



UNIVERSITI PUTRA MALAYSIA

**POLA GRAMMAR FOR AUTOMATED MARKING OF MALAY SHORT
ANSWER ESSAY-TYPE EXAMINATION**

MOHD JUZAIDDIN AB AZIZ

FSKTM 2008 13



**POLA GRAMMAR FOR AUTOMATED MARKING OF MALAY SHORT ANSWER
ESSAY-TYPE EXAMINATION**

By

MOHD JUZAIDDIN AB AZIZ

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia,
in Fulfilment of the Requirements for the Degree of Doctor of Philosophy**

April 2008



DEDICATION

I would like to dedicate my work to my mother, Jawahariah Haji Omar who has passed away during my tenure as a graduate student, may Allah bless you. To my beloved wife, Tengku Siti Meriam bt Tengku Wook, Kak Long Ma, Abg Ngah Addin and Adik Paan. To Abah, Kak Long and family, Kak Baiyah and family, Not and family, Naru and family, and Ad and also to all my family in laws.



Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of
the requirement for the degree of Doctor of Philosophy

**POLA GRAMMAR FOR AUTOMATED MARKING OF MALAY SHORT ANSWER
ESSAY-TYPE EXAMINATION**

By

MOHD JUZAIDDIN AB AZIZ

Chairman: Fatimah Dato Ahmad, PhD

Faculty: Computer Science and Information Technology

The efforts to mark essay-typed examination automatically for English have been started since 1960s. But, there was not many attempted to mark the Malay essay-typed examination automatically. One of works to mark the Malay essay-typed examination was conducted to mark the History subject that focused on the temporal values of the essays rather than the sentence structure of the Malay language. The subjective nature of sentence construction makes the process to identify the important points addressed in the essays difficult to be carried out. Short answer essay-typed examination requires the students to answer the questions with sentences in a short paragraph. While marking the examination scripts manually, the lecturers or teachers have to identify the sentence similarity between the sentences in the answers scripts and answer scheme. The answers scripts have to be carefully read and understood by the examiner in order to award fair marks. Sentence similarity is defined as the sentences that have similar meaning but different from the words used or sentence structure. The problems in this research are solved by using pola grammar techniques where the sentence similarity is identified by a representation of the Malay language structure and a Malay verbs synonymous thesaurus. Pola grammar produces



a Grammatical Relations (GRs) representation. The technique is an enhancement of the four basic Malay language representations. The representations are Noun Phrase + Noun Phrase (NP+NP), Noun Phrase + Verb Phrase (NP+VP), Noun Phrase + Preposition Phrase (NP+PP), and Noun Phrase + Adjective Phrase (NP+AP). In order to recognize the sentence structure, a finite state automata (FSA) is constructed based on the pola grammar rules. The effectiveness of the FSA is computed in an application known as an Automatic Marking System for Short Answer Essay-typed examination (AMS-SAE). There are two tests conducted using AMS-SAE, first, 78 short answer essays in the form of simple, complex and conjoined sentences have been computed for their similarity. The results show that the average scores different to human for simple sentences is 0.032, complex is 0.113 and conjoined is 0.042. Second, the answers from a three questions for a compiler examination is recorded and tested with AMS-SAE and human. Each question which has 30 to 45 answers in the form of short essay-typed has proved that AMS-SAE can be accepted to produce similar marks to human when the Mann-Whitney test and *t*-test have shown that the marks have a strong significant relationship.



Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia
sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

**NAHU POLA UNTUK PEMARKAHAN SECARA AUTOMATIK PEPERIKSAAN
JAWAPAN PENDEK BERJENIS ESEI BAHASA MELAYU**

Oleh

MOHD JUZAIDDIN AB AZIZ

Pengerusi: Fatimah Dato Ahmad, PhD

Fakulti: Sains Komputer dan Teknologi Maklumat

Usaha untuk memarkahkan peperiksaan berasaskan esei secara automatik untuk Bahasa Inggeris telah bermula sejak 1960an. Namun, tidak banyak usaha dilakukan untuk memarkahkan peperiksaan Bahasa Melayu yang berasaskan esei secara automatik. Satu daripada usaha untuk memarkahkan peperiksaan Bahasa Melayu berasaskan esei adalah untuk memarkahkan subjek Sejarah yang focus kepada nilai temporal esei berbanding dengan struktur ayat Bahasa Melayu. Sifat subjektif pembangunan ayat menjadikan proses untuk mengenal pasti fakta penting yang dinyatakan dalam esei sukar dilakukan. Peperiksaan berasaskan esei pendek memerlukan para pelajar untuk menjawab soalan dengan ayat dalam perenggan yang pendek. Semasa menyemak skrip peperiksaan, para pensyarah atau guru perlu mengenal pasti persamaan ayat antara ayat dalam skrip jawapan dan skema pemarkahan. Skrip jawapan perlu dibaca dengan teliti dan difahami oleh pemeriksa untuk memberikan markah yang setimpal. Persamaan ayat didefinisikan sebagai ayat-ayat yang mempunyai maksud yang sama tetapi berbeza dari penggunaan perkataan atau struktur ayat. Masalah dalam penyelidikan ini diselesaikan menggunakan teknik pola grammar iaitu persamaan ayat dikenalpasti menggunakan perwakilan struktur ayat Bahasa



Melayu dan tesaurus Kata Kerja Bahasa Melayu. Kaedah ini merupakan perluasan daripada empat perwakilan asas bahasa Melayu yang dicadangkan oleh ahli Linguistik Bahasa Melayu. Perwakilan tersebut merupakan Frasa Nama + Frasa Nama (FN+FN), Frasa Nama + Frasa Kerja (FN + FK), Frasa Nama + Frasa Hubung (FN + FH), dan Frasa Nama + Frasa Adjektif (FN + FA). Untuk mengenal pasti struktur ayat, satu automata keadaan terhingga (FSA) telah dibangunkan berasaskan kepada peraturan pola grammar. Keberkesanan FSA diuji menggunakan aplikasi yang dikenali sebagai *Automatic Marking System for Short Answers Essay-typed examination* (AMS-SAE). Terdapat dua ujian yang dilakukan menggunakan AMS-SAE, pertama, 78 esei pendek dengan format ayat mudah, kompleks dan tergabung diuji untuk persamaan ayat. Keputusan menunjukkan bahawa purata markah yang berbeza dengan manusia adalah sebanyak 0.032 untuk ayat mudah, 0.113 untuk ayat kompleks dan 0.042 untuk ayat tergabung. Kedua, jawapan daripada tiga soalan peperiksaan untuk subjek pengkompil direkodkan dan diuji menggunakan AMS-SAE dan manusia. Setiap soalan yang mempunyai antara 30 hingga 45 jawapan dengan format esei pendek telah membuktikan bahawa AMS-SAE boleh diterima untuk menghasilkan markah yang sama dengan manusia apabila ujian Mann-Whitney dan *t*-test menunjukkan bahawa markah yang dihasilkan mempunyai hubungan yang amat berkaitan.



ACKNOWLEDGEMENTS

Firstly, I would like to thank Allah swt for giving me support and strength to finish writing this thesis.

I would like to thank my main supervisor Assoc. Prof. Dr. Hajah Fatimah Dato Ahmad for the guidance in every aspect of my thesis at the Faculty of Computer Science and Information Technology, Universiti Putra Malaysia. My grateful thanks also go to my supervisory committee members Assoc. Prof. Dr. Abdul Azim Abdul Ghani and Assoc. Prof. Dr. Ramlan Mahmud. I would like to highlight how fortunate I have been as a graduate student with respect to all my supervisors. The general direction of the work described here stresses the impact of my supervisors' individual backgrounds and strengths, while reflecting the freedom I have had in shaping my own research.

There are many people who have made this research possible, and whom I would like to thank. During my first year, there was a large crop of graduate students who had impact on me such as Khatim, Taufik, Burn and Hakim. During several of my years as a graduate student, Dr. Zukeri and Zaharuddin were great resources of sharing ideas, and Dr. Tg. Nor Rizan who has checked my English.

Finally, I would like to appreciate the Government of Malaysia for my financial support through Kementerian Pengajian Tinggi, study leave from Universiti Kebangsaan Malaysia, and also for providing the IRPA research fund 04-02-02-0053-EA-220.



I certify that an Examination Committee has met on **15th April, 2008** to conduct the final examination of Mohd Juzaidin Ab Aziz on his Doctor of Philosophy thesis entitled “Pola Grammar for Automated Marking of Malay Short Answer Essay-Type Examination” in accordance with Universiti Pertanian Malaysia (Higher Degree) ACT 1980 and Universiti Pertanian (Higher Degree) Regulations 1981. The Committee recommends that the candidate be awarded the relevant degree. Members of the Examination Committee are as follows:

Hamidah Ibrahim, PhD

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Safaai Deris, PhD

Professor
Faculty of Computer Science and Information System
Universiti Teknologi Malaysia
(External Examiner)

Md Nasir Haji Sulaiman, PhD

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)

Masrah Arifah Azmi Murad, PhD

Lecturer
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Internal Examiner)



This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Fatimah Dato Ahmad, PhD

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Chairman)

Abdul Azim Abd. Ghani, PhD

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

Ramlan Mahmud, PhD

Associate Professor
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
(Member)

AINI IDERIS, PhD

Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:



DECLARATION

I declare that the thesis is my original work except for quotations and citations, which have been duly acknowledged. I also declare that it has not been previously and is not concurrently submitted for any other degree at Universiti Putra Malaysia or at any other institutions.

MOHD JUZAIDDIN BIN AB AZIZ

Date:



TABLE OF CONTENTS

	Page
DEDICATION	ii
ABSTRACT	iii
ABSTRAK	v
ACKNOWLEDGEMENTS	vii
APPROVAL	viii
DECLARATION	x
LIST OF TABLES	xv
LIST OF FIGURES	xvii
LIST OF ABBREVIATIONS	xix
CHAPTER	
I INTRODUCTION	
1.1 Introduction	1
1.2 Pola Grammar	2
1.3 Problem Statement	3
1.4 Objectives	5
1.5 Contributions of the Research	6
1.6 Research Methodology	7
1.7 Organization of the Thesis	8
II LITERATURE REVIEW	
2.1 Introduction	10
2.2 Automatic Marking System	11
2.3 Automated Essay Grading: An Application for Historical Malay Text	12
2.4 C-rater	13
2.4.1 Semantic Gap	16
2.4.2 Canonical Representation	17
2.5 UCLES: Automatic Marking of Short Textual Answers	18
2.6 Automark	21
2.7 Automated Text Marker (ATM)	22
2.7.1 ATM Architecture	23
2.7.2 ATM Structured Representation Schemes	25
2.8 Summary	26
III SENTENCE SIMILARITY AND SEMANTIC PROCESSING	
3.1 Introduction	29
3.2 Malay Language Processing	30
3.3 Sentence Similarity	31
3.4 Lexical Distributional Similarity	32



3.5	Shallow Syntactic Analysis	33
3.6	Information Extraction	35
3.7	Word Similarity	38
3.8	Computing Distributional Similarity	40
3.9	Semantic Role	41
3.10	Semantic Relations	42
3.11	Lexical Functional Grammar	43
3.12	Finite State Automata Conversion	44
3.13	Head-Driven Phrase Structure Grammar	46
3.14	Summary	47
IV	THE STRUCTURE OF THE MALAY LANGUAGE	
4.1	Introduction	48
4.2	Malay Sentence	48
4.3	<i>Yang</i> as a pronoun	50
4.4	Sentence	51
4.5	Malay Grammar	53
	4.5.1 Sentence Grammar	54
	4.5.2 Pola Grammar	57
4.6	Parsing the Grammar	61
4.7	Pola Grammar: Analysis of Malay Structure	62
	4.7.1 Adjunct	64
	4.7.2 Subject	65
	4.7.3 Conjunction	66
	4.7.4 Predicate	66
	4.7.5 Verb's position	66
4.8	Summary	67
V	THE POLA GRAMMAR RULES	
5.1	Introduction	68
5.2	Development of Pola Grammar Rules	68
5.3	Empirical Evaluation	77
5.4	Sentence Representation	81
5.5	Parsing and Recognizing of Pola Grammar using Finite State Automata	82
5.6	Finite State Automata (FSA)	84
	5.6.1 Subject and Predicate Recognizer	85
	5.6.2 Verb and Object Recognizer	88
5.7	Pola Grammar System Architecture	93
5.8	Test to Extract the Grammatical Relations	94
5.9	Results	96
5.10	Summary	101
VI	SYSTEM ARCHITECTURE AND DESIGN	
6.1	Introduction	102
6.2	Grading Process	102



6.3	AMS-SAE Architecture	104
6.4	Pola Grammar (PolaG) Module	104
	6.4.1 Tokenizer	105
	6.4.2 Collocation	106
	6.4.3 Reconstructor	106
	6.4.4 Trimmer	107
	6.4.5 Recognizer	107
6.5	Pola Compare (PolaC) Module	109
	6.5.1 Subject Comparison	110
	6.5.2 Verb Comparison	110
	6.5.3 Extra Phrase Comparison	112
	6.5.4 Subject and Object Comparison	112
	6.5.5 Synonym Module	113
6.6	Thesaurus	113
6.7	Pola Grading (polaGr) Module	114
	6.7.1 Assumptions	115
	6.7.2 Score Calculation: An Example	115
6.8	Implementation	117
6.9	Summary	120
VII	RESULTS AND DISCUSSIONS	
7.1	Introduction	121
7.2	Structure Testing	121
7.3	Results	124
7.4	Testing with the Existing Answer Scripts	127
	7.4.1 Statistical Test	127
	7.4.2 Experimental Hypotheses	128
	7.4.3 Mann-Whitney Test on Question 1	130
	7.4.4 <i>t</i> -Test of Computer versus Human for Question 2	134
	7.4.5 Mann-Whitney Test on Question 3	138
7.5	Discussions	141
VIII	CONCLUSIONS AND FUTURE WORKS	
8.1	Introduction	145
8.2	Review of Thesis	145
8.3	Major Contributions	146
	8.3.1 Collocation Determination and Grammatical Relations Extraction	147
	8.3.2 Grammatical Relations Comparison	148
	8.3.3 Scoring Scheme	148
8.4	Future Works	149
8.5	Limitations	150
8.6	Conclusion	151



REFERENCES	152
BIODATA OF THE STUDENT	193
LIST OF PUBLICATIONS	194



LIST OF TABLES

Table	Page
2.1 Example questions that have been scored by c-rater	15
2.2 Tuples for four responses	16
2.3 Summarizations of the Techniques and Modules used in the Existing Applications	27
4.1 Constituents of sentences	52
4.2 Pola, subject, predicate, and object for sentences (4.13), (4.14) and (4.15)	60
4.3 Pola for sentences (4.17) and (4.18)	64
4.4 Example of the Adjuncts	65
5.1 Pola Grammar Extraction of the subject and predicate	78
5.2 Grammatical Relations extracted from the predicate	79
5.3 Subject and predicate of sentence (5.4)	79
5.4 Object and adverb of sentence (5.4)	80
5.5 Verb, conjunction and adverb of sentence (5.4)	80
5.6 FSA transition for sentence (5.1)	87
5.7 FSA transition for sentence (5.2)	87
5.8 FSA_2 transition for Sentence (5.6)	90
5.9 FSA_2 transition for the second phrase of Sentence (5.6)	90
5.10 FSA_2 transition for the third phrase of Sentence (5.6)	91
5.11 FSA_2 transition for the fourth phrase of Sentence (5.6)	92
5.12 Number of instances of GR in the first level of the test set	97
5.13 Number of instances of GR in the test set (predicate)	97



5.14	Results I for PG	98
5.15	Results II for PG	98
5.16	Results I for PSM	100
5.17	Results II for PSM	101
6.1	GRs produced by PolaM	109
7.1	Results of testing AGS-SAE with the language structure	125
7.2	Summary of the results	128
7.3	Rank computation for M_{CH}	132
7.4	f -test Two Sample for Variances	136
7.5	t -Test Two-Sample Assuming Equal Variances	130
7.6	Rank computation for M_{CH}	140



LIST OF FIGURES

Figure	Page
2.1 Automark System Architecture	22
2.2 Conceptual Dependency Groups	24
4.1 Context-Free Grammar for Malay language	55
4.2 Derivation of sentence (4.6)	56
4.3 Derivation of sentence (4.7)	56
4.4 Derivation of the new sentence (modification of sentence (4.6))	57
5.1 FSA recognizing the Subject and Predicate	86
5.2 FSA_2 recognizing the GRs in the Predicate	89
5.3 System Architecture based on Pola Grammar	94
6.1 AMS-SAE organization	103
6.2 Architecture of AMS-SAE	104
6.3 Architecture of PolaG	105
6.4 Process of tokenizing	105
6.5 Reconstructor rebuilding the conjoined sentence	107
6.6 Rebuilt of a passive sentence	107
6.7 Process of recognizing	108
6.8 The output of the pola grammar algorithm is located in the matrices	119
7.1 Comparison between a simple sentence (scheme) with a passive sentence	122
7.2 Comparison between a negative and normal sentence	123
7.3 Comparison of conjoined sentences	123
7.4 Comparison of complex sentence	124



7.5	Scores given by computer and human for Question 1	130
7.6	Properties of Question 1	133
7.7	Computer versus human for Question 2	135
7.8	Properties of Question 2	137
7.9	Computer versus human for Question 3	138
7.10	Properties of Question 3	140



LIST OF ABBREVIATIONS

AMS-SAE	Automatic Marking System for Short Answer Examination
AP	Adjective Phrase
ASDH	Average Score Different to Human
ATM	Automated Text Marker
BNF	Backus Naur Form
CD	Conceptual Dependency
CFG	Context Free Grammar
CL	Computational Linguistic
CLS	Computational Linguistic System
FSA	Finite State Automata
FSA_2	Finite State Automata 2
FSM	Finite State Machine
GRs	Grammatical Relations
HMM	Hidden Markov Model
HPSG	Head Phrase Sentence Grammar
IE	Information Extraction
LR	Left Right
MB	Memory Based
M _{CH}	Mann-Whitney Computer versus Human
MUC	Message Understanding Conference
NLP	Natural Language Processing
NP	Noun Phrase



O	Object
P	Predicate
PG	Pola Grammar
PolaC	Pola Compare
PolaG	Pola Grading
PolaM	Pola Marking
POS	Part Of Speech
PP	Preposition Phrase
PS	Phrase Structure
PSM	Parsing System for Malay language
S	Subject
SR	Semantic Roles
TAG	Tree Adjoining Grammar
TB	Transformation-Based
V	Verb
VG	Verb Group
VP	Verb Phrase

Saya mengesahkan bahawa satu Jawatankuasa Pemeriksa telah berjumpa pada 15 April 2008 untuk menjalankan peperiksaan akhir bagi Mohd Juzaidin bin Ab Aziz untuk menilai tesis Doktor Falsafah beliau yang bertajuk “Nahu Pola Untuk Pemarkahan Secara Automatik Peperiksaan Jawapan Pendek Berjenis Esei Bahasa Melayu” mengikut Akta Universiti Pertanian Malaysia (Ijazah Lanjutan) 1980 dan Peraturan Universiti Pertanian Malaysia (Ijazah Lanjutan) 1981. Jawatankuasa Pemeriksa tersebut telah memperakukan bahawa calon ini layak dianugerahi ijazah Doktor Falsafah.

Ahli Jawatankuasa Pemeriksa adalah seperti berikut:

Hamidah Ibrahim, PhD

Profesor Madya
Fakulti Sains Komputer dan Teknologi Maklumat
Universiti Putra Malaysia
(Pengerusi)

Md Nasir Haji Sulaiman, PhD

Profesor Madya
Fakulti Sains Komputer dan Teknologi Maklumat
Universiti Putra Malaysia
(Pemeriksa Dalam)

Masrah Arifah Azmi Murad, PhD

Pensyarah
Fakulti Sains Komputer dan Teknologi Maklumat
Universiti Putra Malaysia
(Pemeriksa Dalam)

Safaai Deris, PhD

Profesor
Fakulti Sains Komputer dan Sistem Maklumat
Universiti Teknologi Malaysia
(Pemeriksa Luar)

HASANAH MOHD. GHAZALI, PhD

Profesor dan Timbalan Dekan
Sekolah Pengajian Siswazah
Universiti Putra Malaysia

Tarikh: 1 April 2008



I certify that an Examination Committee has met on 15th April 2008 to conduct the final examination of Mohd Juzaidin bin Ab Aziz on his Doctor of Philosophy thesis entitled “Pola Grammar for Automated Marking of Malay Short Answer Essay-Typed Examination” in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the student be awarded the degree of Doctor of Philosophy.

Members of the Examination Committee were as follows:

Hamidah Ibrahim, PhD

Associate Professor

Faculty of Computer Science and Information Technology

Universiti Putra Malaysia

(Chairman)

Md Nasir Haji Sulaiman, PhD

Associate Professor

Faculty of Computer Science and Information Technology

Universiti Putra Malaysia

(Internal Examiner)

Masrah Arifah Azmi Murad, PhD

Lecturer

Faculty of Computer Science and Information Technology

Universiti Putra Malaysia

(Internal Examiner)

Safaai Deris, PhD

Professor

Faculty of Computer Science and Information System

University Teknologi Malaysia

(External Examiner)

HASANAH MOHD. GHAZALI, PhD

Professor and Deputy Dean

School of Graduate Studies

Universiti Putra Malaysia

Date: 25 February 2008



CHAPTER 1

INTRODUCTION

1.1 Introduction

The effort to mark essay-typed examination automatically has started in 1960s. One of the earliest software that is introduced is known as Project Essay Grade (PEG) (Page, 1967). To date, there are many software to mark English essay-typed examinations such as C-rater, E-rater, and Latent Semantic Analysis (Williams, 2001). There are also software to mark Malay essay-typed examination such as the software to mark the History subject examination (Norisma Idris and Syed Malek Fuad Duani Syed Mustapha, 2005).

The essay-typed examination can be categorized into two: long essay answers and short essay answers. The long essay answers are free text essays where the students are given a topic to be discussed in a long essay. This type of essay has common features to be marked by the lecturers such as the style of writing and the contents (Page, 1967). The style includes the punctuation and spelling. The short essay-typed answers are written in short sentences where the style is not important for marking. Marking short answer essay is relying heavily on the contents of the essays only (Pulman and Sukkarieh, 2005). Marking short answer essay-typed examination differs from marking the free test essay, where the score of the latter is the total of the style and contents (Landauer *et al.*, 1998).



The aim of this research is to mark short answer examination automatically and the focus is to investigate techniques to determine whether Malay sentences are similar. Sentences are said to be similar if the meaningful words used in the sentence are found similar. Even if they are not constructed using the same words, but, may be they are using the synonymous words. For example, sentences (1.1) and (1.2) are similar, even though they are formed with different words.

Saya berpuasa di sepanjang bulan Ramadan. --- (1.1)

Saya berlapar dan dahaga di sepanjang bulan Ramadan. --- (1.2)

To find the solution of sentence similarity, the techniques that are based on the language structure will be developed in this research. To examine the accuracy of the techniques, they will be applied to a system that can automatically mark short answers examination. The system will have a marking scheme where the correct sentences are kept. The sentences will then be compared with the answers given by the students.

1.2 Pola Grammar

Pola grammar is a technique to extract syntactic features and grammatical relations (GRs) from the Malay language structure. Language structure, sometimes referred to as language model (Collins *et al.*, 2005; Chelba and Jelinek, 1998), refers to a method for incorporating syntactic features into a language model. Syntactic features break