Phase Transitions in Networks

1. When nodes and links are added to or removed from a network, its connectivity pattern, including the degree distribution $p(k) \equiv N^{-1} \sum_{i=1}^{N} \delta_{k_i,k}$, where $k_i = \sum_j A_{ij}$ is the degree of node i and $A_{ij} \in \{0, 1\}$ is the adjacency matrix, generally changes. However, certain classes of degree distributions retain their functional form under random link removal, acting as fixed points in the space of degree distributions.

   Consider a network with degree distribution $p_0(k)$. A fraction f $(0 < f < 1)$ of links is removed uniformly at random.

   (i) Derive the generating function $g_f(z) = \sum_k p_f(k) z^k$ of the degree distribution after link removal.

   (ii) Find the condition on $p_0(k)$ under which $p_f(k)$ has the same functional form as $p_0(k)$. Give examples.

2. Many real-world networks exhibit power-law degree distributions, implying the presence of a non-negligible fraction of nodes with very large numbers of neighbors.

   Consider a network with a power-law degree distribution

   $$p(k) = \frac{k^{-\gamma}}{\zeta(\gamma)} \text{ for } k = 1, 2, \ldots$$

   with $\zeta(\gamma)$ the Riemann zeta function and $\gamma > 3$. Determine the range of γ for which the considered network has a giant connected component (GCC).

3. Consider an ensemble of networks with mean degree $\bar{k}$ and $p(k) \sim k^{-\gamma}$ for large k, with $\gamma > 3$. In lectures, we learned by taking a branching process approach that the relative size of the GCC, $m = \langle \frac{S}{N} \rangle$, becomes nonzero at a critical mean degree $\bar{k}_c$ as $m = 0$ for $\Delta < 0$ and $m \sim \Delta^{\beta}$ for $\Delta > 0$ with $\Delta = \frac{\bar{k}}{\bar{k}_c} - 1$ and β a critical exponent. Extending the approach, show that the mean finite-cluster size

   $$\bar{s} = \sum_{s<S} sP(s),$$

   where $P(s)$ is the cluster-size distribution, i.e., the probability that a node belongs to a size-s cluster, and the summation runs over all clusters except for the largest (giant) one, diverges near the critical point as

   $$\bar{s} \simeq \begin{cases} \frac{c_-}{|\Delta|} & \text{for } \Delta < 0, \\ \frac{c_+}{\Delta} & \text{for } \Delta > 0, \end{cases}$$

   and determine how the amplitudes c_- and c_+ depend on γ.

4. Cellular metabolism consists of numerous chemical reactions and can be represented as a bipartite network with two types of nodes: reaction nodes and compound nodes. A link connects a reaction node to a compound node if the compound participates in the reaction as a substrate or a product; for example, the reaction R1:A+B $\rightarrow$ C corresponds to links (A,R1), (B,R1), and (R1,C).

   A reaction can occur only when all of its substrates and products are available. A compound attains a stable concentration if it is produced by at least one reaction and consumed by at least one reaction. Accordingly, we define a reaction node as viable if all of its neighboring compound nodes are viable, and a compound node as viable if it is connected to at least two distinct viable reaction nodes. The core is defined as the maximal subgraph of viable reaction nodes and viable compound nodes.

   Consider a metabolic network in which compound nodes and reaction nodes have Poisson degree distributions

   $$p_{\mathrm{c}}(k) = \frac{\bar{k}^k}{k!} e^{-\bar{k}} \text{ and } p_{\mathrm{r}}(q) = \frac{\bar{q}^q}{q!} e^{-\bar{q}},$$

   where $\bar{k}$ and $\bar{q}$ are the mean degrees of compound and reaction nodes, respectively. Suppose that a fraction Q of reaction nodes is removed uniformly at random, representing the inactivation or malfunction of the corresponding enzymes. Derive self-consistent equations for the probability that a compound node reached by following a link is inviable and the probability that a reaction node reached by following a link is inviable, and use them to analyze how the core size decreases with increasing Q.

5. Consider two partially interdependent networks $\mathcal{A}$ and $\mathcal{B}$. A fraction q of the nodes in $\mathcal{A}$ depends on a node in $\mathcal{B}$ and a fraction q of the nodes in $\mathcal{B}$ depends on a node in $\mathcal{A}$ (One node from a network can depend only on one node from the other network). Let $C_{\mathcal{A}}$ and $C_{\mathcal{B}}$ denote the GCCs of $\mathcal{A}$ and $\mathcal{B}$, respectively. The giant mutually-connected component (GMCC) of $\mathcal{A}$ is defined as the set of nodes that belong to $C_{\mathcal{A}}$ and whose dependent nodes, if present, belong to $C_{\mathcal{B}}$. The GMCC of $\mathcal{B}$ is defined similarly.

   Suppose that both networks have Poisson degree distribution $p(k) = \frac{\bar{k}^k}{k!}e^{-\bar{k}}$ with the same mean degree $\bar{k}$. For a given q $(0 < q < 1)$, derive the relative size of the GMCC (e.g., of $\mathcal{A}$) as a function of $\bar{k}$. Does it exhibit a phase transition as $\bar{k}$ increases? If so, is it continuous or discontinuous?