

**EVALUATING THE PERFORMANCES OF DICE, JACCARD,
OVERLAP, AND COSINE COEFFICIENTS IN AUTOMATED
MARKING OF ALGEBRAIC EQUATIONS**



**INSTITUT PENYELIDIKAN, PEMBANGUNAN DAN
PENGKOMERSIALAN
UNIVERSITI TEKNOLOGI MARA
40450 SHAH ALAM, SELANGOR
MALAYSIA**

PREPARED BY:

**HASNAH YUSOF
WAN ROLINI WAN YUSOFF
MASITA JAMAL**

DISEMBER 2006



Surat Kami : 600-IRDC/ST 5/3/1012
Tarikh : 08 Julai 2005

Puan Hasnah Yusof
Fakulti Teknologi Maklumat dan Sains Kuantitatif
Universiti Teknologi MARA
40450, Shah Alam
Selangor

Tuan / Puan

TAJUK PROJEK : EVALUATING THE PERFORMANCE OF THE DICE, JACCARD, OVERLAP, AND COSINE COEFFICIENTS IN AUTOMATED MARKING OF ALGEBRAIC EQUATIONS

Dengan hormatnya perkara di atas adalah dirujuk.

Sukacita dimaklumkan bahawa cadangan penyelidikan yang telah dikemukakan oleh tuan/puan bersama Puan Wan Roslini dan Puan Masita Jamal telah diluluskan.

- i. Tempoh projek penyelidikan ini ialah 12 bulan , iaitu bermula **1 Julai 2005** hingga **30 Jun 2006**.
- ii. Walaubagaimanapun, adalah dimaklumkan bahawa kos yang diluluskan adalah sebanyak **RM 8,000.00** sahaja. Diharapkan perkara ini tidak mematahkan semangat tuan/puan untuk menjalankan penyelidikan dan diharapkan kos tersebut dapat membantu pihak tuan/puan memulakan projek.
- iii. Tuan/puan juga perlu **mengemukakan semula keperluan kewangan J-Series mengikut jumlah peruntukan yang telah diluluskan**. Pembiayaan pembentangan kertas kerja dan perjalanan ke luar negara dihadkan kepada RM 2,000.00 sahaja. Permohonan peralatan penyelidikan adalah dicadangkan dari Fakulti/ Cawangan masing-masing.
- iv. Tuan/puan perlu membelanjakan **50%** daripada geran penyelidikan yang telah diluluskan bagi projek tuan/puan dalam tempoh **6 bulan** pertama projek berjalan. Sehubungan itu , pihak IRDC akan memantau penggunaan geran penyelidikan tuan/puan untuk memastikan **50%** daripada jumlah geran yang diluluskan telah dibelanjakan sehingga bulan **Disember 2005**.
- v. Semua pembelian peralatan yang kosnya melebihi RM500.00 satu item perlu menggunakan Pesanan Jabatan Universiti Teknologi MARA (LO). Pihak tuan/puan juga dikehendaki mematuhi peraturan penerimaan peralatan. Panduan penerimaan peralatan baru dan pengurusannya , dilampirkan.

PENYELIDIKAN, PEMBANGUNAN DAN PENGKOMERSILAN LANDASAN KEWIBAWAAN DAN KECEMERLANGAN

Telefon :
Penolong Naib Canselor (Penyelidikan) : 03-55442094/5
Ketua Penyelidikan (Sains Sosial dan Pengurusan) : 03-55442097
Ketua Penyelidikan (Sains dan Teknologi) : 03-55442091
Ketua INFOREC : 03-55442760
Ketua Perundingan : 03-55442100
Ketua Pengkomersilan : 03-55442750
Ketua Harta Intelek : 03-55442753
Penolong Pendaftar : 03-55442092
Pegawai Eksekutif : 03-55442057
Pejabat Am : 03-55442093/2101
Fax : 03-55442096
Unit Kewangan Zon 17 : 03-55443430



LETTER OF REPORT SUBMISSION

Date : 14 December 2006
Project File No : 600 - IRDC/ST 5/3/1012

Penolong Naib Cancellor (Penyelidikan)
Institut Penyelidikan, Pembangunan dan Pengkomersilan
Universiti Teknologi MARA
40450 Shah Alam

Ybhg. Prof.,

FINAL REPORT ON THE RESEARCH ON “EVALUATING THE PERFORMANCES OF THE DICE, JACCARD, OVERLAP, AND COSINE COEFFICIENTS IN AUTOMATED MARKING OF ALGEBRAIC EQUATIONS”

With reference to the above matter, attach hereby 3 (three) copies of the Final Report on the Research on “ Evaluating the Performances of the Dice, Jaccard, Overlap and Cosine Coefficients in Automated Marking of Algebraic Equations”

Thank You

Your sincerely



Hasnah Yusof
Project Leader

TABLE OF CONTENTS

	Page
LETTER OF APPOINTMENT	ii
LETTER OF REPORT SUBMISSION	iii
PROJECT TEAM MEMBERS	iv
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vi
LIST OF TABLES	viii
LIST OF FIGURES	ix
ABSTRACT	x
CHAPTER I INTRODUCTION	
1.1 Background	1
1.2 Problem Statement	2
1.3 Objective	2
1.4 Significance of Project	2
1.5 Scope of Project	3
1.6 Summary	4
CHAPTER II LITERATURE REVIEW	
2.1 Introduction	5
2.2 Automated Marking of Mathematics	5
2.3 String Similarity and Tokenized Terms	6
2.4 String Similarity Measures	8
2.5 Dice Coefficient	8
2.6 Jaccard Coefficient	9
2.7 Overlap Coefficient	9
2.8 Cosine Coefficient	10
2.9 Manual marking	10
2.10 Summary	11
CHAPTER III METHODOLOGY	
3.1 Introduction	12
3.2 Data Acquisition	13
3.3 Adjustment of Program	15
3.4 Implementation Process	16

ABSTRACT

Computers enable the automation of many tasks and made great contributions in many areas where learning process takes place. Assessment plays an important role in the learning process of many courses and mathematics is no exception. However, the marking activity takes up valuable time and energy, thus automated marking could be a solution. In this research, four similarity measures Dice, Jaccard, Overlap and Cosine (DJOC) coefficients have been employed in the automated marking of four simple algebraic equations in open/ subjective form. The performances of DJOC are compared to the manual marking which is the benchmark. Data was obtained from the solutions to 4 simple algebraic equations given by 180 Universiti Teknologi MARA students of different programmes within the Shah Alam campus. The algorithm used in the automated marking was adjusted to fit the DJOC coefficients. Results for the case of correct worked out solutions and the case of incorrect worked out solutions suggest some modifications needed to be done to the existing algorithm before the case of partially correct worked-out solutions are explored. The automated marking of simple algebraic equations using DJOC coefficients is feasible and further research should be done to obtain threshold values for each of the measure.