CHAPTER 6

CONCLUSION AND SCOPE OF FUTURE EXTENSION

6.1 CONCLUSION

A successful implementation of question answering system occurs only if it produces a correct answer. The backbone of QA system is the understanding of users question in user point of view. The main aim of this thesis is to improve the performance of precision and recall ratios. AQUALOG QA system is a type of QA system based on ontology with semantic web environment. Query processing is done by query triple algorithm with syntactic type of search. Natural language queries are considered as triple set form in AQUALOG QA system and syntactic annotations with string metric algorithm is used.

The first problem in AQUALOG QA system is the syntactic search with string metric algorithm. String metric is a metric which measures the similarity or dissimilarity between two test strings for approximate string matching. In user point of view, a same keyword gives different meaning. It is called ambiguity. The second problem in AQUALOG QA system is the conceptual graph matching with syntactic mapping technique. The problem of getting correct answer from AQUALOG QA system is poor. The third problem is the syntactic indexing mechanism for answer retrieval in AQUALOG QA system. The answer retrieved from this system is poor by using the syntactic indexing
mechanism. The success rate of AQUALOG QA system is 63.5% as an average score. This AQUALOG QA system is not suitable for performance improvement based on syntactic search mechanism. So, the semantic search mechanism is used for query formulation, query processing and answer extraction in QA system.

The main metrics used for evaluation of question answering system are precision, recall, F-Measure and MRR ratios. This thesis proposed an implementation of semantic search concept in QA system based on automatic learning system with ontology and semantic web in closed domain environment. It has been designed with data structure ontology as closed domain and all types of questions have been given as input in this QAAL system. Data structure ontology has been constructed and the terms used in DS ontology have been created as classes in it. This proposed QAAL system has been designed to overcome the drawbacks of AQUALOG QA system for improving the performance of precision and recall ratios.

The first proposed method in this thesis is an improved question to query conversion algorithm. Users question has been analyzed and reformulation of question technique has been suggested for retrieval of answer from QAAL system. Manual mapping of users question and the query representation of ontology in QAAL system has been avoided in this algorithm by using a query template model. Collections of SPARQL query templates have been created as output of Q2Q algorithm and these are referred after reformulation of query in user side.

In the existing AQUALOG QA system, syntactic annotations are used and implemented for conversion of natural language query into query triple set. In this thesis, ambiguity is avoided by using semantic
search based Q2Q algorithm with query template. This problem has been noted and it has been solved in this proposed QAAL system by using Q2Q algorithm model with query template model. The performance based on precision and recall ratios in QAAL system has been improved compared to AQUALOG QA system by using Q2Q algorithm with query template model. Accuracy and recognition ratio performance have also been increased in the proposed QAAL system.

Totally 340 questions have been posted in QAAL system which are based on factoid based, list based, definition based, reason based and explanation based questions. In this set of questions, 270 are predicted type of question, 47 questions are asked from outside domain and the remaining 23 questions are in unpredicted form. Totally 420 possible query template models are created with 250 classes, properties and individuals in DS ontology. After conversion of query from question, there is a need of verifying query in correct format or not. So, sample SPARQL queries are generated answer is verified with ontology. Then, the analysis has been made for the answer and the average performance percentage has been reached to 82.96%.

Precision and recall ratios have been compared in QAAL system with AQUALOG QA system. The QA system performance has been increased to 2% in precision and recall values compared to existing
The second proposed method in this thesis has been an improved QGT algorithm with semantic similarity matching type based on concepts, relations and graphs. This algorithm has been used to map query terms with knowledge terms in ontology by using graph matching technique. In the proposed QAAL system, QGT algorithm has been suggested to improve the performance of precision and recall ratios by implementing graph matching technique based on semantic search. The results are compared in both precision ratio and recall ratio. Totally 6 test sets are used for implementation of QGT algorithm. Precision ratio has been increased to 2.42% in the proposed QAAL system compared with AQUALOG QA system. Recall ratio has been improved to 14% in QAAL system compared with the recall ratio in AQUALOG QA system by using QGT algorithm with semantic similarity matching type in QAAL system.

The F-Measure has been taken based on precision and recall ratios of proposed QAAL system and existed AQUALOG QA system. Six test sets have been taken and the F-measure has been improved 19.3% in QAAL system compared with the ratio in AQUALOG QA system. After implementing QGT algorithm with semantic similarity mechanism, MRR values have been taken based on ranking of answer. Six test sets have been used and the value has been taken in both QAAL system and
The third proposed method in QAAL system has been suggested as the semantic similarity based on integration of semantic indexing with inference mechanism. Compared to syntactic indexing, semantic indexing has gradually increased the performance of proposed QAAL system with 20% in precision, recall, F-measure and MRR values.

Three types of mechanisms are implemented and tested in the proposed QAAL system. These three types of searches are semantic search, semantic search with syntactic indexing and semantic search with semantic indexing mechanism. In this proposed QAAL system, these three types of mechanisms are implemented and the performance is evaluated based on precision, recall, F-measure and MRR values. With the comparison of all types of search, semantic search with semantic indexing and inference mechanism gives a successful answer which is used to improve the proposed QAAL system’s performance. MRR ratio of QAAL system is 84.12% which is high compared to 64% MRR ratio in AQUALOG QA system.

Confidence ratio is another metric used for verifying the retrieved answer is correct or not based on semantic indexing and inference mechanisms. Confidence ratio for semantic search indexing mechanism is taken in both QAAL system and AQUALOG QA system. The proposed QAAL system performance is improved 6.7% (approximately 7%) than the existing AQUALOG QA system by applying semantic search indexing mechanism. After implementing semantic inference in semantic search based on semantic indexing, the performance is improved high in proposed QAAL system. Compared to AQUALOG QA system, QAAL system performance is improved 6% by applying
This thesis illustrates the usage of question answering system in semantic web community with ontology based environment. The main goal of this thesis is to improve the precision and recall values of QA system. It is concluded that the proposed QAAL system by using improved Q2Q algorithm with query template, QGT algorithm with semantic similarity measurement and semantic search with integration of semantic indexing and inference method for answer extraction mechanism have been used to improve the precision and recall ratios in QA system compared with existing AQUALOG QA system successfully.

6.2 SCOPE OF FUTURE EXTENSION

This thesis describes the real time implementation of question answering system in semantic web with closed domain ontology by using improved Q2Q algorithm with query template method, improved QGT algorithm with semantic similarity measurement and semantic search with semantic indexing and inference mechanism for answer extraction. In this thesis, DS ontology is created manually with the collection of classes, methods and objects in it. The time consumption is high for setting class, properties of object and data type model as large setup. Time metric will be considered as a metric in future for automatic creation of ontology. In future, it is proposed to generate automatic creation of ontology in QAAL system in data structure ontology.
In question to query conversion, query template method is used for semantic search implementation. This query template model is created based on user’s question. The user point of analysis is mainly needed for the performance improvement. In future, this type of analysis may be taken as an automatic in query template model for user point of question analysis. This thesis suggests the performance improvement of QA system based on closed domain model. In future, open domain challenges may also be considered for implementing this QAAL system for retrieving answer from all domains. Time may also be taken as a metric for improving the performance of QA system in open domain ontology implementation in future.