

Problem 1. (35 points)

Read the google paper (link available in section C4.3 of the Web-page).

<http://www-db.stanford.edu/~backrub/google.html>

Try to answer after reading the paper the following questions whose answers are in the paper or may need to be further searched upon.

(a) What was the size of google's dictionary/vocabulary around 1998? Justify your answer (5 points)

(b) What is the number of bits used for docID? You might need information beyond the paper for that (5 points).

(c) On the same one page you used for (a), (b) describe the data structures used by google for indexing. (20 points)

(d) Does google (circa 1998) use a hash table for the dictionary? What do they use? Why? (5 points).

Problem 2. (10 points)

Give unary, Elias- γ , Elias- δ , and Golomb $b = 3$ representations for (decimal) integer 12.