ARTIFICIAL INTELLIGENCE(AI)

2023/2024 Semester 2

Introduction: Chapter 1

Artificial Intelligence (AI)

1912-1954

tson, Feb., 2011

来源：Nature

Artificial Intelligence (AI) - deep learning

2019 Turing Award

Yanhong Li, Jan., 2013

Google Brain, New york Times ,June, 2012

Artificial Intelligence (AI) - deep learning

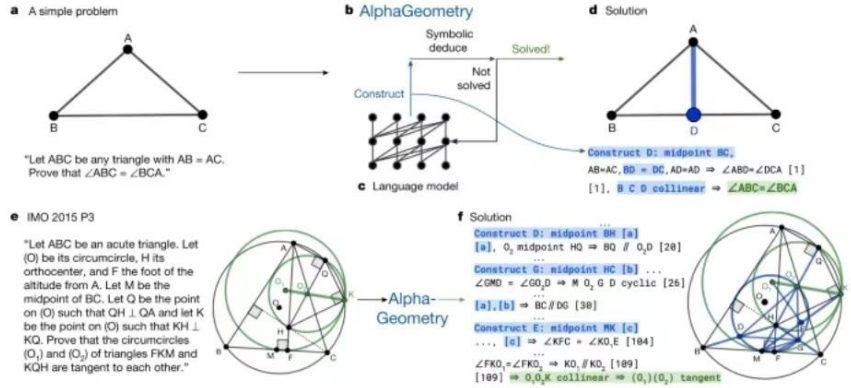
Google AlphaGo（2016）

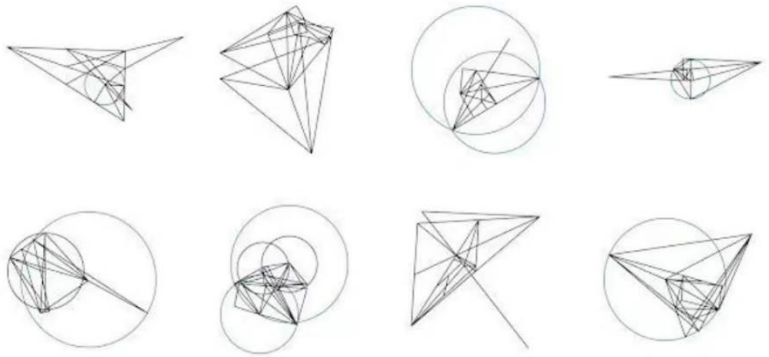
Alpha Zero（2017）

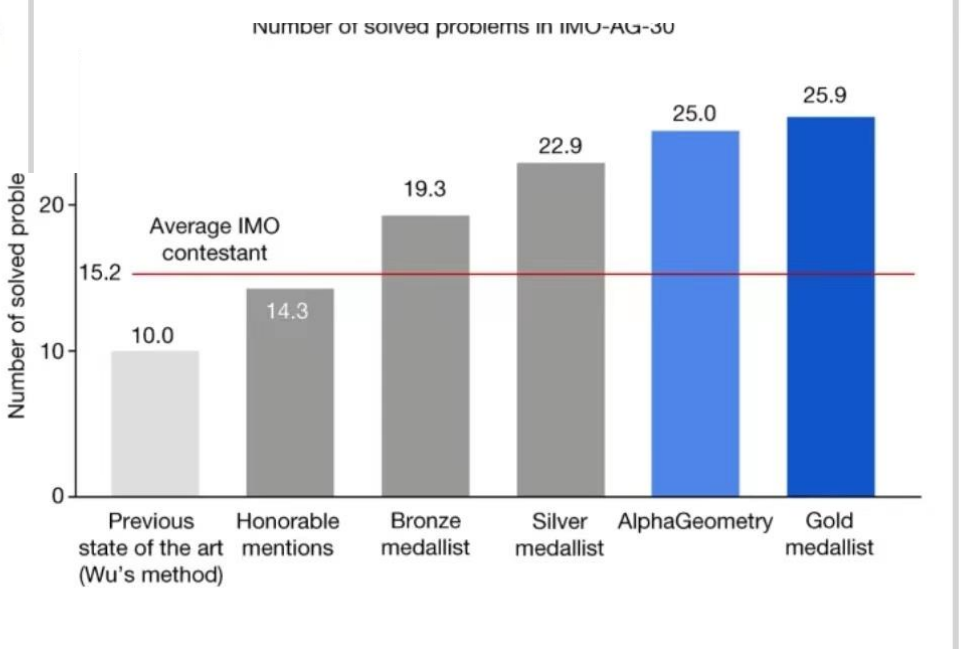
Artificial Intelligence (AI) - deep learning

AlphaFold（2020）…

AlphaGeometry(2024)







- 方法：结合了语言学习和演绎推理的神经符号系统
- 数据量：**10亿**：通过从形状和线条中随机生成图形来创建训练数据，解决了数学中训练数据太少的问题

。。。

Artificial Intelligence (AI) – large language model

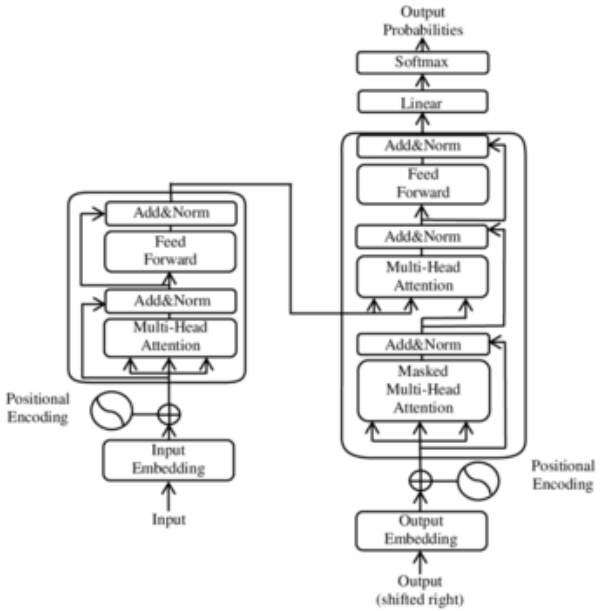
OPEN AI大模型的发展：all in AGI，目的是用计算机模拟真实世界

- **2021年1月5日**
  文生图模型Dall-E发布
- **2022年4月**
  Dall-E 2发布
- 2022年11月30日
  大语言模型 Chat GPT发布
- 2023年3月15日
  GPT-4正式面世
- **2023年9月21日**
  Dall-E 3正式发布
- **2024年3月15日**
  文生视频大模型Sora问世

Artificial Intelligence (AI) – large language model

Chat GPT

Transformer神经网络架构
RLHF 人类反馈强化学习
SFT 有监督微调
IFT 指令微调
CoT 思维链
红蓝对抗
。 。 。



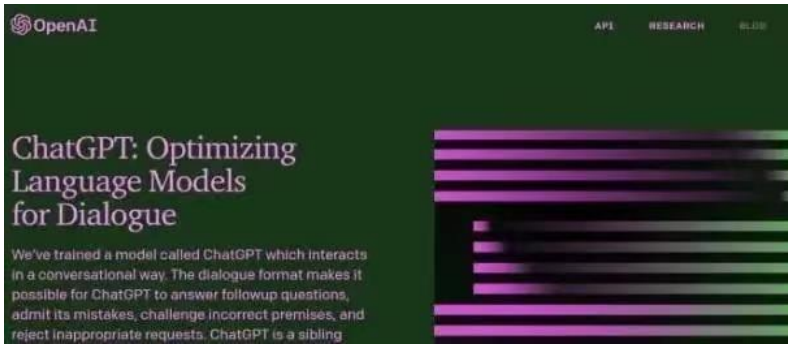
消除反馈(recurrent)机制
Google (2017): Attention is all you need

学习单词与单词之间的共现概率关联

ChatGPT的主要特点

Artificial Intelligence (AI) - large language model

Chat GPT



优化对话的语言模型

模型	发布时间	参数量	预训练数据量
GPT-1	2018年6月	1.17亿	约5GB
GPT-2	2019年2月	15亿	40G
GPT-3	2020年5月	1750亿	45TB
ChatGPT	2022年11月		

GPT家族主要模型对比

- **数据：**训练中使用了45TB数据、近 1 万亿个单词（约1351万本牛津词典所包含单词数量）以及数十亿行源代码。
- **模型：**包含了1750亿参数，将这些参数全部打印在A4纸张上，一张一张叠加后，叠加高度将超过上海中心大厦632米高度。
- **算力：** ChatGPT的训练门槛是1万张英伟达V100芯片、约10亿人民币，模型训练算力开销是每秒运算一千万亿次，需运行3640天（ 3640 PetaFLOPs per day ）。
- **大数据、大模型、大算力下以“共生则关联”原则实现了统计关联关系的挖掘。**

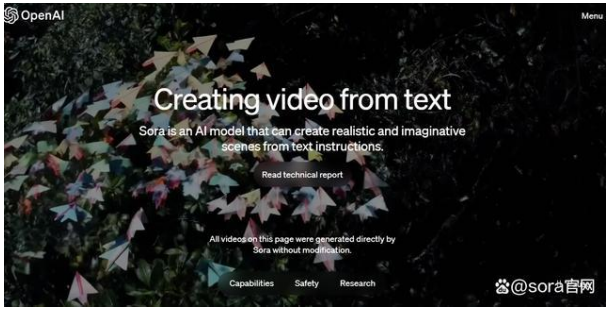
Artificial Intelligence (AI) – large language model

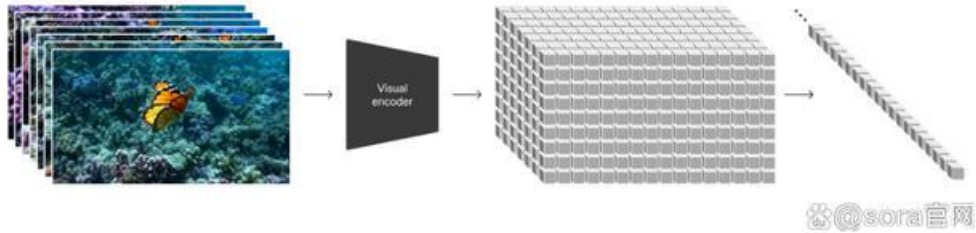
Sora

Transformer + diffusion model架构

能够将随机噪声逐渐转化为有意义的图像或视频内容。

技术背后的原理：对合成内容中的最小单元进行有意义的关联组合，合成以前从未有过的内容





- 从一段文本生成视频的步骤：
  1、语义理解（例如，基于GPT）
  2、生成图像（基于Diffusion）
  3、通过图像序列生成视频

Artificial Intelligence (AI) — large language model

Sora

它能够根据用户输入的文本描述，生成长达一分钟的高质量视频

女子行走在东京街头

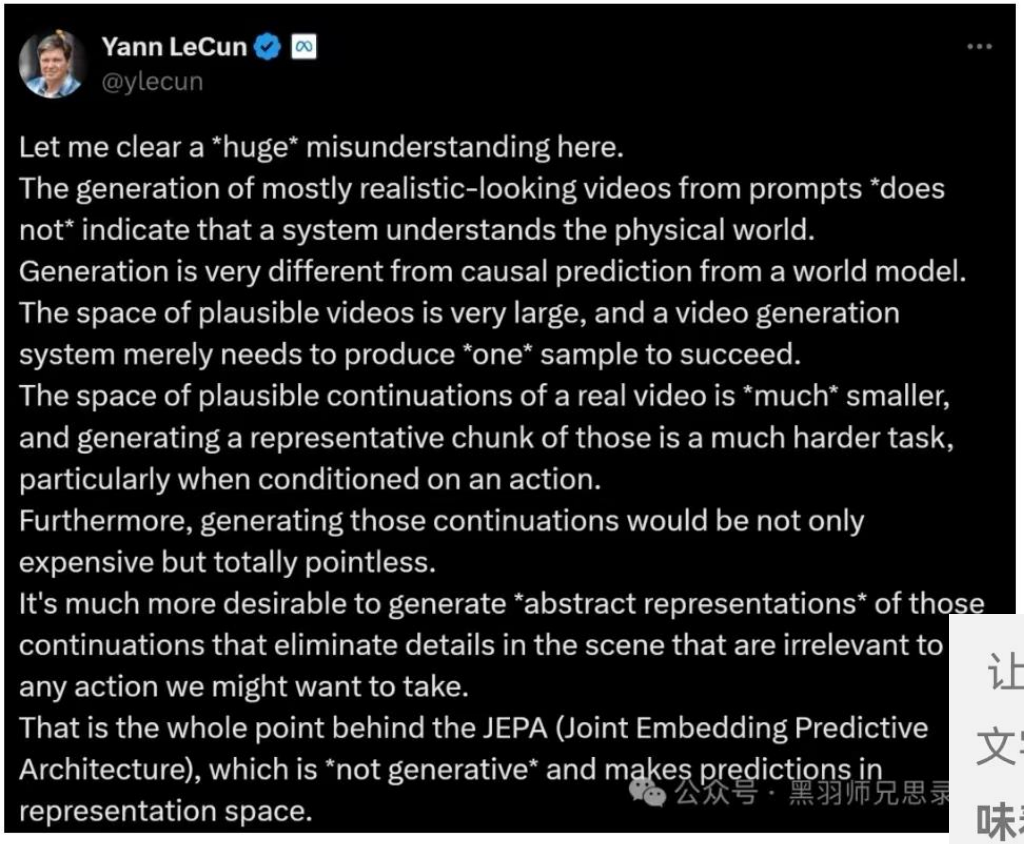
与中国龙一起庆祝农历新年的视频

五只灰狼幼崽在一条偏僻的碎石路上互相嬉戏、追逐（狼的数量莫名改变）

提示词：一位时尚女性走在充满温暖霓虹灯和动画城市标牌的东京街道上。她穿着黑色皮夹克、红色长裙和黑色靴子，拎着黑色钱包。她戴着太阳镜，涂着红色口红。她走路自信又随意。街道潮湿且反光，在彩色灯光的照射下形成镜面效果。许多行人走来走去。

缺陷：还不足以完全模拟所有现实中的物理过程

Artificial Intelligence (AI) — large language model



Yann LeCun
@ylecun

Let me clear a *huge* misunderstanding here.
The generation of mostly realistic-looking videos from prompts *does not* indicate that a system understands the physical world.
Generation is very different from causal prediction from a world model.
The space of plausible videos is very large, and a video generation system merely needs to produce *one* sample to succeed.
The space of plausible continuations of a real video is *much* smaller, and generating a representative chunk of those is a much harder task, particularly when conditioned on an action.
Furthermore, generating those continuations would be not only expensive but totally pointless.
It's much more desirable to generate *abstract representations* of those continuations that eliminate details in the scene that are irrelevant to any action we might want to take.
That is the whole point behind the JEPA (Joint Embedding Predictive Architecture), which is *not generative* and makes predictions in representation space.

让我在这里**澄清一个巨大的误解**。从提示文字，生成看起来相当逼真的视频，**并不意味着，这个系统理解物理世界**。生成一个视频，和基于世界模型的因果预测，大不相同。目前这种生成式的方向，代价高昂，可能还有更好的法子。

Outline

- Why study AI?
- What is AI?
- The foundations of AI
- A brief history
- The state of the art

Why study AI?——大背景

—>智能社会（智能化）

—>信息社会（数字化）

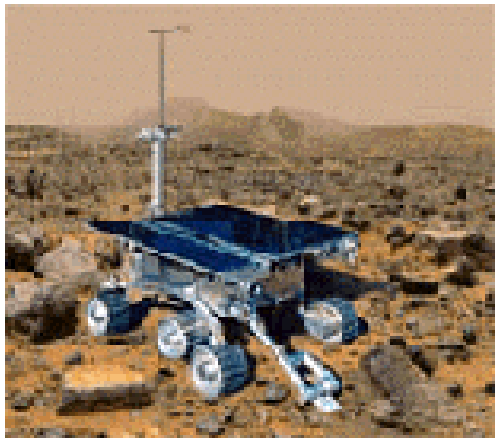
—>工业社会（机械化）

农业社会

AI是智能社会的支撑技术

Why study AI?

Labor



Science

Appliances

Google

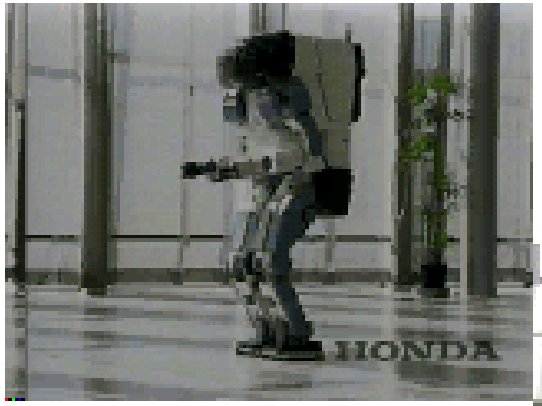
YAHOO!

Search engines

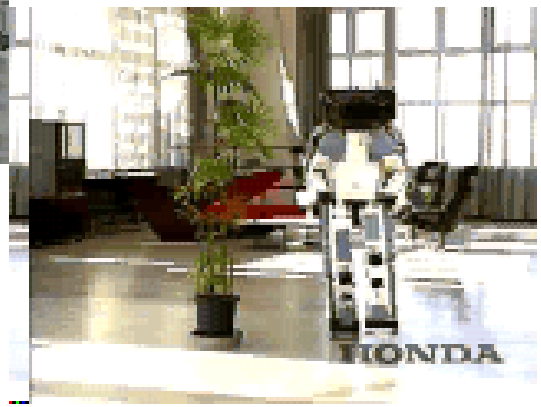
Medicine/
Diagnosis

What else?

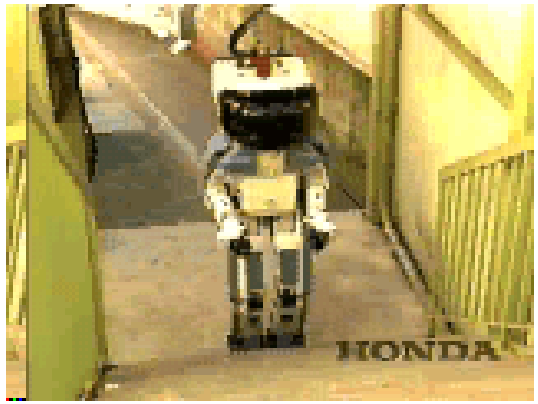
Honda Humanoid Robot(ASIMO)



Walk



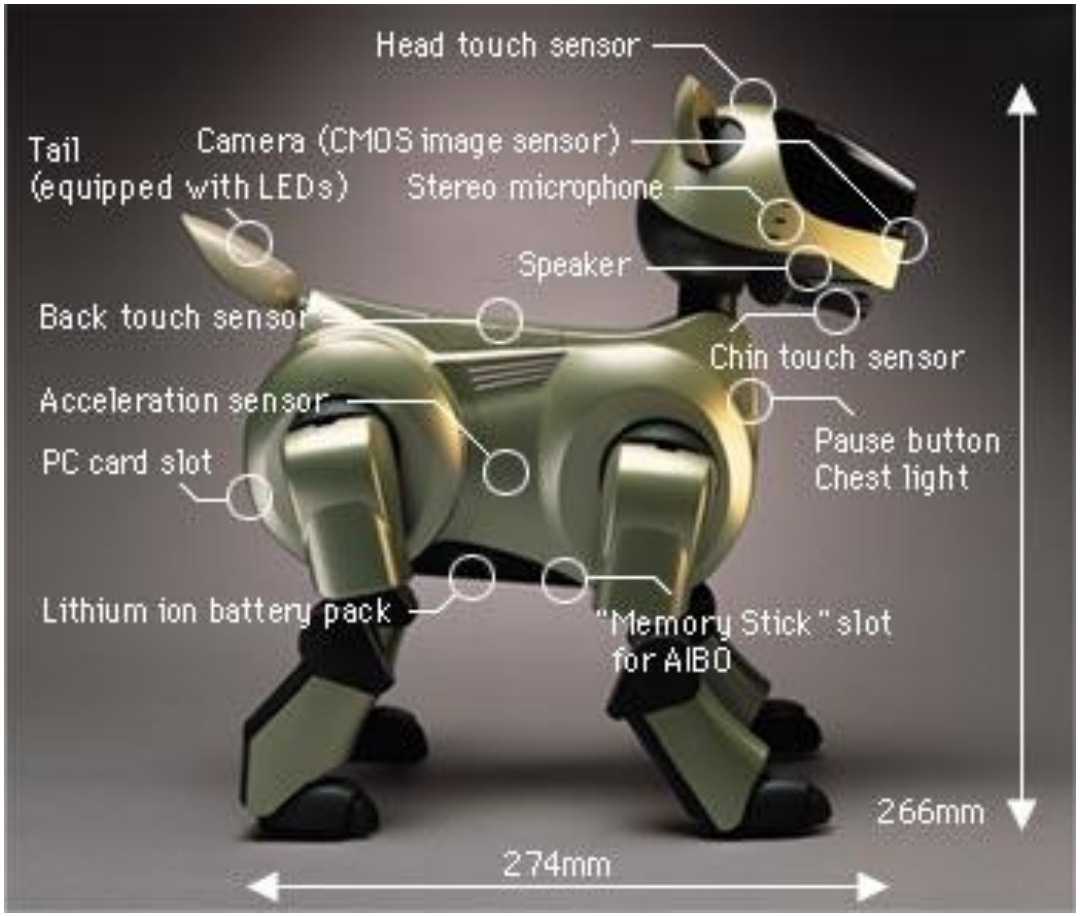
Turn



Stairs

http://world.honda.com/robot/

Sony AIBO



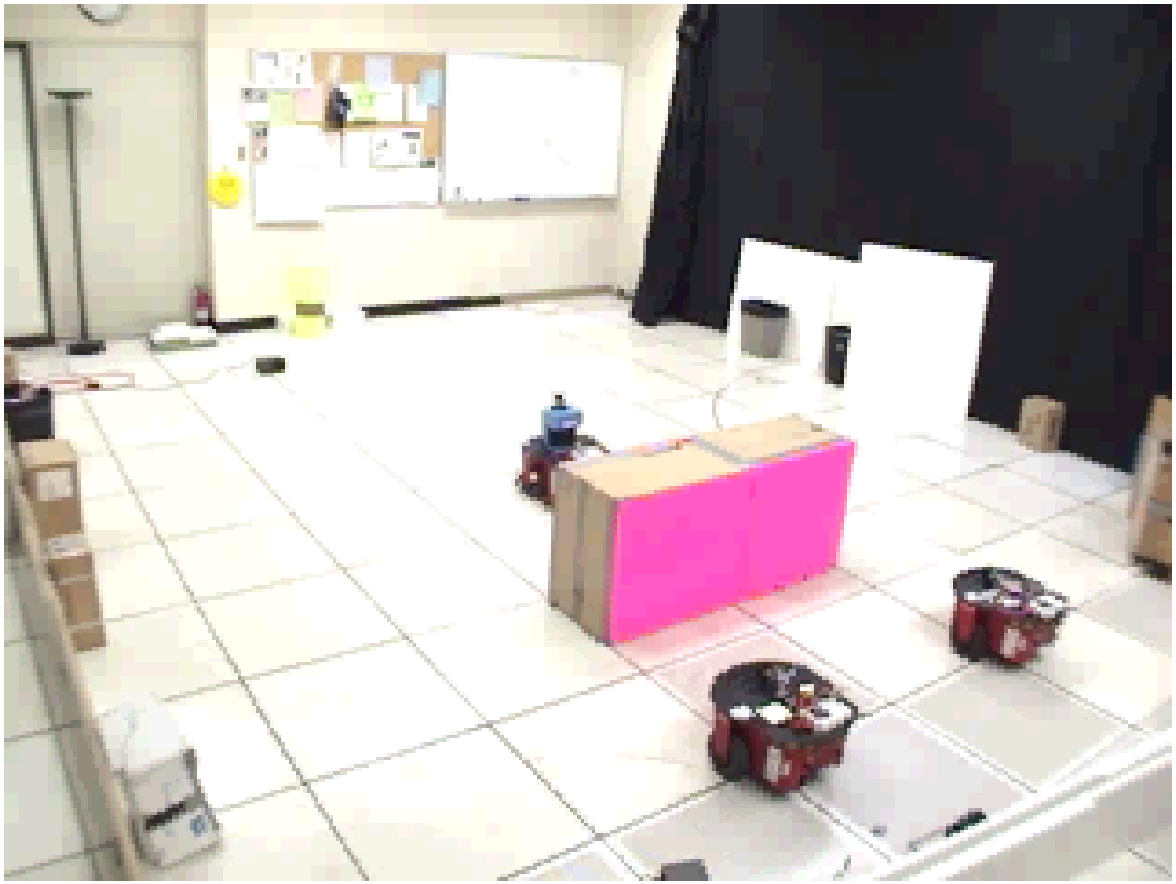
http://www.aibo.com

Natural Language Question Answering

http://aimovie.warnerbros.com

http://www.ai.mit.edu/projects/infolab/

Robot Teams



USC robotics Lab

What is AI?

- **Artificial Intelligence (AI)**
  - Intelligent behavior in artifacts
  - "Design computer programs to make computers smarter"
  - "Study of how to make computers do things at which, at the moment, people are better"
  - ……
- **Intelligent behavior**
  - Perception, reasoning, learning, communicating, acting in complex environments
- **Long term goals of AI**
  - Develop machines that do things as well as humans can or possibly even better
  - Understand intelligent behaviors

What is AI?

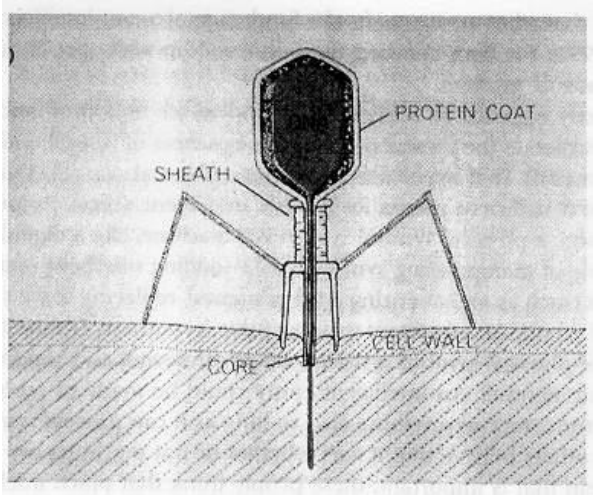
- Can machines think?
  - Depend on the definitions of “machine”, “think”, “can”
- “Can”
  - Can machines think now or someday?
  - Might machines be able to think theoretically or actually?

What is AI?

- "Machine"
  - E6 Bacteriophage:

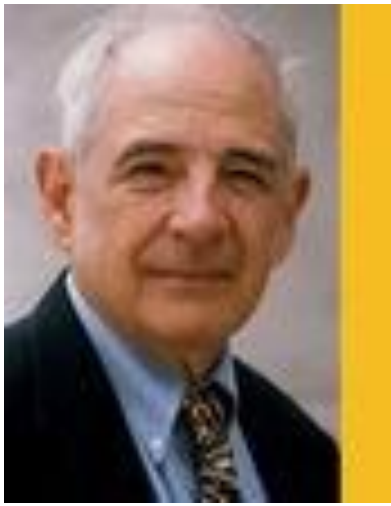
    Machine made of proteins:

    E6噬菌体:由蛋白质构成的机器



  - Searle's belief
    - What we are made of is fundamental to our intelligence
    - Thinking can occur only in very special machines – living ones made of proteins

Chinese Room



What is AI?

Views of AI fall into four categories:

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

The textbook advocates "acting rationally"

Acting humanly: Turing Test

- (Human) judge communicates with a human and a machine over text-only channel,
- Both human and machine try to act like a human,
- Judge tries to tell which is which.
- It is possible to (temporarily) fool humans who do not realize they may be talking to a bot

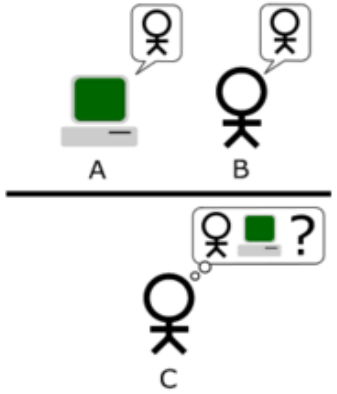
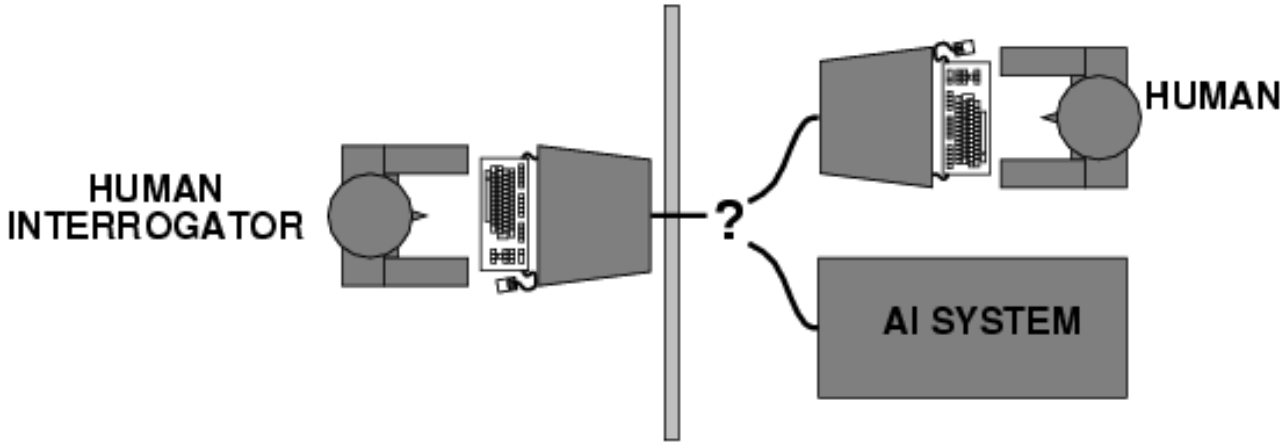


image from http://en.wikipedia.org/wiki/Turing_test

- Current programs nowhere close to passing this
  - http://www.jabberwacky.com/

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

Thinking humanly: cognitive modeling

- 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* and *rules of derivation* for thoughts; may or may not have proceeded to the idea of mechanization

- Direct line through mathematics and philosophy to modern AI

- Problems:
  1. Not all intelligent behavior is mediated by logical deliberation
  2. What is the purpose of thinking? What thoughts should I have?

Acting rationally: rational agent approach

- Rational behavior: doing the right thing

- The right thing: that 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

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:

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

- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance

- Remark: computational limitations make perfect rationality unachievable

  → design best program for given machine resources

Argument on AI—三大学派

- **Symbolism—符号主义**
- **Connectionism—连接主义**
- **Actionism—行为主义**

Symbolicism
—Logicism, Psychlogism, Computerism

AI Origins from mathematic logic.

Cognitive element — symbol.

Core of AI — Knowledge and Knowledge-based theoretical system.

It is based on hypothesis of symbol operation system and principle of limit reasonability.

Representatives: Wiener，Newell，Shaw， Simon and Nilsson

Connectionism —Bionicsism, Physiologism

AI Origins from Bionics.

Thinking element- neuron

Brain working mode

It based on NN and connectionism and learning algorithm between NN.

Representatives: Meculloch-Pitts, Hopfield and Rumelhart

Actionism

—Evolutionism, Cyberneticsism

AI Origins from Cybernetics.

Intelligence depends on perception and actions.

Perception-action working mode

Evolution and interaction

It based on Cybernetics and control principle of perception action.

Representatives: Winner, Brooks

The Foundation of AI

➢ 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(运动神经） control,experimental techniques
• Computer engineering	Building fast computers
• Control theory	Design systems that maximize an objective function over time
• Linguistics	Knowledge representation, grammar

Foundations - Philosophy

- Aristotle (384-322 B.C.) – Author of logical syllogisms
- da Vinci (1452-1519) – designed, but didn't build, first mechanical calculator
- Descartes (1596-1650) – can human free will be captured by a machine? Is animal behavior more mechanistic?
- Bacon(1561-1626)
- Necessary connection between logic and action is discovered

Foundations - Mathematics

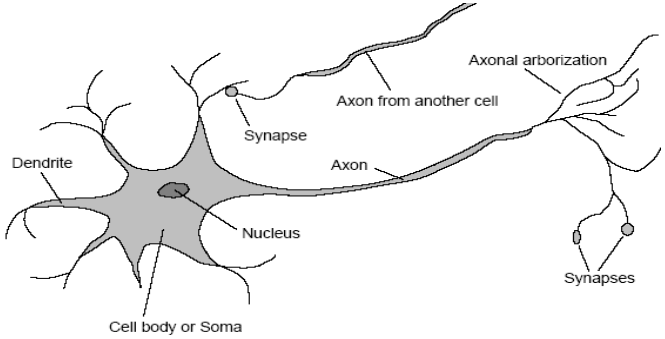
- More formal logical methods
  - Boolean logic (Boole, 1847)
  - First-order logic (Frege, 1879)
- Analysis of limits to what can be computed
  - Intractability 难解性 (1965) – time required to solve problem scales exponentially with the size of problem instance
  - NP-complete (1971) – Formal classification of problems as intractable
- Uncertainty (Cardano 1501)
  - The basis for most modern approaches to AI
  - Uncertainty can still be used in logical analyses

Foundations - Economics

- Humans are peculiar so define generic happiness term: utility
- Game Theory – study of rational behavior in small games
- Operations Research – study of rational behavior in complex systems
- Herbert Simon (1916 – 2001) – AI researcher who received Nobel Prize in Economics(1978) for showing people accomplish satisficing solutions, those that are good enough

Foundations - Neuroscience

- How do brains work?



  - Early studies (1824) relied on injured and abnormal people to understand what parts of brain do
  - More recent studies use accurate sensors to correlate brain activity to human thought (EEG, fMRI)
    - By monitoring individual neurons, monkeys can now control a computer mouse using thought alone

Foundations - Neuroscience

Figure 1.2

	Supercomputer	Personal Computer	Human Brain
Computational units	10^6 GPUs + CPUs	8 CPU cores	10^6 columns
	10^{15} transistors	10^{10} transistors	10^{11} neurons
Storage units	10^{16} bytes RAM	10^{10} bytes RAM	10^{11} neurons
	10^{17} bytes disk	10^{12} bytes disk	10^{14} synapses
Cycle time	10^{-9} sec	10^{-9} sec	10^{-3} sec
Operations/sec	10^{18}	10^{10}	10^{17}

A crude comparison of a leading supercomputer, Summit (Feldman, 2017); a typical personal computer of 2019; and the human brain. Human brain power has not changed much in thousands of years, whereas supercomputers have improved from megaFLOPs in the 1960s to gigaFLOPs in the 1980s, teraFLOPs in the 1990s, petaFLOPs in 2008, and exaFLOPs in 2018 (1 exaFLOP = 10^{18} floating point operations per second).

Computer Sci. and Brain Sci.

Information Processing in Digital Computer

Computing based on Logic

CPU and Storage: Separated

Data Processing & Storage: Simple

Intelligent Information Processing: Complicated and Slow

Cognitive capability: Weak

Information Process Mode:
Logic – Information – Statistics

Information Processing in the Brain

Computing based on Statistics

CPU and Storage: Unified

Data Processing & Storage: Unknown

Intelligent Information Processing: Simple and Fast

Cognitive capability: Strong

Information Process Mode:
Statistics —concepts—logic

Foundations - Psychology

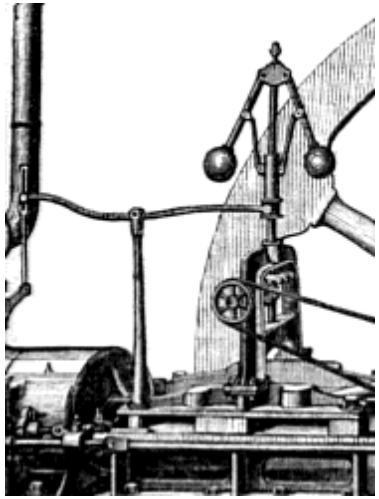
- Helmholtz and Wundt (1821) – started to make psychology a science by carefully controlling experiments
- The brain processes information (1842)
  - stimulus converted into mental representation
  - cognitive processes manipulate representation to build new representations
  - new representations are used to generate actions
- Cognitive science started at a MIT workshop in 1956 with the publication three very influential papers

Foundations – Computer engineering

- How can we build an efficient computer?
  - For artificial intelligence to succeed, we need two things:
    - intelligence
    - artifact
  - The computer has been the artifact of choice

Foundations – Control Theory

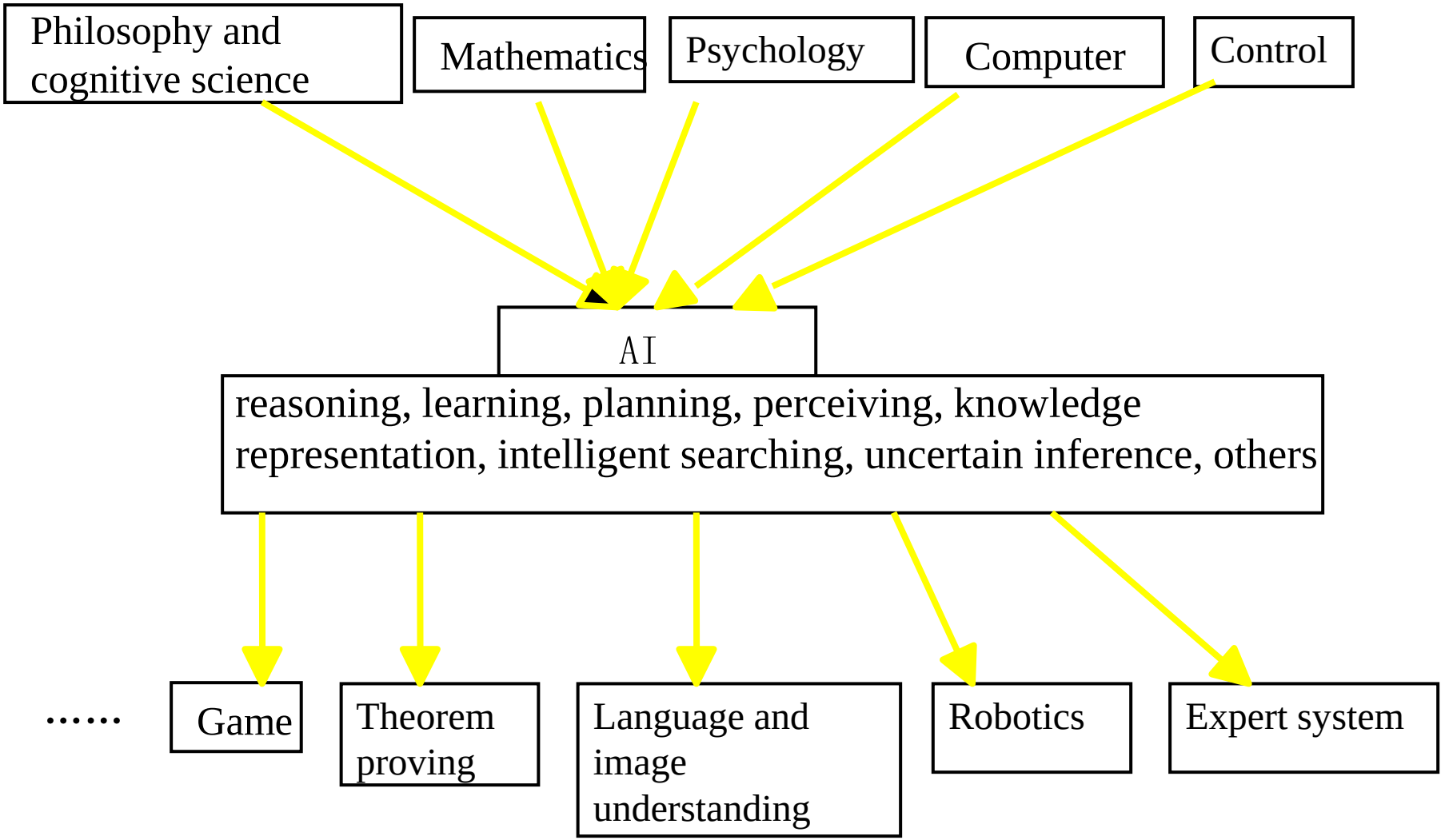
- Machines can modify their behavior in response to the environment (sense / action loop)
  - Water-flow regulator (250 B.C.E), steam engine governor, thermostat
- The theory of stable feedback systems (1894)
  - Build systems that transition from initial state to goal state with minimum energy
  - In 1950, control theory could only describe linear systems and AI largely rose as a response to this shortcoming



Foundations - Linguistics

- How does language relate to thought?
- Speech demonstrates so much of human intelligence
  - Analysis of human language reveals thought taking place in ways not understood in other settings
    - Children can create sentences they have never heard before
    - Language and thought are believed to be tightly intertwined
- A hybrid field: natural language processing

AI and related disciplines



History of AI

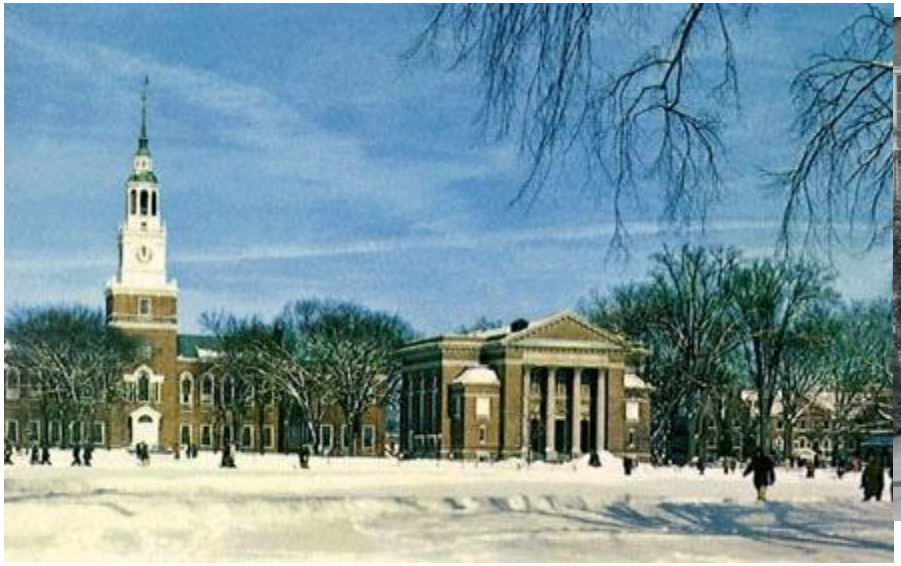
- The gestation of AI (1943-1955)
- The birth of AI (1956)
- Early enthusiasm, great expectations (1952-1969)
- A dose of reality (1966-1973)
- Knowledge-based system (1969-1979)
- AI becomes an industry (1980-present)
- The return of neural network (1986-present)
- AI adopts the science method (1987-present)
- The emergence of intelligent agents (1995-present)
- The availability of very large data sets (2001-present)
- Deep Learning (2006，2011-present)

History of AI: 1943- 1955

- The gestation of AI
  - Warren McCulloch and Walter Pitts proposed a model of artificial neurons. (1943)
  - ENIAC (1946)
  - Hebbian rule. (1949)
  - Marvien Minsky and Dean Edmonds built the first neural network computer. (1950)
  - Turing's "Computing Machinery and Intelligence" (1950)

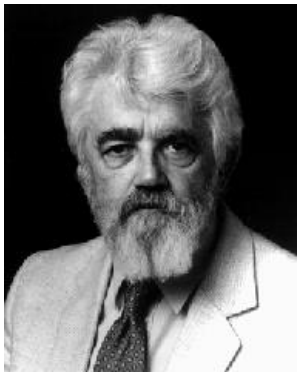
History of AI: 1956

- Dartmouth meeting: "Artificial Intelligence" adopted
  - AI become a separate field
  - Summer of 1956
  - John McCarthy, Minsky, Claude Shannon, and Nathaniel Rochester
  - For the next 20 years, the field would be dominated by these people and their students and colleagues at MIT, CMU, Stanford, and IBM.



History of AI: 1952- 1969 (1)

- Great successes!
  - Logic programs were replicating human logic in many cases
    - Solving hard math problems
    - game playing
  - LISP was invented by McCarthy (1958)
    - second oldest language in existence
    - could accept new axioms at runtime



  - McCarthy went to MIT and Marvin Minsky started lab at Stanford

    Both powerhouses in AI to this day
  - Early AI programs (1950s-1960s): Samuel's checkers program, Newell & Simon's Logic Theorist, GPS ,Gelernter's Geometry Engine

History of AI: 1952- 1969 (2)

- Great successes!
  - Robinson's complete algorithm for logical reasoning (1965)
  - Early artificial neural network: Adalines(1962), Perceptron(1962), perceptron convergence theorem(1962)

History of AI: 1966 - 1973

- A dose of reality (AI discovers computational complexity
  Neural network research almost disappears)
  - Systems fail to play chess and translate Russian
    - Computers were ignorant to context of their logic
    - Problems were intractable
      - algorithms that work in principle may not work in practice
      - Combinatorial Explosion / Curse of Dimensionality
    - Fatal flaw in neural networks was exposed
      - though flaw was first resolved in 1969, neural networks did not return to vogue until late 1980s

机器翻译闹出的笑话举例：

- “***The spirit is willing but the flesh is weak***”，意思是“心有余而力不足”。
  - 机器翻译过程：英语—> 俄语 —> 英语
  - 结果被译为：“***The vodka is good but the meat is spoiled***”，意思是“伏特加是好的，肉变质了”。

- “***Out of sight, out of mind***”，意思是“眼不见心不烦”。
  - 将其翻译成俄语，竟成了：“ 又瞎又疯”。
  - ……

AI History: 1969 - 1979

- Early development of knowledge-based Systems
  - Previous systems knocked because general logical algorithms could not be applied to realistic problems
  - Answer: accumulate specific logical algorithms
    - DENDRAL – infer chemical structure
    - MYCIN, PROSPECTOR
    - Natural language understanding
    - Knowledge representation schemes: Prolog, frame
  - Researchers work on ways to accumulate and store facts for expert systems

AI History: 1980 - present

- AI becomes an industry
  - The demonstrated success of AI invited investments
  - from millions to billions of dollars in 10 years
  - Japan “Fifth Generation” project (1981)
  - extravagant AI promises again led to “AI Winter” when investments in technology dropped (1988)
- Neural networks return and became popular (1986)
  - Hopfield
  - Back propagation

AI History: 1987 - present

- AI becomes a science
  - More repeatability of experiments
  - More development of mathematical underpinnings
  - Show relevance to real-world applications rather than toy examples

AI History: 1995 - present

- Intelligent Agents (1995)
  - AI systems exist in real environments with real sensory inputs

AI History: 2001 - present

- The availability of very large data sets (2001-)
  - Learning algorithm
  - Very large data sets

AI History: 2006 - present

- Deep learning(2006, 2011-present)
  - 2006：Hinton在 Science 发文提出深度神经网络训练中梯度消失的解决方案：无监督预训练+有监督微调,…
  - 2011：ReLu激活函数被提出，有效的抑制梯度消失问题,…
  - 2012：在图像识别领域取得惊人效果，ImageNet错误率从26%降到15%,…
  - 2016：AlphaGo,…
  - 2017：Transformer和注意力机制,…
  - 2018：GPT,…
  - 2022：ChatGPT,…
  - 2024：Sora,…

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 and became popular
- 1987-- AI becomes a science
- 1995-- The emergence of intelligent agents
- 2001-- The availability of very large data sets
- 2006-- The emergence of deep learning

Homework

- To write a short report on your personal interests in the field of AI.