

标下评估了系统的性能。Bejan 和 Harabagiu^[122]运用无监督的非参贝叶斯模型将词汇特征和 WordNet 中的语义相似度引入事件指代消解任务中。2015 年,Araki 等人^[123]首次提出一种联合学习模型,即将事件抽取任务和事件指代消解任务同时研究。随后 Lu 和 Ng^[124]也构建了一个基于一元二元以及三元特征融合的联合学习模型。近年来,神经网络在自然语言处理的各个领域都取得不错的研究成果,Nguyen^[125]通过非连续卷积模型在 KBP^[126]语料上完成事件指代消解任务的研究。同年,Krause 等人^[127]也搭建了卷积神经网络模型,并在 ACE 和 ACE++ 语料进行了相关任务研究。在中文事件指代消解方面,受限于语料,目前只有少量工作,代表性工作包括:Lu 和 Ng^[124]构建的平台不仅汇报了英文事件指代消解的性能,也汇报了 KBP 中文语料上的性能;滕佳月等人^[128,129]基于 ACE 中文语料进行了中文事件指代消解的研究,并提出了基于全局优化进行性能改善的策略。

除指代外,针对篇章意图性的计算模型的研究很少,代表性工作是 Pustejovsky 等人^[130]在 GraphBank 上的相关工作,他们对 GraphBank 进行了分析,认为篇章连接词和两个句子间的跨度距离是高效识别显式和隐式篇章关系的关键因素。

2.4 存在的问题和研究趋势

从上述国内外研究现状的分析中我们可以看到,相比英语,汉语的篇章研究刚刚起步,汉语篇章阅读理解研究鲜有见诸文献。目前汉语篇章理解还存在如下一些主要问题。

(1) 适用于汉语篇章阅读理解的篇章结构理论体系很不完善。有必要借鉴英语的相关篇章理论,并结合汉语特点和复句、句群、广义话题结构等本土理论,逐步建立汉语篇章结构理论体系。

(2) 适用于汉语篇章阅读理解的篇章结构大规模标注资源非常缺乏。虽然有一些研究者,或基于英语篇章理论体系,或基于汉语的复句、句群和广义话题结构等理论,对汉语篇章结构资源库展开了研究,但相关研究比较分散,大多属于探索性工作,有待进一步深入、系统地进行研究。

(3) 适用于汉语篇章阅读理解的篇章结构分析关键技术十分匮乏。由于适用于汉语篇章结构分析的理论体系尚未有效建立,相关标注资源缺乏,因此很难大规模有效地进行关键技术研究。

(4) 篇章理解需要涉及不同视角、不同层次的篇章结构分析结果,各种结构间也存在明显的互补关系,构建统一体系(包括理论体系和资源)进行多视角、多层次的联合分析研究,有待进一步深入。

2.5 机器阅读理解的相关研究

虽然适用于汉语篇章阅读理解的篇章结构分析研究处于起步阶段,机器阅读理解的相关研究却吸引了众多研究者。目前,机器阅读理解方面已经开展了一些工作,具体包括:Hermann 等人^[131]借助爬虫技术从 CNN 和每日邮报新闻网网页爬取数据,构建了一个完形填空类型(cloze-style)的阅读理解数据库 CNN and Daily Mail。2016 年,斯坦福大学通过亚马逊众包平台建立了一个新的阅读理解数据集 SQuAD^[132],它包含 536 篇维基百科文章,100 000 多个问题,而且每篇文章都是经过人工阅读,提出问题并给出答案片段。微软公司选取了 100 000 多名用户通过 Bing 搜索引擎提出的问题,每一个问题都会对应大约 10 篇相关的从网页抽取的文章,相关人员会根据 10 篇文章给出问题的答案,以此构建了 MS MARCO^[133]语料库。随着这些语料的正式发布,各种机器学习方法、深度神经网络方法和 attention 机制都不断被提出并被应用到这一任务中^[134-142]。此外,Cui 等人^[143]发布了第一个中文 cloze-style 阅读理解语料 People Daily News 数据集和 Children's Fairy Tale(CFT)数据集。从 2017 年至今,“讯飞杯”中文机器阅读理解评测已经成功举办两届,从第 1 届以填空题阅读理解问题为主,到第 2 届关注基于篇章片段抽取的阅读理解,评测会议发布了人工标注的中文填空型和篇章片段抽取型阅读理解的数据集^[144],很多的相关研究也在这些数据集上有所展开。但本质上,这些工作只是把篇章看作一个词符号序列,缺乏真正意义上的篇章理解。当然,从另一层面而言,这些研究也大大推动了人们对篇章理解的关注和重视。例如,NSFC 最近几年就批准了多个汉语篇章理解方向的重点项目和人工智能应急重点项目,包括哈尔滨工业大学刘挺主持的篇章级中文语义分析理论与方法,中国科学院自动化研究所宗成庆主持的汉语多层次语篇分析理论方法研究与应用,苏州大学张民主持的面向多层次篇章语义的机器翻译理论、方法与实现,北京理工大学黄河燕主持的中文语义深度计算与阅读理解,以及苏州大学周国栋主持的话题驱动的汉语篇章机器阅读理解等。

3 总 结

综上所述,在自然语言处理领域,与词法分析、句法分析等研究相比,篇章结构分析研究相对滞后.特别是适用于汉语篇章阅读的篇章结构分析研究还处于起步阶段,尚未形成一套有效的理论体系,相应语料库资源建设薄弱,关键技术研究严重滞后.相应地,机器阅读理解的相关研究也刚刚起步,目前主要是基于检索技术的相关片段抽取,缺乏真正意义上的篇章理解.众所周知,与英语等西方语言相比,汉语无论是篇章结构和信息意图表达方式,还是事件描述方式和话题表述方式等方面都有较大的差异.这就迫切需要进一步完善适用于汉语篇章阅读的篇章结构理论体系,建立一定规模的适用于汉语篇章阅读的汉语篇章结构资源库,并在此基础上建立汉语篇章结构分析的计算模型,实现高性能的汉语篇章结构分析和篇章深度理解平台,为自然语言理解和篇章级应用提供基础支撑.

References:

- [1] De Robert-Alain B, Wolfgang DU. Introduction to Text Linguistics. London, New York: Longman Paperback, 1981.
- [2] Schank RC. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 1972,3(4):552–631.
- [3] Halliday MAK, Ruqaiya H. Cohesion in English. Harlow: Longman PubGroup, 1976.
- [4] Hobbs JR. Coherence and coreference. *Cognitive Science*, 1979,3(1):67–90.
- [5] Hobbs JR. Information, intention and structure in discourse. In: Proc. of the NATO Workshop on Burning Issues in Discourse. 1993.
- [6] Mann WC, Thompson SA. Rhetorical structure theory: Toward a functional theory of text organization. *Text-interdisciplinary Journal for the Study of Discourse*, 1988,8(3):243–281.
- [7] Mann WC, Thompson SA. Relational propositions in discourse. *Discourse Processing*, 1986,9(1):57–90.
- [8] Grosz BJ, Sidner CL. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 1986,12(3): 175–204.
- [9] Grosz BJ, Weinstein S, Joshi AK. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 1995,21(2):203–225.
- [10] Halliday MAK. Linguistic function and literary style: An inquiry into the language of William Golding's the inheritors. In: Chatman S, eds. Proc. of the Symp. on Literary Style. New York: Oxford University Press, 1971. 330.
- [11] Marcu D. The rhetorical parsing, summarization, and generation of natural language texts [Ph.D. Thesis]. Toronto: Department of Computer Science, University of Toronto, 1997.
- [12] Marcu D. The Theory and Practice of Discourse Parsing and Summarization. Cambridge: The MIT Press, 2000.
- [13] Halliday MAK. An Introduction to Functional Grammar. 3rd ed., London: Arnold, 2004.
- [14] Daneš F. Functional sentence perspective and the organization of text. In: Functional Sentence Perspective. Prague: Academica, 1974. 106–128.
- [15] Fries PH. On the status of theme in English: Arguments from discourse. *Forum Linguistica*, 1981,6(1):1–38.
- [16] Wu WZ. Chinese Sentence Group. Beijing: The Commercial Press, 2000 (in Chinese).
- [17] Xing FY. Research on Chinese Complex Sentences. Beijing: The Commercial Press, 2001 (in Chinese).
- [18] Yao SY. A research on the collocation of the relation markers of Chinese compound sentences and some relevant explanation [Ph.D. Thesis]. Wuhan: Central China Normal University, 2006 (in Chinese with English abstract).
- [19] Li YC. Research of Chinese discourse structure representation and resource construction [Ph.D. Thesis]. Suzhou: Soochow University, 2015 (in Chinese with English abstract).
- [20] Li YC, Feng WH, Sun J, Kong F, Zhou GD. Building Chinese discourse corpus with connective-driven dependency tree structure. In: Proc. of the EMNLP. 2014. 2105–2114.
- [21] Jiang YR, Song R. Topic clause identification based on generalized topic theory. *Journal of Chinese Information Processing*, 2012, 26(5):114–120 (in Chinese with English abstract).
- [22] Song R. The flowing water model of Chinese discourse generalized topic structure. *Studies of the Chinese Language*, 2013,6: 483–494 (in Chinese with English abstract).
- [23] Shang Y, Song R, Lu DW. Study on the sentence-formability of topic sufficient sentence under the perspective of generalized topic structure theory. *Journal of Chinese Information Processing*, 2014,28(6):107–113 (in Chinese with English abstract).
- [24] Ma JZ. Ma's Grammar. Beijing: The Commercial Press, 1929 (in Chinese).

- [25] Li JX, Liu SR. The Chinese Grammar Textbook. Tianjing: Tianjin Dazhong Press, 1952 (in Chinese).
- [26] Chu XM, Zhu QM, Zhou GD. Discourse primary-secondary relationships in natural language processing. Chinese Journal of Computers, 2017,40(4):842–859 (in Chinese with English abstract).
- [27] Cao FF. Clause and Sentence Structure in Chinese: A Functional Perspective. Student Book Co., 1990.
- [28] Qu CX. Chinese Discourse Grammar. Beijing: Beijing Language and Culture University Press, 1998 (in Chinese).
- [29] Liu LJ. A Contrastive Study of Discourse Structures in English and Chinese. Guangzhou: Sun Yat-sen University Press, 1998. 166–178 (in Chinese).
- [30] Sun K. A contrastive study of discourse constructional patterns between Chinese and English: spective of topic chains. Journal of PLA University of Foreign Languages, 2013,36(3):12–18 (in Chinese with English abstract).
- [31] Wang JJ. On the Continuation of the Topic: Research on Topic-chain-based Chinese-English Discourse. Shanghai: Shanghai Jiaotong University Press, 2013,2013 (in Chinese).
- [32] Zhao Q, Zhou XC. Topic-chain-based coherence annotation scheme for Chinese text. Journal of Chinese Information Processing, 2014,28(5):102–111 (in Chinese with English abstract).
- [33] Xi XF. Research on Chinese discourse topic structure: Representation, resource construction and its analysis [Ph.D. Thesis]. Suzhou: Soochow Univeristy, 2017 (in Chinese with English abstract).
- [34] Xi XF, Sun QY, Zhou GD. Research and prospect of discourse topic structure analysis for discourse intentionality. Chinese Journal of Computers, 2017 (in Chinese with English abstract). <http://kns.cnki.net/KCMS/detail/11.1826.TP.20170601.1917.016.html>
- [35] Carlson L, Marcu D, Okurowski ME. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In: van Kuppevelt J, Smith R, eds. Proc. of the Current Directions in Discourse. New York: Kluwer, 2003. 85–112.
- [36] Prasad R, Dinesh N, Lee A, Miltsakaki E, Robaldo L, Joshi A, Webber B. The Penn Discourse Treebank 2.0. In: Proc. of the Int'l Conf. on Language Resources and Evaluation. 2008.
- [37] Doddington GR, Mitchell A, Przybocki MA, *et al.* The automatic content extraction (ACE) program-tasks, data, and evaluation. In: Proc. of the LREC. 2004,2:1.
- [38] Poesio M, Artstein R. Anaphoric annotation in the ARRAU corpus. In: Proc. of the Int'l Conf. on Language Resources and Evaluation. 2008.
- [39] Hovy E, Marcus M, Palmer M, *et al.* OntoNotes: The 90% solution. In: Proc. of the NAACL 2006. 2006. 57–60.
- [40] Wolf F, Gibson E. Representing discourse coherence: A corpus-based study. Computational Linguistics, 2005,32(2):249–287.
- [41] Yue M. Rhetorical structure annotation of Chinese news commentaries. Journal of Chinese Information Processing, 2008,22(4): 19–23 (in Chinese with English abstract).
- [42] Chen LP. The theoretical framework grounding the annotation of Chinese text structures. Journal of Nanjing University of Aeronautics & Astronautics (Social Sciences), 2008,10(3):68–71 (in Chinese with English abstract).
- [43] Zhou YP, Xue NW. PDTB-style discourse annotation of Chinese text. In: Proc. of the ACL 2012. 2012. 69–77.
- [44] Zhou YP, Xue NW. The Chinese discourse TreeBank: A Chinese corpus annotated with discourse relations. Language Resources and Evaluation, 2015,49(2):397–431.
- [45] Zhang MY, Qin B, Liu T. Chinese discourse relation semantic taxonomy and annotation. Journal of Chinese Information Processing, 2014,28(2):28–36 (in Chinese with English abstract).
- [46] Shu JB. Research of auto-identifying the relation markers of compound sentence for Chinese information processing [Ph.D. Thesis]. Wuhan: Central China Normal University, 2011 (in Chinese with English abstract).
- [47] Zhao Q. Annotation scheme for Chinese Treebank. Journal of Chinese Information Processing, 2004,18(4):2–9 (in Chinese with English abstract).
- [48] Xi XF, Chu XM, Sun QY, Zhou GD. Corpus construction for Chinese discourse topic via micro-topic scheme. Journal of Computer Research and Development, 2017,54(8):1833–1852 (in Chinese with English abstract).
- [49] Hannah R, *et al.* Discourse coherence: Concurrent explicit and implicit relations. In: Proc. of the 56th Annual Meeting of the Association for Computational Linguistics, Vol. 1. 2018.
- [50] Soricut R, Marcu D. Sentence level discourse parsing using syntactic and lexical information. In: Proc. of the NAACL-HLT 2003. 2003. 149–156.
- [51] Hernault H, Bollegala D, Ishizuka M. A sequential model for discourse segmentation. In: Proc. of the CICLing 2010. 2010. 315–326.

- [52] LeThanh H, Abeysinghe G, Huyck C. Generating discourse structures for written texts. In: Proc. of the COLING 2004. 2004. 329.
- [53] Hernault H, Prendinger H, duVerle DA, *et al.* Hilda: A discourse parser using support vector machine classification. *Dialogue and Discourse*, 2010,1(3):1–33.
- [54] Wei FW, Graeme H. Text-level discourse parsing with rich linguistic features. In: Proc. of the ACL 2012. 2012. 60–68.
- [55] Joty S, Carenini G, Ng RT. A novel discriminative framework for sentence-level discourse analysis. In: Proc. of the EMNLP-CoNLL 2012. 2012. 904–915.
- [56] Surdeanu M, Hicks T, Valenzuela-Escarcega MA. Two practical rhetorical structure theory parsers. In: Proc. of the 2015 Conf. of the North American Chapter of the Association for Computational Linguistics: Demonstrations. 2015. 1–5.
- [57] Braud C, Plank B, Søgaard A. Multi-view and multi-task training of rst discourse parsers. In: Proc. of the 26th Int'l Conf. on Computational Linguistics: Technical Papers. 2016. 1903–1913.
- [58] Li Q, Li TS, Chang BB. Discourse parsing with attention-based hierarchical neural networks. In: Proc. of the EMNLP. 2016. 362–371.
- [59] Braud C, Coavoux M, Søgaard A. Cross-lingual RST discourse parsing. In: Proc. of the 15th Conf. of the European Chapter of the Association for Computational Linguistics, Vol. 1. 2017. 292–304.
- [60] Ji YF, Eisenstein J. Representation learning for text-level discourse parsing. In: Proc. of the 52nd Annual Meeting of the Association for Computational Linguistics, Vol. 1. 2014. 13–24.
- [61] Dines N, Lee A, Miltsakaki E, *et al.* Attribution and the (non-) alignment of syntactic and discourse arguments of connectives. In: Proc. of the Workshop on Frontiers in Corpus Annotations II: Pie in the Sky. 2005. 29–36.
- [62] Lin Z, Ng HT, Kan MY. A PDTB-styled end-to-end discourse parser. *Natural Language Engineering*, 2014,20(2):151–184.
- [63] Wellner B, Pustejovsky J. Automatically identifying the arguments of discourse connectives. In: Proc. of the EMNLP-CoNLL 2007. 2007. 92–101.
- [64] Ghosh S, Johansson R, Tonelli S. Shallow discourse parsing with conditional random fields. In: Proc. of the IJCNLP. 2011. 1071–1079.
- [65] Ghosh S. End-to-end discourse parsing using cascaded structured prediction [Ph.D. Thesis]. University of Trento, 2012.
- [66] Kong F, Ng HT, Zhou G. A constituent-based approach to argument labeling with joint inference in discourse parsing. In: Proc. of the EMNLP 2014. 2014. 68–77.
- [67] Pitler E, Louis A, Nenkova A. Automatic sense prediction for implicit discourse relations in text. In: Proc. of the ACL-AFNLP 2009. 2009. 683–691.
- [68] Marcu D, Echihiabi A. An unsupervised approach to recognizing discourse relations. In: Proc. of the ACL 2002. 2002. 368–375.
- [69] Saito M, Yamamoto K, Sekine S. Using phrasal patterns to identify discourse relations. In: Proc. of the NAACL 2006. 2006. 133–136.
- [70] Lin Z, Kan MY, Ng HT. Recognizing implicit discourse relations in the Penn discourse Treebank. In: Proc. of the EMNLP 2009. 2009. 343–351.
- [71] Park J, Cardie C. Improving implicit discourse relation recognition through feature set optimization. In: Proc. of the SIGDIAL 2012. 2012. 108–112.
- [72] Zhou ZM, Xu Y, Niu ZY, *et al.* Predicting discourse connectives for implicit discourse relation recognition. In: Proc. of the COLING 2010. 2010. 1507–1514.
- [73] Biran O, McKeown K. Aggregated word pair features for implicit discourse relation disambiguation. In: Proc. of the ACL 2013. 2013. 69–73.
- [74] Lan M, Xu Y, Niu Z. Leveraging synthetic discourse data via multi-task learning for implicit discourse relation recognition. In: Proc. of the ACL 2013. 2013. 476–485.
- [75] Zhou XP, Hong Y, Che TT, *et al.* An unsupervised approach to inferring implicit discourse relation. *Journal of Chinese Information Processing*, 2013,27(2):17–26 (in Chinese with English abstract).
- [76] Lan M, Wang JX, Wu YB, Niu ZY, Wang HF. Multi-task attention-based neural networks for implicit discourse relationship representation and identification. In: Proc. of the 2017 Conf. on Empirical Methods in Natural Language Processing. 2017. 1299–1308.
- [77] Lin Z, Ng HT, Kan MY. A PDTB-styled end-to-end discourse parser. *Natural Language Engineering*, 2014,20(2):151–184.
- [78] Zhang MY, Song Y, Qin B, *et al.* Chinese discourse relation recognition. *Journal of Chinese Information Processing*, 2013,27(6): 51–58 (in Chinese with English abstract).

- [79] Tu M, Zhou Y, Zong CQ. Automatically parsing Chinese discourse based on maximum entropy. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2014,50(1):125–132 (in Chinese with English abstract).
- [80] Kong F, Zhou G. A CDT-styled end-to-end Chinese discourse parser. *ACM Trans. on Asian and Low-resource Language Information Processing (TALLIP)*, 2017,16(4):26.
- [81] Jia Y, Feng Y, Ye Y, *et al.* Improved discourse parsing with two-step neural transition-based model. *ACM Trans. on Asian and Low-resource Language Information Processing (TALLIP)*, 2018,17(2):11.
- [82] Sun C, Kong F. A transition-based framework for Chinese discourse structure parsing. *Journal of Chinese Information Processing*, 2018,32(12):26–34 (in Chinese with English abstract).
- [83] Vilain M, Burger J, Aberdeen J, *et al.* A model-theoretic coreference scoring scheme. In: *Proc. of the 6th Conf. on Message Understanding*. 1995. 45–52.
- [84] Hirschman L, Robinson P, Burger J, *et al.* Automating coreference: The role of annotated training data. In: *Proc. of the AAAI Spring Symp. on Applying Machine Learning to Discourse Processing*. 1997. 118–121.
- [85] Pradhan S, Ramshaw L, Marcus M, *et al.* Conll-2011 Shared Task: Modeling unrestricted coreference in ontototes. In: *Proc. of the 15th Conf. on Computational Natural Language Learning: Shared Task*. 2011. 1–27.
- [86] Pradhan S, Moschitti A, Xue N, *et al.* CoNLL-2012 Shared Task: Modeling multilingual unrestricted coreference in OntoNotes. In: *Proc. of Joint Conf. on EMNLP and CoNLL-Shared Task*. 2012. 1–40.
- [87] Raghunathan K, Lee H, Rangarajan S, *et al.* A multi-pass sieve for coreference resolution. In: *Proc. of the Conf. on Empirical Methods in Natural Language Processing, EMNLP 2010*. 2010. 492–501.
- [88] Lee H, Peirsman Y, Chang A, *et al.* Stanford’s multi-pass sieve coreference resolution system at the CoNLL-2011 shared task. In: *Proc. of the 15th Conf. on Computational Natural Language Learning: Shared Task*. 2011. 28–34.
- [89] Cardie C, Wagstaff K. Noun phrase coreference as clustering. In: *Proc. of the Joint Conf. on Empirical Methods in NLP and Very Large Corpora*. 1999. 277–308.
- [90] Mccarthy JF, Lehnert WG. Using decision trees for coreference resolution. In: *Proc. of the 14th Int’l Joint Conf. on Artificial Intelligence*. 1999. 1050–1055.
- [91] Soon WM, Ng HT, Lim DCY. A machine learning approach to coreference resolution of noun phrases. *Computational Linguistics*, 2001,27(4):521–544.
- [92] Kong F, Zhou GD. Pronoun resolution in English and Chinese languages based on tree kernel. *Ruan Jian Xue Bao/Journal of Software*, 2012,34(5):1085–1099 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4044.htm> [doi: 10.3724/SP.J.1001.2012.04044]
- [93] Xu R, Xu J, Liu J, *et al.* Incorporating rule-based and statistic-based techniques for coreference resolution. In: *Proc. of the Joint Conf. on EMNLP and CoNLL-Shared Task*. 2012. 107–112.
- [94] Belder JD, Moens MF. Coreference clustering using column generation. In: *Proc. of the COLING 2012: Posters*. 2012. 245–254.
- [95] Wiseman S, Rush AM, Shieber SM. Learning global features for coreference resolution. In: *Proc. of the NAACL-HLT*. 2016. 994–1004.
- [96] Clark K, Manning CD. Deep reinforcement learning for mention-ranking coreference models. In: *Proc. of the 2016 Conf. on Empirical Methods in Natural Language Processing*. 2016. 2256–2262.
- [97] Lee K, He L, Lewis M, *et al.* End-to-end neural coreference resolution. In: *Proc. of the 2017 Conf. on Empirical Methods in Natural Language Processing*. 2017. 188–197.
- [98] Wang HF. Survey: Computational models and technologies in anaphora resolution. *Journal of Chinese Information Processing*, 2002,16(6):10–18 (in Chinese with English abstract).
- [99] Wang HF, He TT. Research on Chinese pronominal anaphora resolution. *Chinese Journal of Computers*, 2001,24(2):136–143 (in Chinese with English abstract).
- [100] Wang HF, Mei Z. Robust pronominal resolution within Chinese text. *Ruan Jian Xue Bao/Journal of Software*, 2005,16(5):700–707 (in Chinese with English abstract). http://www.jos.org.cn/jos/ch/reader/create_pdf.aspx?file_no=20050508&journal_id=jos [doi: 10.1360/jos160700]
- [101] Li GC, Luo YF. Chinese pronominal anaphora resolution via a preference selection approach. *Journal of Chinese Information Processing*, 2005,19(4):24–30 (in Chinese with English abstract).
- [102] Zhou JS, Huang SJ, Chen JJ, *et al.* A new graph clustering algorithm for Chinese noun phrase coreference resolution. *Journal of Chinese Information Processing*, 2007,21(2):77–82 (in Chinese with English abstract).

- [103] Yang Y, Li YC, Zhou GD, *et al.* Research on distance information for anaphora resolution. *Journal of Chinese Information Processing*, 2008,22(5):39–44 (in Chinese with English abstract).
- [104] Wang HD, Hu NQ, Kong F, *et al.* Research on semantic role information in anaphora resolution. *Journal of Chinese Information Processing*, 2009,23(1):23 (in Chinese with English abstract).
- [105] Li YQ, Gan RS, Yng YH, *et al.* Chinese coreference resolution method based on feature respective selection strategy. *Computer Engineering*, 2011,37(18):180–182 (in Chinese with English abstract).
- [106] Zhang MY, Li YB, Qin B, *et al.* Coreference resolution based on head match. *Journal of Chinese Information Processing*, 2011, 25(3):3–9 (in Chinese with English abstract).
- [107] Song Y, Wang HF. Chinese zero anaphora resolution with Markov logic. *Journal of Computer Research and Development*, 2015, 52(9):2114–2122 (in Chinese with English abstract).
- [108] Zhao S, Ng HT. Identification and resolution of Chinese zero pronouns: A machine learning approach. In: *Proc. of the EMNLP-CoNLL 2007*. 2007. 541–550.
- [109] Kong F, Zhou G. A tree kernel-based unified framework for Chinese zero anaphora resolution. In: *Proc. of the EMNLP 2010*. 2010. 882–891.
- [110] Chen C, Ng V. Chinese zero pronoun resolution: some recent advances. In: *Proc. of the EMNLP 2013*. 2013. 1360–1365.
- [111] Chen C, Ng V. Chinese overt pronoun resolution: A bilingual approach. In: *Proc. of the AACL 2014*. 2014. 1615–1621.
- [112] Chen C, Ng V. Chinese zero pronoun resolution: A joint unsupervised discourse-aware model rivaling state-of-the-art resolvers. In: *Proc. of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th Int'l Joint Conf. on Natural Language Processing*, Vol. 2. 2015. 320–326.
- [113] Chen C, Ng V. Chinese zero pronoun resolution with deep neural networks. In: *Proc. of the ACL 2016*. 2016. 778–788.
- [114] Sheng C, Kong F, Zhou GD. Toward better Chinese zero pronoun resolution from discourse perspective. In: *Proc. of the NLPCC-ICCPOL 2017, Lecture Notes in Computer Science*, 2017.
- [115] Kong F, Zhou GD. Chinese zero pronoun resolution: A chain to chain approach. In: *Proc. of the NLPCC-ICCPOL 2017. Lecture Notes in Computer Science*, Springer-Verlag, 2017.
- [116] Yin QY, Zhang Y, Zhang WN, Liu T. Chinese zero pronoun resolution with deep memory network. In: *Proc. of the 2017 Conf. on Empirical Methods in Natural Language Processing*. 2017. 1309–1318.
- [117] Zhang Y, Liu T, Yin QY, Zhang WN. A deep neural network for Chinese zero pronoun resolution. In: *Proc. of the 26th Int'l Joint Conf. on Artificial Intelligence*. 2017. 3322–3328.
- [118] Liu T, Cui YM, Yin QY, Zhang WN, Wang SJ, Hu GP. Generating and exploiting large-scale pseudo training data for zero pronoun resolution. In: *Proc. of the 55th Annual Meeting of the Association for Computational Linguistics*, Vol. 1. 2017. 102–111.
- [119] Yin QY, Zhang Y, Zhang WN, Liu T, Yang WW. Deep reinforcement learning for Chinese zero pronoun resolution. In: *Proc. of the 56th Annual Meeting of the Association for Computational Linguistics*, Vol. 1. 2018. 569–578.
- [120] Ahn D. The stages of event extraction. In: *Proc. of the Workshop on Annotating & Reasoning about Time & Events*. 2006. 1–8.
- [121] Chen Z, Ji H, Haralick R. A pairwise event coreference model, feature impact and evaluation for event coreference resolution. In: *Proc. of the Workshop on Events in Emerging Text Types*. 2009. 17–22.
- [122] Bejan CA, Harabagiu S. Unsupervised event coreference resolution with rich linguistic features. In: *Proc. of the Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics. 2010. 1412–1422.
- [123] Araki J, Mitamura T. Joint event trigger identification and event coreference resolution with structured perceptron. In: *Proc. of the Conf. on Empirical Methods in Natural Language Processing*. 2015. 2074–2080.
- [124] Lu J, Ng V. Joint learning for event coreference resolution. In: *Proc. of the 55th Annual Meeting of the Association for Computational Linguistics*, Vol. 1. 2017. 90–101.
- [125] Nguyen TH, Meyers A, Grishman R. New York University 2016 system for KBP event Nugget: A deep learning approach. In: *Proc. of the 9th Text Analysis Conf*. 2016.
- [126] Joe E, Jeremy G, Dana F, *et al.* Overview of linguistic resources for the TAC KBP. 2015. <https://tac.nist.gov/2015/KBP/>
- [127] Krause S, Xu F, Uszkoreit H, *et al.* Event linking with sentential features from convolutional neural networks. In: *Proc. of the Signll Conf. on Computational Natural Language Learning*. 2016. 239–249.
- [128] Teng JY, Li PF, Zhu QM. Global inference for co-reference resolution between Chinese events. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2016,52(1):97–103 (in Chinese with English abstract).

- [129] Teng JY. Research on Chinese event coreference resolution [MS. Thesis]. Suzhou: Soochow University. 2016 (in Chinese with English abstract).
- [130] Wellner B, Pustejovsky J, Havasi C, Rumshisky A, Sauri R. Classification of discourse coherence relations: An exploratory study using multiple knowledge sources. In: Proc. of the 7th SIGDIAL Workshop on Discourse and Dialogue. 2006. 117–125.
- [131] Hirschman L, Light M, Breck E, *et al.* Deep read: A reading comprehension system. In: Proc. of the 37th Annual Meeting of the Association for Computational Linguistics on Computational Linguistics. Association for Computational Linguistics, 1999. 325–332.
- [132] Rajpurkar P, Zhang J, Lopyrev K, Liang P. SQuAD: 100,000+ questions for machine comprehension of text. In: Proc. of the 2016 Conf. on Empirical Methods in Natural Language Processing. 2016. 2383–2392.
- [133] Bajaj P, Campos D, Craswell N, *et al.* MS MARCO: A human generated machine reading comprehension dataset. arXiv Preprint arXiv: 1611.09268, 2018.
- [134] Ng HT, Teo LH, Kwan JLP. A machine learning approach to answering questions for reading comprehension tests. In: Proc. of the 2000 Joint SIGDAT Conf. on Empirical Methods in Natural Language Processing and Very Large Corpora: Held in Conjunction with the 38th Annual Meeting of the Association for Computational Linguistics, Vol. 13. Association for Computational Linguistics, 2000. 124–132.
- [135] Chen D, Bolton J, Manning CD. A thorough examination of the CNN/daily mail reading comprehension task. In: Proc. of the 54th Annual Meeting of the Association for Computational Linguistics, Vol. 1. 2016. 2358–2367.
- [136] Chu Z, Wang H, Gimpel K, *et al.* Broad context language modeling as reading comprehension. In: Proc. of the 15th Conf. of the European Chapter of the Association for Computational Linguistics, Vol. 2. 2017. 52–57.
- [137] Cui Y, Chen Z, Wei S, *et al.* Attention-over-attention neural networks for reading comprehension. In: Proc. of the 55th Annual Meeting of the Association for Computational Linguistics, Vol. 1. 2017. 593–602.
- [138] Hewlett D, Jones L, Lacoste A. Accurate supervised and semi-supervised machine reading for long documents. In: Proc. of the 2017 Conf. on Empirical Methods in Natural Language Processing. 2017. 2011–2020.
- [139] Long T, Bengio E, Lowe R, *et al.* World knowledge for reading comprehension: Rare entity prediction with hierarchical LSTMs using external descriptions. In: Proc. of the 2017 Conf. on Empirical Methods in Natural Language Processing. 2017. 825–834.
- [140] Shen YL, Liu XD, Duh K, *et al.* An empirical analysis of multiple-turn reasoning strategies in reading comprehension tasks. In: Proc. of the 8th Int'l Joint Conf. on Natural Language Processing, Vol. 1. 2017. 957–966.
- [141] Wang W, Yang N, Wei F, *et al.* Gated self-matching networks for reading comprehension and question answering. In: Proc. of the 55th Annual Meeting of the Association for Computational Linguistics. 2017,1:189–198.
- [142] Xie P, Xing E. A constituent-centric neural architecture for reading comprehension. In: Proc. of the 55th Annual Meeting of the Association for Computational Linguistics. 2017,1:1405–1414.
- [143] Cui Y, Liu T, Chen Z, *et al.* Consensus attention-based neural networks for Chinese reading comprehension. In: Proc. of the COLING 2016, the 26th Int'l Conf. on Computational Linguistics: Technical Papers. 2016. 1777–1786.
- [144] Cui YM, Liu T, Xiao L, Chen ZP, Ma WT, Che WX, Wang SJ, Hu GP. A Span-extraction dataset for Chinese machine reading comprehension. arXiv Preprint arXiv: 1810.07366, 2018.

附中文参考文献:

- [16] 吴为章.汉语句群.北京:商务印书馆,2000.
- [17] 邢福义.汉语复句研究.北京:商务印书馆,2001.
- [18] 姚双云.复句关系标记的搭配研究与相关解释[博士学位论文].武汉:华中师范大学,2006.
- [19] 李艳翠.汉语篇章结构表示体系及资源构建研究[博士学位论文].苏州:苏州大学,2015.
- [21] 蒋玉茹,宋柔.基于广义话题理论的话题句识别.中文信息学报,2012,26(5):114–120.
- [22] 宋柔.汉语篇章广义话题结构的流水模型.中国语文,2013,6:483–494.
- [23] 尚英,宋柔,卢达威.广义话题结构理论视角下话题自足句成句性研究.中文信息学报,2014,28(6):107–113.
- [24] 马建忠.马氏文通.北京:商务印书馆,1929.
- [25] 黎锦熙,刘世儒.中国语法教程.天津:天津大众出版社,1952.
- [26] 褚晓敏,朱巧明,周国栋.自然语言处理中的篇章主次关系研究.计算机学报,2017,40(4):842–859.
- [28] 曲承熹.汉语篇章语法.北京:北京语言大学出版社,1998.

- [29] 刘礼进.英汉篇章结构模式对比研究.广州:中山大学出版社,2011.166-178.
- [30] 孙坤.话题链视角下的汉英篇章组织模式对比研究.解放军外国语学院学报(社会科学版),2013,36(3):12-18.
- [31] 王建国.论话题的延续:基于话题链的汉英篇章研究.上海:上海交通大学出版社,2013.
- [32] 周强,周晓聪.基于话题链的汉语语篇连贯性描述体系.中文信息学报,2014,28(5):102-111.
- [33] 奚雪峰.汉语篇章话题结构:表示体系、资源构建及其分析研究[博士学位论文].苏州:苏州大学,2017.
- [34] 奚雪峰,孙庆英,周国栋.面向意图性的篇章话题结构分析研究与展望.计算机学报(优先出版),2017. <http://kns.cnki.net/KCMS/detail/11.1826.TP.20170601.1917.016.html>
- [41] 乐明.汉语篇章修辞结构的标注研究.中文信息学报,2008,22(4):19-23.
- [42] 陈莉萍.汉语篇章结构标注的理论支撑.南京航空航天大学学报(社科版),2008,10(3):68-71.
- [45] 张牧宇,秦兵,刘挺.中文篇章级句间语义关系体系及标注.中文信息学报,2014,28(2):28-36.
- [46] 舒江波.面向中文信息处理的复句关系词自动标识研究[博士学位论文].武汉:华中师范大学,2011.
- [47] 周强.汉语句法树库标注体系.中文信息学报,2004,18(4):2-9.
- [48] 奚雪峰,褚晓敏,孙庆英,周国栋.汉语篇章微观话题结构建模与语料库构建.计算机研究与发展,2017,54(8):1833-1852.
- [75] 周小佩,洪宇,车婷婷,等.一种无指导的隐式篇章关系推理方法研究.中文信息学报,2013,27(2):17-26.
- [78] 张牧宇,宋原,秦兵,等.中文篇章级句间语义关系识别.中文信息学报,2013,27(6):51-58.
- [79] 涂眉,周玉,宗成庆.基于最大熵的汉语篇章结构自动分析方法.北京大学学报(自然科学版),2014,50(1):125-132.
- [82] 孙成,孔芳.基于转移的中文篇章结构解析研究.中文信息学报,2018,32(12):26-34.
- [92] 孔芳,周国栋.基于树核函数的中英文代词消解.软件学报,2012,34(5):1085-1099. <http://www.jos.org.cn/1000-9825/4044.htm> [doi: 10.3724/SP. J.1001.2012.04044]
- [98] 王厚峰.指代消解的基本方法和实现技术.中文信息学报,2002,16(6):10-18.
- [99] 王厚峰,何婷婷.汉语中人称代词的消解研究.计算机学报,2001,24(2):136-143.
- [100] 王厚峰,梅铮.鲁棒性的汉语人称代词消解.软件学报,2005,16(5):700-707. http://www.jos.org.cn/jos/ch/reader/create_pdf.aspx?file_no=20050508&journal_id=jos [doi: 10.1360/jos160700]
- [101] 李国臣,罗云飞.采用优先选择策略的中文人称代词的指代消解.中文信息学报,2005,19(4):24-30.
- [102] 周俊生,黄书剑,陈家骏,等.一种基于图划分的无监督汉语指代消解算法.中文信息学报,2007,21(2):77-82.
- [103] 杨勇,李艳翠,周国栋,等.指代消解中距离特征的研究.中文信息学报,2008,22(5):39-44.
- [104] 王海东,胡乃全,孔芳,等.指代消解中语义角色特征的研究.中文信息学报,2009,23(1):23.
- [105] 李渝勤,甘润生,杨永红,等.基于特征分选策略的中文共指消解方法.计算机工程,2011,37(18):180-182.
- [106] 张牧宇,黎耀炳,秦兵,等.基于中心语匹配的共指消解.中文信息学报,2011,25(3):3-9.
- [107] 宋洋,王厚峰.基于马尔可夫逻辑的中文零指代消解.计算机研究与发展,2015,52(9):2114-2122.
- [128] 滕佳月,李培峰,朱巧明.基于全局优化的中文事件同指消解方法.北京大学学报(自然科学版),2016,52(1):97-103.
- [129] 滕佳月.中文事件同指消解方法研究[硕士学位论文].苏州:苏州大学,2016.



孔芳(1977-),女,江苏扬州人,博士,教授,博士生导师,CCF 专业会员,主要研究领域为自然语言理解,篇章分析.



周国栋(1967-),男,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为自然语言理解,机器学习.



王红玲(1975-),女,博士,副教授,CCF 专业会员,主要研究领域为自然语言理解,信息抽取.