

Exercise 9.1, 9.2

- 9.1: In Rocchio's algorithm, what weight setting for $\alpha/\beta/\gamma$ does a "Find pages like this one" search correspond to?
 - Slide 14 Part 8
- 9.2: Give some reasons why relevance feedback has been little used in web search.

Exercise 9.3

- Under what conditions would the modified query q_m in Equation 9.3 be the same as the original query q_0 ?
- In all other cases, is q_m closer than q_0 to the centroid of the relevant documents?

$$(9.3) \quad \vec{q}_m = \alpha \vec{q}_0 + \beta \frac{1}{|D_r|} \sum_{\vec{d}_j \in D_r} \vec{d}_j - \gamma \frac{1}{|D_{nr}|} \sum_{\vec{d}_j \in D_{nr}} \vec{d}_j$$

Exercise 9.4

- Suppose that a user's initial query is "cheap CDs cheap DVDs extremely cheap CDs".
- The user examines two documents, d_1 and d_2 . She judges d_1 , with the content "*CDs cheap software cheap CDs*" relevant and d_2 with content "*cheap thrills DVDs*" nonrelevant.
- Assume that we are using direct term frequency (with no scaling and no document frequency).
- Using Rocchio relevance feedback as in Equation (9.3) what would the revised query vector be after relevance feedback? Assume $\alpha = 1$, $\beta = 0.75$, $\gamma = 0.25$.