

陈孝良的个人博客

分享

http://blog.sciencenet.cn/u/brygid

分享声学资讯，传递声学知识

博客首页

动态

微博

博文

相册

主题

分享

好友

留言板

博文

自然语言理解(NLU)概念浅析

精选

已有 11327 次阅读 2016-3-5 11:57 | 系统分类:科研笔记 | 关键词:自然语言理解



近来，人工智能（AI）行业利好消息不断，各大巨头都在积极拼抢这个领域的人才，笔者最近也频繁收到相关方面的咨询和讨论，不得不说，有时候大家真的搞错。人工智能是个非常大的概念，囊括了几乎所有学科，并非单纯计算机、通信、声学、光学等领域，也并非热炒的机器学习(例如深度学习和强化学习)就能实现的。然而，基于大数据和机器学习的狭义人工智能，确实已经给人类生活带来了巨大的变化，比如搜索、电商、广告、社交、语音等等。

一直以来，人们似乎都把语音识别看的极其重要，并认为语音识别就是最接近人工智能的领域。实际上这是一个误区，语音和按键、触摸、手势等方法一样，只是人类交互的一种手段。也有很多人认为语音帮助人类产生了智慧，然而，这一观点并没有有力的证据支撑。目前，这方面的基础研究还没有实质性的进展，我们对人体精密构造的认知还是非常浅薄的，至于智慧更是几乎一无所知。唯一能推断的就是，语音交互确实是我们获取知识的一种重要手段。

鉴于以上的认识和误解，笔者认为应该写篇这方面的科普文章，虽然笔者也不精通这个领域，但是我们可以从底层声学这个范畴来思考和讨论一下人工智能相关的概念，确实这方面通俗易懂的资料太少了，也希望更多学术和产业领域的大牛们来指正和探讨。

在此，笔者要先摆明自己的观点，目前人工智能的相关应用中，语音识别并非是关键点，而且语音识别也并未给我们生活带来多大变化。语音识别经过十多年的高速发展，特别是结合深度学习的融合发展，目前的精度已经相当高，然而达到99%以后，再提升就显得非常困难。但为什么语音识别的精度到达如此程度，我们却对其仍然不满意呢？其实，这和语音识别的关系



陈孝良

加为好友

给我留言

打个招呼

发送消息

扫一扫，分享此博文



作者的精选博文

全部

- 盘点麦克风技术及市场，远场
- Apple HomePod技术解读，为
- 为何最成功的消费级AI产品是
- 如何正确评价声纹识别小度
- 80亿美元巨资，三星为什么要
- 语音交互为何迭代如此之快？

作者的其他最新博文

全部

- 盘点麦克风技术及市场，远场
- Apple HomePod技术解读，为
- 为何最成功的消费级AI产品是
- 如何正确评价声纹识别小度
- 80亿美元巨资，三星为什么要
- 语音交互为何迭代如此之快？

热门博文导读

全部

- 第一个基金要交账了——钱...
- 美国运动员工资制度真相：...
- 米尔扎哈尼：首个女性菲尔...
- 我们为什么要锻炼身体？
- 向民间科学家的精神致敬
- 中国学术奖励制度的真相：...

相关博文

- 《理解：关键在歧义化解》...
- 自然语言理解：语言歧义化...
- 什么是理解（如何做到可形...

不大，是我们误把语言理解的概念强加给了语音识别。实际上，语音识别只是人工智能中的一个小学科，现在也可以算是深度学习中的一个分支。类似于人类，语音识别至多是我们听觉系统内的一个神经单元，只负责将语言转化成人类可听到的信号激励，若在计算机应用中，就是转化成我们常说的文字。

那么关键点在哪？从各大巨头发布的宣传资料和介绍的相关信息来看，实际上大家都在追求：**自然语言处理(Natural Language Process, NLP)或者说自然语言理解(Natural Language Understanding, NLU)的突破**。当然，这也是得益于基础声学 and 语音识别的巨大进步。**基础声学 and 语音识别解决的是计算机“听得见”的问题，而NLP实际上就是要解决计算机“听得懂”的问题**。这么看来，“听得懂”才是目前十年内最为关键的问题，下面，笔者就NLP的概念以及国内外的学习资料进行盘点与分析。

语音和语言有何区别？

我们首先要明确一下语音和语言的区别，这是一个基础性的问题。**语音 (Speech) 是语言 (Language) 的信号载体，语音是人的发音器官发出的，承载一定的语言意义，而语言才承载人类的智慧**。通俗的讲，语音是天生就存在的，婴儿的咿呀咿呀也算是语音，甚至其哭声也代表一定的意义，而语言则是需要学习不断进化的。

人类的进化为什么自然的选择了语音作为交互的手段？一般认为，是因为语音有着得天独厚优越性，比如声音比光更容易传播，而且不受昼夜变化影响，同时也有更广的传播范围。当然，也有许多科学家提出反对的声音，比如很多动物的听觉实际上比人类更为发达，为什么却没有进化出来语音交互呢？这个问题我们暂时放到后面文章探讨，事实上动物界中的很多动物也会发出一定的声音来进行信息的交互，但是仅有语音交互是不够的。

单独强调语音，而不说语言实际上没有任何意义的。当然语言也不仅是只有语音这一种载体，语音是不好保存和搜索的，所以人类还发明了文字这个神奇的东西。文字却又是通过视觉来感受的，不过文字也仅是帮助我们映射而已。象形文字刚开始的时候，还和视觉有直接的关联。后来，不论是汉语，还是其他语言基本上都进行了抽象，与听觉以及视觉的关联不再明显。到此为止，语言就已经完全抽象出来，和我们所理解的概念相关了，已经远远超过了语音、图像等识别的概念，甚至数学也成为了我们的一种语言。

因此当人类使用语言来存储知识和表示概念的时候，计算机和机器人也就必须学会适应这种方式。也就是说，**计算机和机器人在未来的趋势中，必然是理解语言这个极其复杂而又承载人类知识传承的表达方式**。但是，做到这一点是极其困难的，虽然我们看到了很多的进展，但是，机器理解语言在短期内还无法看到实质性的突破。这个领域需要更多优秀的人才加入，笔者写这篇文章目的也是这个原因，人才的涌入才会带来希望，至少我们不断的积累总会带来些许进步。

如何描述和定义NLP？

事实上，自然语言处理 (NLP)，或者说自然语言理解 (NLU)，或者说计算语言学 (CL) 很难有个准确的定义。1999年美国计算机学家Bill Manaris曾这样描述：自然语言处理是一门研究人与人交际中，以及人与计算机交际中的语言问题的学科。自然语言处理 (NLP) 研究表示语言能力、语言应用的模型，通过建立计算机框架来实现这样的语言模型，并提出相应的方法来不断地完善这样的语言模型，还要根据这样的语言模型来设计各种实用系统，并探讨这些

[illegible]

NLP涉及哪些研究内容？

CRF分词 文本推荐 CRF依存句法分析
DA评测工具 CONLL语料预处理 基于互信息和左右信息熵的短语提取
声母 TEXTRANK自动摘要 简繁转换 识别新工具 繁体中文分词 地名识别
TEXTRANK关键词提取 用户自定义词典 多音字 依存句法分析
自动摘要 关键词提取 拼音转换
拼音转换 用户自定义词典 多音字
CONLL语料预处理 简繁转换 语义推荐
语料预处理 简繁转换 语义推荐
简繁转换 语义推荐
索引分词 最短路径分词
词性标注 CRF分词 音译人名识别
声调 地名识别 短语提取
韵母 自动摘要 极速词典分词
文本推荐 BIGRAM统计 MAXENT依存句法分析 繁体中文分词
声字在线

自然语言理解研究的目标自然是希望通过图灵测试，包括了语音、形态、语法、语义以及语用等方面，也要解决人类语言中的因果、逻辑和推理的问题。这是非常重要的，我们知道**大数据仅仅是代表了历史的规律和预测的趋势，如果没有因果关系和逻辑推理，计算机和机器人就始终无法为人类提供可参考的建议，甚至是决策。**

NLP当前难点都有哪些？

这个就太多了，上面我们提到过了，人类的语言融合了语音和文字，实际上也是听觉和视觉的融合，不是割裂开的。但是，现在这两个学科却是互相独立的，人为造成了这种研究的困难。另外，人类的语言多样化，每种语言的表达方式也千差万别。理想主义甚至认为人类的很多语言知识是与生俱来的，这个说法也有一定的实验支持，声学在线曾经报道过这方面的进展。不过，即使经验主义也是认为语言是通过感官和人脑的联想与抽象得来的。这是个大问题，因为这让很多研究人员没有了思路，科学的进展最怕这种困难。很多时候，我们还都把科学的进展归结为哲学的进步，没有了思路也就很难找到突破点，这需要基础研究的进展。

当然NLP现在实际的技术困难还是语义的复杂性，包含因果关系和逻辑推理的上下文等。现在解决这些问题的思路主要还是深度学习（Deep Learning）。深度学习带给了研究人员一种全新的思路，这倒不是最新的，BP神经网络刚兴起的时候，学术和产业界也是非常热闹，因为当时的研究人员认为总算找到了思路，后来BP神经网络几乎没啥实质进展，渐渐的就冷了很长一段时间。深度学习拓展了神经网络的层次，而且大数据的积累和并行计算的增强则给这种方法奠定了基础，这也是最近机器学习非常火热的原因。因此基于大数据、并行计算的深度学习将会给NLP带来长足的进步，但是若能达到人类的这种理解层次，恐怕仅靠这种方法也很难实现。

NLP学习的参考资料有哪些？

笔者一直在说自然语言理解的问题，文字还是最为重要的知识传承方式，因此自然也要提供一些参考资料给大家，但是这个领域的科普书籍实在太少了，笔者也只能推荐一些书籍、网上资料给大家参考学习，姑且先把这篇文章作为入门科普的参考吧。

首先，笔者建议还是要先看斯坦福大学的视频学习课程，哥伦比亚大学的视频学习课程也不错，只是这些课程还是英文的，国内也有几位老师的公开视频可以参考，比如宗成庆老师或者关毅老师的自然语言理解课程，冯志伟老师的计算语言学课程都还不错。微信上不方便放链接，感兴趣的朋友就到我们声学在线的网站上直接点击链接吧。

其次再推荐一些书籍，首推还是Daniel Jurafsky和James H. Martin的《Speech and Language Processing》，已经出第二版本了，也是国内外的主要教材。这两人还写了几本相关方面的书，都还不错，亚马逊等一些线上平台搜索一下就能查到。国内也有冯志伟老师翻译的《自然语言处理综论》可以参考。数学基础稍好的还应该看看《数学之美》（吴军）和《统计学习方法》（李航），计算机编程稍好的更应该看看《机器学习实战》（Peter Harrington）和《Python自然语言处理》（Steven Bird）等。

最后再推荐一些学术会议，通过查询这些学术会议可以获知很多的学术进展，最重要的是，可以通过这些学术组织的会员和文章了解很多相关的研究单位，这些学术会议包括：ACL: The Association for Computational Linguistics，AAAI: Association for the Advancement of Artificial Intelligence，ICCL: The International Committee on

Computational Linguistics , COLIPS: TheChinese and Oriental Languages
Information Processing Society 以及中文信息学会等。

NLP都有哪些知名研究机构？

笔者通过声学在线的众多朋友汇总了一些信息，可能不是很全面，不足的地方还请各位批评指正，这些研究机构（排名不分先后）包括：斯坦福大学自然语言处理研究小组，卡内基梅隆大学语言技术研究院，哥伦比亚大学自然语言处理研究组，约翰霍普金斯大学语言和语音处理研究组，康奈尔大学NLP研究组，宾夕法尼亚大学NLP研究组，伊利诺伊大学香槟分校NLP团队，南加利福尼亚大学自然语言研究组，爱丁堡大学语言与计算研究所，谷歌苏黎世研究院，亚马逊NLP研究组，香港科技大学NLP实验室，中科院计算机所自然语言处理研究组，北京大学计算语言学研究所，清华大学人机交互与媒体集成研究所，中科院声学所语言声学与内容理解重点实验室，中科院自动化所模式识别重点室，百度自然语言处理部，中科大人机语音通信研究评测实验室，哈尔滨工业大学计算机学院智能技术与自然语言处理研究室等。

总之，自然语言理解随着深度学习的发展，还是看到了很多进展，但是NLP所需要解决的语言问题真的是非常复杂，近期内还是难以有所突破。正因为有这些挑战，才更需要更多有才华的人才持续不断的投入这个行业，而且这是未来数十年的新兴行业，现在刚毕业的优秀学生都有可能拿到50万的年薪，其经济收入将是很可观的。当然这也是个问题，虚高的投入或许会给这个行业带来一些阻碍。

声学在线 soundonline



声学在线: soundonline.org

转载本文请联系原作者获取授权，同时请注明本文来自陈孝良科学网博客。

链接地址：<http://blog.sciencenet.cn/blog-1375795-960595.html>

上一篇：[3月3全国爱耳日，关注儿童听力健康](#)

下一篇：[下一场技术革命真的会发生在你的耳朵里吗？](#)

更多

举报

分享

收藏

当前推荐数：**9** 推荐人：[黄永义](#) [赵克勤](#) [文克玲](#) [谢维](#) [强涛](#) [谢平](#) [刘锋](#) [杜江平](#) [icgwang](#)

[推荐到博客首页](#)

评论 (7 个评论)

该博文允许注册用户评论 请点击登录



[7]pksc 2016-4-3 20:11

自然语言理解问题研究进展与战略意义<http://idea.cas.cn/viewdoc.action?docid=138>



[6]saofenghan 2016-3-22 15:37

赞同 

关于人工智能和自然语言理解，鄙人也有一些观点，给您发了消息和邮件，希望有空了可以看一看，讨论一下



[5]徐明昆 2016-3-6 12:09

NLP也不是什么重要的行业，起码几十年也不会有实质进展，我看永远也不会，博主不信，不妨等上几十年再看此言



[4]icgwang 2016-3-5 20:13

感觉语音还是挺特别的，隐藏着声带机械波和脑内电磁波互动间的密码。言为心声，就是它把让看不见的心表达成看得见的能量了！让机器人表情更传神，都不如让机器人说话像人更能打动人。现在的人际系统还是以（视觉）符号为中介进行信息交互的，如果能够直接能够语声波互动，而且非常拟真，那不能不说是极大极大的进步！

博主回复(2016-3-6 16:36): 



[3]wangbin6087 2016-3-5 20:07

语音是一个物理信号，经过“理解”才会有意义。语言是经过“约定”后的语音，是智慧的产物。

博主回复(2016-3-6 16:36): 同意



[2]zjhs111 2016-3-5 16:22

赞同：

如果没有因果关系和逻辑推理，计算机和机器人就始终无法为人类提供可参考的建议，甚至是决策。基础声学 and 语音识别解决的是计算机“听得见”的问题，而NLP实际上就是要解决计算机“听得懂”的问题。现在实际的技术困难还是语义的复杂性，包含因果关系和逻辑推理的上下文等。

.....
可是语义研究领域的学术研究气氛,态势与创新创业的形势实在不相适应啊!
期待博主能进一步做工作!

博主回复(2016-3-6 16:36): 是啊，期望也能冷静一下面向未来



[1]黄永义 2016-3-5 14:34

了不起！自然语言的理解与识别！