

# Artificial Intelligence

## CS 404

Fall 2024

# Overview

- **Course Info**
  - All of the information regarding the course can be found at SUCourse.
- **Introduction to AI**

- **Instructor:** Berrin Yanıkoğlu    Office: MDBF 2056
- **Textbook:** Artificial Intelligence: A Modern Approach (**AIMA**), by Stuart Russell and Peter Norvig., 4<sup>th</sup> edition. Previous editions are a bit simpler and adequate, if you happen to have those books passed on to you. Link to the book's web page, where you can find some useful information: <http://aima.cs.berkeley.edu/>
- **Quizzes:** There will be one quiz every week, starting with the second week (3-5 questions, 3-5 minutes, 10pts), on the Tuesday lecture.
  - **You may skip two quizzes without any problems** (the lowest two quiz grades will be dropped at the end of the course). You will not have any other exception besides this, so no emails about missed quizzes or connection problems etc.

- **Homeworks:** There will be **5** homeworks total, given approximately every other week and there will be no project.
  - **Late homeworks will be accepted for 2 days, with a 5pt penalty per day.**
  - Questions about homeworks etc should be asked at SUCourse discussion boards. Don't ask homework related questions on email, so that the TA's answer to your question helps everyone else who may have the same question.
- **Grading:** Quizzes (20%) + Homeworks (25%) + Midterm (25%) + Final (30%)
  - To pass the course, your grade as calculated above must be at least 40 and your final exam grade should be at least 30/100.
- This tentative plan may change slightly in the first week of the classes.

# What we will cover

## **Artificial Intelligence: A Modern Approach (AIMA) Russell and Norvig**

### **I Artificial Intelligence**

1 Introduction ... 1

2 Intelligent Agents ... 36

### **II Problem-solving**

3 Solving Problems by Searching ... 63

4 Search in Complex Environments ... 110

5 Adversarial Search and Games ... 146

6 Constraint Satisfaction Problems ... 180

### **III Knowledge, reasoning, and planning**

7 Logical Agents ... 208

8 First-Order Logic ... 251

9 Inference in First-Order Logic ... 280

10 Knowledge Representation ... 314

11 Automated Planning ... 344

### **IV Uncertain knowledge and reasoning**

12 Quantifying Uncertainty ... 385

13 Probabilistic Reasoning ... 412

14 Probabilistic Reasoning over Time ... 461

15 Probabilistic Programming ... 500

16 Making Simple Decisions ... 528

17 Making Complex Decisions ... 562

18 Multiagent Decision Making ... 599

### **V Machine Learning**

19 Learning from Examples ... 651

20 Learning Probabilistic Models ... 721

21 Deep Learning ... 750

22 Reinforcement Learning ... 789

### **VI Communicating, perceiving, and acting**

23 Natural Language Processing ... 823

24 Deep Learning for NLP ... 856

25 Computer Vision ... 881

26 Robotics ... 925

### **VII Conclusions**

27 Philosophy, Ethics, and Safety of AI ... 981

28 The Future of AI ... 1012

# DeepBlue (IBM) - Garry Kasparov 1997

**IBM DeepBlue  
won against  
World  
Champion  
Garry  
Kasparov!**

- Winning in games against humans has been a major goal in AI historically.





# IBM Watson (IBM) – Jeopardy! 2011

## IBM Watson Jeopardy! Won two human champions.

“What is the event that started II. World War?”

- Even more important than the chess victory.
- Involves knowledge representation, logical deduction, as well as speech understanding, natural language understanding, logic, decision making, ...
- How did the computer press on the buzzer? 😊



Konu üzerinde kapsamlı makale: <https://thebestschools.org/magazine/watson-computer-plays-jeopardy/>

# AlphaGo (Google DeepMind ) - Lee Sedol 2016

Almost 20 years after DeepBlue, **AlphaGo** won over the human champion in the game of Go.



Image credit: Netflix



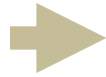
## Big Dog (Boston Dynamics) - 2014



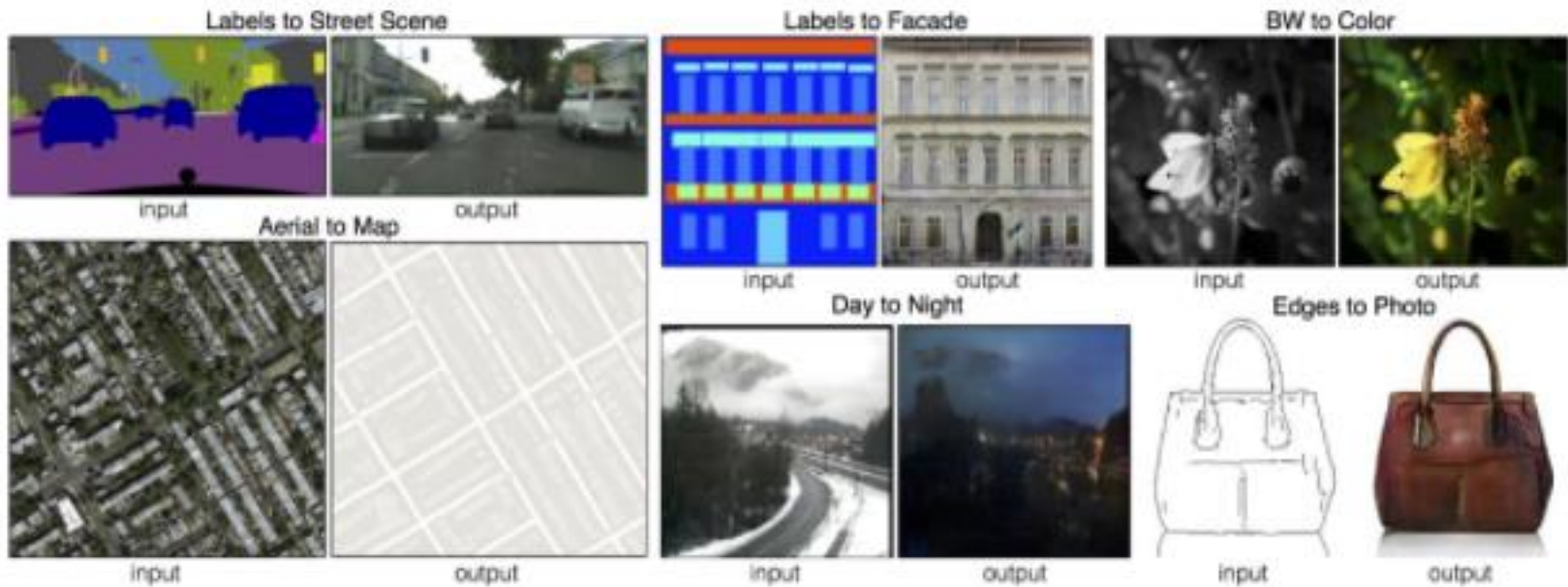
# StyleGANStyleGAN (2018,2020)

Pre-2018

<https://this-person-does-not-exist.com/>



# Neural Style Transfer (2018)



# AlphaFold (Google DeepMind ) - 2020

**AlphaFold**, predicted the 3D shape of a protein, given its amino acid sequence with very high precision.

- Now 200 M proteins are all predicted!

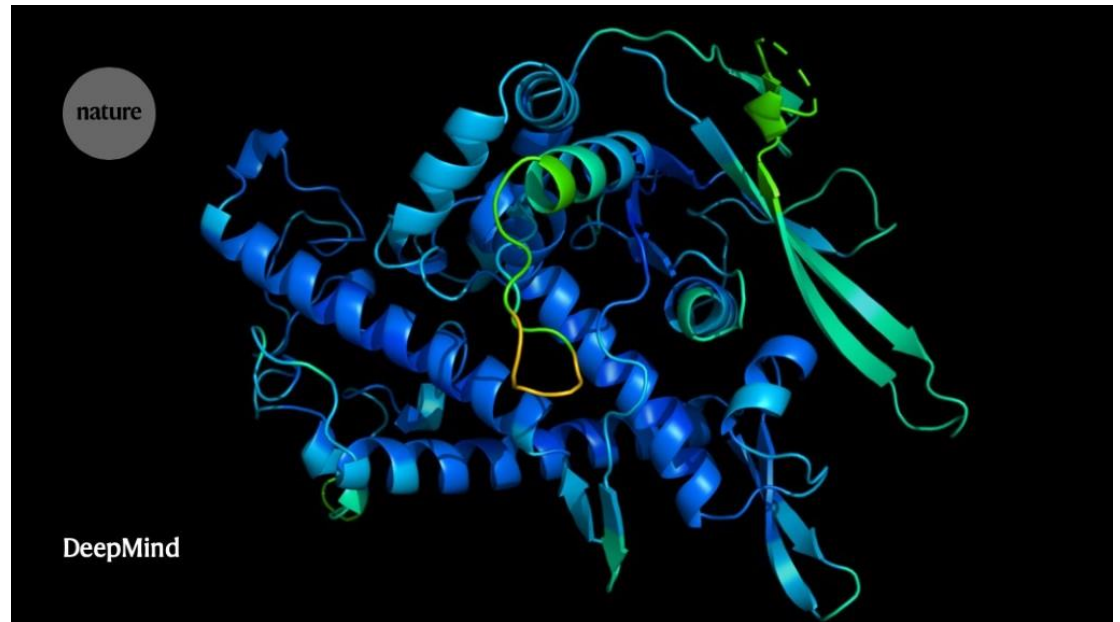


Image credit: Nature

## Dall.E (OpenAI) - 2020

**Produces an image matching the given text.**

- Neural network with 12 billion parameters



## GPT-3 (OpenAI) - 2020

**Continues a given text with suitable continuation.**

GPT-3 on Blockchain. “All cryptocurrency is fraud. It is backed by nothing real and it is made out of thin air (except the electricity used).”



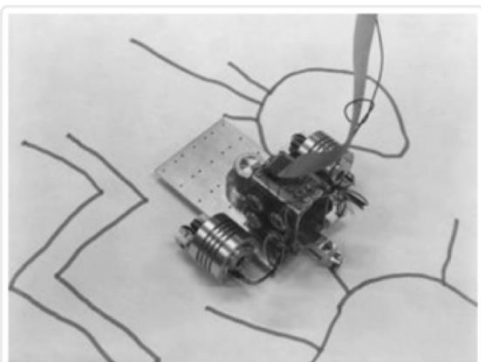


# Aaron - Harold Cohen

## 1980-2012



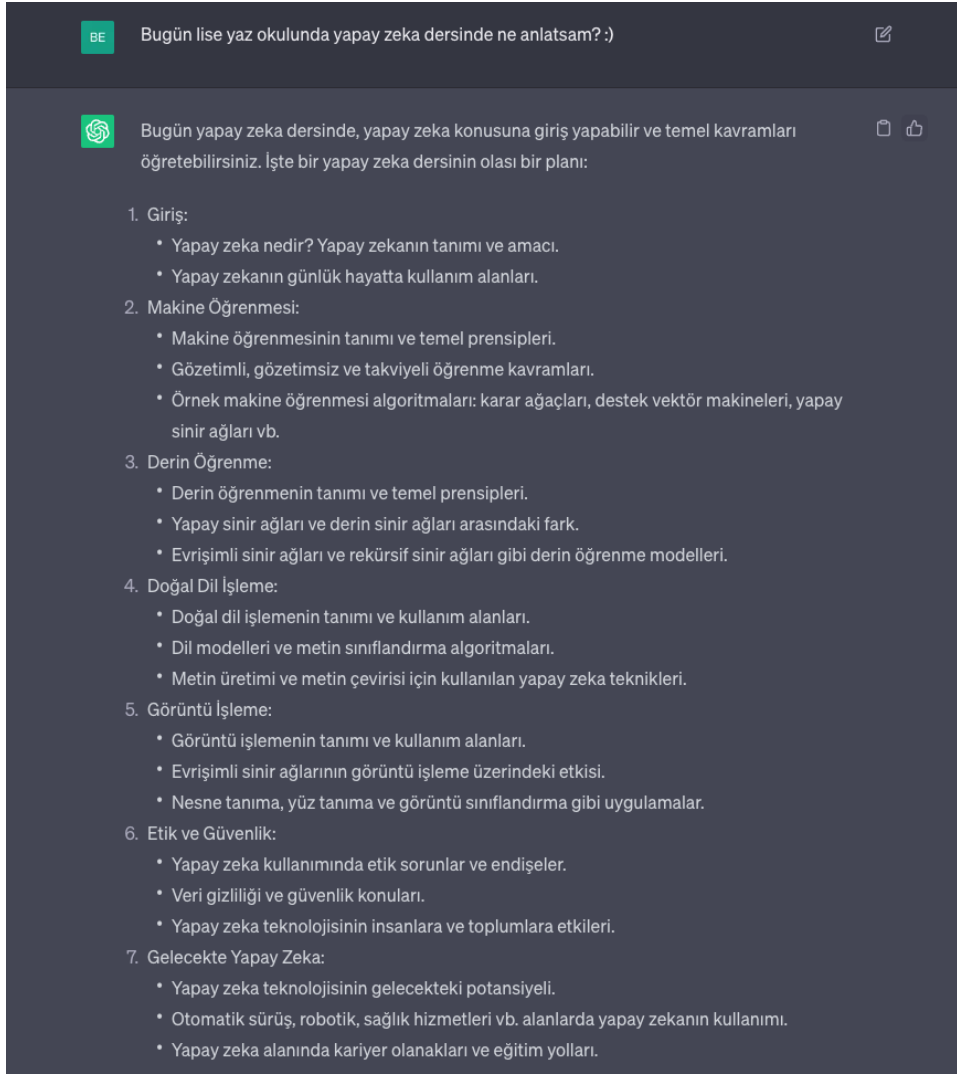
Detail from an untitled AARON drawing, ca. 1980.





# Refik Anadol





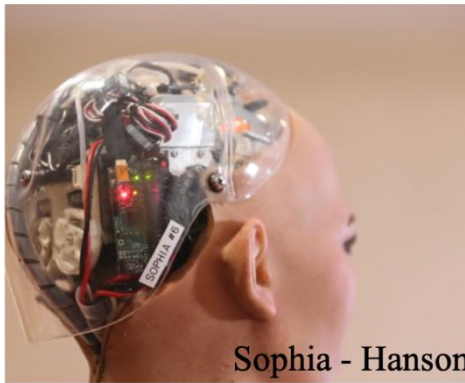
# ChatGPT

and other large language  
models

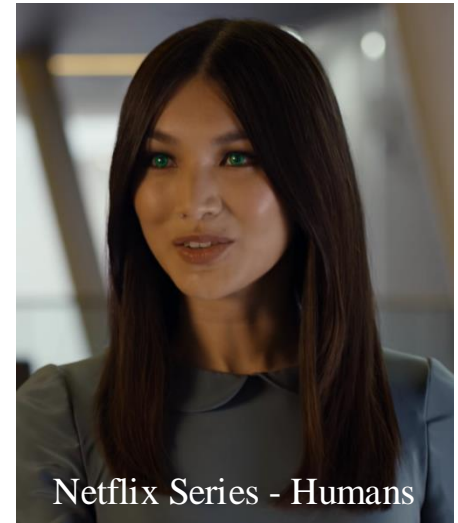
Question answering in natural  
language

# Artificial General Intelligence (AGI)

- Ultimate goal; human-like intelligence.



Sophia - Hanson Robotics 2018



Netflix Series - Humans

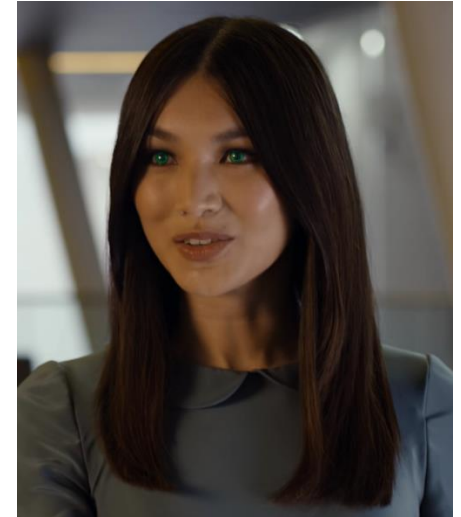
# AGI- When?

Near future is exciting and will bring lots of change & we need to be careful:

- AI learns biases in humans
- AI systems are not explainable
- AI & security

We need to start working on anticipated issues and problems:

- Job loss >> basic living wage
- Living with humanoid robots
- Responsibility in AI systems (e.g.. Autonomous cars)



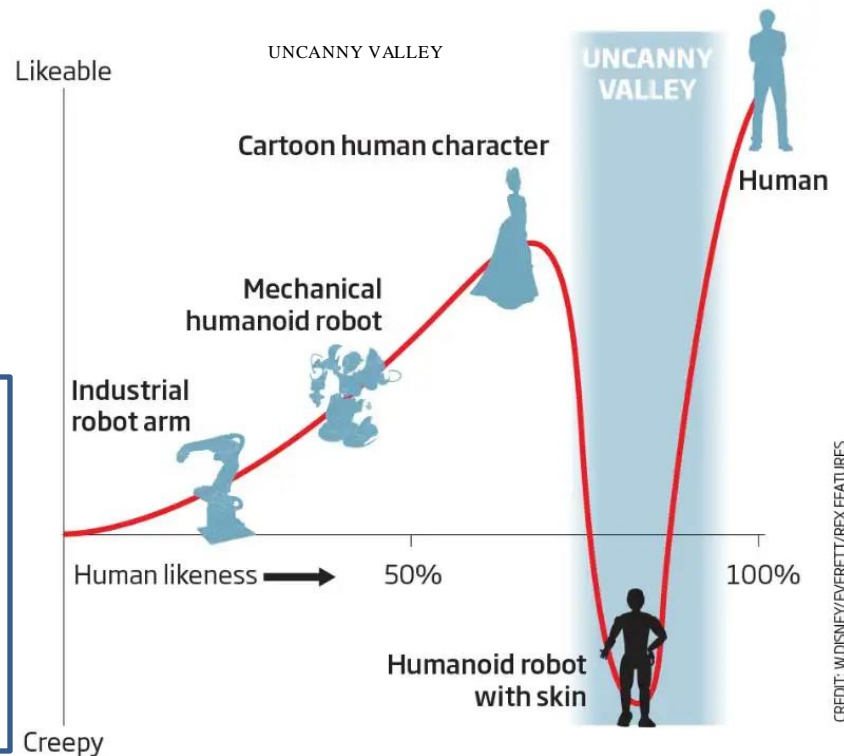
Netflix Series - Humans

# Uncanny Valley

As robots get **similar to humans**, humans first like them more until the resemblance gets uncanny (unbelievable/too much).

## Humans:

Based on the Swedish science fiction drama [Real Humans](#), the series explores **artificial intelligence** and **robotics**, focusing on the social, cultural, and psychological impact of the invention of anthropomorphic robots called "**synths**".



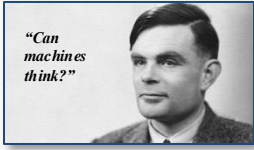
Likeability versus human-likeness



# Brief History of AI

## Erken Yapay Zeka

1950 – Turing Testi



Alan Turing

1956 – Dartmouth Konferansı



Marvin Minsky, Claude Shannon  
Ray Solomonoff ve diğerleri...

## Makine Öğrenmesi

1997 – Deep Blue



## Derin Öğrenme

2020-Dall-E

2016 – AlphaGo



2011 - IBM



1950 – Alan Turing, “Makineler Düşünebilir mi?”

1956 – John McCarthy Yapay Zeka kavramını ilk defa kullanır

1959 – MIT Yapay Zeka Projesi

1982 – İlk Uzman Sistem DENDRAL

1986 – İlk Otonom Araba

1997 – Deep Blue Kasparov’u Yener

2011 – Watson Jeopardy! Şampiyonunu Yener

2016 – AlphaGo Şampiyonu Yener

2017 – Libratus En İyi Poker Oyuncularını Yener

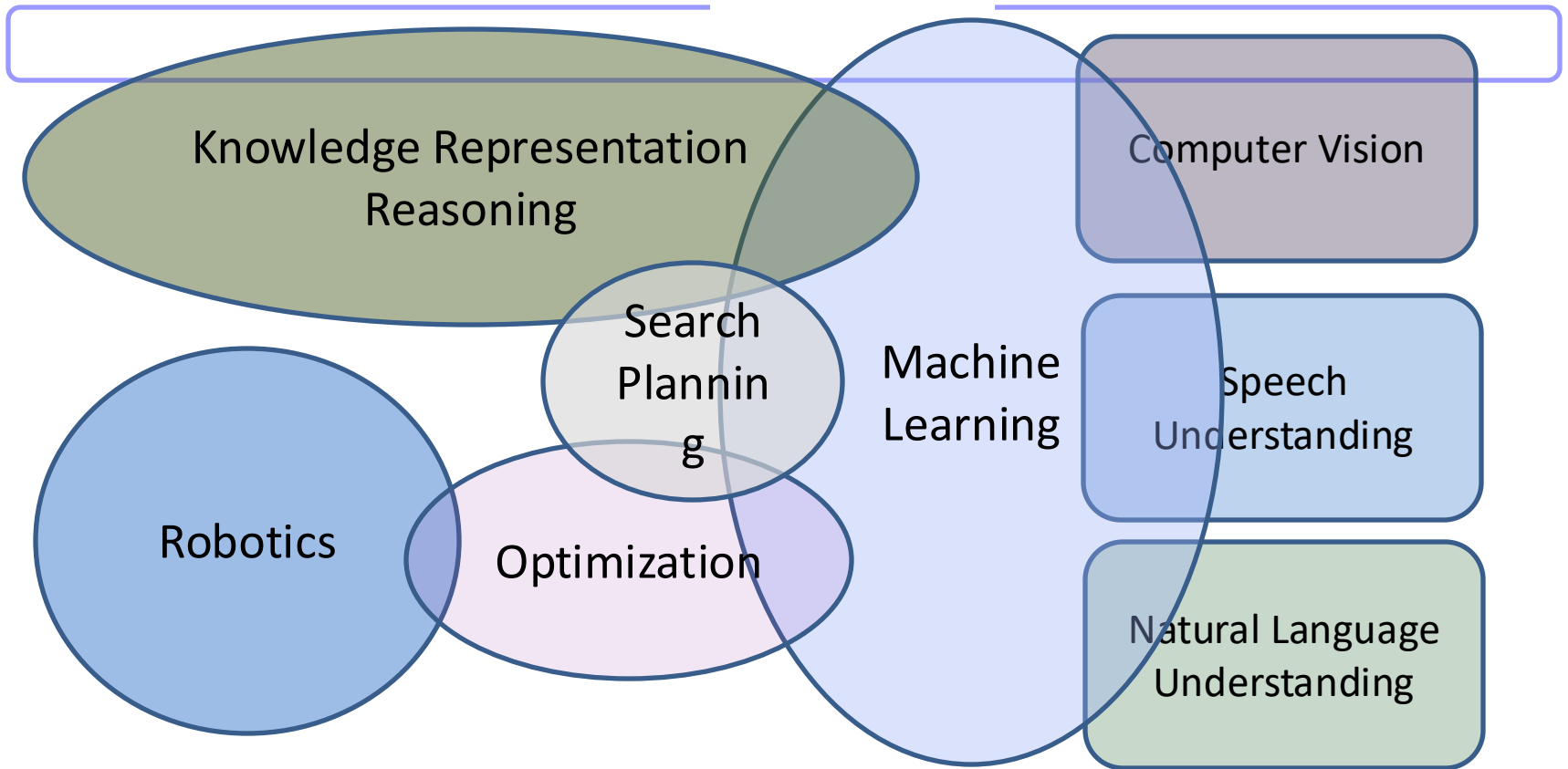
2017 – Deri Kanseri Sınıflandırıcısı

2021 – Protein katlanması problemi



# AI Components

# AI Components



# Machine Learning vs AI

## ARTIFICIAL INTELLIGENCE

IS NOT NEW

### ARTIFICIAL INTELLIGENCE

Any technique which enables computers to mimic human behavior



### MACHINE LEARNING

AI techniques that give computers the ability to learn without being explicitly programmed to do so



### DEEP LEARNING

A subset of ML which make the computation of multi-layer neural networks feasible



1950's

1960's

1970's

1980's

1990's

2000's

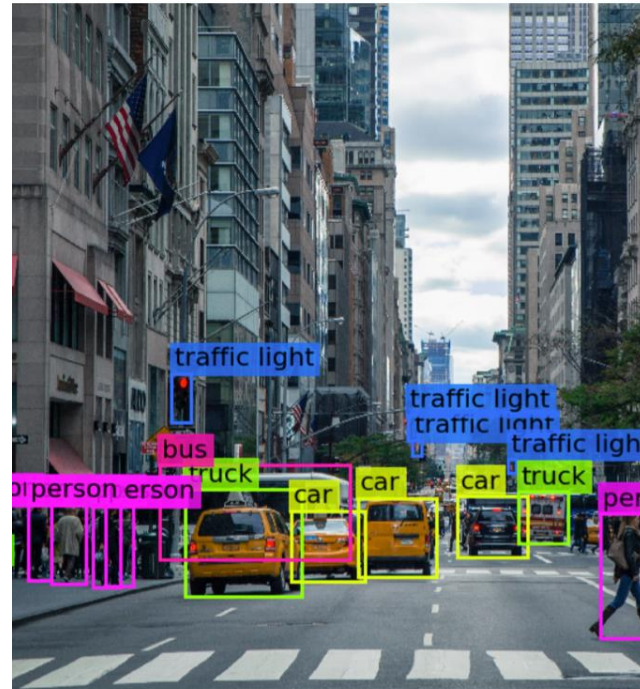
2010s

# Perceiving the World:

## Computer Vision

An intelligent computer must be able to recognize its surrounding environment and adapt to changes in it. To do this it must be able to “see” and “hear” what’s going on.

**Computer vision** is the capability of a computer to mimic the ways that human brains process and interpret light waves to produce a model of reality. Though it’s very easy for people to do that, it’s very difficult for computers to do build and update their models



# Perceiving the World:

## Speech and Natural Language Understanding

- **Speech understanding:** Converting a given speech signal into transcribed text.
- **Natural language understanding:** Understanding what is meant in a given transcribed text.



# Machine Translation

Türkçe▼



Bugünkü konuşmada neler  
anlatacağımı düşünürken, arkadaşım  
makine çevirisindeki başarı oranını  
sordu.

İngilizce▼



When I thought about what I would talk  
about in today's talk, my friend asked  
me about the success rate of the  
machine translation.

Bugünkü konuşmada neler  
anlatacağımı düşünürken arkadaşım  
otomatik çeviri de başarı nasıl diye  
sordu

When I thought about what I would talk  
about in today's talk, my friend asked  
me how to succeed in automatic  
conversation



BE

Bugün lise yaz okulunda yapay zeka dersinde ne anlatsam? :)



Bugün yapay zeka dersinde, yapay zeka konusuna giriş yapabilir ve temel kavramları öğretebilirsiniz. İşte bir yapay zeka dersinin olası bir planı:



#### 1. Giriş:

- Yapay zeka nedir? Yapay zekanın tanımı ve amacı.
- Yapay zekanın günlük hayatta kullanım alanları.

#### 2. Makine Öğrenmesi:

- Makine öğrenmesinin tanımı ve temel prensipleri.
- Gözetimli, gözetimsiz ve takviyeli öğrenme kavramları.
- Örnek makine öğrenmesi algoritmaları: karar ağaçları, destek vektör makineleri, yapay sinir ağları vb.

#### 3. Derin Öğrenme:

- Derin öğrenmenin tanımı ve temel prensipleri.
- Yapay sinir ağları ve derin sinir ağları arasındaki fark.
- Evrişimli sinir ağları ve rekürsif sinir ağları gibi derin öğrenme modelleri.

#### 4. Doğal Dil İşleme:

- Doğal dil işlemenin tanımı ve kullanım alanları.
- Dil modelleri ve metin sınıflandırma algoritmaları.
- Metin üretimi ve metin çevirisi için kullanılan yapay zeka teknikleri.

#### 5. Görüntü İşleme:

- Görüntü işlemenin tanımı ve kullanım alanları.
- Evrişimli sinir ağlarının görüntü işleme üzerindeki etkisi.
- Nesne tanıma, yüz tanıma ve görüntü sınıflandırma gibi uygulamalar.

#### 6. Etik ve Güvenlik:

- Yapay zeka kullanımında etik sorunlar ve endişeler.
- Veri gizliliği ve güvenlik konuları.
- Yapay zeka teknolojisinin insanlara ve toplumlara etkileri.

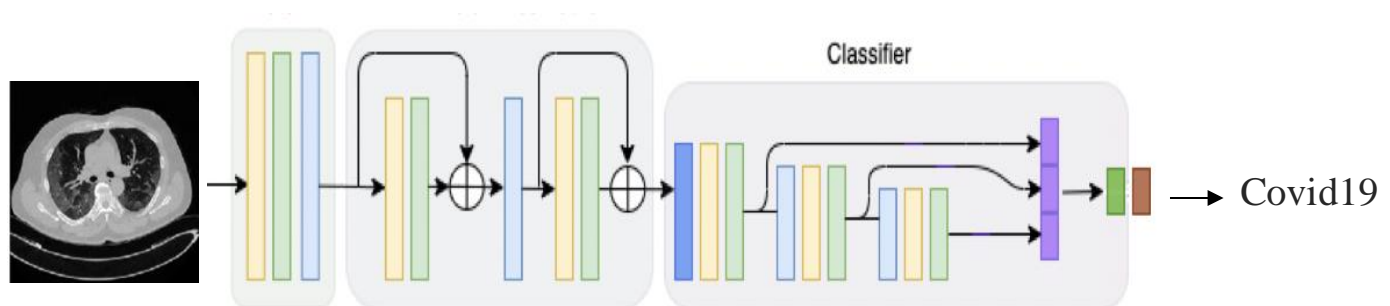
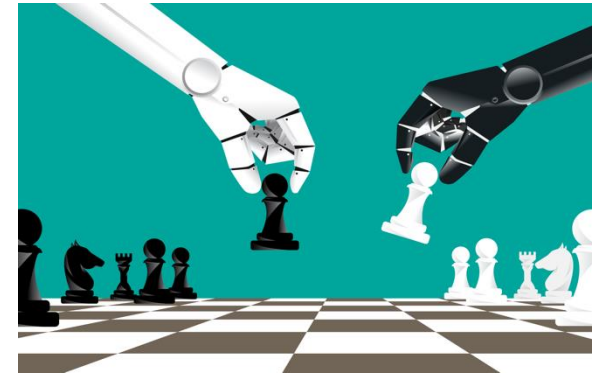
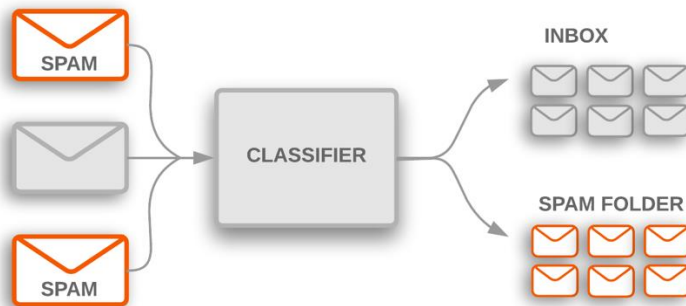
#### 7. Gelecekte Yapay Zeka:

- Yapay zeka teknolojisinin gelecekteki potansiyeli.
- Otomatik sürüş, robotik, sağlık hizmetleri vb. alanlarda yapay zekanın kullanımı.
- Yapay zeka alanında kariyer olanakları ve eğitim yolları.

# ChatGPT

# Machine Learning

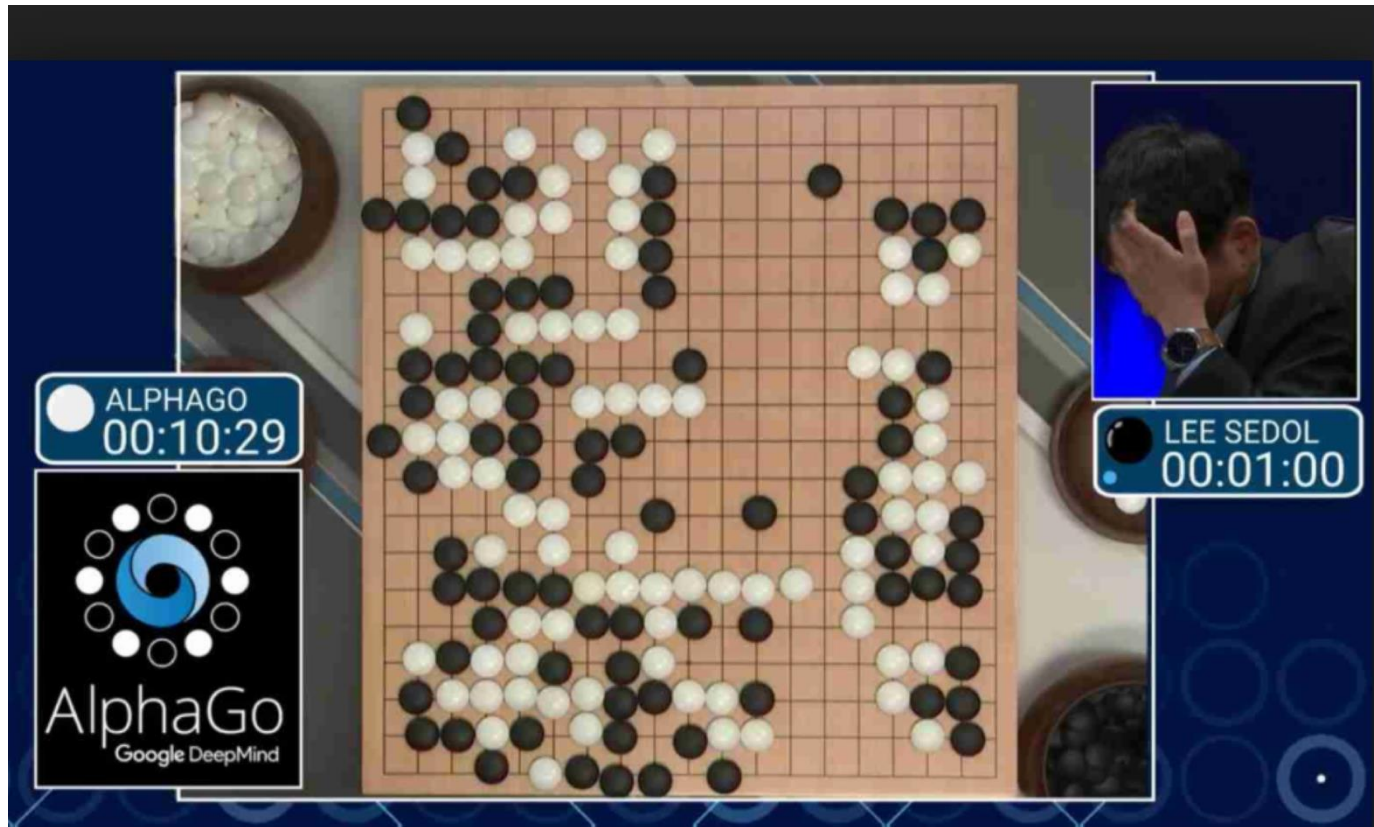
- Improving performance based on past **examples** or **experience**



# Machine Learning: Reinforcement Learning

## Game Playing Formulation:

- Given game board/situation, **choose the best move.**
- Given game board/situation, **assign scores to the reachable states.**



# Search Algorithms

Finding solutions to puzzles/problems by considering alternatives in a systematic way

- Time and space complexity
- Foundation of Game Playing
  - Adversarial search

....

Player 1

Player 2

Player 1

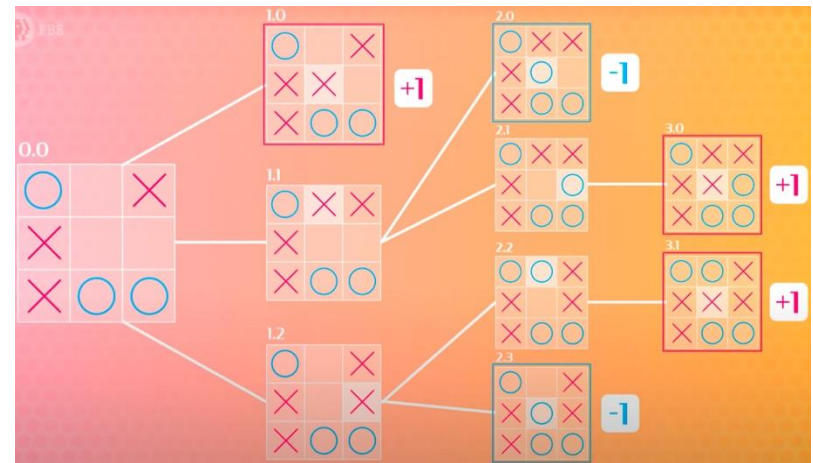
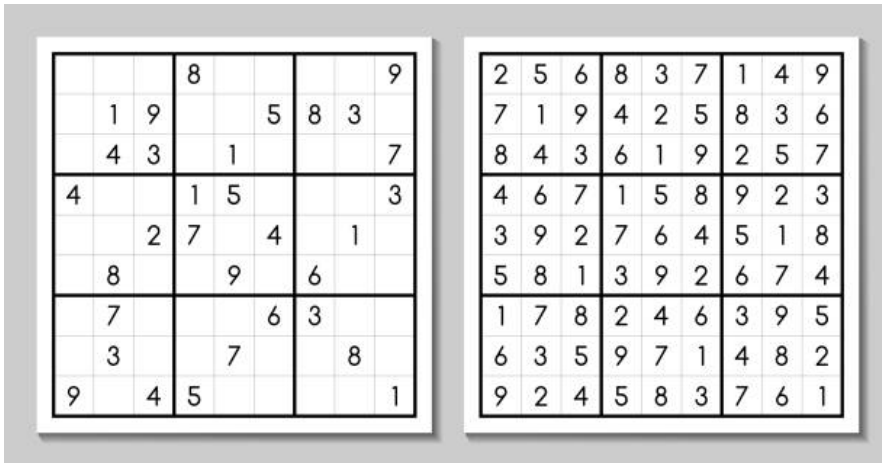


Image: Morioh.com

# Knowledge Representation and Logic

**"You will have fever and body ache in Flu"**

**$\text{Flu} \Rightarrow \text{Fever} \wedge \text{BodyAche}$**

**"Patient has fever, but no body ache"**

**$\text{Fever} \wedge \neg \text{BodyAche}$**

**$\Rightarrow \text{Not Flu}$**

# Robotics



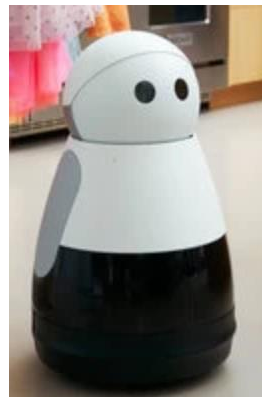
Honda Asimo



Harvard U.'s Insect Bot



SwagBot - Sydney University



Kuri



Samsung Bot Care



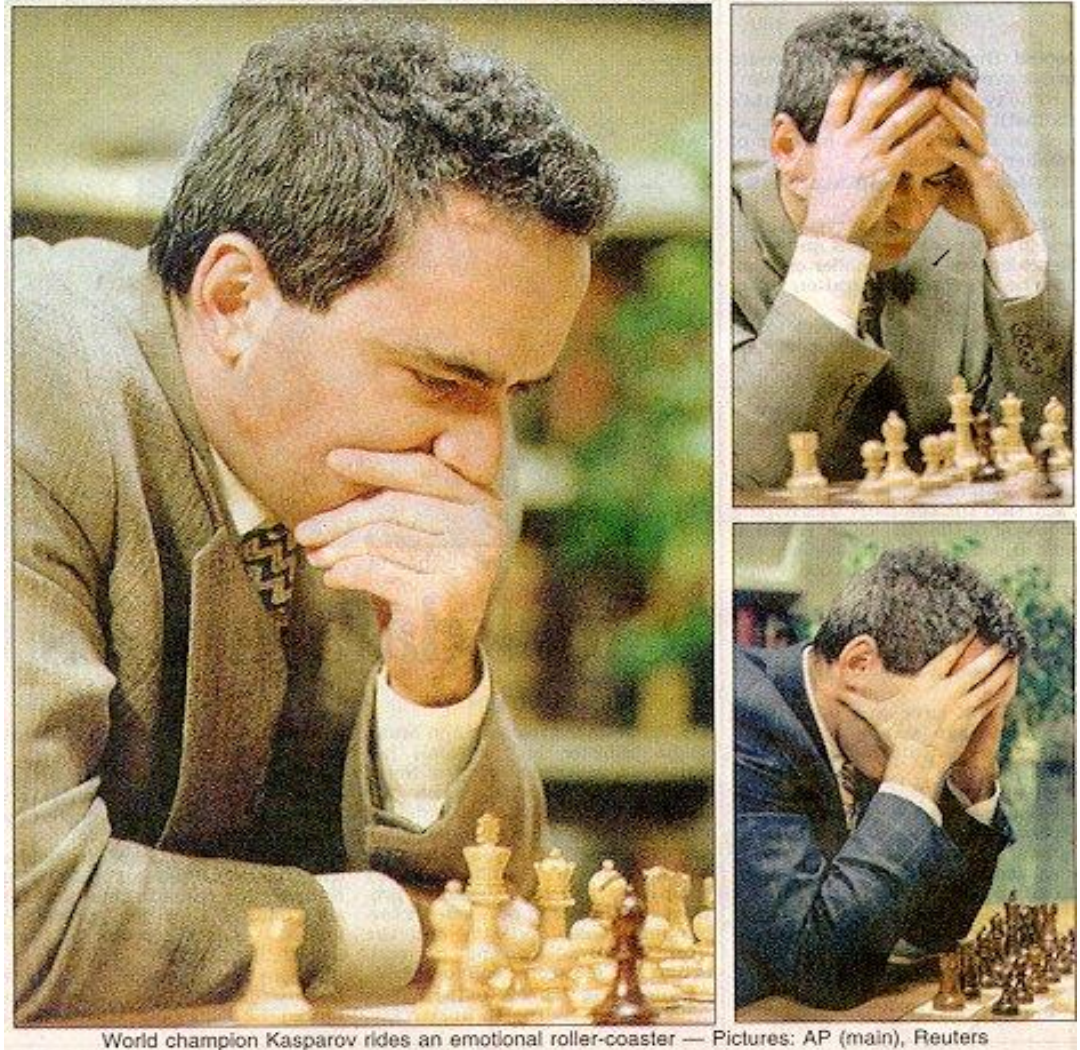
Big Dog – Boston Dynamics



# **AIMA – CHAPTER 1**

On May 12th, 1997, the best chess player in the world, Gary Kasparov, lost a six-game chess match to a computer named "Deep Blue 2"

What was so significant about this event?



World champion Kasparov rides an emotional roller-coaster — Pictures: AP (main), Reuters

Being able to program a computer to defeat a Grand Master level chess player had been a long-standing goal of the science of artificial intelligence — and now it has been achieved

# What is Artificial Intelligence?

In general, artificial intelligence is the field of science devoted to making computers perceive, reason, and act in ways that, until now, have been reserved for human beings.

- The field of Artificial Intelligence goes further and attempts not just to understand but also to build intelligent systems
  - Started out in 1950s
  - The Dartmouth meeting in 1956
- Turned out much more difficult than anyone had imagined
- Currently encompasses a large variety of subfields:
  - from general areas such as perception and logical reasoning to
  - specific tasks such as playing chess, writing poetry...
  - bringing together philosophy, logic, computer science, cognitive science and cognitive neuroscience

# What is Intelligence?

Intelligence is difficult to define and understand, even for philosophers and psychologists who spend their lives studying it. But this elusive quality is, to many people, the characteristic that sets humans apart from other species

**“What is intelligence, anyway? It is only a word that people use to name those unknown processes with which our brains solve problems we call hard. But whenever you learn a skill yourself, you are less impressed or mystified when other people do the same.**

**This is why the meaning of “intelligence” seems so elusive: It describes not some definite thing but only the momentary horizon of our ignorance about how minds might work.”**

**- Marvin Minsky, AI researcher**



# What is Artificial Intelligence?

- “The art of creating machines that perform functions that require intelligence when performed by people” (Kurzweil, 1990).
- “The branch of computer science that is concerned with the automation of intelligent behavior.” (Luger and Stubblefield, 1993)
- There have been many definitions over the history, roughly along 4 lines:

<b>Thinking Humanly</b> “The exciting new effort to make computers think . . . <i>machines with minds</i> , in the full and literal sense.” (Haugeland, 1985) “[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . .” (Bellman, 1978)	<b>Thinking Rationally</b> “The study of mental faculties through the use of computational models.” (Charniak and McDermott, 1985) “The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)
<b>Acting Humanly</b> “The art of creating machines that perform functions that require intelligence when performed by people.” (Kurzweil, 1990) “The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)	<b>Acting Rationally</b> “Computational Intelligence is the study of the design of intelligent agents.” (Poole <i>et al.</i> , 1998) “AI . . . is concerned with intelligent behavior in artifacts.” (Nilsson, 1998)

# Thinking Humanly

- In order to say that a given program thinks like a human, we must have some way of determining **how humans think**
  - Requires scientific theories of internal activities of the brain
    - What level of abstraction? ``Knowledge'' or ``circuits''?
  - How does a slow, tiny brain (biological or electrical) perceives, understands, and manipulates a complex world?
- How to validate?
  - Bring together computational models from AI and experimental techniques from psycho-physics to model the human mind
    - 1) Predicting and testing behavior of human subjects (Cognitive Science; top-down)
    - 2) Direct identification from neurological data (Cognitive Neuroscience; bottom-up)
      - Both approaches are now distinct from AI
  - Most of the machinery of the human mind is not available to our conscious experience

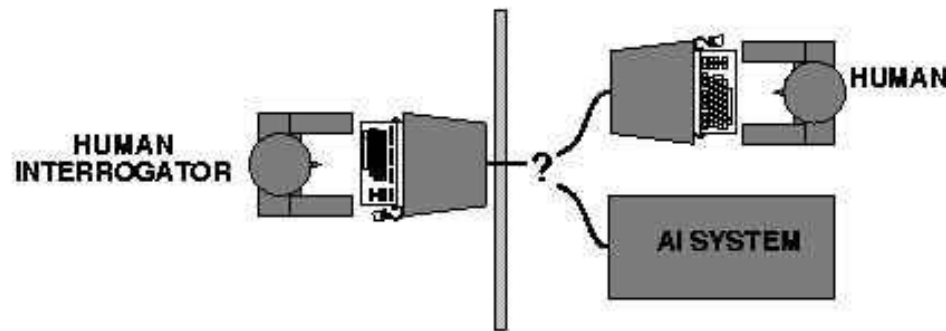
# Thinking Logically

- Let's give up imitating the human thinking process, but try to achieve **logical thinking**
- Laws of formal logic to formalize the thinking process
  - Socrates is a man; All men are mortal => Socrates is mortal
- Difficulties:
  - How to take informal knowledge and state in formal terms especially when knowledge is less than 100% certain
  - Even problems with just a few dozen facts can be prohibitive

# Acting Humanly - Turing Test

Turing (1950) "Computing machinery and intelligence":

- ◇ "Can machines think?" → "Can machines behave intelligently?"
- ◇ Operational test for intelligent behavior: the Imitation Game



- ◇ Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes
- ◇ Anticipated all major arguments against AI in following 50 years
- ◇ Suggested major components of AI: knowledge, reasoning, language understanding, learning

Problem: Turing test is not reproducible, constructive, or amenable to mathematical analysis

- The computer passes the test **if a human interrogator, after posing some written questions, cannot tell whether the written responses come from a person or not**
  - Suggested **major components of AI**: natural language processing, knowledge representation, automated reasoning, machine learning
- **Total Turing test** also requires computer vision and robotics



# Acting Rationally

- **Rationality:** ideal concept of intelligence - doing the right thing
- **The right thing?:** what is expected to maximize goal achievement, given the available information
  - Reflexes do not necessarily involve thinking---e.g., blinking reflex, but they are rational
  - Thinking should be in the service of rational action
- **Caveat:** computational limitations make perfect rationality unachievable
  - ⇒ design best program for given machine resources
- We will emphasize rational agents in this course.

# Why do we want artificial intelligence?

- To **relieve our mental labour**, just as machines relieved our physical labour last century
- It should **make the machines themselves easier to use**
- It might give some insight into the workings of our own minds

# History of AI

- **1943**     **McCulloch and Pitts: Artificial Neuron Model**
- **1950**     **Turing's ``Computing Machinery and Intelligence'‘**
- **1950s**     **Early AI programs**, including Samuel's checkers program, Newell & Simon's Logic Theorist (proving theorems), Gelernter's Geometry Engine, Shannon and Turing writing chess programs
  - **Shortage of computer times => Development of time sharing (=> DEC)**
  - **Creation of LISP (McCarthy)**
- **1956**     **Dartmouth meeting: ``Artificial Intelligence" coined**
- **1965**     **Robinson's complete algorithm for logical reasoning resolution method**
- **1960s**     **Early development of knowledge-based systems; Minsky's microworlds** (blocks as home to various projects: vision, planning, nat. lang. understanding, ...)
  - **ANALOGY** program (what is this figure most similar to?)
  - **Algebra STUDENT** program (one egg costs ... How much does twenty eggs cost?)

# History of AI

- **1966--74 Dose of Reality**
  - Very little domain knowledge:
    - Switching from one domain to another, the programs failed miserably
  - AI discovers computational complexity
    - Early programs worked by representing the basic facts and trying out a series of steps to solve the problem which was only tractable within micro worlds; NP-completeness showed that scaling up to larger problems was not always viable
  - Neural network research almost disappears

# History of AI

- **1980--88 Expert systems industry booms**
  - After all, they work, even if in limited domains
  - An **expert system** is a software designed to replicate the **decision-making process of a human expert, within a narrow topic**. At the heart of every expert system is a *knowledge base* representing ideas from the specific field of expertise
  - A **knowledge-based system** derives knowledge from experts as well as other sources like government regulations, statistical databases, company guidelines, etc.
  - In practice, **the terms *expert system* and *knowledge-based system* are often used interchangeably**
    - While a database contains only facts, a knowledge base also contains a system of if-then *rules* for determining and changing the relationships between those facts



# History of AI

- **1988--93 Expert systems industry start losing its power**
  - Successful only in very narrow domains
  - Building a successful expert system is much more than simply buying a reasoning system and filling it with rules
- **1985--95 Neural networks return to popularity**
- **1988-- With strengthened foundations, AI becomes hot again**
  - Resurgence of probabilistic and decision-theoretic methods, genetic algorithms, belief networks,...
- **2000 -... Deep learning, significant advances in robotics**
  - Now we are talking about the perils of AI

# What to Know

- Basic/Brief history of AI
  - Know some of the important events or at least what happened in different eras
- Difficulty of defining intelligence
  - Fleeting nature of the definition
  - Difference of humanly/rational thinking/acting
- AI current state and goal
  - Artificial General Intelligence: general vs narrow AI
- Important AI concepts
  - Turing test
  - Uncanny valley

# Reading

- Read AIMA Chapter 1 (but slides are sufficient)