ABSTRACT

A Question Answering (QA) system has a major role in web based search to retrieve correct result in current environment. The backbone of QA system is the understanding of user’s question which produces the expected result that should be satisfied by the user. Natural Language Processing (NLP) is used for understanding of user question. So, semantic search comes into a role for implementing the type of searching in real time. Semantic search is implemented for retrieving the data from Knowledge Base (KB) using ontology. Ontology is a technology used to enable the domain knowledge at a high level and improve the query time. Nowadays compared to current web model, semantic web produces the result with the meaning of data. Semantic web is machine readable information and automated services. It is based on text based manipulation but it uses machine based on metadata form. This type of web needs a standard format for storing arbitrary data. Resource Description Framework (RDF) is used to represent semantic web. RDF is a language which is developed for annotation setting and it describes the information resources and their relations among each other. Ontology is used to represent RDF form for implementing semantic web environment in QA system.

An ontoLOGy-portable QUestion Answering system (AQUALOG) is a closed domain QA system based on ontology with semantic web. In this
system, three problems are chosen for proposed work. The first problem in
AQUALOG QA system is query processing which is done by syntactic search
with query triple algorithm. The second problem is the syntactic graph
matching technique. Conceptual Graph (CG) matching technique is used for
query mapping which is also based on syntactic mapping. The third problem
is the syntactic indexing mechanism used in AQUALOG QA system.
AQUALOG QA system is not having high precision and recall ratios. The
aim of proposed QA system is the performance improvement in precision and
recall ratios.

This thesis proposes an improved QA system with high performance
based on semantic search and indexing mechanism. So, a Question Answering
based on Automatic Learning (QAAL) system with ontology and semantic
web in closed domain environment is suggested for implementing semantic
search. It is proposed to design data structure ontology as closed domain. Data
structure ontology is a model used to represent the terms available in data
structure in tree format. All patterns of questions are given as input in this
QAAL system. Questions are asked as factoid, list, definition, reason and
explanation based type. This proposed QAAL system is designed to overcome
the drawbacks of AQUALOG QA system for improving the performance of
QA system based on precision and recall ratios.
The first proposed method in this thesis is an improved Question to Query (Q2Q) conversion based on query template. Mapping of users question and the query representation in ontology is a challenging task and it has been avoided by using semantic search Q2Q with query template model in QAAL system for performance improvement. Question classification and reformulation takes a major role in this proposed Q2Q algorithm. The input of Q2Q algorithm is the different question types such as factoid based, list based, definition based, reason based and explanation based types and collections of Simple Protocol for RDF Query Language (SPARQL) with triple set is the output of Q2Q algorithm. The performance of QAAL system has improved by 2% in precision and recall values compared to AQUALOG QA system by using semantic search based Q2Q algorithm with query template model. Accuracy and recognition ratio performance have been increased by 6% and 3.16% respectively in the proposed QAAL QA system compared to AQUALOG QA system.

The second proposed method is the Question answering for automatic learning Graph Traversal (QGT) algorithm by mapping query terms
improved by 19.3% and 3% respectively in QAAL system compared with AQUALOG QA system.

The third proposed method in QAAL system is the semantic search with integration of semantic indexing with semantic inference mechanism. Inference is a method for knowledge acquisition and expansion. Semantic indexing is an index type which can make use of information extraction and annotation of index terms in ontology. Precision, recall, F-measure and MRR values have been used as performance metrics in this method and confidence ratio metric is used to identify the user satisfaction. The performance has been improved by 7% using semantic search with semantic indexing mechanism and 6% improvement has been made in QAAL system than AQUALOG QA system based on confidence ratio by using semantic search with semantic indexing and semantic inference mechanism.

It is concluded that the proposed QAAL system of Q2Q algorithm with query template, QGT algorithm with semantic similarity measurement and semantic search with semantic indexing and inference method have been