UNIVERSITÉ Grenoble Alpes — Master of Science in Informatics at Grenoble — Grenoble INP ensimag

UE Mathematics for Computer Science

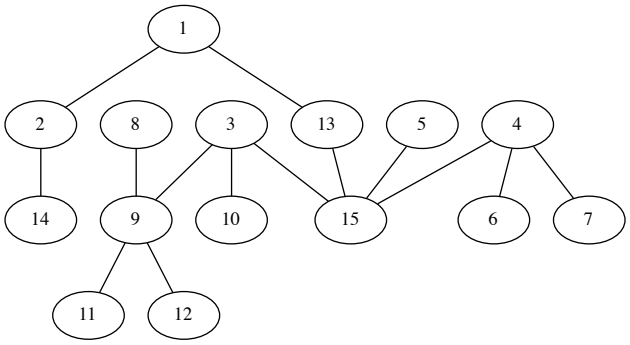
Homework, December 2017

Write on your homework:
I understand what plagiarism entails and I declare that this report is my own, original work.
Names, date and signatures.

- The firm deadline is December Friday 22 midnight (before Saturday).
- The homework should be 3 pages (and appendix)
  in the pdf format (scanned manuscripts in pdf are allowed)
- the filename should be FamilyName1-FamilyName2-Mosig-MfCS-HW2.pdf
- send with your official mail () at Jean-Marc.Vincent@imag.fr with the subject
  `[MOSIG1:MfCS] HW2 FamilyName1 FamilyName2`

Labelled Trees

The problem is to evaluate properties of random trees, the approach is to combine counting techniques, simulation and basic statistics. In the problem we consider labelled trees, where labels are $\{1, 2, \cdots, n\}$.



In this example, the number of nodes is $n = 15$, the labels $\{1, 2, \cdots, 15\}$, the number of leaves is 8, the degree of node 9 is 4 and of node 13 is 2. The diameter of the tree, the length of the longest elementary path in the tree, is 6, for the paths $[14, 2, 1, 13, 15, 4, 6]$ or $[14, 2, 1, 13, 15, 4, 7]$.

1. **Simulation Algorithm**
    1. Propose an algorithm that generates uniformly a random tree with n labelled nodes.
    2. Justify this algorithm and prove it.
2. **Statistical Properties** For a uniformly generated labelled tree compute the following quantities, for each computation you can either simulate and make some statistics or compute it directly.
    1. The average number of leaves and the probability distribution of the number of leaves.
    2. The average degree of a node and the probability distribution of the degree.
    3. The average diameter and and the probability distribution of the diameter.
3. **Synthesis**

    Make a synthesis on the properties of uniform random trees.

Comments

- You are free to choose your programming language.
- The number of nodes n is a parameter of the problem, you have to choose different values for n to provide interesting conclusions.
- For the simulations, take care of the size of the samples and provide confidence intervals on the estimations of average.
- You can use appendix if you wish, but the 3 mandatory pages should be self contained and are to be evaluated.