

学术观点

浅谈大模型时代的知识图谱技术栈

陈华钧
浙江大学

关键词：大型语言模型 知识图谱 知识提示学习

语言VS.知识

ChatGPT：世界知识的神经网络化

人类在认识世界的过程中不断积累世界知识，而语言则是表示世界知识最直接的载体。到目前为止，人类的绝大部分知识都是通过自然语言来描述、记录和传承的。以 ChatGPT 为代表的大型语言模型 (Large Language Model, LLM) 将海量、以序列为主的世界知识预先学习进神经网络中，并在参数化空间对知识进行处理和操作 (见图 1)。与以算法为主的传统模型不同，ChatGPT 训练的是富含知识的模型，是知识模型与算法模型的混合体。

自然语言、知识图谱和语言模型都是表示知识的方法

自然语言以符号序列的方式表示知识，而知识图谱 (KG) 利用图的结构刻画抽象概念以及世界万物之间的关系，代表一类结构化的知识表示方法。典型的结构表示包括：层次结构 (如概念图谱)、关联结构 (如实体关系图谱)、时序结构 (如事理图谱)、逻辑结构 (如逻辑规则) 等。与自然语言一样，知识图谱首先是一种显式的知识表示方法，人可

理解，可解释性好。随着神经网络与知识图谱的不断融合，研究人员进一步发展出知识图谱嵌入、图表示学习和预训练等向量化、数值化或参数化的隐性知识表示方法。

大型语言模型利用神经网络从超大规模的文字语料中学习文字的组合、交互和涌现规律，实现在神经网络参数空间对基本的文字语义编码。与自然语言和知识图谱不同，大型语言模型是完全参数化、数值化的知识表示方法，对人而言难以理解，但更便于机器计算。三者的区别如表 1 所示。

变与不变：新的知识图谱技术栈

大模型普遍面临训练和应用成本高昂、生成内容不可控、知识可信度低、专业性知识处理能力弱



图1 世界知识的神经网络化

表1 自然语言、知识图谱与语言模型都是表示知识的方法

表示方法	可解释性	知识类型	可计算性
自然语言	人可理解	显性知识	不易于计算处理
知识图谱	人可理解	显性知识+隐性知识	较容易计算处理
语言模型	人难以理解	隐性知识	易于计算和处理

等困难，而知识图谱则一直面临知识覆盖度低、长尾知识抽取难、常识推理能力弱等困难。两者有机结合，可以提供综合能力更强的知识处理技术手段。下面分别从大模型技术栈中的知识图谱、知识图谱技术栈中的大模型两个角度展开介绍。

大模型技术栈中的知识图谱

大型语言模型核心技术要素包括预训练、指令精调、提示工程、思维链、奖励模型与人机对齐等，如图2所示。知识图谱相关的技术理念和要素在大模型技术栈的各个环节都可以发挥作用。

知识增强与结构增强

越来越多的研究表明，不论是在预训练阶段，还是在指令提示阶段，提升语料的语义规范性或数据的

结构化水平均有助于增强模型推理能力。例如，代码语言的结构化比自然语言的更好，前者包含大量的运算逻辑，不仅具有代码生成能力，也更利于激活模型推理能力。思维链比普通的自由文本包含更多的逻辑关联性描述，更有利于提升模型推理能力。

我们可以按照结构化水平或语义规范化程度把知识大致分为3个层级：文本序列知识（包括自然文本序列、带思维链的自然文本）、结构化知识（概念层次结构、实体关联结构、时序结构）、逻辑规则知识。语料中知识结构化水平越高，数据质量越高，更有助于提升模型性能。但语义逻辑的规范化程度越高，意味着语料获取的难度越大，规模不容易扩大。因此，提升语料的知识结构化水平可以增强模型能力，但与提升模型规模之间存在权衡。

提示工程即知识工程

在大型语言模型中，提示（prompts）是指在训练或使用模型时使用的一类引导性指令，如问题描述、上下文信息、举例示例、结构模板、思维链等。通过提供明确的提示或指令，模型可以更好地理解人的意图和要求，产生更加遵循人类指令要求

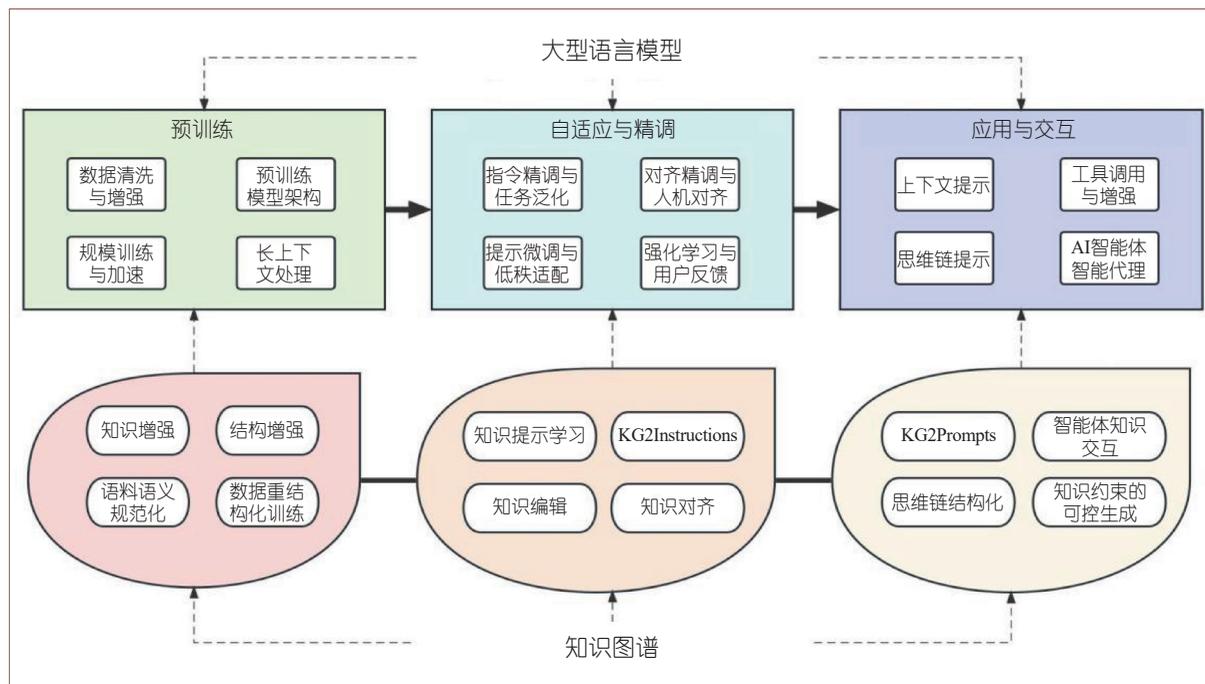


图2 大模型技术栈中的知识图谱

的输出。

大量实践表明，提示设计得好坏和详细程度对模型输出结果影响巨大。在预训练阶段使用的诸如“Next Token Prediction”的监督信号仅能编码基础、通用和浅层的语义。而“提示”则包含大量任务特定的知识，这些知识不只是描述人的意图和问题指令，还可能包括人的思维逻辑（如思维链）、类比逻辑（如举例示例）、与任务相关的实体及关联关系等事实性上下文信息等。构建和获取高质量的提示语料（即提示工程）本质上就是获取人类先验知识的过程，这和传统知识工程的目的本质上是相似的，也都是耗时费力的。事实上，大模型体现出的“智能”很大程度上是由提示工程阶段植入的知识驱动的。

知识图谱与指令精调

指令精调是指在预训练模型的基础上，通过将任务特定的监督数据转化为指令或提示形式，进一步训练和微调以提升模型的性能和准确性。由于各种不同任务的数据都以统一的指令形式输入到模型进行训练，这使得模型具有更好的任务泛化能力。

与自由文本相比，知识图谱是一种逻辑结构关联更加丰富、知识密度更高的数据语料。知识图谱中的知识可用来增强指令提示或直接用于生成指令数据集。例如，知识图谱中的实体属性和关联信息可以用于丰富指令的上下文信息。通过查询知识图谱，可以获取实体之间的关联和属性，并作为提示工程的一部分。更为复杂的知识结构可直接用于辅助构建复杂的提示指令。很多领域的知识图谱数据

由长久人工积累（如 Gene Ontology）或天然的结构化数据转化而来。我们可以充分挖掘和利用知识图谱中包含的知识结构（如层次概念、实体关系、事理时序、规则逻辑等），增强和丰富提示指令或辅助构建逻辑性更强的指令数据集。

图3展示了一个利用知识图谱直接构建知识抽取指令的例子。该模型基于知识图谱转换指令（KG2Instructions）技术产生的大量指令数据，提高语言模型对人类抽取指令的理解，具体基于维基百科和维基数据知识图谱，通过远程监督、模式（schema）约束过滤等方法构建大量的指令数据，并通过随机采样人工指令模板的方式提升指令的泛化性。

知识图谱与思维链结构化

思维链（Chain of Thoughts, CoT）是一类特殊的提示，它模拟了人类在思考和解决问题时，通过联想将不同的概念或知识关联起来的过程，从而引导大模型模仿人的逻辑思维过程。知识图谱显然是表示和生成思维链的好帮手。已有研究表明，采用结构化更好的思维链表示形式（如思维树（tree of thoughts）、代码思维（program of thoughts）、思维图（graph of thoughts）等）比自然语言思维链更有利于提升模型推理能力。随着思维链结构化水平的提升，模型推理能力也将得到提升。

人们可以按自己的兴趣和要求，将不同的概念和思维步骤关联起来，形成逻辑更加清晰的思维链知识图谱（KGCoT），这种结构化的思维链可以直接作为提示的一部分输入到模型中。另外，通过实体链接技术将知识图谱中的概念或实体与手工编写的思维链进行关联，可以便利地扩展和增强自然语言形式的思维链，

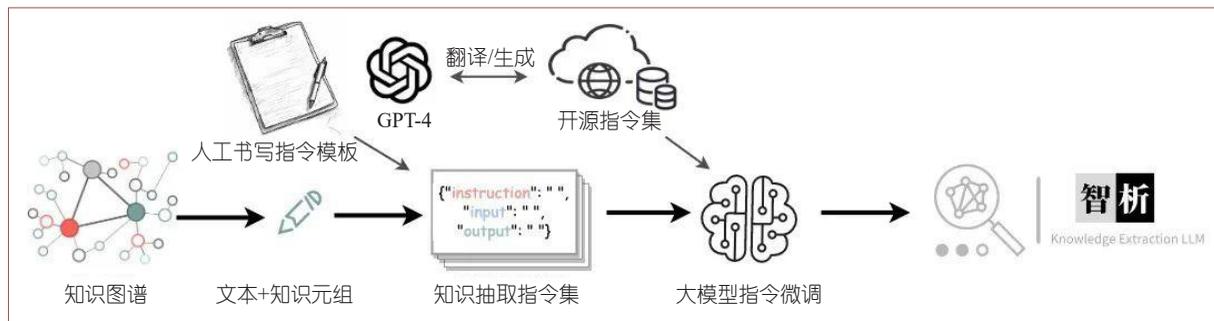


图3 利用知识图谱直接构建知识抽取指令的示例

使得思维链更加接近人的联想思维模式。

奖励模型与知识对齐

人机对齐是指通过训练一个奖励模型，将大模型的输出与人的价值观或道德伦理进行对齐，以对模型的行为进行安全性约束，避免产生对人类不利的内容或行为。常见的做法是通过人工比较打分训练一个奖励模型，再用这个奖励模型进一步优化模型的回答。由于知识图谱中的很多知识都是经过人工审核和校验的，相对标准和正确，可以为奖励模型提供反馈信号。通过将大模型返回的结果与知识图谱中的事实、关联或约束进行比较来提供反馈信号，并改进生成的结果和衡量模型的性能。

知识图谱与工具调用及AI智能体

工具增强的语言模型 (tool-augmented language model) 在模型自身无法解决问题时可以调用外部工具来完成。大模型自主地学习这种复杂的工具调用逻辑仍然是有困难的，而知识图谱可用于复杂 API 组合调用逻辑的建模，这在传统服务计算架构领域已经有很多实践。更一般化的工具调用是 AI 智能体之间的智能交互，即智能体互为工具，当某个智能体无法完成特定任务时，将会向外寻求其他多智能体的协作和帮助。知识图谱和多智能体有着深刻的历史渊源。早期的符号

化知识图谱表示语言 (如 DAML) 本身就是为智能代理之间进行交流和通信而设计的。知识图谱可以和语言模型一起实现更加自主、更加可靠的多智能体协作与交互。

知识图谱技术栈中的大模型

知识图谱的核心技术要素包括表示、存储、抽取、推理、融合、推理、问答、分析等。大模型技术在知识图谱的各个环节均能发挥作用，如图 4 所示。

自动化本体构建与概念抽象

本体 (ontology) 和模式定义了概念、类、属性、关系以及它们的层次结构。在很多领域，构建这类以概念或类目属性的层次体系为主的概念图谱 (concept KG) 具有重要的实用价值。例如电商领域通常需要维护一个细粒度的商品类目及属性体系，以支持商品发布、搜索问答、品类规划、营销预测等一系列下游任务。大型语言模型从海量的文字语料中学习到丰富的有关字词及概念的知识，因而对概念的识别、理解、抽象能力表现突出。这为自动化的概念层次构建、类目扩展、属性补齐、本体对齐、概念归一化及更新处理提供了更强有力的技术手段。

知识抽取大模型

从文本、图像、视频等多种模态的数据中识别实体和对象，并抽取和建立这些实体之间的关

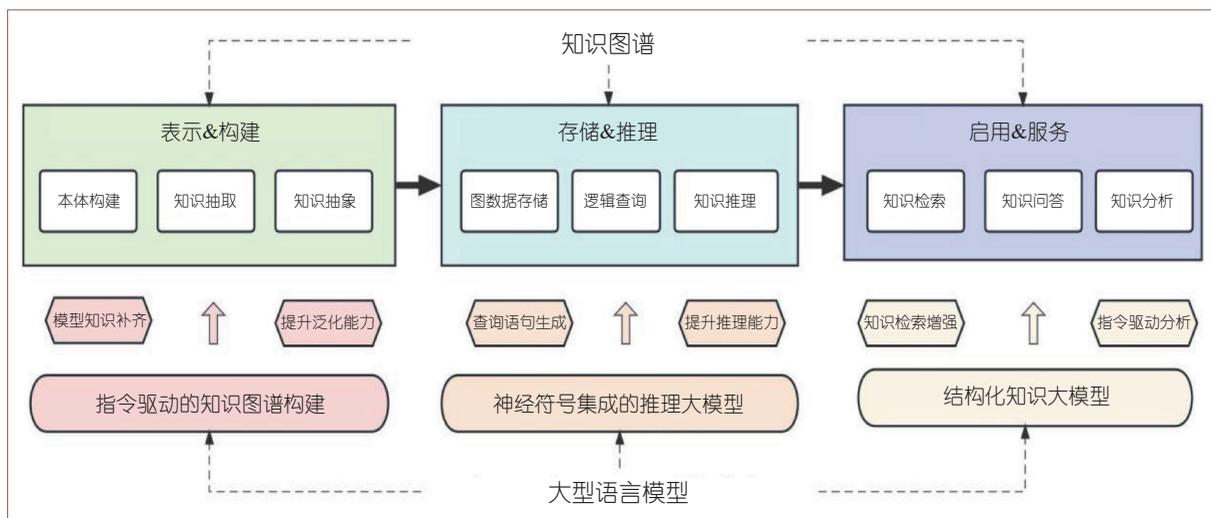


图4 知识图谱技术栈中的大模型

联关系，再进一步提取出各类事件要素，是很多领域处理大数据的基本任务，也是知识图谱构建的基本任务。我们可以基于指令驱动的方法构建专门面向知识抽取的大模型，大量研究表明，知识抽取类大模型在知识图谱构建方面具有独特的能力。(1)模型脑补：大模型学习海量的参数化知识，这使得可以同时从文本语料和参数空间抽取知识，通过模型的“见多识广”对文本抽取的结果直接进行模型脑补，可以大幅提升知识图谱构建的效率；(2)更强泛化：大模型通过指令学习和人类反馈具备了较强的泛化能力，通过指令驱动的抽取大模型具备更强的抽取新类型、新关系、新事件的泛化能力。研究表明，即使对于从未见过的实体或关系类型，大模型同样具有高精度的识别能力；(3)生成补齐：大模型具有较强的生成能力，当模型或文本中缺乏所需的知识时，大模型可利用其生成能力通过知识生成补齐知识图谱。生成式的知识图谱本质上利用了大模型的推理能力，但可能会生成错误知识。

图数据库查询与问答

对于存储在图数据库中的知识图谱，通常需要利用图查询语言（如 SPARQL、Cypher、Gremlin 等）对图数据进行查询、删除、更新等操作。大模型的语言理解能力使我们可以较为精确地实现自然语言到结构化查询语言的自动化翻译，从而实现针对图数据库的自然语言查询调用。这极大地便利了图数据库的管理与维护，以及基于图数据库的应用开发。例如，大模型框架 LangChain 增加了 Cypher 插件，可以方便地从自然语言生成 Cypher 图查询语言。

结构化知识大模型

现有的大模型多基于文本语料训练，因而较为擅长语言理解类任务。一些研究表明，现有大模型在处理天然依赖结构化数据的任务方面，除查询问答类依赖自然语言生成结构化查询的任务表现较好以外，其他诸如链接预测、关联挖掘、子图分析、时序预测等任务的表现仍然无法超过传统小模型。但大模型的知识预先训练、指令驱动的泛化能力、奖励模型与对齐学习等思想均可用来增强结构化知识的处理能力。我们可以训练专门针对结构化知识

的大模型，例如设计更加适合编码结构化知识的基础模型训练机制、将结构化数据指令化实现指令驱动的结构化任务处理、面向结构化数据任务的奖励模型等，提高大模型分析和处理结构化知识的能力，这对于很多依赖结构化数据的大数据分析类应用领域具有重要价值。

神经符号集成的大模型推理

常见的知识图谱推理方法一般分为基于符号表示的推理和基于嵌入或图表示学习的神经网络推理两种方法。符号方法基于显示知识，推理过程明确可解释，答案确定，但泛化能力差；神经网络方法基于隐式知识，推理泛化能力好，但不能给出确定答案。因此，研究人员研究出将两类方法相结合的神神经符号集成的推理方法。

大型语言模型把自然语言形式的符号知识参数化，也是一种在神经网络空间实现的推理。但与传统知识图谱嵌入等推理方法不同，大模型依靠“见多识广”的知识丰富性和生成能力实现更强的泛化能力。因此，大型语言模型可以和传统符号知识图谱推理方法形成更好的互补。例如，一方面，大模型可以将符号知识图谱作为外部可调用的工具，在需要确定答案时，利用符号知识图谱的推理给出更加可靠、明确和确定的答案。另一方面，在符号知识图谱推理因为缺乏知识而推不出答案时，可以利用大模型的见多识广和高泛化能力给出粗粒度的回答。

大模型时代的知识处理：机遇与挑战

语言不等于知识，语言是知识的载体，语言模型则是知识处理的手段。以 ChatGPT 为代表的大型语言模型技术的发展代表人工智能表示和处理知识能力的大幅提升，同时也带来一系列机遇与挑战。

大型知识模型

大型语言模型以处理语言文字序列为输入，以基于“Next Token Prediction”等针对序列结构的自回归模型为基本架构。但人类知识是高度复杂的，

除了包括以字词为基本单元组成的文字序列外，还包括概念的层次结构、图结构、时序结构、本体及规则逻辑等。人们在描述客观世界时，更多采用的是多样化的结构化描述，而非单一的序列化描述。序列结构并不是知识的最好表示结构，自回归模型也并不能有效编码各种维度更高的知识结构。我们最终需要发展的是能处理各种知识表示结构的大型知识模型（large knowledge model）。

大模型知识编辑

幻觉（hallucination）出错以及不确切的回答是大模型智能得以涌现需要付出的代价，同时也是大模型得以实际应用必须要解决的技术挑战。模型编辑技术实现了在神经网络空间对参数化的知识进行增删改，以纠正错误的知识，删除有害的知识，约束错误的生成。但在极大规模的参数空间对知识进行编辑并不是一个简单的工作，这首先需要在神经网络中定位错误的知识、确定修改的边界、更新参数的机制和方式等。由于神经网络的参数之间是高度关联的，改动其中任何一个神经元都会对与之相关的其他神经元产生未知的影响，这些影响有可能产生更多的无法预知的错误。因此需要开发更加鲁棒、可靠、可追溯的模型知识编辑技术。

知识约束的可控AIGC

大模型的人工智能生成内容（AIGC）的能力很突出，但在生成逻辑性更严谨、场景更复杂、更符合人的意图的内容方面仍然面临众多挑战。有研究表明，与单纯从文本生成图片相比，基于知识图谱转化成关联逻辑更丰富的文本，再生成图片，其图片内容更加丰富，关联逻辑性更强。因此，利用知识（图谱）约束和引导大模型的生成能力，减少生成模型的逻辑错乱问题，提升生成内容的丰富和复杂程度，实现知识约束的可控内容生成（KGGC）是未来的一个发展方向。

大模型驱动的多智能体知识协同

早期知识图谱技术（如RDF/OWL等）设计

的本意是为了便于互联网多智能体进行知识获取，并作为多智能体之间的知识交换格式（knowledge interchange format）。例如OWL语言的前身之一DAML正是DARPA Agent Markup Language的缩写。语言模型技术的成熟为多智能体之间进行知识交换和知识交互提供了新的方式。工具增强的大模型可以帮助多智能体自主学习其他多智能体的调用和交互方式，自主寻求其他多智能体的协助，并且无须基于由人工定义的知识交换接口。这使得多智能体之间能更加自主地实现知识协同，建立更加强大的智能体协作社区。

道德知识表示与价值观知识对齐

道德与价值观本质上也是人们通过学习得到的一种评价知识。对大模型做对齐操作本质上是和道德与价值观的知识进行最终对齐。一方面，需要研究如何表示道德与价值观知识，例如不少人工智能伦理工作者正在研究道德与价值观本体的表示与构建方法，很多法律大模型的训练本身也涉及模型与法律知识的对齐。另一方面，需要研究知识对齐的方法和机制，例如如何利用这种知识训练奖励模型，或者约束模型的输出结果等。

知识的表示与处理是人工智能自诞生以来的核心目标，而同样用来处理知识的大型语言模型和知识图谱技术各有优缺点，两者结合必将为人工智能提供更加强大、可靠、可控、负责任的知识处理方法。



陈华钧

CCF 专业会员、CCF 杰出演讲者。浙江大学计算机科学与技术学院教授。主要研究方向为知识图谱、大数据系统、自然语言处理等。
huajunsir@zju.edu.cn

（本文责任编辑：郭得科）