**Information Retrieval**

1. **July 2018**

**Ex 1 [points 3+3]** Consider the Consistent Hashing technique

1. simulate its execution over the urls\_ID = {1, 2, 3, 6, 8, 11, 12, 13} and the crawler\_ID = {2,4,7}, by defining the hash function h(x) = 5\*x mod 11.
2. Show what happens if the crawler\_ID 3 faults.

**Ex 2 [points 3+3]** You are given the two files: F\_old = “cane matto orso”, F\_new = “cane ratto dorso”, 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 [ranks 3+3+3+3]**

Given the set of strings S={aba, abc, baac, babc}.

* Show the (compacted) trie T built on S
* Construct the heap-like representation of the tree structure of T
* Construct the LOUDS representation of the tree structure of T
* Describe the procedure for computing the first-child of node x given LOUDS

**Ex 4 [ranks 3+3]** Given a graph G=(V,E),

* Define the formulas of PageRank and its variant Personalized Page Rank (PPR), and comment on their interpretation.
* Show how PPR is used to compute CoSimRank(u,v) for every pair of nodes u,v in the graph G.