CAP 6629/EEL 6935: Reinforcement Learning
Spring 2025

Prof. Xiangnan Zhong
Email: xzhong@fau.edu

Class time: Wed/Fri 11:00 am - 12:20 pm (in person and distance learning)
Office hours: Wed 2:00-3:00 pm (in person EE431, Zoom meetings per request)

1

1

TA: Andy Cheng
Email: wcheng3@fau.edu
Office Hour: Tuesday, 4-5PM
Office: EE410b

TA: Yiran Pang
Email: ypang2022@fau.edu
Office Hour: Friday, 4-5PM
Office: EE410b

2

2

Textbook:
Andrew Barto and Richard S. Sutton, Reinforcement Learning: An Introduction, MIT Press, 1998 (First Edition)/2017 (Second Edition)
Lecture notes mainly adopted from Prof. Sutton

Supplementary reading:
T. Mitchell, Machine Learning, McGraw Hill, 1997. ISBN 0070428077

Acknowledgement:
Other teaching materials from Dr. S. Sutton (U of Alberta), Dr. T. Mitchell (CMU) and Dr. Levine (UC-Berkeley) are used in this course.

3

3

Grade Methods

Homework - 10 %

Computer projects (1st and 2nd) - 40 %

Computer project (3rd) - 25%

Exam - 25 %

* The instructor reserves the right to change 10% of the grade method

4

4

Important

1. **Plagiarism/Cheating**: Any same or similar assignment

For more information, see **University Regulation 4.001. Code of Academic Integrity**

Plagiarism is unacceptable in the University community. Academic work must be an original work of your own thought, research, or self-expression. When students borrow ideas, wording, or organization from another source, they must acknowledge that fact in an appropriate manner. Plagiarism is the deliberate use and appropriation of another's work without identifying the source and trying to pass off such work as one's own. Any student who fails to give full credit for ideas or materials taken from another has plagiarized. This includes all discussion board posts, journal entries, wikis, and other written and oral presentation assignments.

5

5

Important

2. No late submission: **All assignments** are due by 11:59PM on the due date. However, it is recommended that you submit your work by 9:00pm as no assistance will be provided after 9:00pm should you encounter any technical difficulties.

3. The use of **AI Language is prohibited** for any assignment in this course.

6

6

Important

4. Submit your assignment **on Canvas**. Email submission is not allowed.

5. Recording

7

7

Course Outline

The ***objective*** of this course is to introduce the foundations, principles and algorithms for reinforcement learning (tabular solution methods and approximate solution methods), and apply the knowledge to solve real-world application problems in many other domains. Topics covered in this course include, multi-armed bandits, Markov decision process, dynamic programming, Monte Carlo process, Q-learning, Actor-Critic, temporal difference methods, on-policy and off-policy methods, eligibility traces and others.

8

8

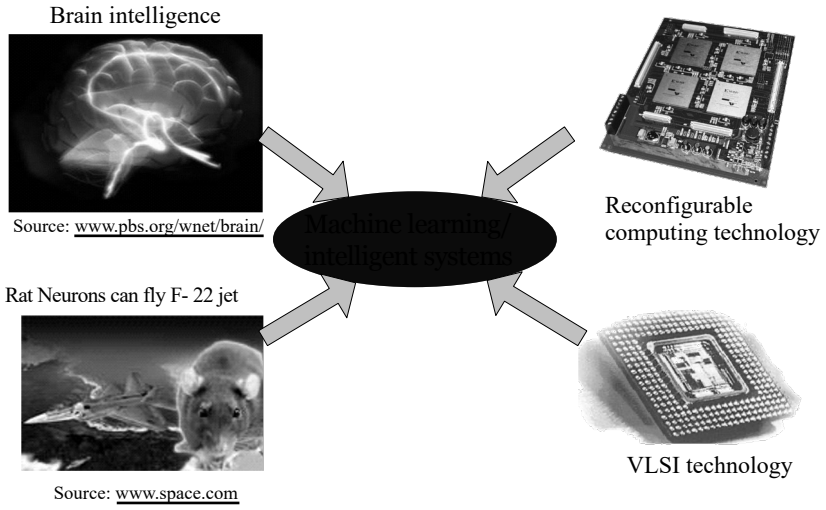
Prerequisites

Basic knowledge of machine learning, neural networks, statistics, and random process.

9

9

Motivations for real-intelligent systems



10

10

One way to understand "intelligence" is by looking at our own capabilities, *which means that* humans are able to:

- think
- understand
- recognize
- perceive
- generalize
- adapt
- learn
- make decisions
- solve daily problems

Source: https://www.flickr.com

Source: https://pixabay.com

11

According to the Oxford and English Dictionaries the word "intelligence" can be defined as follows:

- ability to understand
- reason
- perceive
- quickness in learning
- mental alertness
- ability to grasp relationships
- clever
- information
- news

Source: https://en.wikipedia.org

12

Elements of intelligence

"The ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience." --- (L. S. Gottfredson)
"Four behaviors are commonly used to distinguish intelligent behavior from instinct and stimulus-response associations: (1) prediction, (2) response to change, (3) intentional action, and (4) reasoning. "--- (R. Davis)
"The ability to carry on abstract thinking" --- (L. M. Terman)
"Having learned or ability to learn to adjust oneself to the environment" --- (S. S. Colvin)
"The ability to adapt oneself adequately to relatively new situations in life" --- (R. Pintner)
"The capacity to learn or to profit by experience" --- (W. F. Dearborn)
……

13

13

Intelligent Systems

- To make persons, minds
- Introduction to the Science and Technology of Artificial Intelligence
  - touches on control theory, psychology, operations research, philosophy, and neuroscience
- A technical and conceptual foundation for understanding this large and complex set of issues

14

14

Intelligence is the ability to achieve goals

- "Intelligence is the most powerful phenomena in the universe" —Ray Kurzweil, c 2000
  - The phenomena is that there are systems in the universe that are well thought of as goal-seeking systems
- What is a goal-seeking system?
  - "Constant ends from variable means is the hallmark of mind" —William James, c 1890
  - a system that is better understood in terms of *outcomes* than in terms of *mechanisms*

15

15

How do we design Intelligence?

- Study from biological models (brain, genetic, DNA, life, Molecular biology, …) → neural nets, Genetic algorithm, Artificial Life, DNA Computing, Quantum Computing, Robotics, etc.
- Study from human phenomena (common sense, reasoning, predicting, observing, inference, …)→ fuzzy logic, expert systems, search techniques, etc.
- Need to develop mathematical/logical algorithms based on the above biological models or phenomena

16

16

Challenge for designing of intelligent systems

Dynamic, incremental and adaptive learning algorithms;

Limited understanding of intelligent behaviors;

Lack of system level models and architectures to mimic brain-level intelligence;

Embodiment of machine intelligence hardware within systems that learn through interaction with environment;

Systems design and implementation;

17

17

Intelligence?

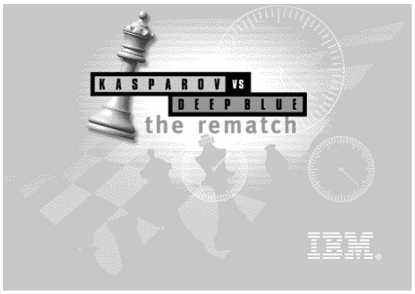
Embodied intelligence;
Dynamically interact with environment;
Dynamic learning;
Associations;
Invariance representation;

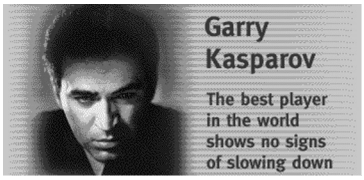
Can a super-computer perform the same way as a one year-old baby?

18

18

Intelligence?





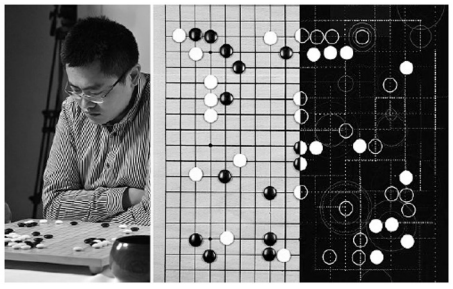
Source: IBM research
http://www.research.ibm.com/deepblue/

19

19

AlphaGo vs. human

Go game



picture from: http://www.dailymail.co.uk/

AlphaGo vs. Fan Hui
(5 : 0)

picture from: http://www.ibtimes.com/

AlphaGo vs. Lee Sedol
(4 : 1)

20

20

The coming of artificial intelligence

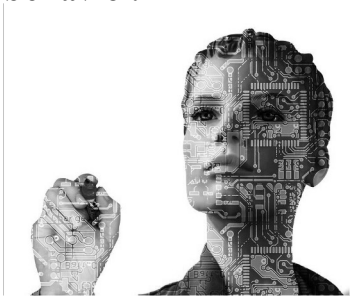
- When people finally come to understand the principles of intelligence—what it is and how it works—well enough to design and create beings as intelligent as ourselves
- A fundamental goal for science, engineering, the humanities, …for all mankind
- It will change the way we work and play, our sense of self, life, and death, the goals we set for ourselves and for our societies
- But it is also of significance beyond our species, beyond history
- It will lead to new beings and new ways of being, things inevitably *much <u>more powerful</u> than our current selves*

21

21

Artificial Intelligence (AI) ?

AI is a study about inventing machines/computers that capable of mimicking human/animal intelligent behavior.



Source: https://pixabay.com

Source: https://pixabay.com

The ultimate objective is to develop a system that can think and act rationally like humans.

22

22

Where AI can/should be applied?

- Data is overwhelming/abundance
- Too many manual operations/procedures
- Optimization is possible
- Parallel/Distributed procedures/architectures are needed
- Decision making is required
- When current techniques are too complicated to be used/designed

23

23

Where AI can/should be applied? .. Cont'd

- Mathematical models are too complex/impossible
- To increase efficiency
- To reduce cost
- To improve performance and reliability

24

24

Some Important Facts, you need to know….

- AI is not the only solution.
- AI is only one part of technology.
- AI is just a tool for improvement.
- You must know your domain/target application.

25

25

Source: http://archive.defense.gov

AI IN INDUSTRIES

26

26

Advantages of Adding Intelligence in Products/ Systems

- Better performance
- Longer Life
- Reliability
- Simpler operation
- Cost effective
- Higher efficiency
- Self-organizing / self-optimization
- Simpler design

27

27

Is there really a need for AI?

- Manufacturers need to improve on their products
- Need to satisfy customers
- Need to improve products' reliability
- Need to improve products' performance
- Need to improve products' features
- Need to distinguish their products away from their competitors

28

28

Is human-level AI *possible*?

- If people are biological machines, then eventually we will reverse engineer, and understand their workings
- Then, surely we can make improvements
  - with materials and technology not available to evolution
  - how could there not be something we can improve?
  - design can overcome local minima, make great improvement, try things much faster than biology

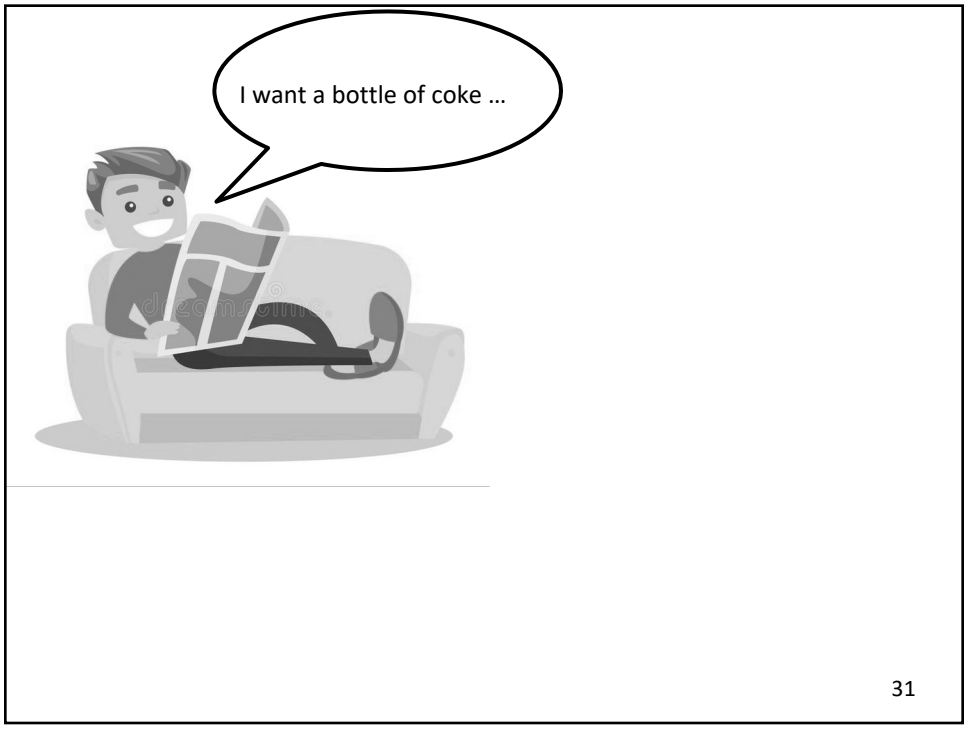
Yes

29

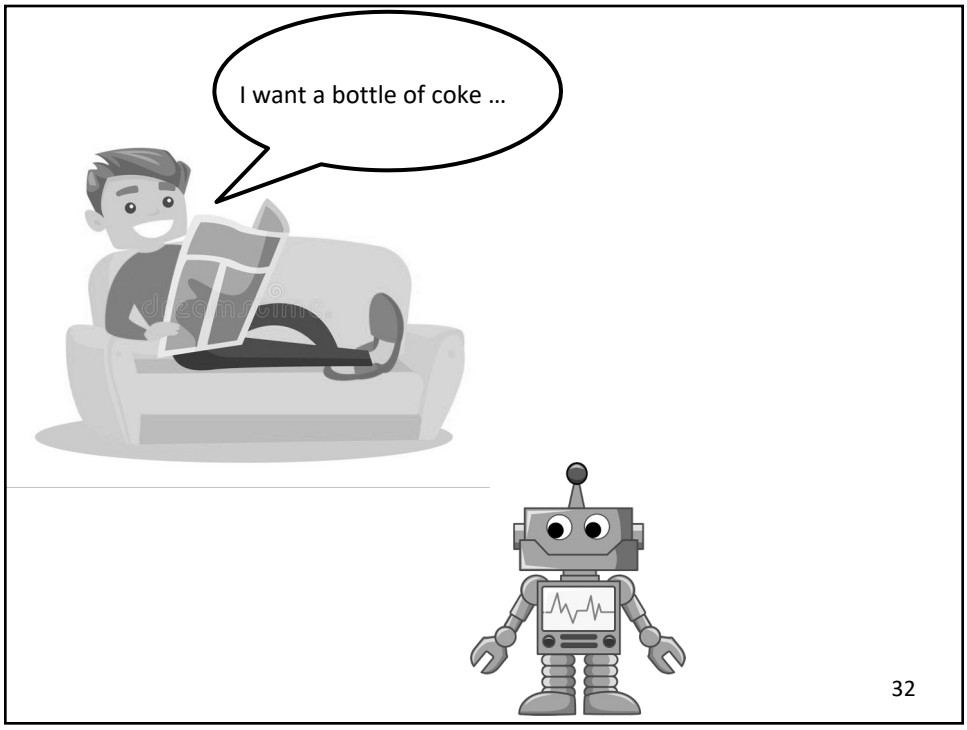
29

30

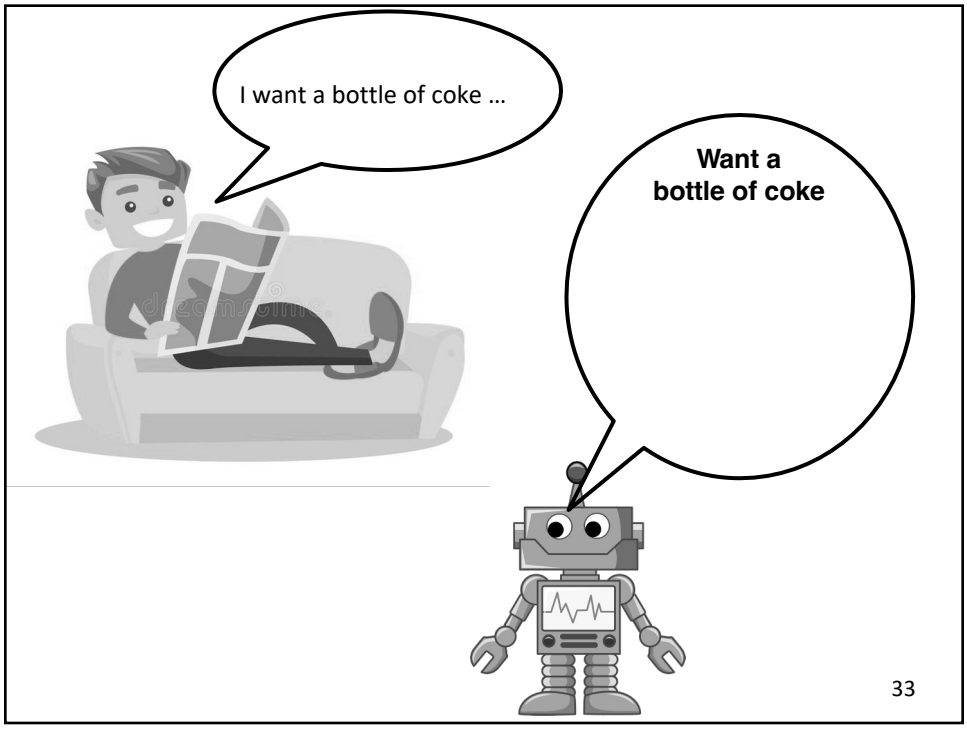
30



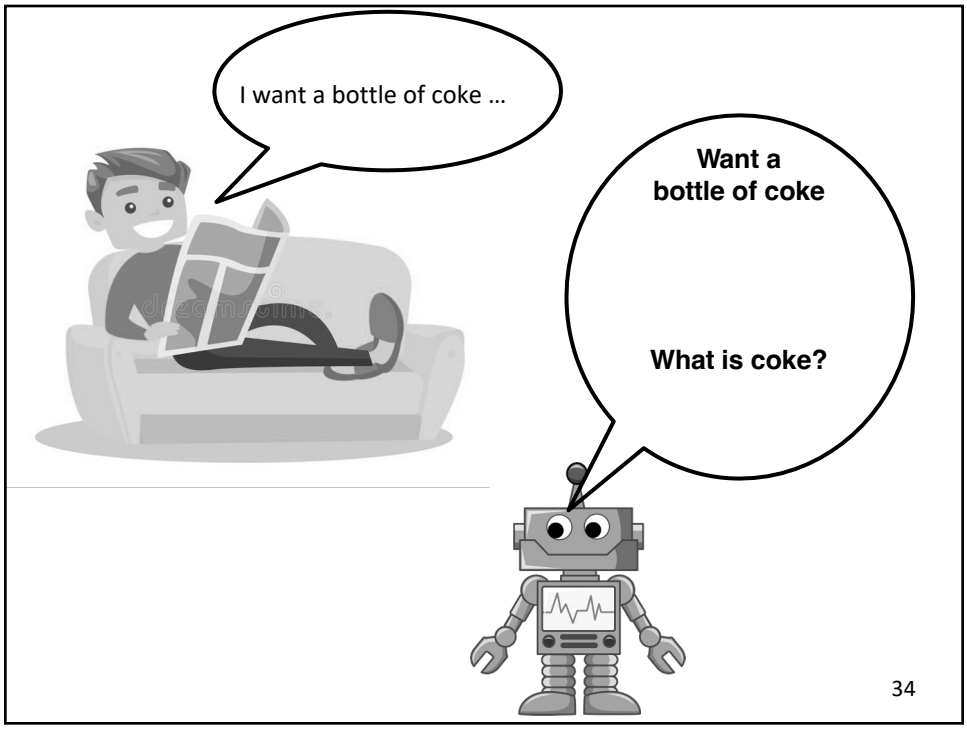
31



32


I want a bottle of coke …
Want a bottle of coke
33

33


I want a bottle of coke …
Want a bottle of coke
What is coke?
34

34

With the availability of …

Large Data & Compute

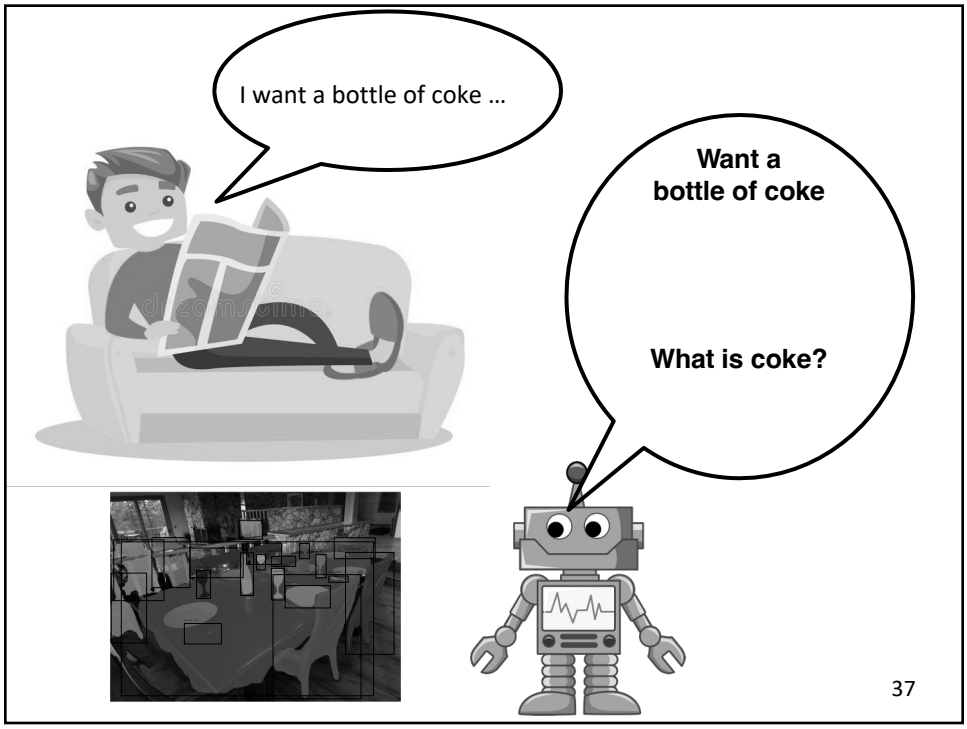
Imagenet

GPUs

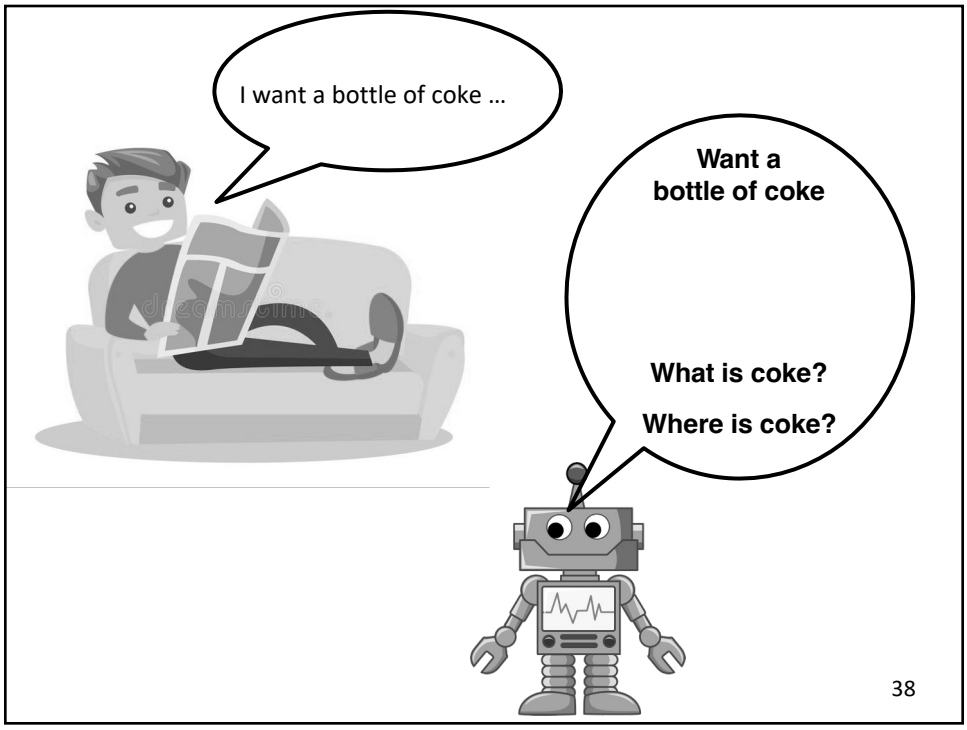
35

35

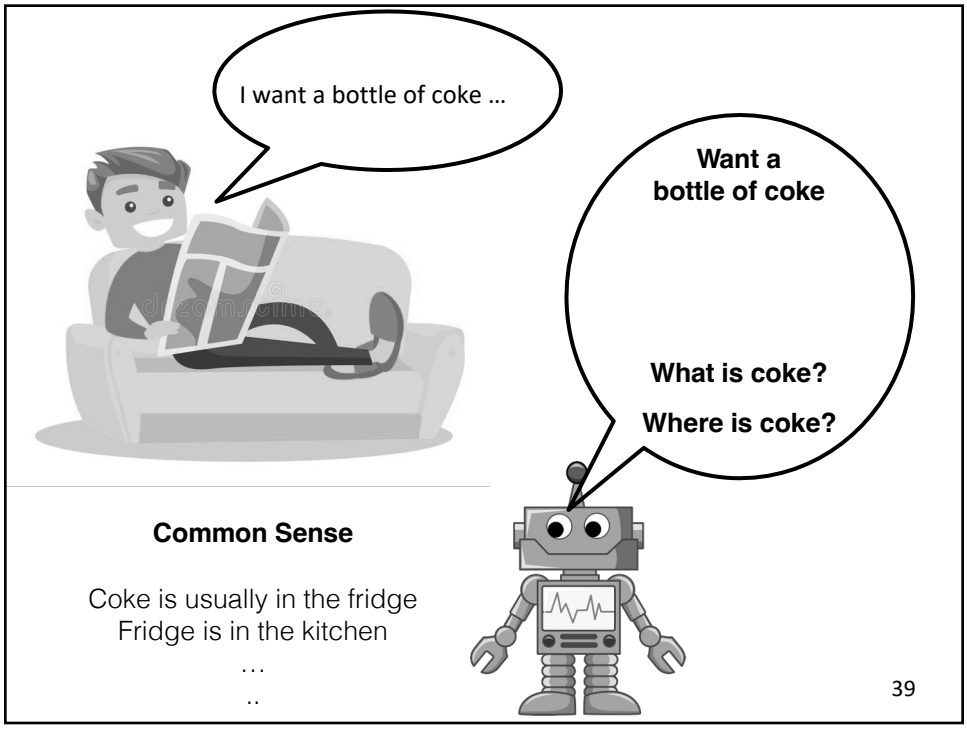
36


I want a bottle of coke …
Want a bottle of coke
What is coke?
37

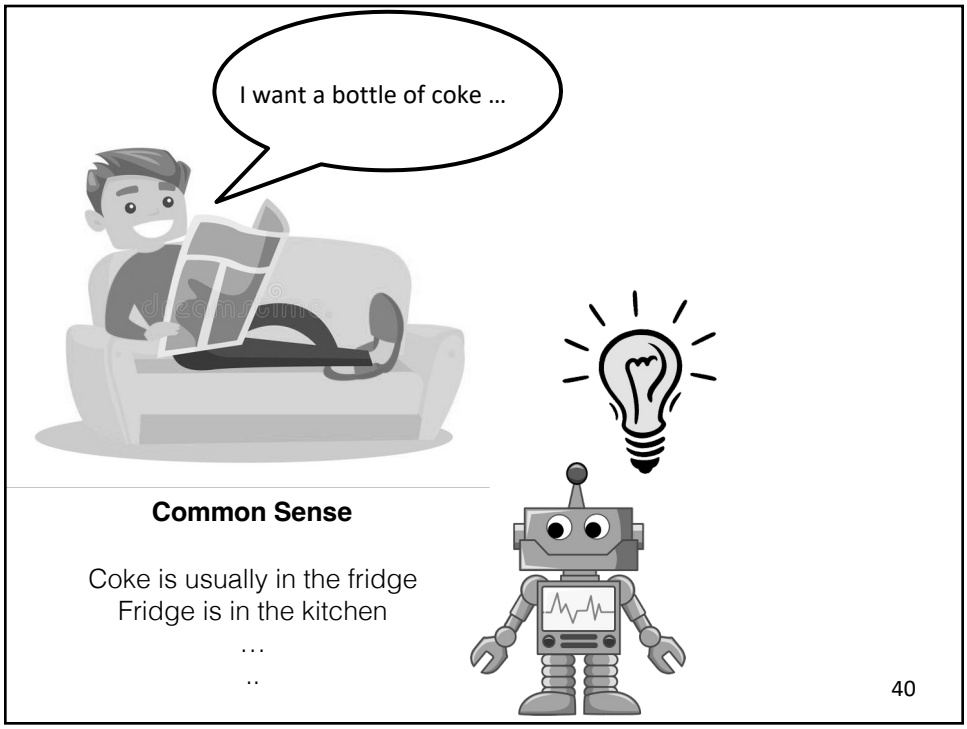
37


I want a bottle of coke …
Want a bottle of coke
What is coke?
Where is coke?
38

38


I want a bottle of coke …
Want a bottle of coke
What is coke?
Where is coke?
Common Sense
Coke is usually in the fridge
Fridge is in the kitchen
…
..
39

39


I want a bottle of coke …
Common Sense
Coke is usually in the fridge
Fridge is in the kitchen
…
..
40

40

When will human-level AI first be created?

- No one knows of course; we can make an educated guess about the probability distribution:
  - 25% chance by 2040
  - 50% chance by 2050
  - 10% chance never
- Certainly, a significant chance within all of our expected lifetimes
  - We should take the possibility into account in our career plans

41

41

Investment in AI is way up

- Google's prescient AI buying spree: Boston Dynamics, Nest, Deepmind Technologies, …
- New AI research labs at Facebook, Baidu, Allen Institute, Vicarious, Maluuba, DeepMind Alberta…
- Also enlarged corporate AI labs: Microsoft, Amazon, Adobe…
- Yahoo makes major investment in CMU machine learning department
- Many new AI startups getting venture capital
- New Canadian AI funding in Toronto, Montreal, and Edmonton
  - The Alberta Machine Intelligence Institute

42

42

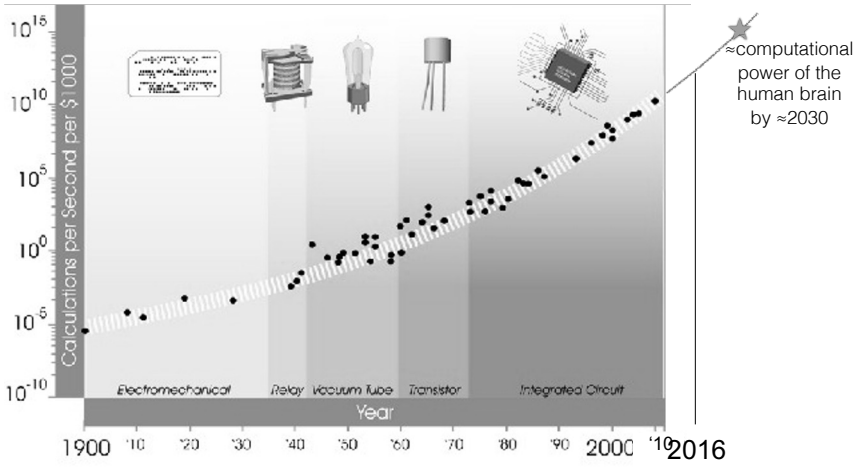
The 2nd industrial revolution

- The 1st industrial revolution was the *physical power* of machines substituting for that of people
- The 2nd industrial revolution is the *computational power* of machines substituting for that of people
  - Computation for perception, motor control, prediction, decision making, optimization, search
  - Until now, people have been our cheapest source of computation
  - But now our machines are starting to provide greater, cheaper computation

43

43

The computational revolution



44
Kurzweil AI

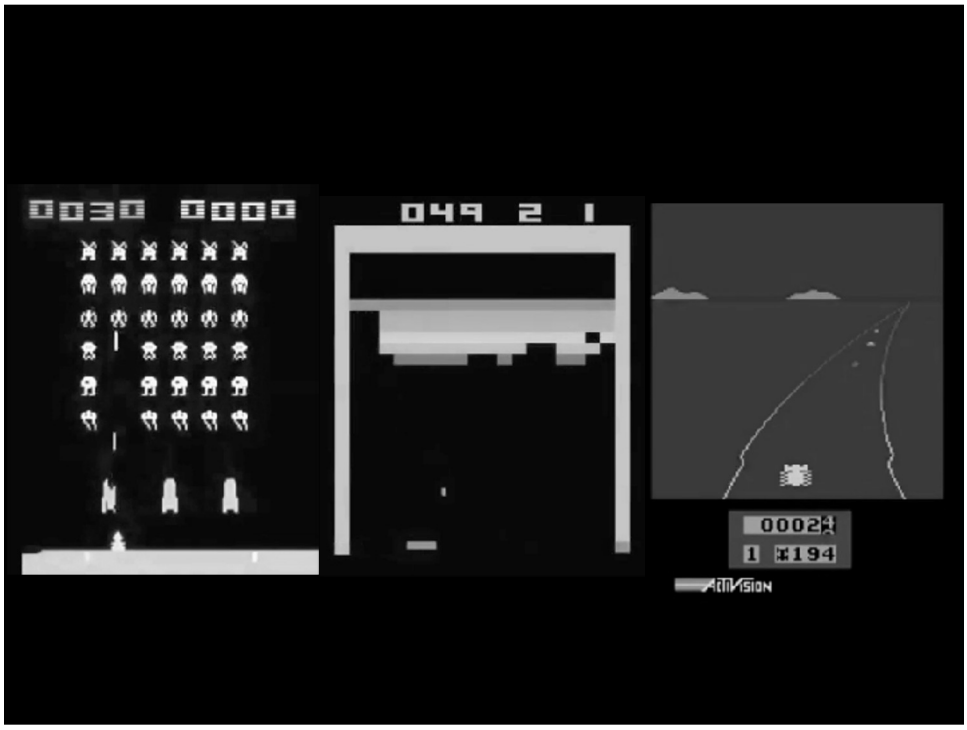
44

Advances in AI abilities are coming faster;

- IBM's Watson beats the best human players of *Jeopardy!* (2011)
- Deep neural networks greatly improve the state of the art in speech recognition and computer vision (2012–)
- Google's self-driving car becomes a plausible reality (≈2013)
- Deepmind's DQN learns to play Atari games at the human level, from pixels, with no game-specific knowledge (≈2014, *Nature*)
- University of Alberta program solves Limit Poker (2015, Science), and then defeats professional players at No-limit Poker (2017, Science)
- Google Deepmind's AlphaGo defeats legendary Go player Lee Sedol (2016, Nature), and world champion Ke Jie (2017), vastly improving over all previous programs

45

45



46