



中国科学院大学
University of Chinese Academy of Sciences

计算机科学导论

学科展望

徐志伟 zxu@ict.ac.cn

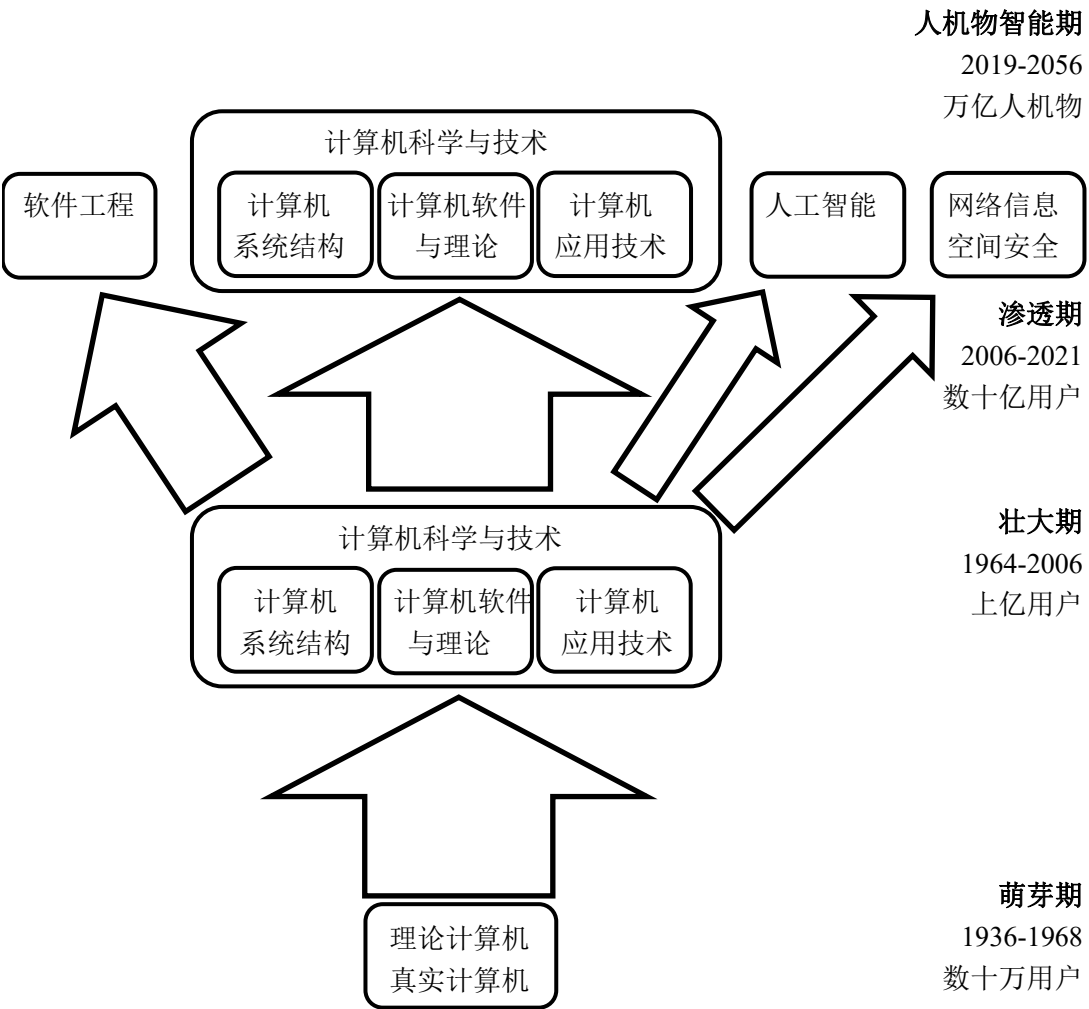
提纲

- 计算机学科展望
 1. 计算机学科发展的演化树
 2. 格雷**12**问题
 3. 科学研究的第五范式
 4. **21**世纪文明需要创新型人才

课件中可能包含素材引用，特此致谢！

1. 计算机学科发展的演化树

- 1936-2056，120年，大体上可分为四个时期



计算机学科发展演化树

- 1936-2022，前三个时期

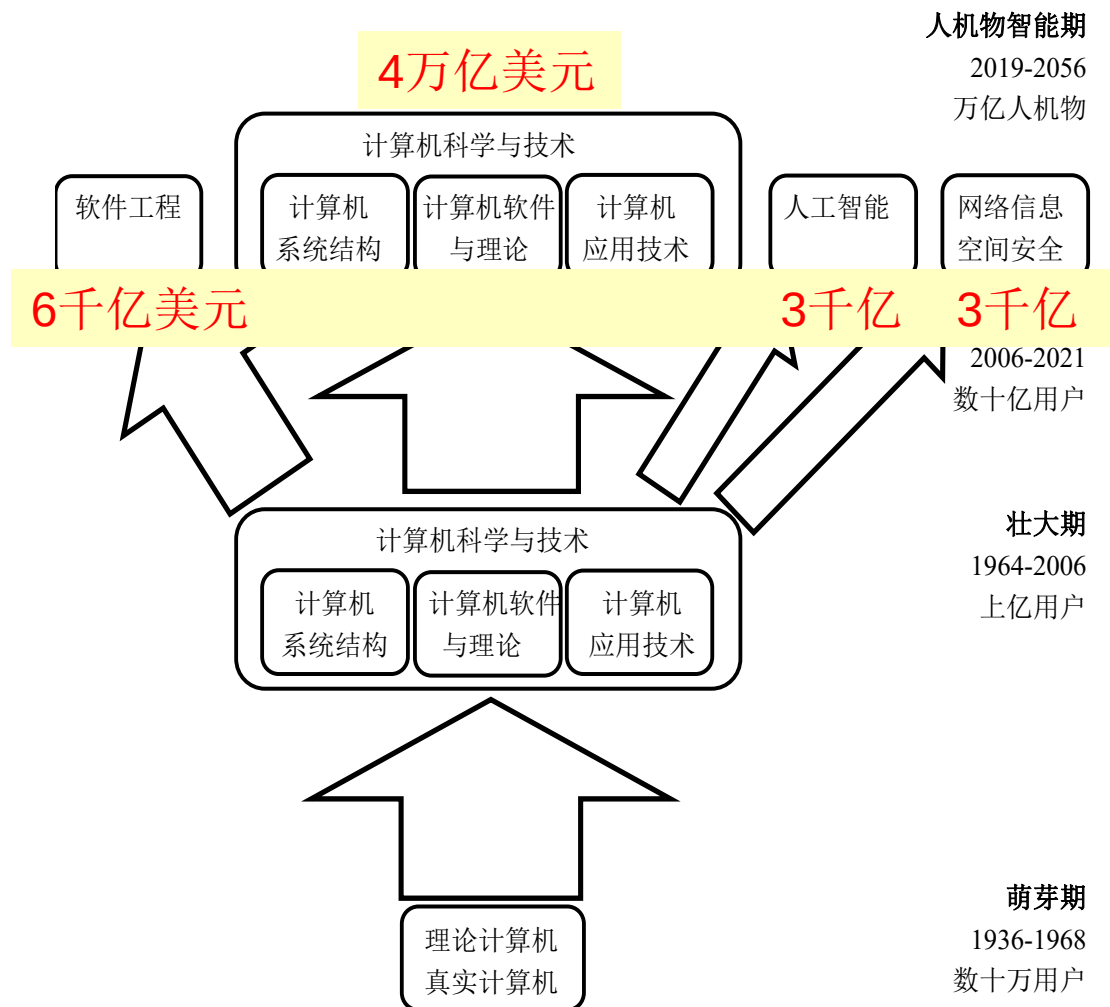
- 全球市场

2022: ~ 5.3 US\$ Trillions
5万亿美元
全球GDP=100万亿美元的5%

2000: ~ 1 US\$ Trillion
1万亿美元
全球GDP=34万亿美元的3%

1960: ~ 1 US\$ Billion, 十亿美元
全球GDP=1.4万亿美元的0.07%

1950: ~ US\$ Millions



计算机学科发展演化树

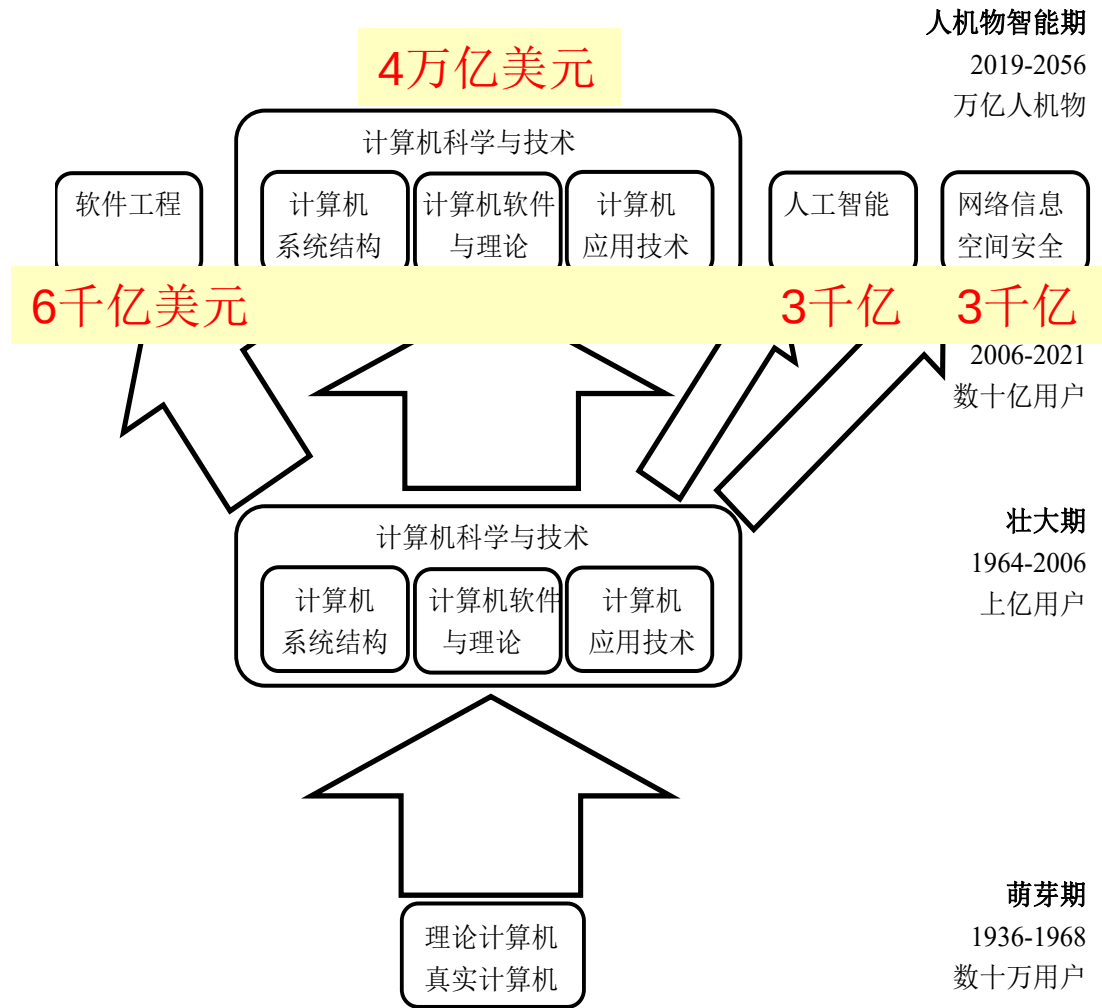
● 1936-2022，前三个时期

● 全球市场

- 2022: ~ 5.3 US\$ Trillions
5万亿美元
全球GDP=100万亿美元的5%
- 2000: ~ 1 US\$ Trillion
1万亿美元
全球GDP=34万亿美元的3%
- 1960: ~ 1 US\$ Billion，十亿美元
全球GDP=1.4万亿美元的0.07%
- 1950: ~ US\$ Millions

服务业提升
是大势所趋

产业占GDP比例	1992	2022
第三产业（服务业）	28%	53%
第二产业（工业）	42%	40%
第一产业（农业）	24%	7%

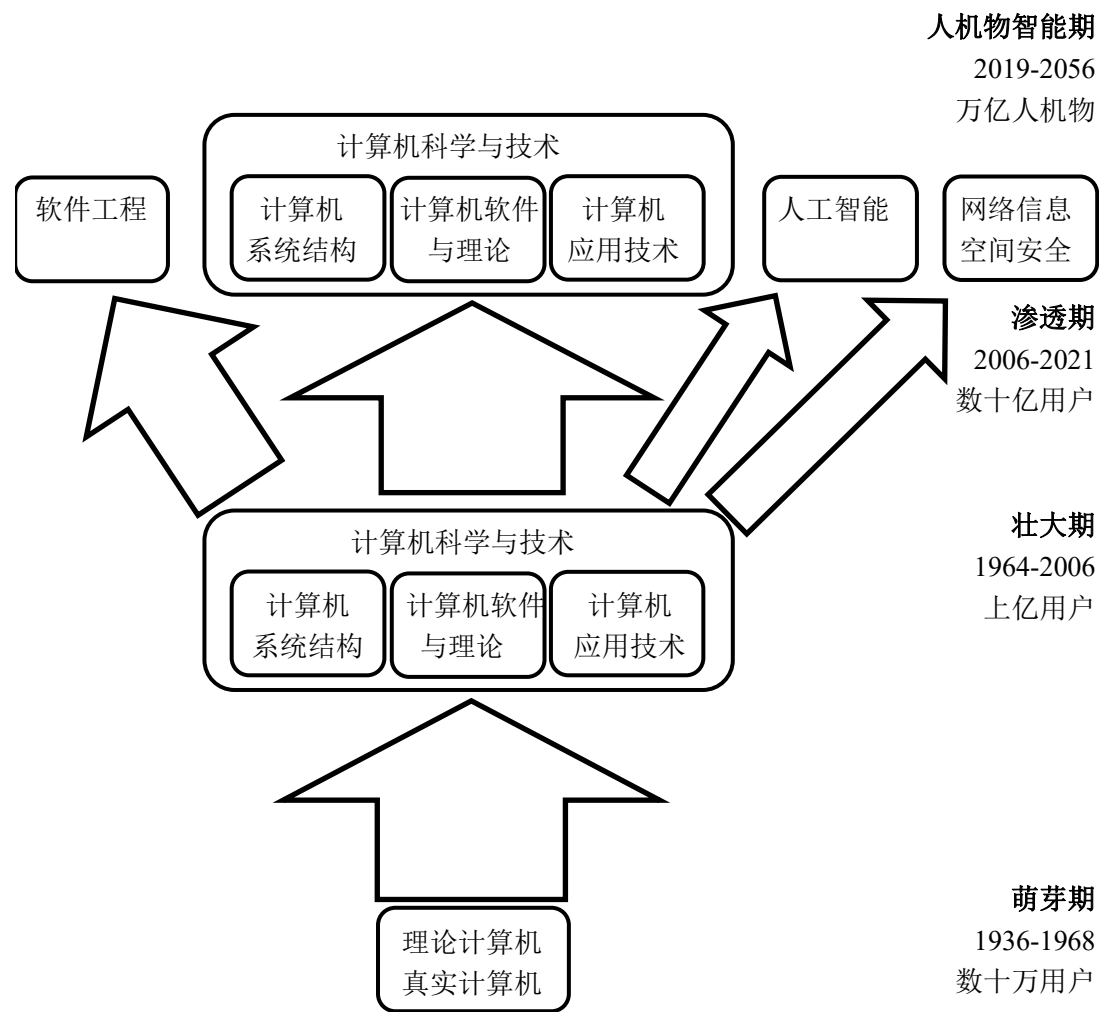


计算机学科发展演化树：萌芽期

- 1936：图灵机论文发表
- 1968：高德纳的《计算机程序设计艺术》教科书出版
 - 20世纪最有影响的12部科学专著之一

- 重要进展：物登明堂，裔裔皇皇

- 通用的真实计算机出现
 - 冯诺依曼体系结构
 - 操作系统
 - 高级程序设计语言
- 计算机应用涌现
 - 科学计算、计算机模拟仿真、计算机图形学、计算机过程控制、计算机信息管理系统等
- 学科生态初步形成
 - 普度大学创立了全球第一个计算机专业
 - ACM、IEEE-CS、中国计算机学会成立



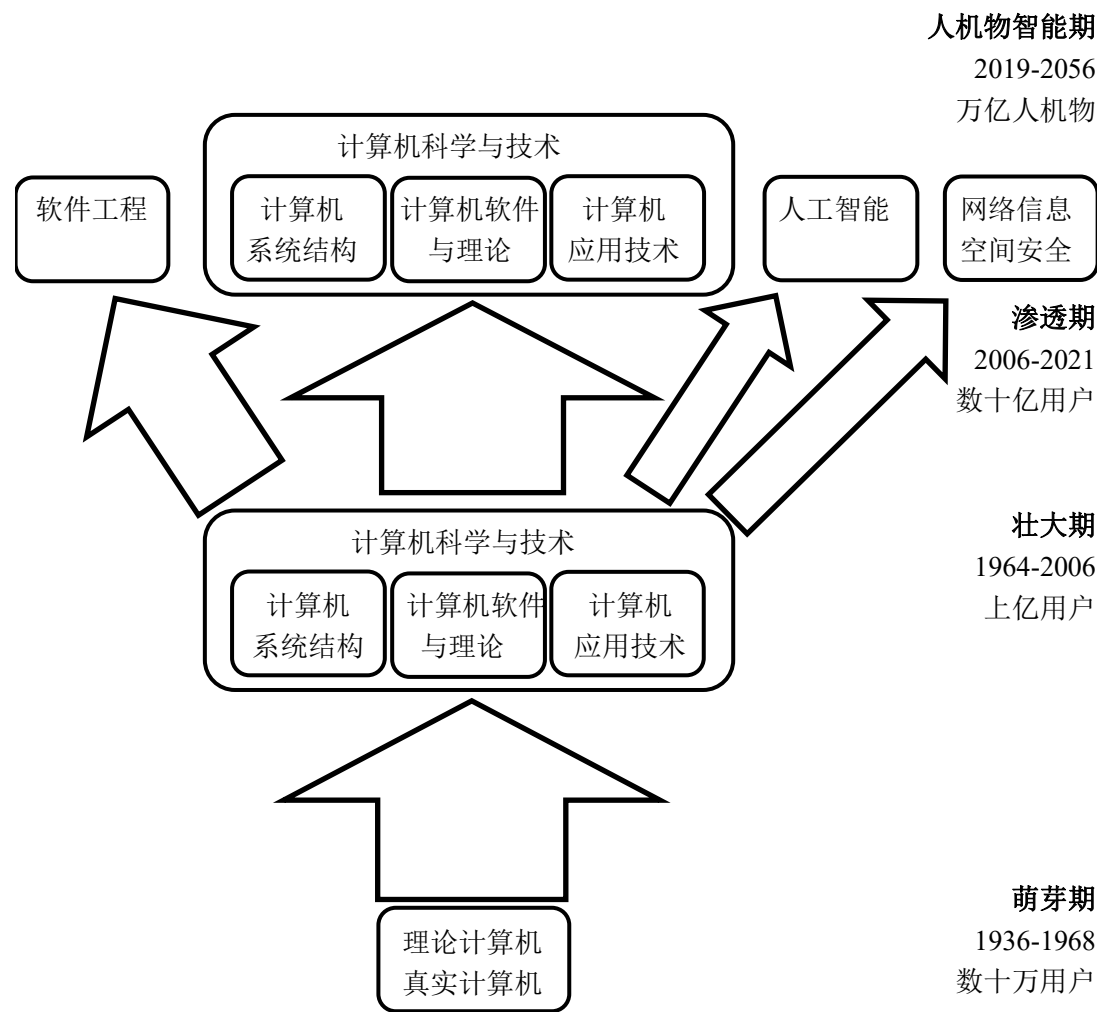
20世纪最有影响的12部科学专著

Philip, and Phylis Morrison. "100 or so BOOKS That Shaped a CENTURY of Science." American Scientist 87, no. 6 (1999): 542–53.

- 狄拉克：量子力学
- 爱因斯坦：瑞士论文集
- 曼德布罗特：分形
- 鲍林：化学键性质
- 罗素和怀特海德：数学基础
- 冯诺依曼和摩根斯坦：博弈论与经济行为
- 维纳：控制论
- 伍德沃和霍夫曼：轨域对称守恒
- 爱因斯坦：相对论的含义
- 费曼：量子电动力学
- 高德纳：计算机程序设计艺术
- Paul Dirac, Quantum Mechanics (1930)
- Albert Einstein, The Collected Papers of Albert Einstein: The Swiss Years: Writings, 1902–09 (1930)
- Benoit B. Mandelbrot, Fractals (1977)
- Linus Pauling, Nature of the Chemical Bond (1939)
- Bertrand Russell and Alfred North Whitehead, Principia Mathematica (1910–13, 3 vols.)
- Cyril Smith, Search For Structure (1981)
- John von Neumann and Oskar Morgenstern, Theory of Games and Economic Behavior (1944)
- Norbert Wiener, Cybernetics (1948)
- R. B. Woodward and Roald Hoffmann, Conservation of Orbital Symmetry (1970)
- Albert Einstein, The Meaning of Relativity (1922)
- Richard Feynman, QED (1985)
- Donald Knuth, The Art of Computer Programming (1968)

计算机学科发展演化树：壮大期

- 1964年：IBM发布S/360通用计算机系列
- 2006年：周以真发表“计算思维”论文
- 重要进展：以一耦万抽象栈成型
 - 计算局部性原理
 - P vs. NP，NP完备性
 - 分布式计算理论
- 计算机系统结构概念
 - 真实计算机**定性研究**
- 计算机体系结构：**量化研究**方法
 - 真实计算机定量研究
- 基础软件与应用软件
- 应对“软件危机”的“软件工程”
- 计算机网络
- 万维网



计算机学科发展演化树：壮大期

- 1964年：IBM发布S/360通用计算机系列
- 2006年：周以真发表“计算思维”论文
- 重要进展：以一耦万抽象栈成型
 - 计算局部性原理
 - P vs. NP，NP完备性
 - 分布式计算理论
 - 计算机系统结构概念
 - 真实计算机**定性研究**
 - 计算机体系结构：**量化研究**方法
 - 真实计算机定量研究
 - 基础软件与应用软件
 - 应对“软件危机”的“软件工程”
 - 计算机网络
 - 万维网

Hennessey & Patterson提出CPI公式

$$\frac{\text{Execution Time}}{\text{Program}} = \frac{\text{Instructions}}{\text{Program}} \times \frac{\text{Clock Cycles}}{\text{Instruction}} \times \frac{\text{Clock Time}}{\text{Clock Cycle}}$$

程序执行时间 = 程序动态指令数 × CPI / 主频

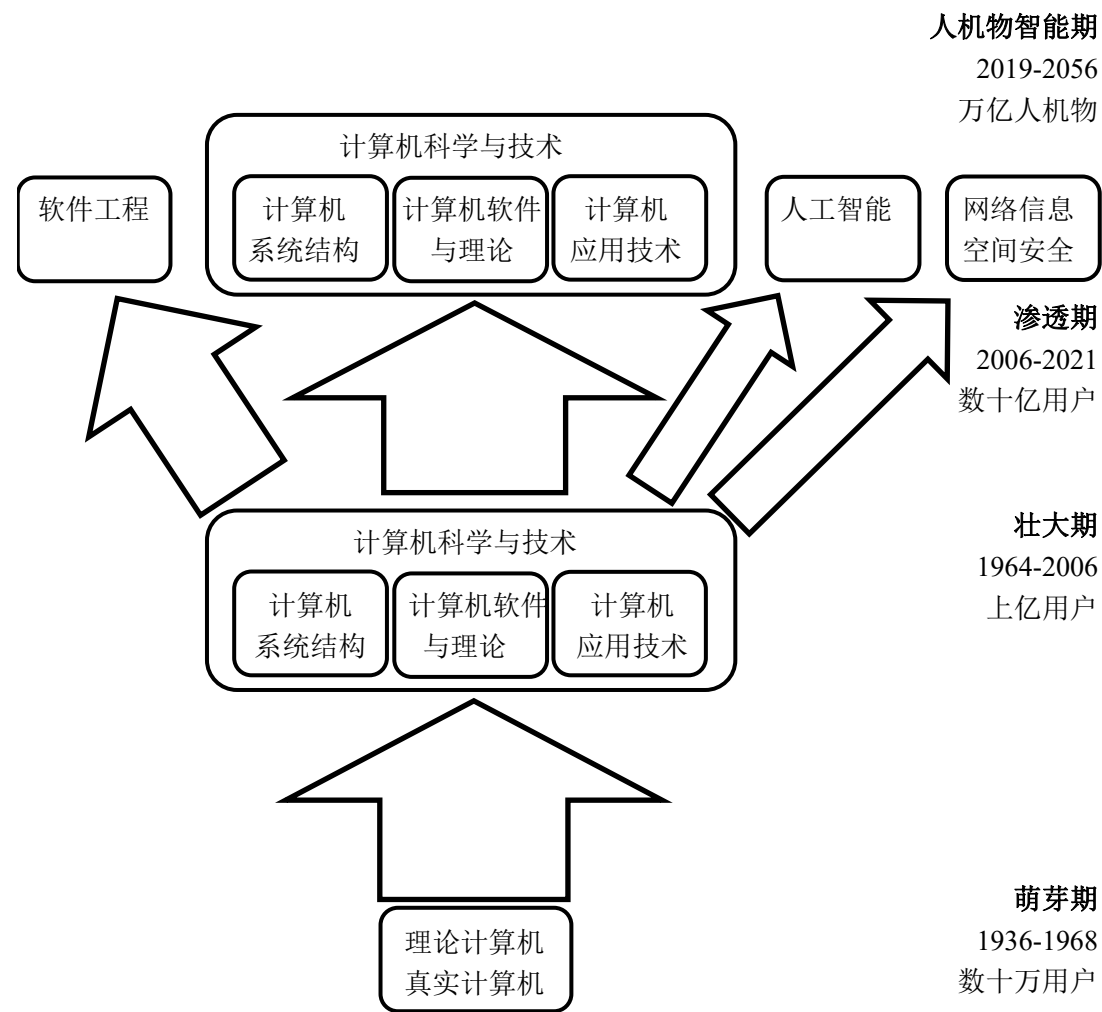
$$\text{处理器体系结构} = \begin{cases} \text{CISC 复杂指令集体系结构} & \text{CPI} > 1 \\ \text{RISC 精简指令集体系结构} & \text{CPI} = 1 \\ \text{Superscalar 超标量体系结构} & \text{CPI} < 1 \end{cases}$$

计算机学科发展演化树：渗透期

- 2006年：周以真发表“计算思维”论文
- 2021年：抖音成为全球访问量最大的网站，超过谷歌
 - 发展中国家也起到重要作用

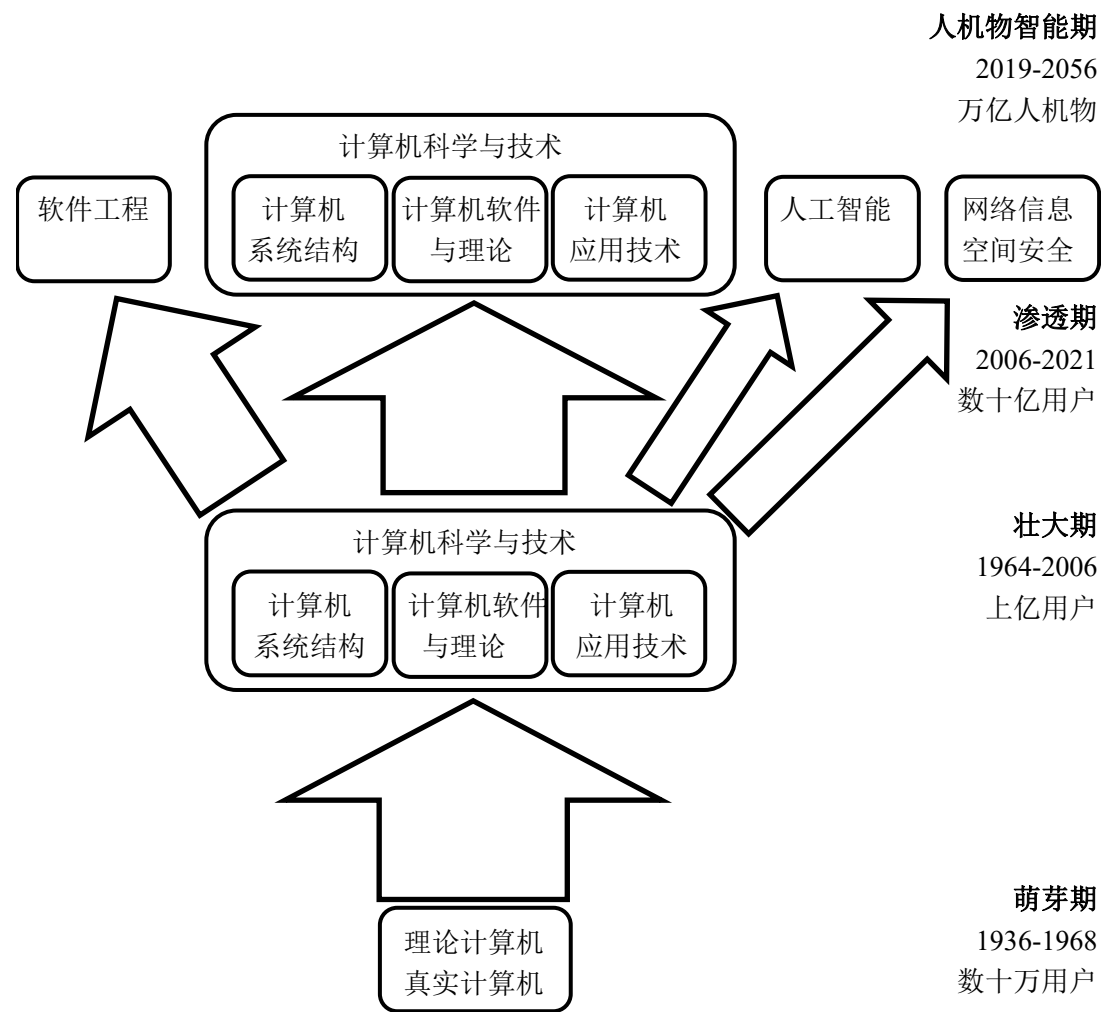
- 重要进展：计算机科学影响全球

- 计算思维愿景
- 移动互联网、云计算、大数据、人工智能等计算机技术迅猛发展，得到了大规模应用
- 渗透到了发展中国家
- 互联网全球渗透率
 - 2000: 6%
 - 2021: 59.5%
- 抖音（TikTok）成为2021年全球访问量最大的互联网网站，超过2020年冠军谷歌



计算机学科发展演化树：人机物智能期

- 2019年：业界提出**万亿设备新世界**愿景
(the Trillion-Device World)
 - 2018年**物端设备**数量超过**人端设备**
- 2056年：图灵机论文发表**120周年**
- 主要特征
 - 计算从赛博空间拓展到人机物三元计算世界
 - 计算机不只是电子计算机，它的部件还包括人和物
 - “人机物”融合的智能万物互联时代



2. 格雷12问题，对标数学的希尔伯特23问题

2050年前做到

- ① 可扩展系统（**Scalability**）：设计出算力可扩展1百万倍的系统结构。
- ② 图灵测试（**Turing Test**）：设计出可通过图灵测试的计算机系统。
- ③ 母语听（**Speech to text**）：构建计算机系统，能够像说母语的人一样听懂人讲话。
- ④ 母语说（**Text to speech**）：构建计算机系统，能够像说母语的人一样讲话。
- ⑤ 真人看（**See as well as a person**）。构建计算机系统，能够像人一样识别事物行为。
- ⑥ 个人数字资产库系统（**Personal Memex**）。记录个人一生所读、所看、所听，并能快速检索，但不做任何分析。消费者个人能够负担其购买成本与使用成本。
- ⑦ 全球数字资产库系统（**World Memex**）。所有文本、声音、图像、视频上网；专家级的分析和摘要能力；快速检索能力。
- ⑧ 远程呈现（**TelePresence**）。模拟人出现在另一个地方，能与当地环境和其他人交互，好像真实出现一样。
- ⑨ 无故障系统（**Trouble-Free Systems**）。构建计算机系统，能够供百万人日常使用，但只需一个兼职人员管理维护。
- ⑩ 安全系统（**Secure System**）。保障上述无故障系统只对授权用户开放、非法用户不能阻碍合法用户使用、信息不可能被窃取。证明这三种安全性。
- ⑪ 系统可用性（**AlwaysUp**）。保障上述无故障系统的可用性高达99.9999999%（9个9），即每100年才出错1秒。证明这种可用性。
- ⑫ 自动程序员（**Automatic Programmer**）。设计出一种规约语言或用户界面，具备5个性质：（1）通用，能够表达任意应用的设计；（2）高效，表达效率提升一千倍；（3）自动，计算机能够自动编译该设计表达；（4）易用，系统较为易用；（5）智能，可针对应用设计中存在的异常与缺失自动做推理、询问用户。

格雷12问题：性能

- ① **可扩展系统（Scalability）**：设计出算力可扩展**1百万倍**的系统结构。
- ② 图灵测试（Turing Test）：设计出可通过图灵测试的计算机系统。
- ③ 母语听（Speech to text）：构建计算机系统，能够像说母语的人一样听懂人讲话。
- ④ 母语说（Text to speech）：构建计算机系统，能够像说母语的人一样讲话。
- ⑤ 真人看（See as well as a person）。构建计算机系统，能够像人一样识别事物行为。
- ⑥ 个人数字资产库系统（Personal Memex）。记录个人一生所读、所看、所听，并能快速检索，但不做任何分析。消费者个人能够负担其购买成本与使用成本。
- ⑦ 全球数字资产库系统（World Memex）。所有文本、声音、图像、视频上网；专家级的分析和摘要能力；快速检索能力。
- ⑧ 远程呈现（TelePresence）。模拟人出现在另一个地方，能与当地环境和其他人交互，好像真实出现一样。
- ⑨ 无故障系统（Trouble-Free Systems）。构建计算机系统，能够供百万人日常使用，但只需一个兼职人员管理维护。
- ⑩ 安全系统（Secure System）。保障上述无故障系统只对授权用户开放、非法用户不能阻碍合法用户使用、信息不可能被窃取。证明这三种安全性。
- ⑪ 系统可用性（AlwaysUp）。保障上述无故障系统的可用性高达**99.9999999%**（**9个9**），即每**100年**才出错**1秒**。证明这种可用性。
- ⑫ 自动程序员（Automatic Programmer）。设计出一种规约语言或用户界面，具备**5个**性质：（**1**）通用，能够表达任意应用的设计；（**2**）高效，表达效率提升一千倍；（**3**）自动，计算机能够自动编译该设计表达；（**4**）易用，系统较为易用；（**5**）智能，可针对应用设计中存在的异常与缺失自动做推理、询问用户。

科学计算系统（超级计算机）已可扩展10万倍

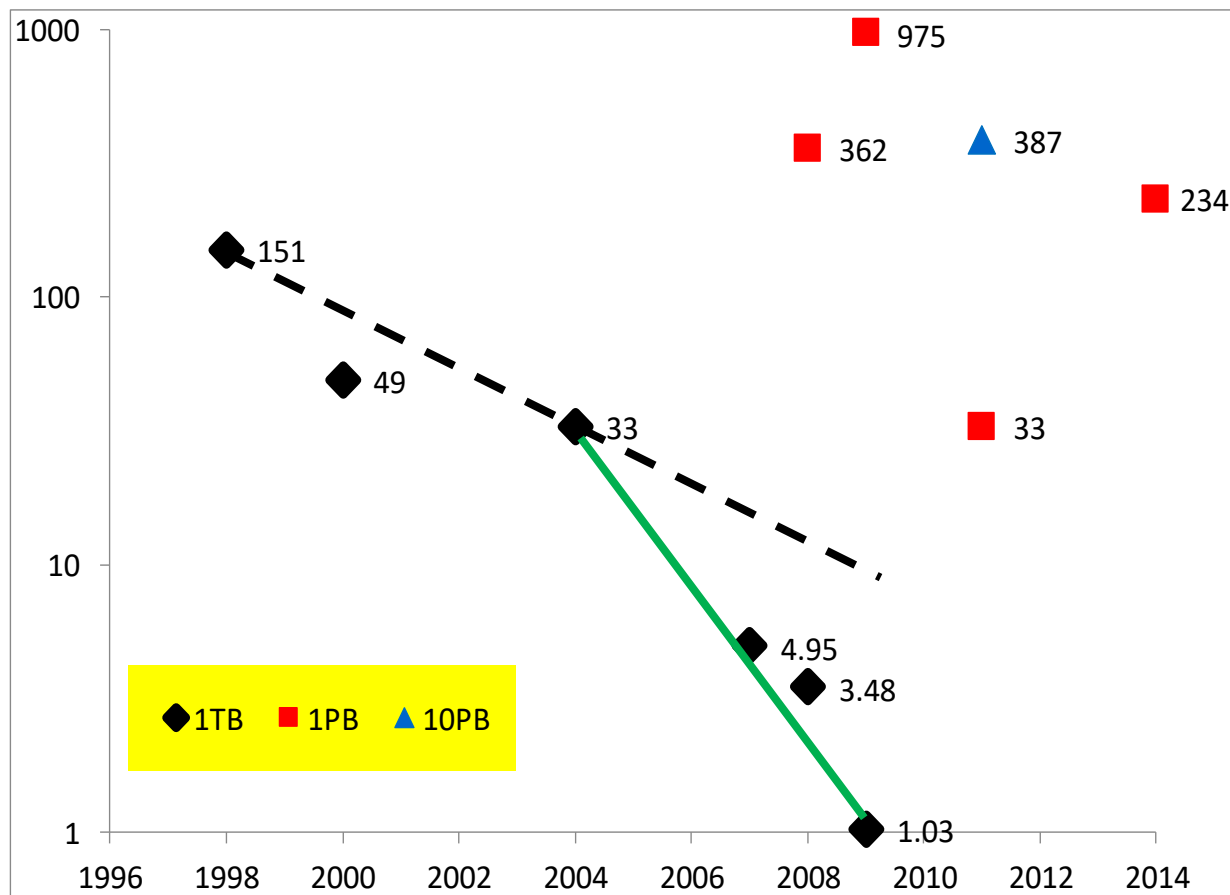
- Linpack基准程序取得的实际计算速度最快，计算速度的单位是每秒执行的64比特浮点运算次数
- Fugaku采用了由多核并行计算节点连接而成的机群体系结构，整机系统结构可扩展，最多可添加到15万个计算节点，将计算速度提升到单节点计算速度的12万倍
 - Fugaku的“多核节点连接而成的机群”体系结构只能扩展12万倍

发布时间	1993年	2020年	1993-2020年增长倍数
冠军名称	Thinking Machine CM-5	Fujitsu Fugaku	N/A
问题规模	N = 52,224	N = 20,459,520	392
计算速度	59.7 GFlop/s	415,530 TFlop/s	6,960,302
主频	32 MHz	2.2 GHz	69
并发度	1,024 cores	7,299,072 cores	7,128
内存容量	32 GB	4,866,048 GB	152,064
功耗	96.5 KW	28,334.5 KW	294
成本	US \$30 million	US \$1 billion	33

大数据计算系统已可扩展上千倍

- 使用TeraSort基准程序，衡量进步的标准是排序1 TB需要多少时间
 - 前期进展很慢，采用“**机群+大数据应用框架**”后进展迅速
 - 数千个节点，整机比单节点快上千倍
 - 腾讯系统保持2016年的TeraSort世界纪录，每分钟排序60.7 TB

基准程序
执行时间
(分钟)



世界纪录
取得的**年度**

格雷12问题：智能

- ① 可扩展系统（Scalability）：设计出算力可扩展1百万倍的系统结构。
- ② 图灵测试（Turing Test）：设计出可通过图灵测试的计算机系统。
- ③ 母语听（Speech to text）：构建计算机系统，能够像说母语的人一样听懂人讲话。
- ④ 母语说（Text to speech）：构建计算机系统，能够像说母语的人一样讲话。
- ⑤ 真人看（See as well as a person）。构建计算机系统，能够像人一样识别事物行为。
- ⑥ 个人数字资产库系统（Personal Memex）。记录个人一生所读、所看、所听，并能快速检索，但不做任何分析。消费者个人能够负担其购买成本与使用成本。
- ⑦ 全球数字资产库系统（World Memex）。所有文本、声音、图像、视频上网；专家级的分析和摘要能力；快速检索能力。
- ⑧ 远程呈现（TelePresence）。模拟人出现在另一个地方，能与当地环境和其他人交互，好像真实出现一样。
- ⑨ 无故障系统（Trouble-Free Systems）。构建计算机系统，能够供百万人日常使用，但只需一个兼职人员管理维护。
- ⑩ 安全系统（Secure System）。保障上述无故障系统只对授权用户开放、非法用户不能阻碍合法用户使用、信息不可能被窃取。证明这三种安全性。
- ⑪ 系统可用性（AlwaysUp）。保障上述无故障系统的可用性高达99.9999999%（9个9），即每100年才出错1秒。证明这种可用性。
- ⑫ 自动程序员（Automatic Programmer）。设计出一种规约语言或用户界面，具备5个性质：（1）通用，能够表达任意应用的设计；（2）高效，表达效率提升一千倍；（3）自动，计算机能够自动编译该设计表达；（4）易用，系统较为易用；（5）智能，可针对应用设计中存在的异常与缺失自动做推理、询问用户。

智能应用已有明显进展

- ChatGPT, GPT-4
- 进步主要得益于深度学习技术的深入发展和应用
 - 算法 + 算力 + 数据
- 但离格雷的目标还差得远



网易拍照翻译（左）、讯飞翻译机（中）与百度智能音箱（右）

格雷12问题：品质

- ① 可扩展系统（Scalability）：设计出算力可扩展1百万倍的系统结构。
- ② 图灵测试（Turing Test）：设计出可通过图灵测试的计算机系统。
- ③ 母语听（Speech to text）：构建计算机系统，能够像说母语的人一样听懂人讲话。
- ④ 母语说（Text to speech）：构建计算机系统，能够像说母语的人一样讲话。
- ⑤ 真人看（See as well as a person）。构建计算机系统，能够像人一样识别事物行为。
- ⑥ 个人数字资产库系统（Personal Memex）。记录个人一生所读、所看、所听，并能快速检索，但不做任何分析。消费者个人能够负担其购买成本与使用成本。
- ⑦ 全球数字资产库系统（World Memex）。所有文本、声音、图像、视频上网；专家级的分析和摘要能力；快速检索能力。
- ⑧ 远程呈现（TelePresence）。模拟人出现在另一个地方，能与当地环境和其他人交互，好像真实出现一样。
- ⑨ 无故障系统（Trouble-Free Systems）。构建计算机系统，能够供百万人日常使用，但只需一个兼职人员管理维护。
- ⑩ 安全系统（Secure System）。保障上述无故障系统只对授权用户开放、非法用户不能阻碍合法用户使用、信息不可能被窃取。证明这三种安全性。
- ⑪ 系统可用性（AlwaysUp）。保障上述无故障系统的可用性高达99.9999999%（9个9），即每100年才出错1秒。证明这种可用性。
- ⑫ 自动程序员（Automatic Programmer）。设计出一种规约语言或用户界面，具备5个性质：（1）通用，能够表达任意应用的设计；（2）高效，表达效率提升一千倍；（3）自动，计算机能够自动编译该设计表达；（4）易用，系统较为易用；（5）智能，可针对应用设计中存在的异常与缺失自动做推理、询问用户。

云账户互联网服务系统

- 为6600万自由职业者提供秒批办照、收入结算、税款代缴、保险保障服务

- 2016年成立，2021年实现收入500多亿元、纳税30多亿元
- 云账户董事长获“全国脱贫攻坚先进个人”荣誉
- 仅有200多名技术人员

- 得益于分布式系统进展，使用了30余种开源应用框架



- 格雷问题进步

- 第9问题：无故障系统，供百万人日常使用，只需一个兼职人员管理维护
 - 云账户系统支持6600多万活跃用户，需要三十余名全职工程师做运行维护

云账户互联网服务系统

- 为6600万自由职业者提供秒批办照、收入结算、税款代缴、保险保障服务

- 2016年成立，2021年实现收入500多亿元、纳税30多亿元；
- 云账户董事长获“全国脱贫攻坚先进个人”荣誉
- 仅有200多名技术人员

- 得益于分布式系统进展，使用了30余种开源应用框架



- 格雷问题进步

- 第10问题：安全系统。保障上述无故障系统只对授权用户开放、非法用户不能阻碍合法用户使用、信息不可能被窃取。证明这三种安全性。
 - 云账户系统实际运营体现了上述三种安全性，但不能数学严密地证明

云账户互联网服务系统

- 为6600万自由职业者提供秒批办照、收入结算、税款代缴、保险保障服务

- 2016年成立，2021年实现收入500多亿元、纳税30多亿元；
- 云账户董事长获“全国脱贫攻坚先进个人”荣誉
- 仅有200多名技术人员

- 得益于分布式系统进展，使用了30余种开源应用框架



- 格雷问题进步

- 第11问题：可用性高达99.9999999%（9个9），每100年才出错1秒。证明这种可用性。
 - 云账户系统已经在五年多时间内无中断地为用户服务
 - 可用性还不能证明

格雷12问题：易用

- ① 可扩展系统（Scalability）：设计出算力可扩展1百万倍的系统结构。
- ② 图灵测试（Turing Test）：设计出可通过图灵测试的计算机系统。
- ③ 母语听（Speech to text）：构建计算机系统，能够像说母语的人一样听懂人讲话。
- ④ 母语说（Text to speech）：构建计算机系统，能够像说母语的人一样讲话。
- ⑤ 真人看（See as well as a person）。构建计算机系统，能够像人一样识别事物行为。
- ⑥ 个人数字资产库系统（Personal Memex）。记录个人一生所读、所看、所听，并能快速检索，但不做任何分析。消费者个人能够负担其购买成本与使用成本。
- ⑦ 全球数字资产库系统（World Memex）。所有文本、声音、图像、视频上网；专家级的分析和摘要能力；快速检索能力。
- ⑧ 远程呈现（TelePresence）。模拟人出现在另一个地方，能与当地环境和其他人交互，好像真实出现一样。
- ⑨ 无故障系统（Trouble-Free Systems）。构建计算机系统，能够供百万人日常使用，但只需一个兼职人员管理维护。
- ⑩ 安全系统（Secure System）。保障上述无故障系统只对授权用户开放、非法用户不能阻碍合法用户使用、信息不可能被窃取。证明这三种安全性。
- ⑪ 系统可用性（AlwaysUp）。保障上述无故障系统的可用性高达99.9999999%（9个9），即每100年才出错1秒。证明这种可用性。
- ⑫ 自动程序员（Automatic Programmer）。设计出一种规约语言或用户界面，具备5个性质：（1）通用，能够表达任意应用的设计；（2）高效，表达效率提升一千倍；（3）自动，计算机能够自动编译该设计表达；（4）易用，系统较为易用；（5）智能，可针对应用设计中存在的异常与缺失自动做推理、询问用户。

云账户互联网服务系统

- 为6600万自由职业者提供秒批办照、收入结算、税款代缴、保险保障服务

- 2016年成立，2021年实现收入500多亿元、纳税30多亿元；
- 云账户董事长获“全国脱贫攻坚先进个人”荣誉
- 仅有200多名技术人员

- 得益于分布式系统进展，
使用了30余种开源应用框架



- 格雷问题进步

- 第12问题：自动程序员。设计出一种规约语言或用户界面，使得表达效率提升一千倍。
 - 云账户仍然主要依赖于产品经理和工程师的合作来设计、开发和测试系统
 - 虽然利用基础组件库、微服务、开发工具链、敏捷开发流程等技术可减少重复工作，但云账户团队主要基于高级编程语言开发分布式应用系统
 - 只在几个限定的应用场景中，才支持最终用户在可视化环境中拖拽完成系统设计
 - 尚缺乏可自动执行的系统抽象

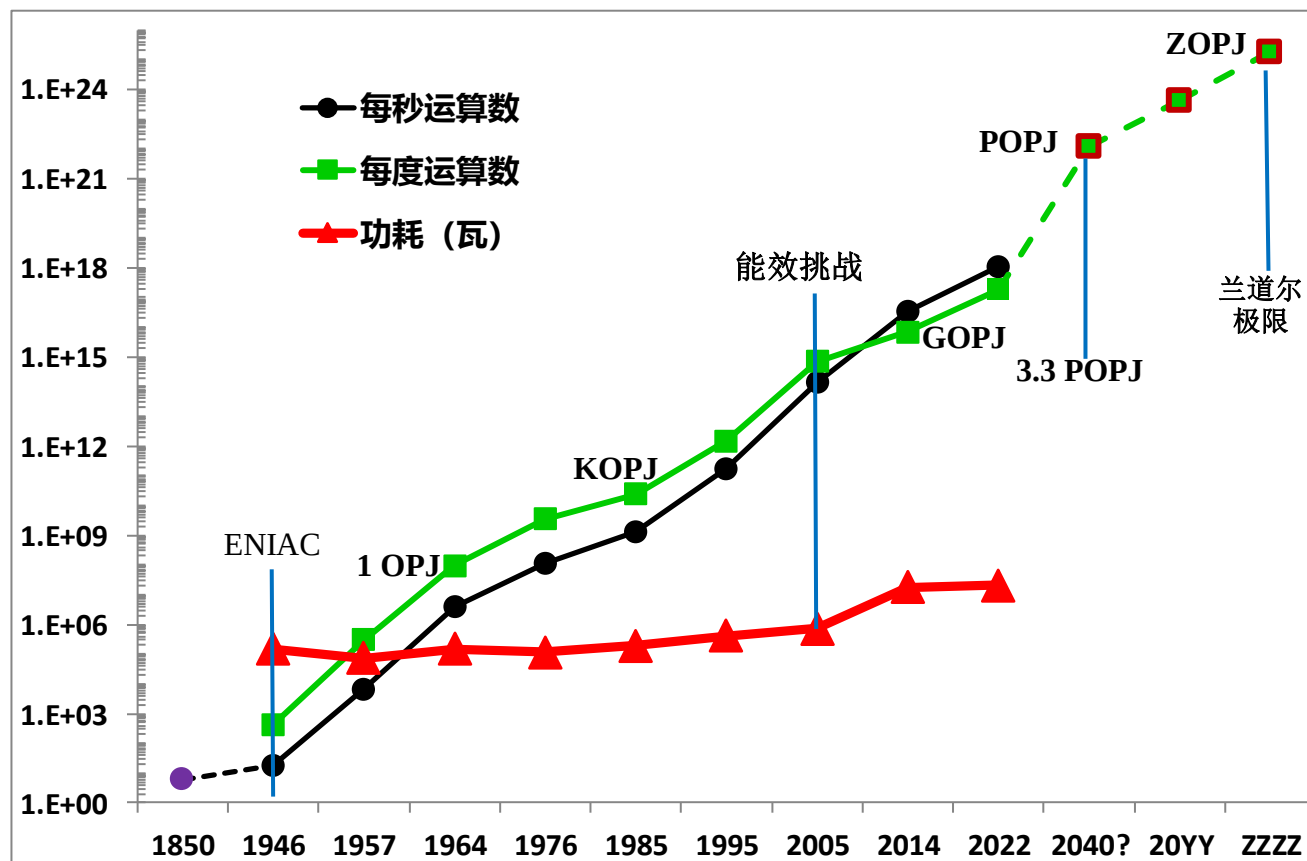
解决格雷问题受制于能效挑战

- OPJ: 每焦耳运算数
- 物理学的兰道尔极限前的三个里程碑
 - ENIAC: 终结了95年缓慢进展, 开启了60年的指数改善
 - 2005年开始, 能效挑战变得严峻
 - DARPA提出3.3 POPJ的目标, 试图大幅提升能效

Z: Zetta, 泽, 10^{21}
P: Peta, 拍, 10^{15}
G: Giga, 吉, 10^9
K: Kilo, 千, 10^3

单套计算机系统的
速度、能效、功耗

发展历史与展望



课堂小测验

姓名：

学号：

针对格雷第1问题，即“设计出算力可扩展1百万倍的系统结构”，下述哪个断言是正确的？请4选1。（）

- A. 科学计算系统已经解决了格雷第1问题，因为2020年的Top1超级计算机已经比1993年的Top1超级计算机速度快了数百万倍
- B. 科学计算系统尚未解决格雷第1问题，因为2020年的Fugaku计算机的算力只能扩展12万倍
- C. 大数据计算系统已经解决了格雷第1问题，因为大数据计算机已经能够排序10PB，是格雷的TeraSort基准程序的1TB要求的1万倍
- D. 大数据计算系统尚未解决格雷第1问题，因为保持TeraSort世界纪录的腾讯系统每分钟排序60.7 TB，其计算速度仅为2009年冠军系统的60.7倍，远未达到1百万倍



计算思维展望

- 数字符号
 - 串行计算过程
 - 图灵可计算性
- ➔ 数字模拟系统（digital and analog system）
 - ➔ 并行分布式系统、全球算力网（信息高铁）
 - ➔ 非图灵可计算性



3. 科学研究的第五范式：AI4R (AI for Research)

- AI4R的别称

- AI4S(AI for Science), AI4SS(AI for Social Science), AI4E(AI for Education)
 - AI for Math, AI for Physics, AI for Chemistry, AI for Biology, ...

- 科学研究的范式演变（paradigm shift）

- 理论，实验，科学计算（第三范式，第四范式，第五范式）

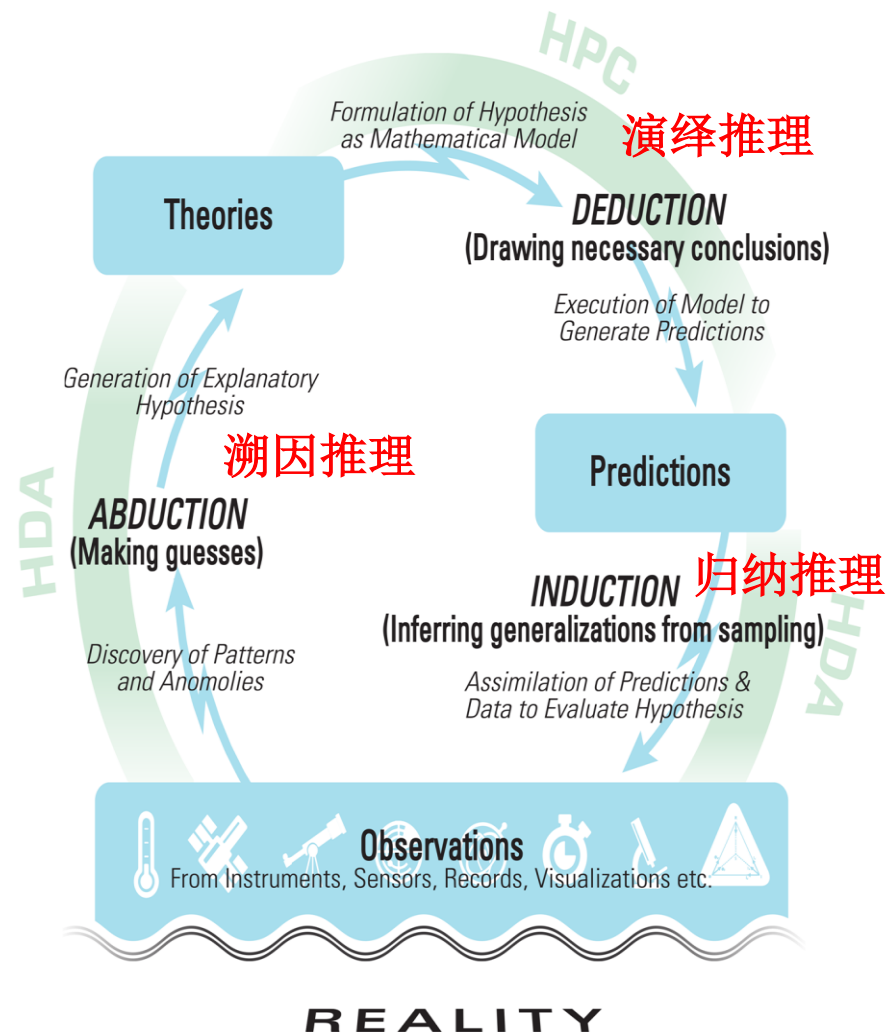
- 第五范式工作

- 观点与综述
- 科学智算实例（中科院计算所贾伟乐团队）

科学研究的范式演变（paradigm shift）

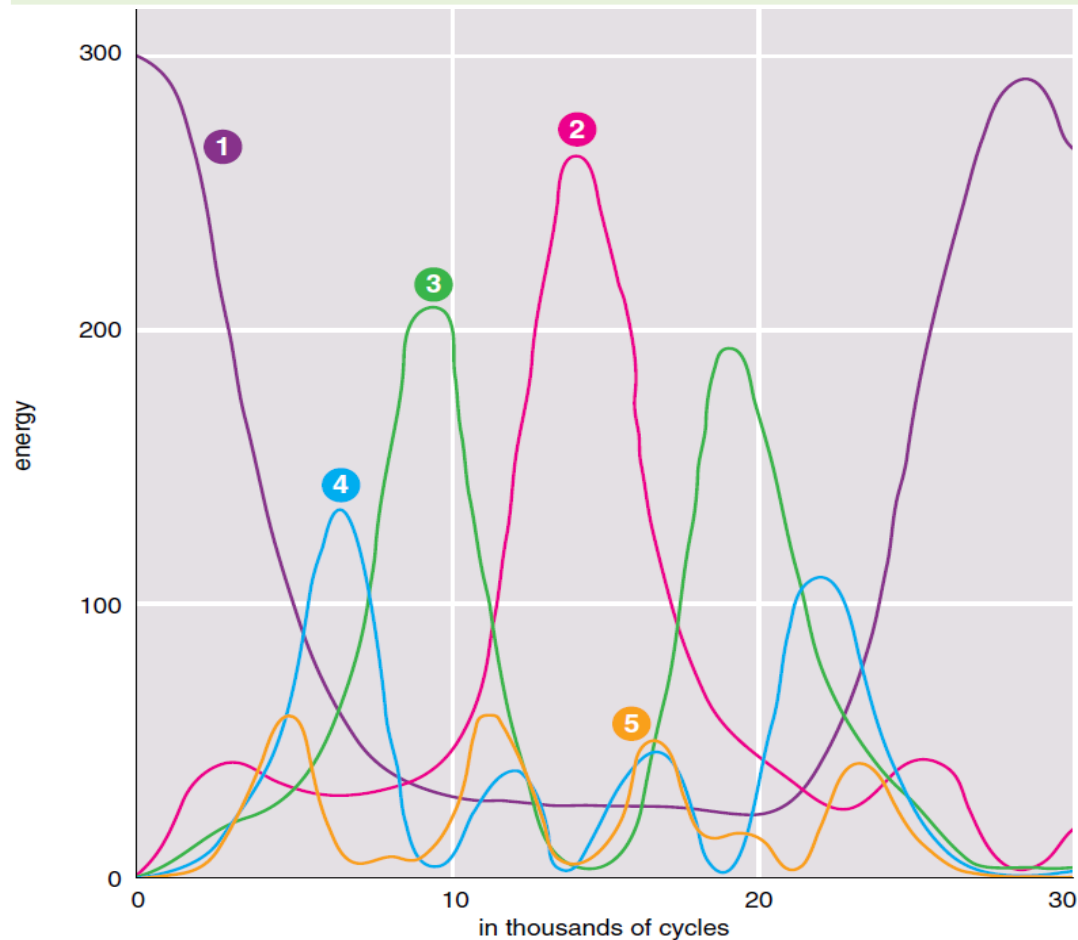
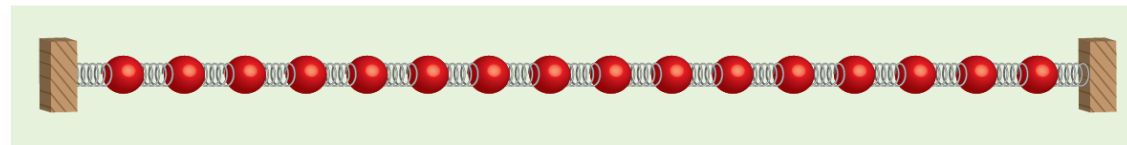
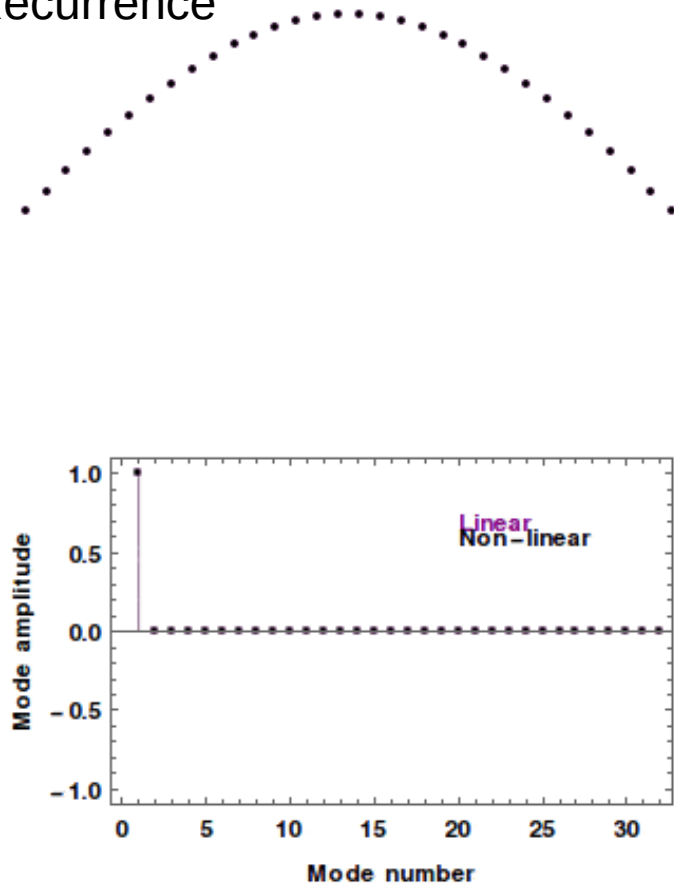
- 第一范式：理论研究
 - 牛顿、爱因斯坦
- 第二范式：实验观察
 - 伽利略
- 第三范式：计算机模拟
- 第四范式：数据处理
- 第五范式：AI辅助科学
 - AI4Science
 - 第一性原理模拟+深度学习+大数据
 - 智能超算：演绎推理+归纳推理+溯因推理

科学计算



3.1 费米等在1953年发明计算机模拟（第三范式）

- 也称为“数值实验”
 - 在Maniac计算机上求解
- Fermi–Pasta–Ulam悖论
 - 发现了FPU Recurrence



Porter, Mason A., et al. "Fermi, Pasta, Ulam and the Birth of Experimental Mathematics: A numerical experiment that Enrico Fermi, John Pasta, and Stanislaw Ulam reported 54 years ago continues to inspire discovery." *American Scientist* 97.3 (2009): 214-221.

3.2 第四范式：数据处理

- 黑洞照片实例

- 天文学依赖望远镜

- 亚米-米级

- LAMOST: <8米

- 郭守敬望远镜

- FAST: 500米

- 中国天眼

- 黑洞拍照需要
直径1万千米的望远镜！



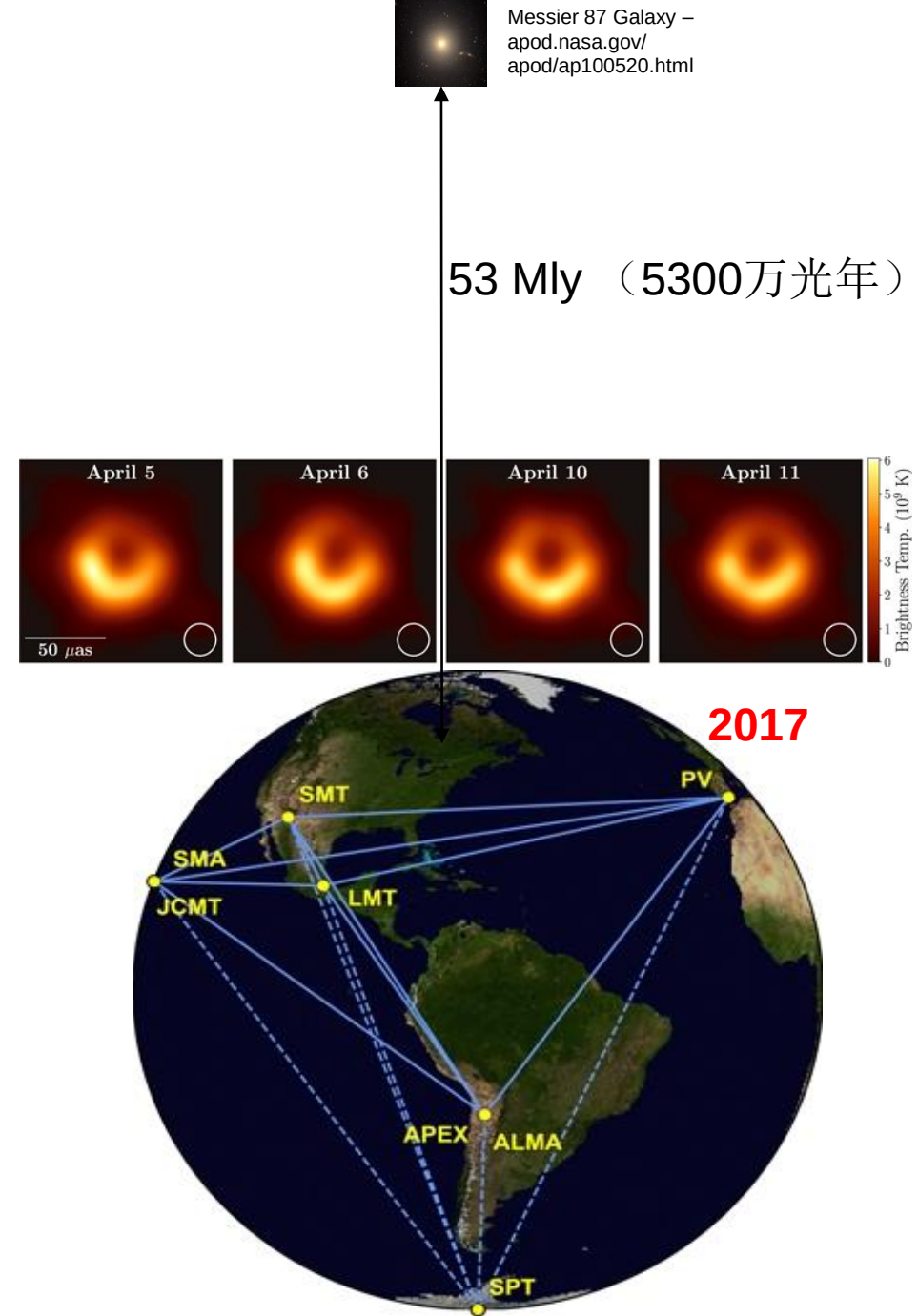
郭守敬望远镜



中国天眼

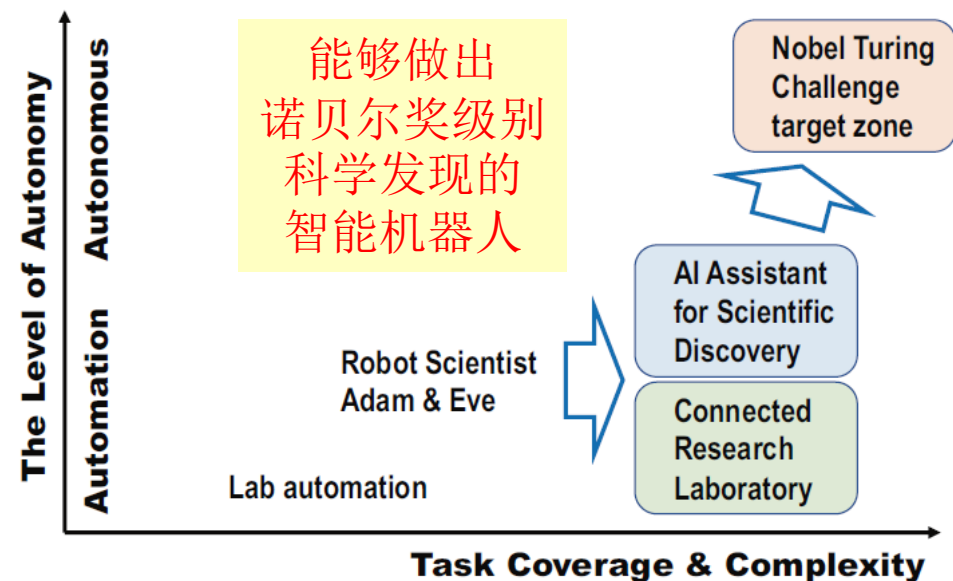
地球规模望远镜

- Event Horizon Telescope
 - 拍照（从曝光到生成照片）
 - 2017年4月，多次曝光
 - 采用原子钟同步
 - 将多点数据硬盘运到超级计算中心
 - 大约2年数据处理，生成黑洞照片
 - 2019年4月，发表论文
 - 验证了爱因斯坦广义相对论



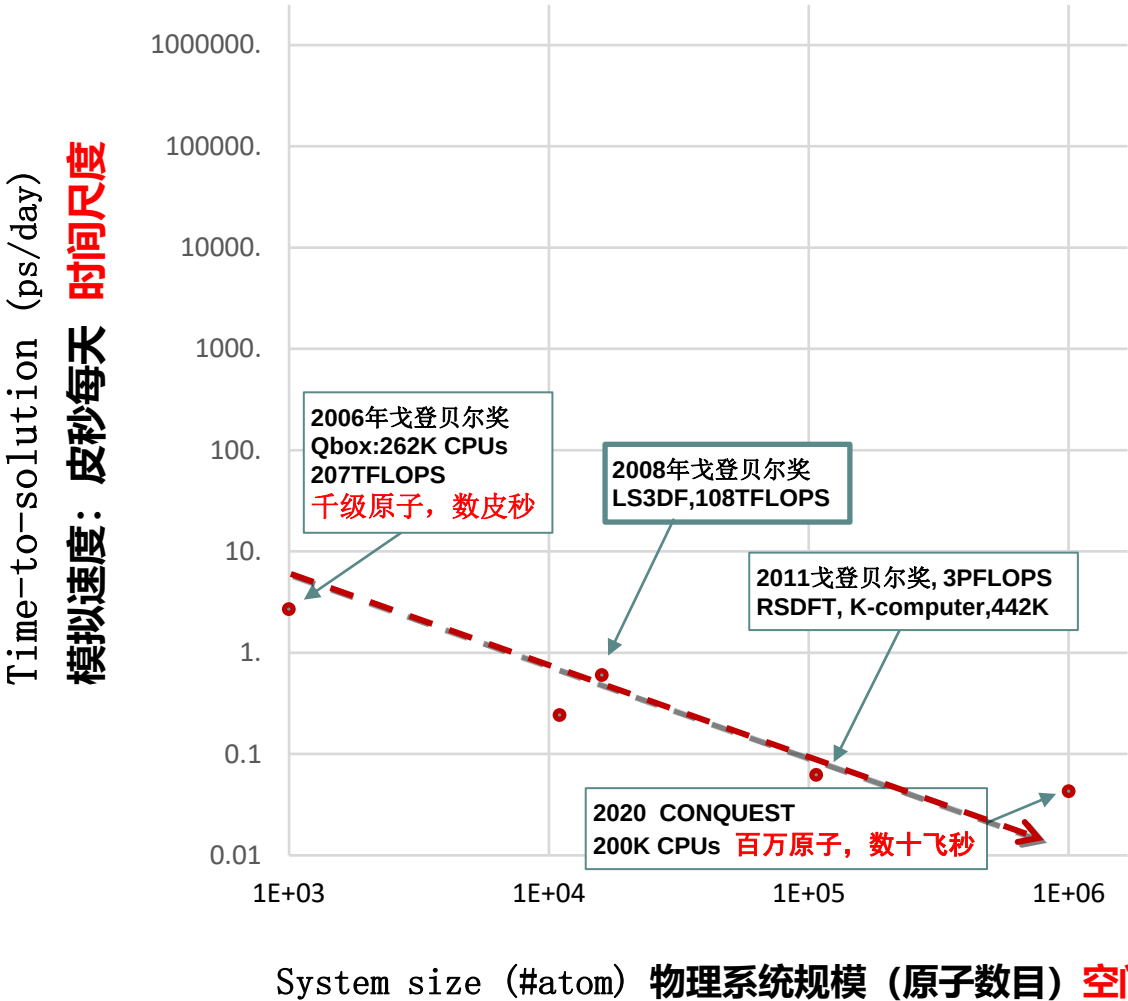
3.3 第五范式：AI辅助科学（AI4Science）

- 非常多的AI4S实际工作，汇聚成了几个主要学术观点
- AI4S带来多种好处
 - 提升计算速度，提高科研效率（人员时间与物质成本），提高科研质量，新的科教渠道
- AI4S正在改变机器的作用：从工具变为伙伴
 - **Nature Reviews Physics社评** 2022: Machines are no longer mere tools, but equal partners in scientific exploration, exchanging ideas, intuition and understanding with the human peers
- AI4S为科学带来新挑战
 - 不懂装懂，威胁多样性
- 五个范式常常混合使用
 - 计算：第一性原理模拟+数据处理+深度学习
- AI4S新倡议：诺贝尔图灵挑战



传统第一性原理分子动力学计算：十四年进展

空间尺度提高3个数量级，但时间尺度降低3个数量级

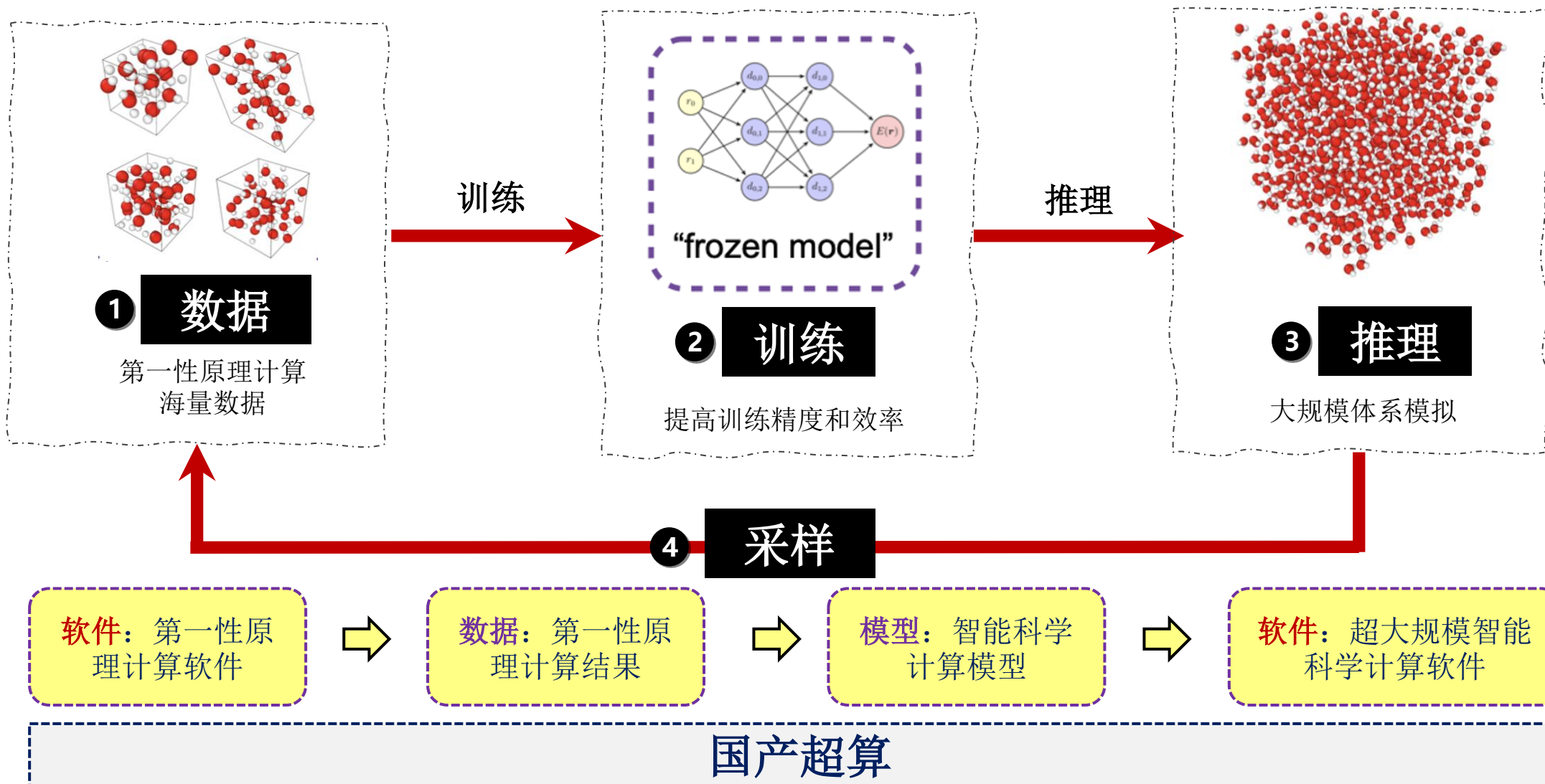


根本原因：计算量= $O(N^3)$
N是原子数目

系统规模N提升1000倍，计算量提升
 $1000*1000*1000 = 10$ 亿倍

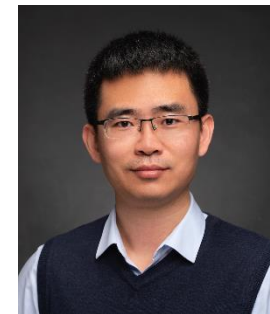
贾伟乐:

HPC+AI驱动的第一性原理分子动力学的四个步骤

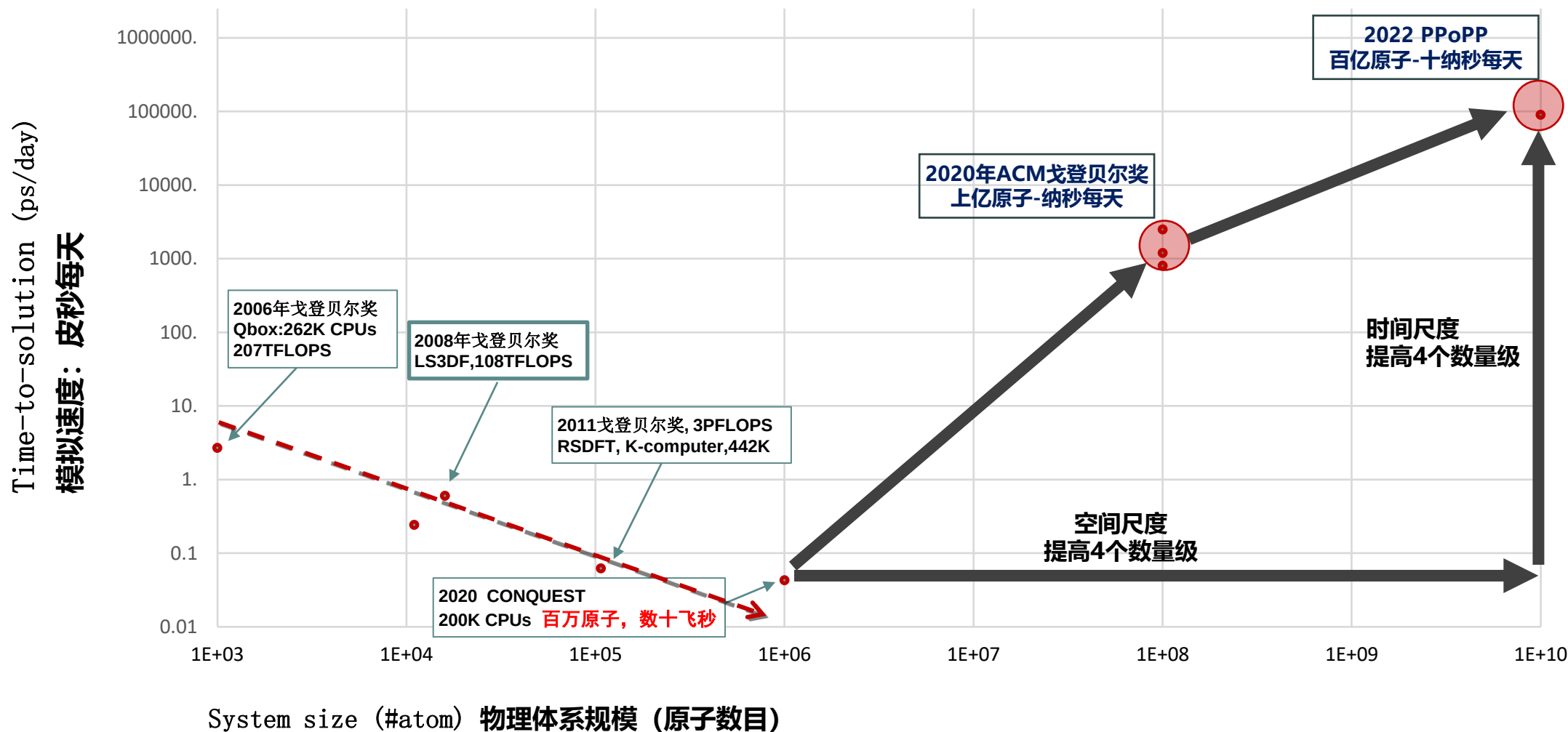


HPC+AI驱动的第一性原理分子动力学

效果：时空尺度提高4个数量级



贾伟乐研究员
中国科学院计算技术研究所
jiaweile@ict.ac.cn
<https://people.ucas.ac.cn/~jiaweile>



参考文献

● 综述与观点

- Hey, T., Tansley, S. & Tolle, K. The Fourth Paradigm: Data-Intensive Scientific Discovery (Microsoft Research, 2009).
- Editorial, Pervasive machine learning in physics, Nature Reviews Physics volume 4, p. 353 (2022).
- Wang, H. et al. Scientific discovery in the age of artificial intelligence, Nature, vol. 620, pp. 47–60 (2023).
- Douglas, MR. How will we do mathematics in 2030? CERN Colloquium, Sep 30, 2021 <https://indico.cern.ch/event/1077061/attachments/2320414/3951088/decade19.pdf>
- 贾伟乐. HPC+AI在第一性原理分子动力学中的应用, 科学计算技术与应用会议, 2023, 中国-三亚
- 诺贝尔图灵挑战倡议 <https://www.nobelturingchallenge.org/>

● AI4Science的新挑战与社会进步

- Editorial, Defining physicists' relationship with AI, Nature Reviews Physics volume 4, page 735 (2022)
- Lisa Messeri & M. J. Crockett, Artificial intelligence and illusions of understanding in scientific research, Nature, volume 627, pages 49–58 (2024)

4. 21世纪文明需要创新型人才

应对“百年未有之大变局”，需要世纪眼光

- 21世纪文明是“发展文明”
- 21世纪文明是“信息文明”

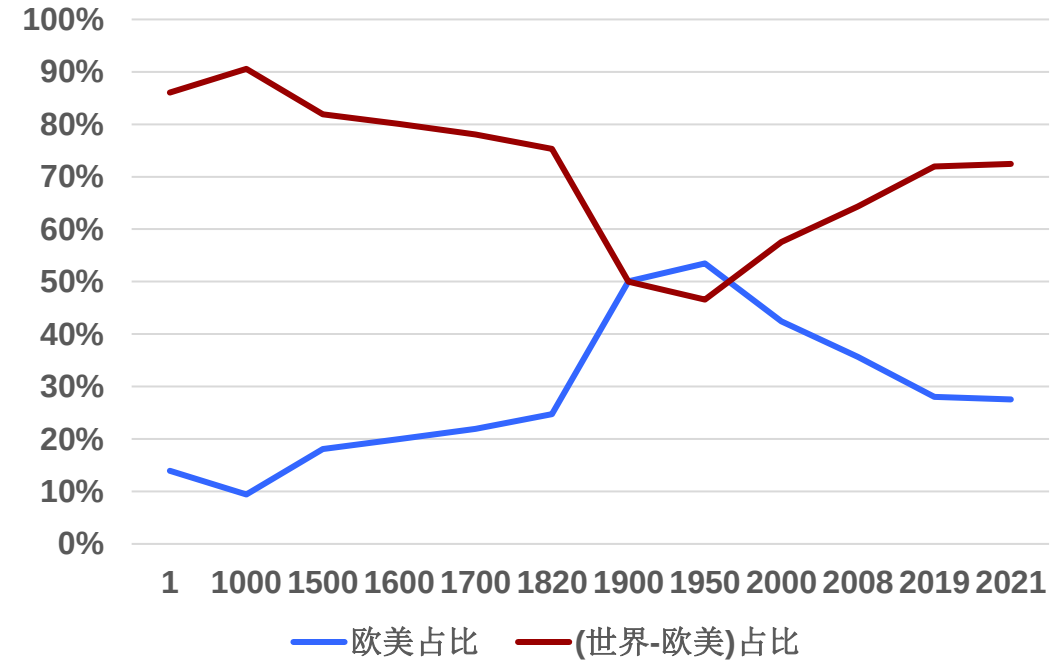
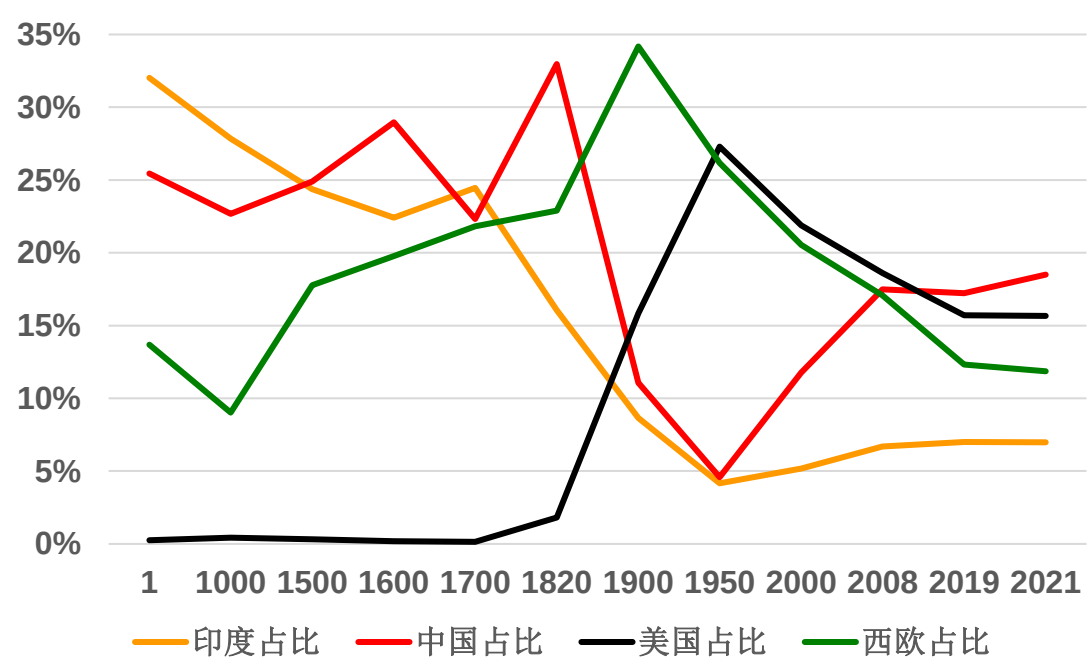
放眼世界，我们面对的是百年未有之大变局。新世纪以来一大批新兴市场国家和发展中国家快速发展，世界多极化加速发展，国际格局日趋均衡，国际潮流大势不可逆转。

习近平，2017

21世纪文明是“国家发展文明”

- 放眼世界， 20世纪文明的最大特征是什么？
- 20世纪文明的最大特征是殖民主义的消亡
 - 是“民族解放”、“民族独立”、“主权自由”

彼得卡赞斯坦的观点



数据来源：世界银行 <https://datatopics.worldbank.org/world-development-indicators/>; Angus Maddison, The World Economy, OECD 2006; <https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2023>

如果认同人类文明发展的长期趋势是“人类不断解放”

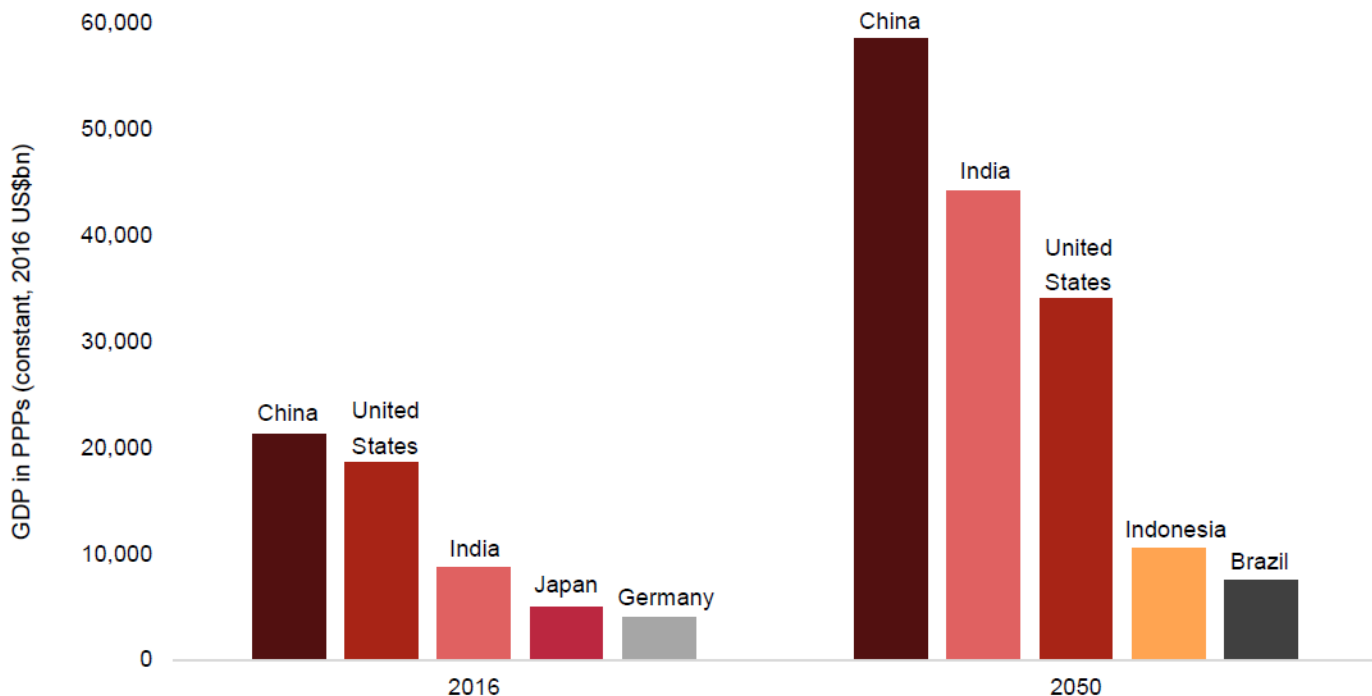
- 自由解放的表征：摆脱 ①野蛮；②神权、皇权、列强、奴役制度；③垄断控制
- 20世纪文明的最大特征是殖民主义的消亡
 - 发展中国家享有了“民族独立”； 20世纪文明是追求“主权自由”的文明
- 21世纪文明的最大特征是所有国家发展，尤其是发展中国家科技兴起
 - “主权自由” → “国家发展的自由”

年度	GDP（万亿国际元）			人均GDP（国际元）			人口平均预期寿命（岁）		
	世界	美国	中国	世界	美国	中国	世界	美国	中国
1820	0.69	0.01	0.23	666	1,257	600	26	39	24
1900	1.97	0.31	0.22	1,261	4,091	545	31	47	24
1950	5.34	1.46	0.24	2,111	9,561	448	49	68	41
2000	36.69	8.03	4.32	6,038	28,467	3,421	68	77	72
2008	50.97	9.49	8.91	7,614	31,178	6,725	70	78	75
2019	130.66	20.51	22.49	16,877	62,471	15,978	73	79	78
2021	134.82	21.13	24.94	17,091	63,636	17,658	71	76	78

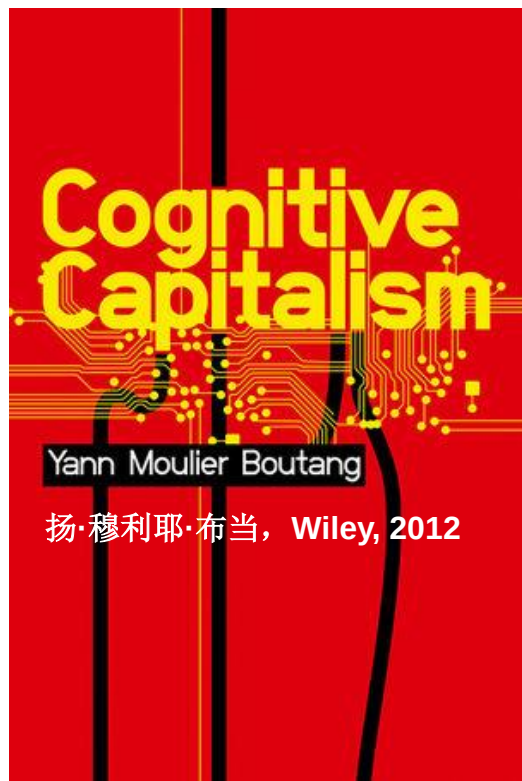
数据来源：世界银行 <https://datatopics.worldbank.org/world-development-indicators/>；Angus Maddison, The World Economy, OECD 2006；
<https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2023>

如果认同人类文明发展的长期趋势是“人类解放事业”

- 21世纪文明的最大特征是所有国家发展，尤其是发展中国家兴起
 - “主权自由” → “发展的自由”；必然要追求“科技发展的自由”
 - 市场的四个阶段：贸易资本主义 → 工业资本主义 → 金融资本主义 → 认知资本主义



数据来源：普华永道，The World in 2050报告，2017



中国进入新发展阶段

尤其需要懂计算思维的创新人才

“0到1创新”是新质生产力的核心

什么是新发展阶段？

跟踪-追赶-贡献的贡献阶段

跟踪阶段：毕业生的主流出口是出国、外企

Follow

恢复高考



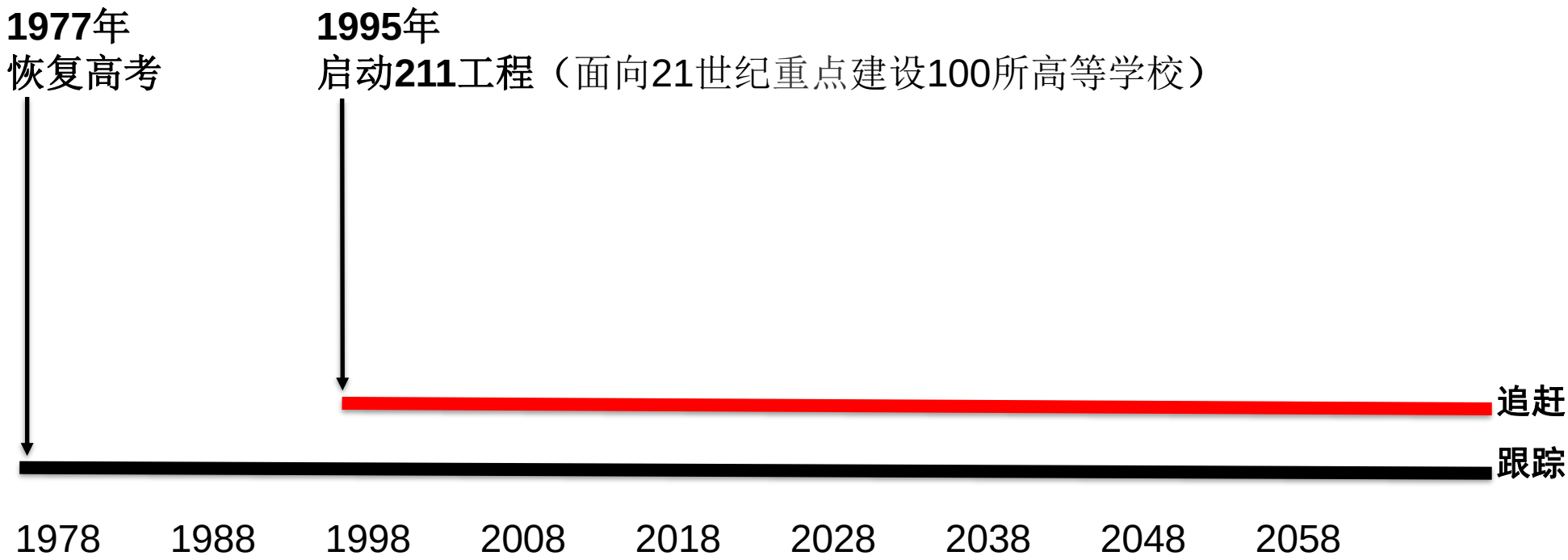
1978 1988 1998 2008 2018 2028 2038 2048 2058

跟踪

什么是新发展阶段？

跟踪阶段：毕业生出口——追求是出国、外企
追赶阶段：

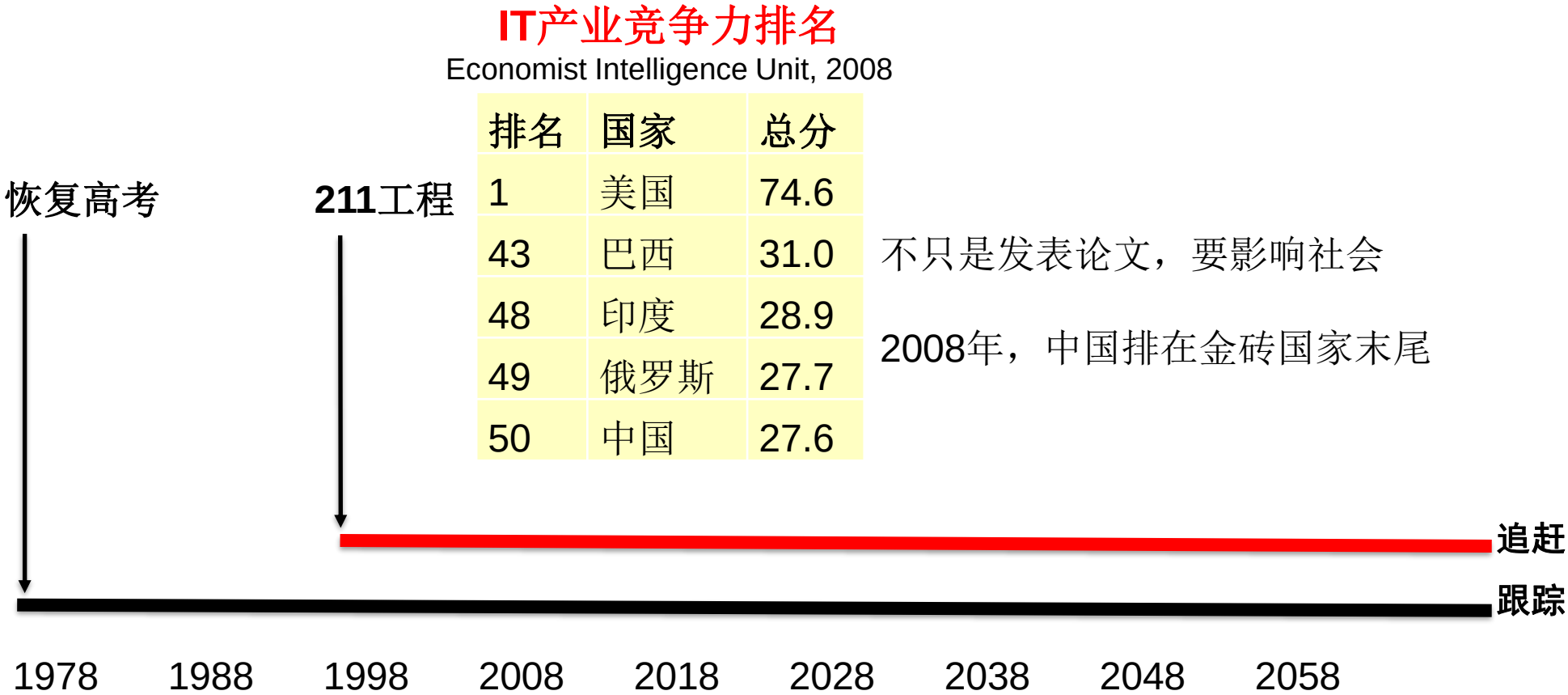
Follow
Catch up



什么是新发展阶段？

跟踪阶段：毕业生出口——追求出国、外企
追赶阶段：毕业生出口——逐渐过渡到大厂、研究所名校

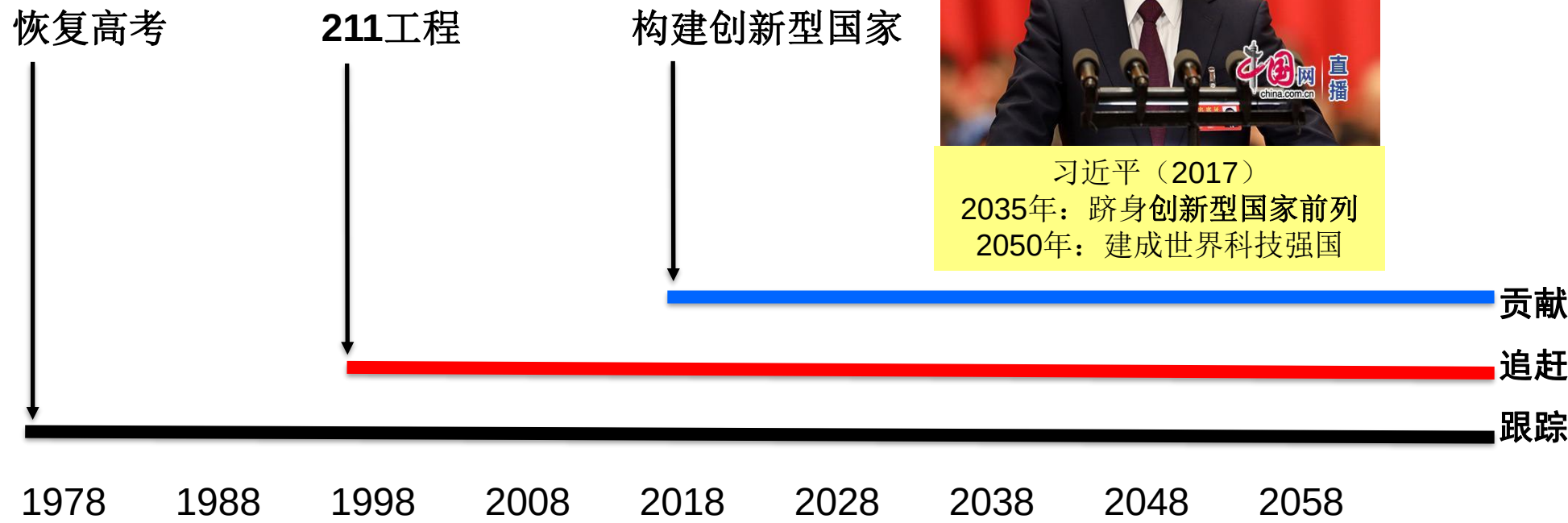
Follow
Catch up



新发展阶段是为全球做贡献的阶段

- 跟踪阶段: 毕业生的主流出口是出国和外企 Follow
- 追赶阶段: 毕业生的主流出口是大厂、名校 Catch up
- 贡献阶段: 毕业生的出口是创新型多样化 Contribute, Lead

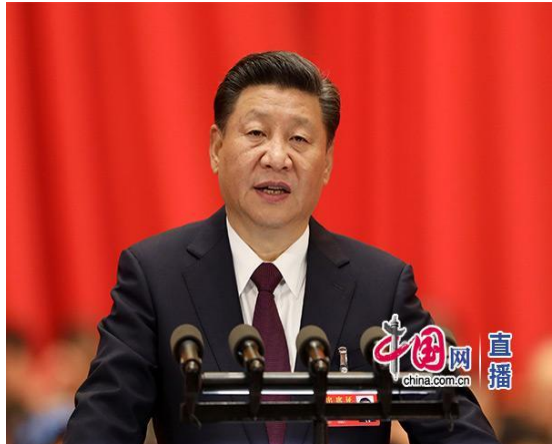
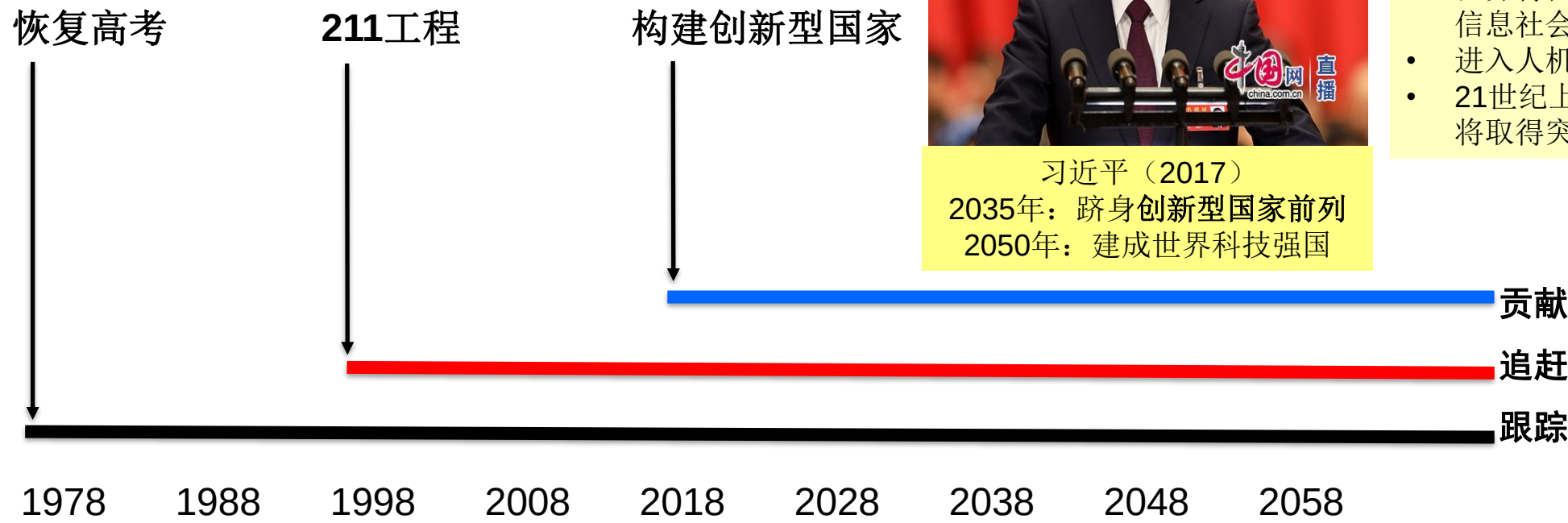
贡献阶段的大学教育，与跟踪追赶的大学教育，是不一样的



新发展阶段是为全球做贡献的阶段

- 跟踪阶段：毕业生的主流出口是出国和外企
- 追赶阶段：毕业生的主流出口是大厂、名校
- 贡献阶段：毕业生的出口是创新型多样化
- Follow
- Catch up
- Contribute, Lead

贡献阶段的大学教育，与跟踪追赶的大学教育，是不一样的
信息社会的大学教育，与工业社会的大学教育，是不一样的



习近平（2017）
2035年：跻身创新型国家前列
2050年：建成世界科技强国



- 世界将在2050年进入信息社会中高级阶段
- 进入人机物三元世界
- 21世纪上半叶信息科学将取得突破性进展