

ChatGPT 时代的知识图谱

东南大学 KGCODE Lab

2023年4月

注：本讲内容基于柯文俊老师的总结

主要内容

1. ChatGPT简介

2. ChatGPT关键技术

3. ChatGPT展望

4. 团队研究

ChatGPT: Optimizing Language Models for Dialogue

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. ChatGPT is a sibling model to InstructGPT, which is trained to follow an instruction in a prompt and provide a detailed response.

Link: <https://chat.openai.com/chat>

November 30, 2022
13 minute read



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



学习深度学习需要掌握数学、计算机科学和机器学习的基础知识。以下是一些学习深度学习的步骤和建议：



1. 学习线性代数和微积分。深度学习中的很多概念都基于这些数学知识，因此，深入学习线性代数和微积分对于理解深度学习至关重要。
2. 掌握编程语言。Python是深度学习中最常用的编程语言之一，因此，建议学习Python语言，并掌握其基本语法和编程概念。
3. 学习机器学习的基础知识。在深度学习之前，应该先掌握机器学习的基础知识，包括常见的算法和技术，例如线性回归、决策树和支持向量机。
4. 学习深度学习的基本概念和架构。深度学习的基础知识包括神经网络、反向传播算法、卷积神经网络和循环神经网络等。
5. 学习深度学习的实践应用。为了更好地理解深度学习，可以尝试完成一些实践项目，例如图像分类、语音识别和自然语言处理等。
6. 阅读最新的研究论文和文献。深度学习是一个快速发展的领域，每天都有新的研究成果。阅读最新的研究论文和文献有助于了解最新的技术和进展。
7. 参加在线课程和学习社区。有许多在线课程和学习社区，例如Coursera、Udacity和Github等，可以帮助学生学习深度学习。
8. 建立自己的深度学习项目。为了更好地理解深度学习，可以尝试自己设计和实现深度学习项目，并应用它们到真实场景中。

总之，深度学习是一个需要耐心和持续学习的领域。只有不断学习和实践，才能不断提高自己的深度学习水平。

PI

请写一个函数把以下数字排序3, 8, 9, 17, -3, 21



以下是一个Python函数，可以将给定的数字排序：



python

Copy code

```
def sort_numbers(numbers):  
    return sorted(numbers)
```

可以将3, 8, 9, 17, -3和21作为参数传递给该函数：

python

Copy code

```
numbers = [3, 8, 9, 17, -3, 21]  
sorted_numbers = sort_numbers(numbers)  
print(sorted_numbers)
```

输出结果为：`[-3, 3, 8, 9, 17, 21]`。

该函数使用了Python内置的`sorted`函数，该函数将列表中的元素按升序排列，并返回一个新的排好序的列表。

PI

请把这段话翻译成中文We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. ChatGPT is a sibling model to InstructGPT, which is trained to follow an instruction in a prompt and provide a detailed response.



我们训练了一个名为ChatGPT的模型，它可以以对话方式进行交互。对话的格式使得ChatGPT能够回答后续问题、更正错误、质疑不正确的前提和拒绝不恰当的请求。ChatGPT是InstructGPT的类似模型，它是根据提示中的指令进行训练，并提供详细的响应。



PI

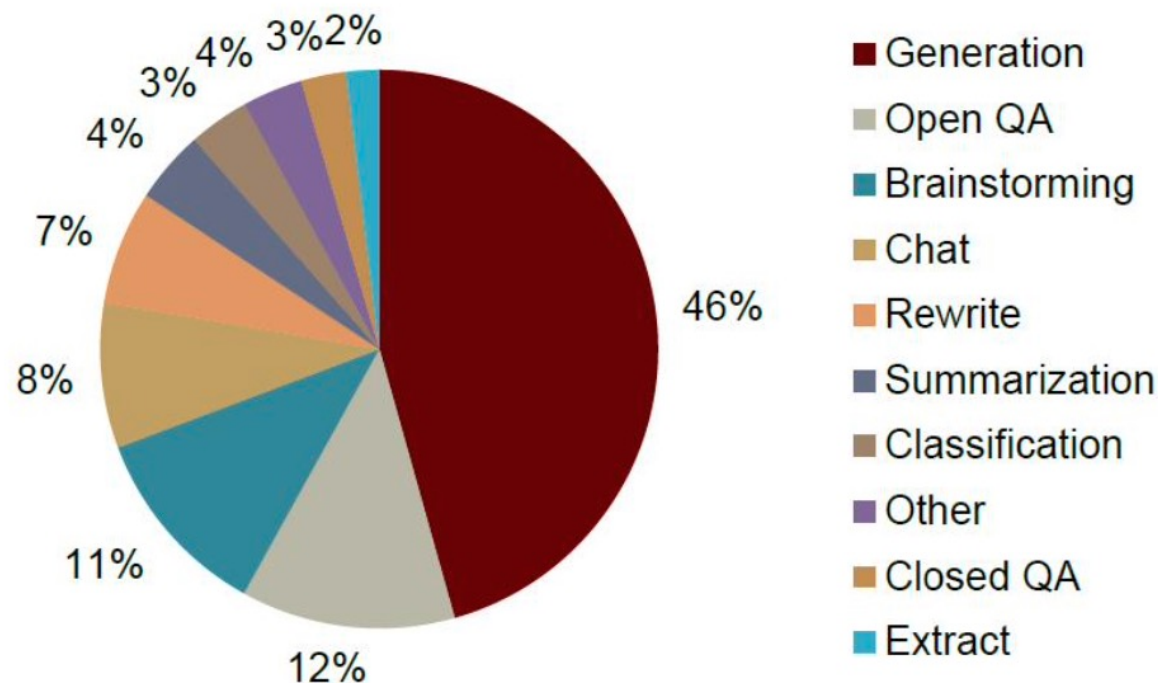
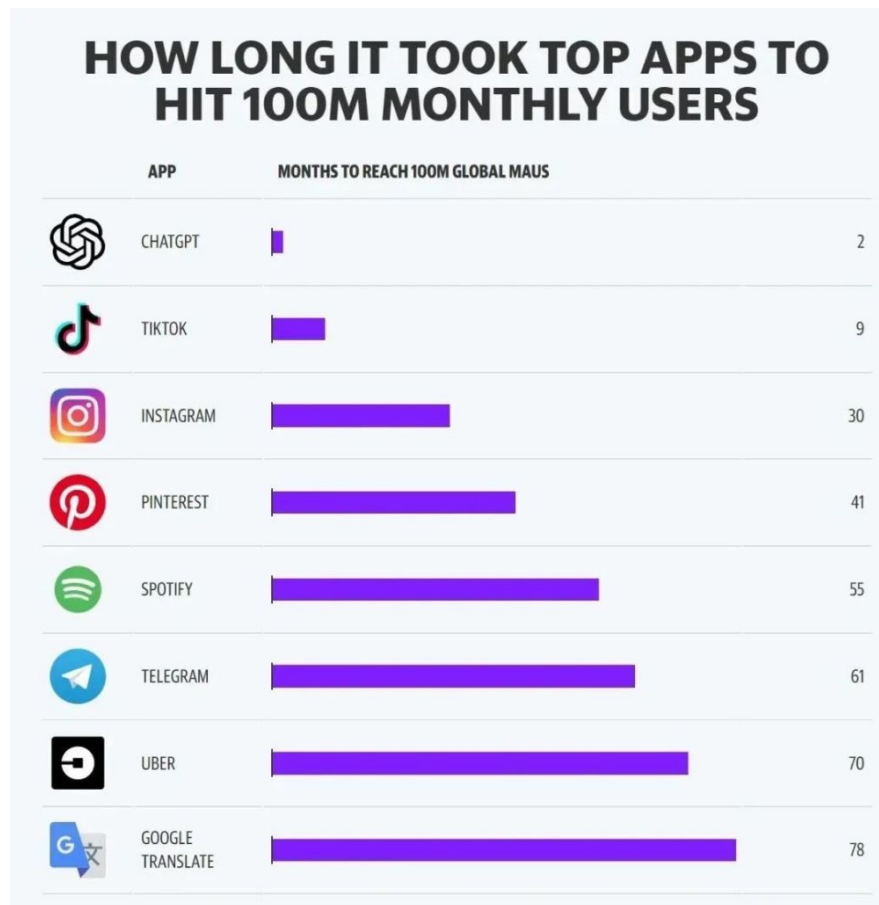
"We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer follow up questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. ChatGPT is a sibling models to InstructGPT, which is trained to follow an instruction in a prompt and provide a detailed response"这句话有语法错误吗



是的，这句话中有一个语法错误。"a sibling models"应该改为"a sibling model"，因为"sibling"是单数形式。所以正确的句子应该是："ChatGPT is a sibling model to InstructGPT, which is trained to follow an instruction in a prompt and provide a detailed response."

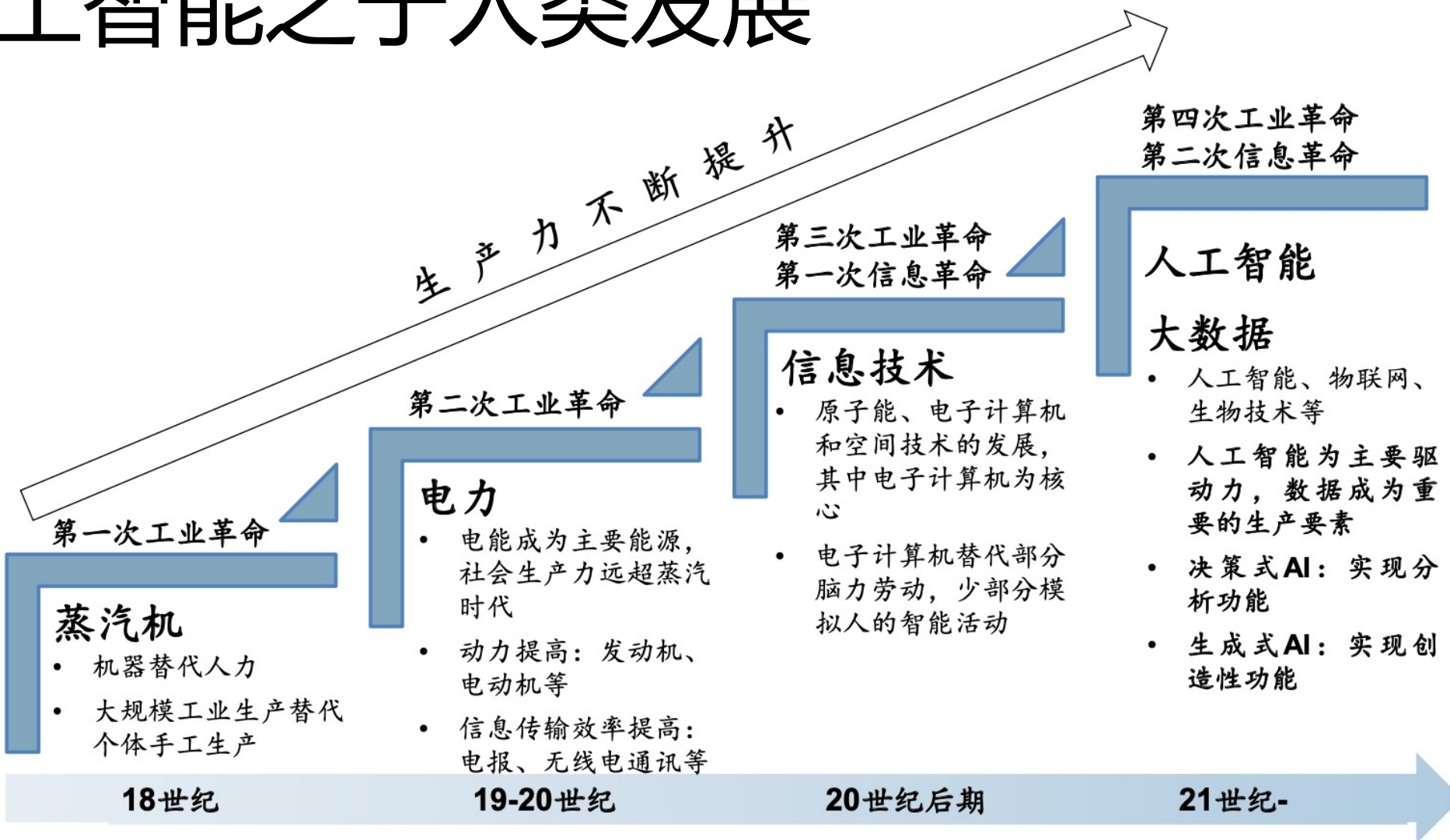


ChatGPT现象



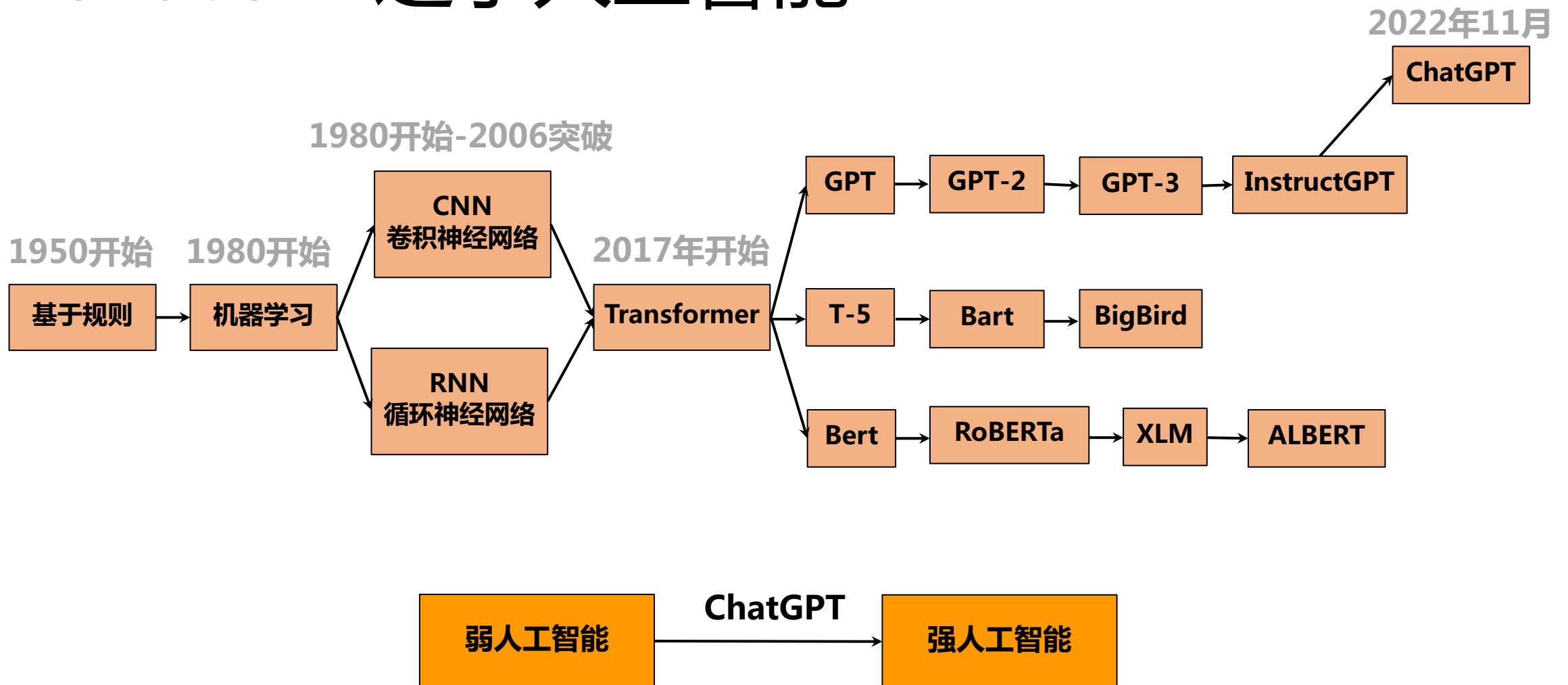
来源：Training language models to follow instructions with human feedback , OpenAI

人工智能之于人类发展



来源：世界银行，联合国教科文组织，科学的历程

ChatGPT之于人工智能



主要内容

1. ChatGPT简介

2. ChatGPT关键技术

3. ChatGPT展望

4. ChatGPT团队研究

ChatGPT关键技术 1

- **Transformer decoder-only**
 - 自回归式生成
- **预训练语言模型**
 - GPT-3 , 175B参数
 - 训练语料3000亿单词
 - 60% 来自于 2016-2019 的 C4
 - 22% 来自于 WebText2
 - 16% 来自于 Books
 - 3% 来自于 Wikipedia

ChatGPT关键技术 2

- **代码训练**

- 带注释，代码与人类语言对齐
- 分模块、分步骤解决问题
- 代码比语言更看重长距离依赖，减弱了上下文学习能力

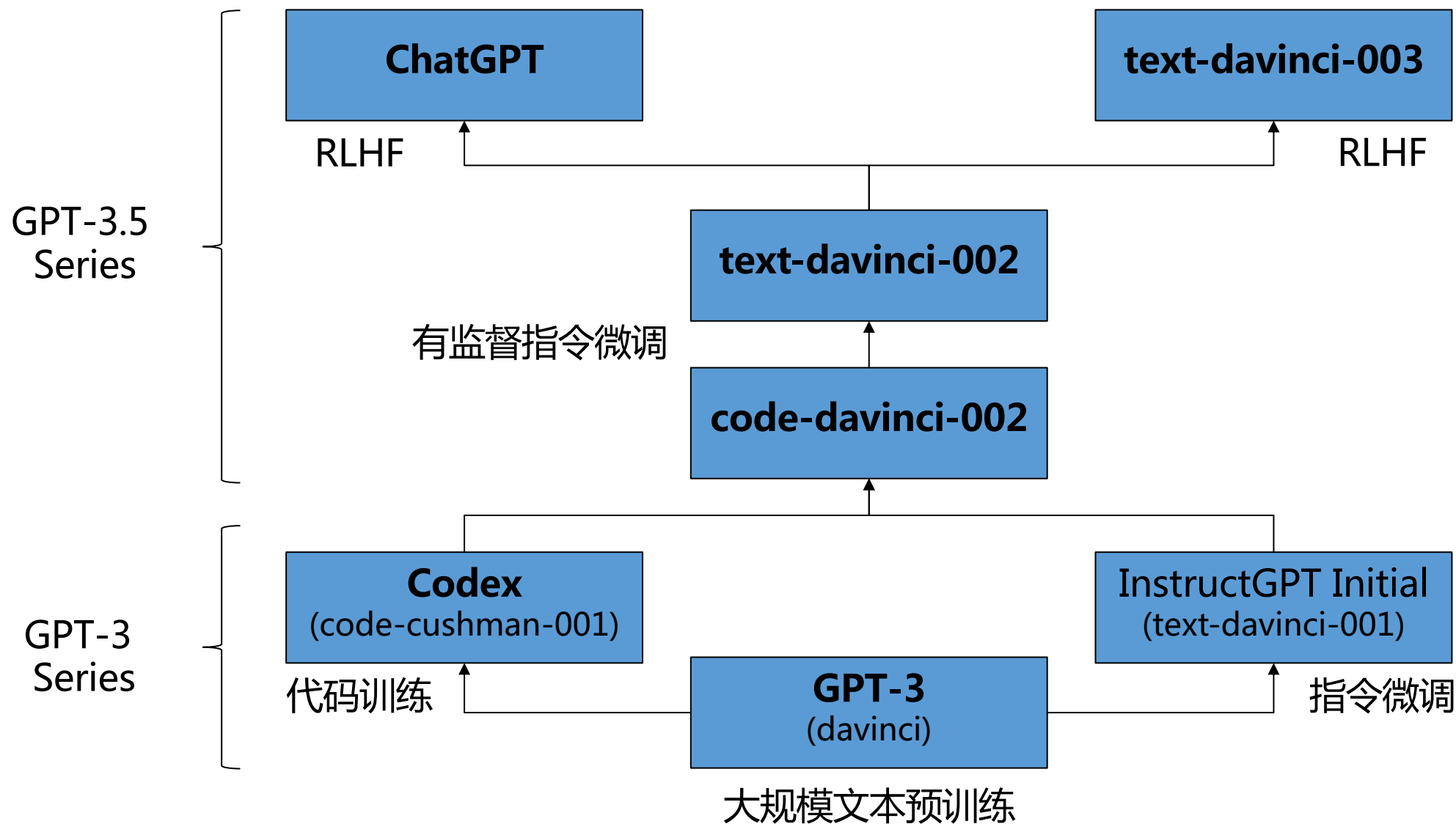
- **指令微调（ Prompt/Instruction Tuning ）**

- 将所有任务统一为指令形式
- 多任务精调

- **RLHF（ 基于人工反馈的强化学习 ）**

- 与人类对齐（ 多样性 ）

ChatGPT技术迭代

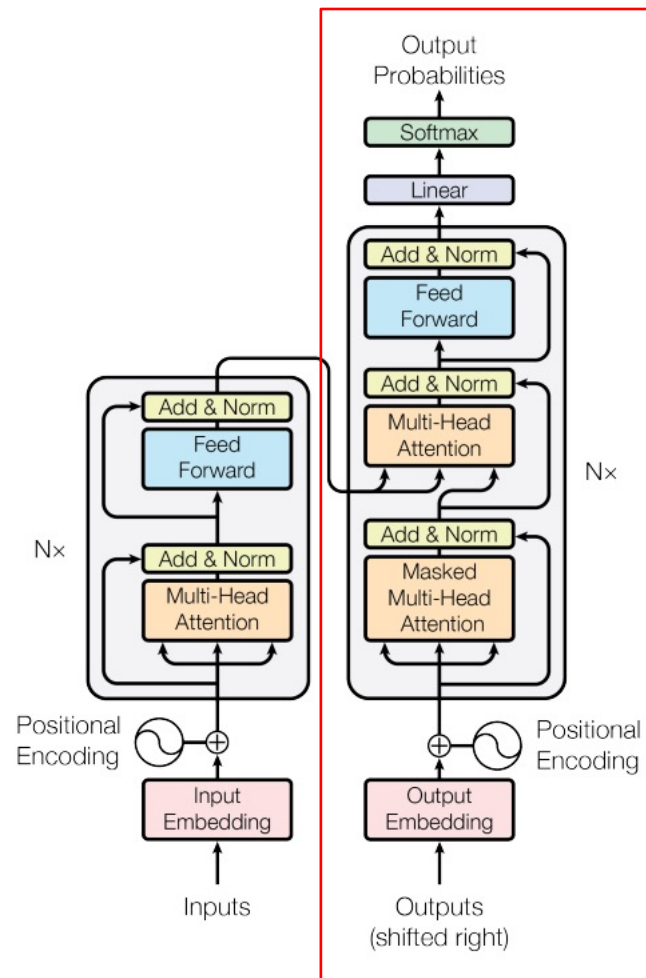


模型结构

• Decoder-only 语言模型

- GPT 系列采用纯 Transformer Decoder 的结构。是生成式模型。
- 语言模型：对于给定词序列，计算该序列是一句话的概率。可以转化为计算下一个词的概率。

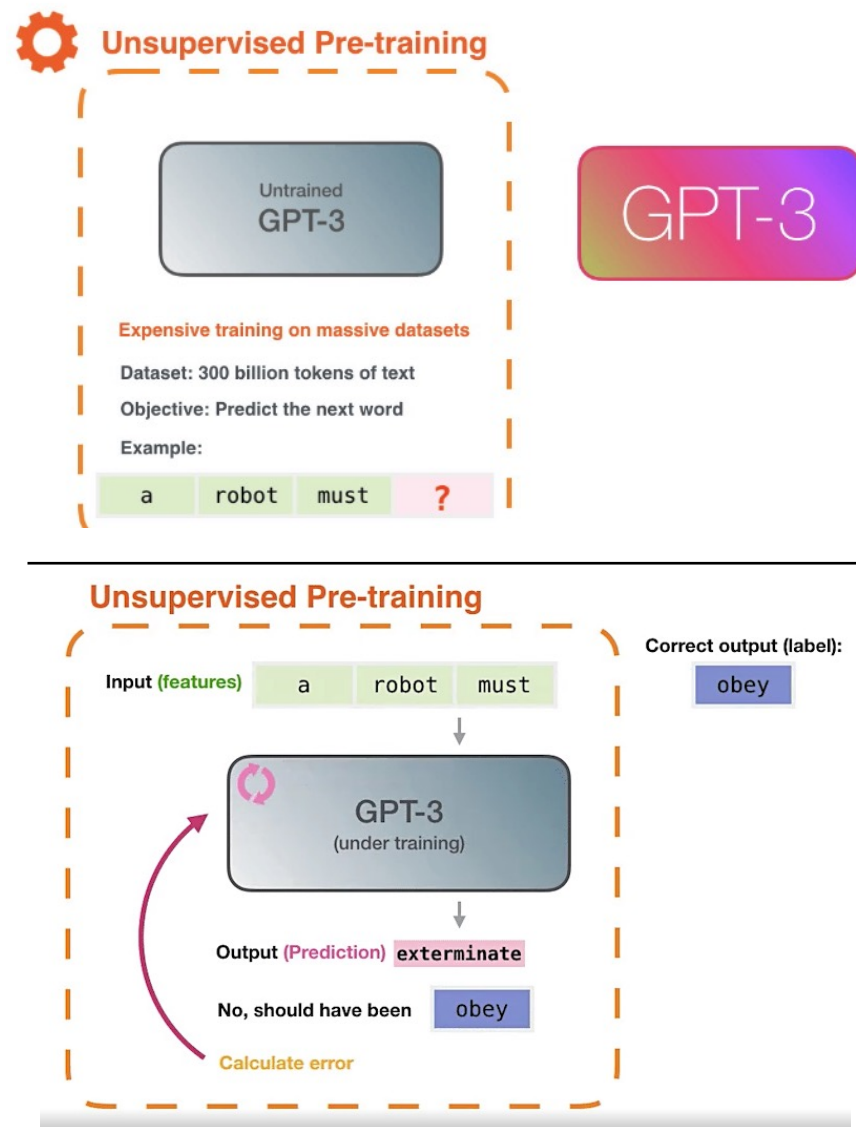
$$p(w_1, w_2, \dots, w_n) = p(w_1) \prod_{i=2}^n \underline{p(w_i | w_1, \dots, w_{i-1})}$$



语言模型预训练

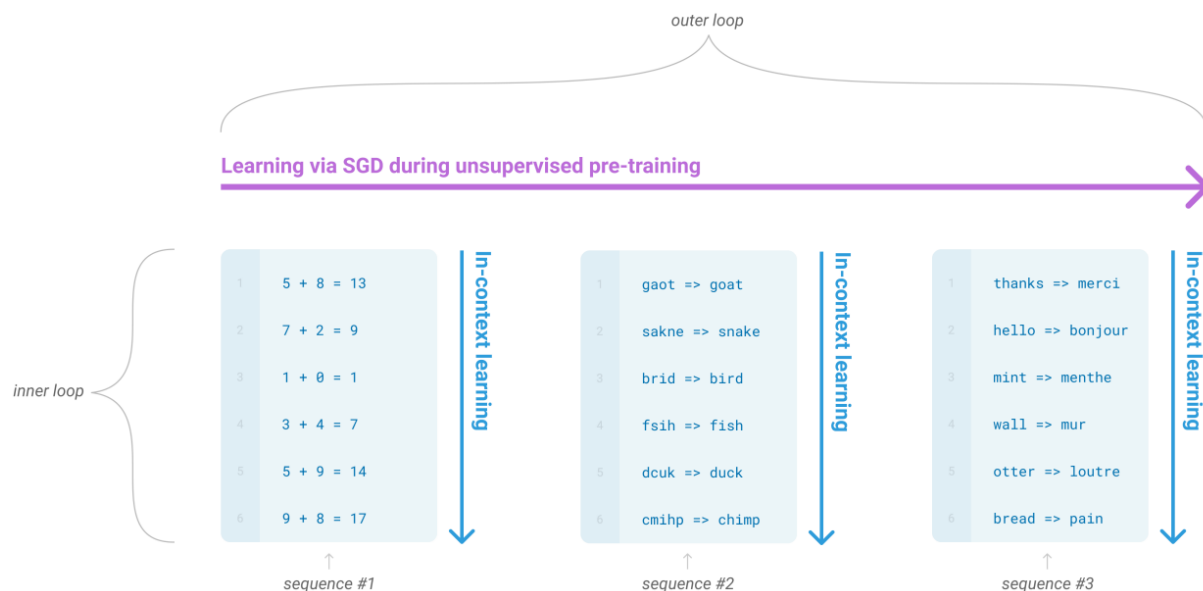
- **GPT-3 预训练**

- 无监督学习
- Language Modeling: 从前述词预测下一个词。如左图中通过 “a robot must” 来预测下一个词 “obey”。



语言模型预训练

• 元学习



- 情境学习 (in-context learning)
 - 从 pretraining-finetuning 范式走向 prompt-based learning

Zero-shot

The model predicts the answer given only a natural language description of the task. No gradient updates are performed.

```
1 Translate English to French:  ← task description
2 cheese => .....           ← prompt
```

One-shot

In addition to the task description, the model sees a single example of the task. No gradient updates are performed.

```
1 Translate English to French:  ← task description
2 sea otter => loutre de mer   ← example
3 cheese => .....             ← prompt
```

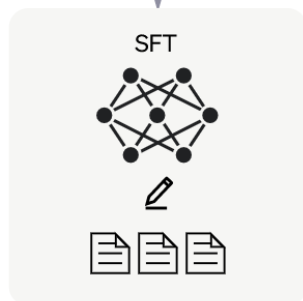
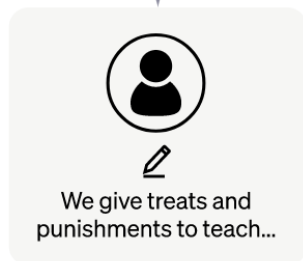
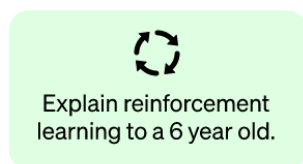
Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.

```
1 Translate English to French:  ← task description
2 sea otter => loutre de mer   ← examples
3 peppermint => menthe poivrée ←
4 plush girafe => girafe peluche ←
5 cheese => .....             ← prompt
```

第1步：指令微调

构建指令数据集，对模型进行有监督的指令微调。目的：使得模型根据指令生成的内容有大致的方向，符合人类的预期



13K prompts

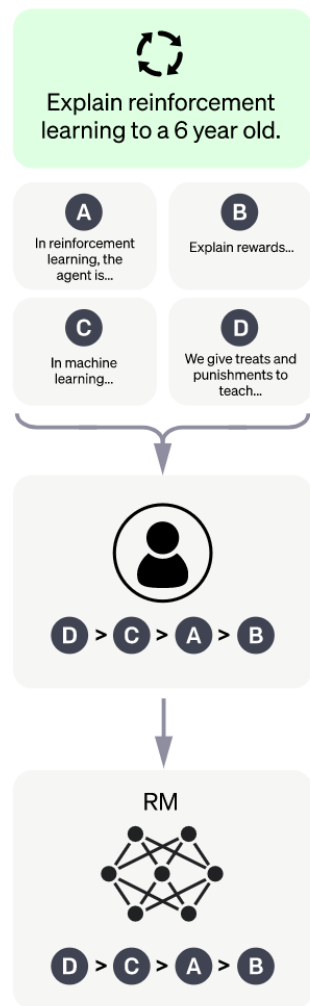
(1)从prompt数据集中采样prompt，如左图这句“向六岁的小孩解释强化学习”

(2)标注员为每条提示（指令）写回复

(3)在GPT3.5模型上，基于 <prompt,回复> 进行有监督训练

第2步：训练reward模型

基于人工反馈训练奖励模型，该奖励模型将用于第三步的强化学习



33K prompts

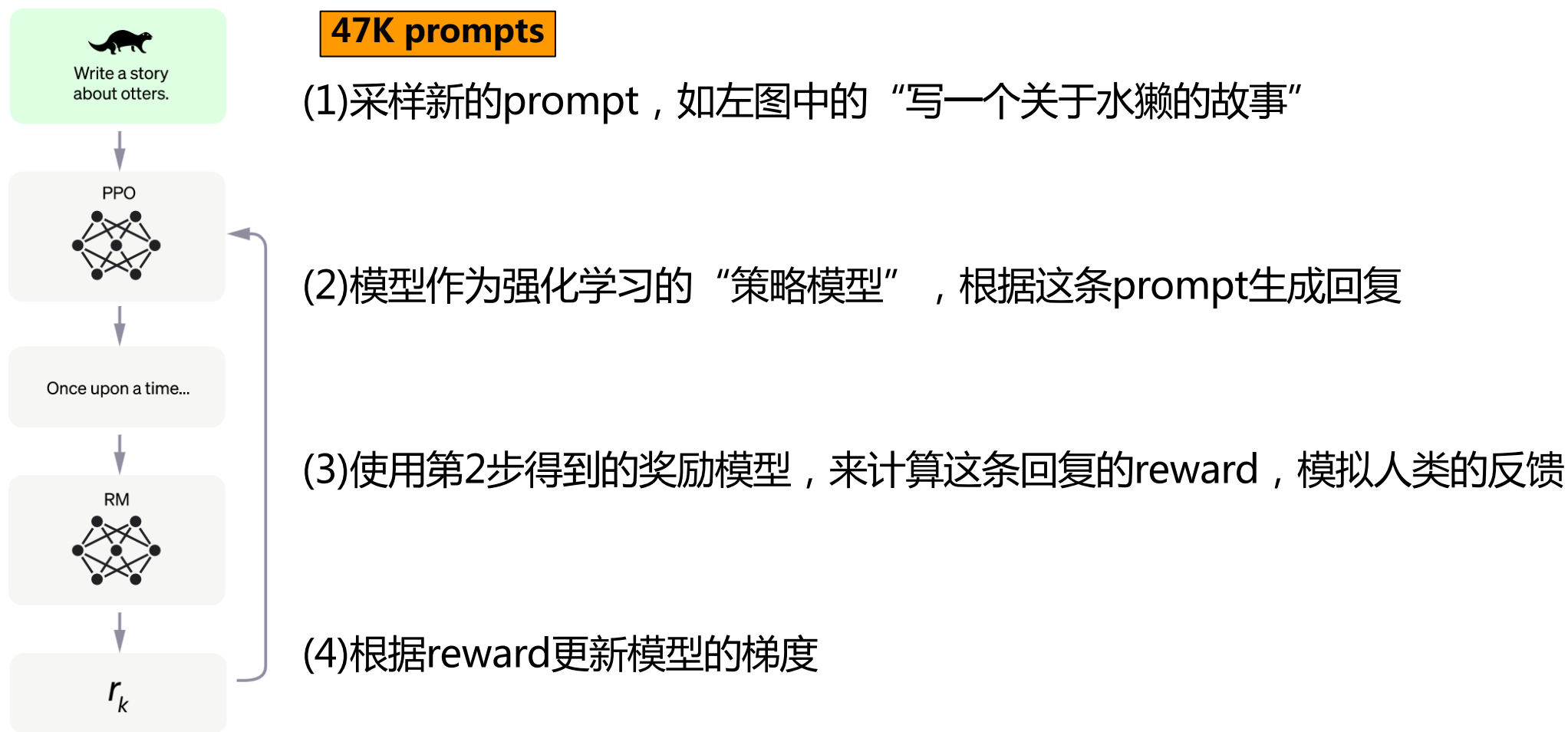
(1)从prompt数据集中采样prompt，并用第1步微调好的模型对同一个提示生成多个回复（左图的A,B,C,D）

(2)人类标注员为同一个prompt的多个回复进行排序（不采用打分的方式是为了去除主观因素的影响）

(3)依据标注员给出的排序结果，来训练奖励模型（reward model）

第3步：强化学习

采用基于PPO算法的方式进行强化学习，继续训练经第1步微调后的模型，使得模型能够接受人类的反馈来输出和人类对齐更好的内容



RLHF：基于人工反馈的强化学习

- RLHF

- 第2、3步多轮循环进行，使模型不断从人类反馈中强化学习，获得与人类更好的对齐能力

基于人工反馈的强化学习

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.



This data is used to fine-tune GPT-3.5 with supervised learning.



Step 2

Collect comparison data and train a reward model.

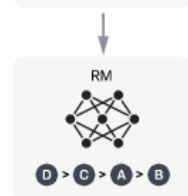
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



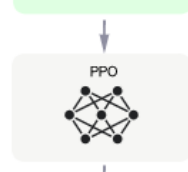
Step 3

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.

A new prompt is sampled from the dataset.



The PPO model is initialized from the supervised policy.



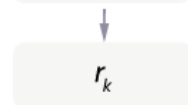
The policy generates an output.



The reward model calculates a reward for the output.

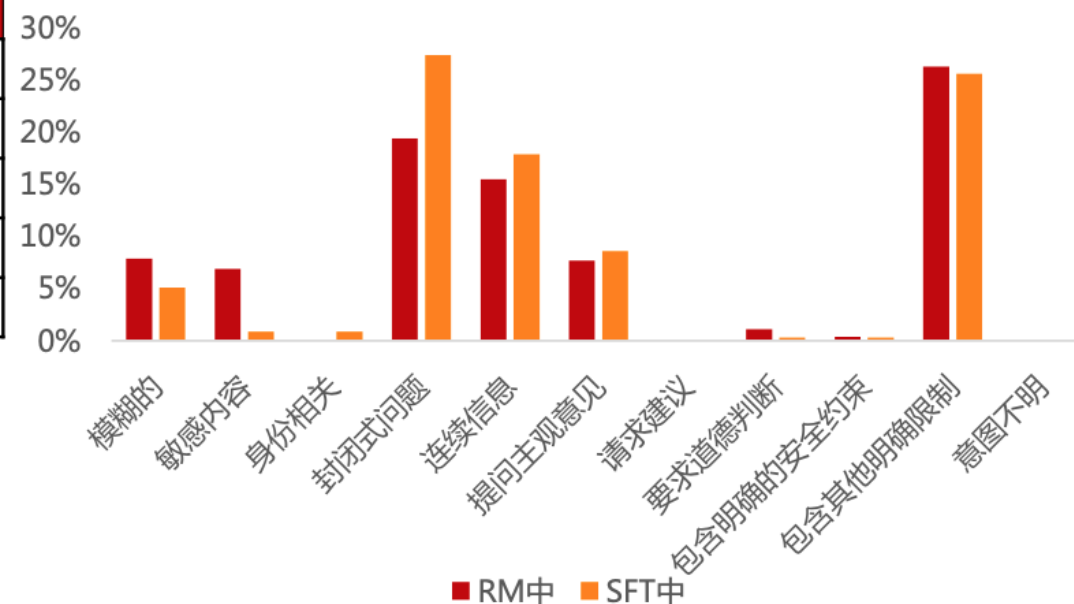


The reward is used to update the policy using PPO.



ChatGPT指令标注

性别		种族认同		国籍		年龄		受教育水平	
男性	50.00%	白人	31.60%	菲律宾	22%	18-24	26.30%	高中以下	0%
女性	44.40%	东南亚	52.60%	孟加拉	22%	25-34	47.40%	高中学历	10.50%
其他	5.60%	原住民	0.00%	美国	17%	35-44	10.50%	本科学位	52.60%
		东亚	5.30%	阿尔巴尼亚	5%	45-54	10.50%	硕士学位	36.80%
		中东	0.00%	巴西	5%	55-64	5.30%	博士学位	0%
		拉丁裔	15.80%	加拿大	5%	65+	0%		
		黑人/非洲裔	10.50%	哥伦比亚	5%				
				印度	5%				
				乌拉圭	5%				
				津巴布韦	5%				



- 指令来源

- 人工编写
- 提取用户输入的指令

数据飞轮（AI正循环）

示例：Chat GPT 的学习四阶段

1. 学习文字接龙

2. 人类老师引导文字接龙的方向

3. 模仿人类老师的喜好

4. 用增强式学习向模拟老师学习

1.学习文字接龙



“跟人类对话”

不完整的句子

“你好”



接一个可能的字



“美”

“跟”

“人”

“跟人”

“类”

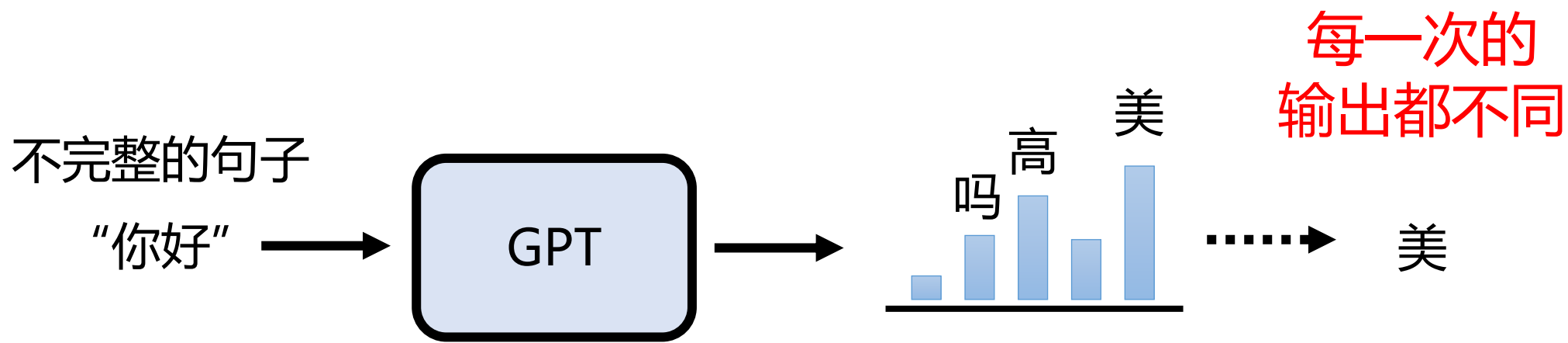
“跟人类”

“对”

GPT = Generative Pre-trained Transformer

不需要人工标注

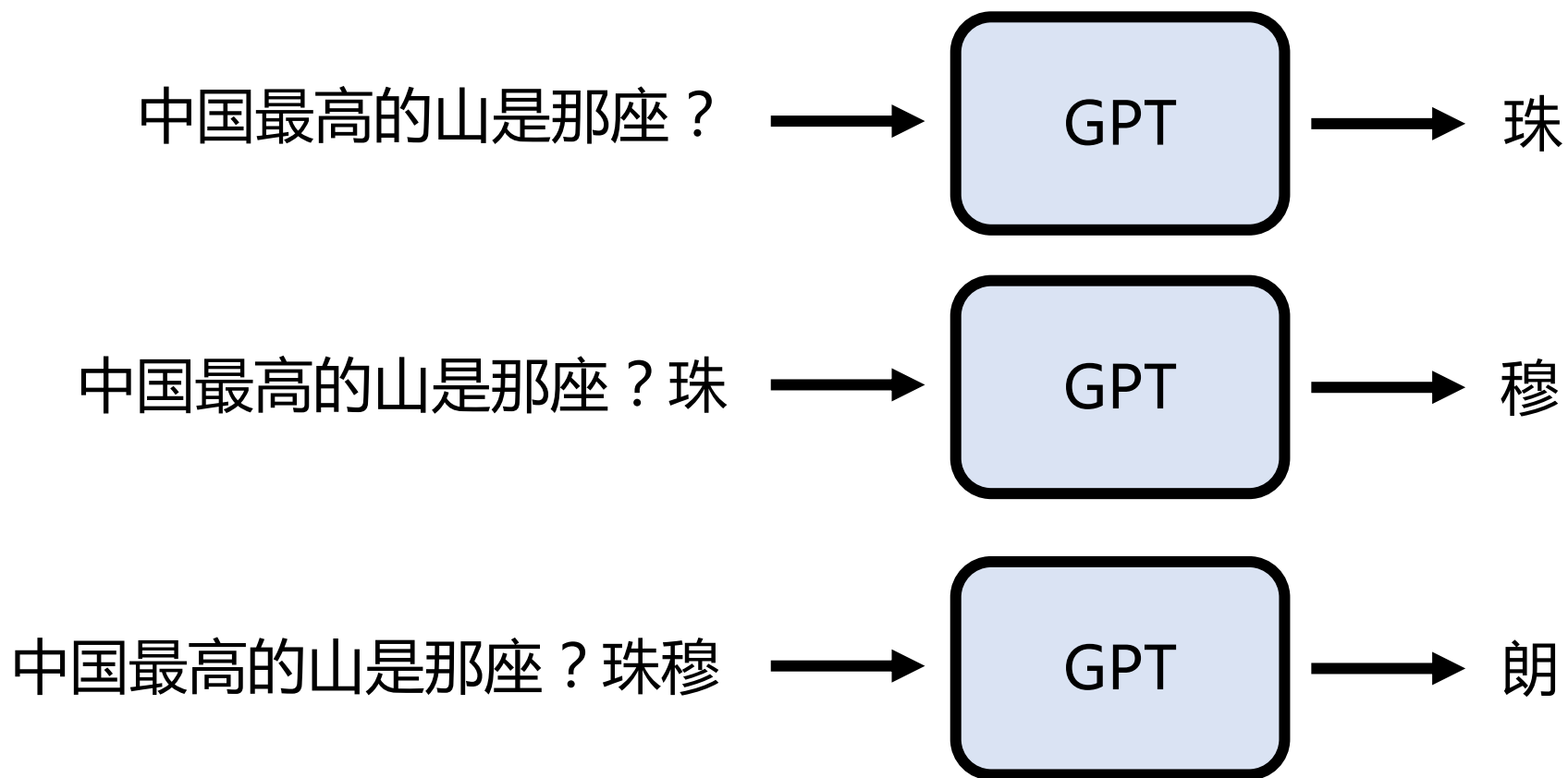
1.学习文字接龙



GPT = Generative Pre-trained Transformer

1.学习文字接龙

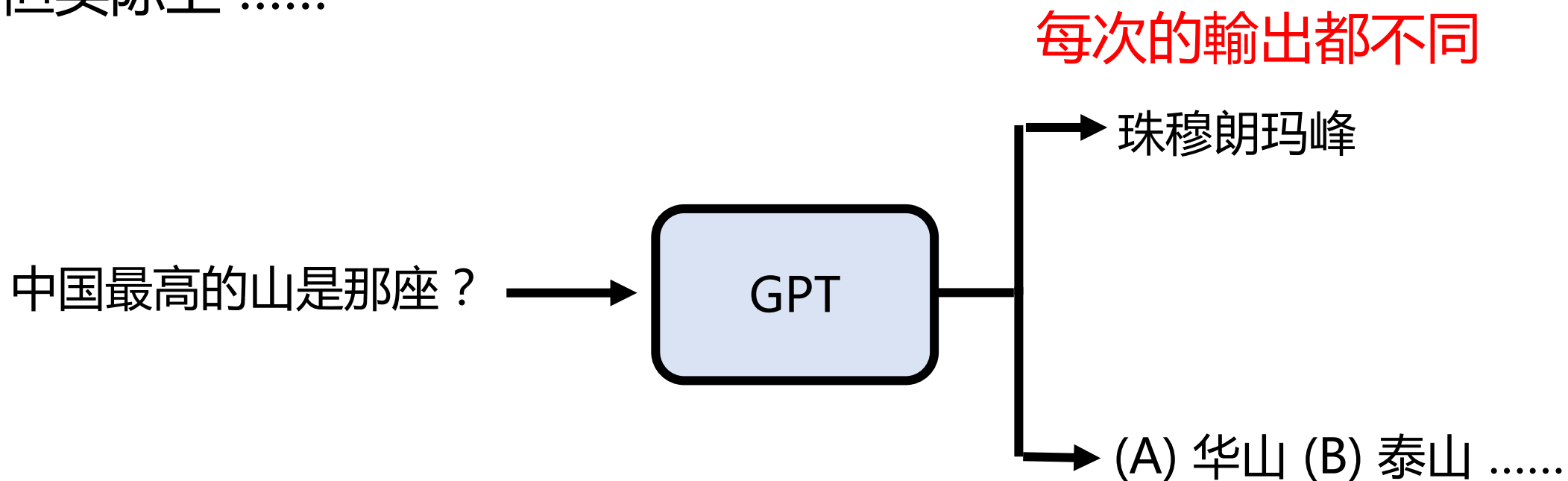
- 文字接龙有什么用？



可以回答问题!

1.学习文字接龙

- 但实际上



如何引导GPT 产生有用的输出呢？

示例：Chat GPT 的学习四阶段

1. 学习文字接龙

2. 人类老师引导文字接龙的方向

3. 模仿人类老师的喜好

4. 用增强式学习向模拟老师学习

2. 人类老师引导文字接龙的方向

- 找人来思考想问 GPT 的问题，并人工提供正确答案

中国最高的山是那座？珠穆朗玛峰

如何学习深度学习？需要先知道基本概念 ...

请把这句话进行翻译

希望GPT多看这些有益的句子，不要再去网络上看杂乱的信息

不需要穷尽所有的問題，我们只是告诉 GPT 人类的偏好

示例：Chat GPT 的学习四阶段

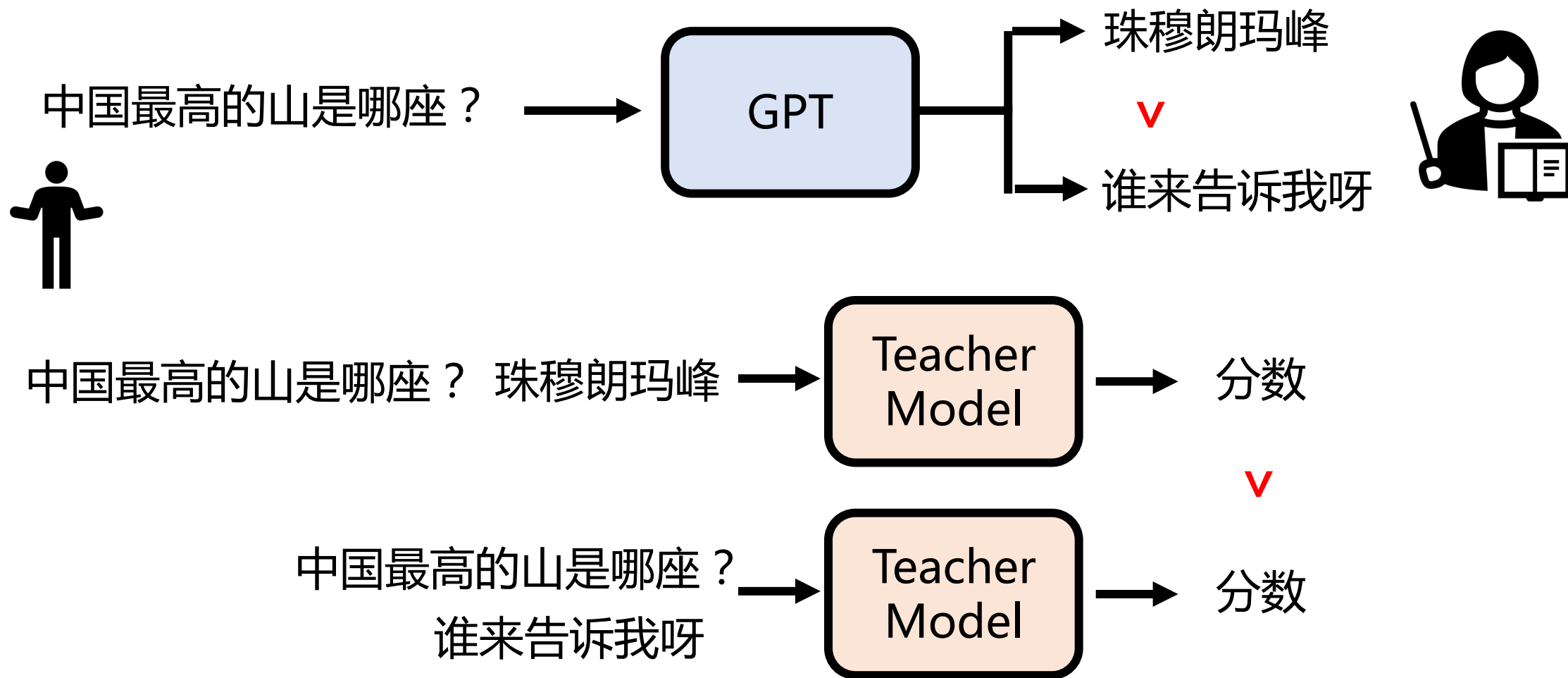
1. 学习文字接龙

2. 人类老师引导文字接龙的方向

3. 模仿人类老师的喜好

4. 用增强式学习向模拟老师学习

3. 模仿人类老师的喜好



示例：Chat GPT 的学习四阶段

1. 学习文字接龙

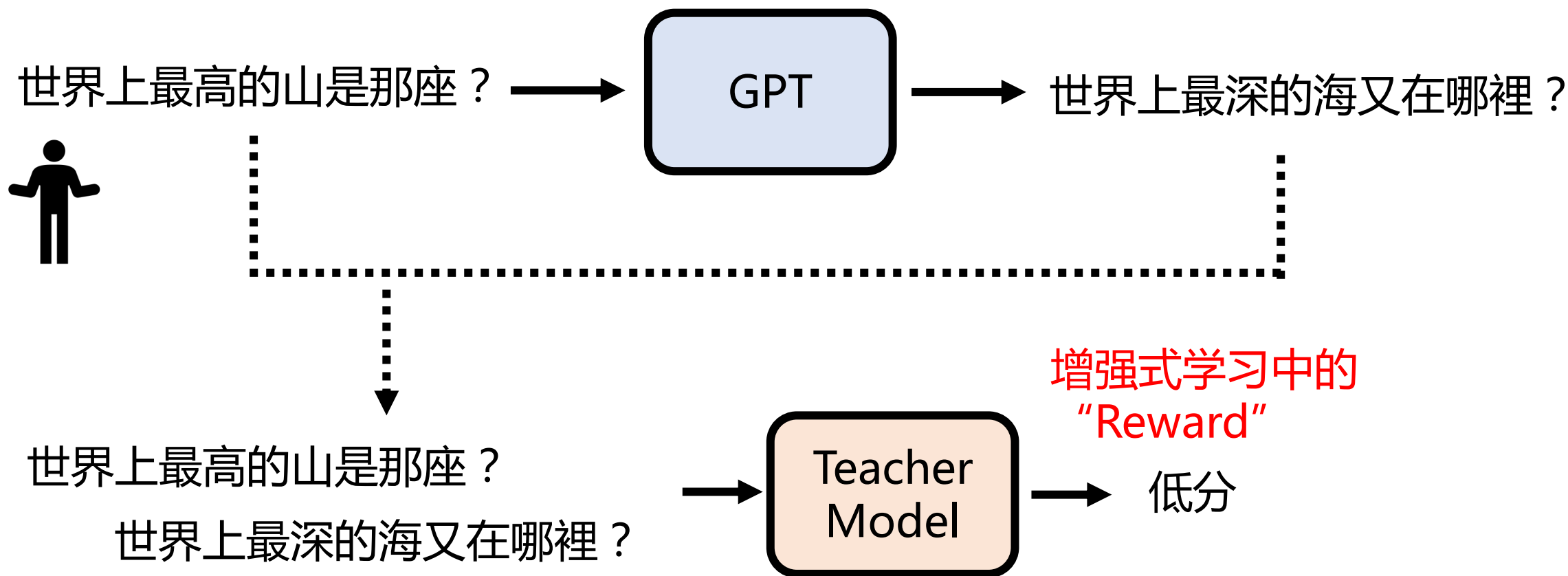
2. 人类老师引导文字接龙的方向

3. 模仿人类老师的喜好

4. 用增强式学习向模拟老师学习

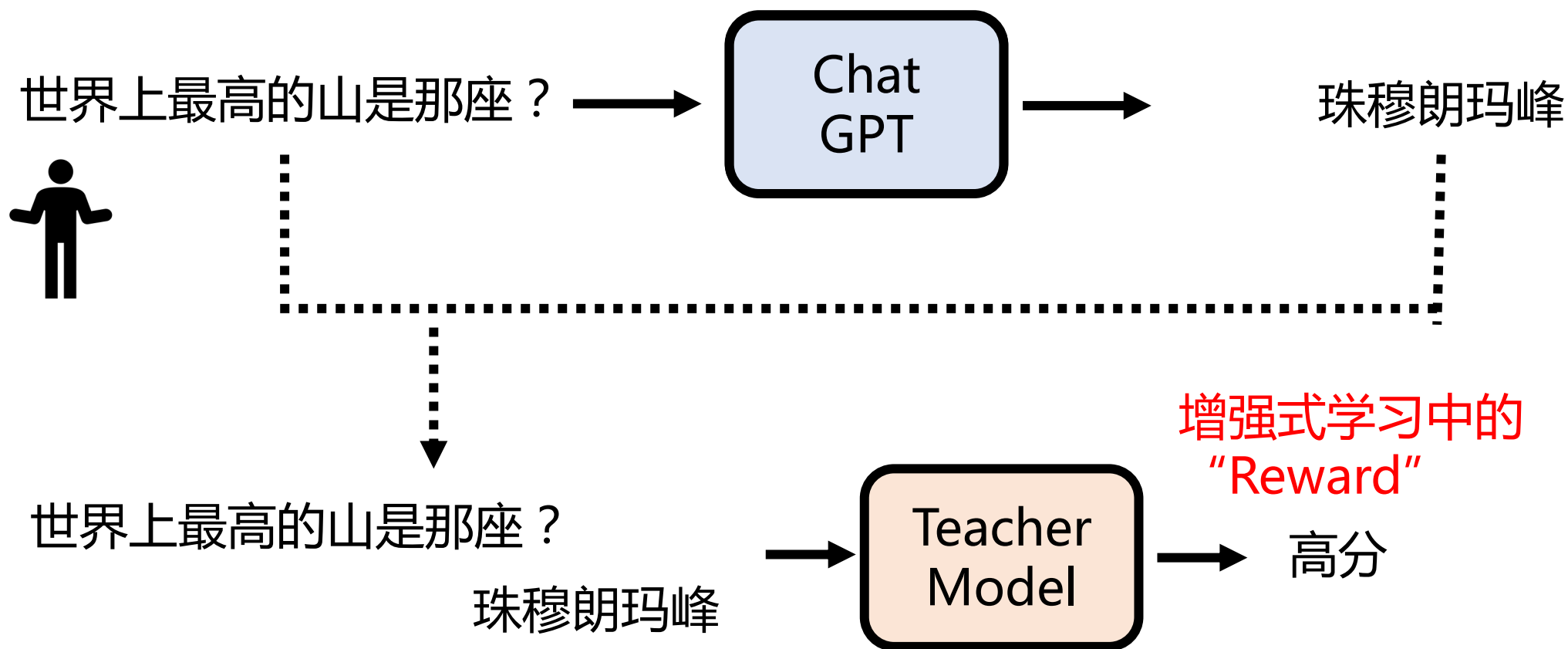
4.用增强式学习向模拟老师学习

增强式学习：调整参数，得到最大的 Reward



4.用增强式学习向模拟老师学习

增强式学习：调整参数，得到最大的 Reward



示例小节

1. 学习文字接龙

想说什么就说什么

2. 人类老师引导文字接龙的方向

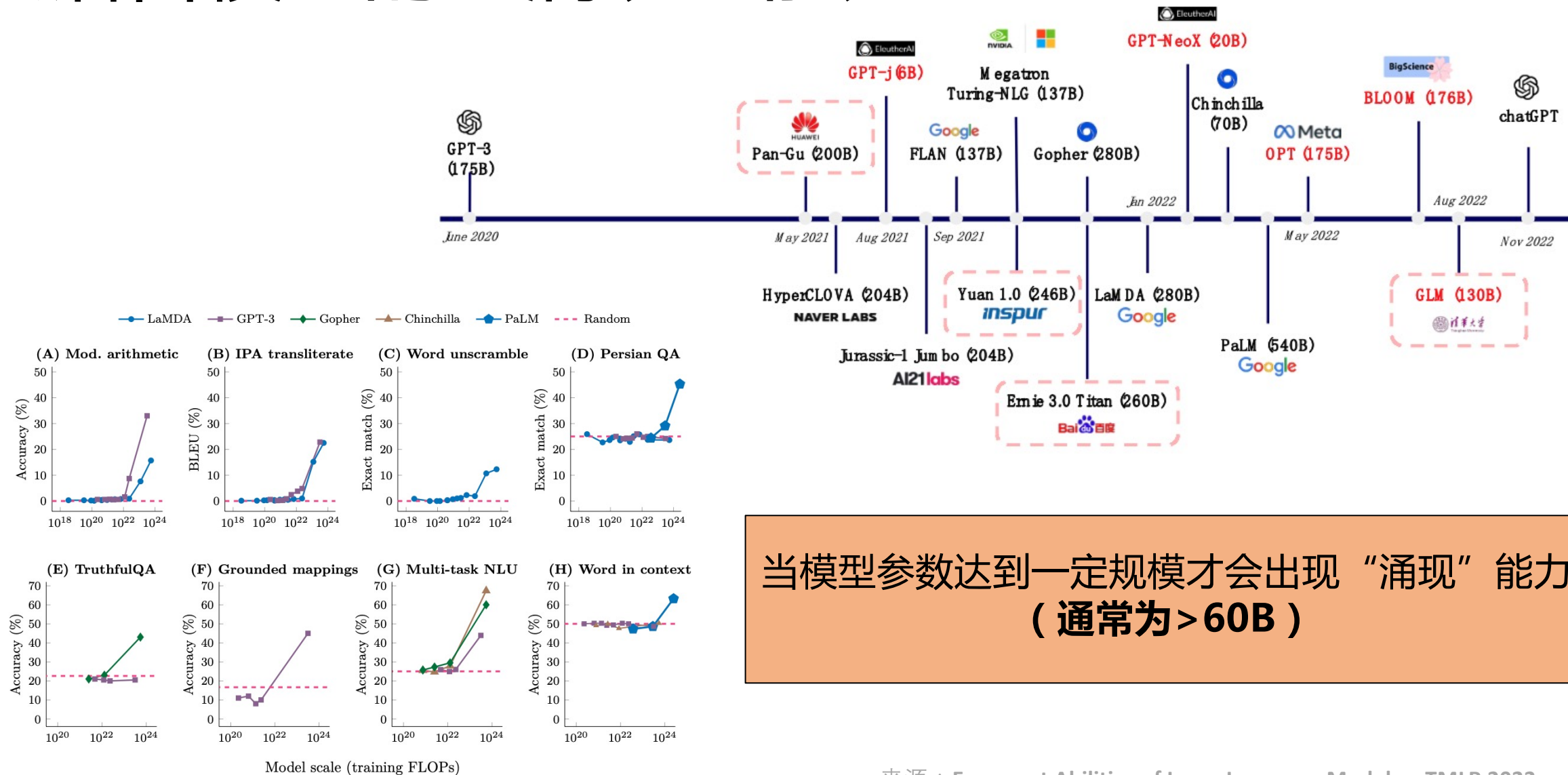
3. 模仿人类老师的喜好

引导 GPT 说人类
要他说的

4. 用增强式学习向模拟老师学习



大语言模型的“涌现”能力



ChatGPT具备的能力

- **预训练语言模型：注入基础能力**

- 语言生成
- 上下文学习
- 世界知识

- **指令微调：解锁/激发能力**

- 遵循指令
- 零样本泛化

- **代码训练**

- 代码生成和理解
- 复杂推理
- 思维链

- **RLHF**

- 公正、翔实回复
- 拒绝不当问题
- Know unknown
- 对话能力

GPT4

- **大型多模态模型**

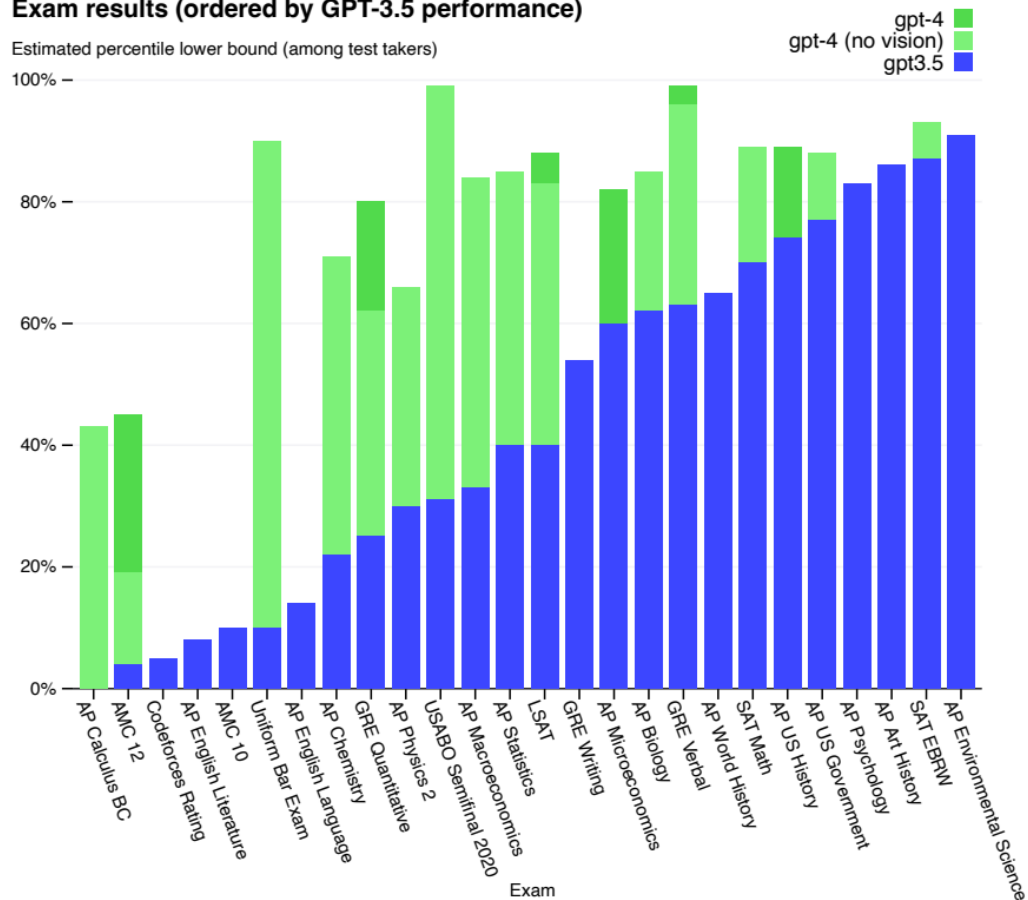
- 可接受图文输入，输出文本
- 专业和学术基准上表现出人类水平

- **和 ChatGPT 相比的进步**

- 更好的推理能力
- 处理更复杂的指令
- 更具创造力

Exam results (ordered by GPT-3.5 performance)

Estimated percentile lower bound (among test takers)



主要内容

1. ChatGPT简介

2. ChatGPT关键技术

3. ChatGPT展望

4. ChatGPT团队研究

ChatGPT的局限

- 无法形式推理
- 事实性错误

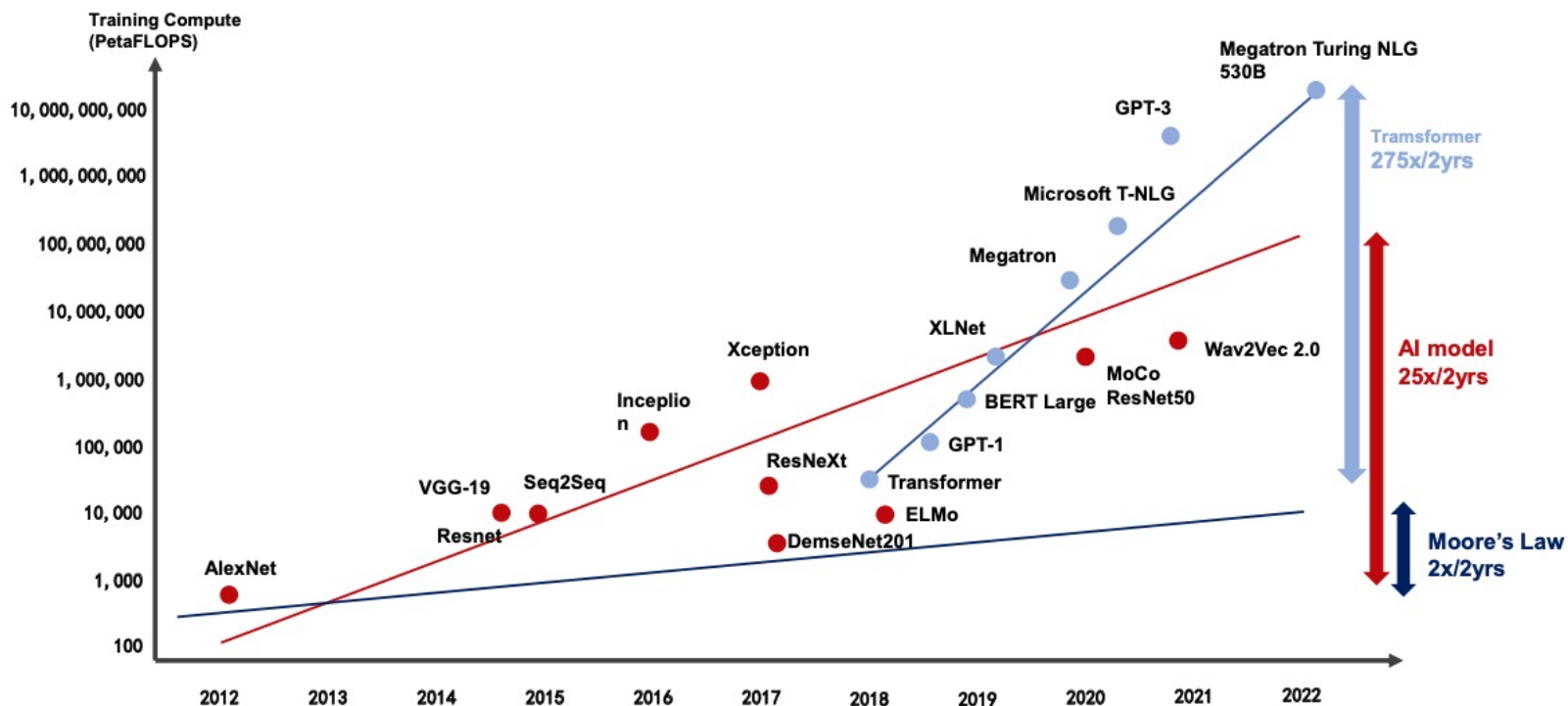
Q: 请问勒布朗·詹姆斯在NBA的总得分是第几名?

A: 截至2021年5月，勒布朗·詹姆斯在NBA的总得分是历史第3名，共得到34,367分。

A: 勒布朗·詹姆斯在NBA的总得分排名第一，他的总得分为34,087分。

大语言模型面临挑战1

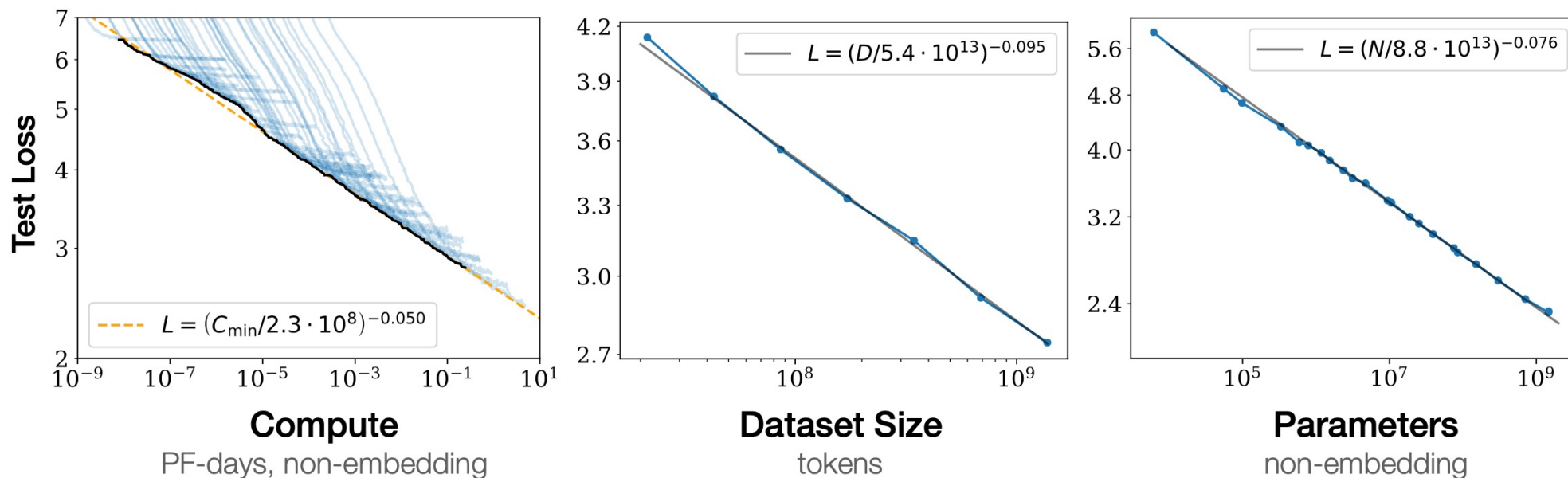
- 训练算力增长速度超越芯片摩尔定律
 - 算力需求3-4个月翻1倍，头部公司每年增长10倍
 - 晶体管每2年增长1倍，深度学习逼近算力的物理极限



大语言模型面临挑战2

- 参数量进入平台期
 - 参数量增大的边际效应递减

Scaling Law

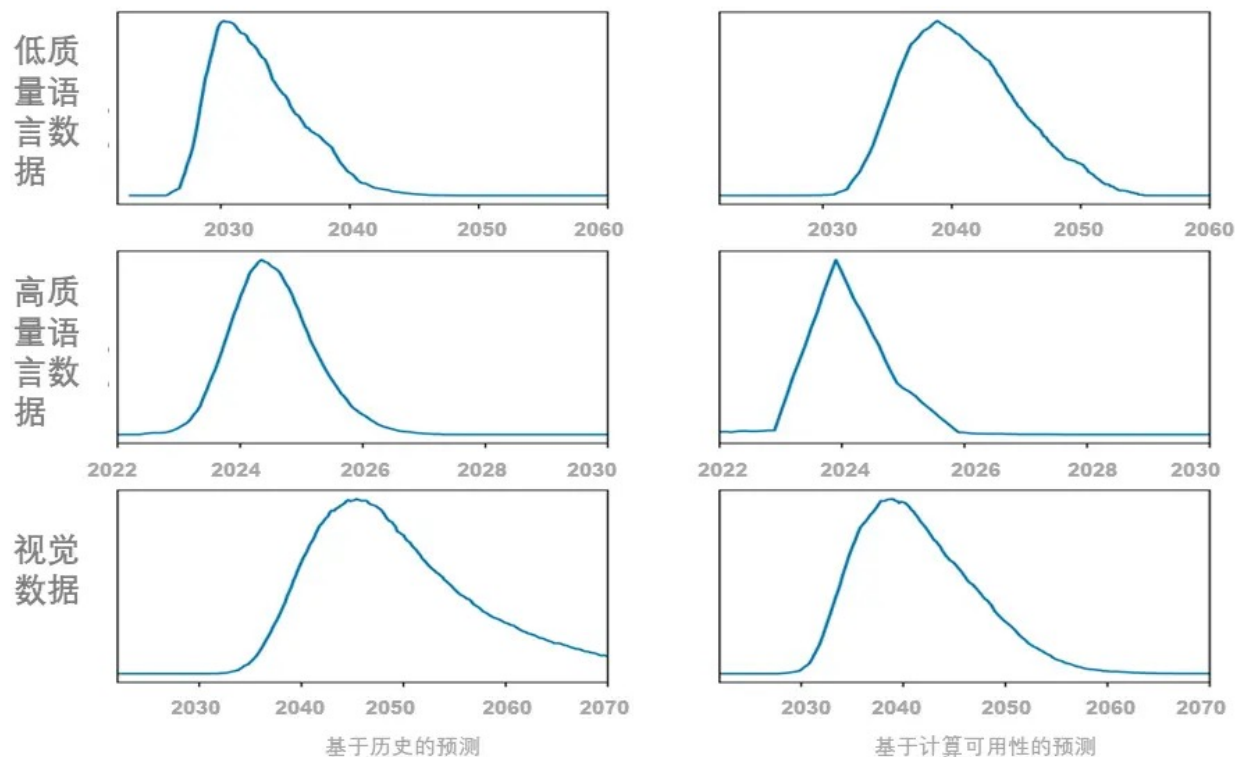


模型能力与参数量成log-linear关系

大语言模型面临挑战3

- 训练语料即将耗尽

不同数据类型的消耗趋势和耗尽日期



来源: Pablo Villalobos等, 未尽研究

说明: 可训练模型数据量有限制, 最大的限制之一就是计算可得性。要对已有模型增加训练数据量, 需要更多额外的计算, 而计算会受到硬件供应以及购买、租用成本的制约。所以, 预测数据集规模时要将计算可得性的考虑进去。

大语言模型面临挑战4

- 自学习

LARGE LANGUAGE MODELS CAN SELF-IMPROVE

Jiaxin Huang^{1*} Shixiang Shane Gu² Le Hou^{2†} Yuexin Wu² Xuezhi Wang²
Hongkun Yu² Jiawei Han¹

¹University of Illinois at Urbana-Champaign ²Google

¹{jiaxin3, hanj}@illinois.edu ²{shanegu, lehou, crickwu, xuezhiw, hongkunyu}@google.com

SELF-INSTRUCT: Aligning Language Model with Self Generated Instructions

Yizhong Wang[♣] Yeganeh Kordi[◇] Swaroop Mishra[♡] Alisa Liu[♣]

Noah A. Smith^{♣+} Daniel Khashabi[♣] Hannaneh Hajishirzi^{♣+}

[♣]University of Washington [◇]Tehran Polytechnic [♡]Arizona State University

[♣]Johns Hopkins University ⁺Allen Institute for AI

yizhongw@cs.washington.edu

- 知识遗忘

RECITATION-AUGMENTED LANGUAGE MODELS

Zhiqing Sun^{1,2*}, Xuezhi Wang¹, Yi Tay¹, Yiming Yang², Denny Zhou¹

¹Google Research, Brain Team

²Language Technologies Institute, Carnegie Mellon University

大语言模型面临挑战5

- 安全性、伦理道德
 - 基于AI反馈的强化学习 (RLAIF)
 - AI助手Claude （来自Anthropic公司）
 - 人工书写规则，基于规则打分

振奋人心之处

- **模型规模增速变缓**
 - 摩尔定律
 - 边际效应递减
- **特定应用不如专门模型**
 - 机器翻译
- **Specialize**只需更小模型规模

未来发展趋势

- 高训练成本和高技术壁垒使得少数组织为主要玩家
 - 头部公司
 - 国家或组织构建大模型作为基础设施，开放下游使用
- 掌握用户和用户数据使得工业界优势明显，形成AI正循环，学术界研究空间受到积压。
- 研究和应用范式改变，部分研究方向需求降低。
- 通用大模型和垂类大模型并存。
- 大模型和知识库结合，如和搜索引擎结合。

学术界如何应对

- 研究新模型
 - There will be better systems that are factual, non-toxic and controllable. They just won't be auto-regressive LLMs. —Yann Lecun.
 - 非自回归模型超越Transformer，获WMT2021德英翻译冠军。
- 研究垂类大模型或具体任务大模型
 - 模型规模将减小
 - 基于领域知识精调
- 扩大交互范围
 - 语言 → 多模态 → 物理世界（具身智能）

主要内容

1. ChatGPT简介

2. ChatGPT关键技术

3. ChatGPT展望

4. ChatGPT团队研究

KG和NLP领域的主要挑战

- 样本低资源情况
 - 标注数据匮乏
 - 整体的数据规模不大
 - 互联网迁移的数据可能存在较多污染的“脏”数据
- 应用场景的实时性
 - 需要在实际的应用场景下是边缘端和国产化的场景
 - 对于情报的分析速度存在较高要求
- 情报分析处置的真实性和可靠性
 - LLMs存在事实性偏差
 - 生成结果存在反事实、违背伦理的情况
 - LLMs的可解释性存在一定的问题

谢谢！