CANKAYA UNIVERSITY

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

MATHEMATICS AND COMPUTER SCIENCE

MASTER THESIS

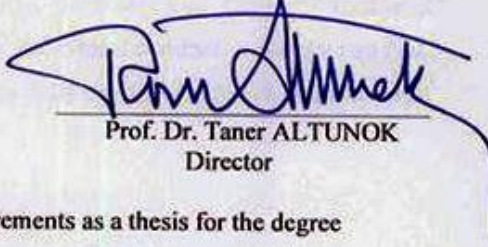**PLAGIARISM DETECTION IN LEARNING MANAGEMENT SYSTEM**

ARKAN KH. SHAKR

SEPTEMBER 2013

Title of the Thesis: **Plagiarism Detection in Learning Management System**

Submitted by: Arkan Kh. SHAKR

Approval of the Graduate School of Natural and Applied Sciences, Çankaya University.

Prof. Dr. Taner ALTUNOK
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

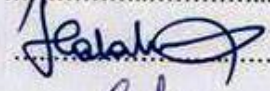Prof. Dr. Billur KAYMAKÇALAN
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Abdul Kadir GÖRÜR
Supervisor

Examination Date: **20/09/2013**

Examining Committee Members:

Assist. Prof. Dr. Abdul Kadir GÖRÜR (Çankaya Univ.) ....

Assoc. Prof. Dr. Hakan MARAŞ (Çankaya Univ.) ....

Assoc. Prof. Dr Fahd JARAD (THK Univ.) ....

STATEMENT OF NON-PLAGIARISM

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Arkan SHAKR

Signature:

Date: 19.10.2013

ABSTRACT

PLAGIARISM DETECTION IN LEARNING MANAGEMENT SYSTEM

Arkan Kh. SHAKR

M.Sc., Department of Mathematics and Computer Science

Supervisor: Assist. Prof. Dr. Abdül Kadir GÖRÜR

SEPTEMBER 2013, 70 pages

Learning Management Systems has became one of the teaching tool same as the pen, pencil, blackboard, book, and notebook. Learning Management System (LMS) is defined and types that are derived from this system such as those focused only on the curriculum and those focuses on the management of the education system, and combination of both are explained. The properties and features that must exist are identified and showed, so that it can be called a learning management system.

The name and the experiences of countries in terms of the use of this system as well as the benefits resulting from the user experiences has motivated schools and educational institutions to start the establishment of such systems.

In this study Moodle system was selected since it is free and open source. This gives an opportunity for schools to work without restrictions or wage. The development and widespread of such systems, programs, and the Internet that use by students in all education levels has became the main reason that students depend on Internet to get solutions to their homework's and share it between them, so it leads to negative use of learning management systems.

Besides benefits and functionalities of these system the major problem regarding to LMS is plagiarism in student homework and school duties in their academic stages. Since the plagiarism detection and controlling process by human is slow process, need for an electronic plagiarism detection system that has been integrated into LMS must be built in order to grade and evaluate electronic student submissions faster and more accurately.

Although there are systems and algorithms to the detect plagiarism, these systems are not open source and they are not integrated in to LMS.

In this study plagiarism detection was developed and integrated into Moodle system. Finally test case is prepared and applied to the developed system in order to observe effectiveness of proposed plagiarism detection approach. The results and error-free application shows that the work of proposed plagiarism system is compatible with Moodle system. In this study our main goal is to demonstrate the applicability of plagiarism detection system in Moodle environment. Therefore performance of plagiarism detection algorithm in terms of accuracy and error write are out of scope of this study and left as future work.

Keywords: Learning Management System, Moodle, plagiarism detection algorithms, plagiarism detection with the Moodle.

ÖZ

İNTİHAL ÖĞRENME İÇİNDE TESPIT YÖNETİM SİSTEMİ

Arkan Kh. SHAKR

Matematik-Bilgisayar Bölümü

Yrd. Doç. Dr. Abdül Kadir GÖRÜR

Eylül, 2013, 70 sayfa

"Öğrenme Yönetim Sistemleri" (LMS) kalem, kurşun kalem, tahta, kitap ve dizüstü bilgisayar gibi bir öğretim aracı haline gelmiştir. (LMS) sisteminin türetilen türleri ve tanımı şöyle yapılır: sadece müfredat odaklı, eğitim sistemi yönetim odaklı ve her ikisinin kombinasyonunu. Bu sisteme "Öğrenme Yönetim Sistemleri" diyebilmemiz için var olması gereken özellikler de tesbit edilmeli.

Bu sistemin ülkeler tarafından kullanılması ve bu kullanımdan kaynaklanan yararlar sonucu okul ve eğitim kurumlarının bu sistemleri oluşturmasına motive olacaktır.

Açık kaynaklı, ücretsiz, gelişmiş ve geniş kullanımı nedeni ile bu çalışmada "Moodle Sistemi" tercih edilmiştir. Bu durumda okullara kayıtsız ve ücretsiz kullanım fırsatı verilmiş olacaktır.

Sistemlerin, program ve internetin gelişmesi ve öğrenciler tarafından yoğun kullanılması neticesinde öğrencilerin de interneti kullanarak sorularına cevap aramaya yada başka bir öğrenciden almalarına sebep oldu, bu da bahsi geçen kesimin "Öğrenme Yönetim Sistemlerini" olumsuz kullanmaları demek oluyor.

"Öğrenme Yönetim Sistemlerinin" yarar ve özelliklerinin yanı sıra ancak intihal önemli bir sorun haline gelmiş ve insan tarafından kontrol edilmesi yavaş ve zor olmuştur. Bu nedenle öğretmenlerin bu olayı daha hızlı ve net bir şekilde tespit etmeleri için elektronik bir sistem ürettik.

İntihalı keşfedecek çok sayıda sistem algoritmalar olmasına rağmen, bu sistemler açık kaynaklı değil ve (LMS) ile entegre edilmemiştir.

Öğrencilerin elektronik cevaplarını daha doğru ve hızlı bir şekilde değerlendirmek için intihali keşfeden sistem LMS ile entegre edilmiştir.

Son olarak İntihal Sistemi ile Moodle Sistemi çalışmalarında hata görülmezken uyum, dakik oran ve hızlı sonuç vermiştir.

Bu çalışmada ana hedefim Moodle ortamında intihal tespit sisteminin uygulanabilirliğini göstermek, bu nedenle de aloritmanın intihali keşfetmekte performansı doğruluk ve hata yazımı bu çalışma kapsamı dışında, gelecekteki çalışmalara bırakılmıştır.

Anahtar Kelimeler: Eğitim Yönetim Sistemi, Moodle, intihal tespit algoritmaları, Moodle ile intihal tespit.

ACKNOWLEDGEMENT

Only, I would like to thank my dear wife for supporting me in completing my study.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

1.1. Introduction

Chapter 1, gives the introduction of the thesis and the main topics of each chapter. Sections 1.2 shows the Moodle System, which provides basis in this thesis and the reason for choosing a Learning Management System. We presents the Moodle System usage statistics and the number of countries that use and access freely as well as the fact that it is available in several languages. In addition, the Moodle provides to communicate with students in terms of how to achieve the objectives of the educational and knowledge performance and even in terms of technological communication with the student and encourage the community to enroll in advance, which displays and clarifies the purpose of this thesis in section 1.3.

Chapter 2, describes the systems for the field of learning Management system in educational institutions. These systems have been designed to help teachers to use Internet in teaching and communicating with learners easily without the need for a deep knowledge of the methods of programming and provides different learner materials which can be obtained from one place, and these systems help a learning environment for a self-managed learner to interact positively with a scientific article. However, Learning Management Systems(LMS) which care administrative matters pertaining to the educational institution does not examine deeper course content.

Section 2.1 shows the scientific definitions and terms related to Learning Management System.

Section 2.2 identify the term, Learning Content Management Systems (LCMS) in which enable to coordinate learning content so that it is easy to access, exchange and transfer between the student and the teacher.

Section 2.3 defines the term, Content Management System (CMS) which is about the management and organization of subject and method of publication or saving, and how to exchange information between the student and the teacher and the methods of communication and exams.

It is well known that the selection of learning management systems are based on specific standards which are all listed in section 2.4.

The experiences of several countries about the application of this system and it is highlighted the positive effects of this system in these countries were listed in section 2.5.

Learning Management Systems provide many opportunities for school administration, the student and teachers in terms of the ease of communication and access to information, regardless of time and place were in section 2.6.

In chapter 3, literature review about plagiarism, algorithms, plagiarism detection softwares, and the reasons for choosing AC were illustrated.

In Chapter 4, AC algorithm is described in detail in terms of algorithm steps with clarified figures and examples in section 4.1 and 4.2 respectively.

In chapter 5, the file types and the role of each file within the Moodle system is identified and shown in 5.1. The steps of implementation are explained with detailed figures, and integration of the developed system in to Moodle is described. The output of the code analysis algorithm in AC system with Moodel in section 5.2.

Test scenarios were developed to check the system operation to ensure the execution of the algorithm code correctly, the system been tested as shown in figures and the steps in Chapter 6, a question is to put in a particular exam and group of students participate in the exam in which there are similar and different answers. After the teacher enters the checked answers, it shows him/her a list of students who participated in the exam with their answers and percentage of similarity in the answers of every student on the basis of the ratios. Thus, the teacher manages the existence of plagiarism in the answers or not.

In Chapter 7, conclusions of the subject of the thesis were discussed in section 7.1. And possible future work suggestions were given in section 7.2.

1.2 Moodle

Moodle is a Learning Management System (LMS) or a Virtual Learning Environment (VLE) or a Course Management System (CMS). It is also known as a free web application with which educators can benefit to create effective online learning websites. Nine years ago, in Australia, a computer science graduate named Martin Dougiamas was testing a web tool which he developed to help teachers to create lessons online. Inspired by his own experiences with "School of the Air", Martin's Modular Object Oriented Dynamic Learning Environment (MOODLE) offered tutors a way to connect remotely with their students in a collaborative and supportive workspace. Moodle has truly changed the face of learning as shown in Figure 1.1. And now used by millions students in thousands websites in many countries, With improved access to the internet, and with commercial companies which are a fast spot for a potential money-earner, many such Learning Management Systems have arisen since then [1]. It is possible to get all the information about this system from the following link https://moodle.org/

Registered sites	82,974
Countries	236
Courses	7,538,212
Users	70,807,568
Teachers	1,293,972
Enrollments	66,747,919
Forum posts	126,056,774
Resources	67,558,919
Quiz questions	183,254,161

Figure 1.1 Moodle Statistics

1.3 Purpose

The main purpose of this study is to clarify the idea of learning management systems and the selection of criteria for this system as well as the benefits of using this system and to get to know the experiences of countries that use this system.

In addition, an online plagiarism detection system will be develped and integrated in in to Moodle so that it can be used by teachers to detect cheating and plagiarism cases in students' submitted answers.

1.4 Thesis Outline

The thesis is organized as follows:

This chapter introduces the thesis and the purpose of the audience. The reader is aimed to understand the research aims of the study after completing this chapter. Second chapter gives an idea of learning management systems, selected standards, benefits and experiences of countries in this area. Third chapter explains the phenomenon of plagiarism and algorithms in this area. The fourth chapter analyzes and converts discuss the implementation of AC plagiarism algorithm. Fifth chapter gives a detailed explanation of how to integrate the output of the code analysis algorithm of plagiarism with System Model. Sixth chapter tests the Moodle system following the merger of the output of the code analysis algorithm and plagiarism. Chapter seven explains future works and conclusions of this study. This chapter is the last chapter of this thesis.

CHAPTER 2

LITERATURE REVIEW

It is essential to clarify what is Learning Management Systems (LMS), Learning Content Management Systems (LCMS) and Course Management Systems (CMS) are. In addition to that, the whole picture of its improvement and to describe the most significant features that build up any system refers to the Learning Management Systems (LMS).

2.1. History and Definition of LMS

The function of the computers for education includes some common titles; for instance, Computer Assisted Instruction (CAI), Computer-based Instruction (CBI), and Computer Assisted Learning (CAL) in its history. Furthermore, it makes clear successively exercise-carry out schedules, much more excelling trainings and more specified education [2].

and makes a shift in the individuals' lives. In terms of schooling, advertisement technology is linked with agreements of converting learning, generally without particular knowledge on the impacts of every one of the technologies once carried out [3,4,5]. Thus, LMS were originated from commercial systems [6,7].

Learning Management System ought to capacitate "the management, transfer and pursuing of combined learning such as online and traditional classrooms for staff, shareholders and buyers [6].

In the e-learning area, numerous networks have aroused as a fourth gap. Three fundamental aspects of it are the retaking of huge totals of information,

the capability to contact with the help of Computer-Mediated Communication (CMC), as well as the progressing competence of the Java language. These typical features assist the production of original e-learning technologies which are named after Learning Management Systems (LMSs) such as Lotus Notes, Moodle, WebCT and Blackboard. Learning Management Systems (LMSs) are noted as Virtual Learning Environments (VLEs) or Course Management Systems (CMSs) as well as being one of the beneficial explanations in e-learning environments for both learners and teachers [8,9,10].

Learning Management System is described as a technology which is a web-based and aids in the estimation of a particular learning progress, designing and delivery. In addition to that, it serves a program between learners and instructors online and a number of exercises can be done with the help of it. Group instruction has important benefits in terms of the instructors and it has chance for the communication in terms of the knowledge about learning to develop instruction about technology and study [11].

A Learning Management System (LMS) generally assists to control a system's training activities and capabilities. The activities controlled by the LMS can be divided from teacher-controlled education to training conferences to online education. According to a user, a LMS helps a useful technique to carry on a personal abilities and capabilities and assists a placement and entering for meaningful educational activities to develop more a student's ability. An LMS aid contact to web-based lessons for the students. Officially, an LMS provides to access, follow, control, and inform upon the instruction and capabilities in a program. Generally, it particularly examines capabilities, instructions, and planning of carrying the activities. It does not concerns with building up, reserving, controlling and development of subject matter only [12].

In an LMS, the general components can be separated to these four aspects: Tools for course administration are devices that are profited to manage and observe the lesson, such as to distribute or record the learners' marks. Tools for distribution are devices that enables instructors to upload papers that learners is able to see clearly. It may be hard copy, but it may be any type of soft copy. This is through instructor-to-student

dissemination of knowledge. Tools for communication allow knowledge to deliver and distribute instructor-to-learner and vice versa or learner-to-learner. E-mail is one of the most known one of them. Tools for interaction are discussion meeting boards, wikis and other devices that bring studies and build up feedback. These devices are of special attraction because they assist learner activity and work together, so that may promote learning [13].

A Learning Management System (LMS) refers to a system which manages teacher-controlled and e-learning lessons and carrying on learner advancement. An LMS can be benefited to observe the efficiency of the program's instruction and education. It is really good in guaranteeing state-and federal-mandated lessons are transferred in an appropriate manner. In addition to that, it is an online program or technology applied to design, carry out, and helps a particular educational progress. Usually, an LMS presents to a teacher with the help of a method to build up and transfer the subject, observe learner attendance, and to help learner achievement. Furthermore, it can supply the learners with the capacity to benefit from shared aspects, such as spontaneous meetings, video chats, and forums.

2.2 Learning Content Management System (LCMS)

Differently, an LCMS assists to build up, benefit, placement, transfer, and develop the learned subjects. Subjects is usually carried out in a unified subject depository in the shape of tiny, self-defining, special recognizable objects, or educational objects, every one of which is enough for one or more described educational objectives. Every educational object might have been built up from mark or by re-planning standing information papers in other arrangement. An LCMS can place and transfer an educational object to the user as a personal section to provide a work-particular need or transfer the educational object as a piece of a bigger lesson, plan, or educational activity described in an LMS [12,14].

A well-developed LCMS carries on the user's contacts with every educational object and benefits from this intense knowledge to transfer advanced individualized educational involvement whereas supply the writers with detailed outlines for

examining the comprehensibility, effectiveness of the subject, thus it may developed on an kept support [12,14].

An LCMS particularly examines building up, re-benefiting, placing, transferring, controlling, and developing the subject. Furthermore, in confident circumstances, the tendency expands to promote information associations and gathering unorganized information in the educational object in a concrete shape. However, an LCMS does not concern with capability control, the expanded managed practicality of educational activities or planning of these activities [12,14].

2.3 Course Management Systems (CMS)

Course Management Systems (CMSs) are software systems that were firstly planned to reinforce formal learning in colleges and high schools. CMSs assists instructors with enabling to practice these works uploading class materials in e web-based program. Most of CMSs brings pre-planned buttons for the lesson outline, its timetable, and materials including particular subjects, for instance, readings, and PPT from the lessons. Learner development is kept with the help of assessment elements, which allows the teachers to provide quizzes and exams online and an online gradebook, where teachers can publish student marks. Discussion board, where teachers and learners can discuss the materials and follow the class discussions. Other tools wich enables teachers to publish announcement to students and contact personally with them. Lock box for students where student can archive their materials in a safe place. Course statistics, which brings information including who accessed the website and which time. Examples of CMSs contain the advertisements like Blackboard and WebCT, and Moodle [15,16].

2.4. Requirements For LMS Systems in an Academic Environment

Requirement which should be accessible in the learning systems. And systems which do not include these needs cannot be thought learning management systems in the right path. These are suggested by [6,17,18,19,20]. These suggestions are classified below:

1. Course Content Management

Course content management is the group of needs about in which way the system controls subject index objects, works, distributing and re-benefit of those objects and other sharing knowledge. It is suggested that:

1. The system must back up memory of special files that are sent by the member.
2. The system must back up class works which are sent by the students to the page.
3. The system must back up re-access and allocating the class materials between instructors.
4. The system must back up an online storage which includes classroom materials and the information can be distributing between all the members.

2. Evaluation

This group is for the needs about in which way the system backs up feedback from the users in terms of sharing opinions in which way the results of them are demonstrated and how those can be evaluated, whether in the program or transported from the system to be evaluated with another application. It is suggested that:

1. The members ought to own the accessibility to give feedback about the lessons after they end up them.
2. The responsible person ought to be able to see the reply on the page.
3. The responsible person ought to be able to evaluate the replies.

3. Third party integration and standards support

This group is for needs about in which way the system deals with contact with third party systems and extrinsic information centers like databases and online libraries. The system ought to back up various online platforms and online improvement standards for LMS systems. It is suggested that:

1. The system ought to be adaptable with other third party software to clarify combination.
2. The system ought to be able to associate with extrinsic research databases and online libraries.
3. The system ought to be able to run on various hardware and software platforms.

4. Third party content support

This group is for the needs about in which way the system deals with the subject improved on other systems or built up by third party organizations benefiting from standing frameworks and subject built up by used third party programs. For instance texts printed with Microsoft word or open office. It is suggested that:

1. The system ought to be able to deal with the subject formed by third party developers.
2. The system ought to be able to deal with the subject from different third parties such as flash, MP3 or video formats… etc.
3. The system ought to back up the subject improved on third party platforms such as MS& open office, photo-shop, video and audio editing software… etc.

5. Usability

This group is for the needs about in which way the member incorporate on the system ought to be planned to simplify to benefit and clear to begin with and easy to learn as well. It is suggested that:

1. The system ought to be simple to use and learn by the learners and the instructors.
2. The system ought to be available via a browser.

6. Configuration and modification

This group is for the needs about in which way the system is to recognize and correct and in which way it can be changed to work in standing server environments.

1. The managers ought to be able to change and correct the systems source code.

7. Technical requirements

This group is for the needs about in which way the system ought to be able to control intense intercourse transfer, assigned works and calculating when the member core amplifies. It is suggested that:

1. The system should be trustworthy under detailed assigned works.
2. The system should be able to deal with the increase at the amount of the members or knowledge.

8. Learning and pedagogical requirements

This group is for the needs about in which way the system backs up the progress of the learning since it is a significant section of class material improvement to contain visual imagery and communication and take back the useless information. It is suggested that:

1. The system ought to back up mutual class subject materials.

2.5. The Use of Learning Management Systems

2.5.1 In New Zealand

The New Zealand Ministry of Education displayed particulars on the schools where LMS is being used in January 2011. Figure1 shows the 25½ of New Zealand schools in which a LMS is used. The LMS from which is benefited displayed in Figure 2.1 are: UltraNet, First Class, KnowledgeNet, Moodle, My Classes, Scholaris, and Other systems. The number of schools which benefit from UltraNet are 227, KnowledgeNet are 229, Moodle are 116, My Classes are 35, Scholaris 7, and other systems are 4. According to this Figureit can be said that KnowledgeNet usage is the highest one and UltraNet follows it. On the other hand, other systems is the lowest one [21].
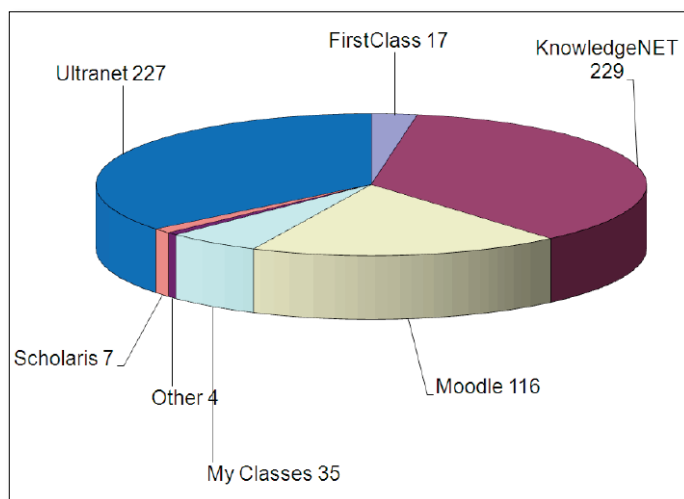
Figure 2.1 Number of Schools Using LMS According to the New Zealand of Education [21]

2.5.2 In the United States

Learning management systems from which are benefited in the USA are well-known and it contains commercial systems and self-developed systems. The results depends on a random sample of 100 of around 2000 schools. According to Figure 2.2, the used learning management systems are Blackboard, WebCT, eCollege, Angel, Ed2Go. According to this figure, the usage of Blackboard is 33%, WebCT's is 18%, eCollege's is 5%, Angel's is 3%, and Ed2Go's is 2%. In addition to that, the percentage of schools which does not use a Learning management system is 21%, the percentage of schools which built their own learning management system is 5% and the percentage of the schools which cannot determine a system is 2%. Thus, it can be said that the most popular LMS from which is benefited at the schools in Usa was Blackboard, and the second most benefited system was WebCT [22].
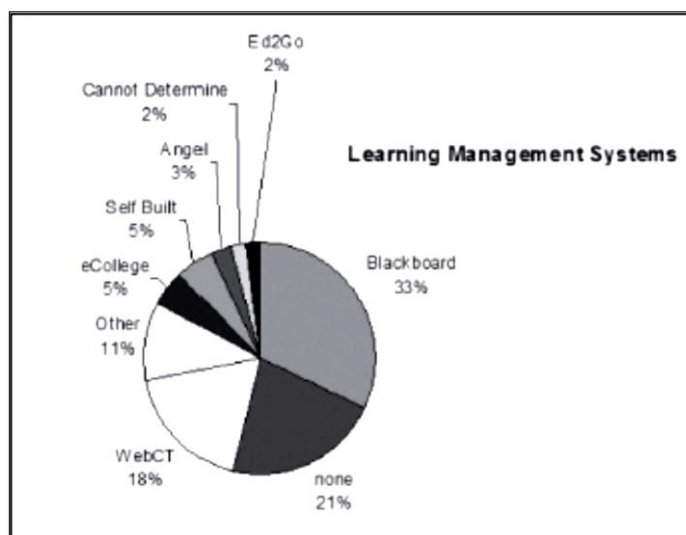
Figure 2.2 The Use of Learning Management Systems in the United States [22]

2.5.1 Sultanate of Oman (Sultan Qaboos University)

It can be seen that the students replies was not very high before according to the table and after the mean of them increased. In Table 2.1, the mean of overall, is 2.66 before, and after it increased 4.69. So it can be said that the common opinion and knowledge of students about the learning management systems (Moodle) in terms of a learning tool was weak according to their replies. However, later on, they realized the positive effects and gave positive feedback for the learning management systems (Moodle) [23].

Dependent Variables	Mean		Size of Change
	Before	After	After-Before
Understanding	2.85	4.73	1.88
Learning	2.88	4.70	1.82
Helped in Exams Preparation	2.25	4.65	2.40
Overall(Helps in course learning)	2.66	4.69	2.03

Table 2.1 Learning Management Systems (Moodle) and Learning [23]

It is seen that the more learning management systems (Moodle) is used and accessible, the more students' attitudes and opinions about learning management systems (Moodle) increases, and it is beneficial for the students' understanding and

learning the course material. The means of the survey are shown in Table 2.1 and in Figure 2.3 [23].
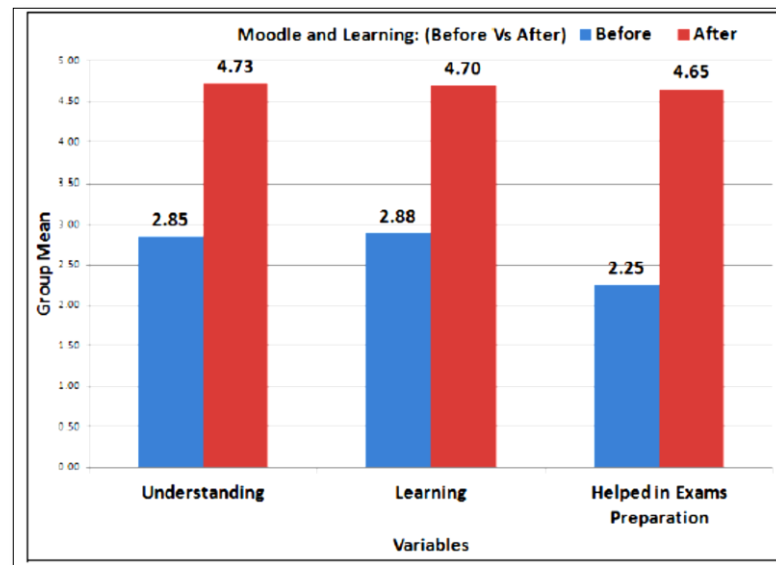

Figure 2.3 Learning Through Learning Management Systems (Before Vs. After) [23]

2.6 The Advantages of an LMS

The five particular advantages of an LMS for the works are listed in the following [24].

1. **Centralized Learning Environment to Ensure Consistency**

Naturally, an LMS has the ability to make all kind of educational content, constructive content, and practice content accessible to persons all the time from any place online. Various members is always able to connect the LMS. The LMS guarantees flexibility in transfer and estimation because every member can access the certain material in the certain manner and can be figured out with the help of shared pre-testing and/or post-testing ways. An LMS enables the members to plan and set up individual educational modules. This component is particularly significant when new material is brought out, standing material is upgraded or determined performances are importantly corrected. Furthermore, this component carries out upgrades to guidelines and performances. Workers cannot deny that they see them.

2. **Tracking and Reporting for Enhanced Performance**

   The LMS enables executive members to check a needed learning way, carry on the process against the learning way, take a look at again the documents, and sign up for extra courses. Workers can suggest these courses via different media such as instructor-led education, online education or webinars. Administration can see the same documents and can examine the data to Figureout the parts of success and the part that should be improved.

3. **Immediate Capabilities Evaluation**

   The LMS enables the members to be assessed before taking a course, while attending the course and ending up a course. Workers is able to evaluate hindrance by routinely managing determined evaluations with the LMS. They can check the documents of the results to decide success levels and real time to finish
   eve ry course and its parts.

4. **Continuous Product and Service Proficiency for Employees who Interact with Customers and Clients**

   The LMS brings a main opportunity for corporations to alter product portrayal, details, needs, shapes, and to enable to uploading of new product or information about the transferring. Member will see the same educational courses, and the same materials to evaluate. Members customize pre-determined course accomplishment dates and observe the total number of participators accomplishing the course in any time. Moreover, an LMS enables works to manage upgrades and analysis in web and help information levels and capabilities.

5. **Regulatory and Legal Compliance**

   Most corporations should meet some allowable and supervisory needs. The medical area and the pharmaceutical area are the manufacturing with very tough needs. Some allowable and supervisory needs are met by one time incidents whereas the others need routinely check or reconfirmation. Industry, generally, can be needed to meet training regulations assorted by corporate or standard agencies such as OSHA and

the EEOC. An LMS guarantees constant transfer of regulated course suggestions, for example, HIPPA, Lockout/Tagout, Sexual Harassment… etc. The LMS, via its kept system, guarantees that the termination of those who took the needed courses can be seen in any time. Corporations can define fields of non-conformance and show remediak actions to decrease the risk of non-compliance with needs.

CHAPTER 3

PLAGIARISM DETECTION ALGORITHMS IN LEARNING MANAGEMENT SYSTEMS

Student plagiarism is in fact huge problem facing academic institutions. Plagiarism is often interpreted as copying or duplicating another's work from the originator such as students or from sources such as textbooks, without indicating the proper acknowledgement of the source.

In 1999, David Gitchell and Nicholas Tran [25] presented implementation of a sim program to measure similarity between two C computer programs which is applicable for detecting plagiarism among a large set of homework programs. This software was used to construct tools to assist the teaching of computer science in academic schools and instaurations.

Nadelson, in 2007, made a survey to collect the observations of 72 academics on issues interested in academic misconduct and reported 570 cases of suspected plagiarism. The majority of cases reported were "accidental/ unwitting plagiarism" with 134 of those cases involving undergraduate students and 39 involving graduate students. Also, the academics reported that a large number of cases involved students presenting papers copied from Internet. Cases ralated to "purposeful plagiarism", "class test cheating", and "take home test cheating" were also reported [26].

Bingfeng Pi et al, 2009, started with the discussion of the seriousness of near-duplicate existing in short messages. Then, they reviewed how SimHash works, and its possible merits for finding near-duplicates. By the way, they demonstrated

a series of findings, including the problem itself and the benefits brought by SimHash-based approach, based on experiments with 500 thousands of real short messages crawled from Internet. This contribution may be considered as a valuable reference for both researchers and applicants [27].

Plagiarism is a real problem in programming courses. In 2001, Culwin et al, made a study on source-code plagiarism in which they got data from 55 higher education computing schools in the United Kingdom. It was founded that 50% out of the 293 schools who participated in this survey believed that plagiarism has been increased in recent last years. additionally, there were 22 of the 49 respondents gave estimates ranging between 20 and 50 percent of students plagiarism in initial programming courses. In academic work case, plagiarism is considered as an academic offence and illegal offence, and is controlled by the rules and regulations of the institutions [28].

Automatic and computer-based plagiarism detection systems are presented to detect plagiarism in student works. Effectiveness of detection in such systems relies on what types of plagiarism they can detect. Such systems offer precious benefits regarding to saving time and efforts of academics to conduct the detection process by themselves. Computer-based plagiarism detection has been interesting to academics in the last two decades due to the fact that using such tools decreases academics workload via automating the comparison procedure and detecting similarity of student works fastly, which the academics need to check for doubtful similarity [29,30].

The use of computer-aided plagiarism detection furthermore concerns a set of ethical and legal issues. These issues are caused both by technical imperfectness of plagiarism detection algorithms (for example, a system might incorrectly suspect a student's work as plagiarized) and by misunderstanding the role of plagiarism detection software in educational process. Due to the importance and the rising interest in ethics of automated plagiarism detection [31].

Daniele Anzelmi,et.al, in 2011, implemented SCAM (Standard Copy Analysis Mechanism) which is a relative measure to detect interfere by comparing on a set of

words that are common between test document and registered document. This system, is evaluated with metrics of precision and recall [32].

There is an increasing and common issue in the scholar process. The plagiarism in the exams. The classic, not automated disclosure of such sort of plagiarism is hard, unreliable, and slow technique. This task addresses to build up a plagiarism detection program which is accessible by Internet and which is able to assist college or school instructors to get a more improved awareness for a learner's task. In present, there are lots of standing detection algorithms.

Since In this study our main goal is to demonstrate the applicability of plagiarism detection system in Moodle AC algorithm is consider to be most appropriate algorithm because of it is simplicity. Since Moodle is developed in PHP environment, in this study the selected algorithm is implemented with PHP.

On the contrary, the program which is introduced here is unpaid, open-source and can be associated with Moodle. Existing Cheating integrated Algorithms and Open Source Software will be introduced in this study.

3.1 Detection Algorithms

There are many algorithms for plagiarism detection programs which have just been improved. This section presents such algorithms and defines their characteristics. These algorithms may be practically categorized into two parts [33].

1. Attribute-counting method.
2. Structure-metric method.

3.1.1 Attribute-Counting Program

This is the first kind of plagiarism detection algorithm. Such programs figures out the degree of the likeness between a pair of works by benefiting from four basic program statistics: [33].

- Estimated values of distinct operators.
- Estimated values of distinct operands.

- Total estimated values of operator occurrences, complete distinct types.
- Total estimated values of operand occurrences, complete distinct types

3.1.2 Structure-Metric Programs

This kind of plagiarism detection algorithms presents more abundant amount of metrics and awareness of likeness for the arising characteristic vector to develop presentation (depending on structure and metric match).

Such algorithms generally depends on transforming the program to a flood of tokens (so not caring simply unstable knowledge for example blank, border gaps, explanations… etc.) and then matching these token floods to discover likeness between them.

The most improved programs in this part (regarding plagiarism detection process) are SIM, MOSS, JPlag, AC and CodeMatch. Here is a short definition of these programs [33,34].

3.1.2.1 Software Similarity Tester (SIM)

It was developed by Gitchell and Tran in 1999 as a program for figuring out the likeness between the text which is written in C, Java, Pascal and general language. **SIM work degrees:**

- Read the program files: read the file and save it in succession.
- Clarify the group of appealing sequences: the algorithm clarifies the comparison between two files.
- Clarify the row figures of the appealing sequences discovers the beginning and last row Figurefor every block.
- Print the index of the sequence in succession

3.1.2.2 Measure of Software Similarity (MOSS)

It was improved by Alex Aiken at Berkeley in 1994 as a program to calculate the likeness of source code written in C, C+, Java or Pascal. MOSS analyzes the source code in the file, indicating it and carrying out the matched algorithm to the indicated shape of the code. Then it matches it with the source code in other files [35,33].

3.1.2.3 JPlag

The total knowledge transferred about JPlag is so rare. JPlag does not match the internet. It is planned to discover likeness between the student homework, which is generally adequate for computer programs. But its basic feature is to transfer the programs to shown rows and matching these rows [36,33].

The valid JPlag website sums up its study by explaining that likeness of 0% or 5% can be shown by the likeness value only, so it is definitely a plagiarism, however, if there is a 40% likeness, these circumstances generally ought to be examined manually for a final decision. The basic features of JPlag can be sum up as:

1. JPlag is online and accessible.
2. JPlag has a strong member participation for getting the results.
3. JPlag is an effective resource and successes to large submissions.
   JPlag has a really efficient falsification detection function.

3.1.2.4 Software Integrity Detection (SID)

It determines the likeness between source codes by calculating the accessible information among them. SID is simple to benefit from the program to find plagiarism in a source code and has indicated to be the most efficient at detecting the ones who cheats. SID presently can work with Java and C++ source codes. Looking at these two programs, SID calculates the accessible information among these two softwares, the accessible information expanse among two programs X and Y is identified as [33,37].

$$(x,y) = \frac{k\,(x) - k(x/y)}{k(xy)}$$

3.1.2.4.1 SID functions in two parts in Figure 3.1

1. In the first part, source programs are separated to produce shown successions by common lexical interpreter.

2. In the second part, Token Compress algorithm is practiced to calculate the accessible information metric d(x,y) among each program couple in the works.
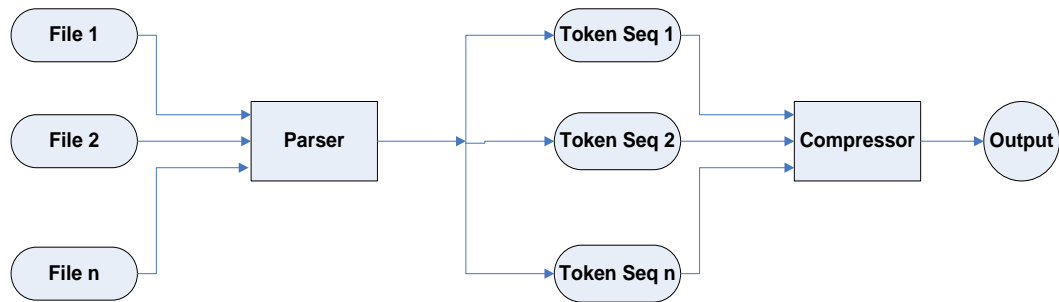

Figure 3.1 SID Phases (Adapted From [37])

3.1.2.5 Code Match

It matches each file in sole guide with each file in another guide, containing all sub-guides if offered. CodeMatch generates an information bank that can be conveyed to an HTML main report that shows the most suitable integrated pairs of files. You can on any specific pair shown in the HTML main report to examine an HTML elaborated report which presents the particular parts in the files (explanations, definitons, instruction orders or comments) that leads to the high integration [33].

3.1.2.6 Anti Copias (AC)

It introduces a web page to find the likeness between the works or projects and any member can freely benefit from it. This site gives statistical tests and a plenty of graphical illustrations assit in the judgment of test results. AC functions are accessible for research and improvement at http://tangow.ii.uam.es/ac [33,38,39].

3.1.2.6.1 AC works with these stages to match students' works:

1. Distance integration

This part sets the letters/symbols in succession and transferring those to sequence of tokens after clearing away the explanations and blanks from the source file.

Token counting similarity distance

This part numbers the tokens between two works by benefiting from decomposer and presents the percentage of the likeness.

3.1.2.6.2 The matching performance in AC can be explained in detail as shown in Figure 3.2

1. Input source code or the symbols are separated in succession.

2. Lexical testing builds up tokens by transferring the succession of the symbols to successions of tokens to make comprehensible characters. For instance, 12*(20+12)/40 is made 12,*,(,20,+,12,),/,40 after lexical testing.

3. Tokens are prepared.

4. Syntactic testing controls if the tokens create available phrase.

5. Semantic separation forms the real separation by matching the successions and provides the outputs.
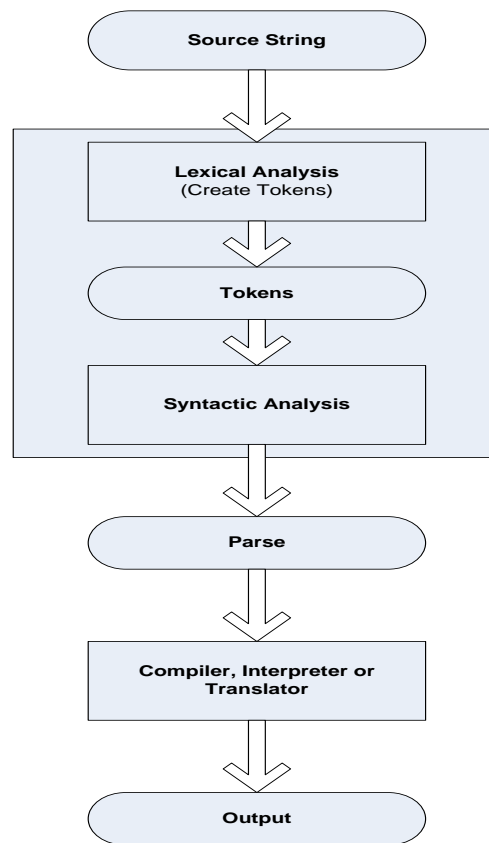
Figure 1.2 Parsing Steps (Adapted From [33])

3.1.2.6.3 AC features can be sum up as:

1. **Locality of the tool:** it may be transforming and benefiting from it specifically, it is definitely sole- standing and can be used in any computer with an appropriate Java runtime context.

2. **Wideness:** a lot of programming languages and alternatives that the function can perform.

3. **Security:** It may exclusively accessible for its owner or it can be seen by the public.

4. **Documentatiton**: standing documentation of the function of AC, source code and integrated programs which works for the understandability.

5. **Algorithms:** the characteristic and alternative likeness distances combined to the function.

6. **Visualization:** standing graphical and chunk function.

7 **Back up:** the accessibility of long time back up of the function.

3.2 Conclusions

Depending on the information, AC is chosen to be benefited for the usage of the task. The aim of this choice can be identified as:

1. AC is simple to use.

2. AC can work with programming and other languages.

These issues makes AC functions a good candidate for this work. In this work, the AC code cannot be used, since it is hard to combine it with model code, thus the code and reprogramming the basic algorithms are analyzed by applying to PHP to benefit from it in the plagiarism detection program.

CHAPTER 4

PRACTICAL SECTION 1

Since, the algorithm is selected, it will be presented here. Implementation of the algorithm so as to facilitate the process of integrating this algorithm with the Moodle system will be explained in the next chapter.

4.1 Algorithm

The algorithm consists of three sections.

A Start

For each document in the collection repeat

1. Clear all the blanks as an example of the extra spaces between words and clear the lines so that the text on one line as an example.
2. Removal of
    a. All the words confined between the A - Z, as well as figures confined between 0-9.
    b. All the symbols and signals, except for the words confined between the A - Z, as well as numbers confined between 0-9.
3. Conversion all text and symbols to-array.
4. Order of the words in the array in alphabetical order to facilitate the comparison process.

B

1. Dictionary creation: Calculate the number of repeated words and numbers, ignore the repeating tokens.

    Note .Steps A, B, applied to the answers (texts) of all students.

End Repeat.

C

Comparison of two documents (Create an array for each documents)

1. Count the number of words and symbols in each array and call this array as token array.
2. Merge token arrays (token array contains words that appears in text)
3. Clear the duplicate words or symbols by keeping only one copy of each duplicate in merged array.
4. Find total words and symbols in the array resulting from merge and duplicate removal operations.
5. Choose the least number as a basis for comparison.
6. Find the total of similar words and symbols by comparing merged(step 4) array with a token array with a less tokens(step 5).
7. Find similarity is calculated dividing the number of tokens of an array with less number of token (obtained in step 5) by number of tokens of a merged array (obtained in step 4).
8. Show results.

4.2 An Example of the Algorithm

Let's assume that we have a text No.1 which is the answer of the first student and text No.2 is the answer of the second student.

Note: These steps are carried out on all the answers.

The answer of the first student is located in Table 4.1. In table 4.3, the answer of the second student is located, Table 4.2 and 4.4 show the texts after removal of extra spaces and new line marks.

Text 1
Arkan   went to (Cankaya University) ? .

Table 4.1 The Answer of the First Student.

Text 1
Arkan went to (Cankaya University) ? .

Table 4.2 The Answer of the Second Student.

Text 2
Arkan   went to   (Cankaya University) ?. But ali existing in (Cankaya University)!

Table 4.3 Delete Blanks Lines of the Answer of the First Student.

Text 2
Arkan went to (Cankaya University)?. But ali existing in (Cankaya University)!

Table 4.4 T Delete Blanks Lines of the Answer of the Second Student.

Table 4.5 and 4.6 show the processes of isolating letters, special symbols for the answers of first and second students, respectively.

Text 1	
Isolate the words and numbers	Arkan went to Cankaya University
Isolate the symbols and signals	()?.

Table 4.5 Isolated Words, Symbols and Prepositions of the First Student's Answer.

ext 2	
Isolate the words and numbers	Arkan went to Cankaya University But ali existing in Cankaya University
Isolate the symbols and signals	()?. ()!

Table 4.6 Isolated Words, Symbols and Prepositions of the Second Student's Answer.

After that, texts are seperated into array of tokens, then sorted alphabetically, as well as the special symbols in order to facilitate the comparison process as shown in the Figure 4.1.
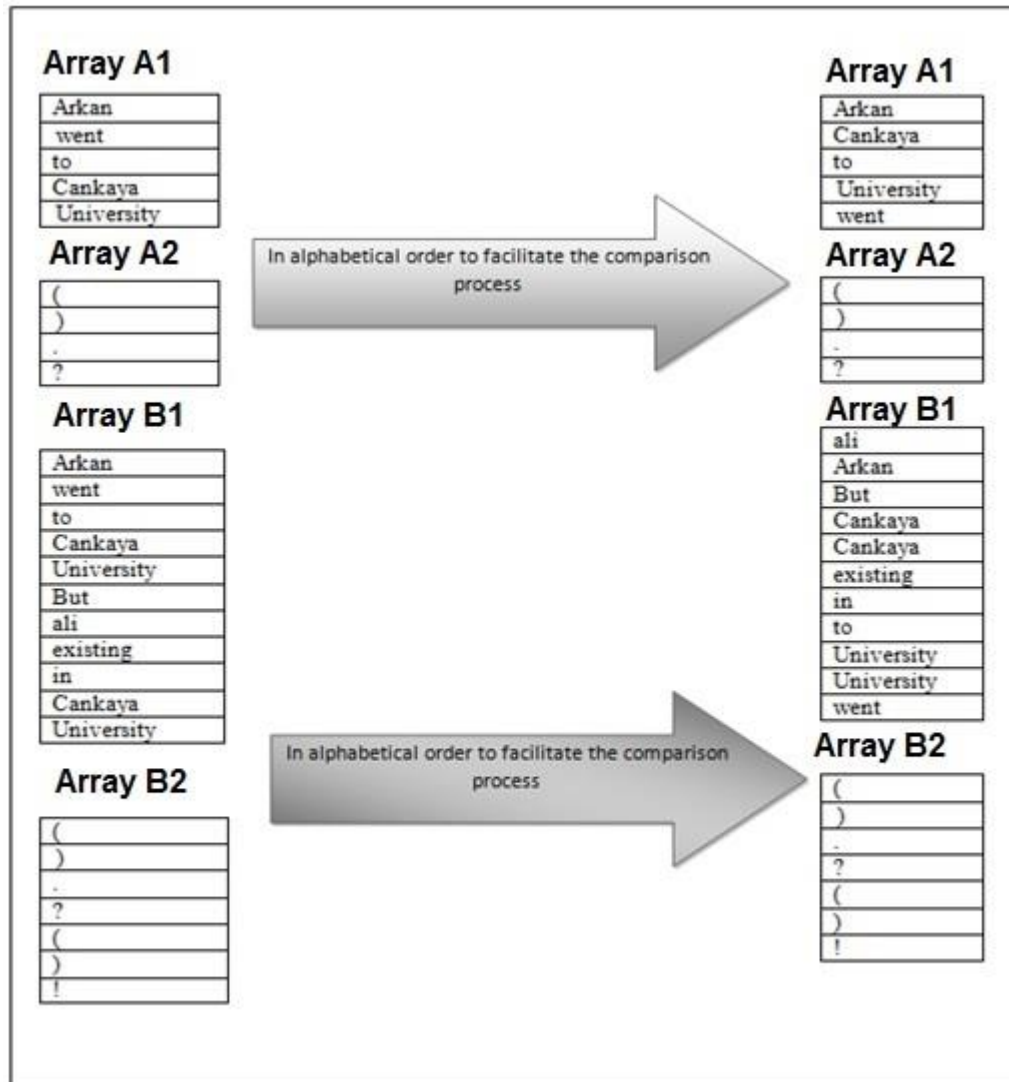
Figure 4.1 Illustrate the Order Answer Text No.1 and No. 2 in Alphabetical.

Then those arrays of tokens of each student's answer will be merged together. After that duplicate tokens will be eliminated as shown in the Figure 4.2.

As shown in Figure 4.2 the alphabet tokens array A1 consists of 5 enries, while the array B1 consists of 11 enteries. The result of merging them is Array C1 will be 16 entries, and after removing duplicates their will be 9 entries in array D1. The same process will be applied on the arrays of special symbol tokens.
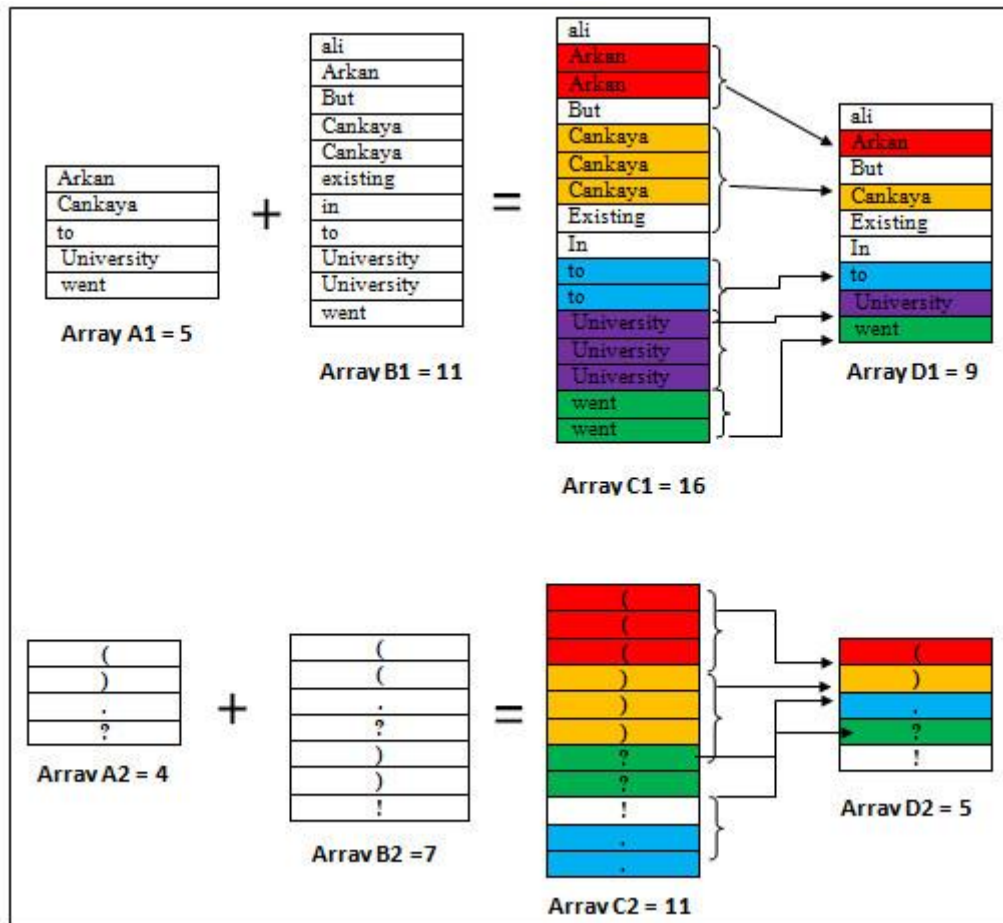
Figure 4.2 Integration Scheme and Delete Duplicate Arrays Text 1and Text 2

When comparing the similarity of array A1 (5 elements) with array D1 (9 elements), the output is array E1 (5 elements). The same process will be applied on the arrays of special symbol tokens as shown in Figure 4.3.
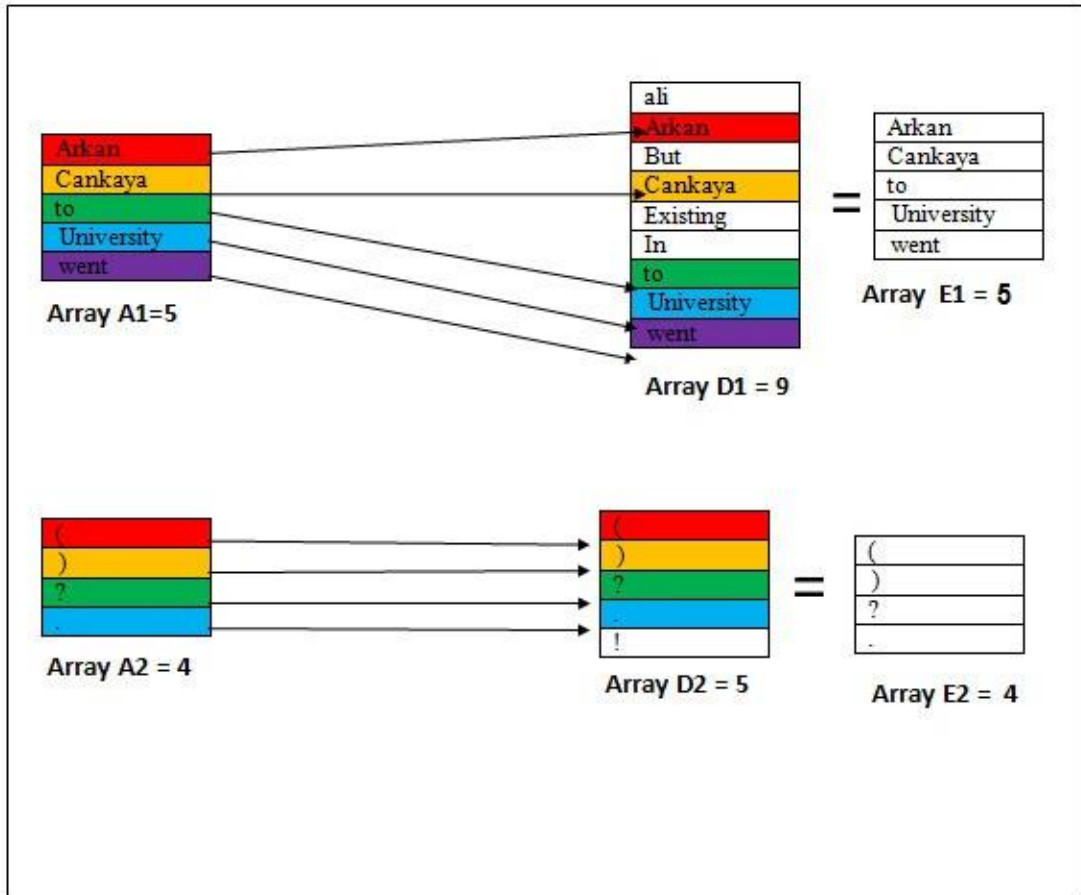
Figure 4.3 Illustrate the Comparison Between Array Words, Symbols for Text 1 and Merged Array for the First Text and Second Text.

To find the percentage of similarity between first text and second text, we use the following equation:

$$(X_i, X_{i+1}) = \frac{X_i}{(X_i + X_{i+1}) - K} * 100 \qquad (1)$$

X_i = number of tokens of ith student answer.

K = number of duplicate tokens in resultant merged array of student's i and i+1 answers.

Numerical example consider text No.1 and text No.2 in Table 4.2 and 4.4 respectively.

$X_1 = 5$ which is number of the token in array A.

$X_2 = 11$ which is number of the token in array B.

$K = 5$ which is number of the token in array D.

By applying (1) the above example, the percentage of similaty in text tokens between the first (array A1)and second (array B1) answers was 55% While the similarity in special symbol tokens was 80% for the same answers.

CHAPTER 5

PRACTICAL SECTION 2

In this part, the features of the code are explained, which includes Moodle program that is chosen as one of the main learning management programs in which it is used to combine code plagiarism detection and identify the variety of the folders in the Moodle program and feature of the way of every one of them. Furthermore, the stages to combine the algorithm which is mentioned in the fourth quarter with program Model upping stages code are explained. In addition to that, the transforming them that will cause mod/quiz in the folders, particularly, in the renderer.php and reviewquestion.php. Remember that this code is combined with Moodle program for school or for colleges and education. To make deeper research for the download and configureit to a school or a college, check this link https://moodle.org/

5.1 Moodle Code

Since PHP is a translated language, the Moodle code is saved as source code files on the web. The PHP translator separates the code on the fly and the output is sent by the web server program when a specific file is offered on the server. The "M" in the Moodle means "Modular", and its control structure shows that every one of the top-level folder indicates a general part of Moodle as explained before. Lots of the basic parts work with plugin modules. Every plugin owns its own folder inside the other part's folder. In some circumstances, modules can work with additional plugins, as well. For example, from the last member's thought, modules are put by copying the module to the suitable folder place on the web. Moodle finds the new module and later, when a manager signs in the program, places the module's SQL code, runs it and lastly, shows the results.

Updates are run same as Moddle, by following the database and uptades the database if it is needed accordingly. This simple integration for last users may be difficult for the developer.

This screenshot is a illustrate list of folders and files after installation of Moodle list of current Moodle installation.

All guides in the basic folder will not be covered not, but the features of some significant folders benefited by developers who modifies Moodle will be explored. Moodle benefits from a basic terminology for modules, in which all modules are covered in their folder. Furthermore, the folder name is the name which Moodle shows in its integration when showing the module in Figure 5.1 [40].
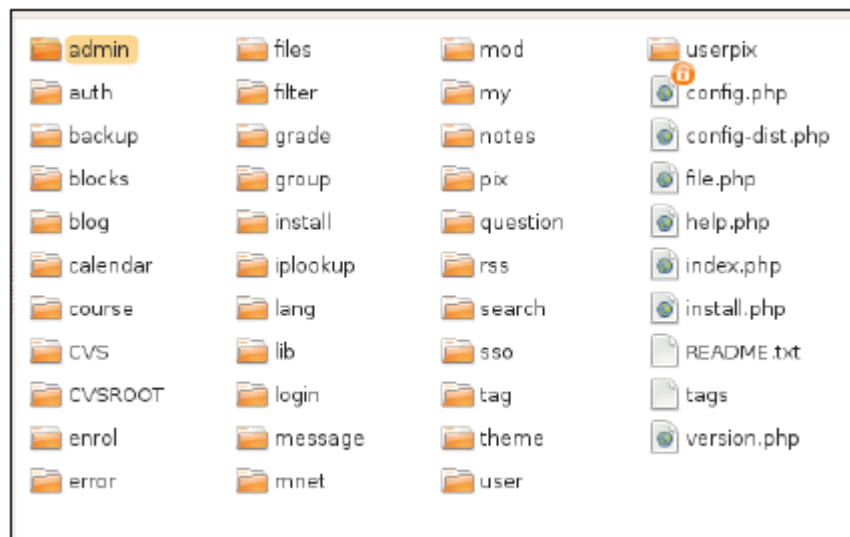
Figure 5.1 Moodle Folders (Adapted from [40])

5.2 Detail Information about Each Folder

1. Admin:it keeps the PHP files that manages the manager user's integration. They contain the cron.php file, and it is run as a set of performance to work program ordering projects for example message transfer and course support. Generally cron.php performance to work a set of operations are examined.

2. Auth: It includes all confirmation modules for Moodle. Every module owns its guide in this folder. Confirmation modules manages the members' makings, member profile information, and a main adaptation work.

3. Backup: It contains the basic course backup equipment for the program. They are the tools for the backup, save, and transfer of the courses. Every personal course module deals with its backup code and benefits from these tools if needed. Every module is self-included, thus developers can add modules to Moodle and they do not need to change the main code.

4. Blocks: they are used to show the boxes of data and they may be in the right side or left side in the Moodle page. It is one of the very basic module kind to create, and inclined to perform with many varieties of Moodle with little or no change.

5. Course: It has a great significance, since Moodle is designed around courses. Developers generally changes course formats and reports. Modifying course formats can help to shift the layout of courses.

6. Enroll: It includes all enrollment modules for Moodle. They manages the making and controlling course-degree role works. They are also another way to find automation.

7. Files: They can let Moodle to integrate files to the program. It contains file installations availability management, and checking the files. Files encounter a large rewrite in Moodle 2.0. Moodle 2.0 can let saving and managing files in external file repositories as Alfresco, Box.net, and Google Docs.

8. Filter: It is a text/order phrase-based search-and-change function. The program improves with members' signing in context from the database while the webpage is made. Filters compares and change the webpage before it is shown. For instance, there is a math filter which can work with auto converting of TEX markup language to match graphics. The Multimedia Plugins filter detects references to main media varieties and rolls the text in a suitable tags, to integrate the media to the webpage. This is a strong ability, but it should be carefully improved and performance suggestions should be kept in mind.

9. Lang: It keeps the main program language sequences. It was founded by Moodle's localization and language back up. All sequences shown to the last member are placed by this function. Language sequence placement are also kept in Moodle info

lang folder. This tool lets for simple local change of language fitlers. This is a tiny part of /lang/en_utf8/moodle.php language file. Aware how every sequence which is shown to the last member is placed to a sequence by benefiting from a key value which is descriptive of the sequence aim.

10. Lib: It keeps the main program library tools. Classes and functions which are identified are used in this folder, since modules and changes are developed.

11. My: It is a light-weight area in Moodle. It supplies a set of courses which a student presents, containing a sum up of future course works. The member can modify the blocks on the own portal. It gives a good place to show individual info with minor customizations in Moodle. For instance, my can be used as dashboard place in lots of change works.

12. Theme: It keeps all Moodle themes and any specific themes uploaded on the program. Themes are an integration of CSS, HTML, and PHP. Every theme owns it folder. The theme program is beneficial for identifying the visual skin, header, and footer of the Moodle page. On the other hand, it is limited to change the number of Moodle page. For instance, definite, types of the Moodle page are coded to show in a definite path.

13. Mod: It keeps activity modules like homework, quizzes, wiki, forum, and course modules. Learning activities are the main ones of any course transferred by benefited from Moodle. Activity modules are more difficult to make rather than blocks, since they support, customize, and keep grades and they should teach something to the student in Figure 5.2.
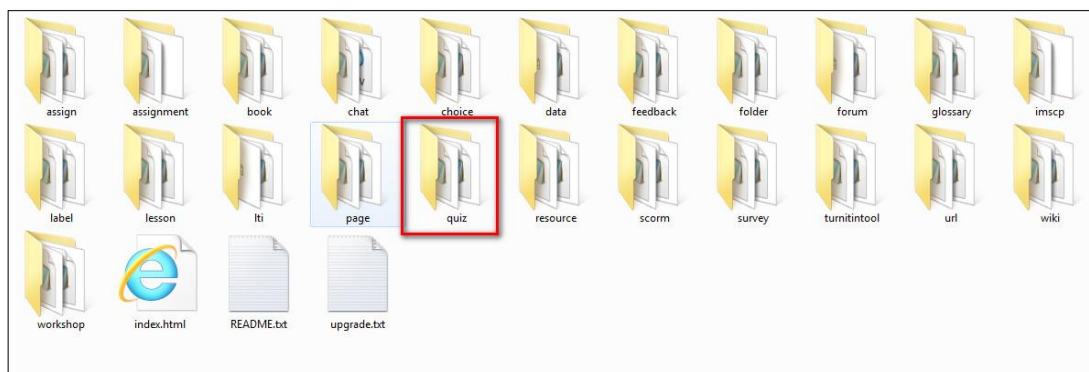
Figure 5.2 Mod Folder Contents

5.3 The Stages to Combine the Code with the Program

The algorithm is converted to php code. After that, the code is stored as similar.class.php file format and saved in the C:\xampp\htdocs\moodle\mod\quiz, as show in Figuer 5.3. Keeping in mind that the details of this code exists in the appendix A.
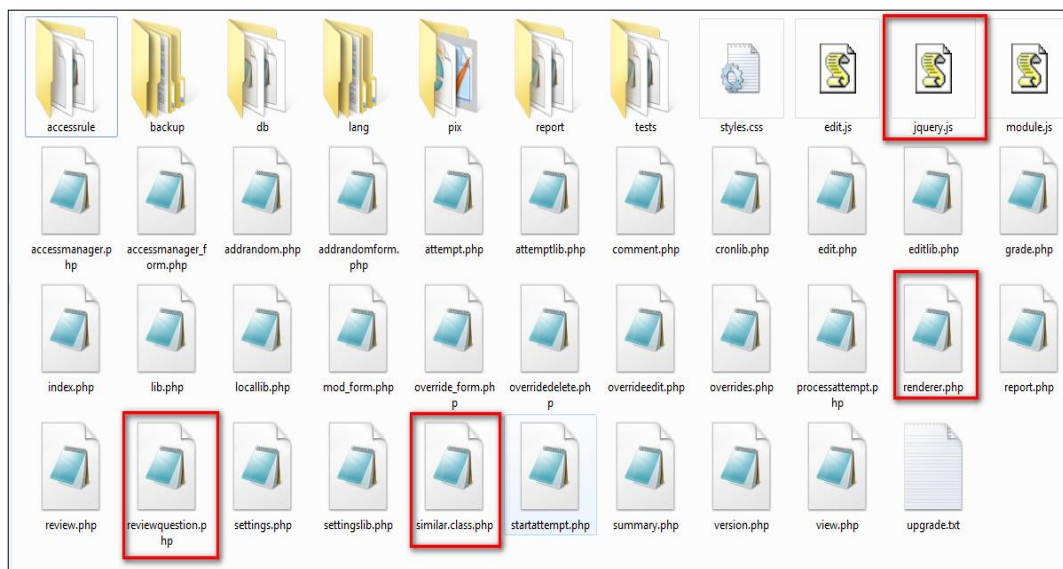
Figuer 5.3 The Contents of the File Quiz

The following is the changes that is applied on file renderer.php and reviewquestion.php.

A- renderer.php file located in C: \ xampp \ htdocs \ moodle \ moodle \ mod \ quiz path, the changes is applied as follows:

1 - Write Code To bring information from the database.

   $qData = $DB->get_record('quiz_attempts', array('id'=>$attemptid));

2 - Writing code to get student's answer.

   $**answer** = **preg_match('#qtype_essay_response.*?\">(.*?)</div>#s', $output, $val);**

3 - Writing code for the first and last name of the student's and the answers by analyzing the student 's user name who has completed the exam.

```php
$uData = $DB->get_records_sql("SELECT u.id, qa.responsesummary,
u.firstname, u.lastname FROM {quiz_attempts} quiza LEFT JOIN
{question_attempts} qa ON quiza.uniqueid=qa.questionusageid LEFT
JOIN {user} u ON quiza.userid=u.id WHERE quiza.userid!= '".$qData-
>userid."' AND quiza.quiz='".$qData->quiz."'");

$output .= '<div class="histroy" style="width:1000px;margin:auto;">
```

4 - write Code to find a similarity ratio.

```php
<h3>Similarity<select id="simRatio">';

for($i=0; $i<=90; $i+=10) $output .= '<option value="'.$i.'">>
'.$i.'%</option>';

$output .= '</select></h3>
```

5 - Write Code To create table. The table is created with four columns, the first and second columns represents the student's first and last name respectivly, the third one represents the answer of student and the last one for the similarity of the words and symbols.

```php
<table class="generaltable"><thead><tr>

<th class="header" style="width:150px">First Name</th>

<th class="header" style="width:150px">Last Name</th>

<th class="header" style="width:300px">Response</th><th
class="header" style="width:300px">Similarity</th>

</tr></thead> <tbody id="ratioList">';$i=0;

foreach($uData as $data){$i++;

if(trim($data->responsesummary)=="") continue;

$sm->resultRatio=0;

$sm->resultText='';
```

```
$sm->text1 = $val[1];

$sm->text2 = trim($data->responsesummary);

$sm->init();

$output .= '<tr data-ratio="'.$sm->resultRatio.'">

<td style="width:150px">'.$data->firstname.'</td>

<td style="width:150px">'.$data->lastname.'</td>

<td style="width:300px">'.trim($data->responsesummary).'</td>

<td style="width:300px">'.$sm->resultText.'</td>

</tr>';   }

$output .= '</tbody>

</table></div></div>';
```

6 - Adding jquery.js function in the path C: \ xampp \ htdocs \ moodle \ moodle \ mod \ quiz.It is one of the special functions in javascript. In order to control the rate of the similarity, it depends on the highest percentage of similarity between the words and symbols.

```
$output .= '<script type="text/javascript" language="javascript" src="jquery.js"></script>';

$output .= '<script>

$("#simRatio").change(function(){

$("#ratioList>tr").each(function(index){

if($("#simRatio").val()!=0  &&  $("#simRatio").val() >= $(this).attr("data-ratio")) $(this).slideUp(500);

else $(this).slideDown(500);});});</script>';
```

B- reviewquestion.php file is located in the path C: \ xampp \ htdocs \ moodle \ moodle \ mod \ quiz, the changes is applied as follows

1 - Adding code that calls similar.class.

require_once('similar.class.php');

$sm = new findSimilar;

CHAPTER 6

TESTING

This part demonstrates in which way all program visuals can be benefited by an instructor. The succession and shown results indicated that all basic program visuals are performing normally. Firstly, the instructor opens the Moodle Website, as an instance, all schools or other colleges benefit from it to discover plagiarism in the program. Then he/she encounters the home page of Moodle that includes transferable items like the names of the courses or the organization and the definition of the web page. To see the program, the instructor writes his/her own ID and password to sign in as shown in Figure 6.1.
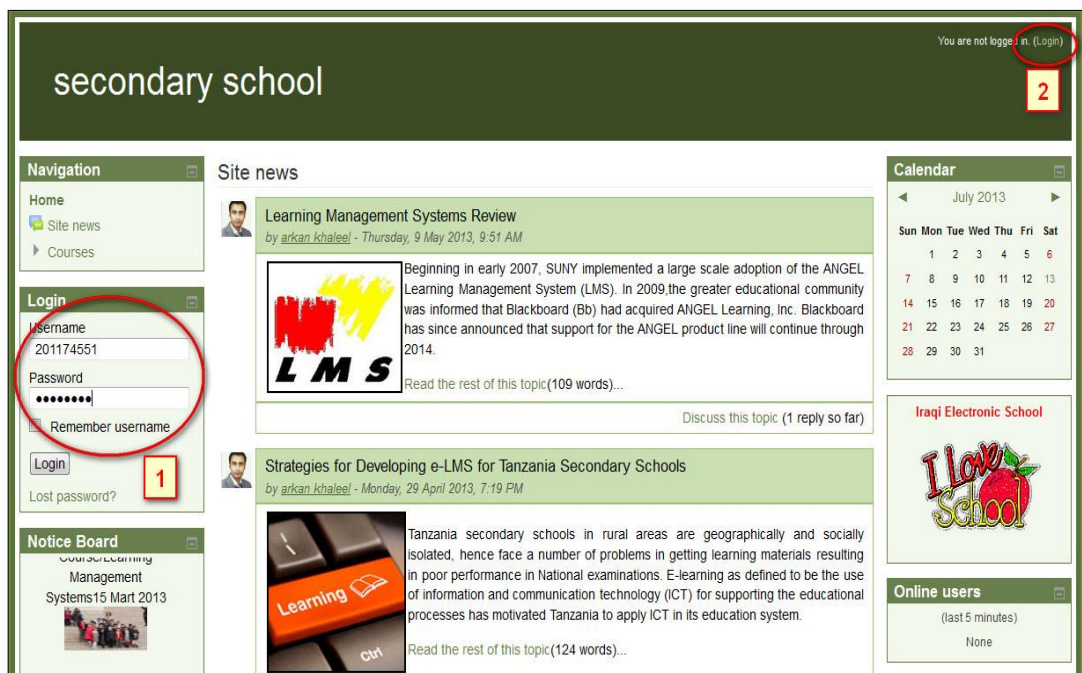
Figure 6.1 Moodle Home Page

Altfdernatively user may click on the login button on the top of the page, he / she will be redirected to the login page; in this page, the teacher should enter his/her own username and password correctly. as shown in the Figure 6.2.
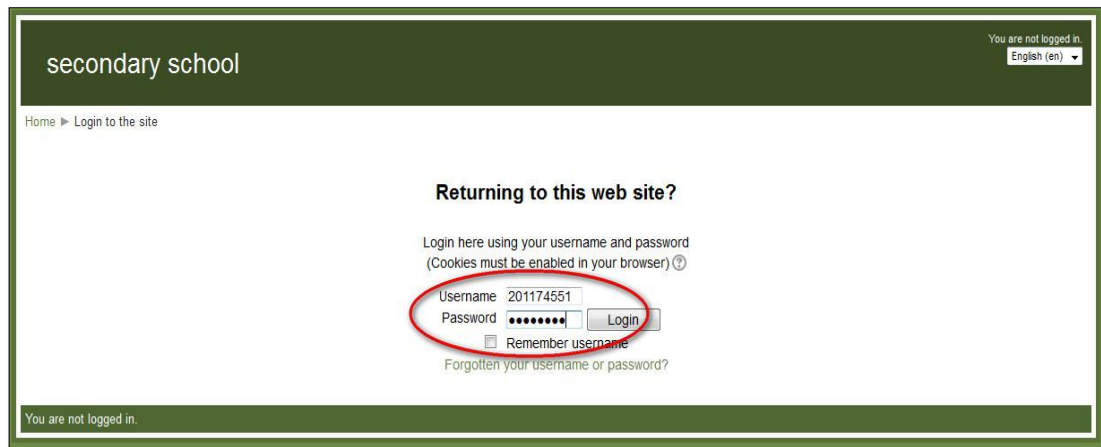
Figure 6.2 Login to the Site

When the username and password are entered correctly, Moodle will redirect teacher to the main page which contains the names of the courses that he/she is teaching, as shown in Figure 6.3.
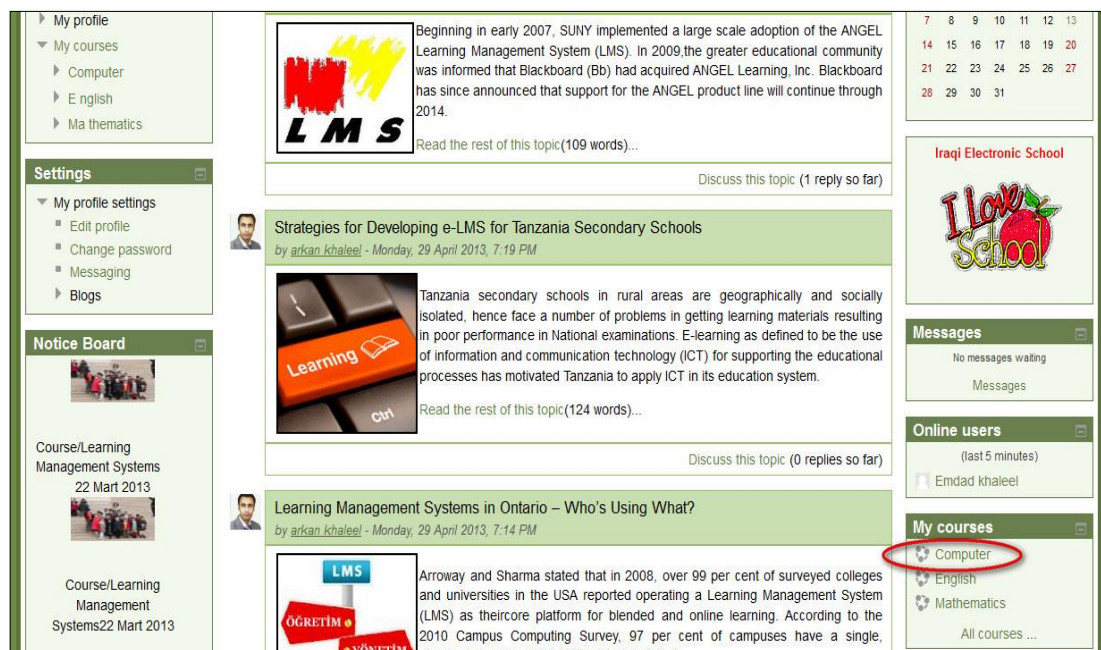
Figure 6.3 Select a Course in the Main Page

When the teacher clicks on a course name, he/she will be redirected to the main page of the selected course. As shown in Figure 6.4.
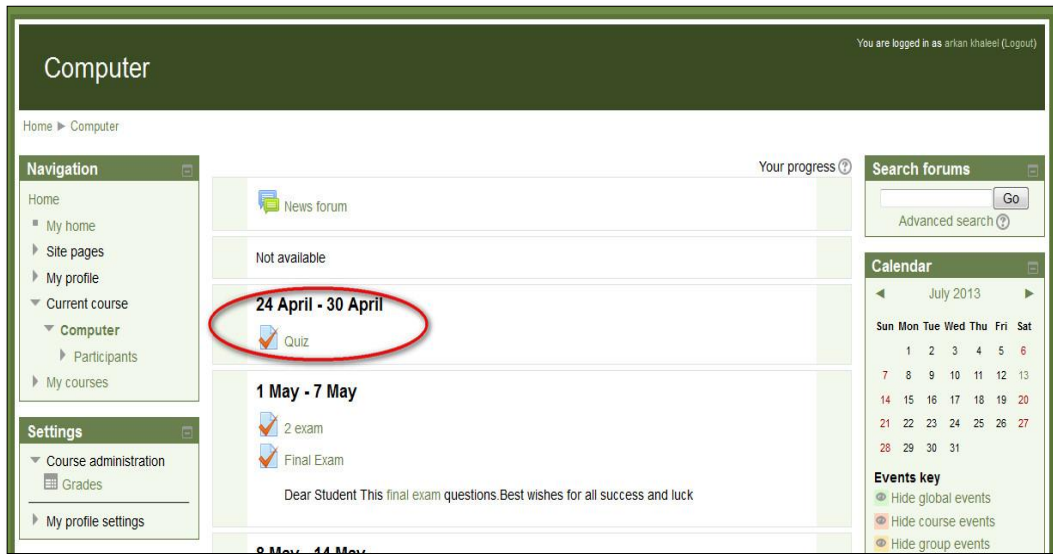
Figure 6.4 Select an Quize

After clicking on the name of the selected course, the main page of that course will be displayed. This page contains course name, course participants, course quiz and other activities.

Then the teacher will be able to click on the name of the available quiz (if any). As shown in Figure 6.4.

The quiz page contains quiz description, number of students who submitted quiz and the link to preview quiz. In order to view the submitted quiz, the teacher will be able to click on the link (Attempts) as shown in Figure 6.5.
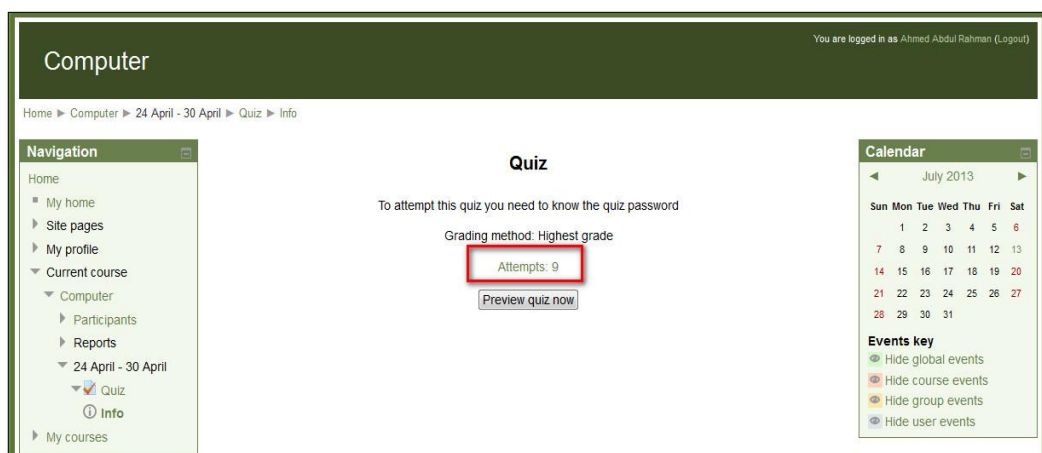
Figure 6.5 Submitted Quiz

After clicking on the (Attempts), the teacher will be redirected to the quiz page that contains submitted quiz, this page contains the names of students who submitted the quiz and some other information. As shown in Figure 6.6.

Figure 6.6 Student Information Page

When the (Requires grading) is clicked as show in Figure 6.7, the developed plagiarism detection module will be trigered and the report will be generated so that the teacher will be able to view and detect plagiarism.

Figure 6.7 Detect Plagiarism Page

Plagiarism detection page (Figure 6.7) consists of a table, which has four columns, the first and second columns are contains student's first and last name respectivly, the third column contains the student's answer and the last one for the similarity of the words and symbols. In addition there is a drop-down list on the right top of the table used to control the ratio of similarity.

After ratios of similarities is displayed to the teacher, system also enables the instructors to write a review or grade about student's submission currently being inspected. Then the user has to click on **make comment or override mark** link as shown in Figure 6.8.
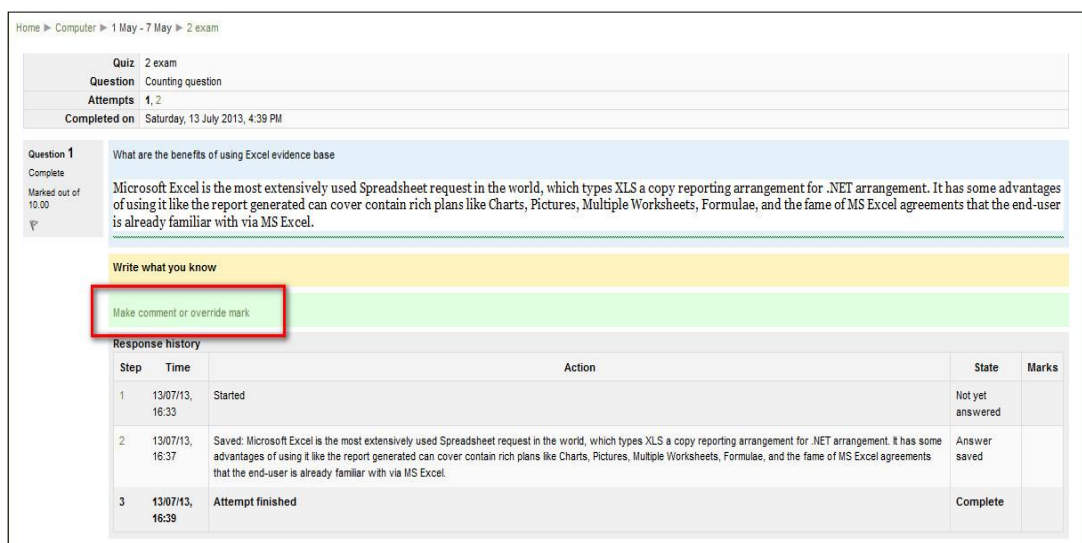

Figure 6.8 Give the Degree and Write a Review Page

After clicking on **make comment or override mark** link it will move to another page which contaians boxes to put the degree and comment and then click on Save button. As shown in Figure 6.9.

Figure 6.9 Give the Degree and Write a Review Page

After writing a degree and clicking on Save button, the student's information page will be displayed as shown in (Figure 6.6). We notice that the paragraph (Requires grading) will be changed to a degree that has given by the teacher, as seen in Figure 6.10.

Figure 6.10 Give the Degree and Write a Review Page

CHAPTER 7

CONCLUSION

In this study we aimed to enhance features of the Moodle. In this sense, plagiarism detection was developed and integrated into Moodle system. The results and error-free application shows that the work of proposed plagiarism system is compatible with Moodle system. An added feature with the Moodle lets the instructors to find out the plagiarism between student's replies. In this study our main goal was to demonstrate the applicability of plagiarism detection system in Moodle environment. Therefore performance of plagiarism detection algorithm in terms of accuracy and error write are out of scope of this study and left as future work.

7.1 Conclusions

1. The program updates Moodle.

2. The plagiarism detection tool of the program depends on a widely known algorithm in this area which is named AC.

3. The program matches between student's replies and provides similarity measure among them.

4. The standing drop-down list has the ability to manage for the instructor which can find the likeness in the replies and the rate that cannot be seen.

5. Presently, the program can work with text files. It also can work with. php, C, C++, java and another languages.

7.2 Future Work

These are a probable guidance for future on the program:

1. Matching between works including wrapped files**.**
2. Matching the student's work with other documents on the internet.
3. Modify more detailed ways to the program to discover some expanded cheating ways.
4. Benefiting from another algorithm.

REFERENCES

[1] **COLE, J., FOSTER, H.,** (2008) "Using Moodle", O'Reilly Media Inc. Press, 2nd Edition, Gravenstein Highway North, Sebastopol, the United States of America

[2] **PARR, J. M., & FUNG, I.** (2006). A review of the literature on computer-assisted learning, particu larly Integrated Learning Systems, and outcomes with respect to literacy and numeracy. available at http://www.minedu.govt.nz/index. cfm?layout=document&documentid=5 499&indexid=6920&indexparentid=1024, visited in May 2013

[3] **BROWN, M.** (2004). Evaluating the Evidence: e-learning Policy, Research and Practice. Computers in New Zealand Schools, 16(2), 3-7

[4] **KORTECAMP, K., & CRONINGER, W. R**. (1996). Addressing Barriers to Technology Diffusion. Journal of Information Technology for Teacher Education, 5 (1/2), 71-82

[5] **TEARLE, P.** (2003). ICT implementation: What makes the difference. British Journal of Educational Technology, 34(5), 567-583

[6] **HALL, J.** (2002). Assessing Learning Management Systems. Chief Learning Officer Magazine. MediaTec Publishing, available at http://clomedia.com/articles/view/91, visited in April 2013.

[7] **PAULSEN, M.** (2003). Experiences with Learning Management systems in 113 European Institutions. Educational Technology & Society, 6(4), 134- 148

[8] **ALTUN, A., GULBAHAR, Y., & MADRAN, O**. (2008). Use of a Content Management System For Blended Learning: Perceptions of Pre-Service Teacher. Turkish Online Journal of Distance Education, 9(4), 11

[9] **CHANG, C. L.** (2008). Faculty Perceptions and Utilization of a Learning Management System in Higher Education. Unpublished doctoral dissertation, Ohio University, USA

[10] **FALVO, D., & JOHNSON, B.** (2007). The Use of Learning Management Systems in the United States. TechTrends, 51(2), 40-45. available at http://dx.doi.org/10.1007/s11528-007-0025-9, visited in May 2013

[11] **JANG, S., J.,** (2008) "Innovations in science teacher edu cation: Effects of integrating technology and team-teaching strategies", Computers & Education vol. 51, no. 2, pp. 646–659,

[12] **RENGARAJAN, R.** (2001) LCMS and LMS – Taking advantage of tight integration

[13] **GARROTE, J., R.; PETTERSSON, T.** (2011) The use of learning management systems: A Longitudinal Case Study. available at http://eleed.campussource.de/archive/8/3145, visited in May 2013

[14] **OAKES, K.** (2002). E-learning: LCMS, LMS— They're not just acronyms but powerful systems for learning. *Training & Development, 56*(3), 73-75

[15] **SAUL, C.,** (2005) Course Management Systems Versus Learning Management Systems, available at www.learningcircuits.org, visited in May 2013

[16] **CARLINER, S.** (2005). Course management systems versus learning management systems. Learning Circuits. available at http://www.learningcircuits. org/2005/nov2005/carliner.Htm, visited in May 2013

[17] **RYAN, K. E.,** (2009) A field guide to Learning Management Systems, Learning Circuits. available at http://www.astd.org/lc, visited in May 2013

[18] **KLUDGE & RILEY,** (2008) Teaching in virtual worlds: opportunities and Challenges, Issues in Informing Science and Information Technology Volume 5. available at http://informingscience.org/, visited in May 2013

[19] **BARCHINO, R., GUTIÉRREZ, J., M., & OTÓN, S.,** (2005) An Example Of Learning Management System, Computer Science Department, University of Alcala

[20] **KLJUN, M., VICIC, J., KAVSEK, B., & KAVCIC, A.,** (2007) Evaluating Comparisons and Evaluations of Learning Management Systems, Proceedings of the ITI 2007 29th Int. Conf. on Information Technology Interfaces

[21] **NEW ZEALAND MİNİSTRY OF EDUCATİON (MOE),** (2011). *Learning* Management Systems.available at http://www.minedu.govt.nz/NZEducation/EducationPolicies/Schools/Initiati ves/ManagedLearningEnvironments/LearningMgmtSys.aspx, visited in Jun 2013

[22] **FALVO, D., A., & JOHNSON, B., F.,** (2007) The Use of Learning Management Systems in the United States, TechTrends, Volume 51, Number 2

[23] **AHMAD, N., & AL-KHANJARI, Z.,** (2011) Effect of Moodle on learning, International Journal of Digital Information and Wireless Communications (IJDIWC) 1(4): 746-752 The Society of Digital Information and Wireless Communications, (ISSN 2225-658X)

[24] **BROWN, A., & JOHNSON, J.,** (2007) Five Advantages of Using a Learning Management System, Microburst Learning. available at www.microburstlearning.com, visited in Jun 2013

[25] **SIGCSE** .(1999) The proceedings of the thirtieth SIGCSE technical symposium on Computer science education, Newyork, USA

[26] **NADELSON, S.** (2007). Academic misconduct by university students: Faculty perceptions and responses. Plagiary, 2(2), 1-10

[27] **BİNGFENG, Pİ., SHUNKAİ, FU., WEİLEİ, W., HAN, S.,** (2009) Proceedings of the Second Symposium International Computer Science and Computational Technology(ISCSCT '09), Huangshan, P. R. China, 26-28,Dec. 2009, pp. 020-025

[28] **LARKHAM P., & MANNS S.** (2002). Plagiarism and its treatment in higher education. Journal of Further and Higher Education, 26(4), 339-349

[29] **CLOUGH, P.** (2000). Plagiarism in natural and programming languages: An overview of current tools and technologies. Internal Report CS-00-05, University of Sheffield, UK

[30] **LANCASTER, T., & CULWİN, F.** (2004). Using freely available tools to produce a partially automated plagiarism detection process. Proceedings of the 21st ASCILITE Conference, Perth, Australia

[31] **FOSTER, A.** (2002). Plagiarism-detection tool creates legal quandary. Chronicle of Higher Education, May 17th, 2002, Section: Information Technology, A37

[32] **DOMENİCO CARLONE, A., RİZZELLO, F., ROBERT THOMSEN, D. M., AKBAR, H.,** (2011). Proceedings of the International MultiConference on Engineers and Computer Scientists

[33] **TRESNAWATI, D., SYAICHU, A., KUSPRIYANTO, R.,** (2011) "Plagiarism Detection System Design for Programming Assignment in Virtual Classroom Base on Moodle", available at http://ac.els-cdn.com/S1877042812052986/1-s2.0-S1877042812052986-main.pdf?_tid=1003e96e-eed5-11e2-96cf-00000aacb362&acdnat=1374061144_87042a6350ab4ec4c72a345055db78f3, visited July 2013

[34] **CHRISTENSSON, P.,** (2009) "Token Definition", available at http://www.techterms.org/definition/token, visited July 2013

[35]   **BOWYER, W., K., LAWRENCE, O., H.,** (2012) "Experience Using "MOSS" to Detect Cheating On Programming Assignments", available at http://fie2012.org/sites/fie2012.org/history/fie99/papers/1110.pdf, visited June 2013

[36] **JPAG**, (2000) "Jplag: Finding Plagiarisms among a Set of Programs", Fakultat für Informatik Universitat, Karlsruhe, Germany, available at http://page.mi.fu-berlin.de/prechelt/Biblio/jplagTR.pdf, visited July 2013

[37] **CHEN, X., MING, L., FRANCIA, B., MCKINNON, B., SEKER, A.,** (2004) "Shared Information and Program Plagiarism Detection", IEEE Transactions on Information Theory, Vol. 50, No. 7, available at http://www1.spms.ntu.edu.sg/~chenxin/paper/IT04a.pdf, visited May 2013

[38] **FREIRE, M., CEBRIAN, M., DEL ROSAL, E.,** (2011) "Uncovering Plagiarism Network", available at http://arxiv.org/pdf/cs/0703136v7.pdf, visited May 2013

[39] **FREIRE, M.,** (2008) "Visualizing Program Similarity in the AC Plagiarism DetectionSystem",availableat http://www.researchgate.net/publication/220944853_Visualizing_program_similarity_in_the_Ac_plagiarism_detection_system, visited July 2013

[40]   **MOORE, J., CHURCHWARD, M.,** (2010) "Moodle 1.9 Extension Development" Packt Publishing Ltd, Lincoln, UK, ISBN: 978-1-847194-24-4, 285 pages

APPENDIX A

SIMILAR.CLASS.PHP (CODE)

```php
<?php
class findSimilar{
        public $strDel = array('the', 'The');
        public $charLimit = 0;
        public $text1 = '';
        public $text2 = '';
        public $spcl1 = '';
        public $spcl2 = '';
        public $textArray1 = array();
        public $textCount1 = array();
        public $text1Array = array();
        public $text2Array = array();
        public $resultText = '';
        public $resultRatio = 0;
        function init(){
                $t1 = $this->text1;
                $t2 = $this->text2;
                $this->text1 = $this->textToArray($this->text1);
                $this->text2 = $this->textToArray($this->text2);
                $this->text1Array = $this->repeatCount($this->text1);
                $this->text2Array = $this->repeatCount($this->text2);
                $this->similarCount($this->text1Array, $this->text2Array, false);
                $this->spcl1 = $this->textToArray($t1, false);
                $this->spcl2 = $this->textToArray($t2, false);
```

```php
        $this->spcl1Array = $this->repeatCount($this->spcl1);

        $this->spcl2Array = $this->repeatCount($this->spcl2);

        $this->similarCount($this->spcl1Array, $this->spcl2Array);

    }

    function textToArray($text, $opp=true){

        $text = preg_replace('/\s+/', ' ', addslashes($text));

        $text = preg_replace('/['.($opp?'^':'').'a-zA-Z0-9öçşığüÖÇŞİĞÜ ]/', ' ',
$text);

        $text = str_replace( $this->strDel, '', $text );

        $text = array_filter( explode(' ', $text) );

        sort( $text );

        return $text;

    }

    function repeatCount($array){

        $textArray = array('text'=>array(), 'count'=>array());

        foreach($array as $value){

            if(strlen($value)<$this->charLimit) continue;

            if(!in_array($value, $textArray['text'])){

                $textArray['text'][] = $value;

                $textArray['count'][] = 1;

            }else{

                $textArray['count'][array_search($value,
$textArray['text'])]++;

            }

        }

        return $textArray;

    }

    function similarCount($textArray1, $textArray2, $textOrChar=true){

        $text1Count = count($textArray1['text']);

        $text2Count = count($textArray2['text']);

        if($text1Count>$text2Count){
```

```php
            $array1 = $textArray1;

            $array2 = $textArray2;

        }else{

            $array1 = $textArray2;

            $array2 = $textArray1;

        }

        $allTextArray = array_merge($array1['text'], $array2['text']);

        $allTextArray=array_unique($allTextArray);

        $allTextCount = count($allTextArray);

        $similar = 0;

        $smlr = array();

        foreach($array1['text'] as $key=>$val){

            if(in_array($val, $array2['text'])){// &&
abs($array1['count'][array_search($val, $array1['text'])] -
$array2['count'][array_search($val, $array2['text'])])<=1){

                $smlr[] = $val;

                $similar++;

            }

        }

        $rat = number_format(($similar/$allTextCount)*100,0);

        if($rat>$this->resultRatio) $this->resultRatio =
number_format(($similar/$allTextCount)*100,0);

        $this->resultText .= '<p><strong>Similar
'.($textOrChar?'Char':'Text').'(s):</strong> '.($similar>0?implode(' | ',
$smlr):'<u><i>none</i></u>').'<br /><strong>Similar Ratio:</strong>
'.number_format(($similar/$allTextCount)*100,2).'% : '.$similar.' in '.$allTextCount.'
'.($textOrChar?'char':'letter').'(s)</p>';

    }
}
?>
```

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: SHAKR, Arkan

Nationality: Iraqi (IQ)

Date and Place of Birth: 05. May. 1985. Kirkuk.

Marital Status: Married.

Phone: +9647701268742

email: arkaniraql@hotmail.com

EDUCATION

Degree	Institution	Year of Graduation
BS	Mousl Univ. Bachelor of Education in Computers Sciences	**2007**
High School	El Walid School	**2003**

WORK EXPERIENCE

Year	Place	Enrollment
2007 – Present	Teacher in Secondary School. Kirkuk	Computer teacher
2008	Technical Institute of Kirkuk. (Lecturer)	Computer teacher
2008	International Reliep and Development (**IRD**). (Lecturer)	Computer teacher
2011	International Reserch and Exchange Boards (**IREX**). (Lecturer)	Computer teacher

FOREIGN LANGUAGES

Arabic, English, Turkish