18.447 Midterm Homework
Due Tues. Nov 1, 2004
no collaboration

Problem 1.

Consider a graph with $3n$ vertices, n colored red, n blue and n green. We add edges randomly and independently. We put an edge between a blue vertex and a green vertex with probability p_1, between a blue vertex and a red vertex with probability p_2, and between a red vertex and a green vertex with probability p_3. Let $p_1 p_2 p_3 = n^{-2}$, so the expected number of trichromatic triangles is n. If we have $p_1 \geq p_2 \geq p_3 \geq \frac{\tau}{n^2}$, prove that with probability at least $1 - \frac{3}{\tau} - \frac{1}{n}$. there is a trichromatic triangle.

Problem 2.

Consider random strings A and B of length n over an alphabet of size L, with the probability distribution where each letter is equally likely. We will consider the asymptotics when n varies while L is fixed. We define the length of the longest common subsequence to be the largest k such that we can find sequences $1 \leq \alpha_1 < \alpha_2 < \ldots < \alpha_k \leq n$ and $1 \leq \beta_1 < \beta_2 < \ldots < \beta_k \leq n$ so that $A[\alpha_i] = B[\beta_i]$ for $1 \leq i \leq k$, where $A[l]$ denotes the l'th letter of A (i.e., the longest common subsequence of SWARM and STEAM is SAM).

(a) If $X(n)$ is the length of the longest common subsequence of two strings of length n, show that $EX(n)+EX(m) \leq EX(n+m)$, and thus $\lim_{n\to\infty} \frac{1}{n} EX(n)$ exists.

The value of this limit is a famous open problem.

(b) Show that for $\lambda > 0$ there are two numbers b_n and c_n with $c_n - b_n \leq 2\lambda\sqrt{n}$ and such that $\mathrm{Prob}(b_n \leq X_n \leq c_n) \geq 1 - 2e^{-\lambda^2/2}$.

Problem 3.

Consider two $n \times n$ matrices A and B filled with symbols chosen independently and uniformly from an L-letter alphabet. Let the size of the largest common square submatrix be the largest k such that we can find four sequences $1 \leq r_1 < r_2 < \ldots < r_k \leq n$, $1 \leq c_1 < c_2 < \ldots < c_k \leq n$, $1 \leq r'_1 < r'_2 < \ldots < r'_k \leq n$, $1 \leq c'_1 < c'_2 < \ldots < c'_k \leq n$, so that $A[r_i, c_j] = B[r'_i, c'_j]$ for $1 \leq i, j \leq k$. Let X_n be the size of the largest common square submatrix.

(a) Show that almost surely (i.e., with probability going to 1 as n goes to infinity) $X_n \leq 5 \log_L n$.

(b) Show that almost surely, $X_n \geq \log_L n$.

(c) Show that there is a constant α such that for large n, there is some value b_n so that $\mathrm{Prob}(|X_n - b_n| \geq \alpha\lambda\sqrt{\log_L n}) < 2e^{-\lambda^2}$.