

Skolem preprocessing using WordNet and lexicon in building effective knowledge representation

ABSTRACT

We are in the information intensive environment in which various forms of digital contents have been growing exponentially. In this era of digital data, knowledge representation has been considered as a crucial component of any information retrieval system. It is also considered as a major problem especially in representing the content of unstructured text in an effective way. Although the mission remains impossible to achieve 100% accuracy, many researchers are indulging themselves in documenting these data in many different techniques so that it can be communicated effectively and easily. Indexing is an important element that determines the success of retrieval. Since we are dealing with multiple documents, preprocessing of data is needed before the data gets indexed. Thus, this paper presents an approach on the preprocessing technique. The semantic data which have been represented in skolem clauses will be preprocessed with the help of automatic lexicon generator output and WordNet. This preprocessing plays an important role in getting rid of redundant data before it gets indexed into the semantic matrix. Besides redundancy, it also helps in dealing with common problem that exists in indexing multiple documents in which similar sentences with more or less the same meaning but have been constructed by using different sets of words. As a conclusion, the integration of WordNet and lexicon leads to better result in terms of building effective knowledge representation.

Keyword: Skolem; Lexicon; Indexing; Preprocessing