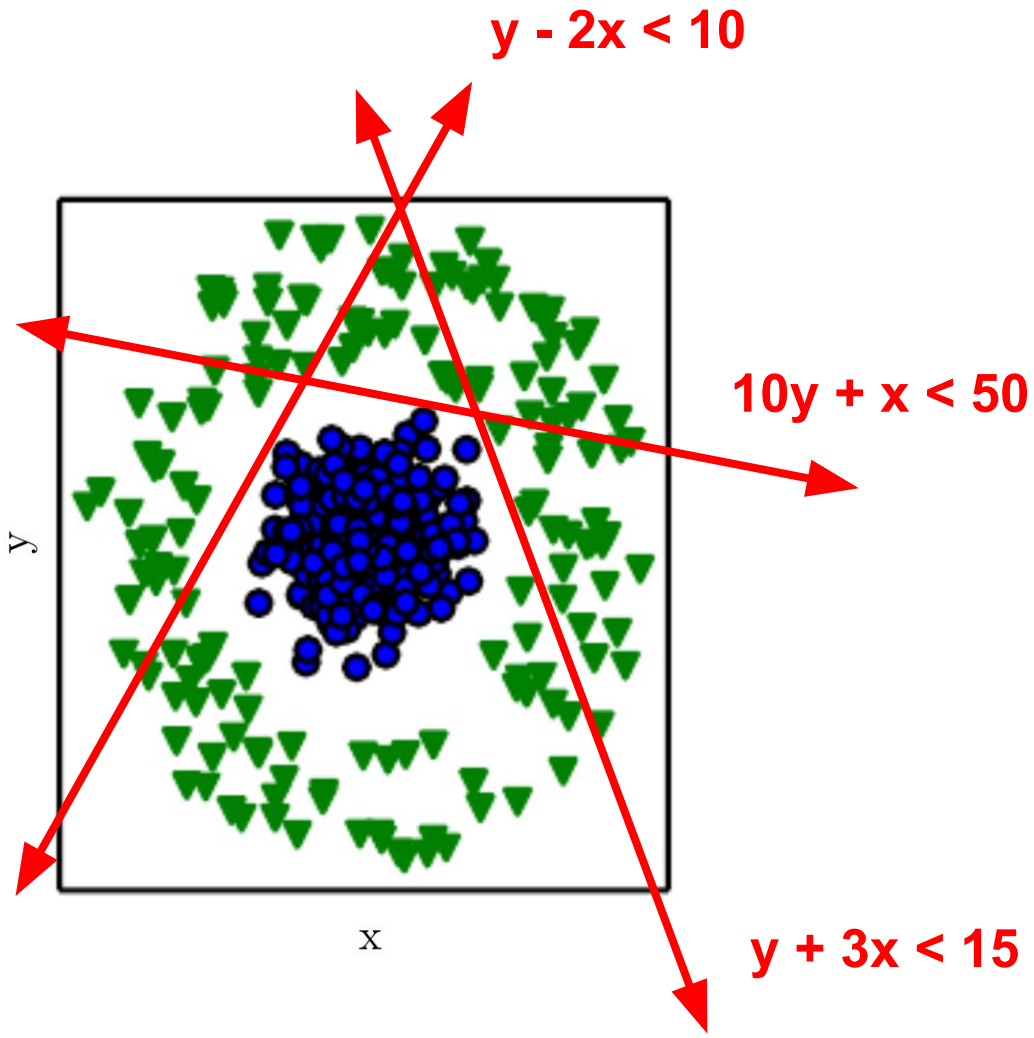
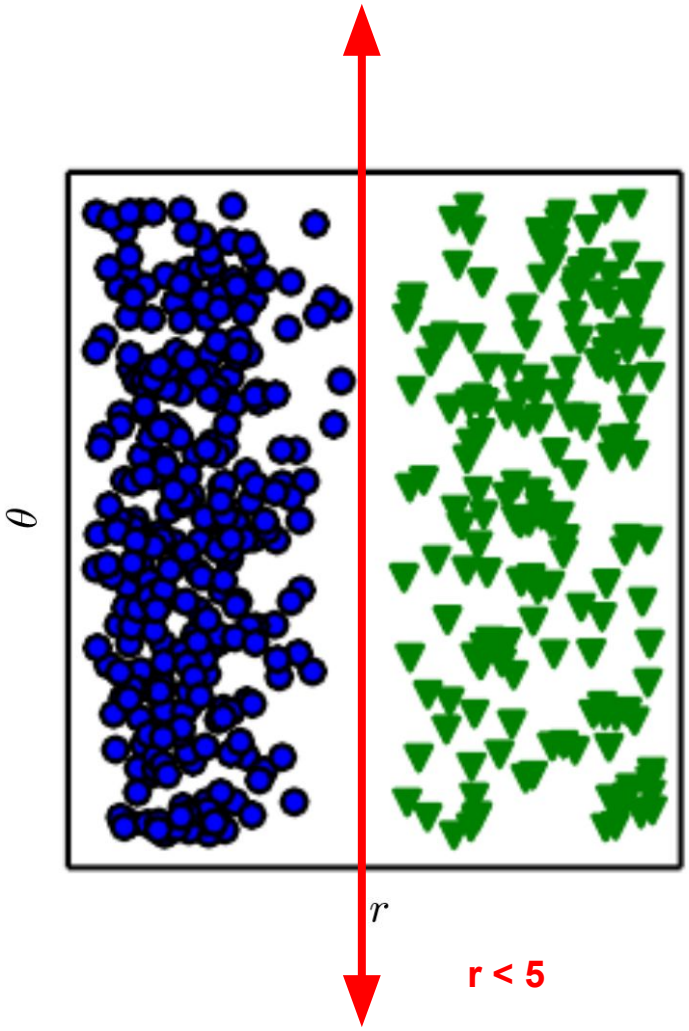


Figure adapted from deeplearningbook.org



$\Theta = \arctan(y / x)$

$r = \mathrm{sqrt}(x^2 + y^2)$

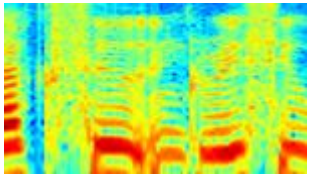
Different features for different tasks

Image

Vision Features

Detection

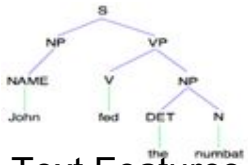
Audio



Audio Features

Identify Speaker

Text



Text Features

Text classification, machine translation, information retrieval

Which Data Features are Relevant?

- Detecting a car in an image
- Cars have wheels ➡ possible feature: presence of a wheel
- Can we describe the pixel values that make up a wheel:
  - Circle-shaped?
  - Dark around the perimeter?

Which Data Features are Relevant?

- Detecting a car in an image
- Cars have wheels ➡ possible feature: presence of a wheel
- Can we describe the pixel values that make up a wheel:
  - Circle-shaped?
  - Dark around the perimeter?
- But what about
  - Occlusion
  - Perspective
  - Shadows
  - White-walled tires
  - ...

Identify factors of variation that explain data

- Unobserved objects or forces that affect observed quantities
- Mental constructs that provide simplifying explanations or inferred causes
- Speech:
  - Age, sex, accent, words being said
- Car:
  - Position, color, angle of sun
- Many factors of variation influence each piece of observed data

The “Representation Learning” Approach

Challenge

Which data features are relevant?

Solution

Learn the features too!

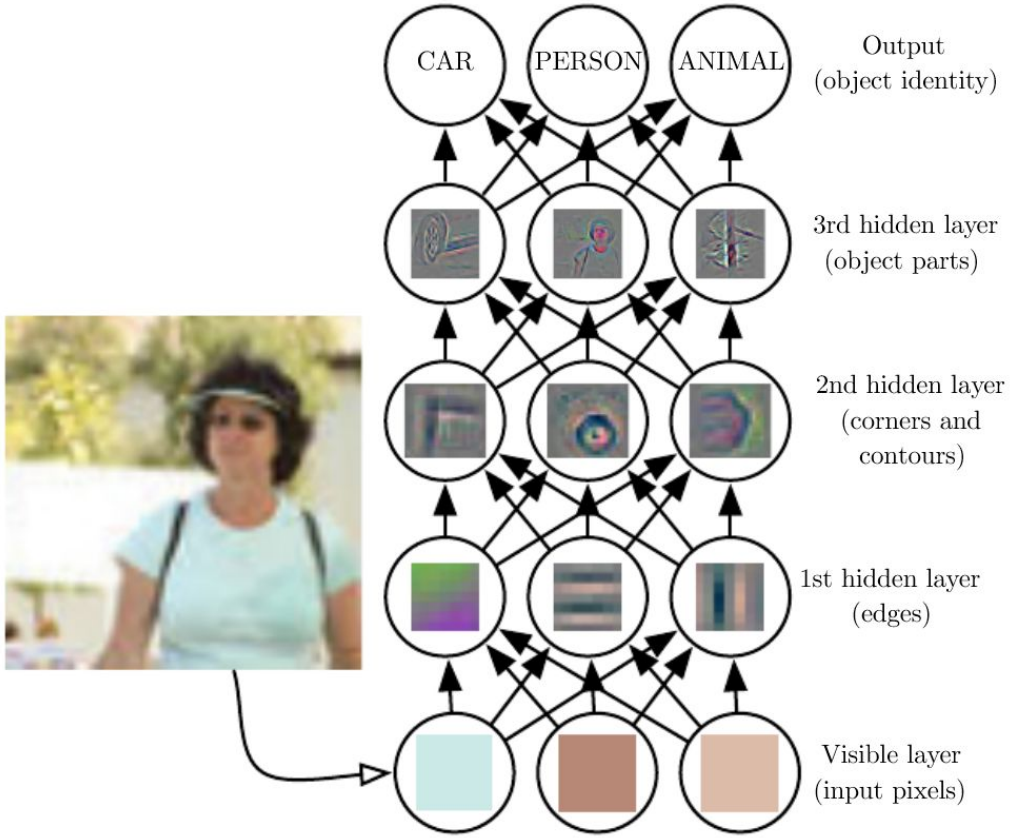
(Looking ahead)
Deep Learning: a deep hierarchy of concepts

Machine Learning

Ability to acquire knowledge by extracting patterns from data

Deep Learning

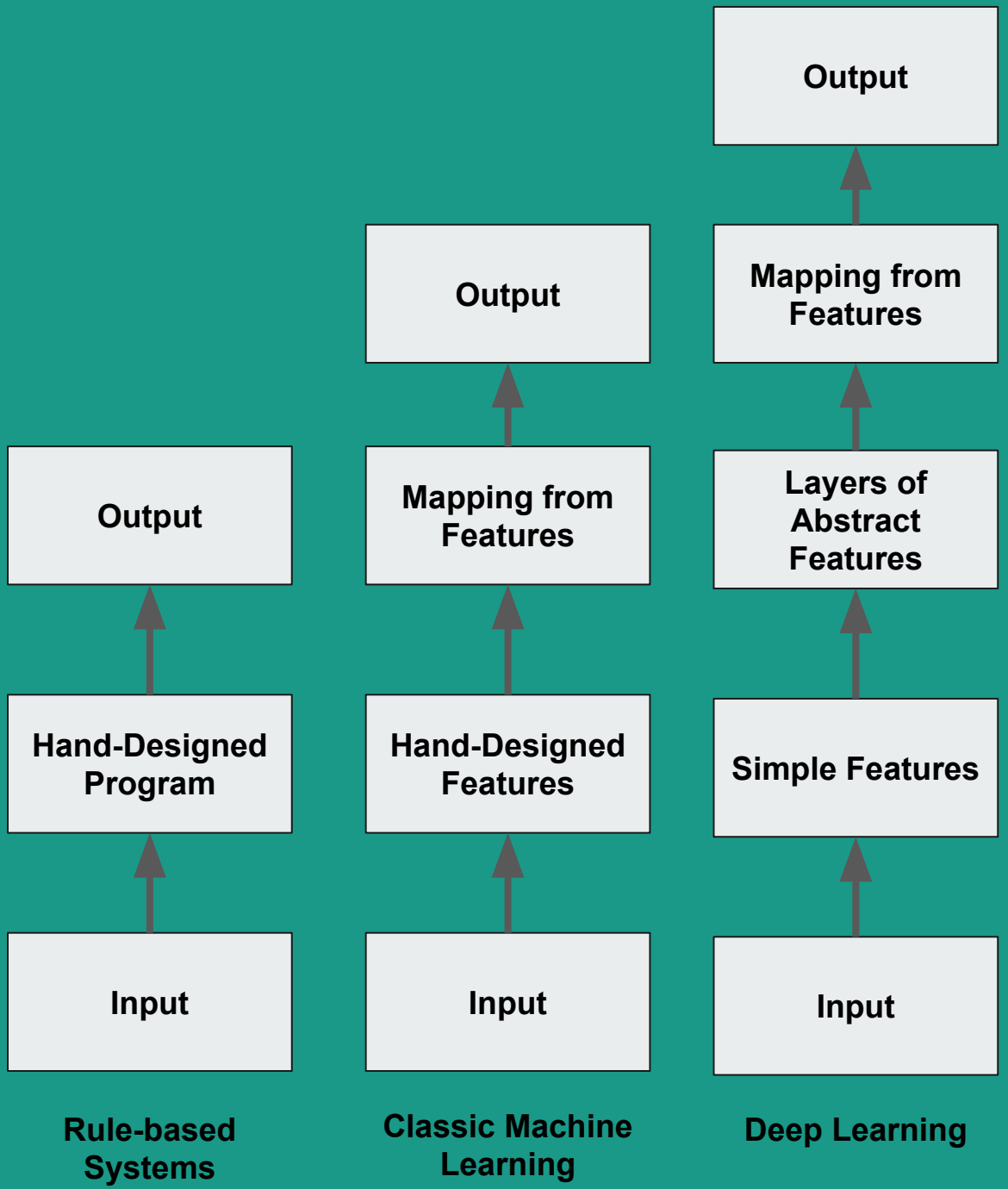
- A type of representation learning
- Representations expressed in terms of other representations



The “Deep Learning” Approach

Challenge	Solution
Hard to formalize the problem	Don’t formalize the problem
	Let the machine learn from experience
Which data features are relevant?	Hierarchy of concepts to capture simple and complicated features
	Learn the hierarchy too!

Summary


Output
Hand-Designed Program
Input
Rule-based Systems
Output
Mapping from Features
Hand-Designed Features
Input
Classic Machine Learning
Output
Mapping from Features
Layers of Abstract Features
Simple Features
Input
Deep Learning

Other Resources

www.deeplearningbook.org

CS440: Artificial Intelligence
CS446: Machine Learning
CS447: Natural Language Processing
CS498 AMO: Applied Machine Learning
CS546: Machine Learning and Natural Language
CS548: Models of Cognitive Processes
CS598 PS: Machine Learning for Signal Processing
CS598 TEL: Machine Learning Theory
ECE448 Intro to Artificial Intelligence
ECE598 NS: Machine Learning in Silicon
ECE598 PV: Learning: Algorithms and Models
STAT542: Statistical Learning