

**An Online Domain-based Portuguese-Chinese Machine Translation System**

**by**

**Wu Fei**

**1999 June**



**Software Engineering Program  
Faculty of Science and Technology  
University of Macau**

---

**TABLE OF CONTENTS**

**LISTING OF FIGURES..... III**

**LISTING OF TABLES..... IV**

**CHAPTER 1 MOTIVATIONS AND GOALS ..... 1**

    1.1 Verifying the interoperability by formal methods ..... 1

    1.2 Learning linguistic knowledge by statistical approach ..... 1

    1.3 Proposing a computing framework for multilingual cyberspace..... 2

    1.4 Specifying a mechanized Hoare logic for distributed computing ..... 2

**CHAPTER 2 INTRODUCTION ..... 4**

    2.1 Overview..... 4

    2.2 Reading the thesis ..... 5

**CHAPTER 3 THE VICISSITUDE OF MACHINE TRANSLATION..... 7**

    3.1 The birth, embryonic and demise of machine translation ..... 7

    3.2 The normal machine translation process ..... 9

    3.3 The renaissance of online machine translation in Cyberspace ..... 12

    3.4 Put it together: Language- the latest frontier in Cyberspace ..... 14

**CHAPTER 4 COMPUTING-FREE THEOREMS..... 18**

    4.1 The retrospect and prospect of computing platforms ..... 19

    4.2 Why and how cross-platform: Interoperability ..... 20

    4.3 Why and how Cross-Language: interlanguage..... 26

    4.4 Put it together: Theoretical or Mechanical oriented computing-free..... 29

**CHAPTER 5 THE SHARED-OBJECT DISTRIBUTED TECHNICAL EXTRACTING TERM  
WORKSTATION ..... 31**

    5.1 The historical review of distributed computing method..... 31

    5.2 A brief introduction to CORBA ..... 35

    5.3 Some topics about distributed computing ..... 38

    5.4 Obtaining computing performance balance using shared-object ..... 43

    5.5 A mechanized Hoare logic for distributed computing..... 46

    5.6 The shared-object distributed technical extracting term workstation ..... 52

        5.6.1 *Background of term extraction*..... 52

        5.6.2 *Related work*..... 54

        5.6.3 *Extracting technical terms from bilingual texts*..... 55

        5.6.4 *Design Issues*..... 59

---

5.6.5 <i>The evaluation of algorithm</i> .....	62
5.7 Put it together: The specter of Amdahl's law .....	63
<b>CHAPTER 6 THE ONLINE DOMAIN-BASED PORTUGUESE-CHINESE MACHINE TRANSLATION SYSTEM</b> .....	<b>65</b>
6.1 Online MT Opportunities and Challenges .....	66
6.2 Converting online hypertext into plain text .....	67
6.3 Retrieving domain-specific information .....	69
6.3.1 <i>Detecting subtopic change via co-occurrence patterns</i> .....	69
6.3.2 <i>Methods of comparing adjacent blocks</i> .....	73
6.3.3 <i>Domain-specific information retrieval algorithm</i> .....	76
6.4 Searching the nearest examples .....	80
6.5 Future Directions .....	88
6.6 Qualitative evaluation .....	90
6.7 Put it together: mobile translator is viable? .....	92
<b>CHAPTER 7 A FRAMEWORK FOR MULTILINGUAL COMPUTING</b> .....	<b>94</b>
7.1 Definition of "Multilingualization" .....	94
7.2 A framework for multilingual computing .....	95
7.2.1 <i>Practical issue</i> .....	96
7.2.2 <i>Former models for multilingual language support</i> .....	98
7.2.3 <i>The Architecture of framework</i> .....	99
7.3 Put it together: Think globally and compute locally .....	102
<b>CHAPTER 8 CONTRIBUTIONS AND FUTURE WORKS</b> .....	<b>104</b>
<b>BIBLIOGRAPHY</b> .....	<b>105</b>
<b>APPENDICES</b> .....	<b>1</b>
Appendix A Listings of Source Code .....	1