Complex Systems 535/Physics 508: Homework 3

1. Consider the random graph $G(n, p)$ with average degree c.

   (i) Show that in the limit of large n the expected number of triangles in the network is $\frac{1}{6}c^3$, where c is the mean degree as before. In other words, the number of triangles is constant, neither growing nor vanishing in the limit of large n.

   (ii) Show that the expected number of connected triples in the network (as in Eq. (7.41)) is $\frac{1}{2}nc^2$.

   (iii) Hence calculate the clustering coefficient C, as defined in Eq. (7.41), and confirm that it agrees for large n with the value given in Eq. (12.11).

2. Starting from the generating function $h(z)$ defined in Eq. (12.26), or otherwise, show that

   (i) the mean-square size of the component in a random graph to which a randomly chosen vertex belongs is $1/(1-c)^3$ in the regime where there is no giant component;

   (ii) the mean-square size of a randomly chosen component in the regime with no giant component is $1/[(1-c)(1-\frac{1}{2}c)]$.

   Note that both quantities diverge at the phase transition where the giant component appears.

3. The cascade model is a simple mathematical model of a directed acyclic graph, sometimes used to model food webs. We take n vertices labeled $i = 1 \ldots n$ and place an undirected edge between each distinct pair with independent probability p, just as in the ordinary random graph. Then we add directions to the edges such that each edge runs from the vertex with numerically higher label to the vertex with lower label. This ensures that all directed paths in the network run from higher to lower labels and hence that the network is acyclic. You can think of the labels as representing, for instance, the trophic level of a species in a food web.

   (i) Show that the average in-degree of vertex i in the ensemble of the cascade model is $\langle k_i^{\mathrm{in}} \rangle = (n-i)p$ and the average out-degree is $\langle k_i^{\mathrm{out}} \rangle = (i-1)p$.

   (ii) Show that the expected number of edges that connect to vertices i and lower from vertices above i is $(ni - i^2)p$.

   (iii) Ecologists study food webs as a way of understanding energy flow in ecosystems. At what trophic levels in this model ecosystem would the greatest energy flow occur, and why?

4. We can make a random graph model of a network with clustering as follows. We take n vertices and go through each distinct trio of three vertices, of which there are $\binom{n}{3}$, and with independent probability p we connect the members of the trio together using three edges to form a triangle, where $p = c/\binom{n-1}{2}$ with c a constant. You can assume that n is very large.

   (i) Show that the mean degree of a vertex in this model network is $2c$.

   (ii) Show that the degree distribution is $p_k = \mathrm{e}^{-c}c^{k/2}/(k/2)!$ if k is even and $p_k = 0$ if k is odd.

(iii) Show that the clustering coefficient, Eq. (7.41), is $C = 1/(2c+1)$. Note that, by contrast with the ordinary random graph, this clustering coefficient doesn't go to zero as n becomes large.

(iv) Show that as a fraction of network size, the expected size S of the giant component, if there is one, satisfies $S = 1 - e^{-cS(2-S)}$.

(v) What is the value of the clustering coefficient when the giant component fills half of the network?

5. **Extra credit:** Write a computer program in the language of your choice that generates a random graph drawn from the model $G(n,m)$ for given values of n and the average degree $c = 2m/n$, then calculates the size of its largest component. Use your program to find the size of the largest component in a random graph with $n = 1\,000\,000$ and $c = 2\ln 2 = 1.3863\ldots$ and compare your answer to the analytic prediction for the giant component of $G(n,p)$ with the same parameter values. You should find good agreement, even though the models are not identical. (As we discussed in class, the models become essentially the same for large values of m and n because the correlations between edges in $G(n,m)$ go to zero and the edges become asymptotically independent.)