

Development Synonym Set for the English Wordnet Using the Method of Comutative and Agglomerative Clustering

Munirsyah^{[1]*}, Moch. Arif Bijaksana^[2], Widi Astuti^[3]

Bachelor of Informatics Engineering, Faculty of Informatics
Telkom University, Bandung, Indonesia

^[1]munirsyahh@student.telkomuniversity.ac.id, ^[2]arifbijaksana@telkomuniversity.ac.id, ^[3]astutiwidi@telkomuniversity.ac.id

Abstract— Wordnet is a collection of words that interpret or present a meaning, in its development Wordnet has an important part, the Synonym Set or Synset. In making Synonym sets, synonyms are needed and the commutative nature of words is needed. To get word synonyms, the English language thesaurus becomes the reference data for taking synonym data. Broadly speaking, the difference between Wordnet and the dictionary is that the meaning of the word is related to other words, to determine the equation requires a commutative process. The process is made easy by using commutative methods that will produce a candidate synonym set. Candidates for the synonym set cannot be used for word syntax, the grouping process of words which produces the Synonym set as the final result must be carried out. The process of grouping words can one of them use clustering techniques, in this study will use Agglomerative Clustering techniques. In the process of agglomerative clustering techniques there is a threshold value to determine the number of repetitions or as a condition to stop the iteration process. The clustering process in this study will use a threshold value of 0.1 to 1 to test the best threshold value to produce the best Synonym set and calculate its accuracy value. Accuracy calculation and evaluation will use the F-measure method to find the best results.

Keywords— Wordnet, Synonym Set, Agglomerative Clustering