Notes: Please work on this with your group-mate(s); just submit one writeup per group. 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. The problem marked (*) may be more difficult than the others.

1. In this problem, we prove that the Janson inequality parameter ∆ is at most $O(\log N)$ in the Garg-Konjevod-Ravi Group Steiner Tree algorithm for trees, as claimed in class. Recall from class that we have a tree $T$ with root $r$ and $n$ nodes. There are $k$ disjoint groups $S_1, S_2, \ldots, S_k$, all of which are sets of leaves; also, $\max_i |S_i| = N$. Specifically, we fix a group $S_i$, and want to show that $\Delta_i$, the Janson inequality parameter for the set of leaves that correspond to $S_i$, is at most $\log |S_i|$. As in class, we let $x_f$ be the fractional value of edge $f$, and if $j$ is a leaf, then let $pe(j)$ denote the unique (“parent edge”) incident on $j$.

(a). Suppose $j, j' \in S_i$. We will say that $j \sim j'$ if and only if (i) $j \neq j'$ and (ii) the least common ancestor of $j$ and $j'$ in $G$ is not the root $r$. If $j \sim j'$, let $lca(j, j')$ denote the least common ancestral edge of $j$ and $j'$ in $T'$. Show that $\Delta_i = \sum_{j,j' \in S_i: j \sim j'} x_{lca(j, j')} x_{pe(j)} x_{pe(j')} x_{lca(j, j')}$. (5 points)

(b). We will now prove the following key fact:

If $x_{pe(j)} > 0$, then $x_{pe(j)} \cdot \sum_{j' \in S_i: j \sim j'} \frac{x_{pe(j')}}{x_{lca(j, j')}} \leq x_{pe(j)} \ln(1/x_{pe(j)})$. (1)

We now take some steps toward proving (1). Suppose $x_{pe(j)} = z \in (0, 1]$. We need some extra notation. Let $e_0, e_1, \ldots, e_t$ be the sequence of edges that we encounter as we walk up the tree starting from $j$; let $y_\ell = x_{e_\ell}$. Thus we have $z = y_0 \leq y_1 \leq y_2 \leq \cdots \leq y_t \leq 1$. Next, for $\ell = 0, 1, \ldots, t$, let $A_\ell = \sum_{j' \in (T(e_\ell) \cap S_i)} x_{pe(j')}$. Then, it is not hard to see that the left-hand side in the statement of (1) equals

$z \cdot \sum_{\ell=1}^t \frac{A_\ell - A_{\ell-1}}{y_\ell}$. (2)

The sum in (2) is clearly bounded by the maximum of the following optimization problem, whose variables are the $y_\ell$ and $A_\ell$. (The optimization problem has a maximum since the domain is a polytope and since the objective function is continuous in the domain.)

$$OPT(z, t): \text{ maximize } \sum_{\ell=1}^t \frac{A_\ell - A_{\ell-1}}{y_\ell} \text{ subject to}$$

$$A_0 = z;$$
$$y_0 = z;$$
$$y_\ell \leq 1;$$
$$A_\ell \leq A_{\ell+1}, \ \ell = 0, 1, \ldots, t-1;$$
$$y_\ell \leq y_{\ell+1}, \ \ell = 0, 1, \ldots, t-1;$$
$$A_\ell \leq y_\ell, \ \ell = 0, 1, \ldots, t.$$ (3)
Constraint (3) holds since the following constraint (4) is a constraint in our LP relaxation:

\[
\sum_{j \in (L(f) \cap S_i)} x_{pe(j)} \leq x_f \quad \text{for every edge } f \text{ and every group } S_i. \tag{4}
\]

Fix any feasible solution \(\{y_\ell, A_\ell : \ell \geq 0\}\) to the above optimization problem.

- If \(v\) is the objective function value of this solution to the optimization problem, show that

\[
v \leq 1 - z/y_1 + \sum_{\ell=1}^{t-1} (1 - y_\ell/y_{\ell+1}). \tag{5}
\]

**(5 points)**

- Take any \(\ell, 2 \leq \ell \leq t - 1\). If we keep all variables but \(y_\ell\) fixed, see when the r.h.s. of (5) is maximized. Start with this idea to show that

\[
v \leq 1 - z/y_1 + \ln(1/y_1). \tag{6}
\]

**(5 points)**

- Use (6) to show that \(v \leq \ln(1/z)\). This will then prove (1). **(5 points)**

(c). Show, using (1), that \(\Delta_i \leq \ln|S_i|\). **(5 points)**

2. We have a set \(V\) of \(n\) elements, and \(m\) distinct subsets \(S_1, S_2, \ldots, S_m\) of \(V\), each having cardinality \(t\). Our goal is to choose a subset \(W\) of \(V\) with “many” elements, subject to the constraint that no \(S_i\) (for \(i = 1, 2, \ldots, m\)) be a subset of \(W\).

Consider the following algorithm \(A\) for this problem. Let \(V\) be the set \(\{1, 2, \ldots, n\}\). Independently for each \(i \in V\), choose a number \(X_i\) uniformly at random from the set \(\{1, 2, \ldots, n^3\}\). Now define a set \(W\) as follows: for each \(i \in V\), \(i \in W\) iff there is no set \(S_j\) such that: (i) \(i \in S_j\), and (ii) for all \(k \in S_j\), \(X_i \geq X_k\).

(a). Show that \(A\) always produces a feasible solution to our problem. **(5 points)**

(b). Suppose \(i \in V\) lies in \(a_i\) of the sets \(S_1, S_2, \ldots, S_m\). Show that the expected size of the set \(W\) produced by \(A\) is at least

\[
\left(1/n^3\right) \cdot \sum_{i=1}^{n} \sum_{\ell=1}^{n^3} \left(1 - (\ell/n^3)^{t-1}\right)^{a_i}. \tag{10 points}
\]

3. Suppose we generate a random graph \(G\) from the \(G(2t, 1/2)\)-model; i.e., we take \(2t\) vertices, and put an edge between each pair of vertices independently, with probability \(1/2\). Prove that the probability that all vertices of \(G\) have degree at most \(t\), is at least \(1/4^t\). **(5 points)**

4. We are given an undirected graph \(G = (V, E)\), where \(|V| = n\) and \(|E| = m\); it is also true that the maximum degree of any vertex in \(G\) is at most twice the minimum degree.

(a) Give a lower bound on the minimum degree and an upper bound on the maximum degree, in terms of \(n\) and \(m\). **(5 points)**
(b) We are given some integer $t \leq m$. Suppose we choose each vertex independently with some probability $p$; let $X$ denote the number of edges, both of whose end-points are chosen. If we want $E[X]$ to be $t$, what should the value of $p$ be? (5 points)

(c) Consider the random process of (b), along with the value $p$ computed there. Give an upper bound on $\Pr[X \leq t/2]$; your upper bound should be of the form $e^{-ab}$, for some positive constants $a$ and $b$. (Hint: The fact that the maximum degree of $G$ is not much more than the minimum degree, may help you.) (10 points)