ABSTRACT

The evolution of knowledge for a domain is intended to solve the domain specific problems. The represented domain knowledge generally provides open world semantics to the endeavor. The assimilated domain knowledge in a representation framework for a specific purpose by the experts could be inoperable for other applications. This may be due to lack of explicitly represented knowledge or correlation between the related information. Hence, a suitable reasoning methodology is required to derive the implicit knowledge. However, the task of achieving automated reasoning is possible only with the sound representation of the elicited knowledge. But in general, the literature shows that the identification and representation of the semantic relationship between the concepts is a complex and tedious task. Thus, there is a need of system to infer the semantic relationships which can be used to enhance the open world semantics of the domain knowledge. This thesis exploits the impact of semantic relationships between the concepts in various applications such as information retrieval, ranking and question generation system.

This thesis is mainly focused on enhancing the open world semantics by integrating the inferred semantic relationships with the existing knowledge. The semantic relationships based reasoning (SemPER) algorithm is proposed for achieving the effective reasoning over semantic relationships. The formal model has been formulated using Description Logic (DL). The SemPER algorithm is designed to infer the semantic relationships for general revelation in other words, it discovers from the represented knowledge. The semantic derivation principle has also been derived for reasoning. In this research work, the consistency and the satisfiability test have been performed over the inferred semantic relationships. The completeness and the soundness of the algorithm have also been proved. The optimizations have been performed to reduce the time complexity of the reasoning algorithm.

Generally in the information search, the number of documents is retrieved on the basis of the given keywords or concepts. However, it could not recognize the potential semantic relationships between the given concepts. Hence, retrieval of information with semantic relationships (RISeR) is proposed which helps to retrieve more relevant documents by identifying the semantic relationships between the given concepts. The retrieved documents are ranked on the basis of number of hits and links. In most of
the scenario, it is unusable. As a result, ranking using semantic relationships (RaSeR) is proposed. It is used to rank the most relevant documents according to the user query using inferred semantic relationships. The prospect of applying the semantic relationships to the question generation system can revolutionize the learning experience. The task of generating questions from the existing information is a tedious task. Thus, the Question generation system based on semantic relationships (Q-Genesis) is proposed to generate more relevant knowledge level questions automatically. It will be useful for the trainer to assess the knowledge level of the learners.

The computer network ontology is developed using protégé and used with other publicly available ontologies and datasets, for the evaluation of the proposed SemPER, RISeR, RaSeR and Q-Genesis algorithms. These algorithms are implemented in Java using eclipse environment. These algorithms are tested with various parameters such as Precision, Recall, F-Measure, Accuracy, Mean Average Precision (MAP), Discounted Cumulative Gain (DCG), Normalized Discounted Cumulative Gain (NDCG), Mathew Correlation Coefficient (MCC) and pseudo relevance feedback method. SemPER algorithm is experimentally proved with precision, recall and F-Measure. The correctness and accuracy of the SemPER algorithm are also verified. The correctness of the SemPER algorithm is statistically proved using MCC. The effectiveness of the RISeR algorithm is experimentally proved with Precision@10, MAP, DCG and NDCG. The significance of the RISeR algorithm is proved with T-Test. The efficiency of the RaSeR algorithm is experimentally proved with precision@10. The accuracy of the generated questions using Q-Genesis is experimentally proved. Finally, comparisons have been made with the existing algorithms to prove the effectiveness of the proposed algorithms.

The conclusion derived from this thesis is that the open world semantics can be enhanced by revolutionizing the reasoning methodology with semantic relationships. Thus it could be used for gen revelation. The semantic relationships are substantially exploited in information retrieval, ranking and question generation system and achieved improved results. It is also proved experimentally and statistically. The future direction of this research includes context and time factors in reasoning and ranking. More number of synthesis and application level questions can be generated by employing the semantic relationships.