Indian Journal of Science and Technology, Vol 9(48), DOI: 10.17485/ijst/2016/v9i48/107625, December 2016
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

A Novel Approach of Collaborative KBQA System using Ontology

R. Mervin* and A. Jaya

Department of Computer Science and Engineering, B.S. Abdur Rahman University, Vandalur, Chennai - 600048, Tamil Nadu, India; mervinvijay@gmail.com, jayavenkat2007@gmail.com

Abstract

The objective of the research is to provide relevant answers for the questions posted by the learners using Collaborative Knowledge Based Question Answering System (CKBQA) of Knowledge Extraction which provide the most suitable answers by sharing the ideas of various learners, when the user needs information. In the modern era numerous information available in the World Wide Web which is too difficult to get the required information. CKBQA system aims at retrieving precise information from a large collection of documents. It provides the most relevant answers rather than the irrelevant information by applying Latent Semantic Analysis. In collaborative learning the users benefit when the learner exposed to diverse viewpoints from the other learners with varied backgrounds and provides the best interactive teaching learning method where the interaction taken place between the learners and also the learners with the exports. The CKBQA system consists of four phases such as Question Preprocessing, Answer Evaluation, Concept Mapping of question with their answer and voting. Ontology plays the vital role for checking the semantic nature of the question and the answer. Sample of 100 questions with their Answers were taken and tested for their various question types and also checked with resoner which gave 89% of correctness.

Keywords: Collaborative Learning, Concept Mapping, Knowledge based Question Answering System, Latent Semantic Analysis, Ontology and Voting

1. Introduction

E-Learning is now becoming a booming technology that can confer many information resources to the end user. Search engines provide tremendous information to the learners, but they have to navigate many pages to find their relevant information. Now a day's retrieval of information becomes simplest task with the help of Question Answering System in which the user get the answers based on their question. The proposed KBQA system helps to give the suitable answer with collaborative learning.

In CKBQA, the users posted questions as they need to learn and the answers were given for the same. The user's questions and answers are analysed both syntactically and semantically with the help of Ontology. Ontology helps to relate the concepts and their subclasses in a structured manner to represent domain-specific concepts in order to promote the semantic capability of a QA system and it gives the semantic relationship between concepts.

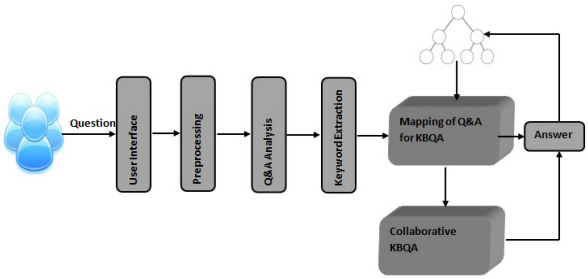

Figure 1. Conceptual diagram of collaborative KBQA.

In Collaborative Learning, the users have the opportunity to converse with peers, present and defend ideas, exchange diverse beliefs and question other conceptual

*Author for correspondence

frameworks. Here the posted question can be answered by other learners with KBQA and all the answers are prioritized based on voting. Then it is passed to the experts check the correctness and relevancy of the answers to the question.

The remaining part of the paper is organized as follows, Section II deals with Related Work, Section III describes Conceptual diagram of Collaborative KBQA and Section IV concludes the paper by giving a brief glimpse into the future directions of research in this area.

2. Related Work

Mervin et al. proposed a novel framework for KBQA mapping system using ontology[8]. It gave the detailed description of ontology construction and discussed the concept mapping of user queries with their answers. In this research paper, collaborative learning is discussed elaborately and is integrated with the KBQA system to make the research into next level.

Khune offered a mapping framework for Multiagent ontology which is having heterogeneous data in web and assigned the semantic relation using semantic relation interpreter which improves response time[1]. Multiagent system mapping system was developed to map the question with their answer. Kumar was proposed a methodology to extract the concept relation from the unstructured text using Naive Bayes Classifier[2]. Hand coded dependency parsing pattern rules and binary decision tree based rules are generated to accomplish the concept mapping to the seed concept. The triple set is used to pull out the feature of data which provides the key for question and answer mapping.

Arai et al. framed[4] an effective question answering system for collaborative learning, which can act not just like a virtual teacher, but also virtual discussion for student. Student can attach their question when they want collaborative using collaborative learning capitalize on one another's resources and skills. Students can ask their questions to the group when they want to collaborate with others, asking one another for information, evaluating one another's ideas. The method also considers that students can communication through Q&A interaction such a discussion forum to support information sharing. The different collaborative techniques[4–7] has been explained and discussed by various author.

3. Conceptual Diagram of Collaborative KBQA

The Figure 1 shows the conceptual diagram of KBQA and it comprises of following phases such as preprocessing of Q&A, Q&A Analysis, Keyword Extraction and Mapping of Q&A for KBQA. This research work explores the semantic mapping of question and suitable answers to the question, which are presented to the users. The query/question from the user will be taken as the input for question preprocessing which will be analyzed both syntactically and semantically. Then the answers posted by the others too analyzed in the same way. The key features of questions and answers are extracted that will be analyzed and correlated with their answer. Then the answer will be moved to the collaborative KBQA and all the answers are prioritized with the help of learners voting and the final answer is stored in the knowledge base.

3.1 Question Pre-processing

The Collaborative KBQA system utilizes, a Parts of Speech (POS) tagger which identifies the question words (why, what, how, where, when, which) and the keywords from the user query. For example, the question "What is Data Mining?" will be defined as follows:

What is Data Mining?

What- ProNoun (WP)

Is- Verb (VBZ)

Data Mining?-Noun (NN)

Once the POS tagger process gets completed, by using the question words, the parse tree was built with the help of Stanford Parser to understand the concept easily. Example of parse tree for question is shown in the Figure 2.

[S [WP What] [VBZ [V is] [^NN Data Mining?]]]

3.2 Answer Evaluation

Using the POS tagger and the Named Entity recognizer to narrow down the search of the keyword. For example, a query with the word who and when could point to person and a date. Now the next step is to find the exact field to match in-order to extract the answer that the user seeks. When question type has been successfully mapped to a top concept, only terms related to this concept will be

added to the term context representation. This way we obtain the terms that made up the context of a unique definition term. The following list of words which are used to identify the answer type for the question posted by the user.

1. who/whom - PERSON
2. when - TIME/DATE
3. where/what place - LOCATION
4. what time (of day) - TIME
5. what day (of the week) - DAY
6. what/which month - MONTH
7. what age/how old - AGE
8. what brand - PRODUCT
9. what - NAME
10. how far/tall/high - LENGTH
11. how large/big/small - AREA
12. how heavy - WEIGHT
13. how rich - MONEY
14. how often - FREQUENCY
15. how many - NUMBER
16. how long - LENGTH/DURATION

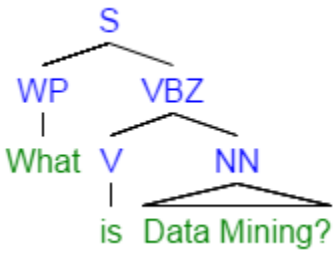

Figure 2. Parse tree for sample question.

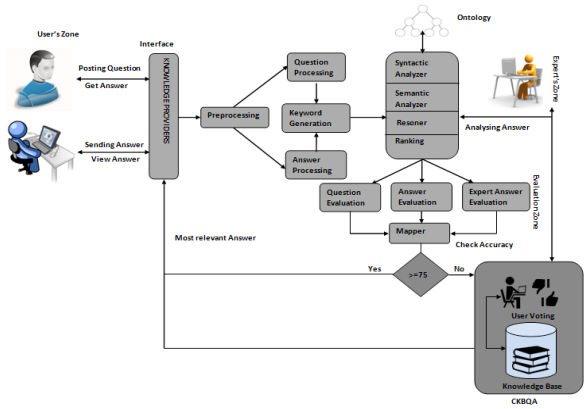

Figure 3. Flow diagram of CKBQA.

Figure 3 shows the flow diagram of CKBQA where the question is posted to the KBQA system. The accuracy of the answer is greater than 75% then is consider as the suitable answer which given back to the user else the same question is passed to the CKBQA system to get the answers with is voting. Again this is checked for accuracy and the process is repeated until the suitable answer is produced by the system. Every time the updated answer and their question is stored in the knowledge base. Figure 4 shows the Asserted Model for question type in KBQA. Initially the asserted model for KBQA was constructed with all possible classes and subclasses. Based on the integrative word, the keyword for the answer has been identified and the it is processed with answering.

Figure 4. Asserted model for question type.

3.3 Concept Mapping

The Figure 5 shows the M:N mapping of Q&A where M is the number of question for a single answer and N is the number of answer for a single question. The mapping of Q & A can be done in three possible way namely 1:M, N:1 and M:N.1:M describes that one question may have Many answers N:1 describes that N answers may have singe question. Finally M:N describes that M question may have N answers. Figure 6 shows the sample mapping of data mining questions and with their answers.

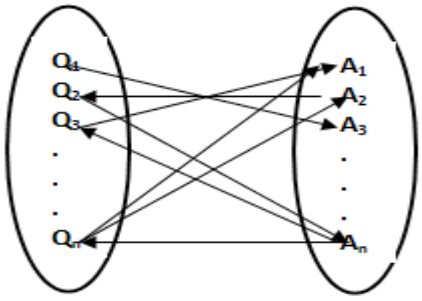

Figure 5. M: N mapping of Q & A.

Let A be the collection of Answers A={ $a_1,a_2,a_3...a_n$ } and K is the set of Keyword K={ $k_1,k_2,k_3...k_n$}.The keyword are extracted based on the count matrix supported to count the frequency.. TF-IDF is used for weight calculation which is done based on the count matrix.

TF(k) = (Number of times keyword 'k' appears in a Answer) / (Total number of terms in the Answers). Eqn (1)

IDF(k) = $\log_e$(Total number of Answers / Number of Answers with keyword 'k' in it) Eqn (2)

From equation (1) and (2), TF * IDF can be calculated as

*TF * IDF (keyword)* =TF(k)*IDF(k) Eqn (3)

where, *TF(term)* is the frequency of a term in the given document, IDF is Inverse Document Frequency. By referring the Equation (1),(2) and (3) the weight calculation was done.

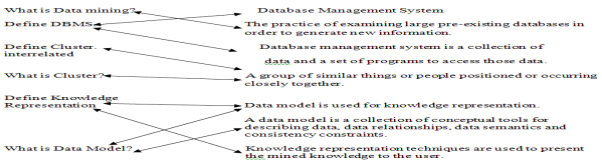

Figure 6. Sample mapping.

3.5 Voting

The answers for each question are analyzed by ever learners and each answer is ranked by the voting system. The voting system consists of "like" and "dislike". Every learner is giving vote for all the answers based on their knowledge and the preference. The different features of voting includes star, vote up, vote down, contributor, member since, activity level, total point, total answer, best answer. Table 1 shows the values which is produced by the KBQA and the CKBQA system and the Comparison chart is given in the Figure 7. Here the accuracy was checked for the different kind of question type with the different QA System and it produces the average of 89% accuracy based on the question type.

$$\text{Accuracy} = \frac{\Sigma rp + \Sigma rn}{rp + rn + wp + wn} \quad \text{Eqn (4)}$$

where, rp = right positive; rn = right negative; wp = wrong positive; wn = wrong negative

Eqn (4) shows that the accuracy calculation of the answers which is posted by the different learners. If the inference is right positive, then the answer is rejected. If the inference is right negative, then the answer is accepted. If the inference is wrong positive, then the answer is rejected. If the inference is wrong negative, then the answer is accepted.

Table 1. Comparison of KBQA

Words	With KBQA	With CKBQA
Data Mining	1.23	1.83
Computer	1.39	1.39
Information	1	1.14
Java	0.94	1.54
Data	1.7	1.61

4. Conclusion

CKBQA has been implemented and this experiment were done with 100 questions with their Answers and also done for the various question types. The syntactic and semantic structure of the question given by the user has been analyzed and it is mapped with the answer. The correctness of question with their answer mapping is also checked with reasoner and produces 89% of correctness which gave more accurate result to the user.

5. References

1. Khune RR. Mapping of semantic web ontology in user query system. International Journal of Computer Applications. 2015 Feb; 111(14).
2. Kumar GS, Zayaraj G. Concept relation extraction using Naı¨ve Bayes classifier for ontology-based question answering systems. Journal of King Saud University – Computer and Information Sciences. 2014:13–24.
3. Pundge AM, Khillare SA, Mahender CN. Question answering system, approaches and techniques: A review. International Journal of Computer Applications. 2016 May; 141(3).
4. Arai K, Handayani AN. Question answering system for an effective collaborative learning. International Journal of Advanced Computer Science and Applications. 2012; 3(1).
5. Mitnik R, Recabarren M, Nussbaum M, Soto A. Collaborative robotic instruction: A graph teaching experience. Computers and Education. 2009; 53(2):330–42.
6. Chiu MM. Effects of argumentation on group micro-creativity. Contemporary Educational Psychology. 2008; 33:383–402.

7. Chiu MM. Flowing toward correct contributions during groups' mathematics problem solving: A statistical discourse analysis. Journal of the Learning Sciences. 2008; 17(3):415–63.
8. Mervin R, Jaya A. A novel approach to mapping for KBQA system using ontology. Proceedings of Fourth International Conference on 'Emerging Research in Computing, Information, Communication and Applications' (ERCICA-2016); 2016. p. 9.
9. Wang C-C, Hung JC, Yang C-Y, Shih TK. An application of question answering system for collaborative learning. 26th IEEE International Conference in Distributed Computing Systems Workshops, ICDCS Workshops; 2006 Jul 04–07. p. 49.
10. Hu H. A study on question answering system using integrated retrieval method. PhD. Thesis submitted to Graduate School of Engineering at the University of Tokushima; 2006 Feb.