iQA: Intelligent Question Answering from the Web

by

Chan Mei Pou

Master of Science in Software Engineering

2005



Faculty of Science and Technology
University of Macau

TABLE OF CONTENTS

LIST OF FIGURES IV
LIST OF TABLES V
GLOSSARY VI
LIST OF ABBREVIATIONS VII

Part I: Introduction on the Development of Question Answering Systems 1

1 INTRODUCTION 1
2 DEVELOPMENT OF QA SYSTEMS 8
2.1 TREC-8 QA Track (1999) 10
2.2 TREC-9 QA Track (2000) 11
2.3 TREC 2001 QA Track 11
2.4 TREC 2002 QA Track 13
2.5 TREC 2003 QA Track 13
2.6 Summary 14
3 DATA REDUNDANCY – IS MORE ALWAYS BETTER? 16

Part II: System Architecture and Implementation Details Of iQA 18

4 MOTIVATION 18
5 SYSTEM OVERVIEW OF IQA 20
6 QUESTION ANALYSIS MODULE 24
6.1 Query formulation 24
6.1.1 Exact query 26
6.1.2 Inexact query 27
6.2 Question classification 28

7 DOCUMENT ANALYSIS MODULE 37
8 ANSWER EXTRACTION MODULE 40
8.1 Extraction of simple potential answers 40
8.1.1 Answers in the form of expected answer type 41
8.1.2 Answers depending on the recognition of question focus 41
8.1.3 Quantitative Answers 42
8.1.4 Description answers 43
8.2 Answer justification 44
9 ABBREVIATION ANALYSIS MODULE 48
9.1 Which are abbreviation terms? 48
9.2 Filter corresponding potential long forms 49
9.3 Work in action 50

Part III: Evaluation and Conclusions **54**

10 EVALUATION 54
11 CONCLUSIONS 57
11.1 Response time 57
11.2 Correctness of answers 58
11.3 Future work 59
BIBLIOGRAPHY 60
APPENDIX A: TOOLS USED FOR IQA 64
APPENDIX B: QUESTION ANSWERING SYSTEMS AVAILABLE ON THE WEB 66
APPENDIX C: MRR OF ANSWERBUS, BRAINBOOST, IQA BASED ON THE FIRST 100 QUESTIONS OF TREC 2003 68
APPENDIX D: SCREEN SHOT ON EVALATUION RESULT FOR QUESTIONS WITH ABBREVIATION TERMS PRESENT IN CONTEXT OF MACAO LOCAL REGION 73

APPENDIX E: TREC 2003 QUESTIONS 85