

ChatGPT 语境下语言学的挑战和出路*

澳门大学 袁毓林

提要: ChatGPT 在语言理解和生成方面的成功,对语言学有关传统观念提出了挑战。本文据此重点讨论了人工智能时代语言学家不可回避的三个问题:1) 怎样看待 ChatGPT 等语言大模型生成的语言? 2) 怎样为语言模型等人工智能系统提供语义学支持? 3) 怎样跟语言大模型等人工智能系统形成人机协同? 对此,我们尝试性的回答是:1) “机造语言”是一种“中介语”; 2) 应该把指称论语义学和分布式语义学结合起来; 3) 人机协同有嵌入、副驾驶和智能体三种模式。

关键词: ChatGPT; 语言大模型; 人工智能; 语义学; 人机协同

[中图分类号] H0-05

[文献标识码] A

[文章编号] 1003-6105(2024)04-0572-06

DOI:10.20071/j.cnki.xdwy.20240523.009

Challenges and Prospects for Linguistics in the Context of ChatGPT

YUAN Yulin University of Macau

Abstract: The success of ChatGPT in language comprehension and generation challenges some traditional concepts in linguistics. This article addresses three unavoidable issues for linguists in the era of artificial intelligence: (1) how to treat the language generated by large language models such as ChatGPT? (2) how to provide semantic support for artificial intelligence systems such as language models? (3) how to form human-machine collaboration with artificial intelligence systems such as large language models? This article concludes with three tentative propositions: (1) “machine language” is a kind of “interlanguage”; (2) denotational semantics and distributional semantics should be combined; and (3) there are three modes of human-machine collaboration, namely, embedding, copilot and agent.

Keywords: ChatGPT; large language models; artificial intelligence; semantics; human-machine collaboration

1. ChatGPT 等大模型对语言学的挑战

长久以来,从事语言学(包括语言教学与研究)的人有一种笃定的信心:只要人类不灭,语言就不会消亡;只要语言不亡,语言学职业就铁定存在。更何况,

* 本文得到国家社科基金专项“新时代中国特色语言学基本理论问题研究”(19VXK06)以及澳门大学讲座教授研究与发展基金“词典资源中句法语义知识的可视化及其运用研究”(CPG2024-00005-FAH)和科研启动基金“面向可解释 AI 的语义推理模式与机制研究”(SRG2022-00011-FAH)的支持。

几千年来,语言学从业者对于语言的结构、功能及其历史演变,对于语言及外语教学的策略与方法,已经积累了丰富的经验,建构了系统的知识,并且形成了专门化的学科建制 (disciplinary institution)和学科规训 (disciplinarity)。虽然语言学内部多少还有点卷,在语言教学与研究的理念与方法上争议不断、流派纷呈,竞争的压力也不容小觑,但从外观上看还算是岁月静好,总体上说景气指数也应该不低。

谁能想到,2022 年底人工智能聊天机器人 ChatGPT 横空出世,以其超强的、接近人类水平的语言理解与生成能力,外加知识学习与常识推理能力,一下子火爆全网;同时也无意中把语言学推到了历史的十字路口,挑战了语言学既有的一系列传统观念。比如:

- 1) 语言是人类独有的现象,只有人类大脑才能产出自然语言;
- 2) 人类语言具有递归性(recursion),所以无法用基于统计的概率模型来刻画;

还向语言学人提出了一系列棘手的问题,比如:

- 1) ChatGPT 等大模型生成的语言是不是人类自然语言? 为什么?
- 2) 语言符号具有任意性 (arbitrary, 即形式和意义没有必然的关系),那么语言模型是怎样跨越这种语义鸿沟的?
- 3) 目前的语言学理论在多大程度上揭示了人脑中实际的语言机制?
- 4) 现代大型语言模型的成功有没有影响到语言学作为一门学科的存在? 语言大模型在自动翻译方面的成功会不会影响到外语专业的存在?

显然,在 ChatGPT 等大模型的背景下思考这些问题,可以拓宽语言学的学术视野。本文集中讨论三个问题:1) 语言学家应该怎样看待 ChatGPT 等大模型生成的语言? 2) 语言学家怎样为语言模型等人工智能系统提供语义学支持? 3) 语言学家怎样跟语言大模型等人工智能系统相处 (互相协作)?

2. 怎样看待 ChatGPT 等大模型生成的语言?

下面,我们先来讨论第一个问题:ChatGPT 等大模型生成的语言是不是人类自然语言? 这个问题的回答,取决于怎样来定义“人类自然语言”。如果规定只有人类(大脑)产出的语言才能叫“人类自然语言”,那么 ChatGPT 等大模型生成的语言就不能算作人类自然语言。但是,这可能会有问题,因为语言大模型通过海量文本的训练,已经基本学会了人类自然语言的句法、语义和语用模式,在人类用户合适的提示 (prompt, 包括问题或命令) 的指引下,能够生成类似于人类自然语言的文本。比如,下面是笔者和 ChatGPT 的问答:

笔者:请向语言学家通俗地介绍 ChatGPT 的工作原理是怎么样的?

ChatGPT:当然,我很乐意向你解释一下。ChatGPT 的工作原理基于人工智能技术中的一种称为“深度学习”的方法。具体来说,它是基于一种称为“神经网络”的模型。(以下删去一小节)

在训练过程中,ChatGPT通过大量的文本数据进行学习,这些文本数据包含了各种各样的语言信息。通过观察输入文本的上下文以及与之相关的输出(回答、建议等),模型试图学习语言的模式、结构和语义。在这个过程中,ChatGPT逐渐建立了对语言的理解能力,并学会了生成自然流畅的文本。(以下删去一小节)

总的来说,ChatGPT的工作原理就是通过大规模的数据训练深度神经网络,使其能够理解和生成自然语言文本。这种模型可以用于各种任务,包括对话、文本生成、问题回答等。

机器的这个回答不仅语句通顺,结构完整,而且内容正确,逻辑自洽,比我心目中的答案还要全面一点。不足之处是篇章布局上略嫌松散,可能是受提示中“通俗地介绍”的影响。如果把这种机器从人类文本中学来的、并且人们可能也会这么说/写的文本,排斥在“人类自然语言”之外,可能是一种短视和狭隘的做法。这有点儿像不承认医院的辅助生殖科在实验室里培育出来的试管婴儿是“人类婴儿”一样,不近情理。其实,务实的人工智能社区已经把机器生成的语言当作“人类自然语言”,来快速有效地扩大训练语料的规模,并且有效地规避使用真正的“人类自然语言”文本可能带来的版权风险。

不过,话要说回来。如果承认机器生成的语言(简称“机造语言”)是“人类自然语言”,那么语言学家在收集语料、提取语法规则时,能不能使用这种“机造语言”来归纳“人类自然语言”的语法规律?更何况,这种“机造语言”可能良莠不齐,有时可能存在语法错误。比如,上例中介词“基于”前面缺核心动词“是”、“建立……能力”这种动宾搭配不同寻常、介词“通过”后面缺核心动词“采用”,等等。就像二语习得者产出的“中介语”(interlanguage)是语言学的一个重要的研究对象一样,“机造语言”(machine language)可能会成为未来语言学研究 and 工程应用的一个重要的研究领域。

3. 语言学家怎样为人工智能提供语义学支持?

接着,我们来讨论第二个问题:ChatGPT等大模型是怎样跨越语言符号的音义任意性这种语义鸿沟的?大家知道,人类自然语言是一种象征性的符号系统,语言符号音义之间的联系是由社会成员在历史的长河中约定俗成的。对于语言学习者来说,这种没有逻辑必然性的音义关系,是他们在一定的文化和语言环境下,通过具身认知,把某种语言符号跟有关的外部世界和内心体验结合起来,日积月累而逐步习得的。对于机器而言,上述这种指称论语义学(Denotational Semantics)不好使了,因为机器没法理解词典释义式的离散性语义表达。于是,怎样对语言符号的意义进行可计算的连续性的数值表达,就成为自然语言处理(natural language processing, NLP)的难题。好在人工智能专家在“意义即用法”、“意义相近的词语有着相近的分布”等语言学思想的指导下,根据词语出现的各种上下文,把词语的意义表示成高维向量(vector)。粗略地说,这是用某个词的

各种上下文来表示这个词的意义和用法。现代大型语言模型的词向量动辄成千上万个维度,相当于在用每一个词的成千上万种语境信息来表示这个词的意义和用法。因为意义相近的词语往往出现在相似的上下文中,所以它们的向量值也相近;或者说,它们在高维的语义空间中挨得很近,并且相关词语的向量之间有一种平行的推导关系。一个著名的例子是“国王”的向量减去“男人”的向量(约等于“royal,王室”的向量)再加上“妇女”的向量约等于“女王”的向量。正是这种基于分布式语义学(distributional semantics)的词语意义的向量化嵌入表达(embeddings),为大模型注入了语义;并且,在大规模训练数据、转换器(transformer)的注意力机制、把在少量样本上学习到的知识迁移到其他领域的的能力、在上下文中学习等各种预训练技术的加持下,使得语言模型能够跨越语言符号的语义鸿沟。

分布式语义学虽然在人工智能语言模型上取得了巨大成功,但是语言学界对于及其前身潜在语义学(Latent Semantics)的关注是严重不够的。我们期望有志于面向人工智能语义学研究的年轻学者,能够从语言学角度去关心和研究分布式语义学。事实上,对于人工智能应用来说,以词向量为核心的分布语义学只是在数字世界(digital world)中解决了意义的表达和计算问题;但是,语言符号没有跟物理世界建立联系,即跟外部环境是脱节的。而语言是交际工具,语言符号最终要在人类生活的物理世界(physical world)中跟有关事物发生指涉和绑定关系。这就是所谓的“符号接地问题”(the symbol grounding problem),即让语言符号的语义解释从人脑或机器中投射到物理世界中,跟语境中有关的事物、事件或状态建立指称关系。比如,一台加载了语言大模型的家用机器人,虽然能够听懂主人的指令“把桌子上的大苹果装进礼品袋里!”的意义,但是实际操作起来必须把“桌子”、“大苹果”和“礼品袋”等符号跟当下环境中的有关物体建立一一对应关系。这里面,涉及到机器人对物品的识别和对语言符号的指称歧义的消解等,是一系列复杂的多模态感知和语义理解问题。这种符号接地问题不解决,人工智能就不能实现“具身智能”(embodied AI),最终无法进入人们的日常生活。而恰恰是在指称消歧和语义理解这一方面,语言学家的语义学知识和智慧是不可缺少的。可见,在人工智能时代,怎样把指称论语义学和分布式语义学有机地结合起来,解决机器人的语言符号的意义和指涉问题,是语言学家和人工智能专家共同的课题。

4. 语言学家跟 ChatGPT 等大模型协作的三种模式

2024年初,人工智能芯片公司英伟达的总裁黄仁勋说:“这是一个令人惊叹的时代,因为我们正处于一场新的工业革命的开始,过去蒸汽机、电力、PC和互联网带来了信息革命,现在是人工智能。”¹ OpenAI 首席执行官 Sam Altman 称人工

¹ 详见 <https://mp.weixin.qq.com/s/QK1CiRvJcT3s2dZ0RaY1kQ>。

智能将“像手机一样”改变世界,彻底颠覆教育、医疗和科学研究等领域;将成为人类建设未来的强大工具,赋予我们表达创造性想象力和意志的能力。他鼓励年轻一代积极利用人工智能工具,实现个人的愿景,创造前所未有的价值。²是的,本轮人工智能也开启了一个学术创新、研究范式转变的伟大时代。

站在这个新时代的门口,我们语言学人也应该积极拥抱人工智能。一方面要思考怎样更深刻地揭示人类自然语言的结构和功能,为人工智能的自然语言处理提供知识资源。另一方面要考虑怎样更好地利用 ChatGPT 等现代大型语言模型,来辅助我们的语言教学和语言学研究。在这里,我们首先要了解和考虑人机协同的三种可能的模式:³

1) 嵌入(embedding)模式,即语言学家通过跟 ChatGPT 等大模型交谈,使用提示词语来设定具体目标,让人工智能系统协助我们完成研究或教学工作中的某一项或几项任务;然后,语言学家自主地结束工作。比如,分别提示 ChatGPT:

找出现代汉语中像“获得、落得”那样以“得”结尾的二字词,越多越好;

造出 10 个像“难道要我下跪不成? → 难不成要我下跪?”一样的句子配对;

仿照“Colorless green ideas sleep furiously.”,造 10 个语义违反的英语句子,并译成汉语;

把下面的段落缩写成 30 字以内的一句话,给下面的文章摘要进行润色和修辞加工;

结果返回以后,语言学家进行检查,看是否符合要求;也可以指出不足,提出更加明确的提示,要求机器重写。在这种模式下,ChatGPT 等大模型充当执行指令和完成特定任务的工具,语言学家则充当任务设计者和指挥者的角色。

2) 副驾驶(copilot)模式,即语言学家跟 ChatGPT 等大模型形成伙伴关系,共同参与到研究工作的流程中,从提出研究计划、收集相关例句、总结有关规律,到编撰论文提纲、写作论文的有关段落,都是人机协同,并且让人机各自发挥自己的作用。比如,分阶段和分批次地提示 ChatGPT:

什么是否定极项(negative polarity items)? 举例说明英语中有哪些典型的否定极项词语? 它们有什么样的句法语义特点? 这方面有哪些参考文献和研究结论?

举例说明否定极项这个概念对于描写和解释语法现象有什么作用?

请举例说明现代汉语中有没有否定极项? 如果有,那么以“现代汉语否定极项的句法语义研究”为主题,进行专题研究和论文写作合适吗?

² 详见 https://mp.weixin.qq.com/s/_DrOy1eOkauTUzIWevh_aA。

³ 关于人机协同的三种模式,详见 https://mp.weixin.qq.com/s/ws3U1NBouPivqatUb_ldQA。

这个主题有哪些重要的参考文献和研究结论? 还可以进行哪些方面的拓展和挖掘?

我从哪几个方面着手现代汉语否定极项的语料调查和例证分析比较合适?

同时,语言学家自己检索真实文本语料库,一方面自己分析语料,总结规律;另一方面把语料给机器,让它也分析语料,总结规律;再对它的分析进行检查,看看是不是能够补充或者修正自己的结论,或者从中获得新的观察角度;再把有关发现告诉机器,咨询它怎样列出论文提纲,乃至请它帮助写作有关部分,再自己亲手修改,等等。在这种模式下,ChatGPT等大模型和语言学家优势互补,相互配合,并且语言学家是主导者,机器是合作者。

3) 智能体 (agent) 模式,即语言学家提供目标和必要的资源(比如,特定的文献和语料、比较明确的观点、结论或立场评价上倾向性),让 ChatGPT 等大模型独立完成某项任务(比如,写作某种科普性的通俗文章),由语言学家监督进程和评估最终结果。比如,根据文章性质或者现实需要,一次性或者分阶段和分批次地提示 ChatGPT:

请以语言学教师的身份,写一篇面向大众的、关于汉语字母词(如“X光片、CT扫描”,等等)对于语言交际的重要性(不能废除、无可取代)的文章,限800字以内;

请以专业语言学家的身份,从下面三对例句(前面打星号“*”表示该句子不合语法)上,概括出一条语法规律,并作出一定的说明、解释和论证,写一篇3000字以内的文章:

- 1) 买这种汽车的人不太多 ~ 这种汽车买的人不太多
- 2) 看见那个学生的人来了 ~ *那个学生看见的人来了
- 3) 我见过买这种汽车的人 ~ *这种汽车我见过买的人

像后面这个高度专业性的论题,可能需要多个回合的人机问答互动,并且需要语言学家给出更多的正面(支持自己的论点)的例句和反面(反驳对方的论点)的例句,最后,语言学家要细致核查和严格把关。在这种模式下,ChatGPT等大模型充分体现出智能体的互动性、自主性和适应性的特点,接近于独立的行动者和代理人;而语言学家则充当执行任务的监督者和验收结果的评估者。也许,目前 ChatGPT 等大模型还没有达到这种自主的水平。

收稿日期: 2023-12-09; 作者修改稿, 2024-03-23; 本刊修订, 2024-04-14

通讯作者: 袁毓林 <yuanyl@pku.edu.cn>

519000 澳门大学人文学院中国语言文学系

Corresponding author: Yuan Yulin, Department of Chinese Language and Literature, Faculty of Arts and Humanities, University of Macau, Macau 519000, P. R. China