

Rationalism and Empiricism in Machine Translation

Jin'an Xu

School of Computer and Information Technology, Beijing Jiaotong University, Beijing

Email: xja2010@gmail.com

Abstract: By introducing the advantages and disadvantages of the rationalism and empiricism in natural language processing, this paper proposes several basic ideas of how to combine the rationalist approach and empiricist approach according to the current state of development in machine translation, and discusses the change of the rationalism and empiricism in word sense disambiguation, finally summarizes the developmental tendency of machine translation.

Keywords: machine translation; nature language processing; computational linguistics; rationalist approach; empiricist approach

机器翻译中的理性主义和经验主义

徐金安

北京交通大学计算机与信息技术学院, 北京, 100044

Email: xja2010@gmail.com

摘要: 本文通过介绍自然语言处理技术中的基于规则的理性主义方法和基于统计的经验主义方法的优缺点; 结合机器翻译的发展现状, 提出了规则 and 统计相结合的机器翻译方法的基本框架, 论述了词义消歧中的理性主义方法和经验主义方法的发展状况; 总结了机器翻译的发展趋势。

关键词: 机器翻译; 自然语言处理; 计算语言学; 理性主义方法; 经验主义方法

1 引言

随着计算机技术和自然语言处理技术研究的发展, 机器翻译先后经历了 60 年代末的低谷期, 70 年代中期的发展期, 80 年代开始进入繁荣期和商品化进程, 90 年代开始进入网络化时期。如今, 机器翻译已经开始进入翻译语言和服务形式的多样化、使用终端小巧化、系统大规模集成化和自动语音翻译系统实用化的实用阶段^[1]。但是, 机器翻译研究现在仍然面临诸多难题, 涉及自然语言处理的多个层面, 诸如, 词性标注、句法结构消歧、语义消歧、知识自动提取等等。在方法论层面, 无论是基于规则的理性主义方法还是基于统计的经验主义方法, 都必须面对自然语言问题的复杂性和多变性, 两种方法既有的理论模型和方法都还不够完善, 发展空间很大, 有待进一步的改进。

本文在第二部分介绍基于规则的理性主义方法和基于统计的经验主义方法的优缺点; 第三部分分析介绍规则和统计相结合的机器翻译研究策略和基本思想; 并

探讨语义消歧中的基于规则和统计相结合的研究策略及其部分成果; 最后, 对机器翻译的发展趋势进行探讨。

2 理性主义方法和经验主义方法的优缺点

冯志伟老师在宗成庆老师的《统计自然语言处理》一书的序言二中^[2], 从哲学和方法论的高度系统地描述了自然语言处理中基于规则的理性主义方法和基于统计的经验主义方法之间水火不相容、充满矛盾和对立的历史发展过程, 透彻地分析了自然语言处理中基于规则的理性主义方法和基于统计的经验主义方法的优缺点, 提倡彼此之间的有机结合, 相互取长补短, 进行优势互补, 为自然语言处理和机器翻译研究指明了方向, 在理论研究和实践应用方面都具有重要的指导意义。

2.1 基于规则的理性主义方法

在自然语言处理的发展过程中, 以乔姆斯基的形式语言为理论基础的基于规则的理性主义方法, 在 20 世纪 60 年代末到 70 年代占据主流地位, 基于统计的经验主义方法在此期间则几乎完全受到了排斥。乔姆斯基主张以一种公式化、形式化的方法, 严格按照一

资助信息: 中央高校基本科研业务费专项资金资助

定的规则来描述自然语言的特征, 试图以有限的文法规则描述无限的语言现象。这种方法从 1970 年前后到 1990 年前后发展很快, 且日趋成熟。但是, 随着研究的深入, 人们也发现: 基于规则的语言处理系统性能有限, 使用范围受限某些特定的小领域, 在应用于不同领域时, 系统的扩展性能很差, 往往要求整套规则重写, 系统精度很难得到进一步提高。

按照冯志伟老师的说法, 基于规则的理性主义方法在自然语言处理中的优点体现在以下几个方面^[2]:

- 规则方法对抽象的语言特征具有很强的形式描述能力和形式生成能力;
- 规则方法对句子结构以及长距离依存关系的处理能力很强;
- 规则方法的语言模型结构清晰易懂;
- 规则方法在本质上无方向性。其语言模型同时适用于分析和生成处理;
- 规则方法在自然语言处理中具有解决浅层到深层多层次问题的能力;
- 规则方法与一些高效算法具有兼容性。

在机器翻译研究领域, 基于规则的机器翻译系统的优点体现在:

- 基于规则的机器翻译系统对不同语种在语法上的相似度的依存度很低、处理能力较强。
- 基于规则的机器翻译系统具有对知识表达的抽象程度高、代表性强的特点;
- 对文法结构具有很强的保持能力;
- 对不同语料的覆盖率高;
- 系统运行占用资源少。

同时, 按照冯志伟老师的说法, 基于规则的理性主义方法在自然语言处理中的缺点体现在^[2]:

- 使用基于规则的理性主义方法研制的自然语言处理系统的鲁棒性和灵活性都很差;
- 规则方法需要语言学家对繁杂的语言现象进行大量的分析, 语法分析任务繁重, 很难用计算机对基于规则的语言模型进行泛化处理;
- 使用基于规则的理性主义方法研制的自然语言处理系统具有很强的针对性, 而且扩展性能差、升级困难;
- 规则方法不具备统计方法的学习能力, 因而领域适应能力较差;

在机器翻译研究领域, 基于规则的机器翻译系统的缺点体现在:

- 基于规则的机器翻译系统的语法分析和生成规

则主要由人工编写, 规则的主观性强, 规则的一致性难以保障;

- 知识获取难、工作量大;
- 不利于系统扩充, 尤其对非规范的特殊语言现象缺乏相应的处理能力等。

2.2 基于统计的经验主义方法

基于统计的经验主义方法从 20 世纪 90 年代开始得到了快速发展, 自然语言处理研究也“重新回到经验主义”的发展道路上来, 特别是大规模语料库的出现, 促使基于统计的经验主义方法成为当前自然语言处理技术的主流。

按照冯志伟老师的说法, 基于统计的经验主义方法在自然语言处理中的优点体现在^[2]:

- 统计方法具有良好的数学模型, 无指导的学习能力和较强的知识自动获取能力;
- 统计方法可以比较容易地采用大规模训练语料不断提高系统的性能;
- 统计方法可以很容易地结合多种规则, 通过处理各种各样的约束条件问题, 不断改善处理效果;
- 统计方法善于处理模糊的语言现象;

在机器翻译研究领域, 基于统计的机器翻译系统的优点体现在:

- 目前, 网上用于开发统计机器翻译系统的资源丰富, 易于实现;
- 良好的统计翻译模型能够融合更多的句法结构和语义语法信息;
- 统计机器翻译的翻译性能可通过大规模语料训练进行改善。

按照冯志伟老师的说法, 基于统计的经验主义方法在自然语言处理中的缺点是^[2]:

- 使用基于统计的经验主义方法的自然语言处理系统, 运行时间和模型参数成比例线性增长, 运行效率相对低下, 消耗的资源巨大;
- 统计方法需要大规模语料库进行训练学习, 对语料库的质量要求较高;
- 统计方法容易出现数据稀疏问题; 且随着训练数据规模的增大呈线性增长, 需要进行平滑处理进行合理解决;

在机器翻译研究领域, 基于统计的机器翻译系统的缺点是:

- 统计机器翻译系统存在系统运行消耗的资源巨大、需要大规模双语平行语料库进行训练学习,

语料的选择和处理工程量也很巨大；而翻译模型和语言参数的精确性直接依赖于语料的多少，翻译质量的高低主要取决于概率模型的好坏和语料库的质量及其覆盖能力等；

- 统计机器翻译的翻译效果对源语言和目标语言间在语法上的相似度存在一定的依赖性，相似度越小，需要调整的参数越多，系统的运行效率越低，改善系统翻译精度的难度系数越大。

3 规则与统计相结合的机器翻译研究策略

3.1 规则与统计相结合的机器翻译基本思想

规则方法和统计方法在自然语言处理中的优缺点各异，使用不同方法研制的机器翻译系统的性能表现也不尽相同，对译文质量的保障程度深浅不一，翻译精度也还不如人意。如何通过规则和统计方法的相互结合，实现能够相互取长补短的新理论和新方法，是机器翻译研究人员普遍关注的热点。

我们认为，在机器翻译研发过程中，要实现规则和统计方法的有机结合，可以本着“具体情况具体分析，分门别类地解决问题”的原则；可以从规则方法和统计方法各自的基本原理出发，结合基于规则的机器翻译系统和基于统计的机器翻译系统各自的主体特征，采用不同的研究方法和研究策略，从理论和实践两方面完善机器翻译理论，实现机器翻译系统的有效改良和升级。

在统计机器翻译系统中，应采用多种策略融合规则，促使规则在统计机器翻译系统中起到关键性辅助作用，根据语言的某些特定语法现象，建立适当的规则体系和模型参数，以在统计机器翻译模型中融合更多的句法结构和语义信息，是当前的统计机器翻译系统改善翻译质量，提高翻译精度的一种思维模式和方法，也是实现规则和统计相结合的机器翻译的一种有效途径。

另一方面，在基于规则的机器翻译系统中，融入多种统计模型，使统计模型在规则型机器翻译系统中起到关键性辅助作用，针对语言中各类复杂文法、非规范文法的具体特征，构建适合其语法特点统计模型；同时，建立特定文法与统计模型之间的相关约束机制，作为规则与统计方法相结合的切入点。这种思维方式是当前基于规则的机器翻译系统改善翻译质量，扩展领域适应能力的一种方法，也是实现规则和统计相结合的机器翻译的另一种有效途径。

以基于转换的规则型机器翻译系统为例，本文所

主张的规则和统计相结合的基本思想框架主要表现在以下几个方面：

1. 针对语言中的简单语法和规范语法，采用以规则为主、统计为辅的策略；
2. 针对语言中的非规范文法现象、复杂程度高的语法现象，采用统计为主、规则为辅的策略；
3. 针对语言中不同性质的非规范文法现象、复杂程度高的语法现象，有针对性地构建适合其语法特点的统计模型来解决具体问题。
4. 建立特定文法与统计模型之间的相关约束机制，作为规则和统计方法相结合的切入点。

我们相信，在对既有基于规则的机器翻译系统进行领域扩展时，采用上述基本思路和方法，可以有效保留既有的翻译规则；可以避免整套翻译规则重写；可以节省造价、缩短研发周期并提高翻译质量。另一方面，在研发新的基于规则的机器翻译系统时，采用上述基本思路，首先对文法的复杂程度进行适当合理的评估，界定其使用规则或统计模型的粗略标准；然后，对相对简单文法和规范文法制定翻译规则；同时，有针对性地对某些特定语法现象建立适合其语法特征的统计模型，根据特定文法与统计模型之间的相关约束机制界定其使用方法。按照这种思路开展研发工作，同样可以简化机器翻译知识获取难度，提高翻译精度和系统适应能力，减少人工编写规则和制作翻译词典的工作量，缩短研发周期，降低研制成本。

3.2 词义消歧中的经验主义和理性主义

词义消歧（Word Sense Disambiguation, WSD）的任务是确定一个多义词在给定的上下文语境中的具体含义，是自然语言理解的基础研究课题之一，在机器翻译、信息检索、文本处理、语音处理、语法和句法分析等领域应用广泛。自机器翻译研究诞生以来词义消歧一直备受计算语言学家们的关注，是自然语言处理中的一项艰巨任务，国内外有众多的学者致力于这项任务的研究工作。

词义消歧方法也经历了基于规则的理性主义方法到基于统计的经验主义方法的变迁，并逐步走上融合规则、义类词典等知识库和多种统计方法相结合的集成消歧策略的道路。

词义消歧方法可分为基于知识的方法和基于统计的方法。基于知识的方法又可以分为基于规则的方法和基于词典的方法。基于统计的词义消歧方法按照机器学习方法可分为有监督(supervised)学习方法和无监

督(unsupervised)学习方法。有监督词义消歧通常被看作词义分类问题,无监督词义消歧则被看作聚类问题。另外,由于语言学家提供的各类词典或知识库是获取词义消歧知识的重要来源,因此,人们往往也把基于词典或知识库的消歧方法单独区分开来进行研究。目前,基于词典的词义消歧方法是一种非常重要的消歧策略,如 WordNet、HowNet、EDR 电子词典等。近年来,运用 bootstrapping、AdaBoost MH、Conditional Random Fields(CRF)、Co-Training 等方法的半监督学习技术在这种方法中起到了重要作用。

典型的有监督学习方法有:基于互信息的消歧方法^[3-9]、基于决策树模型的消歧方法^[10]、基于贝叶斯分类器的消歧方法^[11]、基于最大熵模型的消歧方法^[12]和基于支持向量机^[13](SVM)的消歧方法等。典型的无监督学习方法有:基于双语语料的词义消歧方法^[14-24]、基于 WEB 的词义消歧方法^[25]、基于聚类的词义消歧方法^[26-35]等。其中,典型的基于聚类的词义消歧有:Latent Semantic Analysis (LSA)^[27-29]、Hyperspace Analogue to Language (HAL)^[30,31]、Clustering by Committee (CBC)^[32]等等。典型的基于词典的分析方法有:基于机读词典的词义消歧、基于词典语义定义的方法^[36]、基于义类词典的方法^[37,38]、基于双语词典的方法^[13,14]、基于领域信息的词义消歧方法^[39,40]、基于百科知识(Wikipedia)的消歧方法^[41]等等。其中,基于义类词典的方法又可以细分为:基于概念区域密度的词义分析方法^[42-44]、基于结构化语义关系的图论式词义消歧方法^[45-52]等等。

另外,词义消歧的系统测评是词义消歧研究的重要环节之一,SENSEVAL 是由国际计算语言学联合会(ACL)词汇兴趣小组(SIGLEX)于 1997 年开始组织的关于词义消歧的公共测评任务。第三次 SENSEVAL 评测结果显示,性能表现最好的系统在粗粒度定义下词义消歧的正确率和召回率依旧保持在 79.3%,在细粒度定义的情况下的正确率和召回率约为 72.9%,性能表现排名前几位的系统分别采用了朴素贝叶斯分类器、SVM、最大熵方法,并且结合了多种知识库。从测试结果来看,词义消歧技术还有很大的改善空间^[2]。

目前,很多学者倾向于应用综合多种方法进行集成消歧,包括融合规则、义类词典等知识库、结合多种统计消歧方法,取得了良好的消歧效果。

另外,基于规则和统计相结合的研究策略所取得的成果几乎涵盖自然语言处理任务中的各个层面,比如:单词分割^[53]、词性标注^[54]、句法分析^[55]、语块提

取^[56,57]、自动查错^[58]、口语理解^[59]、自动文摘^[60]、机器翻译^[57,61-66]等等。

4 总结和展望

本文介绍自然语言处理技术中基于规则的理性主义方法和基于统计的经验主义方法的优缺点,以及机器翻译研究中基于规则的方法和基于统计的方法的优缺点。并结合当前机器翻译研究的发展现状,提出了规则和统计相结合的机器翻译方法的基本框架,论述了词义消歧中的理性主义方法和经验主义方法的变迁和发展趋势。

机器翻译依旧面临众多难题,需要研究人员针对机器翻译研究任务中的具体问题进行分析、判断和取舍,拆分、整合和改进;需要研究人员不断进取,推动机器翻译研究的进步和发展。

References (参考文献)

- [1] JinAn Xu, Prospects in Machine Translation, 2010 Cross-Strait Conference on Information Science and Technology, CSCIST 2010, QingDao.
- [2] Ch. Zong, Statistical Natural Language Processing, Tsinghua University Press, May 2008 first edition.
- [3] Brown P. F., Della Pietra S. A., Della Pietra V. J., Mercer R. L., Word-sense Disambiguation Using Statistical Methods. In Proceedings of the 29th ACL. Pp.264-270.
- [4] Brown P. F. et. al, A Statistical Approach to Sense Disambiguation in Machine Translation. In Proceedings of the Fourth DARPA Workshop on Speech and Natural Language. Morgan Kaufman Publishers. Pp.146-151.
- [5] Carpuat M., Shen Y. H., Yu X. F., Wu D. K., Toward Integrating Word Sense and Entity Disambiguation into Statistical Machine Translation. In Proceeding of the International Workshop on Spoken Language Translation (IWSLT), November 27-28, Kyoto, Japan. Pp37-44.
- [6] Carpuat M., Wu D. K., Improving Statistical Machine Translation Using Word Sense Disambiguation. In Proceeding of 2007 Joint Conference on the EMNLP-CoNLL 2007, Prague, Pp.61-72.
- [7] Carpuat M., Wu D. K., Context-Dependent Phrasal Translation Lexicons for Statistical Machine Translation. In Proceeding of Machine Translation Summit XI. Copenhagen, Sept. 10-14, Pp.73-80.
- [8] Carpuat M., Wu D. K., How Phrase Sense Disambiguation outperforms Word sense Disambiguation for Statistical Machine Translation. In Proceeding of the 11th Conference on Theoretical and Methodological issues in Machine Translation (TMI 2007), Sweden, Sept. 7-9, Pp43-52.
- [9] Chan Y. S., Ng H. T., Chiang D., Word Sense Disambiguation Improves Statistical Machine Translation. In Proceeding of the 45th Annual Meeting of the Association for Computational Linguistics (ACL), June 23-30. Pp33-40.
- [10] Y. Wu, M. Wang, P. Jin, Sh. Yu, Ensembles of Classifiers for Chinese Word Sense Disambiguation, Journal of Computer Research and Development, 2008, 45(8):1354-1361.
- [11] Gale W. A., Church K. W., Yarowsky D., A Method for Disambiguating word senses in a large corpus. Computers and Humanities, pp. 415-439.
- [12] G. Chao, M.G. Dyer, Maximum entropy models for word sense disambiguation, in: Proceedings of 19th International Conference

- on Computational Linguistics (COLING-02), 2002, pp.1-7.
- [13] Dagan I., Itai A., Markovitch S., Two languages are more informative than one. In Proc. of the 29th ACL, 1991. pp. 130-137.
 - [14] Dagan I., Itai A., Word sense disambiguation using a second language monolingual corpus. Computational Linguistics, 1994, 20(4), pp. 563-596.
 - [15] Resnik P., Yarowsky D., A perspective On word sense disambiguation methods and their evaluation. In Proc. of the ACL SIGLEX Workshop on Tagging Text with Lexical Semantics: Why, What and How. Morristown: Association for Computational Linguistics, 1997, pp.79-86.
 - [16] Eseudcro G, Marquez L, Rigau G., Boosting applied to word sense disambiguation. In Proc. of the 12th European Conference on Machine Learning. Berlin, Heidelberg, Springer-Verlag, 2000, pp. 129-141.
 - [17] Ide N, Ejavec T., Tufts D., Sense discrimination with parallel corpora. In Proc. of the ACL SIGLEX Workshop on word Sense Disambiguation: Recent Successes and Future Directions. Morristown: ACL, 2002, pp.54-60.
 - [18] Li C., Li H., Word translation disambiguation using bilingual bootstrapping. In Proc. of the 40th ACL, 2002, pp.343-351.
 - [19] Ng H. T., Wang B., Chan Y. S., Exploiting parallel texts for word sense disambiguation: An empirical study. In Proc. of the 41st ACL, 2003, pp455-462.
 - [20] Diab M., Reanik P., An unsupervised method for word sense tagging using parallel corpora. In Proc. of the 40th ACL, 2002, pp.255-262.
 - [21] Diab M., Word sense disambiguation within a multilingual framework, PhD. Thesis, University of Maryland College, 2003.
 - [22] Diab M., An unsupervised approach for bootstrapping Arabic word sense tagging. In Proc. of the Arabic Based Script Languages, COLING 2004.
 - [23] Diab M., Relieving the data acquisition bottleneck in word sense disambiguation. In Proc. of the 42th ACL, 2004, pp.303-310.
 - [24] Bhattacharya I., Getoor L., Bengio Y., Unsupervised sense disambiguation using bilingual probabilistic models. In Proc. of the 42nd ACL, 2004, pp287-294.
 - [25] Klapaftis IP, Manandhar S., Google & WordNet based word sense disambiguation. In Proc. of the 22nd ICML Workshop on Learning & Extending Ontologies. New York: Association for Computing Machinery. 2005.
 - [26] Schutze H., Automatic word sense discrimination. Computational Linguistics, 1998, 24(1):97-123.
 - [27] Deerwester S. et. al, Indexing by latent semantic analysis. Journal of the Amedean Society for Information Science, 1990, 41(6) pp391-407.
 - [28] Landauer T. K., Dumais S. T., A solution to Plato' s problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. Psychological Review, 1997, 104(2) pp211-240.
 - [29] Landauer T. K., Foltz P. W., Laham D., An introduction to latent semantic analysis. Discourse Processes, 1998, 25(2), pp.259-284.
 - [30] Burgess C., Lund K., Modifying parsing constraints with high-dimensional context space. Language and Cognitive Processes, 1997, 12 (2-3), pp177-210.
 - [31] Burgess C., Lurid K., The dynamics of meaning in memory. In Dietrich E., Markman A., eds. Cognitive Dynamics: Conceptual Representational Change in Humans and Machines. 2000, pp117-156.
 - [32] Lin D. K., Pantel P., Concept discovery from text. In: Proc. of the 19th International Conference on Computational Linguistics(COLING-2002). pp. 577-583.
 - [33] Pedersen T., Bruce R., Distinguishing word senses in untagged text. In Proc. of the 2nd Conf. on Empirical Methods in Natural Language Processing. 1997. pp.197-207.
 - [34] Pedersen T., Knowledge lean word sense disambiguation. In Proc. of the 15th National Conf. on AI. 1998, pp.800-805.
 - [35] Purandare A. et. al, Word Sense Discrimination by Clustering Contexts in Vector and Similarity Spaces. In Ng H. T., Riioff E., eds. Proc. of the Conf. on Computational Natural Language Learning. Boston: Quinlan Publishing. 2004, pp 41-48.
 - [36] Lesk M Automated sense disambiguation using machine readable dictionaries: How to tell a pine cone from all ice cream cone. In Proc. of the 1986 SIGDOC Conference, 1986, New York, pp24-26.
 - [37] Walker D., Knowledge Resource Tools for Accessing Large Text Files. In Machine Translation: Theoretical and Methodological Issues, Cambridge University. Pp247-261.
 - [38] Yarowsky D., Word-Sense Disambiguation Using Statistical Models of Roget's Categories Trained on Large Corpora, In Proceedings, COLING-92, Nantes, France, 1992. Pp.241-246.
 - [39] Connroni J. P. et al., Dewey decimal classification and Relative Index. Albany: Forest Press, 1998.
 - [40] Gonzalo J., Verdejio F., Peters C., Caizolari N., Applying EuroWordNet to cross-language text retrieval. Computers and the Humanities, 1998, 32(2-3), pp185-207.
 - [41] Mihalcea R., Using wikipedia for automatic word sense disambiguation. In Computational Linguistics (NAACL). Morristown: Association for Computational Linguistics, 2007, pp196-203.
 - [42] Agirre E., Rigau G., A proposal for word Sense disambiguation using conceptual distance In Mitkov R., Nicotov N., eds, Proc. of the 1st International Conference on Recent Advances in NLP. 1995. pp162-171.
 - [43] Rosso P., Masulli F., Buscaldi D., Pla F., Mofina A., Automatic noun disambiguation. In Proc. of the 4th International Conference on Computational Linguistics and Intelligent Text Processing, Mexico City: Springer Verlag, 2003, pp273-276.
 - [44] Buscaldi D., Rosso P., Masulli F., Integrating conceptual density with WordNet domains and CALD glosses for noun sense disambiguation. In Proceeding of the 4th International Conference on Espana for Natural Language Processing (EsTAL). LNAI 3230, Alicante, 2004. Pp183-194.
 - [45] Litkowski K. C., Use of machine readable dictionaries in word sense disambiguation for Senseval-2. In Proceeding of the ACL/SIGLEX Senseval-2. 2001.
 - [46] McCarthy D., Koeling R., Weeds J., Carroll J., Using automatically acquired predominant senses for word sense disambiguation. In Proc. of the ACL/SIGLEX Senseval-3 Workshop. 2004.
 - [47] Morris J., Hirst G., Lexical cohesion computed by thesaurus relations as an indicator of the structure of text. Computational Linguistics, 1991, 17(1), pp21-48.
 - [48] Galley M., McKeown K., Improving word sense disambiguation in lexical chaining. In Proceeding of the 18th International Joint Conference, on AI (UCAI 2003). Pp1486-1488.
 - [49] Mihalcea R., Tarau P., Figa E., Page Rank on semantic networks with application to word sense disambiguation. In Proceeding of the 20th International Conference on Computational Linguistics (COLING-2004).
 - [50] Navigli R., Velardi P., Structural semantic interconnections: A knowledge-based approach to word sense disambiguation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005, 27(7), pp1075-1086.
 - [51] Agirre E., Martinez D., de Lacalle O. L., Soroa A., Two graph-based algorithms for state-of-the-art WSD. In Proceeding of the Conference on Empirical Methods in Natural Language Processing (EMNLP). Barcelona: Association for Computational Linguistics, 2006, pp583-593.
 - [52] Sinha R., Mihalcea R., Unsupervised graph-based word sense disambiguation using measures of word semantic similarity. In Proceeding of the IEEE International Conference on Semantic Computing (ICSC). Washington: IEEE Computer Society, 2007. Pp363-369.
 - [53] <http://chasen.naist.jp/hiki/ChaSen/>
 - [54] Q. Zhou, Chinese Corpus Tagging Using Rule Techniques and Statistics Techniques, Journal of Chinese Information Processing, 9(3): PP.1-10..

- [55] Y. Liu, Analyzing Chinese with Rule-based Method Combined with Statistics-based Method, *Journal of Computer Engineering And Application.*, 2002 38(7), Pp.3-6.
- [56] B. Jiang, Q. Zhang, B. Kang, B. Chang, Chinese Multi-word Chunks Extraction for Computer Aided Translation, *Journal of Chinese Information Processing*, 2007 21(1), pp.9-16.
- [57] J. Li, Q. Liu, Sh. Bai, Chinese Chunking Parsing Using Rule-Based And Statistics-Based Methods, *Journal of Computer Research And Development*, 2002 39(4), pp.385-391.
- [58] Y. Zhang, Y. Cao, Sh. Yu, A Hybrid Model of Combining Rule-based and Statistics-based Approaches for Automatic Detecting Errors in Chinese Text. *Journal of Chinese Information Processing*, 2006 20(4), pp.1-7.
- [59] G. Xie, Ch. Zong, B. Xu, Chinese Spoken Language Analyzing Facing the Middle Semantic Representation, *Journal of Chinese Information Processing*, 2003 17(1). Pp.1-6.
- [60] J. Fu, Q. Chen, Research on Automatic Summarization Based on Rules and Statistics for Chinese Texts, *Journal of Chinese Information Processing*, 2006 20(5). pp.10-16.
- [61] T. Zhao, E. Xun, Y. Fan, Disambiguation Strategy of E-C Machine Translation system BT863-II. *CJCAI'98*, Pp.207-212..
- [62] H. Guo, G. Hu, Combining Statistical Model with Linguistic Knowledge in Target Language Generation. *The 4th China Workshop on Machine Translation*, 2002, pp.110-115..
- [63] Y. Liu, Syntactic Score and Semantic Score, *Journal of Chinese Information Processing*, 2000 14(4).: 17-24.
- [64] A. Eisele, et. al, Hybrid machine translation architectures within and beyond the EuroMatrix project. In *Proceedings of the 12th annual conference the European Association for Machine Translation (EAMT 2008)*, pages 27-34, Hamburg, Germany, September 2008.