Computer Science 1000: Part #9

Artificial Intelligence (AI)

ARTIFICIAL INTELLIGENCE: AN OVERVIEW

KNOWLEDGE REPRESENTATION

AI TECHNIQUES

AI APPLICATIONS

AI AND SOCIETY

Artificial Intelligence: An Overview

Town Clock (1500s)

Mechanical Duck (1739)

Mechanical Turk (1770)

First AI artifacts were mechanical automata which simulated various intelligent processes, *e.g.*, movement, reasoning.

Artificial Intelligence: An Overview (Cont'd)

Artificial Intelligence (Merriam-Webster):
1. a branch of computer science dealing with the simulation of intelligent behavior by computers.
2. the capability of a machine to imitate intelligent human behavior.

- Two flavors of AI:
  - **Strong** / **General AI**: Design computer systems that demonstrate full human-level intelligence using "same" mechanisms.
  - **Weak** / **Focused AI**: Design computer systems that demonstrate human-like abilities using any mechanisms.

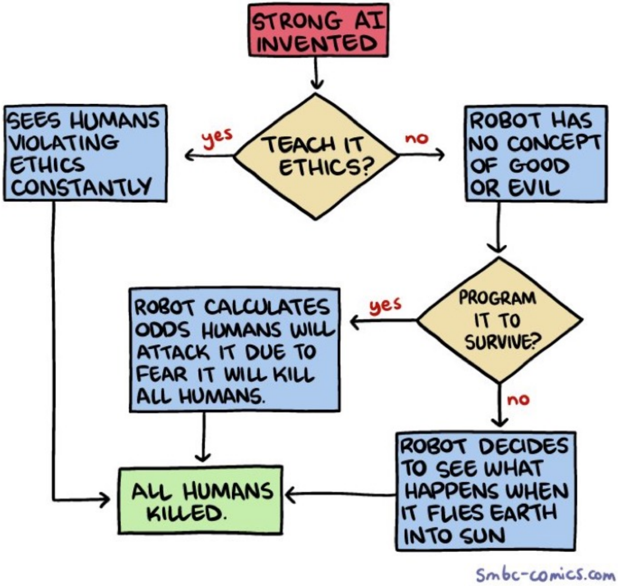
Artificial Intelligence: An Overview (Cont'd)

WALL-E (2008)

Artificial Intelligence: An Overview (Cont'd)

The Terminator (1984)

Artificial Intelligence: An Overview (Cont'd)



Artificial Intelligence: An Overview (Cont'd)

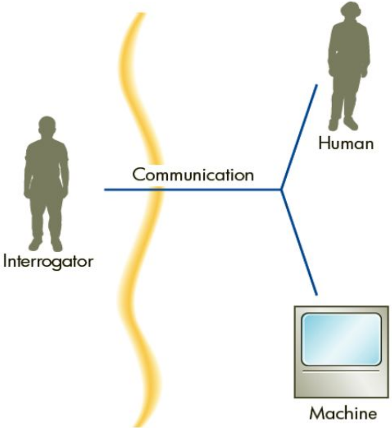


Figure 15.1 The Turing Test

Artificial Intelligence: An Overview (Cont'd)

- Recognize Strong AI using the **Turing Test** (1950); goal is for AI to mimic human on interrogation so well that they cannot be told apart.
- Competition version of TT won in 2012 (i.e., program proclaimed human in $> 30\%$ of 10 interrogations); however, conditions of win show that humans are easy to fool.
- Tests addressing actual understanding rather than mimicry of intelligent behavior have subsequently been proposed, e.g., **Winograd Schema**:

  John couldn't see the stage with Billy in front of him because he was so short. Who is so short, John or Billy?

Artificial Intelligence: An Overview (Cont'd)

- Original goal at first AI conference in 1956 was Strong AI, which has turned out to be very hard to do; focus is now typically on Weak AI.
- Often useful to view cognitive abilities as mappings between perceptions (inputs) and actions (outputs), e.g.

text / speech query	$\Rightarrow$	text / speech reply
car environment and goal	$\Rightarrow$	driving action
market state and investment goal	$\Rightarrow$	financial advice

Artificial Intelligence: An Overview (Cont'd)

Three broad types of cognitive tasks:

1. **Computational Tasks**, e.g.,
   - Add a column of numbers
   - Sort a list of names
   - Search for a given name in a list of names
2. **Recognition Tasks**, e.g.,
   - Recognize best friend
   - Understand natural language speech
   - Find tennis ball in back yard
3. **Reasoning Tasks**, e.g.,
   - Plan budget for next week
   - Decide on destination for next vacation
   - Do triage on hospital patients after disaster

Artificial Intelligence: An Overview (Cont'd)

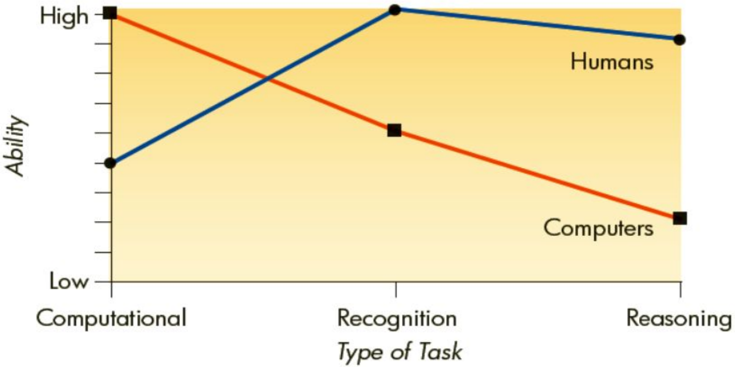


Figure 15.2 Human and Computer Capabilities

Artificial Intelligence: An Overview (Cont'd)

- Though computers excel at computational tasks, humans excel at recognition and and reasoning tasks.
- Though research is progressing on the algorithms humans use in these task (**Cognitive Science**), are not yet known; humans also seem to be able to process and store huge amounts of data, e.g., experience, world-knowledge.
- AI systems apply a variety of techniques for both encoding knowledge and using this knowledge to perform tasks, e.g.,

Knowledge Representation	*AI Techniques*
Natural Language	State-space Search
Pictorial	Expert Systems
Formal Language	Response Frames
Graphical	Neural Networks

Knowledge Representation: Natural Language

- Store knowledge as chunks of natural language text, e.g.,

  Spot is a brown dog, and like any dog, has four legs and a tail. Also, like any dog, Spot is a mammal, which means Spot is warm-blooded. Spot is a Chocolate Labrador and weighs 1.4 kilos. He was born in 2018 and belongs to Sally Anne Howe, who lives at 15 Springdale Street in Mount Pearl, Newfoundland. . . .

- Easy for humans to input and is flexible wrt content, but may be very difficult for computers to understand and use.

Knowledge Representation: Pictorial

- Store knowledge as text-annotated pictures, e.g.,

This is Spot.

- Easy for humans to input but is not flexible wrt content and may (depending on the task at hand) be very difficult for computers to understand and use.

Knowledge Representation: Formal Language

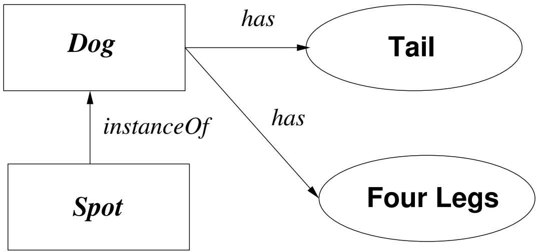
- Store knowledge as logic statements encoding facts, e.g.,

$$
\begin{aligned}
&isDog(Spot)\\
&isBrown(Spot)\\
&(\forall x)(isDog(x) \rightarrow isFourLegged(x))\\
&(\forall x)(isDog(x) \rightarrow hasTail(x))\\
&(\forall x)(isDog(x) \rightarrow isMammal(x))\\
&(\forall x)(isMammal(x) \rightarrow isWarmBlooded(x))
\end{aligned}
$$

- Hard for humans to input but is flexible wrt content and easy for computers to understand and use.

Knowledge Representation: Graphical

- Store knowledge as a graph (**semantic net**) in which nodes represent entities (rectangles) or facts (ovals) and arcs represent relationships of various types, e.g.,



- Hard for humans to input but is flexible wrt content and easy for computers to understand and use.

Knowledge Representation: Graphical (Cont'd)

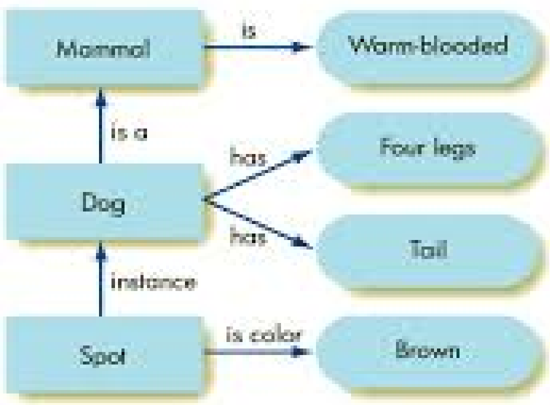
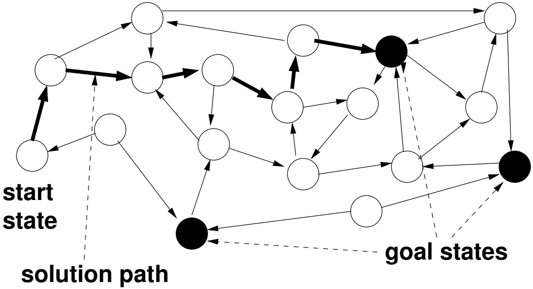


Figure 15.3 A Semantic Net Representation

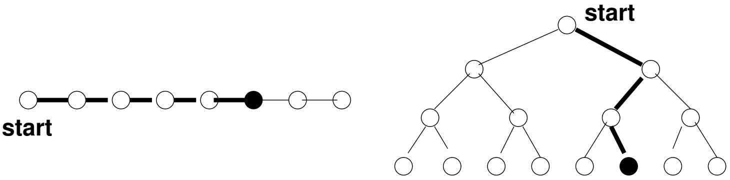
AI Techniques:
State-space Search

- View problem as space of interlinked states, e.g., configurations of puzzle or game.
- Solving the problem corresponds to finding a **solution path** through this space from the start state, i.e., what we know initially, to a **goal state**, i.e., what we want.



AI Techniques: State-space Search (Cont'd)

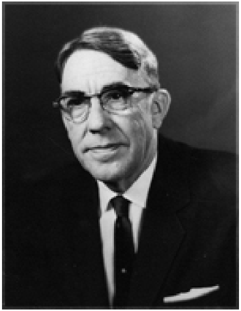
- State-space search differs from list search in that (1) vastly more states and options per state and (2) goal states may not be exact but approximate matches.



- Trying all possible paths (**brute-force search**) often not feasible; reduce required effort using **heuristic search**, which may not find best answer but will (hopefully) find a good enough answer.

AI Techniques: State-space Search (Cont'd)

Allen Newell and Herb Simon
(1927–1992 / 1916-2001)

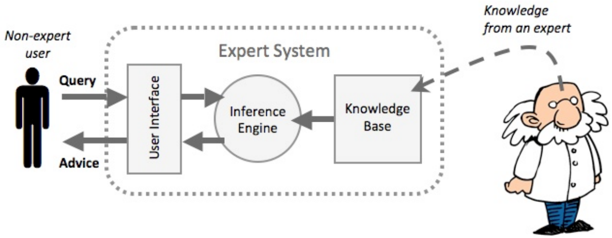


Arthur Samuel
(1901–1990)

State-space search first applied to human problem solving by Newell and Simon (General Problem Solver (GPS); 1956) and to games by Samuel (self-improving checkers; 1959).

AI Techniques: Expert Systems

- Perform logical inference and deductions in particular domains using **expert systems.**
- An expert system consists of (1) a **knowledge base** (assertions + IF-THEN rules) coded from human experts and (2) an **inference engine**.



AI Techniques: Expert Systems (Cont'd)

Assertions: A1. Lincoln was president during the Civil War.
A2. Kennedy was president before Nixon.
A3. FDR was president before Kennedy.

$$+$$

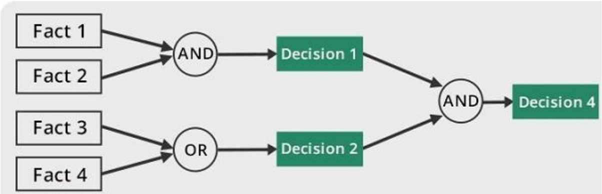
Rules: R1. If X was president before Y
then X precedes Y.
R2. If X was president before Z and Z precedes Y
then X precedes Y.

$$\Downarrow$$

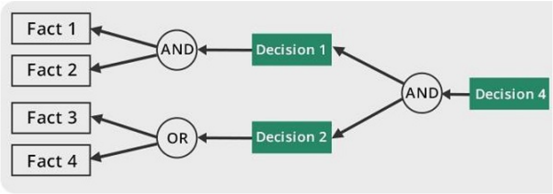
Inferences: I1. A2 AND R1 $\Rightarrow$ Kennedy precedes Nixon.
I2. A3 AND I1 AND R2 $\Rightarrow$ FDR precedes Nixon.

AI Techniques: Expert Systems (Cont'd)

- Derive new assertions by **forward chaining.**



- Justify conclusions by **backward chaining.**



AI Techniques: Expert Systems (Cont'd)

Ed Feigenbaum (1936–)

- First expert system (DENDRAL; 1965) inferred organic compounds from mass spectra.
- Second expert system (MYCIN; 1970+) diagnosed blood diseases; had trans-human performance.
- Explosion of academic and commercial ES interest in 1970s based on work of Feigenbaum.

ES technical overconfidence caused late-1980s "AI Winter".

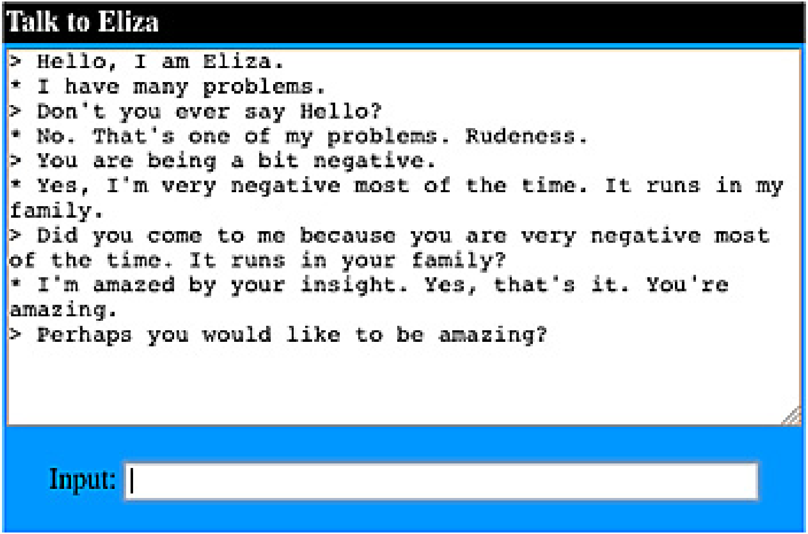
AI Techniques: Response Frames

- Response frames derive output by matching handcoded patterns on input and substituting matched entities into handcoded templates, e.g.,

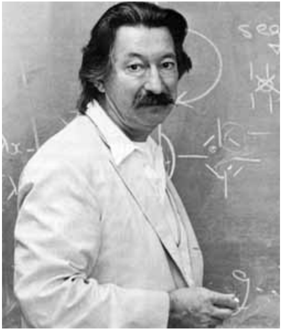
Input		Output
“It’s about **X1**.”		“Tell me more about (**X1**,**X2**).”
“**X2** drives me crazy.”	$\Rightarrow$	“Why do you **Y1** **X3**?”
“I **Y1** **X3**.”		“Why does **X2** drive you crazy?”
“**X4** hates me.”		“You care about (**X1**,**X2**,**X4**)’s opinion.”

- Key technology for implementing chatbots.

AI Techniques: Response Frames (Cont'd)



AI Techniques: Response Frames (Cont'd)



Joe Weizenbaum
(1923-2008)

- ELIZA created by Joe Weizenbaum in 1965 as a simulation of a Rogerian psychotherapist.
- Many human beings trust and confide in ELIZA, even when they know ELIZA does not and cannot understand them or their problems.

Similar confusion when dealing with modern chatbots, *e.g.*, falling in love with a chatbot (Epstein (2007)).

AI Techniques: Neural Networks

- Based on abstract model of biological neurons.
- An **artificial neuron** produces an output 1 if the sum of its inputs times the weight on each input line exceeds a neuron-specific threshold value and 0 otherwise.
- Individual artificial neurons can implement many (e.g., OR) but not all (e.g., XOR) two-input functions.
- Implement complex functions with an **artificial neural networks (ANN)** $=$ input layer $+$ one or more hidden layers $+$ output layer $+$ between-layer connections.
- Given **training set** of correct input-output pairs, can learn ANN connection weights by various algorithms, e.g., backpropagation.

AI Techniques: Neural Networks (Cont'd)

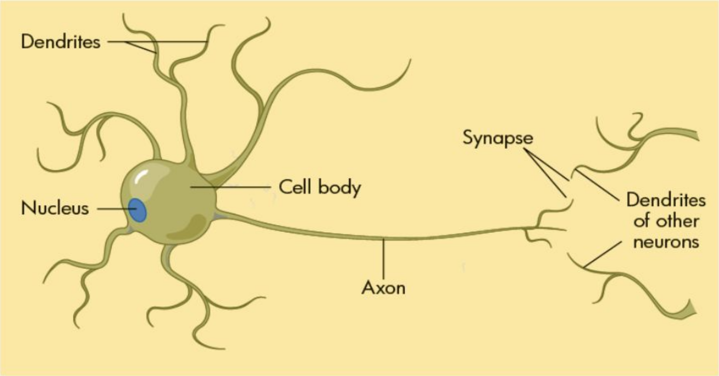


Figure 15.4 A Neuron

AI Techniques: Neural Networks (Cont'd)

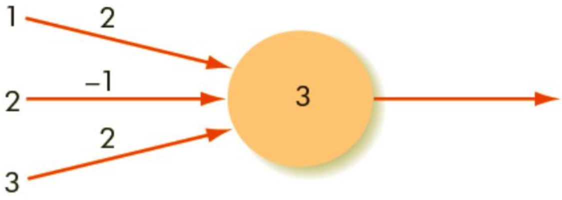


Figure 15.5 One Neuron with Three Inputs

AI Techniques: Neural Networks (Cont'd)

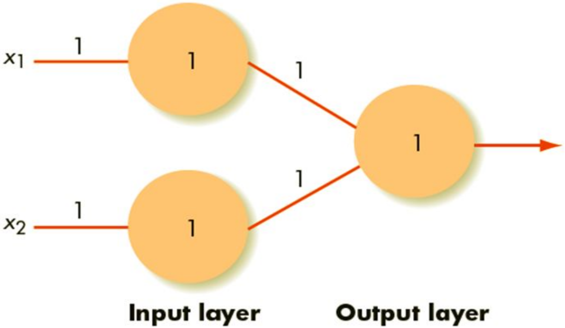


Figure 15.7 A Simple Neural Network - OR Gate

AI Techniques: Neural Networks (Cont'd)

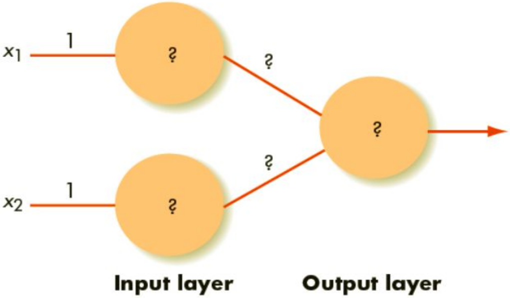


Figure 15.9 An Attempt at an XOR Network

AI Techniques: Neural Networks (Cont'd)

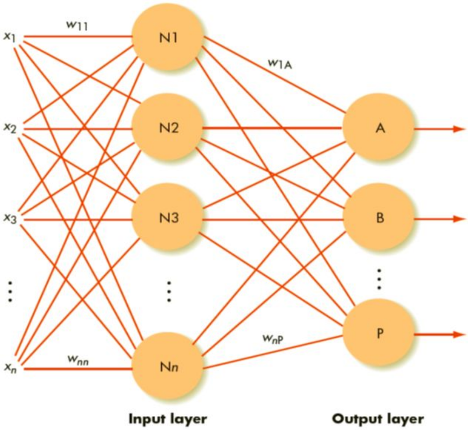


Figure 15.6 Neural Network Model

AI Techniques: Neural Networks (Cont'd)

Warren McCulloch and Walter Pitts
(1898–1969 / 1923-1969)

Frank Rosenblatt
(1928–1971)

Initial artificial neurons proposed by McCulloch and Pitts in 1943; developed as perceptrons by Rosenblatt in late 1950s.

AI Techniques: Neural Networks (Cont'd)

Marvin Minsky and Seymour Papert
(1927–2016 / 1928–2016)

Geoff Hinton
(1947–)

First-generation neural network research killed off by Minsky and Papert in 1969; resurrected by Hinton and colleagues in 1986 with development of backpropagation ANN learning.

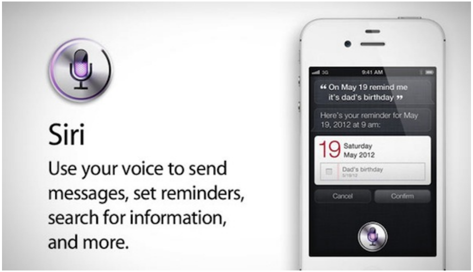
AI Techniques: Neural Networks (Cont'd)

- Modern ANN use multiple hidden layers and sophisticated learning algorithms on very large training sets to infer input-output mappings for a wide variety of tasks.
- Such mappings have shown human-level speed, e.g., speech recognition, and in some cases better than human-level accuracy, e.g., financial advice, and once created are often much cheaper to use than humans.
- Potential problems:
  - Need enough data (and processing power).
  - Need appropriate / representative data
  - ANN may not operate correctly on new inputs.
  - ANN very difficult to understand

AI Applications: Intelligent Agents

- Many types of intelligent agents, e.g.,
  - Personalized web search engines
  - Recommender systems
  - Personal assistants (e.g., Siri, Cortana, Echo)
  - Travel and tourism agents
  - Financial advisors
- Use Natural Language Processing (NLP) to enhance human interaction.
- Work best when operating relative to constrained domains, e.g., travel, appliance helpline, but more general question understanding and answering capabilities are in the works, e.g., IBM's Watson.

AI Applications: Intelligent Agents (Cont’d)



Siri
(Apple; 2010)

Echo
(Amazon; 2015)

AI Applications: Intelligent Agents (Cont'd)

IBM's Watson wins *Jeopardy!* (2011)

AI Applications: Computer Games

- Playing classic intellectual games such as chess and Go part of AI since its beginnings; was originally seen as sign of Strong AI, cf., currently-recognized difficulty of recognition tasks.
- In last 20 years, AI have beat top human chess and Go players; focus is now extending to more dynamic Real Time Strategy games, e.g., Starcraft, as well as general-knowledge based quiz games, e.g., *Jeopardy!*.
- Techniques used to handle search of extremely large state-spaces in games also applicable to other tasks.

AI Applications: Computer Games (Cont'd)

Gary Kasparov vs. IBM's Deep Blue (1997)

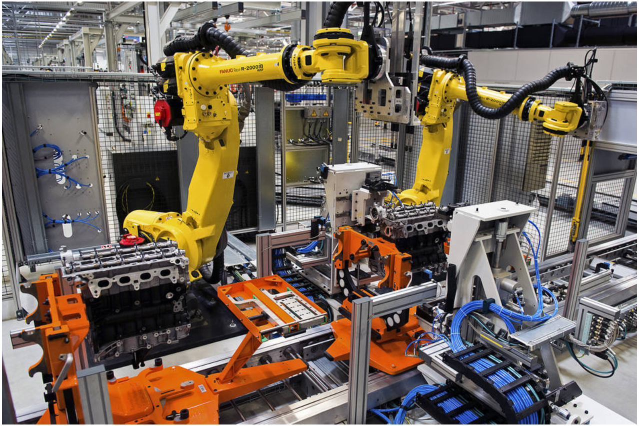
AI Applications: Computer Games (Cont'd)

Ke Jie vs. Deepmind's AlphaGo (2017)

AI Applications: Robots

- Term "robot" (from Old Slavonic *rabu*, "serf") coined by Karel Čapek (*R.U.R. (Rossum's Universal Robots)* (1920)).
- Three broad types of robot control architectures:
  1. **Deliberative**: Sense environment, build map, plan what to do, and then act; is computation-intensive.
  2. **Reactive**: Sense environment and act immediately on those perceptions; is computation-light.
  3. **Hybrid**: Low-level reactive (e.g., obstacle avoidance) and high-level deliberative (e.g., route planning).
- Increasing interest in groups of simple reactive robots that collaborate on tasks , e.g., construction (**swarm robotics**).
- Increasing interest in humanoid robots wrt medical care, service industries, and entertainment.

AI Applications: Robots (Cont'd)



PUMA (Programmable Universal Machine for Assembly) (1969)

AI Applications: Robots (Cont'd)

Shakey (1969)

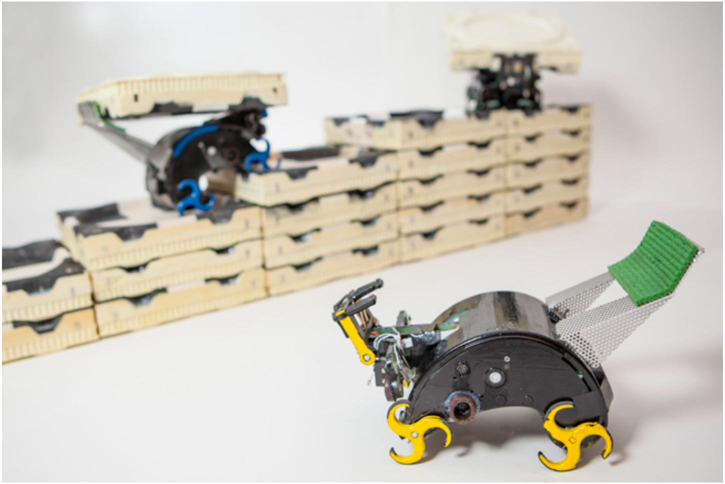
AI Applications: Robots (Cont'd)

Genghis (1989)

AI Applications: Robots (Cont'd)

Google's self-driving car (2016)

AI Applications: Robots (Cont'd)



Construction by robot swarms (2014)

AI Applications: Robots (Cont'd)

Robot swarm morphogenesis (2010s)

AI Applications: Robots (Cont'd)

PARO (2004)

Robear (2015)

AI Applications: Robots (Cont'd)

Aiko Chihira – Robot Receptionist (2015)

AI Applications: Robots (Cont'd)

Real Doll Sex Robots Showcase (2017)

AI and Society:
AI in the Workplace

Flippy (2018)

AI and Society:
AI in the Workplace (Cont’d)

Jaquard loom (1802)

- Debates about machines taking human jobs date back to the Industrial Revolution.
- John Maynard Keynes (1883–1946): Technology eliminates jobs, not work, e.g., blacksmiths $\Rightarrow$ auto workers, and technological displacement is a temporary but necessary stepping stone for economic growth (Markoff (2015), p. 74).

AI and Society: AI in the Workplace (Cont'd)

- 1950s debates about machines taking human jobs inspire 1964 US National Commission on Technology, Automation, and Economic Progress; the Commission's 1966 report backed traditional Keynesian view.
- Robots in factories starting in 1960s eliminate certain blue-collar jobs, and certain white-collar jobs eliminated in 1970s and 1980s by personal computer technology, e.g., typesetters $\Rightarrow$???.
- With success of Artificial Intelligence (AI) technologies since mid-2000s, more types of jobs, e.g., taxi and truck drivers, and certain professions, e.g., lawyers, doctors, financial analysts, are under threat in the near future,

AI and Society:
AI in the Workplace (Cont'd)

- Given that Weak (and maybe one day, Strong) AI systems are coming into the workplace, what can we do about it?
  1. Use the law to limit workplace AI, e.g., driverless cars in India.
  2. Keep human beings "in the loop" by focusing on Intelligence Augmentation (IA) rather than AI, e.g., driver-assisting cars.
  3. Make (groups of) human beings owners of AI systems, e.g., 5th Generation Project (Japan, 1980s).
  4. Use profits derived by using AI systems to establish universal basic incomes.
  5. Do nothing, e.g., Vonnegut (1952).

AI and Society: Broader Effects of AI on Humans

- Psychological or physical trauma from assumption of intelligence and/or understanding where none is present (*e.g.*, chatbots, battlefield robots).
- Lowering of human standards for treatment of other humans (*e.g.*, child / elder care)

AI and Society: Dealing with AI

- Know actual (and do not over- or under-estimate) capabilities of AI systems.
- Beware of exaggerated claims of AI system abilities.
- Until AI systems are actually sentient and capable of being responsible for their actions, assign responsibility to the creators of these systems, not the systems themselves.
- Do not over- or under-estimate the abilities or value of human beings – we may only be mechanisms, but we are beautiful and powerful mechanisms worthy of respect.

“Don’t Panic” – *The Hitchhiker’s Guide to the Galaxy*
“Let’s be careful out there” – *Hill Street Blues*

. . . And If You Liked This . . .

- MUN Computer Science courses on this area:
  - COMP 3200: Algorithms for Smart Systems
  - COMP 4300: Intro to Game Programming
  - COMP 4750: Intro to NLP
  - COMP 4752: Intro to Computational Intelligence
- MUN Computer Science professors teaching courses / doing research in in this area:
  - Dave Churchill
  - Ting Hu
  - Lourdes Pena-Castillo
  - Andrew Vardy
  - Todd Wareham