

Rules frequency order stemmer for Malay language.

ABSTRACT

The importance of stemmer is obvious with the advent of effective information retrieval systems. Unfortunately, Malay stemming problems are difficult to solve due to complexity of words morphology. The Rules Application Order (RAO)stemmer is examined for enhancing performance to minimize the percentage of stemming errors. This paper presents a stemming approach called Rules Frequency Order (RFO). RFO rearranges the stemming rules according to the frequency of their usage from the previous execution. It shows that the approach provides a higher percentage of stemming correctness as compared to RAO stemming approach.

Keyword: Stemming; Stemmer; Information retrieval; Malay language.