



现代图书情报技术 » 2012, Vol. 28 » Issue (5) : 48-54

知识组织与知识管理

[最新目录](#) | [下期目录](#) | [过刊浏览](#) | [高级检索](#)<< [Previous Articles](#) | [Next Articles](#)
>>

面向甲骨文的实例机器翻译技术研究

袁冬¹, 熊晶², 刘永革²

1. 中国海洋大学计算机科学与技术系 青岛 266100;
2. 安阳师范学院计算机与信息工程学院 安阳 455002

Yuan Dong¹, Xiong Jing², Liu Yongge²

1. Department of Computer Science and Technology, Ocean University of China, Qingdao 266100, China;
2. School of Computer and Information Engineering, Anyang Normal University, Anyang 455002, China

- 摘要
- 参考文献
- 相关文章

Download: PDF (850KB) [HTML](#) (1KB) Export: BibTeX or EndNote (RIS) [Supporting Info](#)

摘要 提出基于实例的甲骨文释文机器翻译方案,研究实例库的构建流程、实例句相似度算法和实例检索算法等关键技术,并通过实现一个机器翻译系统,验证所提出方法的有效性。实验结果表明,该方法得到的翻译结果能够满足甲骨文学习者的阅读要求。

关键词: [甲骨文](#) [基于实例的机器翻译](#) [双语平行语料库](#) [句子相似度](#)

Abstract: This paper introduces an Example-Based Machine Translation (EBMT) method for Oracle Bone Inscriptions(OBI). The key technologies in the example library building process, similarity algorithm of example sentences and example retrieval algorithm are proposed. And a machine translation system is developed to verify the proposed method. Experimental results show that the target sentences obtained by the machine translation can meet the demands of OBI researchers.

Keywords: [Oracle bone inscriptions](#), [EBMT](#), [Bilingual parallel corpus](#), [Sentence similarity](#)

收稿日期: 2012-03-13;

基金资助:

本文系国家自然科学基金项目“基于甲骨文语料库的计算机辅助考释技术研究”(项目编号:60875081)和河南省教育厅科学技术研究重点项目“基于本体的甲骨文知识共享平台构建方法研究”(项目编号:12A520003)的研究成果之一。

引用本文:

袁冬,熊晶,刘永革.面向甲骨文的实例机器翻译技术研究[J] 现代图书情报技术, 2012,V28(5): 48-54

Yuan Dong, Xiong Jing, Liu Yongge .Research on Example-based Machine Translation for Oracle Bone Inscriptions[J] , 2012,V28(5): 48-54

链接本文:

<http://www.infotech.ac.cn/CN/> 或 <http://www.infotech.ac.cn/CN/Y2012/V28/I5/48>

- [1] 江铭虎.自然语言处理[M].北京:高等教育出版社,2006.(Jiang Minghu. Natural Language Processing[M]. Beijing: Higher Education Press, 2006.)
- [2] 顾绍通. 甲骨文数字化处理研究述评[J]. 西华大学学报: 自然科学版, 2010, 29(5): 38-42.(Gu Shaotong.Review on Digitization Processing of Jiaguwen [J]. *Journal of Xihua University:Natural Science Edition*, 2010, 29(5): 38-42.)
- [3] 陈光田.古文字与古代汉语学习的关系研究[J]. 新乡学院学报: 社会科学版, 2010,24(4):125-127.(Chen Guangtian. Discussion on the Function of Ancient Words in Learning Ancient Chinese[J]. *Journal of Xinxiang Teachers College:Social Sciences Edition*, 2010,24(4): 125-127.)
- [4] 王宇信, 杨升南, 聂玉海. 甲骨文精粹释译[M].昆明: 云南人民出版社, 2004. (Wang Yuxin, Yang Shengnan, Nie Yuhai.Oracle Bone Inscriptions Pithiness Explanation[M]. Kunming: Yunnan People's Publishing House, 2004.)
- [5] 侯宏旭, 刘群, 那顺乌日图. 基于实例的汉蒙机器翻译[J]. 中文信息学报, 2007, 21(4): 65-72. (Hou Hongxu, Liu Qun, Nasun Urt.Example Based Chinese-Mongolian Machine Translation[J]. *Journal of Chinese Information Processing*, 2007, 21(4): 65-72.) 
- [6] 刘群. 汉英机器翻译若干关键技术研究[M]. 北京: 清华大学出版社, 2008. (Liu Qun. Research on Some Key Techniques in Chinese-English Machine Translation [M]. Beijing: Tsinghua University Press, 2008.) 
- [7] 姜迎春, 雪艳. 词语对齐与机器翻译问题研究——以汉蒙机器翻译为例[J]. 民族翻译, 2010(1): 91-95.(Jiang Yingchun,Xue Yan. Word Alignment and Machine Translation Research-Taking Chinese-Mongolian as Example[J]. *Minority Translators Journal*, 2010(1): 91-95.) 
- [8] Cai H Y, Jiang M H, Deng B X, et al. Method Combining Rule-based and Corpus-based Approaches for Oracle-bone Inscription Information






Service

- ▶ 把本文推荐给朋友
- ▶ 加入我的书架
- ▶ 加入引用管理器
- ▶ Email Alert
- ▶ RSS

作者相关文章

- ▶ 袁冬
- ▶ 熊晶
- ▶ 刘永革

Processing[C]. In: *Proceedings of the 2006 International Conference on Intelligent Computing(ICIC'06)*. Berlin, Heidelberg: Springer-Verlag, 2006: 736-741.

- [9] 马小虎,杨亦鸣,黄文帆,等. 甲骨文轮廓字形生成技术与通用甲骨文字库的建设[J]. 语言文字应用, 2004(3): 105-111. (Ma Xiaohu, Yang Yiming, Huang Wenfan, et al. Research on the Technology of the Automatic Generation "Jiaguwen" Outline Font and Building of Universal Jiaguwen Font [J]. *Applied Linguistics*, 2004(3): 105-111.) 
- [10] 顾绍通,马小虎,杨亦鸣. 基于字形拓扑结构的甲骨文输入编码研究[J]. 中文信息学报, 2008,22(4): 123-128. (Gu Shaotong, Ma Xiaohu, Yang Yiming. Topological Frame Based Input Method Coding of Jiaguwen[J]. *Journal of Chinese Information Processing*, 2008,22(4): 123-128.)
- [11] Wang A M, Wang J P, Ge Y Q, et al. Research on Intelligent Oracle Bone Fragments Rejoining Technology [J]. *Applied Mechanics and Materials*, 2011,50-51: 594-598.
- [12] Gao F, Liu Y G, Xiong J. Ontology-based Semantic Annotation for Oracle Bone Inscriptions[C]. In: *Proceedings of the 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)*. 2011:5087-5090.
- [13] 王爽,熊德兰,王晓霞. 基于实例的古文机器翻译设计与实现[J]. 许昌学院学报, 2009,28(5): 88-91. (Wang Shuang, Xiong Delan, Wang Xiaoxia. The Research and Implementation of Example-based Machine Translation of Ancient Chinese[J]. *Journal of Xuchang University*, 2009,28(5): 88-91.)
- [14] Fang M, Jiang X, Zhao Q, Jiang Y. Automatic Choosing of English Rhymes in Translation of Chinese Ancient Poems[C]. In: *Proceedings of the 2nd International Symposium on Knowledge Acquisition and Modeling(KAM'09)*. 2009: 434-437.
- [15] 郭锐,宋继华,廖敏. 基于自动句对齐的相似古文句子检索[J]. 中文信息学报, 2008, 22(2): 87-91. (Guo Rui, Song Jihua, Liao Min. Ancient Sentence Search Based on Sentence Auto-Alignment in Parallel Corpus of Ancient and Modern Chinese[J]. *Journal of Chinese Information Processing*, 2008, 22(2): 87-91.) 
- [16] 宋继华,胡佳佳,孟蓬生,等. 古今汉语平行语料库的语料构建[J]. 现代教育技术, 2008,18(1): 92-99. (Song Jihua, Hu Jiajia, Meng Pengsheng, et al. The Construction of Corpora in a Classic-Cotemporary Chinese Parallel Corpus[J]. *Modern Educational Technology*, 2008,18(1): 92-99.) 
- [17] 刘一曼. 甲骨文字的特点及主要内容[J]. 档案管理, 2000(1): 40-41. (Liu Yiman. The Characteristics and Main Content of Oracle Bone Inscriptions [J]. *Archives Management*, 2000(1): 40-41.)
- [18] 刘志基. 简论甲骨文字频的两端集中现象[J]. 语言研究, 2010,30(4): 114-122. (Liu Zhiji. On the Two Concentration Features of the Character Frequency of Oracle Bone Inscriptions[J]. *Studies in Language and Linguistics*, 2010,30(4): 114-122.)
- [19] 郑继娥. 甲骨文祭祀卜辞语言研究[M]. 成都: 巴蜀书社, 2007. (Zheng Ji'e. Oracle Bone Sacrifice Inscriptions Language Research[M]. Chengdu: Bashu Publishing House, 2007.) 
- [20] Xiong J, Gao F, Liu Y. Word Segmentation Method for Oracle Bone Inscriptions Based on Dictionary and Syntactic Rules[C]. In: *Proceedings of the 3rd International Conference on Computer Design and Applications (ICCD)*. IEEE, 2011: 592-595.
- [21] 吕学强,任飞亮,黄志丹,等. 句子相似模型和最相似句子查找算法[J]. 东北大学学报: 自然科学版, 2003, 24(6): 531-534. (Lv Xueqiang, Ren Feiliang, Huang Zhidan, et al. Sentence Similarity Model and the Most Similar Sentence Search Algorithm[J]. *Journal of Northeastern University: Natural Science*, 2003, 24(6): 531-534.)
- [22] 贾兆红,陈华平. 基于改进遗传算法的权重发现技术[J]. 计算机工程, 2007,33(5): 156-157. (Jia Zhaohong, Chen Huaping. Weights Finding Based on Improved Genetic Algorithms[J]. *Computer Engineering*, 2007,33(5): 156-157.) 
- [23] 张刚,杨海成,经小川,等. 基于遗传算法的技术成熟困难度计算方法[J]. 北京理工大学学报, 2011, 31(4): 472-476. (Zhang Gang, Yang Haicheng, Jing Xiaochuan, et al. Computation of Advancement Degree of Difficulty Based on Genetic Algorithm[J]. *Transactions of Beijing Institute of Technology*, 2011, 31(4): 472-476.)
- [24] Papineni K, Roukos S, Ward T, et al. BLEU: A Method for Automatic Evaluation of Machine Translation[C]. In: *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*. 2002: 311-318.
- [25] Doddington G. Automatic Evaluation of Machine Translation Quality Using N-gram Co-occurrence Statistics[C]. In: *Proceedings of the 2nd International Conference on Human Language Technology Research*. San Francisco: Morgan Kaufmann Publishers Inc., 2002: 138-145.
- [26] 黄瑾,刘洋,刘群. 机器翻译评测介绍[EB/OL]. [2012-02-05]. <http://lib.ict.ac.cn/libraryict/ITL/data/2007/5/机器翻译评测介绍.pdf>. (Huang Jin, Liu Yang, Liu Qun. Introduction to Machine Translation Evaluation[EB/OL]. [2012-02-05]. <http://lib.ict.ac.cn/libraryict/ITL/data/2007/5/%BB%FA%C6%F7%B7%AD%D2%EB%C6%C0%B2%E2%BD%E9%C9%DC.pdf>.)
- [1] 王志超,翁楠,王宇.基于主题句相似度的标题党新闻鉴别技术研究[J]. 现代图书情报技术, 2011,(11): 48-53
- [2] 何维,王宇.基于句子关系图的网页文本主题句抽取*[J]. 现代图书情报技术, 2009,3(3): 57-61
- [3] 王森,王宇.基于文本结构树的论文复制检测算法[J]. 现代图书情报技术, 2009,(10): 50-55
- [4] 廉站俊,吕学强,张玉杰,施水才.基于句子相似度计算的信息抽取*[J]. 现代图书情报技术, 2007,2(6): 38-41
- [5] 化柏林.基于句子匹配的文章自写度测评系统[J]. 现代图书情报技术, 2007,2(11): 40-44
- [6] 秦新国.基于句子相似度的文档复制检测算法研究[J]. 现代图书情报技术, 2007,2(11): 63-66

