

# 特定领域中的语义校对系统 (YYJDS) 的设计与实现\*

郑逢斌<sup>1,2</sup> 姜保庆<sup>1,2</sup> 郭拯危<sup>2</sup> 邱小平<sup>1</sup> 乔保军<sup>2</sup>

<sup>1</sup>(西南交通大学智能控制开发中心, 四川 成都 610031);

<sup>2</sup>(河南大学计算机与信息工程学院, 河南 开封 475001)

E-mail: zhengfb@henu.edu.cn

**摘要:** 本文介绍了应用于报刊书籍政治性错误校对的一个语义校对软件系统的设计原理和实现过程。本系统先把文本切分成词语, 找出敏感词, 合并成术语, 部分术语通过短语本身或左右片段即可做出正误判断, 另一部分术语要用整个句子语义与知识库进行模糊匹配才能得到它的错误隶属度。

**关键词:** 语义校对; 敏感词语; 术语; 错误隶属度

**作者简介:** 郑逢斌(1963—), 男, 河南开封人, 博士生, 副教授, 主要研究领域为自然语言理解, 软件工程; 姜保庆(1964—), 男, 河南开封人, 博士生, 副教授, 主要研究领域为自然语言理解, 人工智能; 郭拯危(1964—), 女, 河南开封人, 讲师, 主要研究领域为数据挖掘, 计算机网络。

## Design and Execution of Semantic Collation System in a Special Field\*

ZHENG Fengbin<sup>1,2</sup> JIANG Baoqing<sup>1,2</sup> GUO Zhengwei<sup>2</sup>  
QIU Xiaoping<sup>1</sup> QIAO Baojun<sup>2</sup>

<sup>1</sup>(Intelligence Control Development Center, Southwest Jiaotong University, Chengdu 610031, China);

<sup>2</sup>(College of Computer and Information Engineering, Henan University, Kaifeng 475001, China)

E-mail: zhengfb@henu.edu.cn

**Abstract:** In this paper, the design principle and execution process of the new semantic collation system about political error in newspapers and books are introduced. Following contents are included in the system: first, sentences are cut into words; second, sensitive terms are picked up. For these sensitive terms, the correctness of some can be judged by themselves or near segments, whereas for others, their error membership can only be obtained by fuzzy matching between the sentences they belong to and that of knowledge database.

**Key words:** Semantic collation; sensitive word; term; error membership;

\* 本项目受到国家自然科学基金资助(60074014)。

\* Supported by the National Natural Science Foundation of China under Grant No. 60074014