Eksamen 2016, løsningsskisse

SQUEEZE

a) See Section 5.3.1 in textbook.
b) See Section 5.3.2 in textbook.
c) http://www.uio.no/studier/emner/matnat/ifi/INF3800/v15/slides/index_compression_slides.pdf
d) http://www.uio.no/studier/emner/matnat/ifi/INF3800/v15/slides/index_compression_slides.pdf

LUCY IN THE SKY WITH DIAMONDS

a) Using the conservative estimate of the length of unioned postings lists, the recommended order is:

(*kaleidoscope* OR *eyes*) [87K + 213K = 300K] AND
(*tangerine* OR *trees*) [46K + 316K = 362K] AND
(*marmalade* OR *skies*) [107K + 271K = 378K]

However, depending on the actual distribution of postings, (*tangerine* OR *trees*) may well be longer than (*marmalade* OR *skies*) because the two components of the former are more asymmetric. For example, the union of 11 and 9990 is expected to be longer than the union of 5000 and 5000 even though the conservative estimate predicts otherwise.

b) No. For OR-queries we visit all postings in the posting lists of either terms, thus killing the need for skip pointers. Skip pointers are good for AND-queries, but we assume that we don't generate skip pointer data structures for the intermediate results of the OR-queries.

NEW YORK DOLLS

a) See Section 5.1.1 in textbook.
b) Assuming the rate of creation remains the same, we can assume the number of tokens doubles. Thus, there are 800M tokens. Using the equation $M = k\sqrt{T}$ to model the original corpus of 400M tokens, with T = 400M and M = 1M, we solve for k, yielding k = 50. Using this value for k, and T = 800M, we estimate that there are roughly $M = 50\sqrt{800M}$ = 1.41M tokens.

BAYES CITY ROLLERS

a) See Section 13.2 in textbook.
b) See Section 13.2 in textbook.
c) See Example 13.1 in textbook. Substitute *Chinese* with *longmuir*, *Beijing* with *nobby*, etc.
d) See Example 13.1 in textbook.

BEACH BOYS

a) See Section 21.2 in textbook.
b) See Exercise 21.6 in textbook. Substitute *1* with *brian*, *2* with *dennis* and *3* with *carl*. We then have *P* as in Equation 21.3. Note that 1/6 + 1/2 = 2/3, and that 1/6 + 1/4 = 5/12.
c) See Section 21.2.2 in textbook.