Advanced Programming (A.Y. 2014-2015)
M.Sc. Engineering in Computer Science
Sapienza University of Rome

Programming test – September 14, 2015 – Duration 1h 30'

Write a Hadoop MapReduce program that, given as input a directed graph represented as a list of edges, computes the list of all nodes y such that there is an edge (x,y) with x<y and an edge (y,z) with y<z. Assume that nodes are identified by integer values.

Input files must have the form:

```
x1  \t  y1
x2  \t  y2
x3  \t  y3
…
```

where `\t` is the tabulation character (ASCII code 9).

Try as input the following graph:

```
1	2
2	3
3	4
4	1
2	4
3	1
1	1
```

The output on the test graph above should consist of the following two nodes (in any order):

```
2
3
```

For testing the correctness of the program, work with files on HDFS.