

Foundations of Information Retrieval

Exercise 7

Exercise session: 11.12.2014

Questions to: Helge Holzmann, holzmann@L3S.de

1. Relevance feedback & query expansion

1. A user Jon looks for information about populations of banana slugs and issues a query *banana slug* to the search engine *RFSearch*, which supports relevance feedback. For this query, *RFSearch* delivers the following top-3 results:

$D_1 = \text{"banana slug Ariolimax columbianus"}$

$D_2 = \text{"Santa Cruz mountains banana slug"}$

$D_3 = \text{"Santa Cruz Campus Mascot"}$

Jon marks the documents D_1, D_2 as *relevant* and D_3 as *not relevant* for his information need. *RFSearch* uses term frequency and the Rocchio RF algorithm, with $\alpha = \beta = \gamma = 1$. Which search query will be executed next based on the Jon's feedback? Specify the weights for each query term. Sort the terms in an alphabetical order.

2. Specify α, β , and γ values of the Rocchio RF algorithm that facilitate a query for retrieving pages similar to the relevant pages within the search result.