Homework 1

Collaboration and Originality

Your report must include answers to the following questions:

1. Did you receive help of any kind from anyone in developing your software for this assignment (Yes or No)? It is not necessary to describe discussions with the instructor or TAs.

   If you answered Yes, provide the name(s) of anyone who provided help, and describe the type of help that you received.

2. Did you give help of any kind to anyone in developing their software for this assignment (Yes or No)?

   If you answered Yes, provide the name(s) of anyone that you helped, and describe the type of help that you provided.

3. Are you the author of every line of source code submitted for this assignment (Yes or No)? It is not necessary to mention software provided by the instructor.

   If you answered No:
   a. identify the software that you did not write,
   b. explain where it came from, and
   c. explain why you used it.

4. Are you the author of every word of your report (Yes or No)?

   If you answered No:
   a. identify the text that you did not write,
   b. explain where it came from, and
   c. explain why you used it.
Instructions

Instructions are shown in a red italic bold font. Do not include instructions in your report. We will deduct points for leaving instructions in your final report.

1 Structured query set

1.1 Summary of query structuring strategies

Briefly describe your strategies for creating structured queries. These should be general strategies, i.e., not specific to any particular query.

1.2 Structured queries

List your structured queries. For each query, provide a brief (1-2 sentences) discussion of:

1. which strategy (from Question 1.1) was used for that query,
2. any important deviations from your default strategies, and
3. your intent, i.e., why you thought that particular structure was a good choice.

2 Experimental results

Present the complete set of experimental results. Include the precision and running time results described above. Present these in a tabular form (see below) so that it is easy to compare the results for each algorithm.

2.1 Unranked Boolean

<table>
<thead>
<tr>
<th></th>
<th>BOW #OR</th>
<th>BOW #AND</th>
<th>Structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>P@10</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>P@20</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>P@30</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>MAP</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Running Time</td>
<td>mm:ss</td>
<td>mm:ss</td>
<td>mm:ss</td>
</tr>
</tbody>
</table>

2.2 Ranked Boolean

<table>
<thead>
<tr>
<th></th>
<th>BOW #OR</th>
<th>BOW #AND</th>
<th>Structured</th>
</tr>
</thead>
<tbody>
<tr>
<td>P@10</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
3 Analysis of results: Queries and ranking algorithms

Discuss your observations about the differences between the three different approaches to forming queries, and the two different approaches to retrieving documents (i.e., retrieval models) in terms of their retrieval performance and total running time.

Hint: Do not just summarize the results from the previous sections; we can see those results above. You are expected to provide your interpretation of the results based on what you learned in the lectures and readings. This is your chance to show what you learned from this homework assignment - take this section very seriously.

Hint: Probably this section doesn’t need to be longer than ¾ of a page (not counting these instructions).

4 Analysis of results: Query operators and fields

Discuss the effectiveness, strengths, and weaknesses of the query operators and fields, and your success and failure at using them in queries. Did they satisfy your expectations?

Hint: Same hints as above.