人工智能基础

徐林莉 linlixu@ustc.edu.cn

课程简介

- 教材
  - Artificial Intelligence – A modern approach
    S. Russell and P. Norvig
    人工智能—一种现代方法（第三版）
- 参考书
  - 机器学习 周志华 (2016)
- 课程考核
  - 学期总评=期末考试(60%)+书面作业(15%)+实验部分(25%)
- 课件：
  http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

课程大纲

- 第一部分：人工智能概述/Introduction and Agents (chapters 1,2)
- 第二部分：问题求解/Search (chapters 3,4,5,6)
- 第三部分：知识与推理/Logic (chapters 7,8,9,10)
- 第四部分：不确定知识与推理/Uncertainty (chapters 13 -17)
- 第五部分：学习/Learning (chapters 18,19,20,21)

http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

助教

又木 [youmu1998@mail.ustc.edu.cn]

李博诚 [bcli@mail.ustc.edu.cn]

蒋征国 [jiangzhengguo@mail.ustc.edu.cn]

Chapter 1. Introduction

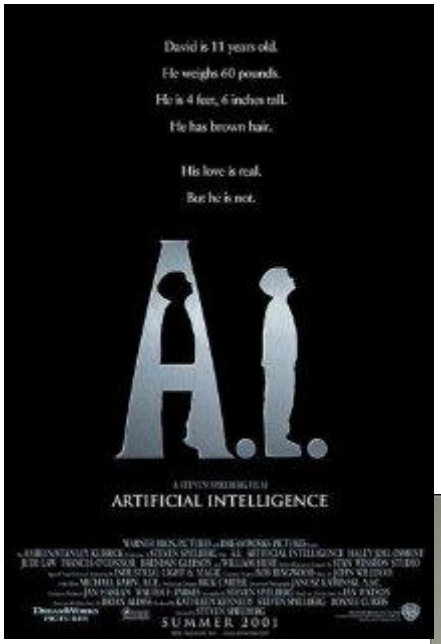
http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

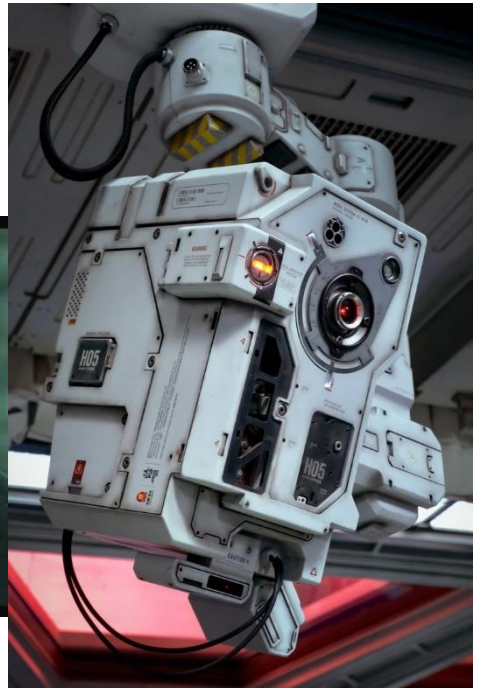
Introduction

- What is AI?
- The history of AI（历史）
- Recent progress in AI（现状）

http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

Sci-Fi AI?


David is 11 years old.
He weighs 60 pounds.
He is 4 feet, 6 inches tall.
He has brown hair.
His love is real.
But he is not.
A.I.
ARTIFICIAL INTELLIGENCE
SUMMER 2001



Early Achievements of AI

- Deep Blue playing chess （国际象棋）

Early Achievements of AI

- Face detection, in most digital cameras for auto focusing

Also blink and smile detection!

http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

Early Achievements of AI

- Feb 2011, Watson（沃森）beat human on the quiz show Jeopardy!. And received the first prize of $1 million.

Threats of AI?

BBC News Sport Weather

NEWS TECHNOLOGY

Home | World | Asia | Australia | India | China | UK | Business | Health | Science

2 December 2014 Last updated at 13:02

Stephen Hawking warns artificial intelligence could end mankind

COMMENTS (1027)

By Rory Cellan-Jones
Technology correspondent

Stephen Hawking: "Humans, who are limited by slow biological evolution, couldn't compete and would be superseded"

Prof Stephen Hawking, one of Britain's pre-eminent scientists, has said that efforts to create thinking machines pose a threat to our very existence.

New Achievements of AI

Silver, D. et al. Nature (2016)

Go, a complex game popular in Asia, has frustrated the efforts of artificial-intelligence researchers for decades.

ARTIFICIAL INTELLIGENCE

Google masters Go

Deep-learning software excels at complex ancient board game.

http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

New Achievements of AI

Li, Y. et al. (2022)

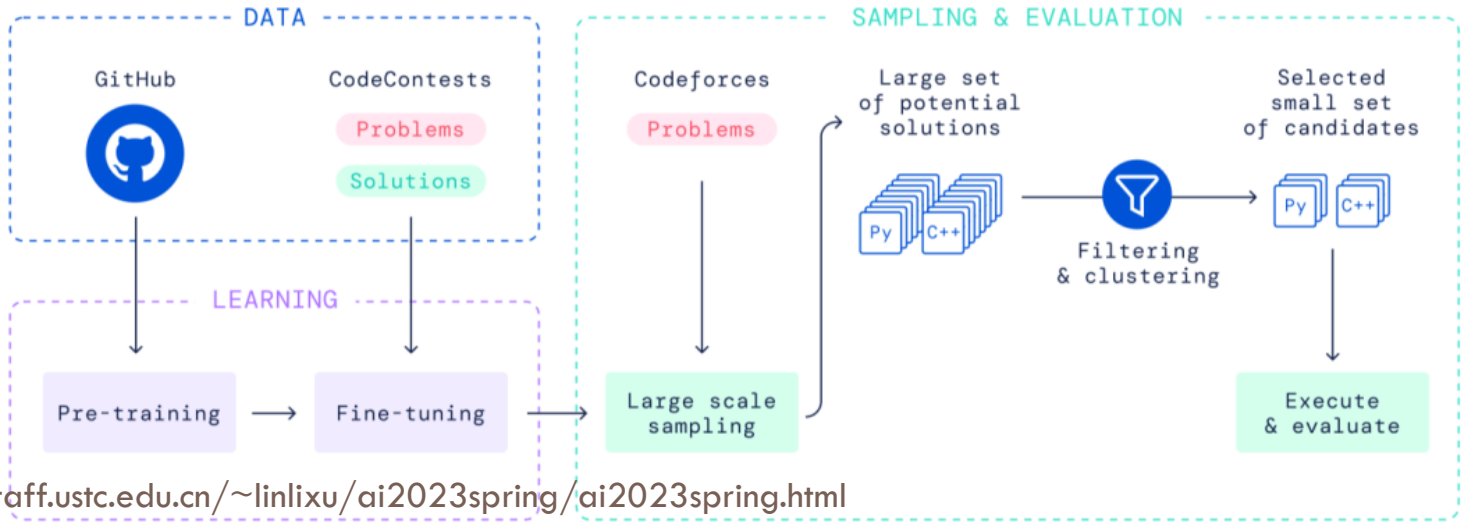
DeepMind

2022-2-19

Competition-Level Code Generation with AlphaCode

Yujia Li*, David Choi*, Junyoung Chung*, Nate Kushman*, Julian Schrittwieser*, Rémi Leblond*, Tom Eccles*, James Keeling*, Felix Gimeno*, Agustin Dal Lago*, Thomas Hubert*, Peter Choy*, Cyprien de Masson d'Autume*, Igor Babuschkin, Xinyun Chen, Po-Sen Huang, Johannes Welbl, Sven Gowal, Alexey Cherepanov, James Molloy, Daniel J. Mankowitz, Esme Sutherland Robson, Pushmeet Kohli, Nando de Freitas, Koray Kavukcuoglu and Oriol Vinyals

*Joint first authors



http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

New Achievements of AI

Agostinelli, A. et al. (2023)

MusicLM: Generating Music From Text

Napoleon Crossing the Alps - Jacques-Louis David

"The composition shows a strongly idealized view of the real crossing that Napoleon and his army made across the Alps through the Great St Bernard Pass in May 1800." By wikipedia

http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

New Achievements of AI

Li, Y. et al. (2022)

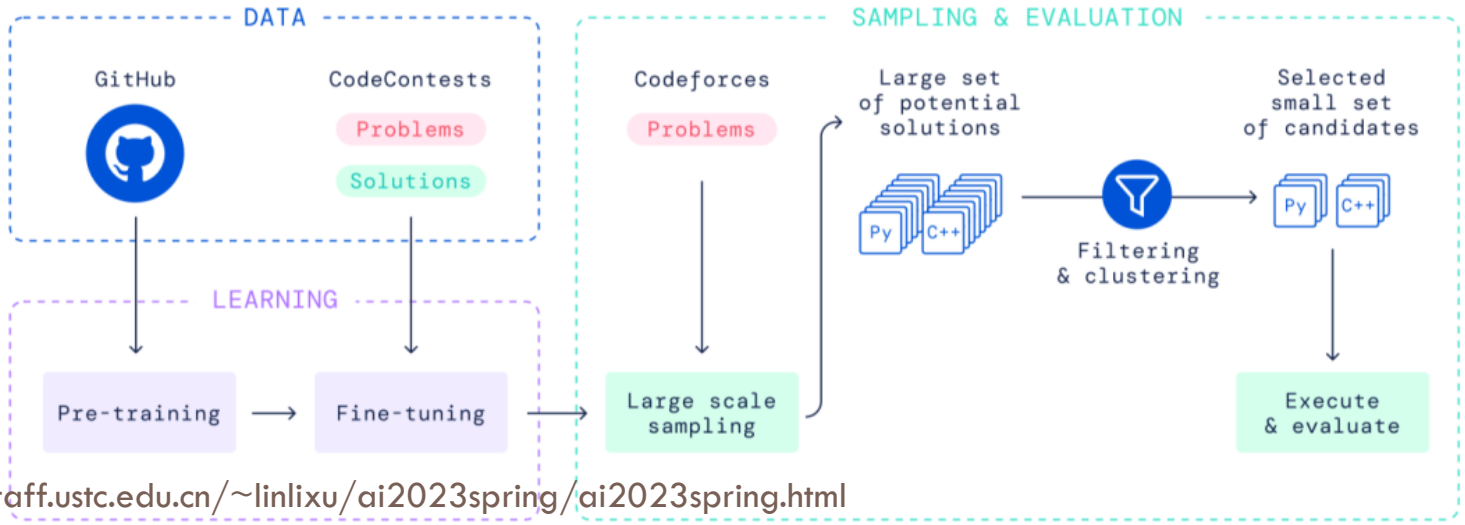
DeepMind

2022-2-19

Competition-Level Code Generation with AlphaCode

Yujia Li*, David Choi*, Junyoung Chung*, Nate Kushman*, Julian Schrittwieser*, Rémi Leblond*, Tom Eccles*, James Keeling*, Felix Gimeno*, Agustin Dal Lago*, Thomas Hubert*, Peter Choy*, Cyprien de Masson d'Autume*, Igor Babuschkin, Xinyun Chen, Po-Sen Huang, Johannes Welbl, Sven Gowal, Alexey Cherepanov, James Molloy, Daniel J. Mankowitz, Esme Sutherland Robson, Pushmeet Kohli, Nando de Freitas, Koray Kavukcuoglu and Oriol Vinyals

*Joint first authors



http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

What is AI?

Different people think of AI differently

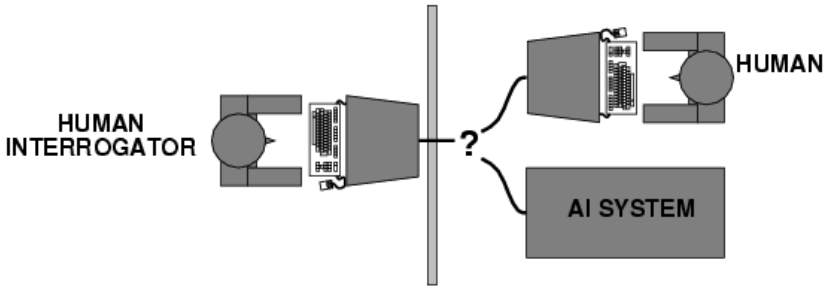
Views of AI fall into four categories:

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

The textbook advocates “acting rationally（理性的）”

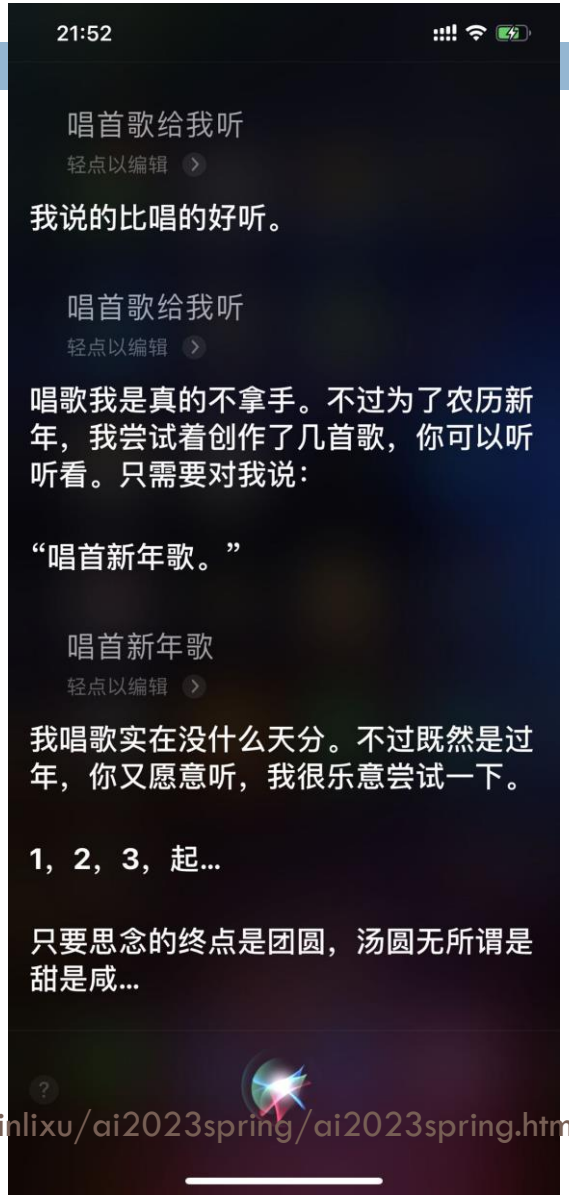
Acting humanly: Turing Test图灵测试

- 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（知识）, reasoning（推理）, language understanding（语言理解）, learning（学习）
- Problem: Turing test is not reproducible or amenable to mathematical analysis

Acting humanly: Turing Test图灵测试



Thinking humanly: cognitive modeling
认知模型

- 1960s "cognitive revolution": information-processing psychology
- Requires scientific theories of internal activities of the brain
- -- How to validate? Requires
  1) Predicting and testing behavior of human subjects (top-down)
  or 2) Direct identification from neurological data (bottom-up)
  - Both approaches (roughly, Cognitive Science and Cognitive Neuroscience) are now distinct from AI

Thinking rationally: "laws of thought"

- Aristotle: what are correct arguments/thought processes?
- Several Greek schools developed various forms of *logic*: *notation*（符号） and *rules*（规则） *of derivation*（推导） for thoughts; may or may not have proceeded to the idea of mechanization
- Direct line through mathematics and philosophy to modern AI
- Problems:
  - Not all intelligent behavior is mediated by logical deliberation
  - What is the purpose of thinking? What thoughts should I have?
  - Logical systems tend to do the wrong thing in the presence of **uncertainty**

Acting rationally: rational agent

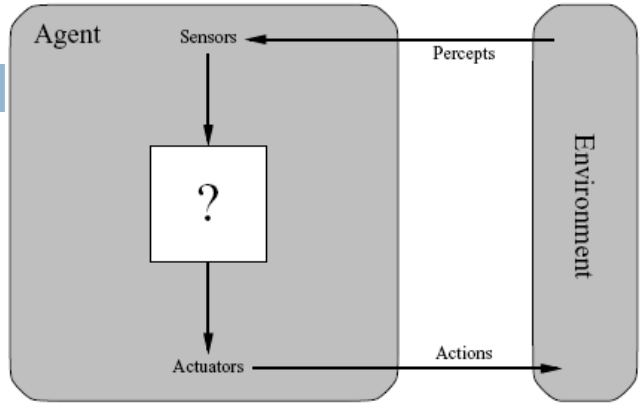
- Rational behavior: doing the right thing
  - The right thing: which is expected to maximize goal achievement, given the available information
  - Doesn't necessarily involve thinking – e.g., blinking reflex – but thinking should be in the service of rational action
  - Entirely dependent on goals!
  - Irrational ≠ insane, irrationality is sub-optimal action
  - Rational ≠ successful

Acting rationally: rational agent

- Our focus here:
  - Systems which make the best possible decisions given goals, evidence, and constraints
  - In the real world, usually lots of uncertainty
    - … and lots of complexity
  - Usually, we're just approximating rationality

Maximize Your Expected Utility

Rational agents



- An agent is an entity that perceives and acts
- This course is about designing rational agents
- Abstractly, an agent is a function from percept histories to actions:

$$[f: \mathcal{P}^* \rightarrow \mathcal{A}]$$

- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance
- Caveat: computational limitations make perfect rationality unachievable
  → design best program for given machine resources

AI prehistory

Philosophy	Logic, methods of reasoning, mind as physical system foundations of learning, language, rationality
Mathematics	Formal representation and proof algorithms, computation, (un)decidability, (in)tractability, probability
Economics	utility, decision theory
Neuroscience	physical substrate for mental activity
Psychology	phenomena of perception and motor control, experimental techniques
Computer engineering	building fast computers
Control theory	design systems that maximize an objective function over time
Linguistics	knowledge representation, grammar

http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

AI history

- The gestation of AI 孕育期 （—1956）
- Reasoning methods 注重推理时期 （1956-1975）
- Knowledge-based system 知识运用时期 （1976-1988）
- Integration 集成运用（1989- present）

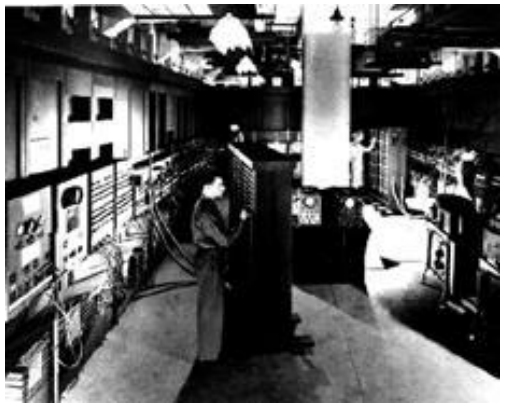
The Gestation of AI

- 古希腊Aristotle（亚里士多德BC 384-322），给出形式逻辑的基本规律 Syllogism(三段论)。
- 英国Bacon（培根1561-1626），系统地给出Induction(归纳法)。
- 德国Leibnitz（莱布尼茨1646-1716）提出Symbolic Logic(数理逻辑)。
- 英国Boole（布尔1815-1864）提出Boolean Algebra(布尔代数)系统，实现了思维符号化和数学化

The Gestation of AI (Cont.)

- 1936 英国Turing（图灵, 1912-1954）:理想计算机模型Turing Machine（图灵机）

- 1946 美国Mauchly（莫克利）, Eckert（埃克特）: ENIAC



- 1948 美国Shannon（香农）: Information Theory(信息论)

- 1950 Turing Test图灵测试

The Birth of AI (1956)

John McCarthy organized a two-month workshop at Dartmouth in the summer of 1956, ten young men were there:

McCarthy, Minsky, Rochester, Shannon, Moore, Samuel, Selfridge, Solomonff, Simon, Newell.

They introduced all the major figures to each other and agreed to adopt the name of Artificial Intelligence for the field.

Abridged history of AI

1943	McCulloch & Pitts: Boolean circuit model of brain
1950	Turing's "Computing Machinery and Intelligence"
1956	Dartmouth meeting: "Artificial Intelligence" adopted
1952—69	Look, Ma, no hands!
1950s	Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
1965	Robinson's complete algorithm for logical reasoning
1966—73	AI discovers computational complexity Neural network research almost disappears
1969—79	Early development of knowledge-based systems
1980--	AI becomes an industry
1986--	Neural networks return to popularity
1987--	AI becomes a science
1995--	The emergence of intelligent agents

http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

AI Today

- Mostly about engineering domain-specific solutions rather than creating general theories
- We don't know how to do most of intelligent things, but the rest can be solved pretty well
- A set of "tools" for representing information and using them to solve specific tasks
  - Neural networks, hidden Markov models, Bayesian networks, heuristic search, logic, …
- There's no magic in AI. It's all about representation, optimization, probability, and algorithms

http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

Well-known AI applications

- Expert systems (organic chemistry, medicine, geology, configuring computers)
- Speech recognition
- Handwriting recognition
- Game playing (chess, checkers, now Go)
- Robots (automated cars, ping pong player, Honda robot)
- Automated theorem proving
- Web search engines
- Natural language understanding (machine translation, Google)
- Logistics scheduling (military --- people, cargo, vehicles)
- Cruise missiles
- Microsoft Answer Wizard

State of the art

- Google language translation services
- Google automatic news aggregation and summarization
- Nuance voice recognition (behind Apple's Siri)
- Face detection and face recognition systems
- Apple Siri question-answering system
- IBM Watson question-answering system
- IBM Deep Blue chess playing program
- Deepmind Alpha Series
- Large Language Models, ChatGPT
- Driverless cars

http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html

Summary

- Applications of AI:
  - high-impact (affect billions of people)
  - diverse (language, vision, robotics)
- Challenges: really hard…
  - computation complexity
  - information complexity
- Paradigm: modeling + algorithms

http://staff.ustc.edu.cn/~linlixu/ai2023spring/ai2023spring.html