



南京大學
人工智能學院

SCHOOL OF ARTIFICIAL INTELLIGENCE, NANJING UNIVERSITY

LAMDA
Learning And Mining from Data
<http://www.lamda.nju.edu.cn>



人工智能跨学科交叉研究初探

钱超

南京大學
人工智能學院



什么是人工智能



人工智能会像人一样思考吗？

人工智能会有意识吗？

人工智能会威胁人类吗？

两种不同的人工智能

□ 强人工智能（“科幻人工智能”）

研制出和人一样聪明，甚至比人更聪明的机器



重要特征：

- 具有自主意识
- 全面达到，甚至超过人类智能水平
- ...

两种不同的人工智能

□ 弱人工智能（“科学人工智能”）

让机器做事时聪明一点

“人工智能就是让机器来完成那些如果由人来做则需要智能的事情的科学”

解读：

- 如果某件事情需要智能，通过机器来做，就是人工智能
- 不要求“全面”达到人类智能水平
- “做事”就行，不必具备“自主意识”“情感”……

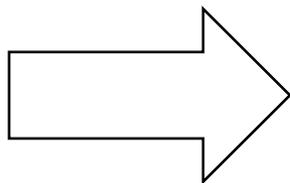


马文闵斯基
(1927-2016)
人工智能奠基者之一
1969年图灵奖

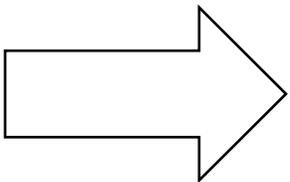
一个类比



人的智能行为



人工智能



人工智能重要，是因为能造出“智能工具”（类比：飞机）

- 造飞机的人不会关心飞机有没有“意识”、会不会“疼”
- 更不会关心飞机是否“全面达到”鸟的能力（例如：下蛋）

Computing Machinery and Intelligence

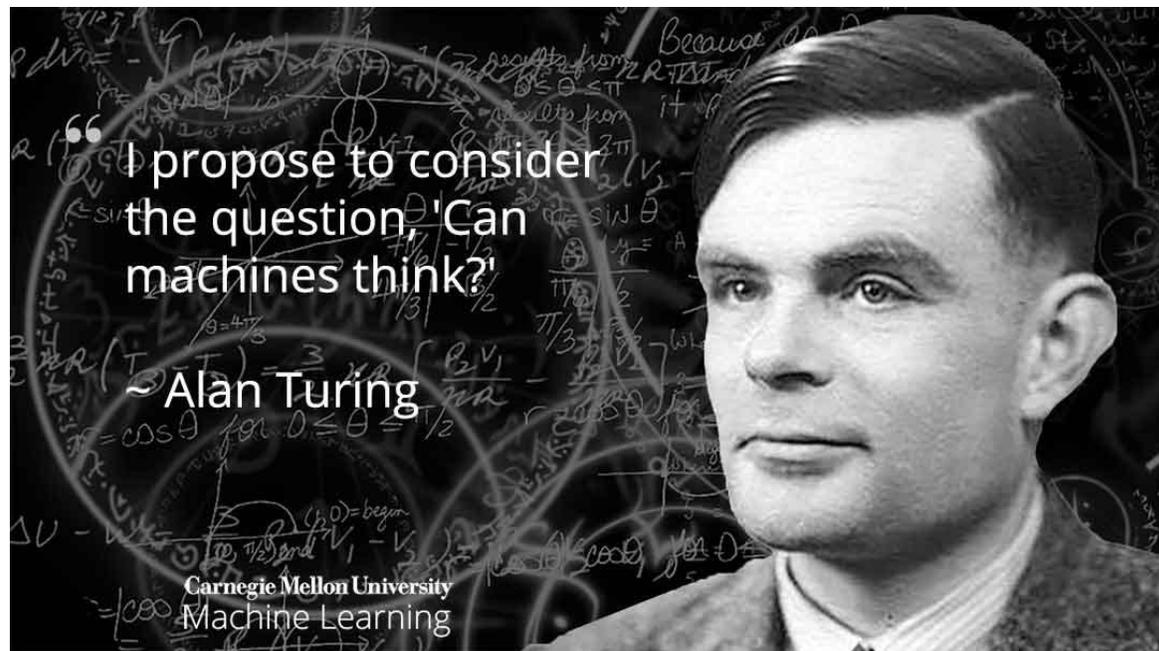
计算机与智能

1950年

艾伦·图灵

“Can machine think?”

机器能思考吗？



艾伦·图灵

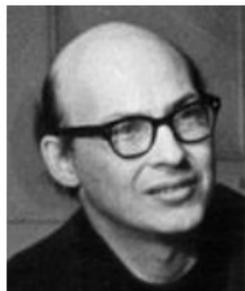
1912-1954

人工智能的起源

1956年的达特茅斯会议标志人工智能这一学科的诞生



J. McCarthy
人工智能之父
图灵奖 (1971)



M. Minsky
图灵奖 (1969)



C. Shannon
信息论之父



H. A. Simon
图灵奖(1975)



A. Newell
图灵奖(1975)

.....

.....



1956年夏 美国达特茅斯学院

第一阶段：推理期

1956-1960s: Logic Reasoning

- ◆ 出发点：“数学家真聪明！”
- ◆ 把人的**思考逻辑**放入电脑
- ◆ 主要成就：自动定理证明系统（例如，西蒙与纽厄尔的 Logic Theorist 系统）

“能够证明数学原理第2章中的前52个定理中的38个，其中定理2.85的证明实际上比罗素和怀特海手抄的证明更优雅”



赫伯特·西蒙
(1916-2001)
1975年图灵奖



阿伦·纽厄尔
(1927-1992)
1975年图灵奖

第二阶段：知识期

1970s-1980s: Knowledge Engineering

- ◆ 出发点：“知识就是力量！”
- ◆ 把人的**全部知识**放入电脑
- ◆ 主要成就：专家系统（例如，费根鲍姆等人的“DENDRAL”系统）



爱德华·费根鲍姆
(1936-)
1994年图灵奖

形成“知识工程”的方法论，包含了对专家知识从获取、分析到用规则表达等一系列技术

第三阶段：学习期

1990s-Now: Machine Learning

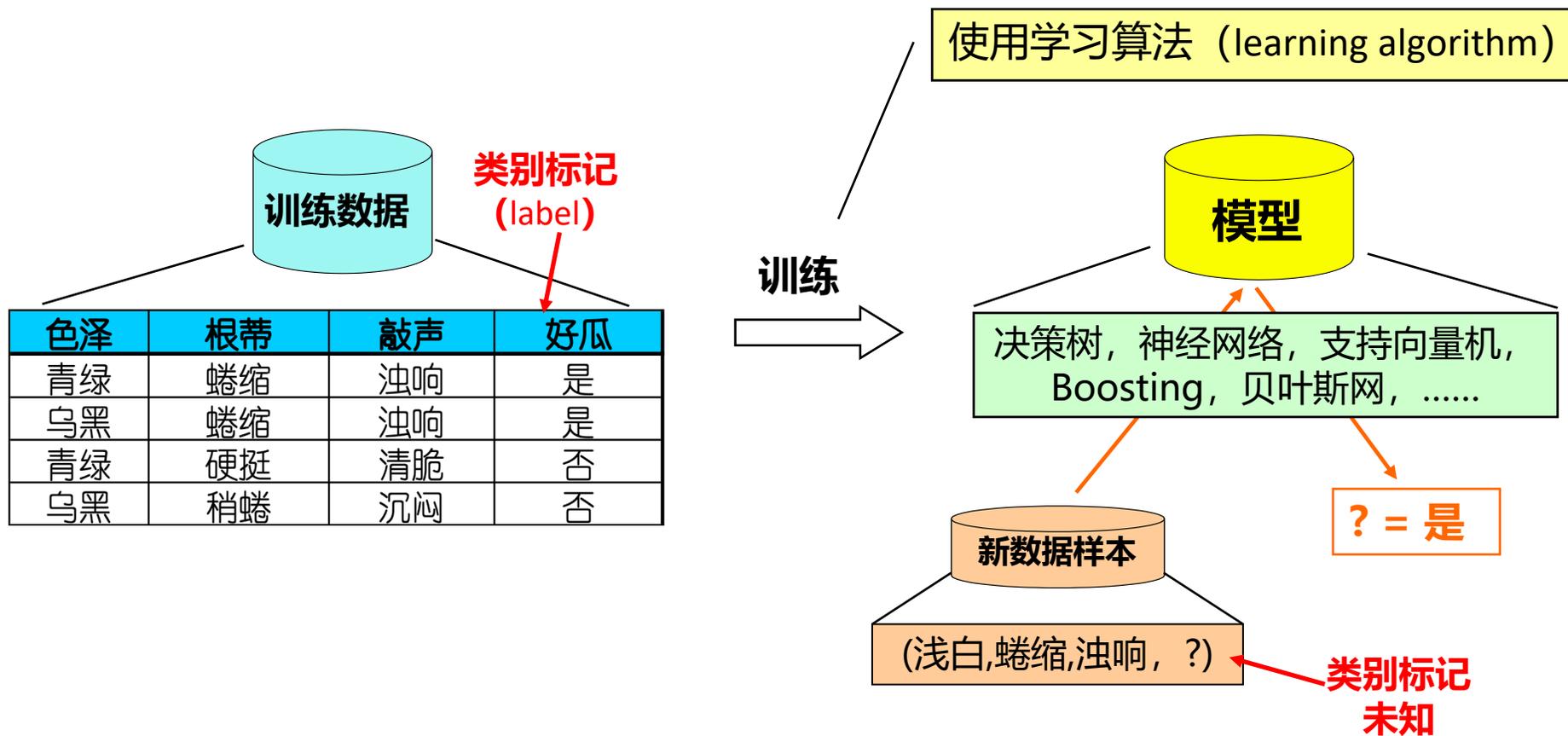
- ◆ 出发点：“让系统自己学！”
- ◆ 把人的**所有看见**放入电脑
- ◆ 主要成就： ...

机器学习是作为“突破知识工程瓶颈”
之利器而出现的



恰好在20世纪90年代中后期，人类发现自己淹没在数据的汪洋中，
对自动数据分析技术 – 机器学习的需求日益迫切

典型的机器学习过程

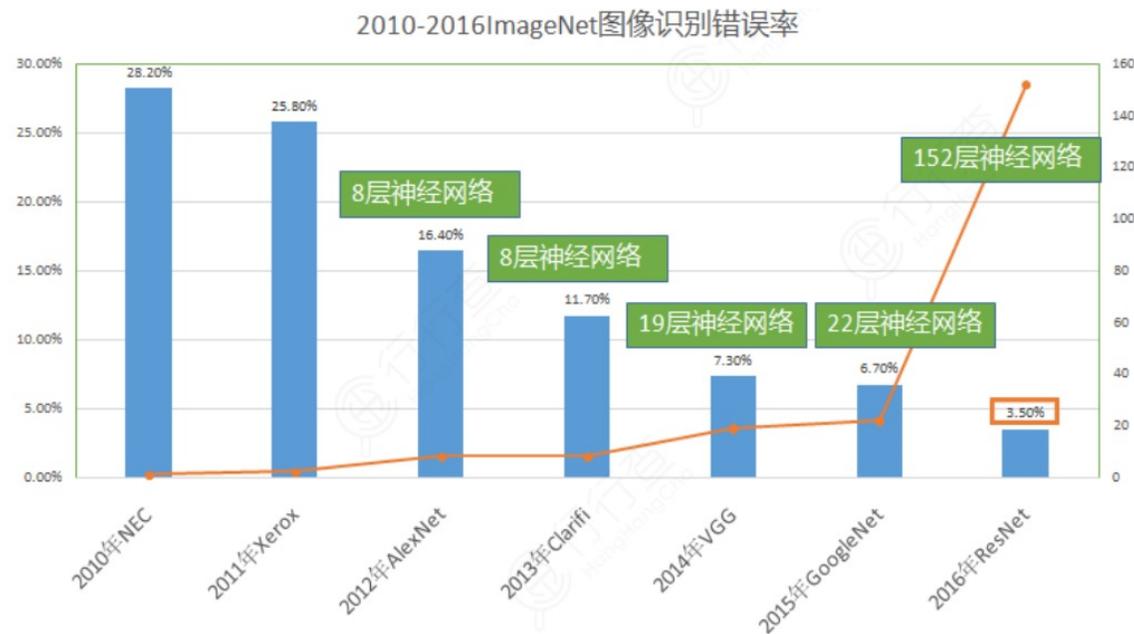


里程碑：AlexNet

ImageNet: Stanford创建的一个包含百万级图片和1000类别的大规模图像分类数据集



2012年深度卷积神经网络将错误率从26%降至16%，几年后降至约3%，超过人类水平

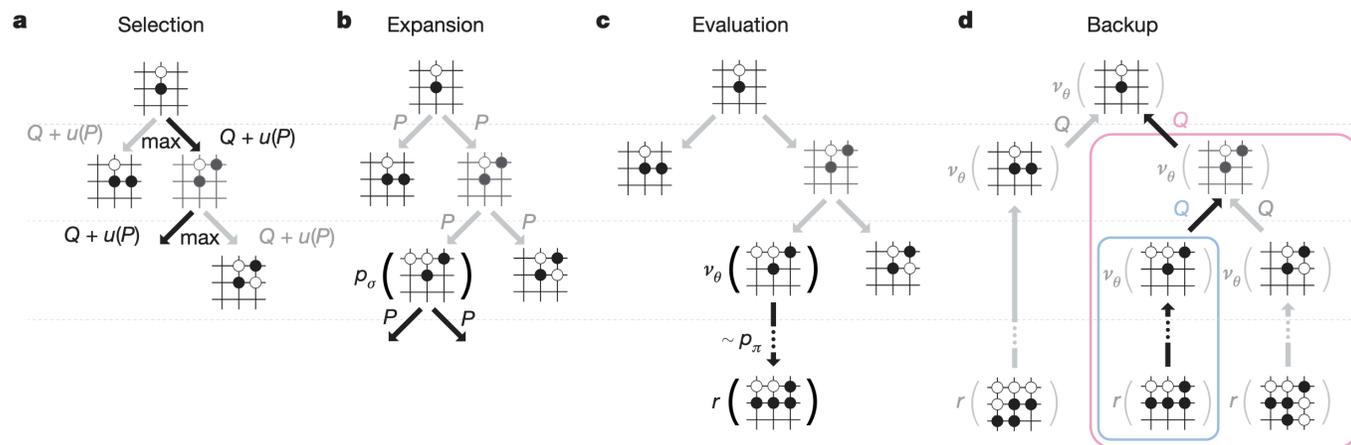


深度学习成为计算机视觉的主流范式

里程碑：AlphaGo

AlphaGo 利用强化学习、蒙特卡洛搜索、
自博弈等技术搜索和学习围棋策略

2016年以4:1战胜李世石九段
2017年以3:0战胜世界冠军柯洁



利用搜索与学习等技术，人工智能可以解决极其复杂的决策问题

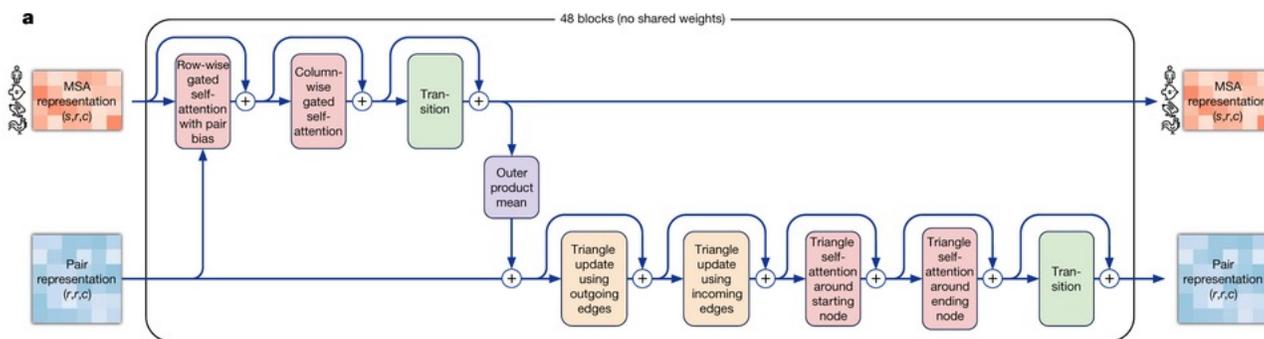
里程碑：AlphaFold

Google DeepMind 开发的 AlphaFold

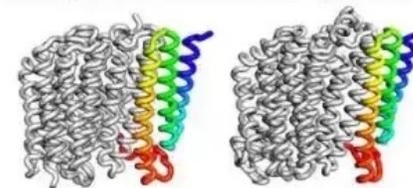
能够根据氨基酸序列预测蛋白质三维结构

2020年在CASP14评测中达到实验精度

随后预测了超过2亿个蛋白质结构

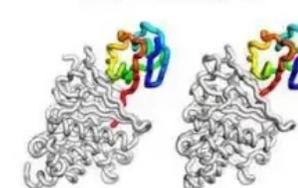


NADH-quinone oxidoreductase (nuoJ/nuoK)



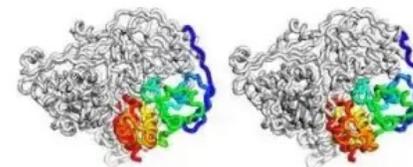
TM-score: 97

thiS/thiF complex



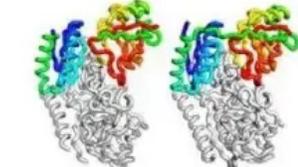
TM-score: 96

Aldehyde oxidoreductase (paoA/paoB)



TM-score: 95

5-oxoprolinase (pxpB/pxpC)



TM-score: 95

将原本需要数月实验的蛋白质结构在几分钟内预测出来，加速科学发现



The Nobel Prize in Physics 2024 was awarded jointly to John J. Hopfield and Geoffrey E. Hinton "for foundational discoveries and inventions that enable machine learning with artificial neural networks"

2024 诺贝尔物理学奖



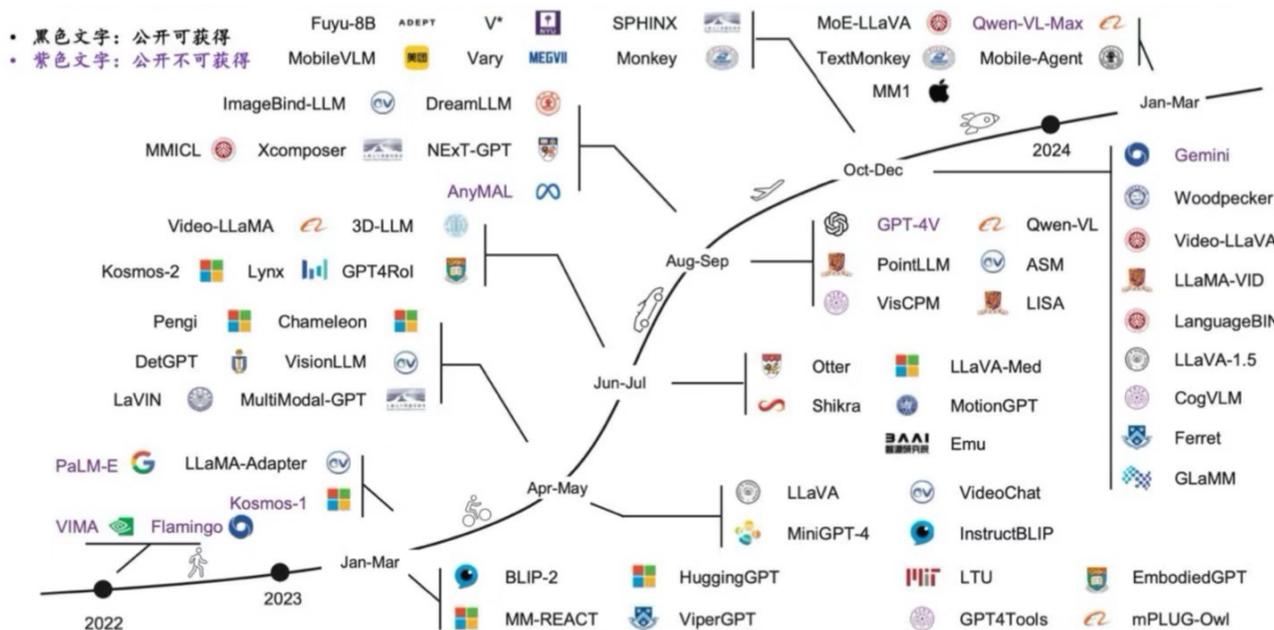
The Nobel Prize in Chemistry 2024 was awarded with one half to David Baker "for computational protein design" **and the other half jointly to Demis Hassabis and John Jumper** "for protein structure prediction"

2024 诺贝尔化学奖

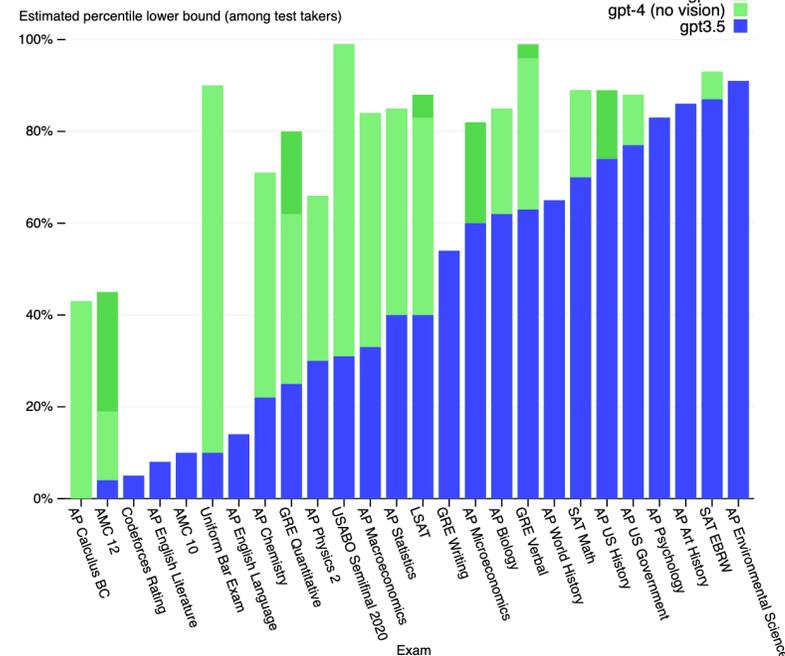
里程碑：ChatGPT

以 ChatGPT 为代表的使用大规模预训练、指令微调和强化学习等技术训练的大语言模型

具有强大的自然语言理解、生成、检索、分析整理、多轮对话能力等



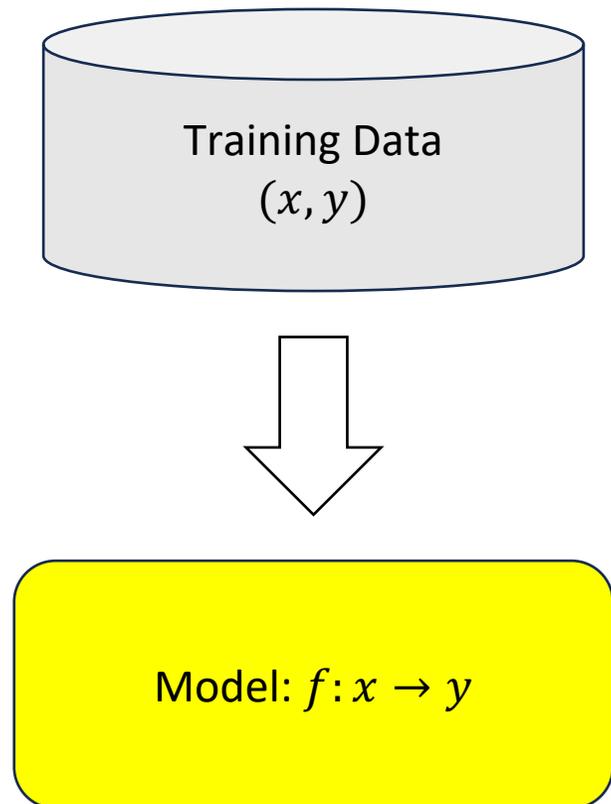
Exam results (ordered by GPT-3.5 performance)



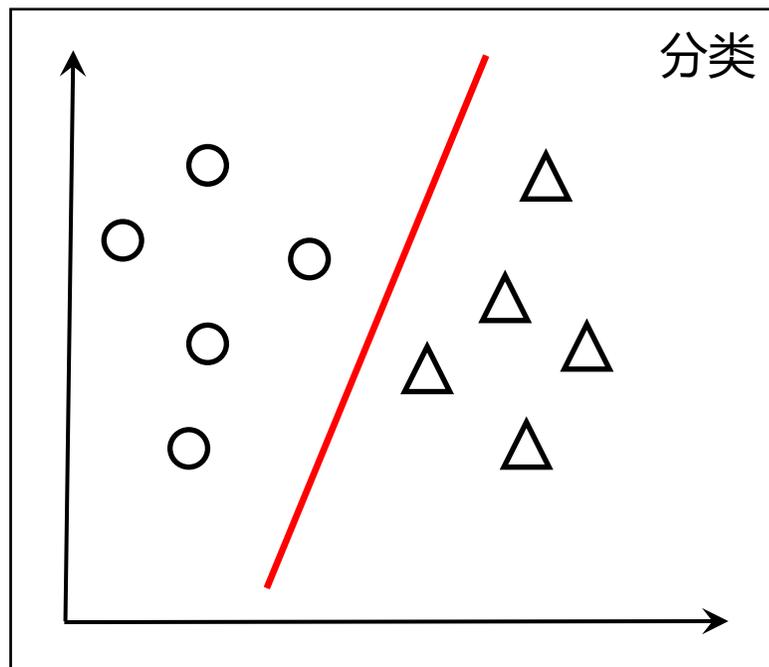
大模型成为通用辅助工具，改变人们的工作和生活方式

人工智能可以帮助解决哪些任务？

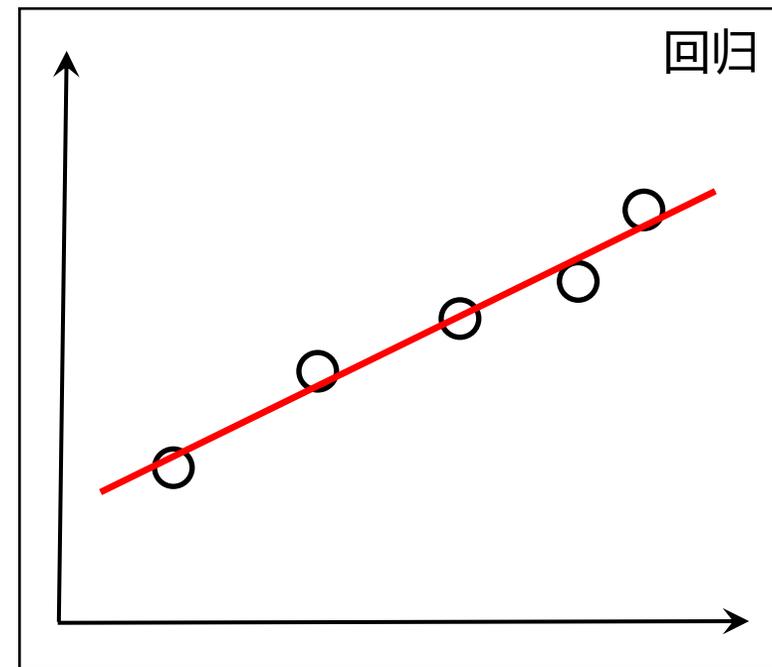
预测任务



分类



回归



人工智能可以帮助解决哪些任务？

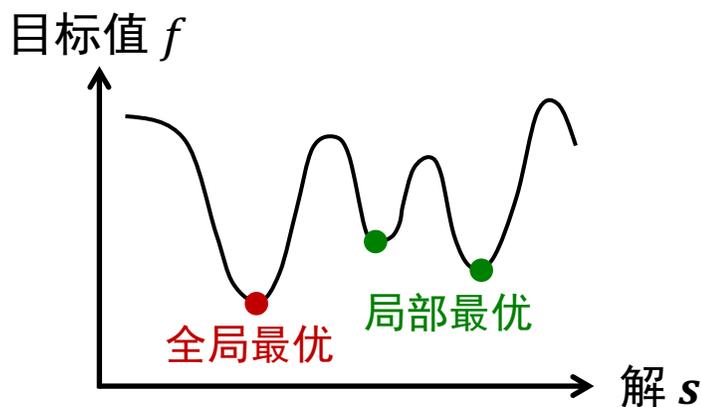
优化任务

$$\min_{s \in S} \text{目标函数} (f_1(s), f_2(s), \dots, f_m(s))$$

$$\text{s. t. } g_i(s) = 0, \quad 1 \leq i \leq q; \quad \text{等式约束}$$

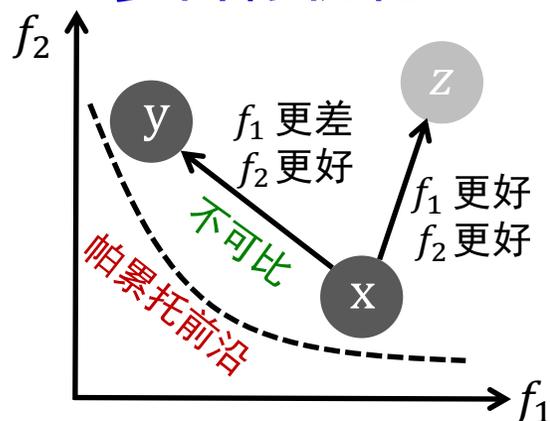
$$h_i(s) \leq 0, \quad q + 1 \leq i \leq m \quad \text{不等式约束}$$

单目标优化



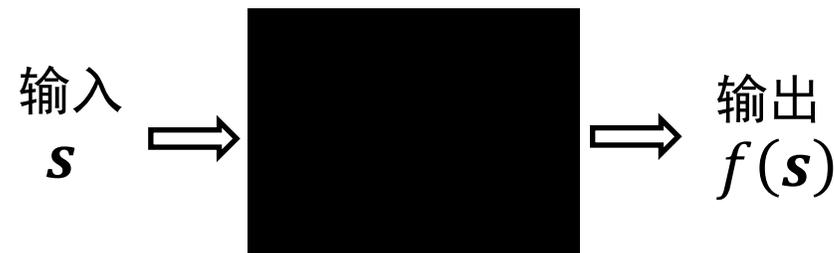
存在许多局部最优

多目标优化



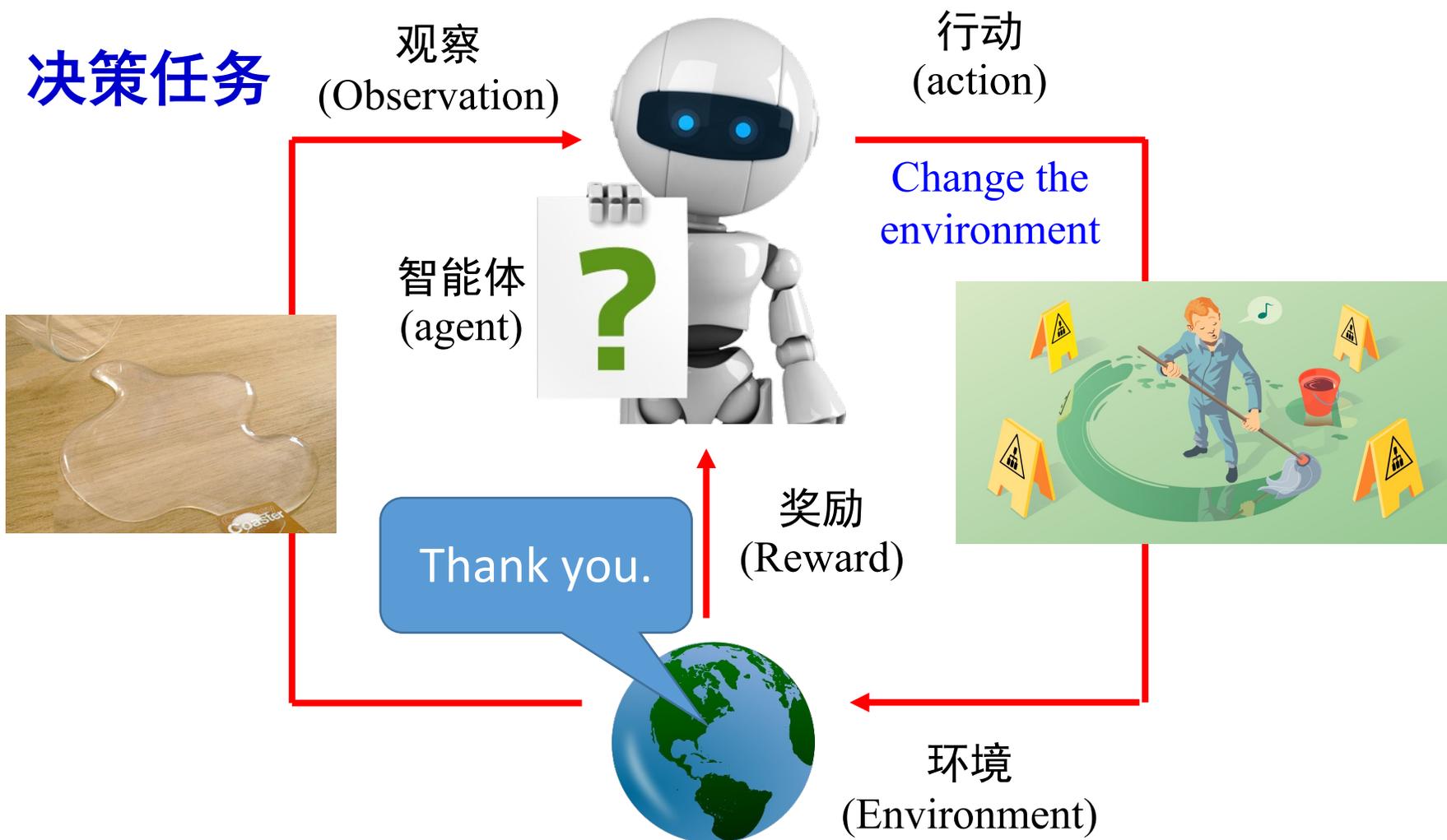
目标函数不唯一

黑箱优化



无法写出目标函数的解析式

人工智能可以帮助解决哪些任务？



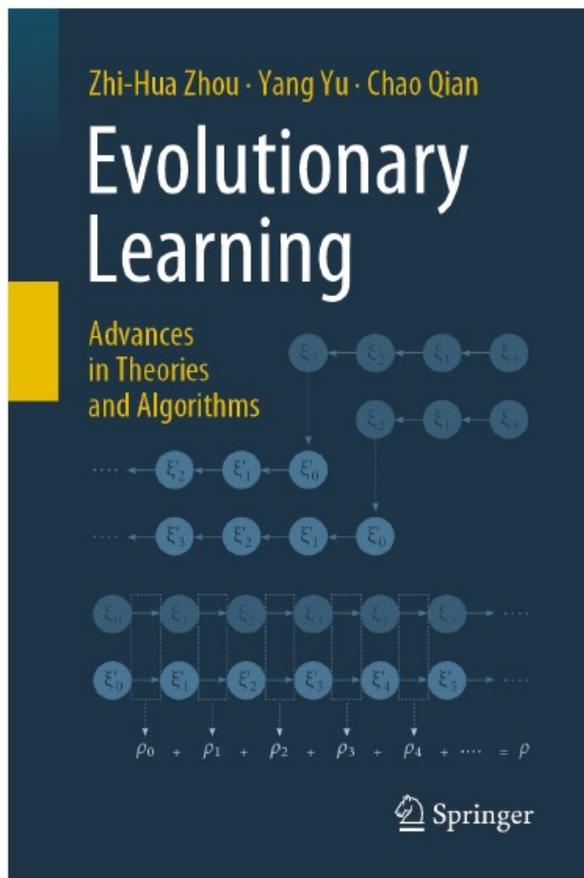
智能体 (Agent) 通过与环境交互，利用环境反馈的奖励信号来学习最优的行为策略

$$\pi: S \rightarrow A,$$

以最大化长期累积奖励

$$\sum_{t=0}^{\infty} \gamma^t r_t$$

本人研究方向：演化学习

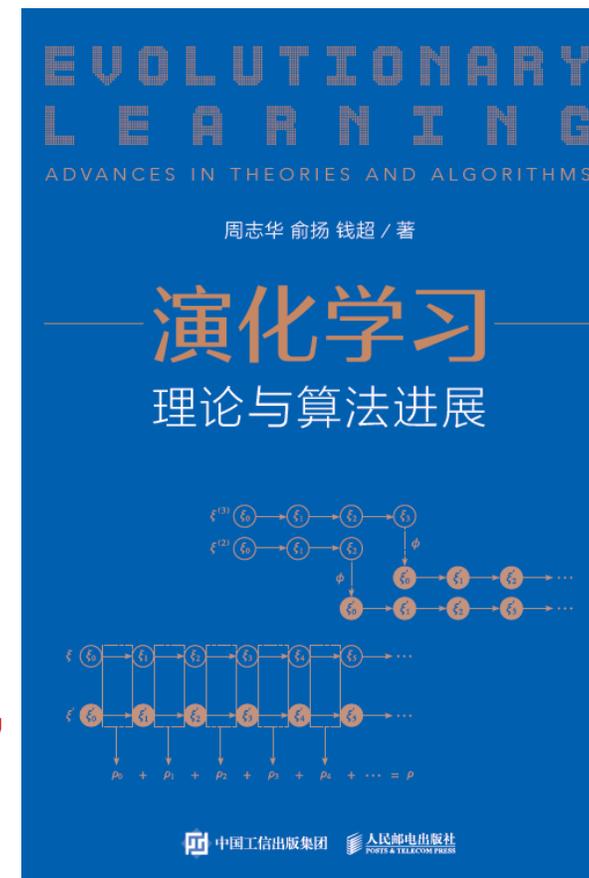


Zhi-Hua Zhou, Yang Yu, Chao Qian

Evolutionary Learning: Advances in Theories and Algorithms

- Presents theoretical results for evolutionary learning
- Provides general theoretical tools for analysing evolutionary algorithms
- Proposes evolutionary learning algorithms with provable theoretical guarantees

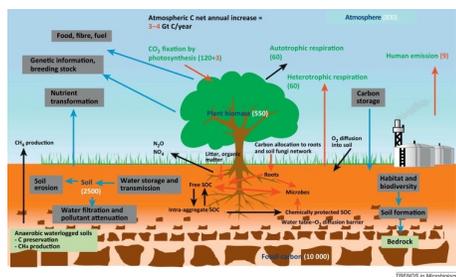
“Springer计算机学科年度最受欢迎作品”



交叉案例—预测任务：人工智能 + 土壤学

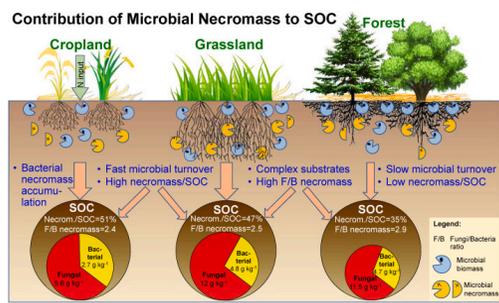
土壤有机碳对全球碳循环有着重大影响

土壤有机碳：最大的陆地碳库



缓解气候变化、保证粮食安全

微生物源碳
土壤有机碳重要组成成分



如何预测其含量？

现有方法：氨基糖分析
存在问题：预测误差大

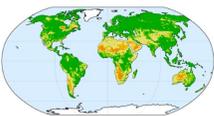
- 数据样本量有限
- 忽略细菌群组成效应

以共同第一作者发表
美国国家科学院院刊

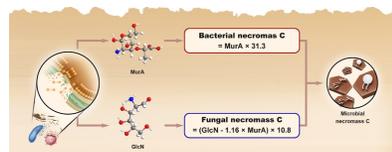
PNAS

与中科院南京土壤所
合作提出基于人工智能的
土壤微生物源碳估算公式

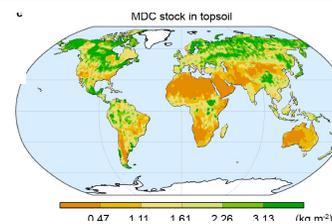
汇编全球数据集



联合建模、降低不确定性



使用机器学习、贝叶斯优化
精准预测关键参数



全球高精度土壤
微生物源碳分布

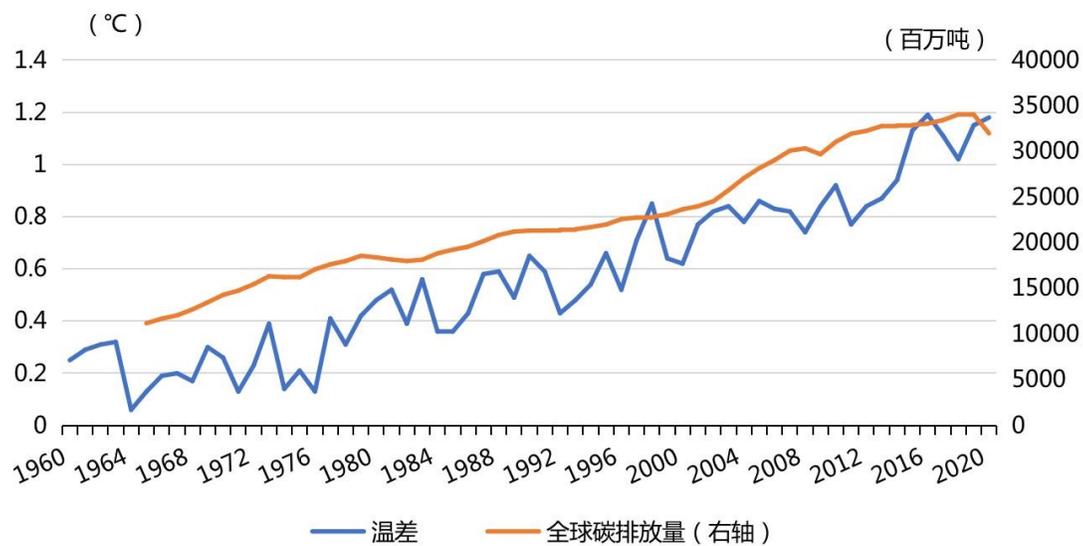
入选
2024年度
中国生态环境
十大科技进展

“双碳”行动：全球变暖下的中国承诺

■ **全球气候现状：**工业革命后，大气CO₂浓度升超50%，全球气温平均上升1.1度

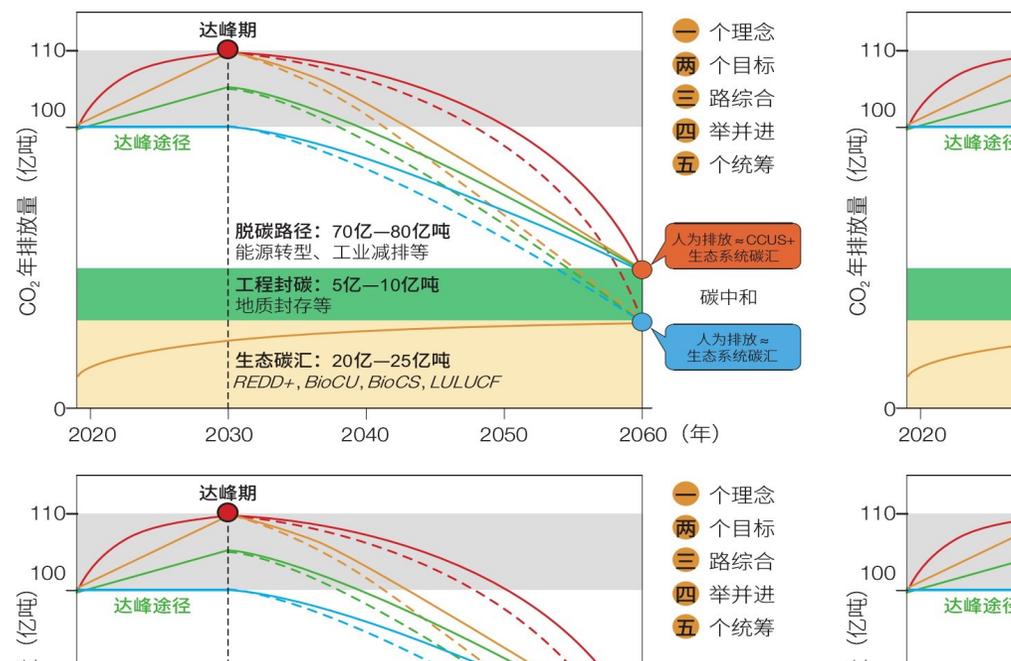
■ **中国“双碳”目标：**2030年前实现碳达峰，2060年前实现碳中和

全球CO₂排放总量和温度呈增长趋势



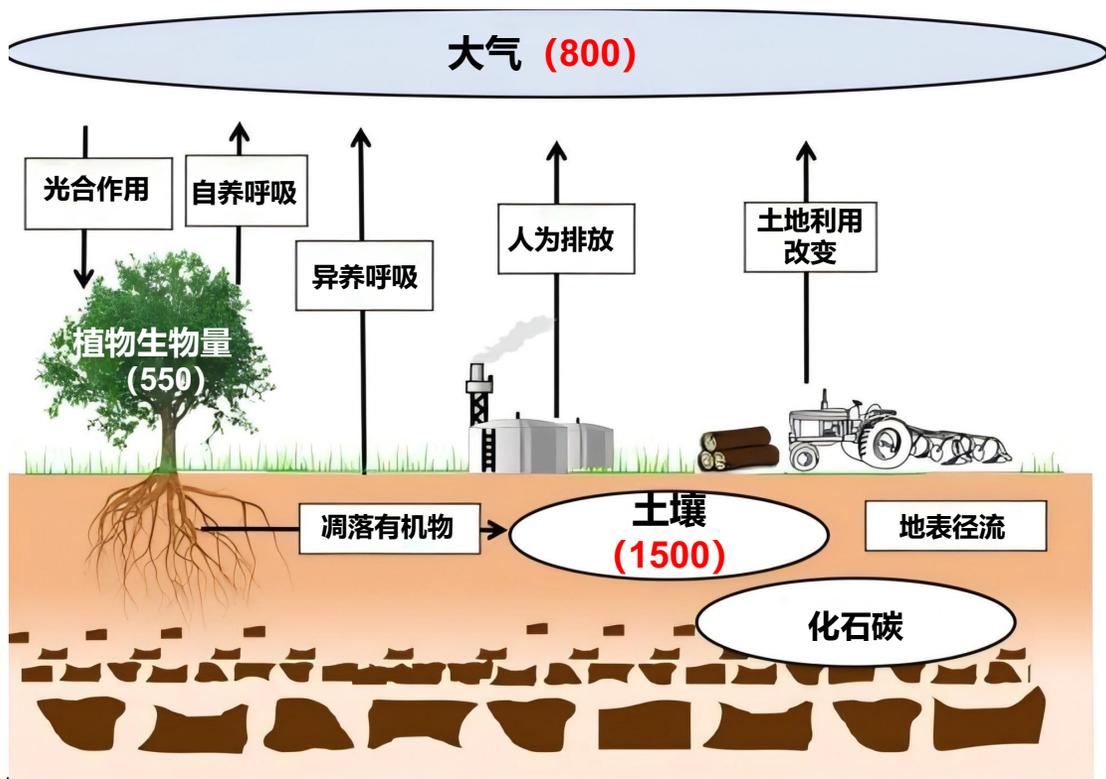
(CMF, 2021; BCAS, 2022)

中国“双碳”目标



土壤碳库：全球碳循环的“稳定器”

- **土壤有机碳库**：陆地生态系统中最大的碳库，其储量高于大气和植被碳库总和
- **千分之四计划**：表层1m土壤有机碳储量增4‰，可抵消当前全球CO₂净排放量

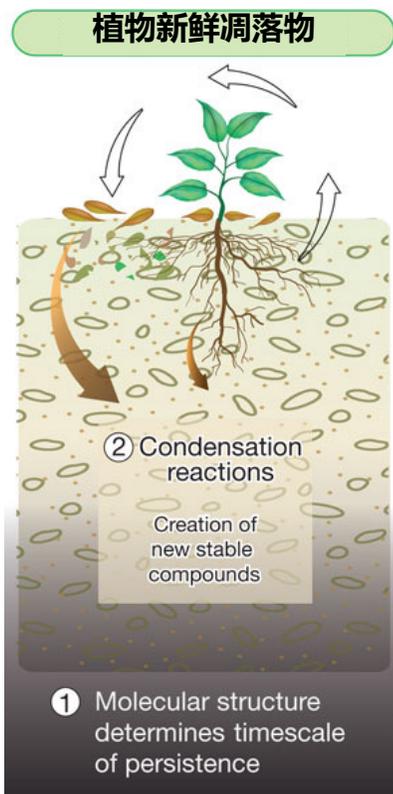


(Soil Carbon Storage, 2018)

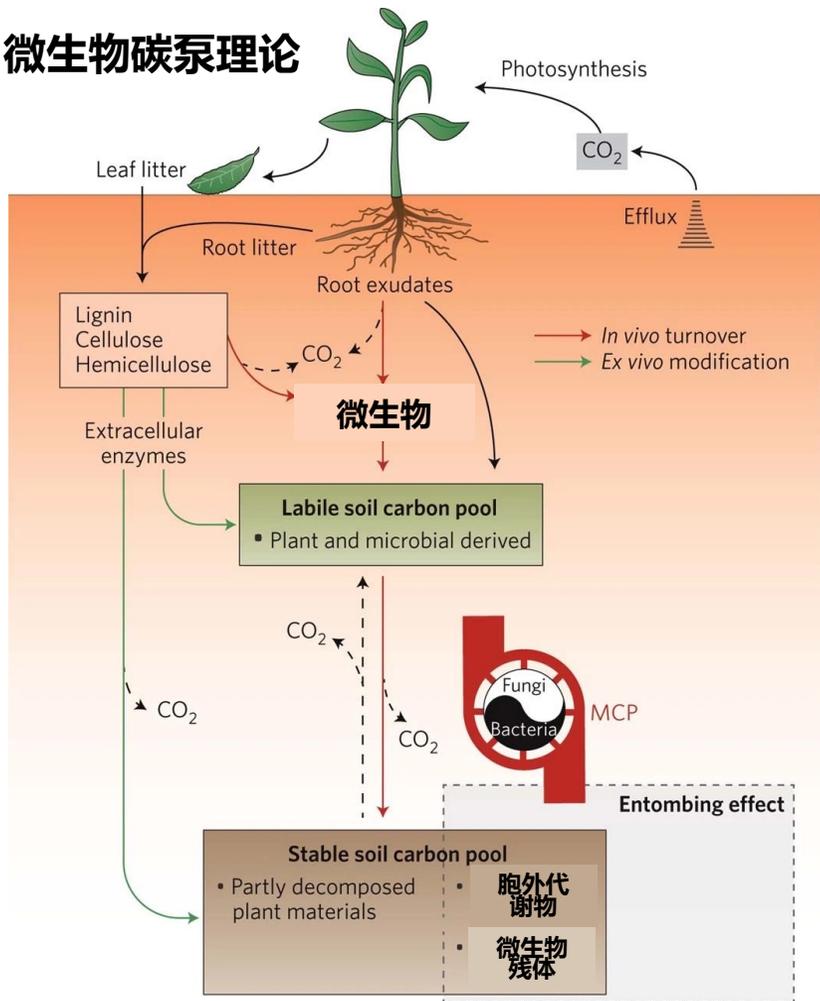


微生物源碳是土壤有机碳的重要组成

传统视角



微生物碳泵理论



传统视角

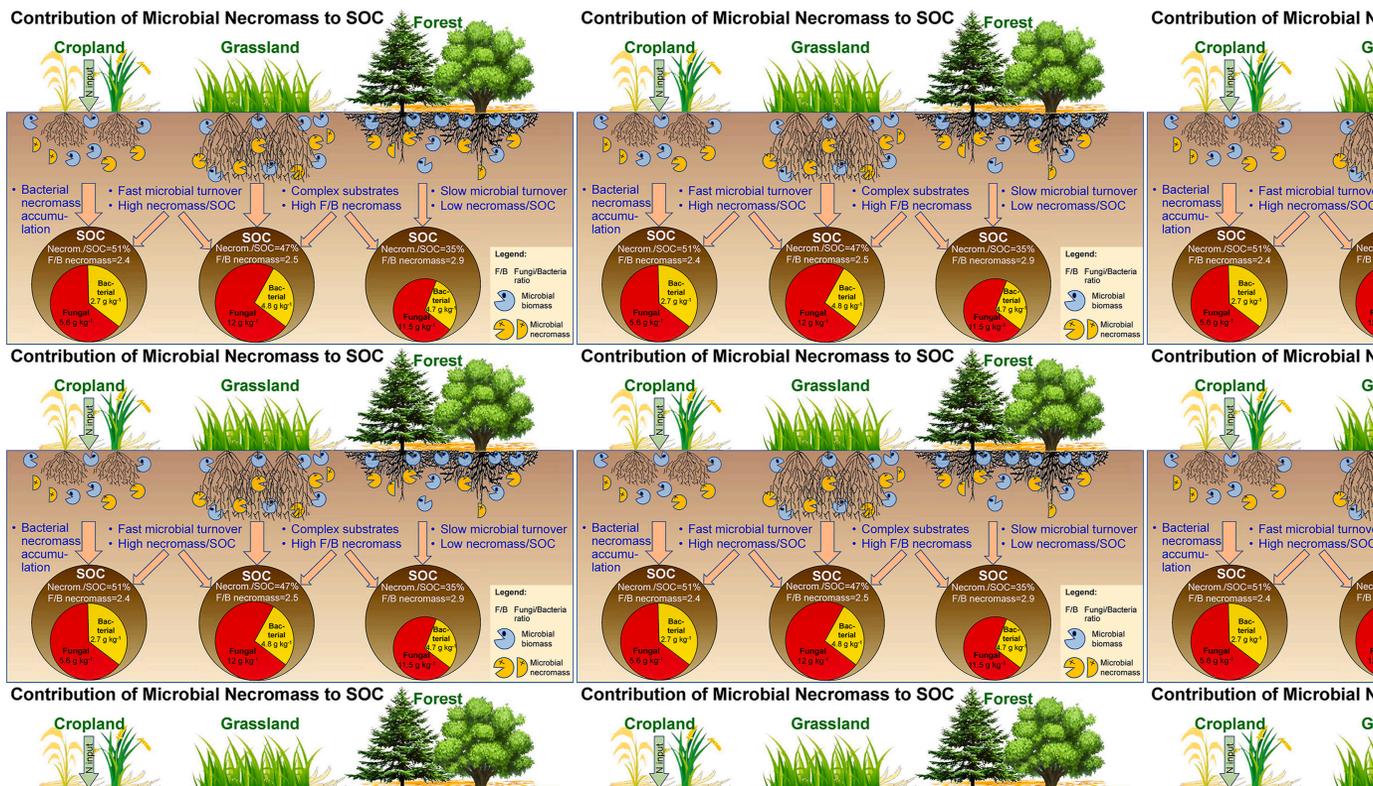
- 土壤的大部分碳直接来自植物碳输入

微生物碳泵理论

- 微生物通过一系列分解代谢和合成代谢将土壤中的碳转化为自身生物量
- 微生物死亡裂解后的残留物与矿物质结合成为稳定的微生物源碳 (MDC)

Nature, 2011, 478:49–56; Nat Microbiol, 2017, 2:17105

微生物源碳是土壤有机碳的重要组成



Q: 目前广泛接受的土壤微生物源碳测定方法?

SBB, 2021, 162:108422

- MDC 大约占总 SOC 库的50%，是持久性 SOC 库的重要组成部分
- 对 MDC 的定量评估对理解 SOC 稳定机制和预测气候变化下的陆地碳储存至关重要

氨基糖分析法



SBB, 2006, 38:1040-1051
GCB, 2022, 28:7167-7185

- 氨基糖是微生物细胞壁中的重要生物标志物，在细胞裂解后会在土壤中积累
- 通过对土壤中可以定量测定的氨基糖：胞壁酸（MurA）和氨基葡萄糖（GlcN）的浓度来反推微生物源碳的浓度

氨基糖分析法

2006年，德国土壤学家Joergensen提出：

- 在微生物死亡、周转及稳定过程中，氨基糖在细胞干重中的占比相对稳定
- 通过将土壤中**MurA**和**GlcN**浓度分别乘以固定**转化系数**，可大致估算土壤中**细菌源碳**和**真菌源碳**浓度



$$\text{细菌源碳} = \text{MurA} \times 45$$

$$\text{真菌源碳} = (\text{GlcN} - 1.43 \times \text{MurA}) \times 9$$

$$\text{微生物源碳} = \text{细菌源碳} + \text{真菌源碳}$$



Soil Biology & Biochemistry 38 (2006) 1040–1051

Soil Biology &
Biochemistry

www.elsevier.com/locate/soilbio

Microbial colonisation of roots as a function of plant species

Astrid Appuhn, Rainer Georg Joergensen *

Department of Soil Biology and Plant Nutrition, University of Kassel, Nordbahnhofstr. 1a, 37213 Witzenhausen, Germany

Received 6 July 2005; received in revised form 3 September 2005; accepted 8 September 2005
Available online 27 October 2005

Abstract

Fifteen plants species were grown in the greenhouse on the same soil and sampled at flowering to obtain rhizosphere soil and root material. In both fractions, the data on fungal and bacterial tissue obtained by amino sugar analysis were compared with the total microbial biomass based on fumigation-extraction and ergosterol data. The available literature on glucosamine concentrations in fungi and on muramic acid concentrations in bacteria was reviewed to prove the possibility of generating conversion values for general use in root material. All microbial properties analysed revealed strong species-specific differences in microbial colonisation of plant roots. The root material contained considerable amounts of microbial biomass C and biomass N, reaching mean levels of 10.9 and 1.4 mg g⁻¹ dry weight, respectively. However, the majority of CHCl₃ labile C and N, i.e. 89 and 55% was root derived. The average amount of ergosterol was 13 μg g⁻¹ dry weight and varied between 0.0 for *Phaseolus* roots and 45.5 μg g⁻¹ dry weight for *Vicia* roots. The ergosterol content in root material of mycorrhizal and non-mycorrhizal plant species did not differ significantly. Fungal glucosamine was converted to fungal C by multiplication by 9 giving a range of 7.1–25.9 mg g⁻¹ dry weight in the root material. Fungal C and ergosterol were significantly correlated. Bacterial C was calculated by multiplying muramic acid by 45 giving a range from 1.7 to 21.6 mg g⁻¹ dry weight in the root material. In the root material of the 15 plant species, the ratio of fungal C-to-bacterial C ranged from 1.0 in mycorrhizal *Trifolium* roots to 9.5 in non-mycorrhizal *Lupinus* roots and it was on average 3.1. These figures mean that the microbial tissue in the root material consists on average of 76% fungal C and 24% bacterial C. The differences in microbial colonisation of the roots were reflected by differences in microbial indices found in the rhizosphere soil, most strongly for microbial biomass C and ergosterol, but to some extent also for glucosamine and muramic acid.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Ergosterol; Glucosamine; Muramic acid; Microbial biomass; Rhizosphere; Fungi; Bacteria

细菌转化因子的估算

细菌细胞碳浓度: $\sim 460 \text{ mg g}^{-1}$

$$\text{细菌源碳} = \text{MurA} \times \frac{460}{13.9 \times 0.65 + 3.7 \times 0.35} = \text{MurA} \times 45$$

土壤MDC
“估不准”

样本数量不足

革兰氏阳性菌 (GP) 平均MurA浓度: 13.9 mg g^{-1}
(基于 26 种细菌菌株)

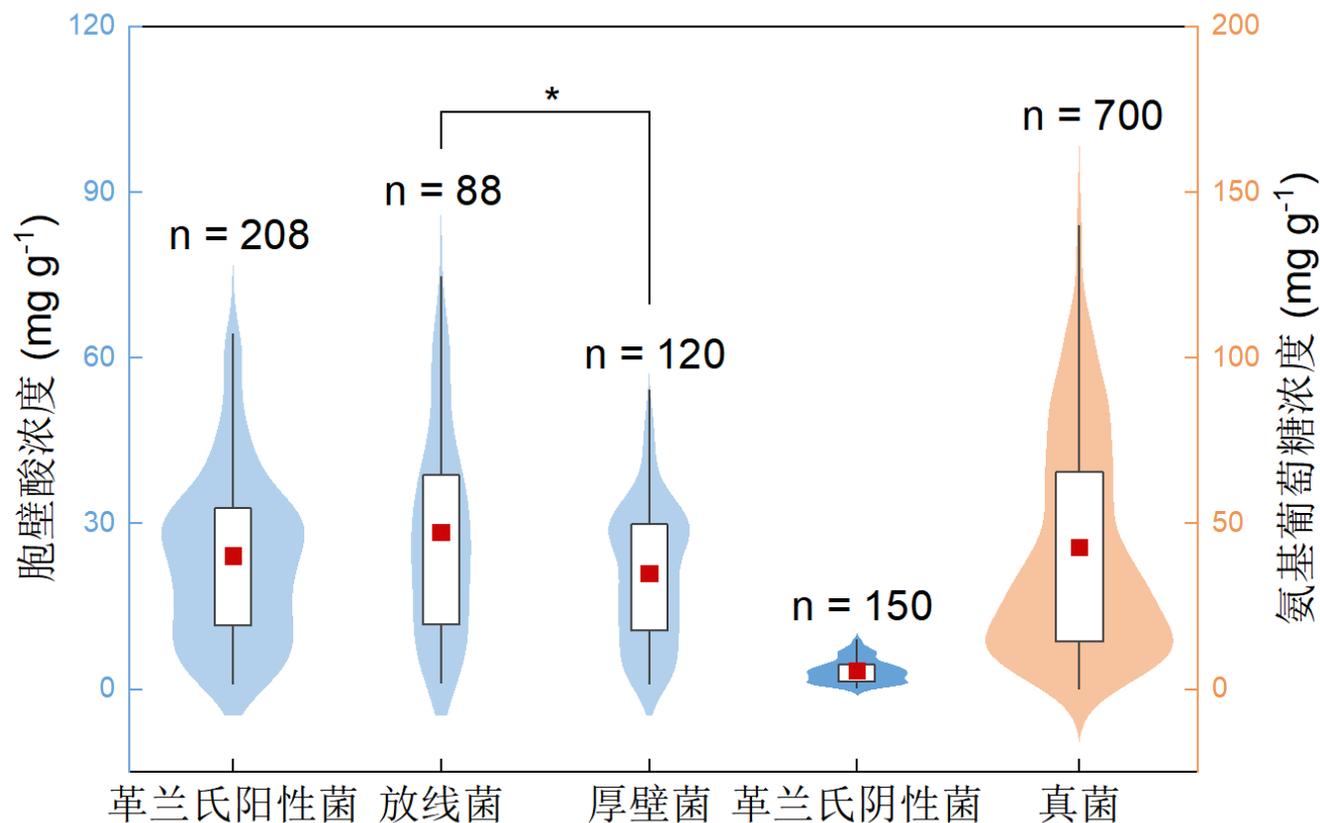
革兰氏阴性菌 (GN) 平均MurA浓度: 3.7 mg g^{-1}
(基于 25 种细菌菌株)

- 细菌群落组成: 测定当地原位土壤磷脂脂肪酸, $\text{GP:GN} = 0.65:0.35$,
并假定该比例全球适用

忽略了细菌组成的全球变化

重估浓度：扩充数据集

细菌菌株的 MurA 数据：51种 扩充至 358种



- 革兰氏阳性菌 (GP) 平均MurA浓度：**24.1** mg g⁻¹ (原先 **13.9**)
 - 放线菌的平均MurA浓度 **28.4** mg g⁻¹
 - 厚壁菌的平均MurA浓度 **20.9** mg g⁻¹
- 革兰氏阴性菌 (GN) 平均MurA浓度：**3.3** mg g⁻¹ (原先 **3.7**)

细菌转化因子的估算

细菌的转换因子需分为放线菌、厚壁菌以及GN三类进行氨基糖浓度的计算

细菌转换因子计算公式

$$\frac{460}{(13.9 \times 0.65 + 3.7 \times 0.35)} = 45$$



$$\frac{460}{(20.9 \times a + 28.4 \times b + 3.3 \times (1 - a - b))}$$

a 和 b 分别代表细菌群落中的厚壁菌和放线菌的比例

测定当地原位土壤磷脂脂肪酸，
GP:GN = 0.65:0.35，并假定该比例全球适用

如何估计 a 和 b ?

全球上高度可变

运行人工智能方法进行预测

构建数据集

- 目标变量：ln (Firmicutes: GN) 、ln (Actinobacteria: GN) 、MurA 和 GlcN
- 输入变量：包括
 - 13 个土壤变量 (SOC、TN、TP、SOC/TN、SOC/TP、MBC、MBN、MBC/MBN、pH、clay、silt、sand)
 - 3 个气候变量 (MAT、MAP、PET)
 - 海拔、纬度和生态系统类型
- 数据集 (规模 3063) : 随机分为训练集和测试集 (70% : 30%)

构建机器学习模型池

回归任务适用模型

经典模型

RandomForest

XGBoost

LightGBM

CatBoost

DeepForest

AutoSklearn

**Multilayer
Perceptron**

深度神经网络模型

CNN

ResNet

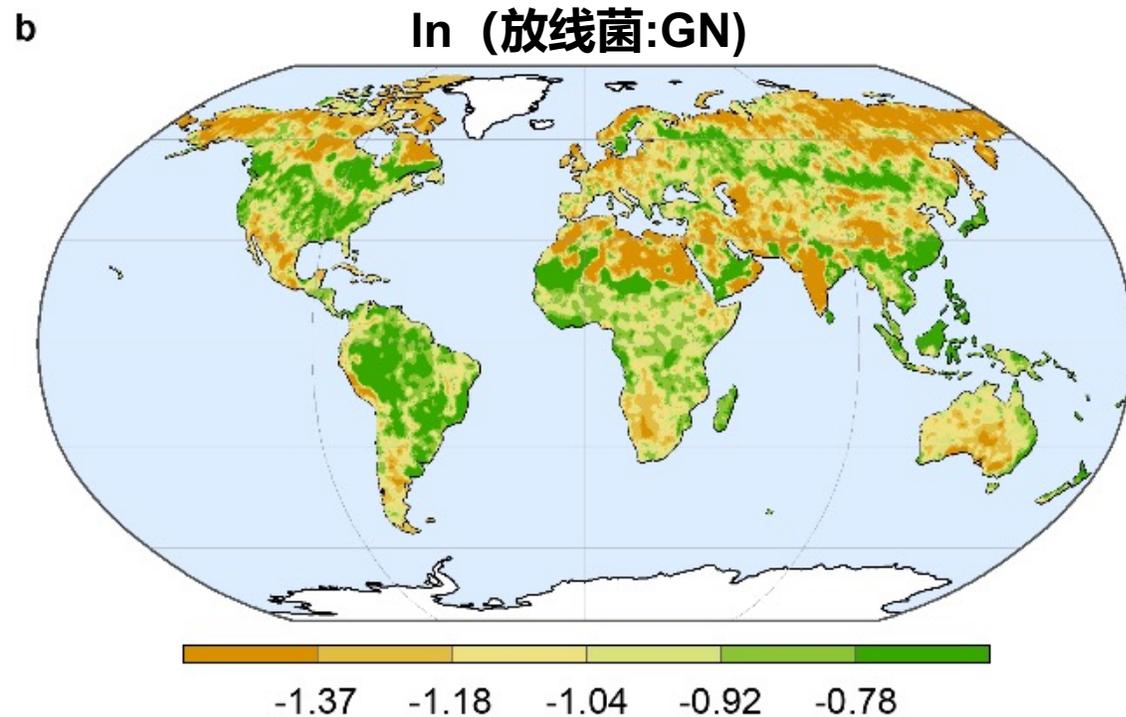
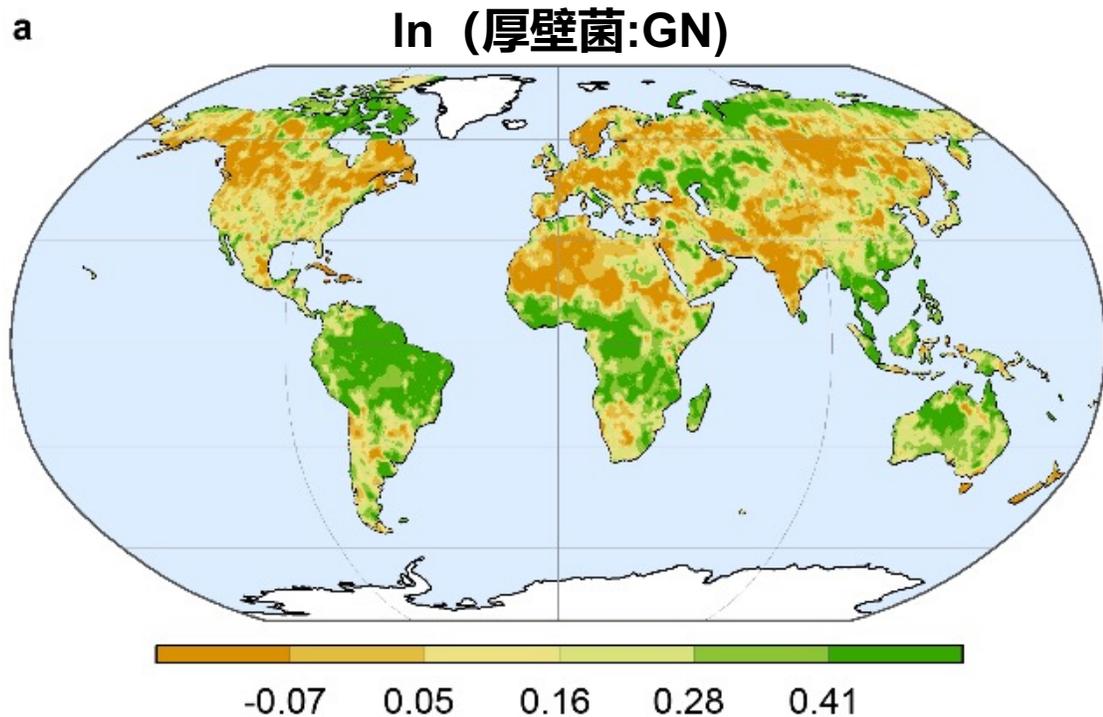
FT-Transformer

选择最佳模型

- 利用机器学习方法估算了土壤中的 $\ln(\text{Firmicutes: GN})$ ($R^2 = 0.82$)、 $\ln(\text{Actinobacteria: GN})$ ($R^2 = 0.76$)、MurA ($R^2 = 0.81$) 和 GlcN ($R^2 = 0.94$) 的全球分布

Model	$\ln(\text{Act:GN})$	$\ln(\text{Firm:GN})$	MurA	GlcN
Test set				
Random Forest	-0.006 ± 0.021	0.381 ± 0.004	0.583 ± 0.007	0.780 ± 0.002
CatBoost	0.127 ± 0.045	0.436 ± 0.023	0.601 ± 0.035	0.851 ± 0.009
LightGBM	-0.031 ± 0.045	0.320 ± 0.025	0.541 ± 0.014	0.798 ± 0.007
XGBoost	-1.779 ± 0.000	0.218 ± 0.000	0.581 ± 0.000	0.832 ± 0.000
Auto-Sklearn	0.091 ± 0.034	0.280 ± 0.007	0.543 ± 0.036	0.833 ± 0.010
Deep Forest	0.001 ± 0.002	0.481 ± 0.004	0.615 ± 0.008	0.832 ± 0.017
Multilayer Perceptron	0.228 ± 0.033	0.501 ± 0.020	0.537 ± 0.035	0.828 ± 0.014
CNN	-0.591 ± 0.296	-0.154 ± 0.138	0.379 ± 0.164	0.664 ± 0.085
ResNet	0.328 ± 0.017	0.565 ± 0.016	0.534 ± 0.026	0.834 ± 0.007
FT-Transformer	0.308 ± 0.044	0.511 ± 0.025	0.615 ± 0.020	0.872 ± 0.008

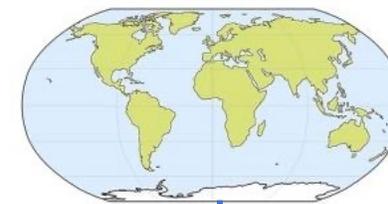
$\ln(\text{放线菌:GN})$ 和 $\ln(\text{厚壁菌:GN})$ 的全球分布



细菌源碳精准估算

机器学习预测得到厚壁菌：放线菌：GN的全球分布

假设的厚壁菌：放线菌：GN比例



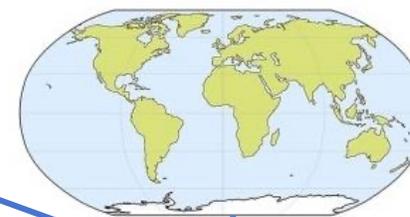
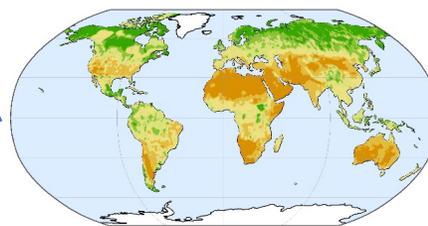
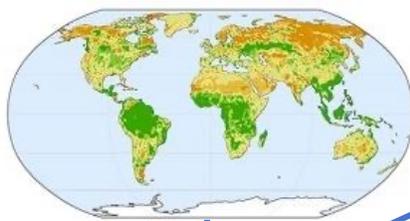
细菌转换因子计算公式

$$\frac{460}{(20.9 \times a + 28.4 \times b + 3.3 \times (1 - a - b))}$$

预测情况的转换因子全球分布

土壤 MurA 浓度的全球分布

假设情况的转换因子全球分布



相乘

相乘

预测情况的细菌源碳全球分布

假设情况的细菌源碳全球分布

结果比较，计算R²

**R²会随着假设参数的变化而变化
R²最大时，即为最优参数**



细菌源碳精准估算

在全球尺度上选取**厚壁菌：放线菌：革兰氏阴性菌=0.478:0.119:0.403**，
R²最大，公式在全球尺度上适用性最高

更新后细菌转换因子计算公式

$$\frac{460}{(20.9 \times a + 28.4 \times b + 3.3 \times (1 - a - b))}$$

此时， $a = 0.478$ ， $b = 0.119$

更新后的细菌转换因子：31.3

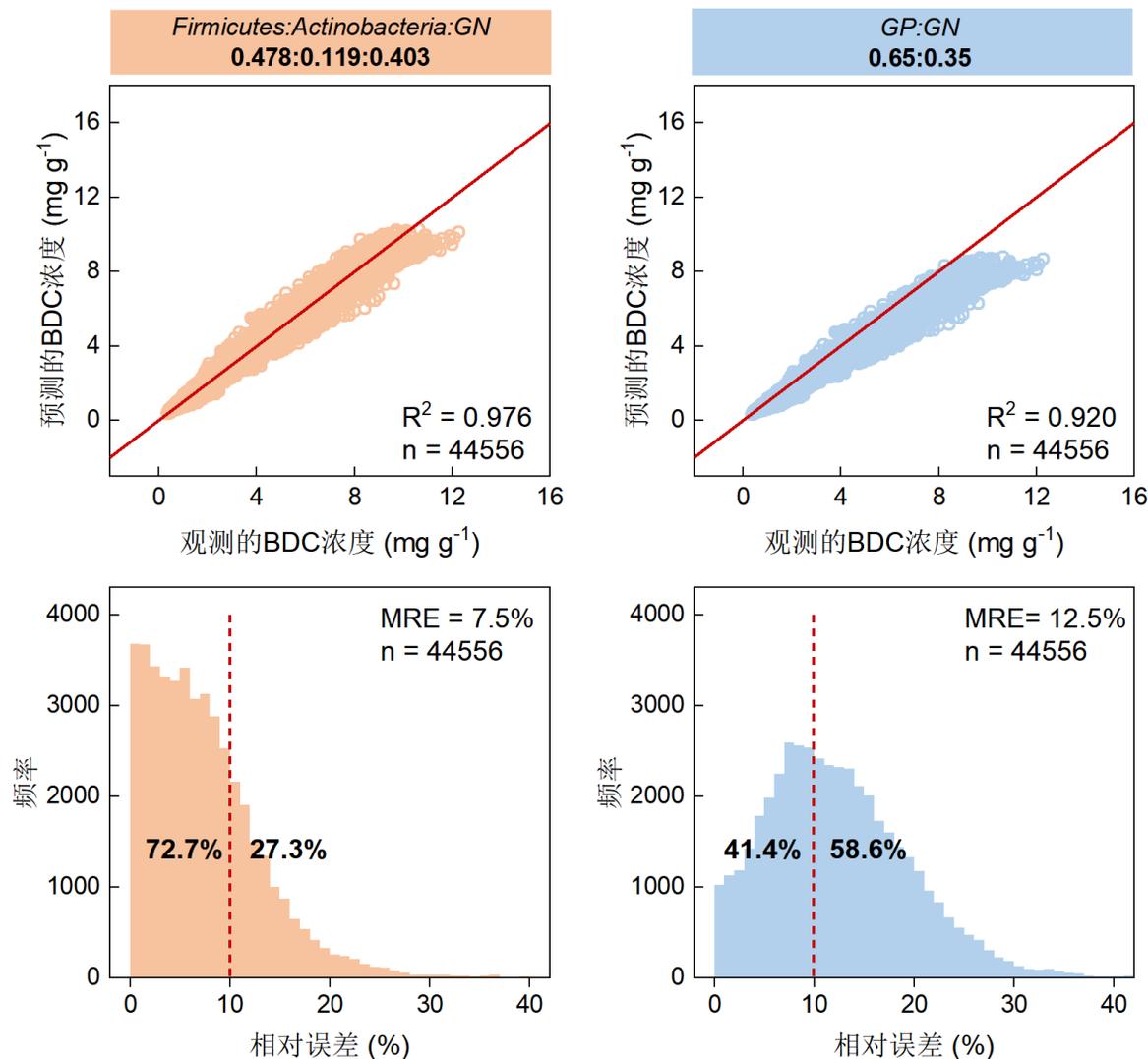
原细菌转换因子计算公式

$$\frac{460}{(13.9 \times 0.65 + 3.7 \times 0.35)} = 45$$

测定当地原位土壤磷脂脂肪酸，

GP:GN = 0.65:0.35，并假定该比例全球适用

细菌源碳精准估算



更新后细菌转换因子计算

- 全球平均误差**7.5%**
- **27%**地区误差超10%

原细菌转换因子计算

- 全球平均误差**12.5%**
- **59%**地区误差超10%

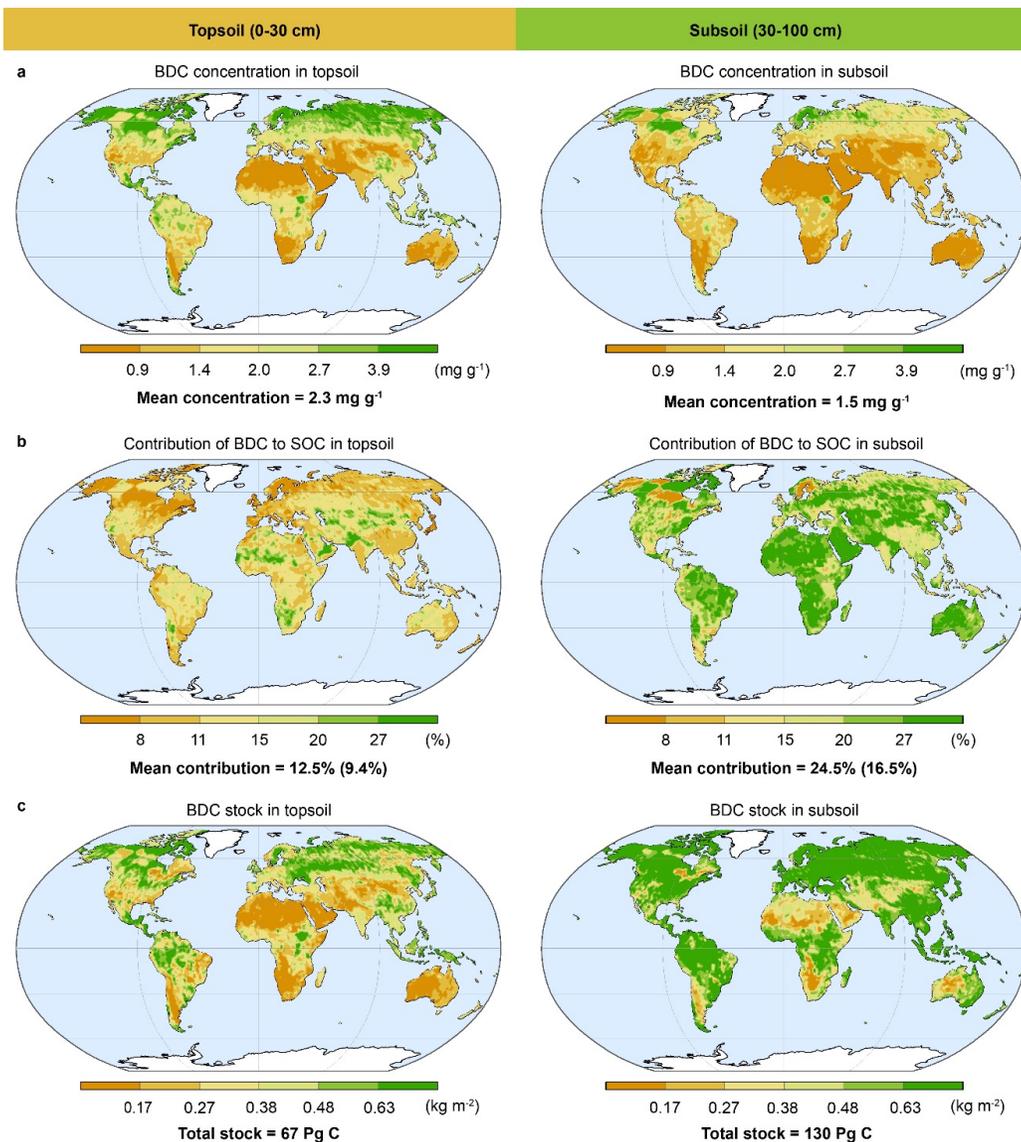
更新后细菌转换因子计算
可大幅提升准确性

细菌源碳精准估算

	热带气候 Tropical	干旱气候 Arid	温带气候 Temperate	寒带气候 Cold	极地气候 Polar
农田 Cropland	30.3	32.7	30.1	30.6	/
森林 Forest	28.9	30.9	29.9	32.1	31.0
草地 Grassland	29.9	31.0	30.7	31.9	29.6
湿地 Wetland	28.8	32.1	30.3	32.5	31.1
沙漠 Desert	/	31.5	/	30.5	29.7
冰川 Glacier	/	/	/	34.2	33.4

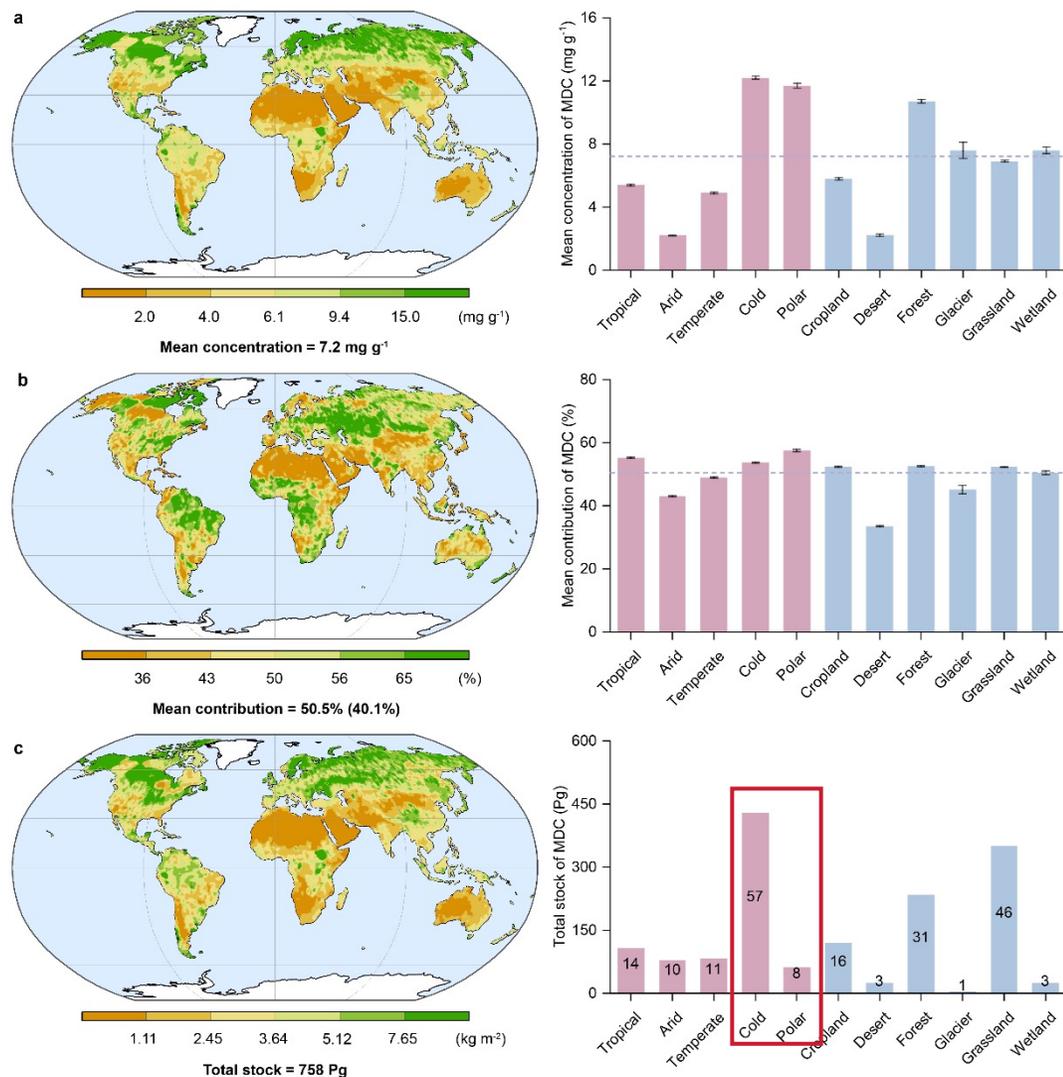
- 针对跨气候带或跨生态系统的研究，建议采用**可变细菌转换因子**来替代通用转化因子，以获得更精确的估算效果

全球BDC储量估计



- 表土和底土中 BDC 总浓度的全球平均值分别为 2.3 mg g⁻¹和1.5 mg g⁻¹
- 估计全球 BDC 总储量约为 197 Pg，其中 67 Pg储存在表土中，130 Pg储存在底土中

全球MDC储量估计



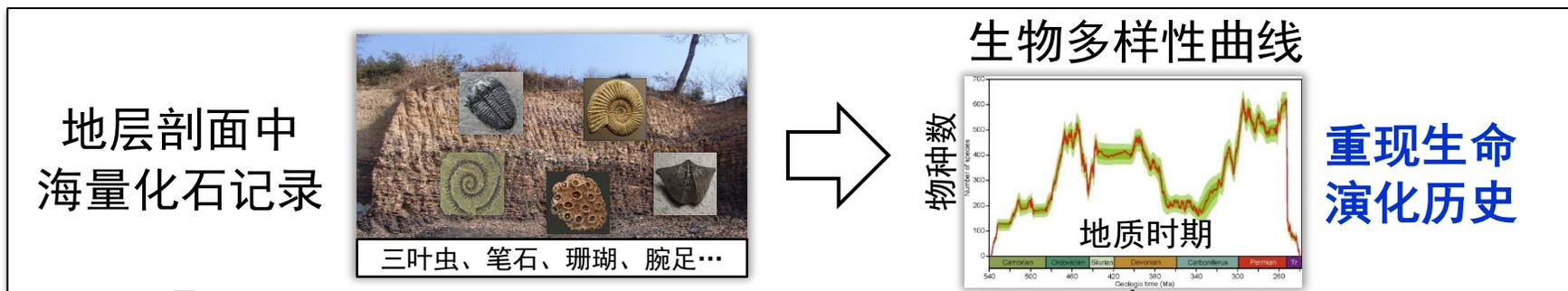
- ❑ MDC 浓度的全球平均值为 7.2 mg g^{-1} ，估计全球 MDC 总储量约为 758 Pg
- ❑ 全球MDC对土壤SOC的贡献平均为 50.5%
- ❑ 寒带和极地气候区储存了约 65% 的 MDC，需警惕变暖导致的MDC分解

微生物源碳浓度 (a)、贡献 (b) 和储量 (c) 的全球分布

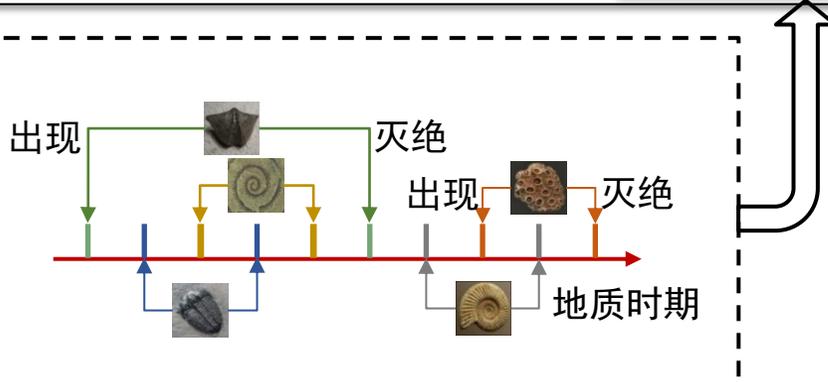
交叉案例—优化任务：人工智能 + 地质学

深时数字地球国际大科学计划核心科学内容

自然科学基本问题 — 生命起源与演化



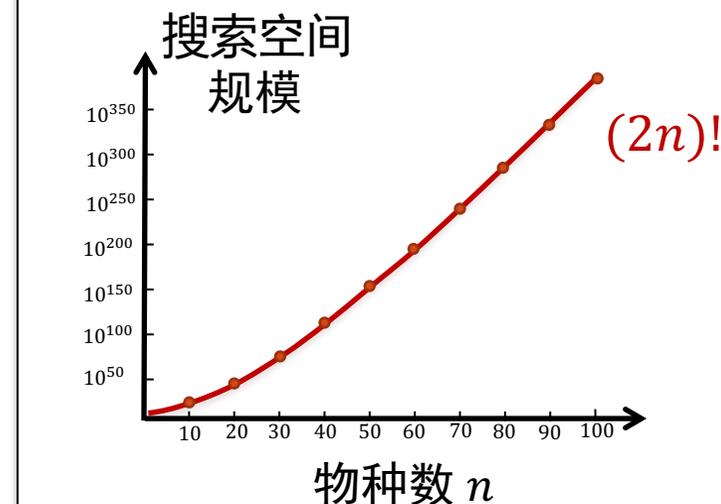
序列优化



对不同物种的“出现”和“灭绝”事件排序，使其与地层剖面中采集到的化石记录尽可能一致

计算复杂
目标黑箱

搜索空间规模
关于物种数呈指数级增长



大规模：11268个物种

交叉案例—优化任务：人工智能 + 地质学

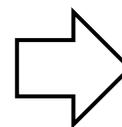
自然科学四大
基础科学问题
之一：
生命起源与演化

地层剖面
海量化石
记录数据



三叶虫、笔石、珊瑚、腕足...

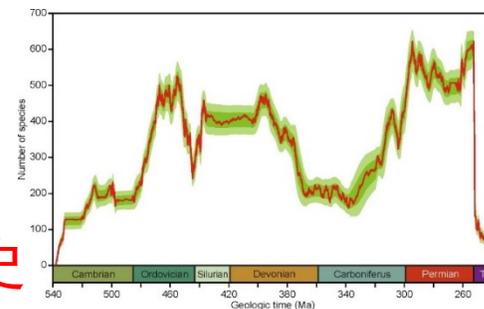
利用化石记录



重现生命演化历史

生物多样性变化曲线

物种数



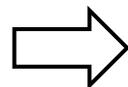
南京大学地球科学与工程学院
研究成果

中国的地层剖面数据

3122个剖面

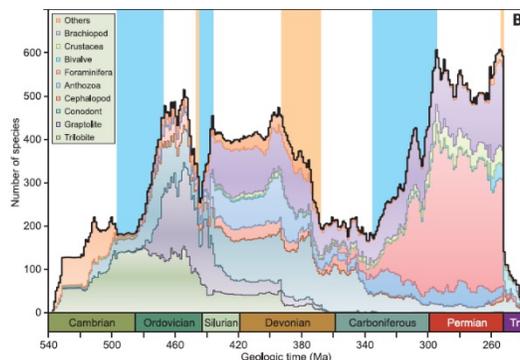
11268个物种

演化算法



“天河2号”
700万核时

全球第一条高精度
海洋生物多样性变化曲线



Science

Contents

News

Careers

Journals

SHARE RESEARCH ARTICLE



A high-resolution summary of Cambrian to Early Triassic marine invertebrate biodiversity

Jun-xuan Fan^{1,2}, Shu-zhong Shen^{1,2,3,*}, Douglas H. Erwin^{4,5}, Peter M. Sadler⁶, Norman MacLeod¹, Qiu-min...

Science: “新的数据集和方法，推动整个演化生物学的变革”

Nature: “古生物学家以惊人的细节绘制地球3亿年历史”

2020 年中国科学十大进展

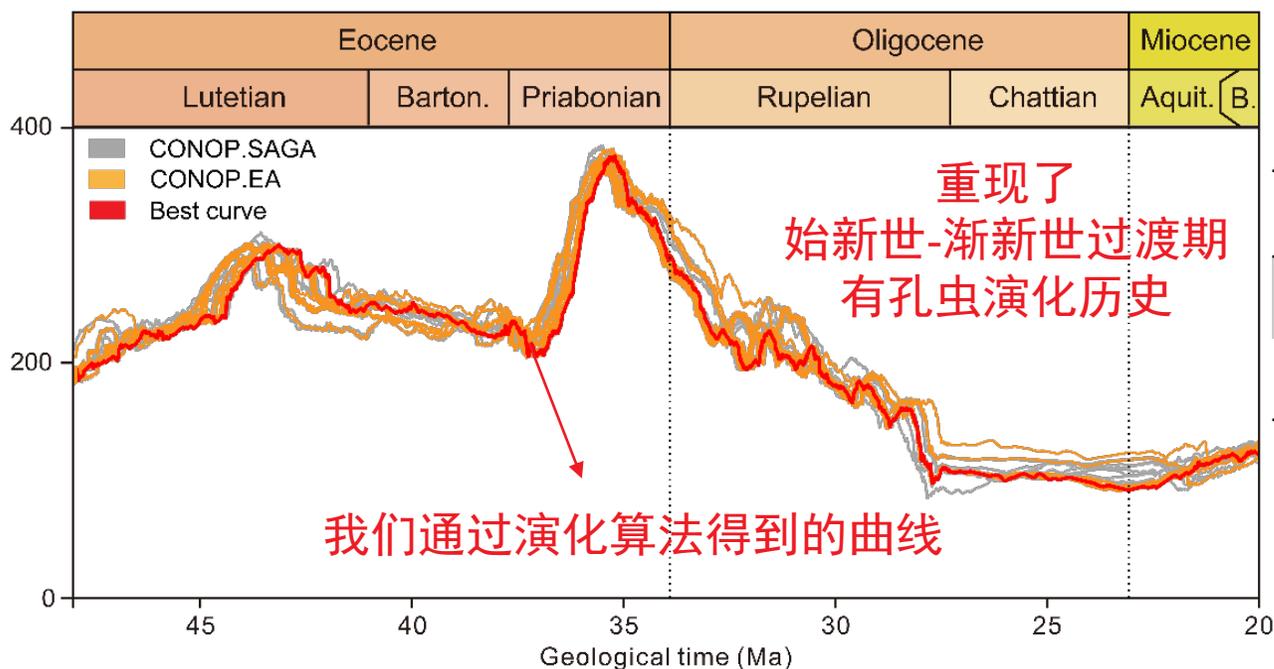
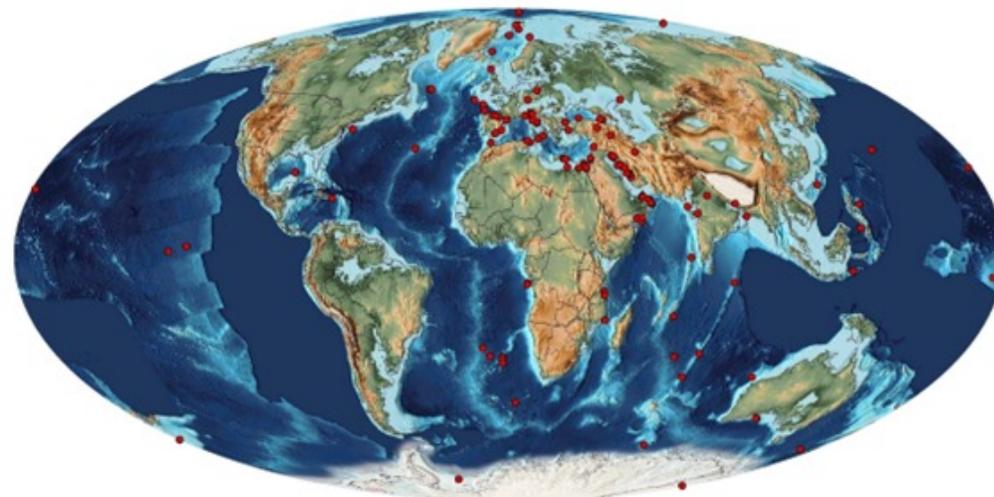
交叉案例—优化任务：人工智能 + 地质学

地科团队收集的数据集

数据来源广：分布于全球范围内多个站点

数据规模大：剖面数: 161；物种数: 1291

该数据集是目前全球最大规模的有孔虫数据集

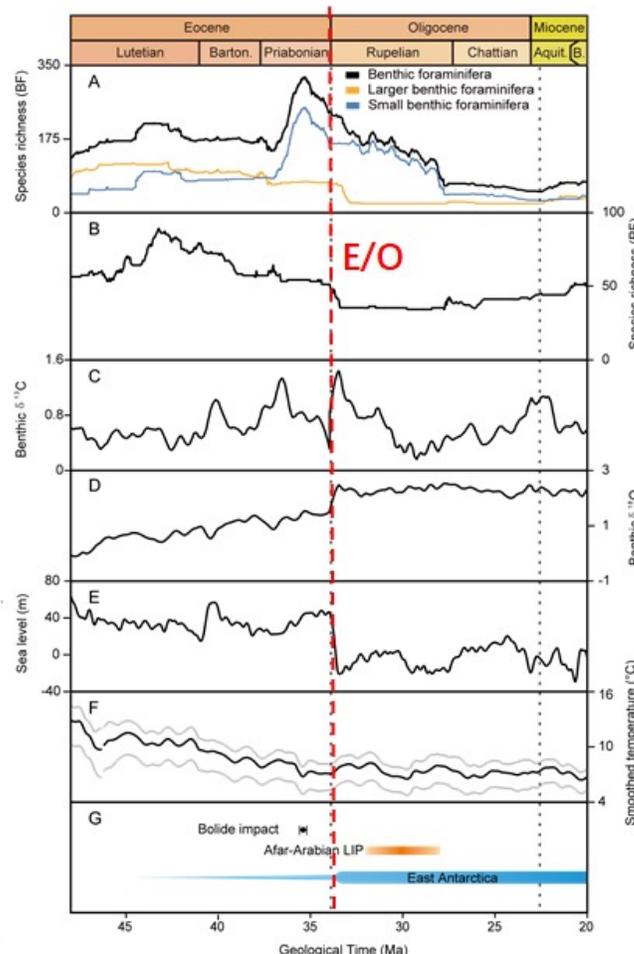
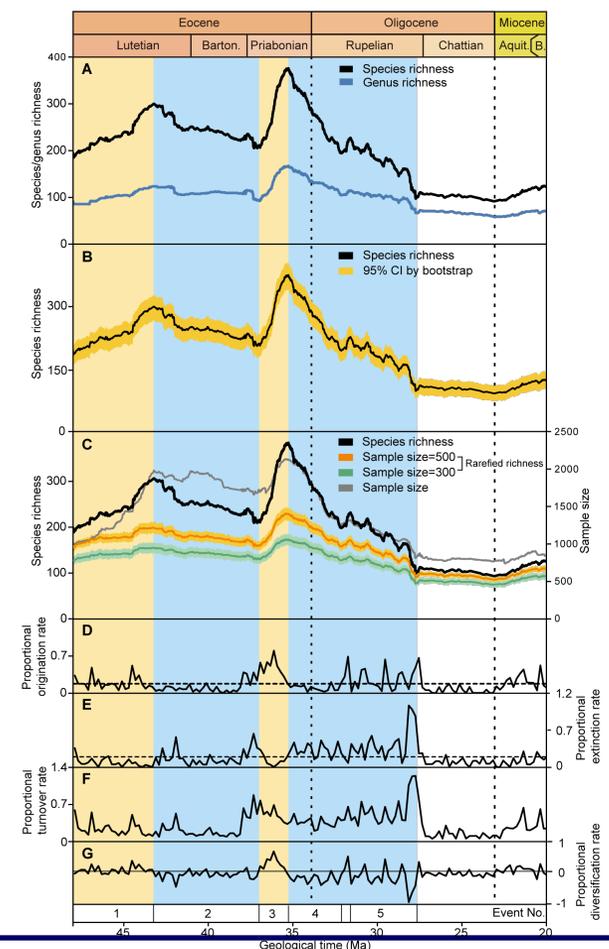


算法	目标值 范围	运行时间
以往算法	39914-41390	~40684s
新算法	39252-39778	~2376s

效率提升 17 倍，目标值更好
精准重现生命演化

交叉案例—优化任务：人工智能 + 地质学

高精度的有孔虫多样性变化曲线有助于精确分析其与气候环境变化的关系



浮游有孔虫和大型底栖有孔虫的灭绝
与快速降温、海平面下降和碳同位素
正漂移有关

揭示E-O过渡期有孔虫
生物多样性变化的驱动机制
帮助预测生物多样性危机

以共同通讯作者发表
《Nature Communications》

交叉案例—决策任务：人工智能 + 芯片设计

功能设计: 设计 RTL 并验证其功能 (文档 → RTL)

逻辑综合: 将RTL设计映射为网表信息 (RTL → 网表)

物理设计: 基于网表设计GDS物理版图 (网表 → GDS)

芯片制造: 通过光刻技术从GDS版图制造芯片 (GDS → 芯片产品)

开发

逻辑综合

物理设计

制造

封装测试

芯片设计文档

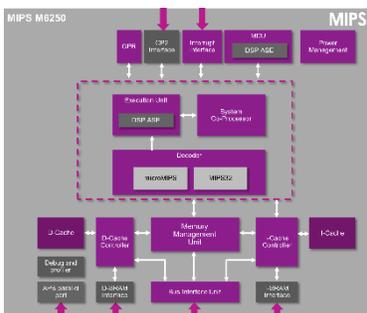
RTL逻辑设计

网表

GDS物理版图

晶圆

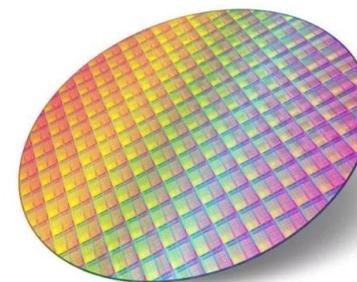
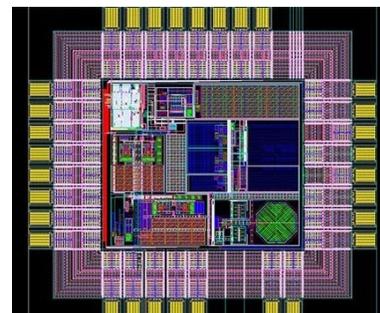
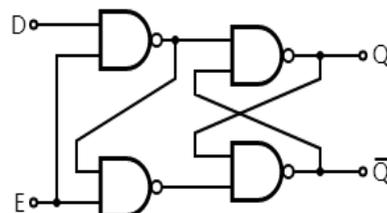
芯片产品



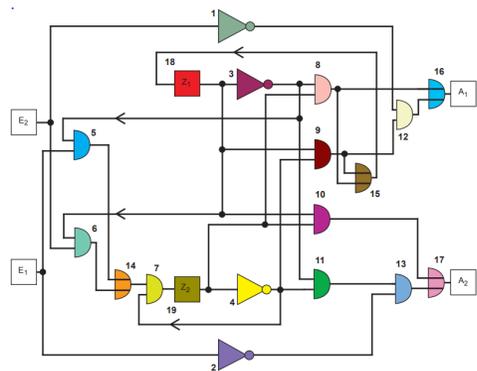
```

module alu32(Result, ALUOp, A, B, Zero) ;
output [^ALULEN:0] Result;
reg [^ALULEN:0] Result;
output Zero;
reg Zero;
input [2:0] ALUOp;
input [^ALULEN:0] A, B;

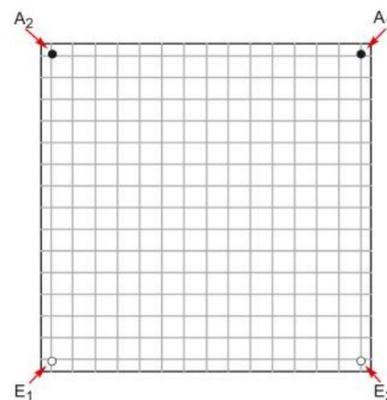
always @ (A or B or ALUOp)
begin
case (ALUOp)
3'b000: Result = A & B ;//and
3'b001: Result = A | B ;//or
// add your code here for addition, subtrac
endcase
// add your code here for Zero detect
end
endmodule
    
```



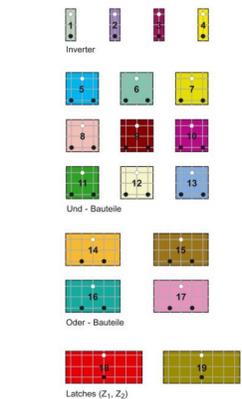
交叉案例—决策任务：人工智能 + 芯片设计



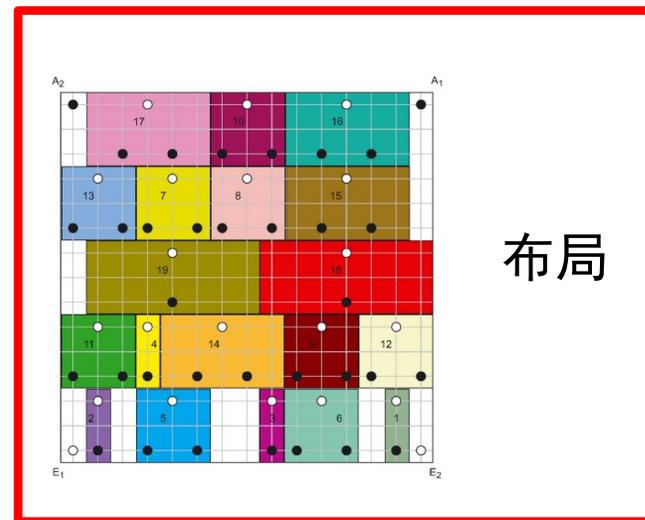
网表



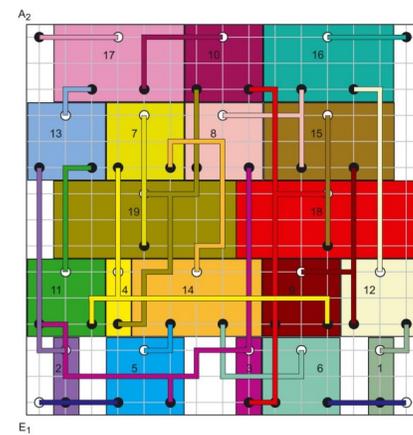
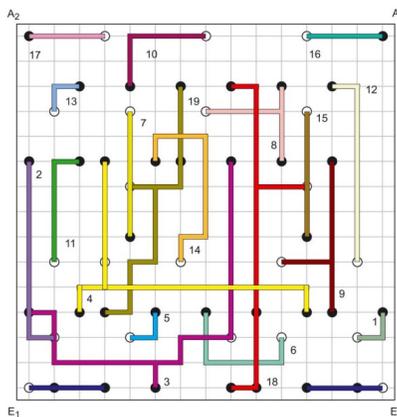
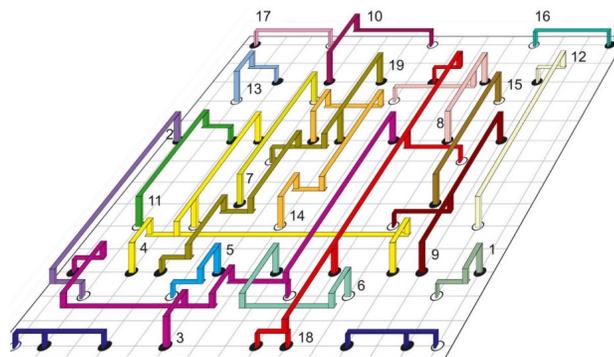
芯片版面



待放置元件



布局



布线

交叉案例—决策任务：人工智能 + 芯片设计

nature

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

[nature](#) > [articles](#) > article

Article | Published: 09 June 2021

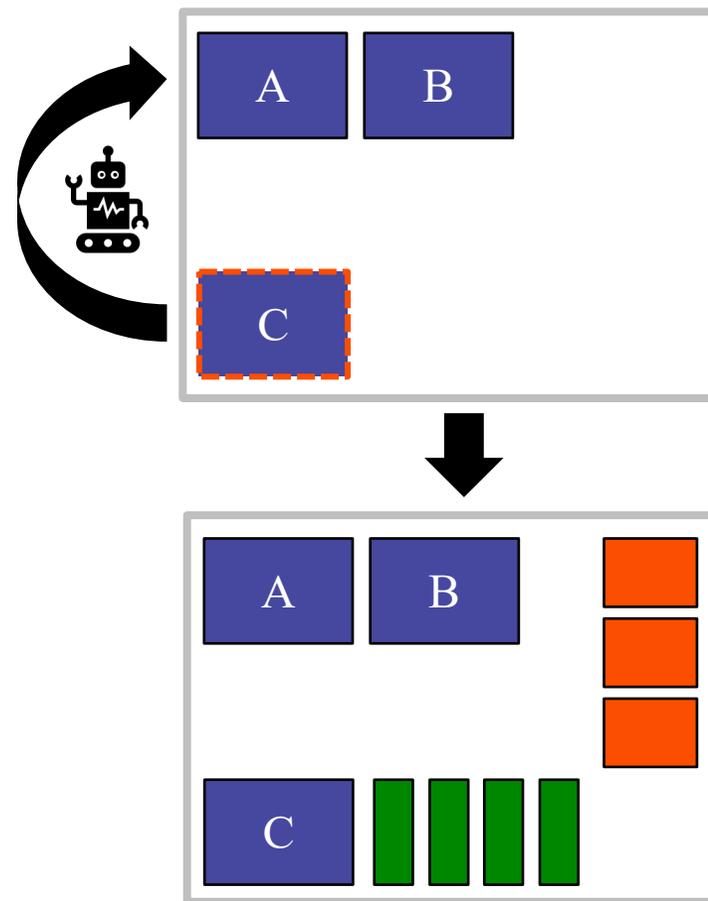
A graph placement methodology for fast chip design

[Azalia Mirhoseini](#) , [Anna Goldie](#) , [Mustafa Yazgan](#), [Joe Wenjie Jiang](#), [Ebrahim Songhori](#), [Shen Wang](#), [Young-Joon Lee](#), [Eric Johnson](#), [Omkar Pathak](#), [Azade Nova](#), [Jiwoo Pak](#), [Andy Tong](#), [Kavya Srinivasa](#), [William Hang](#), [Emre Tuncer](#), [Quoc V. Le](#), [James Laudon](#), [Richard Ho](#), [Roger Carpenter](#) & [Jeff Dean](#)

[Nature](#) **594**, 207–212 (2021) | [Cite this article](#)

[[Google](#), [Nature'21](#)]

Agent每次决定当前模块的放置位置，直到放置完整整个版面



交叉案例—决策任务：人工智能 + 芯片设计

性能指标比较：线长

Table 1: **HPWL values ($\times 10^5$)** obtained by ten compared methods on seven chips. Each result consists of the mean and standard deviation of five runs. The best (smallest) mean value on each chip is bolded. The symbols ‘+’, ‘-’ and ‘ \approx ’ indicate the number of chips where the result is significantly superior to, inferior to, and almost equivalent to WireMask-EA, respectively, according to the Wilcoxon rank-sum test with significance level 0.05.

Method	Type	adaptec1	adaptec2	adaptec3	adaptec4	bigblue1	bigblue3	bigblue4 ($\times 10^7$)	+ / - / \approx	Avg. Rank
SP-SA [30]	Packing	18.84 \pm 4.62	117.36 \pm 8.73	115.48 \pm 7.56	120.03 \pm 4.25	5.12 \pm 1.43	164.70 \pm 19.55	25.49 \pm 2.73	0/7/0	6.86
NTUPlace3 [10]	Analytical	26.62	321.17	328.44	462.93	22.85	455.53	48.38	0/7/0	9.00
RePlace [11]	Analytical	16.19 \pm 2.10	153.26 \pm 29.01	111.21 \pm 11.69	37.64 \pm 1.05	2.45 \pm 0.06	119.84 \pm 34.43	11.80 \pm 0.73	1/6/0	5.28
DREAMPlace [25]	Analytical	15.81 \pm 1.64	140.79 \pm 26.73	121.94 \pm 25.05	37.41 \pm 0.87	2.44 \pm 0.06	107.19 \pm 29.91	12.29 \pm 1.64	1/6/0	4.86
Graph [29]	RL	30.10 \pm 2.98	351.71 \pm 38.20	358.18 \pm 13.95	151.42 \pm 9.72	10.58 \pm 1.29	357.48 \pm 47.83	53.35 \pm 4.06	0/7/0	9.00
DeepPR [13]	RL	19.91 \pm 2.13	203.51 \pm 6.27	347.16 \pm 4.32	311.86 \pm 56.74	23.33 \pm 3.65	430.48 \pm 12.18	68.30 \pm 4.44	0/7/0	8.86
MaskPlace [23]	RL	6.38 \pm 0.35	73.75 \pm 6.35	84.44 \pm 3.60	79.21 \pm 0.65	2.39 \pm 0.05	91.11 \pm 7.83	11.07 \pm 0.90	0/7/0	4.28
WireMask-RS	Ours	6.13 \pm 0.05	59.28 \pm 1.48	60.60 \pm 0.45	62.06 \pm 0.22	2.19 \pm 0.01	62.58 \pm 2.07	8.20 \pm 0.17	0/5/2	2.57
WireMask-BO	Ours	6.07 \pm 0.14	59.17 \pm 3.94	61.00 \pm 2.08	63.86 \pm 1.01	2.14 \pm 0.03	67.48 \pm 6.49	8.62 \pm 0.18	0/3/4	2.86
WireMask-EA	Ours	5.91 \pm 0.07	52.63 \pm 2.23	57.75 \pm 1.16	58.79 \pm 1.02	2.12 \pm 0.01	59.87 \pm 3.40	8.28 \pm 0.25	-	1.43

[Google, Nature 2021]

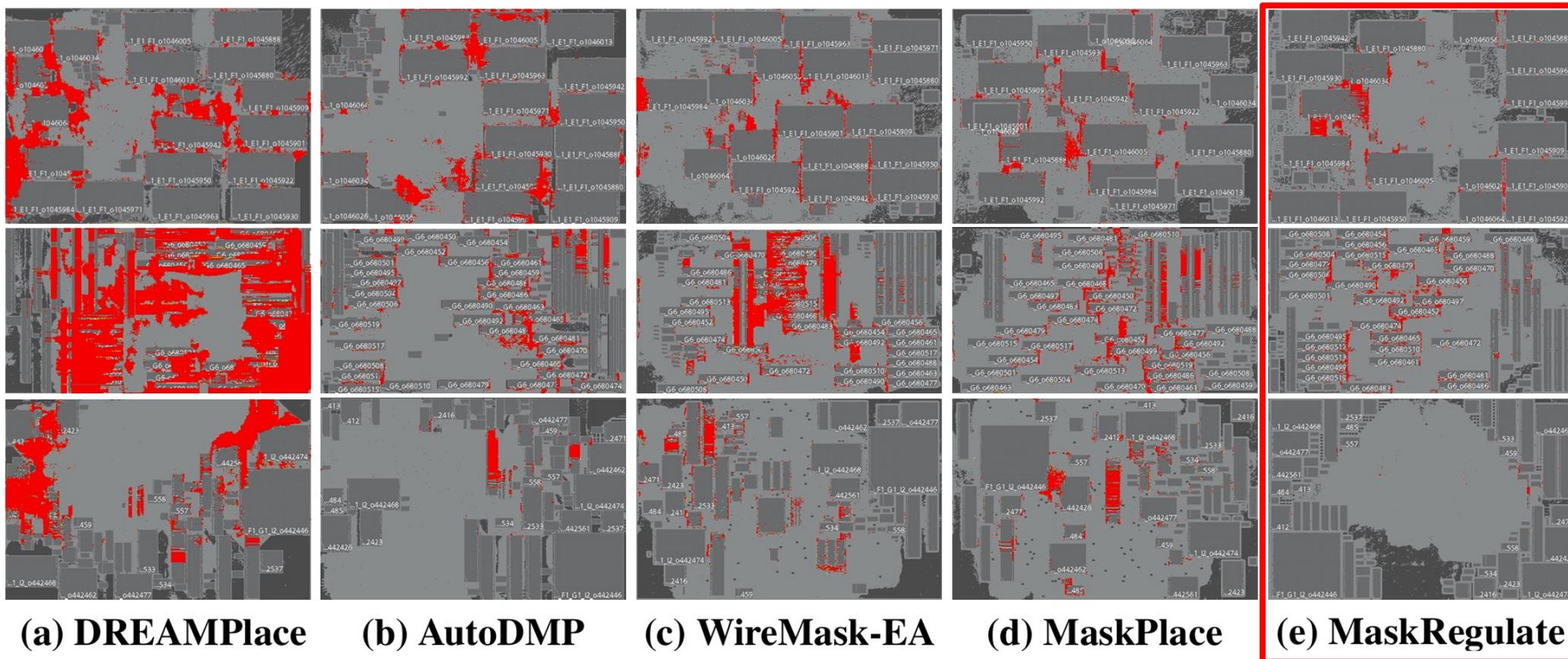
我们方法

在测试的 7 个芯片上，至少 6 个显著好；较 [Google, Nature'21]，线长缩短80%

交叉案例—决策任务：人工智能 + 芯片设计

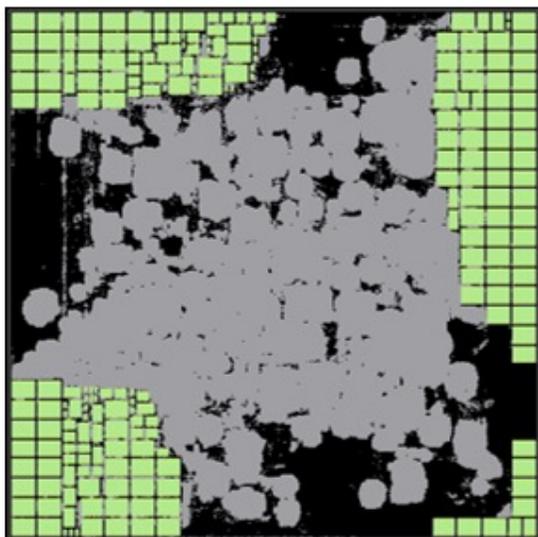
性能指标比较： 拥塞

我们方法

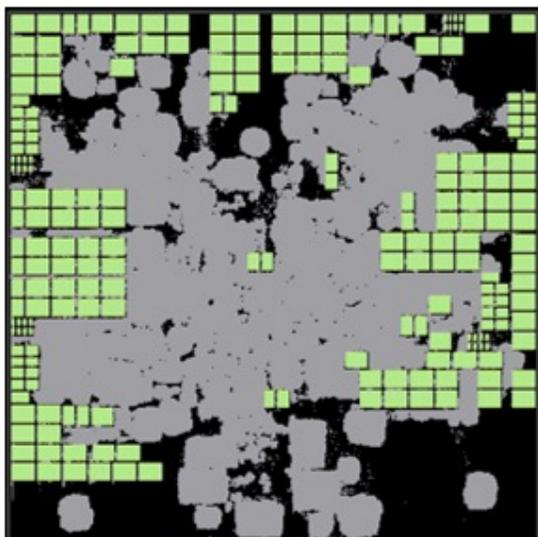


交叉案例—决策任务：人工智能 + 芯片设计

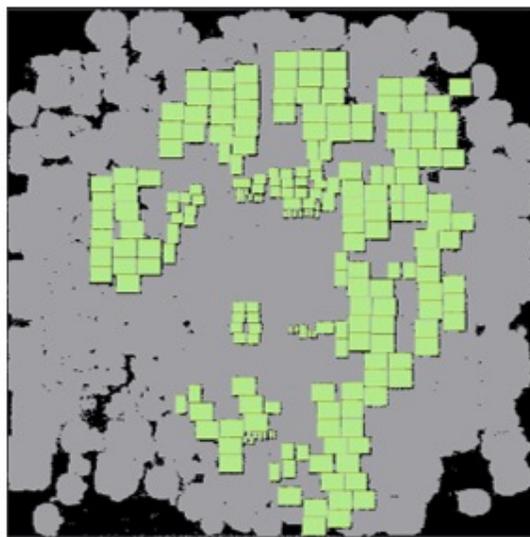
性能指标比较：规整度



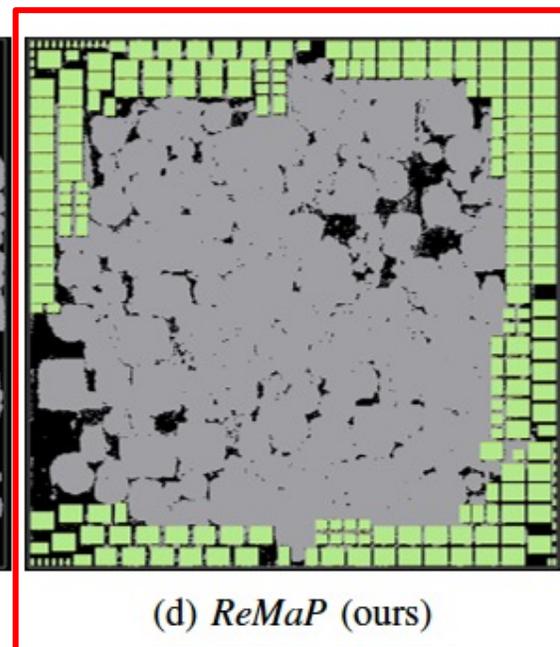
(a) *RTL-MP*



(b) *Hier-RTLMP*



(c) *DREAMPlace*



(d) *ReMaP (ours)*

我们方法

交叉案例—决策任务：人工智能 + 芯片设计

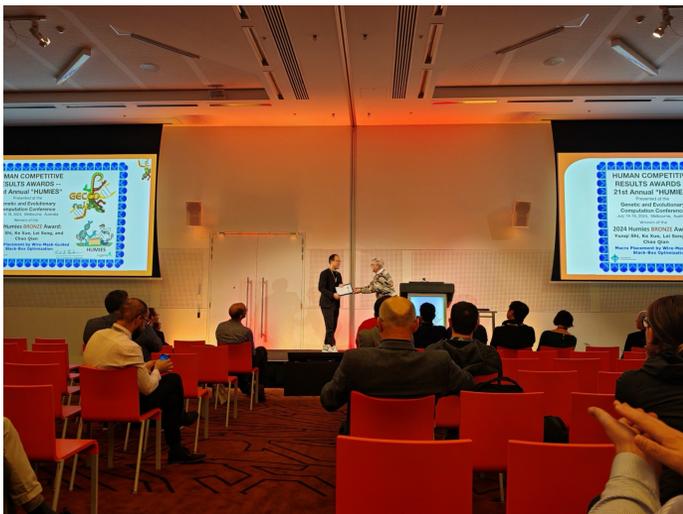
性能指标比较：时序

Benchmark	DREAMPlace* [20]		DREAMPlace 4.0* [18]		Differentiable-TDP† [12]		Distribution-TDP§ [19]		Ours	
	TNS	WNS	TNS	WNS	TNS	WNS	TNS	WNS	TNS	WNS
superblue1	-262.44	-18.87	-85.03	-14.10	-74.85	-10.77	-42.10	-9.26	-17.44	-7.75
superblue3	-76.64	-27.65	-54.74	-16.43	-39.43	-12.37	-26.59	-12.19	-20.40	-11.82
superblue4	-290.88	-22.04	-144.38	-12.78	-82.92	-8.49	-123.28	-8.86	-82.88	-9.17
superblue5	-157.82	-48.92	-95.78	-26.76	-108.08	-25.21	-70.35	-31.64	-62.18	-24.65
superblue7	-141.55	-19.75	-63.86	-15.22	-46.43	-15.22	-95.89	-17.24	-43.52	-15.22
superblue10	-731.94	-26.10	-768.75	-31.88	-558.05	-21.97	-691.10	-25.86	-558.14	-23.08
superblue16	-453.57	-17.71	-124.18	-12.11	-87.03	-10.85	-55.99	-12.21	-22.90	-8.63
superblue18	-96.76	-20.29	-47.25	-11.87	-19.31	-7.99	-19.23	-5.25	-16.16	-6.92
Average Ratio	6.90	2.07	2.75	1.40	2.00	1.09	1.68	1.11	1.00	1.00

总时序 TNS
平均提升 40.5%

最差时序 WNS
平均提升 8.3%

交叉案例—决策任务：人工智能 + 芯片设计



HUMAN COMPETITIVE RESULTS AWARDS -- 21st Annual "HUMIES"

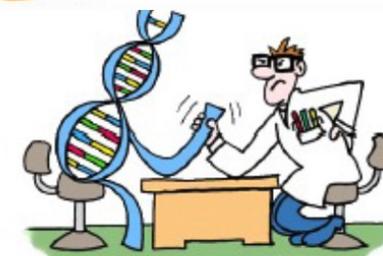
Presented at the
Genetic and Evolutionary Computation Conference
July 14-18, 2024, Melbourne, Australia

Winners of the
2024 Humies BRONZE Award:
Yunqi Shi, Ke Xue, Lei Song, and Chao Qian

Macro Placement by Wire-Mask-Guided Black-Box Optimization

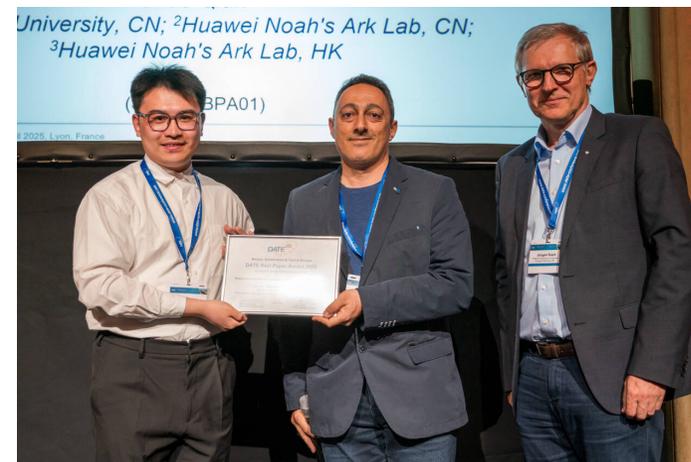
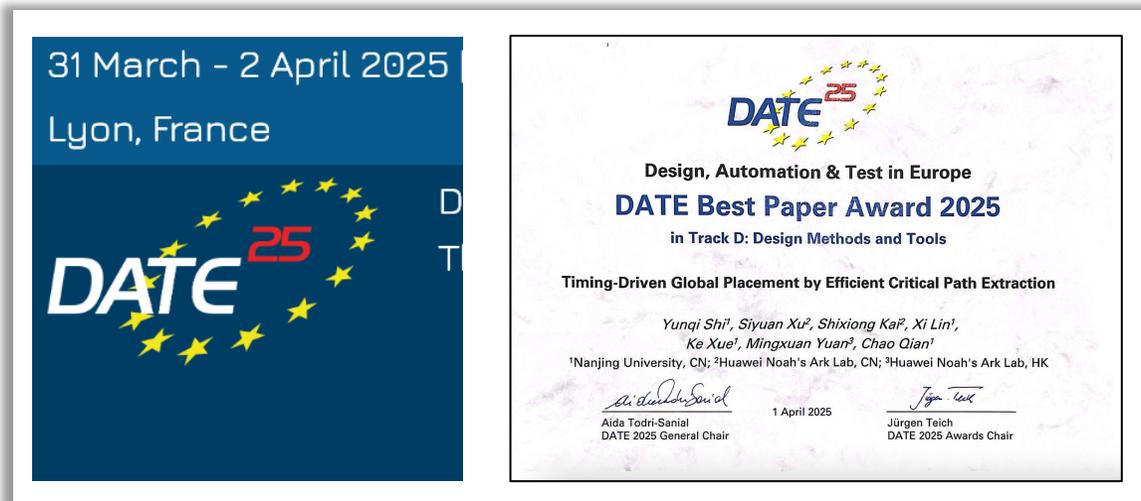


Erik D. Goodman
Erik D. Goodman, Organizer



HUMIES
sigevo

交叉案例—决策任务：人工智能 + 芯片设计



获芯片设计自动化领域顶级国际会议 DATE'25
最佳论文奖

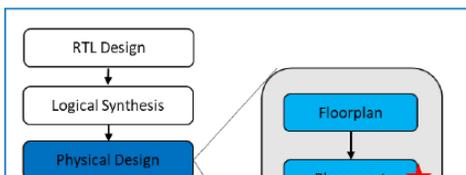
交叉案例—决策任务：人工智能 + 芯片设计

难题4：布局布线优化技术

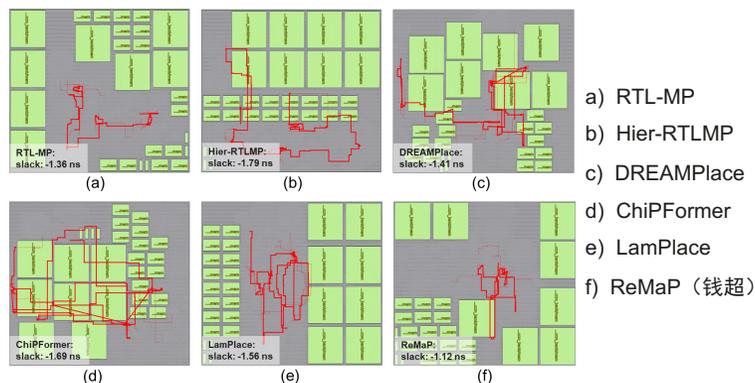
出题组织：海思/诺亚方舟实验室 接口专家：许思源 xusiyuan520@huawei

技术背景

芯片布局起着承上（逻辑综合）启下（布线）的作用，是现代超大规模集成电路设计物理设计流程中的一步，显著影响芯片设计的迭代周期。由于芯片布局被认为是NP-Complete的问题，因此如何快速生成高质量的芯片布局是布局问题的重要挑战。



在 6 个样例上对比SOTA方案平均提升 PPA 性能约 18.9%，平均提升拥塞指标约 46.3%，预测相关性（99%）



已集成进
华为海思
芯片设计工具

火花奖

难题	院校/部门	姓名
布局布线优化技术	南京大学人工智能学院	钱超教授
超长序列视频生成推理加速关键技术	浙江大学计算机科学与技术学院	李玺教授
HPC应用骨架函数自动提取方法	浙江大学软件学院	常志豪研究员

受邀参加华为 CEO
任正非先生的座谈会

交叉案例—决策任务：人工智能 + 芯片设计

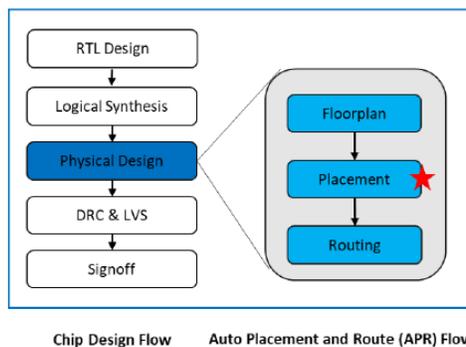
和华为合作攻克2D布局难题

难题4：布局布线优化技术

出题组织：海思/诺亚方舟实验室 接口专家：许思源 xusiyuan520@huawei

技术背景

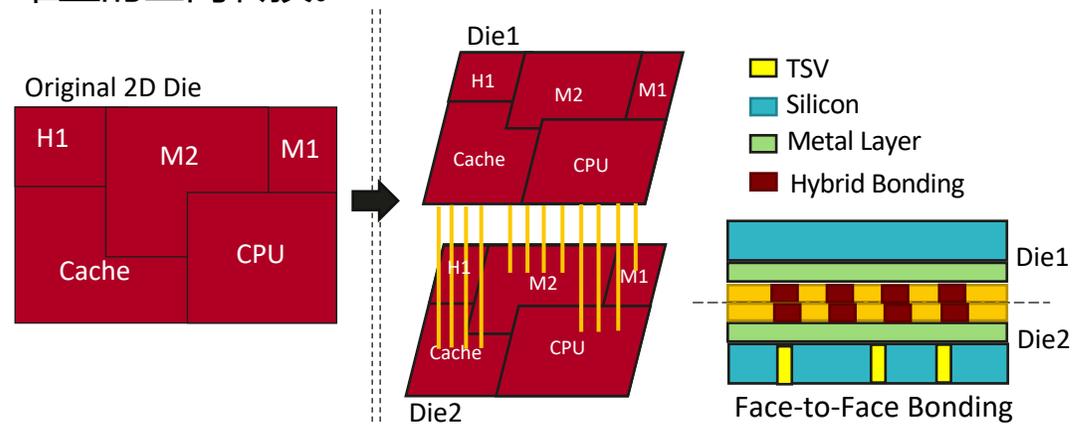
芯片布局起着承上（逻辑综合）启下（布线）的作用，是现代超大规模集成电路设计物理设计流程中的一步，显著影响芯片设计的迭代周期。由于芯片布局被认为是NP-Complete的问题，因此如何快速生成高质量的芯片布局是布局问题的重要挑战。



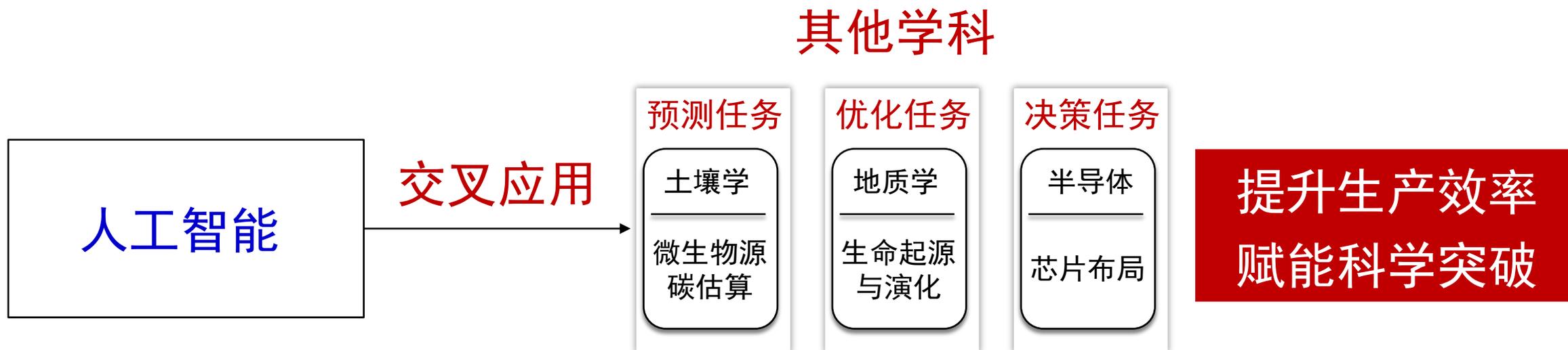
同时合作探索更为前沿的3D堆叠布局

芯片3D堆叠布局技术

芯片3D集成是一种新的设计范式，通过硅芯片或通过嵌入式芯片的多层互连，提供了传统2D平面IC实现到3D堆叠的空间转换。



先进芯片设计 缓解 当前先进制造工艺的局限



谢谢大家!