**Information Retrieval**

**19 January 2019**

**Name: Surname: Matricola:**

**Ex 1 [points 8+5+3]**

* Define what is a Spectral Bloom filter, the operations supported, and show the formula of the error rate (as a function of the space and number of keys) and prove its correctness.
* Define what is a Permuterm index, the operations supported, comment on the occupied space with respect to the indexed set of keys.
* Define the formulas for Precision, Recall and F1

**Ex 2 [points 4+5]** You are given the two files: F\_old = “abracadabra”, F\_new = “abracalabria”, and assume a block size B=3 chars.

* Show the execution of the algorithm rsync. *(comment the various steps)*
* Show the execution of the algorithm zsync. *(comment the various steps)*

**Ex 3 [points 5]** Given the two adjacency lists of a Web graph for the nodes 14 and 15, namely

 14 🡪 3, 10, 12, 13, 14, 17, 19, 21, 25

 15 🡪 5, 10, 11, 12, 14, 17, 19, 21, 24, 33

show how the list of “15” can be compressed given the list of “14” by means of the algorithm adopted in WebGraph via copy-lists and copy-blocks.