Semantic ambiguous query formulation using statistical linguistics technique

ABSTRACT

Natural language query systems mitigate the complexity of structured query. Usually, natural language processing is implemented to solve several problems, such as information retrieval. However, problems such as natural language ambiguity remain unsolved due to the complexity of natural language itself. This issue thus requires further research. Recent studies on semantic query formulation have attempted to resolve ambiguous natural language by proposing different disambiguation approaches. Most such processes are either implemented manually or semiautomated. In the same vein, most recent systems solve ambiguity by using an external dictionary such as WordNet or by providing suggestions manually. The present research proposes a statistical linguistic technique for solving the problem of ambiguity automatically. The proposed technique is experimentally tested on a Quran ontology with queries from the Islamic Research Foundation Website and increases the result of precision and recall by 6% and 10%, respectively.

Keyword: Statistical linguistics technique; Semantic technology; Information retrieval; Ontology; Islamic knowledge