

# 大模型时代知识工程使命的再审视

李涓子

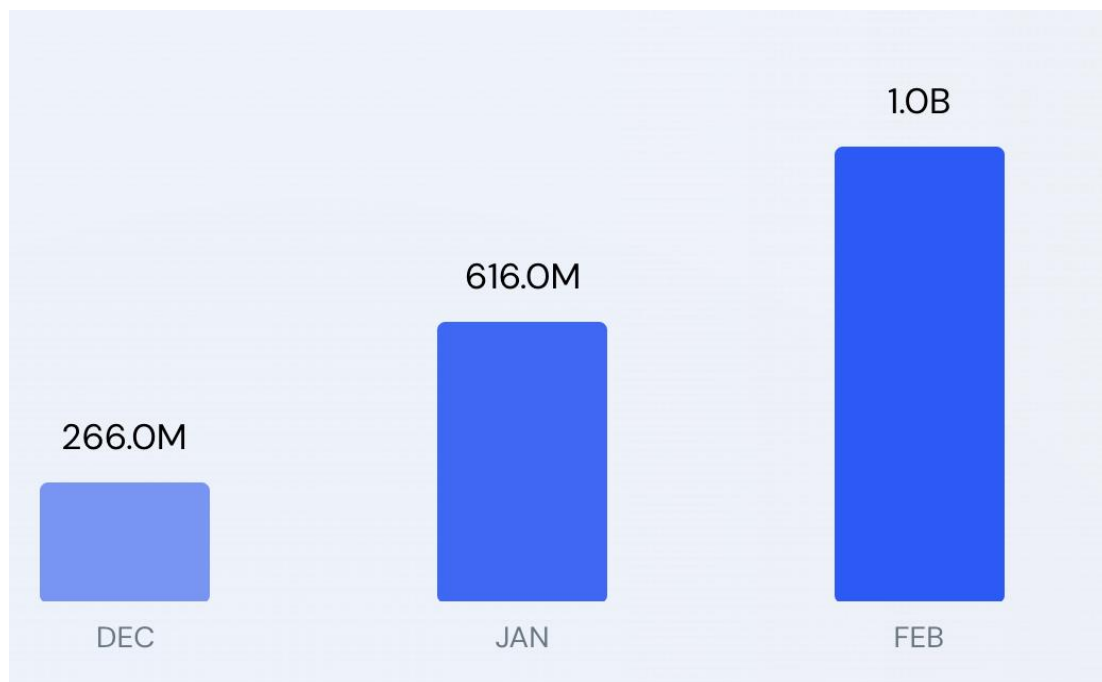
知识工程实验室 (KEG)

清华大学计算机科学与技术系

# 大模型掀起人工智能新浪潮

## ■ OpenAI全球访问量

- 2022.11.30发布ChatGPT
- 单月独立用户量突破**10亿**
- 史上**用户增长最快**网站

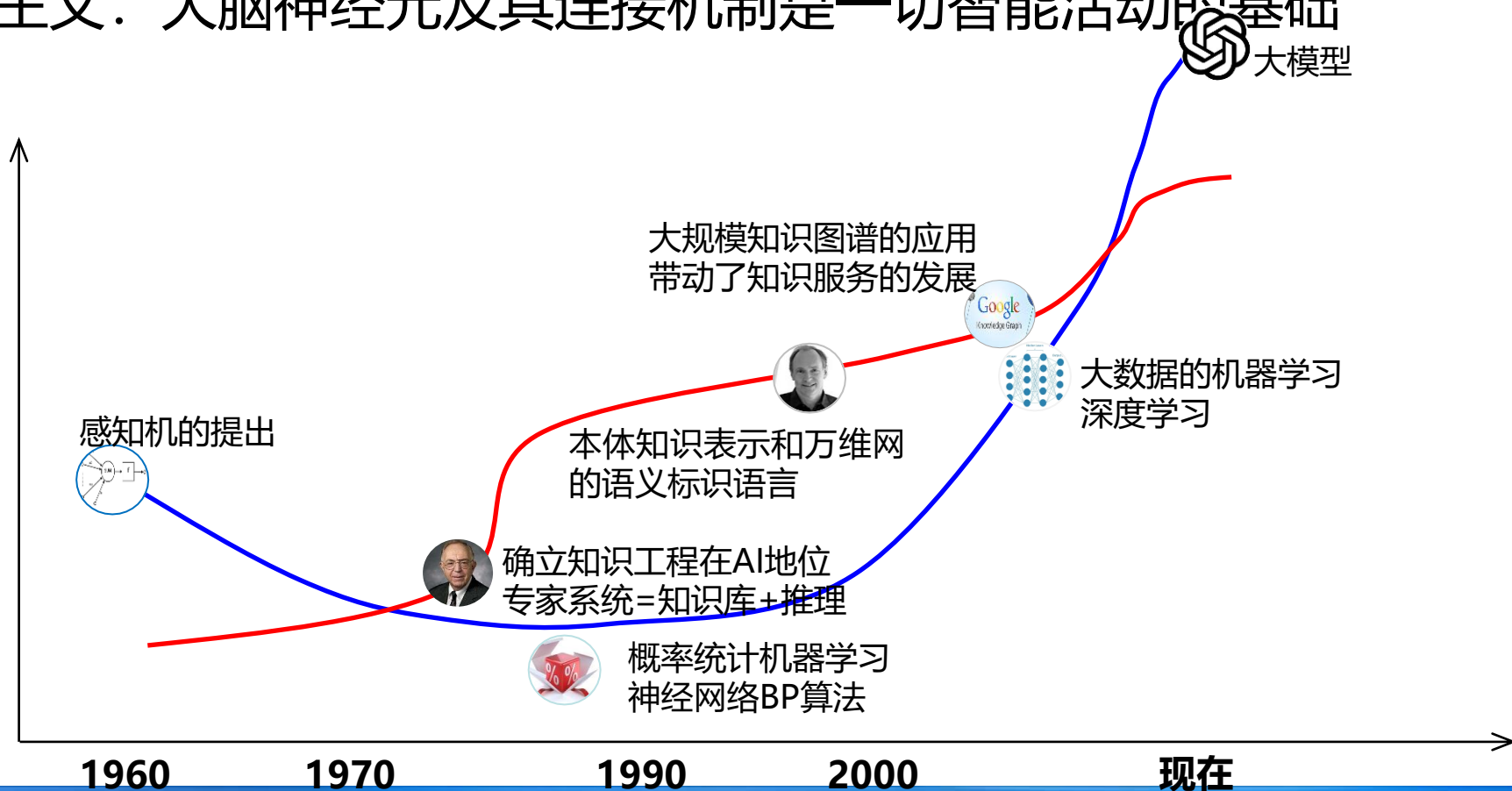


# 历史回顾：人工智能的符号主义与连接主义

■ 人工智能：实现类人层次具有感知、学习、推理和规划能力的系统

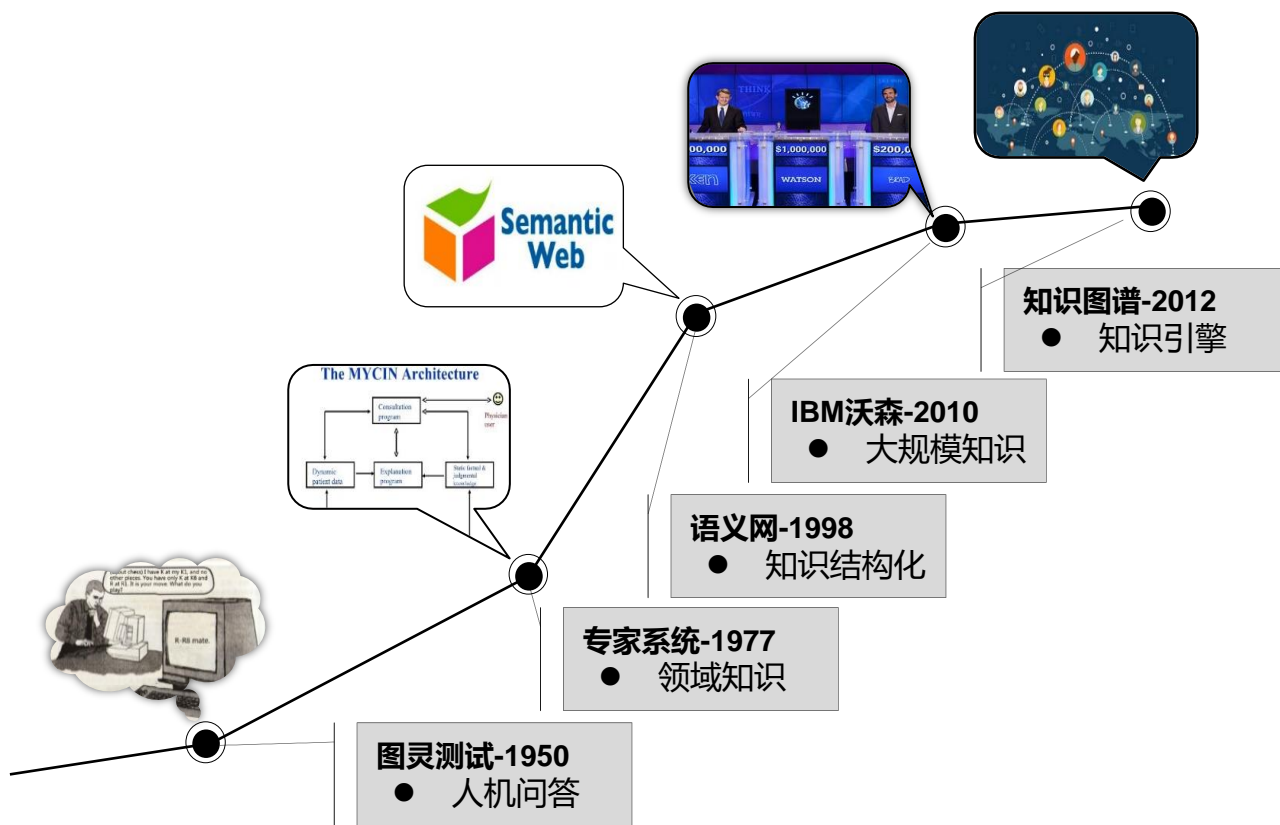
- 符号主义：物理符号系统是智能行为的充要条件

- 连接主义：大脑神经元及其连接机制是一切智能活动的基础



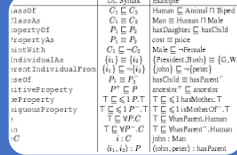
# 符号主义：知识工程

■ **知识工程**：建立基于知识的系统，解决需要专家解决的复杂问题



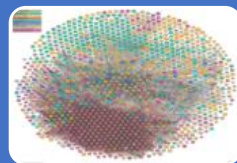
## 语义互操作及多智能体应用

- 图数据库和SPARQL等交互语言（自然语言交互）
- 语义搜索、对话问答、智能推荐和大数据分析



## 逻辑推理

- 谓词逻辑、描述逻辑等
- 归纳、演绎、溯因等推理技术及推理机



## 大规模知识资源建设

- DBpedia、Freebase、Google知识图谱、Wikidata和开放链接数据
- 知识获取和融合系列技术和工具



## 机器可理解计算的语义

- RDF、RDFs、OWL等建模语言
- DBpedia本体、Schema.org等通用和大量领域本体

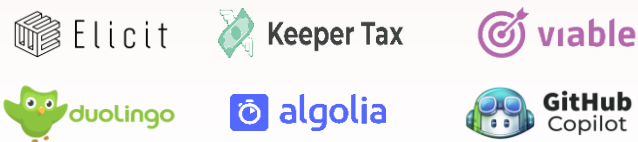
**知识积累和重用、知识推理和决策是知识工程的核心问题**

# 连接主义：大规模预训练模型

## ■ 大规模预训练语言模型是新一代人工智能应用的基础设施

大模型拥有大量参数化知识，强大内容生成能力，支持自然交互，可完成多场景跨域多任务

现有300+ 应用基于OpenAI 的大模型开发



大模型从单模态向多模态发展，从文本逐步发展到视觉、虚实迁移控制

OpenAI与阿里、智源均已推出多模态模型



大模型已成为各国人工智能技术新方向，是新一代人工智能应用的基础设施

GLM-130B大模型  
模型参数  
1300亿

ChatGPT, GPT-4  
ChatGLM

2022年

2023年

GPT  
模型参数  
1.1亿

GPT-2  
模型参数  
15亿

GPT-3  
模型参数  
1750亿

DALL-E  
模型参数  
1.6万亿

悟道2.0  
模型参数  
1.75万亿

2018年

2019年

2020年

2021年

- Transformer应用在OpenAI GPT 1中，Transformer架构是预训练大模型的技术基础

- OpenAI GPT 2.0
- BAIDU Ernie 1.0，基于Bert 模型做进一步优化

- OpenAI GPT 3，跨时代意义的语言大模型，标志着AI文本生成到下一个阶段

- DALL-E / Clip: OpenAI大规模图像模型，可基于短文本生成对应图像和进行图像分类
- AliceMind: 阿里大规模多模态模型，集合语言理解与生成能力

- 出现大规模基于预训练模型的应用
- GLM-130B: 125T中英文文本6亿图+千万级视频千万级别知识图谱
- ChatGPT: 在问答模型中引入人类反馈强化学习，大幅提升生成结果质量

- ChatGLM: 基于GLM-130B底座，具备问答和对话功能的千亿中英双语模型
- GPT-4: 多模态大模型，可同时实现图像和文本输入
- Ernie Bot (文心一言): 百度全新一代知识增强大语言模型

ERNIE GPT-2

GPT-3

DALL-E

ALICEMIND

ChatGPT

智谱·AI ChatGLM

GPT-4 文心一言

# 学者观点1: ChatGPT不能带来真正的智能

- **原理未知**: 知道什么? 能推理吗? 在什么意义上理解了答案?
- **中间任务**: 大基准替代了中间任务, NLP不再研究语言
- **未来研究**
  - 回归由模块化、基于数理逻辑、语义明确定义表示所构成的系统, 但表示不一定由手工设计, 可以从数据中学习
  - ChatGPT应用要区分任务领域, 锚定在事实上, 与规划系统相结合, 将带来更大作用
  - 构建人们能够理解的AI系统



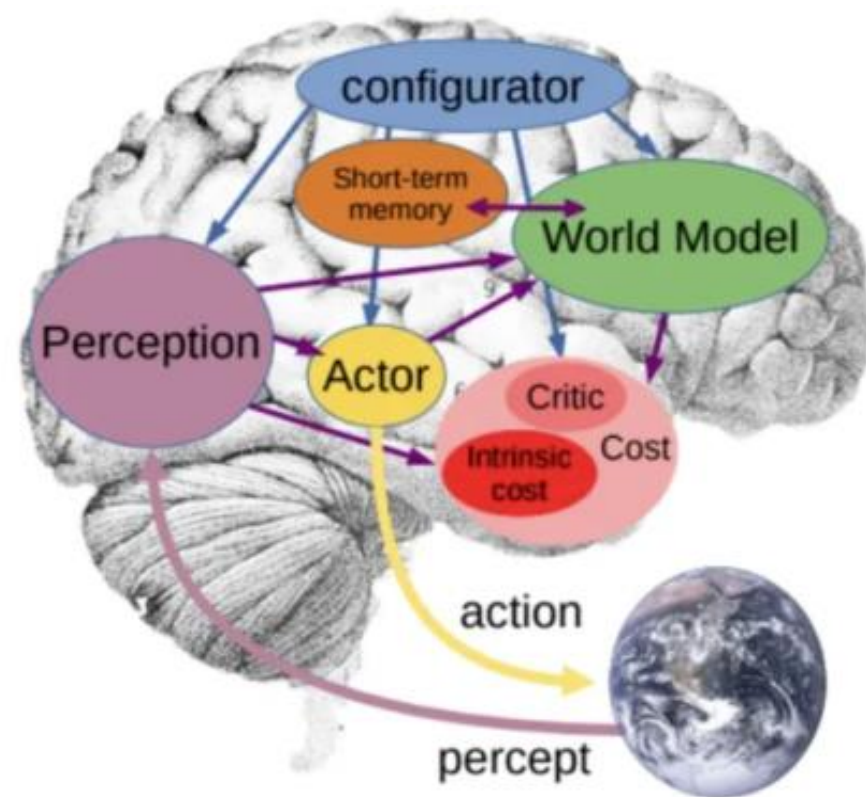
Stuart Russell, UCB教授  
AAAI、ACM 和 AAAS Fellow  
@2023北京智源大会

# 学者观点1: ChatGPT不能带来真正的智能

- Machine learning sucks compare to humans and animals
- chatGPT这样主流路线存在局限
  - 学习世界的表征和预测模型
  - 学习推理, 参考系统1和系统2的概念
  - 如何通过将复杂任务分解成简单任务
- 人工智能应该走向能够学习、推理和规划的大模型

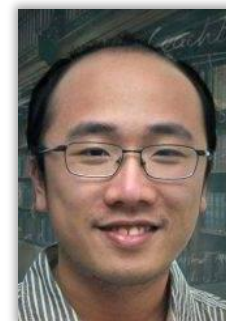


LeCun Yan @ 2023北京智源大会



# 学者观点2: GPT-4通用人工智能的火花

- 系统评测了GPT-4的早期版本, 证实其在数学、编程、视觉、医学、法律、心理学等多样化和高难度的任务中已经接近人类水平
  - 多模态和跨学科能力 (出色)
  - 音乐能力 (能够产生和解释有效旋律, 无法生成非平凡的和声)
  - 编程能力 (出色)
  - 数学能力 (显著提升, 仍距专家有明显差距)
  - 世界交互 (可以适当调用外部工具)
  - 人类共情 (可以建立类人心智模型)
  - 局限: 规划、工作记忆、回溯和推理能力
- 豪华的作者团队: 3位斯隆奖+1位新视野数学奖





# 学者观点2: GPT-4通用人工智能的火花

## ■ Harnessing the Power of AI: A Physicist's Perspective

– 人工智能可以帮助我们更快地实现几乎所有这些目标

## ■ 理解复杂AI系统, 掌控AI (understanding and containing AI)

– mechanistic interpretability

– 从机器学习系统中提取出它们学到的所有知识, 并在其他系统中重新实现它们, 以证明它们将按照我们的意愿行事

– actively participate and contribute to the development of the AI field – an arena filled with **endless potential for acquiring insights to understand, manipulate, and control the formidable power of AI systems.**



Max Tegmark, MIT教授

著名物理学家、宇宙学家和AI研究者

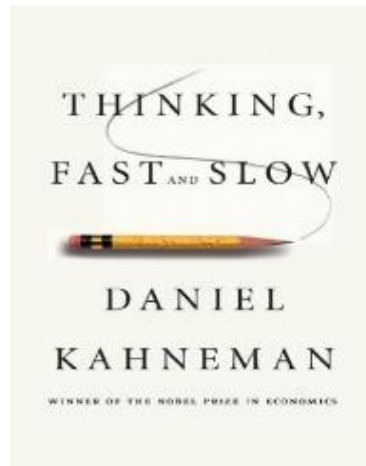
# 重新思考符号主义与连接主义的关系

## ■ 认知动机

- 符号主义：人类依赖符号系统（如语言文字）组织、交流、推理、创造知识
- 连接主义：人脑的物理形式是广泛连接的神经元

## ■ 双过程认知理论

系统1 直觉 (Intuitive)
<ul style="list-style-type: none"><li>• 情感、记忆、经验</li><li>• 无意识、能够快速激活关联对象</li><li>• 容易被欺骗</li></ul>



系统2 理性 (Analytic)
<ul style="list-style-type: none"><li>• 逻辑推理计算</li><li>• 有意识、“循规蹈矩”地深思熟虑</li><li>• “慢诸葛”</li></ul>

连接主义模型是实现基础，符号系统是高层能力特征

# 重新思考符号主义与连接主义的关系

## ■ 知识工程仍然很重要

- 基于知识的系统可以保证知识正确性、可解释性和一致性
- 逻辑推理和规划

## ■ 以连接主义模型满足符号主义特征，实现类人层次智能

- 以符号形式传达、学习知识：Data-efficient、Following Instruction
- 遵循逻辑规则进行推理：严格而非以一定概率遵守
- 可解释地创造新知识：起码做到self-explanation
- 利用元知识：快速适配新符号系统

## ■ 两种方法融合螺旋式上升

# 大模型时代知识工程发展方向

- 大模型中的知识机理研究
- 知识视角下大模型能力评测
- 知识工程和大模型相互增强与协同演化
- 新型知识高效表示和学习
- 知识驱动的可靠、可信与可控应用



2023战略研讨会@广州，4月23-24日

# 方向1：知识工程借力大模型

## ■ Knowledge Engineering **BY** the LLM

– 知识工程全生命周期应该在大模型的加持下全面提升



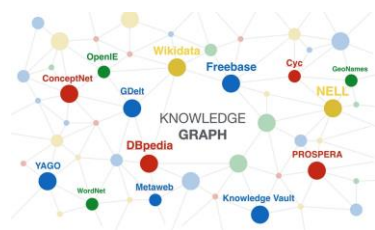
结构化知识丰富

长尾知识丰富



参数化知识匮乏

封闭世界假设



参数化知识丰富

开放世界假设



结构化知识匮乏

长尾知识建模能力差



### ■ 将大模型作为资源

- 数据增强或某种知识资源



### ■ 将大模型作为助推器

- 阅读理解 & QA范式

# 知识工程借力大模型

- 通过提示将大模型存储的大量参数化知识转化为符号知识
- 通过上下文学习或者指令微调从给定资源中获取多种类型的知识

**NER** CoNLL 2003, ACE 2005, Ontonotes, ...

**RE** CoNLL 2004, SciERC, NYT 11, ...

**EE** CASIE, GENIA, ...

Zero-Shot Evaluation Unseen Datasets Mit-Movie ...

ChatIE  
北京交通大学, 阿里巴巴达摩院

**Multi-Task Instruction Tuning**

24 V100 (3 nodes) DeepSpeed ZeRO-3

8 V100 (1 Node) PEFT

Pretrain Dataset: Chinese (5,500K sentences), English (1,500K sentences), Code (900K sentences)

Instruction Dataset: Origin (English, Chinese), Synthetic (Chinese, English), NER/RE/EE, CoNLL/ACE/CASIE/NYT 11, DuEE(IE), MSRA, WEIBONER, PEOPLE DAILY, CLUE, KG2Instruction

LLaMA, ZhiXi, Knowledge, Enhance Chinese

CoT, gsm8K, aqua, creak, Alpaca

General, Alpaca, Handmade, Code, Github, Leetcode

Full Tuning, LoRA, LoRA Tuning, Interactive Update, Future

**Instruct-UIE, 复旦大学**

Multi-Task Instruction Tuning (Tom, person)

LLaMA, ZhiXi, Knowledge, Enhance Chinese

Pretrain Dataset: Chinese (5,500K sentences), English (1,500K sentences), Code (900K sentences)

Instruction Dataset: Origin (English, Chinese), Synthetic (Chinese, English), NER/RE/EE, CoNLL/ACE/CASIE/NYT 11, DuEE(IE), MSRA, WEIBONER, PEOPLE DAILY, CLUE, KG2Instruction

CoT, gsm8K, aqua, creak, Alpaca

General, Alpaca, Handmade, Code, Github, Leetcode

Full Tuning, LoRA, LoRA Tuning, Interactive Update, Future

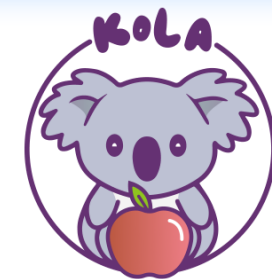
KnowLM, 浙江大学

# 方向2：知识工程助力大模型

## ■ Knowledge Engineering **FOR** the LLM

–以大模型的问题为目标，诉诸合适的知识工程手段

- 系统地评测大模型涌现出来的各种符号能力
- 知识幻觉性
- 知识更新
- 复杂推理过程的逻辑准确与可解释性
- 规模、效率问题



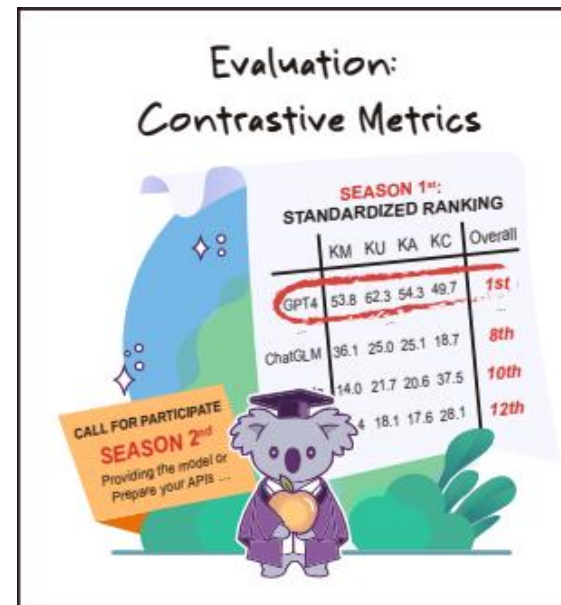
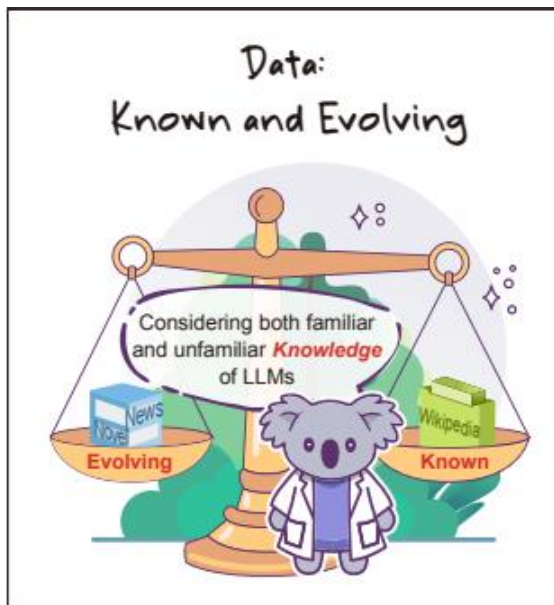
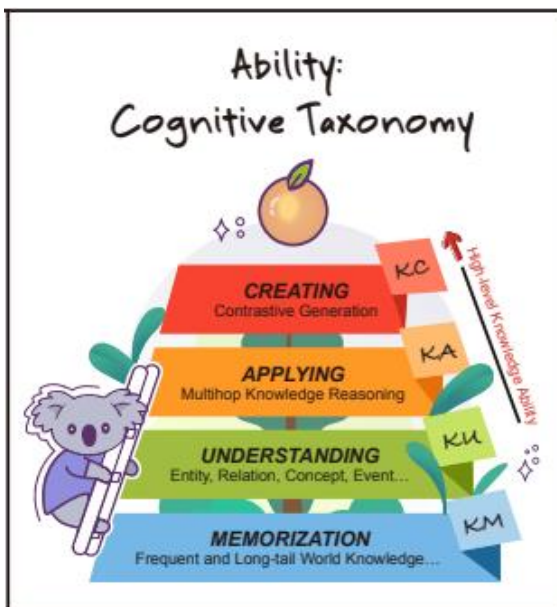
# 大语言模型知识评测

## ■ Knowledge-oriented LLM Assessment (KoLA)

– 目标：全面、公正、适用性强的评测

– 聚焦三要素

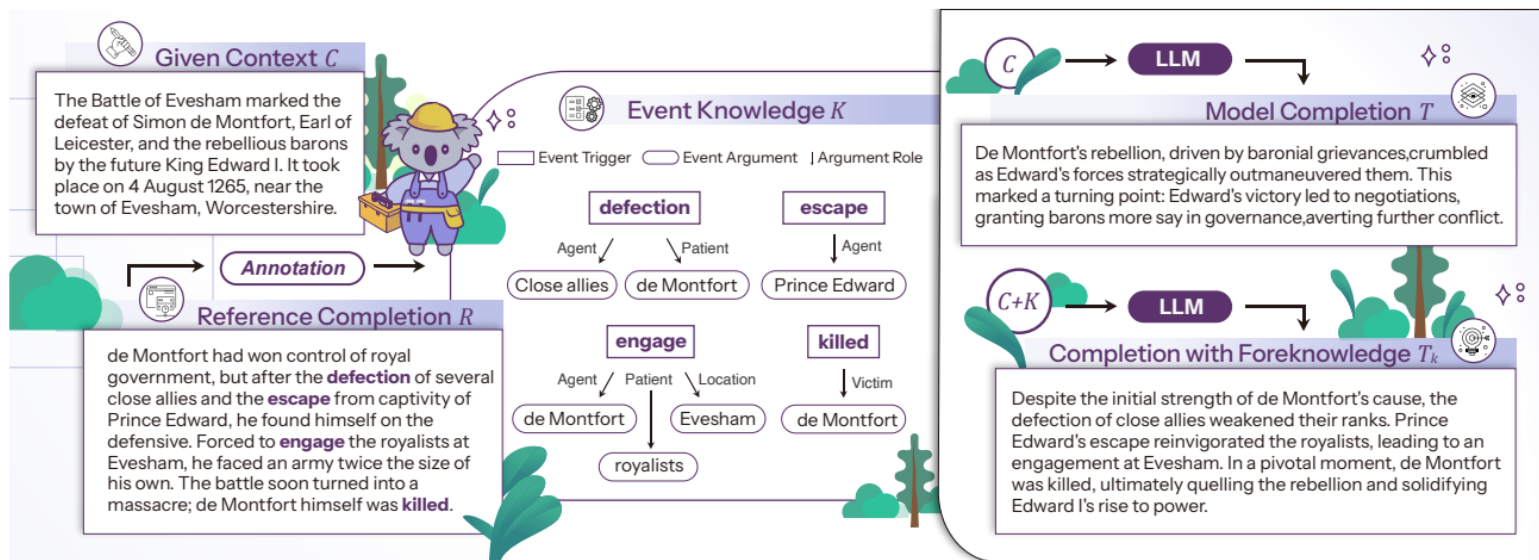
- 能力建模：不仅需要设计各种能力的范畴，还要定义能力间的内在关联
- 评测数据：保证不同模型训练数据的差异不影响评测的公平性
- 评价指标：不仅需要全面客观，还要便于理解和解读评测结果





# 大语言模型知识评测

- 能力建模：参考布鲁姆认知理论，根据知识认知的特点分为四个层次
  - 知识记忆：模型是否能够记住常用和长尾的世界知识
  - 知识理解：模型是否能够理解知识组织结构，包括概念、实体、关系和事件等
  - 知识应用：模型是否能够运用不同来源的知识进行推理
  - 知识创造：模型能否基于给定知识生成新的内容



# 大语言模型知识评测

## ■ 评测数据：支撑能力建模，覆盖已知(Known)和实时(Evolving)知识

Level	ID	Dataset	Metrics	Exclusive	Context Type	Test Set	Pool	Source
KM	1-1	High-Freq.	EM, F1	✓	Triple	100	20.6M	Known
	1-2	Low-Freq.	EM, F1	✓	Triple	100	20.6M	
	1-3	RTM	EM, F1	✓	Triple	100	2.7k	Evolving
KU	2-1	COPEN-CSJ	Acc.	✓	Entity, Concept	100	3.9k	Known
	2-2	COPEN-CPJ	Acc.	✓	Concept	100	4.7k	
	2-3	COPEN-CiC	Acc.	✓	Concept	100	2.3k	
	2-4	FewNERD	F1	✗	Sentence	300	188.2k	
	2-5	DocRED	F1	✓	Document, Entity	100	12k	
	2-6	MAVEN	F1	✓	Document	100	20.4k	
	2-7	MAVEN-ERE	F1	✓	Document(s), Event	199	1.3M	
	2-8	ETU	F1	✓	Document, Entity	100	1.6k	Evolving
KA	3-1	HotpotQA	F1	✗	Document(s)	100	7.4k	Known
	3-2	2WikiMulti.	F1	✓	Document(s)	100	12.6k	
	3-3	MuSiQue	F1	✓	Document(s)	100	2.5k	
	3-4	KQA Pro	F1	✓	KG	100	1.2k	
	3-5	KoRC	F1	✓	Document(s), KG	100	5.2k	
	3-6	ETA	F1	✓	Document(s), KG	49	1.6k	Evolving
KC	4-1	Encyclopedic	BLEU, Rouge	✓	Document, Event	95	4.5k	Known
	4-2	ETC	BLEU, Rouge	✓	Document, Event	95	100	Evolving

高频和长尾知识记忆，来源wikidata5M

概念、实体、关系、事件及其关系理解

实时数据，来源：时事新闻和架空小说

复杂知识推理问答

事件生成能力

# 大语言模型知识评测

## ■ 整体观测结论

- 大模型在知识记忆、理解、应用和创造类任务上普遍还存在很大局限性
- 未进行指令微调或人类对齐的模型，知识记忆能力与模型规模相关
- 指令微调或人类对齐后，高阶能力与模型规模相关性提升，但低阶能力相关性减小（对齐税）

Model	Level 1: KM				Level 2: KU								Level 3: KA						Level 4: KC			Overall (1,2,3,4)			
	1-1	1-2	1-3	Rank	2-1	2-2	2-3	2-4	2-5	2-6	2-7	2-8	Rank	3-1	3-2	3-3	3-4	3-5	3-6	Rank	4-1	4-2	Rank	Avg	Rank
GPT-4	51.4	55.5	54.6	1st	63.5	42.9	46.0	62.3	100.0	72.3	72.8	59.5	1st	56.2	58.3	72.4	26.9	56.5	55.6	1st	47.0	52.5	3rd	2.06	1st
GPT-3.5-turbo	41.7	47.6	42.0	4th	37.5	43.8	44.8	49.2	47.2	44.1	50.5	25.5	2nd	54.7	37.2	48.5	42.5	24.7	24.3	4th	51.1	54.6	2nd	1.32	2nd
InstructGPT davinci v2 (175B*)	30.8	37.2	32.4	7th	26.6	42.5	36.5	36.8	53.1	56.7	34.6	31.2	3rd	23.9	33.8	38.4	15.7	45.3	43.9	6th	53.6	53.3	1st	1.02	3rd
Cohere-command (52.4B)	46.6	42.6	56.8	2nd	33.1	41.2	40.6	21.4	33.5	13.2	40.9	18.6	4th	30.1	36.1	39.5	47.0	49.9	53.8	3rd	11.4	35.4	7th	0.77	4th
FLAN-UL2 (20B)	41.3	31.9	53.0	5th	52.7	41.2	47.8	10.7	18.6	13.2	16.3	18.6	6th	44.9	43.0	33.3	49.3	38.1	51.5	2nd	24.1	15.2	12th	0.55	5th
FLAN-T5 (11B)	44.1	39.9	49.6	3rd	57.0	42.1	43.6	13.4	—	—	—	—	5th	39.8	44.6	26.5	49.3	34.1	—	5th	15.0	17.0	16th	0.38	6th
J2-Jumbo-Instruct (178B*)	23.0	24.0	17.6	11th	20.1	15.8	24.5	32.1	26.3	25.7	45.2	22.0	7th	40.1	24.4	25.2	33.6	22.0	14.3	7th	41.5	42.6	4th	0.29	7th
ChatGLM (130B)	27.8	44.5	36.1	6th	23.3	42.1	46.6	10.7	18.6	15.9	24.4	18.6	8th	30.3	27.2	21.5	31.3	30.8	9.0	9th	19.7	17.8	13th	0.09	8th
InstructGPT curie v1 (6.7B*)	19.0	33.1	33.1	8th	22.3	34.9	35.9	17.1	19.1	14.6	19.9	18.6	9th	25.0	30.8	17.5	22.4	25.0	25.9	10th	23.5	22.6	10th	-0.01	9th
LLaMa (65B)	15.5	16.7	9.9	13th	14.6	10.3	10.7	50.8	25.7	23.0	19.6	18.6	11th	8.2	28.8	35.7	17.9	15.0	18.1	12th	42.6	32.3	5th	-0.09	10th
T0++ (11B)	31.3	28.2	25.3	9th	23.3	32.7	20.9	10.7	—	—	—	—	13th	14.9	14.3	17.1	4.5	34.1	—	16th	12.6	24.7	14th	-0.29	11th
Alpaca (7B)	13.0	16.4	11.0	14th	14.6	10.3	11.3	20.2	20.2	24.3	16.3	18.6	16th	6.7	9.9	14.0	8.9	40.2	25.6	15th	32.5	23.7	6th	-0.39	12th
GLM (130B)	13.4	16.3	9.4	15th	14.6	10.3	10.7	44.7	27.6	22.6	16.3	27.6	10th	16.1	2.5	11.4	13.4	39.6	39.6	13th	25.2	15.7	11th	-0.40	13th
UL2 (20B)	17.6	19.0	10.9	12th	14.6	10.3	10.7	12.8	—	—	—	—	21th	17.9	20.1	19.3	33.6	9.0	—	14th	22.5	23.7	9th	-0.47	14th
ChatGLM (6B)	22.3	23.3	19.5	10th	15.7	39.5	26.3	10.7	18.6	14.6	16.9	18.6	12th	13.5	19.6	15.3	11.2	12.6	18.3	17th	12.1	13.7	18th	-0.49	15th
GPT-J (6B)	12.4	10.6	8.9	18th	14.6	10.3	10.7	17.0	18.6	24.0	—	18.6	19th	33.0	33.4	20.2	47.0	10.5	12.0	8th	26.6	1.4	17th	-0.54	16th
GPT-3 davinci v1 (175B)	10.0	9.8	8.5	20th	14.6	10.7	10.7	25.1	18.6	22.5	16.4	18.6	15th	10.3	3.2	16.0	11.2	14.8	10.7	19th	28.6	17.6	8th	-0.65	17th
GPT-JT (6B)	11.5	10.7	9.2	19th	14.6	10.3	10.7	14.0	18.6	29.9	—	18.6	18th	24.8	32.0	16.2	26.9	11.5	12.3	11th	16.1	0.0	20th	-0.73	18th
GPT-NeoX (20B)	11.6	12.3	9.0	17th	14.6	10.3	10.7	20.6	18.6	25.1	—	18.6	17th	5.5	3.7	10.1	8.9	16.0	12.4	21th	28.5	4.0	15th	-0.77	19th
BLOOM (7B)	12.6	13.4	11.2	16th	14.6	10.3	10.7	25.0	21.9	22.1	16.3	18.6	14th	10.7	13.8	10.1	11.2	21.0	16.9	18th	12.4	3.4	21th	-0.80	20th
GPT-3 curie v1 (6.7B)	9.2	9.6	8.5	21th	14.6	10.3	10.7	16.3	18.6	18.4	18.0	18.6	20th	15.0	5.1	13.6	8.9	12.2	9.2	20th	18.9	6.5	19th	-0.86	21th

注：表中所有得分是标准化评分，绝对得分请查阅原论文：<https://arxiv.org/abs/2306.09296>

# 大语言模型知识推理

## 可解释编程推理框架

### 目标资源 (数据结构)

知识图谱

Input Paragraphs:

The Sum of All Fears is a best-selling thriller novel by Tom Clancy ... It was the fourth of Clancy's Jack Ryan books to be turned into a film ...

Dr. John Patrick Jack Ryan Sr., KCVQ (Hon.), Ph.D. is a fictional character created by Tom Clancy who appears in many of his novels and their respective film adaptations ...

Net Force Explorers is a series of young adult novels created by Tom Clancy and Steve Pinczenik as a spin-off of the military fiction series ...

Question: What fiction character created by Tom Clancy was turned into a film in 2002?

Answer: Jack Ryan

文本语料

大语言模型

### 问题理解 (自然语言→逻辑表达)

Who is taller, Yao Ming or the spouse of Vanessa Laine Bryant?

**Program**

Find  
Yao Ming

Find  
Vanessa Laine Bryant

Relate  
spouse

Select  
height  
greater

### 可解释推理 (问题求解器)

KBQA

IR&MRC

ChatLLM

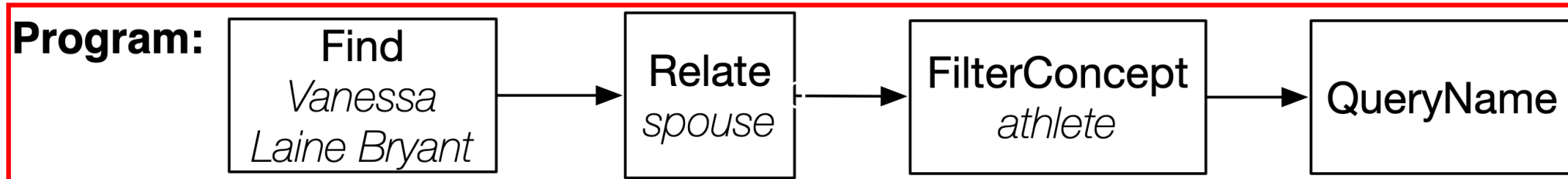
# 大语言模型知识推理

## ■ KoPL: 面向知识的可解释推理编程语言

– 将自然语言问题表示为由基本函数组成的程序，程序运行结果就是问题的答案

**Question:** Who is the athlete that is the spouse of Vanessa Laine Bryant?

**SPARQL:** `SELECT ?e WHERE { ?e <pred:instance_of> ?c . ?c <pred:name> "athlete" . ?e <spouse> ?e_1 . ?e_1 <pred:name> "Vanessa Laine Bryant" . }`



设计理念

- 显式、透明、模块化的推理过程
- 易于理解，便于人机交互
- 面向知识库、文本等不同形式的知识、可扩展性强

# 大语言模型知识推理

## ■ KoPL框架

### - 知识参数

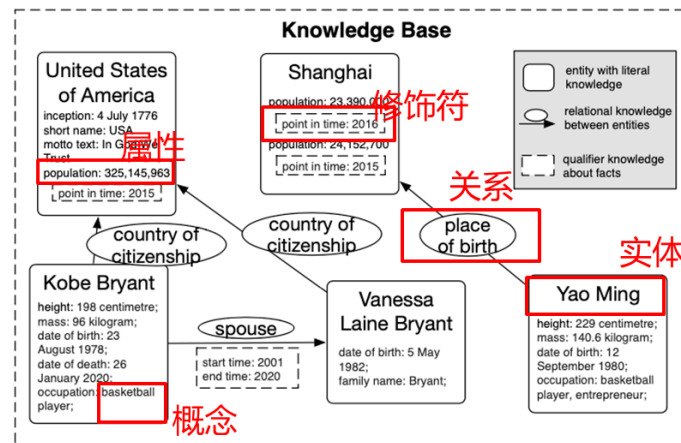
- 知识图谱：多种知识元素

### - 基本函数

- 知识操作函数
- 知识查询函数

### - 推理程序

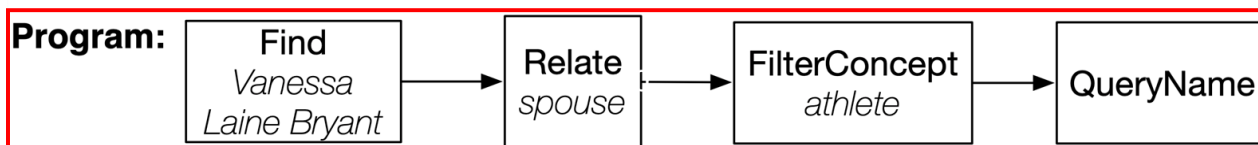
- 由函数组合而成，根据输入参数执行后得到问题答案



基本函数	功能描述	示例 (仅文本型输入)	函数型输入×文本型输入 → 输出	基本函数	功能描述	示例 (仅文本型输入)	函数型输入×文本型输入 → 输出
<i>Find</i>	找出具有特定名字的所有实体	<i>Find</i> (LeBron James)	$() \times (Name) \rightarrow (Entities)$	<i>QueryName</i>	查询实体的名字	---	$(Entity) \times () \rightarrow (string)$
<i>FilterConcept</i>	找出属于特定概念的所有实体	<i>FilterConcept</i> (athlete)	$(Entities) \times (Name) \rightarrow (Entities)$	<i>QueryAttr</i>	查询实体的特定属性值	<i>QueryAttr</i> (height)	$(Entity) \times (Key) \rightarrow (Value)$
<i>FilterStr</i>	找出具有特定属性值的所有实体	<i>FilterStr</i> (gender, male)	$(Entities) \times (Key, Value) \rightarrow (Entities, Facts)$	<i>QueryRelation</i>	查询两个实体间的关系	<i>QueryRelation</i> ()	$(Entity, Entity) \times () \rightarrow (Pred)$
<i>Relate</i>	找出与输入实体有特定关系的所有实体及对应三元组	<i>Relate</i> (father, forward)	$(Entity) \times (Pred, Dir) \rightarrow (Entities, Facts)$	<i>QueryRelationQualifier</i>	查询关系型三元组的特定修饰值	<i>QueryRelationQualifier</i> (drafted by, point in time)	$(Entity, Entity) \times (Pred, QKey) \rightarrow (Qvalue)$
<i>And</i>	返回两个实体集合的交集	---	$(Entities, Entities) \times () \rightarrow (Entities)$	<i>SelectBetween</i>	在两个实体中，查询特定属性值更大/更小的实体	<i>SelectBetween</i> (height, greater)	$(Entity, Entity) \times (Key, Op) \rightarrow (string)$
<i>Or</i>	返回两个实体集合的并集	---	$(Entities, Entities) \times () \rightarrow (Entities)$	<i>SelectAmong</i>	在一个实体集合中，查询特定属性值最大/最小的实体	<i>SelectAmong</i> (height, greatest)	$(Entities) \times (Key, Op) \rightarrow (string)$
				<i>VerifyStr</i>	验证某一属性值是否等于给定的字符串	---	$(value) \times (value) \rightarrow (boolean)$
				<i>Count</i>	查询集合中的实体数量	---	$(Entities) \times () \rightarrow (number)$

**Question:** Who is the athlete that is the spouse of Vanessa Laine Bryant?

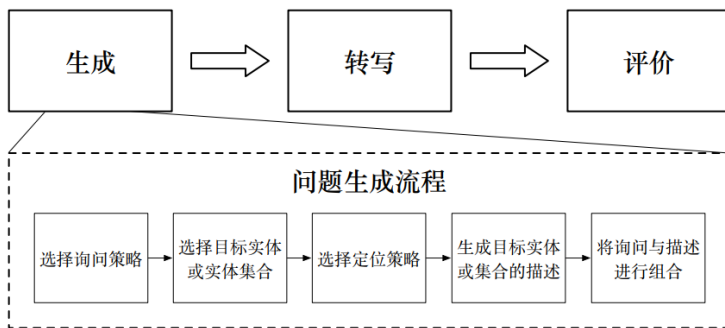
**SPARQL:** `SELECT ?e WHERE { ?e <pred:instance_of> ?c . ?c <pred:name> "athlete" . ?e <spouse> ?e_1 . ?e_1 <pred:name> "Vanessa Laine Bryant" . }`



# 大语言模型知识推理

## ■ 面向自动编程的KoPL标注数据集KQA Pro

- 核心问题：复杂性和多样性
- 解决方案：随机采样+递归组合



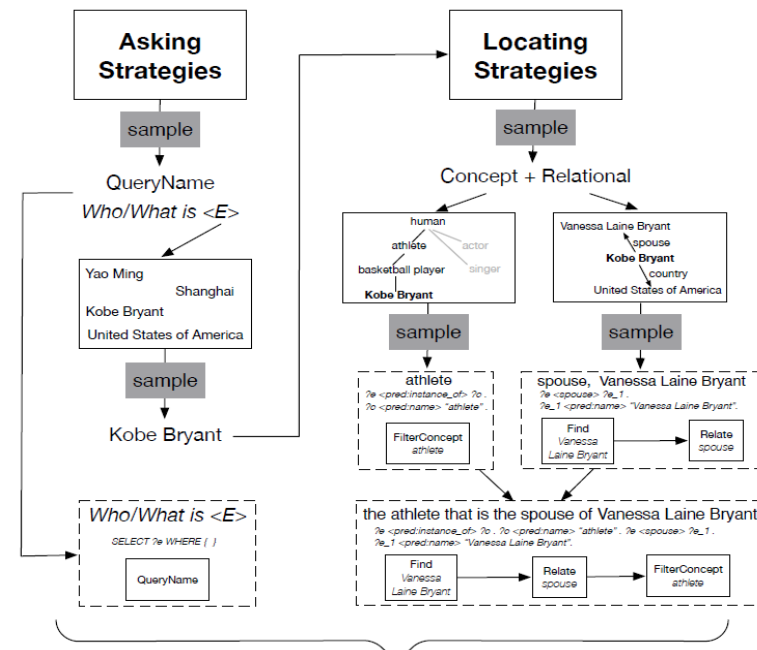
### • 9种关于实体或实体集合的问题

Asking Stage		
QueryName	What/Who is <E>	Who is the basketball player whose height is equal to 229 centimetres?
Count	How many <E>	How many basketball players that were born in Shanghai?
QueryAttribute	For <E>, what is his/hers/its <K> (<QK> is <QV>)	For Shanghai, what is its population (point in time is 2016)?
Relation	What is the relation from <E1> to <E2>	What is the relation from Kobe Bryant to United States of America?
SelectAmong	Among <E>, which one has the largest/smallest <K>	Among basketball players, which one has the largest mass?
SelectBetween	Which one has the larger/smaller <K>, <E1> or <E2>	Which one has the larger mass, Kobe Bryant or Yao Ming?
Verify	For <E>, is his/hers/its <K> <op> <V> (<QK> is <QV>)	For the human that is the spouse of Vanessa Laine Bryant, is his/her height greater than 180 centimetres?
QualifierLiteral	For <E>, his/hers/its <K> is <V>, what is the <QK>	For Shanghai, its population is 24,152,700, what is the point in time?
QualifierRelational	<E1> <P> <E2>, what is the <QK>	Kobe Bryant is the spouse of Vanessa Laine Bryant, what is the start time?

### • 7种描述实体或实体集合的方式

Locating Stage		
Entity Name	-	Yao Ming
Concept Name	<C>	basketball players
Concept + Literal	the <C> whose <K> is <op> <V> (<QK> is <QV>)	the city whose population is greater than 23,000,000 (point in time is 2016)
Concept + Relational	the <C> that <P> <E> (<QK> is <QV>)	the basketball player that was born in Shanghai
Recursive Multi-Hop	unfold <E> in a Concept + Relational description	the basketball player that was born in the city whose population is greater than 23,000,000 (point in time is 2016)
Intersection	Condition 1 and Condition 2	the humans whose height is greater than 190 centimetres and less than 220 centimetres
Union	Condition 1 or Condition 2	the humans that were born in Shanghai or New York

可以递归生成复杂问题



**Question:** Who is the athlete that is the spouse of Vanessa Laine Bryant?

**SPARQL:** `SELECT ?e WHERE { ?e <pred:instance_of> ?c . ?c <pred:name> "athlete" . ?e <spouse> ?e_1 . ?e_1 <pred:name> "Vanessa Laine Bryant" . }`

**Program:** Find Vanessa Laine Bryant → Relate spouse → FilterConcept athlete → QueryName

**Choices:** Yao Ming; Michael Jordan; Kobe Bryant; James Harden; LeBron James; Lionel Messi; Virgil van Dijk; Cristiano Ronaldo; Usain Bolt; Justin Gatlin

**Answer:** Kobe Bryant

Answer Uniqueness Check

# 大语言模型知识推理

- 面向自动编程的KoPL标注数据集KQA Pro
  - 用作评价大模型多类型推理能力（已集成至KoLA）
  - KQA Pro+大模型微调→KoPL自动编程

Model	Overall	Multi-hop	Qualifier	Comparison	Logical	Count	Verify	Zero-shot
<b>Open-Ended Setting</b>								
KVMemNet	16.61	16.50	18.47	1.17	14.99	27.31	54.70	0.06
SRN	-	12.33	-	-	-	-	-	-
EmbedKGQA	28.36	26.41	25.20	11.93	23.95	32.88	61.05	0.06
RGCN	35.07	34.00	27.61	30.03	35.85	41.91	65.88	0.00
RNN SPARQL	41.98	36.01	19.04	66.98	37.74	50.26	58.84	26.08
RNN Program	43.85	37.71	22.19	65.90	47.45	50.04	42.13	34.96
BART SPARQL	89.68	88.49	83.09	96.12	88.67	85.78	92.33	87.88
BART Program	88.55	87.12	82.28	93.87	87.92	85.33	90.95	87.25
<b>Multiple-Choice Setting</b>								
KVMemNet	39.15	36.78	50.09	18.71	38.61	34.69	56.08	6.73
EmbedKGQA	45.19	42.39	54.66	23.01	42.36	36.34	61.95	8.88
RGCN	53.75	51.70	58.24	48.64	55.88	47.93	66.57	1.46
<b>Human Performance</b>								
Human	97.50	97.24	95.65	100.00	98.18	83.33	95.24	100.00



# 大语言模型知识推理

- KoPL自动编程很难保证完全准确，是否有便捷的方式纠错呢？
- 可视化交互式编程VisKoP
  - 图形化元素，支持自动解析和拖拽式自主编程
  - 支持中间结果查看
  - 支持函数类型和参数的修改

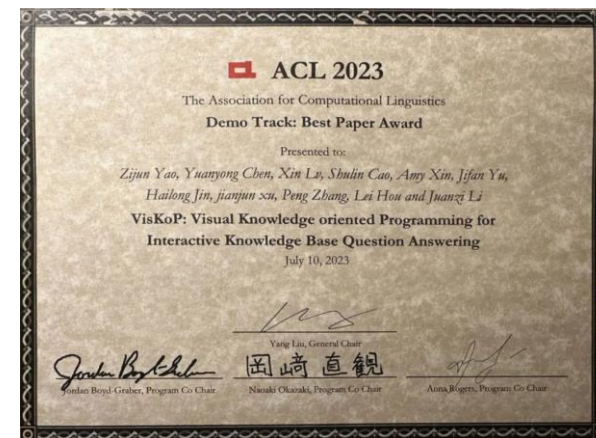
How many countries share borders with both Germany and France? Parse

17 → 4

select a function + Clear All Run on Wikidata Run Code

```
Intermediate Result
```

```
{
  0: {
    entity_label: "United Kingdom"
  },
  1: {
    entity_label: "Kingdom of Westphalia"
  },
  2: {
    entity_label: "Brazil"
  },
  3: {
    entity_label: "Germany"
  },
  4: {
    entity_label: "Andorra"
  }
}
```

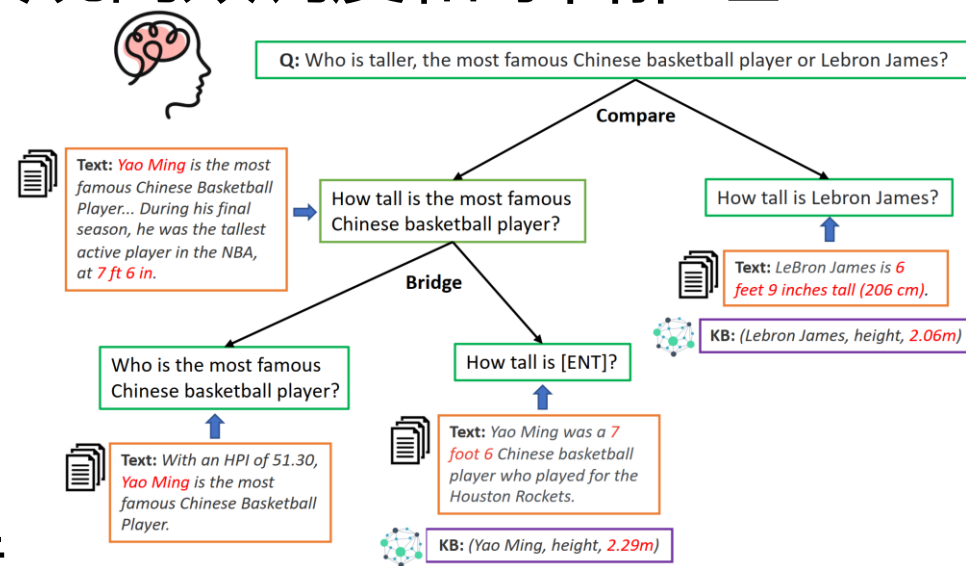


ACL 2023 Demo Track  
Best Paper Award

# 大语言模型知识推理

- KoPL中基本函数不限于知识图谱，可以在多类型资源上实现
- 复杂问题求解需要利用多种知识源，如何实现高效调度和综合推理？

- 知识图谱
- 文本语料
- 大语言模型

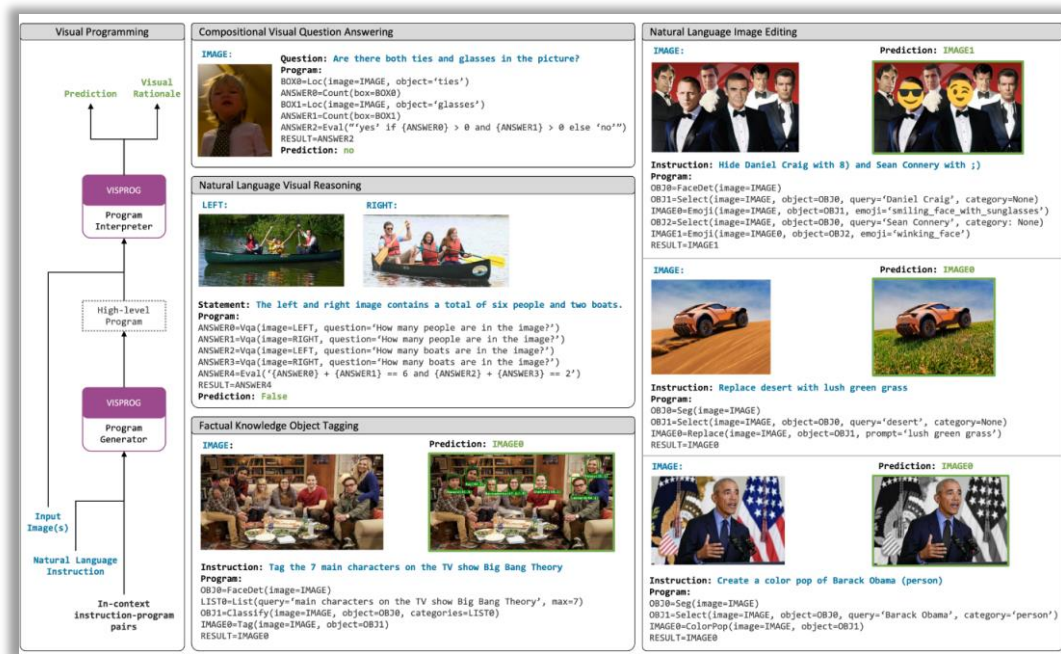
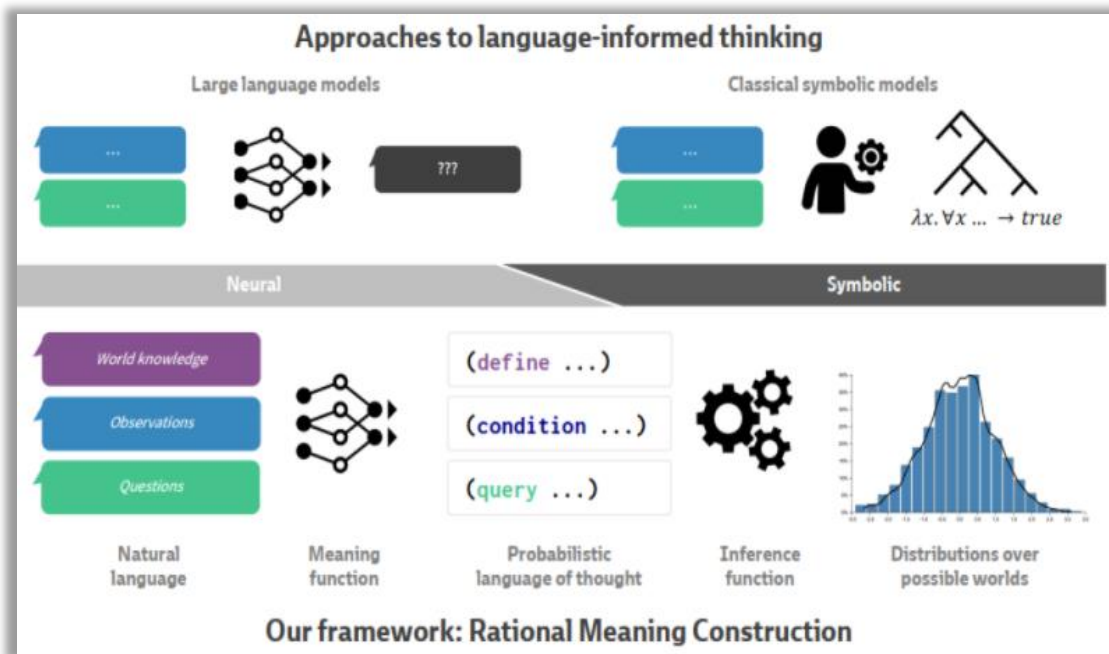


- RoHT: 基于问题层次分解的多源知识推理

- 层次化问题分解树: 可以用于控制不同层次的问题用不同类型的知识源回答
- 概率化推理: 自底向上的概率融合推理

# 大语言模型知识推理

- 可解释编程已经成为大模型推理能力增强的发展趋势
  - 利用大模型的语言理解能力将复杂任务解析成程序
  - 利用程序去操控多类型资源中的知识对象



MIT认知实验室: Word Model → World Model

CVPR 2023最佳论文: Visual Programming

# 方向3：大模型知识获取和应用机理

## ■ Knowledge Engineering **FROM** the LLM

– 如何使大模型具有真正的符号系统特征，并基于此构建符号主义智能

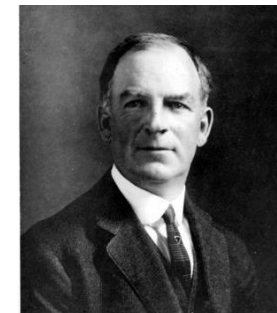
- 应以何种方式回答/探索该问题？
- 现有符号特征/能力的形成机理和边界？
- 还欠缺哪些符号特征/能力？
- 现有体系如何适配大模型可能遵循的新符号系统？

– 大模型的多模态认知、具身智能与世界模型

# 大模型知识获取和应用机理

## ■ 理解大模型机理对于应用非常重要

- Rising AI power necessitates control to prevent potential risks.
- Unveiling AI's inner workings might increase its usability.
- Mechanistic interpretability could improve their trustworthiness.
- **Knowledge representation** and phase transitions are vital for understanding AI.
- Physics could potentially contribute to the understanding of AI systems, leading to new discoveries.
- **Extracting knowledge from AI systems** is crucial for informed decision-making.



——摘自: *Max Tegmark. Keeping AI under control through mechanistic interpretability*

**MIT研讨会: *The impact of chatGPT and other large language models on physics research and education***

<https://indico.mit.edu/event/759/>

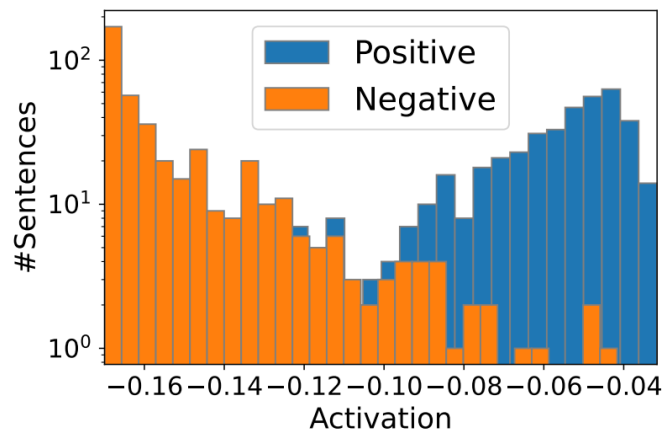
# 大模型知识获取和应用机理

## ■ Physics of AI

- 将大模型视为一个“未知”的系统
- 使用物理学研究方法，从“观测现象”来“推测解构”

## ■ 初步探索

- 模型内是否存在完成特定任务的神经元？
- 提示学习是否具有可迁移性？



Source Task	IMDB	SST-2	laptop	restaurant	Movie	Tweet	MNLI	QNLI	SNLI	deontology	justice	QQP	MRPC	random prompt
IMDB	100	95	65	84	101	52	37	55	37	58	63	43	82	
SST-2	91	100	88	92	93	66	50	59	38	61	62	57	66	
laptop	76	91	100	93	84	74	38	55	37	59	63	43	84	
restaurant	80	92	95	100	81	70	38	55	37	59	62	44	81	
Movie	98	80	70	40	100	54	37	55	37	59	62	62	69	
Tweet	88	94	66	90	94	100	41	55	37	59	62	43	80	
MNLI	55	61	70	62	61	54	100	79	62	60	62	72	81	
QNLI	75	53	3	69	80	54	60	100	65	59	61	65	39	
SNLI	55	53	64	68	58	54	87	82	100	59	62	51	84	
deontology	63	54	5	5	59	58	38	55	38	100	80	48	75	
justice	55	79	64	58	82	46	38	55	37	83	100	49	51	
QQP	55	53	68	8	59	54	43	58	37	59	62	100	78	
MRPC	59	53	3	1	59	54	38	54	36	59	62	78	100	
random prompt	54	52	3	2	59	54	38	55	36	58	62	46	75	

Source Task	IMDB	SST-2	laptop	restaurant	Movie	Tweet	MNLI	QNLI	SNLI	deontology	justice	QQP	MRPC	SQuAD	NQ-Open	Multi-News	SAMSum	random prompt
IMDB	100	96	79	87	98	65	36	52	34	58	54	67	39	0	1	0	0	
SST-2	84	100	89	88	67	69	35	55	35	58	56	45	67	0	0	0	0	
laptop	90	86	100	90	83	76	36	53	36	57	54	41	63	0	0	0	0	
restaurant	90	92	101	100	81	77	36	53	33	57	57	42	68	0	0	0	0	
Movie	100	91	81	87	100	68	38	53	37	62	59	55	46	0	1	0	0	
Tweet	96	92	100	91	84	100	33	53	36	57	56	45	67	0	0	0	0	
MNLI	65	81	60	45	53	43	100	81	98	57	54	41	69	1	2	4	0	
QNLI	62	52	70	73	52	56	59	100	64	57	54	41	69	1	1	1	0	
SNLI	64	66	18	20	53	22	96	76	100	57	54	70	33	0	1	1	0	
deontology	53	60	41	42	53	30	37	56	36	100	74	63	59	0	0	0	0	
justice	51	50	26	19	53	55	44	52	41	58	100	41	69	0	0	0	0	
QQP	51	51	26	20	53	22	36	53	36	58	54	100	78	1	0	0	0	
MRPC	51	50	28	20	53	21	49	56	48	58	54	84	100	0	0	0	0	
SQuAD	73	82	69	73	60	63	40	53	38	58	58	48	62	100	20	33	33	
NQ-Open	73	75	62	47	53	55	42	58	36	56	62	51	50	16	100	23	13	
Multi-News	62	76	26	19	53	21	39	52	36	57	54	70	33	6	25	100	28	
SAMSum	76	77	68	75	51	57	36	53	36	57	54	43	62	14	15	67	100	
random prompt	52	50	26	19	53	22	35	51	35	57	54	41	69	0	0	0	0	

Xiaozhi Wang et al. Finding Skill Neurons in Pre-trained Transformer-based Language Models. EMNLP 2022  
 Yusheng Su, Xiaozhi Wang et al. On Transferability of Prompt Tuning for Natural Language Processing. NAACL 2022

# 总结

## ■ 知识工程的目标

- 知识积累与重用、知识推理与决策

## ■ 大模型对知识工程的冲击

- 知识不局限于符号化表示、符号推理能力能够通过参数化实现

## ■ 未来知识工程研究：不忘初心，拥抱大模型

- **BY**：借力大模型，完成更好的知识积累和推理决策

- **FOR**：大模型知识增强

- **FROM**：探索大模型内部机理，更好的“符号对齐”或建立“新符号系统”

# 迎接大模型时代知识工程的 新机遇、新挑战和新任务

**感谢聆听!**  
**敬请批评指正!**