ChatGPT

从语言知识到知识库

李云聪 (yuncongli)

2023/03/16

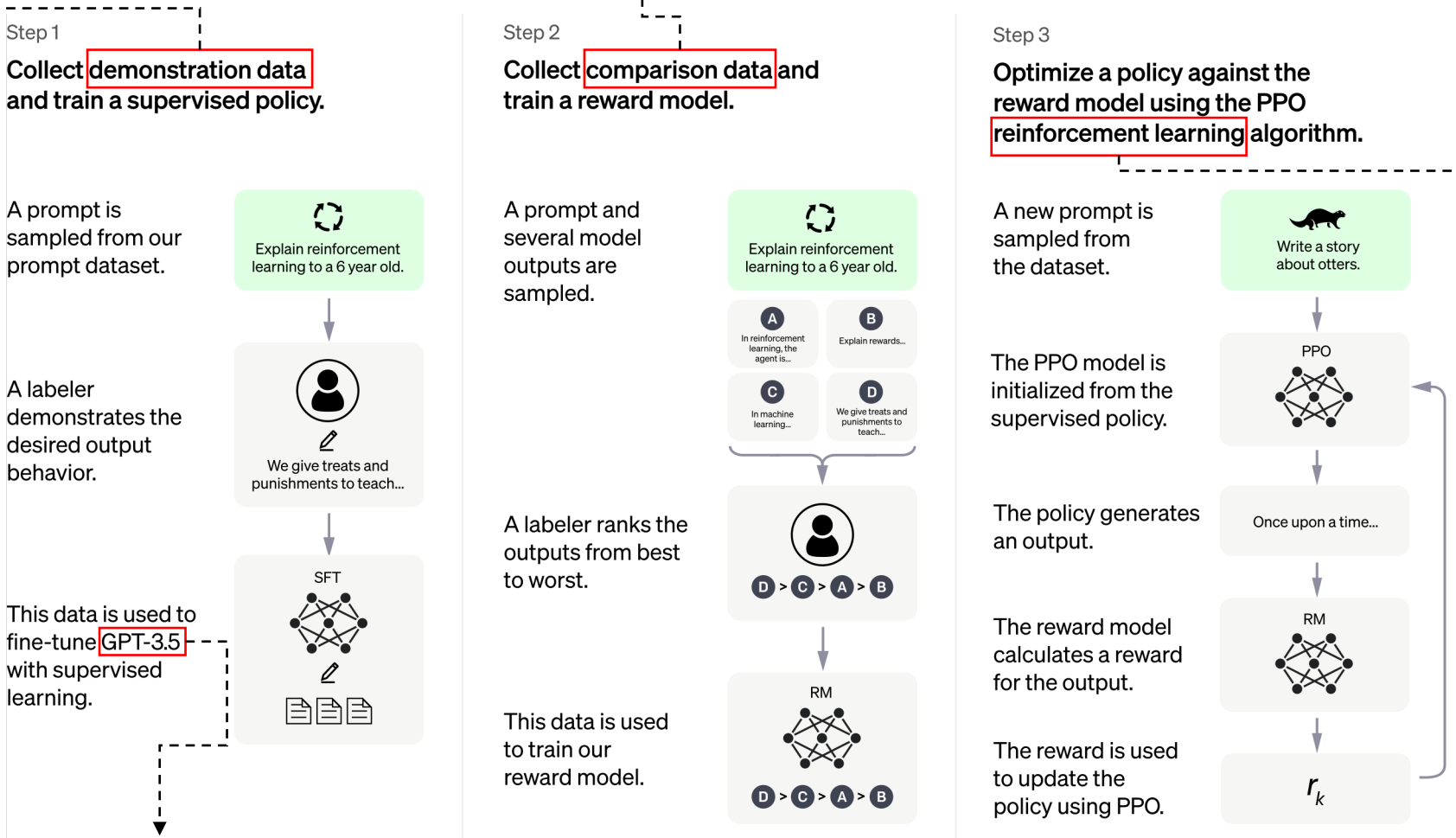
Outline

- ChatGPT是什么
- 语言模型 (Language Models, LM)
- 语言模型是知识库
- 提示学习 (Prompt-Learning)
- 上下文学习 (In-Context Learning)
- 精调语言模型 (ChatGPT)
- ChatGPT评估
- 一些资源

ChatGPT: 通过人类反馈两阶段精调的语言模型

第二阶段精调用的人类反馈

第一阶段精调用的人类反馈



精调使用的方法是强化学习

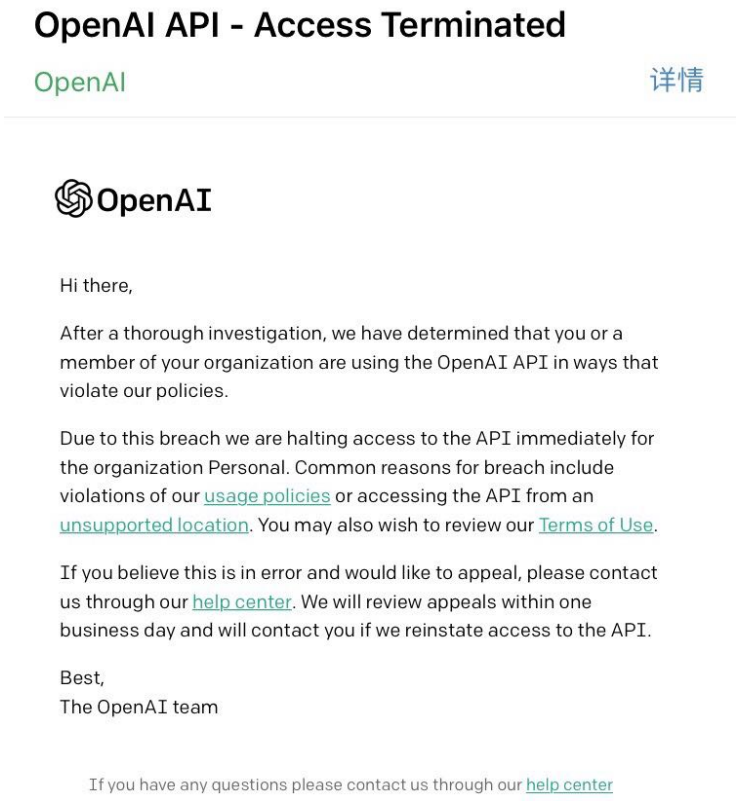
GPT-3.5是预训练好的语言模型

术语中英对照：

- 预训练 Pre-train
- 精调或微调 Fine-tune

试用ChatGPT

- Fanqiang
  - 注册
  - VPN
- 网址
  - 需要账号
  - 需要VPN
- API
  - 需要账号
  - 不需要VPN可以访问，但可能被封号



OpenAI API - Access Terminated

OpenAI 详情

OpenAI

Hi there,

After a thorough investigation, we have determined that you or a member of your organization are using the OpenAI API in ways that violate our policies.

Due to this breach we are halting access to the API immediately for the organization Personal. Common reasons for breach include violations of our usage policies or accessing the API from an unsupported location. You may also wish to review our Terms of Use.

If you believe this is in error and would like to appeal, please contact us through our help center. We will review appeals within one business day and will contact you if we reinstate access to the API.

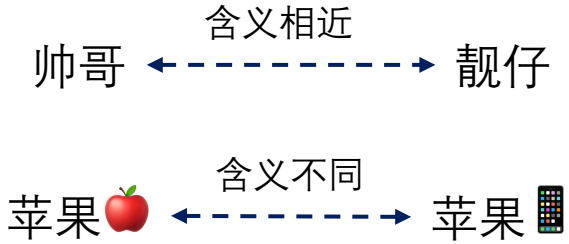
Best,
The OpenAI team

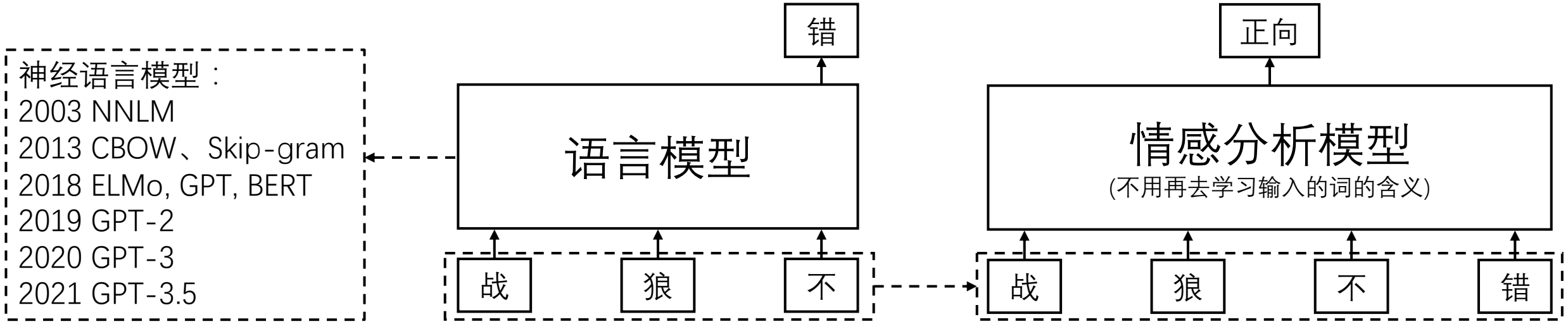
If you have any questions please contact us through our help center

语言模型

- 语言模型预测一串词是合理的句子的概率
  - $p(\text{战狼不错}) = p(\text{战})p(\text{狼}|\text{战})p(\text{不}|\text{战狼})p(\text{错}|\text{战狼不})$
- 语言模型训练：不需要人工标注数据
  - 以之前的字为特征，预测当前字的概率，即$p(\text{错}|\text{战狼不})$
  - 学习文本中语言知识，用于降低其它NLP任务的标注成本

- 语言知识是词义和其它

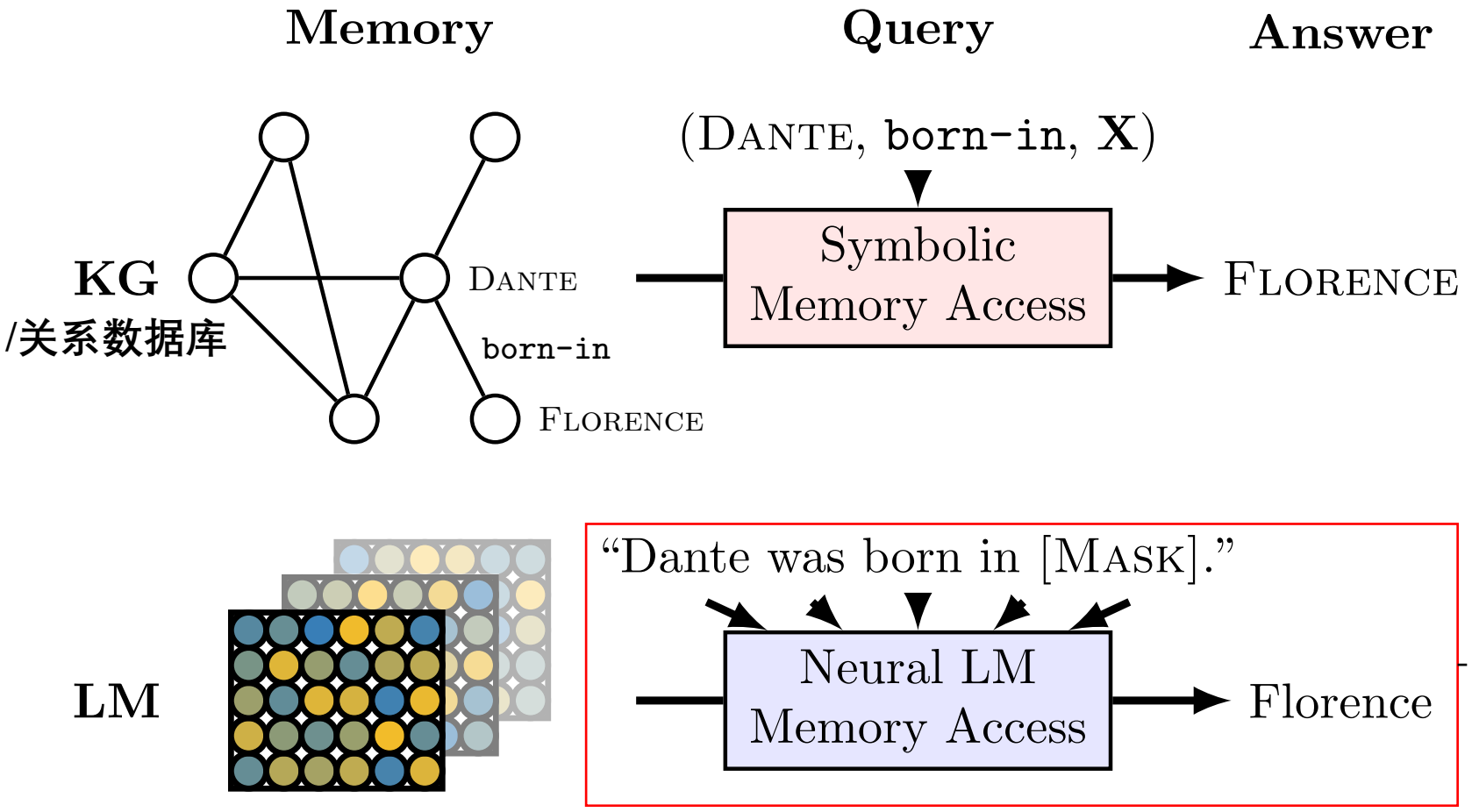




预训练语言模型学习到的语言知识，被其它任务复用，可以降低其它任务的学习难度

语言模型是知识库

语言模型里学到的不只语言知识，还有其它知识，所以语言模型实际上是知识库



不是用于帮助其它任务，而是与训练时一样的方式使用
第一轮：
输入：你在哪儿上班？
输出：腾
第二轮：
输入：你在哪儿上班？腾
输出：讯
第三轮：
输入：你在哪儿上班？腾讯
输出：E (表示结束)

e.g. ELMo/BERT/GPT

2019-EMNLP-Language Models as Knowledge Bases?

语言模型是知识库

知识库对比：知识表示和调用方式的演进

知识表示方式	表示的精确度	知识调用方式	调用方式的自然度	研究领域	代表应用	代表公司
关系型数据库	高	SQL	低	数据库	MySQL	Oracle、微软
互联网	中	Keywords	中	信息检索	搜索引擎	Google、微软
语言模型	低	自然语言	高	自然语言处理	ChatGPT	OpenAI、微软

接下来的工作就是从调用方式和知识表示两个维度来降低从语言模型中检索知识的难度：

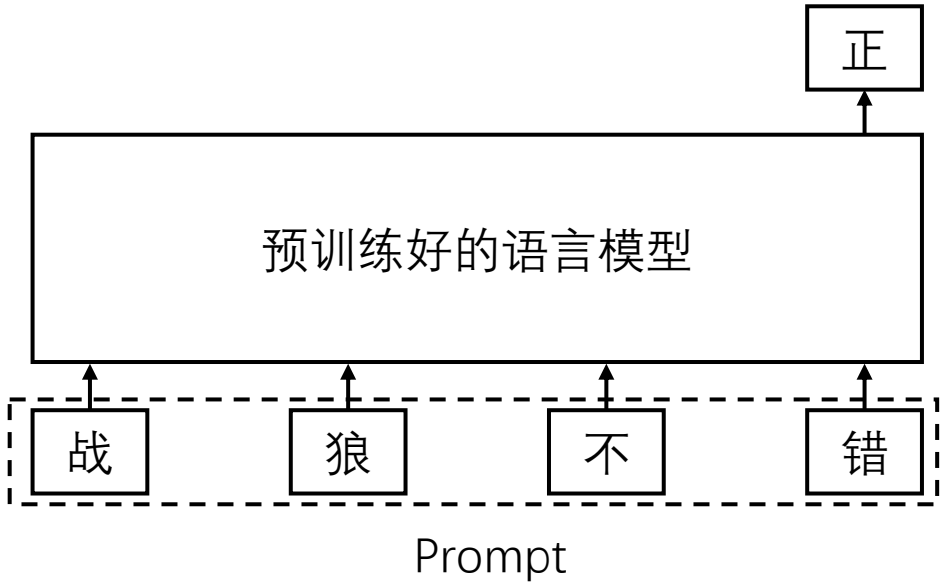
- 提示学习 (Prompt-Learning)，类似于搜索引擎的用户通过改变关键词来从搜索引擎中找到自己想要的信息
- 精调语言模型 (ChatGPT)，类似于Google、百度对搜索引擎做优化

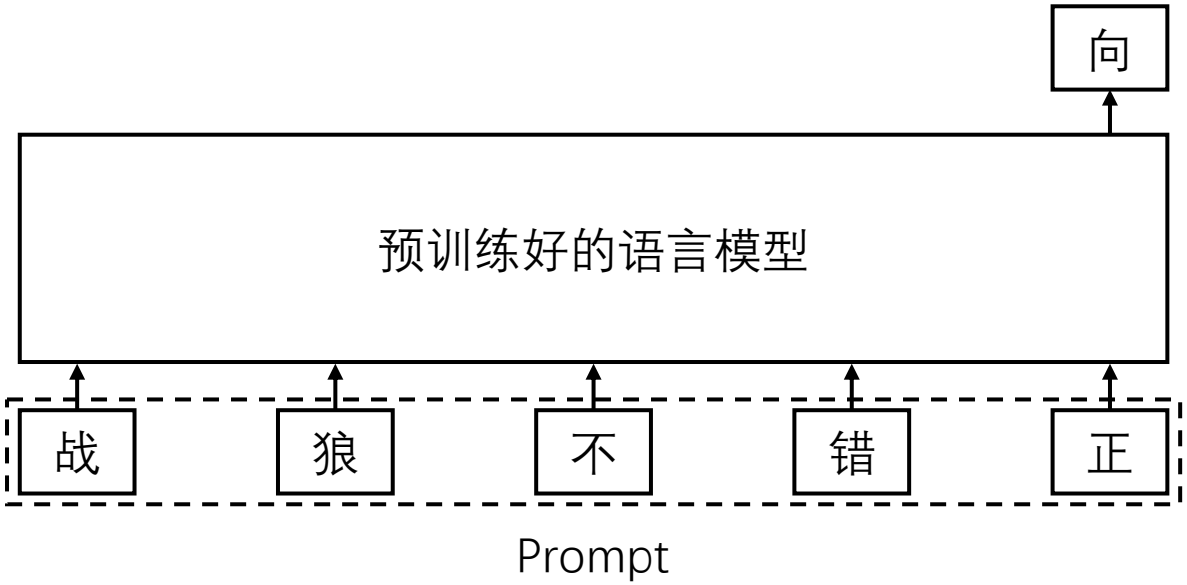
提示学习（Prompt-Learning）

调用方式优化

- Prompt-Learning: 挑选恰当的Prompt从预训练语言模型中查询答案。不用再额外训练模型。
- Prompt只是查询的一个新名字。

注意：下面是推断，不是训练模型

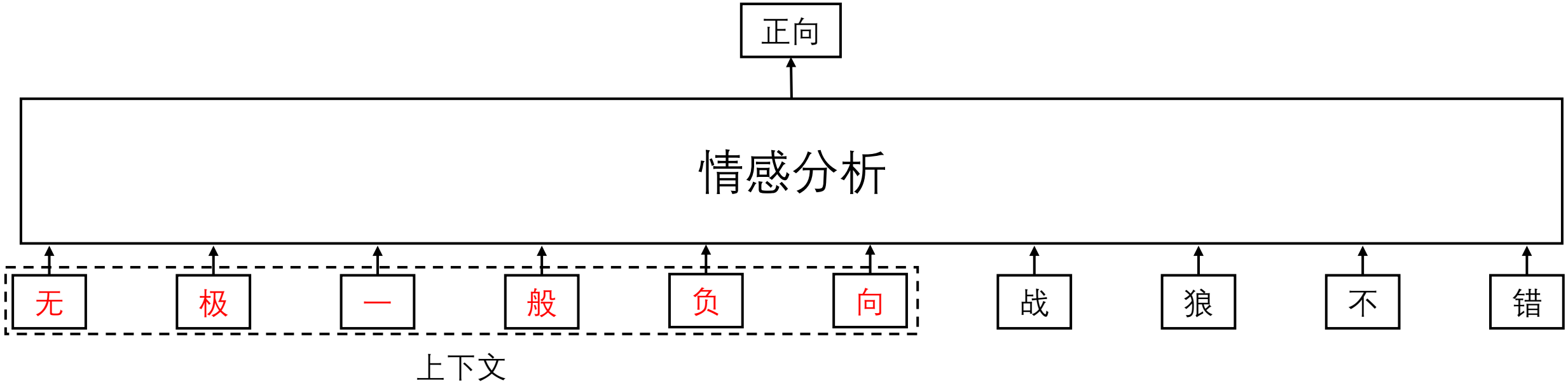




情感分类： 战狼不错 -> 正向

上下文学习（In-Context Learning） 一种提示学习

上下文学习：将其它问题-答案对作为Prompt的一部分，当作上下文，进行Prompt增强



其它上下文：
- 你是一个电影评论家
- 你是一个战争片爱好者

2020-nips-Language Models are Few-Shot Learners

思维链（Chain-of-Thought） 一种上下文学习

将上下文按照大脑思考过程进行拆解

Standard Prompting

Model 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. ✗

Chain-of-Thought Prompting

Model 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 balls. 2 cans of 3 tennis balls each is 6 tennis balls. 5 + 6 = 11. 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 cafeteria had 23 apples originally. They used 20 to make lunch. So they had 23 - 20 = 3. They bought 6 more apples, so they have 3 + 6 = 9. The answer is 9. ✓

类比大脑的COT示例：

- 1/Q: 是一个女演员
- 2/Q: 嘴巴很大
- 3/Q: 和苏大强一起拍过戏
- 4/Q: 这是是谁？
- A: 姚晨

语言模型真能学到示例的推理过程吗？

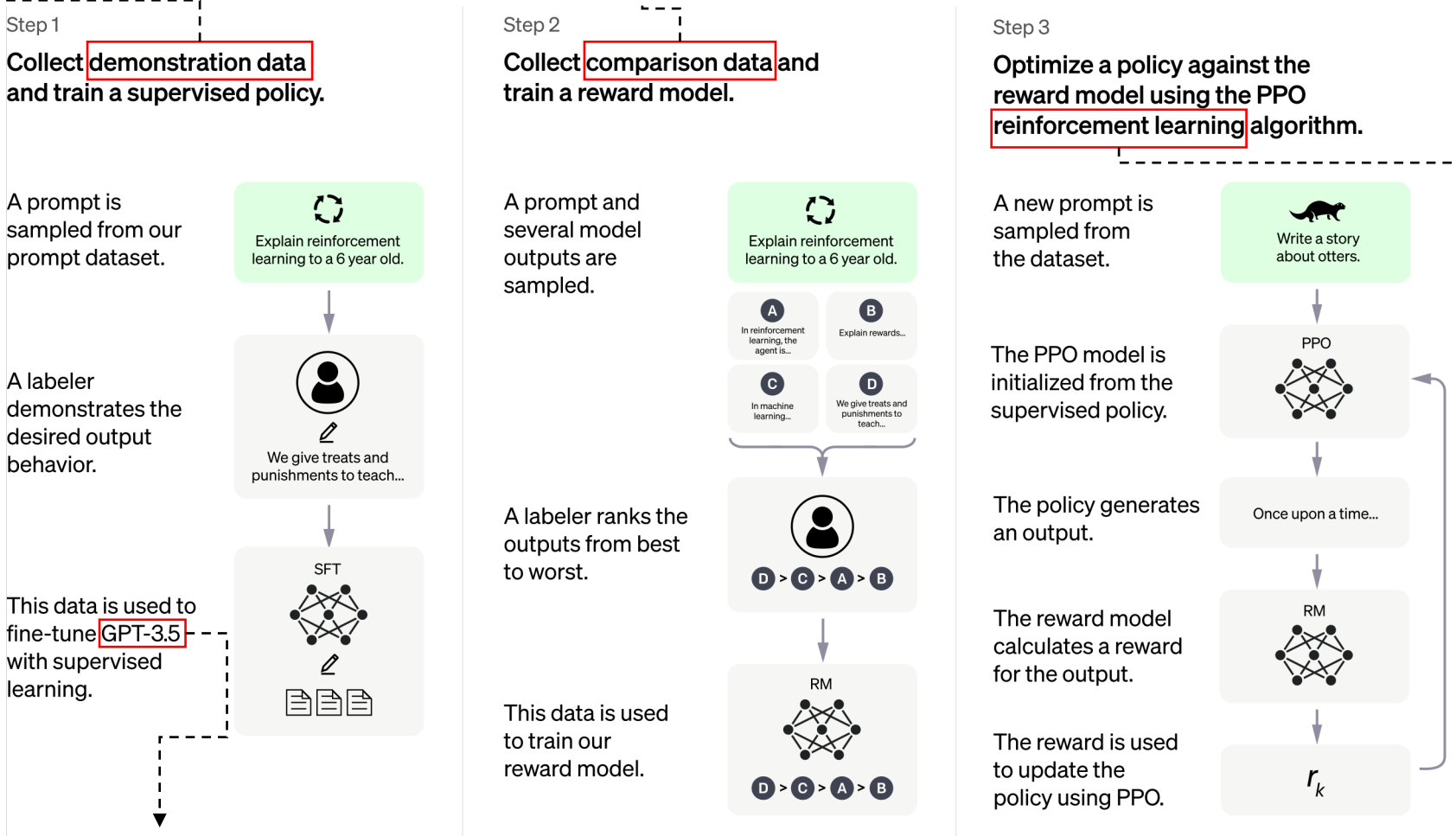
2022-NIPS-Chain-of-Thought Prompting Elicits Reasoning in Large Language Models

精调语言模型-ChatGPT

使用强化学习基于人类反馈精调GPT-3.5 (OpenAI预训练的一个语言模型)

第二阶段精调用的人类反馈

第一阶段精调用的人类反馈



精调使用的方法是强化学习

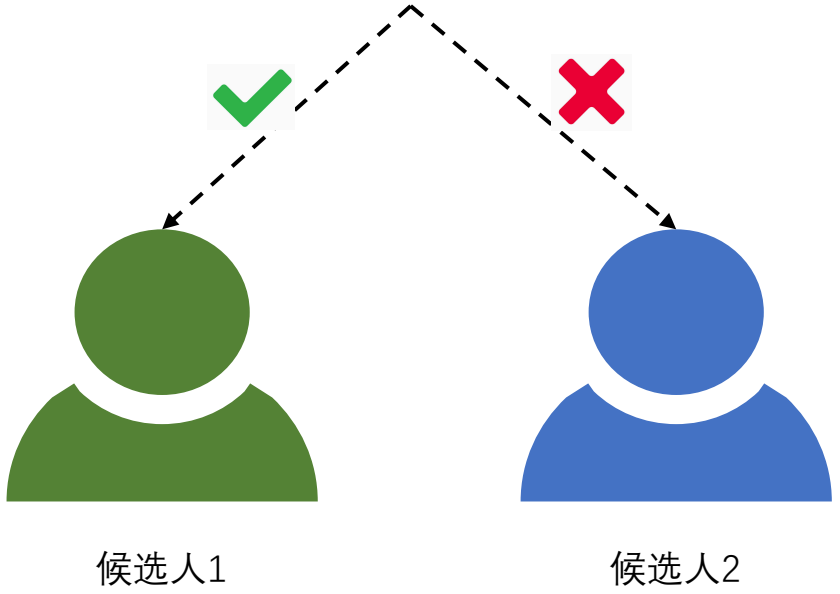
GPT-3.5是预训练好的语言模型

2022-Training language models to follow instructions with human feedback (InstructGPT)

精调语言模型-ChatGPT

第一类人类反馈：专家知识

某公司: HC只有1个，两人都能胜任，但候选人1比候选人2略好

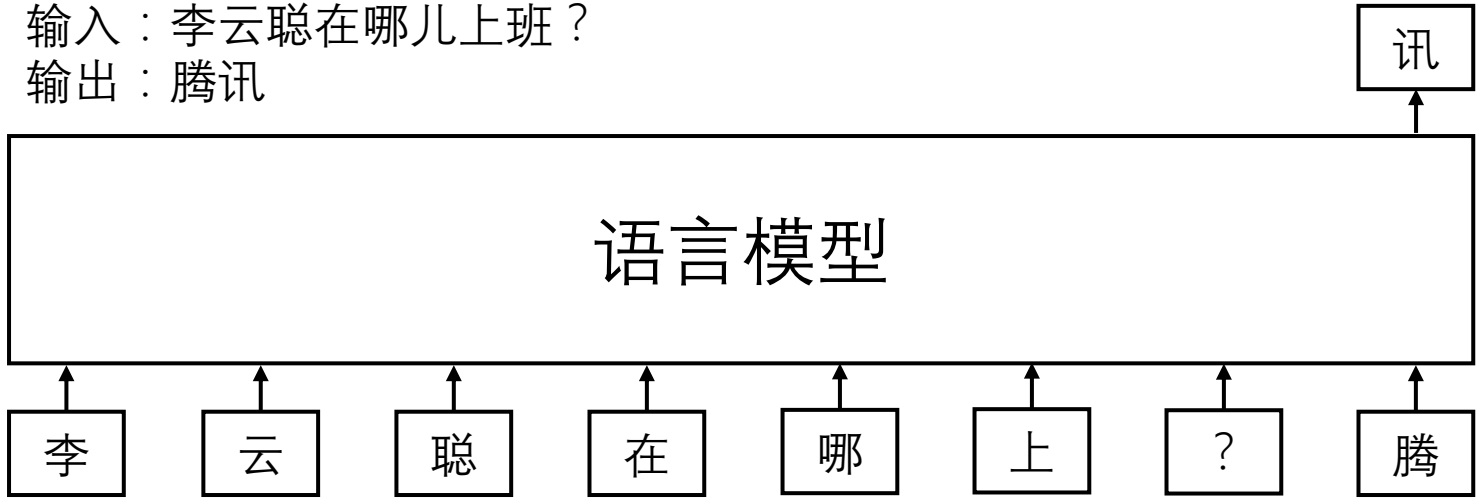


第二类人类反馈：对比

2022-Training language models to follow instructions with human feedback (InstructGPT)

ChatGPT: 精调语言模型

使用第一类人类反馈（demonstration data）精调语言模型
这一步跟原来语言模型的训练过程一样，只是训练数据不一样
第一类人类反馈示例：
输入：李云聪在哪儿上班？
输出：腾讯



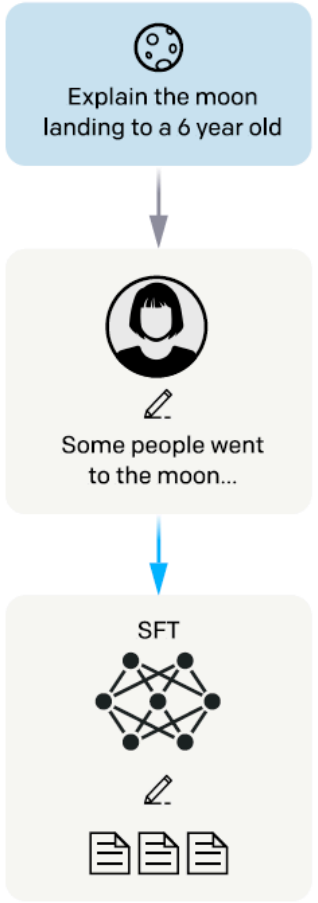
模仿学习 (强化学习的一种)的
训练数据被称为demonstration

Collect demonstration data, and train a supervised policy.

A prompt is sampled from our prompt dataset.

A labeler demonstrates the desired output behavior.

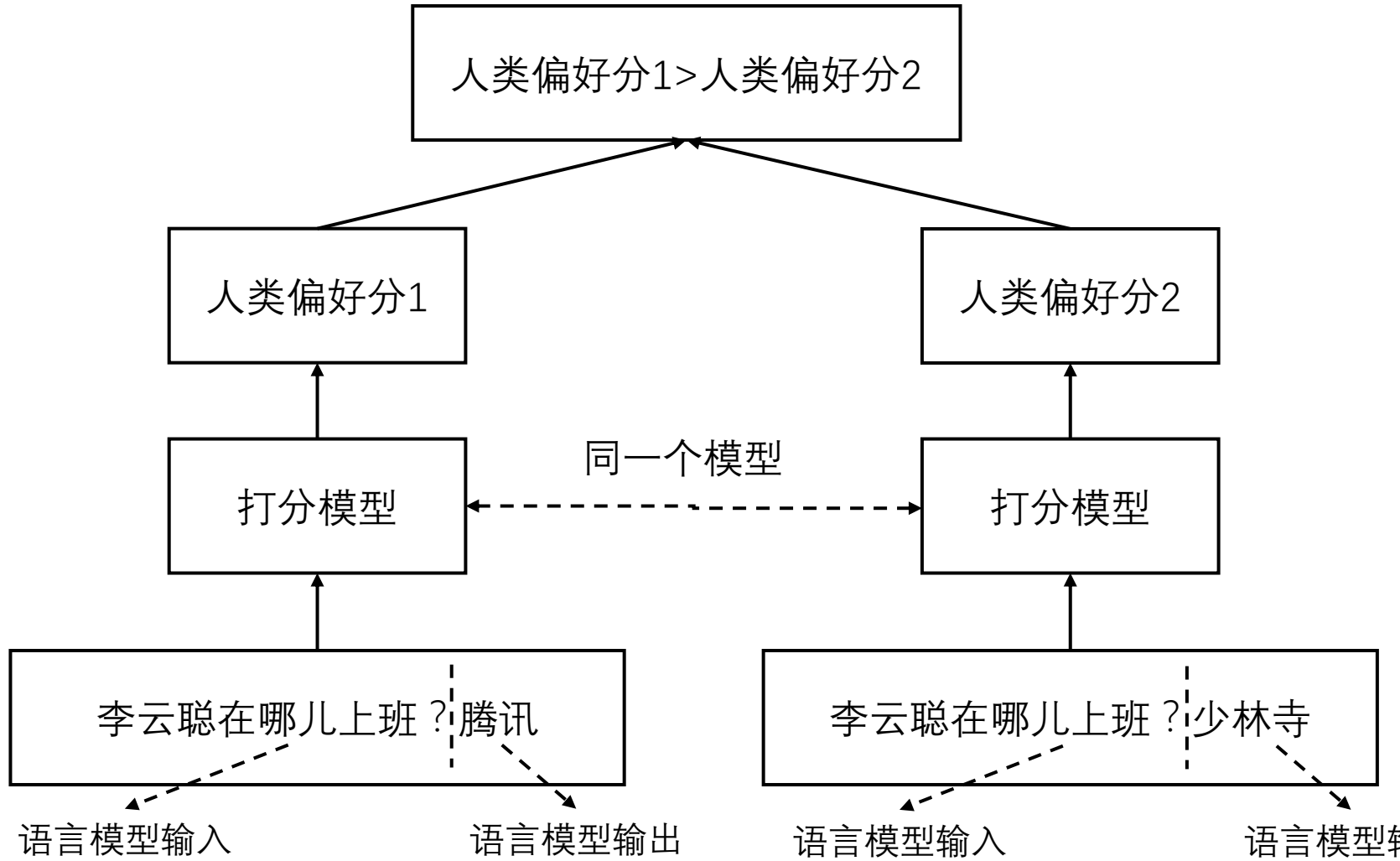
This data is used to fine-tune GPT-3 with supervised learning.

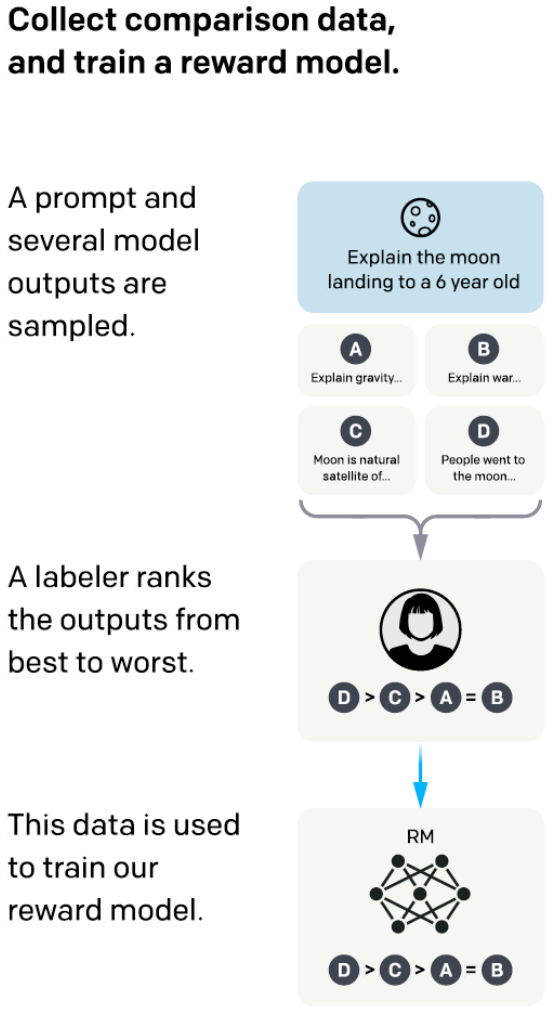


2022-Training language models to follow instructions with human feedback (InstructGPT)

ChatGPT: 训练强化学习打分模型

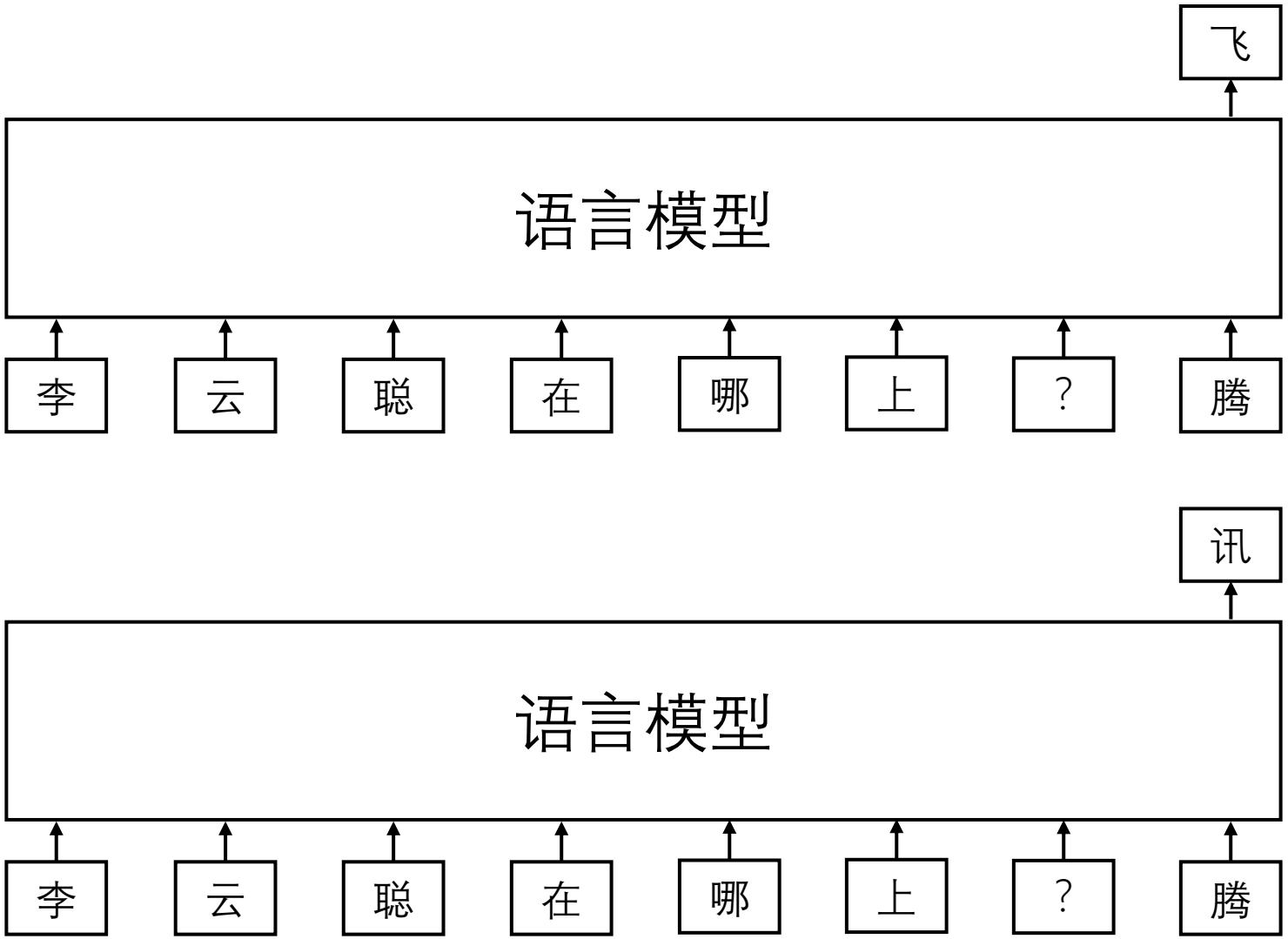
使用第二类人类反馈训练模型输出打分模型





2022-Training language models to follow instructions with human feedback (InstructGPT)

ChatGPT: 用强化学习再精调语言模型



打分模型给低分，
语言模型自己尝试调整策略

打分模型给高分，
语言模型会发扬这个策略

2022-Training language models to follow instructions with human feedback (InstructGPT)

ChatGPT: 用强化学习再精调语言模型

两阶段精调之后的语言模型就是ChatGPT

2022-Training language models to follow instructions with human feedback (InstructGPT)

一些资源

- awesome-chatgpt
- awesome-chatgpt-prompts
- awesome-chatgpt-prompts-zh
- chatgpt-advanced

一些资源

- Introducing ChatGPT
- The Art of ChatGPT Prompting: A Guide to Crafting Clear and Effective Prompts
- How to Make Money with ChatGPT: Strategies, Tips, and Tactics
- ChatGPT技术原理解析：从RL之PPO算法、RLHF到GPT-N、instructGPT
- ChatGPT三问：是什么、从哪来、去往哪？

Thanks

附录：各种复杂系统中的涌现现象