APPENDIX 2

POST-TASK QUESTIONNAIRE

Questionnaire used during real time testing of the system by user objects

Score assigned for Options
Option a: 4; Option b: 3;
Option c: 2; Option d: 1;
Final score has been scaled to 10 point for convenient comparison and evaluation.

Meaning of terminologies used in this questionnaire:
True Positive sentences : Presence of expected sentences in the summary
False Positive sentences : Presence of not-expected sentences in the summary
False Negative sentences : Absence of expected sentences in the summary
True Negative sentences : Absence of not-expected sentences in the summary

1. How is the response time of the summarization system?
   a. Excellent
   b. Good
   c. Satisfactory
   d. Not Satisfactory

2. Are you getting expected content from the web page selected?
   a. 85%-100%
   b. 70%-84%
   c. 55%-70%
   d. below 55%

3. Does the system produce better results than search engine snippets?
   a. Very Good
   b. Good
   c. Similar to snippets
   d. Below the level of snippets

4. Does the system perform better than summarizers (MEAD) without segmentation?
   a. Very Good
   b. Good
   c. Similar to the other summarizer
   d. Lower than other summarizer
5. Are you satisfied with the presentation (Web page) of the composed summary?
   a. Very Good  
   b. Good  
   c. Satisfactory  
   d. Not Satisfactory

6. What is the proportion of true positive sentences in the summary compared to human generated summary?
   a. Above 90%  
   b. 75% to 90%  
   c. 50%-75%  
   d. Below 50%

7. What is the proportion of false positive sentences in the summary compared to human generated summary?
   a. Below 50%  
   b. 50%-75%  
   c. 75% to 90%  
   d. Above 90%

8. Are the sentences arranged as expected?
   a. All sentences are in order  
   b. Above average number of sentence are in order  
   c. Average number of sentence are in order  
   d. Below average number of sentence are in order

9. Important of a sentence is decided according on
   a. Presentation methodology-highlighted contents  
   b. Occurrence of key words  
   c. Query related contents  
   d. All the three

10. Usefulness of the summary
    a. Excellent  
    b. Good  
    c. Satisfactory  
    d. Not satisfactory
User Interface for offline segmentation

Figure A3.1 shows the user interface through which the system administrator performs the offline segmentation process using cosine similarity based segmentation and also frequent terms set and semantic relevance based segmentation approaches.

Figure A3.1 User interface for offline segmentation process
**Topic Block Pre-processing**

Topic blocks of the web documents are stored in relational database table and are processed further to measure the similarity between them. Figure A3.2 shows the pre-topic blocks processing steps implemented as part of the system and Figure A3.3 shows the search engine query interface of the system.

![Figure A3.2 Topic block pre-processing steps](image1)

Figure A3.2 Topic block pre-processing steps

![Figure A3.3 Search engine query interface](image2)

Figure A3.3 Search engine query interface
Search Result and Selection of URLs for comparative summary generation

Search engine component of the system presents the user with a set of matching URLs from which a set of URLs are selected for which the comparative summary is to be generated and the user interface is shown in Figure A3.4. Composed comparative summary is shown in Figure A3.5.

Figure A3.4  Search result and selection of URL

Figure A3.5  Comparative summary