

CS458 - Homework 4  
Due: Friday Oct. 26, by 6pm

*Submission:* E-mail or slide under door

1. (3 points) 8.2
2. (20 points) Evaluation

The list of Rs and Ns below represent relevant (R) and nonrelevant (N) documents in a ranked list of 20 documents retrieved using the same query from two different systems from a collection of 10,000 documents. The top of the ranked list is the leftmost element in the list (I've included whitespace to make counting easier). Assume that there are **eight** relevant documents in the corpus.

System 1: R R N N N    N N N R N    R N N N R    N N N N R  
System 2: R R N R N    N R N N N    N N N N R    N N N R N

- (a) (2 points) What are the precisions of the systems on the top 20?
  - (b) (2 points) What are the precisions of the systems on the top 10?
  - (c) (4 points) What are the  $F_1$ s of the systems on the top 10?
  - (d) (6 points) Assuming the 20 results are the complete set returned by the systems, what are the MAP scores for the query? (make sure to show your work for partial credit)
  - (e) (3 points) Which system does better based on the metrics? Do you agree (why)?
  - (f) (3 points) Assume that System 1 now returns all 10,000 documents. What is the largest MAP score System 1 could have?
3. (3 points) Find a misspelled query where two search engines return different spelling correction (e.g. "Showing results for" from Google

or "We have included" from Yahoo!). Which system did a better job at correcting the spelling?