申请代码	F060401
接收部门	
收件日期	
接收编号	6197021442

国家自然科学基金

申 请 书

（2 0 1 9 版）

资助类别：面上项目

亚类说明：

附注说明：

项目名称：面向成分句法分析的跨领域知识抽取与融合

申 请 人：张岳　　电 话：0571-87381107

依托单位：西湖大学

通讯地址：浙江省杭州市西湖区转塘街道石龙山街18号

邮政编码：310024　　单位电话：0571-86593608

电子邮箱：zhangyue@westlake.edu.cn

申报日期：2019年01月22日

国家自然科学基金委员会

基本信息

申请人信息	姓　　名	张岳	性别	男	出生年月	1980年09月	民族	满族
	学　　位	博士	职称	研究员		每年工作时间（月）		6
	是否在站博士后	否			电子邮箱	zhangyue@westlake.edu.cn		
	电　　话	0571-87381107			国别或地区	中国		
	个人通讯地址	浙江省杭州市西湖区转塘街道石龙山街18号						
	工　作　单　位	西湖大学/工学院						
	主要研究领域	自然语言处理						

依托单位信息	名　　称	西湖大学		
	联　系　人	周奇	电子邮箱	zhouqi@westlake.edu.cn
	电　　话	0571-86593608	网站地址	www.westlake.edu.cn

合作研究单位信息	单位名称

项目基本信息	项目名称	面向成分句法分析的跨领域知识抽取与融合		
	英文名称	Knowledge extraction and fusion for cross-domain constituent parsing		
	资助类别	面上项目	亚类说明	
	附注说明			
	申请代码	F060401.自然语言处理基础理论与方法	F060402.自然语言认知、理解与推理	
	基地类别			
	研究期限	2020年01月01日 -- 2023年12月31日	研究方向：句法分析	
	申请直接费用	71.8000万元		

中文关键词	成分句法分析；谓词论元结构；命名实体识别；语言模型；迁移学习
英文关键词	constituent parsing; predicate-argument structure; named entity recognition ; language modeling; transfer learning

中文摘要	成分句法分析是自然语言处理的一个基础问题，为下游的互联网、电子商务、企业智能、对话机器人等应用提供语言理解信息。近年来，随着深度学习技术的发展，新闻领域成分句法分析已经达到应用水平。然而，科技和社会媒体等领域的精度却有待提高。其中的关键瓶颈是缺少标注语料，难以充分训练机器学习模型。由于人工标注昂贵，而领域众多，如何充分利用一切现有资源，实现跨领域句法分析，成为一个具有重要意义的研究问题。与句法分析相关的可用资源，包括大规模的生文本，和多元异构的相关任务（如命名实体、语义角色等）人工标注语料。现有工作研究了单领域下相关任务对句法分析的促进，以及单任务的句法分析领域迁移。为充分利用资源，本项目提出一个新的研究问题，即如何利用多领域多任务标注数据挖掘融合跨领域成分句法知识。我们计划以深度学习为技术手段，系统研究跨领域稳定句法分析的关键因素和方法。该研究对语言处理技术广泛应用具有潜在价值。
英文摘要	Constituent parsing is a fundamental problem in natural language processing, which provides useful information for downstream tasks such as the Internet, e-Commerce, business intelligence, dialogue robots etc. Over the recent years, with advance in deep learning techniques, constituent parsing has reached practical performance in the news domain. However, for domains such as technology and social media, the accuracies are still relatively low. The key reason is lack of labeled data for training machine learning models. Because manual labeling is expensive and the number of domains is huge, an important research question is how to leverage all resources available for robust cross-domain constituent parsing. Resources that are relevant include large unlabeled text, and labeled data for relevant tasks (e.g., named entity recognition, semantic roles etc.). Existing work has investigated the effectiveness of using relevant tasks for improving constituent parsing in the news domain, and single-task domain adaptation for constituent parsing. For making the most use of resources, this project considers a new research question, namely how to effectively use cross-domain cross-task data for mining and utilizing cross-domain constituent parsing knowledge. We plan to systematically explore key factors and methods for cross-domain robust constituent parsing through the use of deep learning techniques. The research has potential value for wide-coverage natural language processing usage in the industry.

项目组主要参与者（注：项目组主要参与者不包括项目申请人）

编号	姓名	出生年月	性别	职 称	学 位	单位名称	电话	电子邮箱	证件号码	每年工作时间（月）
1	滕志扬	1989-09-10	男	助理研究员	博士	西湖大学	15017052428	tengzhiyang@wias.org.cn	4**********6	8
2	何奇	1986-03-01	男	助理研究员	博士	西湖大学	13880678211	heqi@westlake.edu.cn	5**********9	8
3	张源	1992-11-15	男	技术员	硕士	西湖大学	153114918157	zhangyuan@westlake.edu.cn	2**********2	10
4	崔乐阳	1995-11-01	男	博士生	硕士	西湖大学	13683234080	cuileyang@westlake.edu.cn	1**********8	8
5	王祎乐	1991-11-12	男	博士生	硕士	西湖大学	15068156517	wangyile@westlake.edu.cn	4**********1	8
6	白雪峰	1995-09-09	男	博士生	学士	西湖大学	15754604524	xfbai.hk@gmail.com	1**********5	8
7	贾晨	1995-07-12	男	博士生	学士	西湖大学	18840831231	jiachen@westlake.edu.cn	1**********7	8
8	陈雨龙	1994-10-10	男	博士生	硕士	西湖大学	18248884600	yulongchen1010@gmail.com	4**********3	8

总人数	高级	中级	初级	博士后	博士生	硕士生
9	1	2	1	0	5	0

国家自然科学基金项目资金预算表（定额补助）

项目申请号：6197021442　　项目负责人：张岳　　金额单位：万元

序号	科目名称	金额
	(1)	(2)
1	**项目直接费用合计**	71.8000
2	1、 设备费	15.0000
3	(1)设备购置费	15.00
4	(2)设备试制费	0.00
5	(3)设备升级改造与租赁费	0.00
6	2、 材料费	5.10
7	3、 测试化验加工费	1.50
8	4、 燃料动力费	0.00
9	5、 差旅/会议/国际合作与交流费	26.00
10	6、 出版/文献/信息传播/知识产权事务费	9.40
11	7、 劳务费	12.80
12	8、 专家咨询费	2.00
13	9、 其他支出	0.00

预算说明书（定额补助）

（请按照《国家自然科学基金项目预算表编制说明》的有关要求，对各项支出的主要用途和测算理由，以及合作研究外拨资金、单价≥10万元的设备费等内容进行必要说明。）

设备购置费：

用于购买高性能服务器4台， 12 TFLOPS 独立显卡 Nvidia Titan Xp 5个，网络设备等，预算总计15万左右，以支持数据处理，深度学习计算、系统开发与测试等任务。

材料费：

存储耗材，用于数据保存、备份、服务器容量扩充，包括内存、固定硬盘、移动硬盘等，四年约3万元；
网络耗材，用于支持网络通信，包括上网卡，路由器，网络通信费等，四年共计约1.5万元；
其它办公用品，0.6万元。

测试化验加工费：

句法分析(1万句)及相关任务语料加工，约1.5万元。

差旅费/会议/国际合作与交流费：

项目组成员参加、组织国内召开的学术会议、项目研讨会等费用，预计每年出差4人次左右，四年一共16人次，每人次平均花费约0.45万元，四年共计约7.2万元。

研究人员出国参加国际高水平会议以及评测，预计每年出差2人次左右，四年一共8人次，每人次平均花费约1.6万元，四年共计约12.8万元。

邀请国外相关专家来华访问，四年预计共计4人次，每人次按照平均1.5万元计算，四年共计约6.0万元。

出版/文献/信息传播/知识产权事务费：

申请专利1-3项，共计约0.9万元；
论文版面费：国内外顶级期刊论文2-4篇，注册费共计约2.9万；顶级会议论文6-9篇，注册费约5.0万元；
其它专业相关资料购买4年约0.6万元。

专家咨询费：

用于在项目研究过程中支付临时聘请咨询专家的费用，四年预计约2万元。

劳务费：

用于直接参与项目的技术员与博士研究生的劳务费，四年共计约 12.8 万元：
博士生 5 人，每人每月按 500 元计算，每年发 8 个月； 技术员 1 人，按每人每月 1200 元计算，每年 10 个月。

报告正文

（一）立项依据与研究内容（建议 8000 字以下）：

1. **项目的立项依据**（研究意义、国内外研究现状及发展动态分析，需结合科学研究发展趋势来论述科学意义；或结合国民经济和社会发展中迫切需要解决的关键科技问题来论述其应用前景。附主要参考文献目录）；

1.1. 研究意义

成分句法（constituent grammar）分析 [1, 2] 研究自然语言理解中的层次短语关系结构。如图 1 所示，给定一个句子“上市银行不良贷款率呈现下降趋势”，成分句法分析自动获得其中的名词短语（NP）“上市银行”、动词短语（VP）“呈现下降趋势”等层次短语结构。和其他常见的形式句法 (formal grammar) 相比，成分句法结构标注中含有丰富的结构知识。因此，成分句法结构可以转化出依存句法 [3](dependency grammar)、中心词驱动短语结构文法 [4]（head-lexicalized phrase structure grammar）、词汇功能文法（lexical-functional grammar）、树邻接文法（tree adjoining grammar）等其它结构。作为自然语言处理的基础问题，成分句法分析可以为下游的语言处理任务提供有用信息。例如，图 1 中句法树所包含的事件信息可能给市场预测提供线索。再如，理解社会媒体文本可能给民意分析提供帮助。总体而言，精确和跨领域稳定的成分句法分析对于互联网、电子商务、企业智能、民意检测、家政机器人、智能互动驾驶等诸多应用具有重要价值。

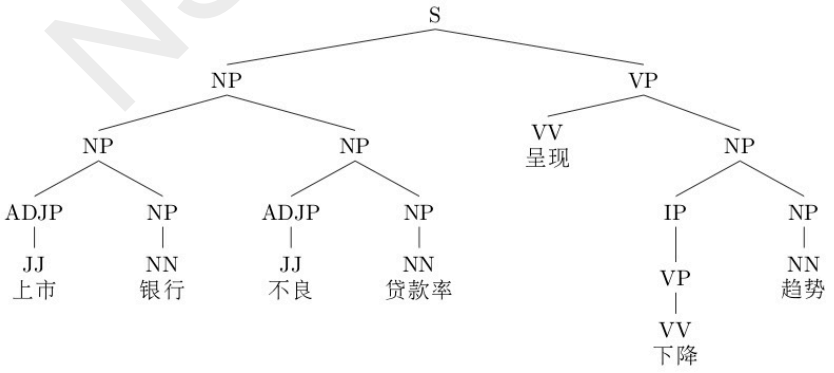

图 1: 成分句法树示例

随着深度学习技术的快速发展，自动句法分析性能不断进步。在研究较多的新闻领域，精确度已经达到了实用的水平。然而，在社会媒体、科技、网络文学、生物医药等不同领域，句法分析模型的精确度明显下降。比如，句法分析在宾州树库 [5, 6] 标准测试集上的精确度已经达到了 95% 以上 [7, 8]，但是在电子邮件、社会媒体、生物医药的某些数据集上的精确度却只有 80% 左右 [9]。一个重要的原因是，在这些领域上缺乏足够的训练数据，难以使机器学习模型达到可靠的精确度。

	成分句法分析	命名实体识别	语言模型
新闻文本	足量标注数据	足量标注数据	大量生文本
科技文本	少量标注数据	无资源	大量生文本
社会媒体文本	无资源	少量标注数据	大量生文本

图 2: 跨领域辅助任务和语料示例

从实用化和工业化的角度来看，在新闻领域句法分析成熟的条件下，跨领域的稳定句法分析成为意义突出的研究课题。由于人工标注成本昂贵，如何有效地将新闻语料中获得的知识迁移到其他领域成为一个有价值的学术研究课题。

传统方法针对句法分析任务本身，研究不同领域成分句法结构之间的区别和联系，例如挖掘跨领域的成分句法特征 [10]，或跨领域深度学习表示 [11]。这些方法对领域迁移起到一定帮助。然而，语言理解是一个综合的过程，很多句法分析之外的任务，也可能包含和句法有关的重要信息。比如，语言模型研究生文本的结构，因而从跨领域的语言建模之中可以获得不同领域行文之间的区别和联系。再如，命名实体识别关注名词短语构成。从跨领域的命名实体识别可以获得不同领域专有名词之间的区别和联系。这些句法分析任务本身之外的丰富信息，可以为挖掘跨领域句法知识提供宝贵参考。

从标注资源的角度，跨领域的句法资源非常有限，但是跨领域的生文本大量存在。另外，跨领域的命名实体、语义角色标注等资源都可能作为可利用的额外资源。已有相关的工作研究单领域的相关任务资源对句法分析的辅助作用 [12, 13, 14, 15]，以及跨领域的单任务句法资源的作用 [53, 54, 55, 56]，但较少有以往工作研究跨任务跨领域知识对跨领域句法结构的作用。**基于以上观察，本项目从一个新的角度研究成分句法分析的领域适应问题，即通过相关辅助任务的跨领域关联，挖掘与融合句法分析的跨领域知识。**

这项研究涉及一个综合的知识挖掘与利用场景，包含跨领域、跨任务两种知识迁移。以语言模型和命名实体识别两个辅助任务为例，在新闻文本、科技文本和社会媒体文本三个领域，存在九个任务和领域的组合，如图 2 所示。其中，句法分析和命名实体识别需要有标注语料进行训练，而语言模型可以在没有人工标注的生文本上训练。在图 2 的场景中，对于新闻领域，所有任务都有足够语料。对于科技文本，存在少量句法标注，但没有命名实体标注。对于社会媒体文本，存在少量命名实体标注，但没有句法标注。如图所示，纵向看，通过语言建模，可以挖掘到全部三个领域之间的关系，而通过命名实体识别可以获得新闻文本和社会媒体文本之间的关系。这些跨领域的知识需要通过任务之间的内在关联传递。而后者可以通过每个领域内部不同任务之间的横向对比自动获得。例如，在新闻文本上，可以挖掘全部三个任务之间的内在联系，而通过科技文本，可以获得句法分析和语言模型之间的联系。综上所述，此场景下跨领域纵向联系和跨任务横向联系相互依赖、相互增强。因此，社会媒体上的命名实体资源，可能通过综合关联知识，

对科技文本上的句法分析起到间接帮助。

如何最有效的在这类场景之下同时利用所有资源挖掘跨领域的句法知识是一个既有实际意义又有技术挑战的科学问题。在我们的初步研究中，已经分别探索了相同领域不同任务之间的知识传递 [16, 12, 17]，以及相同任务不同领域之间的知识传递 [18, 19, 20]，取得了一定成效。我们将以这些初步研究为基础，以相关任务为分析工具和表达形式，深入系统的探索：(1) 成分句法结构在跨领域有哪些共性的本质，(2) 不同领域之间的成分句法结构有何内在的区别和联系，(3) 在深度学习模型之中，如何最有效地表达和融合成分句法结构跨领域的共性和特性知识，实现有监督和无监督条件下的广泛领域迁移。

项目的一个重要的目标是允许尽可能多的同时利用各方面的宝贵人工标注，以及大量没有标注的文本，以资源成本最低的方式提高成分句法分析的跨领域稳定性。这对于构建广泛实用的自然语言分析器，具有直接的潜在意义。而后者有助于下游应用在社会经济中发挥更重要的作用。另外，由于跨领域的性能差异是自然语言处理诸多任务中的普遍问题，本项目的研究对自然语言处理其他任务也可能有借鉴意义。

1.2. 国内外研究现状及发展动态分析

我们分别从问题和方法两个角度介绍国内外相关工作。本项目关注的问题，成分句法分析，是自然语言处理领域的一个核心问题，长期以来一直受到广泛关注。句法分析的领域适应也受到一系列研究关注。从研究思路的角度看，本项目拟采用基于图的深度表示学习方法挖掘结构知识。这个方向近一年多来逐渐受到重视。此外，本项目的跨任务跨领域知识融合问题，在技术层面属于迁移学习与多任务学习。后者随着近些年深度学习技术的发展成为机器学习和自然语言处理领域的一个热点研究方向。

1.2.1 成分句法分析（constituent parsing）

成分句法分析是自然语言处理领域长久以来的一个重要课题。基于语言学规则的句法分析可以上溯到上世纪 60 年代 [21]。上世纪 80 年代，统计机器学习开始兴起时，在计算和人工标注资源受限的条件下，学术界就开始了对局部成分结构的研究工作 [22]。典型的任务包括短语切分 [23]、介词短语附着 [24] 等问题。随着资源的进步，上世纪 90 年代开始了全句结构分析 [25, 26]。统计模型使成分句法分析成为研究热点 [27, 28]。近五年来，深度学习技术在自然语言处理研究中得到广泛应用，使得成分句法分析的性能得到了大幅提升，达到实用程度 [7, 8, 29, 30, 31, 32, 33, 34, 35, 36]。然而，大多数现有工作都是在新闻领域有标注的训练和测试语料 [5, 6, 37, 38, 39, 40, 41, 42] 上进行的。而本项目关注的重点，是

在新闻以外的领域，使成分句法分析可以稳定地应用于更广泛的文体。

也有工作研究**相关任务对句法分析的影响**。Finkel 和 Manning [14] 研究了命名实体识别对成分句法分析的影响。一系列工作研究了词性和句法结构知识的相互影响 [13, 43, 44]。语义角色标注对句法分析的影响也是学术界讨论已久的话题 [45, 46, 47, 48, 12, 49]。朱慕华 [48] 研究了从短语切分等相关任务获取特征，提高统计成分句法分析的方式。孙薇薇 [50] 从多个模型融合的角度提高句法分析。值得注意的是，近一年来，大规模未经标注的生文本上训练的语言模型被广泛应用到基础自然语言处理任务中。作为上下文相关的词的向量表示，这些模型给很多任务带来了性能提升 [8, 51]。这些方法可以看作是跨任务迁移学习的一个案例。与上述工作相同，本项目也关注相关任务对成分句法分析性能的影响。然而，上述工作只在单一领域研究知识迁移。与上述相关工作不同的是，本项目通过辅助任务来挖掘跨领域的句法分析知识。

句法分析领域迁移是一个受到国内外关注的研究课题 [52]。一个最直观的研究思路是进行数据标注 [53, 54, 55, 56, 57]。与这些研究不同的是，本项目专注于从现有资源提取跨领域的相关知识，以减少人工标注的成本。在句法分析领域迁移的相关研究之中，一系列的工作利用自学习 (self-training) 和互学习 (co-training) 等技术挖掘跨领域知识 [58, 59, 60, 61, 62, 63, 64]。Mitchell 和 Steedman[65] 通过聚类方法为新的领域词汇估算模型参数。Wang 等人 [10] 通过研究词汇统计和句法规则之间的关系从新闻文本到医药文本进行领域适应。Yang 等人 [66] 通过深层置信网络（deep belief network）从没有标注的目标领域文本中学习源领域和目标领域之间特征的关系。Mukherjee 等人 [67] 通过无监督的主题模型区分句法特征。Sato 等人 [11] 通过对抗学习（adversarial training）挖掘两个领域之间的共同特点。Mukherjee 等人 [68] 通过修改源领域句法分析器在目标领域上的错误寻找领域之间的特性关系。与本项目不同的是，现有研究着重句法分析任务本身的跨领域关系，却没有利用辅助任务挖掘更丰富的跨领域知识。

综上所述，在新闻领域句法分析成熟的条件下，跨领域的稳定句法分析成为意义突出的研究课题。现有国内外研究关注相关任务对句法任务的帮助作用，也关注句法知识本身的跨领域挖掘，但较少有研究综合利用多方面相关任务的跨领域信息。本项目从这个意义上填补研究空白。

1.2.2 基于图的深度表示学习（deep representation learning for graphs）

自然语言处理的很多任务属于结构问题，比如序列结构、树结构等。给定一个句子，不同任务结构经过融合，其表示形式是一般的图结构。因此，在深度学习技术方面，为充分利用结构性的跨领域知识，就需要对图进行**神经网络编码**。传统的卷积神经网络（convolutional neural network）[69, 70]、循环神经网络（recurrent neural network）[71, 72] 和自注意力机制网络（self attention network）[73] 往往

关注对序列的编码。近些年来，也有工作把网络结构扩展到树状结构 [74, 75] 和有向无环图 [76, 77, 78, 79, 80] 上。这些拓展使得在序列建模方向取得成就的神经元网络结构可以用于句法树结构（syntactic tree structure）[76] 和网格结构（lattice structure）[80] 的建模。但是，它们的序列化本质无法表达有环的复杂图结构。

近一年多，可以表示环状结构的**图网络模型** [81] 逐渐被用于自然语言处理领域。其中，目前应用最多的是基于图的卷积神经网络（graph convolutional neural network）。它的基本思想是相邻结点范围内实现卷积操作，通过迭代卷积逐层推导出具有更广的上下文背景的节点表示。这种卷积神经网络被应用于文本分类 [82, 83, 84]、机器翻译 [85, 86]、语义分析 [87] 和信息抽取 [88, 89, 90] 等任务。循环神经网络也被应用于图的表示学习。它的基本思想是用循环操作代替卷积操作实现逐层抽象的节点表示。这种网络被应用到文本分类 [91]、文本生成 [92, 93]、机器阅读 [94] 以及信息抽取 [95] 等任务。在上述工作中，我们的研究做出了较大贡献。此外，注意力机制网络也被扩展到图的建模任务 [96]。这一系列的前沿工作，尤其是基于图的循环神经网络工作，将被用作本项目结构知识表示的研究方法。

值得注意的是，尽管本项目研究以相关任务结构挖掘的知识，图表示方法也允许我们的模型利用其它现有资源结构，如本体知识库等 [97, 98, 99, 100, 101]。

1.2.3 **迁移学习在自然语言处理中的应用研究** (transfer learning for NLP)

多任务学习（multi-task learning）在神经自然语言处理中有较长的研究历史。Collobert 等人 [102] 在提出早期的神经自然语言处理模型的时候，就考虑到了词性标注、短语切分和命名实体识别任务之间的迁移学习，并且通过**参数共享**（parameter sharing）获得了一定的效果。Miwa 和 Bansal[103] 同样用参数共享的方法研究了实体识别和关系抽取的联合模型。Rei[104] 发现语言模型和序列标注任务之间也可以通过参数共享得到性能提升。Plank[105] 发现键盘的敲击时间预测和短语切分任务可以互相提升。Søgaard 和 Goldberg[106] 研究了不同任务在不同网络层之间共享参数的性能差异，发现选择性的共享参数可以提高迁移学习能力。我们在句法分析和语义角色标注联合学习的任务上也有类似发现 [12, 49]。

除了参数共享，**预训练**（pretraining）也可以实现不同任务之间的迁移学习。Ramachandran 等人 [107] 通过语言模型预训练提高机器翻译的性能。Peters 等人 [108] 通过双向循环神经网络语言模型的预训练提高多个任务的精确度。Devlin 等人 [109] 和 Radford 等人 [110] 通过自注意力机制网络实现了 Peters 等人工作的替代版本。我们通过多任务的预训练提高分词的精准度 [111]。Howard 等人 [112] 研究了语言模型预训练中的一系列实际问题。Kiperwasser 和 Ballesteros[113] 研究了预训练和参数共享之间的区别和联系。预训练也可被视为一种多任务学习。近期，有一系列工作显示，不是所有的任务之间都可以通过参数共享进行迁移学习 [114, 115, 116]。本项目针对传统方法只利用领域共性却不充分利用特性的问题，为

跨任务、跨领域句法知识迁移设计创新的多任务学习方法。

共享-私有网络 (shared-private network) 使用**对抗训练** (adversarial training) 实现不同领域共性和特性分离。Kim 等人 [117] 将对抗训练应用于跨语言的迁移学习。Chen 等人 [118] 将对抗训练应用于跨领域的迁移学习。上述方法被证实有一定效用。然而，当用到本项目的多任务多领域场景下，潜在的一个问题是需要针对每一个任务和领域的组合设定一套特性参数，因此不能解决领域、任务增多时的组合爆炸问题。此外，上述方法要求每种组合都存在训练语料，因而不适用于资源稀缺的实际情境。为解决这个问题，一个可能的方法是**参数分解**，以便拆解领域、任务组合。虽然直接相关工作很少，文献中有相关方法可以借鉴。Ammar 等人 [119] 和 Platanios 等人 [120] 把语言特征存储为嵌入向量 (embedding vector)，解决多语言任务。Stynme 等人 [121] 把树库特征存储为嵌入向量，解决多树库融合问题。我们的前期工作 [122] 把领域特征作为向量，解决单任务多领域问题。我们计划以这些方法作为基础，结合对抗训练的方式，研究领域和任务特性分解的合理方法。

参考文献:

[1] Daniel Jurafsky and James H Martin. Speech and language processing: An introduction to speech recognition, computational linguistics and natural language processing. Upper Saddle River, NJ: Prentice Hall, 2008.

[2] 宗成庆. 统计自然语言处理. 清华大学出版社, 2013.

[3] Natalia Silveira, Timothy Dozat, Marie-Catherine De Marneffe, Samuel R Bowman, Miriam Connor, John Bauer, and Christopher D Manning. A gold standard dependency corpus for english. In LREC, pages 2897–2904, 2014.

[4] Yusuke Miyao, Takashi Ninomiya, and Jun' ichi Tsujii. Corpus-oriented grammar development for acquiring a head-driven phrase structure grammar from the penn treebank. In International Conference on Natural Language Processing, pages 684–693. Springer, 2004.

[5] Marcus Mitchell et al. Treebank-3 (LDC99T42). Linguistic Data Consortium, Philadelphia, PA, 1999.

[6] Marie-Catherine de Marneffe, Timothy Dozat, Natalia Silveira, Katri Haverinen, Filip Ginter, Joakim Nivre, and Christopher D. Manning. Universal stanford dependencies: A cross-linguistic typology. In Proceedings of the Ninth International Conference on Language Resources and Evaluation, LREC 2014, Reykjavik, Iceland, May 26-31, 2014., pages 4585–4592, 2014.

[7] Jiangming Liu and Yue Zhang. In-order transition-based constituent parsing. Transactions of the Association for Computational Linguistics (TACL) , 5:413–424, 2017.

[8] Nikita Kitaev and Dan Klein. Constituency parsing with a self-attentive encoder. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers, pages 2675–2685, 2018.

[9] Dat Quoc Nguyen and Karin Verspoor. From POS tagging to dependency parsing for biomedical event extraction. BMC Bioinformatics, 20(1):72:1–72:13, 2019.

[10] Yan Wang, Serguei Pakhomov, James O. Ryan, and Genevieve B. Melton. Domain adaption of parsing for operative notes. Journal of Biomedical Informatics, 54:1–9, 2015.

[11] Motoki Sato, Hitoshi Manabe, Hiroshi Noji, and Yuji Matsumoto. Adversarial training for cross-domain universal dependency parsing. In Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies, Vancouver, Canada, August 3-4, 2017, pages 71–79, 2017.

[12] Peng Shi, Zhiyang Teng, and Yue Zhang. Exploiting mutual benefits between syntax and semantic roles using neural network. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 968–974, Austin, Texas, November 2016. Association for Computational Linguistics.

[13] Bernd Bohnet and Joakim Nivre. A transition-based system for joint part-of-speech tagging and labeled non-projective dependency parsing. In Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, EMNLP-CoNLL 2012, July 12-14, 2012, Jeju Island, Korea, pages 1455–1465, 2012.

[14] Jenny Rose Finkel and Christopher D. Manning. Joint parsing and named entity recognition. In Human Language Technologies: Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, May 31 - June 5, 2009, Boulder, Colorado, USA, pages 326–334, 2009.

[15] Jie Yang, Shuailong Liang, and Yue Zhang. Design challenges and misconceptions in neural sequence labeling. In Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, Santa Fe, New Mexico, USA, August 20-26, 2018, pages 3879–3889, 2018.

[16] Meishan Zhang, Yue Zhang, Wanxiang Che, and Ting Liu. Chinese parsing exploiting characters. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 125–134, Sofia, Bulgaria, August 2013. Association for Computational Linguistics.

[17] Shaolei Wang, Yue Zhang, Wanxiang Che, and Ting Liu. Joint extraction of entities and relations based on a novel graph scheme. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden., pages 4461–4467, 2018.

[18] Meishan Zhang, Yue Zhang, Wanxiang Che, and Ting Liu. Type-supervised domain adaptation for joint segmentation and pos-tagging. In Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, pages 588–597, 2014.

[19] Yijia Liu, Yue Zhang, Wanxiang Che, Ting Liu, and Fan Wu. Domain adaptation for crf-based chinese word segmentation using free annotations. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 864–874, Doha, Qatar, October 2014. Association for Computational Linguistics.

[20] Likun Qiu and Yue Zhang. Word segmentation for chinese novels. In Proceedings of the Twenty-Nineth AAAI Conference on Artificial Intelligence, Austin, Texas, USA, 2015.

[21] Martin Kay. The tabular parser: A parsing program for phrase structure and dependency. No-4933-PR. RM. RAND CORP SANTA MONICA CALIF, 1966.

[22] Daniel Jurafsky. A probabilistic model of lexical and syntactic access and disambiguation. Cognitive Science, 20(2):137–194, 1996.

[23] Steven P Abney. Parsing by chunks. In Principle-based parsing, pages 257–278. Springer, 1991.

[24] Adwait Ratnaparkhi. Statistical models for unsupervised prepositional phrase attachment. CoRR, cmp-lg/9807011, 1998.

[25] Eugene Charniak. A maximum-entropy-inspired parser. In 6th Applied Natural Language Processing Conference, ANLP 2000, Seattle, Washington, USA, April 29 - May 4, 2000, pages 132–139, 2000.

[26] Michael Collins. Three generative, lexicalised models for statistical parsing. In 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics (ACL), Proceedings of the Conference, 7-12 July 1997, Madrid, Spain., pages 16–23, 1997.

[27] Slav Petrov and Dan Klein. Improved inference for unlexicalized parsing. In Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, April 22-27, 2007, Rochester, New York, USA, pages 404–411, 2007.

[28] Muhua Zhu, Yue Zhang, Wenliang Chen, Min Zhang, and Jingbo Zhu. Fast and accurate shift-reduce constituent parsing. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 434–443, Sofia, Bulgaria, August 2013. Association for Computational Linguistics.

[29] Greg Durrett and Dan Klein. Neural CRF parsing. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China, Volume 1: Long Papers, pages 302–312, 2015.

[30] Taro Watanabe and Eiichiro Sumita. Transition-based neural constituent parsing. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China, Volume 1: Long Papers, pages 1169–1179, 2015.

[31] Mitchell Stern, Jacob Andreas, and Dan Klein. A minimal span-based neural constituency parser. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers, pages 818–827, 2017.

[32] 周青宇. 基于深度学习的自然语言句法分析研究. 硕士论文，哈尔滨工业大学, 2016.

[33] Yikang Shen, Zhouhan Lin, Athul Paul Jacob, Alessandro Sordoni, Aaron C. Courville, and Yoshua Bengio. Straight to the tree: Constituency parsing with neural syntactic distance. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers, pages 1171–1180, 2018.

[34] Carlos Gómez-Rodríguez and David Vilares. Constituent parsing as sequence labeling. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018, pages 1314–1324, 2018.

[35] Zhiyang Teng and Yue Zhang. Two local models for neural constituent parsing. In Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, Santa Fe, New Mexico, USA, August 20-26, 2018, pages 119–132, 2018.

[36] 郭江. 基于分布表示的跨语言跨任务自然语言分析. 博士论文，哈尔滨工业大学, 2017.

[37] Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajic, Christopher D. Manning, Ryan T. McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty, and Daniel Zeman. Universal dependencies v1: A multilingual treebank collection. In Proceedings of the Tenth International Conference on Language Resources and Evaluation LREC 2016, Portorož, Slovenia, May 23-28, 2016., 2016.

[38] Julia Hockenmaier. Data and models for statistical parsing with Combinatory Categorial Grammar. PhD thesis, University of Edinburgh. College of Science and Engineering, 2003.

[39] Nianwen Xue, Fu-Dong Chiou, and Martha Stone Palmer. Building a large-scale annotated chinese corpus. In 19th International Conference on Computational Linguistics, COLING 2002, Howard International House and Academia Sinica, Taipei, Taiwan, August 24 - September 1, 2002, 2002.

[40] 邱立坤，金澎, 王厚峰. 基于依存语法构建多视图汉语树库. 中文信息学报, 29(3):9–15, 2015.

[41] 周强. 汉语基本块描述体系. 中文信息学报, 21(3):21–27, 2007.

[42] 詹卫东. 树库在汉语语法辅助教学中的应用初探. Journal of Technology and Chinese Language Teaching, 3(2):16–29, 2012.

[43] Zhiguo Wang and Nianwen Xue. Joint POS tagging and transition-based constituent parsing in chinese with non-local features. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014, June 22-27, 2014, Baltimore, MD, USA, Volume 1: Long Papers, pages 733–742, 2014.

[44] Dat Quoc Nguyen, Mark Dras, and Mark Johnson. A novel neural network model for joint POS tagging and graph-based dependency parsing. In Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies, Vancouver, Canada, August 3-4, 2017, pages 134–142, 2017.

[45] Swabha Swayamdipta, Miguel Ballesteros, Chris Dyer, and Noah A. Smith. Greedy, joint syntactic-semantic parsing with stack lstms. In Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning, CoNLL 2016, Berlin, Germany, August 11-12, 2016, pages 187–197, 2016.

[46] Stephen A. Boxwell, Dennis Mehay, and Chris Brew. What a parser can learn from a semantic role labeler and vice versa. In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, EMNLP 2010, 9-11 October 2010, MIT Stata Center, Massachusetts, USA, A meeting of SIGDAT, a Special Interest Group of the ACL, pages 736–744, 2010.

[47] Charles A. Sutton and Andrew McCallum. Joint parsing and semantic role labeling. In Proceedings of the Ninth Conference on Computational Natural Language Learning, CoNLL 2005, Ann Arbor, Michigan, USA, June 29-30, 2005, pages 225–228, 2005.

[48] 朱慕华. 基于多数据源的成分句法分析研究. 博士论文，东北大学, 2013.

[49] Peng Shi and Yue Zhang. Joint bi-affine parsing and semantic role labeling. In 2017 International Conference on Asian Language Processing (IALP), pages 338–341. IEEE, 2017.

[50] Weiwei Sun. Learning Chinese language structures with multiple views. PhD thesis, Saarland University, 2012.

[51] Wanxiang Che, Yijia Liu, Yuxuan Wang, Bo Zheng, and Ting Liu. Towards better UD parsing: Deep contextualized word embeddings, ensemble, and treebank concatenation. In Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies, Brussels, Belgium, October 31 - November 1, 2018, pages 55–64, 2018.

[52] Plank Babara. Domain adaptation for parsing. PhD thesis, University of Groningen, Groningen, 2011.

[53] Zhenghua Li, Yue Zhang, Jiayuan Chao, and Min Zhang. Training dependency parsers with partial annotation. CoRR, abs/1609.09247, 2016.

[54] Vidur Joshi, Matthew Peters, and Mark Hopkins. Extending a parser to distant domains using a few dozen partially annotated examples. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers, pages 1190–1199, 2018.

[55] Zheng-Yu Niu, Haifeng Wang, and Hua Wu. Exploiting heterogeneous treebanks for parsing. In ACL 2009, Proceedings of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the AFNLP, 2-7 August 2009, Singapore, pages 46–54, 2009.

[56] Muhua Zhu, Jingbo Zhu, and Minghan Hu. Better automatic treebank conversion using A feature-based approach. In The 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference, 19-24 June, 2011, Portland, Oregon, USA - Short Papers, pages 715–719, 2011.

[57] Xian Li, Wenbin Jiang, Yajuan Lü, and Qun Liu. Iterative transformation of annotation guidelines for constituency parsing. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, ACL 2013, 4-9 August 2013, Sofia, Bulgaria, Volume 2: Short Papers, pages 591–596, 2013.

[58] Xuezhe Ma and Fei Xia. Dependency parser adaptation with subtrees from auto-parsed target domain data. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, ACL 2013, 4-9 August 2013, Sofia, Bulgaria, Volume 2: Short Papers, pages 585–590, 2013.

[59] David McClosky, Eugene Charniak, and Mark Johnson. Reranking and self-training for parser adaptation. In ACL 2006, 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference, Sydney, Australia, 17-21 July 2006, 2006.

[60] Roi Reichart and Ari Rappoport. Self-training for enhancement and domain adaptation of statistical parsers trained on small datasets. In ACL 2007, Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, June 23-30, 2007, Prague, Czech Republic, 2007.

[61] Daisuke Kawahara and Kiyotaka Uchimoto. Learning reliability of parses for domain adaptation of dependency parsing. In Proceedings of the Third International Joint Conference on Natural Language Processing: Volume-II, 2008.

[62] Kenji Sagae. Self-training without reranking for parser domain adaptation and its impact on semantic role labeling. In Proceedings of the 2010 Workshop on Domain Adaptation for Natural Language Processing, pages 37–44. Association for Computational Linguistics, 2010.

[63] Eric Baucom, Levi King, and Sandra Kübler. Domain adaptation for parsing. In Recent Advances in Natural Language Processing, RANLP 2013, 9-11 September, 2013, Hissar, Bulgaria, pages 56–64, 2013.

[64] Juntao Yu, Mohab Elkaref, and Bernd Bohnet. Domain adaptation for dependency parsing via self-training. In Proceedings of the 14th International Conference on Parsing Technologies, IWPT 2015, Bilbao, Spain, July 5-7, 2015, pages 1–10, 2015.

[65] Jeff Mitchell and Mark Steedman. Parser adaptation to the biomedical domain without re-training. In Proceedings of the Sixth International Workshop on Health Text Mining and Information Analysis, EMNLP 2015, Lisbon, Portugal, September 17, 2015, pages 79–89, 2015.

[66] Haitong Yang, Tao Zhuang, and Chengqing Zong. Domain adaptation for syntactic and semantic dependency parsing using deep belief networks. Transactions of the Association for Computational Linguistics (TACL) , 3:271–282, 2015.

[67] Matthias Scheutz, Sandra Kübler, and Atreyee Mukherjee. Creating POS tagging and dependency parsing experts via topic modeling. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017, Valencia, Spain, April 3-7, 2017, Volume 1: Long Papers, pages 347–355, 2017.

[68] Atreyee Mukherjee and Sandra Kübler. Domain adaptation in dependency parsing via transformation based error driven learning. In Proceedings of the 17th International Workshop on Treebanks and Linguistic Theories (TLT 2018), December 13–14, 2018, Oslo University, Norway, number 155, pages 179–192. Linköping University Electronic Press, 2018.

[69] Nal Kalchbrenner, Edward Grefenstette, and Phil Blunsom. A convolutional neural network for modelling sentences. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014, June 22-27, 2014, Baltimore, MD, USA, Volume 1: Long Papers, pages 655–665, 2014.

[70] Yoon Kim. Convolutional neural networks for sentence classification. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL, pages 1746–1751, 2014.

[71] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. CoRR, abs/1409.0473, 2014.

[72] Ming Tan, Bing Xiang, and Bowen Zhou. Lstm-based deep learning models for non-factoid answer selection. CoRR, abs/1511.04108, 2015.

[73] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. In Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA, pages 6000–6010, 2017.

[74] Kai Sheng Tai, Richard Socher, and Christopher D. Manning. Improved semantic representations from tree-structured long short-term memory networks. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China, Volume 1: Long Papers, pages 1556–1566, 2015.

[75] Xiaodan Zhu, Parinaz Sobihani, and Hongyu Guo. Long short-term memory over recursive structures. In International Conference on Machine Learning, pages 1604–1612, 2015.

[76] Nanyun Peng, Hoifung Poon, Chris Quirk, Kristina Toutanova, and Wen-tau Yih. Cross-sentence n-ary relation extraction with graph lstms. Transactions of the Association for Computational Linguistics (TACL) , 5:101–115, 2017.

[77] Xinchi Chen, Zhan Shi, Xipeng Qiu, and Xuanjing Huang. Dag-based long short-term memory for neural word segmentation. CoRR, abs/1707.00248, 2017.

[78] Matthias Sperber, Graham Neubig, Jan Niehues, and Alex Waibel. Neural lattice-to-sequence models for uncertain inputs. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9-11, 2017, pages 1380–1389, 2017.

[79] Jinsong Su, Zhixing Tan, Deyi Xiong, Rongrong Ji, Xiaodong Shi, and Yang Liu. Lattice-based recurrent neural network encoders for neural machine translation. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA., pages 3302–3308, 2017.

[80] Yue Zhang and Jie Yang. Chinese NER using lattice LSTM. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers, pages 1554–1564, 2018.

[81] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. IEEE Trans. Neural Networks, 20(1):61–80, 2009.

[82] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. CoRR, abs/1609.02907, 2016.

[83] William L. Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. In Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA, pages 1025–1035, 2017.

[84] Hao Peng, Jianxin Li, Yu He, Yaopeng Liu, Mengjiao Bao, Lihong Wang, Yangqiu Song, and Qiang Yang. Large-scale hierarchical text classification with recursively regularized deep graph-cnn. In Proceedings of the 2018 World Wide Web Conference on World Wide Web, WWW 2018, Lyon, France, April 23-27, 2018, pages 1063–1072, 2018.

[85] Joost Bastings, Ivan Titov, Wilker Aziz, Diego Marcheggiani, and Khalil Sima'an. Graph convolutional encoders for syntax-aware neural machine translation. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9-11, 2017, pages 1957–1967, 2017.

[86] Diego Marcheggiani, Joost Bastings, and Ivan Titov. Exploiting semantics in neural machine translation with graph convolutional networks. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT, New Orleans, Louisiana, USA, June 1-6, 2018, Volume 2 (Short Papers), pages 486–492, 2018.

[87] Diego Marcheggiani and Ivan Titov. Encoding sentences with graph convolutional networks for semantic role labeling. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9-11, 2017, pages 1506–1515, 2017.

[88] Yuhao Zhang, Peng Qi, and Christopher D. Manning. Graph convolution over pruned dependency trees improves relation extraction. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018, pages 2205–2215, 2018.

[89] Thien Huu Nguyen and Ralph Grishman. Graph convolutional networks with argument-aware pooling for event detection. In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018, pages 5900–5907, 2018.

[90] Xiao Liu, Zhunchen Luo, and Heyan Huang. Jointly multiple events extraction via attention-based graph information aggregation. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018, pages 1247–1256, 2018.

[91] Yue Zhang, Qi Liu, and Linfeng Song. Sentence-state lstm for text representation. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 317–327. Association for Computational Linguistics, 2018.

[92] Linfeng Song, Yue Zhang, Zhiguo Wang, and Daniel Gildea. A graph-to-sequence model for amr-to-text generation. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1616–1626. Association for Computational Linguistics, 2018.

[93] Daniel Beck, Gholamreza Haffari, and Trevor Cohn. Graph-to-sequence learning using gated graph neural networks. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers, pages 273–283, 2018.

[94] Linfeng Song, Zhiguo Wang, Mo Yu, Yue Zhang, Radu Florian, and Daniel Gildea. Exploring graph-structured passage representation for multi-hop reading comprehension with graph neural networks. CoRR, abs/1809.02040, 2018.

[95] Linfeng Song, Yue Zhang, Zhiguo Wang, and Daniel Gildea. N-ary relation extraction using graph state lstm. arXiv preprint arXiv:1808.09101, 2018.

[96] Petar Velickovic, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks. CoRR, abs/1710.10903, 2017.

[97] Deyi Xiong, Shuanglong Li, Qun Liu, Shouxun Lin, and Yueliang Qian. Parsing the penn chinese treebank with semantic knowledge. In in Proceedings of IJCNLP 2005, pages 70–81.

[98] Xiaojun Lin, Yang Fan, Meng Zhang, Xihong Wu, and Huisheng Chi. Refining grammars for parsing with hierarchical semantic knowledge. In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, EMNLP 2009, 6-7 August 2009, Singapore, A meeting of SIGDAT, a Special Interest Group of the ACL, pages 1298–1307, 2009.

[99] Guangyou Zhou, Li Cai, Kang Liu, and Jun Zhao. Improving dependency parsing with fined-grained features. In Fifth International Joint Conference on Natural Language Processing, IJCNLP 2011, Chiang Mai, Thailand, November 8-13, 2011, pages 228–236, 2011.

[100] 孙茂松，陈新雄. 借重于人工知识库的词和义项的向量表示: 以 HowNet 为例. 中文信息学报, 30(6):1–6, 2016.

[101] Lvexing Zheng, Houfeng Wang, and Xueqiang Lv. Improving chinese dependency parsing with lexical semantic features. In Natural Language Processing and Chinese Computing - 4th CCF Conference, NLPCC 2015, Nanchang, China, October 9-13, 2015, Proceedings, pages 36–46, 2015.

[102] Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel P. Kuksa. Natural language processing (almost) from scratch. Journal of Machine Learning Research, 12:2493–2537, 2011.

[103] Makoto Miwa and Mohit Bansal. End-to-end relation extraction using lstms on sequences and tree structures. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016, August 7-12, 2016, Berlin, Germany, Volume 1: Long Papers, 2016.

[104] Marek Rei. Semi-supervised multitask learning for sequence labeling. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers, pages 2121–2130, 2017.

[105] Barbara Plank. Keystroke dynamics as signal for shallow syntactic parsing. In COLING 2016, 26th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers, December 11-16, 2016, Osaka, Japan, pages 609–619, 2016.

[106] Anders Søgaard and Yoav Goldberg. Deep multi-task learning with low level tasks supervised at lower layers. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016, August 7-12, 2016, Berlin, Germany, Volume 2: Short Papers, 2016.

[107] Prajit Ramachandran, Peter J. Liu, and Quoc V. Le. Unsupervised pretraining for sequence to sequence learning. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9-11, 2017, pages 383–391, 2017.

[108] Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2018, New Orleans, Louisiana, USA, June 1-6, 2018, Volume 1 (Long Papers), pages 2227–2237, 2018.

[109] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: pre-training of deep bidirectional transformers for language understanding. CoRR, abs/1810.04805, 2018.

[110] Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. Improving language understanding by generative pre-training. Available: https://blog.openai.com/language-unsupervised/, 2018.

[111] Jie Yang, Yue Zhang, and Fei Dong. Neural word segmentation with rich pretraining. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers, pages 839–849, 2017.

[112] Jeremy Howard and Sebastian Ruder. Universal language model fine-tuning for text classification. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 1: Long Papers, pages 328–339, 2018.

[113] Eliyahu Kiperwasser and Miguel Ballesteros. Scheduled multi-task learning: From syntax to translation. Transactions of the Association for Computational Linguistics (TACL) , 6:225–240, 2018.

[114] Anders Søgaard and Joachim Bingel. Identifying beneficial task relations for multi-task learning in deep neural networks. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017, Valencia, Spain, April 3-7, 2017, Volume 2: Short Papers, pages 164–169, 2017.

[115] Héctor Martínez Alonso and Barbara Plank. Multitask learning for semantic sequence prediction under varying data conditions. CoRR, abs/1612.02251, 2016.

[116] Lili Mou, Zhao Meng, Rui Yan, Ge Li, Yan Xu, Lu Zhang, and Zhi Jin. How transferable are neural networks in NLP applications? In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, EMNLP 2016, Austin, Texas, USA, November 1-4, 2016, pages 479–489, 2016.

[117] Joo-Kyung Kim, Young-Bum Kim, Ruhi Sarikaya, and Eric Fosler-Lussier. Cross-lingual transfer learning for POS tagging without cross-lingual resources. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9-11, 2017, pages 2832–2838, 2017.

[118] Xinchi Chen, Zhan Shi, Xipeng Qiu, and Xuanjing Huang. Adversarial multi-criteria learning for chinese word segmentation. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers, pages 1193–1203, 2017.

[119] Waleed Ammar, George Mulcaire, Miguel Ballesteros, Chris Dyer, and Noah A. Smith. Many languages, one parser. Transactions of the Association for Computational Linguistics (TACL), 4:431–444, 2016.

[120] Emmanouil Antonios Platanios, Mrinmaya Sachan, Graham Neubig, and Tom M. Mitchell. Contextual parameter generation for universal neural machine translation. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018, pages 425–435, 2018.

[121] Sara Stymne, Miryam de Lhoneux, Aaron Smith, and Joakim Nivre. Parser training with heterogeneous treebanks. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15-20, 2018, Volume 2: Short Papers, pages 619–625, 2018.

[122] Qi Liu, Yue Zhang, and Jiangming Liu. Learning domain representation for multi-domain sentiment classification. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2018, New Orleans, Louisiana, USA, June 1-6, 2018, Volume 1 (Long Papers), pages 541–550, 2018.

[123] Meishan Zhang and Yue Zhang. Combining discrete and continuous features for deterministic transition-based dependency parsing. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 1316–1321, Lisbon, Portugal, September 2015. Association for Computational Linguistics.

[124] Jie Yang, Zhiyang Teng, Meishan Zhang, and Yue Zhang. Combining discrete and neural features for sequence labeling. In Proceedings of the CICLing, 2016.

[125] Xavier Carreras and Lluís Màrquez. Introduction to the conll-2005 shared task: Semantic role labeling. In Proceedings of the Ninth Conference on Computational Natural Language Learning (CoNLL-2005), pages 152–164. Association for Computational Linguistics, 2005.

[126] Sameer Pradhan, Alessandro Moschitti, Nianwen Xue, Olga Uryupina, and Yuchen Zhang. Conll-2012 shared task: Modeling multilingual unrestricted coreference in ontonotes. In Joint Conference on EMNLP and CoNLL - Shared Task, pages 1–40, Jeju Island, Korea, July 2012. Association for Computational Linguistics.

[127] Jan Hajič, Massimiliano Ciaramita, Richard Johansson, Daisuke Kawahara, Maria Antònia Martí, Lluís Màrquez, Adam Meyers, Joakim Nivre, Sebastian Padó, Jan Štěpánek, et al. The conll-2009 shared task: Syntactic and semantic dependencies in multiple languages. In Proceedings of the Thirteenth Conference on Computational Natural Language Learning: Shared Task, pages 1–18. Association for Computational Linguistics, 2009.

[128] Gina-Anne Levow. The third international chinese language processing bakeoff: Word segmentation and named entity recognition. In Proceedings of the Fifth SIGHAN Workshop on Chinese Language Processing, pages 108–117, 2006.

[129] Wenliang Chen, Yue Zhang, and Min Zhang. Feature embedding for dependency parsing. In Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, pages 816–826, 2014.

[130] Wenliang Chen, Min Zhang, and Yue Zhang. Semi-supervised feature transformation for dependency parsing. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, EMNLP 2013, 18-21 October 2013, Grand Hyatt Seattle, Seattle, Washington, USA, A meeting of SIGDAT, a Special Interest Group of the ACL, pages 1303–1313, 2013.

[131] Yue Zhang and Stephen Clark. Syntactic processing using the generalized perceptron and beam search. Computational Linguistics, 37(1):105–151, 2011.

2. 项目的研究内容、研究目标，以及拟解决的关键科学问题（此部分为重点阐述内容）；

2.1. 研究内容

本项目通过创新的深度学习模型，研究综合的知识挖掘与利用。我们把问题拆解为四个层面，设置研究内容。

1) 研究有监督领域迁移情景下，多任务结构知识的挖掘与融合。

本项目关注的跨领域知识，具有两种表达形式，包括：(1) 显式的结构化表达形式，例如命名实体短语列表、谓词论元结构关系图等；(2) 隐式的参数化表达形式，例如神经网络参数、向量表达等。其中，前者在传统统计领域适应工作中应用较多，而后者在神经网络工作中应用较广。两种表达形式优势互补。有监督的迁移学习适合前者的利用。在之前图 2 的场景中，科技文本领域是一个有监督的迁移学习场景，可以直接利用一定的特定领域成分句法标注，而社会媒体文本领域是一个无监督的迁移学习场景，没有直接资源。我们计划从相关任务的纵向跨领域对比，以及相关任务和句法分析任务横向对比之中，挖掘公共结构知识。在测试时，我们将研究来自不同任务的多元异构知识的有效融合。这个研究内容将通过创新的深度学习的技术路线开展，解决相应的技术挑战，以达到资源高效利用。（研究方案详见 3.2 节）

2）研究有监督和无监督领域迁移情景下，参数化知识的挖掘与融合。

相比结构化的知识，参数化的知识表达更加灵活和抽象。缺点在于难以直观解释，但优势在于可以自动学习和高效融合。本项目计划研究参数化的知识挖掘在跨任务跨领域关联中的创新作用。具体而言，领域间的知识可以分为共性知识和特性知识。共性知识的提取是深度迁移学习中讨论较多的话题，而特性知识讨论较少。特性知识的跨领域相互关联研究更是有限。我们曾经探索单任务下结构化的 [18, 20] 和参数化的 [112, 123] 跨领域特性。在多任务情况下，知识迁移渠道更多，技术挑战更大，而参数化表达优势可能更大。本项目将通过参数化的特性知识抽取，研究领域之间细粒度的相互关系，促进相关领域之间更有效的迁移学习，同时通过参数化的特性知识研究领域关系可视化。这项研究的另一个创新内容在于，以相关任务为桥梁，学习参数化的成分句法领域差异，从而实现无监督的领域迁移，使目标领域无语料学习（zero-shot learning）成为可能。（研究方案详见 3.2 节 3.3 节）

3) 探索有效融合跨任务跨领域知识的创新算法

我们的多任务多领域场景给迁移学习提出了技术挑战。同时，如何有效地融合显式结构化的和隐式参数化的知识，需要方法层面的创新。我们计划构建一套多任务学习的框架体系，对来自不同领域、通过不同任务、以不同形式表达的特

性知识和共性知识进行分离和融合。其中，共性和特性的分离主要在机器学习中的训练过程，而知识的融合既涉及训练过程又涉及测试过程。传统的多任务学习方法难以解决任务和领域特性之间的组合爆炸问题。因此不能直接应用于本项目。我们的初步工作 [122] 探索到一条解决上述问题的有效途径。我们计划以此为基础深入探索针对本项目更加有效的技术和方法。(研究方案详见 3.3 节)

4) 探讨不同资源条件对迁移学习性能的影响

哪些类型的资源对跨领域成分句法分析帮助最大是一个有价值的研究课题。我们将从不同场景、不同任务、不同数据量等角度研究这个问题。其中场景包括有监督的迁移学习和无监督的迁移学习。通过场景对比，可以挖掘来自句法任务的直接信息和来自相关任务的间接信息对句法迁移的影响。任务方面，不同任务对跨领域句法知识的贡献具有不同形式。此外，相同任务在不同场景下也可能对句法分析有不同影响。我们计划从节省资源、提高效率的角度对不同任务进行对比。(研究方案详见 3.4 节)

2.2. 研究目标

本项目将围绕以下目标进行研究：

1) 通过辅助任务的角度，揭示句法结构在跨领域场景下的定性和定量差别，明确跨领域句法分析的关键难点；

2) 通过多任务多领域的深度迁移学习，挖掘和融合成分句法知识；

3) 有效利用多元异构数据资源，实现跨领域稳定的成分句法分析器。

意义在于从基础研究的角度开拓广泛实用的自然语言理解算法，从而为语言处理技术的实际应用做出贡献。

2.3. 拟解决的关键科学问题

1) 不同相关任务对跨领域成分句法知识挖掘的效用

本项目分析和比较不同结构的相关任务对挖掘跨领域句法知识的效用，从而从相关任务的角度衡量句法结构的领域差异。定量衡量不同领域句法结构的根本差异是一个有挑战的课题。相关任务的跨领域差异可以作为参照，为更深理解领域之间句法关系提供更丰富的衡量角度。通过比较跨领域语言模型知识和语义角色任务对跨领域句法分析性能的影响，可以衡量词汇分布和语义知识对跨领域句法结构的相对重要性。本项目将从各个相关任务的角度定量衡量不同关键因素，找到不同领域和新闻领域之间句法差异，以及它们各自的挑战和瓶颈。阐明这个问题可以为解决跨领域句法挑战提供基础和研究思路。

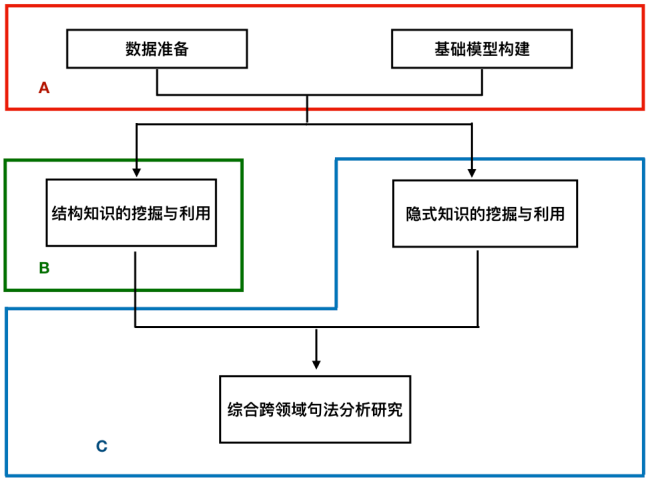

图 3: 研究方案流程图

2）如何有效利用结构化的多元异构跨领域知识

相关任务为句法分析提供结构化的跨领域知识。由于不同任务本质上具有不同结构，再加上领域共性和领域特性的区别，使得结构性知识具有种类多样内容多样的特点，从而给其在成分句法分析中的有效利用提出了技术挑战。数学上，图是融合多元结构的有效表达工具和手段。我们计划通过图的形式容纳不同结构的知识，并通过统一的深度学习表示，探索上述知识的有效利用。由于复杂图存在环状结构，传统的深度表示学习模型 [69, 71, 73, 75, 77, 79] 难以对上述知识结构进行编码。针对图表示的新的神经网络正在成为一个研究前沿问题。我们将以近一年来对复杂图表示学习的探索 [91, 93, 94, 95] 为基础，开展本项目的这个科学问题的研究工作。

3）如何有效融合各类资源实现跨领域稳定的成分句法分析

这个问题直接关系到项目的最终目标。前期工作证明，对于单个任务，结构化的表示和深度学习表示可以提供互补的信息 [123, 124]。本项目探讨如何融合结构化和参数化的知识，并以此融合跨任务跨领域的资源。我们将以近期研究 [122] 作为出发点，通过一个新的多任务学习框架进行探索。除知识挖掘以外，这套框架将用于构建最终的句法分析模型，以便最有效的利用多种跨领域的资源。

3. 拟采取的研究方案及可行性分析（包括研究方法、技术路线、实验手段、关键技术等说明）；

本项目计划基于深度学习算法开展研究工作。研究方案过程图见图 3，相关的深度学习结构框图见图 4。其中，两图标注 A、B、C 的相关部分相互对应。

从实验流程的角度，我们将按照图 3 从上到下的时间顺序开展课题工作。首先，我们将完成数据准备和基础模型的构建，用方便多任务学习的模型结构实现

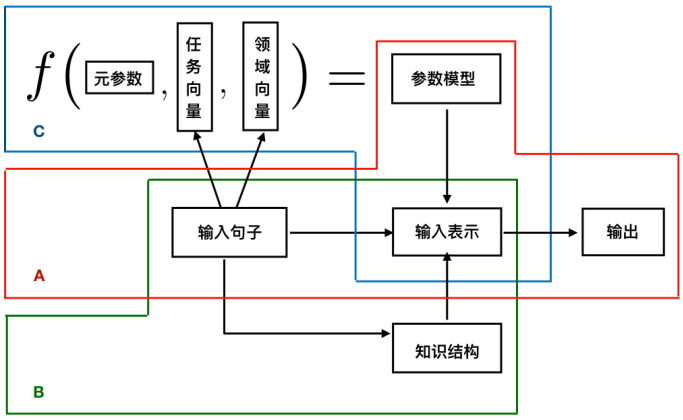

图 4: 技术方案结构图（整体模型）

文献中各自任务前沿水平。3.1 节详细介绍项目相关基础模型的构建、语料的准备以及相关任务的选择。这部分内容对应图 3A。在此基础之上，我们将同时开展结构化的显式知识挖掘和基于多任务学习的隐式知识挖掘研究工作。其中，前者对应图 3B 的内容，在 3.2 节详细介绍。后者对应图 3C 的内容，在 3.3 节详细介绍。最后，我们将综合各类多任务知识，研究跨领域稳定的成分句法分析。这个内容在 3.3 节和 3.4 节做详细介绍。

从技术方案的角度，基本的模型结构如图 4A 部分所示。给定一个输入句子，模型通过一套参数计算出它的向量表示，再根据这个向量表示预测输出结构。这部分内容在 3.1 节介绍。相应图 4B，我们通过相关任务挖掘和表达跨领域的句法知识，并且通过基于图的神经网络结构把知识加入到句子表示之中。这部分研究内容在 3.2 节介绍。相应图 4C，我们研究一种新的多任务学习方法，把跨领域跨任务的知识分解存储到一套元参数、一套领域向量和一套任务向量之中，通过对抗学习（adversarial training）等正则化 (regularization) 方法实现知识分离以及迁移。这套多任务学习的体系结构将被用于融合多方面知识的最终句法分析模型。这部分研究内容在 3.3 节详细介绍。整体模型结构下的定量实验和参数优化在 3.4 节介绍。

最后，上述实施方案的可行性分析在 3.5 节详细介绍。

3.1. 数据准备和基线模型构建

3.1.1 句法分析任务

本项目研究的核心任务是成分句法分析。相同模型结构将被用于中文和英文的成分句法分析。这和当前国内外主流句法分析研究方式是同步的：文献中很多成分句法分析的相关工作同时汇报英文和中文数据集上的结果 [28, 30]。在数据方面，我们将分别使用宾州中文树库（Penn Chinese Treebank）和宾州树库 (Penn

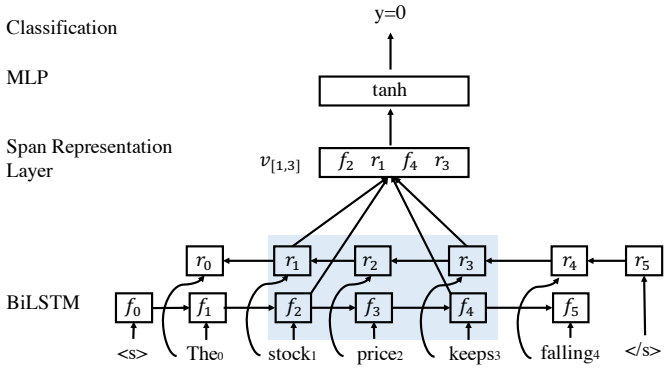

图 5: 基线句法分析模型

Treebank) 作为中文和英文的新闻领域数据集。我们计划在社会媒体、科技、体育、文学、日常对话等五个领域标注少量测试集。其中，每个领域在中、英文各标 1000 个句子。初步计划，在中文，社会媒体语料使用新浪微博为原始文本，科技语料使用生物医药数据为原始文本（http://www.bioon.com.cn/），体育语料使用新浪体育等体育新闻为原始文本，文学语料使用起点中文网（https://www.qidian.com）小说为原始文本，日常对话使用电影对白脚本为原始文本。在英文，社会媒体语料将使用推特 (https://www.twitter.com) 为原始文本，科技语料使用 PubMed 数据库 (https://www.ncbi.nlm.nih.gov/pubmed) 为原始文本，体育语料计划使用 MSN Sports (https://www.msn.com/en-us/sports) 为原始文本，文学语料使用经典英语小说 (https://www.24en.com) 为原始文本，日常对话使用电影对白脚本为原始文本。

方法上，我们将使用我们的近期工作 [35] 作为基线系统。这项工作的模型结构如图 5 所示，其中包括一个多层双向循环神经网络（BiLSTM）编码器和一个非常轻量级的局部预测系统。这种模型结构的一个优势在于，把大量参数集中在了对输入文本的编码上面，而不去过分强调输出结构的非局部依赖关系。以上结构为多任务学习中共享参数提供了有利条件，非常符合本项目的需求。与此同时，该模型结构也取得了文献中标准数据集上非常有竞争力的结果。我们的初步实验显示，在引入上下文相关的词向量表示（如 ELMo, GPT, BERT）[108, 110, 109] 之后，性能还可以大幅提高。此外，除了双向循环神经网络（BiLSTM）编码器，我们还会考虑自注意力机制网络（self-attention network [73]）作为编码器的替换选择。

3.1.2 相关任务

我们计划研究三种不同结构的相关任务，包括语言模型结构、序列标注结构以及二元关系结构。首先，新闻领域和各个目标领域的未经标注语料将被用于训练不同种类型的语言模型。我们将以双向循环神经网络（BiLSTM）和自注意力机

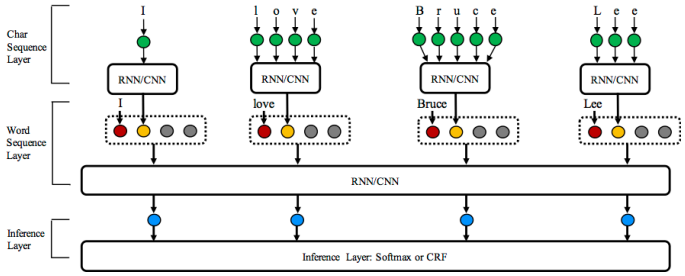

图 6: 序列标注系统

制网络 (self-attention network) 作为主要的语言模型结构进行研究。

第二，我们将以命名实体识别任务作为主要的序列标注任务，用双向循环神经网络（BiLSTM）或自注意力机制网络（Transformer）加条件随机场（CRF）的模型作为基础模型。如图 6 所示，我们的近期工作 [91] 将被作为序列标注基础模型实现。这个实现取得了文献中非常有竞争力的实验精确度。训练语料将包括英文的 CoNLL[125, 126] 数据集、中文的 CoNLL[127] 和 MSRA[128] 新闻领域数据集。我们将标注一定数量的跨领域文献数据集作为跨领域对比基础。

第三，我们将以语义角色标注作为主要的二元关系结构，用双向循环神经网络（BiLSTM）或自注意力机制网络（Transformer）为基础的双仿射（bi-affine）模型 [49] 作为基线模型。我们计划用近期工作 [15] 作为语义角色标注模型实现。这个实现取得了文献中有竞争力的精确度。训练语料包括中英文的 CoNLL[127, 125, 126] 新闻领域数据集。我们也将标注一定数量的跨领域数据集作为跨领域分析的基础。上述各个不同任务模型和我们的句法分析模型基线系统具有相匹配的编码器 (encoder) 结构，方便进行多任务学习。

对于上述所有任务，我们将以大规模训练过的基于上下文的表示 (contextualized representation)，例如 ELMo[108]，BERT[109] 和 GPT[110]，作为输入。

3.2. 结构知识的挖掘和利用

本项目的知识表示框架如图 7 所示。给定一个输入，我们先获得和其相关的结构化的领域公共知识和领域特殊知识，然后将这些知识和输入句子一起进行神经网络编码，得到每个词的基于上下文的隐层向量表示。这些隐层向量表示将被用作 3.1 节提到的各个任务的多任务学习表示。在图 7 中，输入的句子是“该算法通过自注意力机制提高命名实体识别性能”。其中，“自注意力机制”和“命名实体识别”是两个领域相关的命名实体，（“通过”，“机制”），（“机制”，“提高”）和（“提高”，“性能”）是三个跨领域共同的语义角色关系，而（“算法”，“提高”）是一个领域相关的语义角色关系。

本项目通过相关任务的输出结构表示结构化的跨领域知识。对于命名实体识别任务，知识结构可以用一个命名实体词汇表的形式储存。对于语义角色标注任

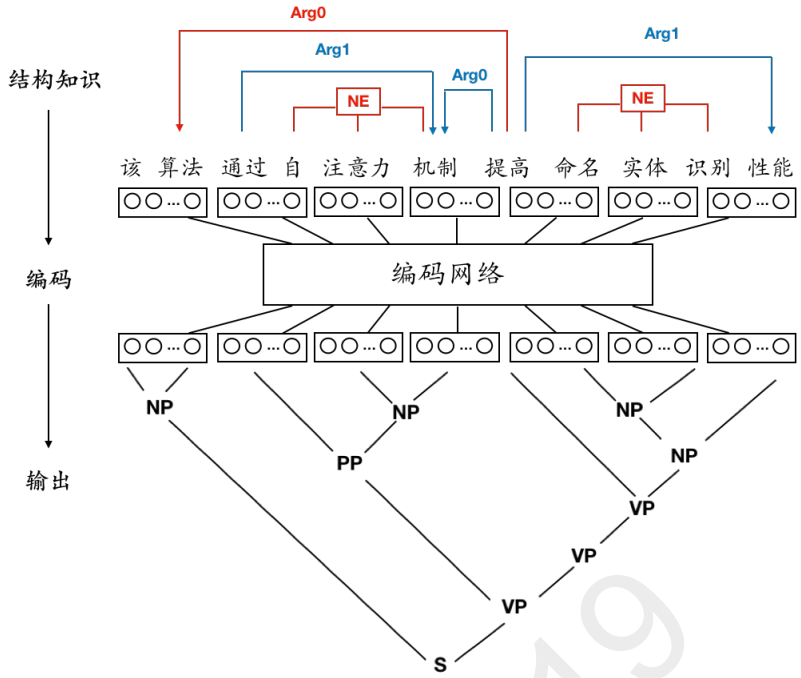

图 7: 结构知识的利用

务，知识结构可以用一个谓词论元关系图谱形式存储。为得到跨领域共同和领域特殊的结构知识，对于每个相关任务，我们通过新闻领域的语料训练 3.1 节所述的基础模型。之后，用所得到的基础模型自动分析各个领域的大量文本。我们将首先通过对比统计的方法，针对每个任务找到高置信度的跨领域公共结构和领域特殊结构。这项工作本身是一个具有研究价值的问题。我们将以前期工作为基础 [129, 130]，构建一个统计模型的基线系统。在此基础之上，我们计划探索深度学习框架下，借助预测和对抗等手段，分离共性与特性结构知识的方案。

如图 7 所示，给定一个输入，我们将结构知识库和输入进行匹配，得到一个以词为结点，知识结构为边的图结构。在这个图中，不同任务和不同性质（共性、特性）的边具有不同标记。如何有效融合这些知识，是一个关键问题。我们计划以近一年来对于图表示的神经网络的探索 [91, 92, 95] 为基础，进行相关研究。具体而言，我们研究了一种基于图的循环神经网络，其基本结构如图 8 所示。给定一个输入和其多元异构知识图，我们把整个图结构当成一个大的状态 (g)，通过反复迭代，更新图中每一个节点的状态 (h) 表示。图的初始状态可以是每个节点的词向量。在每一次更新中，一个节点和它的所有相邻节点进行信息交换。不同种类的相邻边采取不同的信息传递方式传递不同知识结构中的信息。这种机制确保不同来源和不同任务的信息得到区分。和传统的基于序列的和基于树的循环神经网络相比，我们的网络结构可以解决图中的环表示，同时保持循环神经网络的结构优势。我们将通过这种方法对结构知识利用进行深入研究。

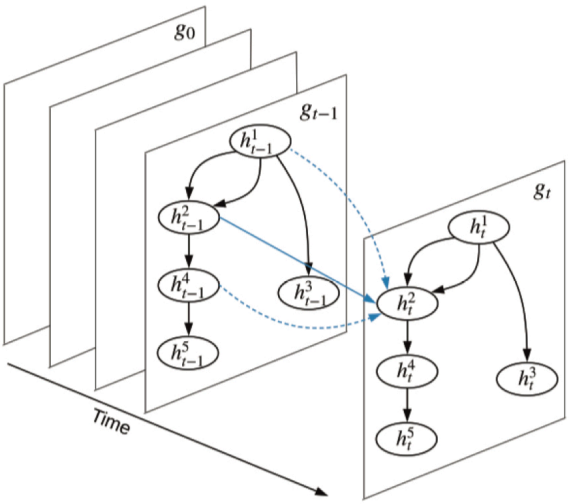

图 8: 图循环神经网络（g：图状态；h：结点状态；t：迭代）

3.3. 融合多种知识的神经网络方法

图 7 中的神经网络承担对给定输入进行编码并且预测输出的作用。如 3.1 节所述，本项目讨论的不同任务，都采用共同的神经元网络编码器。因此，图 7 中的网络结构具有任务共性。然而，对于不同任务和不同领域，神经网络的具体参数并不相同。同时，跨任务跨领域的神经网络参数必须存在联系，以便有效地从不同的任务之间抽取和融合有用的关联知识。这个多领域、多任务的场景给迁移学习提出了技术挑战，也是本项目研究的重点技术问题。

本项目计划以隐式参数化的领域特性表示为出发点对上述问题进行研究。我们的初步研究 [122] 显示，与存储分布式语义表示的词向量类似，领域特性知识可以表示为向量形式。在前期工作 [122] 中，我们把一个输入句子的深度学习表示分解成一个领域无关的表示（common representation）和一个领域向量（domain vector）的组合，以此分离共性知识和特性知识。实验证明，该方法有效地实现了五个不同领域情感分析数据的融合，提高单个模型对每个领域的精确度。

本项目计划将上述研究中的技术深入化，把现有句子表示的分解方法拓展为模型参数的分解方法。给定一个输入，我们根据一套与领域、任务无关的元参数 (meta parameter)，一个领域向量 (domain vector)，以及一个任务向量 (task vector) 共同计算出一套具体的模型参数 (model parameter)：

$$\text{具体参数} = \text{f}\,(\text{元参数}, \text{领域向量}, \text{任务向量})$$

针对每一个任务和领域，图 7 中的神经网络参数都由上式得到。这套框架允许我们把领域特性、任务特性和共性知识分解到相互独立的参数中。其中，元参数存储跨领域跨任务的共性知识，任务向量存储每个任务的特性知识，而领域向量存储每个领域的特性知识。与词向量（word embedding）类似，领域向量和任务向量是

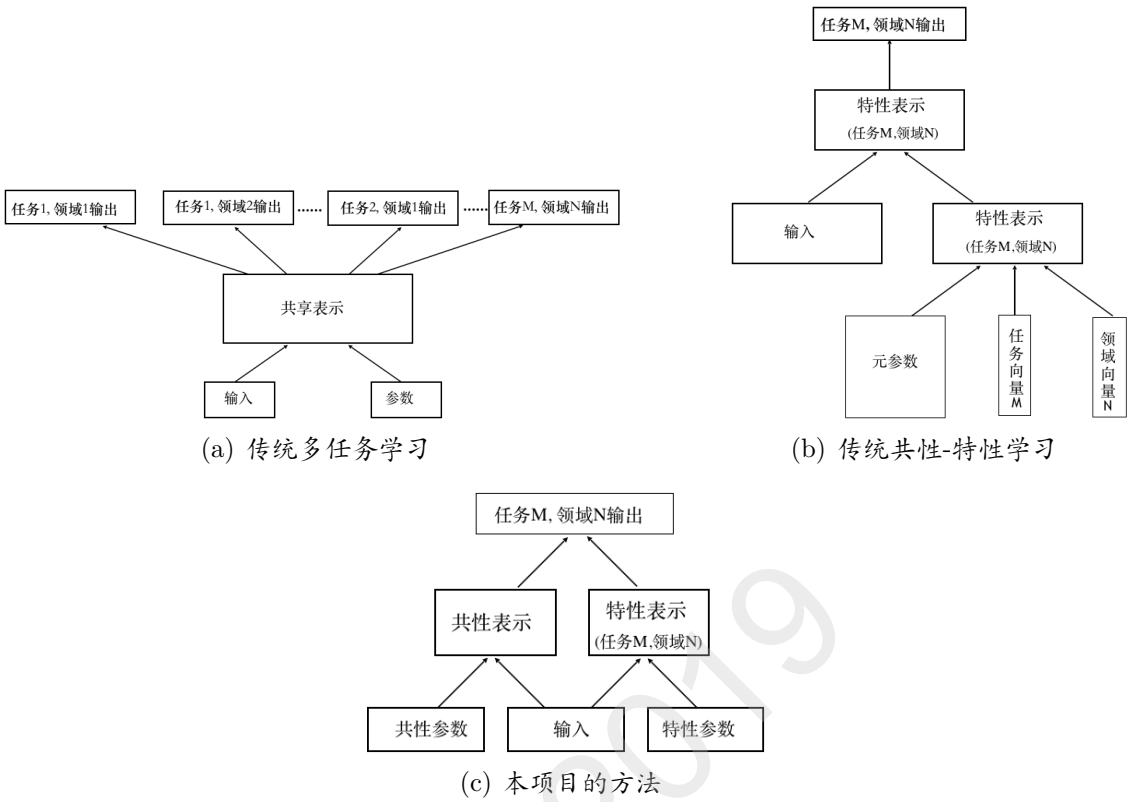

图 9: 跨领域跨任务知识融合方法

一种非稀疏的知识存储方式。同时，向量表示的特性知识便于可视化，有助于分析领域和任务特性之间的相互关系。

对抗学习（adversarial training）等正则化（regularization）手段可以加强共性和特性知识的分离。具体而言，我们可以通过元参数直接预测一个输入的领域或者要解决的任务（把元参数直接转换成具体参数的方法有很多，比如结合全部元素为 1 的任务向量和领域向量，在我们的初步实验中行之有效），通过对抗损失函数的方式确保预测结果具有最大熵，以此确保元参数中只含有领域和任务的共性特征。也可以通过某个任务和领域的具体参数预测该领域和该任务，以此确保特性信息集中的存储到领域向量和任务向量之中。

图 9 展示了本项目的方法（图 9（c））与传统多任务（multi-task）学习方法和传统共性-特性（shared-private）学习方法的对比。其中，图 9（a）展示了传统的多任务学习方法。该方法通过一套共享的表示层和针对每个任务领域组合的输出层提炼共性知识。与之相比，本项目方法的优势在于对共性和特性同时进行建模。虽然多任务学习的方法可以应用到我们的场景中，但是它们难以处理众多任务领域组合之间的信息矛盾冲突和领域任务组合爆炸的问题。

图 9（b）展示了传统的共性-特性学习方法。该方法对每一个任务和领域的组合保存一套特性参数（private parameters），并且为所有领域共性保存一套参数（shared parameters）。针对一个输入，通过两套参数分别得到它的共性表示和特性

表示，并把两个表示组合起来得到最终表示。与这种方法相比，本项目的方法具有三大优势。首先，传统方法每个领域的特性之间不存在联系，因此不能在领域和任务之间进行细粒度的知识传递。本项目的方法允许相关领域通过类似的向量生成相似的参数，因此，可以在任务和领域组合很大的情况下进行细致的知识挖掘。其次，传统方法需要针对每个任务和领域的组合进行训练。对于 M 个领域和 N 个任务，需要 M×N 套训练数据。本项目的方法允许通过少量的任务和领域组合推导出其他任务领域组合所需要的参数。对于 M 个领域和 N 个任务，最小只需要 max(M,N) 套训练数据。最后，相比传统方法，本项目使用向量而非矩阵存储特性知识，因此可以很大程度的节省空间。

3.4. 定量实验与模型优化

在上述各节的技术框架下，我们将进行系统的定量实验，回答 2.3 节中的科学问题。其中，主要的实验课题可以划分为四个大类。第一是有监督和无监督学习场景的影响，即来自句法结构的直接知识和来自相关任务的间接知识的影响对比。第二是不同任务以及任务设置之间的影响对比。第三是结构化的表达和参数化的表达之间的影响对比。第四是不同数据量的影响对比。

此外，参数调节与系统最终优化也是重要的实施步骤。在深度学习技术中，参数的调节与选择对于 2.3 节中的问题 2 和问题 3 可能起到重要作用，是解决科学问题必不可少的工程实践。

3.5. 可行性、可靠性论证

理论上，本项目提出的综合相关资源提高跨领域成分句法分析的思路是可行的。首先，从语言学的角度，成分句法和命名实体、语言模型等任务存在密不可分的联系。这些理论联系支持本项目的研究假设，即相关任务的跨领域知识可以给成分句法分析提供帮助。其次，文献中的相关工作和我们的前期工作表明，在单一领域下，相关任务知识可以提高句法分析的精准度。这给跨领域实验的可行性提供了有力支持。

技术上，团队在完成本项目课题所需要的三个关键环节掌握核心技术。其中，(1) 句法分析和所有相关任务的基线模型均采取我们的近期工作，性能达到文献中有竞争力的水平。(2) 结构知识挖掘和表示的算法，采用我们近期提出的基于图的神经网络。该方法在语言建模、语言生成等任务上展现了利用知识的有效性。(3) 多任务多领域知识融合的神经网络框架，是我们近期跨领域相关工作的扩展。前期工作已经证明了单任务多领域的可行性。我们近期的初步实验结果也证明在多任务多领域下该框架基本可行。

团队上，团队成员具有高级职称和初级职称的合理搭配，核心成员在申请人团

队工作四年以上，熟悉项目所需的关键技术，具有与申请人一起组织协调纵向项目课题的相关经历。此外，学校在基础设施条件方面给予申请人的团队足够支持，有助于顺利开展科研工作。

4. 本项目的特色与创新之处；

1）本项目从相关任务的角度挖掘跨领域的句法知识。与来自句法本身的直接知识相比，相关任务提供了更丰富的信息来源和表达方式。

2）本项目从多领域的角度研究相关任务对句法分析的影响。以语言模型为例，现有方法（如 ELMo, BERT, GPT）研究单个语言模型对句法的知识迁移，而本项目的方法通过对比多个语言模型挖掘跨领域句法知识。

3）本项目研究多元异构数据的融合，探索充分利用一切现有资源的跨领域句法分析。

4）为促进多任务环境下的跨领域句法分析，本项目探索技术层面创新，研究一种新的多任务和多领域融合的深度学习方法。

5. 年度研究计划及预期研究结果（包括拟组织的重要学术交流活动、国际合作与交流计划等）。

5.1. 年度研究计划

2020 年：完成所有基线模型调试；完成各领域句法分析测试集的标注；研究跨领域多个语言模型辅助任务对句法分析任务的影响，并与 ELMo 等基于单个语言模型的方法进行对比；撰写专利、论文等成果。

2021 年：在三个领域和三个任务上研究多任务学习框架对知识融合的影响，以及参数化的知识表示对资源的有效利用；研究基于图的循环神经网络对于跨领域单任务知识的表示，并将其应用到跨领域成分句法分析。撰写专利、论文等成果。

2022 年：把上述三个领域三个任务的研究成果扩展到多个领域和任务组合，并且通过向量表示研究领域和任务的可视化；在三个领域一个相关任务上研究同时融合参数化知识表示和结构化知识表示的方法；研究基于图的循环神经网络对于多个任务跨领域结构知识的融合表示，并将其应用到跨领域成分句法分析。撰写专利、论文等成果。

2023 年：把上述三个领域一个相关任务上的研究成果扩展到多个领域多个任务上；综合前三年定量实验结果，进行总体定量研究分析；研究外部结构化资源（如 WordNet、HowNet）的融入；构建最终的跨领域成分句法分析系统；为了最有效的利用所有已有资源进行参数调节。撰写专利、论文等成果，完成最终系统和成果总结。

5.2. **预期研究成果**

1) 数据：人工标注并公开 1 万句多领域中英文成分句法分析测试集，以促进跨领域成分句法分析研究。

2) 软件系统：开发高性能的多领域成分句法分析系统，分别在中文英文资源上训练模型。

3) 发表论文：国内外顶级期刊论文 2-4 篇，顶级会议论文 6-9 篇，其中 CCF A、B 类期刊或会议论文 4-6 篇。

4) 学术交流活动: 积极参加 ACL、NAACL、EMNLP、COLING 等重要的国际学术会议，以及 CCL、NLPCC 等重要的国内学术会议，扩大学术影响力，并积极邀请国内外相关领域知名学者访问交流。

5) 申请专利：1-3 项

6) 研究生培养：指导培养博士生 3-5 名。

（二）研究基础与工作条件

1. 研究基础 （与本项目相关的研究工作积累和已取得的研究工作成绩）；

申请人在相关领域从事了十年以上的研究工作，积累了本项目关键的核心技术，在句法分析、基于深度学习的自然语言处理、领域适应、多任务学习以及基于图的表示学习方面，做出了一系列贡献。近年来，由申请人指导的课题组，在自然语言处理和人工智能领域发表国际期刊论文十余篇，国际会议论文一百余篇，其中 CCF 列表 A、B 类 95 篇。申请人的谷歌引用超过 3035 次，H 因子 26。在国内外的相关领域的顶级会议 ACL、COLING、EMNLP、NAACL 上，申请人多次担任领域主席 (area chair)，并在相关领域的顶级期刊如 CL、TACL 和 TALLIP 等担任期刊审稿人、编委和副主编 (associate editor)。申请人完成的工作曾获 COLING 2018 和 IALP2017 最佳论文奖。申请人在 2018 年获得 CCF 自然语言处理与中文计算青年新锐奖。更多信息可参考申请人的个人主页 (Github 主页): https://frcchang.github.io/。这些研究成果和研究经验为本项目的开展提供了良好的基础。

本项目的核心技术包括句法分析、结构知识的表示利用和跨任务跨领域迁移学习三个方面。

1)句法分析。申请人提出的一套利用机器学习引导启发式搜索算法的结构预测模型 [131]，在词法分析和句法分析等基础自然语言处理任务取得了文献中领先的准确度与速度。比如，2013 年申请人研发的句法分析系统 ZPar 在宾州中英文树库标准数据集上比主要竞争者 Stanford Parser 和 Berkeley Parser 速度快 15 倍以上，

并具有更高准确度 [25]。开源句法分析器 ZPar (https://sourceforge.net/projects/zpar) 仅在 SourceForge 下载超过 7000 次。

近 5 年来，申请人的团队在成分句法分析任务上一直保持文献中有竞争力的实验结果。此外，在跨领域的句法分析研究方面，团队一直保持着前沿探索。近五年来发表相关 CCF A、B 类论文十余篇，在顶级会议 ACL 2014、CCL 2016 和 IJCNLP 2017 做过相关的前沿技术讲习班。

2）结构化的知识利用。申请人近两年提出基于图的循环神经网络结构。后者在语言建模、语言生成等任务上展现了利用知识的有效性。近一年来发表 CCF A、B 类相关论文 4 篇。

3）在多任务多领域知识融合学习框架的研究方面，申请人近五年来发表 CCF A、B 类相关论文十余篇，在顶级会议 EMNLP 2018、CCL 2017 和 NLPCC 2018 做过相关的前沿技术讲习班。

2. 工作条件 （包括已具备的实验条件，尚缺少的实验条件和拟解决的途径，包括利用国家实验室、国家重点实验室和部门重点实验室等研究基地的计划与落实情况）；

本项目的依托单位西湖大学是一所社会力量举办、国家重点支持的新型高校，以博士生培养为起点，并致力于探索现代科研体制和创新培养模式。西湖大学工学院以高端人才为学科带头人，努力建成国家重大科学技术研究和创新人才培养基地。

申请人为西湖大学工学院文本智能实验室负责人，课题组在自然语言处理领域理论研究和工程实践方面均有良好基础。课题组有副教授 1 名，助理研究员 2 名，科研助理 6 名，博士生 5 名 (另有 3 名即将于 9 月入学)，研究人员的梯队结构合理。

目前，申请人团队具有 40 个 GPU。西湖大学科研设施与公共仪器中心统筹管理大型仪器共享，提供超算平台，提供可用于服务计算开发的服务器集群一套 (40 个 CPU)，为课题提供了便利条件，在研究环境上有较充分的保障。

3. 正在承担的与本项目相关的科研项目情况 （申请人和项目组主要参与者正在承担的与本项目相关的科研项目情况，包括国家自然科学基金的项目和国家其他科技计划项目，要注明项目的名称和编号、经费来源、起止年月、与本项目的关系及负责的内容等）；

申请人张岳作为第三参与人参与自然科学基金面上项目：中文句法语义分析与开放域信息抽取融合技术研究，批准号：61572245，起止年月：2016/01-2019/12。研究探索中文句法在知识领域，主要是在开放领域信息抽取融合技术上的应用问题。主要负责算法设计与研究。与本项目的关系：它是此项目句法分析的一个应

用。

4. **完成国家自然科学基金项目情况**（对申请人负责的前一个已结题科学基金项目（项目名称及批准号）完成情况、后续研究进展及与本申请项目的关系加以详细说明。另附该已结题项目研究工作总结摘要（限 500 字）和相关成果的详细目录）。

无

（三）其他需要说明的问题

1. 申请人同年申请不同类型的国家自然科学基金项目情况（列明同年申请的其他项目的项目类型、项目名称信息，并说明与本项目之间的区别与联系）。

无

2. 具有高级专业技术职务（职称）的申请人或者主要参与者是否存在同年申请或者参与申请国家自然科学基金项目的单位不一致的情况；如存在上述情况，列明所涉及人员的姓名，申请或参与申请的其他项目的项目类型、项目名称、单位名称、上述人员在该项目中是申请人还是参与者，并说明单位不一致原因。

无

3. 具有高级专业技术职务（职称）的申请人或者主要参与者是否具有高级专业技术职务（职称）的申请人或者主要参与者是否存在与正在承担的国家自然科学基金项目的单位不一致的情况；如存在上述情况，列明所涉及人员的姓名，正在承担项目的批准号、项目类型、项目名称、单位名称、起止年月，并说明单位不一致原因。

无

4. 其他。

无

张岳 简历

西湖大学，工学院，研究员

教育经历（从大学本科开始，按时间倒序排序；请列出攻读研究生学位阶段导师姓名）：

（1）2006.09－2009.12，牛津大学，计算机科学，博士，导师：Stephen Clark

（2）2005.09－2006.10，牛津大学，计算机科学，硕士，导师：Stephen Clark

（3）1999.09－2003.07，清华大学，计算机科学与技术，学士，导师：无

科研与学术工作经历（按时间倒序排序；如为在站博士后研究人员或曾有博士后研究经历，请列出合作导师姓名）：

（1）2018.09-至今，西湖大学，工学院，研究员

（2）2012.07-2018.08，新加坡科技设计大学，信息管理与设计学院，助理教授

（3）2010.03-2012.06，剑桥大学，博士后，合作导师：Stephen Clark

曾使用其他证件信息（申请人应使用唯一身份证件申请项目，曾经使用其他身份证件作为申请人或主要参与者获得过项目资助的，应当在此列明）：

主持或参加科研项目（课题）情况（按时间倒序排序）：

横向项目-北京融汇金信信息技术有限公司，中国资本市场文本智能研究，2018/11-2019/10，78万元，在研，主持

横向项目-推文科技，推文实体抽取和机器翻译系统，2018/11-2019/10，30万元，在研，主持

横向项目-Blue Fire AI PTE.LTD, China capital market and unstructured data，2017/03-2018/08，34.4万新加坡元，已结题，主持

纵向项目-新加坡国防部种子项目，Deep learning for Singlish Parsing， 2016/06-2017/05， 4.9万新加坡元，已结题，主持

国家自然科学基金面上项目，61572245，中文句法语义分析与开放域信息抽取融合技术研究，2016/01-2019/12，64万新加坡元，在研，参与

纵向项目-新加坡国防部项目，Cross-functional information systems for decision making，2014/12-2018/06，66万新加坡元，已结题，主持

纵向项目-新加坡教育部Tier2基础科研项目，Statistical machine translation by syntactic transfer，2013/06-2016/05，57.5万新加坡元，已结题，主持

代表性研究成果和学术奖励情况

（请注意：①投稿阶段的论文不要列出；②对期刊论文：应按照论文发表时作者顺序列出全部作者姓名、论文题目、期刊名称、发表年代、卷（期）及起止页码（摘要论文请加以说明）；③对会议论文：应按照论文发表时作者顺序列出全部作者姓名、论文题目、会议名称（或会议论文集名称及起止页码）、会议地址、会议时间；④应在论文作者姓名后注明第一/通讯作者情况：所有共同第一作者均加注上标“#”字样，通讯作者及共同通讯作者均加注上标“*”字样，唯一第一作者且非通讯作者无需加注；⑤所有代表性研究成果和学术奖励中本人姓名加粗显示。）

按照以下顺序列出：①代表性论著（包括论文与专著，合计5项以内）；②论著之外的代表性研究成果和学术奖励（合计10项以内）。

一、代表性论著

(1) **Yue Zhang**(#)(*); Qi Liu; Linfeng Song, Sentence-State LSTM for Text Representation, 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, 2018.07.15-2018.07.20 **（会议论文）**

(2) **Yue Zhang**(#)(*); Jie Yang(#), Chinese NER Using Lattice LSTM, 56th Annual Meeting of the Association for Computational Linguistics , Melbourn, 2018.07.15-2018.07.20 **（会议论文）**

(3) Jie Yang(#); Shuailong Liang; **Yue Zhang**(*), Design Challenges and Misconceptions in Neural Sequence Labeling, 27th International Conference on Computational Linguistics, Santa Fe, New Mexico, 2018.8.20-2018.8.26 **（会议论文）**

(4) Hongmin Wang(#); **Yue Zhang**(*); GuangYong Leonard Chan; Jie Yang; Hai Leong Chieu, Universal Dependencies Parsing for Colloquial Singaporean English, 55th Annual Meeting of the Association for Computational Linguistics, Vancouver , 2017.07.30-2017.08.04 **（会议论文）**

(5) **Yue Zhang**(#); Stephen Clark, Syntactic Processing Using the Generalized Perceptron and Beam Search, Computational Linguistics, 2010.09, 37(1): 105~151 **（期刊论文）**

二、论著之外的代表性研究成果和学术奖励

(1) **Yue Zhang**(*) (3/3), COLING 2018 最佳论文奖(Best Paper Award), Design Challenges and Misconceptions in Neural Sequence Labeling, 27th International Conference on Computational Linguistics, 其他, 其他, 2018.8.26 (Jie Yang(#); Shuailong Liang; **Yue Zhang**(*)) **（科研奖励）**

(2) **Yue Zhang**(*) (2/2), IALP 2017 最佳论文(Best papaer award), Joint Bi-Affine Parsing and Semantic Role Labeling., 21th International Conference on Asian Language Processing, 其他, 其他, 2017.12.5
(Peng Shi(#); **Yue Zhang**(*)) **(科研奖励)**

(3) **张岳**(1/1), 2018 NLPCC 青年新锐奖, 中国计算机学会中文信息技术专业委员会, 其他, 其他, 2018.7.11
(张岳) (科研奖励)

(4) **Yue Zhang**, Deep learning for absolute and abnormal return prediction, The 18th International Conference on Electronic Commerce: ICEC 2016 in conjunction with SmartConnected.(韩国水原国际电子商务会议), Korea, August 17-19, 2016, 2016.8.17-2016.8.19 **(会议报告)**

(5) **Yue Zhang**, Using deep learning for predicting USA stock markets（关于用深度学习技术预测美国股市的特邀报告）, Deep Learning Summit (RE.WORK新加坡深度学习金融峰会）, Singapore, 2017.04.27-2017.04.28 **(会议报告)**

(6) **张岳**, 关于文本挖掘与资本市场的特邀报告, 杭州云栖大会分论坛, 杭州, 2018.9.20-2018.9.23 **(会议报告)**

除非特殊说明，请勿删除或改动简历模板中蓝色字体的标题及相应说明文字

参与者 简历

滕志扬，西湖大学，工学院，助理研究员

教育经历（从大学本科开始，按时间倒序排序；请列出攻读研究生学位阶段导师姓名）：

2015.01-2018.09，新加坡科技与设计大学，信息系统科技与设计，博士，导师：张岳

2011.09-2014.12，中国科学院大学，计算技术研究所，硕士，导师：刘群

2007.09-2011.07，东北大学，软件学院，学士

科研与学术工作经历（按时间倒序排序；如为在站博士后研究人员或曾有博士后研究经历，请列出合作导师姓名）：

2018.09-至今，西湖大学，工学院，助理研究员

曾使用其他证件信息（应使用唯一身份证件申请项目，曾经使用其他身份证件作为申请人或主要参与者获得过项目资助的，应当在此列明）

无

主持或参加科研项目（课题）情况（按时间倒序排序）：

代表性研究成果和学术奖励情况

（请注意：①投稿阶段的论文不要列出；②对期刊论文：应按照论文发表时作者顺序列出全部作者姓名、论文题目、期刊名称、发表年代、卷（期）及起止页码（摘要论文请加以说明）；③对会议论文：应按照论文发表时作者顺序列出全部作者姓名、论文题目、会议名称(或会议论文集名称及起止页码)、会议地址、会议时间；④应在论文作者姓名后注明第一/通讯作者情况：所有共同第一作者均加注上标“#”字样，通讯作者及共同通讯作者均加注上标“*”字样，唯一第一作者且非通讯作者无需加注；⑤所有代表性研究成果和学术奖励中本人姓名加粗显示。）

按照以下顺序列出：

一、代表性论著（包括论文与专著，合计 5 项以内）；

（1）Lei Shi, **Zhiyang Teng**, Le Wang, Yue Zhang, Alexander Binder. DeepClue: Visual Interpretation of Text-based Deep Stock Prediction. TKDE, 2018. （期刊论文）

（2）**Zhiyang Teng**, Yue Zhang. Two local models for neural constituent parsing. COLING, 2018, Santa Fe, New Mexico, USA, 2018.08.20-08.26. （会议论文）

（3）**Zhiyang Teng**, Yue Zhang. Head-Lexicalized Bidirectional Tree LSTMs. TACL, 2017, Volume 5: 163—177. （期刊论文）

（4）**Zhiyang Teng,** Duy Tin Vo, Yue Zhang. Context-Sensitive Lexicon Features for Neural Sentiment Analysis. EMNLP2016, Austin, Texas, USA, 2016.11.01-11.05. （会议论文）

（5）**Zhiyang Teng**, Hao Xiong, Qun Liu. Unsupervised Joint Monolingual Character Alignment and Word Segmentation. CCL&NLP-NABD 2014, Wuhan, Hubei, China, 2014.10.18-2014.10.19. （会议论文）

二、论著之外的代表性研究成果和学术奖励（合计 10 项以内）。

代表性研究成果和学术奖励的格式如下（**仅供规范格式示例使用，不代表排序要求，此部分标题及示例均可删除**）：

授权发明专利

1. Lin Ma, **Zhiyang Teng**, Hao Xiong. Machine Translation Method and Device Thereof. 2015-11-25. CN. Application No. PCT/CN2014/094507.

2. **Zhiyang Teng**, Hao Xiong, Weihua Luo, Shijing Wang. One machine translation processing method and device. 2015-07-01. CN. CN104750676B.

3. **Zhiyang Teng**, Hao Xiong, Qun Liu, Weihua Luo. An intelligent assistant for multilingual machine translation systems. 2015-12-16. CN. CN102968411A.

获得学术奖励

1. 滕志扬（1/3）, Unsupervised Joint Monolingual Character Alignment and Word Segmentation, 中国中文信息学会, NLP-NABD 最佳论文奖, 2014

（**滕志扬**，熊皓，刘群）

除非特殊说明，请勿删除或改动简历模板中蓝色字体的标题及相应说明文字

参与者 简历

何奇，西湖大学，工学院，助理研究员

教育经历（从大学本科开始，按时间倒序排序；请列出攻读研究生学位阶段导师姓名）：

（1） 2013.09-2018.12，电子科技大学，通信抗干扰技术国家级重点实验室，博士，导师：李少谦

（2） 2008.09-2011.06，电子科技大学，通信抗干扰技术国家级重点实验室，硕士，导师：凌翔

（3） 2008.09-2011.06，电子科技大学，通信学院，学士

科研与学术工作经历（按时间倒序排序；如为在站博士后研究人员或曾有博士后研究经历，请列出合作导师姓名）：

2019.01-至今，西湖大学，工学院，助理研究员

曾使用其他证件信息（应使用唯一身份证件申请项目，曾经使用其他身份证件作为申请人或主要参与者获得过项目资助的，应当在此列明）

无

主持或参加科研项目（课题）情况（按时间倒序排序）：

无

代表性研究成果和学术奖励情况

（请注意：①投稿阶段的论文不要列出；②对期刊论文：应按照论文发表时作者顺序列出全部作者姓名、论文题目、期刊名称、发表年代、卷（期）及起止页码（摘要论文请加以说明）；③对会议论文：应按照论文发表时作者顺序列出全部作者姓名、论文题目、会议名称(或会议论文集名称及起止页码)、会议地址、会议时间；④应在论文作者姓名后注明第一/通讯作者情况：所有共同第一作者均加注上标“#”字样，通讯作者及共同通讯作者均加注上标“*”字样，唯一第一作者且非通讯作者无需加注；⑤所有代表性研究成果和学术奖励中本人姓名加粗显示。）

按照以下顺序列出：

一、代表性论著（包括论文与专著，合计5项以内）；

(1) **Qi He**; Tony Q.S. Quek; Zhi Chen; Qi Zhang; Shaoqian Li. Compressive

Channel Estimation and Multi-User Detection in C-RAN with Low-Complexity Methods. IEEE Transactions on Wireless Communications, 2018.03, 17(6): 3931-3944 （期刊论文）

(2) **Qi He**; Zhi Chen; Tony Q.S. Quek; Jinho Choi; Shaoqian Li. Compressive Channel Estimation and User Activity Detection in Distributed-Input Distributed-Output Systems. IEEE Communications Letters, 2018.07, 22(9): 1850-1853 （期刊论文）

(3) **Qi He**; Jun Fang; Zhi Chen; Shaoqian Li. An Iteratively Reweighted Method for Recovery of Block-Sparse Signal with Unknown Block Partition. IEEE International Conference on Acoustics, Speech and Signal Processing, Shanghai, 2016.03.20-2016.03.25 （会议论文）

(4) **Qi He**; Tony Q.S. Quek; Zhi Chen; Shaoqian Li. Compressive Channel Estimation and Multi-User Detection in C-RAN. IEEE International Conference Communication, Paris, 2017.05.21-2017.05.29 （会议论文）

(5) **Qi He**; Zhengchuan Chen; Tony Q.S. Quek, Zhi Chen; Shaoqian Li. A Novel Cross-Layer Protocol for Random Access in Massive Machine-Type Communications. IEEE International Conference Communication Workshop, Kansas, 2018.05.20-2018.05.24 （会议论文）

二、论著之外的代表性研究成果和学术奖励（合计10项以内）。

无

除非特殊说明，请勿删除或改动简历模板中蓝色字体的标题及相应说明文字

参与者 简历

张源，西湖大学，工学院，技术员

教育经历（从大学本科开始，按时间倒序排序；请列出攻读研究生学位阶段导师姓名）：

2015.09-2018.06，中国科学院大学，中国科学院计算技术研究所，硕士，导师：刘群

2011.09-2015.06，华南理工大学，电子与信息学院，学士

科研与学术工作经历（按时间倒序排序；如为在站博士后研究人员或曾有博士后研究经历，请列出合作导师姓名）：

2018.09-至今，西湖大学，研究助理

曾使用其他证件信息（应使用唯一身份证件申请项目，曾经使用其他身份证件作为申请人或主要参与者获得过项目资助的，应当在此列明）

无

主持或参加科研项目（课题）情况（按时间倒序排序）：

无

代表性研究成果和学术奖励情况

（请注意：①投稿阶段的论文不要列出；②对期刊论文：应按照论文发表时作者顺序列出全部作者姓名、论文题目、期刊名称、发表年代、卷（期）及起止页码（摘要论文请加以说明）；③对会议论文：应按照论文发表时作者顺序列出全部作者姓名、论文题目、会议名称(或会议论文集名称及起止页码)、会议地址、会议时间；④应在论文作者姓名后注明第一/通讯作者情况：所有共同第一作者均加注上标“#”字样，通讯作者及共同通讯作者均加注上标“*”字样，唯一第一作者且非通讯作者无需加注；⑤所有代表性研究成果和学术奖励中本人姓名加粗显示。）

按照以下顺序列出：

一、代表性论著（包括论文与专著，合计5项以内）；

Yuan Zhang, Hongshen Chen, Yihong Zhao, Qun Liu and Dawei Yin, Learning Tag Dependencies for Sequence Tagging, International Joint Conference on Artificial Intelligence, 2018.07.13-2018.07.19, Sweden (会议论文)

二、论著之外的代表性研究成果和学术奖励（合计10项以内）。

无

附件信息

序号	附件名称	备注	附件类型
1	zhangyue-acl2018-1.pdf	Sentence-State LSTM for Text Representation	代表性论著
2	zhangyue-acl2018-2.pdf	Chinese NER Using Lattice LSTM	代表性论著
3	zhangyue-coling2018-3.pdf	Design Challenges and Misconceptions in Neural Sequence Labeling	代表性论著
4	zhangyue-acl2017-4.pdf	Universal Dependencies Parsing for Colloquial Singaporean English	代表性论著
5	zhangyue-cl2011-5.pdf	Syntactic Processing Using the Generalized Perceptron and Beam Search	代表性论著
6	ialp17.pdf	IALP 2017 最佳论文奖	其他
7	coling2018.pdf	COLING 2018 最佳论文奖	其他
8	青年新锐奖	2018年NLPCC青年新锐奖	其他
9	icec2016	8th International Conference on Electronic Commerce (ICEC 2016), Korea, August 17-19, 2016	学术会议大会报告或特邀报告邀请信
10	summit2017	Deep Learning Summit, Singapore on 27-28 April, 2017.	学术会议大会报告或特邀报告邀请信
11	云栖大会	2018 杭州云栖大会 特邀报告	学术会议大会报告或特邀报告邀请信

项目名称： 面向成分句法分析的跨领域知识抽取与融合
资助类型： 面上项目
申请代码： F060401.自然语言处理基础理论与方法

国家自然科学基金项目申请人和参与者公正性承诺书

本人**在此郑重承诺：**严格遵守中共中央办公厅、国务院办公厅《关于进一步加强科研诚信建设的若干意见》规定，所申报材料和相关内容真实有效，不存在违背科研诚信要求的行为；在国家自然科学基金项目申请、评审和执行全过程中，恪守职业规范和科学道德，遵守评审规则和工作纪律，杜绝以下行为：

（一）抄袭、剽窃他人科研成果或者伪造、篡改研究数据、研究结论；

（二）购买、代写、代投论文，虚构同行评议专家及评议意见；

（三）违反论文署名规范，擅自标注或虚假标注获得科技计划等资助；

（四）购买、代写申请书；弄虚作假，骗取科技计划项目、科研经费以及奖励、荣誉等；

（五）在项目申请书中以高指标通过评审，在项目计划书中故意篡改降低相应指标；

（六）以任何形式探听尚未公布的评审专家名单及其他评审过程中的保密信息；

（七）本人或委托他人通过各种方式及各种途径联系有关专家进行请托、游说，违规到评审会议驻地游说评审专家和工作人员、询问评审或尚未正式向社会公布的信息等干扰评审或可能影响评审公正性的活动；

（八）向评审工作人员、评审专家等提供任何形式的礼品、礼金、有价证券、支付凭证、商业预付卡、电子红包，或提供宴请、旅游、娱乐健身等任何可能影响评审公正性的活动；

（九）其他违反财经纪律和相关管理规定的行为。

如违背上述承诺，本人愿接受国家自然科学基金委员会和相关部门做出的各项处理决定，包括但不限于撤销科学基金资助项目，追回项目资助经费，向社会通报违规情况，取消一定期限国家自然科学基金项目申请资格，记入科研诚信严重失信行为数据库以及接受相应的党纪政纪处理等。

编号	姓名 / 工作单位名称（应与加盖公章一致）/ 证件号码 / 每年工作时间（月）	签字
1	**张岳** / 西湖大学 / 1**********9 / 6	
2	**滕志扬** / 西湖大学 / 4**********6 / 8	
3	**何奇** / 西湖大学 / 5**********9 / 8	
4	**张源** / 西湖大学 / 2**********2 / 10	
5	**崔乐阳** / 西湖大学 / 1**********8 / 8	
6	**王祎乐** / 西湖大学 / 4**********1 / 8	
7	**白雪峰** / 西湖大学 / 1**********5 / 8	
8	**贾晨** / 西湖大学 / 1**********7 / 8	
9	**陈雨龙** / 西湖大学 / 4**********3 / 8	
10		

项目名称： 面向成分句法分析的跨领域知识抽取与融合

资助类型： 面上项目

申请代码： F060401.自然语言处理基础理论与方法

国家自然科学基金项目申请单位公正性承诺书

本单位依据国家自然科学基金项目指南的要求，严格履行法人负责制，**在此郑重承诺：**本单位已就所申请材料内容的真实性和完整性进行审核，不存在违背中共中央办公厅、国务院办公厅《关于进一步加强科研诚信建设的若干意见》规定和其他科研诚信要求的行为，申请材料符合《中华人民共和国保守国家秘密法》和《科学技术保密规定》等相关法律法规，在项目申请和评审活动全过程中，遵守有关评审规则和工作纪律，杜绝以下行为：

（一）采取贿赂或变相贿赂、造假、剽窃、故意重复申报等不正当手段获取国家自然科学基金项目申请资格；

（二）以任何形式探听未公开的项目评审信息、评审专家信息及其他评审过程中的保密信息，干扰评审专家的评审工作；

（三）组织或协助项目团队向评审工作人员、评审专家等提供任何形式的礼品、礼金、有价证券、支付凭证、商业预付卡、电子红包等；宴请评审组织者、评审专家，或向评审组织者、评审专家提供旅游、娱乐健身等任何可能影响科学基金评审公正性的活动；

（四）包庇、纵容项目团队虚假申报项目，甚至骗取国家自然科学基金项目；

（五）包庇、纵容项目团队，甚至帮助项目团队采取“打招呼”等方式，影响科学基金项目评审的公正性；

（六）在申请书中以高指标通过评审，在计划书中故意篡改降低相应指标；

（七）其他违反财经纪律和相关管理规定的行为。

如违背上述承诺，本单位愿接受国家自然科学基金委员会和相关部门做出的各项处理决定，包括但不限于停拨或核减经费，追回项目经费，取消一定期限国家自然科学基金项目申请资格，记入科研诚信严重失信行为数据库以及主要责任人接受相应党纪政纪处理等。

依托单位公章:

日期： 年 月 日

合作研究单位公章:

日期： 年 月 日

合作研究单位公章:

日期： 年 月 日