Automated Assessment of Short One-Line Free-Text Responses With Identifying Word Ordering

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 With Identifying Word Ordering

Navjeet Kaur & Kiran Jyoti

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

Abstract – Assessment is an important activity in any educational process 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 a technique which also considers grammatical errors during automated evaluation of one-line sentence. In this paper we have define a set of evaluation criteria that covers all the relevant aspects of an essay assessment system and discussion on how this technique find syntactical errors during evaluation of student response.

Keywords - *Automatic assessment, evaluation, examination, text marking, short free text, syntactical errors, grammatical errors, one-line sentence, tenses.*

I. INTRODUCTION

Assessment is used to evaluate the student's understanding of the concepts learned. Computer based assessment can cut down the time that the teachers can take to teach students. A computer can analyze and examine student's answer deeply. It can ease the burden of teachers of assessing large number of questions. Automated assessment can be applied on objective-type questions like follows:-

- Fill in the Blanks

- True/False

- One-word Answer

- Match the columns

- Mental Maths in which student writes resultant value only.

- Multiple Choice Questions

But evaluating Subjective-type questions is a difficult task. While evaluating Subjective-type questions, teacher has to evaluate so many aspects like student's extent of knowledge, number and depth of concepts learned, to evaluate Student's knowledge relevant to that particular topic etc. So Subjective type evaluation is very time-consuming. Much of the teacher time is spent in evaluating Student papers. So in this paper, we will discuss a computer based technique that will evaluate student's answer automatically. This technique can also differentiate between these two answers "She has been gone" and "she gone been has". This technique also checks for grammatical structure and grammatical errors present in student's answer.

Teachers can save their time that was previously spent in marking question papers. They can do their other tasks. This technique is closer to Student and Teacher expectation. This technique can also find misspelt words. It can also find synonyms of words. Synonyms are the collection of words that can be accepted as substitutes of the ones.

II. MODULES OF AUTOMATED ASSESSMENT SYSTEM

- Remove Punctuation: -This module removes all the punctuation symbols present in the text. Punctuation symbols are not relevant to that particular topic so these are first remove before evaluation of subjective type text.

- Check articles: - This module will check whether the article placed before a particular word is correct or not.

- Grammar check: - This module will differentiate between these two sentences "she has been gone" and "she been has gone". I

- Stop-words removal: - It removes all common and inconvenient words from the text.

- Remove Duplicate words: - This module will remove all the duplicate words present in student's answer.

- Size check: - This module will measure the length of Student's answer and the actual correct answer. This value will make contributions to the final marks.

- Spell Checker: - This module will check the spellings of various words present in the student's answer. This module will check the spellings by comparing the words to the inbuilt dictionary that is placed in the database.

- Synonyms check: - This module will check the synonyms. Synonyms are the collection of words that have similar meanings. These words can be replaced with each other.

III. ADVANTAGES OF AUTOMATED MARKING SYSTEM

- These systems are used for deeper assessments.

- Subjective type text can be evaluated with accurate results.

- Unfairness is eliminated.

- The concepts leaned by a student can be evaluated thoroughly.

- Easy to use.

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

- Not vulnerable to guessing of answers by the student.

- This can be used in educational institutions as well as in corporate world.

- Can be used anywhere any time as it is a web based application

- No restriction that examiner has to be present when the candidate takes the test.

- Secure

- Easy to use

- Accurate

- No need of examiner

- In comparison to the present system the proposed system will be less time consuming and is more efficient.

- Analysis will be very easy in proposed system as it is automated

- Result will be very precise and accurate and will be declared in very short span of time because calculation and evaluations are done by the simulator itself.

- The proposed system is very secure as no chances of leakage of question paper as it is dependent on the administrator only.

- The logs of appeared candidates and their marks are stored and can be backup for future use

- Repeatability and consistency

- Reliability

- CAA system can be accepted by all. Biasness in the teachers is removed by using this system.

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.

IV. LITERATURE REVIEW

There are a number of approaches that have been proposed in the past for automatic short answer grading.

C-Rater matches the syntactical information present in student response i.e. subject, object, and verb to that of a set of actual answers.

Another short answer grading system, Auto Tutor is tutoring environment with a graphical and speech recognition features for the improvement of the automated grading system.

The Oxford-UCLES system searches for patterns by starting with a set of keywords and synonyms and searching through windows of a text for new patterns. A later implementation of the Oxford-UCLES system compares several machine learning techniques, including inductive logic programming, decision tree learning, and Bayesian learning, to the earlier pattern matching approach.

Another approach used a machine learning algorithm in which it used combinations of simple features e.g., a pair of nouns appear within 5 words from one another in both texts. This method also considers synonym words, word ordering, text length, and word classes.

Another work found text similarity from the simpler problem of word similarity explores the efficacy of applying WordNet-based word-to-word similarity measures to the comparison of texts and found them generally comparable to corpus-based measures such as LSA.

Project Essay grade PEG is one of the earliest and longest-lived implementations of automated essay grading. It was developed by Page and others. In this an essay is graded on the basis of writing quality, taking no account of content. It uses the concept of "*proxes*", i.e. in*trinsi*c variables of interest within the essay (what a human grader would look for but the computer can't directly measure) to simulate human rater grading. *Proxes* include: essay length, counts of prepositions, relative pronouns and other parts of speech, as an indicator of complexity of sentence structure; variation in word length.

Intelligent Essay Assessor is based on the Latent Semantic Analysis (LSA) technique that was originally designed for indexing documents and text retrieval. LSA represents documents and their word content in a large two-dimensional matrix semantic space. It uses a technique known as Singular Value Decomposition (SVD). A matrix represents the words and their contexts. Each word present in the student's response represents a row of in the matrix, while each column represents the sentences or paragraphs. The cells of the matrix contain the frequencies of the words in each context. This initial matrix is then transformed according to an inverse document frequency weighting approaches. LSA makes no use of word order.

E-Rater uses a combination of statistical and NLP techniques to extract linguistic features of the answers to be graded. Essays are evaluated against a benchmark set of human graded essays.

Finally, explicit semantic analysis (ESA) uses Wikipedia as a source of knowledge for text similarity. It creates for each text a feature vector where each feature maps to a Wikipedia article. ESA outperform LSA on some text similarity tasks.

V. 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. This method also considers grammatical errors during evaluation. It also considers 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 student's answer. It will check whether correct article is placed before that word. If that article is not correct then it will shows the percentage of error.

(4) After checking the articles, it will check for any grammatical mistakes present in the student answer. It can check the 12 types of tenses structure. This software will check whether the position of helping verb is correct or not. I.e. this software can differentiate between these two sentences- "she is not going" and "she not going is". It will also identify pronoun that is replaced by noun by the student. It can also evaluate negative sentence structure. Then after checking all these, it will give the percentage of error occurred in the student answer.

(5) Then after checking grammar, it will remove the unimportant or irrelevant words present in the student answer for finding the words that are irrelevant to the correct answer. It will remove the stop words present in the student answer like is, through, by etc.

(6) Then in this step, it will remove the duplicate words present in the student answer.

(7) In this step, it will refine plurals i.e. covert plurals to singular for words matching purpose.

(8) Then it will count the number words present in the student answer typed by the student and the actual answer. It will do this for checking the size of both student's answer and the actual answer.

(9) This it will check the spellings of words in the student's answer by comparing the words present in the student's answer with the dictionary placed in the database. Then it will show the percentage of correct word spellings in the student's answer.

(10) Then it will check for synonyms in the student's answer. Synonyms words are the collections of words that have similar meanings.

(11) Then after performing all these checking, it will find the average of all the percentages calculated for different aspects and then it will display the final percentage that shows the extent of correctness present in the student's answer.

VI. 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 answers are highly valued, integral to examination and are extensively used by examiners preparing learners for the examination. So we have introduced a technique which is used to automatically evaluate the short text answers. There are various steps that are used in this technique.
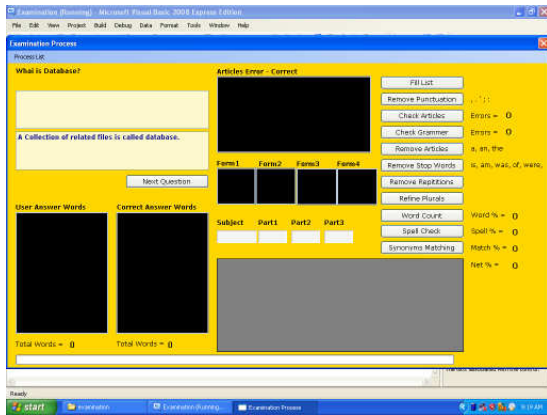

Fig. 1.1 Automated Answer Evaluation System

Figure 1.1 shows two text boxes. In one text box actual correct answer is written and in another text box student's answer is given. There are two lists: user answer words and correct answer words. In this system, there is a menu named "process list". In this there are various submenus named- "add question" by using this we can add more actual question-answers to the database. Next is "add words to dictionary" by using

this we can add more words to the dictionary. Another is "stop words" by using this we can add more stop words to the database. Then there is "add singular plural" by using we can add more words singular and plural form. Then there is "add article words" by using this we can add more words with its corresponding correct article in the database. This system will also guide us during its operation. It gives information about what is the error present in the articles.
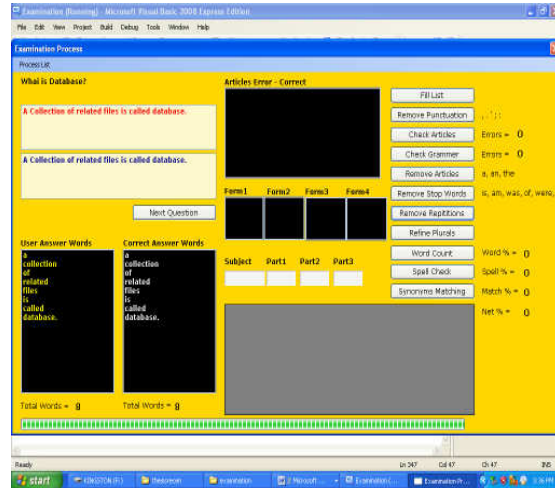

Fig 1.2 : Fill List

In this step as shown in fig. 1.2 we click on fill list button, then both lists user answer words and correct answer words are filled with their corresponding answers.
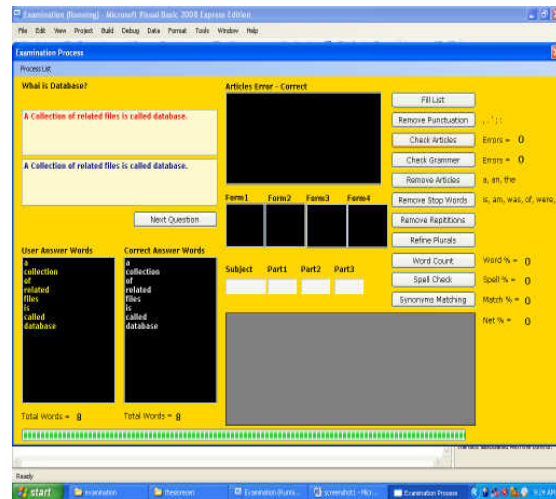

Fig 1.3 : Remove punctuation symbols

Then in this fig.1.3 punctuation symbols that are present in both user answer and correct answer lists are removed because these symbols are not relevant to the actual answer.
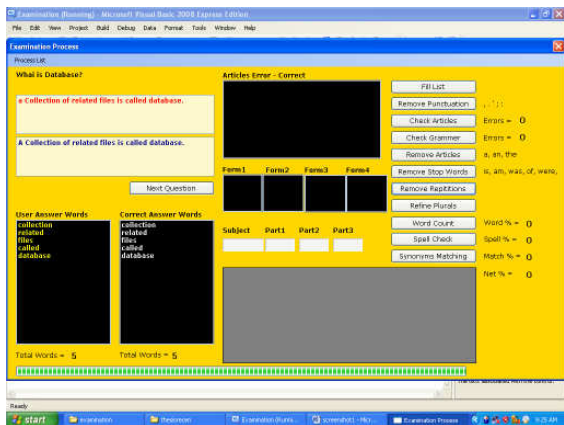

Fig 1.4 : Check articles

In this step, this system will check the articles present in the answer given by the student as shown in fig 1.4. It will check whether the article count placed before a particular word is correct or not. If that article is correct then this system will display 0 errors otherwise it will count the number of wrong articles and display that count value in the errors field.
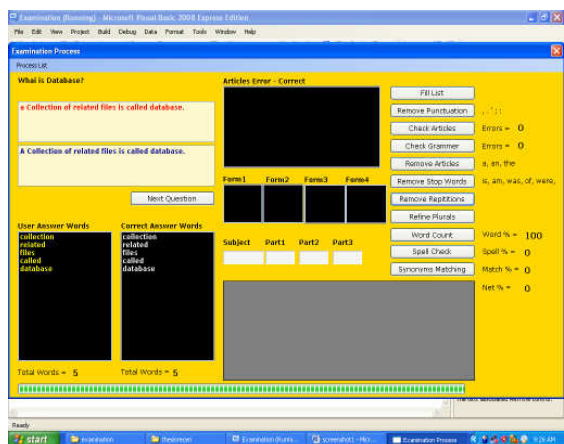

Fig1.5 : Check Grammar

In this step, this system will check the grammar present in the answer given by the student as shown in fig. 1.5. It will check the structure of the answer given by the student. This part will also check the pronoun that is being replaced with the subject by the student. This step will check the twelve types of tenses. It will

measure the extent of correctness present in the structure of the answer given by the student. It will also check the position of helping verbs like had been, have been present in the text. This step will make differentiate between these two sentences "she had been gone" and "she gone been had". Then it will count the number of errors present in the grammar part. If that grammar part is correct then this system will display 0 errors otherwise it will count the number of errors present in the grammar part and display that count value in the errors field.
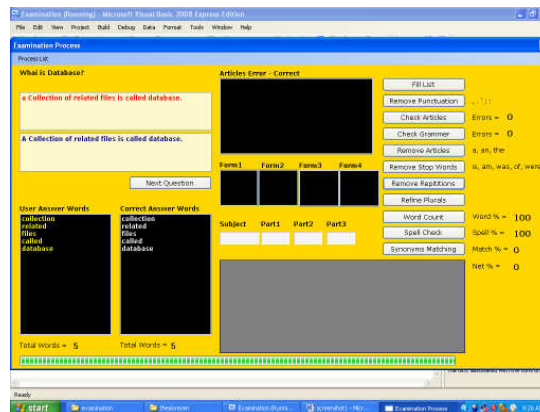

Fig 1.6 : Remove articles

Then after checking the grammar part, this step will remove the articles present in the student's answer and actual answer as shown in fig. 1.6. This step performs filtering operation. Because articles had been checked n the previous step, this step now removes the articles present in the student's and actual answer.
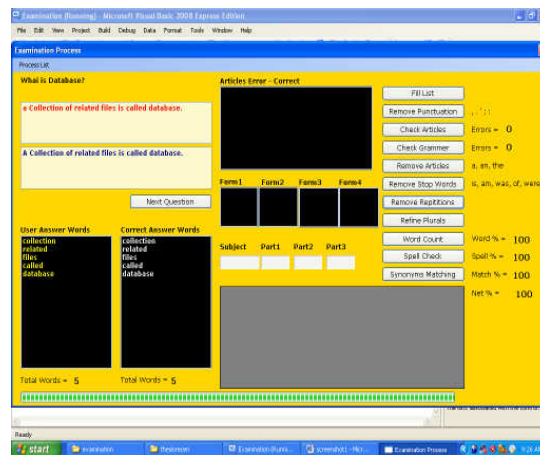

Fig. 1.7 : Remove Stop Words, Remove Repetitions and Refine plurals, Word Count, Spell Check and Synonyms Matching

In fig. 1.7 next step will remove the stop words present in the student's answer and actual answer. This step performs filtering operation. Because stop words are irrelevant to the actual answer, so this step now remove the stop words present in the student's and actual answer.

Then this step will remove the duplicate words present in the student's answer. This step will remove the duplicate words present in the student's answer and replace it with a common word.

Then next step is refine plurals, this step will check the correctness of plurals present in the text.

Then in the next step word count, this system will count the number of words present in the actual answer and the student's answer. By doing this, this system will check the size or length of the answer. Then it will display the percentage according to the size matched.

In the next step, spell check, there is a database which contains dictionary words. So this system will compare each word of a student's answer with the dictionary words. This figure shows the percentage of correct spelled words in a student's answer.

In the next step, it will find the synonyms. Synonyms are the collection of words having same meaning .It will identify root or stem of the words present in the answer. In this system, there is a database in which synonyms of various words are stored. So this system will compare each synonym present in the student's answer with the database and then give then display the percentage of correct words accordingly. This step also considers extra words i.e. words that are not relevant that actual answer during percentage evaluation.

Then in the last step, final percentage of whole answer is displayed. It will display the extent of correctness present in the complete answer. This percentage is evaluated by calculating the average of aspects –spell check percentage, word count percentage and synonyms matching percentage. This final result also considers error value of both grammar part and article part.
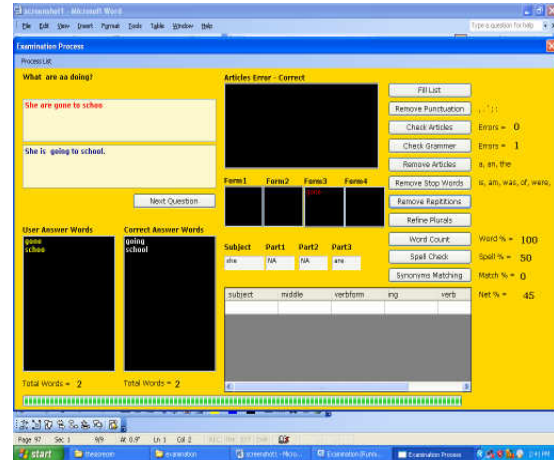

Fig 1.8 : Sample Answer

In this sample answer as shown in fig. 1.8 there is error in spelling "school" and verb form is incorrect "gone" and helping verb is also in correct. This system had displayed the result accordingly.

VII. CONCLUSION

The automated assessment with identifying word ordering of short sentences algorithm is developed in the present research work. This system will automate the process for subjective evaluation. This system will also consider syntactic information during marking process. This system will also check any grammatical errors if any, present in the system. The conclusions for the present research are listed below:-

- A technique that considers all the relevant aspects of text for its automated evaluation.

- This technique performs filtering function before the automated evaluation of student's answer. It removes stop words, identify synonyms, check spelling mistakes, remove punctuation symbols, extra spaces, articles and duplicate words that are present in the answer that are not important for the marking process.

- This technique is used for summative assessment of short one-line sentences.

- This technique considers syntactic information during marking process. it will check the twelve forms of tenses present in the student answer. It will also check whether the positions of helping verb present in the student answer is correct or not. It will also consider subjects that are replaced with pronouns in the student answer. it will also consider verb form that is that comes after the helping verb. It also checks whether the article entered by the

student before subject or object is correct or not. It will also check for any errors in the negative sentences. It will also guide the examiner about the correct type of article that should be present in the student answer.

VII. FUTURE SCOPE

The research work point to following directions of research that are likely to be needed to further enhance the scope of the system.

- Future work could contain a technique for automated evaluation of answers which contains paragraphs.

- To develop a technique that should consider miscellaneous types of sentences. E.g. Do it, Open the windows, let her do it, A stick is to walk with.

- To develop a technique that will consider active and passive voice during evaluation process.

- To develop a technique that considers direct and indirect speech during evaluation process.

- Future work is encouraged which would allow for this technique to consider interrogative sentences that could be present in the paragraphs.

- This technique cannot differentiate between Subject and Object.

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

❖❖❖