

人工智能如何“说人话”？

——对于自然语言处理研究的哲学反思

**How to Make Artificial Intelligence Capable of Speaking Human Language?
Some Philosophical Remarks on Natural Language Processing**

徐英瑾 /XU Yingjin 陈萌 /CHEN Meng

(复旦大学哲学学院, 上海, 200433)
(School of Philosophy, Fudan University, Shanghai, 200433)

摘要:目前的自然语言处理技术的发展, 依然面临如下问题: (1) 不同的自然语言处理机制之间缺乏融合; (2) 自然语言处理技术与人工智能研究的其他技术缺乏彼此融合; (3) 基于大数据的自然语言处理技术的运作必须以“剥削”人类的智能为前提; (4) 基于大数据的自然语言处理技术缺乏灵活处理隐喻、反讽、双关等修辞现象的能力。这些问题所涉及的主要哲学问题有: (1) 语言是外部世界的表征, 还是言说者内部世界的表征? (2) 语言中的规则, 究竟是先验的, 还是经验的? (3) 语言表征与言说者的心理活动之间的关系为何? (4) 与语言表征有关的认知构架, 在多大程度上需要被“具身化”?

关键词: 人工智能 自然语言处理 认知语言学 语言哲学 大数据 图式

Abstract: There are four salient problems in the current state of Natural Language Processing (NLP): (1) There is no real integration among different sub-areas of NLP; (2) There is no real integration among NLP and other branches in AI; (3) The limited successes of big-data-based NLP approaches are based on the exploitation of human labor; (4) Big-data-based NLP approaches do not perform well when they are employed to handle linguistic data with rhetorical features. The most NLP-related philosophical problems include: (1) Do human language primarily represent the external world or the internal world of the language-users? (2) Are linguistic rules innately fixed in human minds or results of the influence of language-learning environments? (3) What about the relationship between linguistic expressions and human cognitive architecture? (4) In what sense and what degree is it necessary to make language-related cognitive architecture “embodied”?

Key Words: Artificial intelligence (AI); Natural language processing (NLP); Cognitive linguistic; Philosophy of language; Big-data; Schema

中图分类号: N0: TP18 文献标识码: A DOI: 10.15994/j.1000-0763.2022.01.002

基金项目: 国家社科重大项目“基于信息技术哲学的当代认识论研究”(项目编号: 15ZDB020); 教育部哲学社会科学重大课题攻关项目“新一代人工智能发展的自然语言理解研究”(项目编号: 19JZD010)。

收稿日期: 2021年1月7日

作者简介: 徐英瑾(1978-)男, 上海人, 复旦大学哲学学院教授, 研究方向为人工智能哲学与英美分析哲学研究。Email: yjxu@fudan.edu.cn

陈萌(1989-)女, 陕西西安人, 复旦大学哲学学院博士研究生, 研究方向为现代外国哲学。Email: 17110160018@fudan.edu.cn

一、导论：为何对人工智能来说，“自然语言处理”非常重要？

概而言之，所谓“人工智能”（“Artificial Intelligence”，简称为“AI”），就是用计算机技术提供的技术手段，对人类智能进行模拟或部分模拟的一门学科。而人类的智能活动的一个非常重要的面相，用非常通俗的话说，就是“会说话”，也就是某种根据在特定语言共同体里通行的词汇表与语法，进行灵活的思想交流的能力。这种意义上的能力无疑是高等智慧生物之“智慧性”的重要指标，而且在我们已知的范围内，人类的确是唯一具备这种能力的物种（当然，这并不是说诸如鹦鹉或者黑猩猩之类的生物无法掌握人类词汇中的一部分，也并不是说它们没有特定物种内部有效的信息交流模式——然而，的确尚且没有足够扎实的证据表明：它们能够像人类那样，通过不同的句法组合方式，创生与理解大量在内容上与其直接生存环境无关的语言表达式）。换言之，如果“会说话”乃是将人类智能与动物智能区分开来的最重要的指标之一的話，那么，完整意义上的AI显然也应当具备这种特征。由此我们就能立即得出这样的结论：对于AI研究来说，让计算机能够“说人话”，就将具有如下的重要理论意义：这样的一项工作，将帮助我们从“人造认知架构”的角度理解语言能力在一个智能体系中所占据的地位，并由此夯实AI研究与广义上的认知科学之间的联系。

而换个角度看，让AI“说人话”的科学与工程学努力，也会带来丰厚的实践红利。很显然，如果经过特定编程的计算机也能够理解人类的语言的话，这些机器就能直接参与人类的信息交流活动，并由此成为人类工作与生活中的好帮手。概而言之，能够“懂人话”的人工智能机器所能胜任的工作将包括（但不局限于）：

（甲）电子邮箱处理。譬如，在面对海量的电子邮件的时候，你会希望你的AI助手能够通过对于邮件内容的分析，鉴别出哪些邮件是需

要迅速处理的，哪些则可暂缓处理，由此提高你的办公效率；

（乙）自动生成阅读摘要。譬如，如果有一篇文章实在太长，让你觉得无法迅速通读，你便会希望你的AI助手能够迅速形成一份内容提要，以便让你迅速把握文章之要点；

（丙）自动翻译。譬如，你要让计算机将一段电子邮件的内容自动翻译成一种你不会的语言——如韩语或者日语——或者将一段你不懂的外语表述翻译成汉语；

（丁）文本自动生成。譬如，你需要为公司的某次年会准备一份发言稿，却除了一些关键词或词组（如“业绩”“维持增长的势头”“优化研发队伍”）之外，什么句子都想不出。这时候，你便会希望你的AI助手能够根据这些提示，给出一些不同方向上的文本生成方案，以供你参考。而当你选定某个方案之后，你的AI助手甚至可以沿着这一路径继续优化相关的文本方案，最后帮你“多快好省”地完成发言稿。

在AI学界，负责完成上述任务的计算机编程研究，都会被打上“自然语言处理”（Natural Language Processing，简称为“NLP”）的标签。顾名思义，“自然语言处理”（NLP）的任务，就是用计算机进行编程，以便让相关程序能够“理解”人类的自然语言（不过，这里的“理解”二字必须要打上引号，因为对于计算机是否可能最终“理解”人类语言，尚且存在着非常复杂的哲学争议）。NLP研究因为涉及的话题非常多，复杂性、综合性特别强，所以一向被视为AI研究的皇冠（参看图1对于NLP所涉及的知识模块的概括）。^[1]

不过，说到这里，爱较真的读者或许会问：“会说话”能算是“具有智能”的充分必要条件吗？

笔者倾向于认为答案是肯定的。换言之，如果某观察对象能够具有我们认可的语言水平（即达到了“会说话”的标准），你就能推出它有智能；反过来说，如果它是有智能的，你就能反过来断定它有比较高的语言水平。举例来说，假设某星系的外星人突然造访地球，还流利地运用英语、汉语、日语等三种语言与我们

进行了长达2小时的富有成效的交谈——在这样的情况下，我们地球是没有任何理由认为它们是缺乏智能的。反过来说，如果外星人的确造访了我们的星球，并在近地轨道悬停了它们的飞碟，但因为某种原因暂时没有与我们建立语言联络的话，那么我们也的确必须假设它们是有语言的——因为倘若没有语言在生产活动的分工协作中所起到的作用，如此复杂的飞碟恐怕是无法被制造出来的。

这里需要提醒读者注意的是，笔者虽然认为“会说话（无论说哪种语言）”是“具有智能”的充分必要条件，但这并不意味着“能说某种特定的人类语言（如汉语或英语）”乃是“具

有智能”的充分必要条件。这就好比说，你与某个异族的人交流的时候，恐怕是不能仅仅因为彼此之间语言不通而假设对方是缺乏智能的。由此外推，我们甚至可以说：即使某些机器人的AI架构所支撑的交流语言的可理解性已经落在了广大人类用户的理解范围之外，我们也不能仅仅以此为据，认定这些机器人缺乏智慧。或说得更学术化一点，是否能够通过“图灵测验”，^[2]并非判断某对象是否具有智能的充分必要条件（而至多只能算是充分条件）。由此我们也就不难推出，虽然我们的NLP研究将不得不具有“为说特定自然语言的人类用户服务”的最终指针，但是作为某种研究的中介

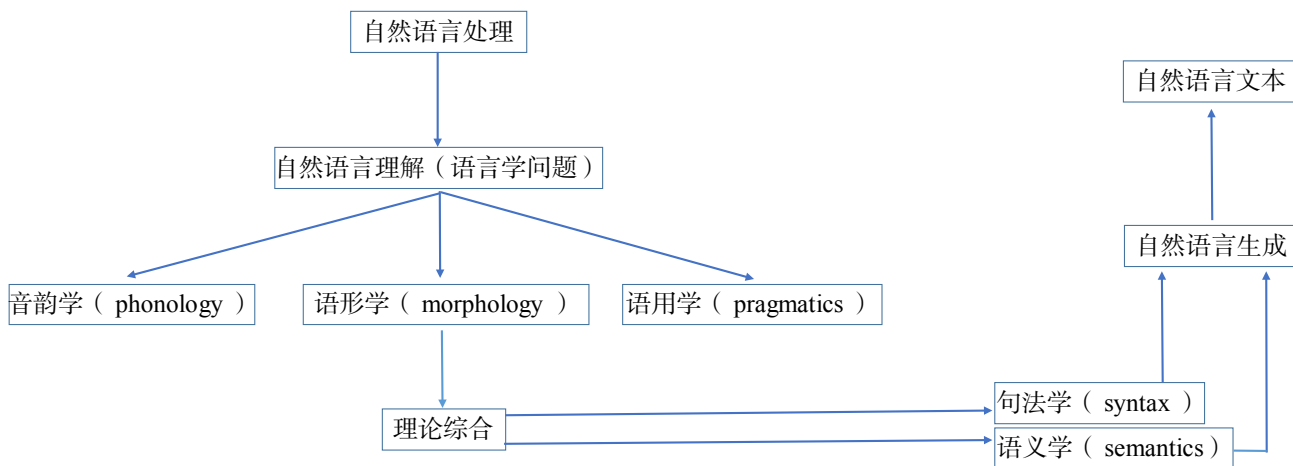


图1 自然语言处理所涉及的知识模块分布

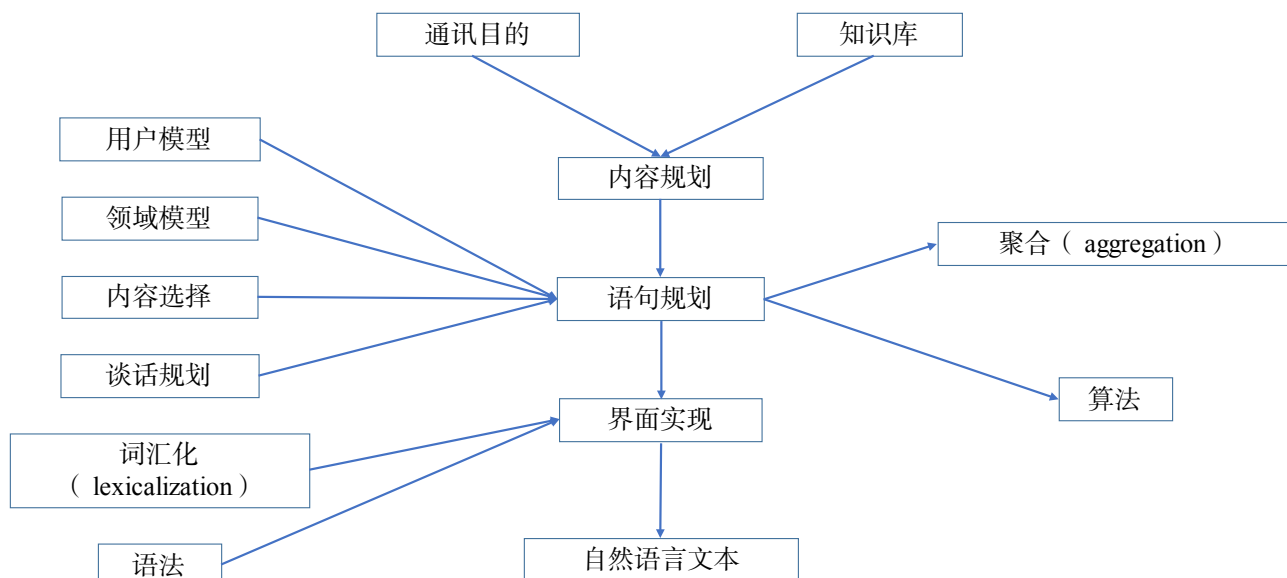


图2 自然语言处理机制的典型信息处理流程

（图1、图2绘制参考来源为 Khurana, D., Koli, A., Khatter, K., Singh, S. 'Natural Language Processing: State of The Art, Current Trends and Challenges'[CS/CL], <https://arxiv.org/abs/1708.05148>.）

阶段，我们所设计的系统所进行的语言表征，可能并不直接就是诸如英语、汉语这样的现成的人类语言。相反，NLP的研究者所要正面的，可是一个从表征的碎片演化为完整的人类符号系统的复杂过程。请参看图2对于目前主流NLP架构的信息处理阶段的概括。

对于本小节所给出的上述讨论，有的读者或许还会反驳说：笔者将语言处理能力视为智能之核心的观点，显得有点过于“逻各斯中心主义”了，并由此忽略了“具身化”(embodiment)在智能构成所起到的作用。换言之，在这些人看来，一个智能体之所以是智能的，首先并不在于其能说话，还是因为其能够通过其身体，而在物理空间中自由移动，感知光线、气味与温度，躲避危险，等等。“会说话”无疑是第二位的。

而在笔者看来，上面这种批评，并没有抓到我的立论的核心。换言之，说“会说话是具有智能的充分必要条件”，并不等于说要去否认：使得“会说话”这一条件本身被满足，还需要大量的前提条件。这就好比说，承认“具有相关行业内三年以上的工作经历，乃是获得某工作职位的最重要条件”，并不意味着要去否认“具有相关行业内三年以上的工作经历”这一条件自身的满足，还需要奠基在大量的前提性条件之上。相反，笔者完全愿意承认“具身性”自身的确构成了“会说话”的一个重要前提。譬如，《庄子·外篇·秋水》所说的“夏虫不可以语于冰者，笃于时也”一语，实际就已涉及到了说话者的身体感受力之局限对于其语言理解力的制约。不过，对于庄子所言的上述阐发，同时也反过来支持了笔者的观点：如果你发现某个对象在言语层面上无法“与之语冰”，这就很可能进一步说明：该对象在感知层面上就无法感受到冰天雪地的时节。这也就是说，语言交流的结果，依然能够有效地反映一个“疑似智能体”的智能架构在非语言层面上所接触的信息的广度与深度。与之相比较，对于某对象的纯粹的非言语身体行为的记录，却往往不能让观察者判断出对象的某些抽象能力的高低。譬如，一位哲学教授肯定无法通过一

位学生的纯肢体动作来判断他是否读懂了康德的《纯粹理性批判》，而只能通过笔试或口试等言语活动来完成这种此类判断。从这个角度看，从言语行为——而不是从身体行为——的角度出发来评判被观察对象的智能水平，是具有其特有的方法论优势的。由此外推，我们也不难得出：就人造智能体而言，其在NLP领域的表现水平，也应当对其整体智能水平具有指标意义。说“NLP研究乃是AI研究的王冠”，毫无夸张之处。

二、目前的人工智能发展水平，是否能够满足“说人话”的需求呢？

前文已经指出，NLP研究乃是AI研究的王冠，具有极大的理论综合性与市场应用价值。目前，也已经有大量的研究资源已经被投入到这个领域，产生了大量的商业产值。譬如，智能语音音箱、手机上装载的各种人机对话应用软件、“百度翻译”“谷歌翻译”，都是此类研究的重要成果。但是需要指出的是，此类产品的表面上的繁荣，并不意味着目下的NLP产品已经达到了“会说人话”的水准。其评判理由是：

第一，诸如“谷歌翻译”这样的机器翻译机制、“亚马逊理解器”(Amazon Comprehend)这样的文本信息挖掘机制、各种自动语音识别机制与各种各样的机器人聊天盒，都是针对不同的NLP任务而被设计出来的特定NLP机制，而不是某种面面俱到的针对所有NLP问题的一揽子解决方案。与之相较，对于一个完整的自然人而言，语义识别、语音识别、翻译等语言功能都是被集成到一个大脑上的，其各自运作背后均有一套统一的心理学与生理学规律予以统驭。从便利角度考虑，我们当然也会期望这种整合能够在NLP中实现。换言之，就像一个仅仅能做翻译，而无法用母语对被翻译文本进行深度解说的翻译者不能算作是已经真正理解了被翻译文字一样，某种仅仅能做浅层翻译，而不能进一步解释被翻译文字的NLP机制，也不算是真正在“说人话”。然而，基于下述两点理由，在NLP中，这样的技术整合非但没有

出现,而且似乎在可以预见的未来,也不太可能出现:(甲)具有不同分工的NLP机制往往分属于不同的公司,因此,知识产权方面的壁垒会导致彼此的融合困难;(乙)具有不同分工的NLP机制的研发往往本身又是基于不同的技术原理的,因此,原理方面的彼此不兼容也会导致彼此的融合困难。

第二,就人类而言,人类的语言能力本身是用来“做事情”的,比如帮助语言使用者在决策活动中进行复杂的信息梳理,或者是帮助语言使用者去说服某人采取某种行动。所以,语言能力天然就与逻辑推理能力、他心感知能力等其他心智能力相互交织。然而,就目前AI工业的学术分工情况而言,NLP的研究与常识推理、非演绎推理等技术模块之间的关系是相对分离的,遑论实现前文所提到的“具身性”条件对于言语理解力的约束机制(图3)。^[3]所以,从“通用人工智能研究”——而不是“专用人工智能研究”——的角度看,目前的NLP研究乃是“无根”的。

第三,传统的AI研究主要分“基于规则的AI”与“基于统计的AI”这两大路数,而随着时间的推移,目前以深度学习技术为代表的“基于统计的AI”正在AI研究的各个领域内大行其道,并在风头上全面压过了“基于规则的AI”(相关内容后文还要详谈)。但需要注意的是,深度学习技术在NLP领域内的运用,往往需要依赖网络提供大量的语料与学习样本,而这些语料与学习样本的最终提供者毕竟还是人类。因此,从哲学角度看,此类技术只能算是

对于人类智慧的“反光映照体”(这就好比月亮无非就是太阳的“反光映照体”一样),而无法真正成为语言智慧的根基。譬如,这样的技术很难在脱离互联网的支持的前提下,自主地创生出对于输入语料的恰当处理结果——而与之相较,具有正常语言智能的人类却能在不借助网络资源的情况下进行流畅的语言交流。因此,目前的主流NLP研究乃是缺乏足够强大的“本地化信息处理能力”的。

第四,也正是因目前的主流NLP技术与大数据的信息采录具有非常密切的关联,一些在原则上就很难通过大数据技术获得的语言材料,也就很难获得目下主流NLP技术的充分处理。这些材料包括:

(甲)人类语言中的双关、隐喻、反讽等修辞手段。具体而言,对于这些表达式的意义的确定往往需要结合当下的语境来进行,而无法通过基于大数据的归纳而被仓促决定。

(乙)缺乏足够网络数据样本的弱势语言,如少数民族语言与地方方言。具体而言,对于这些语料的传统NLP处理方式往往是基于“语料库”的建设的(此类建设需要人类研究者投入大量的精力遴选语料),而不能简单地诉诸于网络资料,因为目前的网络语言乃是由英语、汉语等主流语言所主导的。然而,也恰恰是因为目前基于深度学习的NLP技术对于网络语料的依赖性非常高,所以,此类NLP技术恐怕是很难支持主流语言与弱势语言之间的自动翻译处理的。

与之相较,具有适当语言智力的人类,却

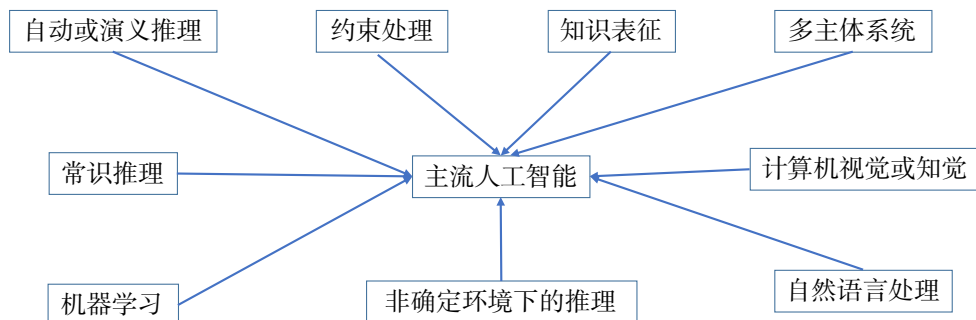


图3 目前主流AI学科内部的学术分工略图

(图3绘制参考来源为Hernández-Orallo, J. *The Measure of All Minds: Evaluating Natural and Artificial Intelligence*[M]. Cambridge: Cambridge University Press, 2017, 148.)

能比较快地通过语境信息提取而理解特定的双关语、隐喻与反讽的含义，或者通过一定时期的努力，仅仅通过少数几个教学者的帮助，学会一门方言。仅就这方面的表现而言，目前NLP的研究所达到的水平，还远远没有达到人类语言智力的平均水平。

笔者认为，目前NLP技术所面临的这些问题，不仅仅基于这样的或那样的工程学问题，而是有着深刻的哲学面相的。换言之，在基本的哲学层面上所遭遇的迷思，是目下NLP研究陷入的种种的工程学难题的总的病根。下面就是笔者对这一问题的几点管见。

三、为何自然语言处理研究需要哲学视角的介入？

从总体上来看，哲学与NLP研究之间的关系，与哲学和一般意义上的理工科研究规划之间的关系，并无本质不同。这也就是说，与很多自然科学研究规划一样，不同的NLP研究规划都已经预设了这样的或者那样的哲学立场，只是相关的实证科学研究者往往没有兴趣对这样的立场进行反思罢了。因此，哲学研究者的任务，就是将NLP研究未及言明的前提予以揭露，并对其进行反思性的评判。需要注意的是，与别的实证研究领域不同，NLP的研究尚且具有很大的前沿性、综合性、探索性与范式层面上的不确定性，这就意味着以综合性反思见长的哲学反思介入NLP讨论的机会，要比其介入物理学、化学等成熟学科的机会大很多。大致而言，由于NLP的研究将不得不预设这样的或者那样的关于语言之本性的看法，作为哲学分支的语言哲学（philosophy of language）与NLP研究之间的关系就会变得异常相关。其中，有如下四个问题是特别值得一提的：

问题一：语言与世界的关系为何？语言表征是对于说话者之外的外部世界的建模，还是对于说话者自己的内部观念世界的建模？

这个问题显然牵涉到语言哲学历史上的一个大争议。像柏拉图、弗雷格、克里普克、普特南这样的带有客观主义倾向的哲学家会倾

向于认为语言的作用是成为外部客观事物的标记符号；而像洛克、胡塞尔、大森庄藏这样的带有主观观念论色彩的哲学家则会认为语言的主要任务乃是表证言说者内部的思想观念，而不是指称外部的对象。此类争议在NLP内部也引发了相关技术路径的分野。其中，客观主义路向的语言观会自然导致诸如“沙德鲁”这样的利用一阶谓词逻辑构造“积木世界”的NLP方案；^[4]而主观主义路向的语言观则会引发丘其兰德的“神经语文学”规划，^[5]以及在“个性化营建”方面走得更远的王培的“纳思”研究规划。^[6]说得更具隐喻色彩一点，这两类研究路线之间的差异，乃是“上帝视角”与“凡人视角”之间的差异：基于“上帝视角”的客观主义的NLP研究路向会预设：程序员已经获得了其关于外部世界的至少某些方面的充分知识；而基于“凡人视角”的主观主义的NLP研究路向则会预设：程序员所知道的，仅仅是被构建的NLP体系内部的表征符号之间的推理关系——至于这些推理关系是否严格对应于外部世界中的诸要素之间的因果关系，则是“未知之事”。

笔者是“凡人视角”的NLP研究路向的支持者，否则，我们就不得不预设NLP系统所储存的某些关于外部世界的知识乃是“不可变的”（因为对于知识的充分性假设，将立即导出对于“知识修正”的必要性的否定），并由于这种预设而使得由此被设计出来的NLP系统失去应有的灵活性。但不幸的是，基于“凡人视角”的NLP研究，目前并非NLP研究的主流。因此，哲学家就特别需要在概念层面上进行相关的“纠偏”工作。

问题二：语言中的规则，究竟是先验的、不可变的，还是经验的，可变的？

前面已经提到，NLP研究素有“基于规则”与“基于统计”这两个分野。但从概念分析角度看，对于这两个分野自身的界定，似乎也就预设了“规则”的确立本身是与经验性的统计工作无关的。但事情果真是如此吗？难道一种语言的语法本身是不能随着时间而发生流变吗（我们不妨就想想近代以来汉语的语法所经历的

“欧化”进程)?关于如何更好的界定“先验”与“经验”之间的分界,大致有以下三种解答方案:

(甲)做大“先验”的范围,即将所有的经验层面上的自然语言语法都视为先验的。不过,这种研究方式由于实在难以配合经验语法在事实层面上的演化现实,而只能被视为某种抽象的可能性。

(乙)与(甲)所提示的方向相反,做大“经验”的范围,即认为所有的先验语法都可以通过统计资料甲乙消化。这是目下主流的基于深度学习的NLP研究的思路。

(丙)与前二者都不同,此路数取其中间值,即在“规则”中又一分而为二:有些规则是“经验的”,如各种语言的表层语法;有些规则是先验的,如某种贯穿于各种表层语法的“深层语法”。乔姆斯基的基于“普遍语法”概念的语言学路数,^[7]以及受到该路数影响的NLP研究,采用的就是该思路。

笔者本人所赞成的立场,乃是路数(丙)的某种更偏向经验论方向的改良版。与乔姆斯基类似,笔者也认为存在着某种贯穿于各种经验的语言形式的先验思想架构,否则我们就很难解释为何任何一个智力正常的人都可能学会任何一门外语;但与乔姆斯基不同的是,笔者并不认为这样的一种先验思想架构必须体现为一种现成的深层语法或普遍语法——它应当只能在某种更抽象的意义上被理解为简单语言符号之间进行“接榫拼接”的各种先验可能性,并因此只能承载最少的语法性质(“语法性质”一词在此是指性、数、格等语法形态)。举个例子来说,印欧语系的语言经常出现的名词的性、数、格的变化、动词词尾的情态与时态变化,都不能在这些最基础的“接榫”形式中出现,而只能被视为这些“接榫”形式的某种后天的复合形式。与之相较,乔姆斯基本人的想法则是这样的:即使在汉语这样的屈折度几乎不可见的东方语言中,上述这些印欧语言的语法“曲折性”特点也是以缄默方式存在的,否则,他心目中的“深层语法”就无法达成其普遍性。或说得概括性更强一点,笔者与乔姆斯基之间的共同点就在于:我们都认为对于所有语言的构成的

终极说明却都可以诉诸于一套统一的语法范畴;而笔者与他的不同之处就在于:在他看来,既然有待说明的诸语言现象有繁、简之分,那么,用以说明它们的语法范畴就必须在“繁”的一头留足“冗余量”,并由此成为了一种预备了所有语法开关的“普遍语法”;而笔者却不赞同这一判断。笔者的反驳如下:倘若上述假设是对的,那么我们就可以由此推出:一个汉语言说者倘若是改去言说某种更复杂的语言(如日语)的话,这种转换应当是不会给他带来更大的心理学负担的,因为他的心智已经具备了言说这种更为复杂的语言的潜在语法开关。然而,这一推理显然难以契合于下述这一朴素的心理学事实:言说在语法上更为复杂的语言一般会让人感到更有心理负担(除非那种更复杂的语言恰好是母语)。与之相较,笔者的假设——简单的卯榫结构能够按照不同经验语言的需要,随时被搭建为特定的语法结构——则可以轻易地解释为何我们在言说语法更简单的语言时会感到更轻松:因为这种言说所需要的卯榫结构的重构工作负荷本来就比较小。此外,笔者的这一研究思路还会带来一个重大的红利:由于笔者所说的这一卯榫结构与逻辑句法结构之间的高度同源性,经由此路数进行的NLP研究,将有机会与AI研究的其他面相(特别是推理与常识表征研究)相互融合。顺便说一句,目前最切合笔者上述思路的NLP编程语言,其实就是前文提到的王培的“纳思”逻辑,因为这种逻辑既具备对各种推理形式与常识经验的表征能力,也可以通过对于自身结构的递归式构造,而去模拟特定经验语言的语法特征。^[8]

问题三:语言与心理架构的关系之间的关系究竟是什么?

前面已经提到,目前基于大数据的NLP研究,基本上乃是与各种各样的认知建模研究相互疏离的。换言之,这些NLP研究者所关心的是,乃是如何在某些特定类型的语料输入与语料输出之间建立起合适的映射关系,而并非是这样的语言现象是从怎样的心理认知架构之中涌现的。与之相较,对于语言与心理活动之间的关系的研究,却成为了战后很多哲学家的

学术聚焦点。譬如，在美国哲学家塞尔（John Searle）看来，诸如“提出一个希望”“表达一个欲望”“表述一个信念”这样的言语行为本身乃是建立在“希望”“欲望”与“相信”这样的“意向性活动”之上的，因此，作为心理学哲学分支的“意向性理论”应当为作为语言哲学分支的“言语行为理论”提供根基。^[9]无独有偶，福多（Jerry Fodor）也在心理学哲学层面上提出过关于“心语”（mentalese）的假设，以便在一个前公共语言的层面上解释心智机器是如何加工处理那些基本信息的。^[10]而在笔者看来，虽然塞尔与福多各自的心理学哲学都有自己特定的问题，但至少他们都正确地看到了“纯粹地停留在言语行为的层面上来研究语言”这一做法的肤浅性，而走出了迈向正确的NLP解决路径的第一步。而之所以说“纯粹地停留在言语行为的层面上来研究语言”这一做法本身乃是肤浅的，则又是基于如下考虑：在言语行为层次上的现象实在是过于繁杂了，因此，对于不同语言现象的输入-输出关系的追索，必然会使得NLP的研究者陷入“以有涯追无涯”的尴尬境地，并由此带来昂贵的数据采集成本与建模成本；而如果我们换一个思路，将复杂的言语行为视为“某种更具有一般性的心智架构在不同外部环境的刺激下而产生的不同的对应输出”的话，我们就能大大降低我们的建模成本，并为相关系统在特定外部条件下的自动升级预留逻辑空间。

但这样的一种研究思路，必然会将主流的NLP研究的进路，进一步升级为一个宏大的通用人工智能的研究规划，因为心智建模本身就意味着对于智能的一般架构的探索。这种带有整体论思维模式色彩的研究路线图恐怕会让一部分研究者感到绝望，因为AI研究的典型操作模式便是针对某个特定应用场景提出的问题进行工程学开发，并将相关的研究成果拓展到别的应用场景上去——而笔者所提倡的研究思路却是先去悬置一切技术应用场景，而在哲学与科学的层面上理清智能推理的一般特征，然后再考虑技术运用的问题。不过，在笔者看来，这里我所提出的研究路线图虽貌似在绕弯路，却实际上更有希望，因为该路线图的执行者能够在最大程度上避免受到特定应用场景的偶然性的影响，而能聚焦于心智架构的某些一般性特征。这就好比是对于牛顿力学体系的纯粹理论研究与其基于该力学体系的各种工程学应用之间的关系：前一类研究虽然具有某种凌驾于各种应用场景的纯理论性，但一旦完成，就可以转变为无穷无尽的应用可能，而起到“四两拨千斤”的作用。

不过，这种面向“通用人工智能”的、带有整体论色彩的研究规划，显然会因为自身的整体论色彩而从心理建模层面自然延展到身体建模层面。这也就自然牵涉到前面我们所已经提到过的那个问题：语言表征与具身性之间的关系究竟是什么？

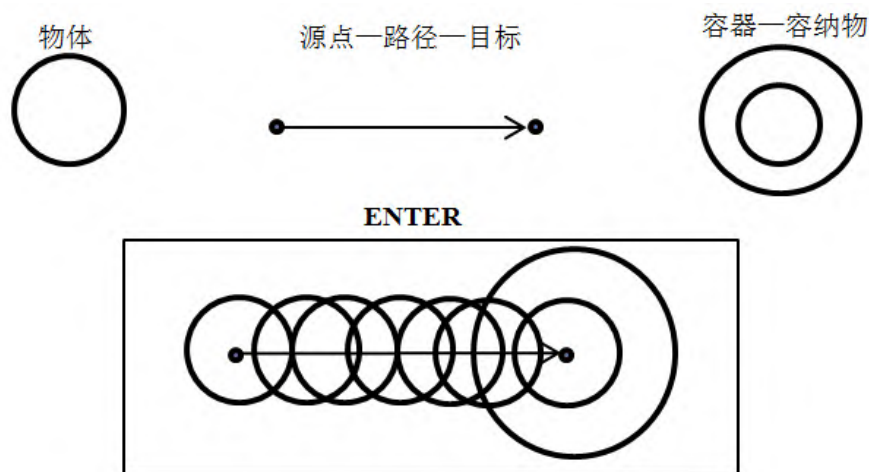


图4 关于“ENTER”的认知图式形成过程的图示

图4 绘制参考来源为 Langacker, R. *Cognitive Grammar: A Basic Introduction* [M]. Oxford: Oxford University Press, 2008, 33.

问题四:自然语言处理所需要的认知架构理论,究竟在多大程度上还需要被“具身化”?

在前文中,通过“夏虫不可语冰”这一案例,笔者已经提出了这样一种观点:语言交流足以让我们判断一个交流对象在身体感知方面的广度与深度,因此,语言交流乃是判断某对象的各方面智能水平的最有效手段。然而,从工程建模的角度看,这并不意味着对于智能体的物理身体的塑造就可以被还原为纯粹NLP性质的问题——这就好比说,在认识论的层面上说什么“美食家的评论乃是判断某餐厅招牌菜之品质的最重要指标”,并不意味着在本体论意义上我们业要承认:如何做出美食料理的问题,就可以被还原为如何撰写美食评论的问题。由此看来,完整意义的通用人工智能研究,将不得不包含对于智能体的感受-运动设备(即人类意义上的“身体”)的设计与制造。

不过,至少从表面上看来,对于AI的感受-运动设备的设计与制造,本身并不会引发任何哲学争议,因为就连最简单的家用计算机都包含着键盘、鼠标等与外部信息环境沟通的媒介,遑论是需要复杂物理环境中行动的AI系统。那么,我们将这个问题予以单列,其意义究竟又为何呢?

其意义就在于对于下述问题的澄清:上述这种“具身化”的工作,究竟对于NLP的研究来说是具有本质性的,还是仅仅具有某种边缘性的意义?说得更清楚一点,在NLP的研究中,架构者是否预先需要思考相关的AI体将被匹配上怎样的感受-运动设备,并为这样的设备而在NLP的界面上预留一些重要的“槽口”?抑或:架构者根本不用关心相关的AI体将被匹配上怎样的感受-运动设备,并完全可以将此类的考虑全部分配给别的领域内的专家?而这个“二选一”问题在近代哲学中的表现形式就是:人类的理性能力,是否能够在悬置各种感官能力的运作的情况下,进行相对独立的运作?对这一问题答“否”的乃是经验派的观点(这种观点的工程学对应者,自然就会强调NLP界面设计与AI体的外部设备设计之间的连续性),而对这一问题答“是”的,则是唯理派的观点(这种

观点的工程学对应者,自然就会强调NLP界面设计与AI体的外部设备设计之间的可分离性)。

笔者对于该问题的解答,则既不是纯粹唯理论的,也不是纯粹经验论的,而是带有康德式的调和意味的:在笔者看来,在纯粹的概念构造与底层的感官信息之间,还有一个重要的中间层被唯理派与经验派所忽略了,这也就是时-空关系的直观形式。一方面,这样的直观形式显然具有一定的前概念性(譬如,对于一个房间的空间感知,不能被还原为对于相关空间的几何学描述),而在另一方面,这样的直观形式又具有针对各种感官道的某种抽象性,并因此更接近于概念(譬如,一位盲人所感知到的教室的内部空间形式,依然会与一个正常人所看到的教室的内部空间形式有着高度的可重叠性)。在现代的认知语言学中,这样的直观形式的不同组合方式,一般称为“图型”(顺便说一句,该术语乃是认知语言学对于康德的“图型”概念进行再包装后的产物),譬如,英语“ENTER”(进入)这个概念就具有如图4像形式。^[11]

由上图看来,英语“ENTER”(进入)这个概念就可以被分析为数个意象图式在时间序列中的组合,包括“物体”(object)“源点-路径-目标”(source-path-goal)与“容器-容纳物”(container-content)。很显然,无论我们所讨论的智能体具有怎样的传感器与运动设备(譬如,无论它是像蝙蝠那样通过回声定位系统来辨别方位,还是像鸽子那样通过磁力线来辨别方位),它们都具有上述关于“ENTER”(进入)的认知图式。换言之,即使他们彼此之间的感官道不同,它们也都能够在NLP的层面上理解“ENTER”(进入)这个概念。

由此我们不难推出:对于NLP的研究来说,我们需要做的是:(甲)列出一系列类似“ENTER”(进入)的与时空感密切相关的概念;(乙)对这些概念进行“图式化”;(丙)对这些图式化的结果进行算法化处理。平心而论,笔者认为在这三个步骤中,最难处理的是(丙),因为目前世界上尚且没有出现比较成熟的针对认知语言学的“图式”概念的算法化方案^[12](某些奠定的处理方案是基于神经网络模型对图

式加以刻画的，但是笔者对相关路径的可靠性有所怀疑。因为篇幅限制，本文将不对相关理由进行展开）。但笔者坚持认为，由“图型论”所代表的康德式调和路线，乃是解决NLP系统之“具身化”问题的必经之路，否则，我们便既无法摆脱极端的唯理论思路所带来的困扰（此困扰即：抽象的符号如何在物理世界中获得意义奠基？），也无法摆脱极端经验论所带来的困扰（此困扰即：具有不同感官道的智能体之间的交流，是如何可能性的？）。换言之，沿着这一中间道路继续进行优化，乃是研究NLP体系之具身化的恰当限度的题中应有之义，因为从哲学角度上看，走别的路径，我们或许就根本没有成功的可能。

总 结

从本“导论”完成的讨论来看，NLP问题的研究的确对整个AI的研究来说具有指标性的意义。但对于该问题的哲学面相的了解，却一直没有被NLP学界所充分地意识到。毋宁说，目前NLP学界研究的话题取向是完全被偶然的工程学需求或商业需求所牵导的，而缺乏哲学（甚至是科学）层面上的整体谋划。更有甚者，在笔者所了解的范围内，语言哲学界目前也缺乏全面介入NLP研究的充分理论冲动。这种“两张皮互不相扰”的状态显然是不能让人满意的。此外，同样令人感到担忧的是，随着国际科技竞争与交流环境的改变，国内很多与AI相关的投资方向都被集中到了精密芯片的制造行业，与之同时，投向貌似更“虚”的AI架构研究的注意力却明显不足。殊不知工程师们对于高性能芯片算力的无休止的索求，在哲学层面上就已经预设了优秀的NLP机制与其它AI机制的运作乃是基于大数据的（因为只有海量的数据才会倒逼人们去寻找巨大的算力与之匹配）。然而，在前文的分析中我们已经看到了，这一预设本身可能就是错误的，因为语言机制运作之本质，便是通过对于少量核心规则与核

心词汇的掌握而具备创生出海量的表达式的潜能——而并非是通过对于海量的现成的表达式的构建方式的模仿，建立一个又一个“特设”（ad hoc）的语言模型，最终陷入“以有涯追无涯”的困境。从这个角度看，目前全球的NLP研究也好，整个AI工业也罢，都尚且处在“盲人摸象”的阶段，而尚且远远没有资格戴上“成熟科学”的王冠。

[参考文献]

- [1] Khurana, D., Koli, A., Khatter, K., Singh, S. 'Natural Language Processing: State of The Art, Current Trends and Challenges' [EB/OL]. <https://arxiv.org/abs/1708.05148>. 2017-08-17.
- [2] Turing, A. 'Computing Machinery and Intelligence' [J]. *Mind*, 1950, 59(236): 433-460.
- [3] Hernández-Orallo, J. *The Measure of All Minds: Evaluating Natural and Artificial Intelligence* [M]. Cambridge: Cambridge University Press, 2017.
- [4] Winograd, T. *Understanding Natural Language* [M]. Waltham: Academic Press, 1972.
- [5] Churchland, P. 'Neurosemantics: On the Mapping of Minds and the Portrayal of Worlds' [A], White, K. (Ed.) *The Emergence of Mind* [C], Fondazione Carlo Elba: Milan, 2001, 117-147.
- [6] Pei, W. *Rigid Flexibility: The Logic of Intelligence* [M]. Dordrecht: Springer, 2006.
- [7] Chomsky, N. *The Minimalist Program* [M]. Cambridge: The MIT Press, 1995.
- [8] 徐英瑾. 如何让电脑真正懂汉语?——一种以许慎的“六书”理论为母型的汉语处理模型 [J]. *逻辑学研究*, 2012, (2): 1-49.
- [9] Searle, J. *Intentionality* [M]. Cambridge: Cambridge University Press, 1983.
- [10] Fodor, J. *The Language of Thought* [M]. New York: Thomas Y. Crowell, 1975.
- [11] Langacker, R. *Cognitive Grammar: A Basic Introduction* [M]. Oxford: Oxford University Press, 2008, 33.
- [12] Manzanares, J. 'Cognitive Linguistics and Computational Modeling' [J]. *Textus*, 2010, 23(3): 763-794.

[责任编辑 王巍 谭笑]