Evaluating Student Descriptive Answers Using Natural Language Processing

Ms. Shweta M. Patil

M.E. (CSE) student, Dept of CSE

G.H.Raisoni Institute of Engg and Management Jalgaon , Maharashtra, India

Prof. Ms. Sonal Patil

Assistant Professor,Dept of Computer Engineering

G.H.Raisoni Institute of Engg and Management, Jalgaon, Maharashtra,India

Abstract— **Computer Assisted Assessment of free-text answers has established a great deal of work during the last years due to the need of evaluating the deep understanding of the lessons' concepts that, according to most educators and researchers, cannot be done by simple MCQ testing. In this paper we have reviewed the techniques underpinned this system, the description of currently available systems for marking short free text response and finally proposed a system that would evaluate the descriptive type answers using Natural Language Processing.**

Keywords—Computer Assisted Assessment,Short free text response, Descriptive type answer, Natural Language Processing.

I. INTRODUCTION

"Computer Assisted Assessment (CAA) is a common term for the use of computers in the assessment of student learning [1]". The idea of using computers to assist learning process has surprisingly changed the field of learning system. The study in the field of CAA started nearly in 70's. The CAA systems developed so far are capable of evaluating only essay and short text answers such as multiple choice questions, short answer, selection/association, hot spot and visual identification. Most researchers in this field agree on the notion that some aspects of complex achievement are complicated to measure using objective type questions. Learning outcomes implying the ability to recall, organize and integrate ideas, the ability to express oneself in writing and the ability to supply merely than identify interpretation and application of data, require less structuring of response than that imposed by objective test items [5]. Due to these students will be evaluated at higher level of Bloom's (1956) Taxonomy (namely evaluation and synthesis) that the essay question or descriptive questions serves it's most useful purpose.

Many researchers claim that the essay's evaluated by assessment tools and by human graders leads to great variation in score awarded to students. Also many evaluations are performed considering specific concepts. If that particular concepts are present then only award grades otherwise answer is marked as incorrect. So to overcome this problem new system is proposed.

Purpose of this paper is to present the new system that can evaluate student's performance at higher level of Bloom's taxonomy by considering the assessment of descriptive type questions. The system can perform grading as well as provide feedback for student's to make improvement in their performance. This paper also discusses various techniques underpinned by computer assisted assessment system as well as current approaches of CAA and utilizes it as a framework for designing our new framework.

The techniques for automatic marking of free-text responses are basically categories into three main kinds, Statistical, Information Extraction and Full Natural Language Processing [2].

A. Statistical Technique

It is only based on keyword matching, hence considered as poor method. It cannot tackle the problems such as synonyms in student answers, nor does it takes into account the order of words, nor can it deal with lexical variability.

B. Information Extraction (IE) Technique:

Information Extraction consists in getting structured information from free text. IE may be used to extract dependencies between concepts. Firstly, the text is broken into concepts and their relationships. Then, the dependencies found are compared against the human experts to give the student's score.

C. Full Natural laguage processing(NLP):

It invloves parsing of text and find the semantic meaning of student answer and finally compare it with instructors answer and assign the final scores.

II. RELATED WORK

Jana and John developed C-rater [3] to score student's answers automatically for evidences of what a student knows about the concepts already described by C-rater. It is underpinned by NLP and Knowledge Representation (KR) techniques. In this system model answers are generated with the help of concepts already given and later student's answers are processed by NLP technique. Later on only concepts detection is done and finally scores are assigned. But there are some disadvantages of this system as No distinct concepts specified, incorrect spelling mistakes unexpected similar lexicons and many more. Then Raheel, Christopher and Rosheena created IndusMarker [4] an automated short answer marking system for Object Oriented Programming (OOP) course. The system was used by instructors to assist the overall performance of student and to provide feedback to the students about their performance. It exploits the structure matching i.e.

matching the pre specified structure with the contents of student response text. Automated Essay Grading (AEG) system was developed by Siddhartha and Sameen in the year 2010 [5]. The aim of the system is to overcome the problems of influence of local language in English essays while correcting and by giving correct feedback to writers. AEG is based on NLP and some of the Machine Learning (ML) techniques. Auto-assessor was developed in year 20111 by Laurie and Maiga [6] with an aim to automatically score student short answers based on the semantic meaning of those answers. Auto-assessor is underpinned by NLP technique. This system consists of component based architecture. The components are created in order to reduce the sentences to their canonical form which are used in preprocessing of both supplied correct answers as well as student response. Later evaluation of student response with the correct answer takes place where each word from correct answer in canonical form is compared with the word from student response which is in canonical form and finally scores are awarded for student response. Ade-Ibijola, Wakama and Amadi developed Automated Essay Scoring (AES) an Expert System (ES) for scoring free text answers [7]. AES is based on Information Extraction (IE). This ES is composed of three primary modules as: Knowledge Base, Inference Engine and Working Memory. Inference engine uses shallow NLP technique to promote the pattern matching from the inference rules to the data contained in knowledge base. The NLP module contains: a Lexical Analyzer, a Filter and a Synonyms Handler module. The correctness evaluation is performed by fuzzy model which generate the scores for student answer with the help of two parameters: the percentage match and the mark assigned.

III. PROBLEM DEFINITION

There are a number of commercial assessment tools available on the market today; however these tools support objective type question such as multiple choice Questions or short one-line free text responses. This will assess student's depth of knowledge only at lower level of Blooms taxonomy of educational objectives. They fail to assess student's performance at higher level of taxonomy of educational objective. Also these systems fail to check spelling & grammatical mistakes made by students. As well as they were unable to check the correct word order. Even the answers with wrong word order were awarded assigned scores by mere presence of words in student response. So to overcome the encountered problems the system is going to be developed that evaluated students descriptive answers by considering the collective meaning of multiple sentences. Also system will mark spelling mistakes made and finally scores will be assigned to student answer. The proposed system will try to provide feedback to the students so to help them to improve their performance in academics.

IV. PROPOSED SYSTEM

The proposed system will implement CAA for descriptive type answer. The existing system checks single line text response without considering word order. So the proposed system will try to avoid this problem by considering collective meaning of multiple sentences. The primary focus of this newly proposed system is to determine the semantic meaning of student answer with a consideration that student responses to question in number of ways. The system basically focuses on multiple sentences response.

The basic architecture of proposed system is depicted in fig [1] below. It is basically composed of following components:

A. Student Module:

It consists of question editor where question will be displayed and response editor to enter student response.

B. Tutor Module:

In this module question as well as correct response to respective question is entered by tutor. Tutor will also identify and enter the keywords from correct answer with their respective weights.

C. Processing Module:

Both answers i.e. student response and correct answer will be processed by initially dividing them into token i.e. words. Later on noun phrase and verb grouping will be assigned to each and every word with the help of Part-Of-Speech (POS) tagger. This task is accomplished by NLP technique.

D. Answer comparison and Grade Assignment Module:

Following text processing module, that actual evaluation of student response with correct answer takes place. Each and every word of student response is compared with correct answer. If exact match is found in word as well as POS tag and word position in sentence the scores are assigned.

After score assignment Final scores are calculated by making summation of assigned scores of all words.

E. Projection of Final Scores:

Final calculated scores assigned to student response are given in report.

Now steps for evaluating the descriptive type answer of the proposed system are:

Step 1: Start.

Step 2: Form correct answer and store it in table.

Step 3: Identify keywords, tag keywords with the help of POS tagger and assigned weights depending upon importance of it presence in sentence.

Step 4: Store synonyms and antonyms into another table.

Step 5: set student score to 0. Input student response and store it in another table i.e.SR table.

Step 6: Now check whether keyword present in SR table if present assigned score=score + already assigned weight.

Step 7: If keyword not found check in synonyms table and on finding assign score=score + already assigned weight.

Step 8: Check if antonyms present in SR table if present then score=score * (-1).

Step 9: Check the position vector of noun and verb combination in input answer and compare it to that of correct answer to verify dependencies of noun and verb in answer.

Step 10: if grammatical mistakes are present deduce 2 marks from net score.

Step 11: Now make summation of assigned scores of Student response.

Step 12: If scores calculated are negative then answer entered by student is INCORRECT, else if scores are in same range with already assigned scores then student response is marked to be CORRECT.

V. CONCLUSION AND FUTURE SCOPE

CAA has been an interesting research area since70's. CAA helps in evaluation of student performance accurately and without wastage of time. Most of the CAA developed provides promising results compared with results provided by human graders.

But the problem of existing CAA developed so far is that they evaluate student performance at lower level by making assessment of only objective type questions or essay questions. So the proposed system will try to overcome this problem by evaluating students at higher level by considering assessment of descriptive type question consisting of multiple sentences. The proposed system will consider the collective meaning of multiple sentences. It will also try to check grammatical as well as spelling mistakes made in student response.

The proposed system will try to provide Report to student giving them detail feed back.

REFERENCES

[1] Perez D. 2007, "Adaptive Computer Assissted Assessment of Free Text Students answers: an approach to automatically generate students conceptual models." PhD Thesis, computer science Department Universidad Autonoma de Madrid, 2007.

[2] Mitchell T, Russell Broomhead P, and Aldridge N. 2002, "Towards robust computerised marking of free text responses." In proceeding of the 6th computer Assissted Assessment Conference, 2002.

[3] Jana Z. Sukkarieh, and John Blackmore," c-rater:Automatic Content Scoring for Short constructed Responses." In proceeding of the 22nd International FLAIRS Conference,2009.

[4] Raheel Siddiqi, Christopher J. Harrison and Rosheena Siddiqi,"Improving Teching and Learning through Automated Short-Answer Marking." IEEE Transactions on learning technologies, vol.3, No.3, July-September 2010.

[5] Siddhartha Ghosh and Dr. Sammen S Fatima,"Design of an Automated Eassy Grading (AEG) system in Indian Context." Internation Journal of Computer Application (0975-8887), vol.1, No.11, 2010.

[6] Laurie Cutrone, Maiga Chang and Kinshuk, "Auto-Assessor: Computerized Assessment System for Marking Student's Short-Answers Automatically", IEEE International Conference on Technology for Education,2011.

[7] Ade-Ibijola, Abejide Olu, Wakama, Ibiba, Amadi, Juliet Chioma, " An Expert System for Automated Essay Scoring (AES) in Computing using Shallow NLP Techniques for Inferencing", International Journal of Computer Applications (0975 – 8887) Volume 51– No.10, August 2012.
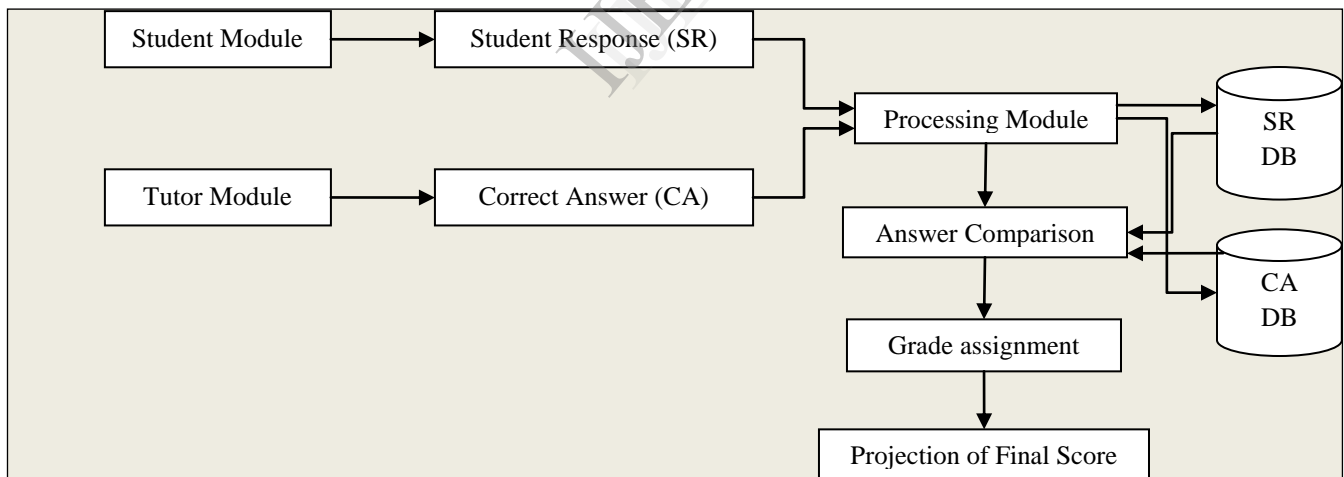

Fig. 1. Architecture of proposed system