北京大学
PEKING UNIVERSITY

本科生毕业论文

题目：**基于语言特征和注意力机制的中文反讽识别研究**

Research on Chinese Irony Identification Based on Linguistic Features and Attention Mechanism

姓　　名：	邱晓枫
学　　号：	1600014401
院　　系：	中国语言文学系
专　　业：	应用语言学专业
导师姓名：	詹卫东

二〇二 零 年 5 月

论文评语页

分数：92

情感倾向性分析是自然语言文本理解的重要任务之一。作为情感倾向表达的反讽语言现象的识别，近年来引起学界越来越多的关注。邱晓枫的论文以反讽表达形式的识别为题展开研究，选题很有意义。他调研了前人相关研究文献，并作了细致的分析，采取了比较合理的研究路线，在对中文社交媒体文本中的反讽用法进行考察基础上提炼了反讽语言特征，并将语言特征融入到深度学习的 LSTM 模型，通过注意力机制的作用，在实验中取得了比传统的单纯基于深度学习数据驱动的模型更好的识别效果。论文的研究思路清晰，特别是结合深度学习模型与语言知识的融合，追求机器学习结果的可解释性，这种思路值得肯定。实验设计合理，论文写作规范、逻辑性强，结构安排得当，是一篇优秀的本科毕业论文。

当前情感分析任务已经从早期的分类模式进化到结构化的要素分析模式，不仅要识别情感的正负类别，还要进一步明确情感的主体和对象等结构化信息。反讽的信息处理也同样。此外，如何从句子层次提升到在篇章层次上进行反讽识别和要素分析，也是值得探讨的问题。

指导老师：詹卫东

2020 年 5 月 20 日

摘要

目前，情感分析是自然语言处理中最活跃的领域之一。反讽是一种特殊的表达情感的修辞手段，通过与文本字面义不一致的隐含义来达到讽刺或幽默的表达效果。反讽的实际语义同字面表达存在反差，因此对于反讽的识别和情感分析具有挑战性。为了提高情感分析的准确度，同时增进对反讽语言现象的认识，本文对中文反讽识别开展研究。

针对中文反讽研究实验数据稀缺问题，本文通过人工标注获取了 1291 条中文反讽语料并以此为基础构建了分布平衡的实验数据集。本文考虑中文社交媒体语言特点，结合反讽理论研究提炼出四种反讽语言的形式特征。在此基础上归纳得到 skip-n 元词组合、标记强烈情感强度的副词、“被+X”构式、特定的标点符号、特定的网络词汇五种具体语言特征。通过卡方统计量选取多种语言特征对应的特征词。本文还从面向计算机识别的角度对反讽小类进行划分。

考虑到反讽识别目标文本的时序性和非连续依赖问题，本文以 LSTM 为基础，提出了一个融合语言特征的注意力机制的中文反讽识别模型（Irony-Feature Enhanced Attention Network, IEAN）。实验结果显示，该模型较基准模型在识别性能上有所提升，F 值达到了 0.8390，证明了该模型能够结合语言特征更好地捕捉文本深层语义。此外，该模型较传统深度学习模型在可解释性上也表现出一定优势。

最后，本文在总结研究结论和不足的基础上提出了对未来工作的展望。

关键词：反讽识别；情感分析；语言特征；注意力机制；深度学习

目录

摘要. 1
目录. 2
表格与图片目录. 3
第一章 绪论. 4
1.1 研究背景及意义. 4
1.2 国内外研究现状. 4
1.2.1 反讽理论研究. 4
1.2.2 反讽识别研究. 5
1.3 当前研究的难点与本文主要工作. 6
1.4 论文框架. 7
第二章 中文社交媒体反讽的语言特征分析. 8
2.1 中文社交媒体反讽的语言特征. 8
2.1.1 表达负面语义的成分+表达非负面语义的成分. 8
2.1.2 提示负面情感的成分+表达非负面语义的成分. 9
2.1.3 提示非字面义的成分+表达非负面语义的成分. 9
2.1.4 居于语义或逻辑矛盾的成分+表达非负面语义的成分. 9
2.1.5 反讽相关的具体语言特征. 10
2.2 实验数据集建立. 12
2.3 卡方统计量. 12
2.4 语言特征选取. 13
2.5 面向计算机识别的反讽小类划分. 14
第三章 融合语言特征的注意力机制的中文反讽识别模型. 15
3.1 词嵌入向量（Word Embedding）. 15
3.2 融合语言特征的注意力机制的中文反讽识别模型. 15
3.3 模型参数设置和实现细节. 18
3.4 实验结果和分析. 18
第四章 结论与展望. 22
4.1 结论. 22
4.2 不足与展望. 22
参考文献. 24
致谢. 26
版权声明. 27

表格与图片目录

表2.1 特征词$\boldsymbol{t}$和文本类别$\boldsymbol{c_i}$关系表. 13
图 2.1 skip-n 元词组合解析示意图. 13
表 2.2 skip-n 元词组合统计值. 13
表 2.3 面向计算机识别的反讽小类划分. 14
图 3.1 融合语言特征的注意力机制的反讽识别模型框架. 16
表 3.1 第一组实验设置. 18
表 3.2 第一组实验结果. 18
表 3.3 神经网络模型主要参数设置. 19
表 3.4 第二组实验结果. 19
表 3.5 两种模型的注意力矩阵可视化. 20
表 3.6 一次实验的识别结果. 20

第一章 绪论

1.1 研究背景及意义

反讽（本文的反讽若不作其他说明则特指话语反讽，即 verbal irony），又称“反语”或“说反话”，作为一种修辞现象，通常是为了达到礼貌、讽刺、幽默、凸显等语用功能，以间接的方式表达说话人真实的意图（涂靖 2002，刘正光 2002）。这种真实意图往往体现了说话人对话题强烈的情感倾向，因此，根据其情感的指向性，可以将反语划分为两类：①反话正说。表达积极含义的反讽，主要用于以幽默、诙谐的手法表达对某些人或事物的调侃或赞赏。例如：“老王也没啥高人之处，除了比我俩多练个十几年。”表面是说老王平庸，实际表达说话人对其技艺的赞赏。②正话反说。表达消极含义的反讽，主要用于以讽刺、夸张的手法表达对某些人或事物的强烈不满或批评。例如：“节假日还要去加班，真是太充实了!”表面是赞扬假日充实，实际表达说话人对节假日加班的不满。在英文和中文中，反话正说的出现频率都远低于正话反说，因此，过去对反讽现象的研究主要是针对后者。

随着社交网络的兴起，反讽成为一种普遍的语言表达方式。针对热门话题及争议话题，用户常常使用反讽表达嘲讽、批评等强烈的负面情感倾向，这对正确分析社交媒体的用户情感提出了挑战。针对反讽识别的可计算化研究也成为研究者关注的热点，并具有重要意义：一方面，能够丰富反讽相关的研究成果，进而推动对反讽的本质意义、认知过程以及区分机制的认识；另一方面，能够提高情感分析、人机对话等自然语言处理任务的准确率，服务于舆情分析。

目前针对反讽识别的研究，主要是面向英文的。中文的语言结构复杂，实现反讽的方式相当丰富，因此面向中文的反讽识别研究是缺少且具有挑战性的。本文在对相关研究进行梳理的基础上，结合社交媒体特点和中文反讽语料实例，对中文反讽现象进行分析，归纳得到了一些与反讽现象相关度高的显式语言特征，再将其融入到本文提出的一种深度学习模型中进行反讽识别。实验表明，新模型能够更有效地捕捉文本的深层语义，较传统深度学习模型具有更好的可解释性。

1.2 国内外研究现状

1.2.1 反讽理论研究

在传统语言理论中，反语受到修辞学家和文学批评家的关注，作为一种修辞现象被研究。修辞学将反讽分为话语反讽（verbal irony）、情景反讽（situational irony）和戏剧反讽（dramatical irony）三类（赵毅衡 2011）。话语反讽指“语言外壳与真实意指之间的对照与矛盾”；情景反讽是由文本的主题立意、情节编排、叙事结构等文体因素所产生的一种内在张力；戏剧反讽则是旁观者上帝视角的全知全能和剧中人物的无知局限之间的张力。三者中，话语反讽具有片段性和局部性，是最常见的一类反讽。相比由复杂的文体因素孕育的情景反讽和戏剧人物、戏剧观众两个层面上展开的戏剧反讽，话语反讽是最容易识别的。同时，话语反讽实现反讽的材料篇幅最小、反讽意蕴最典型，适合构建语料库和作为面向计算机的反讽自动识别任务的研究对象。

随着语用学、心理语言学和认知语言学的兴起，反讽研究通过不同的理论和研究途径取

得了丰富的研究成果。以 Grice 为代表的经典理论指出反讽是隐含义代替明示义（Grice 1975）。为了达到交际意图，言者在提供信息给听者时需要遵从四条准则，即所提供的信息必须：充足、真实、相关和无歧义，如果要达到特殊的表达效果，则必须明显地违反上述“质准则”，以便让听者发现和识别隐含的真实交际意图。Sperber 等提出的提示理论认为，“反讽总是隐含地表达一种态度，反语的关联总是或至少部分地取决于言者对回应观点的态度”（Sperber 1984）。Clark 与 Gerrig 在接受 Grice 观点的基础上提出共同伪装理论（Clark & Gerrig 1984）。该理论认为反讽的言者 S 为了迎合某一特定观众群体，伪装成 S1 给一个想象中的听者 H1 说话，S 对 S1 所说的话持批评态度，H1 可能只理解了话语的字面义，但真正的听者 H 可以从 S1 的无知、考虑不周等识别出伪装意图并推到出意欲表达的反讽义。Giora 提出的间接否定理论则从反讽意义的特征出发，认为反讽是不使用明确的否定标记的否定形式（Giora 1995,1998）。Utsumi 的隐形展示理论认为，构成反讽的前提条件是反讽环境（Utsumi 2000）。假设话语发出前有时间点t_0和t_1，反讽环境由以下三个事件或状态构成：(1)在t_0时刻言者具有某种期望 E； (2)在t_1时刻言者的期望 E 与现实不一致；(3)言者对这种不一致产生了否定的情感态度。

经典理论将反讽视为一种违反“质准则”的特殊语言用法，对反语的生成动机和区分机制缺乏解释。提示理论和共同伪装理论较好地说明了反讽生成的心理动机和环境条件，但对其本质和区分机制的解释鲜有贡献。间接否定理论虽然在描写和解释反语事实方面有所进步，但回避了“反语到底是什么”这个问题，而且用“间接否定”作为区分机制，可操作性差。隐形展示理论在吸收了上述理论的精华部分，较好地解决了区分机制问题，而不能解释反语的功能。此外，还有一些研究启发我们重新审视对于反讽的理解和翻译——将“irony”翻译为“反讽”、“反语”难以准确传达“irony”的全貌。Dews 等在对 irony 语用功能的研究中认为，irony 是用来缓和批评和表扬语气的（Dews et al. 1999）。Holdcroft 认为 irony 并不一定表达一种否定态度，它也可以是“诙谐和深情的” （Holdcroft 1996）。即 irony 的涵盖范围扩大了。

1.2.2 反讽识别研究

反讽识别工作从可计算化的角度，主要研究反语识别的特征构建和分类学习方法。Utsumi 定义话语情境中的三种反讽属性，并以此建立一个识别反讽的计算模型（Utsumi 1996）。Gonlez-Ibanez 等采用字典词语和“@<用户>”等特征识别反语，发现仅通过简单特征的识别效果较差（Gonlez-Ibanez et al. 2011）。Reys 等选取了 n 元文法、词性的 n 元文法、幽默指数、词汇情感极性、情感复杂度和快乐程度六种复杂特征进行实验，发现这些特征在特定领域的反语识别中有明显作用（Reys et al. 2013）。Konstantin 等针对商品评论，将不同特征和分类器进行组合，实验结果显示人工选取的特征在提高分类准确率的同时降低了召回率，而将这些特征与词袋模型结合能够有效解决问题（Konstantin et al. 2014）。Edwin Lunando 等针对印尼社交媒体的数据集，结合负面情感信息和感叹词数量等特征进行讽刺识别，在最大熵模型上达到了 78.4%的最高精确率（Edwin et al. 2013）。David 等基于对 Twitter 数据集的研究，指出包含作者属性、受众和直接语言环境的上下文信息对讽刺识别具有重要作用（David et al. 2015）。以上模型均使用传统机器学习模型，基于特征的统计情况进行分类，难以挖掘深层的语义信息，往往还有严重的跨领域效果退化的问题。

随着深度学习的发展，Aniruddha 等首先将卷积神经网络（CNN）和长短期记忆网络（LSTM）模型应用于反讽识别，实验结果表明深度学习方法较传统机器学习方法有明显优势（Aniruddha et al. 2016）。Yi Tay 等考虑到现有模型难以精确捕捉文本中长期依赖和词对情感极性相反的特征，应用自注意力提取单词对间的信息，提出了一种上下文无关的反讽识

别模型（Yi Tay et al. 2016）。Devamanyu 等考虑用户、主题等上下文信息与文本信息结合，提出了一种复杂的结合上下文识别反讽的模型（Devamanyu et al. 2018）。Lotem 等先将含反讽的英文通过单语机器翻译技术翻译为不含反讽的英文，再对翻译结果的情感倾向打分来判别是否含反讽，并公布了 3000 条反讽的实验语料（Lotem et al. 2017）。

中文的反讽识别研究目前处在初步发展阶段，Tang 等构建了一个繁体字的反讽语料库，并分析了一些反语常用句式，对识别特征和分类算法并未提及（Tang et al. 2014）。邓钊等从新浪微博标记了 300 条反语和 28545 条非反语语料构建实验数据集，构建了基本词汇情感、谐音词、连续的标点符号、微博长度、动词被动化和双引号内外情感模糊度六种特征，在传统机器学习分类器上最高达到了 76.74%的准确率（邓钊 等 2015）。邢竹天等归纳了中文文本中意指义和字面义的偏离、情感的变化张力等特征，在 Logistic 模型上最高达到了 71.2%的召回率和 60.3%的分类准确率（邢竹天 等 2015）。孙晓等构建了一个反讽与非反讽语料各 1000 条的数据集，基于搭配规则和情感语义提出了一种 CNN 与 LSTM 混合的神经网络模型（孙晓 等 2015）。卢欣等将情感词、谐音词、网络词汇和搭配特征作为额外的特征信息输入 CNN 模型，达到了较高的识别效果（卢欣 等 2019）。

1.3 当前研究的难点与本文主要工作

中文反讽识别研究目前主要存在以下难点：

语言差异。不同语言的语言习惯不同，语法结构也有很大差异，英文反讽中的方法和语言特征不能直接对应。同英文相比，中文的语义和语法结构更加复杂，一些反讽现象连人为标注也难以准确识别。面向中文的反讽识别具有不小难度。

缺乏完整、权威的中文反讽语料库。目前中文反讽识别研究领域还没有一个权威的数据集，语料主要靠各研究单位人工标注，语料质量难免会受主观因素影响。已有研究多在 Tang 构建的 950 条繁体字语料库（Tang et al. 2014）的基础上扩充，基于部分规则的构建方法使得该语料库的语料模式比较单一，比如，超过 58%的语料是“很好”或“太好”后接带有负面情感的内容、超过 35%的语料是 “可以再”后接表示负面情感的形容词。如果扩充处理不当，作为反讽识别的数据集，分类器的过拟合风险很高。

社交媒体文本的语言特点。作为一种新兴媒体，社交媒体的文本数据口语化程度高，长度有限，相较新闻等书面语化程度高的文本更贴近人们的日常交流。然而这种文本缺少自然会话中的语气、身体姿势等视听理解的辅助手段，也不含社会背景、个人身份等超文本信息，很难形式化地给出语境的计算表达方式。此外，社交媒体文本的语言具有不规范性，包括错别字、标点符号缺失、不符合标准汉语习惯等，进一步增加了反讽识别的难度。

模型的局限性。传统的机器学习模型提取特征和语言建模的方法丢失了文本的词序信息，难以挖掘深层语义，识别准确度有限。以卷积神经网络（CNN）为代表的深度学习模型在分类效果上有所提升（Aniruddha et al. 2016），但在自然语言处理中，文本具有时序性这一显著特点，词序和语义有密切关系。传统的 CNN 模型从连续的 N-gram 向量矩阵中得到局部特征，无法解决长距离依赖问题。深度学习模型的网络结构复杂、参数量巨大，可解释性不好。

本文关注以上研究难点并尝试针对性解决，主要进行了如下工作：

(1)中文反讽实验数据集构建。目前中文反讽研究较少，实验数据也未公开，本文从微博、博客等社交媒体上收集语料。由于隐形展示理论很好地解决了反语区分机制上的问题，以该理论作为人工标注反讽的指导思想，同时进行情感标注。对数据类别进行平衡化处理。

(2)中文文本的语言特征选择。本文在参考中英文反讽识别研究的基础上，考虑中文语言特点，结合反讽语言学理论进行仔细分析，从概括到具体地归纳了几种与反讽相关的语言特

征。采用卡方统计量选取语言特征对应的特征词。

(3)结合语言特征注意力机制的 LSTM 模型。为了解决文本中长距离依赖问题，更好地对句子进行语义表示，本文提出了一个以 LSTM 为基础，融合语言特征的注意力机制的中文反讽识别模型，并进行了一系列实验。实验结果表明，和其他模型对比，该模型的识别性能有提升。该模型能够输出注意力矩阵查看句子的表示情况，结合特征分析，较一般的深度学习模型具有更好的可解释性。结合分类结果和具体语料，从计算机识别的角度对反讽小类进行重新划分。

1.4 论文框架

本文的主要框架如下：

第一章：绪论。本章介绍了中文反讽识别的研究背景和意义。梳理了反讽语言学理论和反讽识别相关研究情况，总结当前研究的难点并引出本文主要研究内容和组织结构。

第二章：中文社交媒体反讽的语言特征分析。基于研究任务，本章提出了对反讽本体的思考和特征描述，从概括到具体地对中文社交媒体反讽的语言特征进行分析。本文人工构建了实验数据集，再通过卡方统计量选取语言特征对应的特征词，并实验验证了这些特征的有效性。本章最后还从计算机识别的角度对反讽小类进行划分。

第三章：融合语言特征的注意力机制的中文反讽识别模型。本章结合融合语言特征的注意力机制，提出了一个新的中文反讽识别模型。接下来进行了一系列实验来说明新模型在性能和可解释性上的提升，也对模型存在的问题进行了分析、探讨。

第四章：结论和展望。总结了本文的相关研究成果和不足并提出了对未来工作的展望。

第二章 中文社交媒体反讽的语言特征分析

本章结合国内外反讽相关研究对反讽的语言特征进行分析，说明了这些语言特征在反讽使用和理解中的效果。接下来介绍了实验语料情况，然后通过卡方统计量选取相应特征词。本章最后从计算机识别的角度对反讽小类进行划分，希望对未来相关研究的任务细化和实验分析有参考价值。

2.1 中文社交媒体反讽的语言特征

语言学家和心理学家们对反讽理论不遗余力地探索说明反讽不仅仅是一种简单的修辞现象，对它的本质、认知过程以及区分机制的描写和解释也并非一种理论就能达到完美的地步。本文的目标是面向计算机的反讽自动识别，我们希望合理利用各种理论的精华部分，在一个简化而明确的综合理论框架下进行研究。对于反讽本质，不同理论的共同认识是：反讽包含着某种不一致。我们以此为基础，将反讽描述为通过非负面的字面义表达负面的隐含义的修辞方式。相比隐喻、转喻等字面义和非字面义也存在差异的修辞方式，反讽在认知过程中强调这两个层级意义的反差和分化，而隐喻和转喻分别是两个层级基于相似性和相关性的意义扩展。反讽中字面义和非字面义之间的不一致最典型的表现就是正反对立，并试图通过这种反差凸显隐含义来表达批评、讽刺等强烈的负面情感。因此，“反讽”的“反”是形式而“讽”是目的。

本文考虑中文社交媒体语言特点，结合反讽理论研究提炼出以下四种反讽语言的形式特征。

2.1.1 表达负面语义的成分+表达非负面语义的成分

该形式特征必须包含以下两个组成成分：

(a) 表达负面语义的成分 N

(b) 表达非负面语义的成分 P

两个成分的出现顺序没有严格要求，一般是同一句子里的两个小句。两个句子成分可以带有有强烈的语气或情感强度，来进一步提示言者鲜明的主观态度。成分 P 表达的语义通常是正面的，有时也可以是没有明显情感倾向的。言者通过 P 和 N 之间的语义反差来明显地违背“无歧义”的交际准则，达到特殊的表达效果：字面上表达非负面语义的成分 P 具有隐含义，实际也是表达负面语义的。简单来说就是“正话反说”或“明褒暗贬”。例如：

(s1)很好，我语文作业又写错了。

(s2)真山真水拍成了假山假水，这导演太厉害了！

加单下划线的是成分 N，加双下划线的是成分 P。程度副词“很”和“太”分别修饰正向情感词“好”和“厉害”，两句的成分 P 都表达强烈的肯定意味，通过与各自成分 N 的对比，可以发现“很好”的隐含义是“很不好”，“太厉害”的隐含义是“太拙劣”，具有反讽意味。

2.1.2 提示负面情感的成分+表达非负面语义的成分

该形式特征必须包含以下两个组成成分：

(a) 提示负面情感的成分 H

(b) 表达非负面语义的成分 P

两个成分出现顺序没有严格要求，一般来说 P 是句子里的一个小句, H 是同一句子里的词或词组。成分 H 具有提示负面情感的作用，能够向听者提示言者对待某一事件或对象的基本态度，如果句子上下文没有明显表达与之相匹配的负面语义的成分，同样可能出现交际意图的模糊不明，需要听者联系 H 来推导成分 P 隐含的负面语义。例如：

(s3) 废青没口罩活不下去，我还以为他们没脸皮可以省口罩。

(s4) 少看点网络上的公知言论吧，他们脑子里跟塞驴毛一样。

加单下划线的是成分 H, 加双下划线的是成分 P。“废青”和“公知”两个词由于本身经常与批评、讽刺等负面评价相联系，因此具有提示负面情感的作用。相应的，s3 和 s4 中成分 H 实际表达的隐含义分别是“废青没有羞耻心”和“公知的思想、理念毫无价值，不值得相信”，具有反讽意味。

2.1.3 提示非字面义的成分+表达非负面语义的成分

该形式特征必须包含以下两个组成成分：

(a) 提示非字面的成分 T

(b) 表达非负面语义的成分 P

成分 T 是能够提示成分 P 具有某种非字面义的显式标记，既可以是一些特定的标点符号也可以是词。成分 P 具有的非字面义通常是负面的。有些情况下，成分 P 通过负面隐含义表达反讽的用法已经比较固定，可以视为 T 和 P 一体化。例如：

(s5) 一个引导全球经济超百年的金融桂冠不如各位键盘选手眼界高，思路广……

(s6) 别忘了这些战疫英雄：“居里夫人”王某、“通稿复读机”某君、“好院长”蔡某。

(s7) 难怪那么多人喜欢当键盘侠呢，躲在屏幕后面喷人可真爽！

加单下划线的是成分 T, 加波浪线的是成分 P。s5 中省略号表示意在言外，让读者进一步踹度言者表达的隐含义“网络评论者并非真的比金融桂冠眼界高、思路广（而是只会讲大话空话）”。s6 中对部分词语加上引号来暗示另一层含义，如“居里夫人”、“好院长”实际指“王某”和“蔡某”“名不副实（工作水平遭人诟病）”，“通稿复读机”实际指“某君”“像复读机一样无意义地重复内容（以逃避责任）”。s7 中“键盘侠”字面义表示“某种与键盘相关的英雄”，但实际含义与“侠客”、“英雄”等相关的正面情感词相悖，指“现实生活中胆小怕事，在网上占据道德高点肆意发表不负责任言论的一类人”，这种用法已经固定，不需要其他提示听者就能自然地理解该词的反讽意味。

2.1.4 具有语义或逻辑矛盾的成分+表达非负面语义的成分

该形式特征包含以下两个组成成分：

(a) 具有语义或逻辑矛盾的成分 C

(b) 表达非负面语义的成分 P

由于具有语义或逻辑矛盾的成分 C 本身已经涵盖比较丰富的信息，有时也可以不通过

成分 P 而单独地表达负面隐含义，故该形式特征中只有成分 C 是必须的。例如：

(s8) 下午两点半关门，领事馆你可以再懒一点！

(s9) 乱收费的最高境界就是被自愿。

加双下划线的是成分 C。s8 中“可以再懒一点”，表达基于“领事馆懒”这个事实对更坏事实的期待，正常情况下人们都应该期待更好的变化，因此在逻辑上存在矛盾。实际上，这个成分的隐含义是与所要比较的基准相关——“领事馆不可以这么懒（太懒了）”。s9 中“被自愿”并不符合标准的汉语习惯，存在语义上的矛盾（本文在 2.1.5 小节中“被+X”构式部分进行详细说明），其隐含义是“（某种力量）令人们强制性地接受乱收费现象”，表达了对不合理社会问题的讽刺。

2.1.5 反讽相关的具体语言特征

形式特征的实现依赖于语言中更具体的特征，这些具体的语言特征与反讽具有很强的相关性。下文将进一步分析归纳反讽相关的具体语言特征。

(1)skip-n 元词组合。对于一个分词后长度为n的句子，可以构造n^2个组合词语，即“skip-n 元词”。某些共现词组可以连接句法成分，表示递进、转折等特定句法关系，对启发听者理解句子隐含义有提示作用。例如：

(s1) 很好，我语文作业又写错了。

(s10) 很好，不差钱的尼克斯又当了回人傻钱多的扶贫者。

(s8) 下午两点半关门，领事馆你可以再懒一点！

(s11) 拜托，公司的网络可以再慢一点吗？

“…很好…又…”这一固定组合经常出现，“很好”先作为表达正面语义的成分伪装出一种积极、肯定的态度，“又”连接的命题在语义上同这种态度具有转折和冲突，如 s1、s10 分别是表达负面语义的成分“我写错了语文作业”、“球队尼克斯运营不佳”，都是“不好”、“糟糕”的。听者通过语义上的对立或悖逆，可以识别负面的隐含义，如上例中“很好”实际都是表达“糟透了”的隐含义。

在反讽中，“…可以再…”这一固定组合经常出现，后面通常跟随表示负面情感的形容词。“再”后接形容词表示人或事物在性质、状态或属性等方面递升的变化关系，“可以”表达了言者对这种递进变化的许可。这种递进关系首先基于已经存在的基准，在 s8、s11 中分别是“领事馆懒”、“公司的网络慢”的负面现实，进而在字面上表达对更坏事实的期待，不符合常理，于是读者可以推导隐含义分别是“领事馆不可以这么懒”、“公司网络不可以这么慢”，并领会言者对负面现实的不满和讽刺情绪。

(2)标记强烈情感强度的副词。中文反讽中副词使用频率较高，主要是修饰情感强度的程度副词和表达强烈语气的语气副词。例如：

(s2) 真山真水拍成了假山假水，这导演太厉害了！

(s12) 这帮人其他事不会干，贩卖焦虑最在行！

(s13) 你可真是太聪明了，连这道题都做不来。

(s14) 这政策解读起来简直就是魔幻现实主义。

在中文社交媒体反讽中，程度副词修饰具有正面评价意义的情感词能表达强烈的情感倾向，与句子前后文成分表达的负面语义形成更鲜明的对比。如 s2、s12、s13 中的程度副词分别修饰“厉害”、“在行”、“聪明”三个正向情感词，字面义上均是言者对对象的某种强烈肯定，但上下文的信息分别是“导演拍摄水平拙劣”、“这帮人不会干正事”、“你做不来这道题”，

与强烈肯定的部分形成语义上的巨大反差。为了消除信息分歧、明确言者真实的交际意图，听者需要对比推理，发现字面上表示强烈肯定的部分隐含言者实际想表达的负面评价，即反讽中“正话反说”的效果。

s13、s14 中的语气副词“真是”、“简直”修饰比情感词更大的单位，使得句子整体带有一种强烈的语气，表达言者的某种主观态度。s13 中由于字面义“你很聪明”是很明显的肯定，和后文的负面信息对比可以推知言者隐含的、真实的主观态度是表达“你很不聪明”的批评。s14 中字面义“这政策解读起来就是魔幻现实主义（一种艺术手法）”，我们一般不会用艺术手法来形容政策，因此这种表达存在逻辑上的矛盾，言者需要挖掘隐含义。由于“魔幻现实主义”这种艺术手法具有夸张、荒诞的特点，所以隐含义是“这政策很不合理”。Clift 指出，语义焦点的存在性是反讽的必要不充分条件（Clift 1999）。中文环境下，语义焦点可以用音调标识，也可以通过焦点敏感算子来标识，后者在计算语言学中是一种常见的发现语义焦点的方法。从语用学的角度，标记强烈情感强度的副词在反讽中还充当着焦点敏感算子的角色，保证了语义焦点的存在这一反讽的必要条件。

(3)“被+X”构式。在规范的汉语书面语环境下，人们熟悉的介词“被”的被动用法是“被”后面接动作实施者再接及物动词，通常这里的动作实施者可以省略，构成常规的“被+VP”结构。在中文社交媒体语言环境下，“被”后面可以接不及物动词、形容词名词等，这些突破常规的用法能够表达丰富、深刻的内容，与反讽密切相关。例如：

(s9) 乱收费的最高境界就是被自愿。

(s15) 领导的“爱心表演”就是员工被慈善。

(s16) 每年看统计数据觉得自己生活水平很高，结果还不是被平均。

不考虑“被”字的情况下，“被+X”构式中的 X 在不同程度上可以由动作发出者进行控制，从语义特征上来说是[+可控]的（王俊平 2011）。如 s9、s15、s16 中，“自愿”完全是一种自我实施行为；“慈善”是进行慈善行为的人可以部分自控的；“平均”虽然更多地受外界因素影响，但动作发出者也是可以施加影响的，任何人都希望自己的生活水平高，可以通过个人努力、生活态度等因素对所期待的生活有所控制。而在常规的“被+VP”结构中，受事针对 VP 的发生是不可控的（戴耀晶 2009），即[−可控]。因此，“被”的存在使得“被+X”在语义特征上形成一种可控与不可控的冲突，这种冲突反映在另一层面上也表现为施事和受事同为一体的冲突。以 s10 为例说明，“员工”是进行“慈善”行为的发起者，即施事，可以控制是否参与捐款、捐款金额等；但考虑“员工”和“被慈善”，听者按常规被动用法分析出字面义——“员工”进行捐款不是由“员工”自身控制的，“员工”是行为的受事。根据语义特征和语义角色两种矛盾再结合上下文信息，听者可以进一步推导蕴含义：“领导”作为隐含的、实际意义上的施事力量出现，强制地令“慈善捐款”这一 “员工”本来部分可控的行为发生。“被+X”构式凝练地反映了不合理的社会问题，通过语义上的矛盾凸显了弱势群体对强势力量的无奈以及对不合理社会问题的讽刺，具有反讽意味。

(4)特定的标点符号。主要指感叹号、问号、省略号和引号。例如：

(s5) 一个引导全球经济超百年的金融桂冠不如各位键盘选手眼界高，思路广……

(s6) 别忘了这些战疫英雄：“居里夫人”王某、“通稿复读机”某君、“好院长”蔡某。

(s17) 呵呵，你们以为这可能是首次造假吗？

中文社交媒体反讽中经常用带有强烈语气的句子表达言者某种主观态度，s2、s8、s12 中的感叹号和 s17 的问号都起到加强语气的作用。s17 的反问具有强烈感情，因为反问是无疑而问的，反问传递的确定信息就是反讽所要表达的隐含义“这不可能是首次造假”。

(5)特定的网络词汇。社交媒体语言具有凝练化、口语化、创意化的特点，形成了大量以词汇为主体形式的新兴网络用语。一些网络词汇因为经常与反讽联系，可以起到提示反讽的作用，甚至有些网络词汇本身就已经具有负面的隐含义。例如：

(s3) 废青没口罩活不下去，我还以为他们没脸皮可以省口罩。

(s4) 少看点网络上的公知言论吧，他们脑子里跟塞驴毛一样。

(s7) 难怪那么多人喜欢当键盘侠呢，躲在屏幕后面喷人可真爽！

(s18) 一个无视规则闯红灯的人最后竟然得到这种和稀泥的处理，我醉了……

“废青”指一些无理想、不奋斗，对社会没有贡献但自认为有成熟独立思想，并将自身经历的不满盲目归咎于社会、报复社会的年轻人。该词本身就具有负面评价义，大量用于针对近年来进行非法活动扰乱中国香港秩序的年轻人的批评、谴责评论中。在 s3 中，后一小句的字面义是言者以为“废青没有面部故不需要戴口罩”，没有明显体现出与“废青”相匹配的否定意味，但联系“废青”的情感倾向和“脸皮 - 情面；面子 - 羞耻心”这种被人们广泛使用并固化的非字面义延申模式，不难理解隐含义是“废青没有羞耻心”。

“公知”是“公共知识分子”的缩略词，精确定义是“具有学术背景和专业素质的知识者”，单从词义上看该词是中性甚至带有一定褒义的。联系中文社交媒体语言，该词与表达负面情感的语言环境联系紧密，经常用来表达对中国社会实践和公共事务中一些貌似公正博学，实际摇摆不定、自视甚高的知识分子的讥讽。如 s16 中后面成分实际表达了“公知的思想、理念毫无价值，不值得相信”，具有反讽意味。

还有一部分网络词汇除了常用在表达负面评价义的语言环境中外，本身已经具有负面的隐含义，因此不需提示成分就能实现反讽。如 s7 的“键盘侠”和 s18 的“醉了”，后者字面义是“饮酒过量、神志不清”，实际表达言者对不合理现象的无法理解和鄙夷。

2.2 实验数据集建立

目前对于中文反讽的研究较少，也缺乏公开的、权威的反讽语料库，因此需要人工标注。本研究在体育、电影、新闻、娱乐等领域，收集了 1 万条微博数据，并进行了人工标注，共标注得到 19268 个带有情感的句子，其中含有反讽的句子有 1291 个，占带有情感句子总数的 6.58%，说明在中文社交媒体中，反讽的确是一种值得关注的语言现象。

标注过程中，以隐形展示理论（Utsumi 2000）为区分反语的理论指导，如果构成反讽的前提条件——反讽环境存在，那么认为句子是包含反讽的。以 s18 为例，反语发出者期望“无视规则闯红灯的人”受到公正严肃的处罚，而违规者仅仅受到了“和稀泥的处理”，言者的期望落空，所以有理由认为言者对这种不公正现象感到失望和不认同，这种态度具体表现在“醉了”的负面隐含义中，反语环境三要素得以实现。在标签设置上，将反讽标注作为一个二分类问题，若为反讽则标注标签为 1，反之则为 0。情感标注完全通过人工判定，将句子情感类别标注为积极、消极或中性。为了减少人工标注中主观因素造成的判定偏差，在标注过程中采取交叉检验，对于不一致的标注观点开展讨论，统一认识。对数据类别进行平衡化处理，从非反讽数据集中抽取 1291 条句子，使得中文反讽数据集的正负样本比例为 1:1。

2.3 卡方统计量

在文本分类中，特征选取的优劣对于分类性能有直接影响，其思想是从原有特征集合中选出更有代表性的子集，主要方法包括：卡方统计量、信息熵、逻辑回归等。

卡方统计量是一种经典而有效的特征选取方法，它首先假设特征和类别间是相互独立的，再根据观测值和期望值的偏差计算结果选择是否要否定原假设，以此判定特征和类别的相关性。卡方统计量越大，说明特征和类别间相关性越大，即该特征对类别识别的贡献越大，可

以作为类别特征。偏差的基本定义如公式 2.1 所示：

$$\chi^2 = \sum_{i=1}^{k} \frac{(A_i - E_i)^2}{E_i} \tag{2.1}$$

式中$\boldsymbol{A_i}$为观测值，$\boldsymbol{E_i}$为理论值。特别地，在特征选取中特征词$\boldsymbol{t}$和文本类别$\boldsymbol{c_i}$相互独立的假设下，二者的关系表如表 2.1 所示：

表 2.1　特征词$\boldsymbol{t}$和文本类别$\boldsymbol{c_i}$关系表

	属于类别$\boldsymbol{c_i}$	不属于类别$\boldsymbol{c_i}$	总数
包含$\boldsymbol{t}$的文本数	A	B	A+B
不包含$\boldsymbol{t}$的文本数	C	D	C+D
总数	A+C	B+D	N=A+B+C+D

根据表 2.1，可将公式 2.1 具体代入并化简得到特征词$\boldsymbol{t}$和文本类别$\boldsymbol{c_i}$的卡方统计量，如公式 2.2 所示：

$$\chi^2(t, c_i) = \frac{\mathrm{N}\times(\mathrm{A}\times\mathrm{D} - \mathrm{C}\times\mathrm{B})^2}{(\mathrm{A}+\mathrm{C})\times(\mathrm{B}+\mathrm{D})\times(\mathrm{A}+\mathrm{B})\times(\mathrm{C}+\mathrm{D})} \tag{2.2}$$

下文中 skip-n 元词组合、标记情感强度的副词和特定网络词汇均采用卡方统计量选取。

2.4 语言特征选取

(1)skip-n 元词组合

对实验数据集的所有二元词组合进行χ^2统计，再人工选择卡方统计值较高的前 15 个词语组合作为反讽的 skip-n 元词组合。解析过程见图 2.1，部分统计结果见表 2.2。

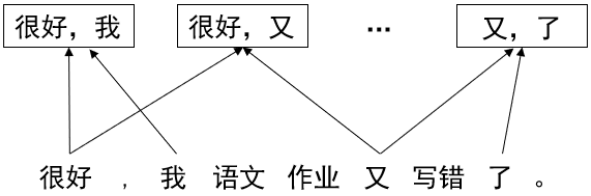

图 2.1　skip-n 元词组合解析示意图

表 2.2　skip-n 元词组合统计值

编号	skip-n 元词组合	卡方统计值
1	…这…很…	3.546
2	…连…都…	3.448
3	…很好…又…	3.258
4	…可以…再…	2.792
5	…又…真是…	2.787

查找卡方分布表，当自由度为 1，χ^2值为 2.71 时，对应的 p 值为 0.1。表 2.2 中χ^2值均大于 2.71，则相应 p 值都小于 0.1，则 skip-n 元词组合与反讽相互独立的概率小于 10%，因此二者具有较高的相关性。

(2)标记情感强度的副词和特定的网络词汇

对实验语料进行去停用词后计算所有词的χ^2值，再人工分别选取前 15 个标记情感强度的副词和特定的网络词汇，有如下结果：

标记情感强度的副词：很、真是、太、有点、挺、完全、非常、那么、超、过于、最、

满、却、反倒、偏偏。

特定的网络词汇：键盘侠、毒鸡汤、五毛、公知、尼玛、作死、废青、醉了、人设、圣母、战狼、沙雕、奇葩、bb、河蟹。

2.5 面向计算机识别的反讽小类划分

过去对反讽小类的划分主要是从本体研究的角度进行的。我们尝试从计算机识别的角度对反讽小类进行划分，希望对未来相关工作有所帮助，如表 2.3 所示。

表 2.3　面向计算机识别的反讽小类划分

反讽	上下文相关反讽	超文本相关的反讽	他那张嘴，靠谱到能把北极熊说成南极物种。
		文本相关的反讽	傻仔去罢课，我先去上课。
			毕竟猪不能抬头看天空。
	上下文无关反讽	显式反讽	你可真是太聪明了，连这道题都做不来。
		隐式反讽	陈一冰用金牌的动作拿到了银牌。
			把个人项目当团体项目完成，是我国的传统。

计算机反讽识别的输入主要是含有反讽的文本。按反讽分析是否依赖输入文本以外的信息，首先可以分为两类，本文分别定义为上下文相关的反讽和上下文无关的反讽。我们人工挑选出两类反讽句子各 20 个输入 IEAN 进行识别，模型成功识别出了 35%的上下文相关反讽句子和 75%的上下文无关反讽句子。这与我们的预期相符，因为同大多数已有的反讽识别模型一样，本文提出的模型没有考虑额外的上下文信息，实质上是上下文无关的模型。若不考虑通过特殊形式的数据构建和网络设计来引入反讽相关的上下文信息，上下文相关的反讽识别显然更加困难。

针对上下文相关的反讽，还可以根据额外输入信息的不同形式分为文本信息相关的反讽和超文本信息相关的反讽，其中，前者的反讽分析依赖于文本形式的信息，即狭义的上下文；后者的反讽分析则需要借助表情、语调、知识图示等非文本形式的信息，前文提到的 Devamanyu Hazarika 等的研究（Devamanyu et al. 2018）就是针对此类反讽进行建模。

针对上下文无关的反讽，可以根据是否存在明显的反讽形式标记分为显式反讽和隐式反讽。这种形式标记是文本明面上出现的能够提示或直接表达反讽意味的成分，包括但不限于本文归纳的语言特征和句子成分表达的强烈情感倾向等。考虑表中对应句子。“用金牌的动作”的预期是“拿金牌”，现实结果是“拿到了银牌”。从这种预期和现实的反差可以推导该句的隐含义是“本来配得上金牌的陈一冰受到了不公正的评判”。“个人项目”理应是独立完成的，在语义特征上是[−合作]，而“团体项目”是合作完成的，在语义特征上是[+合作的]，按句子表述存在语义矛盾，实际表达隐含义“采取以多对少的不当竞争方式是我国一项陋习”。这两句都属于隐式反讽一类。

我们人工挑选出显式和隐式两类反讽句子各 20 个输入 IEAN 进行识别，模型成功识别出了 85%的显式反讽句子和 50%的隐式反讽句子。隐式反讽缺乏明显的标记提示，其非字面义的表示也往往未成固定模式，具有临时性，需要在把握句子整体语义的基础上进行深入分析才能辨识，相比显式反讽不仅对计算机识别更具挑战性，对人正确理解其中的反讽也有不小难度。

这个划分方案仍然是初步的，不可避免地存在一些不足，比如：对上下文相关的反讽采取进一步划分时，根据的是相关信息的具体形式。严格来说这样的划分是侧重于任务特点的划分，目前研究尚不能对计算机识别这两类反讽的难度差异进行直接比较。

第三章 融合语言特征的注意力机制的中文反讽识别模型

传统 CNN 模型从连续的 N-gram 向量矩阵中获取局部特征，存在未考虑文本的时序信息、无法解决长距离依赖问题、可解释性相当有限等缺陷。普通的 RNN 模型能够记录输入的历史信息，根据当前信息对输出信息进行预测从而解决序列依赖的问题。但随着序列增加，RNN 模型会产生梯度消失和梯度爆炸等问题。Hochreiter 等提出了长短期记忆网络模型（LSTM）（Hochreiter et al. 1997），通过对网络隐层的门控制设计，一定程度上解决了梯度消失问题，在自然语言处理的许多任务上得到了广泛使用。注意力机制(Attention)最早是在视觉图像领域提出，核心思想是仿照人类的选择性视觉注意力让模型从众多信息中选取对当前任务目标更关键的信息。Bahdanau 等学者首次将 Attention 机制应用到自然语言处理领域的机器翻译任务中（Bahdanau et al. 2014），提升了模型效果的同时实现了文本的可视化对齐，这一特点对于理解模型工作过程有所帮助。自此类似的结合 Attention 机制的 RNN、CNN 模型逐渐广泛地应用到自然语言处理中的各种任务中。

本文以 LSTM 为基础，引入一种融合语言特征的注意力机制来搭建模型。本文进行了一系列实验来说明新模型在性能和可解释性上的提升，并对存在的一些问题进行了分析、探讨。

3.1 词嵌入向量（Word Embedding）

在自然语言处理领域，将文本转换成向量形式表示，可以作为特定任务算法执行的基础。文本向量化的方法可以分为独热编码（one-hot）和词嵌入表示（word embedding）两种。One-hot 编码规定向量矩阵每一行有且只有一个元素为 1，其余元素均为 0。这种表示方式直观，但构造的词向量往往比较稀疏、浪费空间，而且完全无法体现单词之间的关系。Hinton 率先提出了 word embedding 的思想和表示方法（Hinton 1986），通过线性映射得到了上下文信息在低维向量隐含空间上的表达。Mikolov 等提出了 word2vec 模型来训练特定领域的 word embedding 向量（Mikolov et al. 2013），其中的 Skip-gram 模型通过目标词的词向量来预测上下文词汇的词向量。假设某一词组序列为$w_1, w_2, ..., w_N$，模型目标是最大化公式 3.1 的值。

$$E = \frac{1}{N}\sum_{n=1}^{N}\sum_{-a\le i\le a, i\neq 0} -\log p(w_{n+i}|w_n) \tag{3.1}$$

其中，a是以当前词为中心的窗口大小，表示选取当前词w_n前a个词和后a个词。$p(w_{n+i}|w_n)$表示词w_{n+i}在词w_n已经出现条件下出现的概率，在 Skip-gram 模型中一般按公式 3.2 计算。

$$p(w_{n+i}|w_n) = \frac{\exp\left(v'^{T}_{w_{n+i}} v_{w_n}\right)}{\sum_{t=1}^{|V|}\exp\left(v'^{T}_{t} v_{w_n}\right)} \tag{3.2}$$

其中，v_t和v'_t分别是词w_t的输入和输出向量，|V|是词典大小。语料库足够大时，选择合适的窗口大小，Skip-gram 模型能够快速训练得到高质量的词向量。

3.2 融合语言特征的注意力机制的中文反讽识别模型

为了将反讽本身的语言特征和深度学习方法进行结合，本文构建了一个新的反讽识别

模型。在第二章对语言特征进行分析和特征词提取的基础上，使用词向量建立文本矩阵和特征矩阵作为模型输入。由注意力机制层对句子进行反讽特征强化得到特征注意力表示，在连接层与句子成分表示进行拼接以拓展句子语义，最后将拼接后的向量输入 softmax 分类器来实现对句子的反讽识别。模型的整体框架如图 3.1。

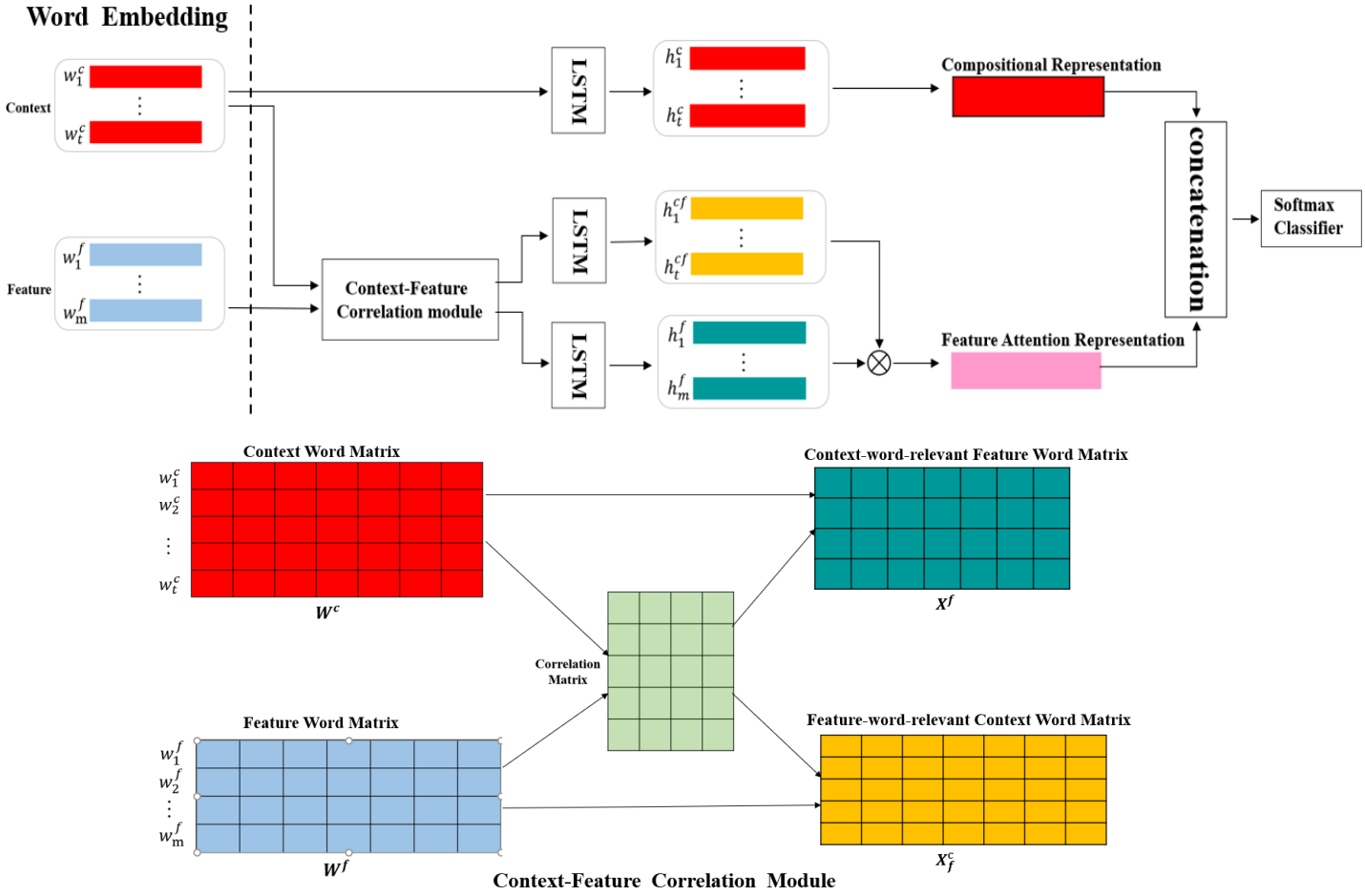

图 3.1　融合语言特征的注意力机制的反讽识别模型框架

(1)输入层

模型的输入层含两个词向量矩阵：句子矩阵（Context Word Matrix）和特征矩阵（Feature Word Matrix）。通过北京大学开源分词工具 pkuseg[1]对数据集的每个句子 S 进行分词，假设句子 S 含有 t 个词，结果按词序表示如式 3.3 所示。

$$S:\{w_1, w_2, ..., w_t\} \tag{3.3}$$

从训练好的词向量中获取词w_i的向量表示，假设向量维度为 d，w_i表示如式 3.4 所示。

$$w_i = (c_{i1}, c_{i2}, ..., c_{id}) \quad i \in [1, t] \tag{3.4}$$

S 的句子矩阵表示是将 S 的每个词向量按词序排列，如式 3.5 所示。

$$S = w_1 \oplus w_2 \oplus ... \oplus w_t \tag{3.5}$$

因此句子 S 被转化成对应的向量矩阵W^c如式 3.6 所示。

$$W^c = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1d} \\ c_{21} & c_{22} & \cdots & c_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ c_{t1} & c_{t2} & \cdots & c_{td} \end{bmatrix} \in \mathbb{R}^{t \times k} \tag{3.6}$$

同理，设提取的特征词数量为 m，特征矩阵$\boldsymbol{W^f}$如式 3.7 所示。

[1] https://github.com/lancopku/PKUSeg-python

$$W^f = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1d} \\ f_{21} & f_{22} & \cdots & f_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ f_{m1} & f_{m2} & \cdots & f_{md} \end{bmatrix} \in \mathbb{R}^{\boldsymbol{m\times k}} \tag{3.7}$$

(2) 融合语言特征的注意力模块

本文将选取的反讽语言特征作为关注句子上下文词的注意来源，以此形成反讽特征增强的句子表示。

本模块的核心是文本−特征相关性模块（Context-Feature Correlation Module）。受到 Xiong 等研究（Xiong et al. 2017）的启发，本文希望在上下文词和特征词之间建立关系。具体的，本文首先通过点积定义句子上下文词和特征词的相关矩阵M^f如式 3.8，表示特征词与上下文词的相关程度。

$$M^f = (W^c)^T \cdot W^f \in \mathbb{R}^{m\times t} \tag{3.8}$$

定义文本相关的特征表示矩阵X^f如式 3.9，代表特征词关于句子的向量表示。

$$X^f = M^f \cdot W^c \in \mathbb{R}^{m\times d} \tag{3.9}$$

定义特征相关的句子表示矩阵X_f^c如式 3.10，代表句子关于特征词的向量表示。

$$X_f^c = (M^f)^T \cdot W^f \in \mathbb{R}^{t\times d} \tag{3.10}$$

接下来，通过三个独立的 LSTM 层编码隐藏状态矩阵H^c, H^f, H_f^c 如式 3.11-3.13。

$$H^c = LSTM(W^c) \tag{3.11}$$

$$H^f = LSTM(X^f) \tag{3.12}$$

$$H_f^c = LSTM(X_f^c) \tag{3.13}$$

在得到隐藏状态矩阵后，反讽特征增强的句子语义表示r_a计算如公式 3.14-3.17。

$$r_a = \sum_{i=1}^{t} a_i h_{if}^c \tag{3.14}$$

$$q_f = \sum_{i=1}^{m} h_i^f / m \tag{3.15}$$

$$s\left(\left[h_{if}^c; q_f\right]\right) = u_s^T \tanh\left(W_s[h_{if}^c; q_f]\right) \tag{3.16}$$

$$\alpha_i = \frac{\exp\left(s\left(\left[h_{if}^c; q_f\right]\right)\right)}{\sum_{i=1}^{t} \exp\left(s\left(\left[h_{if}^c; q_f\right]\right)\right)} \tag{3.17}$$

其中，q_f是对H^f的平均池化以减少参数和防止过拟合。[;]是连接操作。s是计算句子第i个词对于任务重要程度的注意力函数。将s的计算结果经过 softmax 归一化获取句中各词的注意力权重α_i。最终加权平均输出反讽特征增强的句子语义表示r_a。u_s^T和W_s是需要学习的参数。

(3) 全连接层

将普通句子序列化建模表示h^c和对于反讽特征的句内 Attention 表示r_a进行拼接，得到模型下游的句子表示h作为分类器的输入，如式 3.18。

$$h = h^c \oplus r_a \tag{3.18}$$

(4) softmax 分类层

将模型下游的句子表示h输入一个 softmax 层来预测句子的反讽类别标签分布，如公式 3.19。

$$\hat{y} = \frac{\exp\left(\tilde{w}_o^T h + \tilde{b}_o\right)}{\sum_{i=1}^{C} \exp\left(\tilde{w}_o^T h + \tilde{b}_o\right)} \tag{3.19}$$

其中$\hat{y}$是预测的句子反讽类别标签分布，C 是标签的类别数量，$\tilde{w}_o^T$和$\tilde{b}_o$是需要学习的参数。在模型训练过程中，我们的目标是最小化真值类别标签和预测类别标签的交叉熵。同时，为了防止过拟合，我们采用 dropout 策略在每个训练样例中对部分参数进行了随机

省略。

3.3 模型参数设置和实现细节

数据集中最长的句子包含 72 个词，模型输入的最大序列长度（max sequence length）设置为 100，对不足长度的部分用 0 补齐。Word Embedding 维度为 300，由于实验数据集是面向中文社交媒体的语料，预训练的词向量[2]为中文微博上采用 3.1 介绍方法训练得到。为了更好地计算词语相关性，对于输入层的词向量矩阵进行了规范化处理。模型训练中目标函数为交叉熵损失函数，对预测结果采用 L2 正则化项防止过拟合，正则化系数λ为10^{-8}，学习率（learning rate）为10^{-3}，遗忘率（drop out）为 0.5。LSTM 的隐层规模（hidden size）为 256，实验采用 5 次交叉验证，迭代次数（epoch）为 20，批处理大小（batch_size）为 32，更新模型参数采用 Kingma 等提出的 Adam Optimization（Kingma et al. 2015）梯度下降的方法。

3.4 实验结果和分析

本文通过以下几组实验对本文方法的有效性进行验证。

(1)选取语言特征的有效性验证

本实验对比了词袋模型（BOW）和结合反讽语言特征的词袋模型（BOW+Feature）两种语言建模在不同分类器上的效果。分类器采用朴素贝叶斯（NB），支持向量机（SVM）和随机森林（RF）。实验设置如表 3.1 所示，实验结果如表 3.2 所示。

表 3.1 第一组实验设置

方法	说明
BOW	单独使用词袋的文本表示
BOW+Feature	特征和词袋结合的文本表示，按照χ^2统计量大小选取前 2000 维

表 3.2 第一组实验结果

方法	分类器	精确率	召回率	F 值
BOW	NB	0.6971	0.6682	0.6823
	SVM	0.75	0.7058	0.7272
	RF	0.8125	0.6259	0.7071
BOW+Feature	NB	0.7786	0.6994	0.7368
	SVM	0.8182	**0.7578**	**0.7868**
	RF	**0.8361**	0.693	0.7578

结合反讽语言特征的词袋模型在精确率、召回率和 F 值上都有比较明显的提升，说明人工选取的语言特征是有效的。传统的词袋模型仅仅反映了部分句子上下文的信息，由于反讽表达现象本身的复杂性，单纯地通过缺乏词序的词语共现信息难以识别反讽。选取的反讽语言特征是与反讽具有强相关性的词语，对于反讽识别有显著的提示作用，这在一定程度上缓解了上述问题。

(2)融合语言特征的注意力机制的模型有效性验证

[2] https://github.com/Embedding/Chinese-Word-Vectors

为了全面评估本文模型的性能，我们列出了几个句子级反讽识别模型作为基准。

LSTM/Bi-LSTM: Cho 等利用长短期记忆和双向变化的网络模型（Cho et al. 2014）来捕获句子的序列信息。

CNN: Kim 提出的卷积神经网络（Kim 2014）通过从连续的 N-gram 向量中提取局部特征来生成特定任务的句子表示。

Self-Attention: Lin 等提出一种自注意力机制来学习结构化的句子表示（Lin et al. 2017）。Yi Tay 等应用 Intra-Attention 提取单词对间的信息的模型实际上可以视为 Self-Attention 的变种（Yi Tay et al. 2016）。本文的 Self-Attention 参照两个相关模型，在 LSTM 基础上实现。

各模型的主要参数如表 3.3 所示。

表 3.3 神经网络模型主要参数设置

	词向量维度	LSTM 隐层规模	CNN 卷积核	CNN 卷积窗口	CNN 采样窗口	optimizer	epoch
LSTM Bi-LSTM	300	256	——	——	——	Adam	20
CNN	300	——	100	3*3	3*3	SGD	20
Self-Attention	300	256	——	——	——	Adam	20

表 3.4 第二组实验结果

编号	方法	精确率	召回率	F 值
1	LSTM	0.7679	0.7396	0.7535
2	Bi_LSTM	0.7858	0.7656	0.7756
3	Self_Attention	0.8125	0.7864	0.7993
4	CNN	0.8404	0.7907	0.8148
5	IEAN(our model)	**0.8527**	**0.8269**	**0.8390**
6	IEAN w/o pretrained wv	0.7806	0.7380	0.7587

由表 3.4 发现，LSTM 方法的结果最差，对比前面传统的机器学习方法并没有体现出深度学习方法的优势。LSTM 作为一种经过多种任务检验、被公认具有较强学习能力的网络模型，非常依赖训练数据。由于本文人工标注构建的反讽语料集规模相对较小，并不能让 LSTM 的学习能力得到充分发挥。对比 1 和 2，在拼接句子的前后文信息后，双向 LSTM 的效果有一定提升（F≈+2.2%），表明语序信息在反讽语言建模中的重要性；对比 2 和 3，加入自注意力机制后模型效果进一步提升（F≈+2.4%），模型通过对比句子内部词对之间的关系，动态地关注哪些词更有利于反讽句子的识别；CNN 则一向在分类问题上表现不错，通过句子矩阵的池化表示学习某些对于反讽识别有效的隐式特征权重。

与现有模型相比，本研究提出的模型表现出一定优势，尤其是比基准模型中表现最好的 CNN，在精确率上略有提升，在召回率（R≈+3.6%）和 F 值（F≈+2.4%）两项指标上提升明显，表明结合外部语言知识，利用句子上下文词和反讽特征词之间的交互信息分配权重，有利于获取句子整体语义，从而更有效地识别反讽。此外，本文还考察了词向量的有效性，对比抛弃预训练的词向量而采用随机初始化的词向量的方法（标记为 w/o pretrained wv），采用预训练的词向量使得模型性能有较大提升，这是因为基于领域预训练的词向量能够捕捉词语间的关联关系，更好地刻画语言中的词语分布。这相当于间接引入外部数据，一定程度上缓解了过拟合问题。采用预训练词向量的模型在参数优化上更有效率。对于本研究中数据集规模相对较小的问题有所改善。

(3)基于注意力矩阵可视化的模型分析

在本组实验中，本文输出注意力矩阵进行展示，并尝试解释模型表征是如何形成。我

们从 Self-Attention 和 IEAN 两种模型中提取注意力矩阵进行正则化处理，可视化结果如表 3.5 所示（颜色越深，权重越大）。

表 3.5 两种模型的注意力矩阵可视化

类别	模型	句子
反讽	Self_Attention	社会 可以 再 肮脏 一点 吗 ？
	IEAN	社会 可以 再 肮脏 一点 吗 ？
	Self_Attention	公知 们 为了 他们 的 洋爹 真的 够 忍辱负重 了 。
	IEAN	公知 们 为了 他们 的 洋爹 真的 够 忍辱负重 了 。
非反讽	Self_Attention	别 把 某些 东西 看 的 太 重要 ， 有的 时候 简单 一点 才 是 真 。
	IEAN	别 把 某些 东西 看 的 太 重要 ， 有的 时候 简单 一点 才 是 真 。
	Self_Attention	前 七 集 真的 是 太 压抑 了 ！
	IEAN	前 七 集 真的 是 太 压抑 了 ！

在反讽类别的第一个示例中，IEAN 将注意力集中在“可以”、“再”以及表示负面评价的情感词“肮脏”上，这与我们在理解情感冲突时借助的关键词相符合。第二个示例也体现出这种一致性。而在非反讽类别的两个示例中，IEAN 对特征词的关注度也较高。Self_Attention 表现出完全不同的注意力分布。尽管在反讽类别的第一个示例，Self_Attention 重点关注的情感词“肮脏”、表示疑问语气的“吗”和问号对于反讽识别有一定帮助，但在第二个示例中，对于“能”、“了”和句号的关注似乎不太容易解释。对于非反讽的两个示例，Self_Attention 会将注意力集中在一些可能对识别没有意义的词上，比如“东西”、“了”。实际上在相当一部分句子中，Self_Attention 的注意力会集中在句末成分上。我们猜想这可能是因为 Self_Attention 中 LSTM 是对不同时间节点的信息进行组合来获取句子表示的，当文本较短时，第 n-1 个和第 n 个隐藏状态的表示可能非常相似，注意力会集中在最后一个或几个隐藏状态表示上。

整体来看，IEAN 对与反讽具有强相关性的一系列特征词的关注十分到位，这与我们从语言特征中获取提示进而辨识反讽隐含义的过程存在相似性，也帮助模型能够更有效地识别反讽。值得注意的是，IEAN 的注意力矩阵可视化采取了事后解释（Lipton 2017）的方式衡量了神经网络模型特定的概念表示对于任务目标的贡献，较传统模型在可解释性上有所进步。

(4)错误分析

我们从五次交叉验证中选取了一次实验结果，并以此为例对识别错误进行分析。表 3.6 为该次实验的识别结果。

表 3.6 一次实验的识别结果

IEAN		预测	
		1	0
实际	1	222	38
	0	40	220

在预测为反讽，实际为非反讽的 40 条句子中有 14 条至少含有一个本文选取的特征词，例如：

（s19）我 个人 很 喜欢 肖战 在 这 部 戏 里 的 人设 。

这也与我们在第三组实验中观察到的现象一致。和反讽现象具有强相关性的语言特征也能出现在非反讽的文本中，这时集中注意力得到的句子表示对识别非反讽这一目标而言

可能没有帮助甚至有所干扰。

在预测为非反讽，实际为反讽的 38 条句子中，我们发现部分反讽句子的判定严重依赖句外信息，例如：

(s20)他 那 张 嘴 ，靠谱 到 能 把 北极熊 说成 南极 物种 。

(s21)傻仔 去 罢课 ， 我 先 去 上课 。

(s22)毕竟 猪 不能 抬头 看 天空 。

人对于一些反讽句的理解需要一定的背景常识，如 s20 中背景知识是“北极熊不是南极物种”，因为字面义与常识相悖，所以字面上的“靠谱”实际表达“不靠谱”的意思，属于反讽句。但仅靠我们选取的语言特征和相对较小的数据集，模型难以正确理解语义，导致了这类识别错误。

另一类反讽句的识别与整体语境密切相关，如 s21 的微博上文是“这已经是第四位开学的港独头目了。”联系上文，该句字面上从“开学的港独头目”角度直叙对“罢课”者的讥讽，实际表达对祸港青年制造混乱后逃之夭夭的批判。s22 的微博上文是“节目组只在乎眼前热度，欠缺长远考虑，分明是在瞎弄。”联系上文，不难理解“猪不能抬头开天空”的隐含义是前一句的基本字面义，实际表达对“节目组欠缺长远考虑”讽刺。我们将 s21 和 s22 分别与各自的上文拼接后进行测试，其中 s21 能被正确识别为反讽句而 s22 仍旧不能，猜想可能因为 s21 拼接后的文本序列不长，且“开学”与“罢课”、“上课”在词义上具有联系，上文的序列信息有效组合到了最终状态的表示中，有助于模型更好地获取语义表示。

本文选取的语言特征、一条微博的其他句子等，实际上都是同需要识别的文本在长文中具有位置关系的同样形式的信息，即狭义的上下文。广义的上下文还包括说话人的声音、语调、表情、地位，或者社交媒体的转发、回复、评论以及上面提到的背景知识等。因此，原则上加入上下文信息能够使反讽判定更准确。

第四章 结论与展望

针对面向计算的中文反讽识别问题，对本文主要工作得到的结论进行概括并总结不足，在此基础上对未来研究进行展望。

4.1 结论

(1)中文社交媒体的语言特征分析

由于目前研究中权威、公开的中文反讽语料的缺失，本文通过人工标注获取了 1291 条中文反讽语料并以此为基础构建了分布平衡的实验数据集。结合相关研究和中文社交媒体语言特点，本文从概括到具体地对中文反讽的语言特征进行了分析，说明了各种特征与反讽现象本质和认知过程的具体联系。采用χ^2统计方法选取具体的特征词。语言特征有效性验证实验中，使用反讽特征的词袋模型较不使用的词袋模型在精确率、召回率上都有明显提升，同时χ^2统计的绝对数值大于相应阈值，这两点均表明上述具有语言学理论支持的反讽特征也得到了统计意义上的有效性验证。本文还从计算机识别的角度对反讽小类进行划分，新的划分对未来相关工作在任务设置和实验分析上有参考价值。

(2)融合语言特征的注意力机制的中文反讽识别模型

本文采用 Skip-gram 模型训练得到的词嵌入向量，在解决向量稀疏和唯独爆炸问题的同时能够反映词语的分布。考虑到反讽识别目标文本的时序性和非连续依赖问题，本文以 LSTM 为基础，同时引入了一种注意力机制期望模型能结合语言特征更好地实现句子的语义表示。在模型有效性验证实验中，本文提出的模型较基准模型的识别效果有所提升，证明了新模型的有效性。本文展示了模型的可视化注意力分布，对模型内部表示的形成和特定表示对反讽识别的贡献进行了对比说明，再次证明了新模型的有效性和较传统深度学习模型在可解释性上的优势。

4.2 不足与展望

近年来，反讽识别受到了自然语言处理领域的广泛关注并取得了相当丰富的研究成果，但针对中文的相关研究仍然稀缺。本文提出的模型在识别效果和可解释性上得到了一定提升，但囿于研究者的能力和其他客观因素，研究仍存在以下问题和值得努力的方向：

(1)相比英语的反讽识别研究数据集，本文构建的反讽语料规模仍然偏小。尽管花费了相当的精力用于语料收集、整理和人工标注，但仅仅获取了 1291 条反讽语料，这对语言特征分析和深度学习模型训练都有一定限制。实际上大规模、高质量的训练数据才能完全反映 LSTM 的学习能力，有理由相信模型有进一步提升空间。因此，半自动地构建中文反讽语料库是接下来值得重点关注的工作。

(2)本文分析语言现象选取的语言特征主要是词汇层面，尽管这为结合注意力机制带来一定便利，但也因此缺乏对语法等句子深层信息的挖掘和利用。如何进行更深层次的语言特征选择将是一个重要的研究方向。

(3)在识别依赖广义上下文信息的反讽句子时仍然不理想。一方面，用于反讽识别文本材料本身有限，不包含广义的上下文信息；另一方面，人能够通过自身经验或他人言传身教得到的知识对于机器而言是难以有效获取的。特定领域的知识图谱建立将会有力地推动反讽识

别研究。

(4)反讽识别是情感分析的子领域，二者联系紧密。一方面，目前的情感分析任务目标逐渐从文本分类转向信息抽取，这是人们追求计算机具有人类级别语言能力的必然结果。相比之下现有反讽识别研究的任务目标仍然相对单一。对于反讽的言语主体、反讽的对象、反讽的情感强度等反讽结构化信息的抽取也是今后反讽识别工作应该重点关注的任务。另一方面，本文选取的语言特征中有相当一部分能够标记或直接反映情感倾向。本文通过注意力机制融合部分语言特征的做法仅仅是期望结合语言知识和计算模型的简单尝试。人们在对情感分析相当长时间的研究中积累了丰富的语言资源，如何将高质量的语言知识注入到深度学习模型中将是一项有趣且有意义的工作。

参考文献

[1]Aniruddha Ghosh,Tony Veale. Fracking sarcasm using neural network[C]//Proceedings of NAACLHLT, 2016:161-169.

[2] Grice.P.H.Logic and conversation [A] . In Cole, P. & Morgan, J.(Eds)Syntax and Semantics [C] .Vol. 3:Speech Acts:41 - 58.London:Academic Press, 1975.

[3]Bahdanau D, Cho K, Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate[C]//Proceedings of the International Conference on Learning Representations(ICLR), 2014,1-15.

[4]Caiming Xiong, Victor Zhong, Richard Socher. Dynamic coattention networks for question answering[C]//Proceedings of the International Conference on Learning Representations(ICLR), 2017.

[5]Clark H.H&Gerrig R.J.On the pretense theory of irony [J]. Journal of Experimental Psychology:General,113. 1(1984):121-126.

[6]Clift R. Irony in Conversation[J]. Language In Society, 1999, 28(4):523-553.

[7]David Bamman, Noah A Smith. Contextualized sarcasm detection on Twitter[C]//Proceedings of the 9th International AAAI Conference on Web and Social Media, 2015:574-577.

[8]Devamanyu Hazarika, Soujanya Poria, Sruthi Gorantla, Erik Cambria, Roger Zimmermann, Rada Mihalcea CASCADE: Contextual Sarcasm Detection in Online Discussion Forums[C]//Proceedings of the 27th International Conference on Computational Linguistics, 2018.

[9]Dews S&Winner E. Obligatory processing of literal and nonliteral meaning in verbal irony [J]. Journal of Pragmatics, 31(1999):1579-1599.

[10]Edwin Lunando, Ayu Purwarianti. Indonesian social media sentiment analysis with sarcasm detection[C]// Proceedings of the Advanced Computer Science an Information Systems(ICACSIS), 2013 International Conference on IEEE, 2013:195-198.

[11]Giora R. Irony and salience [J]. Metaphor and Symbol, 13(1998):83-101.

[12]Giora R. On irony and negation [J]. Discourse Processes, 19(1995):239-264.

[13]Gonlez-Ibanez R,Muresan S,Wacholder N.Identifying sarcasm in Twitter: A closer look[C]//Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, 2011:581-586.

[14]Hinton G E. Learning distributed representations of concepts[C]//Proceedings of the 8th Annual Conference of the Cognitive Science Society,1986:1-12.

[15]Hochreiter S, Schmidhuber, et al. Long short-term memory[J]. Neural Computation, 1997, 9(8):1735-1780.

[16]Holdcroft D. Irony as a trope, and irony as discourse[J]. Poetics Today, 4. 3 (1983):493-511.

[17]Kingma D P, Ba J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1214.6980v8, 2015.

[18]Konstantin Buschmeier, Philipp Cimiano, Roman Klinger. An impact analysis of features in a classification approach to irony detection in product reviews[C]//Proceedings of the 5th Workshop on Computational Approaches to Subjectivety, Sentiment and Social Analysis. 2014:42-49.

[19]Kyunghyun Cho, Bart van Merrienboer, C¸aglar G¨ulc¸ehre, Fethi Bougares, Holger Schwenk, and

[20]Lotem Peled, Roi Rechart. Sarcasm SIGN: Interpreting sarcasm with sentiment based monolingual machine translation[C]//Proceedings of 55th Annual Meeting of the Association for Computational Linguistics, 2017:1690-1700.

[21]Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space[J]. Computer Science, 2013.
[22]Reyes A, Rosso P, Veale T. A multidimensional approach for detecting irony in Twitter[J]. Language Resources & Evalution, 2013, 47(1):239-268.
[23]Sperber D, Wilson D. Relevance:Communication and Cognition [M]. Oxford: Blackwell, 1986/1995.
[24]Tang Y J, Chen H. Chinese irony corpus construction and ironic corpus construction and ironic structure analysis[C]//Proceedings of the 25th International Conference on Computational Linguistics, 2014:1269-1278.
[25]Utsumi A. A unified theory of irony and its computational formalization[C]//International Conference on Computational Linguistics, 1996: 962-967.
[26]Utsumi A. Verbal irony as implicit display of ironic environment: Distinguishing ironic utterances from nonirony [J]. Journal of Pragmatics, 32 (2000):1777-1806.
[27]Yi Tay, Luu Anh Tuan, Siu Cheung Hui, Jian Su. Reasoning with Sarcasm by Reading In-between[C]//Proceedings of the 56th Annual Meeting of the Association for Computer Linguistics, 2018.
[28]Yoon Kim. Convolutional neural networks for sentence classification[C]. In Proceedings of EMNLP 2014.
[29]Zachary C. Lipton. The mythos of model interpretability[J]. arXiv preprint arXiv:1606.03490v3, 2017.
[30]Zhouhan Lin, Minwei Feng, Cicero Nogueira dos Santos, Mo Yu, Bing Xiang, Bowen Zhou, and Yoshua Bengio. A structured self-attentive sentence embedding[C]//Proceedings of the International Conference on Learning Representations(ICLR), 2017.
[31]戴耀晶. 现代汉语被动句试析[C]. 汉语被动表述问题国际学术研讨会,2009.
[32]邓钊，贾修一，陈家骏.面向微博的中文反语识别研究[J].计算机工程与科学，2015, 37(12):2312-2317.
[33]刘正光. 反语理论综述[J].解放军外国语学院学报,2002(4):16-21.
[34]卢欣，李旸，王素格. 融合语言特征的卷积神经网络的反讽识别方法[J].中文信息学报，2019,33(5):31-38.
[35]孙晓，何家劲，任福继. 基于多特征融合的的混合神经网络模型讽刺语用判别[J].中文信息学报, 2016，30(6):215-233.
[36]涂靖. 反讽的语用特征和限制条件[J].外语学刊,2002(1):77-81.
[37]王俊平.“被+X”构式研究[D]. 吉林大学, 2011.
[38]邢竹天，徐扬. 面向网络文本的汉语反讽修辞识别方法研究[J].山西大学学报(自然科学版), 2015,38(3):385-391.
[39]赵华伦. 网络语言特点浅析[J].语言文字应用,2006(S2):219-221.
[40]赵毅衡. 反讽:表意形式的演化与新生[J].文艺研究,2011(1):18-27.
Yoshua Bengio. Learning phrase representationsusing RNN encoder-decoder for statistical machine translation[C]. CoRR, abs/1406.1078, 2014.

致谢

论文付梓之际，也是我即将向本科时光告别之时。回首四年学习生活，踯躅与奋进同在，付出和收获并存。可敬的老师、可亲的同学以及北大风物人文的浸淫，是我一生的财富。

首先，由衷地感谢北京大学中文系的詹卫东老师。作为我所在专业的负责人，詹老师不仅教授本专业的核心课程，还会关心专业里同学们的选课、学习和生活情况。詹老师也是我学年论文和毕业论文的指导老师，从选题到谋篇给予我悉心指导，使我受益良多。老师一丝不苟的工作作风和严谨求实的治学态度在潜移默化中影响着我，也勉励我在今后的学习生活中不断精进、更上层楼。他工作兢兢业业，待人平易近人，关心、爱护学生，很幸运在未来的研究生阶段能继续跟随詹老师学习和进步！

其次，感谢北京大学计算语言学研究所的刘扬老师。从本科一年级开始，我在刘老师的指导下接触科研实践。这份经历锻炼了我的学习能力，培养了我基本的科研素养，也帮助我更好地了解学科前景。同时也感谢所有于我有授业之恩的良师们。

接下来，感谢所有同门。感谢康司辰师兄、林子师姐和陈龙师兄在学业上对我的帮助和引导。感谢梦夏等专业同级的鼓励、关心和帮助。感谢 28 楼 302 的室友们，让内向又不善交际的我享受到了快乐而难忘的大学时光。还要感谢我的父母以及支持我的朋友们。

北京大学中国语言文学系应用语言学专业是一个计算机和语言学的交叉学科，在这样一个新兴的、发展迅速的专业中求知，是一种特殊的体验。得到其他同学对中文系学生还要学习程序设计、高等数学的震惊的同时，也意味着繁重的课业压力和不低的专业难度。交叉学科融汇的多种可能也曾给尚未对专业面貌形成初步了解的我们带来迷茫。它煎熬着我的煎熬，也成就了我的成就。专业虽小，却像家庭一样紧密。在师长和同窗的陪伴下我坚持下来，学会享受计算语言学和自然语言处理的学习、科研过程。

“嘤其鸣矣，求其友声”，何其有幸！

版权声明

任何收存和保管本论文各种版本的单位和个人，未经本轮为作者同意，不得将本论文转借他人，亦不得随意复制、抄录、拍照或以任何方式传播。否则，引起有碍作者著作权之问题，将可能承担法律责任。