Problem Set 5 Solutions

Problem 1. **(a)** Consider the admissible graph. Let L_i denote the vertices located in the layer i (with distance i from the source). We note that $L_0 = s$ and $L_d = t$. Then we take the summation of number of elements in all pairs of layers

$$\sum_{i=1}^{d} |L_i| + |L_{i+1}| \leq 2n \ .$$

We conclude the inequality because we count each layer at most twice. By pigeonholing, when we add d elements to get $2n$, there must be at least one element of value $\leq 2n/d$. Thus, we conclude that $|L_i| + |L_{i+1}| \leq 2n/d$ for some i, or $(|L_i| + |L_{i+1}|)/2 \leq n/d$.

The number of edges between layers i and $i+1$ is at most[1]

$$|L_i|\,|L_{i+1}| \leq \left(\frac{|L_i| + |L_{i+1}|}{2}\right)^2 \leq \left(\frac{n}{d}\right)^2 \ .$$

Since the full graph has no edges jumping layers that are in the admissible graph (otherwise, the layers are incorrect in the admissible graph), we conclude that the cut between layers L_i and L_{i+1} has a capacity $\leq (n/d)^2$, as each edge has unit capacity. Thus, the max flow is at most $(n/d)^2$.

After k blocking flows, s and t are at a distance of k from eachother. The max flow of the residual graph is $(n/k)^2$ from above, thus we need to do at most $(n/k)^2$ more blocking flows. Setting $k = n^{2/3}$, we get the total number of blocking flows is $O(k + (n/k)^2) = O(n^{2/3} + (n/n^{2/3})^2) = O(n^{2/3} + (n^{1/3})^2) = O(n^{2/3})$.

Problem 2.

(a) We represent the problem as a bipartite graph G with vertex set *Students* $\cup$ *Recitations* and unit-capacity edges representing availability of students to attend recitations. Now use the $\mathcal{O}(m\sqrt{n})$ bipartite matching algorithm to find a maximal matching.

(b) We reduce the problem to a max flow as follows. Expand the graph G by adding a sink s, a source t, unit-capacity edges from s to each student, and edges from each recitation to t with capacity k. Any max flow in this graph gives an optimal assignment of students to recitations.

[1] First step of inequality follows from $\sqrt{ab} \leq (a+b)/2$, or geometric mean is at most arithmetic mean.

(c) Suppose we remove ℓ blocking flows, so that the distance between source and sink is $\geq \ell$. Each path in the residual graph has length $\geq \ell$ and thus passes through $\geq \frac{\ell-1}{2}$ students (since the path must alternate student, recitation, student, recitation, etc., before reaching the sink). Therefore, there exist $\leq \frac{2|\#\text{ students}|}{\ell-1} = \mathcal{O}(\frac{n}{\ell})$ disjoint paths from source to sink, so $\frac{n}{\ell}$ additional blocking flows suffice. Setting $\ell = \sqrt{n}$, this gives $\mathcal{O}(\sqrt{n})$ total blocking flows.

Each blocking flow can be computed in $\mathcal{O}(m)$ time using the same advance-retreat DFS algorithm as for unit capacity graphs. This is because the only non-unit capacity edges are the edges from recitations to the sink, and we only advance to the sink on the last step of the blocking-flow algorithm. The overall running time is thus $\mathcal{O}(m\sqrt{n})$.

Problem 3. Let $A = (a_{ij})$ be an $n \times n$ matrix with fractional entries such that all row and column sums are integral. Without loss of generality, we assume that $0 \leq a_{ij} < 1$ for all $i, j \in \{1, ..., n\}$. (We can subtract $\lfloor a_{ij} \rfloor$ from each a_{ij}, perform the procedure below, and add back the same amount afterwards.)

Consider the graph G with vertex set $\{s, t\} \cup \{v_{ij} : i, j \in \{1, ..., n\}\} \cup \{r_1, ..., r_n\} \cup \{c_1, ..., c_n\}$ where the r's and c's represent the rows and columns of A, respectively. For all $i, j \in \{1, ..., n\}$, let there be edges:

- (s, r_i) with capacity $\sum_{j'} a_{ij'}$,
- (c_j, t) with capacity $\sum_{i'} a_{i'j}$,
- (r_i, v_{ij}) and (c_j, v_{ij}) with capacity $\lceil a_{ij} \rceil$, i.e., 0 if $a_{ij} = 0$ and 1 otherwise.

Consider the s-t cut (S, T) where $S = \{s\}$. Clearly the value of this cut is $\sum_{i,j} a_{ij}$. Now consider the flow f where:

- $f(s, r_i) = \sum_{j'} a_{ij'}$,
- $f(c_j, t) = \sum_{i'} a_{i'j}$,
- $f(r_i, v_{ij}) = f(v_{ij}, c_j) = a_{ij}$.

f has value $\sum_{i,j} a_{ij}$, the same as the min cut. Therefore, f is a max flow.

Note that G has integral capacities, so G has an integral max flow g. Let $B = (b_{ij})$ be the $n \times n$ matrix where $b_{ij} = g(r_i, v_{ij})$. B is the desired integral matrix.