Sensing Textual Plagiarism

by

Nurul Jamilah Binti Abdul Samad

Dissertation submitted in partial fulfillment of

the requirements for the

Bachelor of Technology (Hons)

Business Information System

Jan 2006

Universiti Teknologi PETRONAS

Bandar Seri Iskandar

31750 Tronoh

Perak Darul Ridzuan

1) Parallel programming (computer science)
2) computer algorithms
3) IT/IS -- thesis

CERTIFICATION OF APPROVAL

Sensing Textual Plagiarism

by

Nurul Jamilah Binti Abdul Samad

A project dissertation submitted to the

Business Information System Program

Universtiti Teknologi PETRONAS

in partial fulfillment of the requirement for the

BACHELOR OF TECHNOLOGY (Hons)

BUSINESS INFORMATION SYSTEM

Approved by,

(Izzatdin Bin Abdul Aziz)

UNIVERSITI TEKNOLOGI PETRONAS

TRONOH, PERAK

JANUARY 2006

ABSTRACT

This Final Year Project (FYP) is about Sensing Textual Plagiarism. To realize this, an application that is equipped with the capability of detecting plagiarism from occurring in a textual document is to be developed. The main focus of this project is to perform a study on how to detect plagiarism from a textual document. Word-for-word plagiarism is the most obvious and serious form of plagiarism which can be can be categorized as a form of direct stealing, without significant alteration and consent of another's work. Fact findings are carried out in order to perform the study on plagiarism. This project will incorporate the Smith-Waterman Algorithm which is a classical tool in the identification and quantification of local similarities in biological sequences. As a result, the significance of this project is the availability of the application to sense the wide spread of plagiarism that often occur upon valuable documents, articles, and journals.

ACKNOWLEDGEMENT

First and foremost, I would like to express my gratitude to Allah S.W.T. because of His mercy and blessings had given me the strengths to face challenges in completing this Final Year Project.

I would like to express thousands of appreciations, highest gratitude and sincere thanks to my supervisor, Mr. Izzatdin Bin Abdul Aziz for all the valuable guidance, positive and constructive critics and advice that have been given to me upon completing this project.

I would also like to express my gratitude and thanks to all lecturers and tutors in Information Communication Technology (ICT) and Business Information System (BIS) department who eventually helped me during the project and also in sharing their knowledge and information, which has made this project an unforgettable work. Not to forget, special thanks to all my friends who helped and share their knowledge with me during the development of this project.

Lastly, I acknowledge with greatest appreciation to other personnel not mentioned above whom gave me such great support in completing this project successfully and to UTP for giving me a chance to gain knowledge and experiences during the Final Year Project development. A sincere apology from me for all the problems involuntarily caused. All of your kindness and cooperation are highly appreciated and will be fondly remembered.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATION

ICT	-	Information System Information
BIS	-	Business Information System
FYP	-	Final Year Project
SIRIM	-	Standard and Industrial Research Institute of Malaysia
PDF	-	Portable Document Format

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND OF STUDY

Plagiarism can be elucidated as the use of another's information, language, or writing, when done without proper acknowledgment of the original source. Plagiarism of written text has been widely spreading in the era of Information Technology and aggravated by the internet. Textual Plagiarism can be categorized into several forms:

- Copying directly from the source,

  to take words or sentences verbatim from the original source (with or without footnotes).
- Rewording a sentence (paraphrasing),

  an original sentence is rewritten in a copier's own words, but still no use of quotation marks or referencing is used.
- Submitting someone else's work,

  an obvious example of plagiarism.
- Failing to reference/footnote source material,

  as new facts are presented to people not familiar with the field, a footnote should be presented to reference the source material.
- The Internet "pastiche",

  the copying of a collection of paragraphs from a variety of electronic sources and pasted together in a word processor to form a report.

[Paul Clough, July 2000]

As the project is concern, plagiarism will be detected and intellectual property will be protected as a means of identifying how information has been misused or stolen. The fundamental idea of the software to be developed is that it will be able to verify keyword uses and keyword frequencies in electronic documents and presents a percentage of matches between compared words or paragraph. These findings will be categorized as keyword profiles and amassed in database. The higher the percentage of match will identify whether textual plagiarism exists in the documents or not.

1.2 PROBLEM STATEMENT

1.2.1 Problem Identification

It is essential to protect valuable document such as copyright documents, articles, and journals from plagiarism due to its intellectual property. However, bootlegging still occur for the reason that the availability of internet allowing ease of accessing documents, articles, and journals. With advanced word processors it is much easier to cut-and-paste large amounts of text to create a single work from a number of electronic sources including the Internet, electronic journals, books, newspapers, magazines and etc. This invasion of intellectual property will result in losing rights and authority. Therefore, it is vital to prevent plagiarism from wide spreading.

1.2.2 Problem Significance

With the ability of a system that is capable of sensing textual plagiarism, this will ensure the acknowledgement and recognition of other people's work, thus protect the integrity of a precious document. Plagiarism enforcers such as Standard and Industrial Research Institute of Malaysia (SIRIM), academicians and industry would have the ease to detect and inflict textual plagiarism in industry and academic area from wide spreading.

1.3 OBJECTIVES AND SCOPE OF STUDY

1.3.1 Objectives

- To perform a study on how to detect plagiarism from a textual document.
    - Fact finding is carried out to investigate related matters pertaining textual plagiarism.

- To develop an application that is able to detect plagiarism occurring in textual document.
    - The system is capable in detecting near to accurate textual plagiarism on documents.
    - The system should be able to indicate the other work which has the solely same lines or paragraphs of text, within the same database.

1.3.2 Scope of Study

The scope of this project would be as follows:

- The study would be on how textual plagiarism is detected.
- Performing a plagiarism detection system for textual document.
- The focus of this system will be lingering around the academic environment.
- The system will cater only .txt file formats.
- The system will cater only word-for-word plagiarism.

1.4 PROJECT TIMELINE

This Final Year Project is divided into two parts, Part 1 and Part 2 whereby the first part discusses matters pertaining studies and reports and the latter is about system development and refinement.

Starting with Part 1, project kick-off is on the 1st August 2005; however the confirmed topic is on the 12th September 2005. During this period of time various topics are surveyed and considered. Initials study started on 26th September 2005 by carrying out observations and fact-finding. First formal report submitted on the 7th November 2005 followed by submission of the Interim Report on 21st November 2005. This report summarizes the findings as well as to justify statements. Presentation for Part 1 was viewed on the 5th December 2005. The timeline for Part 1 can be referred to Figure 1 in the Appendices.

As referred to Figure 2 for the timeline of Part 2, system development started on the 30th January 2006 and took about four months to be completed. Prototype of the system was delivered during the Pre-Engineering Design Exhibition (Pre-EDX) on 4th April 2006. The final draft is to be submitted on 16th June 2006 and final presentation with the internal and external examiners is forecasted to be on 19th June 2006.

CHAPTER: 2
LITERATURE REVIEW AND THEORY

Plagiarism in the other words can be copying, bootlegging, stealing, or breach of copyright. The more details of definition for plagiarism could be as below:

"When the work of someone else is reproduced without acknowledging the source, this is known as **plagiarism**. Probably the most frequent cases appear in academic institutions where students copy material from books, journals, the Internet, their peers etc. without citing references. Although sometimes intentional, there are many cases where students actually plagiarize unintentionally simply because they are not aware of how sources should be used within their own work. This problem is not just limited to written text, but also regularly found in software code where chunks are copied and re-used without reference to the original author(s)." [Clough - July 2000]

We can see that plagiarism has it certain significance depending on its scope, strategic location, and the situation in which it occurs. A sentence of paragraph which is plagiarized is not as unpleasant as copying a paper word-by-word. Besides, we need to mull over the overall perspective of plagiarism before we penalize it. Regarding to this situation, a plagiarism detecting tool or system is crucial to be developed.

“The significance of plagiarism can vary widely, depending on its extent, strategic location, and the context in which it occurs. An isolated instance of plagiarism - one sentence or paragraph, for example - would not usually be cause for concern, whereas a paper copied almost *verbatim* would be considered a gross violation of academic norms. Strategic location refers to centrality in an academic presentation. Plagiarism in crucial points of argumentation is more serious than in a largely extraneous literature review. Finally, the overall context of plagiarism must be considered: the nature of the contribution, scholarly or otherwise.” [Martin, October 1984]

“Most students are now expected to submit written work and program assignments in electronic form. Although convenient and easier for both student and lecturer alike, the electronic version provides the student with an easier opportunity to plagiarize. With advanced word processors it is much easier to cut-and-paste large amounts of text to create a single work from a number of electronic sources including the Internet, electronic journals, books, newspapers and magazines etc.” [Clough - July 2000]

This project is concern with detecting text plagiarism within the academic area. The use of electronic form documents among the students such as using Microsoft Word and Notepad in completing assignments makes it easy for textual plagiarism to be occurred in the academic field.

“From a legal point-of-view, proving plagiarism can be very hard. For a start, copyright can only be enforced if the plaintiff can prove that words were copied or trivially transformed (i.e. paraphrased). Even matching verbatim text between two sources does not prove plagiarism. If the two texts are written about the same topic, then it should not be a surprise that some information will be shared. For example names of people, places, technical terms or structure words of the language (i.e. English word classes such as prepositions, conjunctions etc.). In [Susan Finlay, *CopyCatch*, Masters Dissertation, University of Birmingham, 1999], Susan Finlay reports that in her work, independent texts have as much as 50% or more shared vocabulary overlap. With paraphrasing this is even harder because it must be proved the two suspected areas of plagiarism mean the same.” [Clough - July 2000]

It is true that proving plagiarism can be very hard. However, before proving that a document or text is plagiarized, we have to firstly sense the plagiarism. That is the crucial point or purpose of presenting this project, to sense or detect text plagiarism, and at the same time, to protect the rights of intellectual property.

The methodology used in constructing the textual plagiarism detecting system is the Smith-Waterman Algorithm. This algorithm is specifically applied to any form of textual material.

“The method that we propose can be applied to any form of textual material, such as essays, reports and etc. Unlike many existing techniques for collusion detection, it does not depend on statistical properties, such as counts of particular words, but rather on structural similarities between (parts of) texts.” [Irving, 2000]

“The only difference in the approach to this special case is the way in which the source material is parsed. Ordinary textual material will be parsed as a sequence of words, where the term word is given an appropriate precise meaning.” [Irving, 2000]

From the quoted text above, it describes that using this Smith-Waterman Algorithm; the source material will be compared and parsed as a sequence of words or string to identify the significant matches between two compared text materials.

CHAPTER 3
METHODOLOGY/PROJECT WORK

3.1 METHODOLOGY

3.1.1 Formal System Development

The formal systems development model, shown below in Figure 1, utilises a development process that is based on formal mathematical transformation of system models to executable programs. This system development model is similar to the waterfall model; the formal approach has clearly defined (cascading) phase boundaries. The critical distinctions between the two models are:

- The software requirements and specification phases are refined into a detailed formal specification, which is expressed mathematically.
- The design, implementation and unit testing are replaced by a single formal transformation phase.

Detail description of each phase is described on the following page:

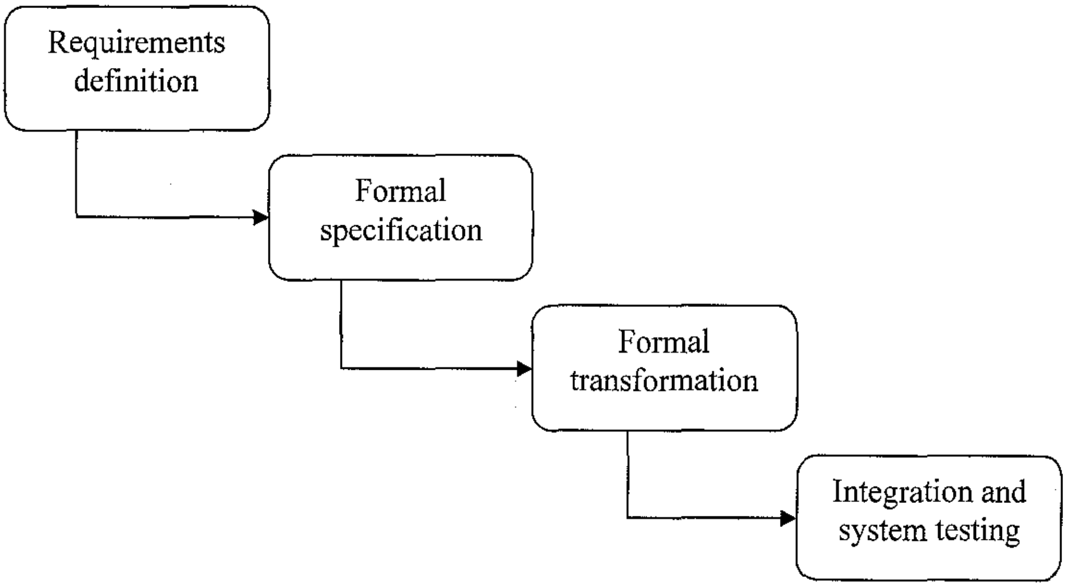


Figure 3: Formal systems development

- Requirements definition

  In this phase, the system's services, constraints and goals are established with system users. As a whole, the requirement from the system is that it should be able to detect plagiarism from textual documents.

- Formal specification

  Formal specification means the more detailed requirements are established from the system. This system involves with the system users, who are the academicians; accessing to the application and opening text files or documents from the specific directory in the computer. After that the users will run the application in order to sense textual plagiarism between two documents by comparing them. Finally, the system will produce a report as the result of the textual plagiarism detection.

- Formal transformation

  During the formal transformation process the mathematical representation of the specifications is systematically refined. As for this project, there are several functions that need to be developed before it can be integrated as a single system. Some of the functions are including the application to detect textual plagiarism itself, and also the report to view results of plagiarism status between two compared text files.

- Integration and system testing

  Finally in this integration and system testing, the individual program units or programs are integrated and tested as a single complete system to ensure that the software requirements have been met. In this project, the application for textual plagiarism detection and the report will be tested and will be ensure that it will be functioning appropriately. Subsequently, both programs will be integrated and tested together as a complete system.

3.1.2 Smith-Waterman Algorithm Methodology

Smith-Waterman algorithm is used in this project as a variant to locate similarities in textual document, with a view to the application in the detection of plagiarism. The Smith-Waterman algorithm is a classical method of comparing two strings with a view to identify highly similar sections within them. It is widely-used in finding good "near-matches", or so-called local alignments. The basis of the method is a dynamic programming scheme.

Basically, the lengths of the given strings X and Y are indicated by m and n respectively. A portion X' of string X aligned with a portion Y' of string Y is allocated a score that, in some sense, represents the "goodness of fit" between X' and Y'. Each matching symbol should make a positive contribution to that score, and each symbol that has to be inserted, deleted or substituted to transform X' to Y' should make a negative contribution.

Let *h* be the (positive) contribution made by a symbol `hit', *d* the (negative) contribution made by a symbol insertion or deletion (an `indel'), and *r* the (negative) contribution made by replacing one symbol by another. The methodology in this context describes only a simple model compare to the model used in computational biology. Even for this simpler model, it is not immediately clear what the relative values of h, d and r should be; the most obvious option is to choose h = d = r = 1, and these values have been shown to work effectively in practice. In much of the following discussion and example, assume that h = d = r = 1.

For example, if X' = abcbadbca and Y' = abbdbda, an optimal alignment has 6 hits, 2 indels, and 1 replacement, as shown in Table 1, a score of 6h-2d- r, or 6-2-1 = 3 in the case where h = d = r = 1, is obtained.

a	b	c	b	a	d	b	c	a
a	b	-	b	-	d	b	d	a

Table 1: An optimal alignment of two substrings

For the implementation to this detection of textual plagiarism application, I have defined a more suitable approach to apply the algorithm by using words or strings instead of using characters or substrings to find the significant "near-matches" as mentioned above. This approach is seemed to be more obvious and clear in order to define the percentage of plagiarism in any textual documents.

She			blows	the	bubbles	at	the	party
He	loves	to	blow	the	bubbles	at	the	park

Table2: An optimal alignment between two strings or words.

For example, if X'= "She blows the bubbles" and Y'= "He loves to blow the bubbles", an optimal alignment has 4 hits, 2 indels and 3 replacement, as shown in Table 2. A score obtained for the figure above is, 4h-2d-3r, or 4-2-3= -1, where h=d=r=1.

When a negative score is obtained after the algorithm is applied, it means that there is no possibility of plagiarism to occur between the two compared texts. In the other words, the documents are not plagiarized each other. However, when a zero or positive score is obtained, it means the other way which there is a possibility that the documents are plagiarized. The higher the score is, the higher the percentage of plagiarism is.

Example	of	an	optimal	alignment	between	two	strings	or	sentences
Example	of	an	optimal	alignment	between	two	substrings	or	characters

Table 3: Another optimal alignment between two strings or words.

An example for the positive score is shown as in Table 3 above. When X'= "Example of an optimal alignment between two strings or sentences" and Y'= "Example of an alignment between two substrings or characters", an optimal alignment has 8 hits, 0 indel, and 2 replacement is obtained with the score is 8h-2r or 8-6= 4, where h=d=r=1.

The objective is to find significant near-matches between substrings of X and Y, where `significant' is defined in terms of some suitably chosen threshold score. Threshold score should be depends on the context, and on the chosen values of h, d and r. The cumulative score of significant near-matches might be taken as an appropriate measure of overall similarity of the two strings in the application context, though more generally, any two strings containing at least one significant near-match might be considered as worthy of further investigation.

3.2 PROJECT WORK

This Final Year Project requires some specifications in terms of hardware and software in order to realize the project work. The project requirements are:

3.2.1 Hardware

- Intel or AMD Processor PC (1.27 GHz or above)
- 256MB RAM (or above)
- 40 Gig Hard Disk (or above)

3.2.2 Software

The software requirements for this project have been redefined from using Java Tools to Microsoft Visual Basic 6. The reason is because of the difficulty in using Java Tools compared to Microsoft Visual Basic 6 which is more comfortable to me.

- Microsoft Visual Basic 6
- Microsoft Office Access (database)
- and others as project progress

CHAPTER 4
RESULT AND DISCUSSION

4.1 RESULTS

As a result, the first objective which is to perform a study on how to detect plagiarism from a textual document has been met. Fact finding has been carried out to investigate related matters pertaining textual plagiarism. For the second objective, it is also completed with the development of application and coding of the system.

4.1.1 System Overview

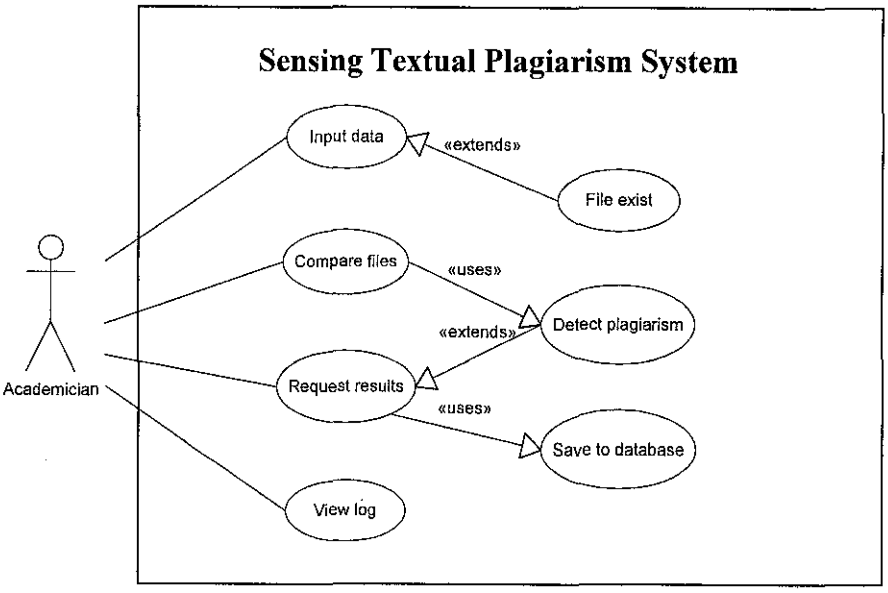


Figure 4: Use Case Diagram for overall system flow

The flowchart of the process flow for data entry, data processing and reporting are illustrated as below:

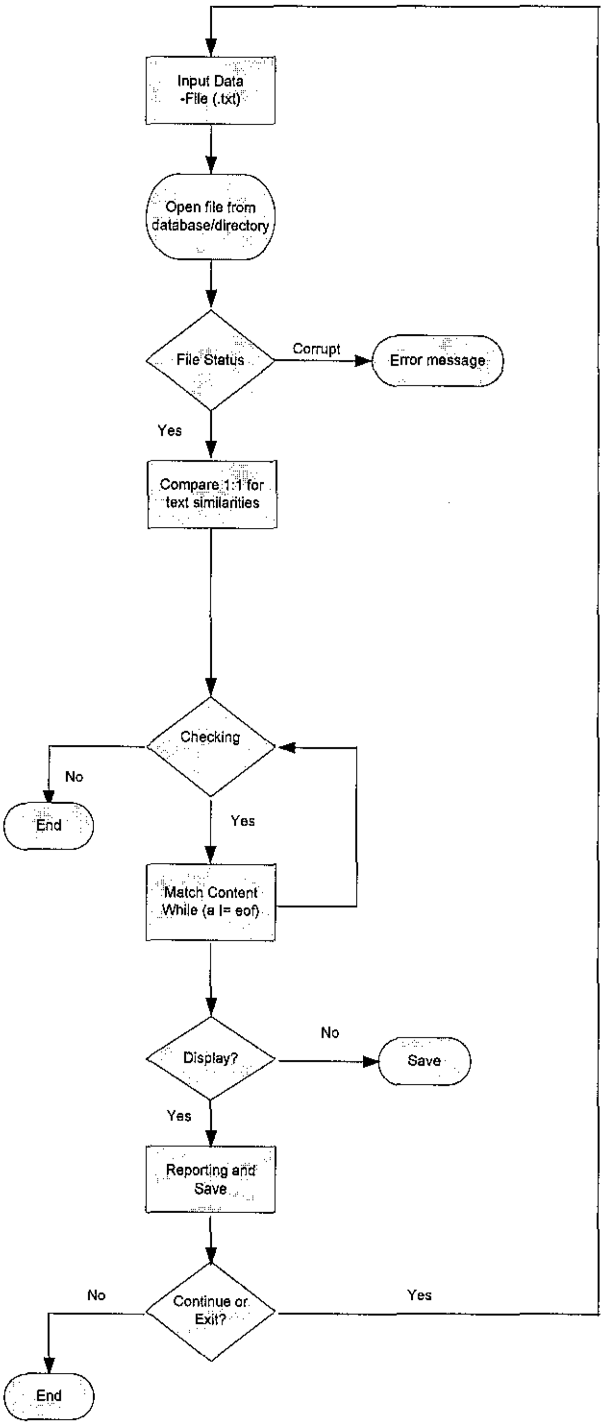


Figure 5: Data Entry, Data Processing and Reporting Process Flow

As illustrated in Figure 5, the flow graph indicates the system flow for data entry, data processing and reporting. First of all the user will input the text data and open text files from the directory or database. The system will display two text files to make it comparable for text similarities. Results of the comparison between two text files will be displayed as the compare result. Status of the result will also be displayed whether those files are suspected as plagiarism or not.

The benchmark for the result would be less than 50% for 'Not suspected as plagiarism' status, and more than 50% for 'Suspected as plagiarism' status. User will be able to view the report of the compared files and save it to the database. Finally, user may choose to continue using the system to compare another text files or exit from the program. A similar representation of system flow is indicated t in terms of nodes presentation with multiple out-going flows on the flow graph (Refer to Figure 6 in Appendices).

4.1.2 Survey Result

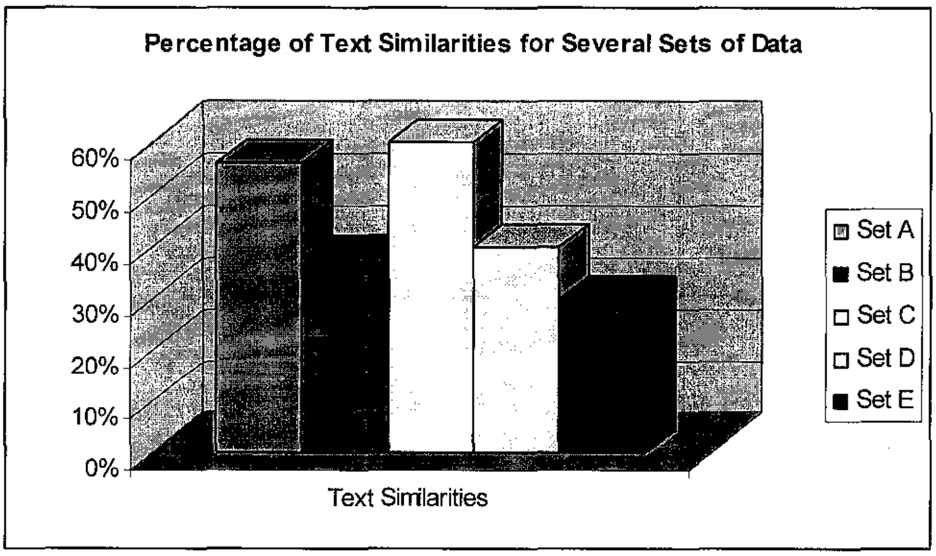


Figure 7: Percentage of Text Similarities for Several Sets of Data

As refer to Figure 7, there has been a survey on the system with five different sets of data. The percentage of text similarities vary with the different sets because of the different text and different total of words in each file. For Set A, the comparison result between two related text files is 56%, which is over the benchmark. The status of the comparison for files in Set A would be 'Suspected as textual plagiarism'. Same with Set C, the result is 60% and also over the benchmark. Whereas for Set B, Set D and Set E, three of them carry the same status which is 'Not suspected as textual plagiarism' since their compare results are 38%, 40% and 30% respectively.

4.2 SYSTEM USER INTERFACE AND FUNCTIONS

4.2.1 Description on System User Interface

User interface for the Textual Plagiarism Detection System has been designed and finalized. A splash screen is created as a welcome screen for using the system (Refer to Figure 8 in Appendices). By pressing any key, the next screen will be appeared which the main window for the system as is showed in Figure 9 in Appendices. From this window, the user will browse for a master text file and also target text file to enable the system to compare the text similarities between them.

After the comparison, the system will display the compare result between the two text files, together with the status of plagiarism. Then, the user may click the 'View Report' button to call for the next window which will display the full report of the compared text files. In the Report window, it will consist of the name of the master and target files which have been compared with each other, together with the compare result and status of plagiarism (Refer to Figure 14 in Appendices). From here, the user may save the report and quit from the system.

4.2.2 System Functions

- Browse Button,

  to browse for Master and Target text files.
- Compare Button,

  to start comparing the Master and Target text.
- Clear All Button,

  to clear the filled field.
- View Report Button,

  to view the full report of compared files.
- Save Button,

  to save the report into the database.
- Quit Button,

  to exit from the system.

CHAPTER 5
CONCLUSION AND RECOMMENDATIONS

The Final Year Project's concern is to accomplish the previous mentioned objectives. In performing the first objectives, fact-finding is carried out to investigate related matters pertaining textual plagiarism. Literature reviews have been refined in order to get more understanding by reviewing previous works pertaining detection of textual plagiarism. In relating to the above, first objective has been accomplished.

The second objective also has been achieved through the development of application and coding of the system. Survey has been done in testing the system to get the textual plagiarism status for several sets of data. The system has allocated a benchmark for determining the status of textual plagiarism. If the result of the text comparison is above 50% then the text files are suspected to be plagiarized. While if the result of text comparison is below 50% then the text files are not suspected to be plagiarized.

Some recommendations have been identified as a future enhancement for this project. The developed textual plagiarism detecting system could be enhanced by:

a) online accessibility
    - make it accessible through web.

b) the ability to handle massive text with better performance, using high performance computer
    - able to handle larger size of text document with faster time and more precise results.

c) the ability to handle multiple file formats (PDF, etc.)

   - able to handle numerous format of files (other than .doc and .text files)

d) the expansion to not only word-for-word plagiarism but also for graphics and images.

   - do a project expansion that is able to detect not only text plagiarism, but as well as for graphics and images.

e) add some security features to the system

   - make a log-in features for different levels of users (lecturers, students, etc).

f) develop the functions to look into the semantic (meaning) of the words to detect textual plagiarism.

REFERENCES

[1] Brian Martin, 1984. *Plagiarism and Responsibility Journal of Tertiary Educational Administration, volume 6, number 2, October 1984, pp. 83-90.*

[2] Christian Collberg, Stephen Kobourov, Joshu, Louie, and Thomas Slattery, 1985. *SPLAT: A Sytem for Self-Plagiarism Detection. Department of Computer Science*, University of Arizona.

[3] M.S. Waterman and M. Eggert, 1987. *A new algorithm for best subsequence alignments with application to tRNA-tRNA comparisons.* Journal of Molecular Biology, 197:723 -728.

[4] Robert W. Irving,1990. *Plagiarism and Collusion Detection using the Smith-Waterman Algorithm.* Department of Computing Science, University of Glasgow, Glasgow G12 8QQ, UK.

[5] X. Huang, R.C. Hardison, and W. Miller, 1990. *A space-effcient algorithm for local similarities.* CABIOS, 6:373 -381.

[6] X. Huang and W. Miller, 1991. *A time-effcient linear-space local similarity al-gorithm.* Advances in Applied Mathematics, 12:337 -357.

[7] Jeff Prosise, 1996. *The Boyer-Moore Algorithm.*

[8] D. Gusfield,1997 *Algorithms on strings, trees, and sequences. Computer Science and Computational Biology.* Cambridge University Press.

[9] Lisa Hincliffe, 1998. *Cut-and-Paste Plagiarism: Preventing, Detecting and Tracking Online Plagiarism.*

[10] Paul Clough, July 2000. *Plagiarism in natural and programming languages: an overview of current tools and technologies.* Department of Computer Science, University of Sheffield.

[11] JJ Kuhl, 2002. *Project Lifecycle Models: How They Differ and When to Use Them.*

[12] T. Lancaster, 2003. *Effective and Effcient Plagiarism Detection. PhD thesis, School of Computing*, Information Systems and Mathematics, South Bank University.

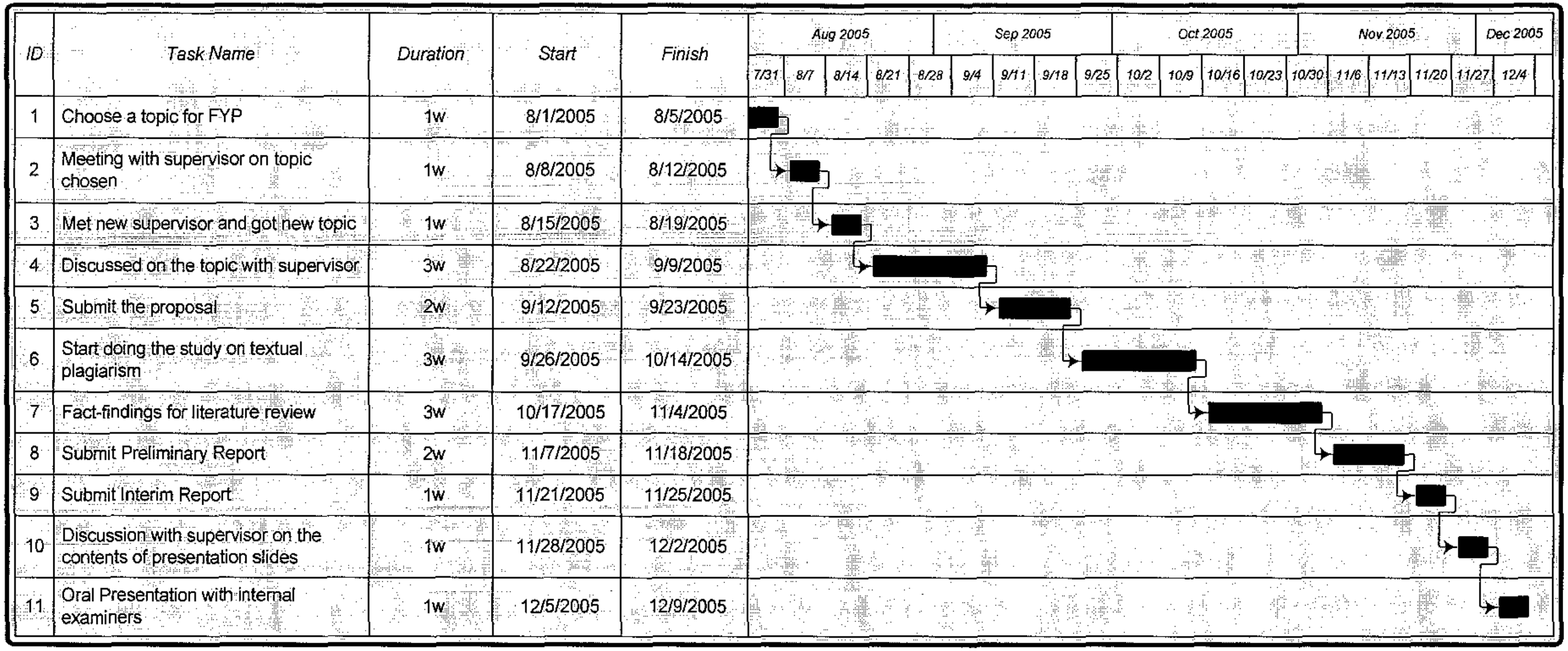
[13] Adrian Als & Paul Walcott, 2004. *Formal System Development.*

[14] Juha-Peleka Tolvanen, 2004. *Domain-Specific Modeling for Full Code Generation.*

[15] Martining & Associates, 2004. *Methods and Tools.*

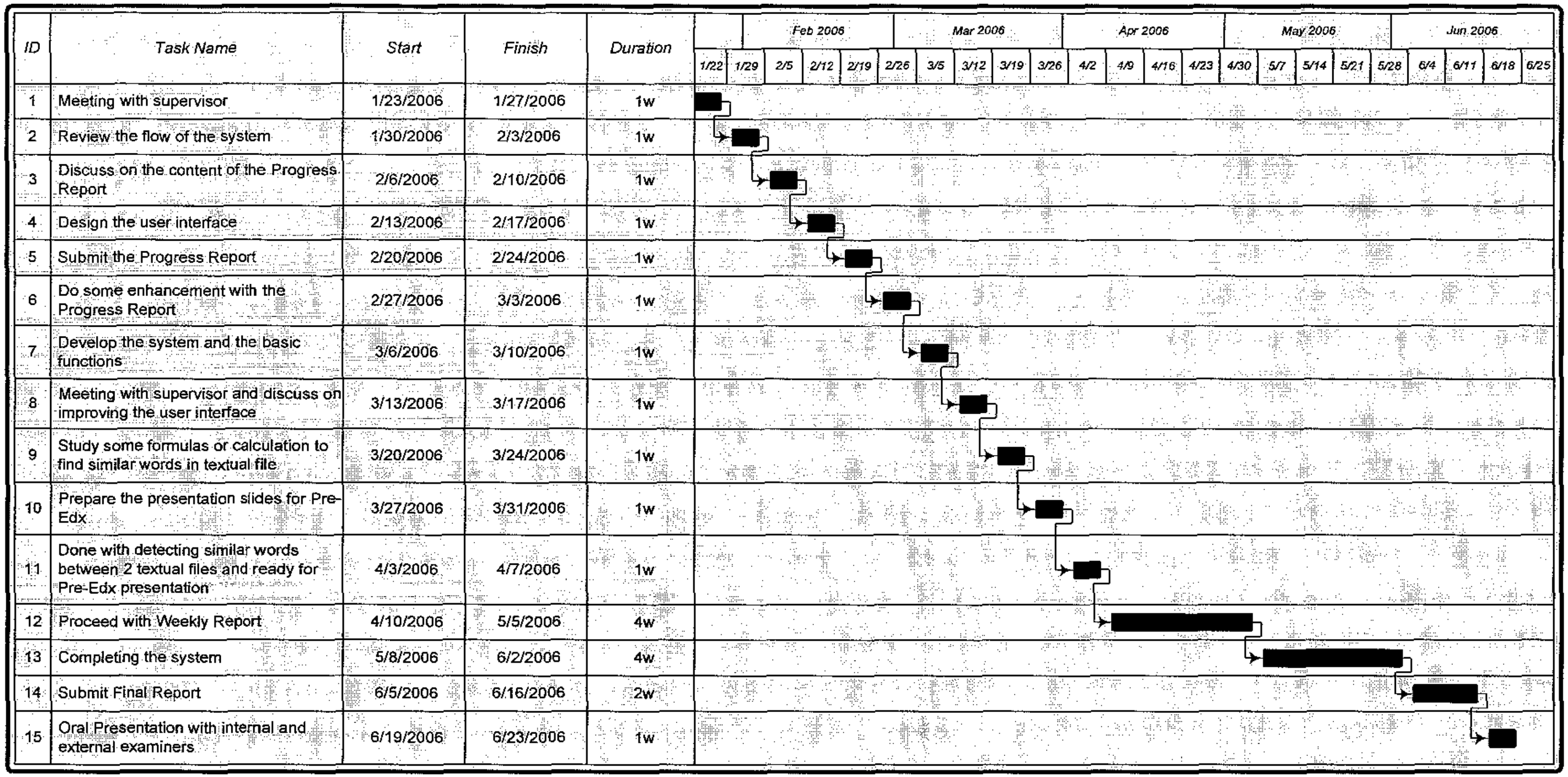
[16] Mike Cohn, 2004. *Estimating with Use Case Points.*

APPENDICES



ID	Task Name	Duration	Start	Finish
1	Choose a topic for FYP	1w	8/1/2005	8/5/2005
2	Meeting with supervisor on topic chosen	1w	8/8/2005	8/12/2005
3	Met new supervisor and got new topic	1w	8/15/2005	8/19/2005
4	Discussed on the topic with supervisor	3w	8/22/2005	9/9/2005
5	Submit the proposal	2w	9/12/2005	9/23/2005
6	Start doing the study on textual plagiarism	3w	9/26/2005	10/14/2005
7	Fact-findings for literature review	3w	10/17/2005	11/4/2005
8	Submit Preliminary Report	2w	11/7/2005	11/18/2005
9	Submit Interim Report	1w	11/21/2005	11/25/2005
10	Discussion with supervisor on the contents of presentation slides	1w	11/28/2005	12/2/2005
11	Oral Presentation with internal examiners	1w	12/5/2005	12/9/2005

Figure 1: Timeline for FYP Part 1



ID	Task Name	Start	Finish	Duration
1	Meeting with supervisor	1/23/2006	1/27/2006	1w
2	Review the flow of the system	1/30/2006	2/3/2006	1w
3	Discuss on the content of the Progress Report	2/6/2006	2/10/2006	1w
4	Design the user interface	2/13/2006	2/17/2006	1w
5	Submit the Progress Report	2/20/2006	2/24/2006	1w
6	Do some enhancement with the Progress Report	2/27/2006	3/3/2006	1w
7	Develop the system and the basic functions	3/6/2006	3/10/2006	1w
8	Meeting with supervisor and discuss on improving the user interface	3/13/2006	3/17/2006	1w
9	Study some formulas or calculation to find similar words in textual file	3/20/2006	3/24/2006	1w
10	Prepare the presentation slides for Pre-Edx	3/27/2006	3/31/2006	1w
11	Done with detecting similar words between 2 textual files and ready for Pre-Edx presentation	4/3/2006	4/7/2006	1w
12	Proceed with Weekly Report	4/10/2006	5/5/2006	4w
13	Completing the system	5/8/2006	6/2/2006	4w
14	Submit Final Report	6/5/2006	6/16/2006	2w
15	Oral Presentation with internal and external examiners	6/19/2006	6/23/2006	1w

Figure 2: Timeline for FYP Part 2

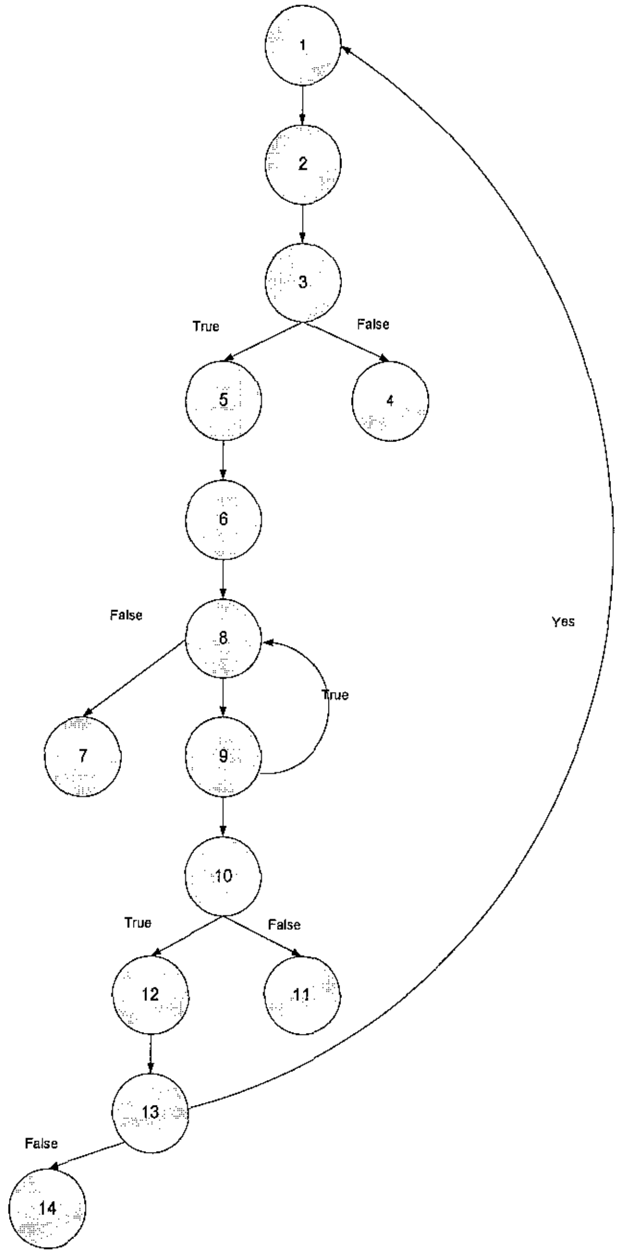


Figure 6: Data Entry, Data Processing and Reporting Flow Graph

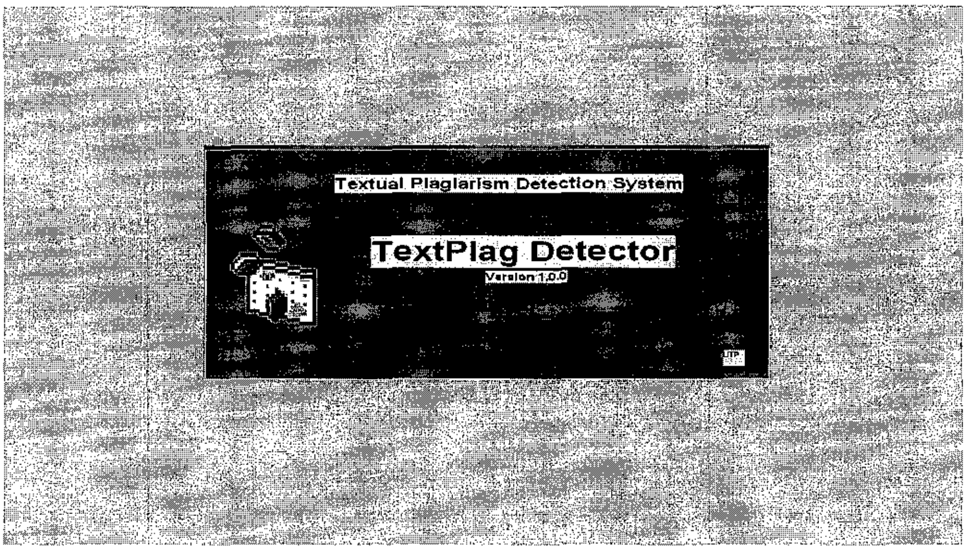


Figure 8: Splash Screen

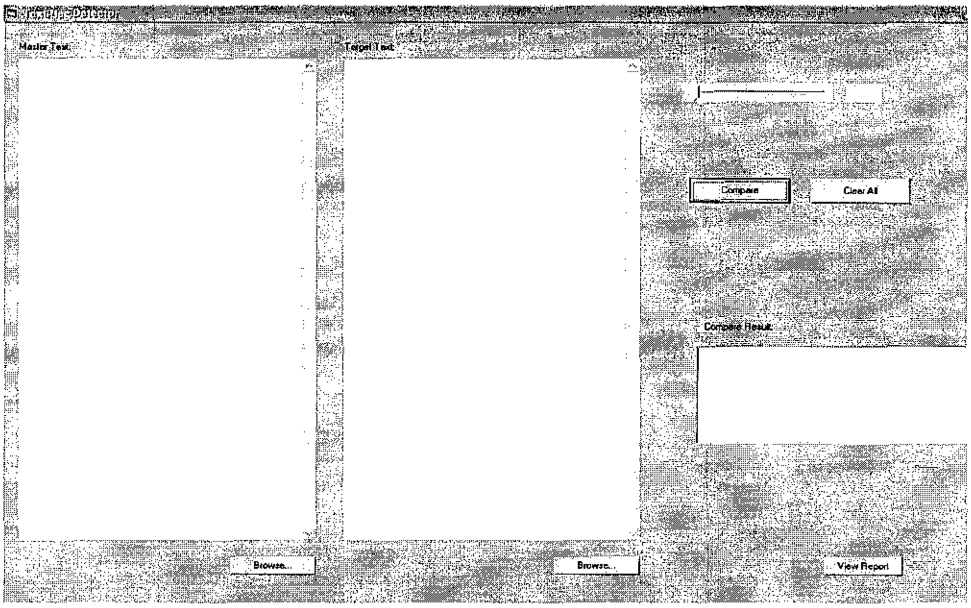
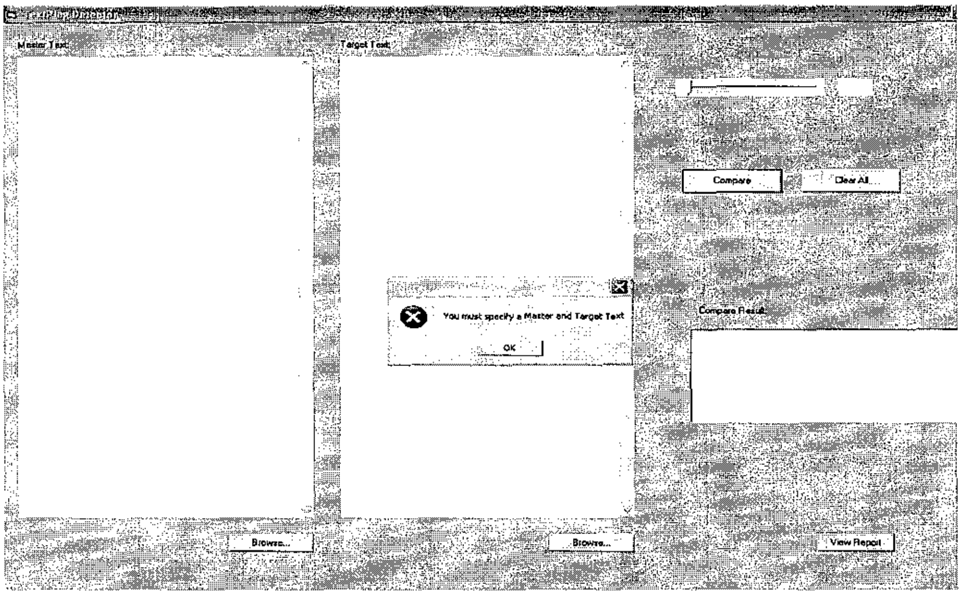
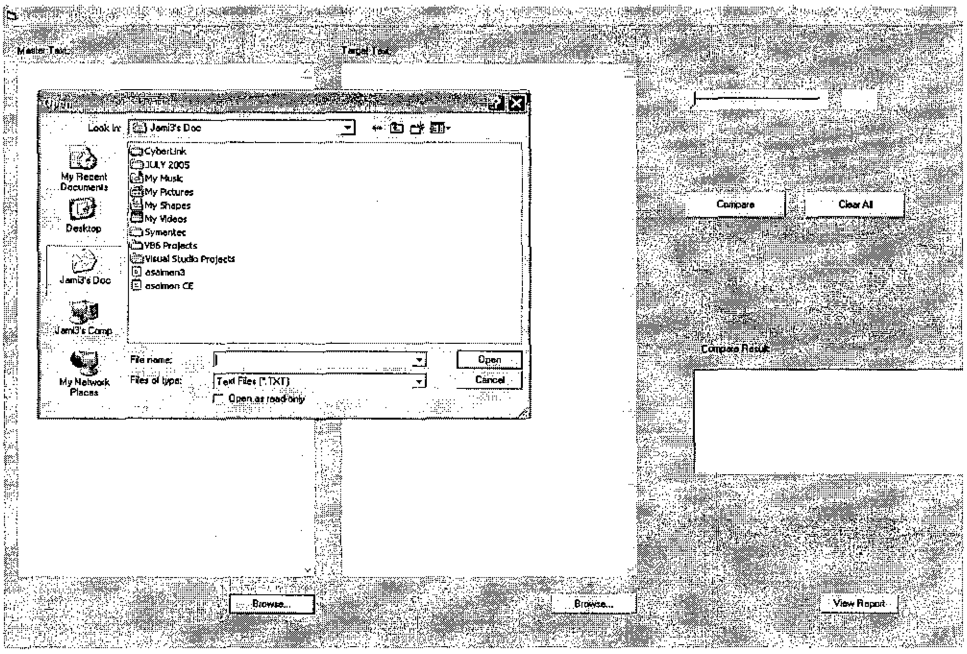


Figure 9: Word by Word Comparison Screen



You must specify a Master and Target Text

Figure 10: Specify Master and Target Text Dialog Box



Look In: Jami3's Doc

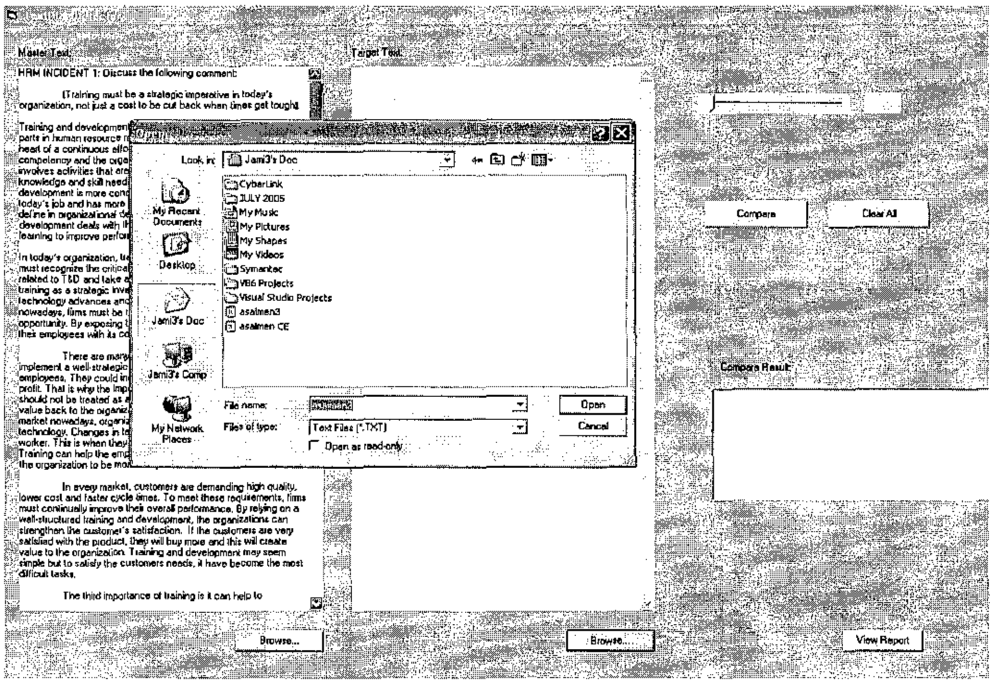
CyberLink
JULY 2005
My Music
My Pictures
My Shapes
My Videos
Symentec
VB6 Projects
Visual Studio Projects
asalmon3
asalmon CE

File name:

Files of type: Text Files (*.TXT)

□ Open as read-only

Figure 11: Browse Master Text Dialog Box



HRM INCIDENT 1: Discuss the following comment:

(Training must be a strategic imperative in today's organization, not just a cost to be cut back when times get tough)

Training and development
parts in human resource
heart of a continuous effo
competency and the orga
involves activities that are
knowledge and skill need
development is more conc
today's job and has more
define in organizational de
development deals with th
learning to improve perfor

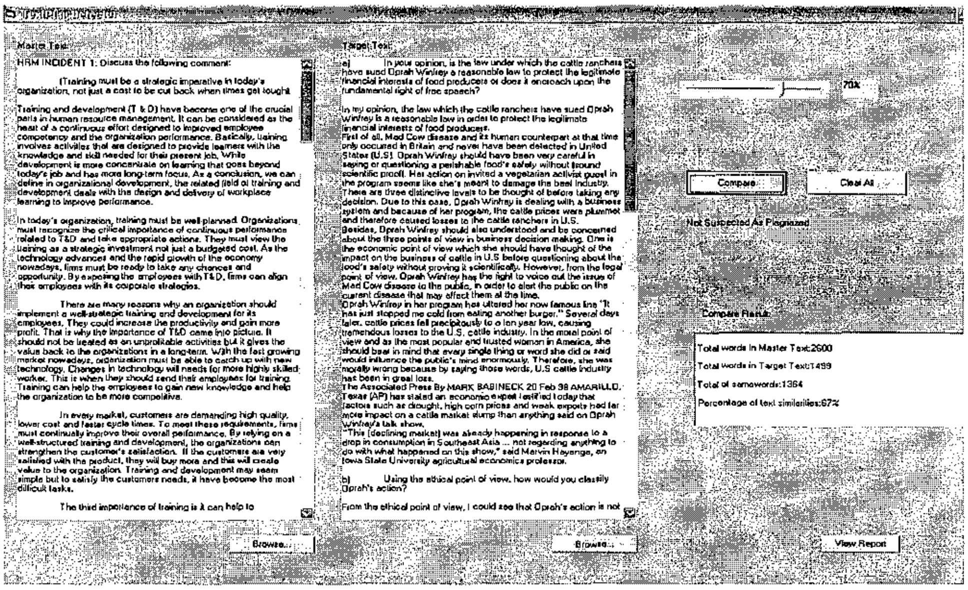
In today's organization, tr
must recognize the critical
related to T&D and take a
training as a strategic inve
technology advances and
nowadays, firms must be r
opportunity. By exposing t
their employees with its co

There are many
implement a well-strategic
employees. They could inc
profit. That is why the imp
should not be treated as a
value back to the organiz
market nowadays, organiz
technology. Changes in te
worker. This is when they
Training can help the emp
the organization to be mor

In every market, customers are demanding high quality, lower cost and faster cycle times. To meet these requirements, firms must continually improve their overall performance. By relying on a well-structured training and development, the organizations can strengthen the customer's satisfaction. If the customers are very satisfied with the product, they will buy more and this will create value to the organization. Training and development may seem simple but to satisfy the customers needs, it have become the most difficult tasks.

The third importance of training is it can help to

Figure 12: Browse Target Text Dialog Box



HRM INCIDENT 1: Discuss the following comment:

(Training must be a strategic imperative in today's organization, not just a cost to be cut back when times get tough)

Training and development (T & D) have become one of the crucial parts in human resource management. It can be considered as the heart of a continuous effort designed to improved employee competency and the organization performance. Basically, training involves activities that are designed to provide learners with the knowledge and skill needed for their present job. While development is more concentrate on learning that goes beyond today's job and has more long-term focus. As a conclusion, we can define in organizational development, the related field of training and development deals with the design and delivery of workplace learning to improve performance.

In today's organization, training must be well-planned. Organizations must recognize the critical importance of continuous performance related to T&D and take appropriate actions. They must view the training as a strategic investment not just a budgeted cost. As the technology advances and the rapid growth of the economy nowadays, firms must be ready to take any chances and opportunity. By exposing the employees with T&D, firms can align their employees with its corporate strategies.

There are many reasons why an organization should implement a well-strategic training and development for its employees. They could increase the productivity and gain more profit. That is why the importance of T&D came into picture. It should not be treated as an unprofitable activities but it gives the value back to the organizations in a long-term. With the fast growing market nowadays, organization must be able to catch up with new technology. Changes in technology will needs for more highly skilled worker. This is when they should send their employees for training. Training can help the employees to gain new knowledge and help the organization to be more competitive.

In every market, customers are demanding high quality, lower cost and faster cycle times. To meet these requirements, firms must continually improve their overall performance. By relying on a well-structured training and development, the organizations can strengthen the customer's satisfaction. If the customers are very satisfied with the product, they will buy more and this will create value to the organization. Training and development may seem simple but to satisfy the customers needs, it have become the most difficult tasks.

The third importance of training is it can help to

a) In your opinion, is the law under which the cattle ranchers have sued Oprah Winfrey a reasonable law to protect the legitimate financial interests of food producers or does it encroach upon the fundamental right of free speech?

In my opinion, the law which the cattle ranchers have sued Oprah Winfrey is a reasonable law in order to protect the legitimate financial interests of food producers.
First of all, Mad Cow disease and its human counterpart at that time only occured in Britain and never have been detected in United States (U.S). Oprah Winfrey should have been very careful in saying or questioning a perishable food's safety without (sound scientific proof). Her action on invited a vegetarian activist guest in the program seems like she's meant to damage the beef industry.
There are three distinctive levels to be thought of before taking any decision. Due to this case, Oprah Winfrey is dealing with a business system and because of her program, the cattle prices were plummet and therefore caused losses to the cattle ranchers in U.S.
Besides, Oprah Winfrey should also understood and be concerned about the three points of view in business decision making. One is the economic point of view which she should have thought of the impact on the business of cattle in U.S before questioning about the food's safety without proving it scientifically. However, from the legal point of view, Oprah Winfrey has the right to voice out the issue of Mad Cow disease to the public, in order to alert the public on the current disease that may affect them at the time.
Oprah Winfrey in her program has uttered her now famous line "It has just stopped me cold from eating another burger." Several days later, cattle prices fell precipitously to a ten year low, causing tremendous losses to the U.S. cattle industry. In the moral point of view and as the most popular and trusted woman in America, she should bear in mind that every single thing or word she did or said would influence the public's mind enormously. Therefore, she was morally wrong because by saying those words, U.S cattle industry has been in great loss.
The Associated Press By MARK BABINECK 20 Feb 98 AMARILLO, Texas (AP) has stated an economic expert testified today that factors such as drought, high corn prices and weak exports had far more impact on a cattle market slump than anything said on Oprah Winfrey's talk show.
"This (declining market) was already happening in response to a drop in consumption in Southeast Asia ... not regarding anything to do with what happened on this show," said Marvin Hayenga, an Iowa State University agricultural economics professor.

b) Using the ethical point of view, how would you classify Oprah's action?

From the ethical point of view, I could see that Oprah's action is not

Not Suspected As Plagiarized

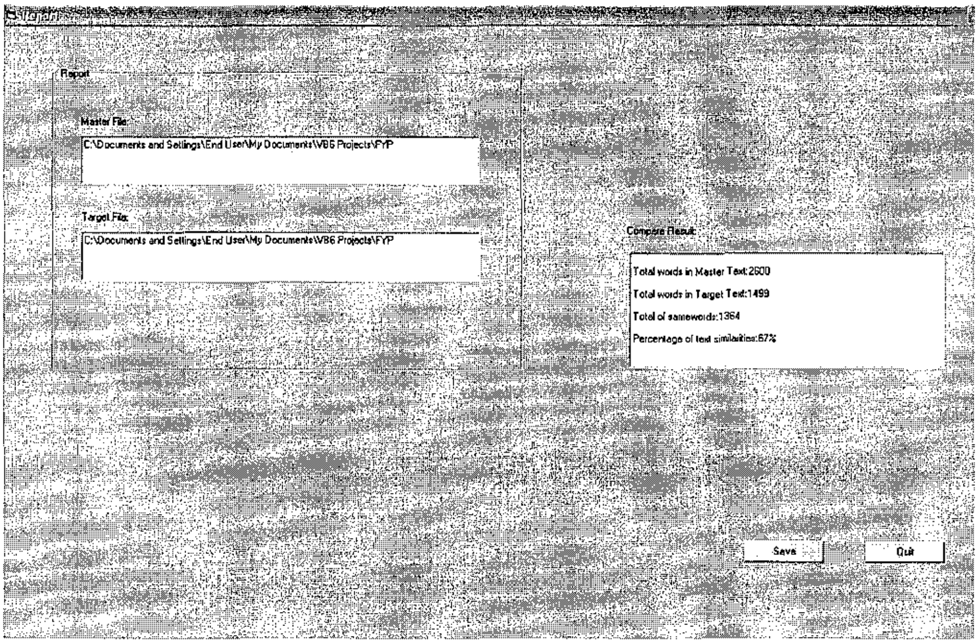
Total words in Master Text:2600

Total words in Target Text:1499

Total of samewords:1364

Percentage of text similarities:67%

Figure 13: Compare Result Screen



Master File: C:\Documents and Settings\End User\My Documents\VB6 Projects\FYP

Target File: C:\Documents and Settings\End User\My Documents\VB6 Projects\FYP

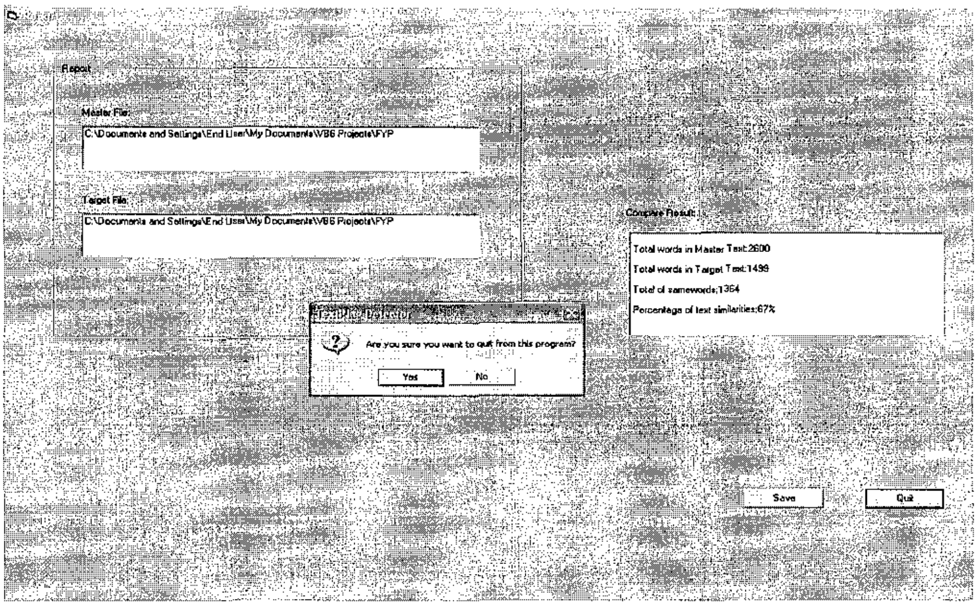
Total words in Master Text:2600

Total words in Target Text:1499

Total of samewords:1364

Percentage of text similarities:67%

Figure 14: View Report Screen



Master File: C:\Documents and Settings\End User\My Documents\VB6 Projects\FYP

Target File: C:\Documents and Settings\End User\My Documents\VB6 Projects\FYP

Total words in Master Text:2600

Total words in Target Text:1499

Total of samewords:1364

Percentage of text similarities:67%

Are you sure you want to quit from this program?

Figure 15: Quit Dialog Box