April 2013

Automated Assessment of Short One-Line Free-Text Responses In Computer Science

Navjeet Kaur
Guru Nanak Dev Engineering College, Ludhiana, Punjab, navjeet.kaur23@gmail.com

Kiran Jyoti
Guru Nanak Dev Engineering College, Ludhiana, Punjab, kiranjyotibains@yahoo.com

Automated Assessment of Short One-Line Free-Text Responses
In Computer Science

Navjeet Kaur & Kiran Jyoti

Guru Nanak Dev Engineering College, Ludhiana, Punjab, India
E-mail : navjeet.kaur23@gmail.com, kiranjyotibains@yahoo.com

Abstract - Assessment is used to evaluate the learner's knowledge about the concepts learnt. Evaluation through objective testing is common in all evaluation system, where Multiple Choice Questions, Fill in the Blanks, Matching etc. are used for evaluation. The method of objective testing is not sufficient to completely verify all the concepts learnt by the learner. Thus computer assisted assessment of short text answers is developed. Here we present an approach to assess the short answers of computer science automatically. In this paper we have defined a set of evaluation criteria that covers all the relevant aspects of a short text evaluation system.

Keywords - *Automatic assessment, evaluation, examination, text marking, short text evaluation.*

I. INTRODUCTION

A large amount of research has been done on the evaluation and assessment of students' knowledge. Unfortunately, time spent evaluating students, is more than spent instructing them. As teaching is the most highly valued activity for teachers , assessments take time away from students' learning and from teachers' teaching.

e-assessment: It is a powerful tool that is used to solve the problems regarding teaching, providing high quality education. A large part of the research on assessment has been devoted to developing the abilities of automatic marking systems.

Assessment is an important activity in any educational process. It is used to evaluate the learner's knowledge about the concepts learnt. Evaluation through objective testing is common in all learning system, where Multiple Choice Questions, Fill in the Blanks, Matching etc. are used for evaluation but this method of objective testing is not sufficient to completely verify the concepts learnt by the learner. While the Objective type of questioning are very easy to mark, they do not provide the opportunity for deeper assessment. Structured questions and multiple choice tests are vulnerable to guessing.

II. TYPES OF ASSESSMENT

There are mainly two of assessment formative and summative assessment.

Formative Assessment **:** It provide direction, focus and guidance. It can be used concurrently with the learner engaging in some learning process. This type of assessment allows interaction between the assessor (possibly human or machine) and the student while the question is being solved. It is like a tutoring process. It will help the students that lead him/her towards the final solution. These systems acts as an aid in the learning process.

Summative Assessment : It occurs at the end of the session. It is the activity that happens at the conclusion of a learning activity. It evaluates a learner's knowledge. In this teacher will mark a paper when the student finished his work. Then teacher will provide a final mark that is equivalent to the extent of a pupil's knowledge. Summative systems are used to measure the understanding of students.

III. BUILDING BLOCKS OF ASSESSMENT:

When marking [Richard Klein] unstructured texts, teachers look for the following general properties:

1) The concepts covered by the students relevant to a particular topic.

2) Understanding of various concepts.

3) Evaluation by considering critical thinking and analysis skills.

4) Misconceptions.

Content versus Style : There are many researches in this field. Some researches consider content only, others

consider only the style of the submission. There is a also research that consider both.

Content based system : Content assessment is used to measure a student's understanding of a topic. It is more important than style. There are various implementations of content approaches like Bayesian Networks (BETSY) and Latent Semantic Analysis and natural language processing methods that use grammars to parse and interpret input.

Style based system : This approach use proxes. There are intrinsic variables that a human assessor would consider when rating a submission. These variables cannot be measured directly. These systems use multiple regression analysis on data to calculate the proxes and then their evaluation contribute to the final mark.

IV. PRE-PROCESSING

There are a number of words in a text that do not actually affect the meaning of text. For this reason, the input text is to be normalized before the process analyses the input. There a number of techniques that can be used to removing and normalizing these words.

Some common natural language preprocessing techniques are outlined below.

1) Stop-word removal : It is a technique, which removes all common and inconvenient words from the text.

2) Spelling and Grammar : Spelling is a major problem for automatic marking because the system will take it as different word with different meaning. Spelling and grammar mistakes can recognized by many of the dictionary tools available.

3) Stemming : Stemming reduces a word to its stem or root word. Firstly, the pre-processing module parses the text. It then removes all the punctuation and stop words in the text.

To complete this research we first introduce work done by various persons. We then describe our research methodology. Then we will discuss results and future scope. Finally we make our conclusion and raise directions for future research.

Advantages of Automated Marking System

1) These systems provide opportunity for deeper assessments.

2) Subjective type text can be evaluated by using these systems.

3) Unfairness is eliminated.

4) The concepts leaned by a student can be evaluated thoroughly.

5) Easy to use.

6) Teachers can utilize their time in teaching. It leads to time saving.

7) Not vulnerable to guessing of answers by the student.

V. LITERATURE REVIEW

In 1994 page et. al developed a technique Project Essay Grade (PEG) In this technique essay is graded on the basis of text quality only. It uses the concept of "proxes", i.e. computer approximations or measures of variables of interest within the block of text.

Foltz, Laham. has developed a The Intelligent Essay Assessor (IEA) was developed at Knowledge Analysis Technologies (KAT)In this technique LSA (Latent Semantic Analysis) represents documents and their word content in a large two-dimensional matrix semantic space. Using a matrix technique new relationships between words and documents are uncovered and existing relationship are modified to represent more accurately their true significance. LSA is used to compare students' answers to model answers by calculating the distance between their corresponding vector projections. The LSA technique does not take into account any syntactic information. It cannot differentiate between "Class is an entity" and"Entity is a class"

C-Rater:

It is an automated scoring engine, C-rawter™, has been developed at the Educational Testing Service (ETS) in 2001.

Approach :

It uses automated natural language processing techniques to determine whether a student response contains specific linguistic information required. E-Rater assigns a score for writing skills rather than for specific-content.

E-rater :

E-rater is a system developed by the Education Testing Service (ETS) in 1997, 1998.

Approach :

It has been used to rate essays. In this essays are graded against a bench mark of human graded essays. The system uses parsing techniques to identify syntactic features. Content is checked by vectors of content words.

TANGOW & ATENEA

TANGOW

The TANGOW (Task-based Adaptive learner Guidance on the Web) system delivers adaptive websites.

ATENEA

Atenea is a Computer-Assisted Assessment system for automatically scoring students' short answers.

It uses combination of natural language processing (NLP) techniques and statistically based evaluation techniques.

VI. METHODOLOGY

This method accesses a text by computing a percentage based on keywords matching between the student answer and the actual correct answer. The irrelevant words are removed before the final evaluation. Irrelevant words are those words which are not important for the evaluation of answer typed by a student. So in this filtering process this method will remove extra spaces, stop words, punctuation symbols, articles and duplicate words that are present in the actual answer and student's answer. It also consider spell checking and synonyms and extra words while evaluating the student's answer. This method contains the following steps or modules.

1) It creates a list of student's answer and the actual answer.

2) Then it checks for any punctuation symbol that are present in both answers i.e. student's answer and the actual answer. It removes those punctuation symbols.

3) Then it searches for any articles present in both lists. Then it removes the articles present in both lists.

4) After removing articles, it performs search operation for stop words present in both list.

5) Then it will search for duplicate words present in both lists. Then it will replace duplicate words.

6) Then it will count word length by comparing the number of words present in both lists.

7) Then this method will check spellings in both lists.

8) Then after considering all these aspects, this method calculates the percentage of correct part in the answer. It will also find synonyms of a word during this marking process. This method will also consider extra words i.e. the words that are not

relevant to the actual answer, in the marking process.

VII. IMPLEMENTATION

Many academic examinations make heavy use of questions that require students to write at least one or two sentences. For example, questions that requires the learner to state, to suggest, to describe, or to explain. These short answer questions are highly valued, integral to the examinations and are extensively used by educators preparing learners for the examinations. So, in this paper we will introduce a technique which is used to automatically evaluate the short text answers. There are various steps that are used in this technique.
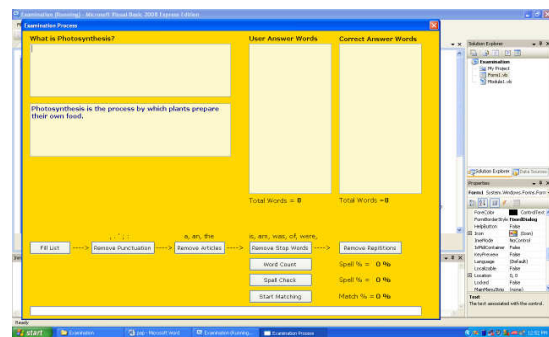

Fig. 1 : Automated Answer Evaluation System

Fig 1 shows two text boxes. In one text box actual correct answer is written which is taken from database and in another text box student will type the answer. There are two lists user answer words and correct answer words. When we click on fill list button then both list are filled with their corresponding answers.
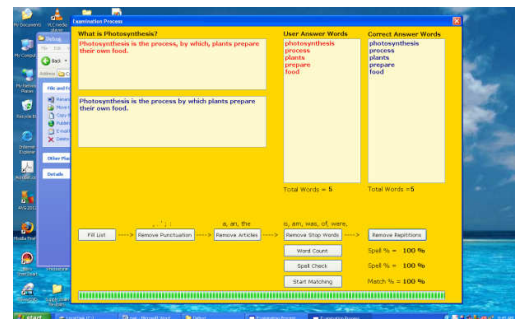

Fig. 2 : Punctuation Symbols

Then in the fig2 , this system will remove punctuation marks which are present in both student answer and correct answer list. In this module, this system will remove punctuation marks, if any present in both answers actual as well as answer typed by the student. Then this system will remove stop words present in both student answer and the actual answer.
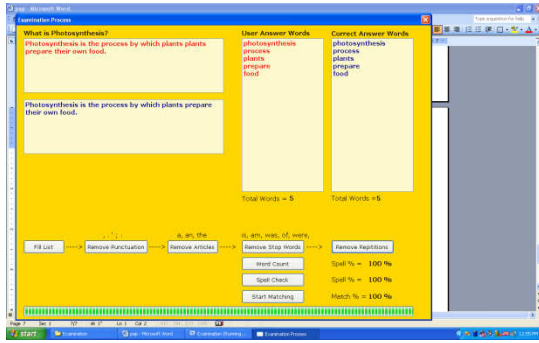
Fig. 1: Duplicate words

In single line sentence, the possibility of duplicate words is rare. So in the fig 3 , it will remove the duplicate words in both lists and replace it with a common word.
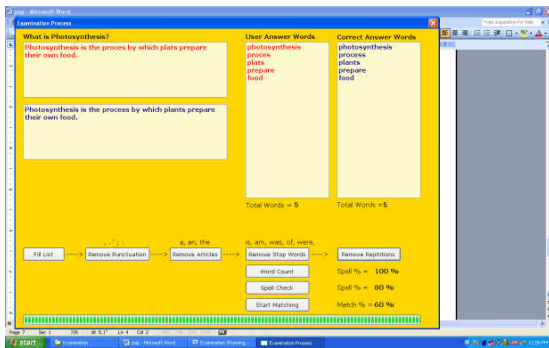

Fig. 4 : Spelling mistake

In this module, there is a database which contains dictionary words. So this system will compare each word in a student answer with a database. Then fig 4 will show a percentage of correct spelled words in a student answer.
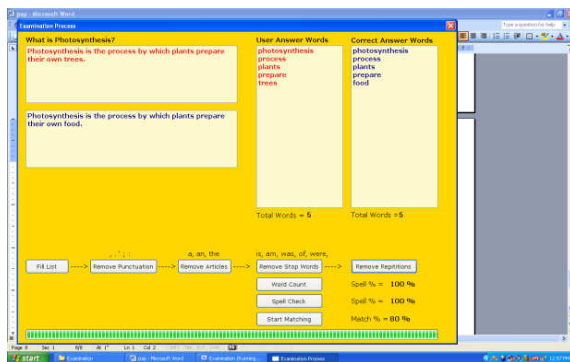

Fig. 5 : Extra words

At the final stage, this system will aggregate the percentage of word count and spell checker module. So

fig 5 shows the aggregation it will also consider extra words that are not relevant to the actual answer.

VIII. CONCLUSION

In this paper we have proposed a technique for evaluating all the relevant aspects of automatic assessment systems for free-text responses. Objective type questions provide limited insight into students' knowledge. So teachers prefer to use unstructured questions in assessments to evaluate the concepts or ideas learned by a student thoroughly. This paper outlined the implementation and analysis of technique used to automatically evaluate short one-line sentence response. This technique is used for summative assessment of short text responses. This technique performs filtering process before evaluation. It removes unimportant data that is present in the answer before the actual evaluation of a response typed by a student. It remove stop words, identify stem or root of a particular word , punctuation symbols, extra spaces, articles, duplicate words if any present in the answer. It also considers spelling mistakes and synonyms while evaluation process. This system will also take into account synonyms of words present in the text during evaluation process.

IX. FUTURE SCOPE

This paper proposed a technique to automatically evaluate short one-line sentence. It evaluates subjective type answers typed by a student. In objective type questions, the teacher cannot thoroughly evaluate the concepts relevant to any particular subject, learned by a student. So we have implemented a technique that will evaluate the subjective type questions. Future work is encouraged which would allow for this technique for word ordering i.e. grammar checking. Future work could contains a technique for automated evaluation of answers which contains paragraphs and should also consider grammar checking while evaluating the short text responses.

REFERENCES

[1]   Richard Klein(2010)" Automated Assessment of Short Free-Text Responses in Computer Science using Latent Semantic Analysis"IEEE Research paper

[2]   Sachin Saxena, Poonam Rani Gupta" Automatic Assessment of Short Text Answers from Computer Science

[3]   Domain through Pattern Based Information Extraction"

[4]   Anne-Laure Ligozat, Brigitte Grau, Anne Vilnat, Isabelle Robba, Arnaud Grappy(2007)" Towards

an automatic validation of answers in Question Answering" 19th IEEE International Conference on Tools with Artificial Intelligence

[5] Victor Gonzalez-Barbone, Martin Llamas-Nistal(2008)" eAssessment of Open Questions: an Educator's Perspective" 38th ASEE/IEEE Frontiers in Education Conference

[6] Adidah lajis, NORMAZIAH ABDUL AZIZ(2009)" ASSESSMENT OF LEARNERS' UNDERSTANDING: AN EXPERIMENTAL RESULT" International Conference on Computer Technology and Development.

[7] Laurie Cutrone and Maiga Chang(2010) " Automarking: Automatic Assessment of Open Questions" 10th IEEE International Conference on Advanced Learning Technologies

[8] Xinming Hu, Huosong Xia(2010)" Automated Assessment System for Subjective Questions Based on LSI" Third International Symposium on Intelligent Information Technology and Security Informatics

[9] Raheel Siddiqi, Christopher J. Harrison, and Rosheena Siddiqi (2010)" Improving Teaching and Learning through Automated Short-Answer Marking" IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES, VOL. 3, NO. 3, JULY-SEPTEMBER 2010

◈ ◈ ◈