

人工智能概述

主讲：王亚星、刘夏雷、郭春乐
南开大学计算机学院

致谢：本课件主要内容来自浙江大学吴飞教授、
南开大学程明明教授

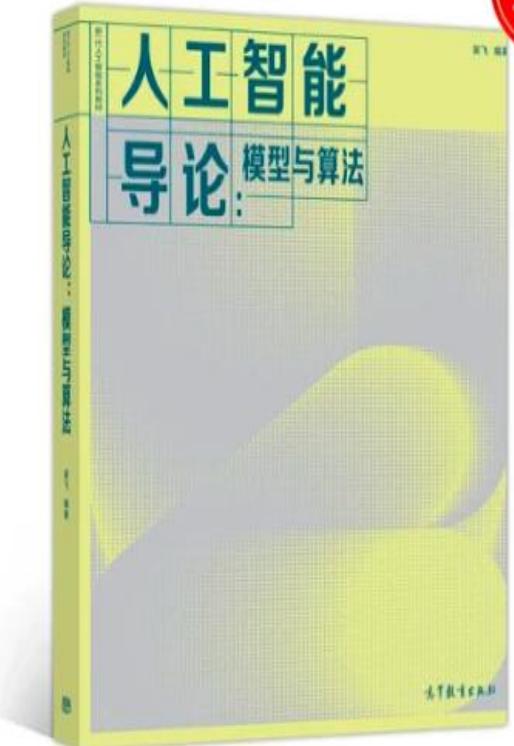


课程信息

- 教师
 - 程明明 cmm@nankai.edu.cn
 - 王亚星 yaxing@nankai.edu.cn
 - 刘夏雷 xialei@nankai.edu.cn
 - 郭春乐 guochunle@nankai.edu.cn
- 助教
 - 胡泰航
 - 张鑫
- 地点：津南公教楼**B区510**，津南实验楼**A区306&308**
- 时间：1-17周讲授，3-17周实训
- 成绩：课上随堂测试考核(10%)、研讨内容(10%)、实验内容考核(40%)和期末考试(40%)



官方
正版



重要提示

- 我们计划采用浙大的在线平台及算力 (<https://mo.zju.edu.cn/>) 进行课程实验，我们已经给大家注册了账号。

The screenshot shows the ZJU MOOCs platform interface. At the top, there is a blue header bar with the text '课程 / Python 基础'. Below the header, the course title 'Python 基础' is displayed along with a '17 章节' badge and a '加入课程' button. The main content area is titled '课程目录' and contains a list of eight lessons:

- 第一课 01 做计算
- 第二课 02 循环计算
- 第三课 03 字符串
- 第四课 04 列表
- 第五课 05 在数列上的计算
- 第六课 06 使用列表的计算
- 第七课 07 元组、集合与字典
- 第八课 08 使用容器的计算

前四周为大家安排了Python基础学习课程，两周自学，两周实训学习。
大家尽可能多学习基础知识，对后续实训题目帮助很大。

提纲

- 人工智能历史
- 人工智能课程体系
- 课程内容概要
- 趋势与发展
- 实训题目安排

荀子的智能观：《荀子·正名》

知之在人者 谓之 知	知觉 ：人所固有认识外界客观事物本能，如视觉、听觉和触觉等能力
知有所合 谓之 智	智慧 ：知觉对外界事物的认知
所以能之在人 者为之 能	本能 ：人身上所具用来处置事物能力
能有所合 谓之 能	智能 ：对外界所产生的认知和决策



荀子

战国末期赵国人
(约公元前313年 - 公元前238年)

从感知、到理解、到认知、到决策与行动

计算的诞生：从手工计算到自动计算



结绳计数

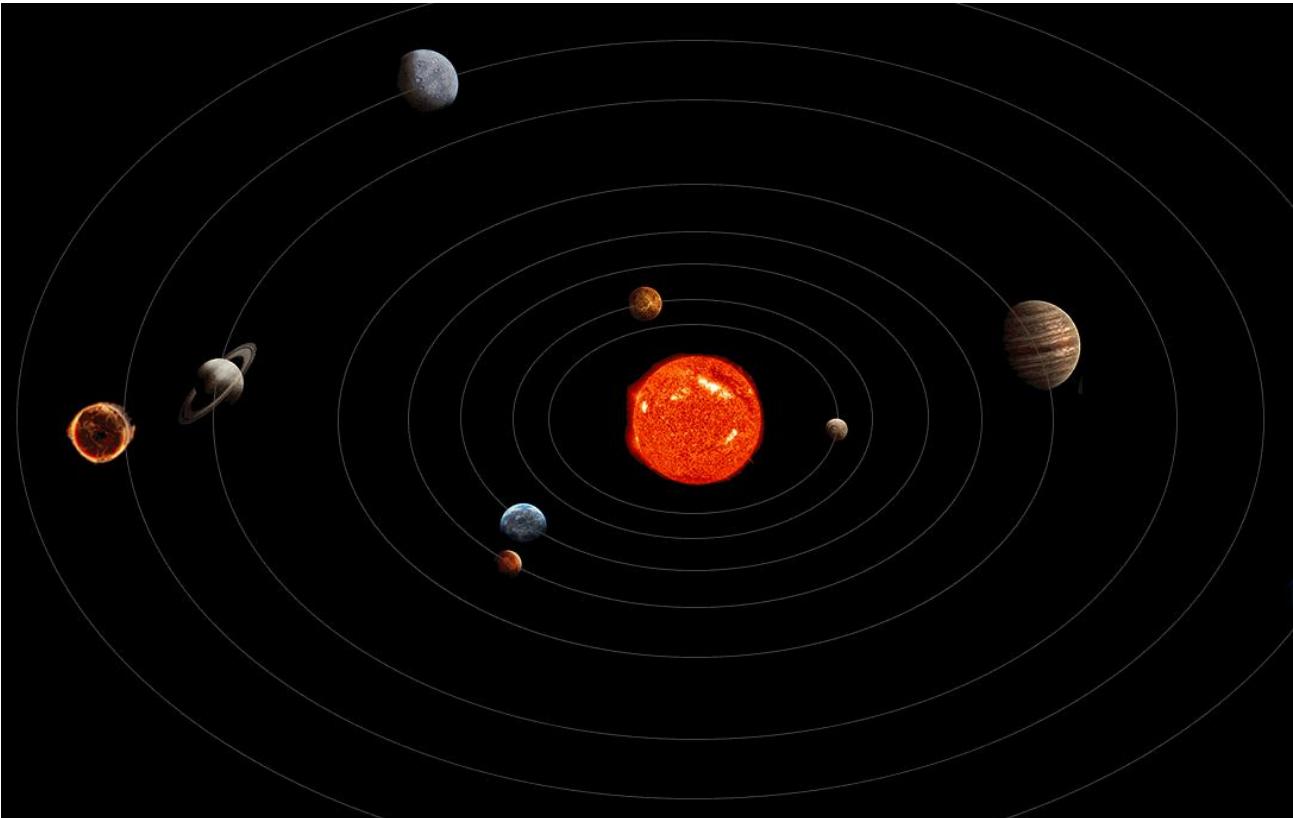


《清明上河图》中赵太丞家药铺中的算盘

上古结绳而治，后世圣人易之以书契，百官以治，万民以察。
——
《周易·系辞》

计算的诞生：从可计算到不可计算

- 20世纪初，人们发现有许多问题无法找到解决的方法。于是开始怀疑，是否对这些问题来说，根本就不存在算法，即不可计算。



计算的诞生：从可计算到不可计算

当整数 $n > 2$ 时，关于 x, y, z 的方程 $x^n + y^n = z^n$ 没有正整数解

• 从费马猜想到费马定理



费马
(Pierre de Fermat)
法国人
律师，业余数学家
(1601-1665)



1637年，费马在书本空白处提出费马猜想。

1770年，欧拉证明 $n=3$ 时定理成立。

1823年，勒让德证明 $n=5$ 时定理成立。

1832年，狄利克雷试图证明 $n=7$ 失败，但证明 $n=14$ 时定理成立。

1839年，拉梅证明 $n=7$ 时定理成立。

1850年，库默尔证明 $2 < n < 100$ 时除 37、59、67 三数外定理成立。

1955年，范迪维尔以电脑计算证明了 $2 < n < 4002$ 时定理成立。

1976年，瓦格斯塔夫以电脑计算证明 $2 < n < 125000$ 时定理成立。

1985年，罗瑟以电脑计算证明 $2 < n < 4100000$ 时定理成立。

1987年，格朗维尔以电脑计算证明了 $2 < n < 10^{180000}$ 时定理成立。

1995年，怀尔斯证明 $n > 2$ 时定理成立。

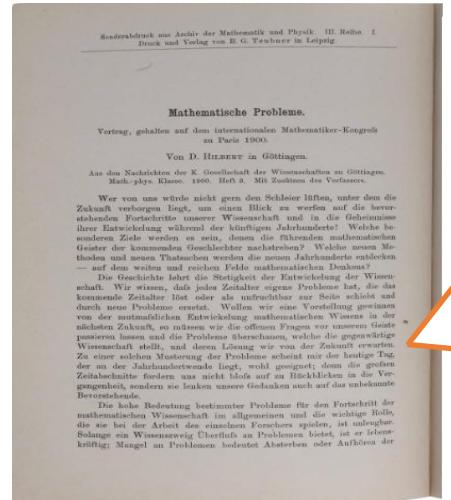
计算之力从何而来：以机器为载体的智能

• 工欲善其事，必先利其器：智能之器从何而来



David Hilbert

(1862-1943)
德国著名数学家



Mathematical Problems
(23个数学问题)

Lecture delivered before the International Congress
of Mathematicians at Paris in 1900

问题2：证明算术公理的相容性

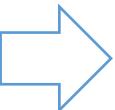
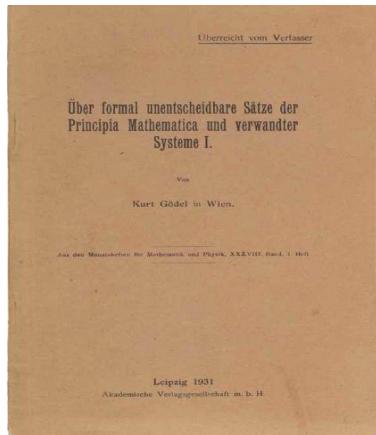
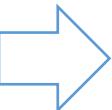
(the compatibility of the arithmetical axioms)

- **完备性**：所有能够从该形式化系统推导出来的命题，都可以从这个形式化系统推导出来。
- **一致性**：一个命题不可能同时为真或为假
- **可判定性**：算法在有限步内判定命题的真伪

如何证明（计算）这一问题？可用怎样的“计算载体”来实现？

计算之力从何而来

• 工欲善其事，必先利其器：智能之器从何而来



Kurt Friedrich Gödel
(1906-1978)

《论数学原理及有关系统
中不可判定命题》
(On Formally Undecidable Propositions
of Principia Mathematica and Related
Systems)

任何表达力足够强的(递归可枚举)形式系统都不可能同时具有一致性和完备性，而且这个系统本身的一致性不能在系统内被证明。

哥德尔不完全性定理
(1931年)

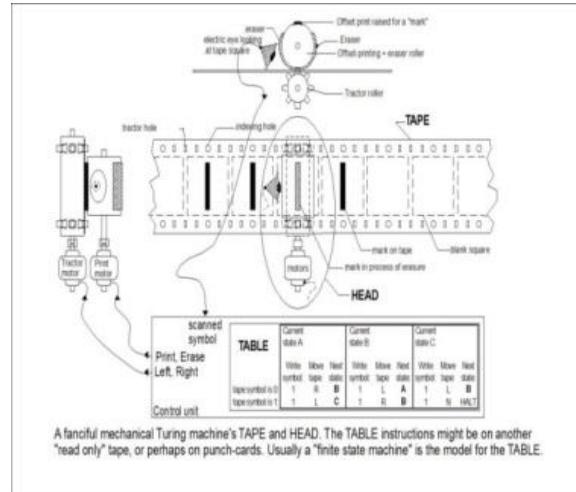
计算之力从何而来

• 工欲善其事，必先利其器：智能之器从何而来

图灵机模型



Alan
Turing
(1912-1954)



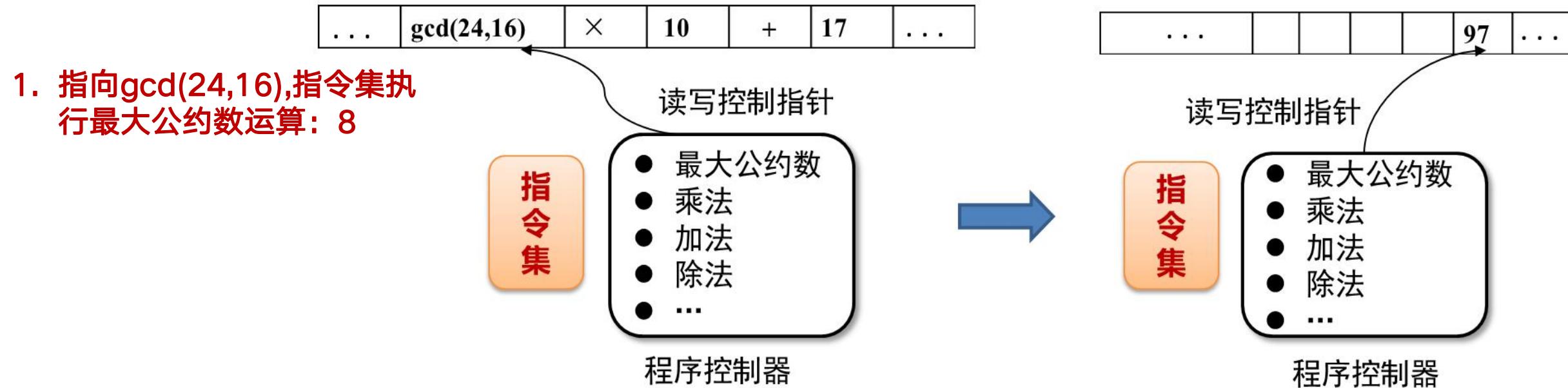
1. 方格（数字或者信息）
2. 控制指针
3. 指令集

“判定性问题”是无法解决的，即有些数学问题是不可求解的（图灵不可停机的），比如费马猜想

On Computable Numbers, with an Application to the Entscheidungsproblem, 1937
《论数字计算在决断难题中的应用》

计算之力从何而来

• 工欲善其事，必先利其器：智能之器从何而来



图灵机模型计算“ $\text{gcd}(24,16)*10+17$ ”示意

计算之力从何而来

• 工欲善其事，必先利其器：智能之器从何而来

计算载体	提出学者	计算角度
原始递归函数	哥德尔 Godel	数学的形式
λ - 演算 (λ - calculus)	丘奇 Church	数理逻辑的形式
图灵机	图灵 Turing	机械的形式

- 图灵论题：凡是可计算的函数都可以用图灵机计算
- 邱奇论题：任何计算，如果存在一个有效过程，它就能被图灵机实现

邱奇 - 图灵论题

(Church–Turing thesis,
computability thesis)

图灵机模型使得人类得以迈入自动计算时代

- 图灵奖：计算机界最高奖（1966年设立）

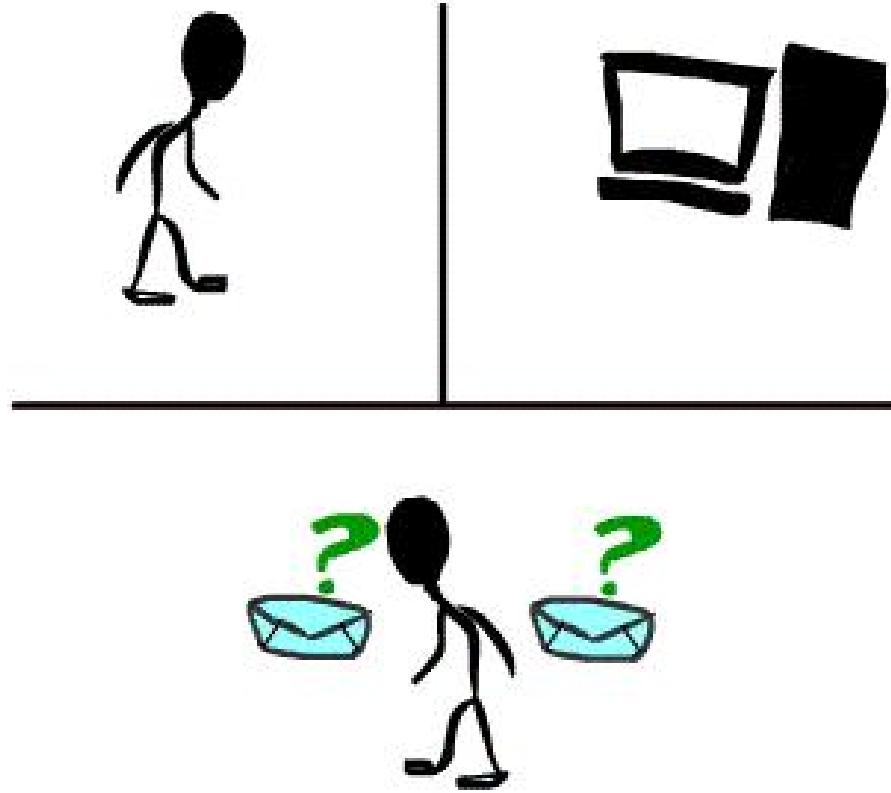


Alan Turing
(1912-1954)



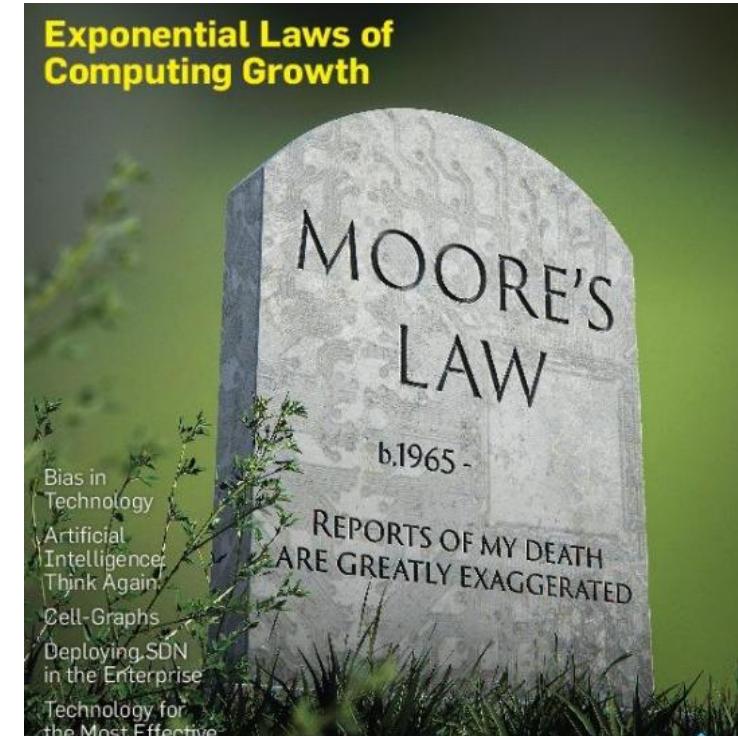
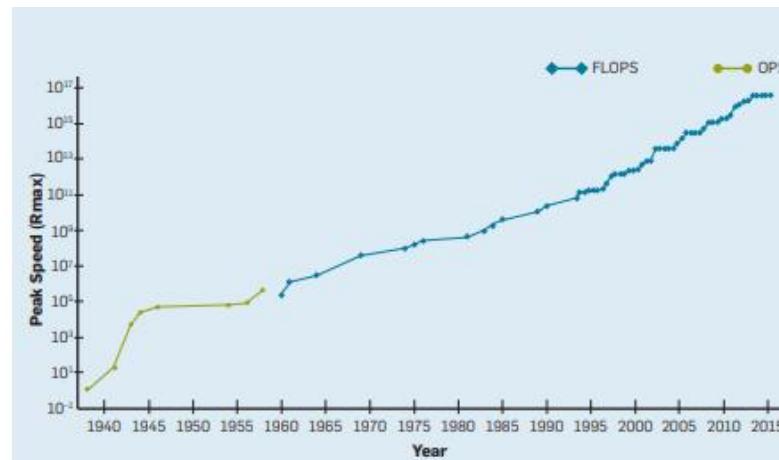
二战提前五年结束
的解码器Enigma

对智能的测试：图灵测试



摩尔定律支撑自动计算时代快速发展

- 摩尔定律（计算机速度1年半增长1倍）



人工智能的诞生

- 四位学者在1955年提出了人工智能这一术语及研究范畴
 - 麦卡锡 (John McCarthy, 时任Dartmouth数学系助理教授, 1971年度图灵奖获得者)
 - 明斯基 (Marvin Lee Minsky, 时任哈佛大学数学系和神经学系Junior Fellow, 1969年度图灵奖获得者)
 - 香农 (Claude Shannon, Bell Lab, 信息理论之父)
 - 罗契斯特 (Nathaniel Rochester, IBM, 第一代通用计算机701主设计师)

让机器能像人那样认知、思考和学习，即用计算机模拟人的智能。

人工智能的诞生

- 四位学者在1955年提出了人工智能这一术语及研究范畴

A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence

August 31, 1955

John McCarthy, Marvin L. Minsky,
Nathaniel Rochester,
and Claude E. Shannon

The 1956 Dartmouth summer research project on artificial intelligence was initiated by the August 31, 1955 proposal, authors by John McCarthy, Marvin Minsky, Nathaniel Rochester and Claude E. Shannon. The original typescript consisted of 17 pages plus a title page. Copies of the typescript are being made available at the Computer Science Department Archives at the Computer Science Department, Stanford University. The first 5 pages state the proposal, and the remaining pages give qualifications and interests of the four who proposed the study. In the interests of brevity, this article reproduces only the proposal itself, along with the short autobiographical statements of the proposers.

guage, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer. The following are some aspects of the artificial intelligence problem.

1. Automatic Computers
If a machine can do a job, then an automatic calculator can be programmed to simulate the machine. The speeds and memory capacities of present computers may be insufficient to simulate many of the higher functions of the human brain, but the major obstacle is not lack



当年提出AI概念的建议书

AI概念提出50年后，建议人合影

Research Project on Artificial Intelligence, August 31, 1955, Dartmouth

让机器能像人那样认知、思考和学习，即用计算机模拟人的智能。

人工智能的诞生

- 四位学者在1955年提出了人工智能这一术语及研究范畴

A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence

August 31, 1955

John McCarthy, Marvin L. Minsky,
Nathaniel Rochester,
and Claude E. Shannon

The 1956 Dartmouth summer research project on artificial intelligence was initiated by the August 31, 1955 proposal, authors by John McCarthy, Marvin Minsky, Nathaniel Rochester and Claude E. Shannon. The original typescript consisted of 17 pages plus a title page. Copies of the typescript are being made available at the Computer Science Departmental Archives at the Stanford University. The first 5 pages state the proposal, and the remaining pages give qualifications and interests of the four who proposed the study. In the interests of brevity, this article reproduces only the proposal itself, along with the short autobiographical statements of the proposers.

guage, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer. The following are some aspects of the artificial intelligence problem.
1. Automatic Computers
If a machine can do a job, then an automatic calculator can be programmed to simulate the machine. The speeds and memory capacities of present computers may be insufficient to simulate many of the higher functions of the human brain, but the major obstacle is not lack

当年提出AI概念的建议书



AI概念提出50年后，建议人合影

人工智能（Artificial Intelligence）是以机器为载体所展示的人类智能，因此人工智能也被称为机器智能（Machine Intelligence）

人工智能的诞生

- 会议报告列举了**Artificial Intelligence**值得关注七个问题
 - Automatic Computers : 自动计算机
 - How Can a Computer be Programmed to Use a Language : 计算机编程
 - Neuron Nets : 神经网络
 - Theory of the Size of a Calculation : 复杂性
 - Self-improvement: 自我学习与提高
 - Abstractions: 归纳与演绎
 - Randomness and Creativity: 随机性和创造力

从典型应用对人工智能分类

- 机器定理证明（逻辑和推理），仿解题者
- 机器翻译（自然语言理解），仿译者
- 专家系统（问题求解和知识表达），仿专家（如医生）
- 博弈（搜索树），仿弈者
- 模式识别（多媒体认知），仿认识者
- 学习（神经网络），仿初学者
- 机器人和智能控制（感知与控制），仿生物者

符号学派、链接学派、行为学派

从智能角度对人工智能的分类

- 领域人工智能

- 依葫芦画瓢、任务导向，
如Deep Blue和AlphaGo

- 通用(或跨领域)人工智能

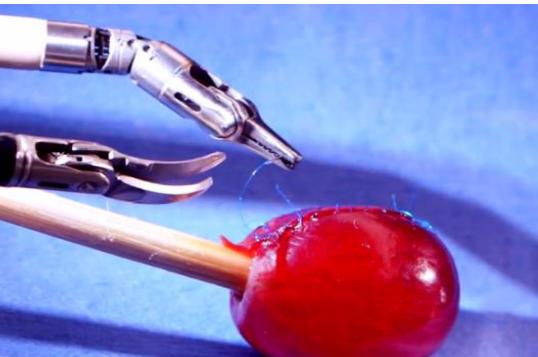
- 举一反三、从经验中学习，
如“人类”智能

Traditional AI	Artificial General Intelligence
Focus on having knowledge and skills	Focus on acquiring knowledge and skills
Action acquiring via programing	Ability acquiring via learning
domain-specific ability via rule-based and exemplar-based	general ability via abstraction (intuition) and context (common sense)
Learning by data and rules 数据与规则驱动	Learning to learn 学会学习

从智能角度对人工智能的分类

- 混合增强人工智能：多种智能体的混合形式

- 外骨骼机器人
- 人类智能+机器智能
 - 如达芬奇外科手术机器人
- 人、机、物、网互联
 - 如智慧城市系统



Da Vinci手术机器人及其缝制葡萄皮



电子设备将大脑与腿部神经元
“重新连接”

人工智能螺旋式发展之路：三次低谷

- 第一次低谷：1973年英国发表**James Lighthill** 报告
 - 该报告评判AI基础研究中**A自动机、B机器人和C中央神经系统**。
 - 报告得出结论：A和C的研究有价值，但进展令人失望。B的研究没有价值，进展非常令人失望。建议取消B的研究。
 - 批评后，AI开始了严冬（AI Winter）

(Sir James Lighthill, *Artificial Intelligence: A General Survey*, Science Research Council, 1973)

教训：AI尚属婴儿期，难以测算准确

人工智能螺旋式发展之路：三次低谷

- 第二次低谷：日本智能（第五代）计算机研制失败
 - 1982年开始，日本通产省主持第五代计算机。
 - 动机：计算机从计算与存储数据向能直接推理与知识处理的新型结构过渡。
 - 目标：构成一个具有1000个处理单元的并行推理机，推理速度比常规高1000倍。连接10亿信息组的数据库和知识库，具备听说能力。
 - 1992年因失败而告终（开支US \$850 million）

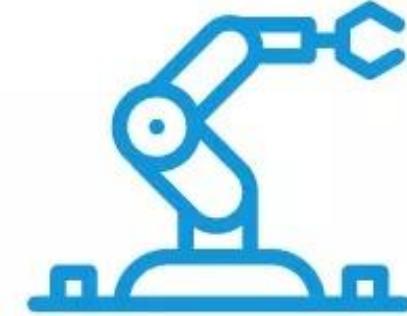
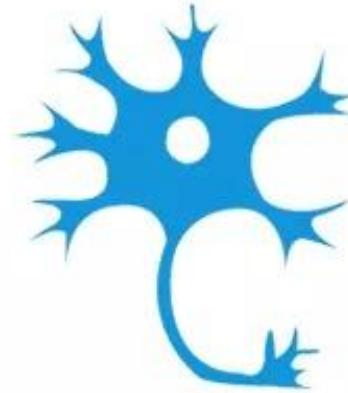
教训：驱动AI的发展要靠软件、数据和知识，而非只依靠硬件

人工智能螺旋式发展之路：三次低谷

- 第三次低谷：知识词典日趋势微、网络百科兴起
 - 斯坦福1984年通过专家来建设知识百科全书Cyc。截止2015年11月（其建设也时断时续），包括23万多个概念、实体和200多万个三元组，如：place: ~19,000 organization: ~26,000 business related thing: ~28,000
 - Cyc在90年代后期衰败，因搜索引擎崛起，显示互联网威力。
 - Cyc也开始链接外部知识库：Dbpedia, UMBEL, CIA World Factbook等等。

教训：知识不能靠专家表达，要自动学习

从典型应用对人工智能分类



符号学派、联结学派、行为学派

人工智能主流方法：符号主义人工智能

- 符号主义人工智能(**Symbolic AI**)为核心的逻辑推理

物理符号系统(即符号操作系统)假设和有限合理性原理,关系和属性符号化。

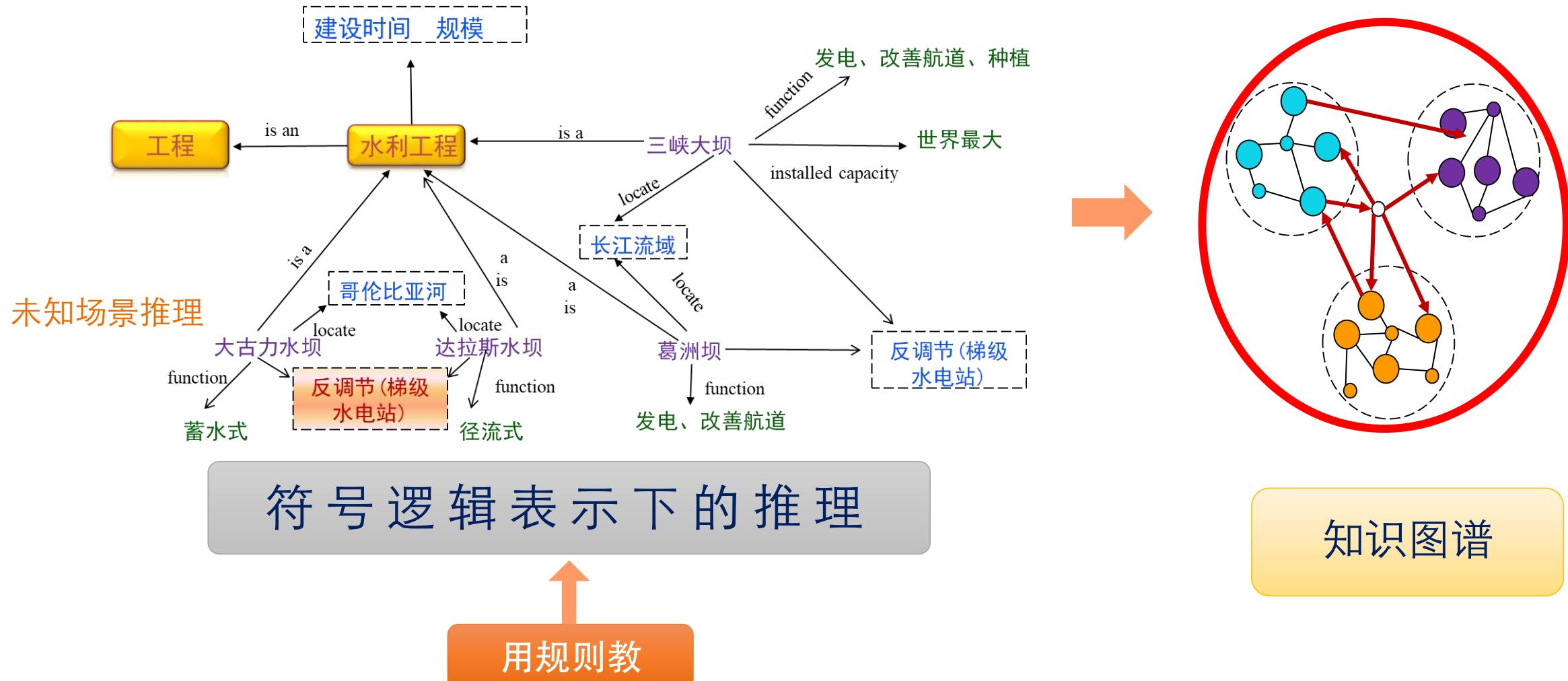
符号逻辑表示下的推理



用规则教

人工智能主流方法：符号主义人工智能

- 符号主义人工智能(Symbolic AI)为核心的逻辑推理



人工智能主流方法：符号主义人工智能

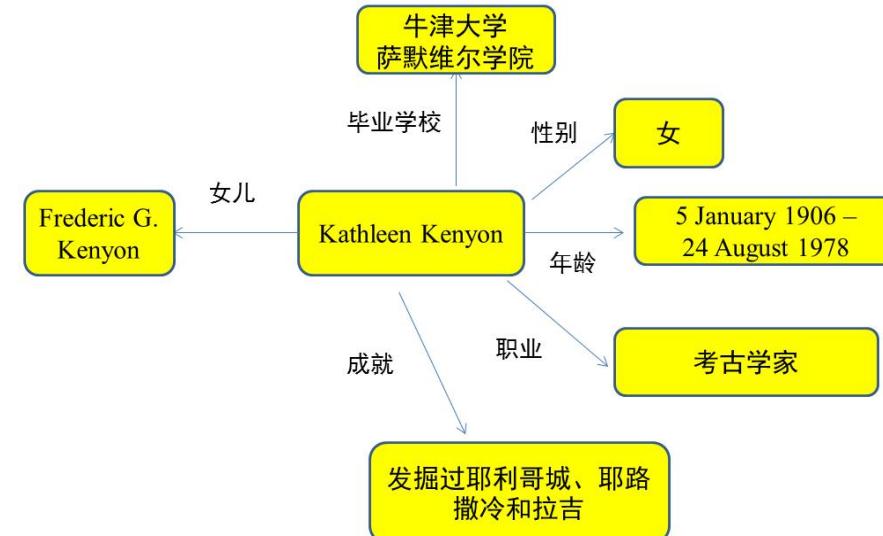
• 符号主义人工智能(Symbolic AI)为核心的逻辑推理

例子：IBM “沃森” 的推理

主持人问：**Kathleen Kenyon's**
excavation of this city mentioned in
Joshua showed the walls had been
repaired 17 times. (Kathleen Kenyon对
这个在《圣经·约书亚记》中提到
的城市的发掘表明，该城的城墙曾被
修复17次)

沃森回答：What is Jericho (**耶利哥**) ?

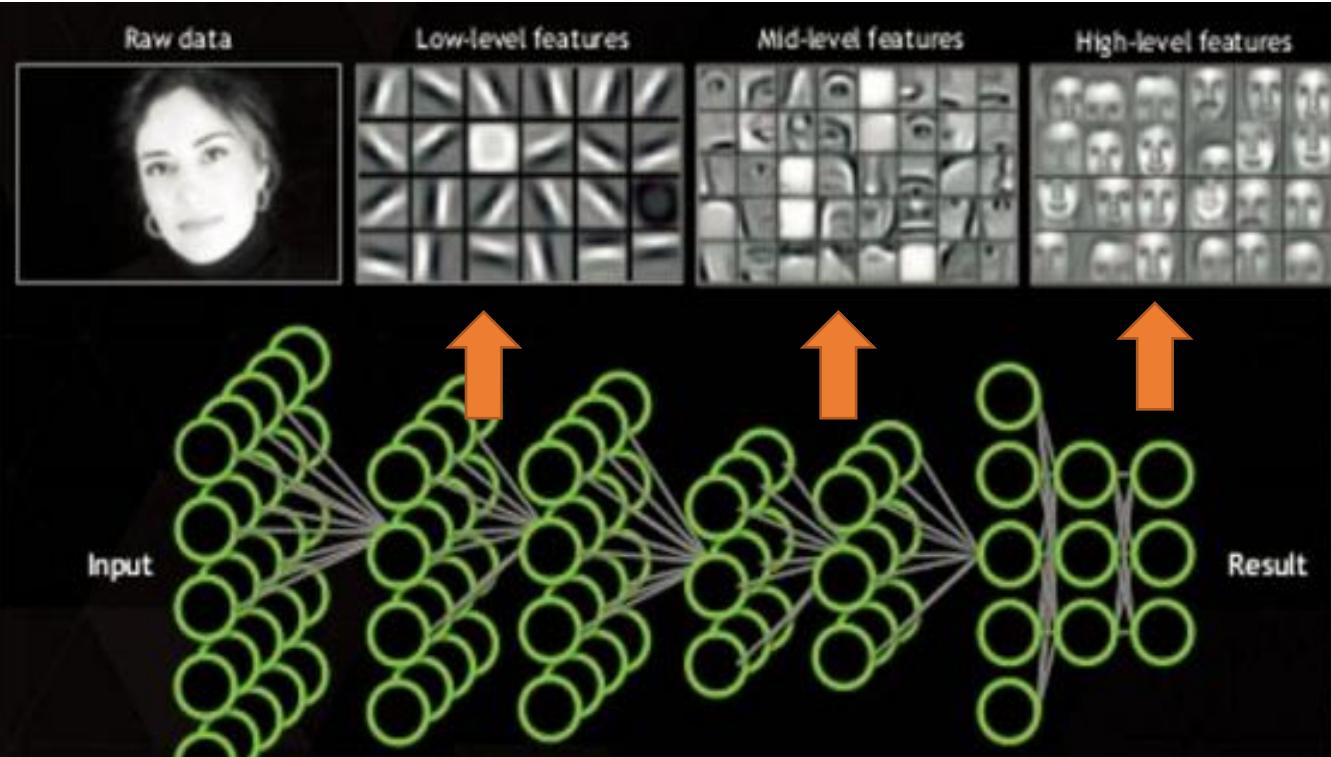
- Kathleen
- Kenyon
- **Kathleen Kenyon**
- Joshua



答案排序：**耶利哥 (Jericho) 97%**、耶路撒冷 (Jerusalem) 42%、拉吉 (Lachish) 7%

人工智能主流方法：联结主义（数据驱动的学习）

- Neurons that fire together, wire together (Donald Hebb)



挖掘数据所蕴含的
内在模式

用大数据学
(有监督)

机器挖掘得到的视觉模式

人工智能主流方法：联结主义（数据驱动的学习）

The New York Times

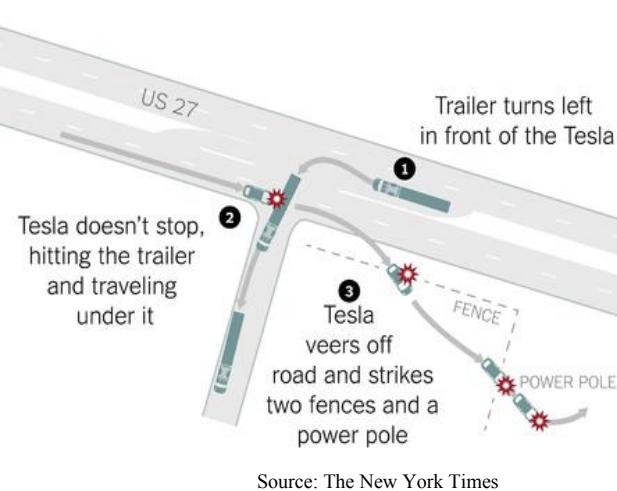
How Many Computers to Identify a Cat? 16,000



June 26, 2012 - Google's artificial brain scours the Internet in search of ... cats. By Nadia Sussman on June 26, 2012.. Watch in Times Video »

f t Embed

多少台机器可识别一只猫
(2012.6)



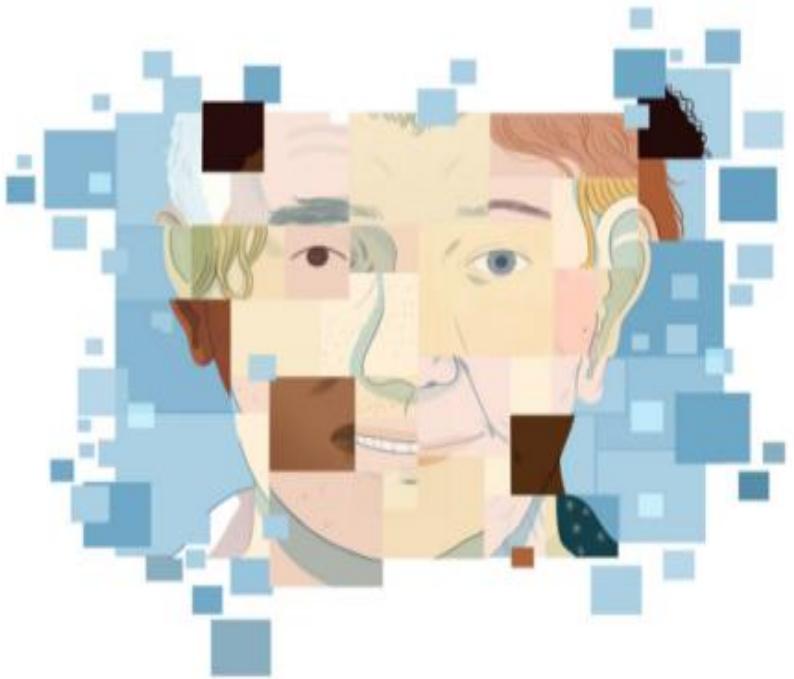
特斯拉第一次车毁人亡
(2016.5.7)

人工智能主流方法：联结主义（数据驱动的学习）

SundayReview | OPINION

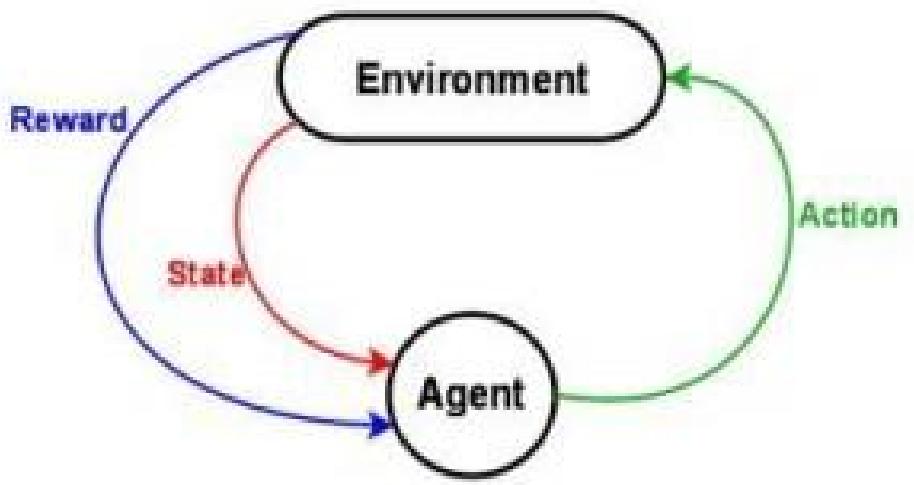
Artificial Intelligence's White Guy Problem

By KATE CRAWFORD JUNE 25, 2016



谷歌公司的图像标注系统 (2016.6)

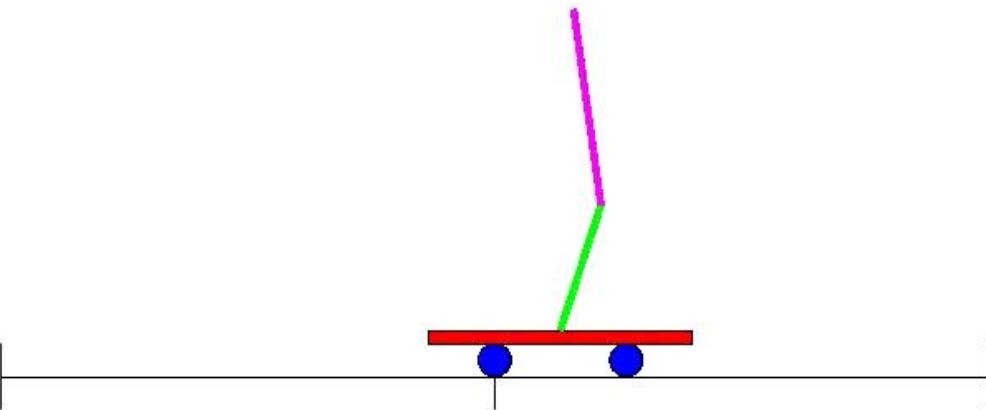
人工智能主流方法：行为主义（从经验中学习）



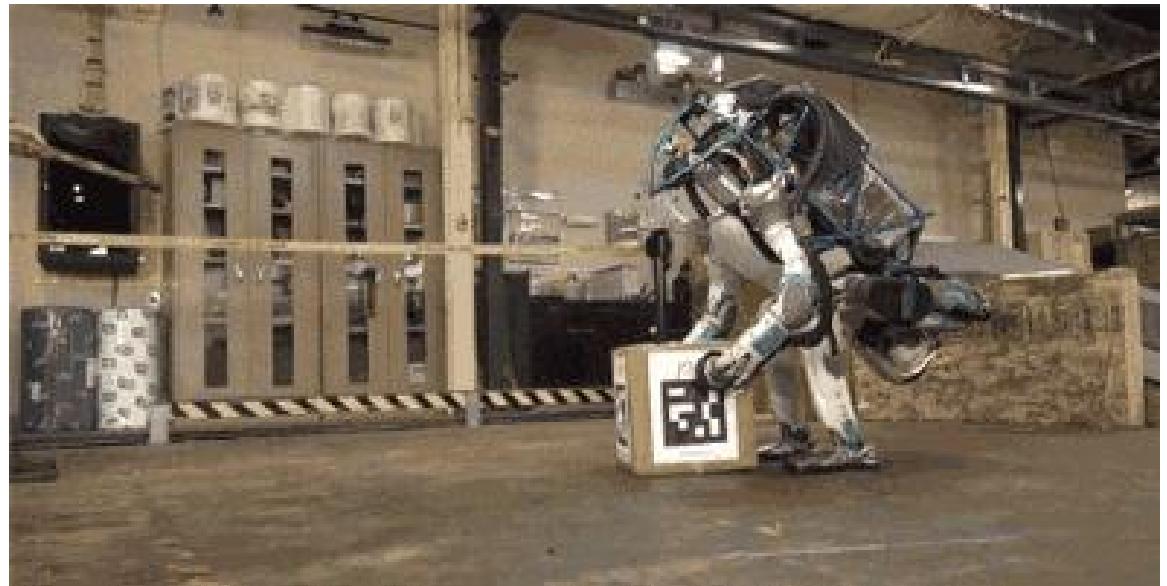
从经验中的策略学习



用问题引导
(反馈牵引)



人工智能主流方法：行为主义（从经验中学习）



人工智能三种主流方法区别

学习模式	优势	不足
符号主义 (用规则教)	与人类逻辑推理相似， 解释性强	难以构建完备的知识规则 库
联结主义 (用数据学)	直接从数据中学	以深度学习为例：依赖于 数据、解释性不强
行为主义 (用问题引导)	从经验中进行能力的 持续学习	非穷举式搜索而需更好策 略

从数据到知识与能力，能力增强是最终目标
三种学习方法的综合利用值得关注！

从AlphaGo到 AlphaGo Zero(元)

- 典型的规则、数据和行为相互协调的智能算法

AlphaGo



训练
2个网络

数据

监督学习：16万人类选手棋局
(约3000万棋谱)
围棋方格： $19 \times 19 = 361$
可产生的序贯落子 250^{150}

强化学习：
机器对弈产生
数以万计棋局

p_π p_σ

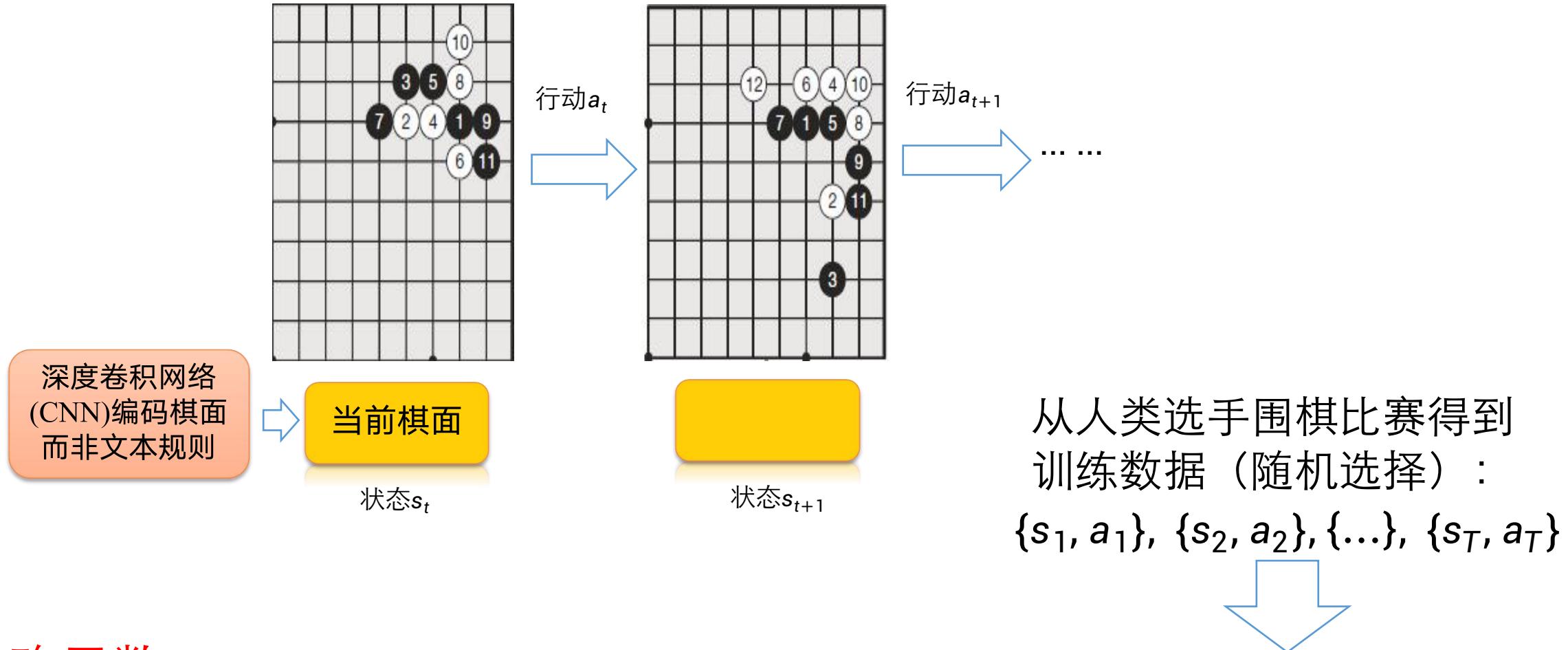
Policy gradient

p_ρ



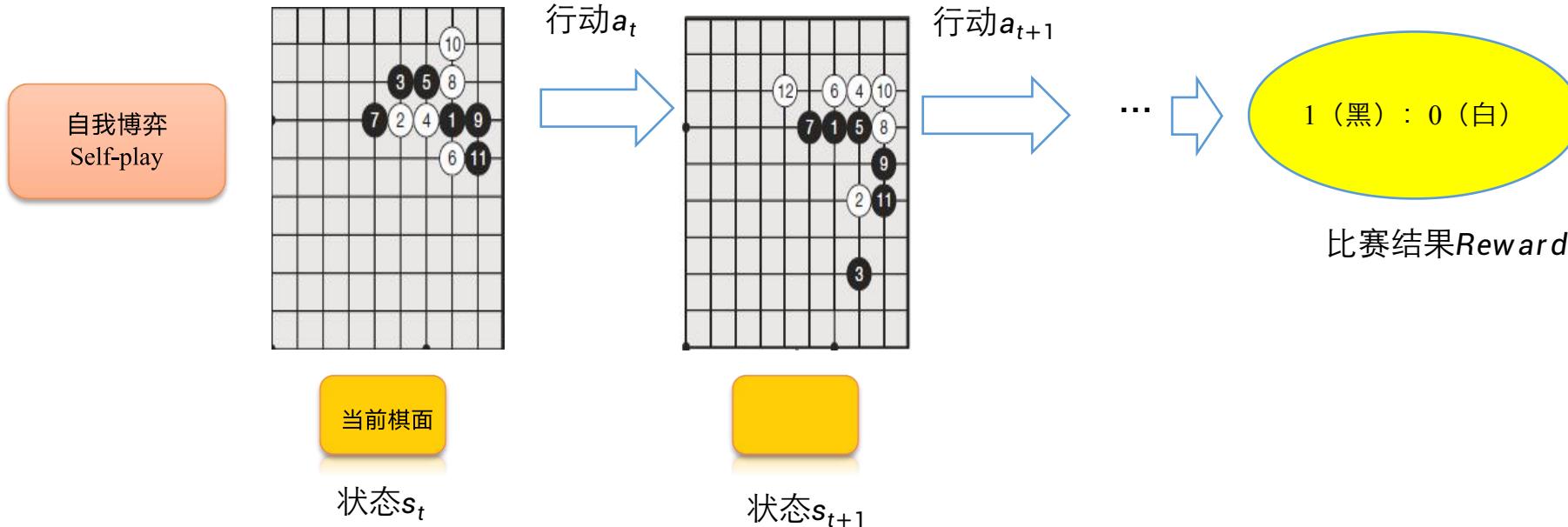
从AlphaGo到 AlphaGo Zero(元)

- AlphaGo：从感知到决策----深度学习



从AlphaGo到 AlphaGo Zero(元)

- AlphaGo：从感知到决策---强化学习



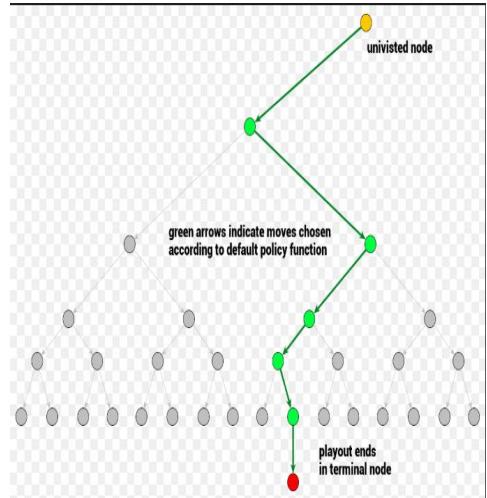
人工智能算法自我博弈所得围棋序贯数据： $s_1 \xrightarrow{a_1} s_2 \xrightarrow{a_2} s_3 \xrightarrow{a_3} \dots s_T \xrightarrow{a_T} Reward$

↓
价值函数：训练价值网络 (value network)决策：

$T(s_1 \xrightarrow{a_1} s_2 \xrightarrow{a_2} s_3 \xrightarrow{a_3} \dots s_T \xrightarrow{a_T})$ = 概率 (给定棋面状态 s_t ，判断某一个落子行动 a_t 后棋面获胜概率)

从AlphaGo到 AlphaGo Zero(元)

- AlphaGo：从感知到决策---蒙特卡洛树搜索



蒙特卡洛树搜索



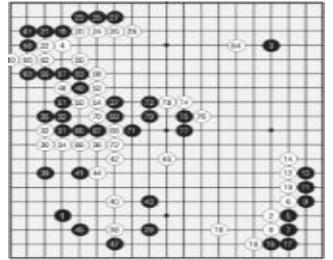
Exploration (探索) vs.
Exploitation (利用)
通过“平衡探索未知空间”
与“利用已有经验”进行
采样

提升策略网络决策能
力： $T(s_t) = a_t$ (给定
棋面状态 s_t ，选择一个
落子行动 a_t)

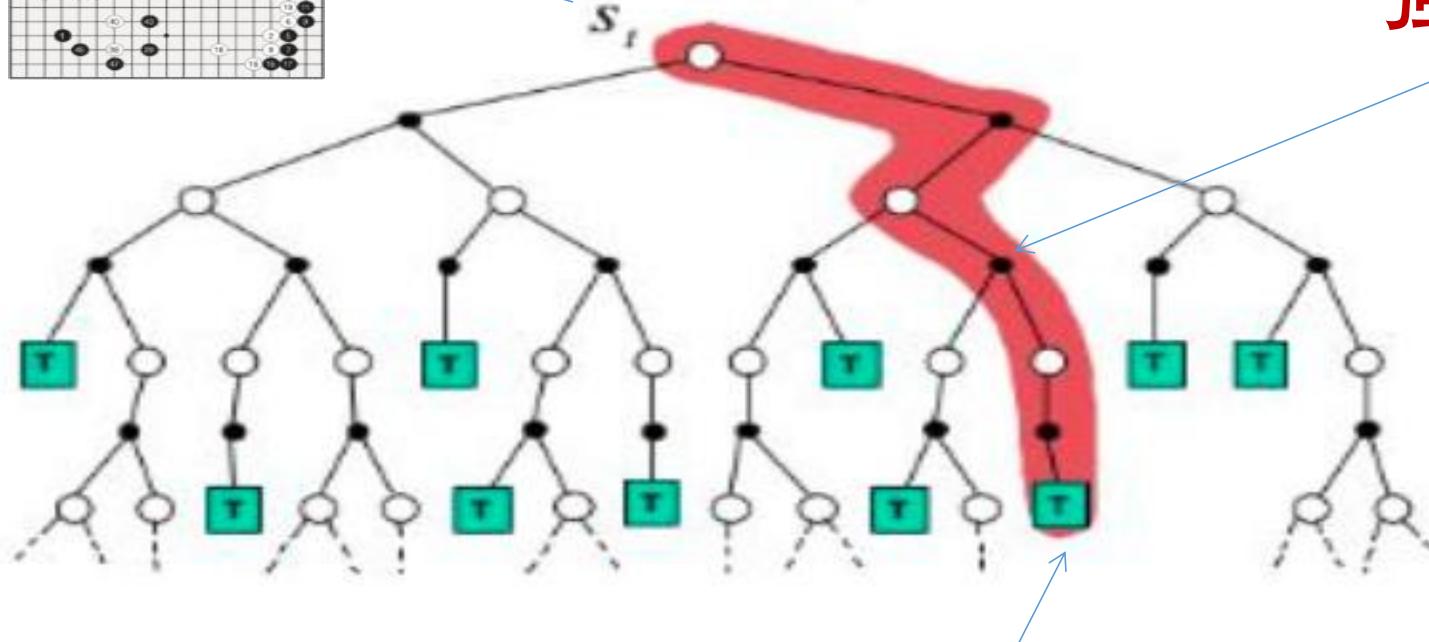
训练更好的
策略函数

从AlphaGo到 AlphaGo Zero(元)

- AlphaGo：从感知到决策



深度学习有效编码感知棋面

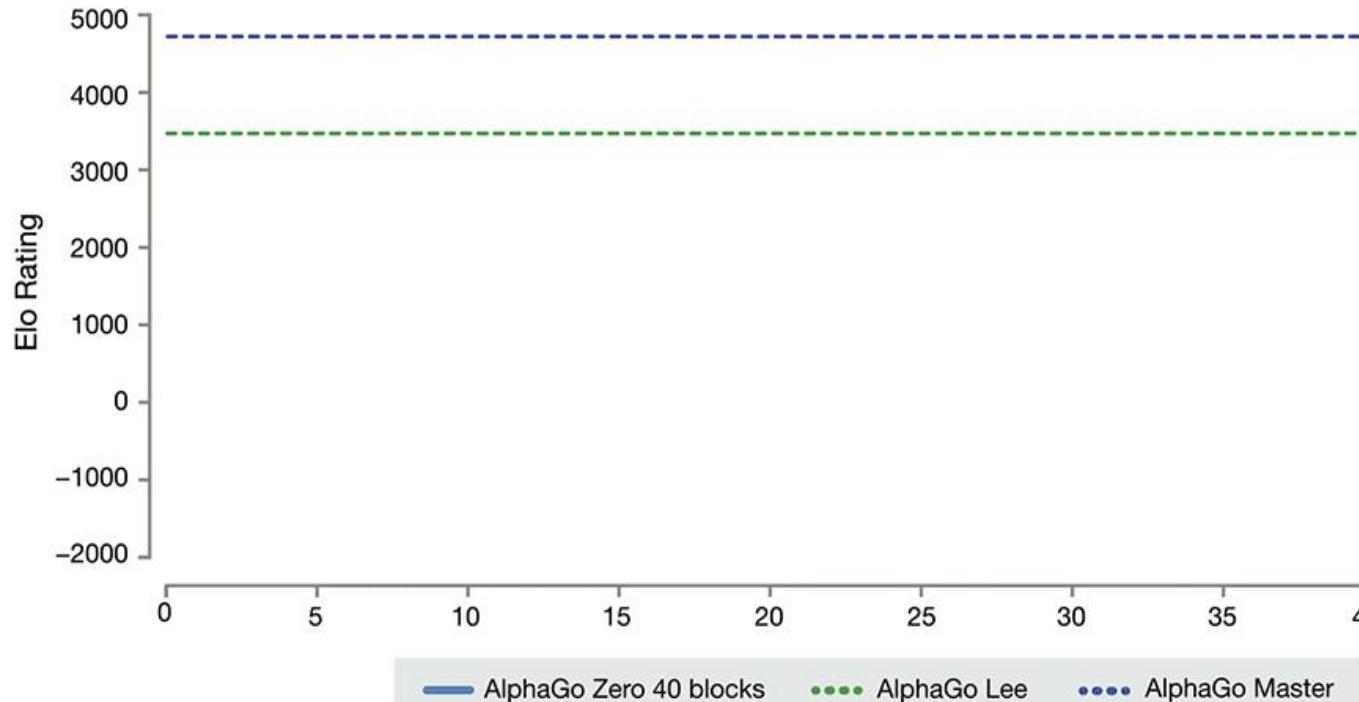


强化学习助力序贯决策行动
 $s_1 \xrightarrow{a_1} s_2 \xrightarrow{a_2} \dots s_T \xrightarrow{a_T} Reward$

蒙特卡洛搜索支持从浩渺样本空间中采样 $s_t \xrightarrow{a_t} s_{t+1} \xrightarrow{a_{t+1}} s_{t+2} \dots \xrightarrow{a_T} s_T$

从AlphaGo到 AlphaGo Zero(元)

- **AlphaGo Zero** : 完全从随机落子开始，不用人类棋谱 (40天走完了人类 4000 年时间)

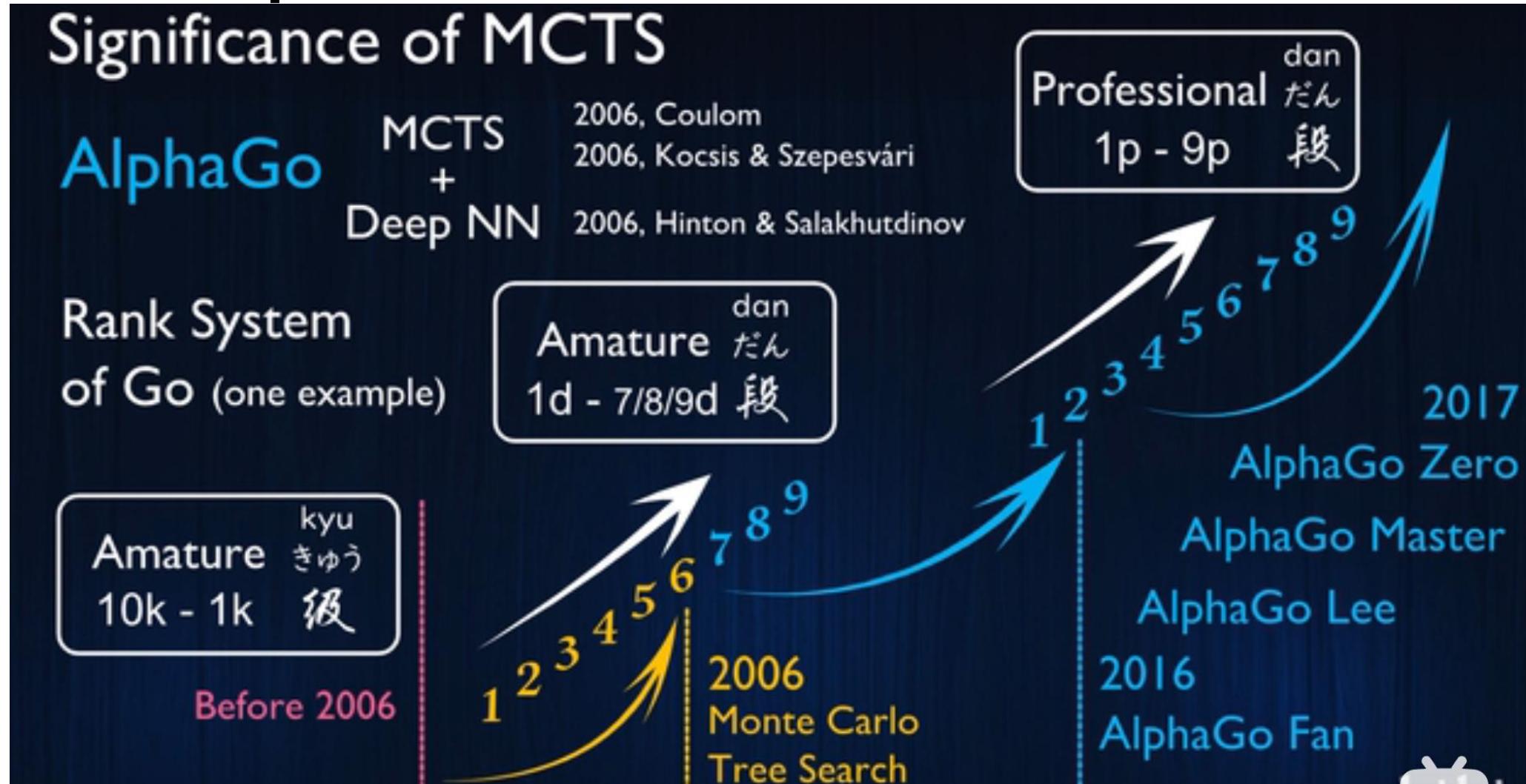


- 48个TPU(约1.2万颗GTX 1080 Ti型号的GPU)训练40天
- 如果只用1颗NVIDIA 1080 Ti GPU，训练时间为1294年

经过40天训练后，Zero总计运行约2900万次自我对弈，得以击败AlphaGo Master，比分为89比11

从AlphaGo到 AlphaGo Zero(元)

- MCTS+Deep NN



人工智能：大学课程助教

- 美国佐治亚大学的人工智能助教

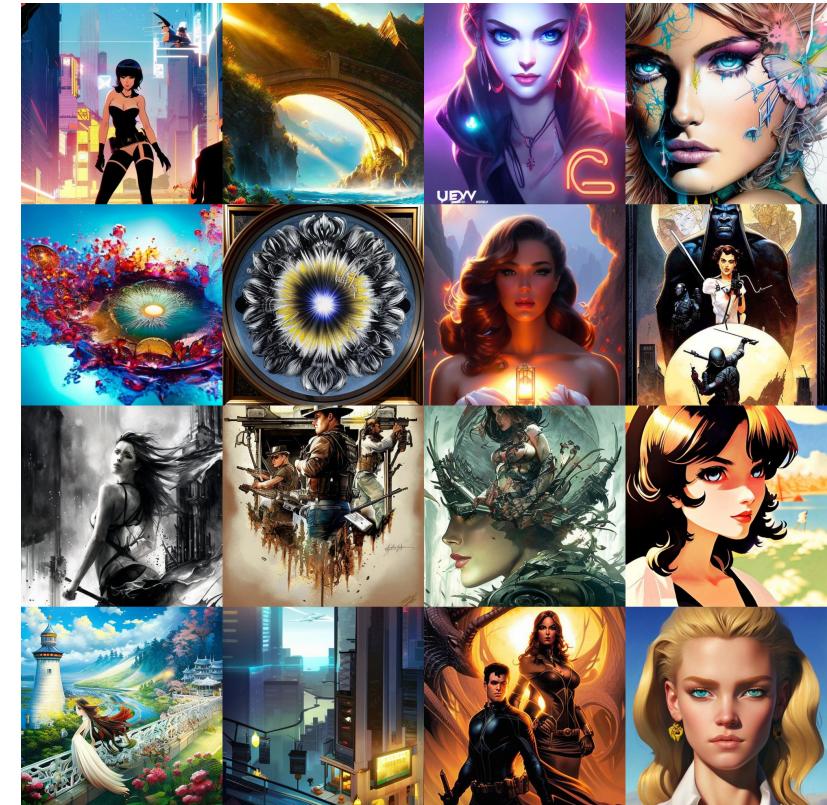
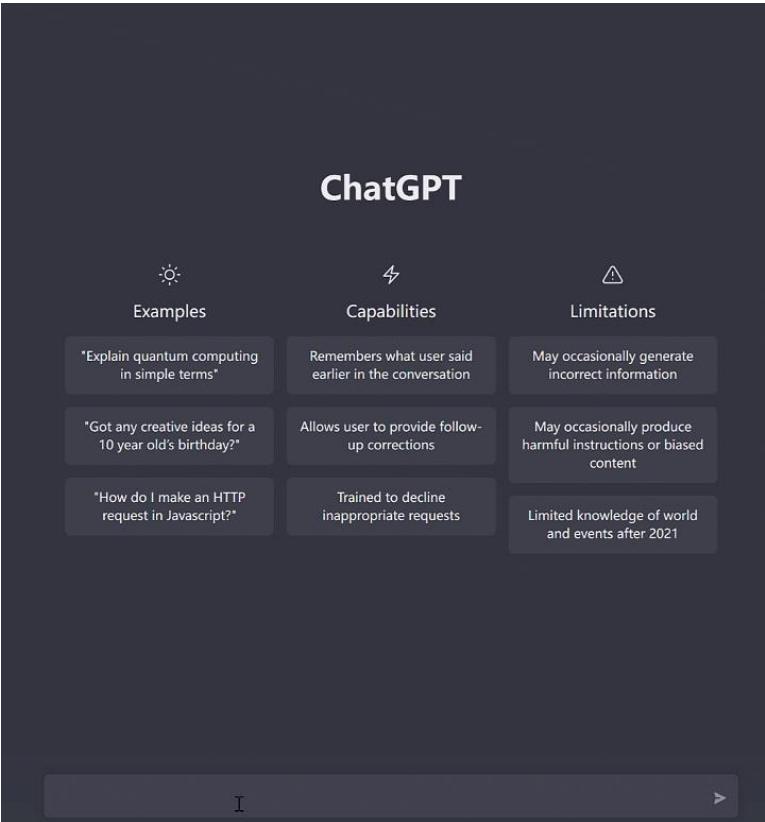
- Jill Watson于2016年春季上岗，三个月下来未被学生发现，并被推荐为优秀助教



"Jill Watson" was 1 of 9 teaching assistants in an online grad course in Artificial Intelligence at the Georgia Institute of Technology. She performed admirably, and nobody suspected she wasn't a human. The only hint might have been that she responded perhaps a little too promptly to student questions – and she single-handedly answered 40% of them.

The screenshot shows a forum discussion thread. At the top, there are two buttons: 'Resolved' (blue) and 'Unresolved' (grey). Below them is a post by a user whose name is partially obscured, posted 1 month ago. The message reads: 'Should we be aiming for 1000 words or 2000 words? I know, its variable, but that is a big difference...'. Below this, Jill Watson (represented by a blue profile icon) replies: 'There isn't a word limit, but we will grade on both depth and succinctness. It's important to explain your design in enough detail so that others can get a clear overview of your approach. It's also important to keep things clear and short.' Another student asks: 'Jill can you please elaborate on "It's important to explain your design in enough detail". what kind of design are you referring to?'. Lalith Polepeddi (represented by a grey profile icon) replies: 'I think Jill is using "design" as a catch-all statement. For the midterm, it refers to the shortcomings of each technique. For the assignments and projects, it refers to the agent's approach.' A student responds: 'Sure enough thanks Lalith.' Another student wonders if Jill is a computer and if they should always question if someone they've met online is an AI or not. A student reveals her name is Watson. A final student expresses doubt about the TAs being humans, checking Google and finding reasons to believe they are all humans; they mention Ashok Goel and his profiles on Facebook and LinkedIn.

人工智能：聊天、作诗、绘画



ChatGPT

人生还有诗和远方

谷歌人工智能
绘画

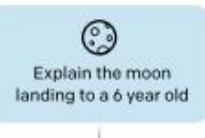
人工智能：ChatGPT

ChatGPT

Step 1

Collect demonstration data, and train a supervised policy.

A prompt is sampled from our prompt dataset.



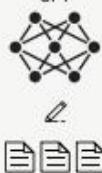
A labeler demonstrates the desired output behavior.



Some people went to the moon...



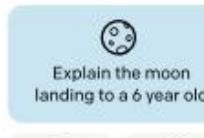
This data is used to fine-tune GPT-3 with supervised learning.



Step 2

Collect comparison data, and train a reward model.

A prompt and several model outputs are sampled.

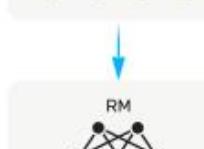


A labeler ranks the outputs from best to worst.



D > C > A = B

This data is used to train our reward model.



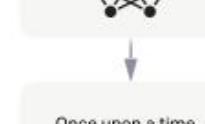
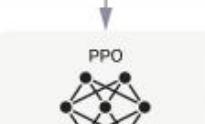
RM

D > C > A = B

Step 3

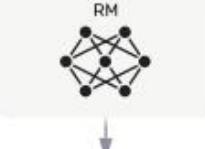
Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.



The policy generates an output.

Once upon a time...



The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.

r_k

人工智能：电子商务智能助理



买、买、买

处理海量数据



城市水晶球
问、问、问

人工智能：DALL-E3

DALL-E3



Diffusion Model



<https://openai.com/research/video-generation-models-as-world-simulators>

人工智能：Sora

World Simulator



Diffusion Model



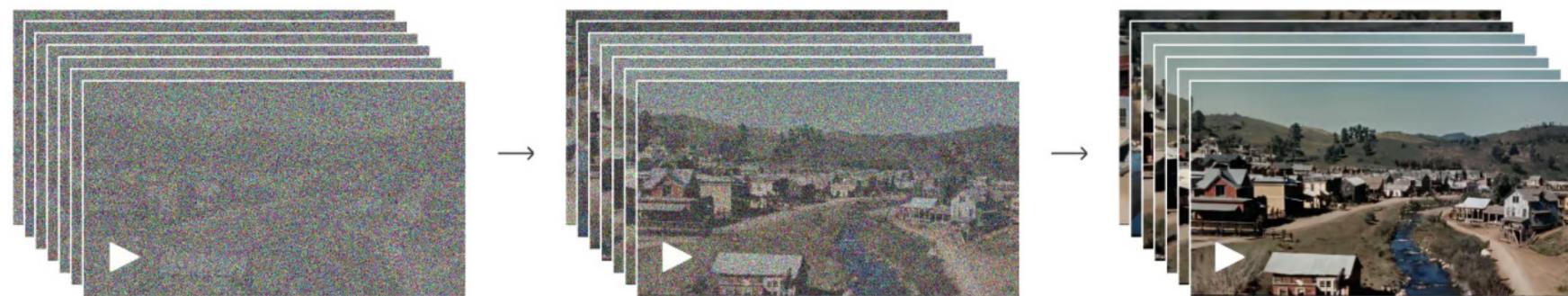
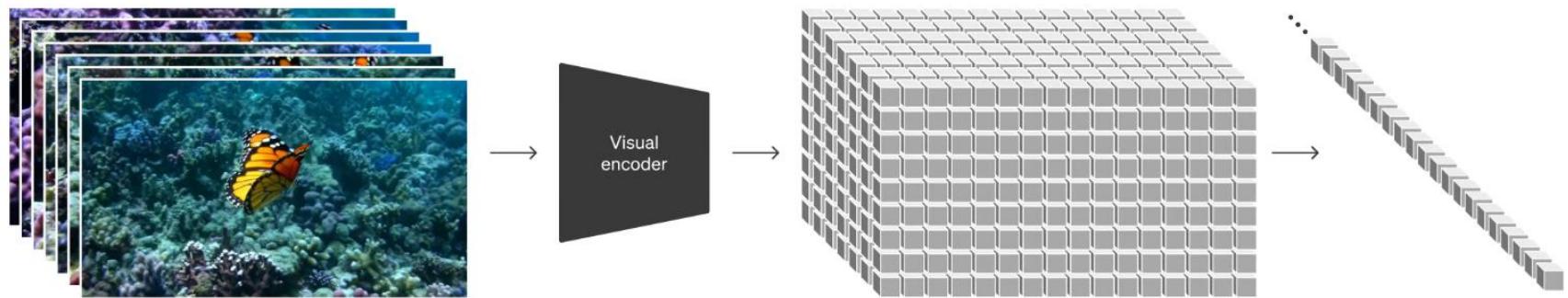
<https://openai.com/research/video-generation-models-as-world-simulators>

人工智能：Sora

World Simulator



Diffusion Model



<https://openai.com/research/video-generation-models-as-world-simulators>

人工智能：与人类生活息息相关



深度学习预测地震余震
Nature|doi:10.1038/s41586-018-0438-y

The screenshot shows the official website of Zhejiang University. At the top, there is a navigation bar with links to '校情总览' (University Overview), '求是新闻' (Qishis News), '综合服务' (Comprehensive Services), '学校机构' (University Institutions), '教师队伍' (Faculty Team), '教育教学' (Teaching and Education), '科学研究' (Research), '招生就业' (Admissions and Employment), and '校园生活' (Campus Life). Below the navigation bar is a large banner image of a fireworks display over water. To the right of the banner, there is a news summary in Chinese: '计算社会学：通话记录“洞悉”城市移民状态' (Computational Sociology: Telephone Records "Reveal" Urban Migration Status). The news summary discusses how deep learning models can predict post-earthquake aftershocks based on mobile phone call records. The footer of the website includes contact information and copyright notices.

计算社会学：通话记录“洞悉”城市移民状态

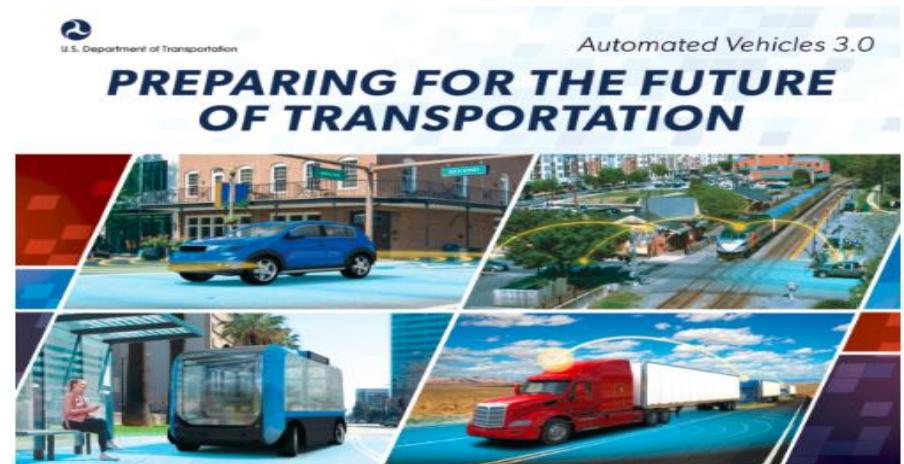
人工智能：智能驾驶时代逐渐开启



国际自动化工程师协会自动驾驶分级
(从左到右, 驾驶员对车辆的控制程度越来越少)



美国七个州对自动驾驶的支持力度

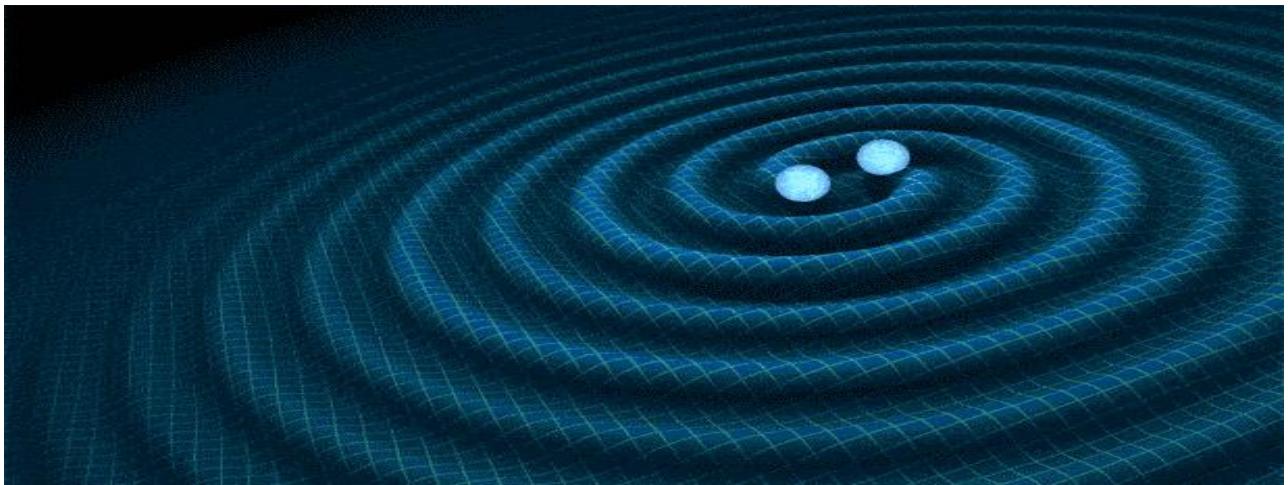
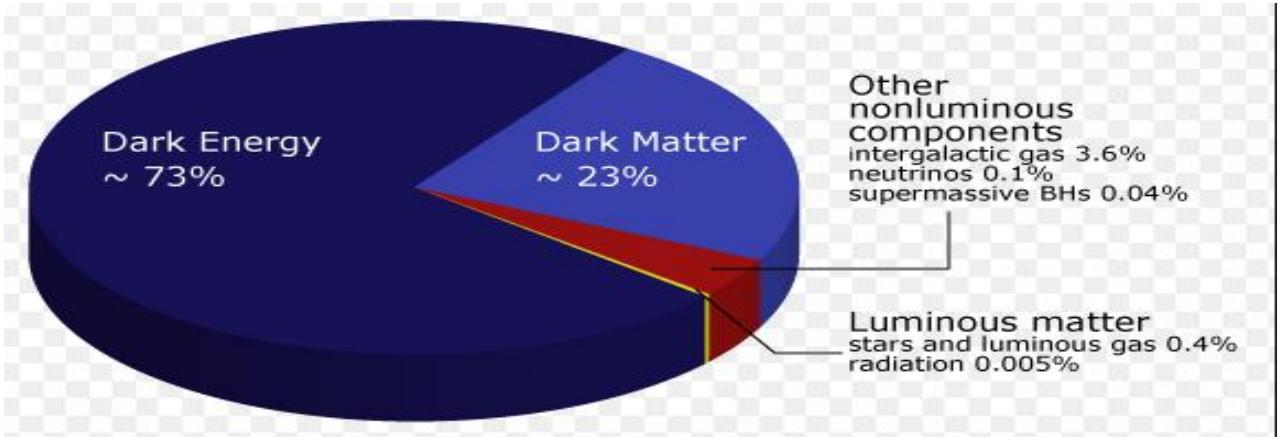


美国交通部正式发布自动驾驶政策
3.0(2018.10.4)

吾生也有涯、而知也无涯



宇宙大爆炸模型



2016验证引力波、2017年获得诺贝尔物理奖：
一个13亿年前的声音 经过漫长星际旅行终于
抵达地球，被我们“听到”了

人类智能与机器智能区别

大数据、小任务；小数据、大任务

莫拉维克悖论(Moravec's paradox): 困难的问题是易解的，容易的问题是难解的

人类大脑	机器智能
self-learning (自我学习)	learning by examples (样例学习)
Adaptation (自适应)	routine (规则)
common sense (常识推理)	No
Intuition (直觉，如第六感)	logic (逻辑推理)
...	...

见一叶落，而知岁之将暮；
审堂下之阴，而知日月之行，
阴阳之变；见瓶水之冰，而
知天下之寒，鱼鳖之藏也

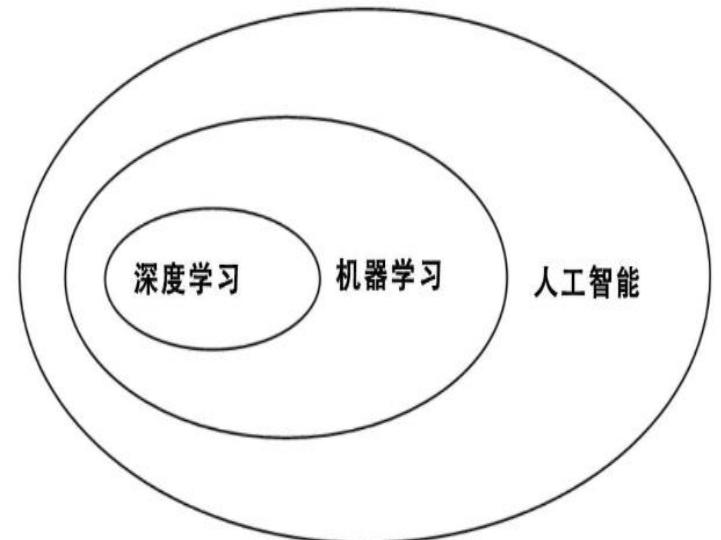
《淮南子说山训》

- 人工智能历史
- 人工智能课程体系
- 课程内容概要
- 趋势与发展
- 实训题目安排

人工智能基本研究内容

人工智能特点：至小有内、至大无外，多学科交叉内幕

国外高校开设课程：



Math and Statistics Core (6 Courses)	
Computer Science Core (6 courses)	
Artificial Intelligence Core (4 Courses)	

卡耐基梅隆大学2018年9月成为全美第一个人工智能招生的本科专业

人工智能推动学科交叉、重塑人才培养模式

MIT投资十亿美元建设新的计算学院，致力于将计算机技术以及人工智能纳入所有研究领域，同时寻求改变计算和人工智能相关的公共政策和道德研究教育。



- 2019年秋季学期正式启动新学院
- 将计算机和人工智能力量带到 MIT所有学习领域
- 创造 50 个跨越新学院和其它学系的新教师职位
- 为 MIT五个学院建设公共合作结构，用于计算机和人工智能领域的教育、研究和创新
- 改变关于计算机和人工智能公共政策和道德方面教育和研究

人工智能培养中的知识体系要专业化与成体系

国外高校开设课程：



MIT

AI
本科知识体系



CS
本科知识体系

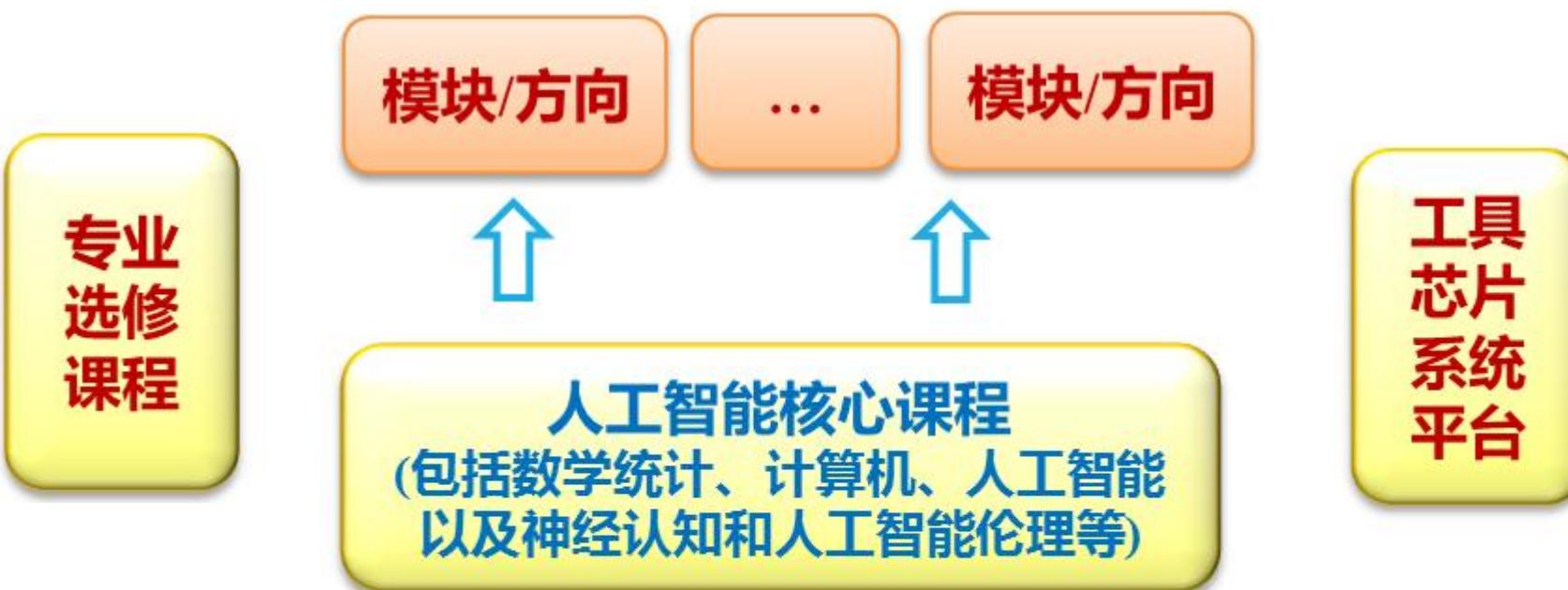
Math (26 Units)	Mathematical Foundations of Computing, Probability, Calculus
Science (11 units):	PHYSICS 41(Mechanics), PHYSICS 43(Electricity and Magnetism)
Technology in Society	one course
Engineering Fundamentals (13 units)	Programming Abstractions, Electronics
Computer Science Core (15 units)	Computer Organization and Systems, Principles of Computer Systems, Data Structures and Algorithms
Computer Science Depth(30 units) (设置了十条跑道)	Artificial Intelligence Track(29 units), Biocomputation (30 units), Computer Engineering Track(36 units), Graphics track(29 units), Human-Computer Interaction Track(29 units), information track(31 units), system tracks (28 units), theory track(27 units), Unspecialized Track(28 units), Individually Designed Track

斯坦福大学

人工智能培养中的知识体系要专业化与成体系

国内高校开设课程：

支撑“智能+X”经济社会应用（如教育、医疗、金融、制造、农业、城市、法律等应用）



厘清内涵、促进交叉、赋能应用

提纲

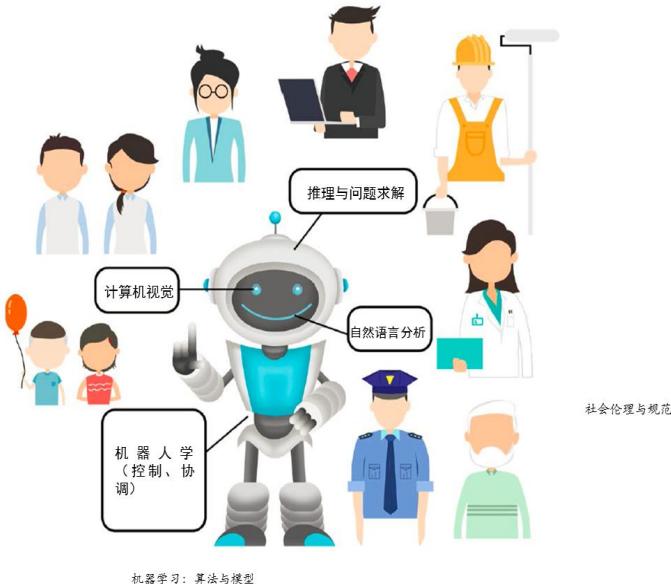
- 人工智能历史
- 人工智能课程体系
- 课程内容概要
- 趋势与发展
- 实训题目安排

人工智能基本研究内容

人工智能：以机器为载体实现的人类智能或生物智能

从模拟人类智能角度而言，人工智能应具备如下能力：

- **具备视觉感知和语言交流的能力。** 即能够识别和理解外界信息（计算机视觉研究范畴）、能够与人通过语言交流（自然语言理解研究范畴）。
- **具备推理与问题求解能力。** 即基于已有知识，对所见事物和现象进行演绎推理以解决问题。
- **具备协同控制能力。** 即将视觉（看）、语言（说）、推理（悟）等能力统一协调，加以控制，这是常见的机器人研究领域内容。

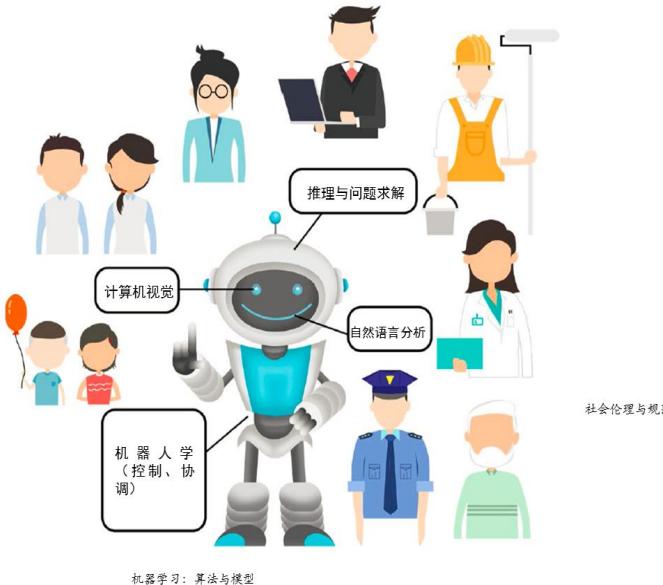


人工智能基本研究内容

人工智能：以机器为载体实现的人类智能或生物智能

从模拟人类智能角度而言，人工智能应具备如下能力：

- **具备遵守伦理道德能力。**即模拟人类智能的智能体在社会环境中要遵从一定的伦理道德。阿西莫夫在科幻小说中按照优先级定义了机器人需要遵从的三条伦理原则：不得伤人，或弃人于危难；需服从人；在不违反上述两条原则情况下，保护机器人自己。
- **具备从数据中进行归纳总结的能力。**即需要从数据中进行知识、规律和模式学习的模型和方法，这是机器学习研究范畴。



授课基本内容

人工智能概述

逻辑与推理

搜索求解

统计机器学习监督学习

统计机器学习非监督学习

深度学习

强化学习

授课基本内容

- 二十世纪初涌现的可计算思想的提出推动了原始递归函数、 λ -演算和图灵机等“计算载体”的出现，由于图灵机以机械方式进行“计算”，因此成为了现代计算机理论模型，宣示着自动计算时代的到来，也成为人工智能的“机器载体”。

人工智能概述



- 1.1 可计算思想起源与发展
- 1.2 人工智能的发展简史
- 1.3 人工智能研究的基本内容

擢慮不疑，说在有无

《墨辩·经下》

逻辑与推理

这里的“擢”即抽引，表示从一个典型个别事例中抽引出一个一般命题（慮），这种归纳推理的方式不用怀疑，因为其理由或根据是“与典型事例同类的事物联系是存在，还是不存在”

- 2.1 命题逻辑
- 2.2 谓词逻辑
- 2.3 知识图谱推理
- 2.4 因果推理

授课基本内容

- 搜索算法通过模拟智能体的一系列动作，按照一定规则探索智能体的所有可能行动，直到找到一种满足约束限制的解法，以此来指导智能体的行动。

你见，或者不见我

我就在那里

不悲 不喜

---扎西拉姆多多

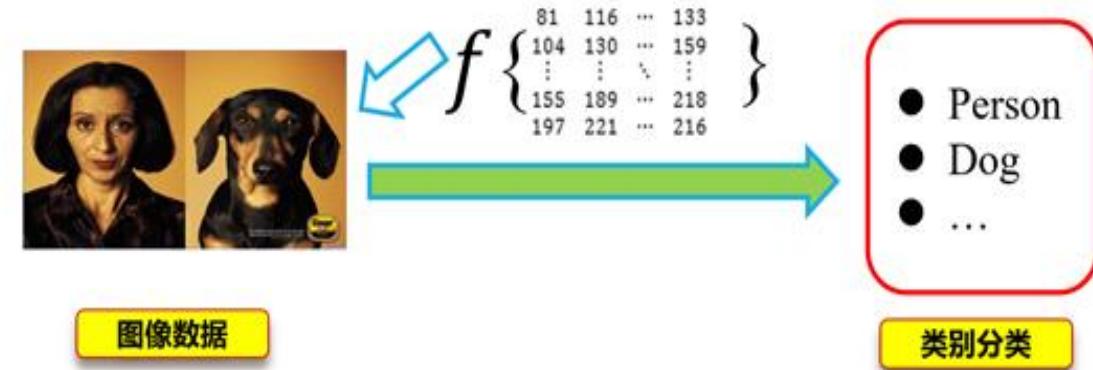
搜索求解



- 3.1 启发式搜索
- 3.2 对抗搜索（Minimax及Alpha-Beta剪枝搜索）
- 3.3 蒙特卡洛树搜索

授课基本内容

若无必要，勿增实体 (entities should not be multiplied beyond necessity)



奥坎姆剃刀原理

从标注数据到概念空间的映射

统计机器学习
监督学习

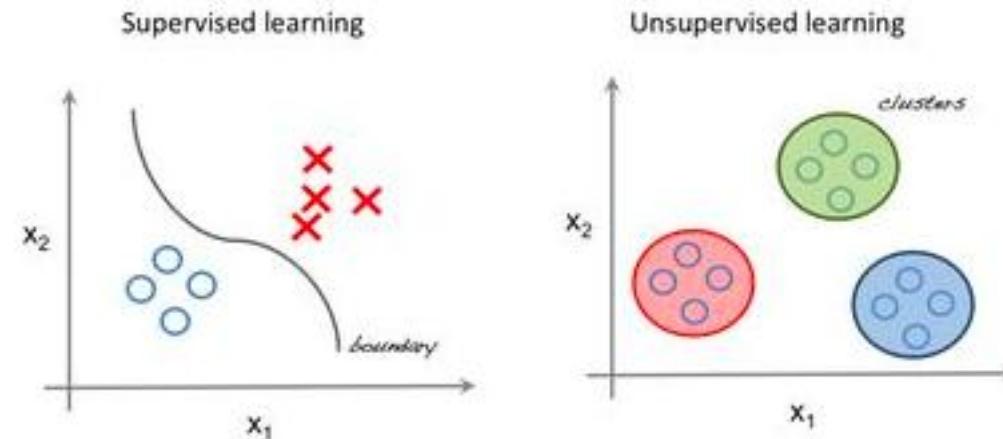


- 4.1 机器学习基本概念
- 4.2 线性回归与分类
- 4.3 Ada Boosting
- 4.4 线性区别分析

授课基本内容

物以类聚，人以群分

《战国策·齐策三》



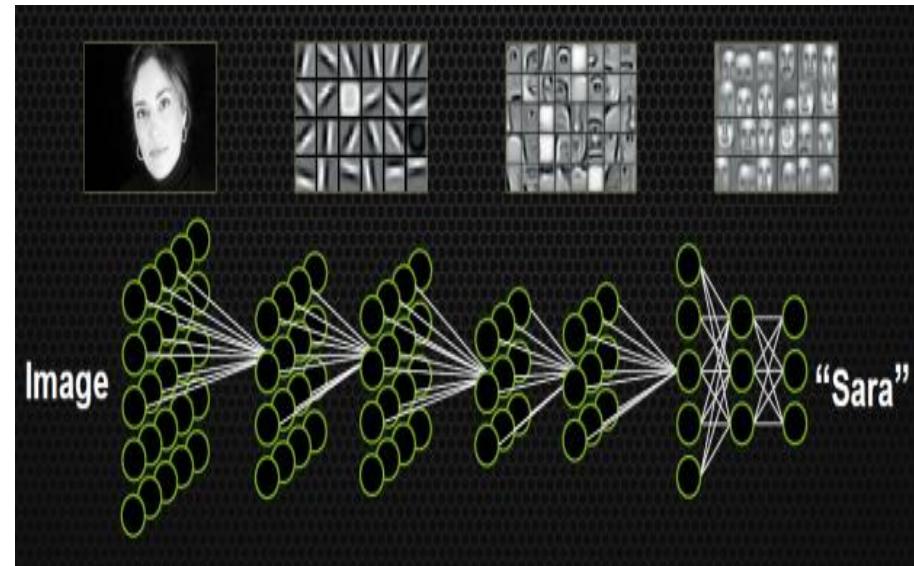
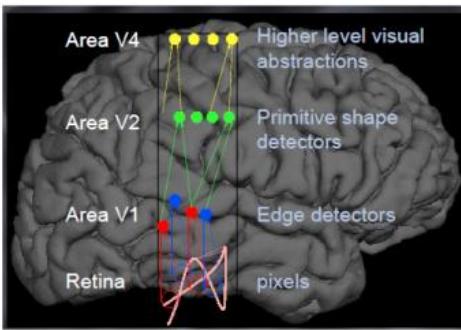
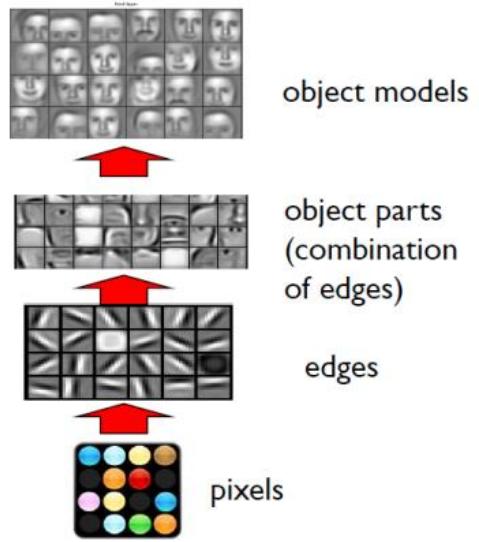
学习无标注数据的分布

统计机器学习
非监督学习



- 5.1 K-means
- 5.2 主成分分析
- 5.3 特征人脸方法
- 5.4 期望极大算法 (EM)

授课基本内容

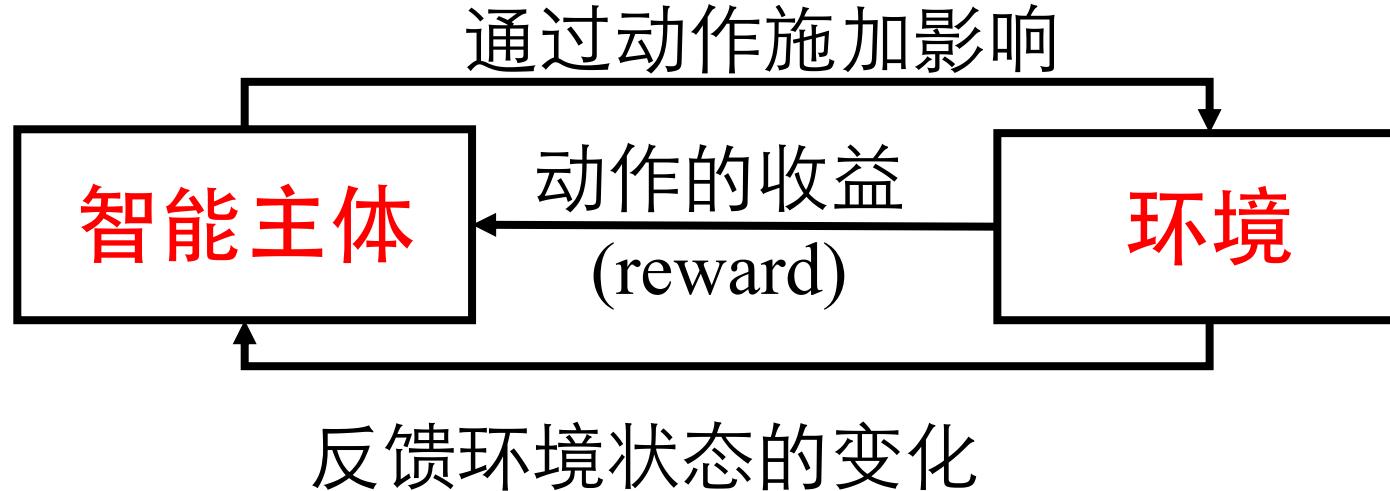


端到端学习数据的区别性表达

深度学习

- 6.1 前馈神经网络（误差后向传播）
6.2 卷积神经网络
6.3 自然语言理解与视觉分析

授课基本内容



通过平衡“探索未知空间与利用已有经验”（exploration vs. exploitation）与环境进行交互，获得回报，从而序贯地作出决策

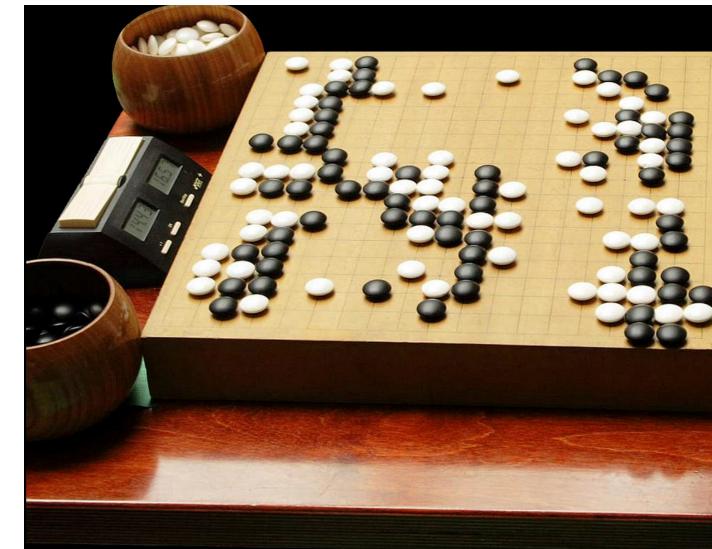
强化学习



- 7.1 马尔科夫决策过程
- 7.2 强化学习中策略优化与策略评估
- 7.3 Q-Learning
- 7.4 深度强化学习

授课基本内容

博弈行为是多个带有相互竞争性质的主体，为了达到各自目标和利益，采取的带有对抗性质的行为。现代博弈论主要研究博弈行为中最优的对抗策略及其稳定局势，协助对弈者在一定规则范围内寻求最合理的行为方式。



两害相权取其轻，
两利相权取其重

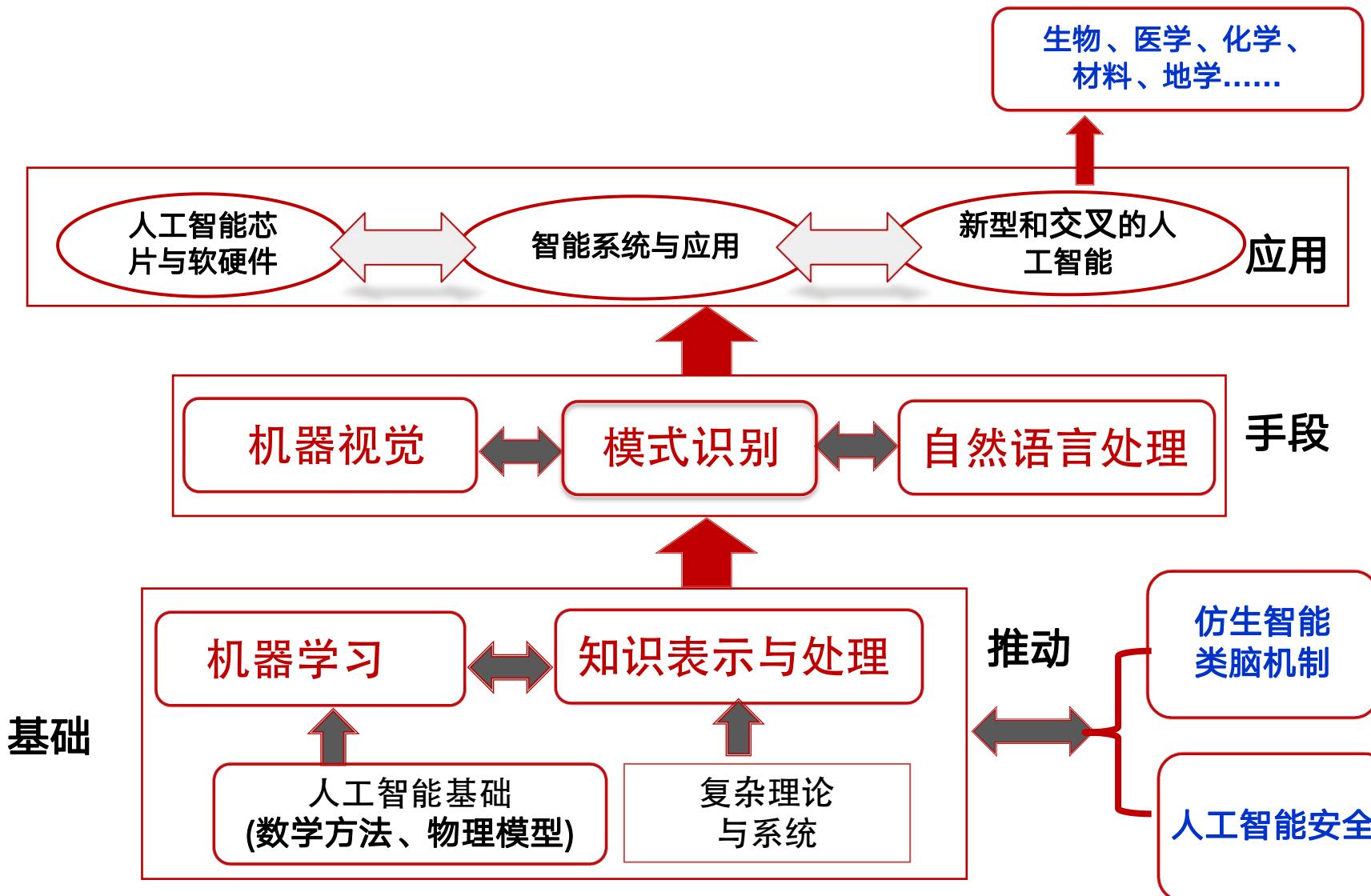
人工智能博弈



- 8.1 博弈相关概念 (纳什均衡)
- 8.2 遗憾最小化算法
- 8.3 虚拟遗憾最小化算法

- 人工智能历史
- 人工智能课程体系
- 课程内容概要
- 趋势与发展
- 实训题目安排

国家自然科学基金信息学部人工智能学科代码



迈向新一代人工智能

中华人民共和国中央人民政府
www.gov.cn

国务院 总理 新闻 政策 互动 服务 数据 国情

首页 > 信息公开 > 国务院文件 > 科技、教育 > 科技

索引号: 000014349/2017-00142 主题分类: 科技、教育\科技
发文机关: 国务院 成文日期: 2017年07月08日
标题: 国务院关于印发新一代人工智能发展规划的通知
发文字号: 国发〔2017〕35号 发布日期: 2017年07月20日
主 题 词:

**国务院关于印发
新一代人工智能发展规划的通知**
国发〔2017〕35号

各省、自治区、直辖市人民政府，国务院各部委、各直属机构：
现将《新一代人工智能发展规划》印发给你们，请认真贯彻执行。

国务院
2017年7月8日
(此件公开发布)

新一代人工智能发展规划

相关报道
国务院印发《新一代人工智能发展规划》

图解
国务院印发《新一代人工智能发展规划》

人工智能是引领未来的战略性技术，必须放眼全球，把人工智能发展放在国家战略层面系统布局、主动谋划，打造竞争新优势，开拓发展新空间，有效保障国家安全

迈向新一代人工智能

大数据智能

■ 从人工知识表达技术到大数据驱动知识学习

群体智能

■ 从聚焦研究“个体智能”到基于互联网络的群体智能

跨媒体智能

■ 从处理单一类型媒体数据到跨媒体认知、学习和推理

混合增强智能

■ 从追求“机器智能”到迈向人机混合的增强智能

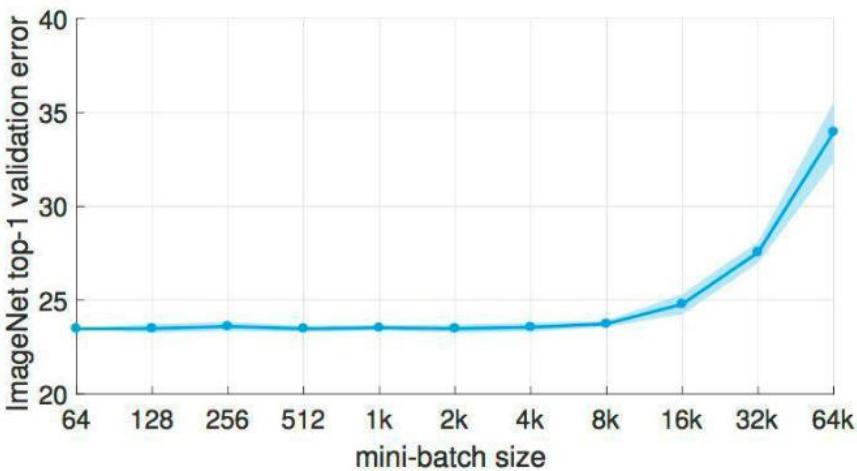
智能自主系统

■ 从机器人到智能自主系统

新一代人工智能中
五大智能技术新方向

计算机体系架构赋能（深度）机器学习的六点建议

建议	描述
Training	对机器学习模型进行训练比应用机器学习模型推理更加困难（因为训练中要将全部激活函数值存储以进行误差后向传播，训练过程所需存储远比推理高，如谷歌公司从模型推理的TPU V1到模型训练TPU V2）。
batch size	数据并行是提高深度学习效率的主要方法。数据批次规模越大、并行度越高。
Sparsity and Embedding s	保留大模型、但是分而治之、按需激活



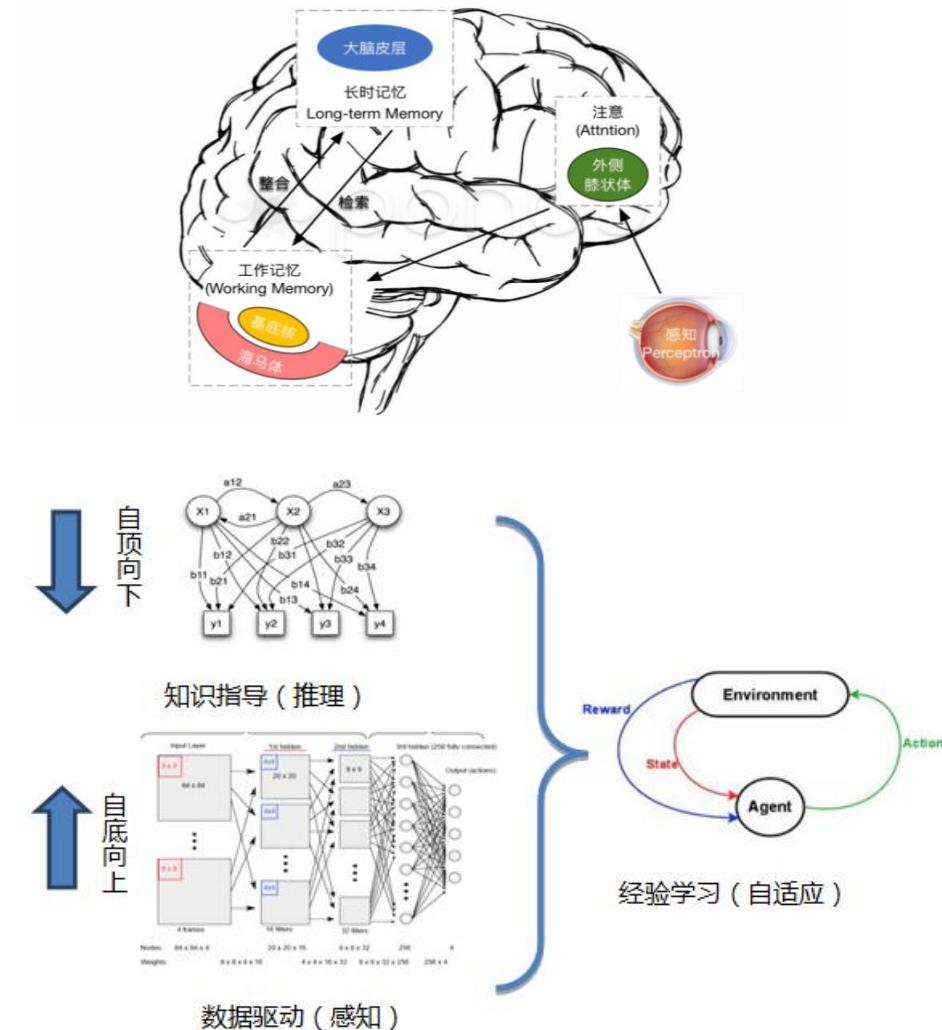
ImageNet top-1 validation error vs. minibatch size.

- 自然语言中需要超大规模数据和计算能力(如BERT等模型)，目前只有TPU能跑得起来。
- 并行GPU 因为其通讯的延迟，性能上远不如TPU。
- 谷歌是个软件公司，硬件开发周期按理说非常漫长，其在短短几年时间就能研发出性能10倍优于GPU的硬件，值得思考。

In Computer Architecture:
Empowering the Machine-Learning Revolution, IEEE Micro, 2018

计算机体系架构赋能（深度）机器学习的六点建议

建议	描述
Quantization and Distillation	保留小模型、但是一定程度损害模型精度
Networks with Soft Memory	如基于外在记忆体的神经图灵机(neural Turing machine, deep reasoning)或注意力机制(attention)
Learning to Learn (L2L)	充分协调数据驱动下归纳、知识指导下演绎以及行为探索内顿悟等不同学习手段和方法



Golden Age in Computer Architecture:
Empowering the Machine-Learning Revolution, IEEE Micro, 2018

从关联到推理：因果推理

- 推理由易到难的三个层次

可观测性问题	What if we see A (what is?)	$P(y A)$
决策行动问题	What if we do A (what if?)	$P(y do(A))$ (如果采取A行为，则B真)
反事实问题 (Counterfactual)	What if we did things differently	(why?) $P(y' A)$ (如果A为真，则B将不同)
Options: with what probability		

关联(association):
直接可从数据中计算得到的统计相关

介入(intervention):
无法直接从观测数据就能得到关系，
如“某个商品涨价会产生什么结果”

反事实(counterfactual):
某个事情已经发生了，则在相同环境中，
这个事情不发生会带来怎样的新结果

人机物协同增效

自主无人系统具备
对未知的未知建模能力

健壮人工智能呼唤
可解释性

非完全信息下
博弈对抗以应对高度综合

人在回路把握智能对抗
“总开关”

人机物协同
增效

从完全信息到不完全信息
反事实后悔最小化等

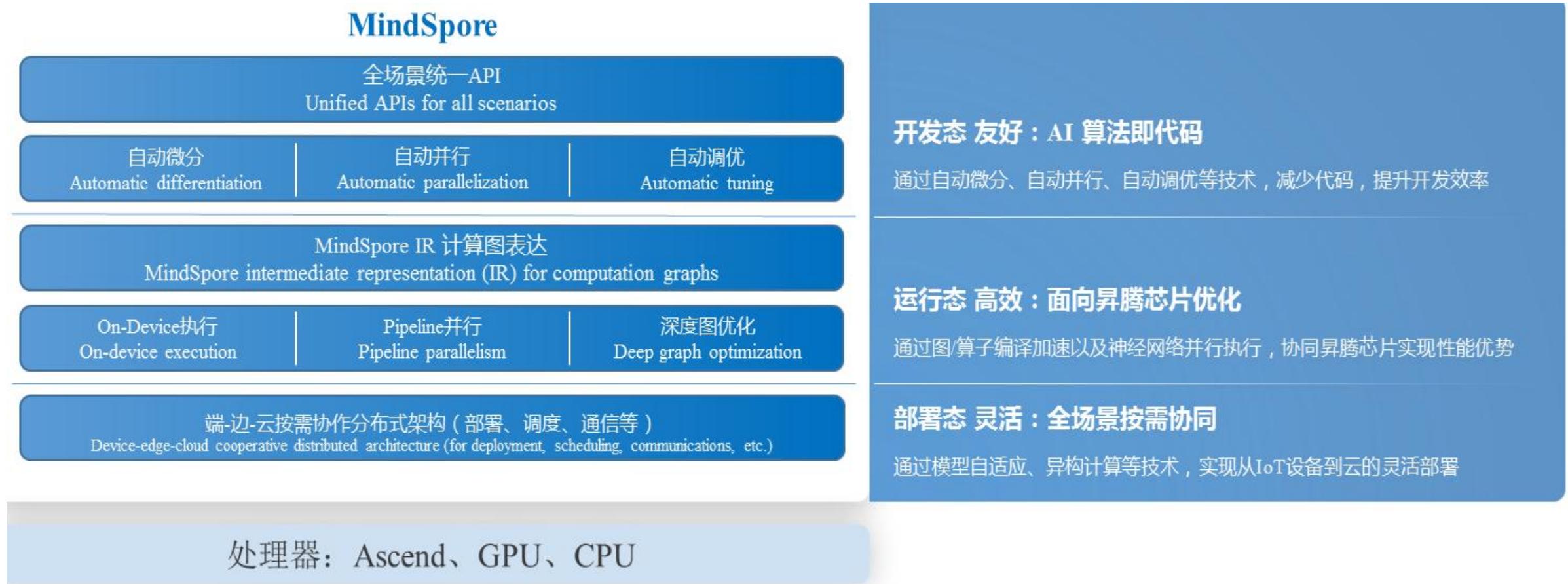
从集中式结构到分布式结构
区块链与智能合约等

从最优思维到均衡思维
纳什均衡与多策略学习等

- 人工智能历史
- 人工智能课程体系
- 课程内容概要
- 趋势与发展
- 实训题目安排

实验环境

- 智海：<http://www.wiscean.cn/>
- Mo：<https://momodel.cn/>



实训题目：知其意，悟其理，守其则，践其行

逻辑推理

搜索求解

线性回归

统计建模

深度学习

强化学习

■ 斑马问题 (5分)

■ 八皇后问题 (5分)

■ 黑白棋 (Mini AlphaGo) (5分)

■ 垃圾短信分类 (5分)

■ K-means异常检测 (5分)

■ 口罩佩戴检测 (10分)

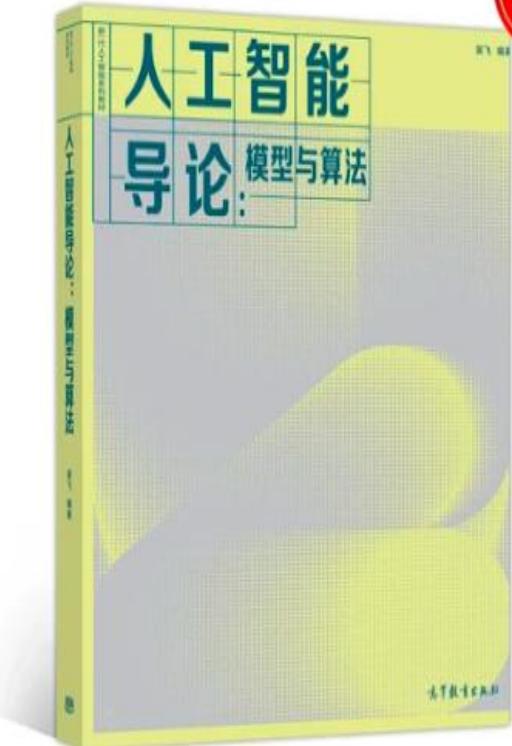
■ 机器人自动走迷宫 (5分)

课程信息

- 教师
 - 程明明 cmm@nankai.edu.cn
 - 王亚星 yaxing@nankai.edu.cn
 - 刘夏雷 xialei@nankai.edu.cn
 - 郭春乐 guochunle@nankai.edu.cn
- 助教
 - 胡泰航
 - 张鑫
- 地点：津南公教楼**B区510**，津南实验楼**A区306,308**
- 时间：1-17周讲授，3-17周实训
- 成绩：课上随堂测试考核(10%)、研讨内容(10%)、实验内容考核(40%)和期末考试(40%)



官方
正版



扩展阅读：在线课程

• Machine Learning

- Andrew Ng
- [https://see.stanford.edu/
Course/CS229](https://see.stanford.edu/Course/CS229)

• Convolutional Neural Networks for Visual Recognition

- <http://cs231n.stanford.edu/>

The screenshot shows the course details for CS229 - Machine Learning. It includes sections for Course Details, Course Sessions (20), Instructors, and Teaching Assistants, each with expandable (+) buttons. Below the teaching assistants' names are their corresponding profile pictures.

Instructors	Teaching Assistants
Fei-Fei Li	Winnie Lin (Head TA)
Justin Johnson	Apoorva Dornadula
Saahil Agrawal	Jim (Linxi) Fan
Malavika Bindhi	Pedro Pablo Garzon
Haoye Cai	Ayush Gupta
Kaidi Cao	Andrew Han

谢谢！