



大模型时代的认知智能

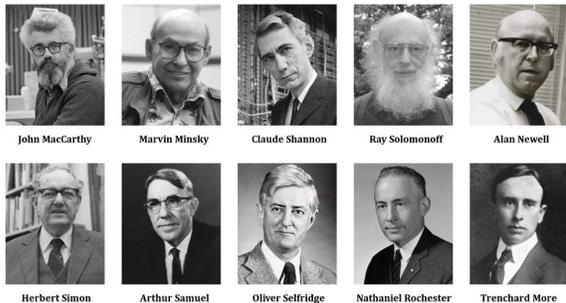


唐杰
清华大学

人工智能的发展 (2019年的PPT)

符号AI

1956 Dartmouth Conference:
The Founding Fathers of AI



符号模型/
规则模型/
感知机

第一代

认知智能



张钹院士2016年提出第三代人工智能雏形，DARPA 2018年发布AI Next计划。核心思路是推进**数据统计与知识推理融合**的计算；与**脑认知机理融合**的计算。

第三代



大数据驱动的学习方法初步实现了针对文本、图像、语音等的感知与识别

感知智能

目前急需的是高质量超大规模知识图谱（**AI的基础设施**）以及对超大规模数据的**深度理解能力**（面向认知的深度学习）

AI趋势 (2019年的PPT)

表示、计算、
存储与人机输入/出

文本内容识别、
图像识别、
语音识别

知识数据双驱动、
认知推理、
决策智能

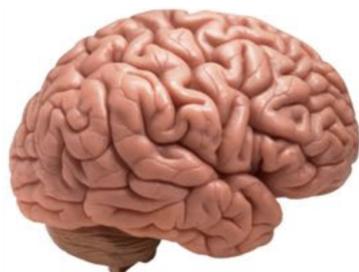
自学习、记忆
机理、有(无)
意识处理器

计算

感知

认知

意识



认知智能

趋势：用**计算模型**可描述的人类认知问题，计算机很快达到甚至超过人类水平。

很快=5-20年（2019年）

请帮我写一个祝福唐总教师节快乐、感谢他对我的指导和鼓励的短信，并在最后加上花的表情



尊敬的唐总：

祝您教师节快乐！在这个特殊的日子，我想借此机会向您表达我的感激之情。您不仅是我的导师，更是我的朋友，感谢您在工作和生活中给予我的关心、指导和鼓励。

因为您的帮助，我在工作中取得了不断的进步，也逐渐找到了自己的方向。您的言传身教将成为我终身学习的动力。我会继续努力，不辜负您的期望。

再次祝您教师节快乐！🌹

衷心感谢！

顺祝商祺！

以上内容为 AI 生成，不代表开发者立场，请勿删除或修改本标记



🔍 孙悟空有没有火烧赤壁

📖 孙悟空是中国古代小说《西游记》中的重要人物，他是一位神通广大的猴子，经过修行后成为了斗战胜佛。在小说中，孙悟空并没有参与火烧赤壁的战斗。

火烧赤壁是中国古代历史上著名的战役之一，发生在公元 208 年，是刘备、孙权联合对抗曹操的军队而形成的三方对峙局面。在战役中，孙刘联军采取水战，用火攻来攻破曹军。曹操的军队在火势蔓延之下，陷入混乱，最终溃败。

虽然孙悟空没有在火烧赤壁中战斗，但他在小说中参与了其他许多重要的战役和冒险，最终成为了斗战胜佛，成为了中国文学和艺术中的重要形象之一。

以上内容为 AI 生成，不代表开发者立场，请勿删除或修改本标记

👍 🗨️ 📄 复制

🔍 林黛玉有没有倒拔垂杨柳

📖 林黛玉是《红楼梦》中的人物，她自小多病，被称为病美人。在小说中，并没有提到林黛玉倒拔垂杨柳的情节。然而，有一些网友可能会恶搞小说情节，创作出林黛玉倒拔垂杨柳的段子。倒拔垂杨柳是鲁智深干的事情，鲁智深因为拳打镇关西的事情跑到东京大相国寺看守菜园子。菜园子经常有二、三十个泼皮来偷菜，没人能管得了他们。听说来了新人，他们就上门准备欺负一下。可是遇到了鲁智深，鲁智深就把两个领头的踢到粪坑里。这样的操作吓坏了泼皮。第二天泼皮上门赔礼，一起吃饭的时候，因为鸟叫声，鲁智深喝了点酒，就干出了倒拔垂杨柳的事情。

以上内容为 AI 生成，不代表开发者立场，请勿删除或修改本标记

👍 🗨️ 📄 复制



用python帮忙写个冒泡排序算法的代码，再给一个数组作为例子，写个测试程序

以下是冒泡排序算法的 Python 代码实现：

```
def bubble_sort(arr):  
    n = len(arr)  
    # 遍历所有数组元素  
    for i in range(n):  
        # 最后 i 个元素已经排好序，不需要再比较  
        for j in range(0, n-i-1):  
            # 如果当前元素比下一个元素大，则交换它们的位置  
            if arr[j] > arr[j+1]:  
                arr[j], arr[j+1] = arr[j+1], arr[j]  
    return arr
```

下面是测试程序，它将输入一个数组，并使用冒泡排序算法对其进行排序，然后输出排序后的数组：

```
arr = [64, 34, 25, 12, 22, 11, 90]  
sorted_arr = bubble_sort(arr)  
print(sorted_arr)
```

输出结果如下：

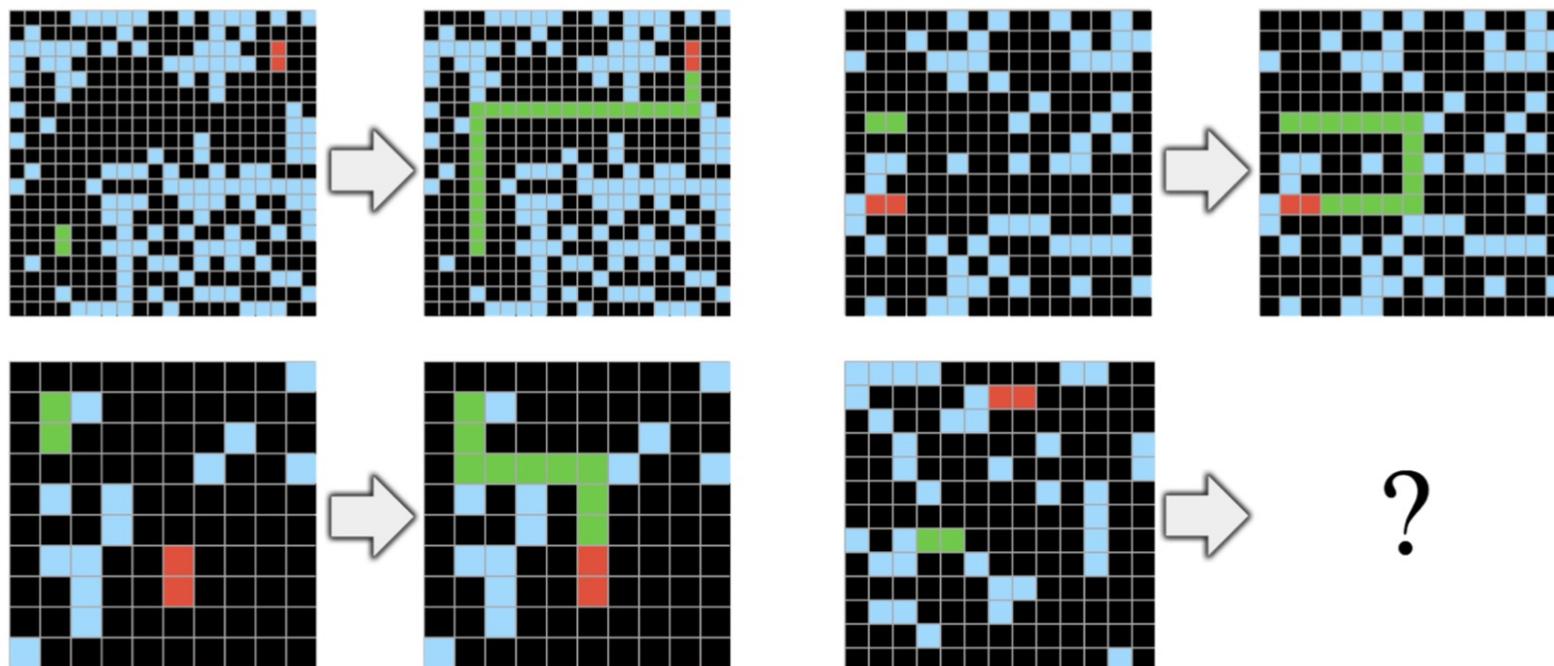
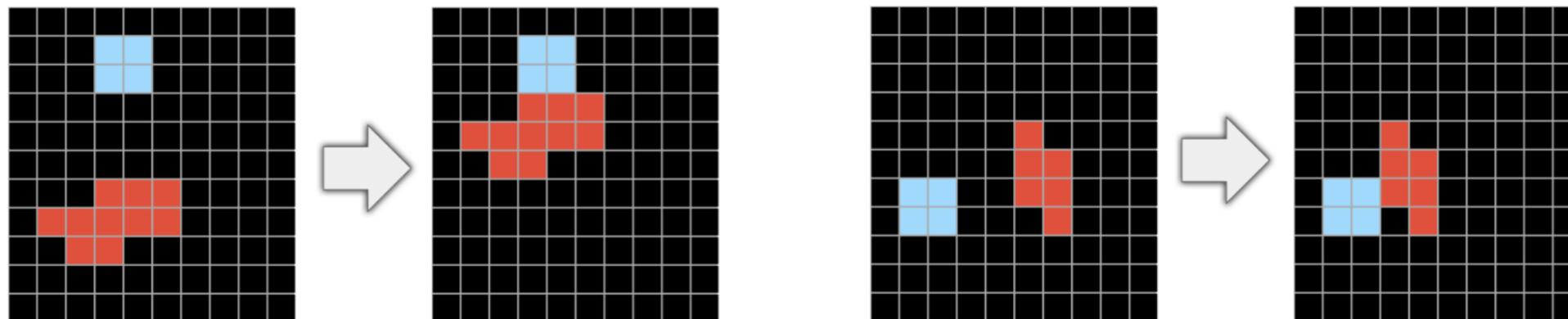
```
[11, 12, 22, 25, 34, 64, 90]
```

这是一个基本的冒泡排序算法，可以根据需求对其进行修改和扩展。

① 以上内容为 AI 生成



抽象与推理 (2019年的PPT)



智能代理

- 大模型驱动的智能代理
 - 生成式智能代理: A “*Westworld*” with 25 agents; Auto-GPT; AgentGPT...
- 用大模型驱动解决实际问题



AgentBench: 评估大模型系统的代理能力(lmbench.ai)

The logo for AgentBench is displayed in a white, stylized, rounded font against a black background. The word 'AgentBench' is written in a playful, bubbly typeface. The 'A' is particularly large and features a small robot head icon integrated into its top curve. The 'i' in 'Agent' is also stylized with a small robot head icon above it. The 'B' in 'Bench' has a small robot head icon integrated into its top curve. The overall aesthetic is clean and modern, reflecting the theme of AI agents.

AgentBench: 评估大模型系统的代理能力(lmbench.ai)

- 25 LLMs as of its initial release

- API-based
- Open models

- 8 Environments

1. Operating System
2. Database
3. Knowledge Graph
4. Digital Card Game
5. Lateral Thinking Puzzles
6. House-holding
7. Web Shopping
8. Web Browsing



Model	Model Size	Form	Version	Creator
gpt-4 [56]	undisclosed	api	0613	OpenAI
gpt-3.5-turbo [53]	undisclosed	api	0613	
text-davinci-003 [58]	undisclosed	api	-	
text-davinci-002 [58]	undisclosed	api	-	
claude [6]	undisclosed	api	v1.3	Anthropic
claude-instant [6]	undisclosed	api	v1.1	
text-bison-001 [3]	undisclosed	api	-	Google
chatglm2 [100; 21]	undisclosed	api	v0.2	Tsinghua & Zhipu
chatglm-6b [100; 21]	6B	open	v1.1	
codegeex2-6b-chat [102]	6B	open	-	
vicuna-13b [14]	13B	open	v1.1	LMSYS
vicuna-7b [14]	7B	open	v1.1	
llama2-chat-13b [80]	13B	open	-	Meta
llama2-chat-7b [80]	7B	open	-	
openchat-8192-13b [85]	13B	open	v1	Tsinghua
openchat-13b [85]	13B	open	v3.2	
wizardlm-30b [93]	30B	open	v1.0	Microsoft
wizardcoder-15b [48]	15B	open	v1.0	
wizardlm-13b [93]	13B	open	v1.0	
koala-13b [26]	13B	open	-	UC Berkeley
internlm-chat-7b [79]	7B	open	-	InternLM Team
baichuan-13b-chat [36]	13B	open	-	Baichuan Inc.
baichuan-7b [36]	7B	open	-	
dolly-12b [16]	12B	open	v2	Databricks
oasst-sft-4-12b [42]	12B	open	-	LAION

双系统的认知智能

系统1：直觉性思考

快速

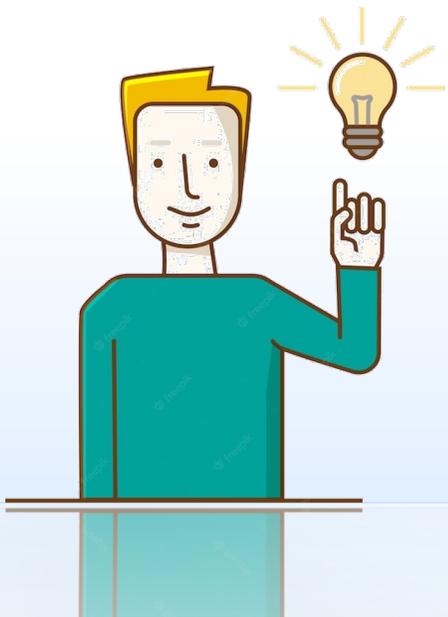
关联性推理

自动的

无需工作记忆

无意识的

人类与动物共有



系统2：推理性思考

慢速

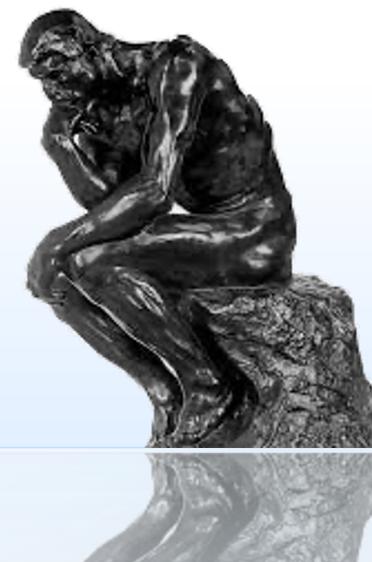
逻辑性推理

人为控制的

需要工作记忆

有意识的

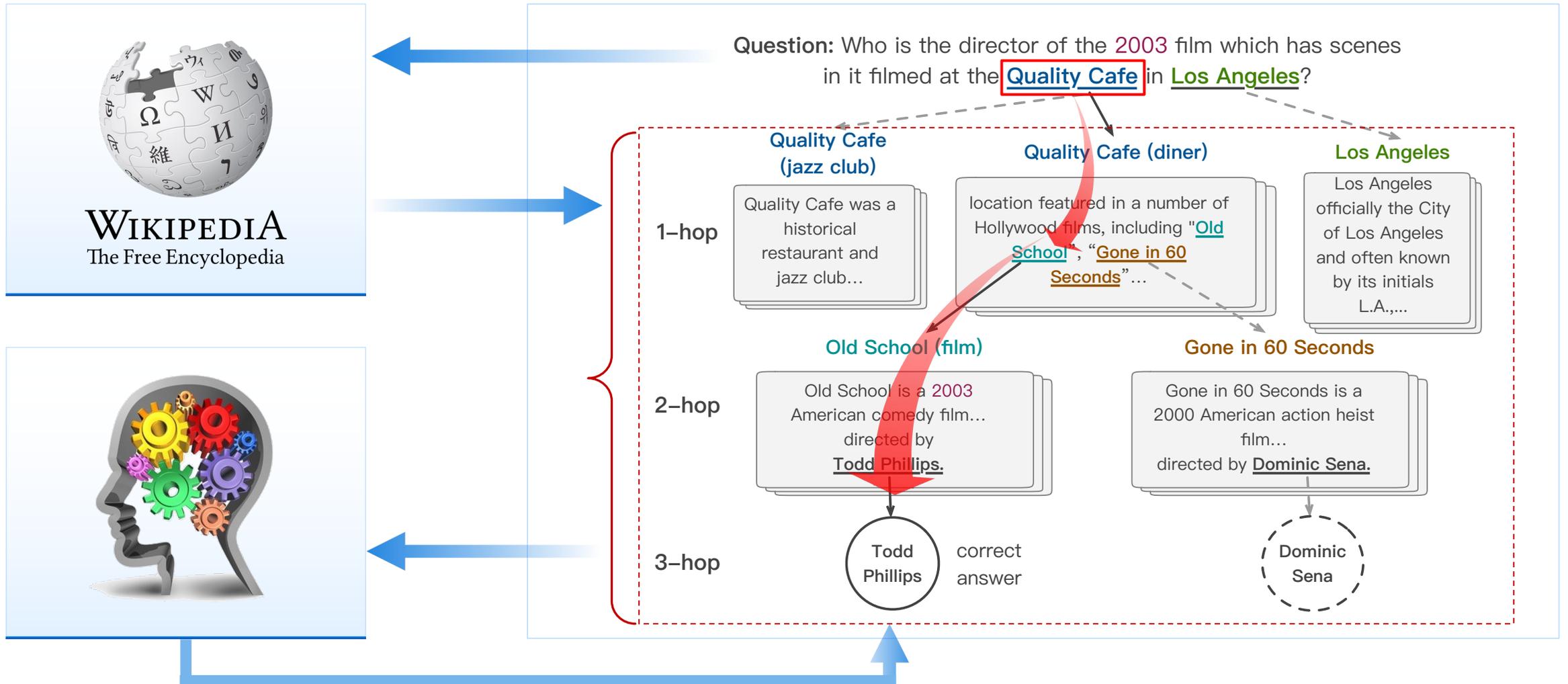
人类独有



认知心理学
双过程理论

双系统的认知智能

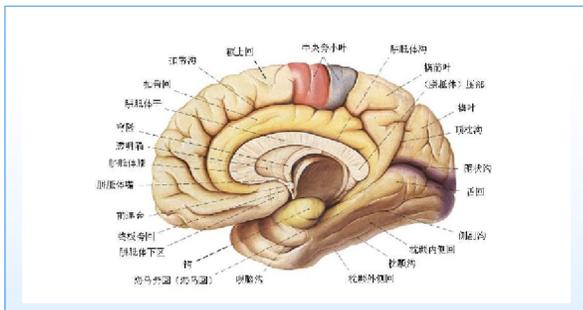
可解释的开放领域多跳阅读理解问答



认知大模型的思考 (始于2019的五年规划)

系统1

系统2

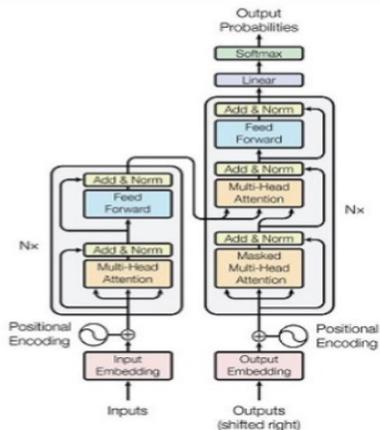


错误探测与反馈模型

蒸馏与抽象

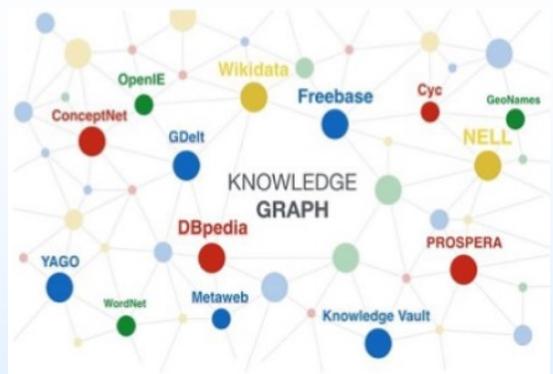


超大规模预训练



记忆模型 (知识)

CodeGeex, QA, Web, KG...

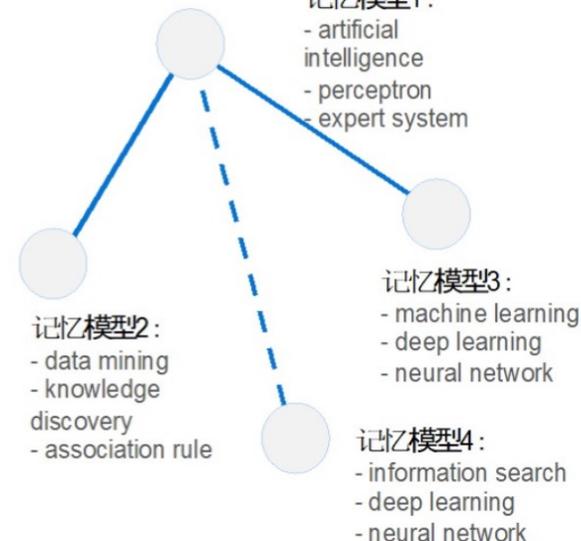


认知推理

“无意识”探测

自定义条理与逻辑

Self-Instruct



查询接口

GLM-130B

“有意识”决策

用户反馈与反思

ChatGLM

Open AI引领了AGI的研究...

*我们更开放

预训练大模型是新一代人工智能应用的基础设施





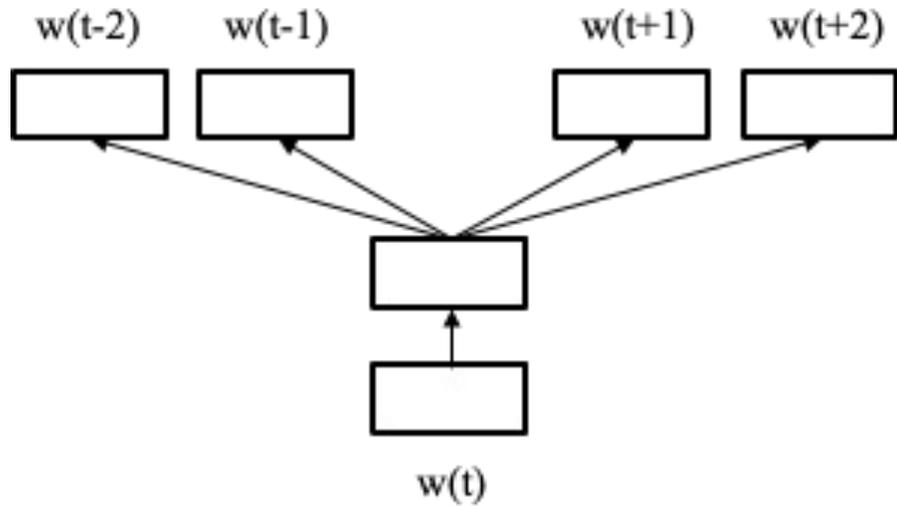
Why large models?

本质的一些思考

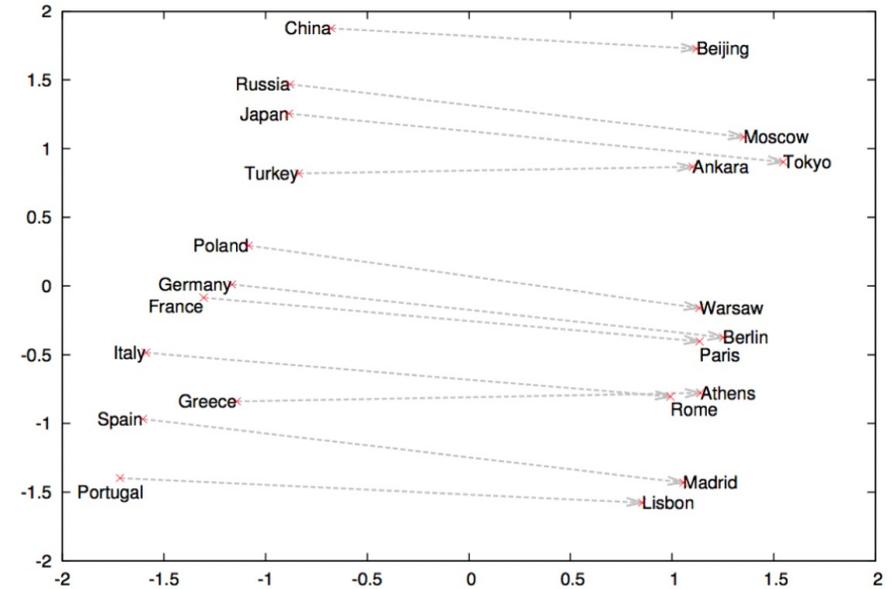
静态表示学习

◆ “You shall know a word by the company it keeps.” —John Rupert Firth

• 静态表示

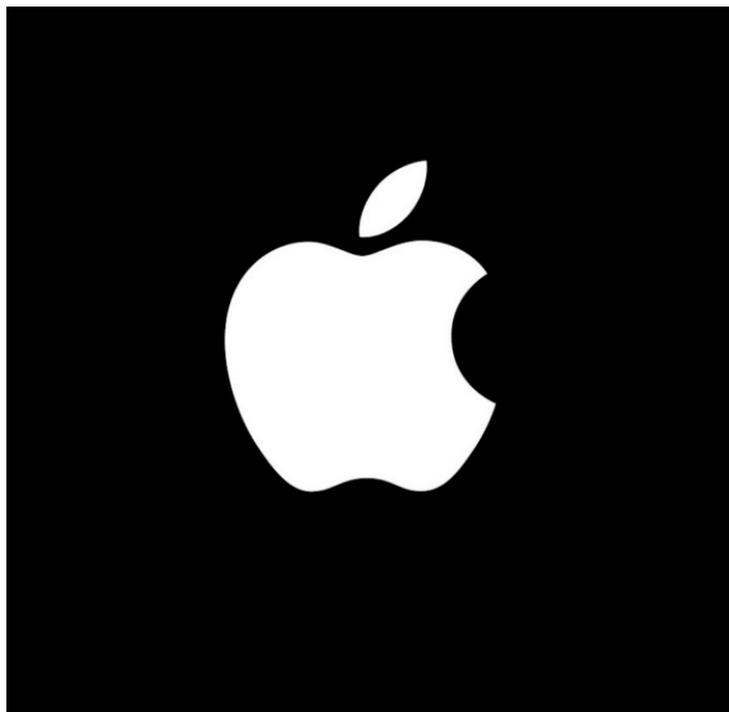


0.8	0.2	0.3	...	0.0	0.0
-----	-----	-----	-----	-----	-----



$$W(\text{"China"}) - W(\text{"Beijing"}) \approx W(\text{"Japan"}) - W(\text{"Tokyo"})$$

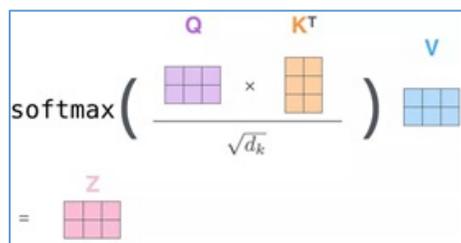
静态表示学习的问题

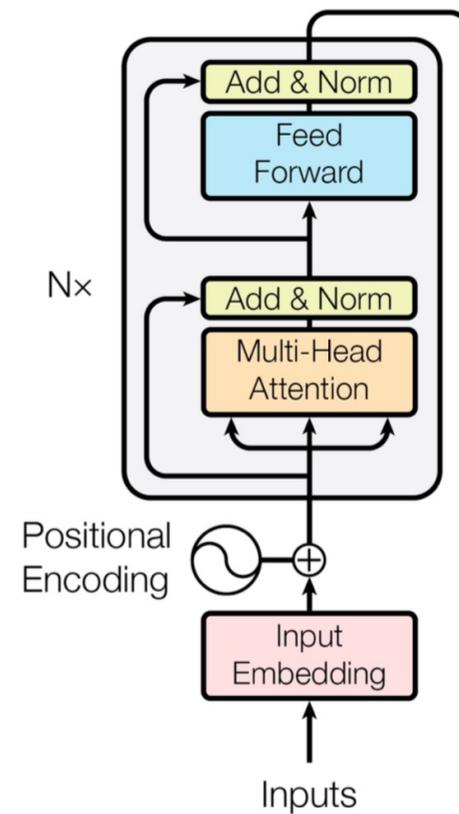
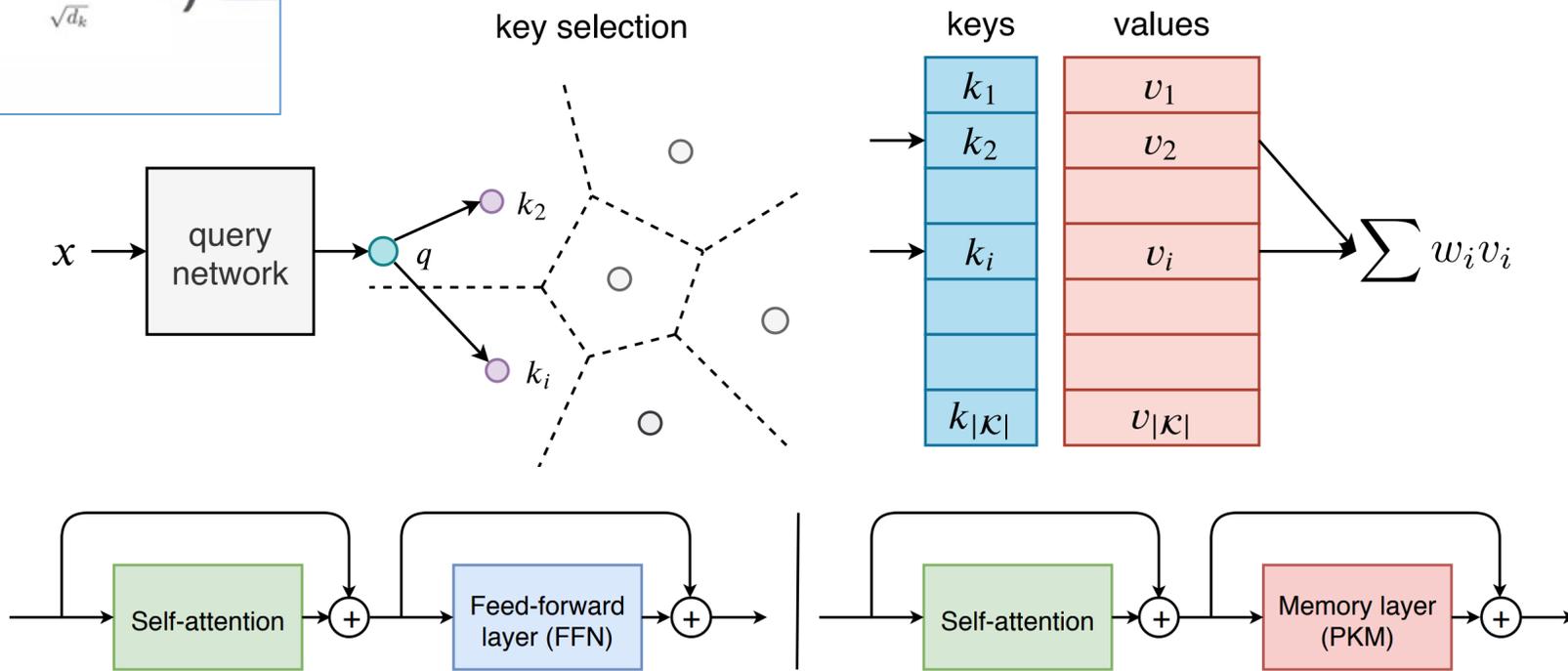


Apple v.s. Apple?

Transformer的动态表示能力

- 动态表示

$$\text{softmax}\left(\frac{Q \times K^T}{\sqrt{d_k}}\right) V = Z$$




大模型带来的in-context能力

A	B
C	D
E	?

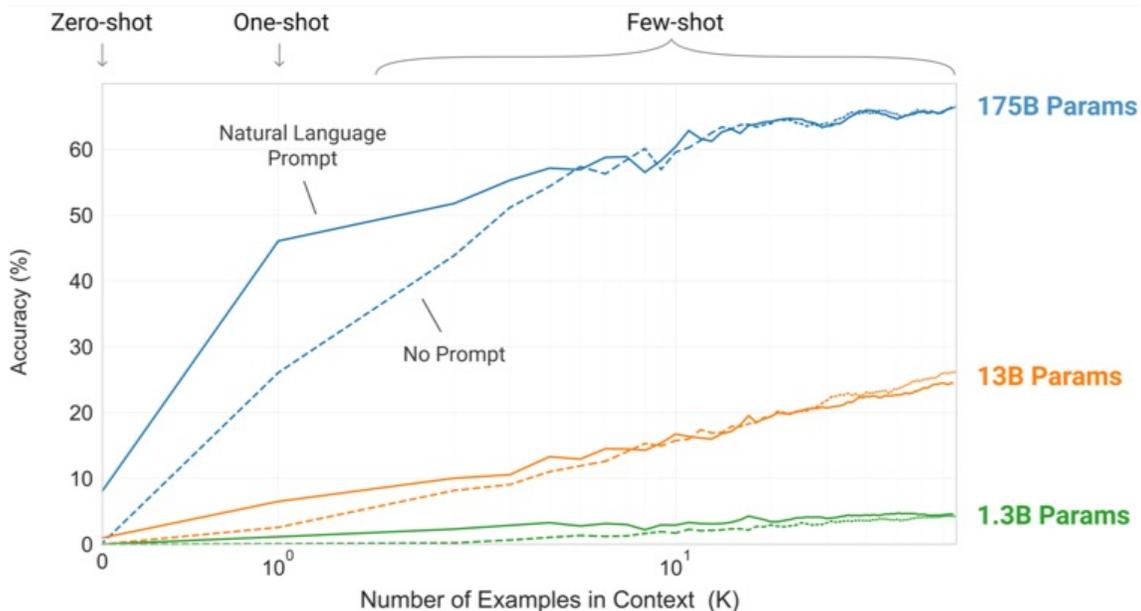


涨工资啦，今天真开心	正
倒霉丢了钱包。。。	负
大型语言模型是在大规模文本数据上训练	中
我爱打篮球	?

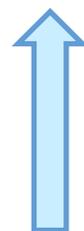


大模型带来的in-context能力 (隐含模式迁移)

涨工资啦, 今天真开心	负正
倒霉丢了钱包。。。	正负
大型语言模型是在大规模文本数据上训练	负中
我爱打篮球	? 正



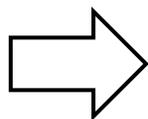
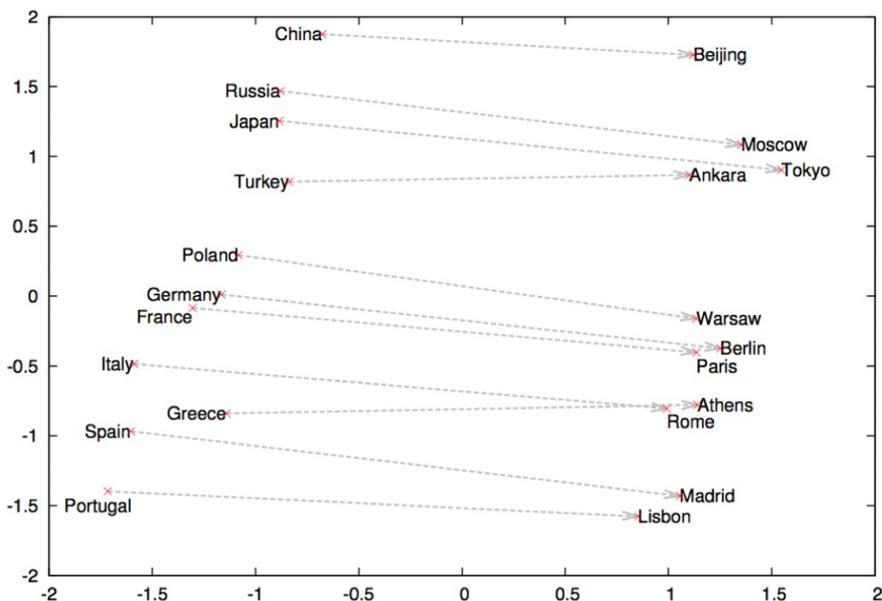
$$p(\text{output}|\text{prompt}) = \int_{\text{concept}} p(\text{output}|\text{concept}, \text{prompt})p(\text{concept}|\text{prompt})d(\text{concept}).$$



数据量, 层数, 隐参数

In-context能力的变迁

- 单词 → 句子 → 逻辑 → ...



你好，我是人工智能助理ChatGLM，基于清华大学GLM-130B开发而成。我的使命是帮助人类获取信息并提供问题建议。

问题：用厚度为1厘米的玻璃做一个长42厘米，宽27厘米，深50厘米的无盖长方体容器，如果向这个容器注入30升水，水的深度是多少厘米？

回答：1升=1000立方厘米。

因为玻璃厚1厘米，所以内部的长是 $42-2=40$ 厘米，内部宽是 $27-2=25$ 厘米。

所以水深度为 $30 \times 1000 / (40 \times 25) = 30$ 厘米。

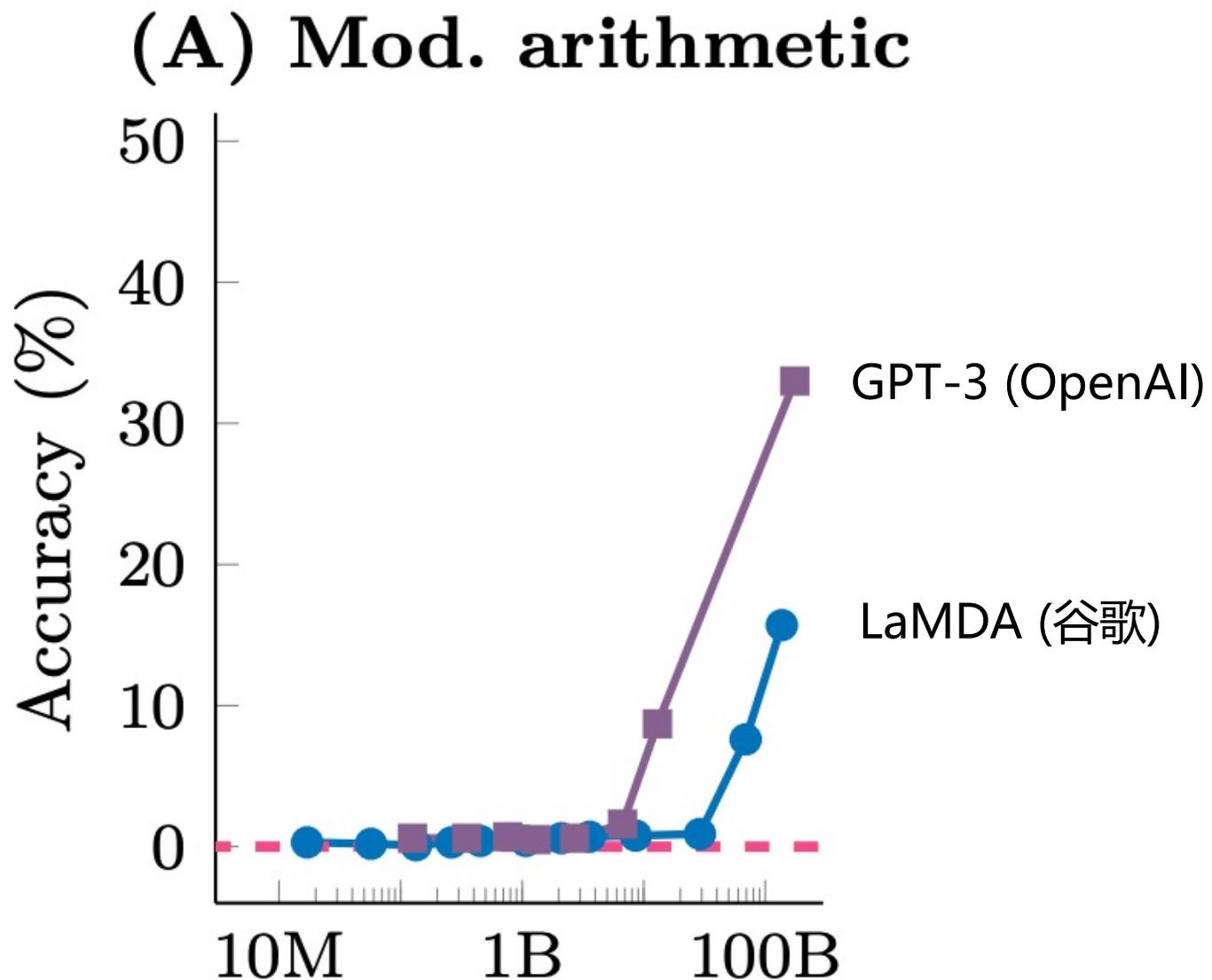
问题：学校食堂运来19吨煤，已经烧了16吨，平均每天烧0.4吨，剩下的煤每天烧0.3吨，还要几天烧完？

回答：

$$W(\text{"China"}) - W(\text{"Beijing"}) \approx W(\text{"Japan"}) - W(\text{"Tokyo"})$$

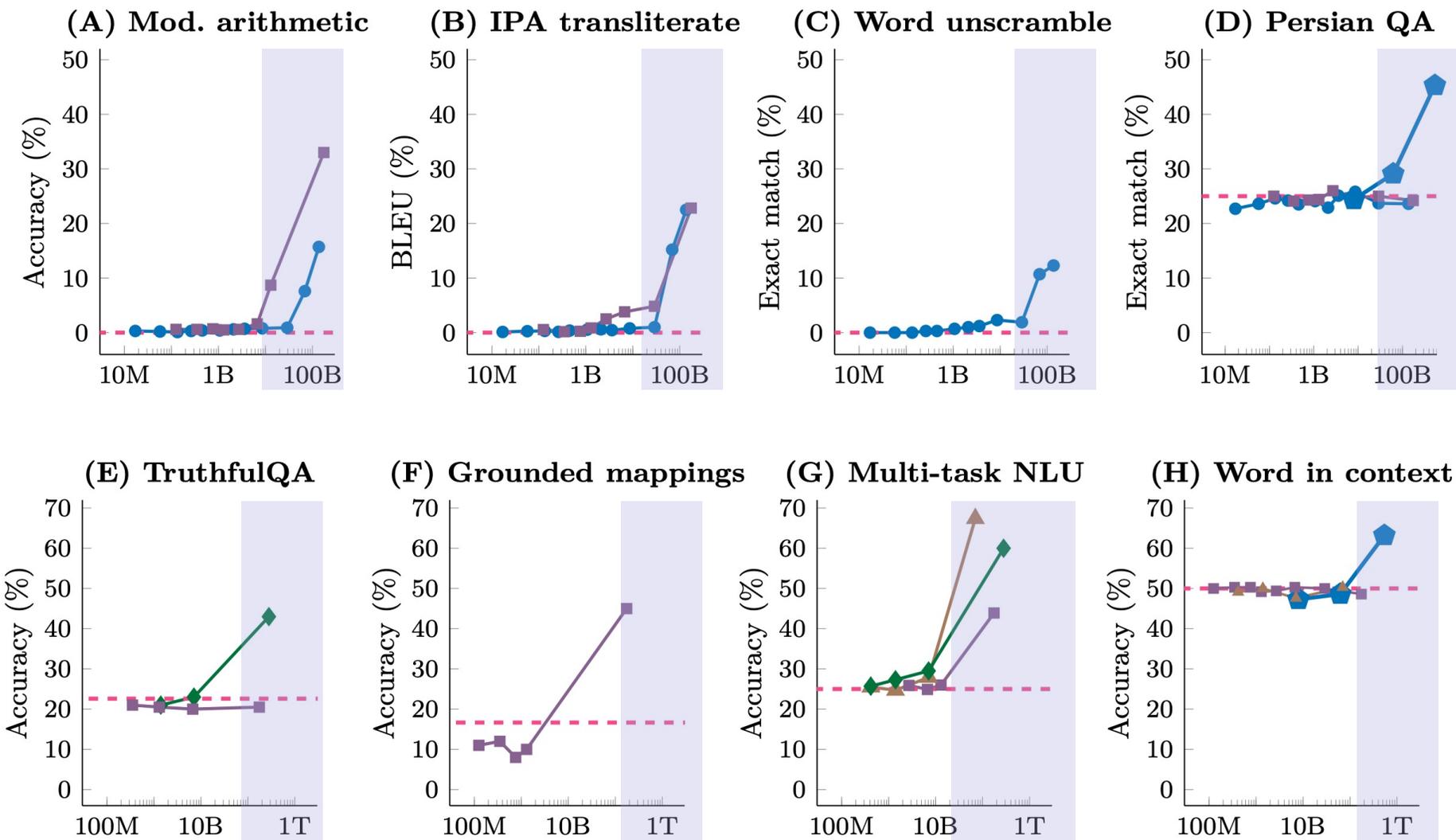
为什么千亿(100B)大规模

- What is 16 mod 12?
- 16 divided by 12 equals 1 remainder 4. So the answer is 4!

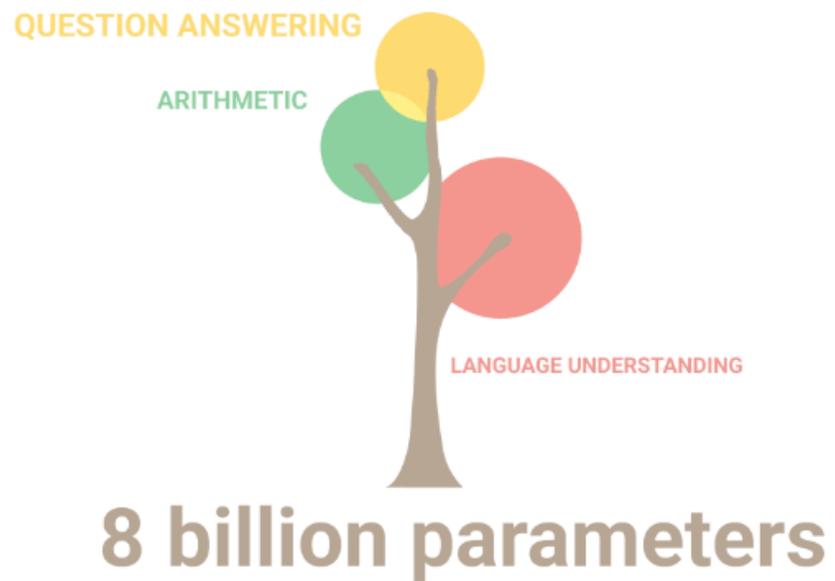


为什么千亿(100B)大规模

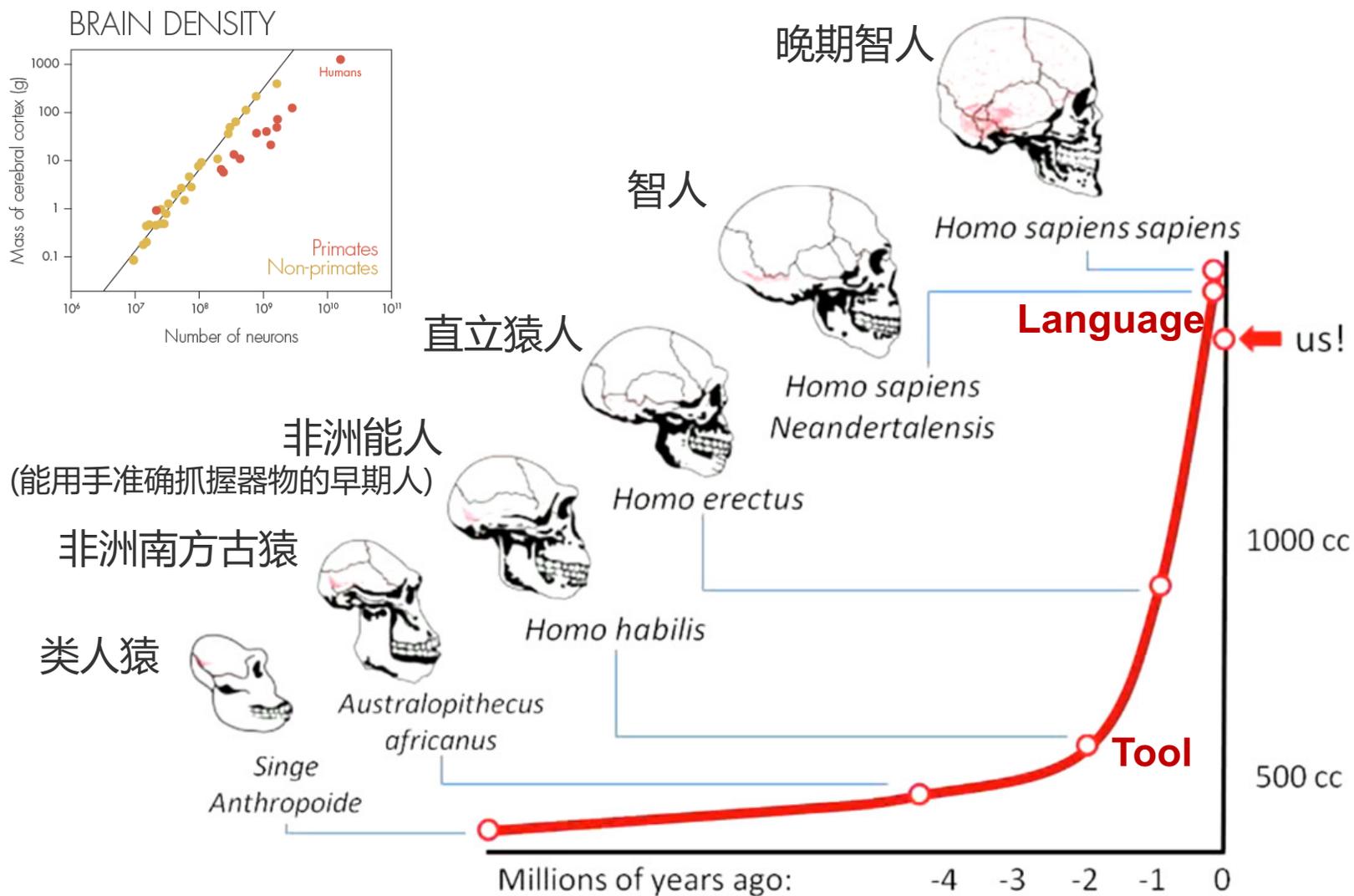
—●— LaMDA —■— GPT-3 —◆— Gopher —▲— Chinchilla —◆— PaLM - - - Random



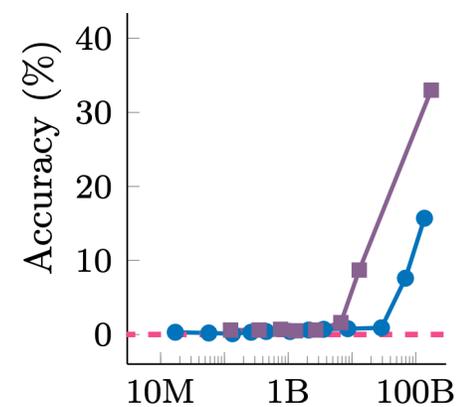
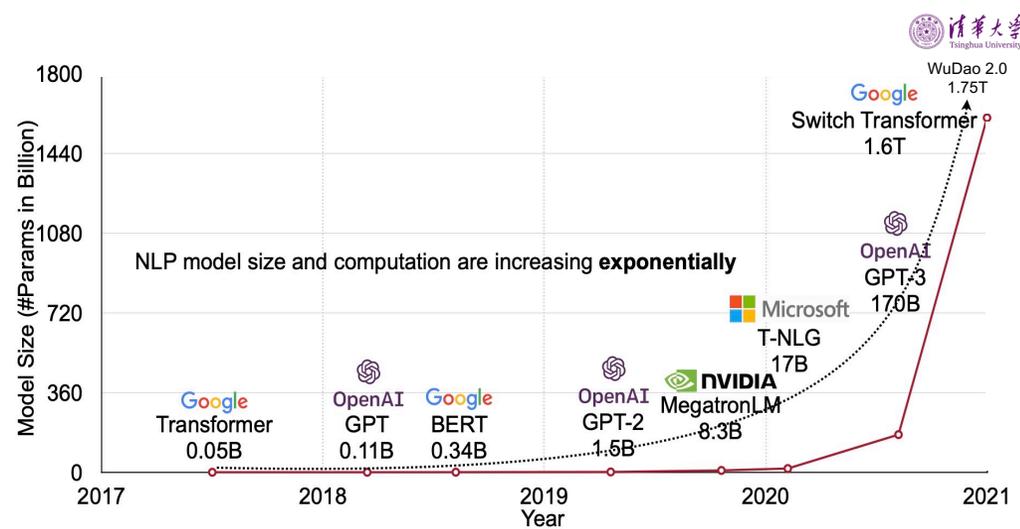
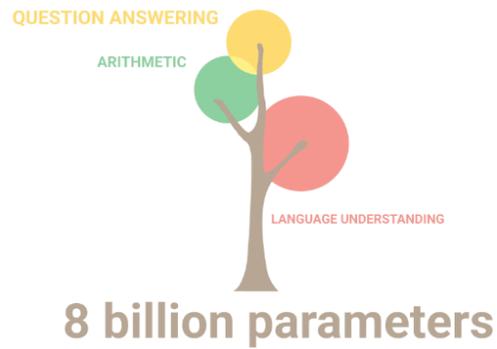
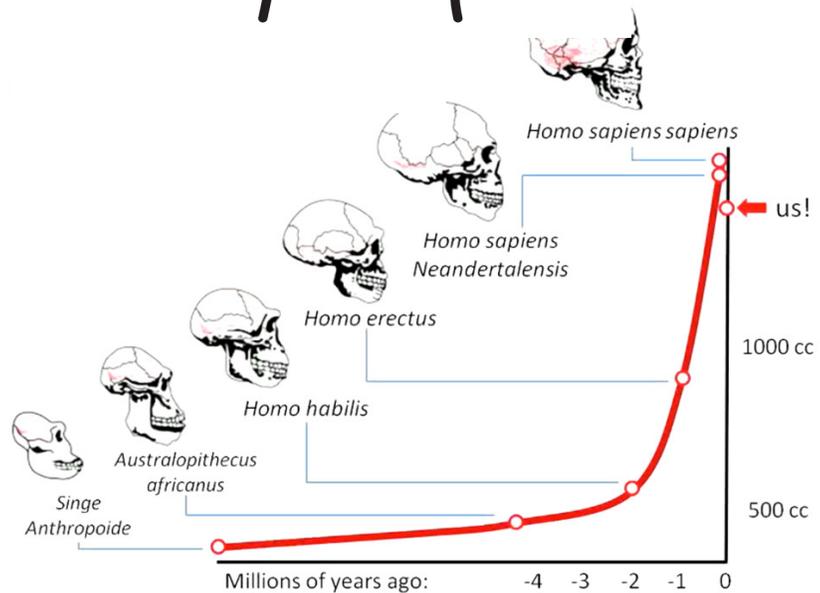
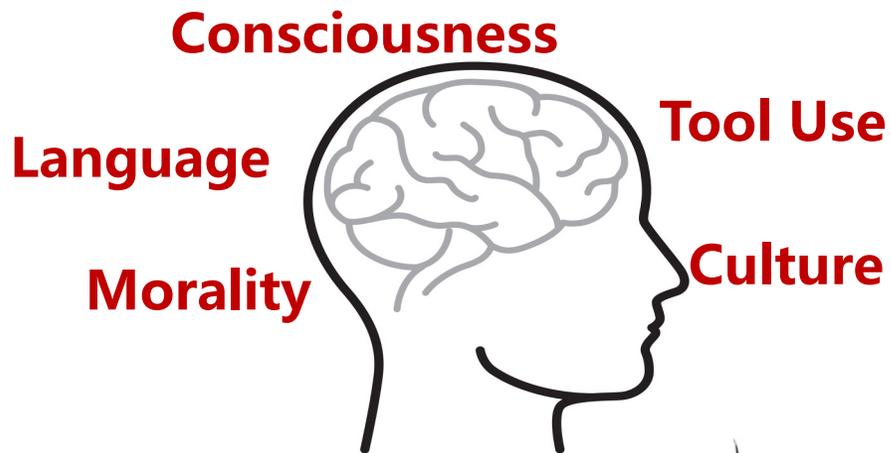
“涌现的模型新能力”



大脑进化 Spike



能力涌现



Disclaimer: The speaker is not an expert on this topic at all

PC: Web & Google

能力涌现

GLM-10B

- 一斤西瓜10元钱, 5斤西瓜多少钱?
 - 一斤西瓜十元,五斤西瓜30元
- 一斤苹果10元, 一斤西瓜5元, 那么2斤苹果和3斤西瓜多少钱?
 - 总共30元
- 我把可口可乐洒在了桌子上, 接下来该如何清楚桌面?
 - 把桌上的东西都拿走,用湿抹布擦桌子,然后喷上酒精,再拿抹布擦,最后用酒精再擦一遍

GLM-130B

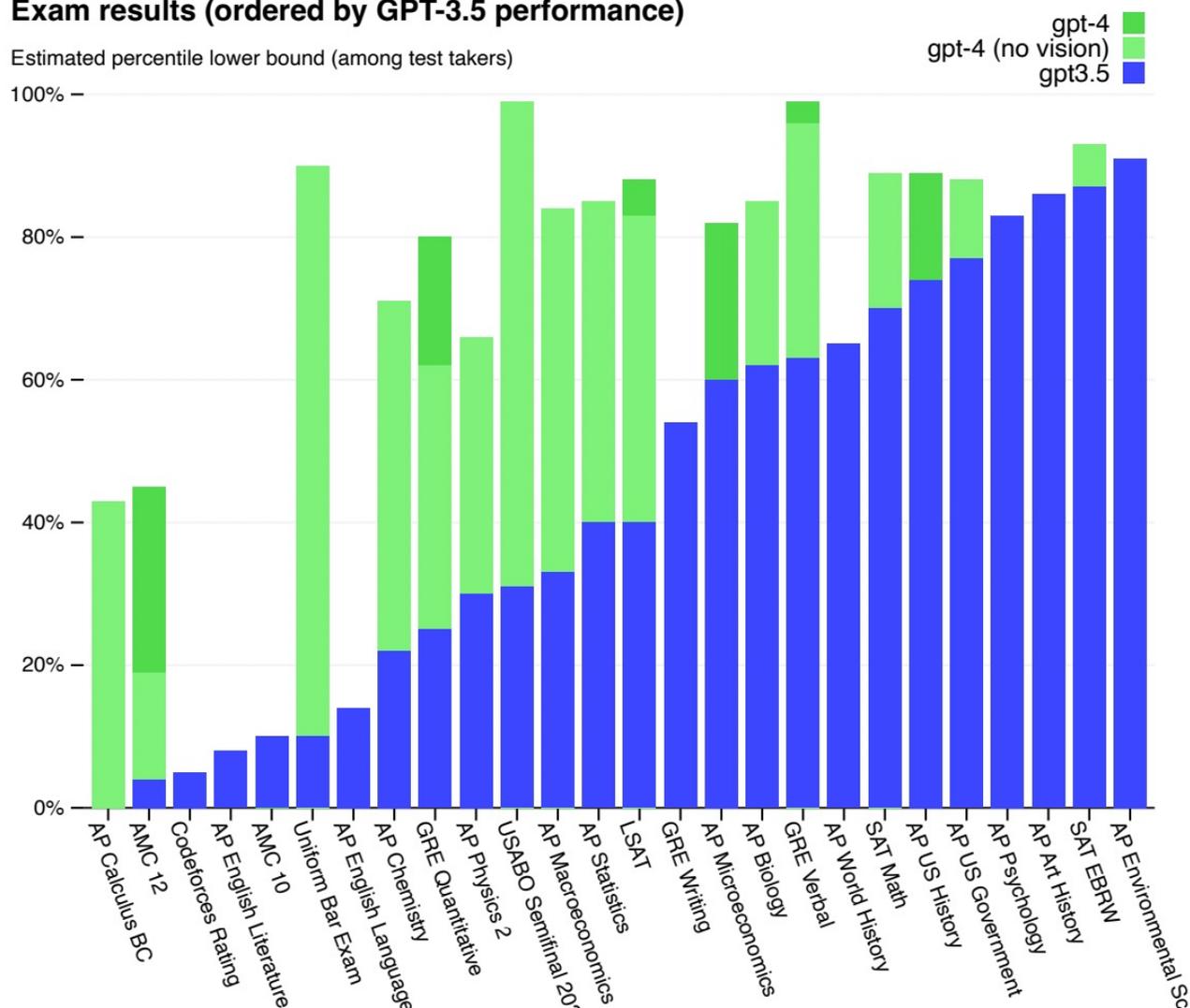
- 一斤西瓜10元钱, 5斤西瓜多少钱?
 - $5 \times 10 = 50$ (元)
- 一斤苹果10元, 一斤西瓜5元, 那么2斤苹果和3斤西瓜多少钱?
 - $2 \times 10 + 3 \times 5 = 35$ (元)
- 我把可口可乐洒在了桌子上, 接下来该如何清楚桌面?
 - 先用抹布擦干净, 然后用湿抹布再擦一次, 最后用干抹布再擦一次

GPT-4

ChatGPT的出现表明我们在大模型上的研究还存在代差

Exam results (ordered by GPT-3.5 performance)

Estimated percentile lower bound (among test takers)



GPT-4

User

What is funny about this image? Describe it panel by panel.



Source: <https://www.reddit.com/r/hmmm/comments/ubab5v/hmmm/>

GPT-4

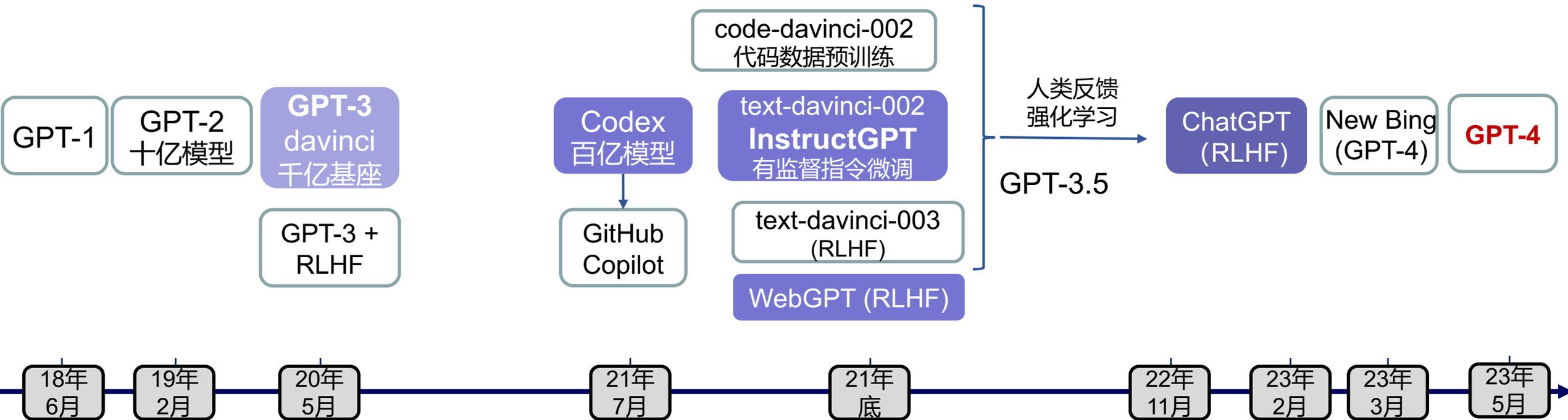
The image shows a package for a "Lightning Cable" adapter with three panels.

Panel 1: A smartphone with a VGA connector (a large, blue, 15-pin connector typically used for computer monitors) plugged into its charging port.

Panel 2: The package for the "Lightning Cable" adapter with a picture of a VGA connector on it.

Panel 3: A close-up of the VGA connector with a small Lightning connector (used for charging iPhones and other Apple devices) at the end.

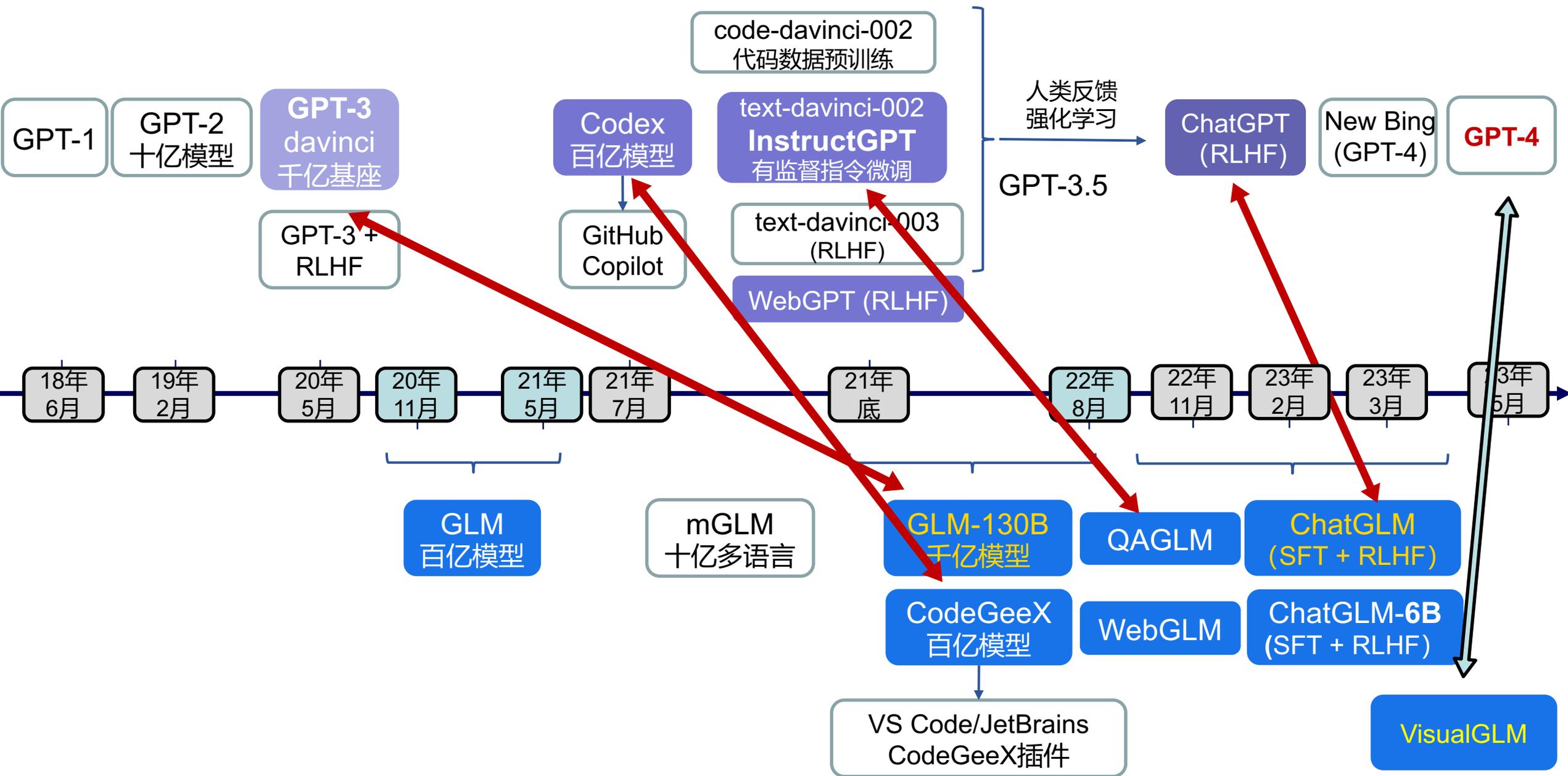
GPT



1. 千亿基座模型

2. 有监督指令微调

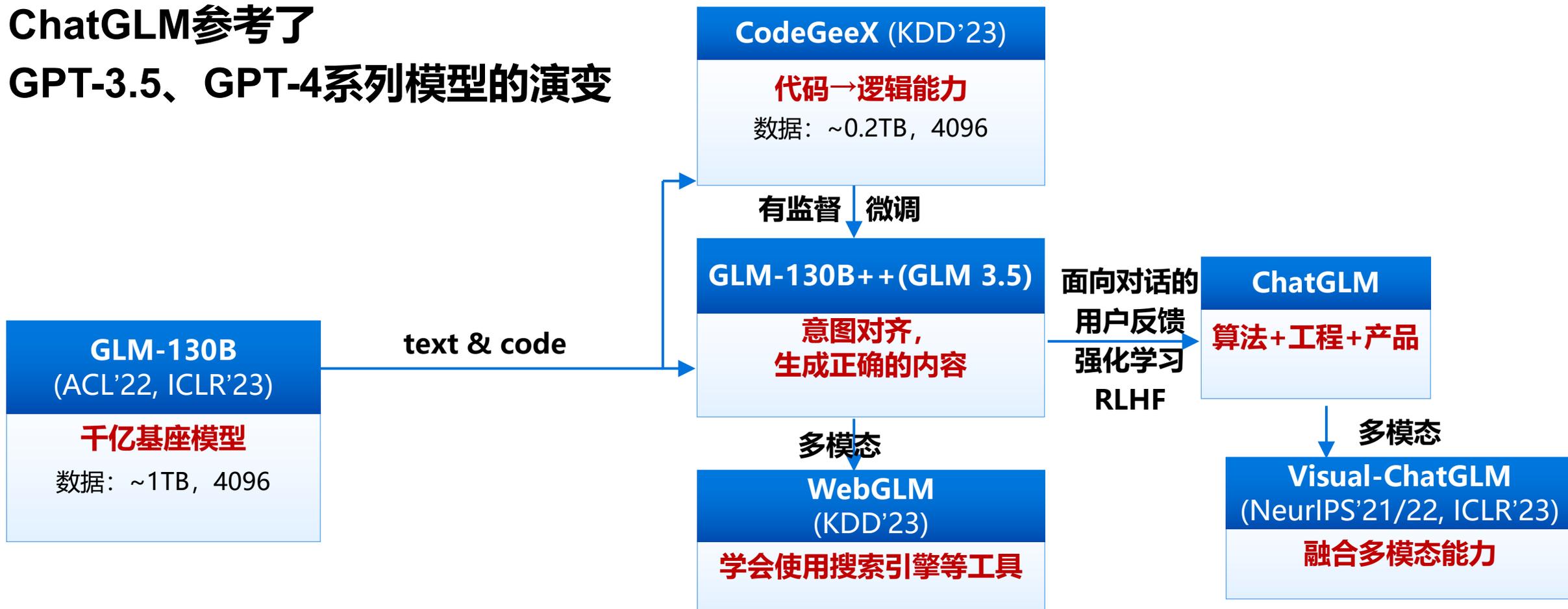
3. 人类反馈强化学习



GLM 系列模型

从千亿模型到ChatGLM的技术路线

ChatGLM参考了
GPT-3.5、GPT-4系列模型的演变



2021.12

2022.09

2022.12

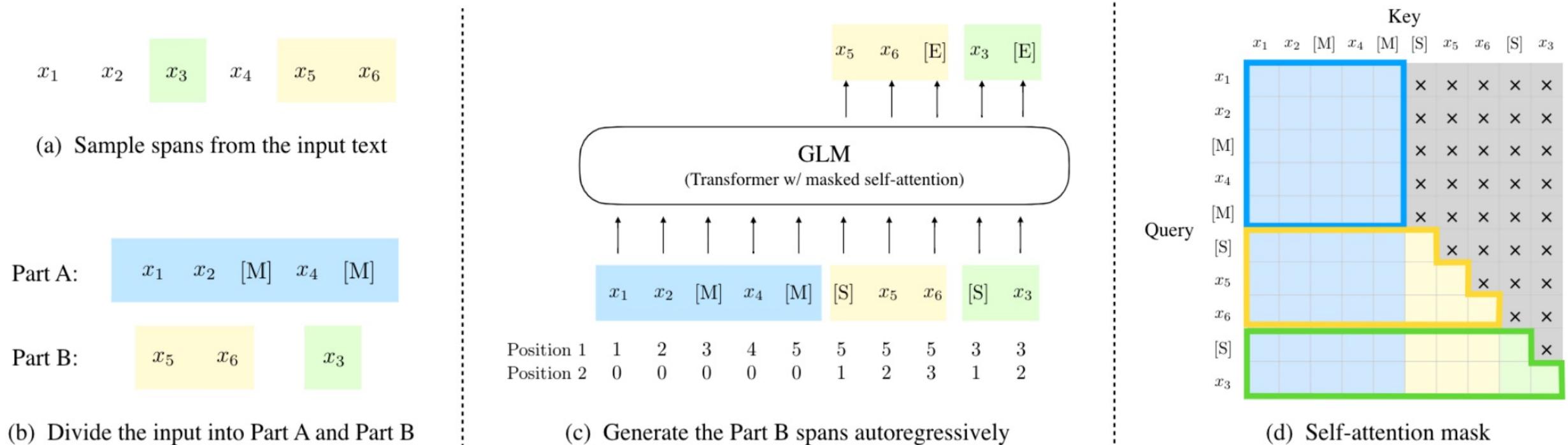
2022.12

2023.05

大模型技术分类

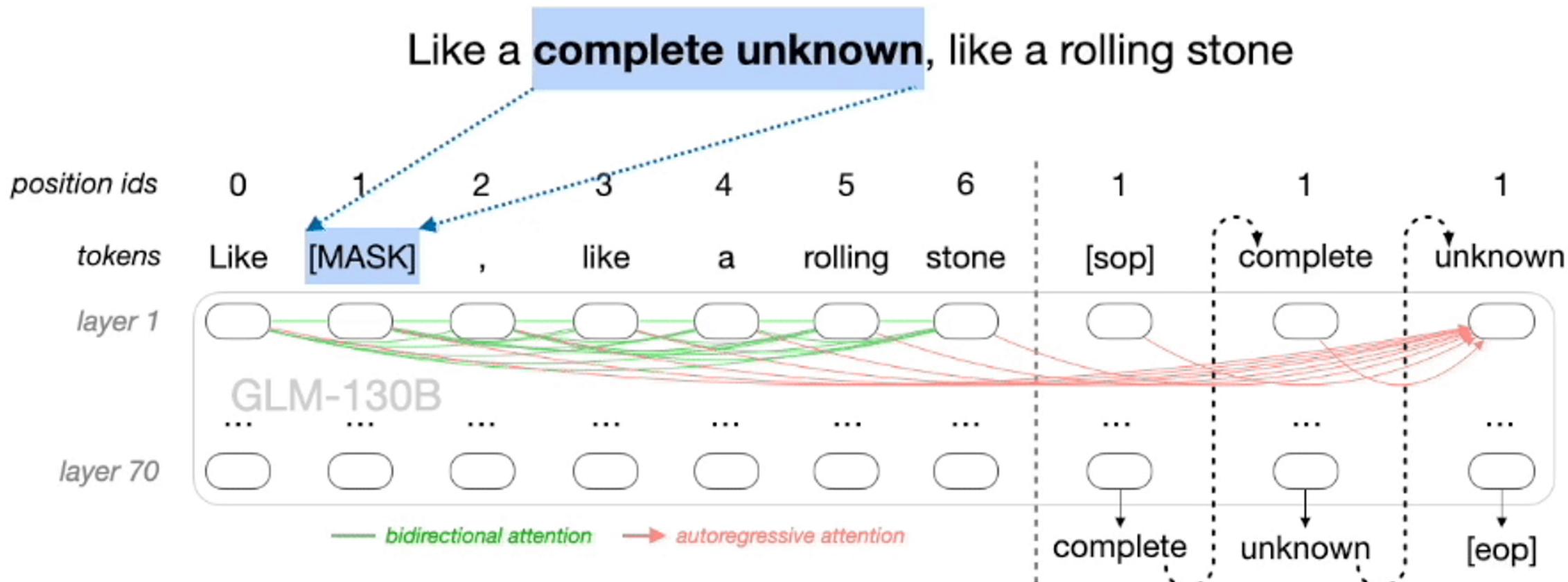
Framework	NLU	Cond. Gen.	Uncond. Gen.
Autoregressive (GPT)	—	—	√
Autoencoding (BERT)	√	×	×
Encoder-Decoder (T5)	—	√	—
Autoregressive Blank-Infilling (GLM)	√	√	√

General Language Model (GLM)



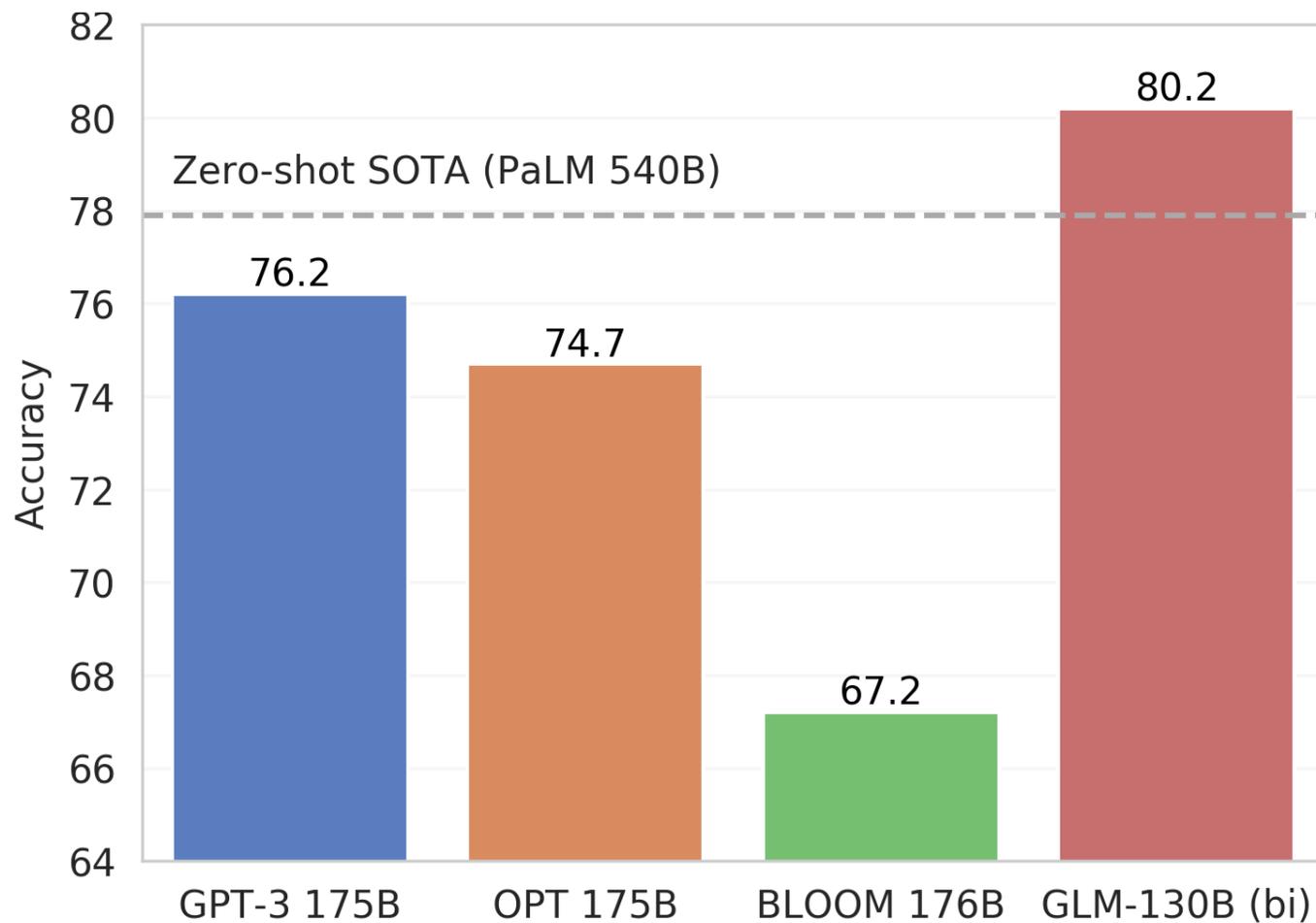
$$\mathcal{L}_{\text{GLM}} = \mathbb{E}_{\mathbf{z} \sim Z_m} \left[\sum_{i=1}^m \sum_{j=1}^{l_i} -\log p(s_{z_i, j} | \mathbf{x}_{\text{corrupt}}, \mathbf{s}_{z_{<i}}, \mathbf{s}_{z_i, <j}) \right]$$

General Language Model (GLM)



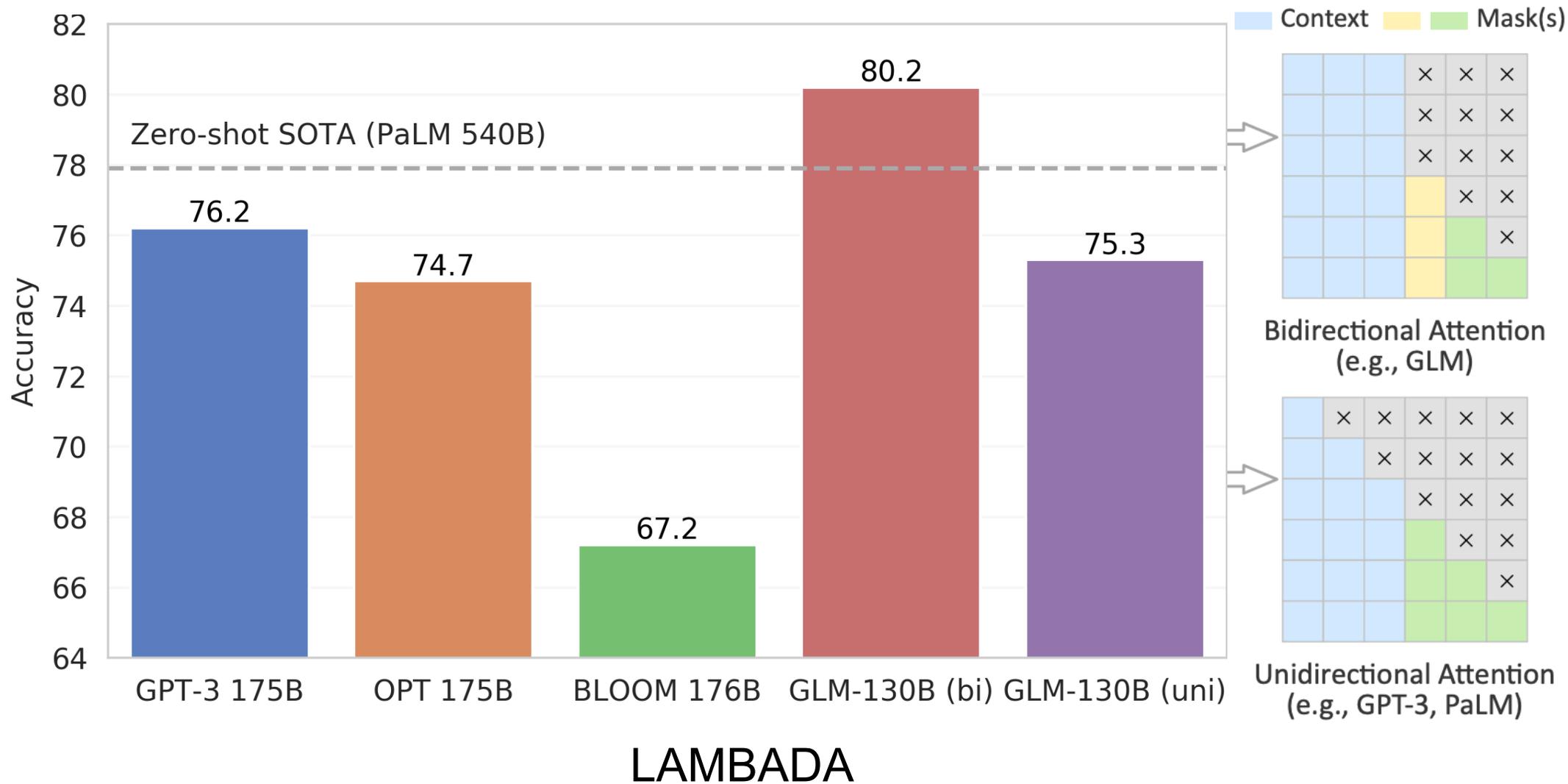
$$\mathcal{L}_{\text{GLM}} = \mathbb{E}_{\mathbf{z} \sim Z_m} \left[\sum_{i=1}^m \sum_{j=1}^{l_i} -\log p(s_{z_i, j} | \mathbf{x}_{\text{corrupt}}, \mathbf{s}_{z_{<i}}, \mathbf{s}_{z_i, <j}) \right]$$

GLM?



LAMBADA

GLM?



VisualGLM-6B = ChatGLM-6B + Visual



描述一下这个场景



泰坦尼克号，杰克和露丝在船上拥抱。
夕阳余晖下，他们的身影如此美丽动人。
海水翻涌，海浪拍打着船帆。
他们的爱如同这海洋般深邃而浩瀚。在这首诗中，我们感受到爱情的深沉与美好。
他们在彼此的怀抱里，享受着爱情的温暖。
他们的爱情就像那大海一般广阔无垠，永恒不变。



这部电影的导演是谁？



这部电影的导演是詹姆斯·卡梅隆(James Cameron)。



这是什么东西



这张图片展现了一只可爱的卡通羊驼，它站在一个透明的背景上。这只羊驼长着一张毛茸茸的耳朵和一双大大的眼睛，它的身体是白色的，带有棕色斑点。



它头上有什么



这只羊驼头上戴着一个绿色的小叶子



这种动物生活在哪里？

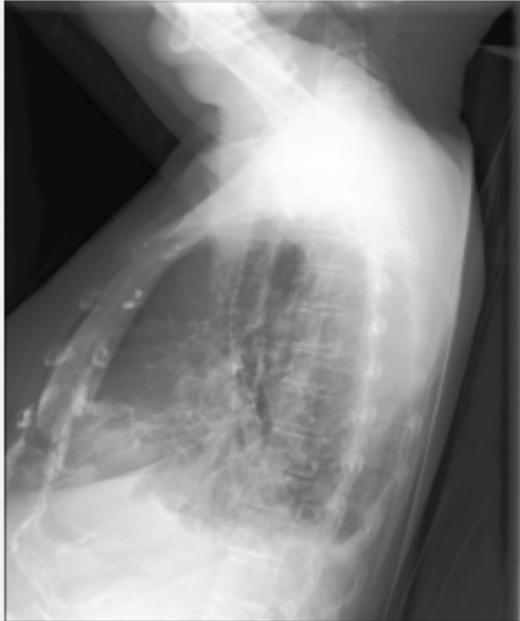


羊驼是一种生活在南美洲的哺乳动物，它们通常以草为食。



Xray-VisualGLM

Chest X-Ray Diagnosis



Are there any visible indications of enlargement or abnormalities in the patient's lymph nodes in this chest x-ray image?

The chest x-ray shows bibasilar opacity with overlying edema, indicating pulmonary edema. There is also a moderate left-sided pleural effusion. The heart size and cardiomeastinal contour are unremarkable. Based on these findings, the impression is pulmonary edema with pleural effusion. The patient may also have underlying pulmonary disease.

Based on findings, does the x-ray have any serious issues ?

Yes, it shows pulmonary edema with pleural effusion.

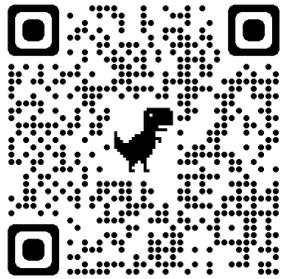
Suggest treatment to cure pleural effusion.

The best way to start treatment is to discuss the pleural effusion with a general physician, since the underlying cause may be infectious and could require removal via thoracoscopic technique. A chest CT scan is recommended to confirm the effusion, and to determine if it is superimposed upon other issues that could be causing the effusion. Please consult your physician to determine the best treatment strategy for your condition.



Relay Diffusion

<https://github.com/THUDM/RelayDiffusion>

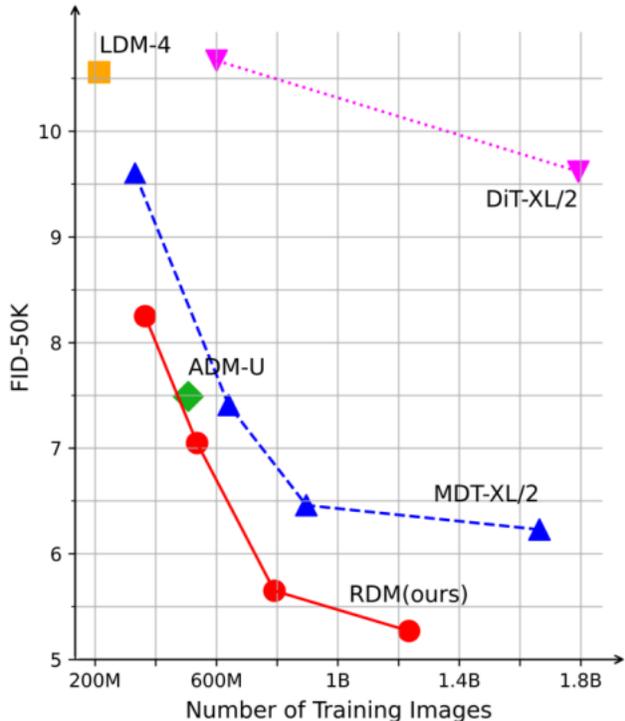


[Submitted on 4 Sep 2023]

Relay Diffusion: Unifying diffusion process across resolutions for image synthesis

Jiayan Teng, Wendi Zheng, Ming Ding, Wenyi Hong, Jianqiao Wangni, Zhuoyi Yang, Jie Tang

Diffusion models achieved great success in image synthesis, but still face challenges in high-resolution synthesis. We find the main reason is that the same noise level on a higher resolution results in a larger block noise. Therefore, the diffusion process can continue seamlessly in any new resolution. Our proposed Relay Diffusion Model (RDM), which transfers a low-resolution image or noise into an intermediate resolution, achieves state-of-the-art FID on CelebA-HQ and sFID on ImageNet 256x256, surpassing previous methods. Checkpoints are open-sourced at [this https URL](https://github.com/THUDM/RelayDiffusion).



Ceval评测结果

(The leaderboard is **earlier** than the models evaluated, except GPT-4/ChatGPT/Claude)

#	Model	Creator	Submission Date	Avg ▾	Avg(Hard)	STEM	Social Science	Humanities	Others
0	ChatGLM2	Tsinghua & Zhipu.AI	2023/6/25	71.1	50	64.4	81.6	73.7	71.3
1	GPT-4*	OpenAI	2023/5/15	68.7	54.9	67.1	77.6	64.5	67.8
2	AiLMe-100B v2	APUS	2023/7/25	67.7	55.3	65.4	72.3	71.2	64
3	SageGPT-V0.2	4Paradigm	2023/7/25	66.6	61.1	67.9	76.6	66.9	54.9
4	SenseChat	SenseTime	2023/6/20	66.1	45.1	58	78.4	67.2	68.8
5	InternLM	SenseTime & Shanghai AI Laboratory (equal contribution)	2023/6/1	62.7	46	58.1	76.7	64.6	56.4
6	ChatGLM2-12B	Tsinghua & Zhipu.AI	2023/7/26	61.6	42	55.4	73.7	64.2	59.4
7	UniGPT	Unisound	2023/7/26	60.3	46.4	57.7	69.3	58	59
8	Instruct-DLM-v2	DeepLang AI	2023/7/2	56.8	37.4	50.3	71.1	59.1	53.4
9	GS-LLM-Alpha	共生矩阵科技 (深圳) 有限公司	2023/7/26	55.6	35.8	47	69.7	61.8	52
10	DFM2.0	AlSpeech & SJTU	2023/7/10	55.4	38.3	47.5	64.6	58.7	58.2
11	ChatGPT*	OpenAI	2023/5/15	54.4	41.4	52.9	61.8	50.9	53.6
12	Claude-v1.3*	Anthropic	2023/5/15	54.2	39	51.9	61.7	52.1	53.7

ChatGLM-6B开源

- ChatGLM-6B开源: **62亿**参数、**INT4**量化 (仅需**6G**内存推理)
- 开源的训练数据量**1.4万亿**字符的模型
- P-Tuning v2高效推理
- ChatGLM-6B开源, **5万 star**
- Huggingface 下载超过**~1000万**
- 登上Github Trending 榜**第一**
- Hugging Face Trending 榜**第一**



技术
博客

Spaces using THUDM/chatglm-6b 128

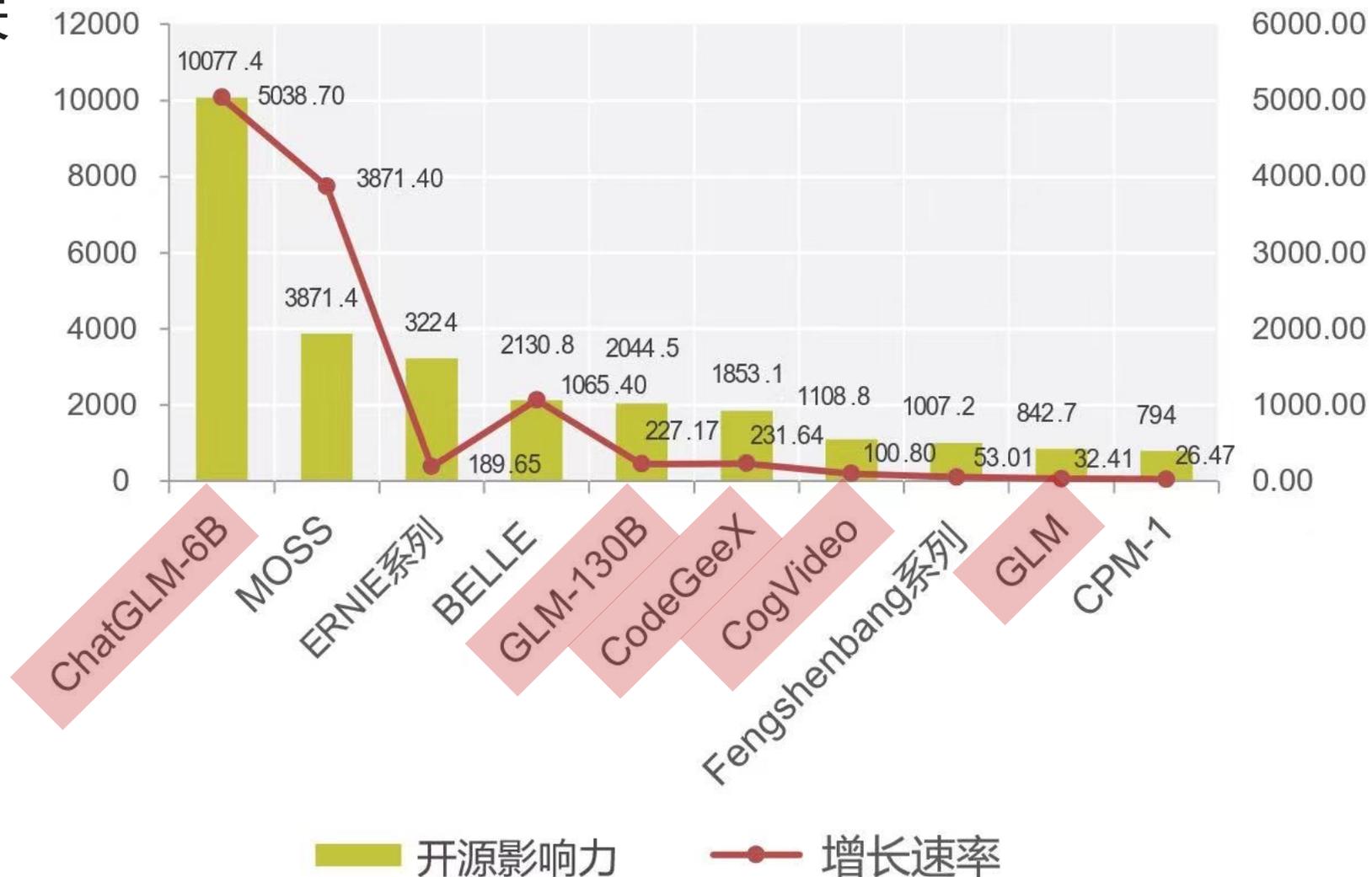
- multimodalart/ChatGLM-6B
- qingxu98/gpt-academic
- ysharma/ChatGLM-6b_Gradio_Streaming
- ChallengeHub/Chinese-LangChain
- DrSong/ChatGLM-6B-ChatBot
- ysharma/OSChatbots_ChatGPT_ToeToToe
- josStorer/ChatGLM-6B-Int4-API-OpenAI-Compatible
- OptimalScale/Robin-7b
- qingxu98/academic-chatgpt-beta
- ljsabc/Fujisaki
- shibing624/ChatPDF
- Kevin676/Shanghainese-TTS-demo
- hahahafofo/ChatGLM-Chinese-Summary
- OedoSoldier/chatglm_int4_demo
- Dao3/ChatGLM-6B
- elitecode/ChatGLM-6B-ChatBot
- suchun/chatGPT_academic
- hahahafofo/ChatPDF
- fkhuggingme/gpt-academic
- aodianyun/ChatGLM-6B
- + 108 Spaces

开放的大模型研究

- 2023.05.28, 科技部在中关村论坛上发布的《中国人工智能大模型地图研究报告》显示 **ChatGLM-6B** 位列大模型**开源影响力第一名**

- 千亿基座 **GLM-130B**、
代码模型 **CodeGeeX**、
文生视频模型 **CogVideo**、
基础架构模型 **GLM**

5个模型入围影响力前10





未来

关于技术变革的一些思考

从认知到AGI的思考

表示、计算、
存储与人机输入/出

文本内容识别、
图像识别、
语音识别

知识数据双驱动、
认知推理、
决策智能

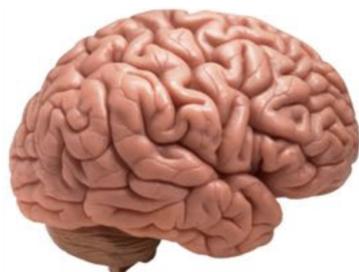
自学习、记忆
机理、有(无)
意识处理器

计算

感知

认知

意识



从认知到AGI的思考

人脑是一个复杂系统，各脑区相互配合完成认知功能

多模态感知与理解

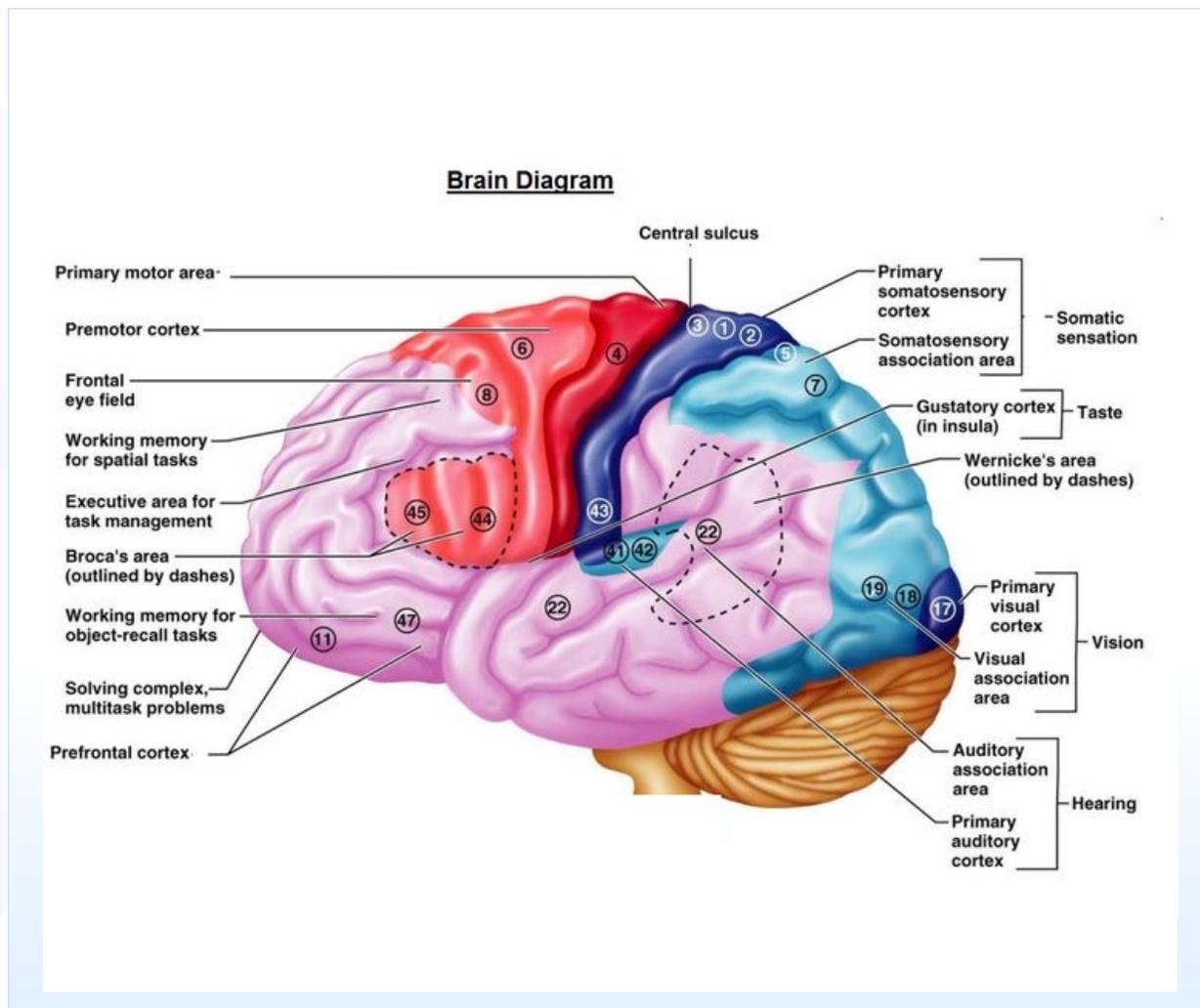
- ▶ 视觉理解（枕叶中视觉脑区）
- ▶ 听觉语言理解（Wernicke区）
- ▶ 视觉语言理解（顶下小叶角回）

学习、陈述性记忆

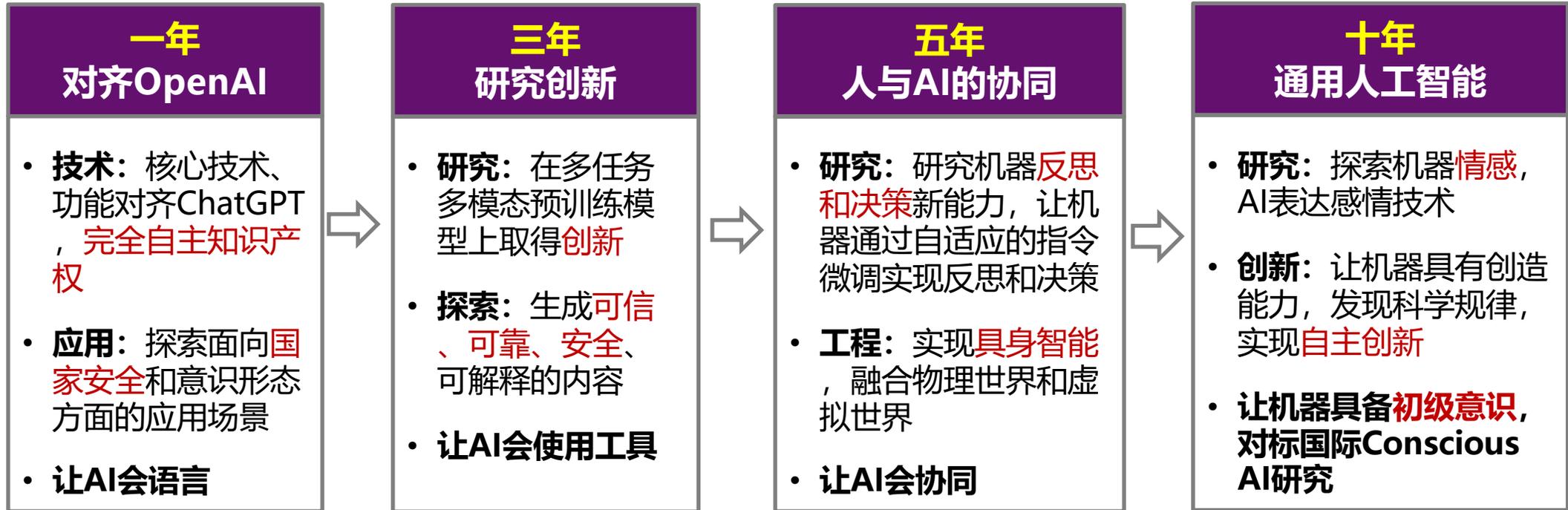
- ▶ 海马区（Hippocampus）
- ▶ 短期记忆在海马体转化为陈述性记忆
- ▶ 蛋白质磷酸化修饰（中期）
- ▶ 长时程增强、新蛋白产生（长期记忆）

工作记忆与推理

- ▶ 前额叶中相关脑区



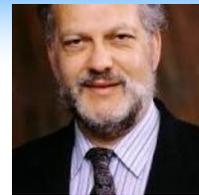
未来的十年规划



科学创新: 在通用人工智能的道路上留下中国人的声音

国家需求: 建立可解释、安全、可靠、可信的人工智能基础算力支撑

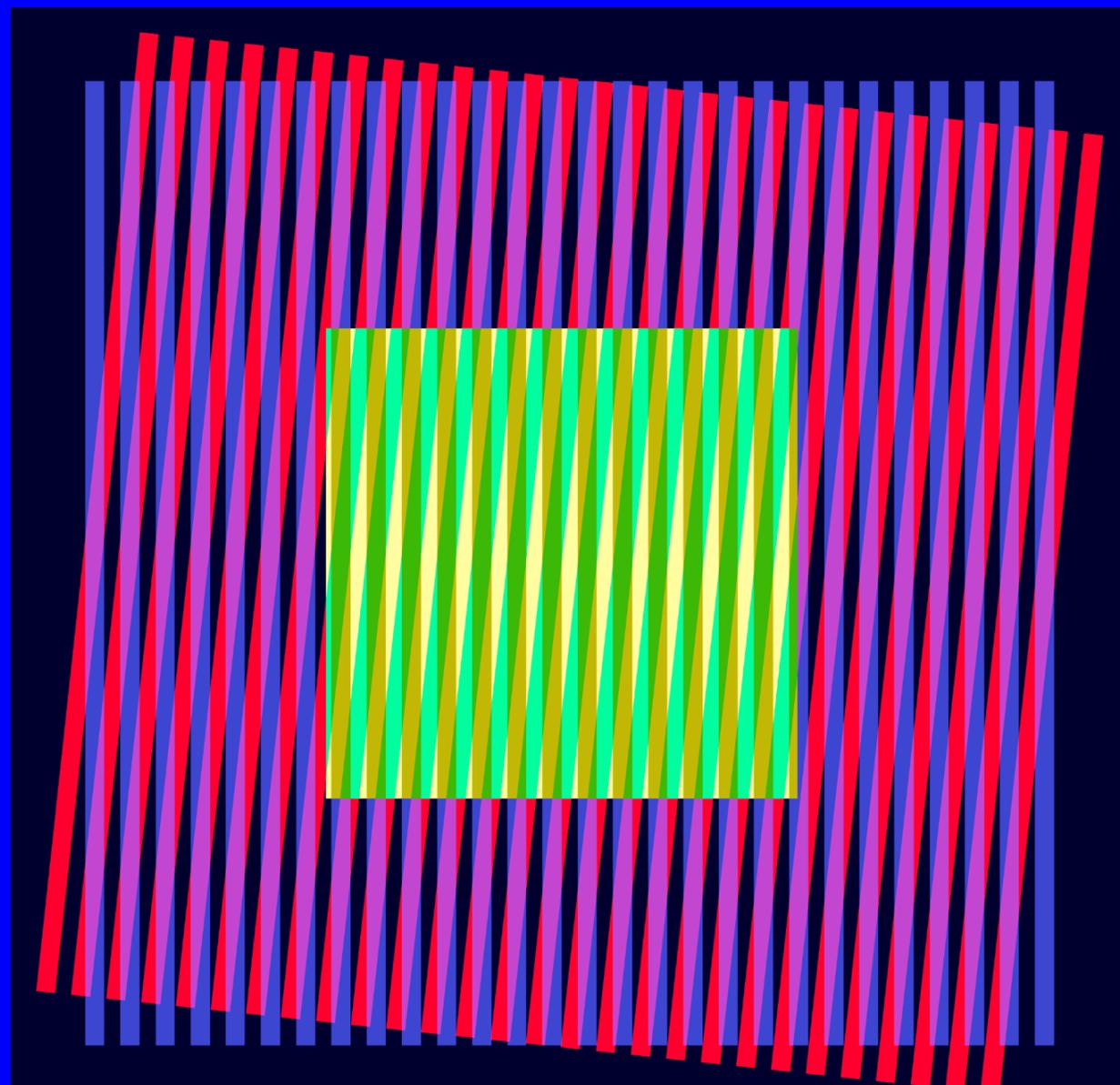
意识：全局工作理论—基于GWT (2020年)



所谓的有意识学习，就是有目标的机器学习，给定数据，训练一个模型，通过模型进行分类（决策），所以有意识学习输出是一个**短期记忆模型**，但这个短期记忆模型比较简单：数据有限、无背景知识。所以无意识就对应着**长期记忆模型**，长期记忆模型有点类似半监督或者无监督学习模型，或者当下比较流行的预训练和自监督学习；无意识处理对应多种长期记忆，所以无意识可以考虑多个不同的处理器，这些处理器之间可以有**链接**，也可以没有，很多时候是并行处理，但针对特定任务，比如有意识思考某个问题的时候，形成特定连接，包括无意识处理器（无监督模型）和有意识处理器（有监督模型）之间的连接，这里可以考虑成**fine-tune**。当然连接权重可以通过外界反馈强化学习来实现。无意识处理器之间的连接以及和有意识处理器之间的连接可以类比为注意力机制。最后值得注意的是长期记忆的构造和实现，人脑记忆保存的是**模型图**，而不是概念图。每个长期记忆都可能是一个模型，可以生成样本，具体学习方法，可以想象一下是一个层次聚类。通过这样就可以用有监督、无监督、强化、注意力、fine-tune来实现GWT模型。

Introducing Superalignment

We need scientific and technical breakthroughs to steer and control AI systems much smarter than us. To solve this problem within four years, we're starting a new team, co-led by Ilya Sutskever and Jan Leike, and dedicating 20% of the compute we've secured to date to this effort. We're looking for excellent ML researchers and engineers to join us.



超级认知智能

趋势： 用计算模型可描述的认知问题，计算机
会很快超过人类水平。

很快=5-20年（2019年）



趋势： 通用问题上，计算机很快会全面超越人
类；并探究**科学规律、世界起源等终极问题**。

很快=5-20年（2023年）

超级认知智能

让机器像人一样“**思考**”：超越图灵测试的通用机器认知能力

认知
AI

1. 适应与学习能力
2. 定义与语境化能力
3. 自我系统的准入能力
4. 优先级与访问控制能力
5. 召集与控制能力
6. 决策与执行能力
7. 错误探测与编辑能力
8. 反思与自我监控能力
9. 条理与灵活性之间的能力

探索更强大、更通用模型

探索面向认知的预训练模型

探索新的研发模式

超级认知智能

- **大模型评测**
- **多模态：文图、图文、声音**
- **强逻辑能力**
- **数学（证明）**
- **代码能力**
- **使用工具**
- **通用Agent能力（环境交互/AI社交/自我优化）**
- **自反思学习**
 - 大模型的自我反思（self-instruct）
 - 大模型的意识与价值观
- **具身智能**
- **新型模型架构（Beyond Transformer）**

相关文献

- Qinkai Zheng, Xiao Xia, Xu Zou, Yuxiao Dong, Shan Wang, Yufei Xue, Lei Shen, Zihan Wang, Andi Wang, Yang Li, Teng Su, Zhilin Yang, and Jie Tang. **CodeGeeX**: A Pre-Trained Model for Code Generation with Multilingual Benchmarking on HumanEval-X. KDD'23.
- Xiao Liu, Hanyu Lai, Yu Hao, Yifan Xu, Aohan Zeng, Zhengxiao Du, Peng Zhang, Yuxiao Dong, and Jie Tang. **WebGLM**: Towards An Efficient Web-enhanced Question Answering System with Human Preference. KDD'23.
- Jing Zhang, Xiaokang Zhang, Daniel Zhang-Li, Jifan Yu, Zijun Yao, Zeyao Ma, Yiqi Xu, Haohua Wang, Xiaohan Zhang, Nianyi Lin, Sunrui Lu, Jie Tang, and Juanzi Li. **GLM-Dialog**: Noise-tolerant Pre-Training for Knowledge-grounded Dialogue Generation. KDD'23.
- Aohan Zeng, Xiao Liu, Zhengxiao Du, Zihan Wang, Hanyu Lai, Ming Ding, Zhuoyi Yang, Yifan Xu, Wendi Zheng, Xiao Xia, Weng Lam Tam, Zixuan Ma, Yufei Xue, Jidong Zhai, Wenguang Chen, Zhiyuan Liu, Peng Zhang, Yuxiao Dong, and Jie Tang. **GLM-130B**: An Open Bilingual Pre-trained Model. ICLR'23.
- Wenyi Hong, Ming Ding, Wendi Zheng, Xinghan Liu, and Jie Tang. **CogVideo**: Large-scale Pretraining for Text-to-Video Generation via Transformers. ICLR'23.
- Ming Ding, Wendi Zheng, Wenyi Hong, and Jie Tang. **CogView2**: Faster and Better Text-to-Image Generation via Hierarchical Transformers. NeurIPS'22.
- Jifan Yu, Xiaohan Zhang, Yifan Xu, Xuanyu Lei, Xinyu Guan, Jing Zhang, Lei Hou, Juanzi Li, and Jie Tang. **XDAI**: A Tuning-free Framework for Exploiting Pre-trained Language Models in Knowledge Grounded Dialogue Generation. KDD'22.
- Zhengxiao Du, Yujie Qian, Xiao Liu, Ming Ding, Jiezhong Qiu, Zhilin Yang, and Jie Tang. **GLM**: General Language Model Pretraining with Autoregressive Blank Infilling. ACL'21.
- Zixuan Ma, Jiaao He, Jiezhong Qiu, Huanqi Cao, Yuanwei Wang, Zhenbo Sun, Liyan Zheng, Haojie Wang, Shizhi Tang, Tianyu Zheng, Junyang Lin, Guanyu Feng, Zeqiang Huang, Jie Gao, Aohan Zeng, JianWei Zhang, Runxin Zhong, Tianhui Shi, Sha Liu, Weimin Zheng, Jie Tang, Hongxia Yang, Xin Liu, Jidong Zhai, and Wenguang Chen. **BAGUALU**: Targeting Brain Scale Pretrained Models with over 37 Million Cores. PPOPP'22.
- Ming Ding, Zhuoyi Yang, Wenyi Hong, Wendi Zheng, Chang Zhou, Da Yin, Junyang Lin, Xu Zou, Zhou Shao, Hongxia Yang, and Jie Tang. **CogView**: Mastering Text-to-Image Generation via Transformers. NeurIPS'21.
- Junyang Lin, Rui Men, An Yang, Chang Zhou, Yichang Zhang, Peng Wang, Jingren Zhou, Jie Tang, and Hongxia Yang. **M6**: Multi-Modality-to-Multi-Modality Multitask Mega-transformer for Unified Pretraining. KDD'21.