Social and Technological Networks Edinburgh, 2015

Lecture 3. Graph Properties & Random Graphs, continued.

TA: Maria Astefanoaei *Solutions to Exercises*

* **Exercise 1.1.** *Show that an ER random graph of n nodes has an expansion rate larger than some constant.*

I realized later that the problem stated this way is more complicated than what I intended to ask. So let us instead look at a smaller version of the problem:

* **Exercise 1.1.** *Show that in an ER random graph of n nodes, the number of edges between subsets $S \subset V$ and $V \setminus S$ is larger than $\alpha|S|$ for a suitable constant α.*

Answer. Let S be a random set of nodes in an ER random graph $G(n, p)$ and $C(S)$ the number of edges between S and $V \setminus S$. We are going to show that $C(S)$ is large enough in expectation for $p = (\ln n)/n$.

Let set S have size k and set $V \setminus S$ size $n - k$, where without loss of generality, we can assume $k \leq \lfloor n/2 \rfloor$ (since the problem is symmetric in S and $V \setminus S$). Then $C(S)$ is a binomial random variable with $k(n-k)$ trials, and probability of success p. Therefore, the expected value for $C(S)$ is $\mathbb{E}[C(S)] = k(n-k)p$.

The expansion of S is defined as $h(S) = \frac{C(S)}{|S|}$, therefore:

$$h(S) \geq \frac{k(n-k)p}{k} \geq \frac{n \cdot \ln n}{2 \cdot n} \geq \frac{\ln n}{2}.$$

Thus, for sufficiently large n, in fact, $h(S) \geq \alpha$ for any constant α.

Exercise 2.1. *Prove that the total number of all triads (open and closed) in a connected graph is $\geq c'n$ for some constant c', assuming $n \geq 3$.*

Answer. Note that here we are considering arbitrary connected graphs, and not ER graphs. Let us prove the property by induction.

First of all, observe that the given statement is true for a minimally connected triad, that is, a tree on 3 nodes. In this case, $c' = 1/3$. This is our base case.

Now, for induction, assuming that the property is true for a graph G of n nodes with m edges, we want to show that it is true when either a node or an edge is added. Let us use T to represent the count of triads, and by hypothesis, we have $T \geq n/3$. We want to show that in any $G' \supset G$ with T' triads.

- If and edge is added to G, this does not destroy any existing triads and does not increase n. Thus $T' \geq T \geq n/3$ holds.
- If a node i is added to G, for the new graph G' to remain connected, it must also have an edge ij. Since G was already connected (with 3 or more nodes), an edge jk must have existed. Thus, the triad ijk exists in G', which did not exist in G. Thus $T' \geq T+1 \geq n/3+1 \geq \frac{1}{3}(n+1)$.

Exercise 2.2. *For what sort of values of k are the neighborhoods likely to have cycles? That is, when does the above deduction fail, and when do cycles become likely?*

Answer. For the first part, that is, the condition for the deduction to fail, we can proceed as follows. The claim that a cycle is unlikely rests on the fact that $\frac{\mathrm{poly}(\ln n)}{n}$ is arbitrarily small for sufficiently large n. The claim of "unlikely" can be said to fail when the bounding probability is only a constant. Suppose we want to say that $\frac{\ln^k n}{n} \geq q$, where q is a constant. Then transposing and taking log on both sides gives $k = (\ln(qn))/\ln\ln n$. Which is the point where the probability bound is only as good as a constant q. That is, the above deduction no longer says ccyles are unlikely when k is at least that large.

Although, this may be simply a failure of the deduction, and the cycles actually may not be as likely as q. The second part asks you to find the point where we can say that cycles are actually likely. We omit the proof of that for now. You should try that as an exercise.