CS533

Assignment No. 5

April 4, 2009 (2nd version)

April 2, 2009 (1st version)

Due Date: April 9, Thursday, 23:59.

This assignment is on query-oriented summarization for Turkish. Another possible description for it is query-biased summarization.

We plan to use a summarization system in our Turkish Bilkent News Portal. For example, documents matching a query will be summarized by paying attention to query words (your query words are the keywords to be used during summarization – see below-).

- 1. Please summarize ten different news articles.
- 2. Use the ROSS system for summarization. For this purpose please use the link Summarization User Interface provided on the following web page http://www.ug.bcc.bilkent.edu.tr/~cakar/. Please do this every time you use the system since the IP address may change.
- 3. Your user id (name) is your first name with small case letters (no Turkish character is allowed: "erdoğan" becomes "erdogan") and your password is your Bilkent id.
- 4. You have to summarize a chosen news article by extracting sentences. Please perform your summarization from the view point of keywords (anahtar kelimeler). What does this mean? For example, the news article that you are summarizing may be retrieved by a search engine by using a set of query terms. Within the context of this summarization task such query terms are referred to as keywords (anahtar kelimeler).
- 5. Does it mean that the extracted sentences must include the keywords? Not necessarily, if it is done like that then generating a summary would be mechanical (remember Google snippets).
- 6. For selecting a sentence, please click the tick box that comes before the sentence.
- 7. Is it possible to summarize an article that does not contain any of the user keywords? Yes. I had that experience and it was quite meaningful.
- 8. After summarizing a news article (by sentence extraction) please enter approximate amount of time you spent for the summary.
- 9. Please note that after submitting a summary you cannot undo that. Therefore, please do your best at your first attempt.
- 10. For more information about this project please refer to the Web site at the following address: ttp://www.ug.bcc.bilkent.edu.tr/~cakar/
- 11. We have identified a security problem and are trying to resolve this issue. If the system is unavailable please come back later and try again.

Please turn the page for some ROSS screenshots.



