ABSTRACT

Automatic summarization is a system used to help human to extract core information of a long text instantly. System capable to help generate a summary automatically. There's a lot of system developed, but there's still problem to solve. One of the problem is word ambiguity sometimes neglected, thus a different word with same meaning treated as a two different word. This word ambiguity is decreasing automatic summarization system performance. In this research, document index graph algorithm proposed as a summarization method alongside TextRank algorithm that used to score nodes in graph. Texrank algorithm adapted from pagerank to score a directed weighted graph. Document index graph algorithm is combined with lesk, a word sense disambiguation algorithm. Lesk algorithm adapted to user wordnet as a knowledge based. Based on this research result, word sense disambiguation implementation in automatic summarization resulting a decerase in overall system performace. Highest decrease is happening on <5 zipf frequency on 80% compression.

Keyword*: text summarization, word sense disambiguation, document index graph, TextRank