Pola grammar technique for grammatical relation extraction in Malay language

ABSTRACT

A basic sentence in Malay language is either a combination of NP+NP, NP+VP, NP+PP, or NP+AP. The language is a structure phrase grammar. The Context Free Grammar was developed by Nik Safiah (1993). However, in order to derive a parse tree for a syntactic process, the CFG was found to be complicated due to many ambiguities for part of speech. This paper shall introduce a pola-grammar technique that does not require lexical process of retrieving the part of speech for each word. The techniques used are the automata and the finite states. During the process, sentences will be grouped into an adjunct, a subject, a post-subject, a conjunction and a predicate. The predicate consists of a verb, a conjunction, an object and an adverb. An adverb consists of a verb, a conjunction and an adverb. These components which are in sequence order are called pola. The subject, the object and the verb can be identified by making use of the pola in the sentence. This can be done by splitting the sentence into its pola. The information in the predicate will be processed to get the verb and object.

Keyword: Grammar; Grammatical relation; Malay language; Pola grammar