怎样利用语言知识资源进行语义理解和常识推理

袁毓林[1]　卢达威[2]
（1：北京大学 中文系/中国语言学研究中心/计算语言学教育部重点实验室，北京 100871
2：中国人民大学文学院，北京 100872）

摘要：本文讨论怎样利用语言知识资源来帮助机器进行语义理解和常识推理。首先，指出人类生活在常识和意义世界中，人工智能机器人必须理解自然语言的意义、能够进行常识推理。接着，简单梳理了基于知识和基于统计两种自然语言处理路线各自的优长和短缺。然后，说明完全绕开知识的统计方法和深度学习，都不能真正理解概念和语言。文章通过具体案例说明，《**实词信息词典**》已经配备了有关词项的语义角色关系及其句法配置信息；把这种语言知识加入知识图谱和内容计算中，可以为人工智能提供理解和解释。由于“物性角色”描述了名词所指事物的百科知识，可用以回答相关事物是什么（形式角色）、有哪些部件（构成角色）、用什么做的（材料）、怎么形成的（施成）、有什么用途（功用）等常识性问题。

关键字：语言知识资源　语义理解　常识推理　基于知识/统计　语义角色　物性角色

How to Use Semantic Knowledge Resources for Semantic Understanding and Commonsense Reasoning

YUAN Yulin[1]　LU Dawei[2]
(1. Dept. of Chinese Lang. & Lit., Peking University / Research Center of Chinese Linguistics / Key Laboratory of Computational Linguistics, Ministry of Education, Beijing, 100871, China;
2. School of Liberal Arts, Renmin University of China, Beijing, 100872, China)

Abstract: This paper discusses how to use semantic resources to assist computer in semantic understanding and commonsense reasoning. Firstly, we point out that human beings live in a world with common sense and meaning, and that artificial intelligence robots are required to understand the meaning of natural language to make commonsense reasoning. Then, we briefly summarize the advantages and disadvantages of two approaches of natural language processing based on knowledge and statistics. Then, we explain that neither concepts nor language can be truly understood with statistical methods and Deep Learning can hardly account for any knowledge. The paper shows with specific cases that *Information Dictionary of Notional Word* has been equipped with semantic role information and syntactic configuration of the words, which can be employed in the knowledge graph and the content computing and served for the improvement of the artificial intelligence. As the "Qualia Role" describes the encyclopedic knowledge of nouns, it can be used to answer commonsense questions such as what it is (formal role), what it consists of (constitute role), what it is made of (material role), how it is created (agentive role), and what it is used for (telic role).

Key words: semantic knowledge resources; semantic understanding; commonsense reasoning; knowledge based / statistics based; semantic role; qualia role

1. 人工智能呼唤语义理解和常识推理

我们从小到大一直生活在一个由常识（commonsense）构筑的世界中：脚下是大地、头顶为青天，早晨日出东方、傍晚夕阳西下、夜空星辰闪耀，春来草树斗芳菲、秋去叶落千山

枯……。在日常生活中，我们习惯于在常识框架内思考和谈论事物的形状、材质、构成、功用、来源等等。比如，我们认为水是一种无色、无味、透明的液体，可以用来喝和解渴，当然还可以用来降温、灭火、灌灌、洗涤、游泳、漂浮船只等等；猫是一种小型的、驯化的哺乳动物，身上有柔软的皮毛，长着锋利的爪子、尖尖的耳朵，拖着一条毛茸茸的长尾巴，会捉老鼠，被人当作宠物饲养；铁路是一种由钢轨等材料铺成的、在上面行走火车的道路，可以用来运输人员和物质等等；词语和句子是从人的口中发出的、有意义指称的声音，可以用来分类命名、描述事物、发表意见、人际交流等等；政府是一种权力机构，用以推行法律、执行管治、组织防御、控制暴力、保障人们权利、提供公共服务、满足人民需求等等。[1-2]

根据 Daniel[3]的见解，人类长着一颗贪婪的大脑，具有一个明确的特性是对事实永不满足的追求。……通过发现大自然的隐蔽规则，通过将两种完全不同的思想根据它们潜在的、共同的信息结构联系起来，我们的大脑创造了一个广阔的意义世界。这种不懈努力的结果之一就是：当我们看到一张椅子时，看到的不只是椅子基本的外部特征。当然，我们会认出这是一张椅子，然后马上会想到与这个物体相关的一系列意义：椅子有什么形状，具有何种功用，跟其他家具的关系如何，放在哪幢大楼哪个房间内，等等。事实上，当我们观看周围世界时，无意识可能忙着处理一些基本的感觉特性，但是在意识的大本营内，每一项内容都要经过我们掌握的知识结构的严密筛选。我们看到的任何物体，都会触发理解的意识波，即该物体不同层次的意义。[1]

可见，常识和意义如影随形般地跟我们人类生活纠缠在一起。如果人工智能要更好地服务于人类、更多地介入人类的日常生活，那么人工智能机器人就必须理解人类自然语言的意义、掌握常识并且据此进行推理。但是，不管是关于世界的常识还是关于事物的意义，它们都是十分模糊、难以定义的。于是，怎样教人工智能机器人理解语义和掌握常识，就提到人工智能进一步发展的议事日程上来了。据《纽约时报》报道，微软联合创始人保罗·艾伦（Paul Allen）正在为他的非营利性计算机实验室——艾伦人工智能研究所（AI2）投资 1.25 亿美元，并计划未来 3 年投资预算翻倍。这笔资金将用于现有项目，以及“亚历山大项目”——一项聚焦于教授机器人“常识概念”的新计划。艾伦在新闻发布会上指出：在人工智能研究早期阶段，人们对常识概念有很多关注，但是这项工作仍停滞不前。人工智能机器人仍缺少多数 10 岁儿童所具有普通常识概念，我们希望启动这项研究，并在该领域获取重大突破。如果机器人非常先进，那么它们可以模拟人类完成任务；比如，定位和识别物体、攀爬、出售房屋、提供灾难援助等。然而，即使是这些先进的机器人，现在也无法处理简单的问题和指令，无法应对一个不寻常的处境，无法使用“普通常识”去校正行为和反应。AI2 研究所执行总裁奥伦·埃齐奥尼（Oren Etzioni）说：“目前没有一个人工智能系统准确地回答一系列简单问题。例如：如果我将袜子放在抽屉里，明天它还会在那里吗？或者：你怎么知道一个奶瓶是否满了？”他还强调称，2016 年当 AlphaGo 人工智能程序打败世界排名第一的围棋棋手时，AlphaGo 却并不知道围棋是一种棋盘游戏。[1]

闻到了备战的气息，加上自己的研究和从业经验，上海阡寻信息科技公司董事长白硕博士直言：自然语言处理从浅层到深层面临范式转换，还处在对接情感计算与常识计算的战略性要地的关键位置。谁能拔得头筹，谁就能在当下的人工智能“军备竞赛”中处于有利地位。……如果说自然语言处理是人工智能的王冠，那么语义表示和理解技术就是王冠上的明珠。目前人工智能领域的发展态势，在语义这一块已经到了重兵集结的程度。[4]

显然，大家已经充分地认识到：人工智能的下一步发展和实用化，必须突破语义理解和常识推理这一瓶颈。我们认为语义理解和常识推理研究的进展，依赖于全新的自然语言处理技术和理念。为了找到这种技术和理念，下面我们先梳理和检讨一下既有的各种自然语言处理技术，在此基础上尝试提出我们的技术路线和方法论观念。

[1] 详见 http://tech.sina.com.cn/d/i/2018-03-13/doc-ifyscsmu9166662.shtml。

2. 自然语言处理的两种路线：基于知识 vs.基于统计

自然语言处理（natural language processing, NLP）的智能技术是当前人工智能热潮的一个支脉，应该放在当前整个人工智能技术路线和方法论取向的大背景上来看待和理解。

经典的人工智能基本的技术路线是基于知识：首先调查人类解决问题的途径和技巧，然后尝试用可执行的方式对这些途径和技巧进行编码。由于人类理解和生成语言依赖词汇、句法、语义等语言知识和相关的关于外部世界的百科知识，因而学者们就为计算机理解自然语言建造了各种知识库：比如，词汇知识库（如 WordNet）、句法标注库（如 TreeBank）、语义关系知识库与标注库（如 VerbNet, PropBank, FrameNet）、常识知识库（如 Cyc, ConceptNet, DBpedia: Wikipedia 的数据库化）、常识与词汇结合的知识库（如 YAGO: WordNet 和 DBpedia 的结合，IBM 公司的 Watson 系统以此作为知识库，参加知识竞赛节目 Jeopardy，战胜了人类冠军）、关于概念分类体系的本体知识库（如 SUMO: Suggested Upper Merged Ontology）、词汇-常识-本体相结合的知识库（如 YAGO-SUMO）……，不一而足。

这种技术路线的困难是：除了人工构造各种知识库代价昂贵之外，人们也不可能把各种相关知识都弄清楚、并且明确完整地表示出来和巧妙灵活地组织起来。虽然人类是用其全部的经验与知识来理解和生成语言的，[5]但是我们无法把全部的世界知识编码进入计算机；更何况，常识往往还是模糊不清、难以定义的呢。因此，我们暂时还不能指望一个聊天机器人（chatbot）能够回答这种问题：“为什么小鸡仔不会下蛋？”

跟基于知识的方法相对的是基于统计的方法：从大量数据中学习概率分布。在自然语言处理上，最常用的统计方法是建立各种“词袋”（bag-of-words）模型：把每一个文档看作一个词频向量，把文本信息转化为易于建模的数字信息。比如，通过统计文本中所用的不同感情色彩的词语（褒义词、贬义词等）的数量，来判定用户对产品的情感评价（sentiment）。再复杂一点，通过给每一个词指派一个反映其在给定文档中的出现次数的指数（index number），从而把一个给定文档表示为一个向量（vector）。这样，如果一种语言的词汇规模是 5 万个词，那么表示文档的矢量就有 5 万个维度（dimensions）；其中，许多维度的指数是 0，因为相应的词没有在这个给定文档中出现。于是，可以利用一个词在全部文档中的稀疏性（sparsity）来为每一个词设定权重。比如，信息检索上使用的词项频率-倒排文档频率（term frequency-inverse document frequency, TF-IDF）方法就是一种为每一个单词分配权重的算法，该算法在分配权值时不仅考虑文档中的词频，而且考虑了倒排文档频率。用这种方法可以快速地计算出不同文档的相似度。稍微复杂一点的是潜在语义索引（Latent Semantic Indexing, LSI）模型，它通过海量文本找出词汇之间的关系：当两个词或一组词大量出现在同一个文档中时，就认为这些词是语义相关的。又如，潜在狄利克雷分布（Latent Dirichlet Allocation, LDA）文档主题生成模型。这是一种由词、主题和文档三层结构组成的三层贝叶斯概率模型。其朴素的假设是：一篇文章的每个词都是以一定的概率选择了某个主题，并从这个主题中以一定概率选择某个词语。据此，可以把每一个文档表示为一些主题所构成的一个概率分布，而每一个主题又可以表示为很多单词所构成的一个概率分布。它可以识别大规模文档集或语料库中潜藏的主题信息，从而发现特定文档的文本内容所属的主题类型。可见，“词袋”方法不考虑词与词之间的顺序等结构信息，简化了问题的复杂性；但是，不能发现“狗咬人”与“人咬狗”这两个文本之间的意义差别。

跟基于统计的向量化方法不同的是词嵌入（word embedding）方法。这种模型以向量形式给每一个词指派一长串数字，从而把每一个词表示为一个低维实数向量。通过词向量的距离来计算不同的词之间的语义距离。比如，“run”和“jog”的词向量的距离比较接近，它们跟“Chicago” 的词向量的距离比较遥远。每一个词的词向量有相同的维度，通常是

300 维左右。为了学到词向量，Skip-gram 算法首先给每一个词向量赋予一个随机值，然后在所有的文档中，不断地循环，推动词-1 和跟它分布（搭配环境）相近的词-2 在词向量上接近，同时推动词-1 和跟它分布不同的其他词在词向量上相差较大。还可以用循环神经网络（recurrent neural network, RNN）把句子编码成向量，并且用另外的 RNN 来反向地把它解码为不同的句子。这种串对串的编码器-解码器（encoder-decoder）模型，可以在双语（源语言-目标语言）对齐语料库上进行训练，从而形成基于神经网络的机器翻译模型，就像谷歌翻译（Google Translate）那样。这种基于多层次神经网络的研究路线，近年来被称为深度学习（deep learning）。但是，人类语言可以对无限的概念组合进行编码，形成无限多的话语。而双语对齐语料库之类的训练集总是有限的。更何况，人类的语言理解是植根于对外部世界的感觉和跟外部世界的互动行为的。比如，“鸡仔”对于人来说，不仅意味着它是一种鸟类，有各种鸟类的行为；而且还意味着我们可以对它做的一切事情，还有它在我们的文化中所代表的一切东西。[4][6]显然，这些属于人类的常识范畴的知识，都是不容易通过训练来让机器掌握的。

3. 深度学习能否带领自然语言处理突出重围？

当前的人工智能研究和开发，主要采用基于大数据的统计方法和基于多层神经网络的深度学习技术；在语音识别和合成、机器翻译、图像（人脸）识别等领域取得了一定的成功，但是在抽象概念及其关系、语义理解和常识推理等内容领域尚未取得太大的成功。有人断言深度学习方法对于概念、语义等内容领域的处理难以有成就。甚至有人对人工智能目前的研究方向表示怀疑和否定。比如，1956 年在著名的“达特茅斯会议”（Dartmouth Conference）上提出“人工智能”（artificial intelligence）概念的美国麻省理工学院教授马文·明斯基（Marvin Minsky, 1927—2016）。他虽然一直认为人类的思维可以用机器模拟，并且有一句广为流转的话：“大脑无非就是肉做的机器而已”（the brain happens to be a meat machine）。但是，明斯基曾参加过智囊机构 TTI/Vanguard 赞助的一些会议，TTI/Vanguard 的主管史蒂文·彻丽（Steven Cherry）说：

他发现最近几年的一些发展方向出现了偏差，谷歌和脸谱网正在利用深度学习技术开发它们的庞大数据集。明斯基认为，这只是短期的成果，其代价是真正的机器智能问题得不到解决。[7]

就自然语言处理而言，情况也是这样：许多人暂时放弃基于规则和知识等可靠解决方案（solid solution），而是尝试采用统计学习方法等讨巧的快速解决方案（smart solution）。原因是目前的理论语言学研究还不能为自然语言处理等应用语言学提供足够的支撑。正如德国爱尔兰根—纽伦堡大学的计算语言学教授罗兰德·豪塞尔（Roland Hausser）所说的：

实用语言学的例子有语音识别、桌面出版、文字处理、机器翻译、内容提取、文本分类、互联网查询、自动辅导、对话系统和其他所有的自然语言的应用。这些实际应用催生了对实用语言学方法的巨大需求。

但是，现有的实用语言学方法还远远不能满足用户的需求和期待。到今天为止，最成功的实用语言学方法是基于统计学和元数据标注的方法。这些是快速解决的方法（smart solution），不需要关于自然语言交流过程的一般性理论支持，其目的是最大限度地挖掘每一次应用或者每一类应用的特殊性及其本质上的局限性。[8]

粗略地浏览相关文献和媒体报道，我们就可以看到这样一幅纠结的学术画面：一方面，深度学习是驱动最新一波人工智能热潮的关键技术。由于深度 CNN 在图像和语音任务中展现出的卓越性能，催生了大量实验性、开发性的工作，人们希望将其应用到许多其他的问题和工程产品当中。另一方面，人们发现虽然可以用深度网络来解决一些问题，但

这都是在过度的试错和参数调整之后才实现的。更何况深度学习的理论基础不清楚，还不能解释深度网络如何有用以及为什么有用。也就是说，深度学习无论是作为一门基础科学、还是作为一门工程学科，都不够成熟。以至于纽约大学的心理学和神经科学教授、几何智能公司创始人 Gary Marcus 在 arXiv 上传了一篇论文，列举深度学习十大局限，说深度学习其实并没有解决什么问题。[9] 下面，我们挑跟自然语言处理有关的来引述：

（1）深度学习目前缺少通过明确的、言语定义学习抽象概念的机制，而且机器却必须经过成千上万的训练才能发挥最好效果。

（2）深度学习并没有理解抽象的概念。DeepMind 用深度强化学习玩“打砖块”游戏，但系统并不知道什么是隧道、什么是墙，它所学会的，只是特定场景下的一个特定动作。深度学习目前没有足够的能力进行迁移。

（3）深度学习还不能自然地处理层级结构。当前大多数基于深度学习的语言模型，都将句子视为词的序列。在遇到陌生的句子结构时，循环神经网络（RNN）无法系统地展示句子的递归结构。深度学习习得的特征之间的关联是平面的，没有层级关系。

（4）深度学习目前还无法进行开放式推理。系统无法理解“John promised Mary to leave”和“John promised to leave Mary”之间的细微差别，机器也就无法推断出谁要离开谁，或者接下来会发生什么。

（6）深度学习还没有很好地与先验知识相结合，部分原因是深度学习系统中表示的知识主要涉及特征之间的（很大程度上是不透明的）相关性，而不是像量化的陈述那样的抽象（例如，“每个人都有死亡的一天”）。深度学习适合的问题更多与分类有关，而与常识推理相关的问题几乎都超出了深度学习的解决范围。

（8）深度学习假设世界是大体稳定的，但实际并非如此。深度学习在高度稳定的世界中表现很好，例如“围棋”这类有固定规则的棋盘游戏，但在政治和经济等不断变化的系统中，深度学习的表现并不好。

Marcus 的文章引发了不小的讨论，著名机器学习专家、AAAI 前主席 Thomas Dietterich 连发 10 条 Twitter，一一驳斥 Marcus 列出的“十大罪状”，并且对深度学习中的关键技术反向传播（back propagation）和权重绑定（weight-tying）进行了拓展和延伸，从而强调了一种新的编程范式——可微分编程（differentiable programming）。[10] 纽约大学终身教授、纽约大学数据科学中心的创始人、以及 Facebook 人工智能研究部门（FAIR）负责人 YannLeCun 在 Facebook 个人主页上写了一篇短文，不仅支持可微分编程，还说：好，深度学习作为一个流行词，现在时效已过（Deep Learning has outlived its usefulness as a buzz-phrase.）。深度学习已死，可微分编程万岁！（Deep Learning est mort. Vive Differentiable Programming!）。[11] 没错，“可微分编程”不过是把现代这套深度学习技术重新换了个叫法，这就跟“深度学习”是现代两层以上的神经网络变体的新名字一样。这位被人们称为卷积神经网络（convolutional neural network, CNN）之父的法国科学家认识到：人工智能发展的一大难题就是怎么样才能让机器掌握人类常识，这是让机器和人类自然互动的关键。想要做到这一点，它需要拥有一个内在模型，以具备预测的能力。LeCun 用一个公式简洁地概括了这种人工智能系统：预测+规划=推理。而研究人员现在要做的，就是不需依赖人类训练，让机器学会自己构建这个内在模型。关于机器视觉如何与常识相联系，LeCun 说，就连 Facebook 内部也有很大分歧。“一些人认为可以与智能系统只进行语言交流，但是语言是一个相当低带宽(low bandwidth)的渠道，信息密度很低。语言之所以能承载很多信息，是因为人们拥有大量的背景知识，也就是常识，来帮助他们理解这些信息。”LeCun 解释道。看来，他暂时也拿自然语言理解没辙。

总的来说，看起来对于人工智能和自然语言处理来说，相关领域知识和语言知识的挖掘、整理和表示还是不可缺少的，完全绕开知识的统计方法和机器学习，都难以真正理解概念和语言。

4. 语义资源帮助知识图谱赋能 AI 理解和解释

知识图谱（Knowledge Graph）用可视化技术呈现知识，把以往各种线性的、离散的、非结构化的知识，用图（graph）这种数据结构形式组织起来，从而描述关于世界万物的实体（entities）、概念（concepts）、事件（events）及其之间的关系。知识图谱实质上是一种语义网络（semantic network），其结点代表实体或概念，边代表实体/概念之间的各种语义关系。它通过对海量数据中各种个体/概念及其盘根错节的关系的梳理，使得原本模糊的信息世界（sybernetic world）、乃至现实世界（virtue world）变得更加脉络清晰。这种数据的组织和呈现形式，可以为当前人工智能实现进一步的突破提供基础。正如上文所引述的，当前这波人工智能热潮得益于以深度学习为代表的大数据处理方法。但是，深度学习之机理的不透明性、不可解释性已成为制约其发展的障碍。因此，“理解”与“解释”是人工智能需要攻克的下一个挑战，而知识图谱为“可解释的 AI”提供了全新的视角和机遇。[12]下面是两个通过把语义知识加入知识图谱，来为人工智能提供理解和解释的构想性案例。

清华大学李涓子教授在跟笔者学术交流时说：开发知识图谱，光是在连结两个实体节点的边上标定表示其关系的动词是不够的，最好还得有这两个结点相对于动词的语义角色。例如：对于“特朗普—辞退了—联邦调查局局长科米”来说，如果能够让机器“懂得”或“知道”：“特朗普”是辞退行为的发出者，“科米”是辞退行为的受影响者，就比较理想。[2]问题是，能不能利用语言知识资源，来生成或给出这种语义角色？查询了袁毓林教授主持研制的《**北京大学现代汉语实词句法语义功能信息词典**》（简称《**实词信息词典**》），我们发现这个语义知识资源基本上可以满足这种需要。下面是“解雇”这个词条的部分信息：

表 1 词条“解雇”的句法语义功能信息

词目：	解雇
汉语拼音：	jiěgù
词类属性：	体宾动词
词义解释：	停止雇用；解除雇佣关系。跟“辞退、开除”相近，跟“聘请”相对。
近义词：	辞退、开除
反义词：	聘请
语义角色：	施事 A: 停止雇用他人的人或机构； 受事 P: 被施事解雇的人。
句法格式：	S1: A + __ + P 如: 校长～了两名代课教师。丨董事会～了新来的市场部经理。丨那家企业～了不少闲人。丨 桂林丝厂有 2500 多名工人，从来没有～过一个人。 S2: A + 把 P + __了 如: 校长把那位代课老师～了。丨公司把闲人都～了。丨董事会把新来的市场部经理～了。 S3: P + 被（A）+ __了

[2] 在“语言资源——理论、方法与应用国际研讨会”（2017 年 11 月 5 日）上的个人交流，和同年 11 月 27 日双方团队在北京大学中文系就事件分析知识图谱与语义角色关系的正式讨论。

	如：涉事老师已经被～了。丨新来的公关经理被公司～了。丨那些闲人们都被公司～了。

可见，机器系统通过调用上述词典信息，如根据句法格式，可以分别把“特朗普”绑定到“施事 A”、“科米”绑定到“受事 P”这两个语义角色上；从而推定“特朗普”是“停止雇用他人的人”，“科米”是“被施事[=特朗普]解雇的人”。更何况，这个词典中除了进行多重释义之外，还给出了“解雇”的同义词（辞退、开除）和反义词（聘请）。利用这些意义关系和句法格式（论元角色的配置方式），再查询我们的另一个资源（《动词蕴涵型式库》）就可以进行语义（蕴涵）推理了。例如：

特朗普 解雇 科米

→ 特朗普 辞退/开除 科米

→ 特朗普 不再雇用/聘请 科米

→ 特朗普 把 科米 解雇/辞退/开除了

→ 科米 被 特朗普 解雇/辞退/开除了

无独有偶，白硕[4]指出：

的确，知识图谱就是当代最通用的**语义知识表示形式化框架**。它的节点就是语义学里面说的“符号根基”（symbol grounding），即语言符号与真实或想象空间中的对象的对接，在计算机中体现为语言符号与数字化对象的对接。它的边则是语义学里面说的“角色指派”（role assignment），在计算机中体现为每个数字化对象与其他数字化对象之间的语义关系标签。节点和边，这恰恰是知识图谱所支持的要件。

但是，事情并没有完结。语义结构表示框架中，现有的知识图谱可以完美描述**实体、关系、属性（状态）及其值**这三类要素。但是剩下的还有**事件、时间、空间、因果条件、逻辑模态等**，我们必须对现有的知识图谱结构进行改造，才能适应这些语义要素的表示。

先看**事件**。事件可以改变关系和属性。比如“撤销职务”的事件真正的语义效果是改变相应实体的“职务”属性的取值，其他一切操作，如果不落到这上面，都是耍流氓。此外，一个事件可以触发其他事件（例如“国王去世”触发“王储继位”），一串事件可以是一个大事件的细粒度展开（比如“立案侦查”、“调查取证”、“拘捕”可能是某个“案件”事件的细粒度展开）。这些具有动态特性的操作如何与静态知识图谱的结构和工具融为一体，是一个非常有挑战性的问题。我们注意到哈工大有关“事理图谱”的相关研究成果，但要成体系地解决事件的表示问题，目前成果还是很不够的。

检索《**实词信息词典**》，发现其中已经配备了有关词项的语义角色关系标签，还有这些语义角色的常见的句法配置。下面以“立——案件——调查——取证——拘捕”这几个词条为例：

表 2 词条“立”的句法语义功能信息

词目：	立 03
汉语拼音：	Lì
词类属性：	体宾动词
词义解释：	创立；创建；建立；制定。跟“建”相近。
近义词：	建
语义角色：	施事 A: 建立或制定某种事物的人； 结果 R: 施事所建立或制定的事物； 与事 D: 施事为他建立或制定结果的人； 量幅 EXT: 结果建立起来以后的时间长度。

句法格式：	S1: A+__+R 如: 我单独～了一个户头。| 韩秀又～了一个账户。 S2: A+给（D）+__+R 如: 您先给～个户头吧！| 学校给他们每个人都～了一个档案。| 这孩子，我得给他～～规矩。 S3: R+（A+）__+EXT 如: 这个户头（我已经）～了一年了。| 那个账户韩秀已经～了多年了。 S4: R+（A+）__+起来 如: 怎么这么久了规矩（你）还没～起来？！| 户头我们先～起来。 S5: A+把 R+__+起来 如: 你先把规矩～起来，大家好照着做。| 韩秀把那个账户～起来了。

表 3 词条“案件”的句法语义功能信息

词目：	案件
汉语拼音：	ànjiàn
词类属性：	名词
感情色彩：	消极
词义解释：	有关诉讼和违法的事件。
语义角色：	形式 FOR: 人造物、工具； 构成 CON: 通常由犯罪主体、犯罪的主观方面、犯罪的客观方面、犯罪客体等构成；根据种类的不同分为：刑事案件、民事案件、行政案件、经济案件，等等； 单位 UNI: 个体：个，等等；集合：批、类、种，等等； 评价 EVA: 重大、大、小、恶性、特大，等等； 施成 AGE: 制造、谋划、实施、策划出，等等； 功用 TEL: 违法、破坏社会、危害公众，等等； 行为 ACT: 震惊、发生、惊动，等等； 处置 HAN: 审理、处理、查证、核实、侦破、裁决，等等。
句法格式：	S1: CON +__ 如: 刑事～	民事～	行政～	民事～	死刑～	反革命～	犯罪～	误杀～	违纪～	妇女权益～	法律援助～ S2: Num + UNI + __ 如: 一个～	一类～	一种～	一系列～	一批～ S3: EVA +（的+）__ 如: 大～	小～	重大～	恶性～	重大～	特大～ S4: AGE + __ 如: 谋划～	实施～	制造～| 设计～| 策划出～ S5: __ +TEL 如: ～破坏社会	～危害公众 S6: __ +ACT 如: ～震惊	～发生	～惊动 S7: HAN + __

	如: 审理～ | 处理～ | 查证～ | 核实～ | 侦破～ | 裁决~

表 4 词条“调查”的句法语义功能信息

词目：	调查
汉语拼音：	diàochá
词类属性：	弱谓宾动词/名动词
词义解释：	为了了解情况而到现场进行实地考察。带有书面语色彩。跟“检查、考查”相近。
近义词：	检查、考查
语义角色：	施事 A: 为了了解情况而到现场进行实地考察的人； 与事 D: 施事向他调查的人； 系事 RE: 施事所调查的事情。
句法格式：	S1: A +（作/进行 +）__ 如: 大力曾经～过。| 警方～过。| 他曾经作过～。| 警方进行了～。 S2: A + __ + D/RE 如: 大力曾经～过当地的居民。| 警方～过附近的群众。| 公司总部～了这起事件的起因。| 警方～了起火的原因。| 大力曾经～过小王在三月五号那天和谁在一起。| 大力～小王在三月五号那天是不是和他女朋友在一起。 S3: A + 向 D + __ + RE 如: 小王向大力～了这件事件的起因。| 警方向附近居民～了事故的原因。 S4: D/RE + A + __ 如: 这些人小王都～过。| 附近的居民警方～过了。| 这件事老王～过了。| 事故的原因警方～了一下。 S5: A + 对 D/RE + 作/进行 +__ 如: 小王对这些人作了～。| 警方对事故的原因进行了～。 S6: A + 把 RE + __ + 清楚 如: 小王把这一事件～清楚了。| 警方已经把事故的原因～清楚了。

表 5 词条“取证”的句法语义功能信息

词目：	取证
汉语拼音：	qǔzhèng
词类属性：	不及物动词/名动词
词义解释：	寻找并取得证据。
语义角色：	施事 A: 寻找并取得证据的人和组织； 与事 D: 施事向其获取证据的人； 源点 SO: 证据所在的地方； 量幅 EXT: 所获取的证据的数量或次数； 范围 RA: 取得的证据所涉及的事件。
句法格式：	S1: A + (从 SO +) __ 如: 公安机关正在调查～。| 相关部门会认真～，秉公办案。| 控方律师从现

	场～。| 警方从各种渠道～。 S2: A + 对/向 D + (进行) + __ 如: 美国向黑手党要人～。| 专案组再次对钱某人进行调查～。 S3: A + __ + EXT 如: 检察机关调查上千人次，～200 余份。| 律师～九次，次次受到阻扰。 S4: RA + 由 A + __ 如: 这一案件由公安人员～。｜ 受贿案由检察机关～。

表 6 词条“拘捕”的句法语义功能信息

词目：	拘捕
汉语拼音：	jūbǔ
词类属性：	体宾动词
词义解释：	逮走；捉拿。跟“逮捕”相近，跟“释放”相对。
近义词：	逮捕
反义词：	释放
语义角色：	施事 A: 逮捕、捉拿他人的人； 受事 P: 被施事所拘捕的人。
句法格式：	S1: A + __ + P 如: 警方立即～了犯罪嫌疑人。| 第一小队在行动中～了 3 名男子,他们涉嫌贩卖毒品。 S2: A + 把 P + __ 如: 警方已经把这几个嫌疑人～了。| 公安机关把一名怀疑与此案有关的 55 岁男子～。 S3: P + 被（A）+ __ 如: 许多爱国青年学生被～。｜ 所有卖淫嫖娼人员都被警方～了。

对此，白硕的回应是：“仔细学习了一下你的词条，的确很靠近我的想法了”。[3]

白硕[4]总结说：

自然语言的语义的确是一个博大精深的体系。知识图谱为语义计算准备好了基本的框架，但要全面推进到实用，还要做许多基础性的工作，包括资源建设和理论模型创新。我们期待在这一领域能有重量级的成果出现，将语义表示和计算的工作推向深入。

我们希望语义资源建设能够更好地为知识图谱和语义计算服务，并且在这个过程中逐步完善语义描述体系和词典构架。

5. 语义资源帮助机器人回答常识性问题

袁毓林教授的《**实词信息词典**》主要描述名词、动词和形容词的语义角色及其句法配置，同时突出相关词语所反映的常识概念和百科知识。特别是其中的《**汉语名词句法语义功能信息词典暨检索系统**》（简称《**名词信息词典**》），借鉴生成词库论（generative lexicon theory）关于词项的语义表达、特别是物性结构的有关学说，从服务于中文信息处理这种应用需求出发，来设计汉语名词的物性结构的描述体系。通过“物性角色”来描述名词所指的事物（简

[3] 2018 年 2 月 28 日 E-mail 通信。

称“事物”）的语义结构和相关的百科知识。调用这种语义资源，可以回答名词所指的事物（简称“事物”）的有关常识性问题。比如，是什么（形式角色）、有哪些部件（构成角色）、由什么材料做的（材料）、怎么形成的（施成）、有什么用途（功用），等等。这样，本文第1节中“围棋是什么？”，可以通过查询名词“围棋”的形式角色来回答。词条如下：

表 7 词条“围棋”的句法语义功能信息

词目：	围棋 01
汉语拼音：	wéiqí
词类属性：	名词
感情色彩：	中性
词义解释：	棋类运动的一种。棋盘上纵横各十九道线，交错成三百六十一个位，双方用黑白棋子对着，互相围攻，吃去对方的棋子。以占据位数多的为胜。
语义角色：	形式 FOR: 过程、事件，一种棋盘游戏； 构成 CON: 围棋，一种策略性两人棋类游戏，中国古时称“弈”，西方名称“Go”。流行于东亚国家（中、日、韩、朝），属琴棋书画四艺之一。围棋起源于中国，传为帝尧所作，春秋战国时期即有记载。隋唐时经朝鲜传入日本，流传到欧美各国。围棋蕴含着中华文化的丰富内涵，它是中国文化与文明的体现，等等；有起源时间、起源地、历史、棋具、规则、等级、头衔等；可以根据朝代分为:春秋战国时期、南北朝时期、明朝时期等；可以根据时代分为：古代、现代等；还可以根据国家分为：中国、日本、韩国等；阿尔法狗（AlphaGo）是第一个击败人类职业围棋选手、第一个战胜围棋世界冠军的人工智能程序； 单位 UNI: 个体：盘，等等；动量：场、次，等等； 评价 EVA: 精彩、惊心动魄、传统、著名、复杂、普通、有趣、知名、公平、乏味、激烈、正规、古老、热门，等等； 施成 AGE: 发明，等等； 材料 MAT: 围棋、棋手，等等； 功用 TEL: 下、玩、用以娱乐、用以锻炼头脑、用以比赛智慧高下，等等； 行为 ACT: 发展、衰退，等等； 处置 HAN: 擅长、玩、迷恋、学习、学会、提到、开设、组织、研究、爱上、喜欢、关心、倡导、爱好、喜爱、通晓、讲解、倡导、开展、扶持、探索、介绍、推广、普及、熟悉，等等；
句法格式：	S1: __（的＋）＋CON 如: ～（的）起源时间｜～（的）起源地｜～（的）历史｜～（的）棋具｜～（的）规则｜～（的）等级｜～（的）头衔 S2: CON＋（的＋）__ 如: 春秋战国时期（的）～｜南北朝时期（的）～｜明朝时期（的）～｜古代（的）～｜现代（的）～｜中国（的）～｜日本（的）～｜韩国（的）～ S3: Num＋UNI＋__ 如: 一场～｜一盘～｜一次～ S4: EVA＋（的＋）__ 如: 乏味（的）～｜激烈（的）～｜正规（的）～｜古老（的）～｜热门（的）～｜精彩（的）～｜惊心动魄（的）～｜传统（的）～｜著名（的）～｜复杂（的）～｜普通（的）～｜有趣（的）～｜知名（的）～｜公平（的）～

	S5：AGE＋__ 如：发明～ S6：__＋ACT 如：～发展丨～衰退 S7：（用＋）__＋TEL 如：下～ ｜ 玩～ ｜ 用～娱乐 ｜ 用～锻炼头脑 ｜ 用～比赛智慧的高下 S8：HAN＋__ 如：擅长～丨长玩～丨迷恋～丨学习～丨学会～丨提到～丨开设～丨组织～丨研究～丨爱上～丨喜欢～丨关心～丨倡导～丨爱好～丨喜爱～丨通晓～丨讲解～丨倡导～丨开展～丨扶持～丨探索～丨介绍～丨推广～丨普及～丨熟悉～

更加重要的是，该语义资源还可以跟计算机视觉技术相结合；来帮助机器人基于词典进行常识推理，并且回答常识性问题。比如，下列任务（图 1）原本是一个机器人智能推理的实验[4]。每一组任务（纵列看）中，上图是一些工具，下图是要求完成任务（铲土）；让机器人判断用什么工具来完成(下图中绿色的是判断机器人抓手的地方,红色是判断土的位置)。任务 1 的上图正常的铲土工具（机器人选择了铲子和刷子作为铲土的第一、第二选择），任务 2 是拿走铲子、刷子，仅提供其他家庭用品（机器人选了平底锅和杯子来铲土），任务 3 是一般的石器（机器人选了两个不同形状的石头）。

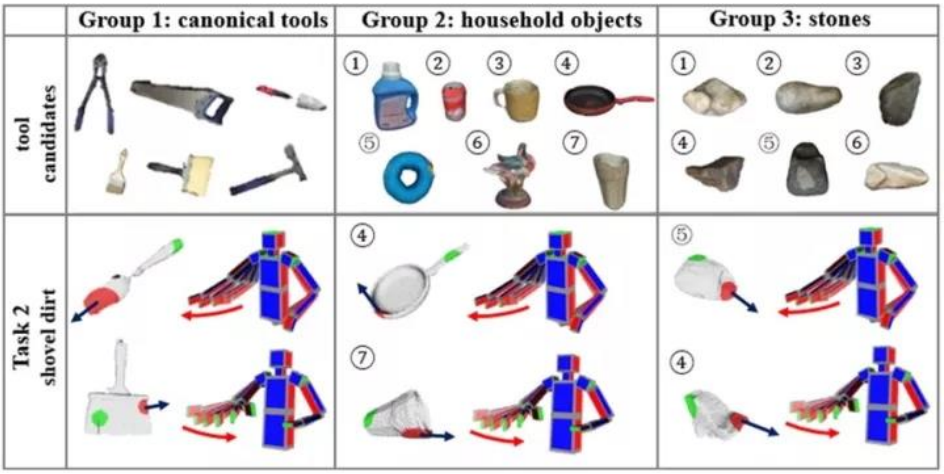

图 1 机器人智能推理任务

我们设想，完成这个任务，如果结合基于名词的物性角色进行推理和验证，那么效果也许更好。比如，先验地设定诸如下面这一类启发式规则（heuristic rules）：

（1）要了解事物是什么，就查相应名词的形式角色；

（2）要了解事物是有哪些部件，就查相应名词的构成角色；

（3）要了解事物是由什么材料做的，就查相应名词的材料角色；

（4）要了解事物是怎么形成的，就查相应名词的施成角色；

（5）要了解事物是有什么用途，就查相应名词的功用角色。

下面是名词“铲子”的词条：

表 8 词条“铲子”的句法语义功能信息

词目：	铲子
汉语拼音：	chǎn • zi

[4] 该任务和图片选自文献[13]。

词类属性：	名词
感情色彩：	中性
词义解释：	撮取或清除东西的用具。
语义角色：	形式 FOR: 人造物、工具； 构成 CON: 铲子由铲头或是铁柄、木柄等构成，根据种类的不同分为：炒菜用的、掘土用、儿童玩具，等等； 单位 UNI: 个体：个、把，等等；集合：类、部分，等等；不定：些，等等；容器：车厢、屋子，等等； 评价 EVA: 好、高质量、一流、干净，等等； 施成 AGE: 制造、做、生产、加工，等等； 材料 MAT: 不锈钢、木头、铁、塑料，等等； 功用 TEL: 铲土、翻东西、挖掘、抛掷东西，等等； 行为 ACT: 不见了、消失，等等； 处置 HAN: 抓、握、拿、摔、砸、使用、挥动、玩、操起，等等；
句法格式：	S1: CON +（的）__ 如: 炒菜～ | 挖土～ | 煤～ | 兵工～ S2: Num +NUI + __ 如: 一个～ | 一把～ | 一部分～ | 一些～ | 一类～ | 一屋子～ | 一车厢～ S3: EVA + __ 如: 好～ | 高质量～ | 整洁～ | 一流的～ | 干净的～ S4: MAT + __ 如: 塑料～ | 不锈钢～ | 铁～ S5: __ +TEL 如: ～能铲土 | ～能挖东西 | ～抛掷东西 S6: AGE + __ 如: 制造~| 做～ | 生产～ | 加工～ S7: __ +ACT 如: ~不见了 | ～消失了 S8: HAN + __ 如: 抓～ | 握～ | 拿～ | 摔～ | 砸～ | 使用～ | 挥动～ | 玩～ | 操起～

从“铲子”的功用角色中，我们可以发现铲子的用途之一是能够铲土。通过这种功用角色，能够类推出其他家庭物品也作为替代品，从而完成铲土的任务。

另外一个应用场景也是基于计算机视觉的。现在，计算机读下面这一个图[5]：

[5] 图片选自百度图片“客厅”。

图 2 场景识别任务“客厅”

机器能够识别出里面的物品，但是它不知道这些物品背后的含义。而人是知道这些物品都是干什么的，所以就能判断出这个图的深层含义。比如，判断出上图是客厅，能够会客等等，或者还能推理出其他功用。因为人看一个物体，就能知道他的功用是什么、通常放在什么房间中。通过图中“桌子、椅子、茶几、电视”等物品跟各种房间的匹配，可以发现，这个房间跟“客厅”最接近。下面是我们词典中“客厅”这一词条。

表 9 词条“客厅”的句法语义功能信息

词目：	客厅
汉语拼音：	kètīng
词类属性：	名词
感情色彩：	中性
词义解释：	主人与客人会面的房间，也是房子的门面。
语义角色：	形式 FOR: 人造物、处所、房间； 构成 CON: 外面常常有阳台，里面有桌子、椅子、沙发、茶几、电视，上面有吊灯等。 单位 UNI: 个体：个、间，等等； 评价 EVA: 宽敞、明亮、漂亮、豪华、狭小、简谱、雅洁、舒适、典雅、整洁、中式、地中海式、视野开阔、古朴、空荡荡，等等； 施成 AGE: 搭建、建，等等； 功用 TEL: 会面、招待客人、休闲，等等； 处置 HAN: 冲进、布置、装修、装饰、迈入、穿过、打扫，等等；
句法格式：	S1: __ + 的 + CON 如: ～的阳台 | ～的沙发 | ～的电视 | ～的茶几 | ～的吊灯 S2: Num + UNI + __ 如: 一个～ | 一间～ | 一屋子～ S3: EVA +（的）+__ 如: 宽敞的～ | 明亮的～ | 漂亮的～ | 豪华的～ | 狭小的～ | 简谱的～ | 雅洁的～ | 典雅的～ | 整洁的～ | 中式（的）～ | 地中海式（的）~ | 视野

	开阔的~ | 古朴的~ | 空荡荡的~ S4: AGE + __ 如: 搭建～ | 建～ S5: TEL + __ 如: 在～会面 | 在～招待 S6: HAN + __ 如: 冲进～ | 布置～ | 装修～ | 装饰～ | 迈入～ | 穿过~| 打扫~

当然，上面这两个任务似乎都比较大，需要结合计算机视觉、基于深度学习的分类，再加上基于词典资源的常识推理等多方面的协同，才能高质量地完成。

参考文献

[1] Chomsky, Noam. The Science of Language: Interviews with James McGilvray. Cambridge University Press, 2011.《语言的科学：詹姆斯·麦克吉尔弗雷访谈录》，曹道根、胡朋志译，北京：商务印书馆，2015: 226-238.

[2] Pinker, Steven. The Stuff of Thought: Language as a Window into Human Nature[M]. New York: Penguin Groups, Viking Press, 2007.《思想本质：语言洞察人类天性之窗》，张旭红、梅德明译，杭州: 浙江人民出版社，2015: 332.

[3] Daniel, Bor. The Ravenous Brain: How the New Science of Consciousness Explains Our Insatiable Search for Meaning[M]. Basic Books, 2012.《贪婪的大脑：为何人类会无止境地寻求意义》，林旭文译，北京：机械工业出版社，2013: 2-3.

[4] 白硕. 人工智能的诗与远方：一文读懂 NLP 起源、流派和技术[EB/OL]. http://mp.weixin.qq.com/s/VpWabo7_ekA7j_fw2ZuDdA. 2018-1-11.

[5] Winograd, Terry. Language as a Cognitive Process[M]. Addison-Wesley Publishing Company, Inc. 1983.

[6] Mugan, Jonathan. The Two Paths from Natural Language Processing to Artificial Intelligence[EB/OL]. https://medium.com/intuitionmachine/the-two-paths-from-natural-language-processing-to-artificial-intelligence-d5384ddbfc18. 2017-2-9.

[7] Hassler, Susan 马文·明斯基与人工智能之探[J]. 科技纵览（IEEE Spectrum），2016(3):8.

[8] 罗兰德·豪塞尔. A Computational Model of Natural Language Communication--Interpretation, Inference and Production in Database Semantics[M]. Springer Science & Business Media, 2006.《自然语言交流的计算机模型——数据库语义学下的语言理解、推理和生成》，冯秋香译、冯志伟审校，北京：商务印书馆，2016: 前言，第 xi 页。

[9] Marcus, Gary. Deep learning: A critical appraisal[J]. arXiv preprint arXiv:1801.00631, 2018.

[10] 文强，刘小芹. AAAI 前主席回怼马库斯[EB/OL]. https://mp.weixin.qq.com/s/Z0KEPCz3U51EtxBbzfh_jQ. 2018-1-5

[11] Yann LeCun. 深度学习已死，可微分编程万岁[EB/OL]. https://mp.weixin.qq.com/s/xyjrr5uWGP-oYsRWCqqTDg. 2018-1-6.

[12] 金丝猴. 为什么知识图谱终于火了？[EB/OL]. http://baijiahao.baidu.com/s?id=1585311312955034812&wfr=spider&for=pc. 2017-11-28.

[13] 朱松纯. 浅谈人工智能：现状、任务、构架与统一[EB/OL]. https://mp.weixin.qq.com/s/-wSYLu-XvOrsST8_KEUa-Q. 2017-11-2.

袁毓林（1962——），教授、博士生导师，主要研究领域为理论语言学和汉语语言学，特别是句法学、语义学、语用学、计算语言学和中文信息处理等。

Email：yuanyl@pku.edu.cn

卢达威（1983——），博士，主要研究领域为中文信息处理。
Email：wedalu@163.com