

从三种复杂句看 ChatGPT 是不是随机鹦鹉?*

——语言大模型能不能理解语言意义的测试与讨论

袁毓林

提要 本文通过考察 ChatGPT 在指称歧义句、花园幽径句和递归嵌套句等三种复杂句的语义理解上的表现,来讨论下列问题:ChatGPT 等现代大型语言模型到底能不能理解语言的意义?它们真的只是一种不顾意义的随机鹦鹉吗?文章发现,ChatGPT 在指称歧义句和一般递归结构的理解方面表现出色,在花园幽径句的理解方面表现平凡;在中心内嵌的递归结构的理解方面,跟人类一样表现不佳。由此,文章指出,ChatGPT 等语言大模型能够理解语句的语言性意义,但是未必能够理解语句背后隐微的交际意义和具身意义。最后,文章展望,语言学可为建设关于多种意义的理解的新型智能科学做出独特的贡献。

关键词 指称歧义句;花园幽径句;递归嵌套句;ChatGPT;语言大模型;语义理解;语言性意义;交际意义;具身意义

一 引言:ChatGPT 等语言大模型是不是随机鹦鹉?

大家知道,ChatGPT 等现代大型语言模型能够围绕着某个主题,生成流畅、连贯和切题的文本;并且,在 SuperGLUE 等自然语言理解大基准(benchmark)的测试上表现卓越。不过,它们是通过不断地预测下一个单词来生成和理解语言的,迥异于人类的语言生成(从经验、概念、语义到词句、段落)和语言理解(从词句、段落到语义、概念、甚至具身体验)方式。^①因此,尽管有不少人坚信语言大模型能够理解意义和概念,并且能够进行常识推理;但是,也有不少人认为:语言大模型只学会了语言表面上的序列形式,却没有世界模型(world model),即对于世界的经验或心智模型(mental model),所以并没有学会语义,并不具备真正的语言理解能力。

根据 Michael et al. (2022)的一项调查,在自然语言处理社区内受访的 480 人中,约有半数(51%)的学者相信:如果提供足够大的数据和算力,那么仅基于文本训练的语言模型能够真正理解自然语言;但是,另外一半(49%)的学者则不同意这种观点。有人强调,任何认为语言

作者简介:袁毓林,男,江苏昆山人,博士,曾任北京大学中文系教授,现为澳门大学人文学院中国语言文学系讲座教授,主要研究理论语言学 and 汉语语言学,特别是句法学、语义学、语用学、计算语言学和中文信息处理,电邮:yuanyl@pku.edu.cn。

* 本文得到澳门大学讲座教授研究与发展基金(CPG2024-00005-FAH)和启动研究基金(SRG2022-00011-FAH)支持,承蒙周韧、卢达威、孙晓雪与《语言教学与研究》编辑部和匿名审稿专家的指正,谨此谢忱。

①关于语言生成与理解的心理过程,详见 Levelt (1989)和 Gazzaniga et al. (2009/2011:362-363)。

大模型具备理解力或意识的人,其实都陷入了伊莱莎效应(Eliza effect):在看到机器展现出微弱的类人语言或行为的迹象时,就会倾向于认为它们具有理解力或主体性。这是一种人类的天然倾向,得名于韦岑鲍姆(Joseph Weizenbaum)在1960年代创建的聊天机器人“伊莱莎”,其角色被设定为一个心理医师。尽管它的程序极为简单,却仍然能够诱使人们相信,它能理解他们并且富有同情心。(详见 Weizenbaum 1976;袁毓林 2021;Mitchell & Krakauer 2023) Bender & Koller (2020)指出,从语言学理论上讲,语言是一个由形式和意义配对的符号构成的系统;但是,语言模型所用的训练数据只是形式,没有通达意义。我们认为,这种观点并不全面;因为,语言模型所用的词向量表示,已经比较充分地反映了词语在高维度的语义空间中跟有关词语之间的语义距离;而且,转换器构架中的多头注意力机制,又能够抓住有关词语之间的语义关系的紧密程度。(详见袁毓林 2024b)因此,现代大型语言模型在一定程度上是能够通达意义的。

Bender et al. (2021)的观点比较犀利:看上去语言模型具有连贯的表达能力,实际上这只是一种“情人眼里出西施”(coherence/beauty in fact is in the eye of the beholder)。^② 因为,人类对于连贯的语言表达能力的理解,来自我们对于特定语境中说话人的信念、意图等的认识能力。人类的语言使用发生在共享一定的背景知识的个体之间,双方往往共处于某种说或写的交际情景中;他们有用语言来进行交流的意愿,并且各自揣摩对方的心智状态。人际沟通是一种合作建构的活动,依赖于对个体之间所传递的隐性意义的解释。但是,我们听读者跟通过机器来生成语言的主体(说话人)并不共处于现场(co-present),而是存在着一定的时间和空间距离。当我们遇到这种通过机器生成的语言并且要解释它们时,我们并不知道发出这种语言的人是谁。我们只能建立一种说话人是谁、他们跟我们共享什么知识的不完全的模型,并且以此来解释他们的话语。显然,由语言模型生成的语言并不植根于交际意图、任何的世界模型或读者的心智状态模型。因为,训练数据中并不包含跟听读者共享的思想;并且,机器也没有这种能力。我们只是借助于自己的语言能力和天然倾向,把交际行为解释为传递连贯的意义和意图,而不管它们本来是否如此。这样,问题就来了:既然交际的一方(语言模型)并没有意义,那么,另一方(人类)对隐性意义的理解就是一种错觉(illusion)。显然,这种错觉源于我们人类单方面的语言理解,跟语言模型没有什么关系。事实上,语言模型只是一种草率地把语言形式序列缝合在一起的系统;当然,这是基于对庞大的训练数据的观察,根据词语怎样组合起来的概率信息;但是,它并不参考任何意义。就此而言,语言模型是一种随机鹦鹉(stochastic parrot)。^③

那么,ChatGPT 等现代大型语言模型到底能不能理解语言的意义呢?它们真的只是一种不顾意义的随机鹦鹉吗?显然,答案不会是简单的“是”或“否”;而是取决于我们对“什么是语言的意义?”“什么是对语言的理解?”等问题的回答。为了帮助回答这些问题,下面专门考察 ChatGPT 对于指称歧义句、花园幽径句和递归嵌套句等三种复杂语句的语义理解问题。

二 ChatGPT 在代词的歧义消解方面的出色表现

因为代词回指涉及句法、语义、语用等语法知识和常识、信念、推理等认知因素,所以从 ChatGPT 等现代大型语言模型在代词的歧义消解方面的表现,就可以在一定程度上看出它们

^②这实际上就是上文“伊莱莎效应”的文学性表达。

^③根据 Bender et al. (2021) § 6.1 编译,还加入了笔者自己的理解。如要引用,务请核对原文。

是不是真正理解人类自然语言,还是只是随机的鹦鹉学舌。例如:

- (1) The research and development of the homemade shield tunneling machines is a showcase for what can be achieved through China's unwavering commitment to innovation. It has provided solid support for rapid development of urban transportation, reducing travel times and improving people's quality of life.

上例第二句中的代词“It”到底指“盾构机”还是指“盾构机的研发”,曾经在“中国语言学教师微信群”(2024-05-23)中引起讨论。为了考察 ChatGPT 对这个问题的理解程度,我们向它布置了下面三个任务:

- 1.把上面这段英语文章翻译成汉语;2.用 30 个字以内的一句话概括这段文章的中心思想;3.第二句话开头的代词“It”指什么?为什么?

结果,ChatGPT 给出的答案是:

- 1.这些自制盾构机的研发是中国对创新坚定不移承诺的一个典范。它为城市交通的快速发展提供了坚实支持,缩短了出行时间,提高了人们的生活质量。
2.中国自制盾构机展示了中国创新的决心,为城市交通快速发展提供了坚实支持。
3.第二句话开头的代词“It”指代的是前文提到的“homemade shield tunneling machines”(自制盾构机)。这个代词引导了一个对前文内容的进一步说明,指的是自制盾构机所提供的支持。

我们不满意 ChatGPT 对于 3 的回答,于是跟它进行启发式的对话:

笔者:不对吧,“it”是单数形式的代词,而“homemade shield tunneling machines”(自制盾构机)是复数形式?朋友,你再想一下!

ChatGPT:你说得对,我错了。让我重新思考一下。在这种情况下,“It”指代的应该是整个前文提到的概念,即“the research and development of the homemade shield tunneling machines”(自制盾构机的研究与开发)。因此,正确的理解是:第二句话开头的代词“It”指代的是“the research and development of the homemade shield tunneling machines”(自制盾构机的研究与开发)。

可见,ChatGPT 可以在用户的引导下,逐步找到有关问题的正确的答案。从这一点来看,它应该是理解笔者的提示的意义的;并且,它还应该具备笔者提示中没有明确说出的“代词必须跟其先行语在数上保持一致”之类的英语语法知识。否则,它怎么能够修正自己先前的答案,给出正确的回答呢?

2024 年 5 月,我在网上读到关于托尔金(Christopher Tolkien)的下面这段文字:

- (2)身为最知名的畅销小说《魔戒》的作者、现代奇幻文学的奠基人,托尔金生前一直反对为自己作传,他认为真正值得记录的只有他的作品,自己的生活与作品并无直接关系,也不重要。(《牧冬读〈托尔金传〉|走近奇幻作家的真实人生》,上海书评,2024-05-20, <https://mp.weixin.qq.com/s/DuNYMOWExaYY59uKcTWYHQ>)

我认为其中的“托尔金生前一直反对为自己作传”,至少有三重歧义;这跟“为自己作传”中的零形式(记作 $\emptyset$)的无声主语(silent subject)及其跟“托尔金”与“自己”的同指(co-reference)或异指(dis-reference)关系有关。试分析如下:

- (2-1)托尔金_i生前一直反对 $\emptyset_i$ 为 自己_i 作传

(=托尔金不赞成[别人劝他]自己为他自己作传,但是不反对别人为他自己作传)

- (2-2)托尔金_i生前一直反对 $\emptyset_j$ 为 自己_i 作传

(=托尔金不赞成[别人]为他自己作传,但是不反对自己亲自为他自己作传)

(2-3)托尔金_i生前一直反对为_{o_i}自己_i作传

(=托尔金不赞成[别人]自己为自己作传,但是不反对别人由他人为他自己作传)

于是,我们就这个问题考问 ChatGPT:

笔者:“托尔金生前一直反对为自己作传”这句话的意思是什么,这里的“自己”指代谁?

ChatGPT:这句话的意思是托尔金在世的时候反对有人为他写传记。这里的“自己”指代托尔金本人。

笔者:在“托尔金生前一直反对为自己作传”这句话中,“反对为自己作传”的人是谁?是托尔金还是其他人?这里的“自己”指代谁?是托尔金还是其他人?

ChatGPT:这句话中的“反对为自己作传”的人是托尔金,反对为托尔金写传记。这里的“自己”指代托尔金本人。

笔者:“托尔金生前一直反对为自己作传”这句话有几种可能的意思?为什么?

ChatGPT:这句话基本上有两种可能的理解:

1. 托尔金反对别人为他写传记:托尔金本人不希望在他生前有人为他写传记。
2. 托尔金反对自己写自传:托尔金本人反对给自己写传记(自传)。

在通常情况下,这句话更常见的理解是托尔金反对别人为他写传记,因为“作传”通常指别人为某人写传记,而自传通常明确为“写自传”。但严格说来,第二种理解也是语法上可能的解释。

可见,ChatGPT 没有考虑到上面(2-3)这种虽然并不常见、当时依然有可能的语义解读。但是,它对于其他两种意义的理解基本正确;而且,还能够理解这个歧义句的优先理解是什么。而这除了需要反身代词回指先行语的有关语法知识之外,还需要理解“作传”“传记”“传主”“传记的作者”等概念及其关系、以及跟这些概念相关的词汇语义;并且,还得有“作传”既可以是别人写传记,也可以是自己写传记(自传,作者就是传主)等默会的(tacit)知识。

三 ChatGPT 在理解花园幽径句方面的平凡表现

台湾歌手罗大佑于 1982 年发表风靡一时的流行歌曲《穿过你的黑发的我的手》,其中排比使用了三组通过关系化造成的复杂名词短语,让许多人觉得不容易理解:

- (3-1-1)穿过你的黑发的我的手(←我的手 穿过你的黑发)
- (3-1-2)穿过你的心情的我的眼(←我的眼 穿过你的心情)
- (3-2-1)牵着我无助的双手的你的手(←你的手 牵着我无助的双手)
- (3-2-2)照亮我灰暗的双眼的你的眼(←你的眼 照亮我灰暗的双眼)
- (3-3-1)留不住你的身影的我的手(←我的手 留不住你的身影)
- (3-3-2)留不住你的背影的我的眼(←我的眼 留不住你的背影)

从语法分析的角度看,上例这种名词化是主谓结构的关系化,即原来基础结构中的领属性定中式主语作中心语,原来基础结构中的谓语加“的”而名词化为同位性定语。结果,原来基础结构中的领属性定语后的中心语前出现两个组合式定语;但是,它们跟中心语的潜在语法关系和语义关系(论元角色关系)不同。并且,初听乍看之下,很容易走上另外的理解轨道(比如,“穿过你+的+黑发……”),最终发现走进了错误解读的死胡同(因为其底层形式“黑发+穿过你”是没有意义的)。也就是说,这种句子属于语法理论上所谓的“花园幽径句”(garden path sentence):一开始提供了多种结构解析与词语组合的可能性,但是多数都是死胡同,往往只有

一种解析是最终可以得到正确的语义解释的。许多人往往一开始选择了错误的结构解析与词语组合方式,最终得不到合理的语义解释;于是,只得折回来重新分析和理解。这种语义关系相对复杂的歌词,当时让许多人觉得既新鲜有趣,又朦胧费解。

我们为这三组排比句设置了不尽相同的问题,来跟 ChatGPT 进行问答对话:

笔者:“穿过你的黑发的我的手”说的是:什么穿过了什么?为什么?

ChatGPT:我的手穿过了你的黑发。描述的是一种亲密的动作,用手穿过对方的黑发,表示爱意或关心。

笔者:“穿过你的心情的我的眼”说的是:什么穿过了什么?为什么?

ChatGPT:我的眼穿过了你的心情。用眼睛注视或关注对方的情感和内心状态,表示深刻的理解和共鸣。

笔者:“牵着我无助的双手的你的手”说的是:谁牵着谁的手?是用一只手去牵还是用两只手去牵?被牵着的是一只手还是两只手?为什么?

ChatGPT:你的手牵着我无助的双手。用的是你的手(可以是一只手或两只手)去牵我的双手。因为歌词中特别提到了“无助的双手”,所以被牵的是两只手。

笔者:“留不住你的背影的我的眼”说的是:什么留不住什么?背影的数量是几个?眼的数量是几个?为什么?

ChatGPT:我的眼留不住你的背影。背影是一个,而眼是两个。因为提到的是“背影”(单数)和“眼”(一般理解为双眼)。

可见,ChatGPT 对付这种语句没有什么障碍,并且还对常识有足够的关注。比如,知道“背影是一个,而眼是两个”,“你的手(可以是一只手或两只手)去牵我的双手”。但是,对于隐藏在歌词背后的失落、无奈等主观情绪,则未必能够理解和体会到。

林斤澜“瘾”系列小说中有《电话》一文,文前有近于绕口令的两句“梦话”:

(4)骂你的嘴里有我的嘴,打你的手里有我的手。

这个对偶句当时引起读者和批评家的关注和争议,有人认为不好懂,是“尚奇求异”。为什么?从语法理论的角度看,这两个句子类似于多重“花园幽径句”。试分析如下:

一上来,读者容易把“骂你”识解为述宾结构;但是,“你”后面出现助词“的”,迫使他打碎原来的述宾结构“骂你”,让“你”跟“的”组成“的”字结构“你的”;“你的”可以跟后面的名词“嘴”组成同位式偏正结构“你的嘴”;接着,黏着性方位词“里”有向前附着的要求,所以“你的嘴”跟“里”组合后形成处所短语“你的嘴里”。这样一来,问题马上来了:不仅接下来的“你的嘴里有我的嘴”在语义上是无解的,而且还把句子开头的动词“骂”晾在一旁没有着落了。因此,还得打碎这个处所短语,重新进行能够导致有意义的句法解析。既然这“嘴”不是“你的”了,那么可以寻找它跟“骂”之间的概念语义关系(论元角色关系):骂人是动嘴的,所以“骂”和“嘴”之间可能存在“动作—工具”关系,于是,“你”就可能是“骂”的对象(受事角色);也就是说,是“某一个人或某一群人动嘴骂你”,底层结构是“[某个/群人]动嘴骂你”,通过关系化表达,让工具格“嘴”作中心语,形成“骂你的嘴”。然后,再后附方位词“里”,处所化为“骂你的嘴里”;并且,表示范围的“里”还使“骂你的嘴”在语义上复数化,即不止是一个人的一张嘴骂你。接续下文“有我的嘴”,表示我也是骂你的一员。这种山重水复、峰回路转的语义表达方式,的确别致有趣,也十分烧脑费神。因此,引起争议也在所难免,尽管作者对此不以为然。^④

^④记得林斤澜自己评论这两句话时说:“有什么难懂的?真是的!”出处失记。

我们要求 ChatGPT 先阅读上面例(4)中的语句,然后问答下列问题:

笔者:谁骂了谁?谁打了谁?为什么?

ChatGPT:这句话的意思是,骂人和打人的行为是通过某人进行的,但实际上是另一个人在背后指使。也就是说,有人通过他人骂人和打人。

笔者:骂人的是一个人还是许多人?被骂的是一个人还是许多人?打人的是一个人还是许多人?被打的是一个人还是许多人?为什么?

ChatGPT:根据这句话,骂人的行为和打人的行为都可以是由一个人进行的,也可以是由许多人进行的。这句话没有明确指出人数,但暗示了这种行为是由某个人或某些人操控的。

— 骂人的是:一个人或许多人(“嘴里有我的嘴”暗示说话者可以是单数或复数)。

— 被骂的是:可以是一个人或许多人(“你”可以指单数或复数)。

— 打人的是:一个人或许多人(“手里有我的手”暗示打人的人可以是单数或复数)。

— 被打的是:可以是一个人或许多人(“你”可以指单数或复数)。

笔者:骂人所用的工具是什么?是单数(一个)还是复数(许多)?打人所用的工具是什么?是单数(一个)还是复数(许多)?为什么?

ChatGPT:骂人所用的工具是嘴巴。由于“嘴里有我的嘴”这句话,说明骂人用的是“嘴”。这里的“嘴”虽然是一个器官,但因为每个人都有嘴巴,所以可以理解为单数的工具(一个人的嘴)或复数的工具(许多人的嘴)。打人所用的工具是手。“手里有我的手”这句话表明打人用的是“手”。和嘴一样,虽然“手”是一个器官,但每个人都有手,所以也可以理解为单数的工具(一个人的手)或复数的工具(许多人的手)。

总结:这句话隐含了指使者和执行者的关系,实际操作可以由一个人或许多人进行,具体人数和工具数量取决于实际情况。

可见,ChatGPT 跟许多人初次瞥见这个花园路径式的对偶句时的反应一样,有点儿发懵;最后,竟然自作聪明地认为“这句话的意思是,骂人和打人的行为是通过某人进行的,但实际上是另一个人在背后指使”。即便如此,我们也不能指责 ChatGPT 是“随机鹦鹉”;因为,我们许多人在这种句子的理解上也困难重重。说句公道话:人犹如此,机何以堪?

四 ChatGPT 在理解一般递归结构方面的出色表现

美国知名经济学家杰弗里·萨克斯在谈美国对中国的“神经质的反应”(neurotic reaction)和恐惧时说(2024-5-24 从网上短视频中记录):

(5) Everything is about China, not because China's threatening us. But because China is showing we don't run the world the way we like to think we run the world.

笔者:请把上面的英语句子翻译成汉语。

ChatGPT:一切都关乎中国,不是因为中国在威胁我们,而是因为中国正在展示我们并不像我们所想的那样掌控世界。

可见,ChatGPT 对于复杂的宾语从句(其中有部分的主谓结构“we [don't] run the world”的递归)能够正确地理解,并且翻译为流畅的汉语。

在汉语和英语等 SVO 型语序的语言中,宾语小句的递归性嵌套是右向分枝,逐步从左向右展开,所以比较容易生成和理解。例如:

(6)我们知道他们在说谎,他们也知道他们在说谎,他们知道我们知道他们在说谎,我们也知道他们知道我们知道他们在说谎,但是他们依然在说谎。(俄·索尔仁尼琴《古拉格群岛》)

其中的“我们也知道他们知道我们知道他们在说谎”,是一个由三重“主语—述语—宾语”结构层层嵌套构成的递归结构:最底层的“主语—述语—宾语”结构($a_3 \frown b_3 \frown c_3$)^⑤,充当上一层的“主语—述语—宾语”结构($a_2 \frown b_2 \frown c_2$)的宾语(即 $c_2 = a_3 \frown b_3 \frown c_3$);这中层的“主语—述语—宾语”结构($a_2 \frown b_2 \frown c_2$),又充当最外层的“主语—述语—宾语”结构($a_1 \frown b_1 \frown c_1$)的宾语(即 $c_1 = a_2 \frown b_2 \frown c_2$)。如果纯粹从左向右的句法成分的表层序列来看,我们看到的是一个“ $a_1 \frown b_1 \frown a_2 \frown b_2 \frown a_3 \frown b_3 \cdots$ ”式的主谓结构右向递归结构。正如 Chomsky(1956)所指出的,这种低层次的主谓结构逐层充当上一层的主谓结构的宾语的嵌套结构,是有限状态马尔可夫过程模型(the model of finite state Markov processes)无法处理的。

特别需要注意的是,其中的谓语核心“知道”是正叙实动词(factive verb),例(6)是一种叙实性的多重递归表达。为了增加复杂性,我们还分别用反叙实动词(counter factive verb)“以为”和非叙实动词(non-factive verb)“相信”替换“知道”;然后,让 ChatGPT 翻译成英语,并且回答问题:

(6-1)我们知道他们在说谎,他们也知道他们在说谎,他们知道我们知道他们在说谎,我们也知道他们知道我们知道他们在说谎。

ChatGPT: We know they are lying, they also know they are lying, they know we know they are lying, and we know they know we know they are lying.

笔者:他们说谎了吗?

ChatGPT: 回答:是的,他们在说谎。

(6-2)我们以为他们在说谎,他们也以为他们在说谎,他们以为我们以为他们在说谎,我们也以为他们以为我们以为他们在说谎。

ChatGPT: We think they are lying, they also think they are lying, they think we think they are lying, and we think they think we think they are lying.

笔者:他们说谎了吗?

ChatGPT: 不确定,因为这只是双方的猜测或认为。

(6-3)我们相信他们在说谎,他们也相信他们在说谎,他们相信我们相信他们在说谎,我们也相信他们相信我们相信他们在说谎。

ChatGPT: We believe they are lying, they also believe they are lying, they believe we believe they are lying, and we believe they believe we believe they are lying.

笔者:他们说谎了吗?

ChatGPT: 不确定,因为这只是信念或主观的看法,并不代表事实。

可见,ChatGPT 不仅能够正确地理解多重递归的宾语从句,还能够翻译为流畅的汉语;但是,理解不同类型的动词的叙实性差别的能力有所不足:理解“知道”是正叙实动词,预设其宾语小句所表示的命题为真;理解“相信”是非叙实的信念动词,表示说话人对宾语小句所表示的

^⑤这种句法成分的符号化与关系链接表示,参考 Chomsky (1956)。

命题可能为真的一种主观态度;但是,不理解“以为”除了有“认为、相信”一样的非叙实用法之外,还有一种反叙实用法:表示说话人预设宾语小句所表示的命题为假。^⑥

五 ChatGPT 在理解中心内嵌结构方面的不佳表现

下面这个复杂句改编自 19 世纪英格兰童谣(nursery rhyme),是语言学文献上经常引用的宾语受定语从句修饰的嵌套结构:

(7) This is the dog that chased the cat that killed the rat that ate the malt that lay in the house that Jack built.(这是那只追赶咬死吃了放在杰克建造的房子上的麦芽的老鼠的猫的狗)

对于英语这种允许修饰语后置的语言来说,在造句时具有开放性的右向递归的优势;即在宾语(或表语, predicate)之后附加定语从句,逐层递归,不断打开新的修饰语窗口;结果,从高层宾语(或表语,比如下面的 c_1)到内嵌近于最深的底层宾语(或表语,比如下面的 c_5),逐层展开其定语从句,形成从上到下的阶梯形语句结构。为了展示语言线性序列背后的层次构造,可以用括号和分行,另加句法成分的符号化与关系链接,分别表示为^⑦:

(7-1-1) This is the dog { that chased the cat [that killed the rat (that ate the malt < that lay in the house ‘ that Jack built ’ >)] }.

(7-1-2) This is the dog($a_1 \cap b_1 \cap c_1$)
that chased the cat($\# \cap a_2 = \emptyset \cap b_2 \cap c_2$)
that killed the rat($\# \cap a_3 = \emptyset \cap b_3 \cap c_3$)
that ate the malt($\# \cap a_4 = \emptyset \cap b_4 \cap c_4$)
that lay in the house($\# \cap a_5 = \emptyset \cap b_5 \cap c_5$)
that Jack built.($\# \cap a_6 \cap b_6 \cap c_6 = \emptyset$)

粗略地说,这是一个由六重“主语—述语—宾语”结构层层嵌套构成的递归结构:最底层的“主语—述语”结构($a_6 \cap b_6$),在关系代词“that”(记作 $\#$)的引导下,充当上一层的“述语—宾语”结构($b_5 \cap c_5$)的宾语的定语从句(即 c_5 受 $\# \cap a_6 \cap b_6$ 修饰);这次底层的“述语—宾语”结构($b_5 \cap c_5$),又在关系代词“that”的引导下,充当近中层的“述语—宾语”结构($b_4 \cap c_4$)的宾语的定语从句(即 c_4 受 $\# \cap b_5 \cap c_5$ 修饰);再从这近中层到中层、从中层到次高层、直到最高层,逐层作上一层的宾语的定语从句。如果纯粹从自左向右的句法成分的表层序列来看,我们看到的是一个“ $a_1 \cap b_1 \cap c_1 \cap \# \cap b_2 \cap c_2 \cap \# \cap b_3 \cap c_3 \cap \# \cap b_4 \cap c_4 \cap \# \cap b_5 \cap c_5 \cap \# \cap a_6 \cap b_6$ ”式的“(主语—)述语(—宾语)”结构作定语从句的右向递归结构。显然,这种低层次的主谓/述宾结构逐层充当上一层的述宾结构的宾语的定语从句的嵌套结构,也是马尔可夫过程模型或 n-grams 模型所无法处理的。

但是,对于汉语这种修饰语前置的语言来说,在造这种句子时只能封闭性地左向递归;即从最外围的高层宾语(或表语)到内嵌最深的底层宾语(或表语),在它们之前逐层附加定语从

^⑥ 吕叔湘主编(1980/2021:619)指出“以为”的一种用法:用“以为”做出的论断往往不符合事实,用另一个小句指明真相。例如:“我以为有人敲门,其实不是。| 原来是你,我还以为是老王呢。”

^⑦ 我们设定从外层到内层依次是:花括号、方括号、圆括号、尖括号和单引号,即{[(< ‘ >)]};另外,我们用井号“#”标记“that”等引出定语从句的关系代词,用“ $\emptyset$ ”标记零形式;并且忽略不同的定语从句内部的结构差异,把它们都分析成“主语—述语—宾语/表语”结构。

句(述宾短语+“的”);结果,修饰语中的述语从高层到底层被从左向右地挤成一个邻接的序列,接着修饰语中的宾语连同“的”从底层到高层被从左向右地挤成一个邻接的序列;最终,形成一个先从上到下、再从下到上的喇叭口形的层次结构。可以用括号和分行两种方式分别表示为:

(7-2-1)这是那只{追赶[咬死(吃了<放在‘杰克建造的房子’的麦芽>的老鼠)的猫]的}狗。

(7-2-2)这是那只($a_1 \wedge b_1 \wedge \dots \wedge c_1$)狗
 追赶($b_2 \wedge \dots \wedge c_2 \wedge de$)猫的
 咬死($b_3 \wedge \dots \wedge c_3 \wedge de$)老鼠的
 吃了($b_4 \wedge \dots \wedge c_4 \wedge de$)麦芽的
 放在($b_5 \wedge \dots \wedge c_5 \wedge de$)房子里的
 杰克建造的($a_6 \wedge b_6 \wedge de$)

粗略地说,这是一个由六重“主语—述语—宾语”结构层层嵌套构成的递归结构:最底层的“主语—述语”结构($a_6 \wedge b_6$),在具有名词化功能的助词“的”(记作 de)的作用下,形成相当于英语关系化小句的名词性短语,充当上一层的“述语—宾语”结构($b_5 \wedge c_5$)中核心名词的修饰语(即 c_5 的核心“房子”受 $a_6 \wedge b_6 \wedge de$ 修饰);这次底层的“述语—宾语”结构($b_5 \wedge c_5$),又在助词“的”的作用下,充当近中层的“述语—宾语”结构($b_4 \wedge c_4$)中宾语的修饰语(即 c_4 受 $b_5 \wedge c_5 \wedge de$ 修饰);再从这近中层到中层、从中层到次高层、直到最高层,逐层作上一层的宾语的相当于定语从句的修饰语。如果纯粹从自左向右的句法成分的表层序列来看,我们看到的是一个“ $a_1 \wedge b_1 \wedge b_2 \wedge b_3 \wedge b_4 \wedge b_5 \wedge a_6 \wedge b_6 \wedge de \wedge c_5 \wedge de \wedge c_4 \wedge de \wedge c_3 \wedge de \wedge c_2 \wedge de \wedge c_1$ ”式的“(主语—述语—宾语)的”结构作定语从句的右向递归结构。显然,这种低层次的主谓/述宾结构逐层充当上一层的述宾结构的宾语的定语从句的嵌套结构,也是马尔可夫过程模型或 n-grams 模型所无法处理的。

从话题结构的角度看,例(7)是以“狗”为表语性话题展开的话语,如果直接用“麦芽”为主语性话题来展开话语,那么语句结构会怎么样呢?我们尝试把例(7)改写为下面的例(8),并且分别用括号和分行两种方式,来显示其层次构造:

(8)The malt the rat the cat the dog chased killed ate was laid in the house that Jack built.(那条狗追赶的那只猫杀死的那只老鼠偷吃的麦芽放在杰克建造的房子。)

(8-1-1)The malt {the rat [the cat (the dog chased) killed] ate} was laid [in the house that Jack built].(那条狗追赶的那只猫杀死的那只老鼠偷吃的麦芽放在杰克建造的房子)

(8-1-2)The malt($a_1 \wedge \dots \wedge b_1$)was laid in the house that Jack built.
 the rat($a_2 \wedge \dots \wedge b_2$)ate
 the cat($a_3 \wedge \dots \wedge b_3$)killed
 the dog chased($a_4 \wedge b_4$)

粗略地说,这是一个由四重“主语—述语”结构层层嵌套构成的递归结构:最底层的“主语—述语”结构($a_4 \wedge b_4$),直接充当上一层的“主语—述语”结构($a_3 \wedge b_3$)的主语的定语从句(即 a_3 受 $a_4 \wedge b_4$ 修饰);次底层的“主语—述语”结构($a_3 \wedge b_3$),又充当次高层的“主语—述语”结构($a_2 \wedge b_2$)的主语的定语从句(即 a_2 受 $a_3 \wedge b_3$ 修饰);再从次高层到最高层,逐层作上一层的主语的定语从句。如果纯粹从自左向右的句法成分的表层序列来看,是一个“ $a_1 \wedge a_2 \wedge a_3 \wedge a_4$

$\neg b_4 \neg b_3 \neg b_2 \neg b_1$ ”式的“(主语—)述语”结构作定语从句的右向递归结构。显然,这种低层次的主谓结构逐层充当上一层的主谓结构的主语的定语从句的嵌套结构,也是有限状态马尔可夫过程模型或 n-grams 模型所无法处理的。

可见,对于英语这种定语从句后置的语言来说,这种语句构造在向右逐层递归,不断打开新的定语从句的窗口时,定语从句中的主语从高层到底层也被从左向右地挤成一个邻接的序列,接着定语从句中的述语从底层到高层被从左向右地挤成一个邻接的序列;最终,形成一个先从上到下、再从下到上的喇叭口形(凹槽式)的层次结构。相反,相应的汉语句子结构,则利用定语从句前置的特点,从左向右逐层打开定语从句(主谓短语+“的”)的窗口;结果形成了作为主语核心的名词的、从最外围的定语到内嵌最深的贴身定语的序列。可以用括号和分行分别表示为:

(8-2-1){那条狗追赶的[那只猫杀死的(那只老鼠偷吃的)]}麦芽放在杰克建造的房子
里。

(8-2-2)那条狗追赶的($a_1 \neg b_1 \neg de$)

那只猫杀死的($a_2 \neg b_2 \neg de$)

那只老鼠偷吃的($a_3 \neg b_3 \neg de$)

麦芽 放在杰克建造的房子里($a_4 \neg b_4$)。

粗略地说,这是一个由四重“主语—述语”结构层层嵌套构成的递归结构:最上层的“主语—述语”结构($a_1 \neg b_1$)在助词“的”的作用下,形成相当于英语关系化小句的名词性短语,直接充当下一层的“主语—述语”结构($a_2 \neg b_2$)的主语的定语从句(即 a_2 受 $a_1 \neg b_1 \neg de$ 修饰);次上层的“主语—述语”结构($a_2 \neg b_2$),又在助词“的”的作用下,形成名词性短语,充当次底层的“主语—述语”结构($a_3 \neg b_3$)的主语的定语从句(即 a_3 受 $a_2 \neg b_2 \neg de$ 修饰);再从次底层到最底层,逐层作下一层的主语的定语从句。如果纯粹从自左向右的句法成分的表层序列来看,是一个“ $a_1 \neg b_1 \neg de \neg a_2 \neg b_2 \neg de \neg a_3 \neg b_3 \neg de \neg a_4 \neg b_4$ ”式的“(主语—)述语”结构作定语从句的右向递归结构。显然,这种高层次的主谓结构逐层充当下一层的主谓结构的主语的定语从句的嵌套结构,也是有限状态马尔可夫过程模型或 n-grams 模型所无法处理的。

比较汉语和英语,可以发现,对于表达连环的施受事件链条来说,汉语之类定语(从句)前置型语言,适合表达最末端受事作话题的命题;而英语之类定语从句后置型语言,适合表达最初始端施事作话题的命题。当我们要求 ChatGPT 把例(7)译成汉语时,结果是:

ChatGPT:这只狗追赶那只猫,那只猫杀死了那只老鼠,那只老鼠吃了那放在杰克盖的房子里的麦芽。

可见,ChatGPT 不擅长处理多重定语从句嵌套的结构,所以把外围的前面两个定语从句翻译成独立的主谓句。结果,事件之间的时间先后关系模糊了,好像是在说一只被猫杀了的死耗子吃了麦芽似的。事实上,在自然语言的真实文本中,很少有例(7)这种人为刻意杜撰的多重定语从句的嵌套结构。有意思的是,例(7)所从改编的童谣,原来也是避免一开篇就使用多重定语从句;而是从内层到外层,先使用“系词+受单重定语从句修饰的表语”的结构,再使用“系词+受多重定语从句修饰的表语”的结构,逐步地增加语句的结构层次的复杂度。例如:⑧

(9) This is the house that Jack built.(这是杰克建造的房子。)

This is the malt that lay in the house that Jack built.(这是放在杰克建造的房子

⑧ 详见 https://en.wikipedia.org/wiki/This_Is_the_House_That_Jack_Built。

麦芽。)

This is the rat that ate the malt.(这是吃那麦芽的老鼠。)

That lay in the house that Jack built.(它[=那麦芽]放在杰克建造的房子里。)

This is the cat.(这是那只猫。)

That killed the rat that ate the malt.(它[=那只猫]杀死了吃那麦芽的老鼠。)

That lay in the house that Jack built.(它[=那麦芽]放在杰克建造的房子里。)

This is the dog that worried the cat.(这是让那只猫担心的狗。)

That killed the rat that ate the malt.(它[=那只猫]杀死了吃麦芽的老鼠。)

That lay in the house that Jack built.(它[=那麦芽]放在杰克建造的房子里。)

在语法学上,像(7-2)和(8-1)这种递归方式叫中心嵌套(centre-embedding),大概是最复杂的一种嵌套结构。其特点是述宾(或主谓)结构中多批次地内嵌述宾(或主谓)结构,形成多重述宾(或主谓)结构的递归性嵌套。比如,(7-2)的构造特点是:述宾结构(记作“V+O”)的宾语逐层地被前置的由述宾结构名词化造成的成分修饰,结果形成了近于“ $V_1 V_2 \cdots V_{n-1} V_n O_n -de_n O_{n-1} -de_{n-1} \cdots O_2 -de_2 O_1 -de_1$ ”的句法成分序列(这里用“-de”表示名词化标记“的”)。简而言之,是内嵌最深的述宾结构把内嵌次深的述宾结构挤断,就这样下一层的述宾结构依次挤断上一层的述宾结构,形成定语从高到下和宾语从下到高邻接排列的格局。而(8-1)的构造特点是:主谓结构(记作“S+P”)的主语逐层地被后置的由主谓结构造成的定语从句修饰,结果形成了“ $S_1 S_2 \cdots S_{n-1} S_n P_n P_{n-1} \cdots P_1$ ”的句法成分序列。简而言之,是内嵌最深的主谓结构把内嵌次深的主谓结构挤断,就这样下一层的主谓结构依次挤断上一层的主谓结构,形成主语从高到下和谓语从下到高邻接排列的格局。

六 语言学家在理解中心内嵌结构方面的不佳表现

上述这种多层次的中心嵌套结构,在自然语言的真实文本中极其罕见;通常只有在语言学文献上,才能偶尔看到。有意思的是,由于这种多重嵌套结构的人为性和不自然性;即使是专业的语言学家,也会造出顾此失彼、前后照应不周的句子。例如(见 Everett 2017: 236/2020: 259):

(10) A man that a woman that a child that a bird that I heard saw knows loves sugar.

这个句子可以解析为:“ $a_1 \cap \# \cap a_2 \cap \# \cap a_3 \cap \# \cap a_4 \cap \# \cap a_5 \cap b_5 \cap b_4 \cap b_3 = \emptyset \cap b_2 \cap b_1 \cap c_1$ ”式的“(主语—)定语”结构作定语从句的右向递归结构。当我们努力把它翻译成中文时,却出现了(实际上是“发现了”)问题:这是一个“主语+谓语(定语+宾语)”多重中心内嵌的结构,“A man loves sugar”是主句;无论从最高层的主语“A man”开始,还是从内嵌最深的主谓结构的主语“I”开始,主语和谓语都无法一一匹配,也就是说少了一个谓语动词。其主要的句法语义关系,可以用中文表示如下:

(10-1) 我听见(heard)的鸟[叫]→那只鸟看到(saw)的孩子→那个孩子认识(knows)的女人→那个女人???的男人→那个男人喜欢(loves)糖。

可见,多重中心内嵌是一种极其耗费心力(psychological effort)的语法结构。以至于 Everett(2017:236/2020:260)要特别说明:“实际上,有人声称它们只存在于语言学家的脑海里,尽管我认为这一说法太绝对了。”

为了检测 ChatGPT 对于多重中心内嵌结构的理解能力,我们通过对对话的方式,让它完成一系列任务:

笔者:请把上面的英语句子(10)翻译成汉语。

ChatGPT:一个男人,那个女人,那个孩子,那个鸟,我听说看见的,知道的,爱吃糖。

笔者:请把(10)这个英语句子中发生语法关系的成分列出来,比如:

I heard a bird, a bird saw a child, a child knows a woman, a woman...

ChatGPT:— I heard a bird; — a bird saw a child; — a child knows a woman; — a woman knows a man; — a man loves sugar

笔者:检查一下(10)这个句子有没有语法错误? 有没有缺少动词? 如果缺少,请补充出来。

ChatGPT:这个句子没有语法错误,但是有些部分可以更加清晰地表达,补充动词的方式如下:— I heard a bird; — a bird saw a child; — a child knows a woman; — a woman knows a man; — a man loves sugar

笔者:请调整(10)这个句子的结构,使其更容易理解:

ChatGPT:A man, whom a woman knows, whom a child saw, whom a bird I heard knows, loves sugar.

可见,ChatGPT 翻译像例(10)这种多重主谓结构嵌套的句子的能力并不强,因为不能很好地发现不同层次上的主谓结构及其从句化表达,只能按照原句的顺序来逐个铺排不同层次上的多个主语和多个谓语及其名词化形式(“的”字结构)。我们通过例示(I heard a bird, a bird saw a child, a child knows a woman),来让它找出其他潜在的主语—谓语关系,结果它不仅正确地找到了“a man loves sugar”,而且还为本来遗漏谓语的主语“a woman”补充了谓语动词([a woman] knows [a man])。我们不知道这种增补算不算幻觉(hallucination),但是能够感觉到大模型的确对句法结构及其构成成分之间的关系是比较敏感的。正因为它自动地补充进去了一个动词,所以也就不能发现其中的语法错误。在改写和调整句子时,把在“a bird saw a child”和“a child knows a woman”等关系搞混了。

事实上,上面例(10)的语法上正确的版本应该是类似下面例(11)这种形式:

(11)A man that a woman that a child that a bird that I heard saw knows likes loves sugar.(我听见[它啼叫]的鸟看到的孩子所认识的女人喜欢的男人爱吃糖。)

可以用短竖线的多少来表示结构上内嵌层次的深浅,大致地表示如下:

(11-1)A man || that a woman |||| that a child |||||| that a bird |||||||| that I heard ||||||| saw ||||| knows ||| likes | loves sugar.

(11-2)我听见[它啼叫]的 ||||| 鸟看到的 |||| 孩子所认识的 ||| 女人喜欢的 || 男人 | 爱吃糖。

于是,我们又以语法上正确的句子例(11)为讨论的对象,跟 ChatGPT 进行对话,让它完成一系列任务:

笔者:请把上面的英语句子(11)翻译成汉语。

ChatGPT:一个男人,一个女人,一个孩子,一只鸟,我听说看见的,知道的,喜欢的,爱吃糖。

笔者:请把上面(11)这个英语句子中发生语法关系的成分列出来。比如:I heard a bird, a bird saw a child, a child knows a woman, a woman...

ChatGPT:1. I heard a bird. 2. A bird saw a child. 3. A child knows a woman. 4. A woman likes a man. 5. A man loves sugar.

所以,完整的句子结构是:— 我听说(I heard) — 一只鸟(a bird) — 看见了(saw) — 一个孩子(a child) — 知道(knows) — 一个女人(a woman) — 喜欢(likes) — 一个男人(a man) — 爱吃糖(loves sugar)

可见,对于例(11)这个语法结构完整的多重内嵌的英语句子,ChatGPT 虽然可以把其中发生语法关系的成分都列出来,但是依然不能翻译成像(11—2)那样的层层内嵌的汉语句子。究其原因,ChatGPT 是一种基于统计的生成式语言模型;它可以通过学习数据背后的概率分布,来对数据分布和数据的生成过程进行建模,从而生成跟观察数据类似的新数据。^⑨ 而像例(7)(10)和(11)之类多重中心内嵌结构在真实文本中极其罕见,在训练语料中也不可多得;结果,ChatGPT 等语言模型在训练时没有办法学到这种句子的分布模式。因此,我们相信,如果用一定规模的多重中心内嵌结构语料去训练和微调语言模型,那么 ChatGPT 等语言模型应该可以更好地理解和生成这一类对于人类来说也有相当的挑战性的句子。

七 结语:新型智能科学视野下多种形式的语言理解

从上面的测试和分析来看,ChatGPT 对于指称歧义句、花园幽径句和递归嵌套句等三种复杂句的语义理解,在某些方面似乎达到了人类的水平。更加公平的说法是,ChatGPT 等语言大模型基本上可以理解句子表面的语言性意义;但是,未必能够理解句子中隐微的主观情绪和人际互动方面的社会性含义,更是缺乏人类接触语言符号时所产生的肉身体验和具身认知。比如,人听到“挠痒痒”会激发想要笑出来的感觉;但是,语言模型不具有肉身,不可能有这种切身感受和具身认知。(详见袁毓林 2024a)这说明,问“机器/语言大模型能不能理解语言的意义?”,就像问“潜水艇会不会游泳?”一样^⑩,可能是使用了“错误的范畴”。(详见 Gopnik 2022; Mitchell & Krakauer 2023)因为,我们人类对于语言符号的意义的掌握,并不是源于我们处理自然语言的能力;而是源于我们通过在外部世界采样(sampling)、跟物理世界和人类群体互动(interacting),藉此而积累起来的对生活世界(lived world)的更加基本的理解。显然,处理自然语言文本的转换器等人工智能系统,并不跟赋予词语和句子以意义的互动发生直接的关系。(详见 Pezzulo et al. 2024)

尽管如此,探讨“大语言模型能否像人类一样理解自然语言,从而理解语言所描述的物理场景和社会场景?”这种问题,依然具有紧迫的现实意义。因为,对这个问题的不同回答,将决定我们是否信任机器驾驶车辆、诊断疾病、看护老人和教育儿童(详见 Mitchell & Krakauer 2023),以及人工智能产品能不能进入我们人类的日常生活。鉴于 ChatGPT 等语言大模型是通过高维度的词向量、多头注意力机制和前馈神经网络来表示和发现词语和句子之间的关系,从而抓住语句的表面意义的(详见袁毓林 2024b),“语义理解”这个概念(范畴)需要重新界定,更需要拓宽其使用范围,并且进行细致的划分,厘清不同用法之间的关系。正如 Mitchell & Krakauer (2023)所设想的:“应该发展这样一种智能科学,使其能够为不同形式的‘理解’、这些不同形式理解的长处和局限、以及如何整合不同形式的认知等问题提供启发。”在这种新型智能科学的视野下,超越经典的简单化的“是/否(能够理解)”二分,更好地讲述生命的自然智

^⑨ 关于生成式语言模型,详见刘伟(2024)。

^⑩ 荷兰计算机科学家 Edsger Wybe Dijkstra(1930—2002)有一句名言:“The question of whether a machine can think is no more interesting than the question whether a submarine can swim.”(问机器会不会思考比问潜水艇会不会游泳更无趣)。

能和机器的人工智能两种系统的不同的理解能力(详见 Pezzulo et al. 2024),更加全面和合理地回答“ChatGPT 等语言模型是不是随机鹦鹉”之类的问题。

更进一步,考虑到生成式人工智能已经在学术界渐趋普及,并且逐步进入公众视野,相关学科的专家有必要向大家解答下列问题:1)如何合理可靠地评估生成式人工智能系统的能力? 2)如何理性地避免陷入“伊莱莎效应”? 3)如何避免把先进的人工智能系统的行为拟人化的倾向?^⑩ 从而开创智能科学的新局面。正是在这种发展新型智能科学的学术潮流中,传统的语言学科如果积极地投身其中,那么也许可以获得浴火重生的重大机遇。比如,认知科学一般认为:人类主体跟外部环境的互动,对于形成理解和引导意义是十分重要和必要的。但是,现代大型语言模型显然不具备跟文本相关的外部环境进行互动的条件与可能性,它主要是通过词向量来表示词语和句子的意义的。于是,自然地就会产生下面一系列相关的问题:1)词向量的矩阵中到底包括多少语言意义赖以生成的主体与环境互动的信息? 2)对于语言大模型来说,这种推动语言意义生成的主体与环境的互动,有哪些(或多少)是必要的? 3)从语言大模型的卓越表现来看,对于语言意义的生成来说,主体与环境互动的作用有没有被现有的具身认知理论所夸大? 4)生成式人工智能系统能不能在没有跟世界进行感觉运动交流的初始环境下,获得对语言的真正的理解? 5)除了具身的跟世界进行感觉运动交流,人类沟通与语言交流行为是不是也能产生意义和理解? 6)单纯地用人类沟通行为所产生的文本来训练语言模型,是不是也能产生足够的意义和理解?(参考 Pezzulo et al. 2024)对于诸如此类的问题,语言学家应该是有兴趣且有能力和能力参与讨论和进行探索的。如果能够这样,那么语言学也必将发生脱胎换骨般的变迁和飞跃。

参考文献

- 刘伟 2024 生成式、启发式与产生式,微信公众号“人机与认知实验室”2024-07-20, <https://mp.weixin.qq.com/s/aJ--z08ABKFqCv1PXkrtw>。
- 吕叔湘主编 1980 《现代汉语八百词》(增订本),北京:商务印书馆,2021年。
- 袁毓林 2021 “人机对话—聊天机器人”与话语修辞,《当代修辞学》第3期。
- 袁毓林 2024a 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,《语言战略研究》第1期。
- 袁毓林 2024b ChatGPT 等大模型的语言处理机制及其理论蕴涵,《外国语》第4期。
- Bender, Emily M. & Alexander Koller 2020 Climbing towards NLU: On meaning, form, and understanding in the age of data. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*: 5185-5198.
- Bender, Emily M., Timnit Gebru, Angelina McMillan-Major & Shmargaret Shmitchell 2021 On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610-623. FAccT '21, 2021-03-01. New York: Association for Computing Machinery.
- Chomsky, Noam 1956 Three models for the description of language. *IRE Transactions on Information theory* 2(3): 113-124.
- Everett, Daniel L. 2017 *How Language Began: The Story of Humanity's Greatest Invention*. New York / London: Liveright Publishing Corporation. 何文忠、樊子瑶、桂世豪译《语言的诞生:人类最伟大发明的故事》,北京:中信出版集团,2020年。

^⑩这里的“拟人化”指人们可能会把人类自己的行为 and 特征投射到 AI 模型等非人类实体上。

- Gazzaniga, Michael S., Richard B. Ivry & George R. Mangun 2009 *Cognitive Neuroscience: The Biology of the Mind* (3rd edn.). New York: W. W. Norton & Company, Inc. 周晓林、高定国等译《认知神经科学——关于心智的生物学》,北京:中国轻工业出版社,2011年。
- Gopnik, Aliso 2022 What AI still doesn't know how to do. *Wall Street Journal*, July 15, <https://www.wsj.com/articles/what-ai-still-doesnt-know-how-to-do-11657891316>.
- Levelt, J. M. Willem 1989 *Speaking: From Intension to Articulation*. Cambridge, MA: The MIT Press.
- Michael, J., A. Holtzman, A. Parrish, A. Mueller, A. Wang, A. Chen, D. Madaan, N. Nangia, R. Y. Pang, J. Phang & S. R. Bowman 2022 What do NLP researchers believe? Results of the NLP community metasurvey. arXiv: 2208.12852.
- Mitchell, Melanie & David C. Krakauer 2023 The debate over understanding in AI's large language models. *PNAS* 120(13), e2215907120. 大语言模型真能“理解”语言吗?,微信公众号“神经现实”2023-09-19, <https://mp.weixin.qq.com/s/7YZR2BROlvapsiWZiLnwPg>.
- Pezzulo, Giovanni, Thomas Parr, Paul Cisek, Andy Clark & Karl Friston 2024 Generating meaning: Active inference and the scope and limits of passive AI. *Trends in Cognitive Sciences* 28(2): 97-112.
- Weizenbaum, Joseph 1976 *Computer Power and Human Reason: From Judgment to Calculation*. New York: W. H. Freeman & Company.

Tests and discussions on whether large language models can understand linguistic meanings from the perspective of three complex sentences:

Is it the case that ChatGPT is a stochastic parrot

YUAN Yulin

Abstract This paper tests ChatGPT's performance in understanding the semantic meanings of three types of complex sentences: referential ambiguity sentences, garden-path sentences, and recursive sentences, especially center-embedding recursive sentences, for answering the following questions: Can modern large language models like ChatGPT truly understand the meaning of natural languages? Are they merely stochastic parrots that disregard meaning? The paper finds that ChatGPT excels in understanding referential ambiguity sentences and normal recursive structures but performs averagely in garden-path sentences. And similar to humans, it struggles in confusion to understand center-embedding recursive structures. Based on these tests, the paper argues that while large language models like ChatGPT can comprehend the linguistic meaning of sentences, they may not fully grasp the subtle communicative and embodied meanings behind them. Finally, the paper looks ahead, suggesting that linguistics study can make a unique contribution to the development of a new intelligent science that understands the multifaceted dimensions of meaning.

Keywords referential ambiguity sentences; garden-path sentences; recursive-embedded sentences; ChatGPT; large language model (LLM); semantic understanding; linguistic meaning; communicative meaning; embodied meaning

(袁毓林 519000 澳门大学人文学院中国语言文学系/
100871 北京大学中文系/中国语言学研究
(责任编辑 高晓虹)