Graph Theory Applications EPFL, Spring 2014

Solution to Graded Problem Set 4

Date: 13.03.2014 Due by 18:00 – 20.03.2014

Problem 1. Let V be the set of vertices, x be the number of leaves in the tree and y the number of all other vertices. Let e be the number of edges, then $e = x + y - 1$. Also, $\sum_{v\in V} deg(v) \geq x + 3y$ since every non-leaf vertex has degree ≥ 3. Using these two observations we have:

$$2(x + y - 1) = 2e = \sum_{v\in V} deg(v) \geq x + 3y \geq x + 3y - 2,$$

from which $y \leq x$ follows.

Problem 2. If G is not a tree then it contains a cycle. There is least one edge (u, v) of this cycle which is not in G'. Obviously, for these vertices $d_G(u, v) = 1$, while in G' they are not neighbors, thus $d_{G'}(u, v) > 1$.

Problem 3. The "only if" direction is easy. Let $T = (V, E)$ be a tree with n vertices. Since a tree has exactly $n - 1$ edges $\sum_{i=1}^{n} d_i = 2e = 2(n - 1)$. For the "if" direction, we will give a construction which given a degree sequence satisfying the given condition will produce a tree with that degree sequence.

Suppose that $d_1 \leq d_2 \ldots \leq d_n$ is the given degree sequence. We proceed by induction on the length of the sequence.

Base step: $n = 2$. We have 2 nodes connected by an edge: this is a tree.

Induction step. Assume that the claim holds for all degree sequences of length less than n. Now for a degree sequence of length n, since d_1 is the lowest degree, $d_1 = 1$. Indeed, if $d_1 \geq 2$, then $\sum_{i=1}^{n} d_i \geq 2n$, which results in a contradiction. Also, since d_n is the highest degree, $d_n \geq 2$. Indeed, if $d_n \leq 1$, then $\sum_{i=1}^{n} d_i \leq n$, which results in a contradiction.

Consider the degree sequence $d_2, d_3, \ldots, d_n - 1$. These are $n - 1$ numbers summing to $2(n - 2)$. By the induction hypothesis, we have a tree T' corresponding to it. Now, construct a new tree T with n vertices by gluing a single vertex to the vertex of degree $d_n - 1$ in T'. This completes the proof.

Problem 4. Suppose that G is connected. If not, all his connected components have vertices with degree ≥ 3 and we consider one of those. Let k be the number of edges and n the number of nodes. The proof is by contradiction, namely, we suppose that there is no cycle of even length. Let us denote by N_c the number of cycles, which are all of odd length.

First of all, let us show that $N_c \leq \frac{k}{3}$. If two cycles share an edge, then we can build a cycle with even length and we are done. Hence, all the cycles are disjoint and since their length is at least 3, their number cannot be $> k/3$.

Then, let us show that $N_c \geq k - (n - 1)$. Each connected graph has a spanning tree. Start from the spanning tree which has $n - 1$ edges. Every time we add an edge, we create at least one new cycle. Hence, there are at least $k - (n - 1)$ cycles.

Now, let us conclude. Since $k - (n - 1) \leq N_c \leq \frac{k}{3}$, we have that $\frac{k}{3} \geq k - (n - 1)$, from which we get that $k \leq \frac{3}{2}(n - 1)$. However, $k \geq \frac{3}{2}n$, because each vertex has degree at least 3. So, we have found a contradiction and the proof is complete.

Problem 5. In $G(n, p)$ every edge exists independently and with the same probability p. Hence, the probability of drawing a graph with m edges is

$$\binom{\binom{n}{2}}{m} p^m (1-p)^{\binom{n}{2}-m},$$

which gives a binomial distribution. The mean value is given by

$$\binom{n}{2} p.$$

As concerns the last point, what you should observe is the following behavior. Suppose that n is large enough and let $c = pn$.

- If $c < 1$, then a graph in $G(n, p)$ will almost surely have no connected components of size larger than $O(\log(n))$. This behavior is exemplified by the graph in Figure 1 which is obtained taking $n = 10000$ and $c = 0.7$.
- If $c = 1$, then a graph in $G(n, p)$ will almost surely have a largest component whose size is of order $n^{2/3}$.
- If $c > 1$, then a graph in $G(n, p)$ will almost surely have a unique giant component containing a positive fraction of the vertices. No other component will contain more than $O(\log(n))$ vertices. It is possible to show that the fraction of the vertices contained in the giant component is given by the non-zero solution to the equation

  $$x + e^{-cx} = 1. \tag{1}$$

  This situation is schematized in Figure 2, which is obtained taking $n = 10000$ and $c = 1.1$.

If you compute the size of the largest connected component divided by the total number of nodes as a function of $c = p/n$, you should observe a behavior similar to that obtained in Figure 3.

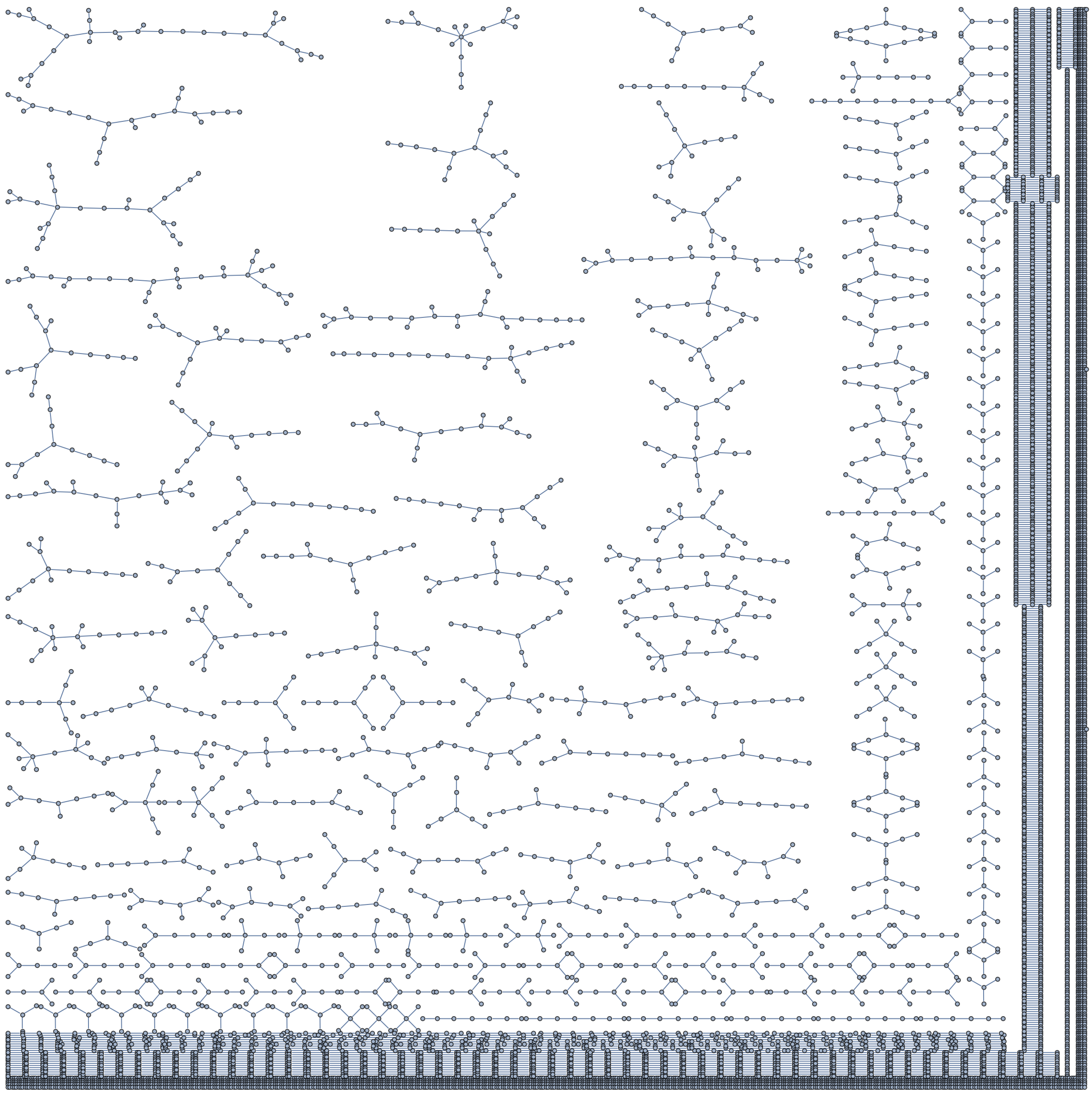

Figure 1: Random graph $G(n, p)$ with $n = 10000$ and $p = 0.7$. As $p < 1/n$, there is no giant connected component.

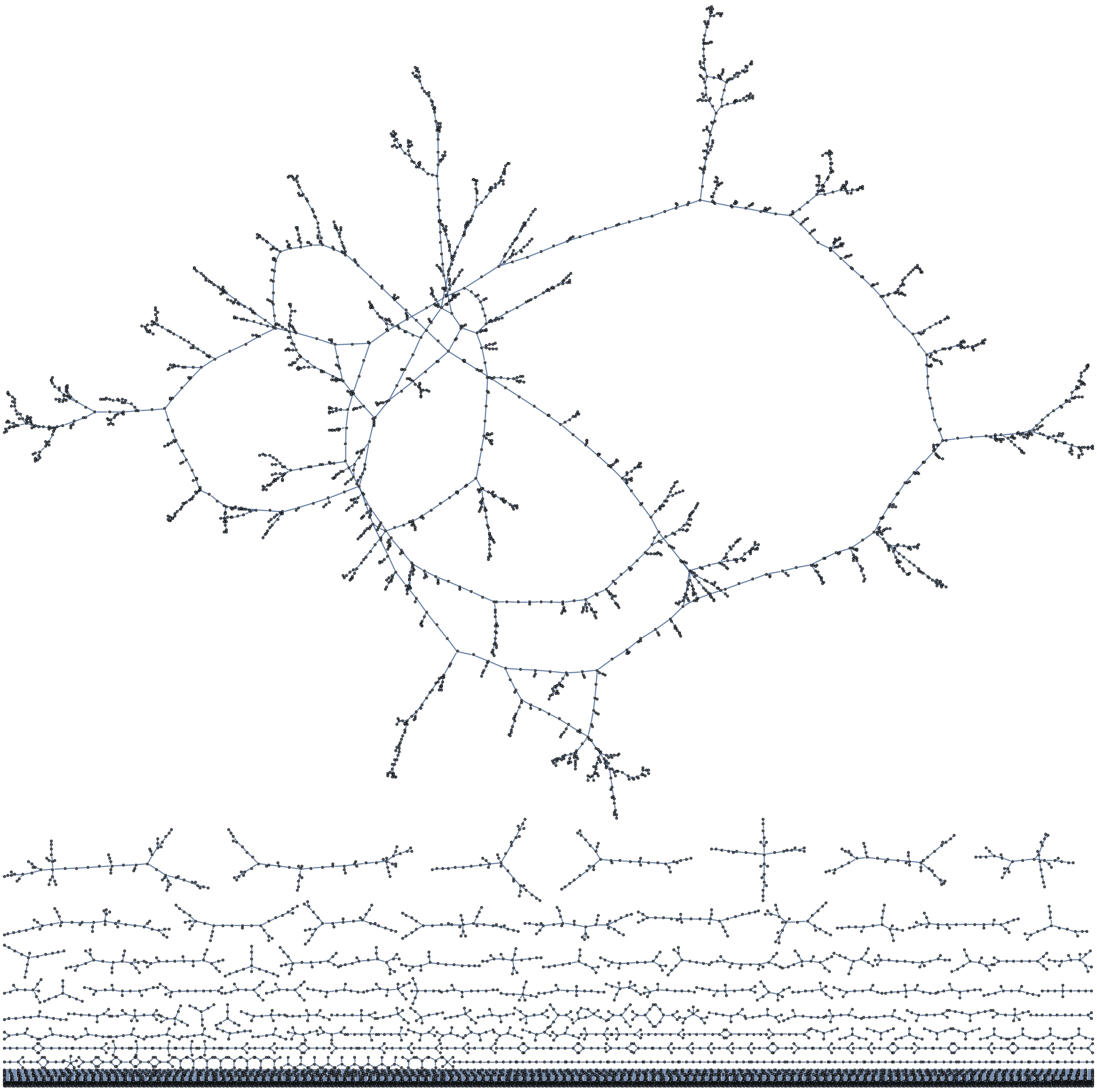

Figure 2: Random graph $G(n, p)$ with $n = 10000$ and $p = 1.1$. As $p > 1/n$, the giant component appears.

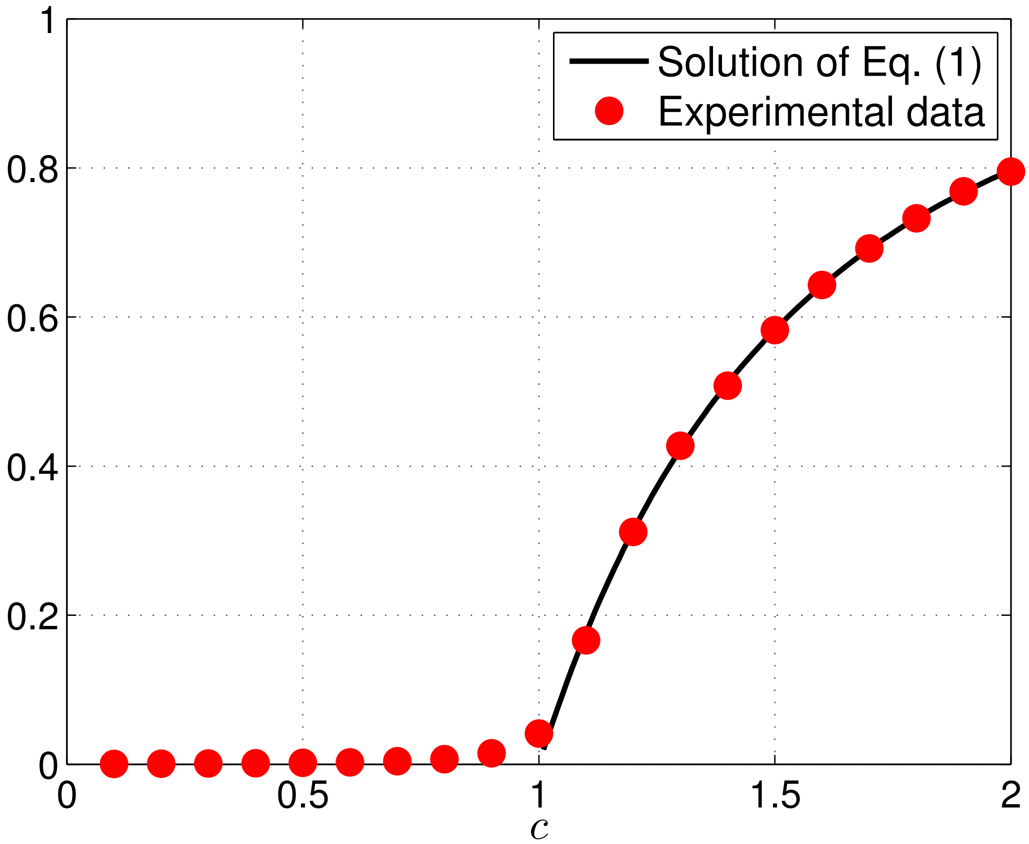

Figure 3: Normalized size of the largest connected component as a function of c with $n = 10000$. If $c < 1$, there is no giant component. If $c > 1$, the giant component contains a fraction of the nodes of the graph which is given by the solution to (1).