Analysis of Biological Networks **Spring 2009**

Assignment 1

The exercise should be submitted in pairs by 26/3/09 to Nir Yosef's mailbox in Schreiber.

1 Scale free networks (20 pts)

A degree distribution $P(\cdot)$ is scale free if there exists a function g such that for all a, $P(a \cdot k) = g(a) \cdot P(k)$. Prove that a degree distribution is scale free if and only if it is power-law. You can assume that both $P(\cdot)$ and $g(\cdot)$ are differentiable.

2 Estimating the clustering coefficient (50 pts)

For a given degree sequence, let Φ be the collection of all graphs on the same vertex set V with this degree sequence. Let $n = |V|$ and let m be the number of edges in a graph from Φ.

a. Prove that the chance of having an edge between u and v in a graph from Φ is approximately $\frac{d(u)d(v)}{2m}$, where $d(v)$ is the degree of v.

b. Let $P(\cdot)$ be the degree distribution of graphs in Φ (as defined by the degree sequence). Use your result from item (a) to prove that the clustering coefficient of such graphs is: $C = \frac{M_1}{n} \cdot (\frac{M_2 - M_1}{M_1^2})^2$, where M_1, M_2 are the first and second moments of $P(\cdot)$.

For the rest of the question, assume that $P(\cdot)$ is scale free, $P(k) \propto k^{-c}$.

c. Show that the maximum degree in a graph from Φ is roughly $n^{\frac{1}{c-1}}$, by assuming that there exists a single vertex with this maximum degree.

d. Conclude that for power law distributions with $2 < c < 3$, the clustering coefficient can be estimated as $n^{\frac{3c-7}{c-1}}$.

Hint: Estimate M_2 using the largest term in the sum defining it.

3 MCMC (30 pts)

a. Let M be an n-state, irreducible, aperiodic Markov chain with a stationary distribution $\{p_i\}$ and transition probabilities $\{p_{ij}\}$. Show that the detailed balance equations hold if and only if $p_{ij} = r_{ij}$ where r_{ij} are the transition probabilities of the reversed-time chain.

b. Consider an asymmetric proposal mechanism $\{q_{ij}\}$, where $q_{ij} > 0$ for all i, j and $\sum_{j=1}^{n} q_{ij} = 1$ for all i. Show how to modify the Metropolis algorithm (which is designed for symmetric proposals) so as to create an irreducible, aperiodic Markov chain whose stationary distribution is $\{p_i\}$.
Hint: Specify the resulting transition probabilities p_{ij} (also for the case $i = j$) and use them to prove that the resulting chain is irreducible and aperiodic, and that the detailed balance equations hold.