# Assignment 5

CS 410/510: Information Retrieval on the Internet
Due Tuesday February 13, 2007
8 points

**Goal**: The goal of this assignment is work through some calculations of evaluation metrics for a toy-sized test collection and experimental results so that you can be sure you understand the metrics.

**Details**:
Assume the Homework Test Collection consists of the following 40 documents, three queries, and relevance assessments for the three queries.

| **Documents in the Homework Test Collection** | | | |
|---|---|---|---|
| **Document ID** | **Document ID** | **Document ID** | **Document ID** |
| 0011 | 0023 | 0035 | 0047 |
| 0012 | 0024 | 0036 | 0048 |
| 0013 | 0025 | 0037 | 0051 |
| 0014 | 0026 | 0038 | 0052 |
| 0015 | 0027 | 0041 | 0053 |
| 0016 | 0028 | 0042 | 0054 |
| 0017 | 0031 | 0043 | 0055 |
| 0018 | 0032 | 0044 | 0056 |
| 0021 | 0033 | 0045 | 0057 |
| 0022 | 0034 | 0046 | 0058 |

| **Relevant Documents For Each Query** | | |
|---|---|---|
| **Query 1** | **Query 2** | **Query 3** |
| 0012 | 0011 | 0014 |
| 0015 | 0012 | 0033 |
| 0022 | 0021 | 0041 |
| 0031 | 0023 | 0044 |
| 0045 | 0026 | 0048 |
| 0056 | 0032 | 0052 |
| | 0044 | 0054 |
| | 0051 | |

When you ran the queries in your experimental system, SuperFinder, you got the following results:

| **Documents Retrieved For Each Query** | | | |
|---|---|---|---|
| **Rank** | **Query 1** | **Query 2** | **Query 3** |
| 1 | 0017 | 0027 | 0044 |
| 2 | 0012 | 0016 | 0031 |
| 3 | 0031 | 0055 | 0041 |
| 4 | 0052 | 0023 | 0054 |
| 5 | 0047 | 0045 | 0028 |
| 6 | 0056 | 0036 | 0014 |
| 7 | 0022 | 0031 | 0013 |
| 8 | 0028 | 0011 | 0051 |
| 9 | 0033 | 0044 | 0021 |
| 10 | 0048 | 0014 | 0042 |
| 11 | 0021 | 0052 | 0047 |

| | | | |
|---|---|---|---|
| 12 | 0015 | 0041 | |
| 13 | 0053 | 0025 | |
| 14 | | 0012 | |
| 15 | | 0038 | |
| 16 | | 0051 | |
| 17 | | 0032 | |
| 18 | | 0026 | |
| 19 | | 0018 | |
| 20 | | 0021 | |

For each of the three queries:

1. Calculate recall and precision (regardless of ranking)
2. Calculate (interpolated) precision at the 11 standard recall levels
3. Calculate average precision (AP). Show your work (i.e. show the precision at each seen document as well as your final values).
4. Calculate precision at a cutoff of 5 retrieved documents (precision@5, or P@5)

For the group of three queries:

1. Calculate the interpolated recall-precision averages (for the 11 standard recall levels)
2. Plot the recall-precision curves (from your interpolated values) for the three individual queries and for the averages of the three queries. Plot them on the same graph so that you can compare them.
3. Calculate the mean average precision (MAP)
4. Calculate the mean precison@5.

What can you say about the performance of SuperFinder on the three queries? Did it perform better on some queries than others? Justify your answer.

Turn in your answers as hard copy at the beginning of class on Tuesday February 13, 2007. Make sure your answers are legible.