Abstract

On the development of Indonesian WordNet, synonym set is an important part that represents the similarity of meaning between words. Synonym sets are built using the Indonesian Thesaurus as the lexical database used. After going through the extraction process from the Indonesian Thesaurus, we will get a synonym set that has a similarity or word sense between words. In general, the difference between WordNet and the dictionary is its main focus, the dictionary usually focuses on just one word, while in WordNet the focus is on the meaning of words and connectedness with other words. In previous studies the construction of synonym sets was carried out using several approaches, one of which is clustering to produce synonym sets and WSD (Word Sense Disambiguation). In this research the approach way used to produce synonym sets is clustering using ROCK (Robust Clustering Using Links) algorithm which uses similarity and link values. The resulting synonym set will then be used for lexical database development. Therefore the main focus of this research is to produce synonym sets through the clustering process and calculate their accuracy, for performance calculation and evaluation using the F-measure method involving the gold standard.

Keywords: WordNet, ROCK Clustering, Synonim Sets.