

# INDONESIAN LANGUAGE QUESTION ANALYSIS FOR QUESTION ANSWERING SYSTEM (QAS)

**Kartina, Julio Adisantoso**

Department of Computer Science

Faculty of Mathematics and Natural Sciences

Bogor Agricultural University

julioipb@gmail.com

2010

## **Abstract**

Question analysis is the first step in QAS. This step decides the final result of QAS process because the result of question analysis is used to retrieve relevant document and answer entity correctly. The used question query is limited to question type: WHO, WHERE, WHEN, and HOW MANY or HOW MUCH. The question word on query is used to obtain the answer candidate, while other words beside the question word are used to analyze the question. Question analysis process is started by parsing the keyword become tokens. The parsing is conducted with regarding token's pair possibility as a phrase. This phrase formation is expected to be able to keep the semantic aspect of question sentence. The question sentence that has parsed is used to retrieve document and top passage. Top passage is obtained through heuristic scoring. The answer extraction is conducted by calculating the nearest distance between each answer candidate in top passage and each word in keyword. Answer correction is evaluated by using these criteria: right, unsupported, wrong, and null.

The evaluation result of system showed the more correct answer for the less number of documents. The result of 106 tested documents was 75.56 % for criteria right, 2.22 % for criteria unsupported, 17.78 % for criteria wrong, and 4.44 % for criteria null. The result of 200 tested documents was 73.33 % for criteria right, 22.22 % for criteria wrong, and 4.44 % for criteria null. The result of 300 tested documents was 71.11 % for criteria right, 22.22 % for criteria wrong, and 6.67 % for criteria null. This decreasing of percentage is caused by the less correct of passage scoring method that was used.