University of Indonesia Participation at Query Answering-CLEF 2005

Mirna Adriani and Rinawati
Faculty of Computer Science
University of Indonesia
Depok 16424, Indonesia
(mirna@cs.ui.ac.id, rinaw101@mhs.cs.ui.ac.id)

Abstract. We present a report on our participation in the Indonesian-English question-answering task of the 2005 Cross-Language Evaluation Forum (CLEF). We chose to translate an Indonesian query set into English using a commercial machine translation tool called *Transtool*. We used a linguistic tool, the Monty Tagger, to find the answer to the question in a passage that has the same tagging as the query.

1   Introduction

The University of Indonesia IR-group participate in the bilingual Query-Answering (QA) task in Cross Language Evaluation Forum (CLEF) 2005, i.e., the Indonesian-English QA. We use commercial machine translation software called *Transtool*[1] to translate an Indonesian query set into English. We learned from our previous work [2] that freely available dictionaries on the Internet did not provide sufficiently good translation terms, as their vocabulary was very limited. We hoped that we could improve the result using machine translation.

2   The Question Answering Process

There were a number of steps to process the queries that we received from CLEF. Since what we got was English queries, we manually translated the 200 original English queries from CLEF into Indonesian. We then applied a tagging algorithm [2, 3] that we developed to categorize Indonesian queries according to the type of question, such as location, person, date, measure, etc. So, the question word '*dimana'* (*where*) triggers the algorithm to insert a location tag to the query. After applying the tagger, the Indonesian query was then translated back into English using *Transtool*.

2.1   Finding the Answer in the Document

The translated queries were run using an information retrieval system to retrieve the top 100 documents. These top 100 documents were then split into passages each containing two sentences. Each passage was then ran through Monty Tagger[2] to identify the name entity of each word such as NN, NNP for nouns, and CD for numbers, etc. Based on the category of the query, i.e., based on its query tag: location, person, etc., we specified that terms with NN tags are the answers to location type questions, and terms with CD tags are the answer to date type questions, and so forth.

In order to give the score to the passages, we use a similar scoring technique as the one used by Li and Croft [4] in their QA work. The scoring rules are as follows:

---

[1] See http://www.geocities.com/cdpenerjemah/.
[2] See http://web.media.mit.edu/~hugo/montytagger/.

1. Give 0 to a passage if its tag is not the same as the query tag.
2. Give 0 to a passage if the number of words in the query is smaller than some specified threshold, otherwise give the passage a score equal to the number of words in the query (count_m).
3. Add 0.5 if all words in the query are in the passage.
4. Add 0.5 if the order of words in the passage is the same as the query.
5. Calculate the final score of the passage:

    Score = score + count_m / passage_size

    where:

        count_m = the number of the words in a query.

        passage_size = the number of the words in the passage.

Once the passages obtained their scores, get the top 20 passages that have the highest scores. The answer to the question in the passage can be estimated by calculating the distance between the candidate words and the queries words. The passage that has the smallest distance is the answer.

3   Experiment

The collection contains English documents from the *Glasgow Herald* and the *Los Angeles Time*s. There are 200 queries that are being used in this year QA task. We use *Lemur*[3] information retrieval system to index and retrieve the documents.

4   Results

Our work focused on the bilingual task using Indonesian queries to retrieve English documents. The result that we submitted is the translated queries from using the machine translation tool. The average number of words in the original English queries is 8.50 words, in the Indonesian queries is 7.89 words, and in the translated English queries is 8.94 words. Transtool failed to translate 8 Indonesian words into English.

Our result shows that only two correct answers are being found (Table 1). There were 36 inexact (ambiguous) answers and 162 wrong answers. One of the reasons why the result was so poor was because our tagger did not provide specific enough tagging to the passages. As a result, the tagging in most passages was too general. For example NN tags could be the answer to questions about location or about organization. We did not have enough time to analyze the results before submitting them and correct this problem.

Task : Bilingual QA	Evaluation
W (wrong)	162
U (unsupported)	0
X (inexact)	36
R (right)	2

Table 1. evaluation of the QA result.

4   Summary

Our participation in the QA-CLEF this year provided us with invaluable experience in working on the QA task. The result of using the algorithm that we developed in this work was relatively poor. Our plan for the future is to use better linguistic tools and improve the scoring algorithm to produce more accurate answers.

---

[3] See http://www.lemurproject.org/.

5 References

1. Adriani, M. Bilingual CLEF for English and Dutch. In *CLEF 2001 Working Note Workshop.* Germany, September 2001.
2. Clarke, C. L. A., G. G. Cormack, D. I. E. Kisman and K. Lynam. Question Answering by Passage Selection: *The 9th Text retrieval Conference (TREC-9).* 2000.
3. Hull, David. Xerox TREC-8 Question Answering Track Report: *The 8th Text Retrieval Conference (TREC-8).* 1999.
4. Li, Xiaoyan dan Croft, Bruce. *Evaluating Question-Answering Techniques in Chinese.* In *NIST Special Publication: The 10th Text Retrieval Conference (TREC-10).* 2001.