《物理与人工智能》

1. 人工智能概览

授课教师：马滟青

2026/09/07（第一周）



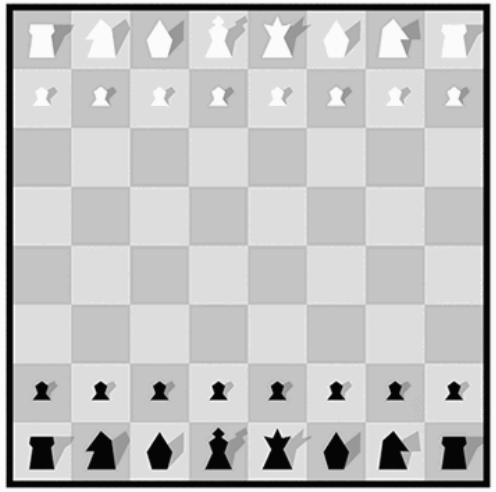
大纲

I. 疯狂的人工智能

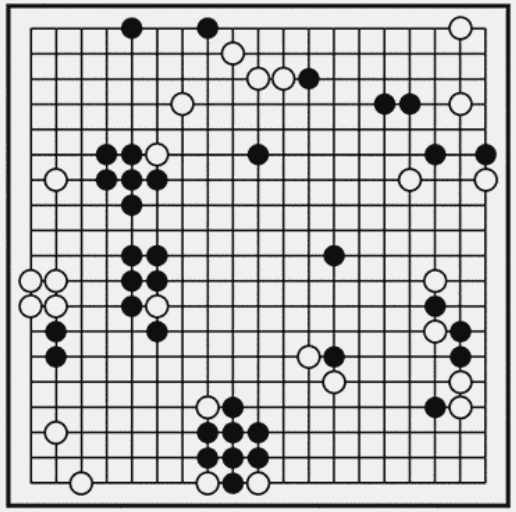
II.什么是人工智能

III. 人工智能的基础及发展

人工智能——认知推理



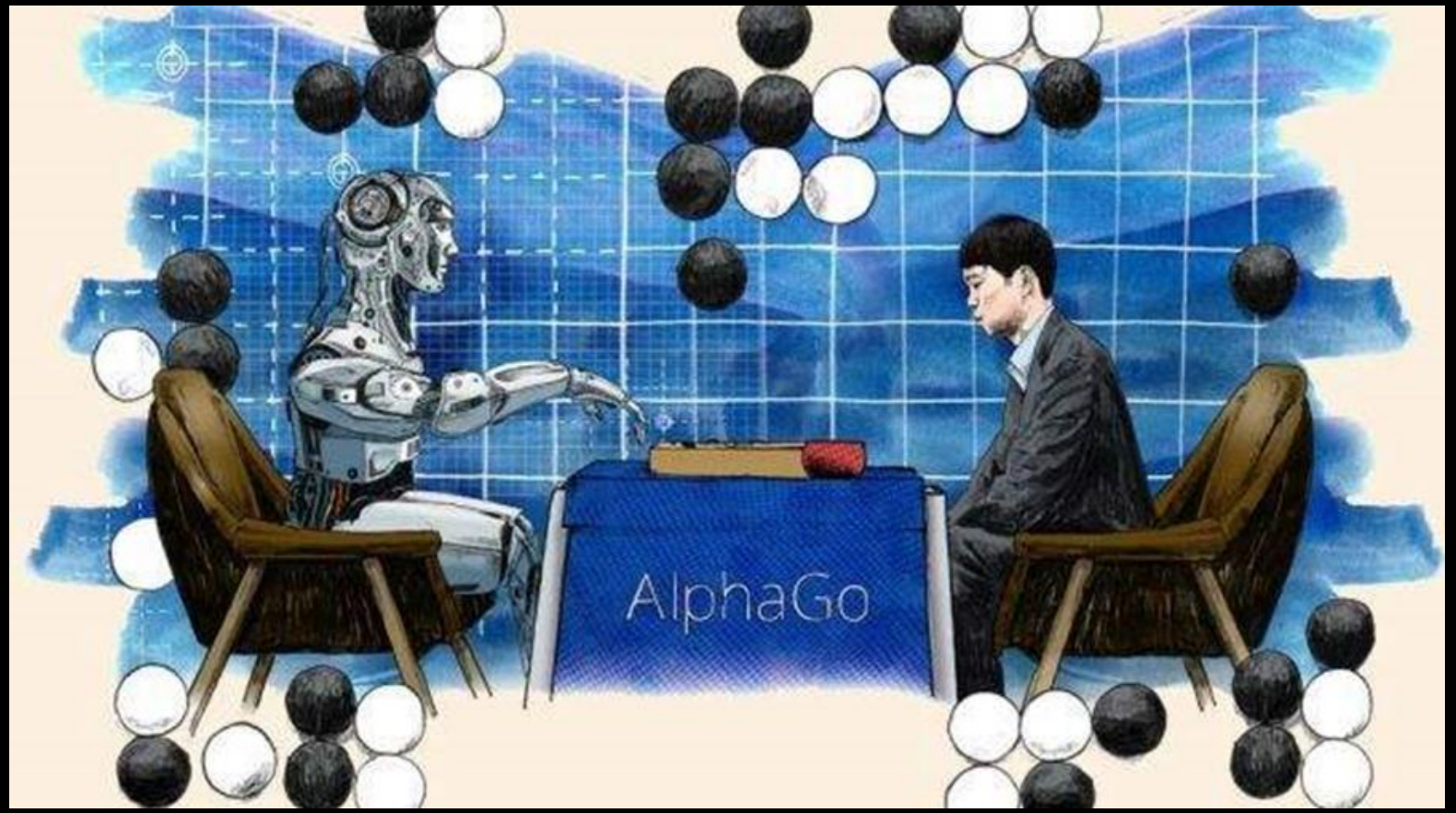
1997年， IBM深蓝战胜世界冠军



$2^{361} \approx 4 * 10^{108}$，巨大的搜索空间

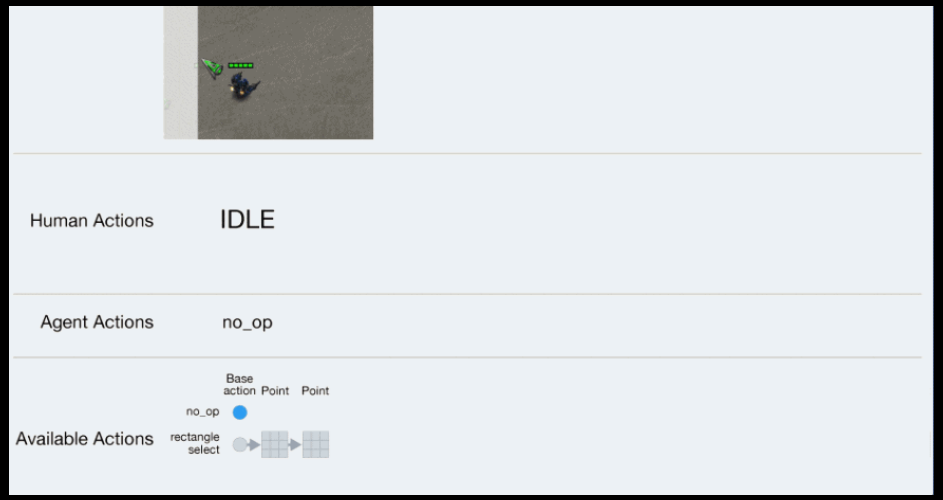
人工智能——认知推理





2016年，AlphaGo战胜世界冠军，蒙特卡洛树搜索+深度学习

人工智能——多智能体







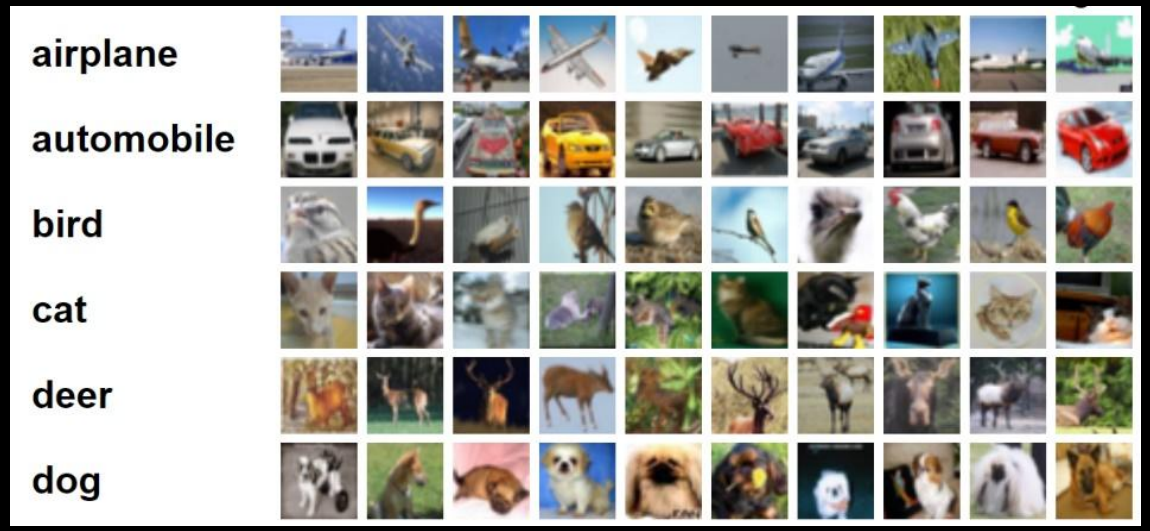


2019，AlphaStar人机大战以10:1战胜职业选手

Grandmaster level in StarCraft II using multi-agent reinforcement learning, Nature 2019

人工智能——计算机视觉

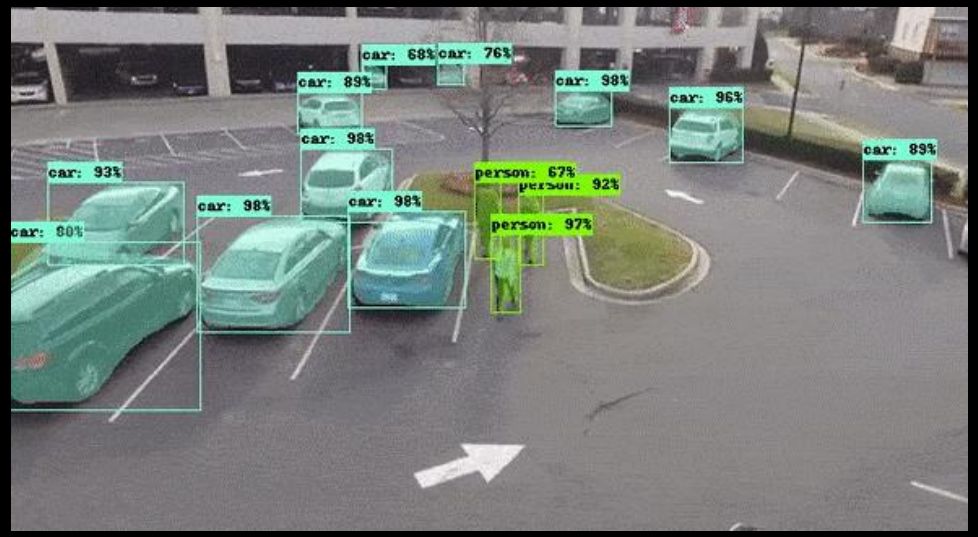




2009年, ImageNet数据集引发了计算机视觉领域的一场革命，
当前很多情况下AI的识别正确率超过人类

人工智能——计算机视觉

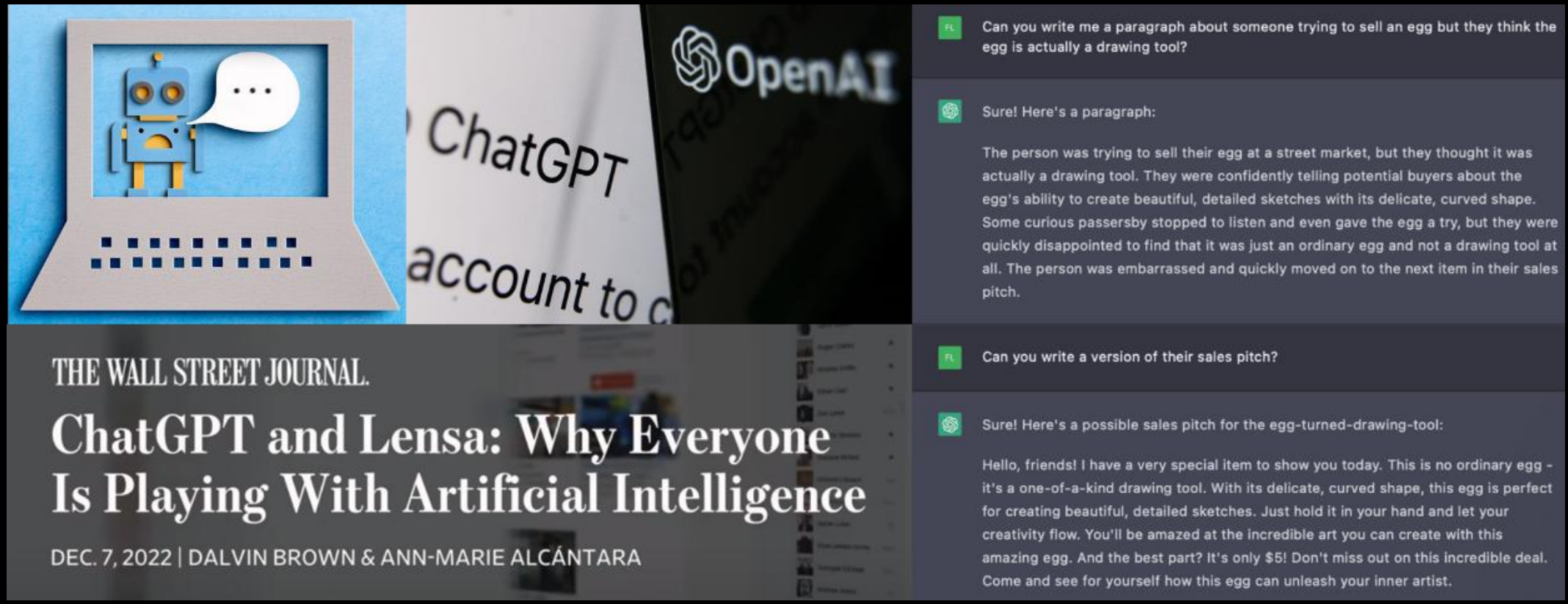

Speed: 37.5 km/h
1:16.7
1. G. DALRYMPLE
Lap: 10.01
2. A. BICANOVA
Lap: 10.10
3. S. CIRTI
Lap: 9.92


car: 68%
car: 76%
car: 89%
car: 98%
car: 98%
car: 96%
car: 89%
car: 93%
car: 98%
car: 98%
person: 67%
person: 92%
person: 97%


人脸识别
年龄：28
性别：男


48 %
D
40
15
1:29 PM
65°F

人工智能——自然语言处理



ChatGPT

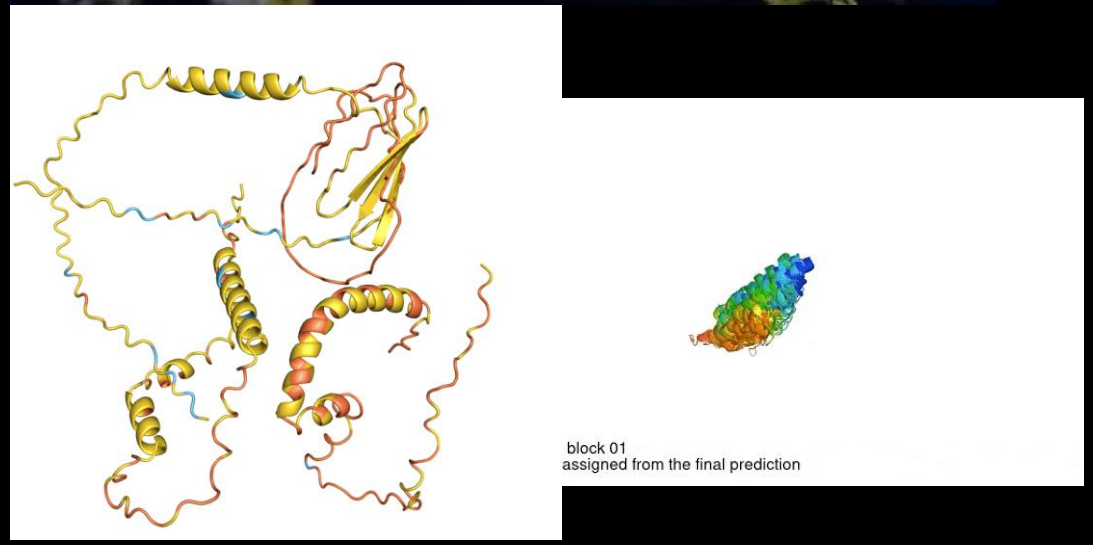
Training language models to follow instructions with human feedback, Arxiv 2022

https://openai.com/blog/chatgpt/



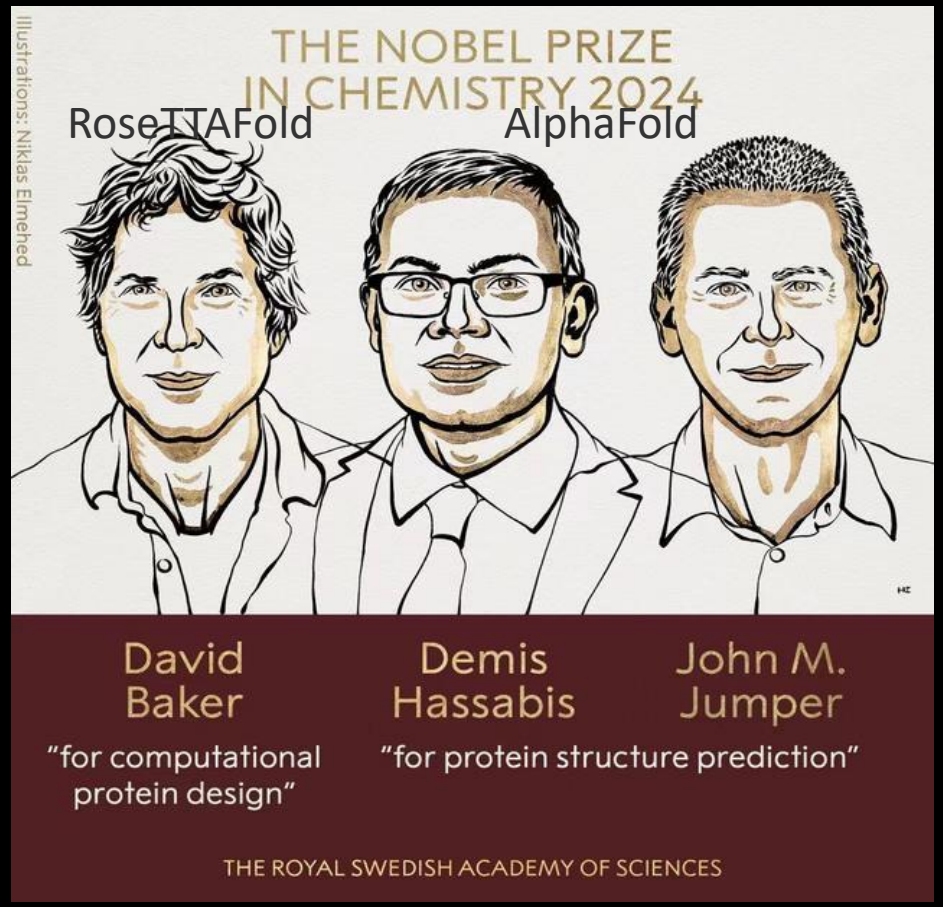
人工智能——科学研究





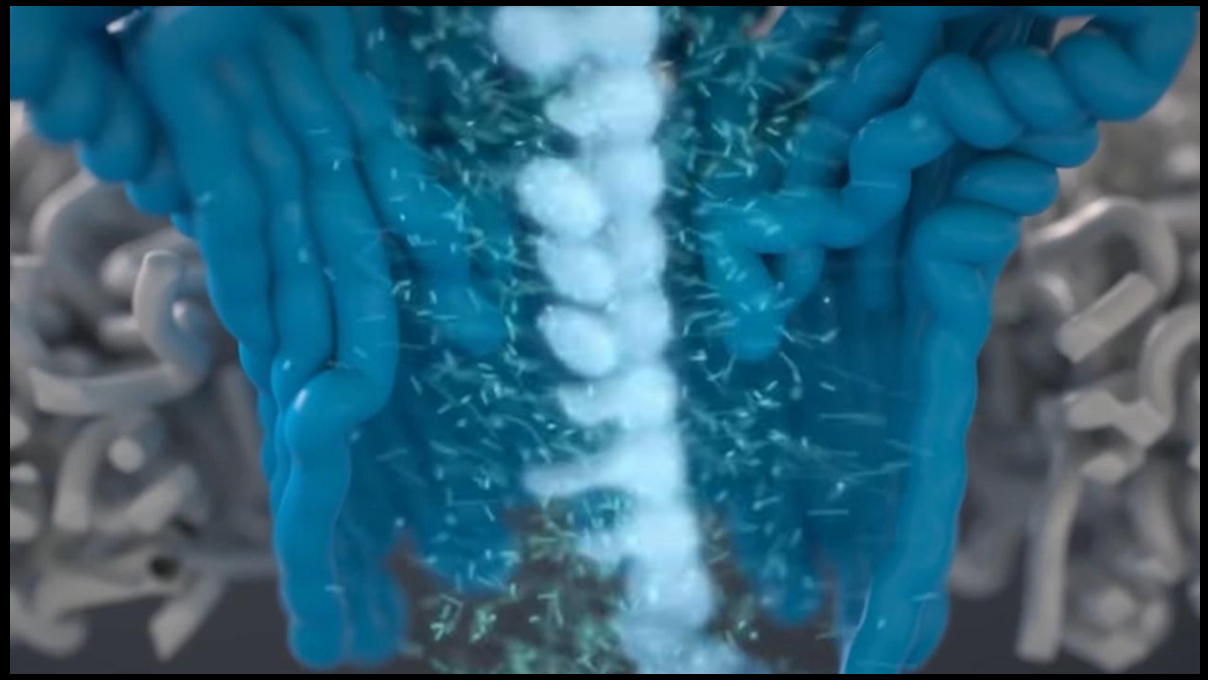
2020年11月30日，AlphaFold 2 蛋白质结构预测

1.8亿个蛋白质结构：0.1% -> 98.5%

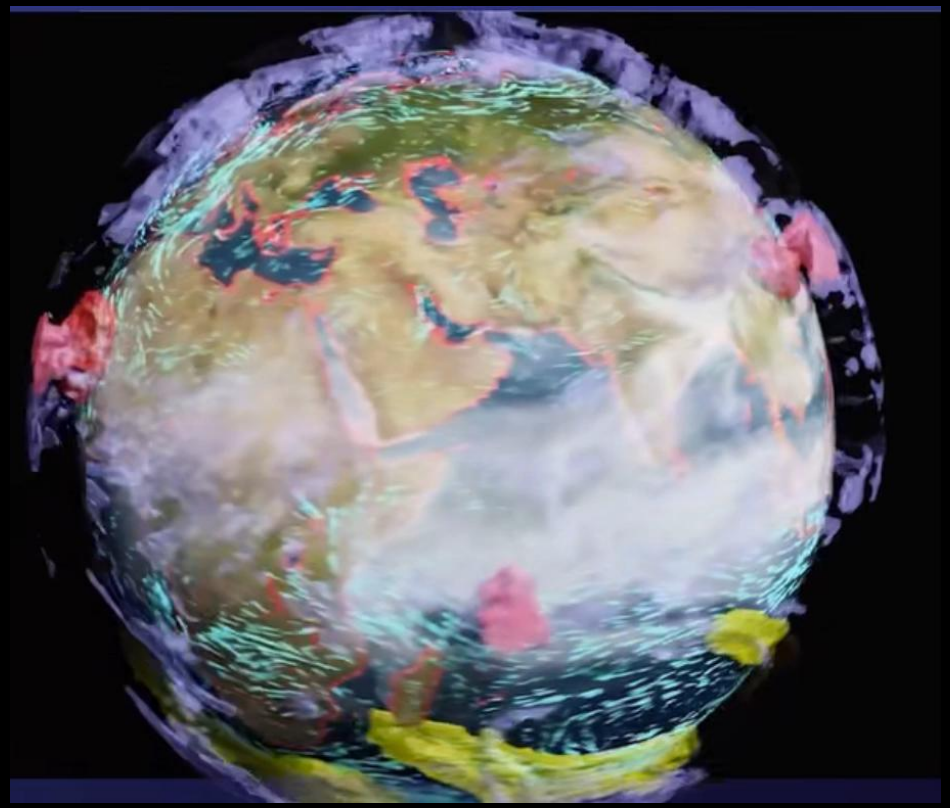


（基于人工智能）进行蛋白质设计和结构预言

人工智能——科学研究



2022年1月，DNA 测序的世界纪录：5小时2分钟



英伟达正在创建地球的数字孪生，Earth-2，
希望提前几十年来预测气候的变化

AI攻城略地

结束的开始:

2022

- Chat client

CHERRY STUDIO

NextChat

Chatbox

Cici

2023

- Translation

DeepL



极端模糊

2024

- Coding

Cline

Claude

GEMINI

GitHub Copilot

2025

- Mathematics

Google AlphaEvolve

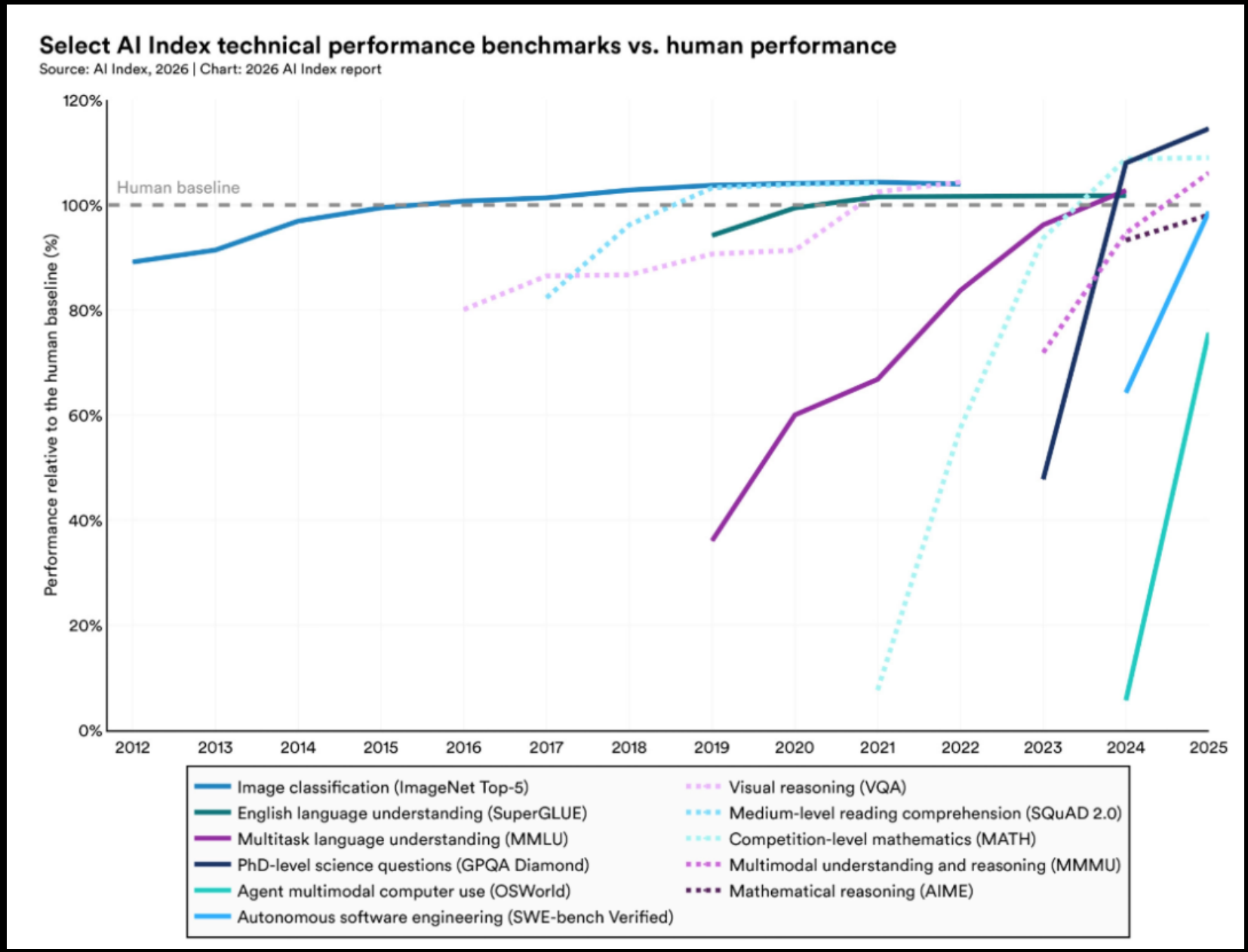
MadEvolve

MAGELLAN

SkyDiscover

极端精确

与人类能力比较


Select AI Index technical performance benchmarks vs. human performance
Source: AI Index, 2026 | Chart: 2026 AI Index report
Performance relative to the human baseline (%)
120%
100%
80%
60%
40%
20%
0%
Human baseline
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
Image classification (ImageNet Top-5)
English language understanding (SuperGLUE)
Multitask language understanding (MMLU)
PhD-level science questions (GPQA Diamond)
Agent multimodal computer use (OSWorld)
Autonomous software engineering (SWE-bench Verified)
Visual reasoning (VQA)
Medium-level reading comprehension (SQuAD 2.0)
Competition-level mathematics (MATH)
Multimodal understanding and reasoning (MMMU)
Mathematical reasoning (AIME)

AI：所向披靡

5 : 0



K.O！

大纲

I. 疯狂的人工智能

II.什么是人工智能

III. 人工智能的基础及发展史



什么是智能？

- **智能**（intelligence）：是智力和能力的总称。从感觉到记忆到思维这一过程，称为“智慧”，智慧的结果就产生了行为和语言，将行为和语言的表达过程称为“能力”，两者合称“智能”
- **多元智能理论**：H. Gardner提出
  ① 语言智能
  ② 逻辑智能
  ③ 空间智能
  ④ 肢体运动智能
  ⑤ 音乐智能
  ⑥ 人际智能
  ⑦ 自我认知智能
  ⑧ 自然探索智能



什么是智能？

$$y = f(x)$$

输出
（响应）

输入
（状态）

- 棋盘：下子位置
- 局势：兵种操作
- 氨基酸排序：蛋白质结构
- 当前天气：未来天气
- 已有语句：后续语句
- ……

什么是智能？

$$y = f(x)$$

输出
（响应）

输入
（状态）

预言（泛化）能力？

不能简单地枚举

信息压缩产生智能

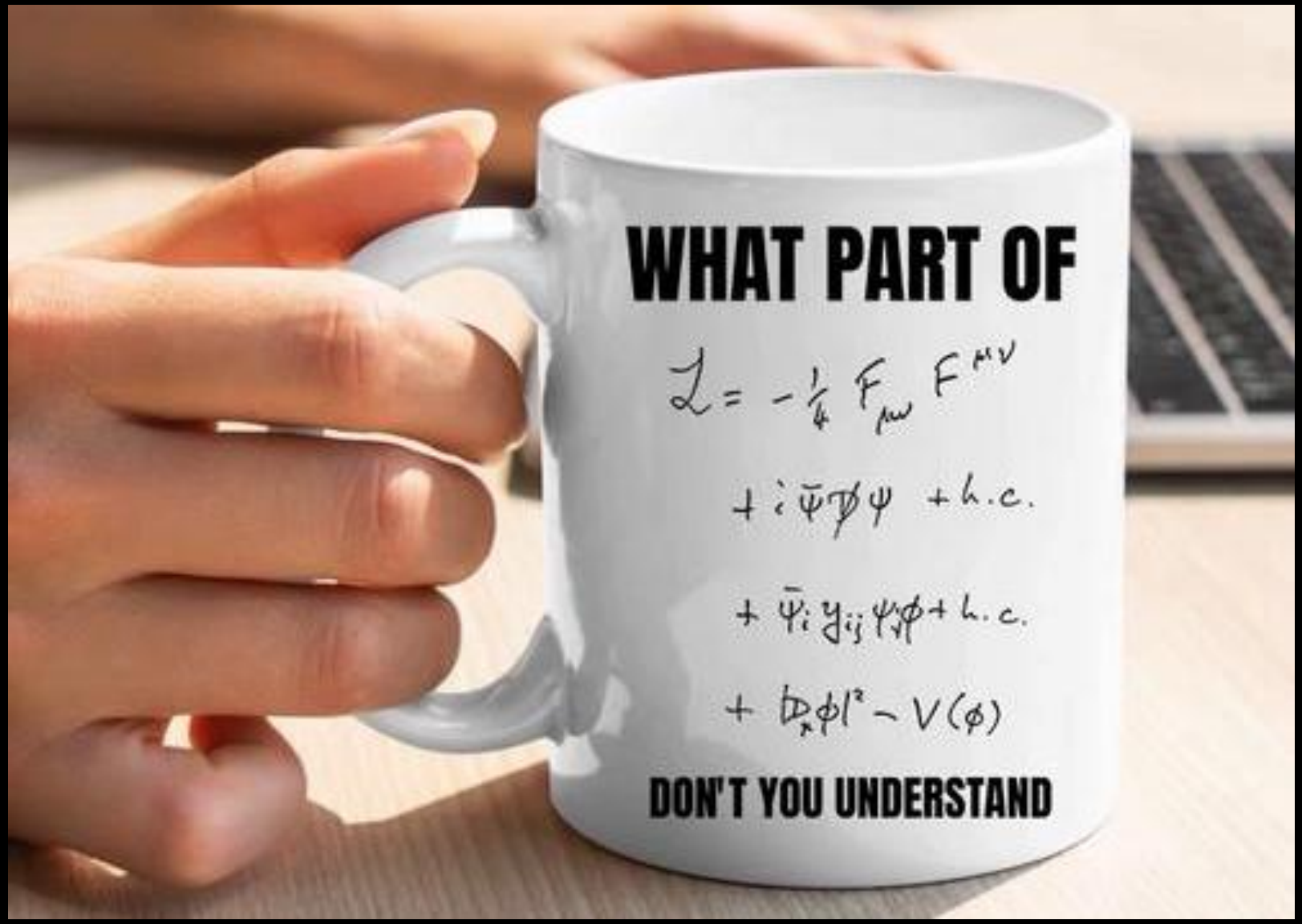
- 发现共性，产生规律
- 信息压缩 → 简单函数 → 智能

人类智能举例

经验和规律：

- 见到乌云预测将下雨
- 见到台阶要抬脚
- 围棋定式
- 1+1=2
- ……

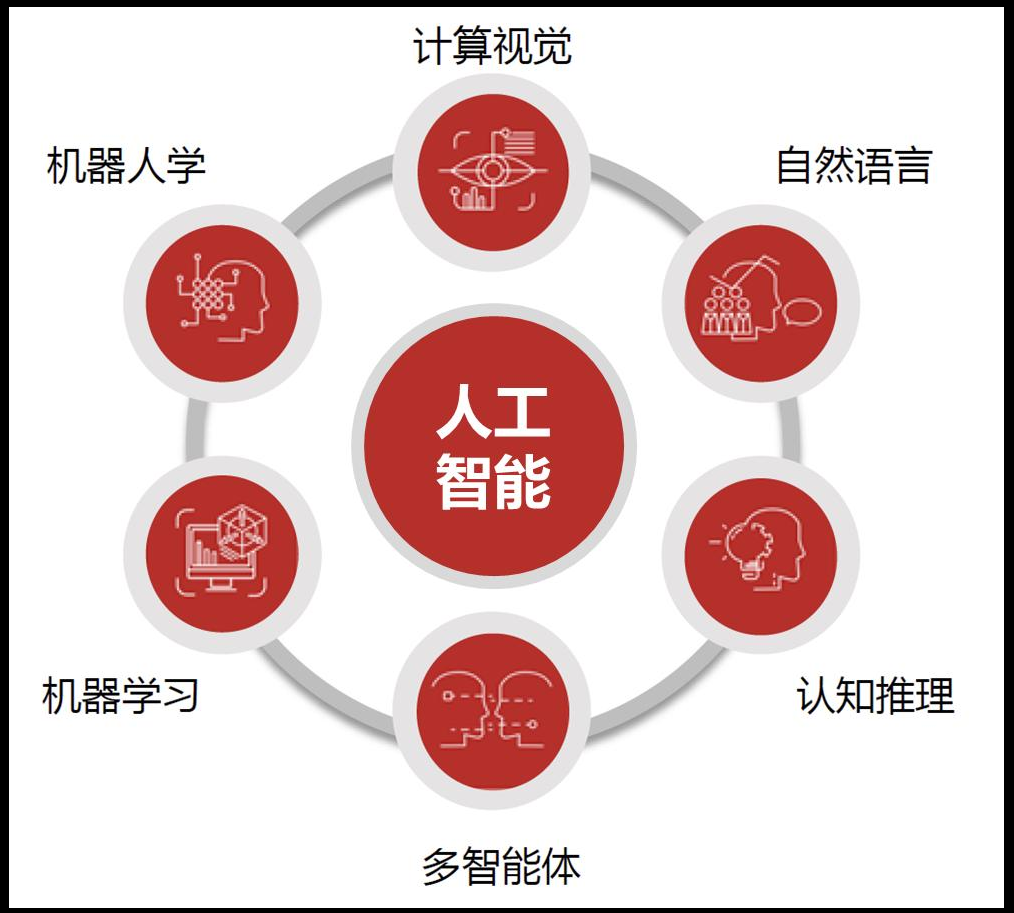
人类智能：科学（物理）规律探索



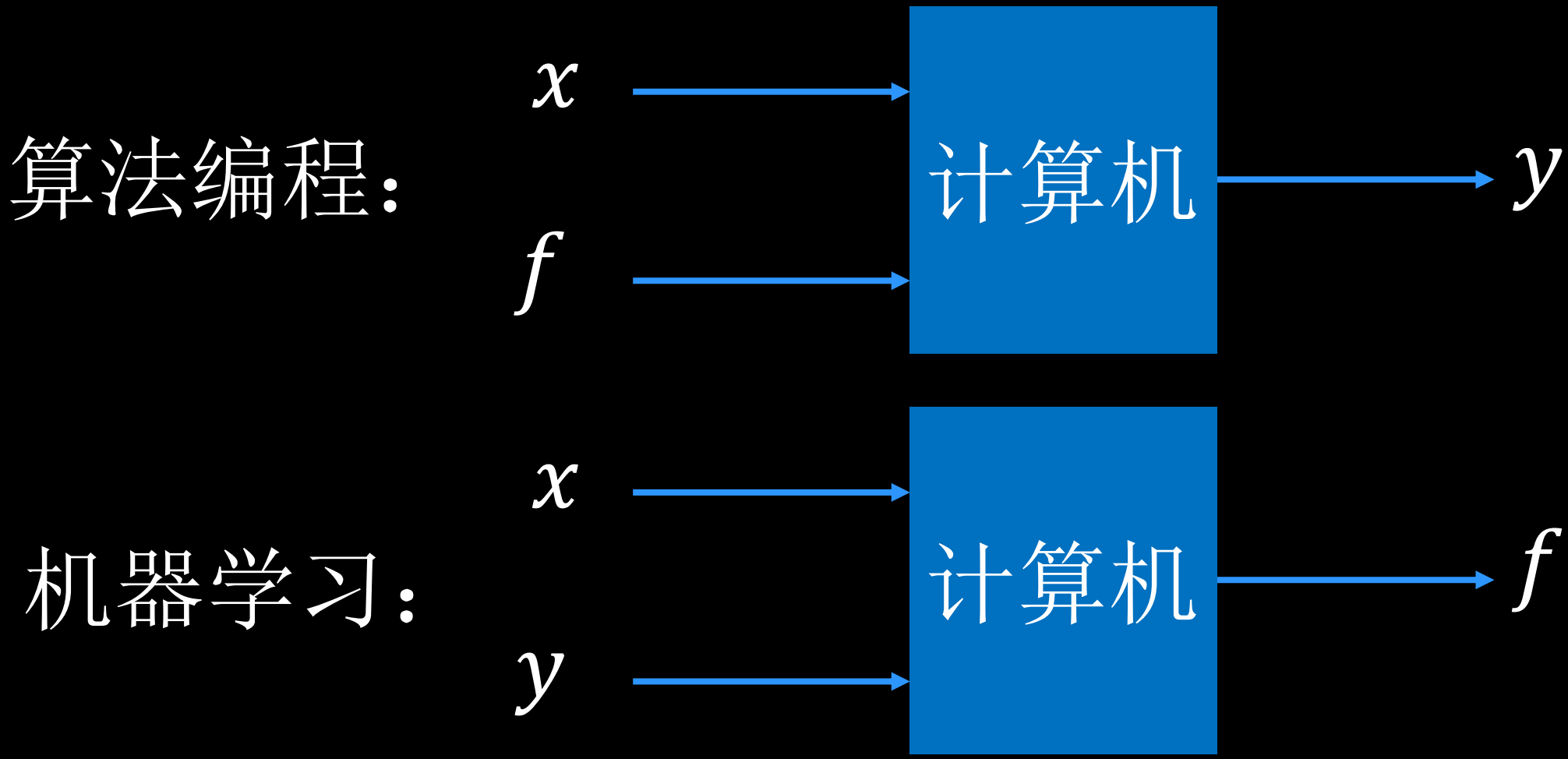
什么是人工智能？

- **人工智能**：是研究模拟、延伸和扩展人类智能的理论、方法、技术及应用系统的学科。





智能函数的来源



类比：解析计算方法 vs 数值模拟方法

算法编程

✓ 准确性、可靠性

× 使用局限性：需对问题有流程化的解决方案
× 人力成本高：每个问题需单独编程

机器学习

✓ 原则上可用于任意问题、人力成本低

× 缺少解释性、可靠性差

训练/学习

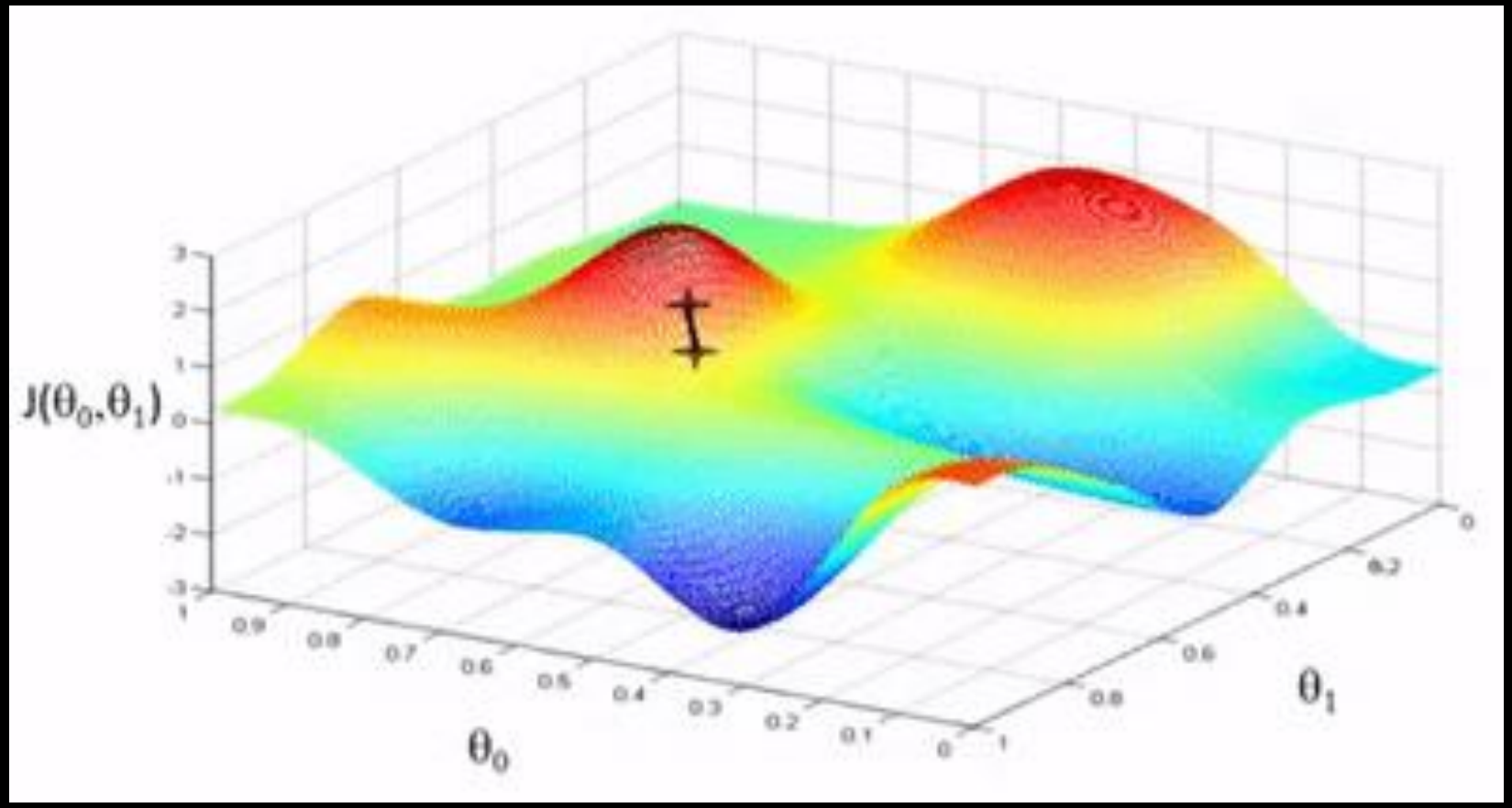
- 优化目标(Loss function)，如最小二乘法：

$$L(y, f_\theta(x)) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - f_\theta(x_i)\right)^2$$

- 从而$f_\theta(x) \approx f(x))$

梯度下降(GD)

每次沿梯度的相反方向走一个步长，直至目标函数收敛，或者达到预设的最大迭代次数。



简化模型：正则化

- 实际中一般在损失函数后加一项一起优化：

$$\min_{f} \frac{1}{n} \sum_{i \in [n]} L(f(x_i), y_i) + \lambda \cdot R(f)$$

- $\lambda \cdot R(f)$ 称为正则化项 (regularization term)，用于惩罚过于复杂的模型
- 模型越简单，发掘的共性越多，越有泛化能力
- 防止过拟合 (overfitting)，即测试误差远远高于训练误差

大纲

I. 疯狂的人工智能

II.什么是人工智能

III. 人工智能的基础及发展史

人工智能的流派

- **符号主义**(symbolism)（逻辑主义或计算机学派）：认为人工智能源于数理逻辑
  - 将符号作为人工智能的基本元素，认为人工智能的运行建立在由符号构成的数理逻辑之上
- **行为主义**(behaviorism)（进化主义或控制论学派）：认为人工智能源于控制论
  - 专注于主体与环境的相互作用，并将这种相互作用看作是智能行为
- **联结主义**(connectionism)（仿生主义或生理学派）：认为人工智能源于仿生学
  - 试图使机器模拟大脑，通过建立一个类似于人脑中神经元的模拟节点网络来处理信号
  - 人工神经元模型(1943)：联结主义的**开端**
  - 误差反向传播(BP)算法(1986)：联结主义的彻底**复苏**
  - 深度学习(2011，先语音识别后图像识别)：联结主义的迈向**高潮**

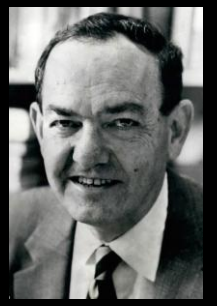
人工智能的基础——1、哲学

- 可以使用形式化规则得出有效结论吗？
  - 亚里士多德(BC384-BC322)-三段论
  - 鲁尔(1232-1315)试图用一组旋转纸盘构成的机械来实现一种推理系统
- 思维如何从物质大脑中产生的？
  - 笛卡尔(1596-1650)二元论：人类思维一部分处于自然之外，不受物理定律约束
  - 唯物主义(materialism)：大脑根据物理定律的运作构成了思维
- 知识从何而来？
  - 经验主义；逻辑实证主义；确证理论-将思维视为计算过程
- 知识如何导致行为？
  - 逻辑规划以实现确定目标；基于期望效用最大化的理性决策
  - 基于规则：“做正确的事情”不是由结果决定的，而由管理可行行为的普适社会法则决定

人工智能的基础——2、数学

- 得出有效结论的形式化规则是什么？
  - 形式化逻辑(formal logic)：古代中国、古希腊、古印度的哲学家
  - 布尔(1815-1864)：命题和布尔逻辑；弗雷格(1848-1925)：一阶逻辑
- 什么可以被计算？
  - 第一个非平凡的算法(algorithm)：欧几里得算法，计算最大公因数
  - 哥德尔(1906-1978)不完全性定理：演绎的限制确实存在
  - 图灵(1912-1954)：存在任何图灵机都无法计算的函数
  - 可计算性(computability)；易处理性(tractability)；NP完全性(NP-completeness)理论
- 如何使用不确定的信息进行推理？
  - 概率论(probability)：信息不确定下的广义逻辑；贝叶斯法则
  - 统计学(statistics)：概率的形式化结合数据的可用性

人工智能的基础——3、经济学

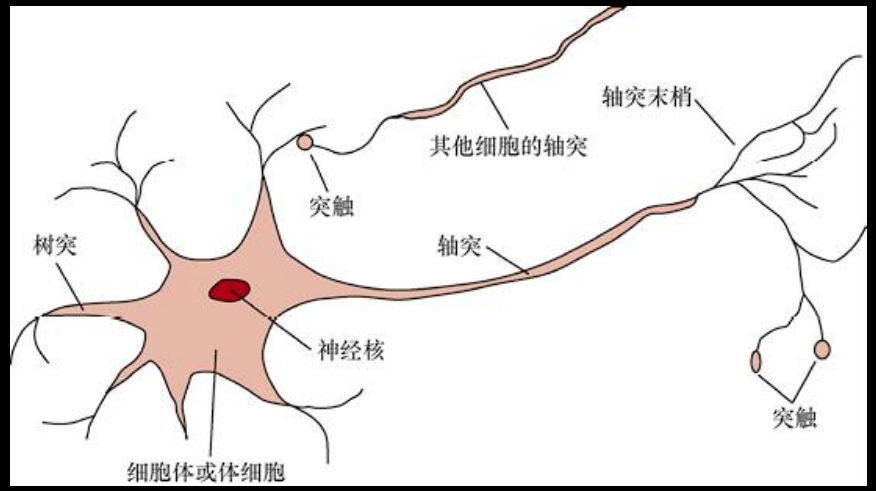


Herbert Simon
(1975年图灵奖)
(1978年诺贝尔经济学奖)

- 我们应该如何根据自己的偏好做出决定？
  - 经济学起源于1776年，亚当.斯密(1723-1790)《国富论》
  - 阿尔诺(1662)最大期望值公式；伯努利(1738)基于期望效用最大化准则，解释大众投资选择
  - 《博弈论与经济行为》(1944)：经济学不再是研究金钱的学科，而是对**欲望**和**偏好**的研究
- 当其他人可能不支持时，我们应该怎么做？
  - 决策论(decision theory)：概率论+效用理论，适用于“大型”经济体
  - 博弈论(game theory)：不对行为的选择提供明确的指示，适用于“小型”经济体
  - 多智能体系统(multiagent system)
- 当收益可能在很遥远的未来时，我们应该怎么做？
  - 运筹学(Operations research)：二战英国对雷达安装的优化
  - 强化学习(reinforcement learning)：马尔可夫决策过程，形式化贯序决策问题

人工智能的基础——4、神经科学

- 大脑如何处理信息？
  - 直到18世纪中叶，大脑才被广泛认为是意识所在地
  - 布卡罗(1861)发现大脑左半球一个局部区域负责语音的产生，开启大脑功能组织的研究
  - 高尔基，卡哈尔(1906诺奖)提出“神经元学说”：神经元结构的电化学反应产生认知功能
  - 伯杰(1929)脑电图仪(EEG)；功能磁共振成像(fMRI)；光遗传学(optogenetics)
  - 超级计算机summit(2017)在某些指标上已经与大脑相当，但需正确的理解智能的理论指导



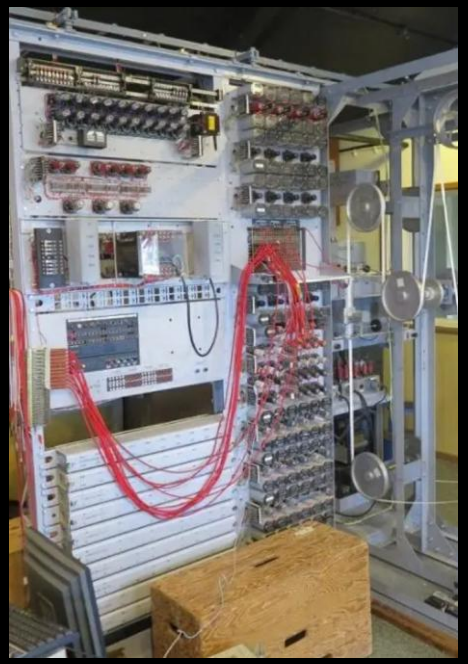
	超级计算机	个人计算机	人类大脑
计算单元	10^6 个 GPU + CPU	8 个 CPU 内核	10^6 列
	10^{15} 个晶体管	10^{10} 个晶体管	10^{11} 个神经元
存储单元	10^{16} 字节（10 PB）RAM	10^{10} 字节（10 GB）RAM	10^{11} 个神经元
	10^{17} 字节（100 PB）磁盘	10^{12} 字节（1 TB）磁盘	10^{14} 个突触
周期时间	10^{-9} 秒	10^{-9} 秒	10^{-3} 秒
运算/秒	10^{18}	10^{10}	10^{17}

人工智能的基础——5、心理学

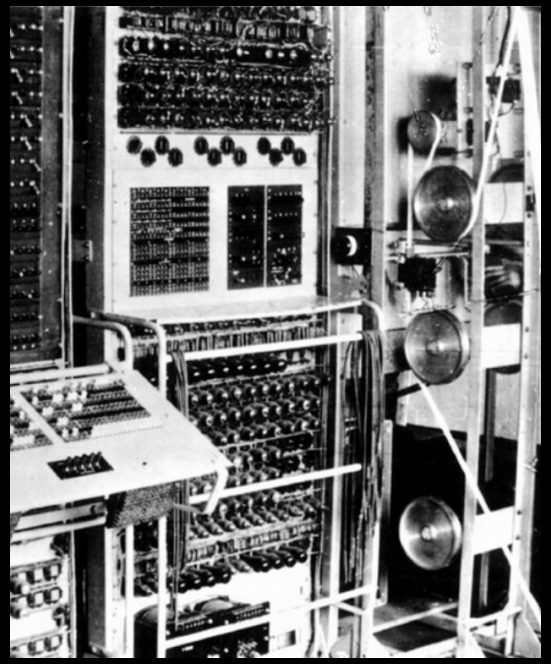
- 人类和动物是如何思考和行为的？
  - 科学心理学的起源：赫尔姆霍茨(1821-1894)和温特(1832-1920)，坚持严格控制的实验，进行感知/联想任务时，内省思维过程
  - **行为主义**(behaviorism)：只研究施加动物的感知/刺激及其产生的行为/反应的客观度量
  - **认知心理学**(cognitive psychology)：大脑是信息处理设备，感知涉及无意识的逻辑推断
  - **认知科学**(cognitive science)：从信息处理的角度来描述认知功能的运作，1956年9月MIT的一个研讨会，3篇论文展示了如何用计算机模型处理记忆、语言和逻辑思维的心理学问题
  - **人机交互**(human-computer interaction, HCI)先驱恩格巴特倡导**智能增强**(intelligence augmentation, IA)的理念而非AI
  - IA强调人类控制，AI强调机器智能

人工智能的基础——6、计算机工程

- 如何构建高效的计算机？
  - 二战期间3个国家的科学家们独立且几乎同时发明了现代数字电子计算机



Heath Robinson



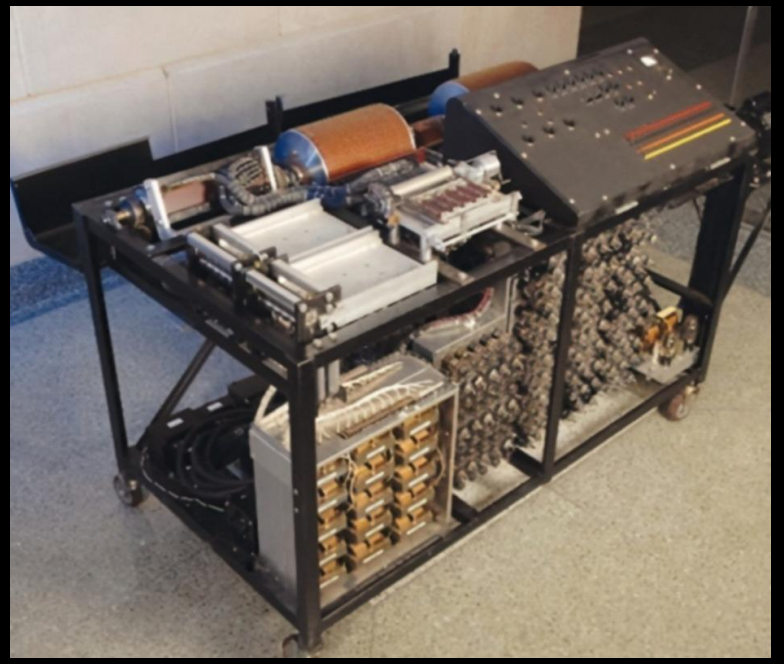
Colossus

英国(1943)



Z-3

德国(1941)

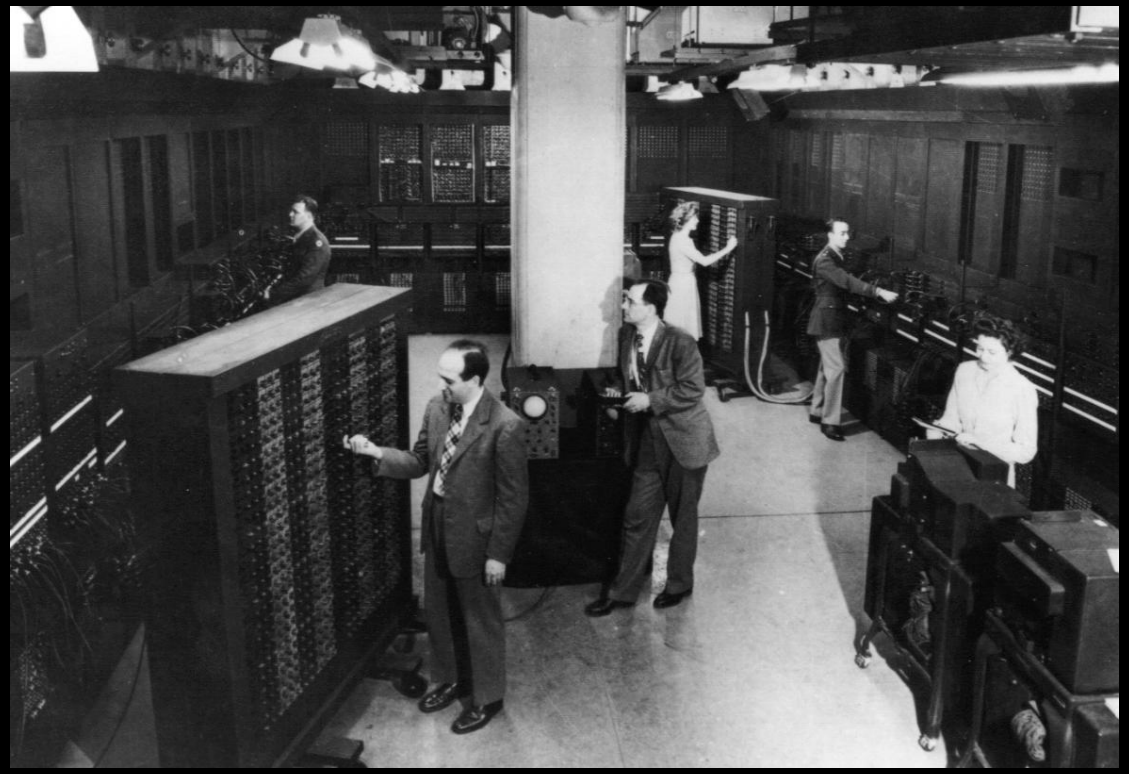


Atanasoff－Berry Computer, ABC

美国(1940-1942)

人工智能的基础——6、计算机工程

- 如何构建高效的计算机？
  - 第一台通用电子计算机ENIAC（Electronic Numerical Integrator And Computer）

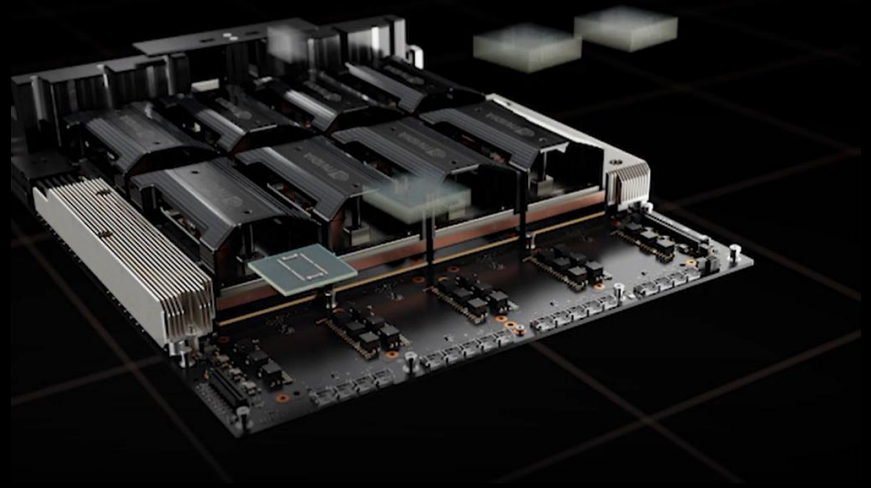


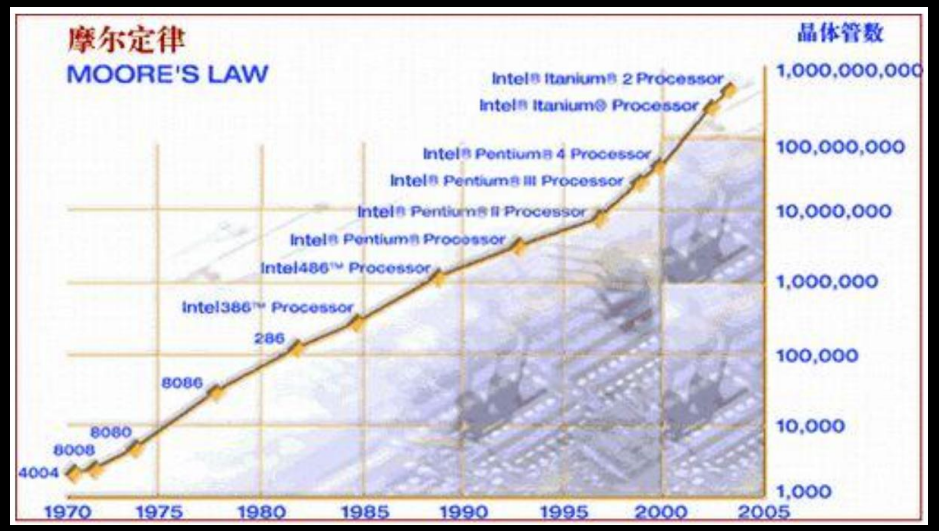
- ✓ 长30.48米，宽6米，高2.4米
- ✓ 占地面积约170平方米，30个操作台
- ✓ 重30英吨，耗电量150千瓦
- ✓ 包含17468根真空管（电子管）7200根水晶二极管，1500 个中转，70000个电阻器，10000个电容器，1500个继电器，6000多个开关
- ✓ 计算速度是每秒5000次加法或400次乘法

美国宾夕法尼亚大学
（1946年2月14日）

人工智能的基础——6、计算机工程

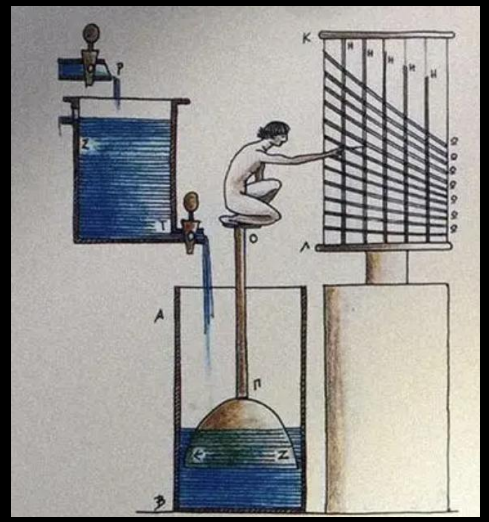
- 如何构建高效的计算机？
  - **摩尔定律**(Moore’s law)：GPU性能18个月翻一番
  - 2005年之后增加CPU核数而非提高CPU时钟频率，并行性对于性能提升更加重要
  - 图形处理单元(GPU)、张量处理单元(TPU)、晶圆级引擎(WSE)
  - 训练机器学习应用的计算能力：2012年后，约100天翻一番
  - AI还得益于计算机软件技术的发展：操作系统、编程语言和工具等

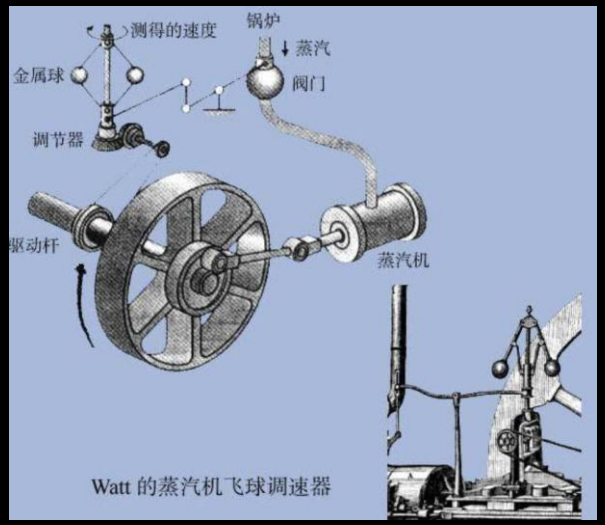




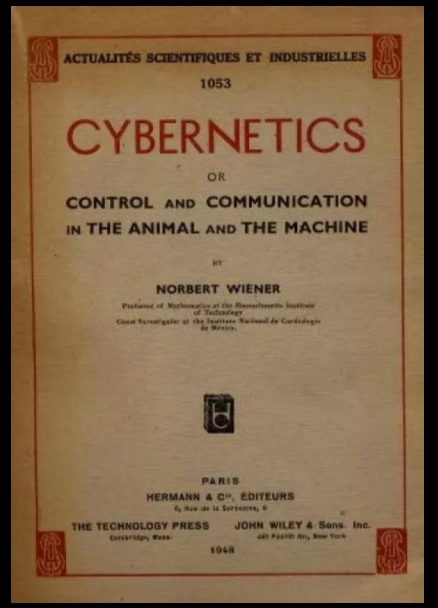
人工智能的基础——7、控制论

- 人造物如何在它们自己的控制下运行？
  - 古希腊克特西比乌斯(公元前250)建造了第一个自我控制的机器：水钟；自调节反馈控制
  - **控制理论**(control theory)：维纳《控制论》(1948)，有目的的行为源于试图最小化“错误”的调节机制，即当前状态和目标状态之间的差异
  - 艾什比《大脑设计》：通过自我平衡(homeostatic)设备来实现智能，反馈回路
  - **现代控制理论**：设计随时间最小化**代价函数**(cost function)的系统





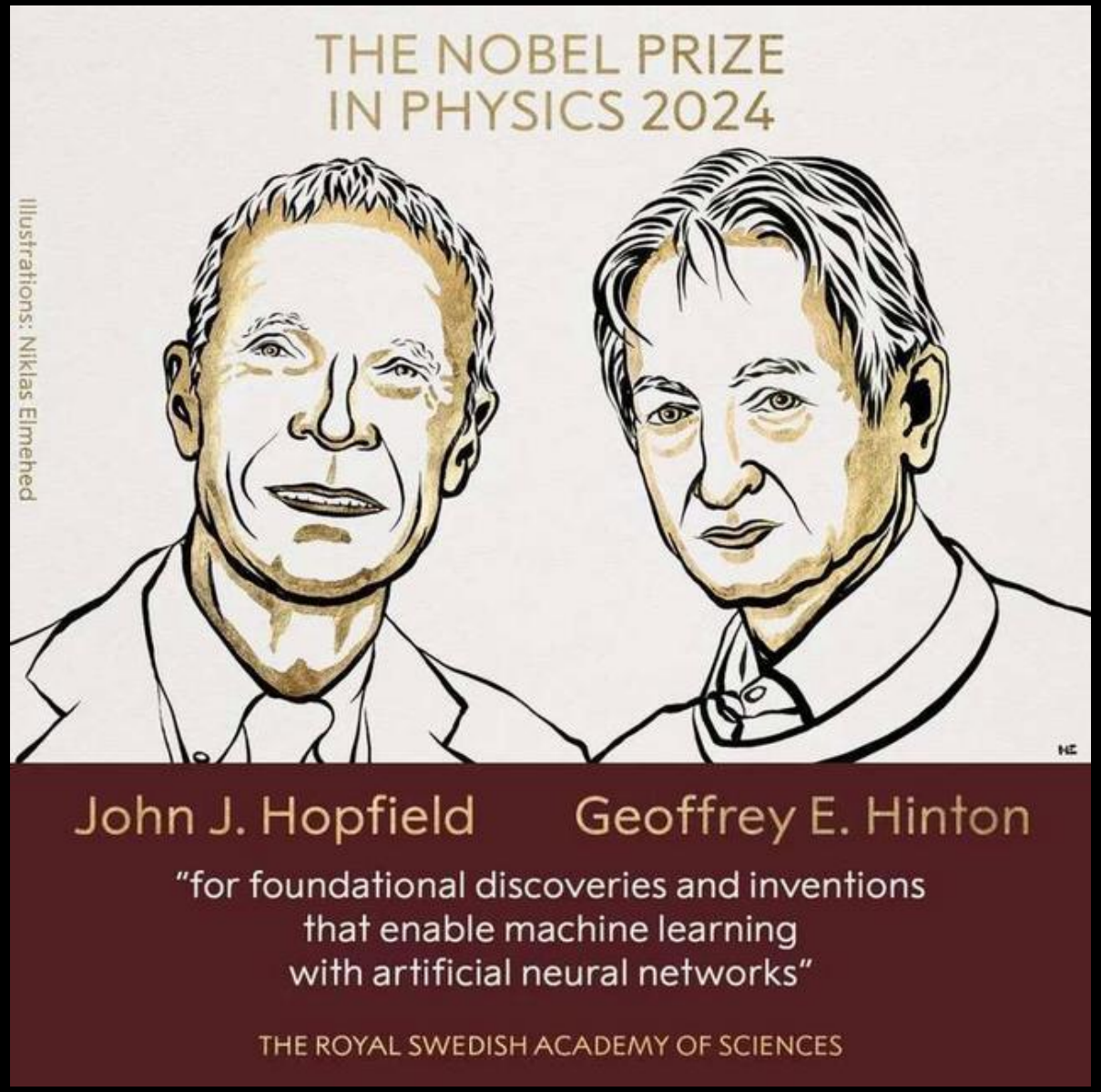
Watt 的蒸汽机飞球调速器



人工智能的基础——8、语言学

- 语言是如何与思维联系的？
  - 乔姆斯基《句法结构》(1957)：句法模型为基础的理论，可以解释语言创造力，足够形式化，原则上可被程序化
  - 现代语言学
    - **计算语言学**(computational linguistics)
    - **自然语言处理**(natural language processing)
  - 理解语言需要理解主题和上下文，而不仅仅是理解句子结构
  - **知识表示**(knowledge representation)：如何将知识转化为计算机可以推理的形式的研究

人工智能的基础——9、物理学？？？



在人工神经网络机器
学习方面的基础性贡献

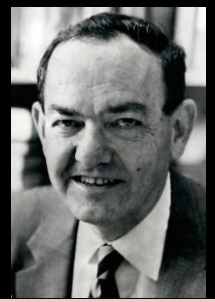
人工智能的历史——AI图灵奖得主

提出关于问题求解和
人类认知的**符号模型**

Allen Newell (1975年获奖)



Herbert Simon (1975年获奖)



提出处理不确定性的
概率因果推理技术

Judea Pearl (2011年获奖)





Marvin Minsky (1969年获奖)



John McCarthy (1971年获奖)

定义了基于**表示**和**推理**
的领域基础



Ed Feigenbaum (1994年获奖)



Raj Reddy (1994年获奖)

开发对知识编码
来解决真实问题的**专家系统**

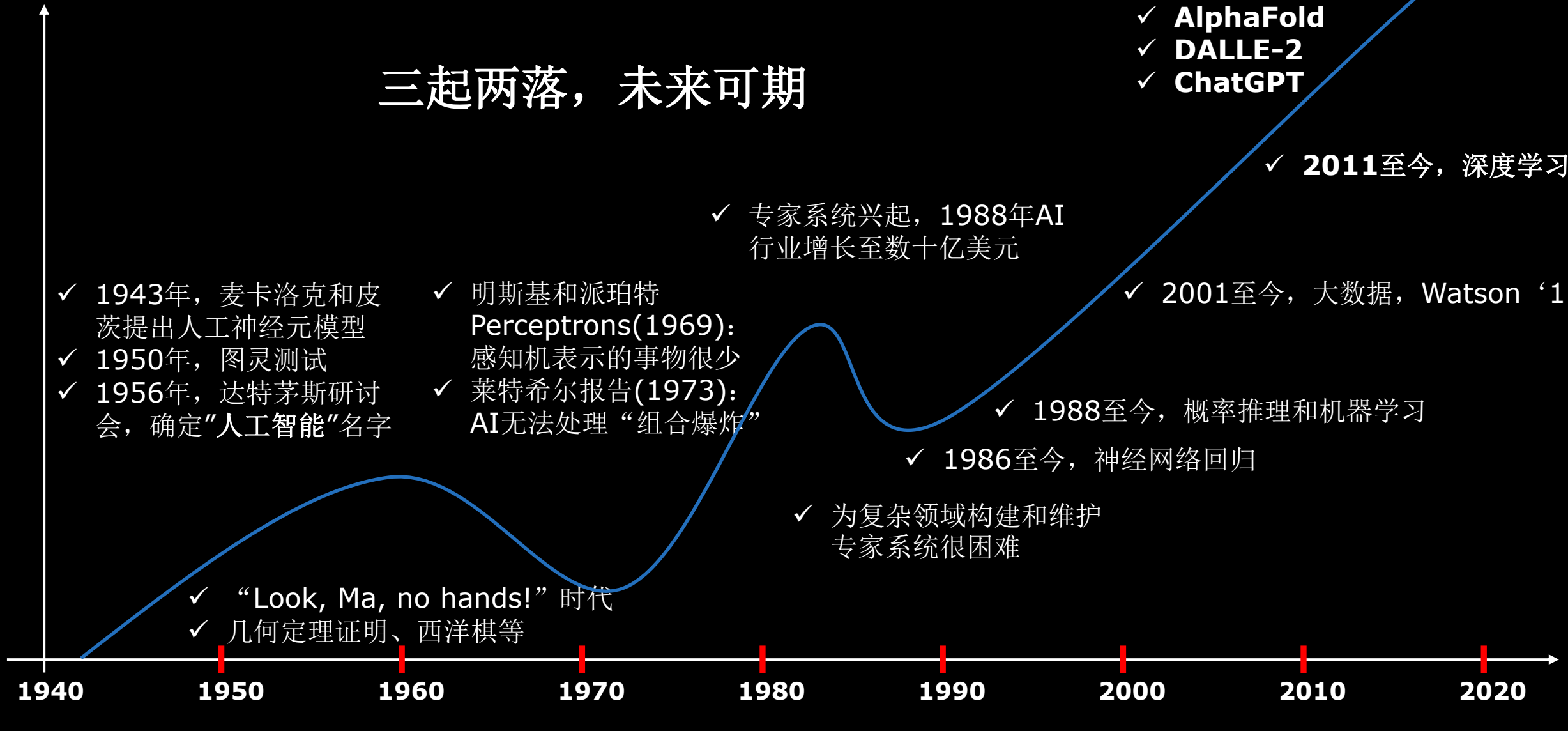


Y. LeCun G. Hinton Y. Bengio (2018年获奖)

深度神经网络
深度学习

人工智能的历史

三起两落，未来可期



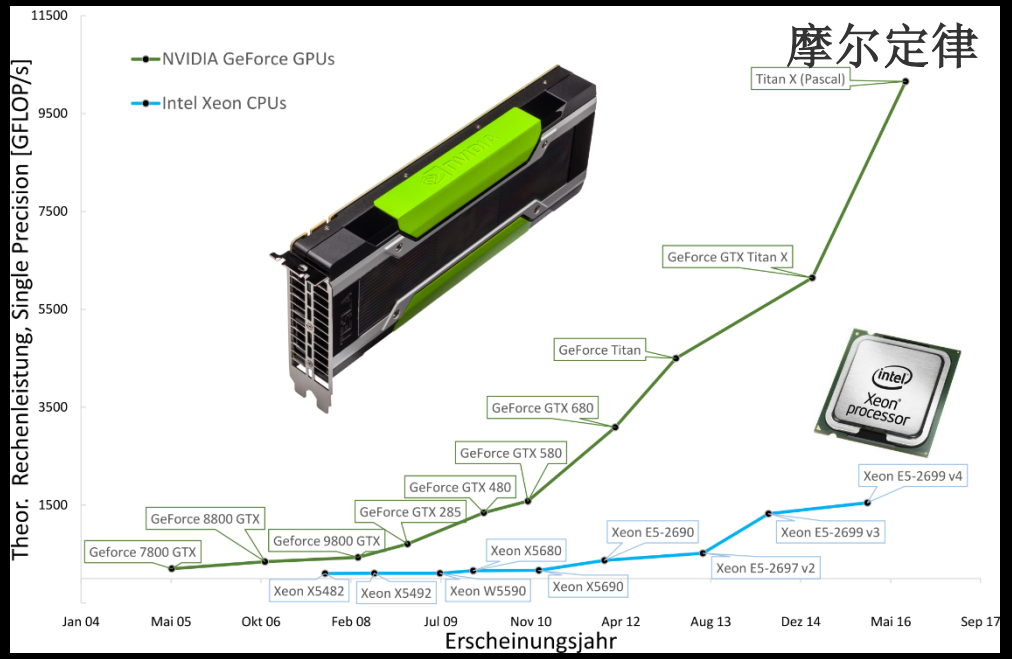
早期智人复原图（约20万年前）
智人脑容量已有1300-1600毫升



大数据+大算力



2009年, ImageNet数据集引发了计算机视觉领域的一场革命



CPU: 10^9~10^{10}次运算/s
GPU/TPU上的DL: 10^{14}~10^{17}次运算/s

语言基础：AI前景——deepseek

结论：**AI的未来不是一条直线，而是一条在反复校准中持续向上的曲线。2026年的主题是“去虚向实”，淘汰浮夸，沉淀价值，为下一轮质的飞跃积蓄深厚基础。**

- 整体态势：AI已跨越“概念验证”阶段，进入“工程化”和“价值兑现”的关键期。技术迭代速度未减，但发展重心从“能做多酷”转向“能用多好”。
- 技术拐点：参数规模的边际收益显著递减，行业焦点转向世界模型和多模态融合。AI开始学习物理规律，具备空间想象和因果推理能力，这为机器人、自动驾驶和科学研究打开了新空间。
- 资本流向：投资总量保持高位，但内部出现剧烈分化。通用大模型的融资热度降温，资金加速涌入具身智能、AI4S（科学智能） 和行业垂直模型。泡沫正在被挤压，资源向能解决实际问题的方向集中。
- 落地加速：在制造业，AI已深度介入产线质检、工艺优化和供应链调度，成为“新质生产力”的核心引擎；在医疗健康，AI辅助诊断和新药研发进入临床实用阶段，部分场景已取代传统人工流程；在科研领域，AI正以“加速器”角色推动材料、生物、化学等基础学科的突破。
- 效率与安全并重：小模型和端侧AI崛起，用更低的能耗和成本实现接近大模型的性能。同时，可解释AI、对抗性防御和事实校验机制成为刚需，安全从“附加项”变为“必选项”。
- 人机关系重塑：AI Agent从“助手”进化为“执行者”，能独立完成跨应用、多步骤的复杂任务。人与AI的关系正从“使用工具”转向“组建团队”。

GDP理论——智能时代生存法则

➢ 核心判断

- 在未来几年，AI会保持研究伙伴的角色（未达到全自动化）
- 用户使用AI做研究，与导师带研究生做研究类似

➢ GDP is all you need——马滟青

G—沟通能力：与AI沟通讨论、布置任务

D—洞察能力：问关键问题、从大量产出中捕捉到潜在的错误、合理路线

P—品味能力：决定做什么事

谢谢！

《物理与人工智能》课程

- **课程概括**
  - 研本课程：3学分
  - 考核含：AI基础的平时作业（40%）+专题部分的考核（40%，每个专题5分，任课老师可采用签到、随堂练习、课后作业等形式）+期末大作业（20%）
- **课程内容**
  - 人工智能基础
  - 物理与人工智能的关联（物理中的应用及物理对其的推动）

课程网站：https://aiphy.pku.edu.cn/

课程安排

第1周（9月7日）AI概览、LLM概览、API配置和Python开发实践	第2周（9月14日）LLM的局限性、Agent工作流、RAG架构	第3周（9月21日）提示工程、Tool calling、项目实践	第4周（9月28日）机器学习基础、逻辑回归、决策树与随机森林	第5周（10月12日）神经网络与优化、CNN、GAN、扩散模型
第6周（10月19日）RNN、Transformer、项目实践	第7周（10月26日）盲目搜索、启发式搜索、对抗搜索、蒙卡搜索	第8周（11月2日）AI与凝聚态物理-陈基	第9周（11月9日）AI 代理模型与天体物理-王力乐	第10周（11月16日）AI与核技术-林晨
第11周（11月23日）AI与量子多体物理-张亿	第12周（11月30日）AI在对撞机实验中的应用-朱华星	第13周（12月7日）AI 与高能核物理-宋慧超	第14周（12月14日）AI与大规模数值模拟-冯旭	第15周（12月21日）AI for discovery：进展、困难及未来