**Information Retrieval**

1. **June 2016**

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

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

**Ex 2 [points 3+3+3]**

* Construct a 2-gram index for the set of strings S={bat, baz, bit, box, zox}.
* Design an algorithm that finds all strings of S at 1-edit distance from P[1,p].
* Simulate the proposed algorithm over the string P=bix.

**Ex 3 [ranks 4+4]** Given the list of integers S=(2,4,8,10,13,14,15,21).

* Show how to compress it via Elias-Fano coding.
* Describe how to select the 5th integer from S, by assuming to have a Select\_1 data structure over the array H.

**Ex 4 [ranks 5+2]** 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.