

P.O.Box 8718, Beijing 100080, China	Journal of Software, Nov 2005,16(11):1909-1919
E-mail: jos@iscas.ac.cn	ISSN 1000-9825, CODEN RUXUEW, CN 11-2560/TP
http://www.jos.org.cn	Copyright © 2005 by The Editorial Department of <i>Journal of Software</i>

# 基于语义单元表示树剪枝的高速多语言机器翻译

高小宇, 高庆狮, 胡 玥, 李 莉

[Full-Text PDF](#) [Submission](#) [Back](#)

高小宇<sup>1</sup>, 高庆狮<sup>1,2</sup>, 胡 玥<sup>1,2</sup>, 李 莉<sup>2</sup>

<sup>1</sup>(中国科学院 计算技术研究所, 北京 100080)

<sup>2</sup>(北京科技大学 智能、语言与计算机研究所, 北京 100083)

作者简介: 高小宇(1976—), 男, 福建漳州人, 硕士, 主要研究领域为数据库, 自然语言与机器翻译, 模糊集合, 网络安全; 高庆狮(1934—), 男, 教授, 博士生导师, 中国科学院院士, 主要研究领域为大型、巨型计算机体系结构, 并行算法与并行处理系统, 自然语言及其处理, 网络安全, 模糊集合, 人工智能及其应用; 胡玥(1963—), 女, 博士, 副教授, 主要研究领域为并行算法, 自然语言与机器翻译, 网络安全; 李莉(1980—), 女, 博士生, 主要研究领域为自然语言处理.

联系人: 高庆狮 E-mail: qsgao@public.bta.net.cn

Received 2004-08-10; Accepted 2005-02-03

## Abstract

In this paper, a high speed multi-language machine translation approach based on pruning on tree representations of semantic elements is proposed. This is the multi-language machine translation with the following several characteristics: Chinese segmentation before translation into another languages is not necessary, and the translation time is  $O(L)$  rather than general  $O(LN)$ , where  $L$  is the length of text,  $N$  is the number of semantic elements (i.e. number of language patterns) in SER-base, even if  $N$  is hundreds of thousands or millions.

Gao XY, Gao QS, Hu Y, Li L. High speed multi-language machine translation based on pruning on the tree of representations of semantic elements. *Journal of Software*, 2005,16(11):1909-1919.

DOI: 10.1360/jos161909

<http://www.jos.org.cn/1000-9825/16/1909.htm>

## 摘要

提出一种基于语义单元表示树剪枝的高速多语言机器翻译方法. 此方法是一种将汉语翻译到其他语种不需要先进行汉语切分的多语言机器翻译方法. 而且翻译时间为 $O(L)$ 而不是 $O(LN)$ , 其中,  $L$ 是文本的长度,  $N$ 是语义单元库中语义单元的数量, 一般有数十万或者数百万.

基金项目: Supported by the National Natural Science Foundation of China under Grant No.60343010 (国家自然科学基金); the Foundation of Institute of Computing Technology, the Chinese Academy of Sciences under Grant No.20016250 (中国科学院计算技术研究所创新基金)

## References:

- [1] Simon HA. Cognitive Psychology, Lecture. Beijing: Peking University, 1983.
- [2] Gao QS, et al. The principle of human-like machine translation. Computer Research and Development, 1989,26(2):1-7 (in Chinese with English abstract).
- [3] Brown RD. Example-Based machine translation in Pangloss system. In: Proc. of the COLING'96. Copenhagen, 1996. 169-174.
- [4] Brown RD. Automated generalization of translation examples. In: Proc. of the COLING 2000. Sarbrucken, 2000. 125-131.
- [5] Zhao TJ. The principle of machine translation. Harbin: Harbin Institute of Technology, 2001 (in Chinese).

- [6] Och FJ, Weber H. Improving statistical natural language translation with categories and rules. In: Proc. of the 35th Annual Conf. of the Association for Computational Linguistics and the 17th Int'l Conf. on Computational Linguistics. Montreal, 1998. 985-989.
- [7] Och FJ, Tillmann C, Ney H. Improved alignment models for statistical machine translation. In: Proc. of the Joint SIGDAT Conf. on Empirical Methods in Natural Language Processing and Very Large Corpora. College Park: University of Maryland, 1999. 20-28.
- [8] Kaji H, Kida Y, Morimoto Y. Learning translation templates from bilingual texts. In: Proc. of the COLING'92. 1992. 672-678.
- [9] Abney SP. Parsing by chunks. In: Berwick R, Abney S, Tenny C. Principle-Based Parsing. Kluwer Academic Publishers, 1991. 257-278.
- [10] Cicerkli I, Guvenir HA. Learning translation templates from examples. Information Systems, 1998,23(6):353-363.
- [11] Miller GA. WordNet. 1998. <http://www.wordnet.princeton.edu>
- [12] Dong ZD, Dong Q. How Net. 1999 (in Chinese). <http://www.keenage.com>
- [13] Gao XY, Gao QS, Hu Y. The multi-language machine translation approach based on Semantic language. Report of Invent Patent, ZL01131689.6, 2001 (in Chinese with English abstract).
- [14] Gao QS, Hu Y, Li L, Gao XY. Semantic language and multi-language MT approach based on SL. Journal of Computer Science & Technology, 2003,18(6):848-852.

附中文参考文献:

- [2] 高庆狮,等.类人机译原理.计算机研究与发展,1989,26(2):1-7.
- [5] 赵铁军.机器翻译原理.哈尔滨:哈尔滨工业大学出版社,2001.
- [12] 董振东,董强.知网.1999. <http://www.keenage.com>
- [13] 高小宇,高庆狮,胡玥.基于语义语言的机器翻译系统和方法.发明专利报告, ZL01131689.6, 2001.