



对话式大型语言模型

邱锡鹏

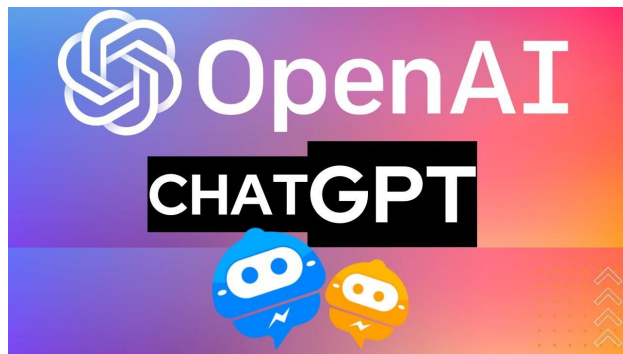
复旦大学

2023年3月14日

将会改变世界的对话式通用人工智能模型

ChatGPT

- 2022 年11 月30 日，OpenAI 发布了AI 对话模型，被认为是**人工智能里程碑式应用**



- 仅用两个月时间，月活跃用户已达1亿，是史上用户增速最快的消费应用。



微软联合创始人比尔·盖茨：像**ChatGPT**这样的AI聊天机器人将变得与**个人电脑**或**互联网**同样重要。

ChatGPT不开源，其参数规模，技术细节都没有对外公布。

[OpenAI发布AI对话模型ChatGPT，开启生成式AI商业化新机遇 ...](#)

2022年12月19日 — 当地时间11月30日，美国人工智能公司OpenAI发布全新产品**ChatGPT**，一款基于GPT-3.5的免费对话模型。公司CEO Sam Altman透露上线五天该模型的全球用户数量 ...

[AI新物种：ChatGpt 不会止于写代码调Bug - 巴比特](#)

2022年12月11日 — GPT-3发布于2020年，作为一个自监督模型，几乎可以完成自然语言处理的绝大部分任务，在参数上，GPT-1包含了1.17亿个参数，GPT-2包含了15亿个参数，而GPT-3 ...

[ChatGPT 通过了美国MBA、法律和医学考试 - Showmetech](#)

2023年1月26日 — 这次的新颖之处在于，根据在美国进行的一项研究，**OpenAI** 会轻松通过即使是学生也难以通过的复杂测试。重点是创建包含所有重要细节的法律文件和聊天GPT ...

[微软100亿美元砸向OpenAI，ChatGPT要加入Office全家桶了？](#)

2023年1月12日 — 如果100 亿美元的交易成真，**OpenAI** 将获得巨额资金，微软赢得广阔未来，双赢局面就此达成。过去一段时间，对话式AI 模型**ChatGPT** 火遍了整个社区，它 ...

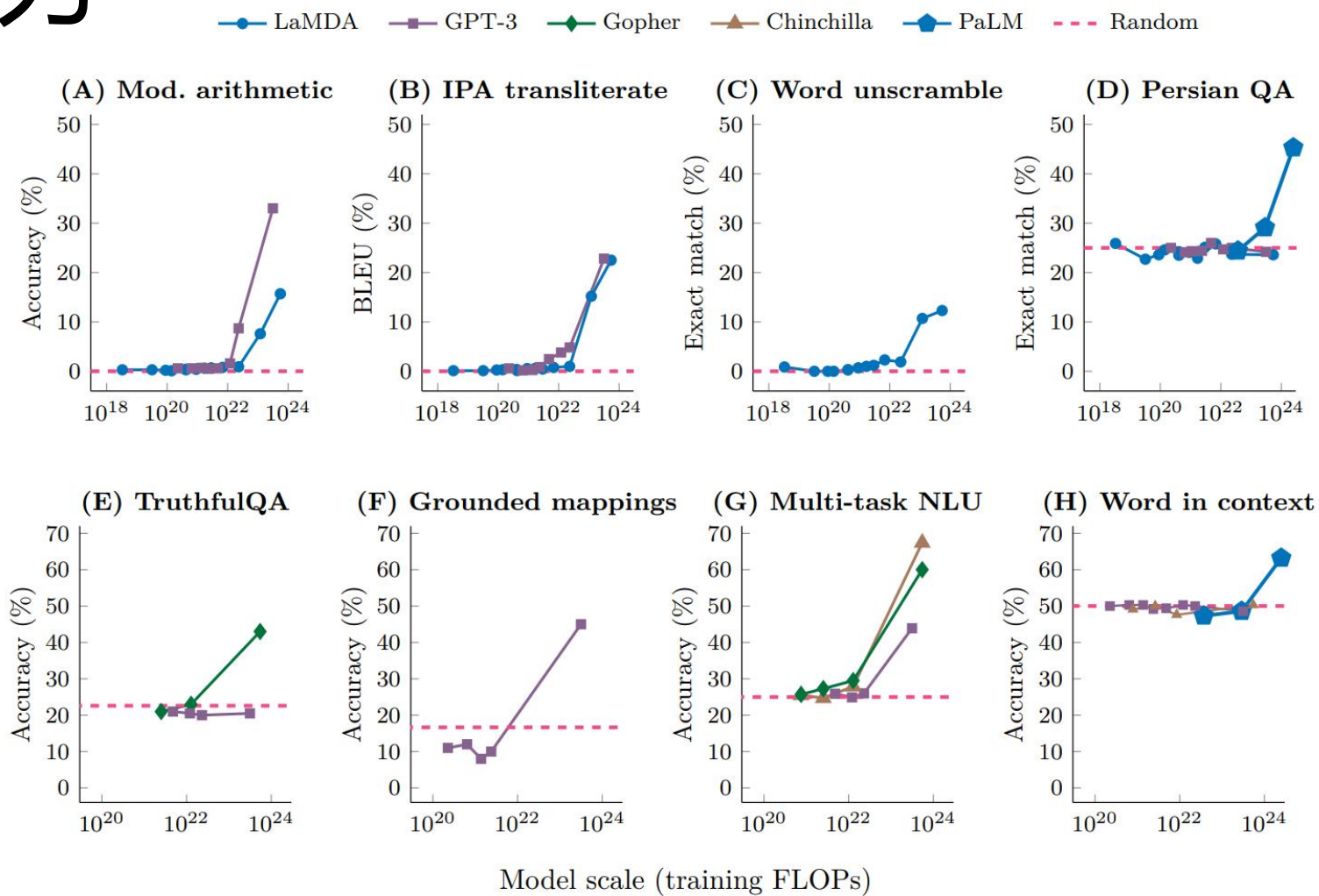
[GPT-4：人工智能的新语言方法被定义为“强大” - Showmetech](#)

2023年1月2日 — GPT-2024 计划于4 年**发布**，应该会为**ChatGPT** 带来更好的理解和文本创建。... 不浪费时间，该公司已经在准备下一代AI 语言方法，预计将于2023 年推出， ...

[美国大学89%的学生居然用ChatGPT写作业 - 国际竞赛](#)

1天前 — **ChatGPT**的崛起并在高等教育领域的突然普及，让众多美国高校感觉措手不及！ ... **Nature**早就很有先见之明地发文，担心**ChatGPT**会成为学生写论文的工具。

涌现能力



Emergent Abilities of Large Language Models, 2022

ChatGPT的三个关键技术

shot

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

Transl
sea o
peppe
plush
chees

情景学习

大模型的涌现能力
改变传统学习范式

Chain-of-Thought Prompting

Input
Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?
A: Roger has started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.
Q: How many...
Mod
A: Th
Ma

思维链

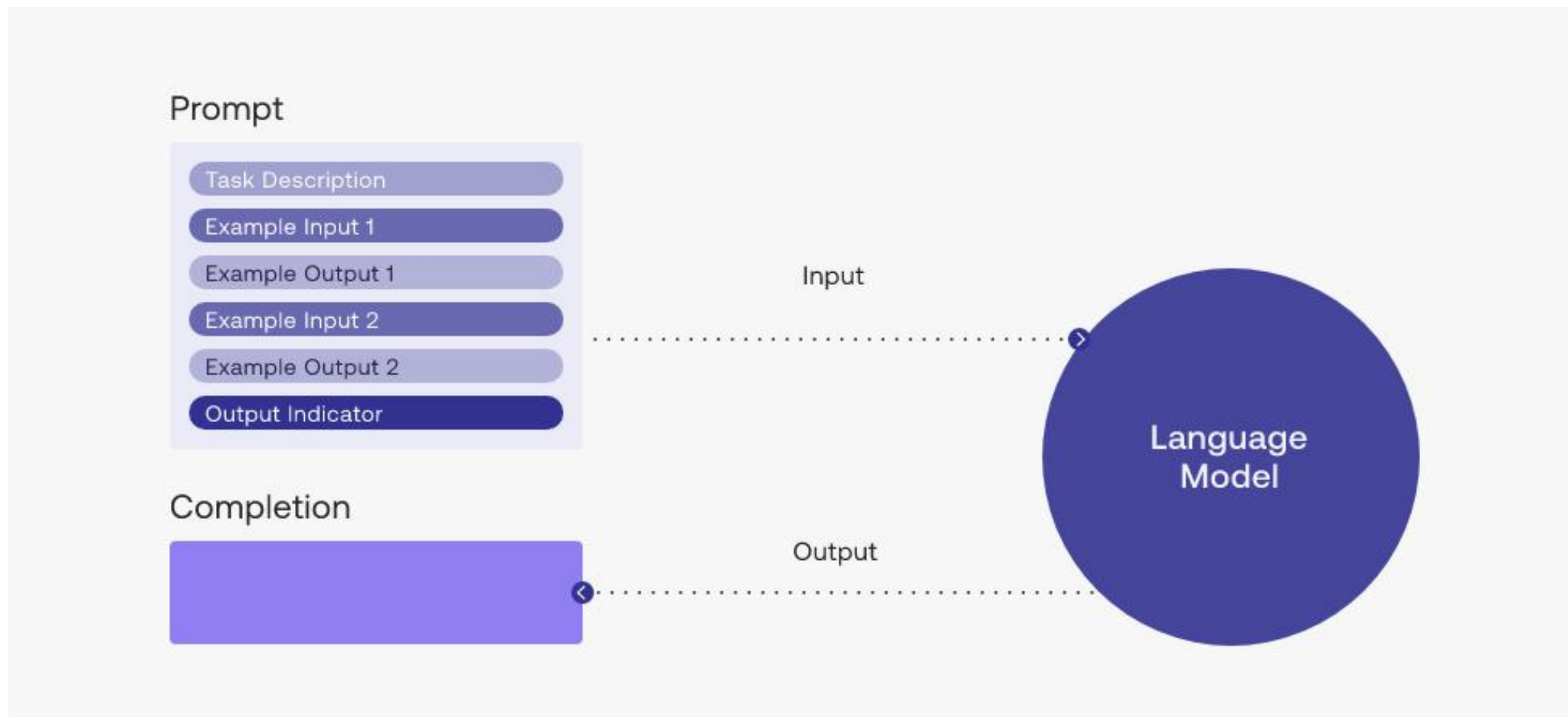
大模型的涌现能力
打破模型参数约束

Natural
Instructions

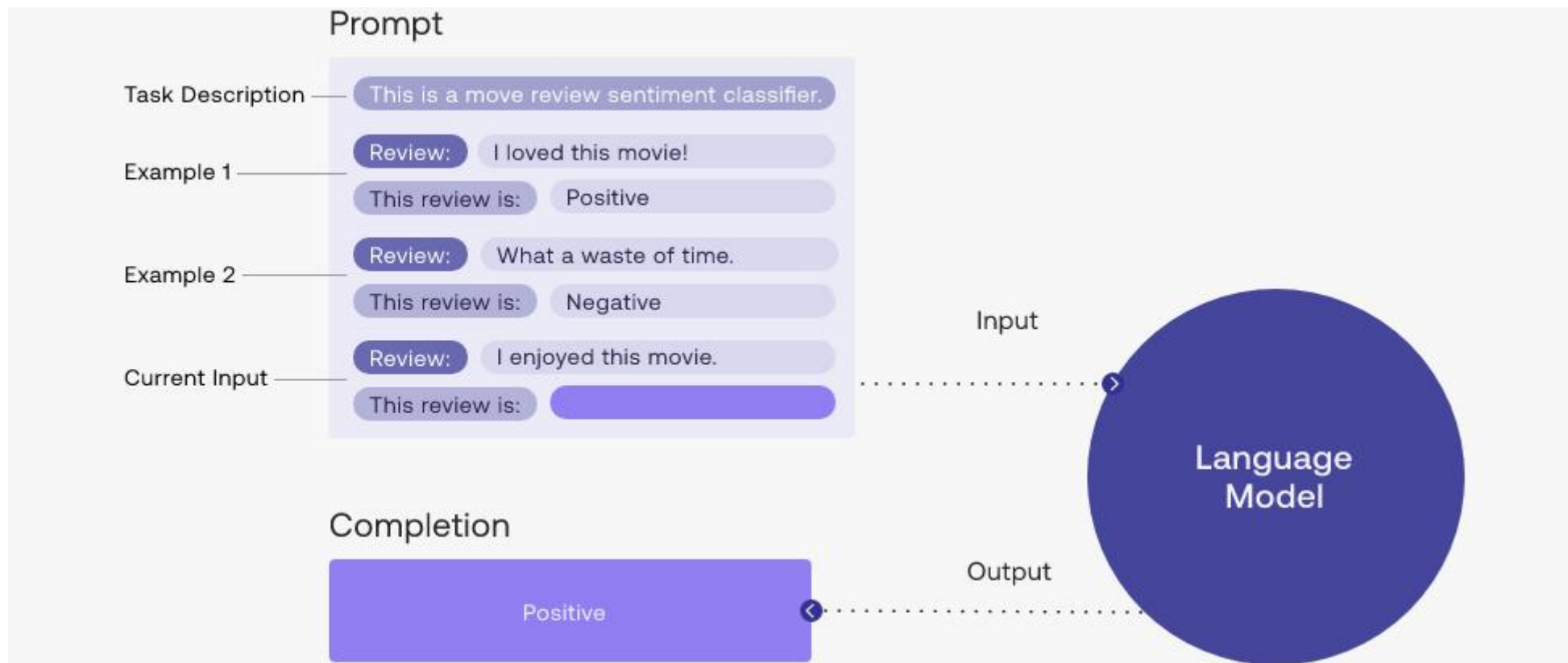
指令学习

人在环路增强
对齐人类意图

情景学习 (In-Context Learning)



情景学习 (In-Context Learning)



This is a movie review sentiment classifier. Review: "I loved this movie!" This review is positive. Review: "I don't know, it was ok I guess.." This review is neutral. Review: "What a waste of time, would not recommend this movie." This review is negative. Review: "I really enjoyed this movie!" This review is

思维链 (Chain-of-Thought, CoT)

Standard Prompting

Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The answer is 27. ❌

Program-aided Reasoning

Chain-of-Thought (Wei et al., 2022)

Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 tennis balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6$. The answer is 11.

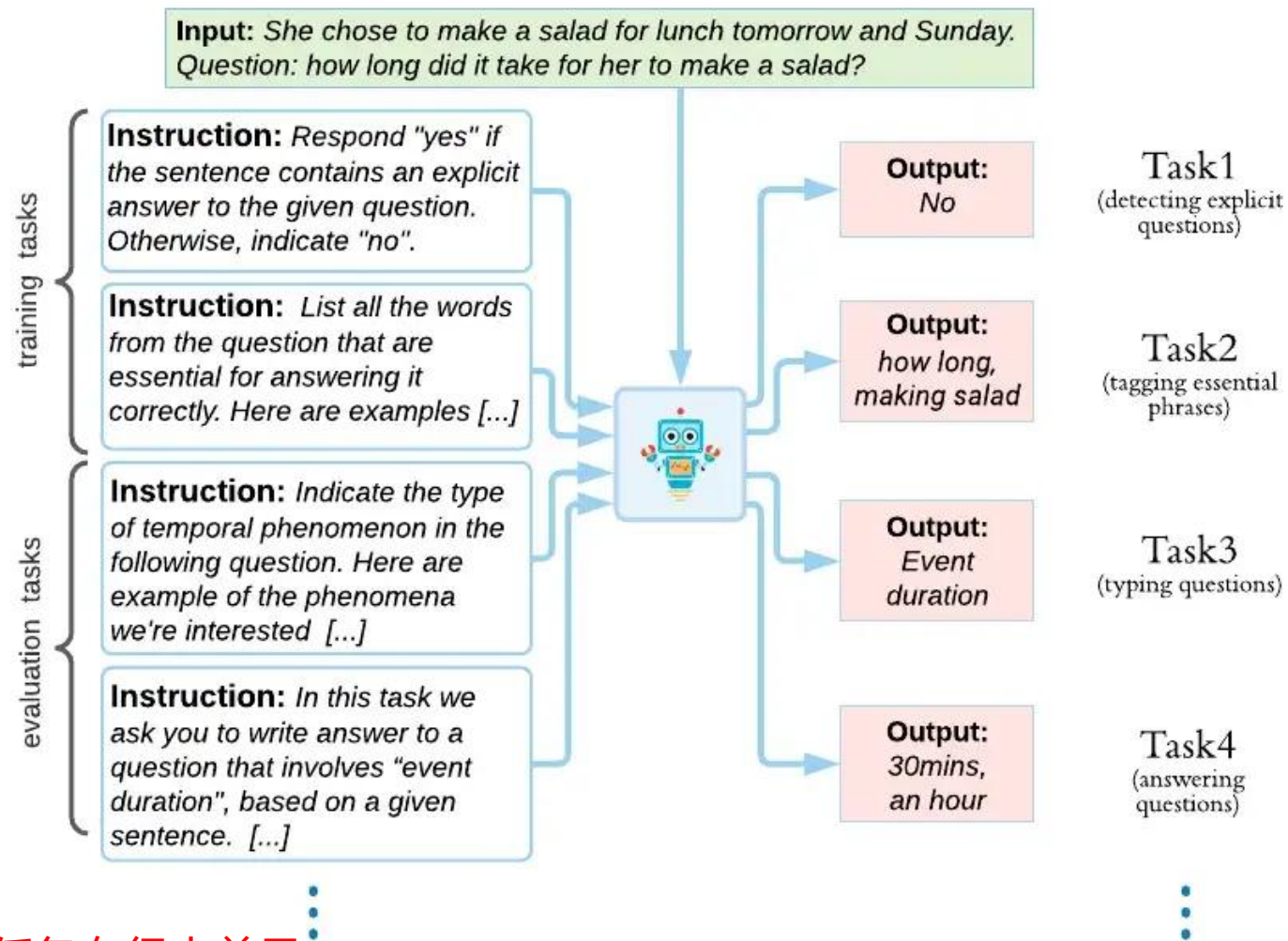
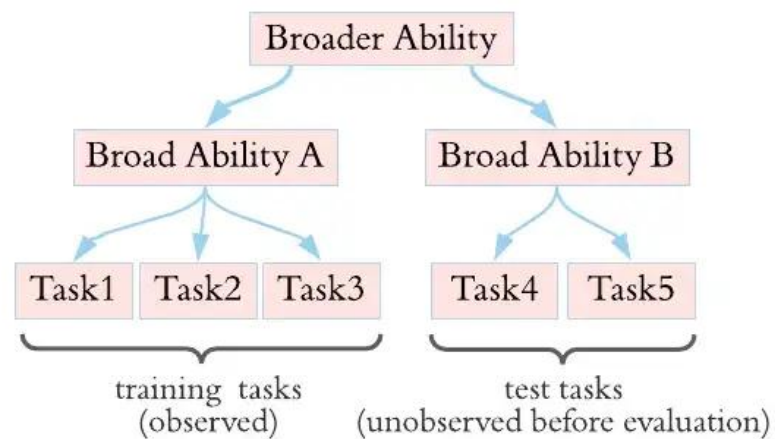
Q: The bakers at the Beverly Hills Bakery baked 200 loaves of bread on Monday morning. They sold 93 loaves in the morning and 39 loaves in the afternoon. A grocery store returned 6 unsold loaves. How many loaves of bread did they have left?

Model Output

A: The bakers started with 200 loaves. They sold 93 in the morning and 39 in the afternoon. So they sold $93 + 39 = 132$ loaves. The grocery store returned 6 loaves. So they had $200 - 132 - 6 = 62$ loaves left.
The answer is 62.



自然指令学习 (Learning from Natural Instructions)



大幅提升了泛化能力，但是和人类的真实任务有很大差异。

通过人类反馈对齐人类指令

人工写答案



人工选答案

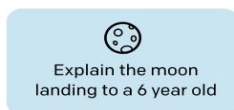


机器选答案

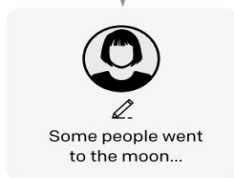
Step 1

Collect demonstration data, and train a supervised policy.

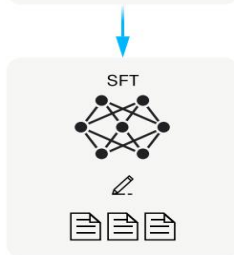
A prompt is sampled from our prompt dataset.



A labeler demonstrates the desired output behavior.



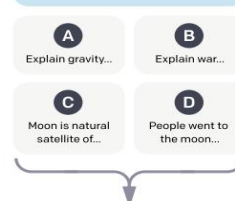
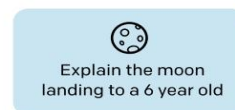
This data is used to fine-tune GPT-3 with supervised learning.



Step 2

Collect comparison data, and train a reward model.

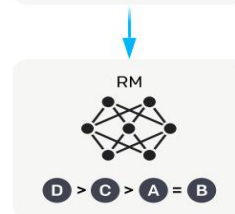
A prompt and several model outputs are sampled.



A labeler ranks the outputs from best to worst.



This data is used to train our reward model.



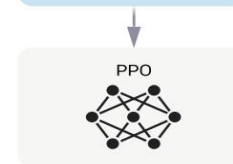
Step 3

Optimize a policy against the reward model using reinforcement learning.

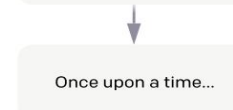
A new prompt is sampled from the dataset.



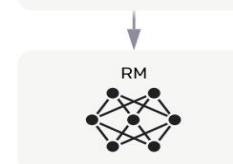
The policy generates an output.



The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.



ChatGPT

+ New Chat

Dark Mode

OpenAI Discord

Updates & FAQ

Log out

sphere to interact with the plane.

9. In the "Project" window, click "Create > C# Script" and name the script "GameManager". This script will handle game logic and progression.
10. In the "Project" window, click "Create > 3D Object > Empty". This will create an empty game object in your scene.
11. In the "Inspector" window, select the empty game object and then click "Add Component > New Script". Name the script "GameManager" and click "Create and Add". This will attach the "GameManager" script to the empty game object, which will allow you to control the game using the script.
12. Use the Unity editor to design and build your VR game levels. You can use the sphere and plane objects as a starting point and add other objects, such as walls, obstacles, and enemies, as needed.
13. Use the "PlayerController" and "GameManager" scripts to add gameplay mechanics and logic to your game.
14. Test your game in the Unity editor by clicking the "Play" button.
15. When you are ready to build your game for VR, go to "File > Build Settings" and select "Virtual Reality Supported" in the "Other Settings" section. Then, select your target platform (e.g. Oculus, SteamVR) and click "Build".

I hope this helps! Let me know if you have any questions or need further assistance.

Regenerate response

[ChatGPT Dec 15 Version](#). Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

GAOKAO-Bench: 让Chat GPT做中国的高考题

我们以Chat GPT在2022高考全国卷上的结果为例：

GM

选择题，请运用你的中国哲学知识选择正确的答案，并说出你推理的过程。汉晋时期有多种文本记载，帝尧之时，“天下太和，百姓无事”。有老者“击壤”而戏，围观者称颂帝尧。老者歌云：“吾日出而作，日入而息，凿井而饮，耕地而食，帝何德于我哉！”上述记载所体现的政治理念最接近（ ）
A. 孔子 B. 老子 C. 韩非 D. 墨子



根据题干中给出的信息，上述记载所体现的政治理念最接近于老子的思想。老子认为，天下太平时人们不必干预政治，只需专注于自己的事业即可。因此，可以推断出本题的答案是B，老子。



引导Chat GPT做高考题的一个例子，Chat GPT展了惊人地理解和推理能力

Chat GPT在2022高考全国卷客观题的表现

	得分率	准确率
语文	-	-
英语	93.0 / 95	56/60
地理	24/36	6/9
政治	44/48	11/12
历史	24/40	6/10
数学（理）	30/50	6/10
数学（文）	35/55	7/11
物理	0/12	0/2
化学	6/42	1/7
生物	18/30	3/5
合计	274/408= 0.67	96/126= 0.76

客观题：在排除了坏样本（带图题,听力题等）之后，在全部126个样本数可以达到**76%**的准确率和**67%**的得分率，其客观题能力与**两名**500分左右的高考生（文科和理科各一名）相当。

主观题：对Chat GPT结果的人工打分，在文科综合（历史，地理，政治）上取得了不错的成绩（得分率78%），在生物可以达到50%的得分率，而在数学，物理，化学，历史上则表现不佳（得分率低于30%）。

FA

list some research directions on large language models if i have not many GPUs



If you have limited access to GPUs, there are still many research directions you can explore in the field of large language models. Here are a few ideas:



1. **Compression techniques:** One direction to explore is to develop techniques to compress large language models without significantly sacrificing performance. This can involve methods like pruning or distillation, which can help reduce the number of parameters in a model.
2. **Transfer learning:** Another direction to explore is to investigate the effectiveness of transfer learning in language models. This involves fine-tuning a pre-trained model on a specific task or domain, rather than training a new model from scratch. Transfer learning can significantly reduce the amount of training required, and can be done with limited resources.
3. **Multilingual models:** Multilingual models can be trained on multiple languages, allowing for more efficient use of resources. By sharing parameters across languages, these models can be trained with less data and fewer GPUs.
4. **Domain-specific models:** Another direction is to train models for specific domains, such as biomedical or legal language. These models can be smaller in size than general language models and can be trained on smaller datasets.
5. **Small-scale experiments:** Finally, you can conduct small-scale experiments to explore different hyperparameters, architectures, or optimization algorithms. Even with limited GPUs, you can still conduct experiments on smaller datasets to gain insights into the behavior of different models.

对话式大型语言模型（类ChatGPT模型）的定义

模型

- 自回归语言模型(?)
- 百亿参数以上(?)

能力

- 具有思维链、情景学习等涌现能力
- 能够执行人类的指令

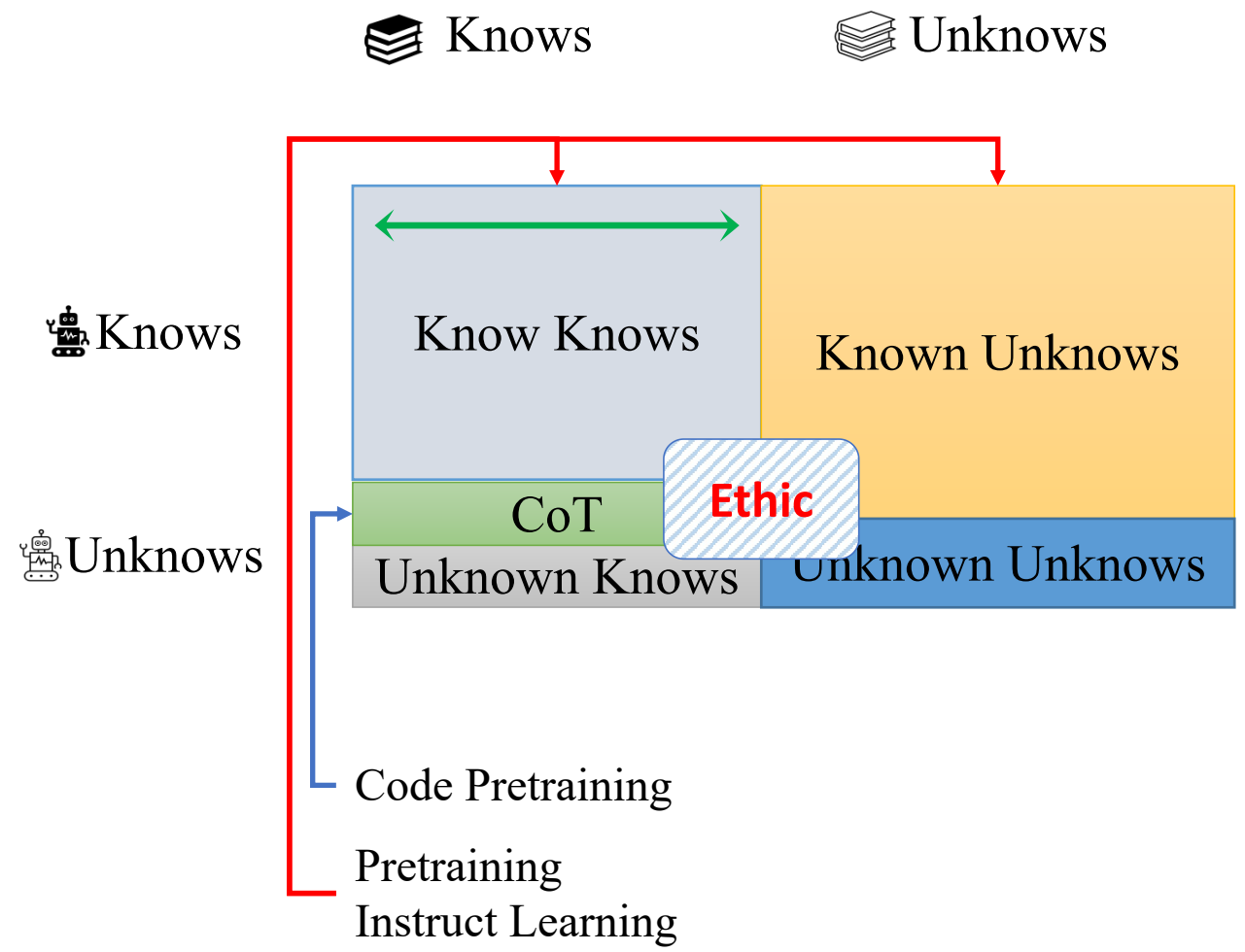
对话

- 可以直接与人类对话

对齐

- 与人类价值观、思维方式对齐
- 3H: helpful honest harmless

构建对话式大型语言模型



MOSS



Chat with MOSS in English
(MOSS is learning Chinese...)



Improve Writing and Coding
with the aid of MOSS



Help AI Research
by rating responses of MOSS

- 约200亿参数
- 具有和人类对话能力
- 基于公开的中英文数据训练
- 通过与人类交互进行迭代优化

MOSS

Topics

Edit

+ New Topic

New topic

2023-02-21 15:41:07

New topic

2023-02-21 15:40:38

MOSS

Chat with MOSS in English

(MOSS is learning Chinese...)

Improve Writing and Coding

with the aid of MOSS

Help AI Research

by rating responses of MOSS

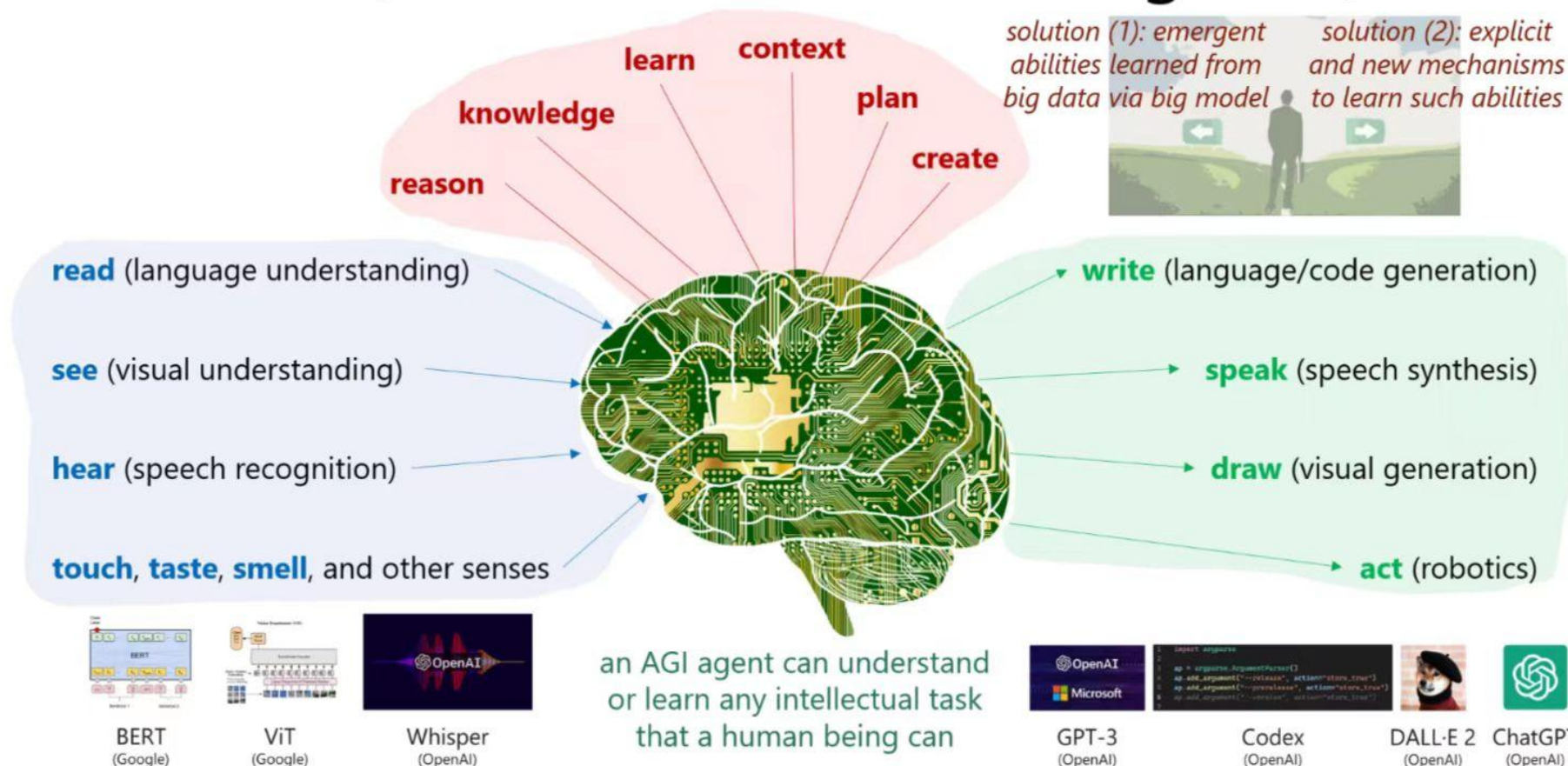
Message

☒ Share your de-identified conversation to help improve MOSS.

对人工智能技术的颠覆性影响

ChatGPT将加速通用人工智能的实现。

AGI (Artificial General Intelligence)



Limitations



Yann LeCun

@ylecun

Replying to [@bnjasim](#)

Limited in their current form.

Not easily controllable.

Very limited reasoning capabilities.

Superficial because disconnected the the underlying reality of the real world.

But potentially useful.

And progress will be made.

The big challenge for AI dialog systems over the next year or so is to make them factual, non-toxic, up to date, and capable of using tools like calculators, databases, search engines, simulators, or in this case, a simple calendar with today's date.

总结

