**Information Retrieval**

**18 December 2017**

**Name: Surname: Matricola:**

**Ex 1 [points 3+4]** Given the sequence of integers (11, 14, 16, 20, 21, 22), show how to encode them based on

* Gamma code
* PForDelta Code, by assuming b=2.

**Ex 2 [points 5+5]** Given the set of strings S = {abaa, abbb, abbc, b}, build a compacted trie for S and then

* show the two succinct representations of the structure of the trie based on the Heap-like notation and on the LOUDS notation.
* Show the formulas that allow to navigate both tree representations for computing the left/first child of a given trie node.

**Ex 3 [points 4+4+5]** Given the directed graph G consisting of nodes {A, B, C, D} and edges {(B,A), (A,C), (D,C), (A,B), (C, B), (D,A)}:

* Compute one step of the PageRank by assuming that the teleportation step occurs with probability 0.5 and the starting probability distribution is uniform.
* Define, in formulas only, how it is computed the authoritative score of node A and its hubness score as a function of the authoritative/hubness scores of its adjacent nodes.
* Comment how the similarity between node B and all the other G’s nodes can be estimated by using Personalized PageRank. Show your proposal over G for one step only, with the teleportation step occurring with probability 0.5, and with starting probability (0, 1, 0, 0).