ABSTRACT

E-Learning is the trending technology that can confer many information resources to the end user. MOOC, Moodle, Coursera, IIT-Bombay and many open sources are available for doing online courses. Learners who are doing courses may have doubts and that may be clarified through internet resources. But the web sources are highly flooded with relevant and irrelevant information. There are many search engines available to provide tremendous information to the user. But the user has to navigate many pages to find their authenticated relevant information for their doubtful queries. Question Answering System (QAS) provides answers for the query posted by the user with the support of Information Retrieval (IR), Natural Language Processing (NLP) and ontology.

The reasoning based question answering system is relying on the KBQA (Knowledge Based Question Answering System) for providing the most relevant answer to the query posted by the user. This research work consists of four phases which include the semantic analyzer for questions and answers, Collaborative forum, Rule based mapping of questions and answers, and the ranking of answers.

The given user query is pre-processed and the keywords are extracted for retrieving the answers either from the repository or the web. In the semantic analyzer phase, the questions are analyzed both syntactically and semantically and the keyword is extracted using term frequency method. The answers for the questions are collected from the various sources like web, KBQA repository, and the collaborative forum. The retrieved answers are needed to analyze both syntactically and semantically. The semantic analysis helps to improve the accuracy of the system. The KBQA repository is updated periodically to serve the need of the user.
The correctness of the mapped question and their answers must be improved with the aid of reasoner. In Semantic Reasoner phase, Case Based Reasoning (CBR) is applied in case of retrieving the previously stored answers from the repository. If the answer for the question is not available in the repository, then it is passed to the collaborative forum to collect the answers from the various learners. All the answers are analyzed with its weight and their closure is measured. In the Collaborative forum, Send – A – Problem method is proposed where the users are allowed to give more than one answer for a single query posted by other users. The mapped question and answers are weighted with membership function of fuzzy reasoner based on the semantic closure.

The member function is applied to every question and their answers where the fuzzification and defuzzification are completed to validate and improve the correctness of the answers. All the rules are processed with the help of inference engine and the authenticated relevant answer will be stored in the knowledge base. The accuracy of the mapped question and answers are analyzed and the process is repeated until the suitable answer is produced by the KBQA system. Every time the updated answer and their question are stored in the knowledge base and it will be updated.

Consequently, in the mapping phase, the keywords of the question and the retrieved answers are conceived as the rules which are mapped into the ONTO-KBQA with the support of LP-LSA (Logarithmic Probabilistic Latent Semantic Analysis). LP-LSA is the concept mapping technique which helps for mapping the question with their related answers based on the similarity. It is integrated with SPARQL (Simple Protocol And RDF Query Language), SWRL (Semantic Web Rule Language) and Description Logic to create a rule based mapping system. With the support of KBQA inference engine, the questions and the answers are passed through the ‘N’ number of rules for the validation process which ensure its correctness of mapping.
Finally, all the answers need to be prioritized for the given query to cater the most relevant answer to the user. In the ranking phase, all the answers are ranked based on the threshold value, NDGC method and the user's voting is considerable into the account. The answers are said to be relevant when it crosses the threshold value. If the answer is not meeting the threshold value, then it is discarded from the ranking process. The selected relevant answers are evaluated by the NDCG method and prioritized based on the value. If there is a tie between the two answers then the user's voting is considered to provide the most relevant answer. The user can vote for the answer based on their opinion. The answers which get the highest value have been declared as the most relevant answer and the paired question and answer will be stored in the KBQA repository and the need of the user's requirement is fulfilled by the KBQA system.