

自然语言处理：让机器善解人意

华为诺亚方舟实验室语音语义首席科学家刘群指出，随着人类进入智能时代，智能设备和数据量都空前增长，通过语音和语言进行人机交互的需求也在急速增长，语音和自然语言处理领域将涌现更多有影响力的前沿研究和技术创新。

文/许胜蓝 薛桦

自然语言是人类智慧的独特表现。自然语言处理（Natural Language Processing, NLP）是计算机科学领域与人工智能领域的一个重要研究方向，旨在研究人机之间用自然语言进行有效通信的理论和办法。根据Gartner发布的《2018世界人工智能产业发展蓝皮书》，到2021年，全球自然语言处理市场的价值预计会达到160亿美元。

自然语言处理是AI皇冠上的明珠

《营赢》：从整个科学领域的角度来考虑，您觉得自然语言处理的意义是什么？

刘群：语言学是一门古老的学科，人类为什么会有语言？动物为什么没有发展出人类这样复杂高级的语言？语言机制是人类大脑中先天就有的，还是像其他能力一样后天获得的？语言是如何形成和发展的？语言本身服从一些怎样的规律？无数的不解之谜等待着科学家来回答。

计算语言学，或者自然语言处理，既是一门科学，也是一门应用技术。从科学角度说，像其他计算机科学一样，它是一种从模拟角度来研究语言的学科。自然语言处理并不直接研究人类语言的机制，而是试图让机器去模拟人类的语言能力。如果说计算机拥有了像人一样的语言能力，从某种角度，我们就可以说，我们理解了人类的语言机制。由于理解自然语言需要关于外在世界的广泛知识以及运用操作这些知识的能力，所以自然语言处理是一个人工智能完备（AI-complete）的问题，并被视为人工智能的核心问题之一。





自然语言理解和处理也是人工智能中最难的部分。比如一幅图像，改变像素，或者一个局部，对整个图像的内容影响并不太大。但文字就不一样了，很多情况下，一句话中改变一个字，意思会完全不一样。



《营赢》：有人认为“自然语言处理应该是实现通用人工智能的关键挑战，得语言者得天下”，您怎么看？

刘群：这句话有一定的道理。有人把人的智能分为三大类：感知智能、运动智能和认知智能。

第一，感知智能，包括听觉、视觉、触觉等；最近两年，深度学习的引入大幅度提高了语音和图像的识别率，所以计算机在感知智能层面已经做得相当不错了，在一些典型的测试下，达到或者超过了人类的平均水平。

第二，运动智能，指能够在复杂的环境中自由行动的能力。运动智能是机器人研究的核心问题之一。

第三，认知智能属于最高级的智能活动。动物也具有感知智能和运动智能，但在认知智能方面，却明显低于人类。认知智能是包括理解、运用语言的能力，掌握知识、运用知识的能力，以及在语言和知识基础上的推理、规划和决策能力。认知智能中最基础也是最重要的部分就是语言智能，研究语言智能的学科就是自然语言处理。

自然语言处理的研究对象是人类语言，如词语、短语、句子、篇章等。通过对这些语言单位的分析，我们不仅希望理解语言所表达的字面含义，还希望能理解说话人所表达的情感，以及说话人通过语言所传达的意图。没有成功的自然语言处理，就不会有真正的认知智能。

自然语言理解和处理也是人工智能中最难的部分。比如一幅图像，改变像素，或者一个局部，对整个图

像的内容影响并不太大。但文字就不一样了，很多情况下，一句话中改变一个字，意思会完全不一样。

很多人工智能/机器学习技术，在图像识别领域，已经获得了很大的成功，但在自然语言处理领域，还处于起步的阶段。比如说，风格的生成。现在人们很容易把一幅画转变成梵高的风格，但要把一段话转换成莎士比亚风格，还没有技术能够做得很好。

所以，有些学者把自然语言处理称为“人工智能皇冠上的明珠”，是很有道理的，我非常认同这种说法。

自然语言处理的赛道

《营赢》：您认为目前自然语言处理面临的最大问题是什么？

刘群：我认为自然语言处理今后面临的主要问题是两个。

一是语义理解，或者说知识的学习，或常识的学习问题。这是自然语言处理技术如何变得更“深”的问题。尽管常识的理解对人类来说不是问题，但是它却很难被教给机器。比如我们可以对手机助手说“查找附近的餐馆”，手机就会在地图上显示出附近餐馆的位置。但你如果说“我饿了”，手机助手可能就无动于衷，因为它缺乏“饿了需要就餐”这样的常识，除非手机设计者把这种常识灌入到了这个系统中。但大量的这种常识都潜藏在我们意识的深处，AI系统的设计者几乎不可能把所有这样的常识都总结出来，并灌入到系统中。

二是低资源问题。所谓无监督学习、Zero-shot学习、

“

深度学习的应用，使得自然语言处理达到了前所未有的水平，也使得自然语言处理应用的范围大大扩展。可以说，自然语言处理的春天已经来临。

”

Few-shot学习、元学习、迁移学习等技术，本质上都是为了解决低资源问题。面对标注数据资源贫乏的问题，譬如小语种的机器翻译、特定领域对话系统、客服系统、多轮问答系统等，自然语言处理尚无良策。这类问题统称为低资源的自然语言处理问题。对这类问题，我们除了设法引入领域知识（词典、规则）以增强数据能力之外，还可以基于主动学习的方法来增加更多的人工标注数据，以及采用无监督和半监督的方法来利用未标注数据，或者采用多任务学习的方法来使用其他任务，甚至其他语言的信息，还可以使用迁移学习的方法来利用其他的模型。这是自然语言处理技术如何变得更“广”的问题。

《营赢》：过去十年，自然语言处理领域影响最远的研究是什么？

刘群：答案很明确，就是深度学习技术的引入。

基于深度神经网络的深度学习方法从根本上改变了自然语言处理技术的面貌，把自然语言处理问题的定义和求解从离散的符号域搬到了连续的数值域，导致整个问题的定义和所使用的数学工具与以前完全不同，极大地促进了自然语言处理研究的发展。

在深度学习技术引入自然语言处理之前，自然语言处理所使用的数学工具跟语音、图像、视频处理所使用的数学工具截然不同，这些不同模态之间的信息流动存在巨大的壁垒。而深度学习的应用，把自然语言处理和语音、图像、视频处理所使用的数学工具统一起来了，从而打破了这些不同模态信息之间的壁垒，使得多模态

信息的处理和融合成为可能。

总之，深度学习的应用，使得自然语言处理达到了前所未有的水平，也使得自然语言处理应用的范围大大扩展。可以说，自然语言处理的春天已经来临。

《营赢》：自然语言处理对框架乃至硬件层面提出了哪些独特的需求呢？

刘群：AI研究对硬件的需求是无止境的。可以说，贫穷限制了想象力，硬件限制了模型。只有在更好的硬件条件下，科学家才能尝试更复杂的模型，才能发展出更新更好的方法。

我不觉得自然语言处理对框架和硬件有独特的需求，自然语言处理的需求跟其他AI研究的需求应该是类似的，更多的内存、更高的带宽、更多的并行、更高的速度，永远都是需求的。有了这些，针对自然语言处理问题做一些特定的优化，并不是太大的问题。

《营赢》：2018年，学术界有哪些新的方法或者趋势，让您觉得会对自然语言处理的应用产生新的影响？

刘群：2018年自然语言处理研究领域最令人惊艳的成果是预训练语言模型，这包括基于RNN的Elmo和基于Transformer的GPT和Bert。预训练语言模型的成功充分证明了我们可以从海量的无标注文本中学到大量潜在的知识，而无需为每一项自然语言处理任务都标注大量的数据。

在应用方面，Google演示的Duplex技术让人耳目一新。与此同时，国内几个公司提供的会议同声翻译技术

也令人印象深刻，语音同传虽然还有很多错误，与人类同声传译相比还有很大差距，但其实用性已经毋庸置疑。这项技术的实际应用在几年前还难以想象，在这么短时间内就达到了现在这样可以初步实用的水平，真是非常出乎意料。

华为发力自然语言处理

《营赢》：华为诺亚方舟实验室在自然语言处理领域的技术布局包括哪些？取得了哪些进展？

刘群：华为诺亚方舟实验室的自然语言处理研究主要包括三个大的方向：语音技术、机器翻译和对话技术。

诺亚方舟的自然语言处理技术，已经开始在华为的各种产品和服务中获得广泛的应用。以大家日常使用的华为手机为例，华为的手机语音助手就集成了诺亚方舟的语音识别和对话技术。诺亚方舟的机器翻译技术支持了华为内部海量的技术资料翻译。诺亚方舟基于知识图谱的问答技术为华为的全球技术支持系统（GTS），提供了快速准确地回答复杂技术问题的能力。

诺亚方舟实验室的自然语言处理在研究方面也取得了突出的成果。我们在自然语言文本匹配、对话生成、神经网络机器翻译方面的研究成果被研究者广泛引用。在NIPS近五年来引用最多的50篇论文中就有一篇是我们的论文，在ACL近五年来应用最高的20篇论文中也有我们的一篇论文、引用最高的30篇论文中有我们的3篇论文。

《营赢》：听觉、视觉结合的多模态是否会自然语言处理未来很有前途的研究方向？华为是否已有这方面的研究了？

刘群：是的。我们已经开始在这一领域开展研究了。

自然语言处理场景应用产业化

《营赢》：自然语言作为人类社会信息的载体，使得自然语言处理不只是计算机科学的专属。自然语言处理如何能成为其他领域（如金融、法律、医疗健康等）的重要支撑技术？

刘群：自然语言处理技术早已深入我们的日常生活。很多人没有意识到，我们每天都在享受自然语言处理技术提供的便利，比如，拼音输入法。很多人大概

有印象，20年前流行过五笔字型输入法。但后来，拼音输入法几乎完全取代了五笔字型。

其实当年五笔字型流行的时候并非没有拼音输入法，那时的拼音输入法智能程度非常低，用户需要为每一个拼音从大量同音字中选择正确的汉字，联想功能也非常弱，输入速度非常慢。只有到后来自然语言处理技术取得了长足的进步，统计语言模型使得我们能够为一长串的拼音自动选择最可能的汉字序列，才使得拼音输入法最终取代五笔字型，并占据了汉字输入法的主导地位。

另外，现在的搜索引擎也大量使用了自然语言处理技术。你在百度或者谷歌搜索“四川的人口”，搜索引擎除了给你一系列相关的网页以外，还会直接给出一个具体的答案，这就用到了自然语言问答技术。

在金融、法律、医疗健康等领域，自然语言处理技术也得到了越来越广泛的应用。在金融领域，自然语言处理可以为证券投资提供各种分析数据，如热点挖掘、舆情分析等，还可以进行金融风险、欺诈识别等。在法律领域，自然语言处理可以帮助进行案例搜索、判决预测、法律文书自动生成、法律文本翻译、智能问答等等。在医疗健康领域，自然语言处理技术更是有着广阔的应用前景，如病历的辅助录入、医学资料的检索和分析、辅助诊断等等。现代医学资料浩如烟海，新的医学手段、方法发展迅猛，没有任何医生和专家能够掌握所有的医学发展的动态，自然语言处理可以帮助医生快速准确地找到各种疑难病症最近的研究进展，使得病人最快地享受医学技术进步的成果。

《营赢》：自然语言的研究对普通人有什么影响，会给我们的生活带来什么方便或冲击？

刘群：会让大家的生活越来越方便。比如打客服电话，你不用再选择一大堆的语音菜单。语音助手可以理解你的需求，贴心地帮你完成日常生活中的各种任务。机器甚至可以帮你写报告、写诗、写情书等等。与此同时，技术的进步也会给我们的生活带来一些冲击。比如就业方面，机器取代人工会造成一些人失业。但新技术的应用让一些职业消失的同时，又创造出大量新的就业机会。作为个人应该积极主动地想办法去适应这种变化，而不是消极等待和抱怨。winwin