Towards an automated system for short-answer assessment using ontology mapping

V Senthil kumaran and A Sankar
PSG College of Technology, Coimbatore, India

Abstract: *A key concern for any e-assessment tool (computer assisted assessment) is its efficiency in assessing the learner's knowledge, skill set and ability. Multiple-choice questions are the most common means of assessment used in e-assessment systems, and are also successful. An efficient e-assessment system should use variety of question types including short-answers, essays etc. and modes of response to assess learner's performance. In this paper, we consider the task of assessing short-answer questions. Several researches have been performed on the evaluation and assessment of short-answer questions and many products are deployed to assess the same as part of e-learning systems. We propose an automated system for assessing short-answers using ontology mapping. We also compare our approach with some existing systems and give an overall evaluation of experiment results, which shows that our approach using ontology mapping gives an optimized result.*

Keywords : *e-assessment, short answer assessment, Ontology mapping, automatic assessment*

Received July 25, 2013; Accepted January 29, 2014

1. Introduction

Recent developments in e-learning systems have changed the way of teaching and learning process [26, 27]. E-learning is not just concerned with providing easy access to learning resources but also concerned with some supporting features such as personalized learning, customized pedagogy, understanding learner's behavior and attitude, recommendation to learners, collaborative learning and assessing the learner's knowledge in the concept learnt. A large amount of research has been performed on e-learning and its supporting features [12]. Our research focuses on the e-assessment.

Assessment is the most important and critical activity in any educational system. Generally, automatic assessment is preferred over manual assessment to avoid bias errors, human errors and also conserves teacher's time. Evaluation through objective tests like multiple choice questions, fill in the blanks, matching, true / false is common and successful in all educational systems, but it is not sufficient to completely verify the knowledge acquired by the learner. The reason is that they lack deeper assessment. Learners can easily guess the answer. Short –answer type question will make the learners think and write the answer for the given question in 3 to 4 sentences. One can more effectively assess the learner's knowledge using short-answer type questions. Our research focuses on assessing short-answer type questions.

E-assessment is a powerful tool that automates the assessment. A good amount of research has been done on assessing short-answer type questions and there are at least 12 products such as ETS (Educational Testing Service) c-Rater, e-Rater (automatic essay scoring system), IAT (Intelligent Assessment Technologies) etc. deployed to assess short-answer as part of self tutoring or as a component of e-learning systems [4, 31].

1.1. Ontology

The underlying data models in our process are ontologies. There are several research fields in computer science that have embraced ontologies, including knowledge engineering, knowledge representation, qualitative modeling, language engineering, database design, information retrieval, information extraction, knowledge management and knowledge organization. It is possible to define ontology with a sextet of the form:

$$O = [C, P, R^C, R^P, A, I]$$

Where C is the concepts, P is the properties, R^C and R^P are the relations between concepts and relation between properties, A is the set of axioms and I is the instance of concepts and properties. Standard languages like RDF and OWL are used to create the ontology documents.

1.2. Ontology Mapping

Ontology mapping seeks to find semantic correspondences between similar elements of different ontologies. We describe our understanding of the term "mapping": Given two ontologies O1 and O2, mapping one ontology onto another means that for each entity (concept C, relation R, or instance I) in ontology O1, we try to find a corresponding entity, which has the same intended meaning, in ontology O2. The mapping between two ontologies (M) can be defined as follows:

$$M = [C1, C2, R, S]$$

In this, C1 is a concept in O1, C2 is a concept in O2, R is the relation between C1 and C2 and S is the similarity between C1 and C2. So, ontology mapping is the creation of a mapping function M.

There are several approaches [16, 20, 23 and 32] that create ontology from plain text. These approaches are either Schema-Driven or pattern discovery Relationship Discovery. They can extract only the pre-specified information from the plain text. This is sufficient if the extracted information is used to create a knowledge domain that follows a predefined schema, but not for any assessment purpose. Hence, we build RDF sentence followed by ontology for the entire answer in the plain text form.

This paper is organized as follows. In section 2 we give a brief overview about the related work. In section 3 we give the data set used and section 4 discusses the overall architecture of the system and describes the procedure in assessing short answers, while in section 5 we evaluate the system. Finally, section 6 concludes the paper and suggests future works.

2. Related Work

To find relevant literature on e-Assessment in general, and assessing short answer in specific and in RDF graph, a systematic search method was applied. We have reviewed the related work in short answer assessment and text semantics, and carefully selected the articles that are most relevant to our research. Systems for automating the assessment of textual answers have been available commercially since the mid 1990's. [27]. Many e-assessment tools for assessing short answers have significantly spread in recent years. Perhaps the most well known system for the e-assessment of free text is e-rater [2], an automatic essay scoring system employing a holistic scoring approach. The system is able to correlate human reader scores with automatically extracted linguistic features, and provide an agreement rate of over 97% for domains where grading is more concerned with writing style than the content. Since e-rater considers only the writing style, it can assess essay questions but not suitable for short answers.

Siddiqi and Harrison [26, 27] emphasized that content is important rather than writing style in order to assess short answers. Poor writing style is normally tolerated. He assessed the short answers based on content. C-Rater was developed by Leacock et al. [15] at the Educational Testing Service (ETS) and it matches syntactical features to assess the short answers. They find the similarity between two concepts based on the shortest path between two concepts in the lexical database using node counting.

A different technique, which shows high promise, is that of Latent Semantic Analysis (LSA) [14, 11]. LSA has been applied to essay grading, and high agreement levels obtained. These techniques are more suited to evaluate essays than short-answer questions, since they focus on metrics, which broadly correlate with writing style, augmented with aggregate measures of vocabulary usage. Term co-occurrences are captured by means of dimensionality reduction through singular value decomposition on the term document matrix. Pulman and Sukkarieh [22] have come up with a system using information extraction techniques for evaluating short-answer questions. Similarity between two concepts is based on the longest common subsequence between two concepts using the node depth in the lexical database such as WordNet. Mohler and Rada [20] explored a set of unsupervised techniques for automatic grading of short answer. They assigned grade based on a measure of text-text semantic similarity between student answer and model answer.

The software developed by Intelligent Assessment Technologies (IAT) and used by the Open University is most closely related to the system developed by Pulman and Sukkarieh [22]. The main strength of the IAT system is that it provides an authoring tool which enables a question author with no knowledge of natural language processing (NLP) to use the software. Recent advancements in Natural Language Processing (NLP) and Machine Learning (ML) have motivated the researchers to use these techniques in assessing short answers. ML techniques can automate evaluation of texts without having to understand the student answer. Therefore, no context and semantics are considered. Using ML technique, machine can learn to recognize complex patterns in student answer. Alotaibi et al. [1] used a hybrid approach that combines Information Extraction, Decision Tree Learning and Machine Learning to assess short answers. AutoMark, an automatic assessor developed by the Intelligent Assessment Technologies [19] uses IE techniques to assess short answer. AutoMark matched the student answer with the set of templates to evaluate the answer. It requires manually created templates for each correct answer.

Several measures are compared, including knowledge-based and corpus based measures, with the best results being obtained with a corpus-based measure using Wikipedia combined with a "relevance feedback" approach [12, 20] that iteratively augments the instructor answer by integrating the student answers that receive the highest grades. Use of ontology in e-learning concentrates on formalization of learning objects, instructional process and learning design. Cubric M et al. [6] describe various ontology-based strategies for automatic generation of Multiple Choice Questions (MCQ), from arbitrary knowledge domains. The generation is based on the basic meta-ontology relations between a ‘class’ and an ‘individual’.

Methods like Computational Linguistics [28], combining keyword based methods [13], pattern matching techniques [16], breaking the answers into concepts and their semantic dependencies [6], LSA with syntactic, semantic information [14 and 11], and graph matching techniques [9] are the other techniques used for the assessment of student’s free text answers. The semantic annotation of texts consists of extracting semantic relations between domain relevant terms in texts [3, 18 and 24]. Several studies address the problem of capturing complex relations from texts [10]. They combine statistical and linguistic analyses. Cherfi H. et al. [5] have semantically annotated text using RDF graph. Thomas, Ani et. al., [30] explored the text semantics by extracting noun phrases in subject and object roles.

3. Data Set

In order to evaluate our approach for assessing short answer and to compare it with some existing approaches, we have used the data set released by Mohler, Rada [20]. The data set consists of three assignments of seven questions each, in plaintext format. Each assignment includes a question, teachers’ answer (model answer), and set of student answers with the average grades of two annotators included. Thus, the data set we used consists of a total of 630 student answers (3 assignments x 7 questions/assignment x 30 student answers. The format for the assignment files is as follows. Table 1 shows the format of the assignment files with line numbers and < > for reference:

Table 1. Format of the assignment files with line numbers and < > for reference.

```
1 | ##################################
2 |        Question: <QUESTION1>
3 |        Answer: <ANSWER1>
4 |
5 | <Grade1:1> [<Student1>] <StudentAnswer1:1>
6 | <Grade1:2> [<Student2>] <StudentAnswer1:2>
...
32| <Grade1:28> [<Student28>] <StudentAnswer1:28>
33|
34| ##################################
35|        Question: <QUESTION2>
36|     Answer: <ANSWER2>
37|
38| <Grade2:1> [<Student1>] <StudentAnswer2:1>
...
```

The answers were graded by a human judge, using an integer scale from 0 (completely incorrect) to 5 (perfect answer). Table 2 shows a sample data taken from the data set released by Mohler and Rada [20].

Table 2. Sample Question, correct answer and 6 students answers taken from the data set.

Sample Question, Model Answer and 6 Students answers		**Grade by human expert**
Question: What is a variable? Correct Answer: A variable is a location in memory that can store a value.		
Student Answer 1:	A variable is a location in memory where a value can be stored.	5
Student Answer 2:	In programming, a structure that holds data and is uniquely named by the programmer. It holds the data assigned to it until a new value is assigned or the program is finished.	4.5
Student Answer 3:	Variable can be a integer or a string in a program.	2
Student Answer 4:	Variable is a location in the computer's memory, in which a value can be stored and later can retrieve that value.	5
Student Answer 5:	A named object that can hold a numerical or letter value	3.5
Student Answer 6:	a value/word that can assume any of a set of values	3

4. Short Answer Assessment

There are three stages pipelined in assessing short answer, as outlined in Figure 1.

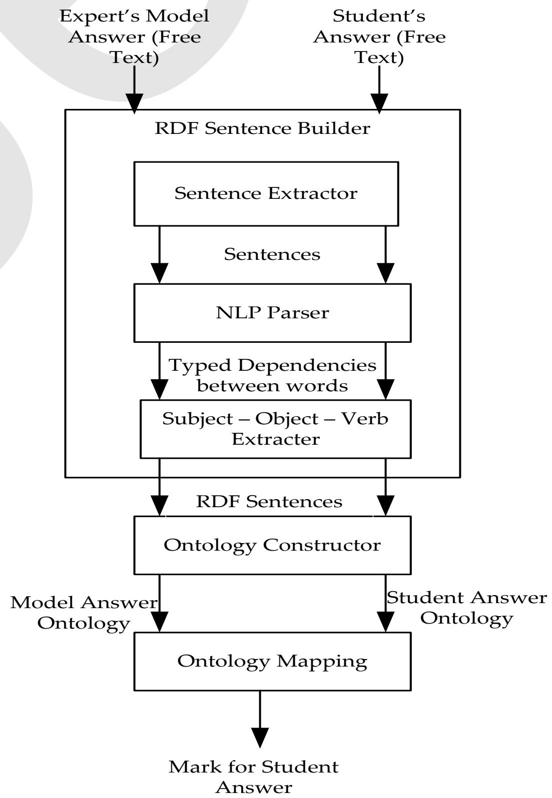


Figure 1. Architecture of automatic assessment of short answer.

In the first stage, the system reads the student answer and model answer as input in plain text format and builds the RDF sentence for each sentence in the model answer and student answer. Ontology for student answer and model answer is constructed in the second stage, from the RDF sentences built in the previous stage. Third stage is the Ontology Mapping, which matches the Model Answer Ontology with student answer Ontology and returns the mark for student answer based on the weightage and similarity score.

4.1. RDF Sentence Builder

There are four steps involved in the RDF Sentence Builder stage. First, system extracts the sentences from the model answer and student answer. Next, the system parses each sentence and builds the typed dependency representation for each sentence. We use Stanford typed dependency parser [7] to build the typed dependency representation. Stanford typed dependency parser provides the grammatical relationships in a sentence. It gives only the textual relation. Marie-Catherine de Marneffe and Christopher De Marneffe [7] define typed dependency parser as follows: The current representation contains approximately 53 grammatical relations. The dependencies are all binary relations: a grammatical relation holds between a governor (also known as a regent or a head) and a dependent. The definitions make use of the Penn Treebank part-of-speech tags and phrasal labels.

The typed dependency representation for the model answer and for student answer 1 in the sample data is given in Figure 2. The dependencies generated by the Stanford parser for each of the sentence carry abbreviated relation name, word-position numbers along with their arguments.

det(variable-2,A-1) nsubj(location-5,variable-2) cop(location-5,is-3) det(location-5,a-4) root(ROOT-0,location-5) nsubj(store-10,location-5) prep_in(location-5,memory-7) aux(store-10,can-9) rcmod(location-5,store-10) det(value-12,a-11) dobj(store-10,value-12)	det(variable-2,A-1) nsubj(location-5,variable-2) cop(location-5,is-3) det(location-5,a-4) root(ROOT-0,location-5) prep_in(location-5,memory-7) det(value-10,a-9) nsubjpass(stored-13,value-10) aux(stored-13,can-11) auxpass(stored-13,be-12) rcmod(location-5,stored-13)
(a) for Model Answer	(b) for Student Answer 1

Figure 2. Typed dependency relation.

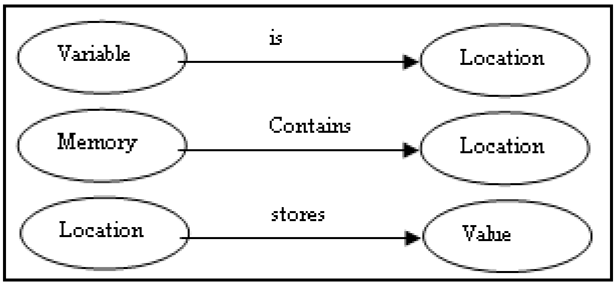
Third step in RDF sentence builder is removing the unwanted dependencies. The unwanted dependencies are the ones that do not give any meaningful relation. For example, det gives the articles and demonstratives in the sentence, which are not important for identifying RDF triple. The next operation is to transform the relations into RDF triples by removing and/or combining unwanted relations. To transform into RDF triple we follow the following strategy on the typed dependency relation generated in the previous step.

- A determiner is the relation between the head of an NP (Noun Phrase) and its determiner. Articles and demonstratives are not important in identifying S-P-O. The determiner relation is given as det(X-n1, Y-2), where Y is the determiner for X and n1 & n2 are the positions of the words X & Y in the sentence. We delete the det relation (i.e. the determiner Y) from the sentence. At the same time if the determiner is a quantifying determiner the relation is given as predet(X-n1, Y-n2). Quantifying determiners such as some, all, etc. are important and they can change the meaning of the NP. So, it is not removed. We combine both words into a single word and give the position as the lower and replace every occurrence of the X and Y with XY.
- Next we consider the nn (noun compound modifier) dependency relation. A noun compound modifier of an NP is any noun that serves to modify the head noun. If nn relation is given for two continuous words, we combine both word into single word and assign the position of the word as the lower position and then the relation is deleted after replacing every occurrence of both words with the combined word.
- advmod and amod relations are also treated as like nn relation.
- agent relation is a complement of a passive verb. We take the dependent as the subject and the governor as the predicate.
- aux and auxpass relations give non-main verb of the clause and the passive information, which is taken for predicate.
- cc, coordination is the relation between the elements of a conjunct and the coordinating conjunction word of the conjunct.
- neg, negation modifier is used to specify the negation word.
- nsubj, nsubjpass, csubj, csubjpass relations specify the subject of the sentence.
- dobj, iobj, root relations give the object of the sentence.

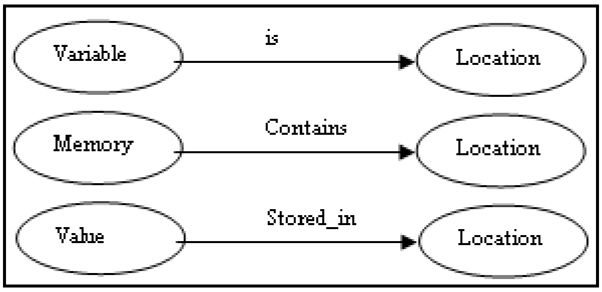
The RDF triples for the typed dependency relation given in Figure 2 is given in Figure 3.

4.2 Ontology Construction

In this stage, we construct ontology for the RDF sentences built in the previous stage. Zhang Xiang et al., [33] have mentioned that there are two types of links between two RDF sentences based on common term they shared. 1. Sequential Link 2. Coordinate Link. Sequential link exists when predicate or object is common between two RDF sentences. Coordinate link exists when subject of the both RDF sentences are same. Xiang Zhang et al., [33] have defined an RDF Graph, as a weighted and directed graph, characterizing the links between RDF sentences from the viewpoint of a user. We construct RDF graph for the RDF sentences using sequential and coordinate links. RDF graph for the RDF sentence triples (given in Figure 3) is given in Figure 4.

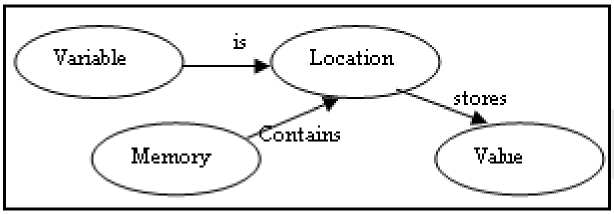


(a) for Model answer

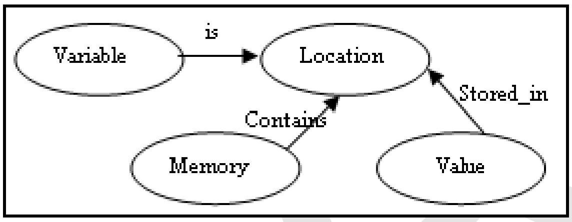


(b) for Student Answer 1

Figure 3. RDF triple.



(a) for Model Answer.



(b) for Student Answer 1.

Figure 4. Ontology.

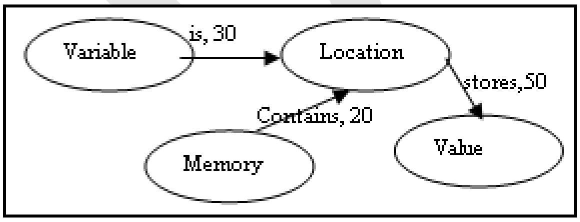


Figure 5. Ontology with weight for the model answer.

The scheme for evaluation of model answer given by the human expert is taken as the weight in the ontology of model answer. The weight of each link in the graph will be in the range of 0 to 100 based on the level of importance of that sentence in the answer and the total weight of all links will be 100. The ontology with weight for the model answer is given in Figure 5.

4.3. Ontology Mapping

In this section we map the RDF graph constructed for model answer and student answer. The mapping value we get is the score assigned to the student answer. Let H be the graph constructed for model answer and T be the graph constructed for student answer. The edge set of H, E(H) and edge set of T, E(T) is:

- $E(H) = \{h_1, h_2, h_3 \ldots h_n\}$ h_i, i = 1, 2, .. n
- $E(T) = \{t_1, t_2, t_3 \ldots t_m\}$ t_j, j = 1, 2, .. m

To find the matching between every edges in H with every edges in T, we take the Cartesian product of E(H) and E(T).

- $E(H) \times E(T) = \{(hi, tj) / i = 1, 2 \ldots n, j = 1, 2 \ldots m\}$

Let the end vertex of h_i be (v_{i1}, v_{i2}) and the end vertex of t_j be (u_{j1}, u_{j2}).
Let the edge weight of h_i is W_i denotes the weight assigned by the teacher in the model answer.
The matching between hi and tj is as follows:

$$m(hi,tj) = \frac{Sim(vi1,uj1)+Sim(vi2,uj2)}{2} * wi \qquad (1)$$

where $m(h_i, t_j)$ is an integer value, $Sim(v_{i1},u_{j1})$ is the similarity between the two vertices v_{i1} and u_{j1} and $Sim(v_{i2}, u_{j2})$ is the similarity between the two vertices v_{i2} and u_{j2}. We use the wordnet based similarity measure [17] to find the similarity between two vertices. The total matching of all edges is:

$$k = \sum_{j=1}^{m}\sum_{i=1}^{n} m(hi,tj) \qquad (2)$$

k is the score assigned to the student answer.

5. Experimental Evaluation and Discussion

In this section we discuss the experimental evaluation of our proposed system and the results are compared against the other existing approaches for representing text for assessing content. We use the dataset described in section 3 to evaluate our system. The efficiency of the system depends on how far system assigned score correlates with human assigned score. To evaluate our system we use scatter graph and Pearson correlation. The scatter graph between the score assigned by human expert and the score assigned by our system for each question (per assignment) is given in Figure 6. From Figure 6, it is clear that system assigned score correlates well with human expert assigned score. The Pearson correlation between the score assigned by human expert and the score assigned by our system per assignment is given in Table 3.

Table 3. Pearson correlation between the score assigned by human expert and the score assigned by our system.

Assignment No.	Pearson Correlation
1	0.816927
2	0.758832
3	0.796056

From Table 3, we understand that our system works well for short answers.

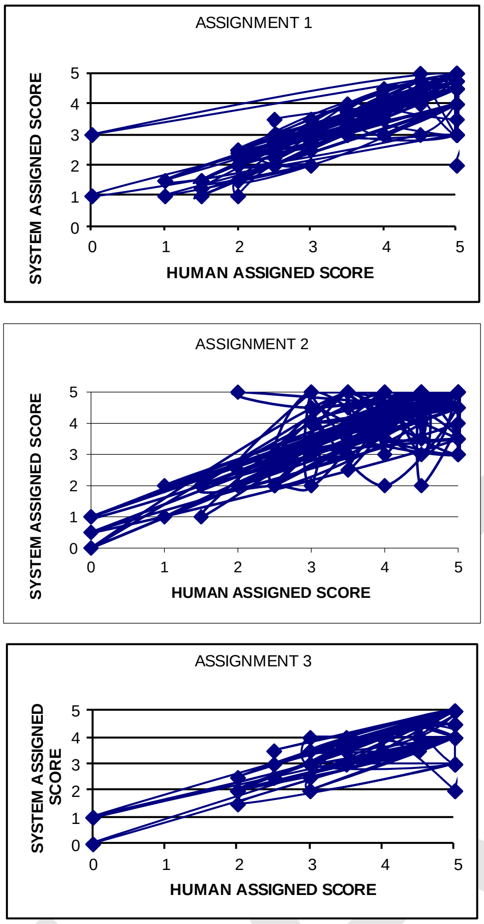


Figure 6. Scatter graph between the score assigned by human expert and the score assigned by our system (per assignment)

5.1. Comparison with other approaches

The most common representations of text for assessing content is LSA (Latent Semantic Analysis), vector based, WSD (Word Sense Disambiguation), Graph, BLEU, Ontology. These approaches can be classified into corpus based or knowledge based. Corpus based representations lose syntactic and semantic information which is essential to determine entailment. For our application, it is important to preserve the syntactic and semantic relationships between words as well as words. The dependency between the words is to be preserved. The knowledge-based representation depends on the particular domain.

Hu Xinming [11] and Klein Richard [14] present a system for assessing free text responses using LSA. LSA approach is more suited to assess essays than short-answer questions, since they focus on metrics, which broadly correlate with writing style, augmented with aggregate measures of vocabulary usage.

Gabrilovich and Markovich [8] proposed Explicit Semantic Analysis (ESA) method for fine-grained semantic representation of unrestricted natural language texts. They used knowledge encoded in Wikipedia to define vector space.

The Word Sense Disambiguation (WSD) approach described by Tacoa Francisco [29] to convert Natural Language into Common Semantic Description is a semi automated one. So, human intervention is required to select the meaning of a word. They use Universal Networking Language (UNL) Ontology as the domain knowledge. Some verb contains ambiguous words in the ontology. This approach fully depends on the UNL ontology. This is not suitable for assessing short answers.

Haghighi Aria et. al, [9] converted the text into graph and used graph matching algorithm to decide if the hypothesis is "entailed" by the text. P´erez [21] used BLEU approach along with LSA for assessing free text answers. Initially they applied BLEU method for assessing free text answers. But BLEU approach performs only syntactic analysis of the answers. Hence BLEU was combined with LSA to perform semantic analysis, even though it does not consider the dependency relation between concepts. Background knowledge in terms of Ontology was used for text clustering [10]. Matching percentage of nouns, verbs, adverbs, adjectives in the model answer and student answer is used to assess the unstructured text answers [25]. It does not consider the semantic relation between words. Comparison of our approach with other approaches is given in Table 4.

Table 4. Comparison with other approaches.

Approach	Pearson Correlation with Human grade
LSA [11], [14]	0.60
ESA [8]	0.72
WSD [29]	0.701205
Graph [9]	0.698201
BLEU [21]	0.686185
BLEU with LSA [21]	0.715257
Ontology [10]	0.652961
NLP with OntoMappping(Proposed Model)	0.790605

6. Conclusion

Assessment is used to assess the learners' understanding on the concept learnt. E-assessment is a powerful tool that automates the assessment task. There are several tools, e-learning systems and Intelligent Tutoring Systems available in the market which uses e-assessment to assess their learners' knowledge level. Most of the systems use only Multiple Choice testing to test their learners. But it is not enough to assess the student thoroughly. For in-depth assessment we need to use subjective type questions. In this paper, we proposed and implemented a technique to assess the short answer automatically using ontology mapping. We used the data set to evaluate our system and compared our system with some existing systems. Experimental results show that our system outperforms some existing systems.

Results are encouraging. In the future work, we would like to improve the accuracy of results. We are planning to construct the knowledge base and use that knowledge base to evaluate the answer.

References

[1] Alotaibi, Shaha, T. & Mirza, Abdulrahman, A., "Hybrid Approach for Automatic Short Answer Marking," *in proceedings of 2012 Southwest Decision Sciences Institute Conference (SWDSI 2012)*, pp. 581-589, 2012

[2] Attali, Y., & Burstein, J., "Automated essay scoring with e-rater® V. 2," *The Journal of Technology, Learning and Assessment*, vol. 4, no. 3, 2006.

[3] Augenstein, I., Padó, S., & Rudolph, S., "LODifier: generating linked data from unstructured text.," *In The Semantic Web: Research and Applications*, pp. 210-224. Springer Berlin Heidelberg, 2012.

[4] Briscoe, T., Medlock, B., & Andersen, Ø., "Automated assessment of ESOL free text examinations.," *University of Cambridge, Computer Laboratory, Technical Report UCAM-CL-TR-790* , 2010.

[5] Cherfi, H., Corby, O., Faron-Zucker, C., Khelif, K., & Nguyen, M. T., "Semantic Annotation of Texts with RDF Graph Contexts," *In ICCS Supplement*, pp. 75-82, 2008.

[6] Cubric, M., & Tosic, M., "Towards automatic generation of e-assessment using semantic web technologies," *International Journal of e-Assessment*, 2011.

[7] De Marneffe, M. C., & Manning, C. D., "Stanford typed dependencies manual," *URL http://nlp. stanford. Edu /software/ dependencies manual. pdf* , 2008.

[8] Gabrilovich, E., & Markovitch, S., "Wikipedia-based semantic interpretation for natural language processing," *Journal of Artificial Intelligence Research*, vol. 34, no. 2, pp. 443, 2009.

[9] Haghighi, A. D., Ng, A. Y., & Manning, C. D., "Robust textual inference via graph matching," *In Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pp. 387-394. Association for Computational Linguistics, 2005.

[10] Hotho, A., Maedche, A., & Staab, S., "Text clustering based on good aggregations.," *In Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pp. 607-608. IEEE, 2001.

[11] Hu, X., & Xia, H., "Automated Assessment System for Subjective Questions Based on LSI," *In Intelligent Information Technology and Security Informatics (IITSI), 2010 Third International Symposium on*, pp. 250-254. IEEE, 2010.

[12] Jordan, S., & Mitchell, T., "e-Assessment for learning? The potential of short-answer free-text questions with tailored feedback," *British Journal of Educational Technology*, vol.40, no. 2, pp. 371-385, 2009.

[13] Kaur, N., & Jyoti, K., "Automated Assessment of Short One-Line Free-Text Responses In Computer Science.," *International Journal of Computer Science and Informatics*, vol.2, no. 1,pp.2, 2012.

[14] Klein, R., Kyrilov, A., & Tokman, M., "Automated assessment of short free-text responses in computer science using latent semantic analysis," *In Proceedings of the 16th annual joint conference on Innovation and technology in computer science education*, pp. 158-162. ACM, 2011.

[15] Leacock, C., & Chodorow, M., "C-rater: Automated scoring of short-answer questions," *Computers and the Humanities*, vol.37, no. 4, pp. 389-405, 2003.

[16] Li, Yunyao, et al., "Facilitating pattern discovery for relation extraction with semantic-signature-based clustering," *In Proceedings of the 20th ACM international conference on Information and knowledge management*, pp. 1415-1424. ACM, 2011.

[17] Liu, H., Bao, H., & Xu, D., "Concept vector for semantic similarity and relatedness based on WordNet structure," *Journal of Systems and Software*, vol.85, no. 2, pp.370-381, 2012.

[18] Miarka, R., & Zácek, M., "Knowledge patterns for conversion of sentences in natural language into RDF graph language," *In Computer Science and Information Systems (FedCSIS), 2011 Federated Conference on*, pp. 63-68. IEEE, 2011.

[19] Mitchell, Tom, et al., "Towards robust computerized marking of free-text responses," *In 6th International Computer Aided Assessment Conference*, Loughborough, 2002

[20] Mohler, M., & Mihalcea, R., "Text-to-text semantic similarity for automatic short answer grading," *In Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 567-575. Association for Computational Linguistics, 2009.

[21] Pérez, Diana, et al., "Automatic Assessment of Students' Free-Text Answers Underpinned by the

Combination of a BLEU-Inspired Algorithm and Latent Semantic Analysis," *In FLAIRS Conference*, pp. 358-363. 2005.

[22] Pulman, S. G., & Sukkarieh, J. Z., "Automatic short answer marking," *In Proceedings of the second workshop on Building Educational Applications Using NLP*, pp. 9-16. Association for Computational Linguistics, 2005.

[23] Ramakrishnan, C., Kochut, K. J., & Sheth, A. P., "A framework for schema-driven relationship discovery from unstructured text," *In The Semantic Web-ISWC 2006*, pp. 583-596. Springer Berlin Heidelberg, 2006.

[24] Rus, V., Graesser, A., & Desai, K., "Lexico-syntactic subsumption for textual entailment," *amsterdam studies in the theory and history of linguistic science series,* vol.4, no. 292, pp.187, 2007.

[25] Sami, B., Yasin, H., & Mohammad Yasin, M., "Automated Score Evaluation of Unstructured Text using Ontology," *International Journal of Computer Applications*, vol.39, no. 18, pp.19-22, 2012.

[26] Siddiqi, R., & Harrison, C. J., "Improving teaching and learning through automated short-answer marking," *Learning Technologies, IEEE Transactions on,* vol.3, no. 3, pp.237-249, 2010.

[27] Siddiqi, R., & Harrison, C. J., "On the automated assessment of short free-text responses," *In IAEA Conference Paper*, 2008.

[28] Sukkarieh, J. Z., Pulman, S. G., & Raikes, N., "Auto-marking: using computational linguistics to score short, free text responses," *In the annual conference of the International Association for Educational Assessment (IAEA)*, Manchester, UK. 2003.

[29] Tacoa, F., Uchida, H., & Ishizuka, M., "A Word Sense Disambiguation Approach for Converting Natural Language Text into a Common Semantic Description," *In Semantic Computing (ICSC), 2010 IEEE Fourth International Conference on*, pp. 478-486. IEEE, 2010.

[30] Thomas, Ani, et al., "Extracting Noun Phrases in Subject and Object Roles for Exploring Text Semantics," *International Journal on Computer Science and Engineering*, vol. 3, no. 1, 2011.

[31] Valenti, S., Neri, F., & Cucchiarelli, A., "An overview of current research on automated essay grading," *Journal of Information Technology Education*, vol.2, pp.319-330, 2003.

[32] Wojtinnek, Pia-Ramona, et al., "Conceptual knowledge acquisition using automatically generated large-scale semantic networks," *In Conceptual Structures: From Information to Intelligence*, pp. 203-206. Springer Berlin Heidelberg, 2010.

[33] Zhang, X., Cheng, G., & Qu, Y., "Ontology summarization based on rdf sentence graph," *In Proceedings of the 16th international conference on World Wide Web*, pp. 707-716. ACM, 2007.

V. Senthil Kumaran received his Master degree in Computer Science from Madurai Kamaraj University, Madurai, India. He is currently working as Senior Assistant Professor at the Department of Applied Mathematics and Computational Sciences, PSG College of Technology, Coimbatore, India. Since 2008, he is working towards his PhD in Computer Science from the Faculty of Science and Humanities, PSG College of Technology. His research is focused on semantic web, data mining, intelligent information retrieval and e-learning.

A. Sankar received his PhD in Computer Science from Bharathiar University, Coimbatore, India, in 2003. He is currently working as an Associate Professor at the Department of Computer Applications, PSG College of Technology, Coimbatore, India. He has more than 24 years of teaching and ten years of research experience. His research interests include agile software engineering, data mining, e-learning and networks.