



COMPUTER BASED ASSESSMENT SYSTEM FOR EVALUATING SUBJECTIVE QUESTIONS

*Abraham Eseoghene Ewwiekpaefe and Yahya Umar Muhammad

DEPARTMENT OF COMPUTER SCIENCE, NIGERIAN DEFENCE ACADEMY, KADUNA, NIGERIA

*Corresponding author's email: aeevwiekpaefe@nda.edu.ng

ABSTRACT

With the increase in the number of students in our tertiary institutions coupled with the number of courses offered in the various universities, the management of examination processes has become more complex in terms of resources, time and manpower. This has resulted to the rapid adoption of a computerized means of conducting examinations instead of the traditional pen and paper. However, this computer-based technique currently used in our tertiary institutions in assessing students' examination handles only the multiple choice questions (MCQ). This paper therefore is aimed at designing and implementing a simplified algorithm that assesses subjective (descriptive) type questions. Keyword and synonym matching techniques were used as the method for creating the simplified algorithm for the assessment of the subjective answers. The algorithm was designed using the Unified Modeling Language. The application was built in the NetBeans IDE version 8.2 with Java as the main programming language. Spring Boot Framework with Thymeleaf View Engine was used as the Web development Framework. MySQL was used for the database while HTML, CSS and JavaScript were used for building the web page interface. For the synonyms, grammar and spell checking, Merriam-Webster Dictionary API and Language Tool API were used. A prototype system was consequently developed, tested and various results shown.

Keywords: Computer Based Assessment (CBA), Grammar, Keyword, Subjective Questions, Synonyms.

INTRODUCTION

For years especially in educational institutions, written examination has been used as a major means of evaluating one's knowledge or proficiency in a subject or skill. It is largely accepted that assessment determines the degree at which learning goals have been attained and to what extent educational institutions have provided the society with the needs that it requires (Shah, 2002). Qureshi (2015) declared that electronic assessment system is not a new idea in educational institutions as it has been used in the assessment of objective type questions for many years now. In Nigerian schools as reported by Agbonifo (2014), the approval of computer-based assessment systems is rapidly increasing among teachers and students.

While electronic assessment is vastly commended due to its usefulness and efficiency, however its constraint depends on the depth of knowledge that can be evaluated. It has been noted that the demand for the assessment of a comprehensive knowledge of a subject compelled the need for electronic examination systems that support free-text (subjective) questions such as descriptive type or definition type question, one word or one sentence answer, fill in the blanks etc. Wang *et al.* (2013) asserted that an electronic grading system that has

most of its questions as MCQs is found to be insufficient in assessing learning outcome as it cannot completely test the student's capability. For instance, a student can choose an answer at random when he or she does not actually comprehend the real meaning of the questions and still gets the correct answer.

Although the development of subjective assessment systems is not a novel phenomenon in the world, the algorithms and even the tools used in the development of such system could be very challenging to adopt. For example, Dharma *et al.* (2017) designed an automated evaluation for subjective answers called ApTeSa. Smart and systematic techniques were what the system used in evaluating its answers. The system however, was designed to work in a semi-automated mode or in full-automated mode making it a rather cumbersome process to adopt. Again, Alrehily *et al.* (2018) designed an e-examination system that could be used for subjective questions. The system finds a matching ratio for the keywords in the preferred (model) and student answers when assessing the descriptive answers. Semantic and text similarity were used in attaining the matching ratio. Preprocessing, keyword expansion, matching, and grading were the four modules used in the system. PHP and JavaScript languages were used to implement

the system interfaces, while MySQL was used in the system database implementation. The assessment algorithm for exam was implemented using python. This system again had a different structure from that applicable in Nigeria and handled both objective and subjective questions.

Some universities in Nigeria have adopted the Computer Based Assessment (CBA) approach in evaluating their students' performances. However, the existing CBA system adopted by these universities handle only multiple choice (also known as objective) questions used primarily for the first year or second year courses. Also, although these universities have been using e-examination for the objective questions, these schools still use the traditional means of assessment for the subjective examination where lecturers manually mark the answer script written by their students. This system is an error-prone process, as many lecturers are bound to award varied marks to identical answers. Marks awarded in this manual evaluation are subject, not only, to the contents of the answer script, but also to other factors such as the student's hand writing, the elegance in which the student is answering the question, and even the opinion of the one marking the script.

In view of this, the objective of this paper is to develop a simple framework (algorithm) based on keywords and

synonyms matching technique that is capable of evaluating descriptive or subjective questions.

MATERIALS AND METHOD

System Methodology

In this paper, keywords and synonyms matching technique was adopted for the new system. Thesaurus dictionary API provided synonyms of the keywords and also checked the grammar. A marking guide or scheme is decided based on the number of keywords match.

System Description

The new system is intended to accept and evaluate subjective or free-text answers by comparing the student's answer with the model answer stored on the system by the examiner. The standard answer is stored in the database along its keywords and marks. Each of the student's answer is evaluated by matching the keywords or its synonyms with the model answer. Sentence formation (grammar) of the answers is also checked. After the evaluation, the system grades the answer based on the answer's accuracy. The whole process is made up of four units which include Login unit, Information Extraction unit, Weighting unit and Score Generation unit as shown in Figure 1

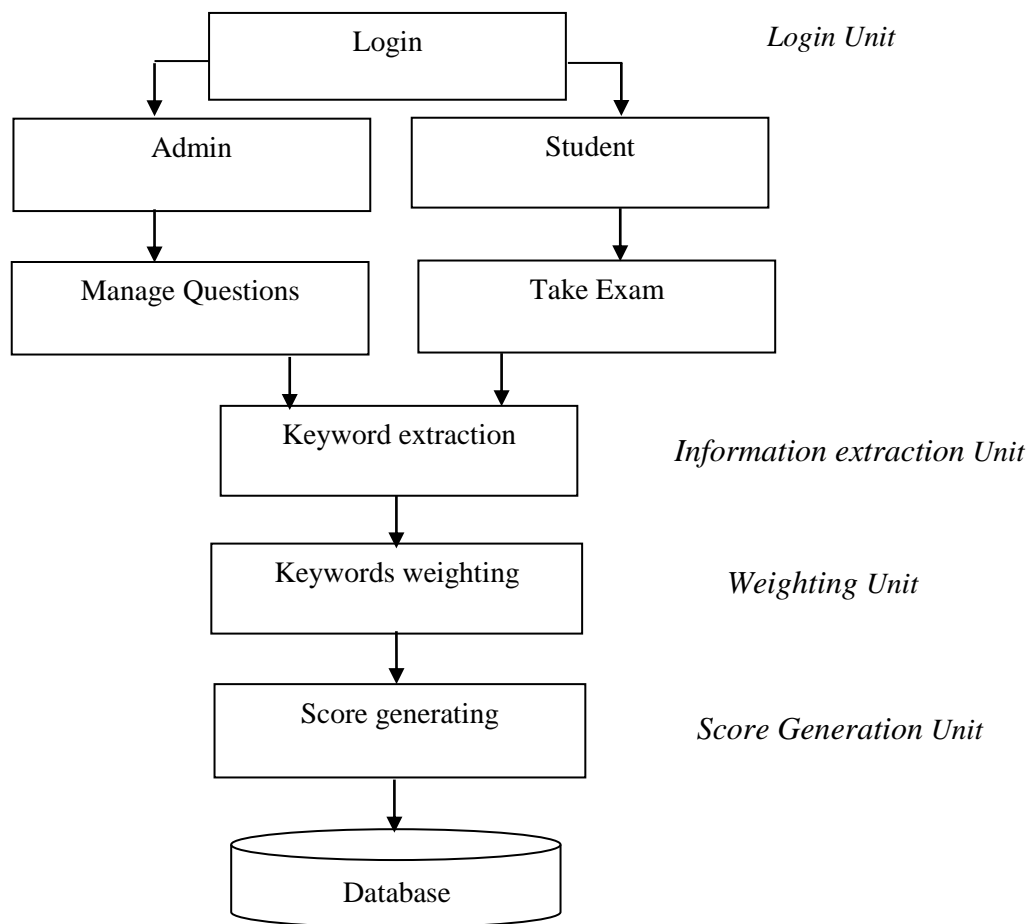


Figure 1: Flow of data in the System Unit

Login Unit: Login unit is where users of the system, i.e. admin and students, are authenticated in order for them to gain access to the system. Once users (admin and students) are authenticated, then they can go on with their respective activities.

Information Extraction Unit: This is where keywords on the model answer as well as the answer submitted by the students are extracted. Only the relevant words (keywords) in the answer to the question are extracted. Repeated keywords are discarded as the system returns unique keywords only. The extraction is based on the concept that keyword in the model answer of every question should be entered separately while being stored in the database.

Weighting Unit: Under the weighting unit, the system works on two aspects: keywords (or their synonyms) and sentence formation (grammar). For the keyword, the system compares the keywords found in the student's answer with the model answer and allocates marks depending on the number of the matched keywords or the synonyms present in the student's answer. For the grammar, students must provide the keywords in an appropriate sentence formation to earn highest marks. No matter the number of keywords contained in the student's answer, sentences with incorrect grammar will earn fewer marks as certain percentages are assigned for the sentence formation. It is only when the sentences are correctly formed that marks for grammar are awarded.

Score Generation Unit: The student's answer final score is attained by adding up the marks from the section of keywords and that of grammar. According to significance, keyword has 75% priority over grammar which has 25%. The marks for sentence formation are awarded only if keywords are matched. Accordingly, marks for sentence formation are not added when a given student's answer has no keywords in it.

System Marking Algorithm

After a student successfully submits his/her answer, the following steps are taken by the system to mark and assign scores to the answer of each question:

1. Start
2. Get the student's answer
3. Find in the student's answer each keyword.
4. If keywords found, award mark based on the marking scheme, go to step 8.
5. If keywords not found, then check for the keywords synonyms from the dictionary.
6. If synonyms found, then award mark for it based on the marking scheme

7. If synonyms not found, then award zero mark and go to step 12
8. Check if at least half of the keywords is found.
9. If half of keyword found, check the grammar of the answer from Language Tool.
10. If the grammar is correct, award mark for grammar base on the marking scheme, if not then award zero mark, go to step 12.
11. If half of the keyword not found, award zero marks for grammar, go to step 12
12. Calculate the total mark scored for the question which is the summation of keyword marks and the grammar mark.
13. Calculate the total mark of the examination which is the summation of each answer score.
14. Stop

System Marking Score Algorithm

This further explain in detail how marks are assign to each answer based on the keywords and sentence formation (grammar).

- a. 75% of the Question's mark goes to Keyword Score (if mark of a given question is 10, the total keyword score is 7.5), while 25% of the Question's mark goes to Grammar Score (if mark of question is 10, the total grammar mark is 2.5).
- b. The 75% (i.e. 7.5 marks of 10 marks) of Keyword Score is divided across the list of keywords of the question. If there are 5 keywords, then each keyword will have a score of 1.5marks ($7.5/5 = 1.5$).
- c. The Grammar Score is only awarded when the answer's sentence is grammatically correct (with correct spelling) and the keyword scored is more that 50% (at least half of the keyword is available).
- d. The total examination score is the total score of each answer all over the total mark of questions.

System Modeling using UML

The Unified Modeling Language (UML) is a standardized modeling language consisting of an integrated set of diagrams, developed to help system and software developers for specifying, visualizing, constructing, and documenting the artifacts of software systems. The following UML models were used in the designing of the new system.

System Use case

A use-case diagram defines a system's functional requirements in terms of use cases. Figure 2 illustrates the use case diagram for the new system

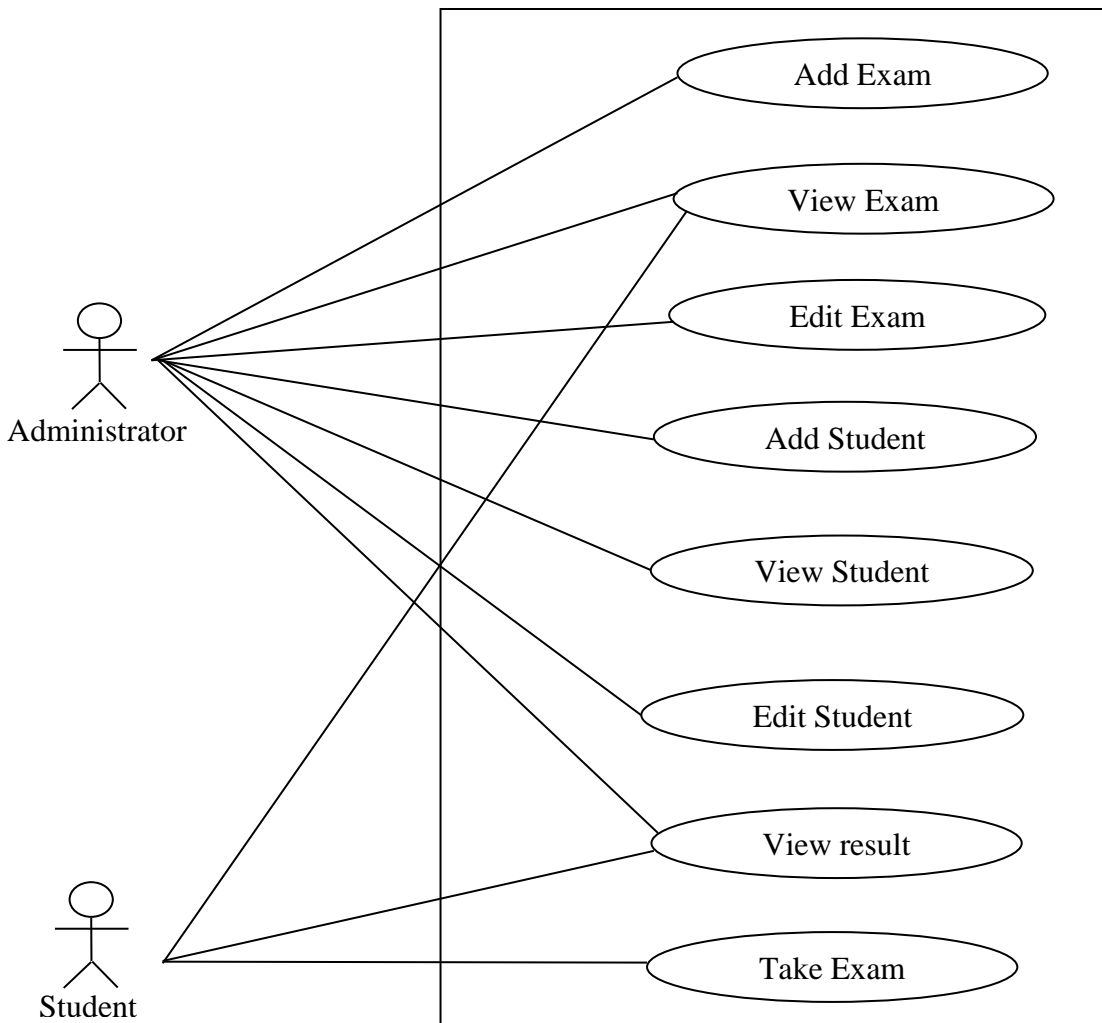


Figure 2: System Use Case Diagram

System Sequence Diagrams

The Sequence Diagram models the relationship of objects based on a time sequence. It shows how the objects interact with one another in a particular situation of a use case. Figure 3 and Figure 4 represents the System Sequence Diagrams for Admin and student module in the new system.

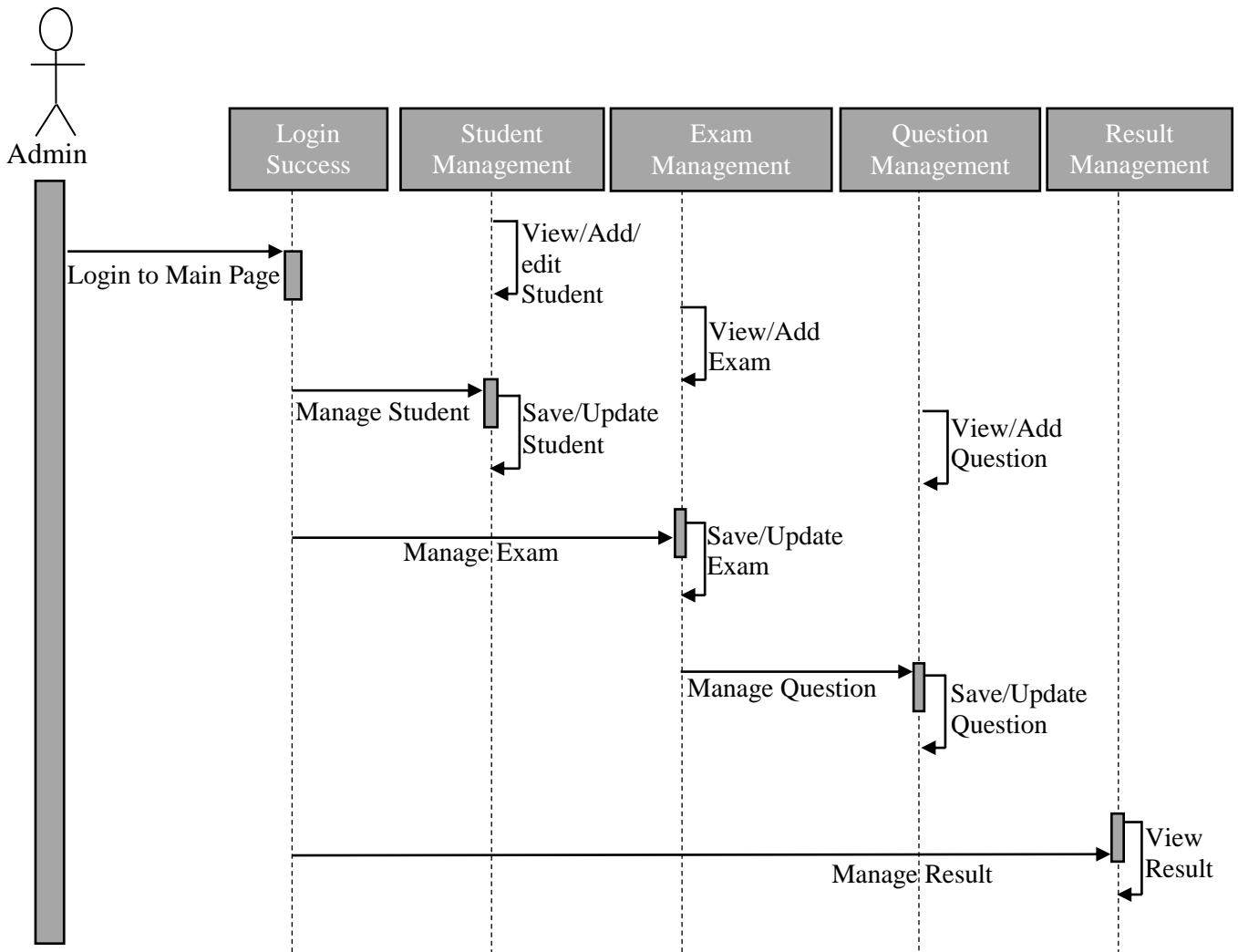


Figure 3: Sequence Diagram for Admin Module

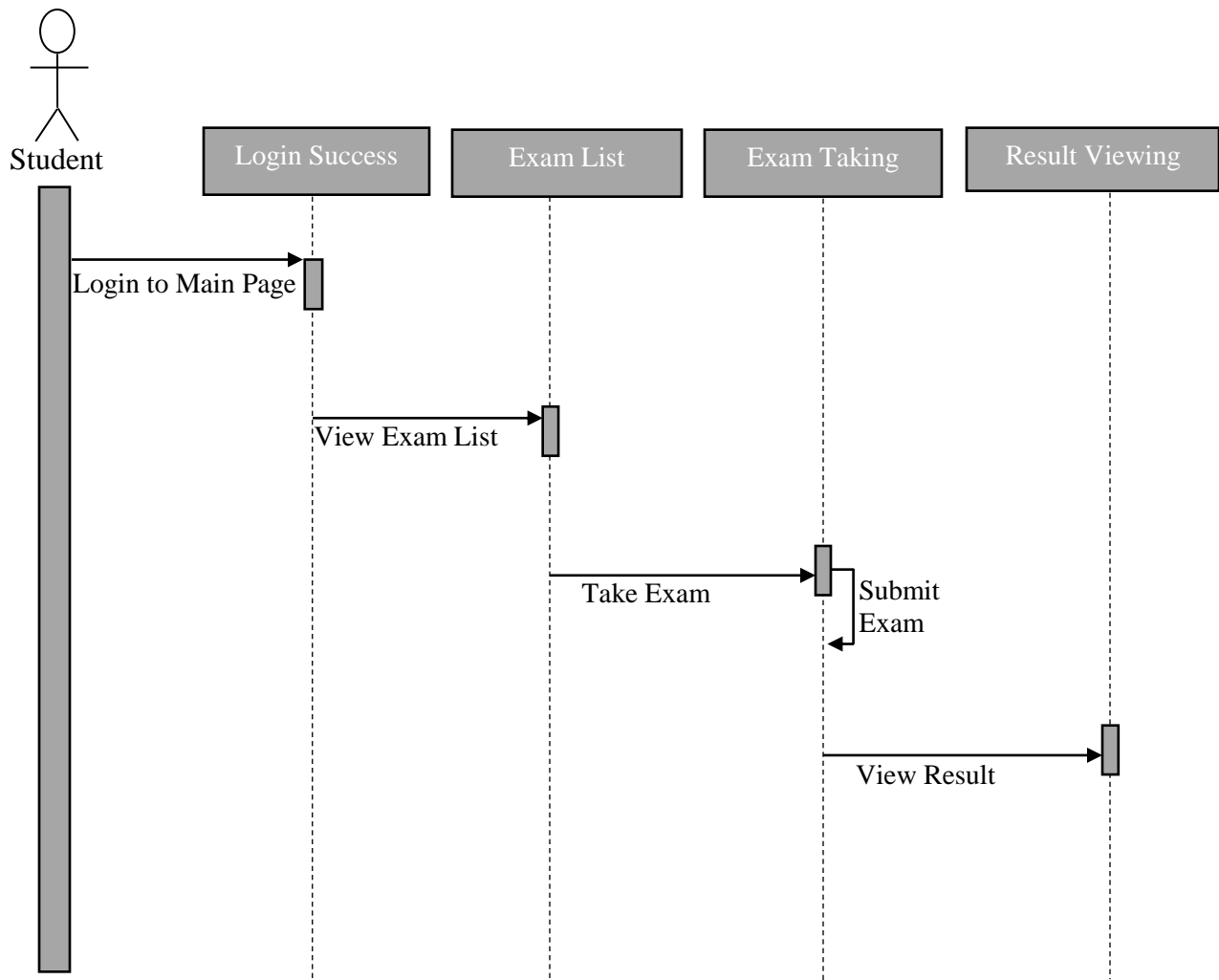


Figure 4: Sequence Diagram for Student Module

System Class Diagram

Class diagram describes the types of objects in the system and various kinds of static relationships which exist between them. Class diagram here describes the structure of the new system classes, their attributes, operations or methods, and the relationships among objects as shown in figure 5. The main classes of the proposed system are Admin, students, Exam, Questions, Answer, Keywords, and Result.

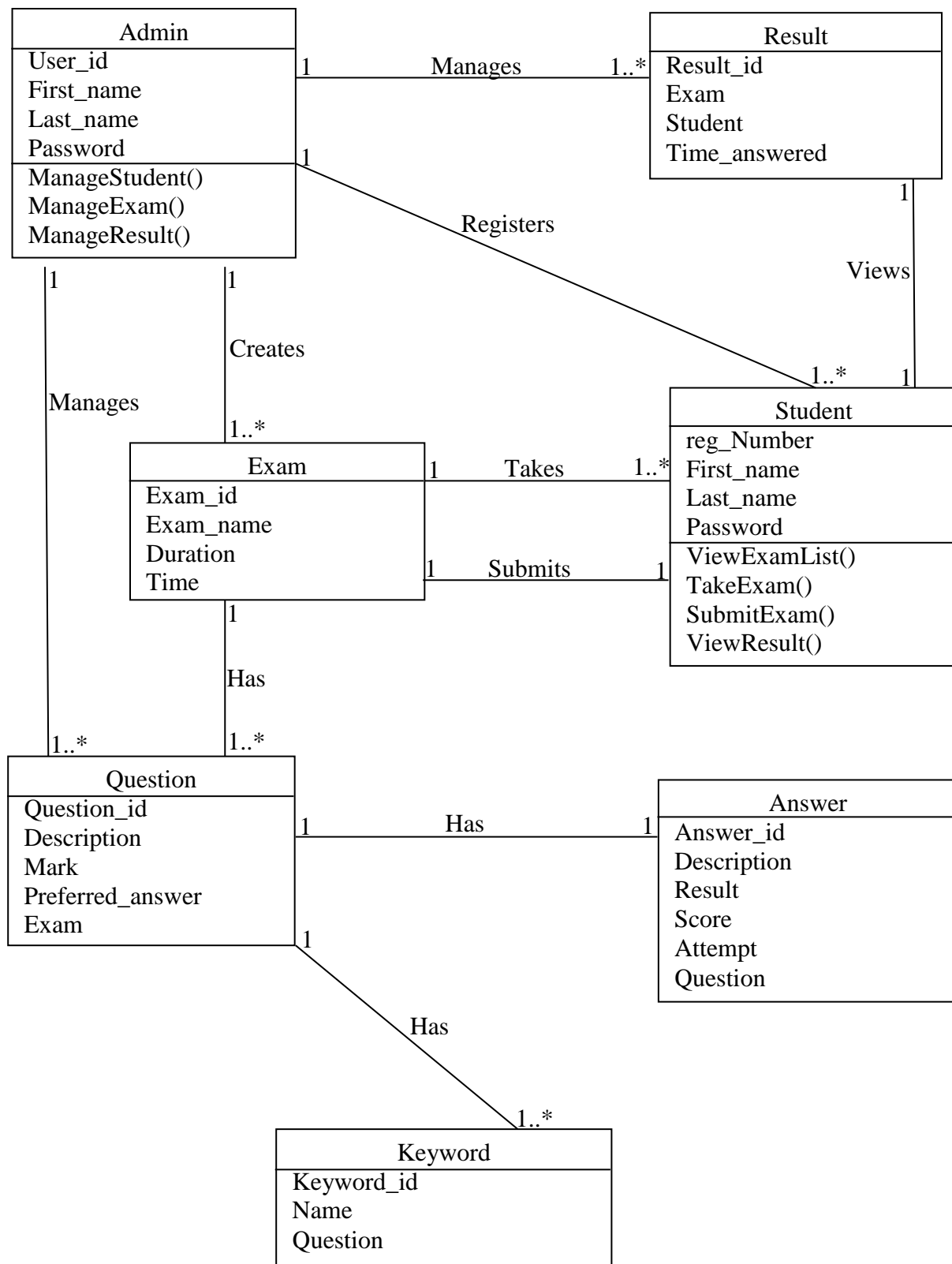


Figure 5: System Class Diagram

RESULTS AND DISCUSSION

System Implementation

The application was built in the NetBeans IDE version 8.2 and deployed on the built-in Tomcat server, which listens on port 8080. Java was chosen as the main programming language for the development of the new system. Spring Boot Framework with Thymeleaf View Engine was used as the Web development Framework. *Spring* is a popular Java application framework. Thymeleaf on the other hand, is a modern server-side Java template engine for both web and standalone environments. MySQL was used as a database tool as it is freely available open source Relational Database Management System (RDMS) that uses Structured Query Language (SQL). For an easy to use and interactive web pages (interface), Hyper-Text Markup Language (HTML), Cascading Style

Sheets (CSS) and JavaScript were used. For the synonyms, sentence grammar and spell checking, Merriam-Webster Dictionary API and Language Tool API were used. The results of designed and implemented system are discussed here.

Login Interface

The users of the system are the system administrator and the students. Each of the users gets access to the system by logging into the main (Login) page of the system. This is done by typing the address of the system on the address bar of either Mozilla Fire Fox, Google Chrome browser or any other web browser. Students log in using their registration or matric numbers as USERID and the password issued to them by the system admin as shown in Figure 6 and Figure 7 respectively.

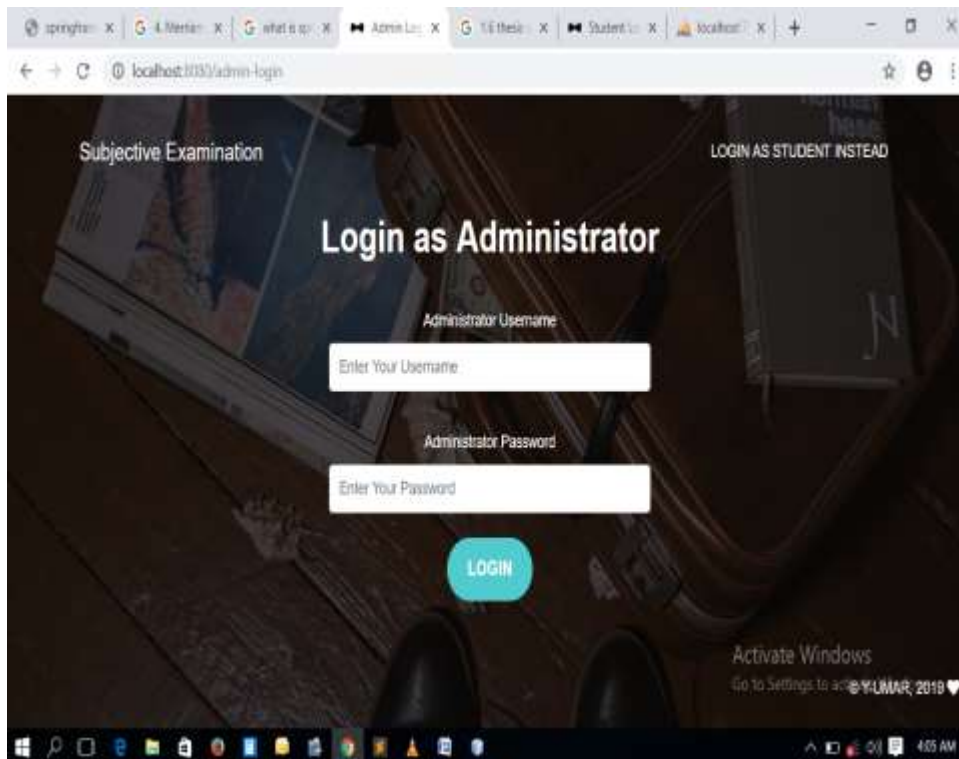


Figure 6: Admin Login Page

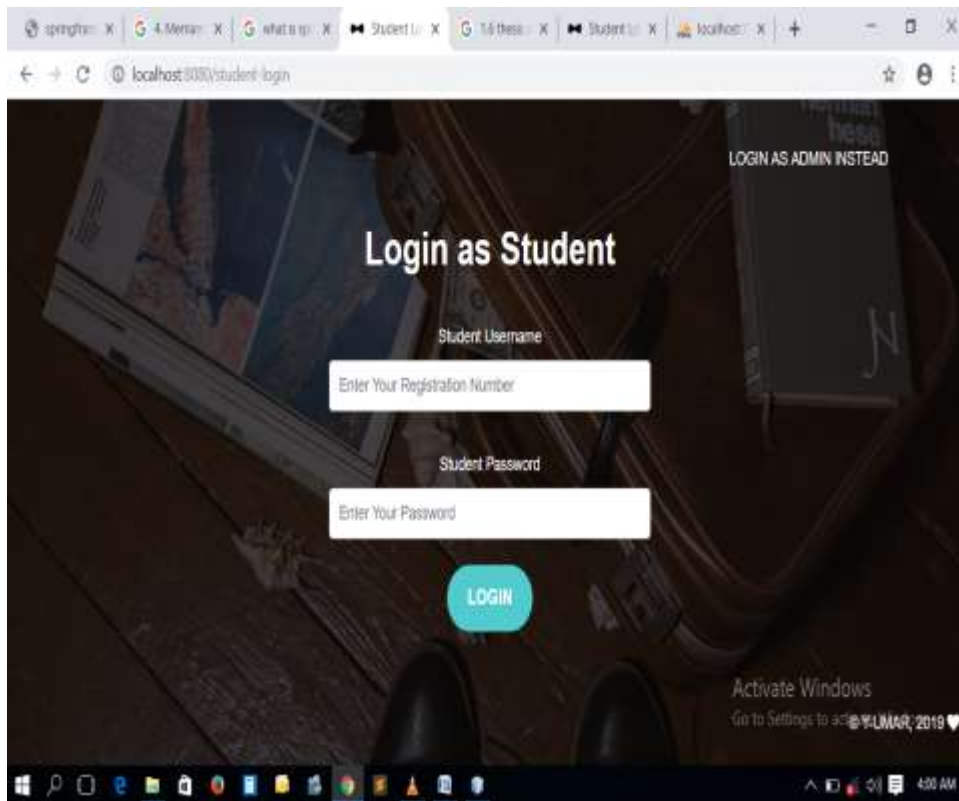


Figure 7: Student Login Page

The Students' Module

Students, who are the main users, can login into the system using their Registration Number and take the exam. Students on their module can basically do the following activities after login: view the list of Examinations, take an examination and view examination result as shown in Figures 8, 9 and 10.

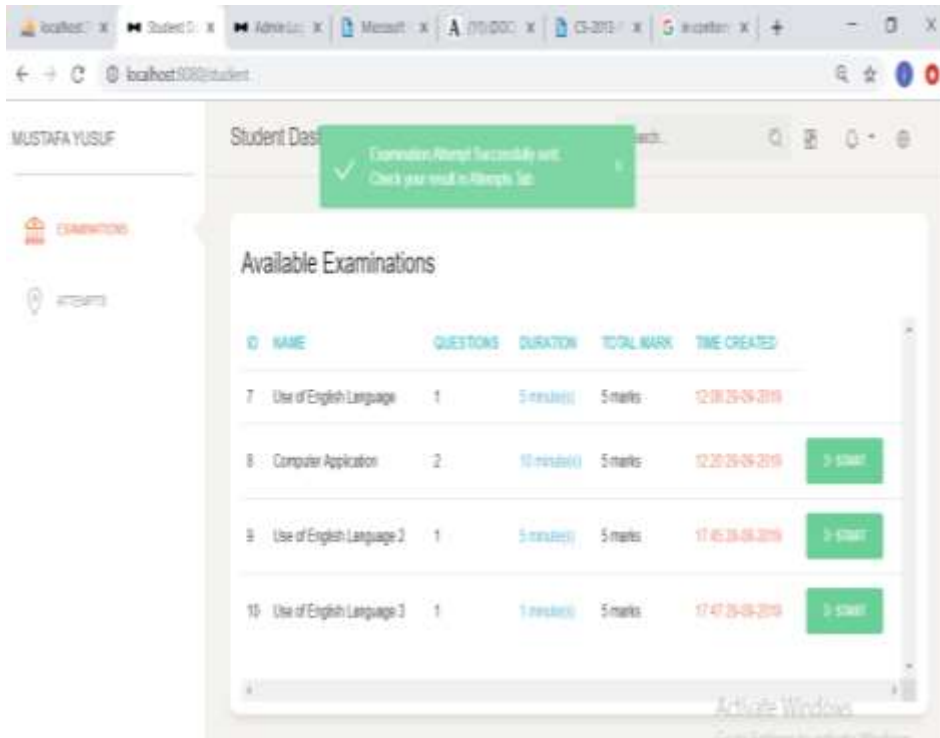


Figure 8: List of available examinations

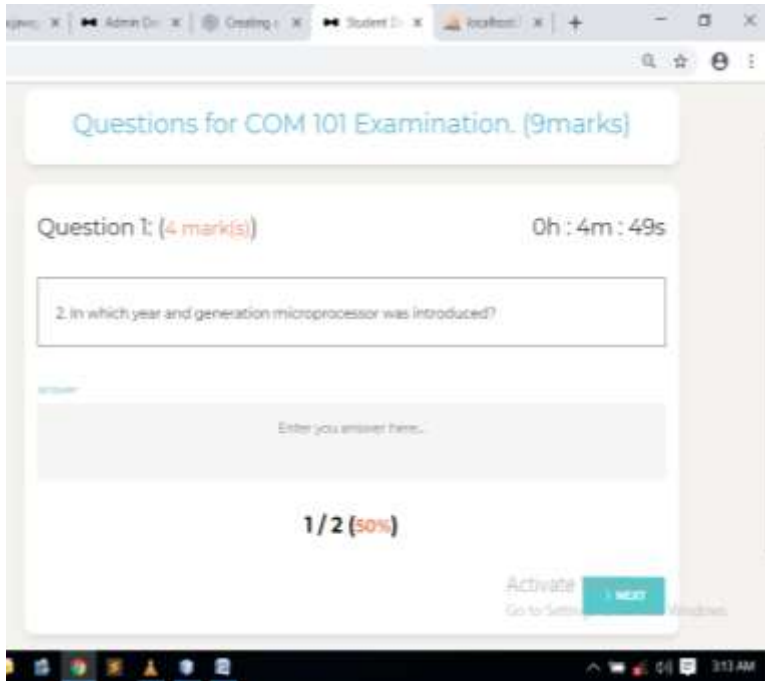


Figure 9: An examination session

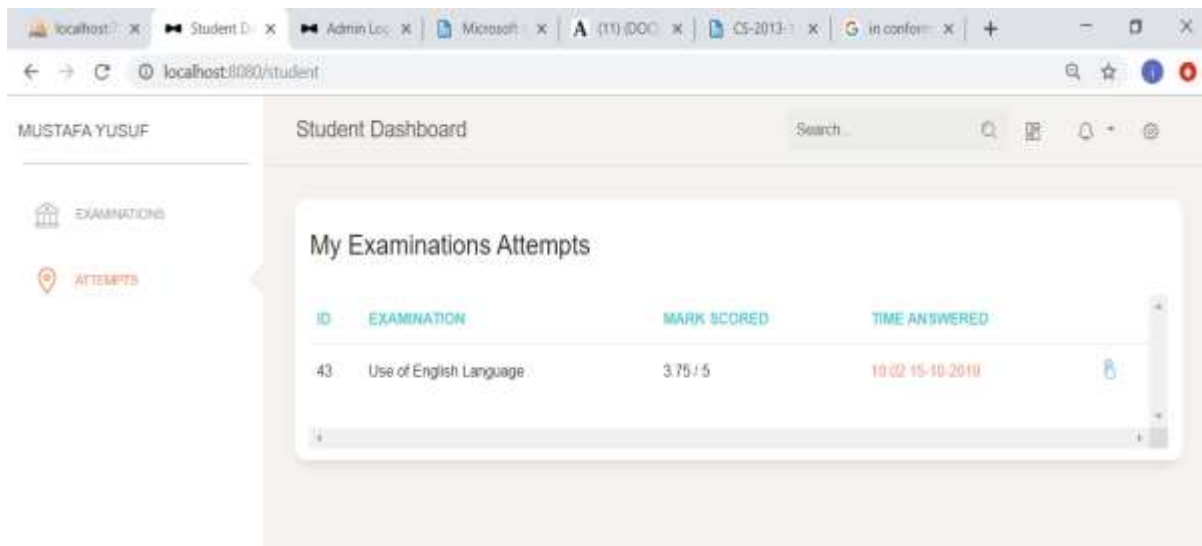


Figure 10: Result display

Performance Evaluation

The developed system was evaluated against the existing system to ascertain its suitability. The evaluation of the performance of the developed and existing system was based on two aspects: the marking speed (response time) and the marking accuracy.

A set of ten (10) different questions were drawn from the various courses offered as the General Studies Programme (GSP) at the Al-Qalam University, Katsina. The test was first conducted using the existing system, for five students, where the answers were manually marked by a lecturer. The same set of questions was then executed on the developed system. The evaluation revealed that the developed system had better response (marking) time than the existing system. This indicated that the developed system is faster than the existing manual system in terms of the marking speed as shown in Table 1.

Table 1: Time taken to mark the manual system against that of the developed system for five students.

Marking Time	Manual system (seconds)	Developed CBA system (seconds)
1	300	0.6
2	600	1
3	480	0.9
4	420	0.8
5	720	0.5
Total	2520	3.8

On the accuracy of marking, Table 2, Table 3 and Figure 11 illustrated the results of the questions after being evaluated using both the existing manual system and the developed system. The mark accuracy of the developed system had an improved accuracy.

Table 2: Developed System Mark Accuracy in Comparison with that marked by the Lecturer

Question	Exam marked by Lecturer	Exam marked by developed system
Q1	3	3.5
Q2	1.5	1.5
Q3	5	5
Q4	3	3
Q5	2.5	3
Q6	5	5
Q7	2	2
Q8	1	2
Q9	4.5	4.5
Q10	2	2
Total	29.5	31.5

Table 3: Developed System Mark Accuracy in Comparison with that of the Lecturer for five students

Question	Exam marked by Lecturer	Exam marked by developed system
Student1	29.5	31.5
Student2	29.5	29.5
Student3	28	30
Student4	30	31
Student5	27	32

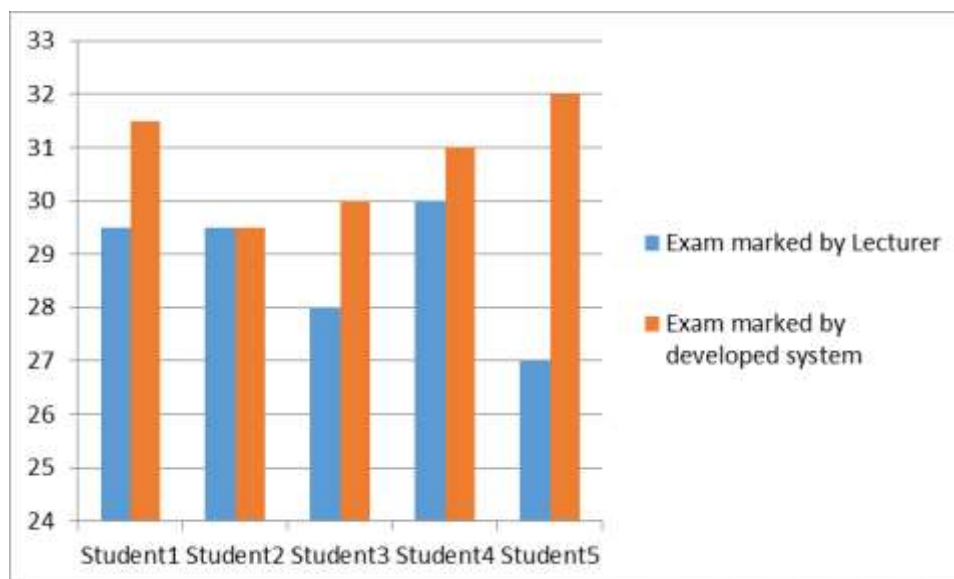


Figure 11: Developed System Mark Accuracy in Comparison with that of the Lecturer

CONCLUSION

In this paper, a simple algorithm was created using keyword and synonym matching technique. The new system was designed using the Unified Modeling Language and a prototype was developed. Java programming language and MySQL were used for the system implementation and database respectively. HTML, CSS and JavaScript were used for the web page interface while Merriam-Webster Dictionary API and Language Tool API were used for the synonyms and sentence grammar. The developed system, though a prototype, demonstrating the designed algorithm, actually serves as a model system of how a comprehensive and fully functional system could be when developed for the evaluation of the subjective questions. Although, the developed system works satisfactorily, it only evaluates descriptive questions written in English language. Therefore, it could be enhanced in future to accommodate other languages used in the Nigerian university system such as Hausa, French and Arabic Languages.

REFERENCES

- Agbonifo, O. C. (2014). An Online Model for Assessing Students' understanding with stepwise solving of calculus questions, *Nigerian Journal of Technology*, 33(1): 86 – 96.
- Alrehily, D. A., Siddiqui, A. M. and Buhari, S. M. (2018). Intelligent Electronic Assessment for Subjective Exams. *Computer Science & Information Technology*. 8 (8): 47-63.
- Dharma R. T., Kiran K. G. and Lakshmi R. (2017). A Python tool for evaluation of Subjective Answers (ApTeSa). *International Journal of Mechanical Engineering and Technology (IJMET)*, 8, 247–255.
- Qureshi, M. Rizwan. (2015). A Proposal of Electronic Examination System to Evaluate Descriptive Answers. *Science International-Lahore*, 27, 2143-2146.
- Shah, J. H. (2002). Validity and Credibility of Public Examinations in Pakistan. An Unpublished Thesis Submitted for the Degree of Ph. D., in the Department of Education, Islamia University Bahawalpur, Pakistan.
- Wang, N., Xu, L., Li, L. and Luxiong X. (2013). Design And Implementation of an Automatic Scoring Subjective Question System Based on Domain Ontology. *Advanced Materials Research*, 7: 3039-3042 available online @www.scientific.net/AMR.753-755.3039



©2021 This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license viewed via <https://creativecommons.org/licenses/by/4.0/> which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is cited appropriately.