CMSC 858S: Homework 1, due on February 9th (Friday) at the start of class.

Notes: Please work on this with your groupmate. Consulting other sources (including the Web) is not allowed. Write your solutions neatly; if you are able to make partial progress by making some additional assumptions, then state these assumptions clearly and submit your partial solution.

1. Prove the union bound using Markov's inequality. **(5 points)**

2. (a) For integers r, n such that $1 \leq r \leq n$, show that $\sum_{i=0}^{r}(-1)^i \binom{n}{i} = (-1)^r \binom{n-1}{r}$. (As usual, if $r = n$, we take $\binom{n-1}{r} = 0$.) **(5 points)**

(b) Prove that truncated inclusion-exclusion alternately lower- and upper-bounds the probability of a union $\Pr[A_1 \vee A_2 \vee \cdots \vee A_n]$, starting with the following definition. For each i, let X_i be the indicator random variable for A_i; for any n-bit string $b = b_1 b_2 \cdots b_n \in \{0,1\}^n$, define

$$f(b) = \Pr[(X_1 = b_1) \wedge (X_2 = b_2) \wedge \cdots \wedge (X_n = b_n)].$$

(5 points)

3. You are given a Boolean formula ϕ with m clauses in conjunctive normal form, with at least 3 literals in each clause. For instance, ϕ could be

$$(X_2 \vee \overline{X_5} \vee \overline{X_8}) \wedge (X_1 \vee X_5 \vee \overline{X_6} \vee \overline{X_8}) \wedge (\overline{X_2} \vee X_6 \vee \overline{X_9}) \wedge (X_3 \vee X_4 \vee \overline{X_7} \vee \overline{X_{10}});$$

there are $m = 4$ clauses here, and the underlying Boolean variables are $X_1, \ldots, X_{10}$.

Given any such ϕ, show that there exists an assignment of truth values to the underlying Boolean variables that satisfies at least $7m/8$ of the given m clauses. **(5 points)**

4. Recall that Chebyshev's inequality proceeds as follows. We are given some random variable X with mean μ and standard deviation σ. In order to upper-bound $\Pr[|X - \mu| \geq t]$ for some $t > 0$, we introduce the quadratic function $f(X) = (X - \mu)^2$, observe that

$$\Pr[|X - \mu| \geq t] \leq \Pr[f(X) \geq t^2], \qquad (1)$$

and apply Markov's inequality to the latter, in order to get the bound σ^2/t^2. (Equality in fact holds in (1), but the given inequality is sufficient for us.) Now suppose we only want to upper-bound $\Pr[X - \mu \geq t]$ for some $t > 0$, and therefore want a better bound. So we again plan to start with a general quadratic function $f(X) = aX^2 + bX + c$, use some bound broadly along the lines of (1), then apply Markov's inequality etc.

(i) What conditions would be required of the function $f(X) = aX^2 + bX + c$ that you introduce? **(5 points)**

(ii) Choose optimal values of a, b, c; what best bound on $\Pr[X - \mu \geq t]$ do you get? **(5 points)**

5. Let $G = (V, E)$ be a graph with n vertices and minimum degree $\delta \geq 11$. Show that there is a partition of V into two subsets A and B such that $|A| \leq O(n \log(\delta + 1)/(\delta + 1))$, and such that each vertex of B has at least one neighbor in A and at least one neighbor in B. **(10 points)** (**Hint:** Start with a simple random process followed by a carefully-designed alteration.)