

**A Portuguese-Chinese Corpus-Based
Machine Translation System**

by

Pun Chi Man
(m-95-6325-7)

This thesis is submitted in partial fulfillment
of the requirements for the Degree of

Master of Science
in
Software Engineering

Supervisor:

Prof. Li Yi Ping

**FACULTY OF SCIENCE AND TECHNOLOGY
UNIVERSITY OF MACAU**

November 1997

Contents

- Acknowledgment..... (i)**
- Abstract..... (ii)**
- Contents..... (iii)**
- Listing of Figures..... (v)**
- Listing of Tables..... (vi)**

- 1 Introduction..... (1)**
 - 1.1 Overviews..... (1)
 - 1.2 Reading the Thesis..... (2)

- 2 Background and Previous Work..... (3)**
 - 2.1 Translator’s Workstation..... (4)
 - 2.2 Corpus-based Machine Translation..... (4)

- 3 Theoretical Aspects of Corpus-based Machine Translation**
..... (7)
 - 3.1 Introduction..... (7)
 - 3.2 Corpus-based Machine Translation..... (10)
 - 3.3 Building bilingual corpus-base by
aligning Portuguese-Chinese parallel texts..... (12)
 - 3.4 Nearest Example Retrieval..... (26)
 - 3.5 Phrase Substitution Technique in Translation..... (27)

- 4 System Architecture and Design..... (30)**
 - 4.1 System Architecture..... (30)
 - 4.2 System Design..... (31)

4.3	Top Level Design.....	(32)
4.4	Second Level Design.....	(34)
4.5	Third Level Design.....	(37)
5	Evaluation and Discussion.....	(65)
5.1	Evaluation of Translator's Workstation.....	(65)
5.2	Evaluation of CBMT Central.....	(70)
5.3	Discussion.....	(72)
6	Contributions and Future Work.....	(74)
6.1	Contributions.....	(74)
6.2	Limitations and Future work.....	(75)
	References and Bibliography.....	(77)

Appendices

A	Sample Portuguese-Chinese Parallel Texts.....	(79)
B	Listing of Source Code.....	(85)