

文章编号: 1003-0077(2024)12-0001-17

面向自然语言理解的语义表示方法综述

张翔¹, 何世柱^{1,2}, 张元哲^{1,2}, 刘康^{1,2}, 赵军^{1,2}

(1. 中国科学院大学 人工智能学院, 北京 100190;

2. 中国科学院 自动化研究所 复杂系统认知与决策实验室, 北京 100190)

摘要: 语义是自然语言理解的核心研究对象, 让机器掌握语义有多种途径, 因此产生了不同的语义表示方法。但是, 这些不同的途径之间往往缺乏联系, 分别散落于不同的研究领域和研究任务中, 如知识库问答中的查询图和 SparQL、表格问答中的 SQL、句子语义分析中的框架语义和 AMR 图等。虽然形式相近, 但相关研究却缺乏协同。随着研究深入, 语义表示之间缺少对比、具体任务中语义表示难以选择且性能差异大等缺点也显露出来。为了缓解这个问题, 该文综述了常见于各类任务中的语义表示, 并以世界和语言的关系为主线将它们重新划分为“外延世界语义表示”和“语言内部语义表示”两大类。总结了前者的技术研究热点及新型语义表示的设计和对比方法, 并简单探讨了近期围绕后者是否真正含有语义的辩论。最后, 该文综述了结合外延和内部两类语义表示的研究, 发现这类方法有较强潜力。

关键词: 语义表示; 语义解析; 自然语言理解

中图分类号: TP391

文献标识码: A

A Survey of Semantic Representation Methods for Natural Language Understanding

ZHANG Xiang¹, HE Shizhu^{1,2}, ZHANG Yuanzhe^{1,2}, LIU Kang^{1,2}, ZHAO Jun^{1,2}

(1. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100190, China;

2. The Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Science, Beijing 100190, China)

Abstract: Semantic is the core research object of NLU studies. There are multiple approaches to capture semantics, leading to various semantic representations. These approaches are scattered across domains and tasks, such as the query graphs and SparQL for knowledge graph question answering, the SQL for table question answering, and the Frame and AMR for parsing sentence semantics. This paper reviews common semantic representations and categorizes them into two types: External World Semantics (EWS) and the Language Interior Semantics (LIS), following the notable world-language relation. For the former, this paper discusses the research trendings, the design of new representations, and the comparison metrics. For the latter, this paper briefly summarizes the debate around whether LIS truly contains semantics. Finally, this paper introduces researches combining both types, revealing their strong potential.

Keywords: semantic representation; semantic parsing; natural language understanding

0 引言

知识表示和知识推理是人工智能和自然语言处理领域的两个关键研究内容^[1]。一个理想的自然语言处理系统应当掌握的知识就是关于自然语言的知

识, 本文关注该系统中的一个转换构件: 如何提取自然语言表达的知识内容并转换为其对应的语义表示, 以便更好地支撑后续任务(如问答、推荐、对话等)。朴素地说, 语义就是人们理解了某个语言中的每个言语后所掌握的那个东西^[2]。因此, 对于自然语言处理系统, 知识表示的核心就在于设计语义表

收稿日期: 2024-02-27 定稿日期: 2024-04-01

基金项目: 国家重点研究与发展计划(2022YFF0711900); 国家自然科学基金(62376270, 62276264)

示。为了验证这一形而上学的推论,本文从“如何表示语义”这一问题出发,对常见的语言理解系统进行综述,并通过对比研究对未来工作进行一些展望。

设计一种语义表示方法的首要工作就是对语义内容进行规划,这是极具挑战的任务。一方面,由于研究背景不同,“语义”一词在上下文中的侧重点不同,有可能只是作形容词使用,而不是提出一种语义理论。例如,谓词逻辑中语义是一种真值;语义角色中“语义”标签比词汇和句法更深入;认知科学强调语义是有别于感知的心灵印象;语义网的“语义”则强调为万维网加上机器可读的形式逻辑基础设施。另一方面,即使是语义学中明确提出的语义理论,相互之间也存在较大差异,它们既受不同的哲学思潮影响^[3],又有鲜明的时代地域特征^[4]。为了应对这一困难,Block 将语义相关工作分为语言学语义和形而上学语义两类^[5]。本文专注前者,即“特定自然语言中特定表达式的意义,以及它们怎样组合出更大的表达式的意义”。因此,本文中“语义(semantic)”和“自然语言的意义(meaning)”等价,并且“语义表示”仅限于自然语言理解系统中使用的表示。对于一个自然语言句子,设计语义表示需要回答这一研究问题(Research Question,RQ):如果语义是一种实体,它是什么?如果不是,它从哪里来?

为了回答该问题,本文将从两个方向选取用于分析讨论的语义表示。一个方向是为解决实际需求而设计使用的语义表示,相关工作涉及语义解析(semantic parsing)、知识问答、词嵌入、大语言模型等自然语言处理的任务和技术。这些系统具有一定代表性,且往往主张自己刻画了语义的某些属性,即“理解了语义”。另一个方向是从理性角度出发的意义理论,它们曾为语义的自动处理提供了参考^[6],尽管有的工作更多地是以语义为媒介讨论哲学内部问题^[2-3,7]。通过两个方向的相互溯源和结合,借助“语言和世界的关系”这一组经典概念,对大量的不同语义表示方法进行分类和理性比较便成为了可能。新的分类既有助于综述当前自然语言理解尤其是大语言模型等技术的发展成果,同时也便于展望以不同语义表示为基础的将来研究。

本文从学术搜索引擎收集文献素材。具体步骤包括:①通过“semantic parsing”关键词在谷歌学术中搜索从 2021 年起的文献,将标题输入 Semantic Scholar API,保留最相关的 3 条结果及其摘要和引用量,共获得 2 258 篇文献。②使用 meaning repr、semantic pars、AMR、text-to-sql 等前缀短语过滤摘

要,并排除 video scene、semantic segmentation 等图像领域术语,同时对 ACL Anthology 中近 3 年的文献应用同样的过滤规则,二者合并获得 2 887 篇文献。③针对 2023 年的所有文献和更早的高引文献共 373 篇进行人工标注,最终获得与语义表示研究有关的 213 篇文献。本文针对它们进行讨论。

相比其他相关综述^[8-14],本文的区别和贡献如下:

(1) 以句子的语义表示而非任务的视角进行综述;

(2) 将任务中的具体语义表示沿着理论方向追溯,便于对比多样化语义表示之间有何异同;

(3) 联系当前大语言模型的突出表现,对符号式和向量式系统怎样结合的老问题,从语义表示的新角度进行探讨。

本文组织结构如下:第 1 节综述从需求出发的常见语义表示;第 2 节综述从理性角度出发的语义表示并介绍两种分类;第 3 节讨论两个方向的语义表示的关联,并据此回答前述问题 RQ;第 4 节介绍基于特定语义表示的近期进展和遇到的问题;第 5 节介绍结合多种语义表示的主要工作;第 6 节做出总结。

1 需求出发的语义表示

在实际应用中,一般并不需要完善的语义理解能力,只要能自动处理有限的特定场景输入的句子,该系统的语义表示就可以纳入对比。因此,本节根据研究背景不同,分别考虑语义表示的三种方法论途径:逻辑途径、认知途径、分布途径。

1.1 逻辑途径的语义表示

逻辑途径的语义表示常见于语义解析、知识库问答、代码生成等任务,是最具代表性的语义表示。主要包括基于编程语言的、基于逻辑系统的、基于图的语义表示等三类^[9-10]。一般需要人类专家介入,结合具体使用场景,定义语义表示中的所有符号,并提供适当的标注数据集。

基于编程语言的语义表示主要应用于有具体的可执行环境的任务,依赖语言的运行时,所得表示可以直接执行并得到最终结果。常见的选择有 Java、Python、SQL、SparQL 等。基于逻辑系统的语义表示有命题逻辑和一阶逻辑等。它们无法表达对集

合的操作,一般与 λ 演算结合使用^[15]。实际使用时需要依赖 Lisp 等通用编程语言为其定义运行时,因此又能视为领域特定语言(Domain Specific Language,DSL)。基于图的语义表示(Graph-based Representation)泛指采用图数据结构呈现的语义。其中一类以语言学主要理论为指导,设计机器可读

的标签并尽可能囊括重要的语义学概念,如谓词论元结构、时体气态(Tense-Aspect-Mood, TAM)标注、共指关系等,代表工作有 MRS^[16]、AMR^[17]、UMR^[18]。另一类则从特定领域需要为出发点定义查询图,例如,将知识库的实体和关系定义为图的节点和边。三类语义表示的示例如表 1 所示。

表 1 逻辑和认知途径的语义表示样例

类别	语义表示	样例
逻辑途径 (1.1 节)	一阶逻辑	例句: 西游记电视剧中有多少人出演过唐僧? $\text{count}(\lambda x. \exists y. \text{cast}(\text{JourneyToTheWest}, y) \wedge \text{actor}(y, x) \wedge \text{character}(y, \text{TangSanzang}))$
	SQL 语言	例句: 西游记有多少人出演过唐僧? $\text{select count(actor.name) from actor}$ $\text{inner join cast on actor.name=cast.actor_name}$ $\text{inner join drama on drama.name=cast.drama_name}$ $\text{where drama.name="西游记" and cast.character="唐三藏"}$
	AMR	例句: the destruction of the city by the God (d / destroy-01 :ARG0 (g / God) :ARG1 (c / city))
	查询图 ^[19]	例句: 86 版西游记中谁最先出演唐僧?
认知途径 (1.2 节)	框架网络 ^[20-21]	例句: 王老师安排李四参加明天的学术研讨会。 [Frame1: 安排 [施动者: 王老师][事件: Frame2][计划时间: 明天]] [Frame2: 出席 [施动者: 李四][事件: 学术研讨会]]
	构式语法 ^[22]	例句: The teacher gives students exercises. [SPECIALIZE <i>ditransitive-cx</i> ARG-ST $\langle \text{NP}_x, \text{NP}_y, \text{NP}_z \rangle$ SEM [FRAMES [$\left[\begin{array}{l} \text{transfer-fr} \\ \text{AGENT } x_{[\text{teacher}]} \\ \text{RECIPIENT } y_{[\text{students}]} \\ \text{THEME } z_{(\text{exercise, homework...})} \end{array} \right] \rangle]]$

1.2 认知途径的语义表示

认知途径源于认知语言学的相关探索,代表性工作有框架、构式语法等。框架语义以框架(Frame)为基本单元进行组织。每个框架是一个能被认识和掌握的意义整体,语言的词汇必须基于框架才能被理解。例如表 1 中的“安排”框架表示安排某事件在特定时间、地点发生,并标记了框架的元素。框架主要从实践中获得,也可通过学习、语言应用等方式获得^[23]。

构式语法以构式(Construction)为核心。一般将构式定义为形式-意义对,从而使构式的设计也包含了语义表示。表 1 中构式形似词汇功能语法的 F 结构,属于基于语符的构式语法(Sign-based Construction Grammar, SBCG)。其他构式理论表示略有不同,本文不再赘述。

1.3 分布途径的语义表示

分布途径以学习分布结构为出发点^[24],主张将语义定义为语料中分布的函数(Meaning as a

Function of Distribution),从而语义理解问题就变为从语料学习每个词的分布的问题。而一个最直接的学习词汇分布的方法就是假设相似的语境下有相似的语义,即分布式假说。因此,基于分布途径的语义表示难以表征成人类能阅读的符号系统,一般只用数值向量表达。学习数值向量的方法主要分为基于统计的方法和基于预测的方法^[8,11,25]。基于统计的方法包括信息检索领域的向量空间模型,以及依赖于分解共现矩阵的 LSI、LDA 等经典方法。基于预测的方法一般使用神经网络进行训练,通过给定上下文预测词,包括 Word2Vec、GloVe 等基于固定词向量的方法,以及 ELMo、BERT、GPT 等可变上下文向量方法。

2 理性出发的语义表示

需求出发的语义表示分类有助于划分不同的语义理解任务,但随着相关研究的深入,该分类体系的局限性也愈发凸显。至少有三个问题有待解决:第一,不同语义表示是否多余?由于句子 s 的语义一般不变,服务于 s 的语义表示应当一致,尤其逻辑途径中诸多表示更是在图灵完备的意义上等价。第二,现有工作发现不同语义表示的性能差异很大^[26],基于大模型的 SOTA 技术却也不能弥补^[27],从语义表示的角度是否有改进的空间?第三,以大语言模型为代表的基于分布途径语义表示的系统已经取得了非常显著的成效,是否说明定义符号式语义表示已无必要?这些问题似乎动摇了语义表示的研究基础,本节试图综述理性出发的语义表示以寻求问题的解答。

2.1 世界和语言的关系

陈嘉映在讨论语言任意性时总结称^[7]:“我们不再以语词和现成事物的一一对应为前提,而是把语言看作一个整体系统,与现实在整体上相联系。”可见,自语言转向(Linguistic Turn)以来,“世界和语言”的对应关系取代了自笛卡尔以来的“世界和心灵”的关系成为了相关哲学思想的辩论中心^[3],亦有语言学研究者主张区分词义的“指称能力和推理能力^[28]”。因此,为了完成对语义表示进行分类这一目标,选取世界和语言的关系为主线是可行的。

不同的语义表示可以根据它们对这一关系所持的态度,分为从外延世界定义的语义表示和从语言内部定义的语义表示两类。

2.2 外延世界语义表示

外延世界语义表示(EWS)通过现实世界(即语言之外)定义语义^[3],人工设计了语义符号的系统,典型代表是源于数理逻辑的语义表示,包括指称论、涵义论、描述主义、条件真值语义等^[2]。

指称论特点是将语义定义为事物本身,如“孔子”的语义就是历史上的孔子本人。弗雷格通过把专名语义定义为指称,谓词、连接词、量词的语义定义为一阶和二阶函数,句子语义为真值,从而得到由谓词逻辑表达的语义表示^[2]。涵义论(Sense)将语义定义为“涵义”并由它来确定真值,试图表达指称相同但明显语义不同的词汇,如“启明星”和“长庚星”等。因此其语义表示仍然可以使用谓词逻辑,但其符号代表了涵义。描述主义(Theory of Descriptions,或摹状词理论)则认为专名的语义也是二阶函数,基于奥卡姆剃刀原则删除了涵义概念。条件真值语义使用自然语言本身书写语义表示,用形如“ s 为真当且仅当 s ”的条件真值元语言句子,作为句子 s 的外延刻画从而定义自然语言的语义。作为对比,表 2 给出了这几种语义表示的样例。

表 2 外延世界语义表示样例

自然语言	李四参加了全部会议。
指称论	$\forall x, \text{meeting}(x) \wedge \text{attend}(\text{LiSi}, x)$
涵义论	$\forall x, \text{meeting}(x) \wedge \text{attend01}(\text{LiSi}, x)$
描述主义	$\exists y, \text{is}(y, \text{LiSi}) \wedge (\forall x, \text{meeting}(x) \wedge \text{attend}(y, x)) \wedge (\forall z, \text{is}(z, \text{LiSi}) \rightarrow z = y)$
条件真值语义	“李四参加了全部会议” is true if and only if LiSi has attended every meeting.

其他基于外延的定义方法则有维特根斯坦的使用论(Use)、以布龙菲尔德为代表的行为主义语义论、场景论等。但是这些外延方法规定的是语义获取的机制,不直接对应具体的语义表示。

2.3 语言内部语义表示

语言内部语义表示(LIS)源于共时语言学和结构主义的发展,认为语义源于语言内部的结构。但相关研究主要聚焦于词汇语义。例如,关系语义学^[29]通过词汇之间的形式关系来探究语言“结构”,包括同义词、反义词、上下位词、整体部分关系等,并发展出 WordNet 等重要语义资源。语义场理论^[6]将语言结构看成场,不同的词存在于一个场中获取了各自的语义,并区分了研究聚合关系的词汇场和

组合关系的联想场,并在词汇场中讨论与关系语言学类似的场内关系。类似地,概念角色语义^[5](Conceptual Role Semantics)认为概念间关联赋予了词的角色语义,例如,“加速度”概念虽然没有物质实体,但通过与牛顿定律的表述中其他概念相联系,人脑即可掌握“加速度”一词的语义。人本语义学^[30]主张回归索绪尔的任意性原则和结构主义思想,提出了便于使用、清晰表达等以人为本的语言观,区分了索绪尔意义上的系统性语义和非系统性乃至半系统性语义。

也有语义表示主张定义更基本的元语义^[31],并组合成每个词的语义,例如义素分析、概念依存等方法^[6]。对应语义表示则需定义基本的元语义符号,已有其他词汇语义综述介绍,本文从略。

3 两个方向的对齐

世界和语言的关系这一主线不仅贯穿了语言哲学的大量讨论,在科学研究中亦存在相似的“形式语言能力”和“功能语言能力”二分体系^[32],用于评估语言模型。同时认知科学中的经验研究如对成人大脑的功能研究和失语症患者的行为研究等^[32],也在一定程度上支持了该二元区分。因此,将本文前两节所设计的语义表示在“世界和语言的关系”中找到位置并简单对齐是有可能的。

在此基础上对前述语义表示进行分类的结果如图 1 所示。指称论和条件真值语义均是外延的,谓词逻辑、编程语言也依赖定义好的本体,所以它们更靠近世界一侧。与之相反,关系语义、语义场、概念角色语义、预训练语言模型等使用语料数据学习语言内部模式,并不关心外延,更靠近语言一侧。



图 1 世界-语言关系下语义表示汇总

而位于中间的框架网络、构式语法、人本语义学、涵义论和描述主义等表示方法则同时具有两方

面特点。具体地,如表 1 中框架语义使用的“安排”框架包含施动者、事件、安排时间等框架的核心元素,还包括受益人、频度、方式等非核心元素。这一定义与语义场类似,都引入了一个“框架”或“场”的整体性概念,并通过该概念为场中的词汇或框架元素赋予语义。因此,框架表示具有图 1 左侧表示的一些特点。但是,框架表示从认知研究出发,是人类对真实世界的提炼,因而相比语义场又更强调了外延世界的重要性^[23]。类似地,具体的语义表示可以借鉴不同语义表示理念进行设计。

该分类不精确但简洁地回答前述问题 RQ: 靠近世界一侧的语义表示将语义定义为外部实体,并设计符号表示;靠近语言一侧的语义表示认为语义存在于语言内部的系统中,并不直接可见,故通过研究系统来探讨语义。更进一步,要分析如描述主义等位于中间的语义表示,可以将 RQ 分为关于“语义是什么”的一阶问题和“语义从哪里来”的二阶问题^[33]。这一拆分也有助于分析前文提到的使用论,尽管概念角色语义^[5]和基于语言游戏和博弈论的语义处理^[34]都被认为是基于使用论的,但在二阶问题上,前者主张语义从语言联系中来,后者主张语义源于智能体之间交互,因此有别。对于描述主义,虽然基于真值和外延定义,但它认为语义完整包含在语言之内并由语词表达,而外在论者认为至少有一部分语义是在语言之外的^[33,35]。语义内在、外在之争已超出本文范围,但可以看出本文借用的这一组关系对于深入分析依然是有益的。

同时,对于第 2 节所提出的不同的语义表示取舍问题,从图 1 的分类能看出它们之间存在较大的理念分歧,依照现有实验结论也无法支持语义表示的取舍,但可以考虑语义表示的结合,具体方法将在第 5 节介绍。对于同种类型语义表示之间的性能差异,图 1 给出了一个可能解释即表达风格不同。例如,同样是一阶谓词逻辑,选择不同的风格后可对应表 2 诸多写法,在简洁度和复杂性上有明显区别,本文将在 4.2 节中讨论外延世界语义表示的改进。

4 单种语义表示的应用和挑战

4.1 外延世界语义表示的研究热点

将自然语言转换为外延世界语义表示(EWS)大致上等价于语义解析(Semantic Parsing)这一术语的主要所指任务^[36-37]。从自然语言形式和语义

表示之间的对应关系入手,可以将技术演进分为三个阶段:①建模词项和语义的最小一一对应,如组合范畴语法(CCG)驱动的方法^[38-41];②在词项和语义集合的任意映射上应用统计模型^[15,42-43];③使用神经网络端到端建模样本的完整映射过程,以兼容任意类型和格式的语义表示^[44-46]。最近的研究工作主要围绕着第三类技术展开,包括领域应用、与 LLM 结合、组合泛化、多语言研究、评价指标等热点内容。

4.1.1 领域应用

EWS 由于其符号化特点,能广泛用于各类能提炼出符号的领域。Cunningham 等人^[47]将使用自然语言描述的初等数学证明转换为形式化语言 Coq,从而实现自动进行定理验证。Yin 等人^[48]将建筑领域的本体信息融入语义解析中,实现了自然语言到领域查询语言的转换。Frisoni 等人^[49]则利用了事件抽取和 AMR 表示等语义解析技术增强生物医药论文的生成式摘要性能。AMR 表示还用于风格转换和情感分析之中^[50-51]。此外,也有新提出的不同领域数据集,如基于 SparQL 语言回答学术领域问答的 DBLP-QuAD^[52]、基于 SQL 语言查询电子病历的 EHRSQL^[53]、针对机器人手术设计的框架库 RSPF^[54]、对交通或旅游领域中可移动对象进行查询的 NALMO^[55]等。由于 SQL 的用途和影响力较广,Spider 已经不能满足需要,有研究者提出超大规模 SQL 数据集 BIRD^[56]和 UNITE^[57]。

4.1.2 与 LLM 结合

最近的相关研究趋势之一是换用大语言模型(LLM)来完成符号式语义表示的解析任务,和过去的深度学习方法同样属于外延世界语义表示的技术应用。其中,Ni 等人^[58]利用符号语义表示的可执行特点,在取得 LLM 生成的候选 SQL 语句之后,将实际执行结果通过额外训练模型进行验证,从而输出最终的 SQL 代码,提升了 LLM 在此任务上的精确程度。Wang 等人^[59]则将语义表示符号的句法规则以 BNF 范式的形式加入到 LLM 的提示(Prompt)中,从而提升 LLM 在复杂语义解析上的性能。Liu 等人^[60]则利用 LLM 中思维链(Chain-of-Thought)的特点,设计了对 SQL 语句生成进行分治处理的方法。但是,LLM 目前尚无法很好地生成如领域特定语言(DSL)等语义表示,其成功应用仅限于语料中常见的编程语言^[27],在 LLM 的框架下如何实现任意的 DSL 生成技术还是一个开放问题。

4.1.3 组合泛化

组合性(Compositionality,或称系统性,即 Systematicity)常被视为自然语言的一大特点,也是近年来语义解析的研究热点。过去的辩论中有研究者主张其必须使用基于外延世界的语义表示才能刻画^[2,61]。Yin 等人^[62]提出了一种模型无关的正则化方法,通过约束词在不同上下文的表示一致性和样本整体的表示一致性,显著提升了模型在组合泛化上的性能。Lin 等人^[63]期望将校准误差(ECE)扩展到结构化输出上,提出组合期望校准误差,用于评估 Seq2Seq 模型预测图结构的可靠性。Drozdov 等人^[64]在 LLM 基础上提出使用 least-to-most 的特殊提示词构造方法,成功在 SCAN 和 CFQ 两个数据集上验证了组合泛化能力。Rai 等人^[65]发现通过专门化的 token 预处理、加入块标签等简单的方式,让基于 LLM 的解析器能更好地理解 SQL 等语义表示的分割边界,从而提升其在组合泛化上的性能。Rongali 等人^[66]提出了对领域概念进行预训练的方法,使得基础模型可以在新领域以 zero-shot 或 few-shot 的低资源方式迁移。Crouse 等人^[67]设计了一个自底向上的解码器,通过自底向上的方式生成 S 表达式并惰性地放到候选池中,直到执行特定动作时才用于构造输出。通过配合 T5 和 LSTM 等不同编码器,该方法都有效提升了组合泛化的性能。

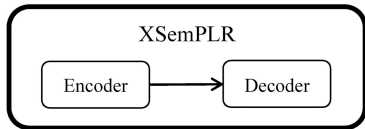
除了方法改进之外,Li 等人^[68]提出了 SLOG 数据集,在 COGS^[69]数据集的基础上细分了 17 种不同的组合泛化形式,便于对组合性进行更细致的研究。Wu 等人^[70]从相反的方向提出了对 COGS 的批评,认为它对组合性评估的设计以及现有模型无法学习组合泛化能力的结论过于草率。他们仅通过语义等价的简单修改就发现基线模型组合泛化性能有较大提升,因此提出了一个修改后的 ReCOGS 数据集,旨在排除次要因素,更好地评估组合性。总的来说,由于语义表示的多样性,且不同技术选型受限于数据集构造、标注量等其他现实约束,这些针对组合泛化的特定方法依旧缺少对不同表示的充分验证,亦难以揭示本质上基于分布语义表示的 Seq2Seq、LLM 等方法如何习得基于外延世界的语义表示的组合性。

4.1.4 多语言研究

多语言问题随着近年来语言模型的技术进步也成了语义表示的热点问题之一。研究者先后提出了多语言的语义解析 MTOP^[71]和对话理解 PRESTO^[72]数据集,覆盖了 6 种语言的语料,并重点考虑了在同

时使用多语种的人群中常见的语码转换现象(Code Switching)。此外,其他研究者也提出了将多语言与其他研究热点相结合的数据,如结合多语言和组合性的 MCWQ^[73]、结合多语言并探索多种外延世界语义表示的 XSemPLR^[74]等(图 2)。对于低资源语言,Ruder 等人^[75]提出了包括语义解析在内的 9 种任务的单一数据集,覆盖了 88 种欠表示的语言。而 Sherborne 等人^[76]提出利用最优传输最小化跨语言概率隐变量的散度,从而减少了对低资源语言的样本量需求,也对缺乏多语平行语料的场景有提升作用。

英 what players made less than three assists over a season
德 Wie viele japanische Restaurants gibt es in Paris
中 有多少摇滚歌手
西 monstrami i ristoranti con piu recensioni
... ..



λ演算 (call SW.listValue (call SW.getProperty ((lambda s (call SW.filter (vars) (call SW.ensureNumericProperty (string num_assists)) (string <)) (call SW.ensureNumericEntity (number 3 assist)))) (call SW.domain (string player)))) (string player)))
FunQL query(area(keyval('name','Paris'), keyval('is_in:country','France')), nwr(keyval('cuisine','japanese'), qtype(count)))
SQL SELECT count(*) FROM singer WHERE genre='Rock'
... ..

图 2 多语言多语义表示(XSemPLR)数据样例

4.1.5 评价指标

对于如何评估模型性能,不同语义表示有所不同。对于编程语言、一阶逻辑等,可使用精确匹配(Exact Match, EM)作为指标,即要求输出序列和期待目标完全一致。

对于有语言运行时或知识库等模型的场景,亦有很多工作选用执行结果(Execution),即编程语言的输出和逻辑表达式的指称(Denotation)作为评估标准。用 EM 评估虽然精确但缺少表示风格上的灵活性,用执行结果评估灵活性好但可能有错误的语义表示碰巧蒙对指称。

而 AMR 等图表示一般使用 SMATCH 指标作为评估,通过对齐图的节点并计算图三元组的匹配程度(如 F_1 值)来估计输出的图结构和目标图结构的差异。然而 Opitz 最近指出^[77],该方法不够精确且在指标计算中引入了图对齐这样的 NP 难问题,因此提出了 SMATCH++ 指标,将原指标拆分为预

处理、对齐、打分等三部分分别计算。类似地,Shou 和 Lin 针对 AMR 中 SMATCH 计算的对齐匹配问题提出了 AMRSim^[78]指标,通过模型学习输出的图结构和目标之间的向量表示,再用向量相似度估计二者的相似度。最后,Groschwitz 等人^[79]认为尽管很多 AMR 解析器已经在 SMATCH 指标上取得了不错的成绩,但实际使用中依然和预期相差较远,因此提出了 9 大类 36 小类的细分标准,以便从实用性、技术性、语言学特性等多方面指导 AMR 解析器的提升。

4.2 外延世界语义表示的改进

改进 EWS 之前,应当明确性能指标,以评估是否有改进空间。Guo 等人首先证实了表示之间存在显著差异^[26]。他们为 GeoQuery 等经典数据集标注了不同语义表示的平行语料,并验证了其他条件不变时选用不同表示的性能差别较大,而使用形式语言的抽象句法树(AST)进行解码时,不同的句法关系也会显著影响性能。这一现象在以 LLM 的 In-Context Learning 为基础时也同样被观察到^[27]。改进语义表示的研究主要包括旧表示的简化、新表示的设计、中间表示的设计、表示之间的量化对比等。

4.2.1 语义表示的简化

基于不同表示的差异,Bogin 等人^[27]主张使用 LLM 较为擅长的 Python、Scala、JavaScript 等编程语言,而避免使用原有的领域特定语言。Meron 针对 SMCaFlow^[80]进行了大幅简化^[81]获得了显著提升。类似地,对 COGS 默认 λ 表达式的简化^[70]、对 SparQL 语言的简化^[82]、对 λ-DCS 的简化^[27]等相关工作均验证了语义表示的简化能明显改善性能。部分简化的表示如图 3 所示。Zhang 等^[83]提出了一种基于 LLM 和最优传输的指标 DisStruct 用于评估表征的不同句法结构影响并进行自动化结构搜索。但受限于对神经网络模型本身的理解,如何简化表示更契合于现有建模技术还是一个开放问题。

4.2.2 设计全新语义表示

除了对现有表示进行简化外,从理论语言学和实际应用需求出发设计全新的语义表示也是一个重要方向。Abend 和 Rappoport^[84]综述了如 AMR、UCCA、UDS 等语义表示的发展特点,如跨语言的统一表达能力、句法语义界面的抽象能力等。最近, Van Gysel 等人^[18]针对 AMR 在跨语言、逻辑推理、跨句子时序和模态等需求上的弱点,提出了一致意

句子	person whose gender is male and whose birthdate is 2004
λ -DCS	(call SW.listValue (call SW.filter (call SW.filter (call SW.getProperty (call SW.singleton en.person) (string !type)) (string gender) (string =) en.gender.male) (string birthdate) (string =) (date 2004 -1 -1))))
简化 λ -DCS	(listValue (filter (filter (getProperty en.person !type) gender = en.gender.male) birthdate = 2004))
句子	Make an appointment in Central Park on Friday.
SMCalFlow	(Yield :output (CreateCommitEventWrapper :event (CreatePreflightEventWrapper :constraint (Constraint[Event] :location (?= # (LocationKeyphrase "Central Park")) :start (Constraint[DateTime] :date (?= (NextDOW :dow # (DayOfWeek "FRIDAY"))))))))
简化 SMCalFlow	CreateEvent(AND(at_location(Central Park) , starts_at(NextDOW(FRIDAY))))

图3 语义表示简化样例

义表示(Uniform Meaning Representation, UMR)。例如,UMR 引入了 5 种体态(Aspectual)标注,即状态(State)、惯常(Habitual)、进行(Activity)、未完成(Endeavor)和完成体(Performance)等,弥补了 AMR 的缺陷。图 4 给出了 UMR 表示的一个样例。

He denied any wrongdoing.
(s3d / deny-01
:Aspect Performance
:ARG0 (s3p / person
:ref 3s)
:ARG1 (s3t / thing
:ARG1-of (s3d2 / do-02
:ARG0 s3p
:ARG1-of (s3w / wrong-02))))
(s3 / sentence
:temporal ((s3d :before DCT))
:modal ((s3d :AFF AUTH)
(s3d2 :NEG (s3h :AFF AUTH)))
:coref ((s3p :same-entity s1p)))

图4 支持体态标注和跨句推理的 UMR 表示样例

Martinez Lorenzo 等人^[85]同样试图对 AMR 进行多语言推广,提出了 BabelNet 意义表示,即 BMR。BMR 以构建语言无关的意义表示为目标,用 BabelNet^[86]和 VerbAtlas^[87]等跨语言词汇语义和语义角色资源,替代了原本 AMR 中刻画谓词论元结构的 PropBank,并提供了 4 种语言的数据和标注工具。Garg 等人则提出了统一语义表示 USR^[88],他们借鉴古梵语中的整体性思想,认为词并非独立而是与句中其他词有语义联系的单元,从而为 USR 设计了词汇-概念层、句法-语义层、篇章层等多层表示结构,且便于表现说话人的主观意向(vivaksha)。

4.2.3 设计中间表示

不难看出,上述立足理论语言学的新语义表示

对既有表示的影响较大,解析器技术层面不能很快迁移适应,因此也有很多新的语义表示以一种“中间表示”的形式出现,作为获得最终语义表示前的中继,或不同语言理解任务的转换格式。典型的代表 SemQL^[89]就是为了抽离 SQL 目标语言的复杂性而设计的一种任务特定的上下文无关语言。最近,Nie 等人^[90]设计了一种标注语言 GraphQ IR,相比 SparQL、Cypher、Lambda-DCS 等经典图数据查询语言,GraphQ IR 更为简洁,且更接近自然语言,便于模型处理。Hsu 等人^[91]提出了一种序列标注形式的框架,能统一处理事件抽取、关系抽取、任务型语义解析等语言理解任务。Li 等人^[92]提出将三元组、语义标注视为一种虚拟二维表格,该表格和数据库表格一起线性化,转换为一种标记语言,实现了多数据来源的少样本文本生成。Cao 等人^[93]基于表格问答任务的特点,提出使用 Python 生态中 Pandas 库的数据框(Data Frame)作为中间语义表示,使用 Python 作为查询语言并借助当前 LLM 对于 Python 代码的优秀性能,从而将表格问答转换为基于 Python API 的代码生成任务。

4.2.4 语义表示的量化对比

语义表示如何设计在学术共同体内尚未形成一致的意见。Li 等人^[94]设计了 ISS 指标,并证实它与模型性能有显著相关性。ISS 通过抽样数据集中差别为 1 个条件的两个同表示样本,再比较它们经过预处理后的 AST 之间的树编辑距离(Tree Edit Distance),从而提前预估一种语义表示的性能。实务之中可能很难观察到某种句法或字面统计量可以影响表示的性能,而模型缺少对 DSL 中个别操作符语义的理解可能是关键点所在^[27]。

Pavlova 等人^[95]沿着这个方面在前语义问题、语义特征、语义应用等三个维度上分别总结了一系列对比项(如语义特征包括谓词论元结构、时序、体态模态、空间信息、作用域等 11 种语义学话题)。这些项目进一步用于对包括 AMR、UMR 在内的 8 种语义表示进行横向对比。作为最新的相关工作之一,Pavlova 等人并不认为这是对语义表示的完备的考察,仅希望所选取的特征集对未来工作在选择语义表示时有所帮助。因此,目前的证据并不支持某一种语义表示统一相关领域。而根据本文第 3 节的分析,很可能对语义表示的评估和改进已超出语义表示的范围,而依赖于将来对于语言本质的探索。

4.3 语言内部语义表示的挑战

在基于语言内部的语义表示(LIS)中,大语言

模型(LLM)改变了语言处理诸多任务的旧有范式^[96]。由于LLM技术发展迅速,本文不再赘述其改进,仅关注语义表示的视角。如何表征语义无疑是语义表示的根本问题,围绕这一问题LLM引发了两种观点的激烈辩论^[97]: 正方认为LLM已经走上了正确的道路上,它计算的向量表征和生成的句子已经掌握了语义,剩下的无非是技术细节比如规模、计算效率等问题;反方则认为现有技术没有解决语言理解问题,无论如何增加模型大小都不可能有望。

4.3.1 对 LIS 的批评

首先分析反方,包括三种代表性的观点。第一种反方观点是LLM没有掌握与人类相当的语言能力^[98]。Dentella等人^[99]设计了包括共指、比较、形容词、副词顺序等8种语言现象的样本,并发现LLM在不符合语法规则的样本上表现较差。Sinha等人^[100]则发现在蕴含任务的样本中加入特定噪声会影响人类的判断,但却并不影响LLM的判断。虽然这些经验证据表明LLM和人类的语言能力确有差异,既难以为LLM提供后续改进思路,但也难以揭示LLM与人类语言能力的区别究竟是什么。

第二种反方观点源自指称论,认为LLM训练过程因缺少真实世界的指称而无法理解语言。如前文所述,语义的一种主要传统是将其定义为世界中的实体,并用数理逻辑的成果进行系统化。用Bender和Koller^[101]的诙谐例子来说: 两座小岛之间的通信被海底章鱼截获,但由于不知道词语究竟指代陆地上的什么实体,它将永远无法理解所截获消息的语义。但是,目前看来指称论并不能揭示语义的全部侧面: 认知科学中提出的表示多元化^[102-103]、语言具身假说^[104]等观点认为概念并不由感知体验所独立地表征,也依赖于语言的表征或语言经验的模拟;涵义论(Sense)等经典语义理论也表明^[2],只通过指称定义的语义无法解释如神话人物、“启明星就是长庚星”“PyTorch长期支持版已被废弃”等词汇和句子的语义;由于真实世界的多样性,即使是“椅子”这样的常用词汇,我们也很难写出其精确的指称或词典释义^[105]。因此,强求LLM应有具身体验并断言其没有意义理解能力可能是过分的苛责^[106]。在工程上,如DALLE-2^[107]等多模态大模型尝试把文本和图像结合,似乎就简单地补上了文本落实(Grounding)缺失的一环,显得指称论反驳并不严重。

第三种反方观点围绕组合性展开。戴维森纲领

明确提出了自然语言的语义理论应有组合性^[2],符号式AI的部分支持者认为神经网络等连接主义方法只能实现类比泛化,无法实现组合性的泛化^[61,108]。然而,这种观点将组合性当作自然语言的金科玉律,同样有待商榷: 构式语法研究中发现有的构式具有部分能产性(Partial Productivity)^[109],组合性规则有可能生成不被接受的句子;Rabovsky和McClelland提出语义具有伪组合性^[110],即同时存在常规的(Regular)和特异的(Idiosyncratic)用法;最后,组合性虽然呼应了“有限的手段产生无限的句子”这一观察,但难以应对词汇可能存在的无限多的涵义(Sense)^[111]。因此,“组合性不存在于神经网络中”这一论点在前件预设上有瑕。

总的说来,上述三种批评指出了LLM尚存在的具体问题,可以作为短期技术的优化方向,但难以构成对“LLM确实刻画了语义”的有效反驳。

4.3.2 为 LIS 的辩护

下面分析正方。正方的首要辩护无疑是大量的经验证据。以LLM为核心的新范式在自然语言处理中发展迅速^[96],比过去的诸多技术都更好地“理解了语义”。

正方第二种辩护集中于语言学知识。LLM能较好地掌握由规则覆盖的形式语言能力^[32],也学到了一定的句法关系^[112],且与脑电实验测量到的大脑皮层计算句法关系有相似趋势^[113]。

正方第三种辩护源自第二节中所述概念角色语义(CRS)。Piantadosi和Hill^[114]将CRS作为LLM的解释,并据此主张LLM可以学到一定的语义侧面。支持该辩护有可能成立的理论基础包括: 复杂科学中的嵌入定理^[115-116]允许从系统的低维投影中重建一定的高维状态;借助于邱奇编码(Church Encoding)可以让神经网络通过参数训练编码目标领域的概念角色^[117]。对于该辩护的实证层面,即LLM产生的向量表示是否的确一定程度上包含了CRS所预期的概念间关联,也存在一定的佐证^[114]: LLM可以完成一些需要概念关联才能完成的任务,如事实问答^[118-119]、Winograd^[120]、定量推理^[121-122]等;LLM学习的数值向量能还原出一定类比关系^[123]和词汇在一定特征上的渐变关系^[124],如动物的体型大小、危险性等;LLM可以学到一定的语言之外世界的概念间结构^[125],即使其并没有用实际观测进行训练。但是,尽管Word2Vec开始的LIS就通过形如“king-man+woman=queen”的类比关系展示出对概念的掌握能力,但这只适用于较少

的情形^[126]。

总的说来,虽然 LLM 的理论远落后于实践,但通过联系语言内部语义表示(LIS)的概念角色语义(CRS),可以较为确信地主张 LLM 不是“老练的鹦鹉”,而确实可以掌握一定的语义能力。虽然 LLM 尚不能很好地组织语义空间^[97],但这更多地是技术所限,而非方针错误。

5 不同语义表示的结合

5.1 结合的动机

前文区分了两种语义观,即外延世界语义(EWS)和语言内部语义(LIS)。但两类表示均被发现存在不同缺陷,例如,指称论难以解释非物质实体的概念、语义场理论难以解释缺失部分专名而语义保持等。因此这一组概念区分只能是暂时的,更多的语义表示散落在两类观点之间,语义表示的研究者为了克服相应缺点而不断进行二者结合的尝试。

以词汇关系语义的发展^[29]为例,Lyons 等^[127]学者首先主张研究语言结构内部的系统意义,避开百科知识或指称层面的东西,以处理语义学中讨论同义反义关系的困难。但在实际语境中,这样的区分很难具体地执行。Cruse 等^[128]针对系统意义定义了同义、反义、上下位、整体和部分等四类关系。但实际语料中,反义关系常由语境或感知所确立,例如,art、civilization、artifact 等词均可能是 nature 的反义词、“听”可能是“看”的反义词;上下位关系则往往被科学分类、俗称等真实世界用法所影响;整体部分关系也可能是实体的组构,并非是词语之间的纯粹形式关系。也就是说,纯粹的形式定义不好操作且难以与真实世界解耦,而纯粹基于真实世界的定义又无法解释语料中的所有词汇关系现象。这引发了两种语义观念的结合。

5.2 语义结合方法

常见的结合方法是在技术实现层面,选取基于一种表示的算法,再将另一种表示的信息加入算法流程。例如,在基于神经网络的语义解析器中加入层次句法标签^[129],以及在知识库问答中通过 embedding 向量放松对于候选答案的邻近节点的约束^[130]等。但这类结合方法是一般性的机器学习方法,被加入的新语义表示与其他新特征、新资源、外

部先验知识等并无二致。

第二种结合方法则以评价为目的,明确地考虑了语义表示的差异性,通过一种语义表示的视角来分析基于另一种语义表示的技术。例如,从构式语法的观点来构建语言模型的评价基准^[131-132],目前的发现证实了语言模型能够掌握构式的句法,但是否掌握了构式的语义尤其是非组合语义还缺少证据^[133-134]。

第三种结合方法有别于第一种,针对不同语义表示的特点进行设计。例如,对构式语法和语言模型的结合,CxLM 等工作^[135]提出使用构式资源来精调语言模型,Dunn 等人把词向量等 LIS 技术作为构式挖掘的特征^[136-137]。但两种表示分属于理论语言学和自然语言处理各自的前沿,其结合作还还处于起步阶段。例如,Rambelli 等人^[22]设计了一种构式的表示,在语义部分融合了 FrameNet 和分布式语义,但该分布式语义或基于 Word2Vec 的词向量表征,或基于构式的一系列高频动词的向量来表征,并未用到基于预训练语言模型的新技术。

除构式语法之外的其他语义表示也可与 LLM 结合。由于 LLM 完全基于分布式假说训练,在逻辑否定等情况下易生成不一致文本。研究者尝试从概念角色语义(CRS)的角度来理解 LLM,运用词汇释义资源(如 Wikitionary 等人工词典)精调 LLM,令其判断一条释义是否能刻画一个词^[138-139]。词汇释义是人工为真实世界的实体撰写的表征,有助于增强 LLM 的词汇关联能力,符合 CRS 的预期。与之类似,Bogin 等人^[27]为 LLM 加入了编程语言书写的领域描述(如领域中 Class 的 Python 定义等),显著提升了 LLM 转换自然语言为领域特定语言的性能。有的研究则考虑结合不同来源但同类的语义表示,如专注事件本体对齐的 EventOA^[140]结合了 FrameNet 和 WikiData。而 EDeR^[141]则考虑了事件之间除常见的时序、因果之外的条件关系和论元关系,扩充了对事件对象关联的研究。

可见,不同语义表示结合,尤其是分属本文所述 EWS 和 LIS 的两类语义表示之间的结合,有望给现有以 LLM 为主体的研究带来新的提升,但还需要更多的研究跟进。例如,根据 LIS 模型输出的向量对于不同维度投影^[124],可以对模型进行系统性的白盒检验,若发现模型在区分系统性语义上做出了错误预测,则可以有针对性地补充 EWS 表示数据进行精调。

6 总结和展望

本文以语义表示为核心讨论语言处理系统,并沿着从需求和理性出发的两个方向进行综述。随着相关研究深入,语义表示的重要性逐渐凸显,而旧的分类方法也逐渐引发了如语义表示如何对比和选择等新问题。本文工作主要有三方面:首先,借鉴理论语言学中关于世界和语言这一组对比概念,将不同语义表示按照其在二者之间的位置进行重新分类,分别为“外延世界语义表示(EWS)”和“语言内部语义表示(LIS)”。其次,借助这一分类,综述了相关研究的进展,包括EWS的应用研究热点、EWS语义表示的改进,以及围绕以LLM为代表的LIS相关辩论。最后,本文综述了结合上述两类语义表示的相关研究,尽管这些工作尚处于起步阶段,但已展现出蓬勃发展的潜力。从这三方面工作中,本文得出的主要结论有:①LLM很有可能学到了一定语义侧面,而不是机械地复述语料;②两种分类的语义表示之间存在显著的观念分歧,在表征句子意义时并不等价,不能互相代替;③针对使用一种语义表示的系统,与另一种差异较大的语义表示结合,要比同种类的语义表示更为有效。

参考文献

- [1] VAN HARMELEN F, LIFSCHITZ V, PORTER B. Handbook of knowledge representation: Vol. 1[M]. Amsterdam, Oxford: Elsevier, 2008: 1-2.
- [2] MILLER A. Philosophy of language[M]. London: Routledge, 2018: 8-378.
- [3] 王寅. 语言哲学研究[M]. 北京: 北京大学出版社, 2014.
- [4] 邱雪玫, 李葆嘉. 语义解析方法的形成过程及其学术背景——揭开“结构主义语义学”的第二个谜[J]. 江海学刊, 2016(3): 65-73.
- [5] BLOCK N. Conceptual role semantics[M]. CRAIG E. Routledge Encyclopedia of Philosophy. London: Routledge, 1998: 242-256.
- [6] 冯志伟. 自然语言处理的形式模型[M]. 合肥: 中国科学技术大学出版社, 2010.
- [7] 陈嘉映. 简明语言哲学[M]. 北京: 中国人民大学出版社, 2013: 49-51.
- [8] 袁书寒, 向阳. 词汇语义表示研究综述[J]. 中文信息学报, 2016, 30(5): 1-8, 27.
- [9] 仇韞琦, 王元卓, 白龙, 等. 面向知识库问答的问句语义解析研究综述[J]. 电子学报, 2022, 50(9): 2242-2264.
- [10] KAMATH A, DAS R. A survey on semantic parsing [C]//Proceedings of Automated Knowledge Base Construction. Amherst, MA, 2019.
- [11] PATIL R, BOIT S, GUDIVADA V, et al. A survey of text representation and embedding techniques in NLP [J]. IEEE Access, 2023, 11: 36120-36146.
- [12] APIDIANAKI M. From word types to tokens and back: a survey of approaches to word meaning representation and interpretation[J]. Computational Linguistics, 2023, 49(2): 465-523.
- [13] HERSHCOVICH D, DONATELLI L. It's the meaning that counts: The state of the art in NLP and semantics[J]. KI - Künstliche Intelligenz, 2021, 35(3): 255-270.
- [14] MAO R, HE K, ZHANG X, et al. A survey on semantic processing techniques[J]. Information Fusion, 2024, 101: 101988.
- [15] LIANG P. Learning executable semantic parsers for natural language understanding[J]. Communications of the ACM, 2016, 59(9): 68-76.
- [16] COPESTAKE A, FLICKINGER D, POLLARD C, et al. Minimal recursion semantics: An introduction[J]. Research on Language and Computation, 2005, 3(2-3): 281-332.
- [17] BANARESCU L, BONIAL C, CAI S, et al. Abstract meaning representation 1.0 specification [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Seattle: ACL, 2012: 1533-1544.
- [18] VAN GYSEL J E L, VIGUS M, CHUN J, et al. Designing a uniform meaning representation for natural language processing [J]. KI-Künstliche Intelligenz, 2021, 35(3): 343-360.
- [19] YIH W, CHANG M W, HE X, et al. Semantic parsing via staged query graph generation: Question answering with knowledge base [C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, 2015: 1321-1331.
- [20] 刘开瑛. 汉语框架语义网构建及其应用技术研究[J]. 中文信息学报, 2011, 25(6): 46-52, 62.
- [21] LI J, YAN Z, SU X, 等. CCL23-Eval 任务 3 总结报告: 汉语框架语义解析评测 [C]//Proceedings of the 22nd Chinese National Conference on Computational Linguistics. Harbin, China: Chinese Information Processing Society of China, 2023: 113-123.
- [22] RAMBELL G, CHERSONI E, BLACHE P, et al. Distributional semantics meets construction grammar.

- towards a unified usage-based model of grammar and meaning [C]//Proceedings of the 1st International Workshop on Designing Meaning Representations. Florence, Italy: Association for Computational Linguistics, 2019: 110-120.
- [23] FILLMORE C J. Frames and the semantics of understanding[J]. *Quaderni di Semantica*, 1985, 6(2): 222-254.
- [24] HARRIS Z S. Distributional structure [J]. *Word*, 1954, 10(2-3): 146-162.
- [25] PAVLICK E. Semantic structure in deep learning[J]. *Annual Review of Linguistics*, 2022, 8(1): 447-471.
- [26] GUO J, LIU Q, LOU J G, et al. Benchmarking meaning representations in neural semantic parsing[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Online: Association for Computational Linguistics, 2020: 1520-1540.
- [27] BOGIN B, GUPTA S, CLARK P, et al. Leveraging code to improve in-context learning for semantic parsing[C]//Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, 2024: 4971-5012.
- [28] MARCONI D. Lexical competence[M]. Cambridge, Massachusetts: MIT Press, 1997: 1-7.
- [29] 刘慧, 李葆嘉. 关系语义: 基于指称意义的特定关系联想——揭开“结构主义语义学”的第三个谜[J]. *南京师范大学文学院学报*, 2014(2): 142-153.
- [30] 程雨民. “人本语义学”十论[M]. 上海: 上海外语教育出版社, 2010.
- [31] SAEED J I. Semantics[M]. New York: Wiley, 2011: 5-9.
- [32] MAHOWALD K, IVANOVA A A, BLANK I A, et al. Dissociating language and thought in large language models[J/OL]. *Trends in Cognitive Sciences*, 2024. [https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613\(24\)00027-5](https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(24)00027-5) [2024-01-10].
- [33] KALLESTRUP J. Semantic externalism[M]. London: Routledge, 2012.
- [34] TRIPODI R, PELILLO M. A game-theoretic approach to word sense disambiguation[J]. *Computational Linguistics*, 2017, 43(1): 31-70.
- [35] GASPARRI L, MARCONI D. Word meaning [G]. ZALTA E N. *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, 2021.
- [36] ZELLE J M, MOONEY R J. Learning semantic grammars with constructive inductive logic programming[C]//Proceedings of the 11th National Conference on Artificial Intelligence, 1993: 817-822.
- [37] ZELLE J M, MOONEY R J. Learning to parse database queries using inductive logic programming[C]//Proceedings of the National Conference on Artificial Intelligence, 1996: 1050-1055.
- [38] ZETTLEMOYER L S, COLLINS M. Learning to map sentences to logical form: Structured classification with probabilistic categorial grammars[C]//Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence, Arlington, Virginia, United States: AUAI Press, 2005: 658-666.
- [39] ZETTLEMOYER L S, COLLINS M. Online learning of relaxed CCG grammars for parsing to logical form [C]//Proceedings of the EMNLP-CoNLL, 2007: 678-687.
- [40] KWIATKOWSKI T, ZETTLEMOYER L, GOLDWATER S, et al. Inducing probabilistic CCG grammars from logical form with higher-order unification [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2010: 1223-1233.
- [41] KWIATKOWSKI T, CHOI E, ARTZI Y, et al. Scaling semantic parsers with on-the-fly ontology matching[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Seattle, Washington, USA: Association for Computational Linguistics, 2013: 1545-1556.
- [42] LIANG P, JORDAN M I, KLEIN D. Learning dependency-based compositional semantics[J]. *Computational Linguistics*, 2013, 39(2): 389-446.
- [43] PASUPAT P, LIANG P. Compositional semantic parsing on semi-structured tables[C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Beijing, China, 2015: 1470-1480.
- [44] VINYALS O, KAISER L, KOO T, et al. Grammar as a foreign language[C]//Proceedings of the 28th International Conference Neural Information Processing Systems, 2015: 2773-2781.
- [45] XIAO C, DYMETMAN M, GARDENT C. Sequence-based structured prediction for semantic parsing [C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, 2016: 1341-1350.
- [46] DONG L, LAPATA M. Language to logical form with neural attention [C]//Proceedings of the 54th

- Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 2016: 33-43.
- [47] CUNNINGHAM G, BUNESCU R C, JUEDES D. Towards autoformalization of mathematics and code correctness: Experiments with elementary proofs[C]// Proceedings of the 1st Workshop on Mathematical Natural Language Processing, Abu Dhabi, United Arab Emirates (Hybrid): Association for Computational Linguistics, 2023: 25-32.
- [48] YIN M, TANG L, WEBSTER C, et al. An ontology-aided, natural language-based approach for multi-constraint BIM model querying[J]. *Journal of Building Engineering*, 2023, 76: 107066.
- [49] FRISONI G, ITALIANI P, SALVATORI S, et al. Cogito ergo summ: Abstractive summarization of biomedical papers via semantic parsing graphs and consistency rewards[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023, 37 (11): 12781-12789.
- [50] MA F, HU X, LIU A, et al. AMR-based network for aspect-based sentiment analysis[C]// Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, Stroudsburg, PA, USA, 2023: 323-337.
- [51] SHI K, SUN X, HE L, et al. AMR-TST: Abstract meaning representation-based text style transfer[C]// Proceedings of the Association for Computational Linguistics, Stroudsburg, PA, USA, 2023: 4231-4243.
- [52] BANERJEE D, AWALE S, USBECK R, et al. DBLP-QuAD: A question answering dataset over the DBLP scholarly knowledge graph[C]// Proceedings of the 13th International Workshop on Bibliometric-enhanced Information Retrieval, Dublin, Ireland, 2023.
- [53] LEE G, HWANG H, BAE S, et al. EHRSQL: A practical text-to-SQL benchmark for electronic health records[C]// Proceedings of the 36th Conference on Neural Information Processing Systems Tracks on Datasets and Benchmarks, 2022: 15589-15601.
- [54] BOMBIERI M, ROSPOCHER M, PONZETTO S P, et al. The robotic-surgery propositional bank[J]. *Language Resources and Evaluation*, 2024, 58(3): 1043-1071.
- [55] WANG X, LIU M, XU J, et al. NALMO: Transforming queries in natural language for moving objects databases[J]. *GeoInformatica*, 2023, 27(3): 427-460.
- [56] LI J, HUI B, QU G, et al. Can LLM already serve as a database interface? A big bench for large-scale database grounded text-to-SQLs[C]// Proceedings of the 37th International Conference on Neural Information Processing Systems, 2023: 42330-42357.
- [57] LAN W, WANG Z, CHAUHAN A, et al. UNITE: A unified benchmark for text-to-SQL evaluation[J/OL]. arXiv preprint arXiv: 2305.16265, 2023.
- [58] NI A, IYER S, RADEV D, et al. LEVER: Learning to verify language-to-code generation with execution[C]// Proceedings of the 40th International Conference on Machine Learning, PMLR, 2023: 26106-26128.
- [59] WANG B, WANG Z, WANG X, et al. Grammar prompting for domain-specific language generation with large language models[C]// Proceedings of the 37th Conference on Neural Information Processing Systems, 2023: 65030-65055.
- [60] LIU X, TAN Z. Divide and prompt: Chain of thought prompting for text-to-SQL[J/OL]. arXiv preprint arXiv: 2304.11556, 2023.
- [61] FODOR J A, PYLYSHYN Z W. Connectionism and cognitive architecture: A critical analysis[J]. *Cognition*, 1988, 28(1): 3-71.
- [62] YIN Y, ZENG J, LI Y, et al. Consistency regularization training for compositional generalization[C]// Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, Stroudsburg, PA, USA, 2023: 1294-1308.
- [63] LIN Z, PHAN D, PASUPAT P, et al. On compositional uncertainty quantification for seq2seq graph parsing[C]// Proceedings of the 11th International Conference on Learning Representations, 2023.
- [64] DROZDOV A, SCHÄRLI N, AKYÜREK E, et al. Compositional semantic parsing with large language models[C]// Proceedings of the 11th International Conference on Learning Representations, 2023.
- [65] RAI D, WANG B, ZHOU Y, et al. Improving generalization in language model-based text-to-SQL semantic parsing: Two simple semantic boundary-based techniques[C]// Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, Toronto, Canada, 2023: 150-160.
- [66] RONGALI S, SRIDHAR M, KHAN H, et al. Low-resource compositional semantic parsing with concept pretraining[C]// Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics, Dubrovnik, Croatia, 2023: 1410-1419.
- [67] CROUSE M, KAPANIPATHI P, CHAUDHURY S,

- et al. Laziness is a virtue when it comes to compositionality in neural semantic parsing[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics. Toronto, Canada, 2023: 8434-8448.
- [68] LI B, DONATELLI L, KOLLER A, et al. SLOG: A structural generalization benchmark for semantic parsing[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Singapore, 2023: 3213-3232.
- [69] KIM N, LINZEN T. COGS: A compositional generalization challenge based on semantic interpretation[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing, Online: Association for Computational Linguistics, 2020: 9087-9105.
- [70] WU Z, MANNING C D, POTTS C. ReCOGS: How incidental details of a logical form overshadow an evaluation of semantic interpretation[J]. Transactions of the Association for Computational Linguistics, 2023, 11: 1719-1733.
- [71] LI H, ARORA A, CHEN S, et al. MTOP: A comprehensive multilingual task-oriented semantic parsing benchmark[C]//Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics, Online: Association for Computational Linguistics, 2021: 2950-2962.
- [72] GOEL R, AMMAR W, GUPTA A, et al. PRESTO: A multilingual dataset for parsing realistic task-oriented dialogs [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Singapore: Association for Computational Linguistics, 2023: 10820-10833.
- [73] CUI R, ARALIKATTE R, LENT H, et al. Compositional generalization in multilingual semantic parsing over wikidata[J]. Transactions of the Association for Computational Linguistics, 2022, 10: 937-955.
- [74] ZHANG Y, WANG J, WANG Z, et al. XSemPLR: Cross-lingual semantic parsing in multiple natural languages and meaning representations[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics. Toronto, Canada, 2023: 15918-15947.
- [75] RUDER S, CLARK J, GUTKIN A, et al. XTREME-UP: A user-centric scarce-data benchmark for under-represented languages[C]//Proceedings of the Association for Computational Linguistics, Singapore, 2023: 1856-1884.
- [76] SHERBORNE T, HOSKING T, LAPATA M. Optimal transport posterior alignment for cross-lingual semantic parsing[J]. Transactions of the Association for Computational Linguistics, 2023, 11: 1432-1450.
- [77] OPITZ J. SMATCH ++ : Standardized and extended evaluation of semantic graphs[C]//Proceedings of the Association for Computational Linguistics: EACL 2023. Dubrovnik, Croatia, 2023: 1595-1607.
- [78] SHOU Z, LIN F. Evaluate AMR graph similarity via self-supervised learning[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics. Toronto, Canada, 2023: 16112-16123.
- [79] GROSCWITZ J, COHEN S, DONATELLI L, et al. AMR parsing is far from solved: GrAPES, the granular AMR parsing evaluation suite[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Singapore, 2023: 10728-10752.
- [80] ANDREAS J, BUFE J, BURKETT D, et al. Task-oriented dialogue as dataflow synthesis[J]. Transactions of the Association for Computational Linguistics, 2020, 8: 556-571.
- [81] MERON J. Simplifying semantic annotations of SMCaFlow[C]//Proceedings of the 18th Joint ACL - ISO Workshop on Interoperable Semantic Annotation within LREC. Marseille, France, 2022: 81-85.
- [82] FURRER D, VAN ZEE M, SCALES N, et al. Compositional generalization in semantic parsing: pre-training vs. specialized architectures [J/OL]. arXiv preprint arXiv: 2007.08970, 2020.
- [83] ZHANG X, HE S, LIU K, et al. On the effects of structural modeling for neural semantic parsing[C]//Proceedings of the 27th Conference on Computational Natural Language Learning. Singapore, 2023: 38-57.
- [84] ABEND O, RAPPOPORT A. The state of the art in semantic representation[C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017, 1: 77-89.
- [85] MARTÍNEZ LORENZO A C, MARU M, NAVIGLI R. Fully-semantic parsing and generation: The BabelNet meaning representation[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. Dublin, Ireland, 2022: 1727-1741.
- [86] NAVIGLI R, BEVILACQUA M, CONIA S, et al. Ten years of BabelNet: A survey[C]//Proceedings of the 29th International Joint Conference on Artificial Intelligence. 2021: 4559-4567.
- [87] DI FABIO A, CONIA S, NAVIGLI R. VerbAtlas: A novel large-scale verbal semantic resource and its application to semantic role labeling[C]//Proceedings

- of the Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Hong Kong, China, 2019: 627-637.
- [88] GARG K, PAUL S, SUKHADA S, et al. Evaluation of universal semantic representation (USR) [C]// Proceedings of the 4th International Workshop on Designing Meaning Representations. Nancy, France, 2023: 13-22.
- [89] GUO J, ZHAN Z, GAO Y, et al. Towards complex text-to-SQL in cross-domain database with intermediate representation [C]// Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy, 2019: 4524-4535.
- [90] NIE L, CAO S, SHI J, et al. GraphQ IR: Unifying the semantic parsing of graph query languages with one intermediate representation [C]// Proceedings of the Conference on Empirical Methods in Natural Language Processing. Abu Dhabi, United Arab Emirates, 2022: 5848-5865.
- [91] HSU I H, HUANG K H, ZHANG S, et al. TAGPRIME: A unified framework for relational structure extraction [C]// Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics. Toronto, Canada, 2023: 12917-12932.
- [92] LI A H, SHANG M, SPILIOPOULOU E, et al. Few-shot data-to-Text generation via unified representation and multi-source learning [C]// Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics. Toronto, Canada, 2023: 16171-16189.
- [93] CAO Y, CHEN S, LIU R, et al. API-assisted code generation for question answering on varied table structures [C]// Proceedings of the Conference on Empirical Methods in Natural Language Processing. Singapore, 2023: 14536-14548.
- [94] LI Z, GUO J, LIU Q, et al. Exploring the secrets behind the learning difficulty of meaning representations for semantic parsing [C]// Proceedings of the Conference on Empirical Methods in Natural Language Processing. Abu Dhabi, United Arab Emirates, 2022: 3616-3625.
- [95] PAVLOVA S, AMBLARD M, GUILLAUME B. Structural and global features for comparing semantic representation formalisms [C]// Proceedings of the 4th International Workshop on Designing Meaning Representations. Nancy, France, 2023: 1-12.
- [96] SUN T X, LIU X Y, QIU X P, et al. Paradigm shift in natural language processing [J]. Machine Intelligence Research, 2022, 19(3): 169-183.
- [97] LENCI A. Understanding natural language understanding systems [J]. Sistemi Intelligenti, 2023, 35(2): 277-302.
- [98] DÜRR O, SEGESSENMAN J, STEINMANN J J. Meaning, form and the limits of natural language processing [J]. Philosophy, Theology and the Sciences, 2023, 10(1): 42-72.
- [99] DENTELLA V, GÜNTHER F, LEIVADA E. Systematic testing of three language models reveals low language accuracy, absence of response stability, and a yes-response bias [C]// Proceedings of the National Academy of Sciences, 2023, 120(51): e2309583120.
- [100] SINHA K, PARTHASARATHI P, PINEAU J, et al. Unnatural language inference [C]// Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. Online, 2021: 7329-7346.
- [101] BENDER E M, KOLLER A. Climbing towards NLU: on meaning, form, and understanding in the age of data [C]// Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Online, 2020: 5185-5198.
- [102] DOVE G. Beyond perceptual symbols: A call for representational pluralism [J]. Cognition, 2009, 110(3): 412-431.
- [103] BINDER J R. In defense of abstract conceptual representations [J]. Psychonomic Bulletin & Review, 2016, 23(4): 1096-1108.
- [104] DOVE G O. Rethinking the role of language in embodied cognition [J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 2022, 378(1870): 20210375.
- [105] ELBOURNE P. Meaning: A slim guide to semantics [M]. Oxford, New York: Oxford University Press, 2011: 1-30.
- [106] 袁毓林. 如何测试 ChatGPT 的语义理解与常识推理水平? ——兼谈大语言模型时代语言学的挑战与机会 [J]. 语言战略研究, 2024, 9(1): 49-63.
- [107] RAMESH A, DHARIWAL P, NICHOL A, et al. Hierarchical text-conditional image generation with CLIP latents [J/OL]. arXiv preprint arXiv: 2204.06125, 2022.
- [108] BERENT I, MARCUS G. No integration without structured representations: Response to pater [J]. Language, 2019, 95(1): e75-e86.

- [109] GOLDBERG A E. Explain me this: Creativity, competition, and the partial productivity of constructions [M]. Princeton: Princeton University Press, 2019: 57-70.
- [110] RABOVSKY M, MCCLELLAND J L. Quasi-compositional mapping from form to meaning: A neural network-based approach to capturing neural responses during human language comprehension[J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 2019, 375(1791): 20190313.
- [111] PUSTEJOVSKY J. The generative lexicon [J]. Computational Linguistics, 1991, 17(4): 409-441.
- [112] CHI E A, HEWITT J, MANNING C D. Finding universal grammatical relations in multilingual BERT [C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Online, 2020: 5564-5577.
- [113] KUMAR S, SUMERS T R, YAMAKOSHI T, et al. Shared functional specialization in transformer-based language models and the human brain [J]. Nature Communications, 2024, 15(1): 5523.
- [114] PIANTADOSI S, HILL F. Meaning without reference in large language models [C]//Proceedings of the NeurIPS Workshop on Neuro Causal and Symbolic AI, 2022.
- [115] PACKARD N H, CRUTCHFIELD J P, FARMER J D, et al. Geometry from a time series [J]. Physical Review Letters, 1980, 45(9): 712-716.
- [116] TAKENS F. Detecting strange attractors in turbulence [G]//Dynamical Systems and Turbulence, Warwick 1980. Berlin, Heidelberg: Springer, 1981: 366-381.
- [117] PIANTADOSI S T. The computational origin of representation [J]. Minds and Machines, 2021, 31(1): 1-58.
- [118] MIN B, ROSS H, SULEM E, et al. Recent advances in natural language processing via large pre-trained language models: A survey [J]. ACM Computing Surveys, 2023, 56(2): 30: 1-30: 40.
- [119] JIANG Z, ARAKI J, DING H, et al. How can we know when language models know? On the calibration of language models for question answering [J]. Transactions of the Association for Computational Linguistics, 2021, 9: 962-977.
- [120] KOCIJAN V, LUKASIEWICZ T, DAVIS E, et al. A review of Winograd schema challenge datasets and approaches [J/OL]. arXiv preprint arXiv: 2004.13831, 2020.
- [121] SAWADA T, PALEKA D, HAVRILLA A, et al. ARB: Advanced reasoning benchmark for large language models [C]//Proceedings of the 3rd Workshop on Mathematical Reasoning and AI at NeurIPS, 2023.
- [122] LEWKOWYCZ A, ANDREASSEN A, DOHAN D, et al. Solving quantitative reasoning problems with language models [C]//Proceedings of the 36th International Conference on Neural Information Processing Systems 35, 2022: 3843-3857.
- [123] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality [C]//Proceedings of the 26th International Conference on Neural Information Processing Systems, 2013: 3111-3119.
- [124] GRAND G, BLANK I A, PEREIRA F, et al. Semantic projection recovers rich human knowledge of multiple object features from word embeddings [J]. Nature Human Behaviour, 2022, 6(7): 975-987.
- [125] PATEL R, PAVLICK E. Mapping language models to grounded conceptual spaces [C]//Proceedings of the International Conference on Learning Representations, 2022.
- [126] LENCI A, SAHLGREN M, JEUNIAUX P, et al. A comparative evaluation and analysis of three generations of distributional semantic models [J]. Language Resources and Evaluation, 2022, 56(4): 1269-1313.
- [127] LYONS J. Linguistic semantics: An introduction [M]. Cambridge: Cambridge University Press, 1995: 3-124.
- [128] CRUSE D A. Lexical semantics [M]. Cambridge: Cambridge University Press, 1986.
- [129] DO D T, NGUYEN M P, NGUYEN L M. GRAM: Grammar-based refined-label representing mechanism in the hierarchical semantic parsing task [G]//Natural Language Processing and Information Systems. Cham: Springer Nature Switzerland, 2023: 339-351.
- [130] SAXENA A, TRIPATHI A, TALUKDAR P. Improving multi-hop question answering over knowledge graphs using knowledge base embeddings [C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Online, 2020: 4498-4507.
- [131] PANNITTO L, HERBELOT A. CALaMo: A constructionist assessment of language models [C]//Proceedings of the 1st International Workshop on Construction Grammars and NLP. Washington, D.

- C.: Association for Computational Linguistics, 2023: 21-30.
- [132] VEENBOER T, BLOEM J. Using collocation analysis to evaluate BERT's representation of linguistic constructions[C]//Proceedings of the Association for Computational Linguistics, Toronto, Canada, 2023: 12937-12951.
- [133] WEISSWEILER L, HOFMANN V, KÖKSAL A, et al. The better your syntax, the better your semantics? Probing pretrained language models for the english comparative correlative[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Abu Dhabi, United Arab Emirates, 2022: 10859-10882.
- [134] WEISSWEILER L, HE T, OTANI N, et al. Construction grammar provides unique insight into neural language models[C]//Proceedings of the 1st International Workshop on Construction Grammars and NLP. Washington, D.C., 2023: 85-95.
- [135] TSENG Y H, SHIH C F, CHEN P E, et al. CxLM: A construction and context-aware language model [C]//Proceedings of the 13th Language Resources and Evaluation Conference, Marseille, France; European Language Resources Association, 2022: 6361-6369.
- [136] DUNN J. Learnability and falsifiability of construction grammars[J]. Proceedings of the Linguistic Society of America, 2017, 2(1): 1-15.
- [137] DUNN J. Exploring the construction: Linguistic analysis of a computational CxG[C]//Proceedings of the 1st International Workshop on Construction Grammars and NLP. Washington, D. C., 2023: 1-11.
- [138] JANG M, MTUMBUKA F, LUKASIEWICZ T. Beyond distributional hypothesis: Let language models learn meaning-text correspondence[C]//Proceedings of the Association for Computational Linguistics: Seattle, United States, 2022: 2030-2042.
- [139] JANG M, LUKASIEWICZ T. Improving language models' meaning understanding and consistency by learning conceptual roles from dictionary[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Singapore, 2023: 8496-8510.
- [140] GUO S, WANG C, CHEN Y, et al. EventOA: An event ontology alignment benchmark based on FrameNet and Wikidata[C]//Proceedings of the Association for Computational Linguistics. Toronto, Canada; Association for Computational Linguistics, 2023: 10038-10052.
- [141] LI R, HASLUM P, CUI L. EDeR: Towards understanding dependency relations between events[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Singapore, 2023: 14969-14983.



张翔(1989—), 博士研究生, 主要研究领域为语义解析、知识图谱。

E-mail: xiang.zhang@nlpr.ia.ac.cn



张元哲(1986—), 博士, 副研究员, 主要研究领域为自然语言处理、知识工程、可解释性研究。

E-mail: yuanzhe.zhang@ia.ac.cn



何世柱(1987—), 通信作者, 博士, 副研究员, 主要研究领域为问答系统、对话系统、知识推理、知识图谱、自然语言处理。

E-mail: shizhu.he@nlpr.ia.ac.cn