Solution to CSE 531 Homework Assignment 5

Prepared by Hung Q. Ngo*

November 22, 2007

Problem 1. We define the *Escape Problem* as follows. We are given a directed graph $G = (V, E)$ (picture a network of roads). A certain collection of nodes $X \subset V$ are designated as *populated nodes*, and a certain other collection $S \subset V$ are designated as *safe nodes*. (Assume that X and S are disjoint.) In case of an emergency, we want evacuation routes from the populated nodes to the safe nodes. A set of evacuation routes is defined as a set of paths in G so that (i) each node in X is the starting point of one path, (ii) the last node on each path lies in S, and (iii) the paths do not share any edges. Such a set of paths gives a way for the occupants of the populated nodes to "escape" to S, without overly congesting any edge in G.

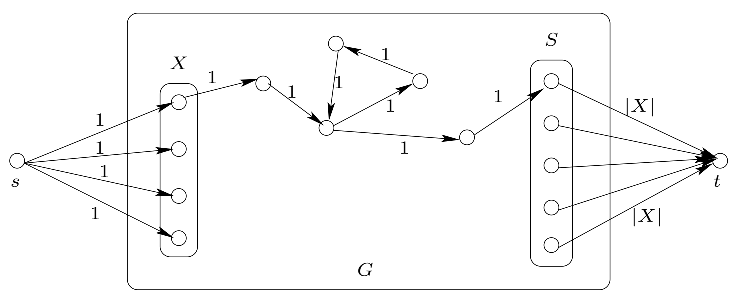
(a) Given G, X, and S, show how to decide in polynomial time whether such a set of evacuation routes exists.

(b) Suppose we have exactly the same problem as in (a), but we want to enforce an even stronger version of the "no congestion" condition (iii): we change (iii) to say "the paths do not share any *nodes*."

With this new condition, show how to decide in polynomial time whether such a set of evacuation routes exists.

(c) Provide an example with the same G, X, and S, in which the answer is yes to the question in (a) but no to the question in (b).

Solution Sketch. (a) Construct a flow network D as shown in Figure 1. All edge capacities are set to 1. Only the edges from S to t are of capacities $|X|$.



The flow network D

Figure 1: Construction of the flow network D from the graph G

Claim. The flow network has max-flow value equal to $|X|$ if and only if a set of evacuation routes exists.

*Let me know ASAP of any typos/mistakes you find. Thanks!

Proof.

($\Rightarrow$) Suppose the flow network has a max-flow value of $|X|$. (It cannot be more than $|X|$ since the cut separating the source s has capacity $|X|$.) Because all edge capacities are integral, the integrality theorem tells us that there exists an integral flow f. This flow f has $f(e) \in \{0, 1\}$ for any edge e other than edges from S to t. Similar to our proof of Menger's theorem, trace a walk of edges with flow values a starting from s. This walk has to end at t due to flow conservation constraints. Remove all cycles from this walk and we get the first X, Y-path. Then, set reduce all flow values on edges of this walk by 1, and we get a new feasible flow with value $|X| - 1$. Repeat the procedure recursively until we get $|X|$ paths.

($\Leftarrow$) If a set of evacuation routes exists, we simply set all flow values on edges of these routes to be 1, flow values of edges from the source s to each vertex in X to 1, and flow values of an edge (v, t) from S to t to be the number of evacuation routes which end at t. It is straightforward to check that the flow is feasible and has value $|X|$.

(b) Before constructing the flow network, replace each vertex v of G by v_1 and v_2, add an edge between v_1 and v_2, let all in-edges of v point to v_1 and all out-edges of v go out from v_2. The procedure is shown in Figure 2. The source s is connected to all the x_1 of vertices

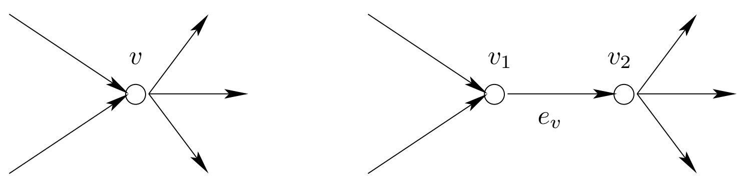


Figure 2: Replace a vertex by two vertices.

$x \in X$. The v_2 of all nodes $v \in S$ are connected to t. The capacities are set up just like in part (a). We re-run the algorithm in part (a). Since no two paths can share an edge, no two of them can share the "internal edges" e_v. We thus can "shrink" the pairs (v_1, v_2) back to obtain a set of node-disjoint paths.

(c) See Figure 3.

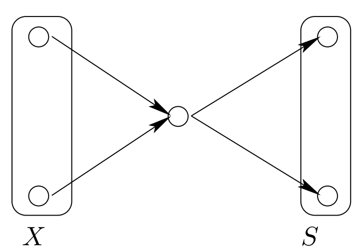


Figure 3: An example with YES for (a) but NO for (b)

□

Problem 2. Consider the problem faced by a hospital that is trying to evaluate whether its blood supply is sufficient.

The basic rule for blood transfusion is the following. A person's own blood supply has certain *antigens* present (we can think of antigens as a kind of molecular signature); and a person cannot receive blood with a particular antigen if their own blood does not have this antigen present.

Concretely, this principle underpins the division of blood into four *types*: A, B, AB, and O. Blood of type A has the A antigen, blood of type B has the B antigen, blood of type AB has both, and blood of type O has neither. Thus, patients with type A can receive only blood types A or O in a transfusion, patients with type B can receive only B or O, patients with type O can receive only O, and patients with type AB can receive any of the four types.

(a) Let s_O, s_A, s_B, and s_{AB} denote the supply in whole units of the different blood types on hand. Assume that the hospital knows the projected demand for each blood type d_O, d_A, d_B and d_{AB} for the coming week. Give a polynomial-time algorithm to evaluate if the blood on hand would suffice for the projected need.

(b) Consider the following example. The typical distribution of blood types in U.S. patients is roughly 45% type O, 42% type A, 10% type B, and 3% type AB. The hospital wants to know if the blood supply it has on hand would be enough if 100 patients arrive with the expected type distribution. There is a total of 105 units of blood on hand. The table below gives these demands, and the supply on hand.

Blood type	supply	demand
O	50	45
A	36	42
B	11	10
AB	8	3

Is the 105 units of blood on hand enough to satisfy the 100 units of demand? Find an allocation that satisfies the maximum possible number of patients. Use an argument based on a minimum-capacity cut to show why not all patients can receive blood. Also, provide an explanation for this fact that would be understandable to the clinic administrators, who have not taken a course on algorithms. (So, for example, this explanation should not involve the words *flow, cut* or *graph* in the sense we use them in our course.)

Solution Sketch. (a) Construct a flow network as shown in Figure 4. All edges in the middle have infinite (or if you want to be precise, just set it to the sum of supplies and demands). The supply suffices for the projected demand iff the max-flow value is $d_A + d_O + d_B + d_{AB}$. I will omit this typical proof here, but you should use the integrality theorem to show an "if and only if."

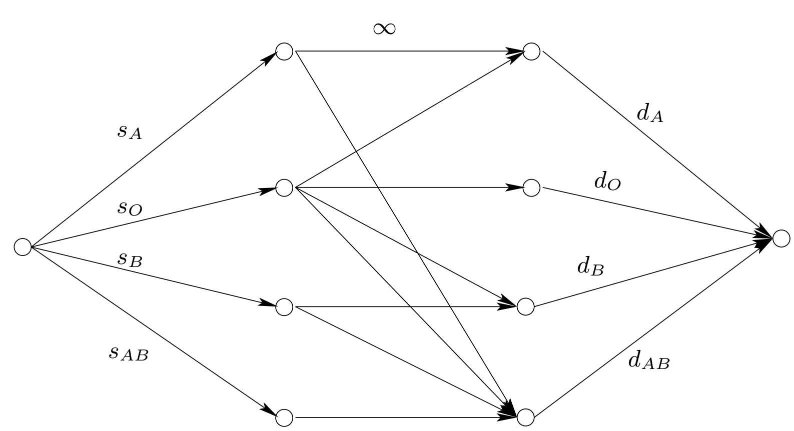


Figure 4: Flow network to check if blood supply is sufficient

(b) In this particular case, the capacities are shown in Figure 5. The cut shown has capacity $50 + 36 + 10 + 3 = 99$ which is less than the total demand of 100. Hence, the problem is infeasible.

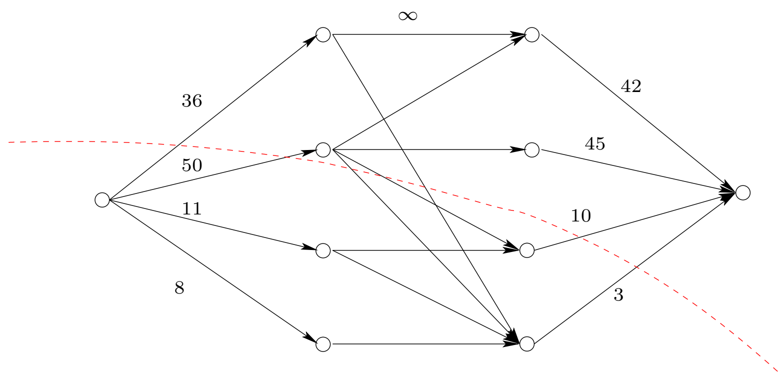


Figure 5: Flow network with particular capacity values of (b)

The cut also gives a layman explanation: the total A and O supply is only 86, which cannot satisfy the total A and O demand of 87, since B and AB cannot be used to supply A and O demands.

□

Problem 3. Let m and k be positive integers. Let $G = (L \cup R; E)$ be a bipartite graph satisfying the following conditions: (a) all vertices in L have degree m, (b) all vertices in R have degree mk.

Show that we can color the edges of G using only m colors (each edge gets one color) such that vertices in L are incident to edges with different colors, and vertices in R are incident to exactly k edges of each color.

Solution Sketch. There are several ways to prove this result. I will present here the most "pure" in terms of network flows. As usual, construct a flow network D from the graph G by adjoining a source s and a sink t as shown in Figure 6. The main idea is as follows.

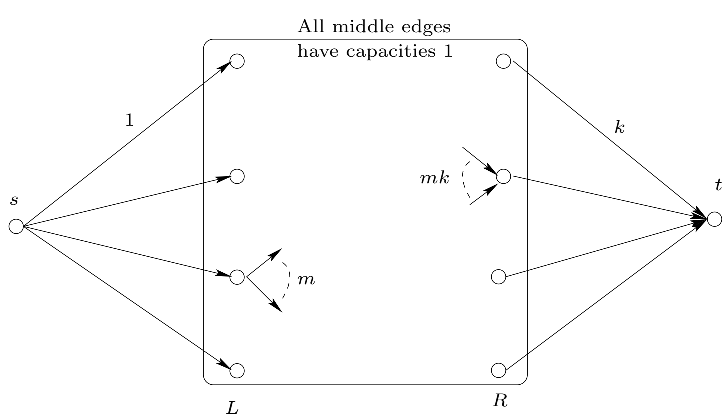


Figure 6: Flow network D constructed from the bipartite graph G

- The total number of edges of G is $|L|m = |R|mk$, hence $|L| = k|R|$.
- We will show that the flow network has max flow value $|R|k = |L|$.

- By the integrality theorem, there exists an integral flow f with value $|L| = |R|k$. Since the total flow out of s is $|L|$, it must be the case that $f(s, v) = 1$ for all $v \in L$. Similarly, it must be the case that $f(u, t) = k$ for all $u \in R$. Consider the set S of edges e in the middle with $f(e) = 1$. Due to flow conservation, every vertex on the left is incident to exactly one edge in S and every vertex on the right is incident to exactly k edges in S. This means we can color all edges in S with one color.

- Remove all edges in S, we are left with a graph whose left degree is $m-1$ and right degree is $k(m-1)$. Induction finishes the job.

We are left to show that the max flow value of the network is $|R|k = |L|$. The min cut capacity is *at most* $|R|k = |L|$, because the cut $(\{s\}, V \cup \{t\})$ has this capacity. It suffices to show that any cut has capacity *at least* $|R|k = |L|$.

Consider a generic cut $(\{s\} \cup X \cup Y, \{t\} \cup X' \cup Y')$ as shown in Figure 7. Note that X could be $\emptyset$ or L and Y could be $\emptyset$ or R.

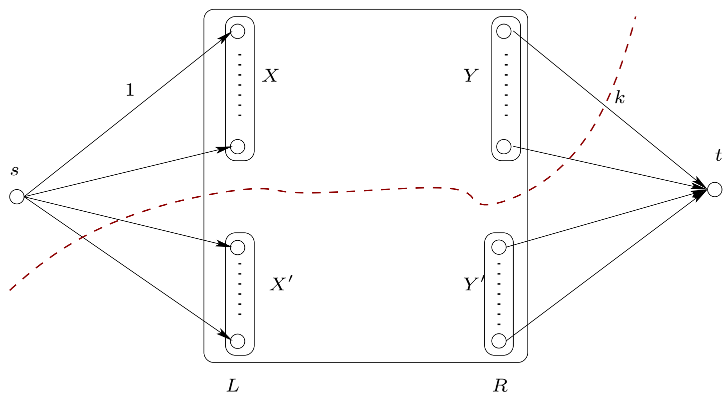


Figure 7: A generic cut

This generic cut's capacity is equal to

$$|X'| + |Y|k + \text{number of edges from } X \text{ to } Y'.$$

We want to show that this capacity is at least $k|R|$. If $|X'| + |Y|k \geq k|R|$ then we are done. Suppose $|X'| + |Y|k \leq k|R|$, or equivalently $k|Y'| - |X'| \geq 0$.

The number of edges pointing into Y' is $mk|Y'|$, of which at most $|X'|m$ can come from X'. Hence, the number of edges from X to Y' is at least $mk|Y'| - |X'|m$. In total, the cut capacity is at least

$$|X'| + |Y|k + mk|Y'| - |X'|m = (m-1)(k|Y'| - |X'|) + k(|Y| + |Y'|) \geq k|R|.$$

$\square$

Problem 4. There are n students in a certain university in the Midwest. The ith student can only take at most c_i courses in a semester. (For simplicity, let's assume that every student can take as many courses as she/he wants subject to the upper bounds c_i.) There are m courses that are going to be offered next Spring. Every course needs at least k registered students, otherwise it will be canceled. Moreover, the ith student has a set S_i of courses that she/he prefers. No student registers for a course she/he does not prefer.

(a) Describe a polynomial-time algorithm which evaluates whether the students can register for Spring courses so that no course will be canceled.

(b) Now, suppose each student belongs to one of l ethnic groups. For cultural diversity, the university imposes the constraint that a course will be canceled if all registered students are from the same ethnic group. Describe a polynomial-time algorithm to evaluate whether there is a way for students to register for courses so that no course is canceled.

Solution Sketch. (a) Construct a flow network as shown in Figure 8.

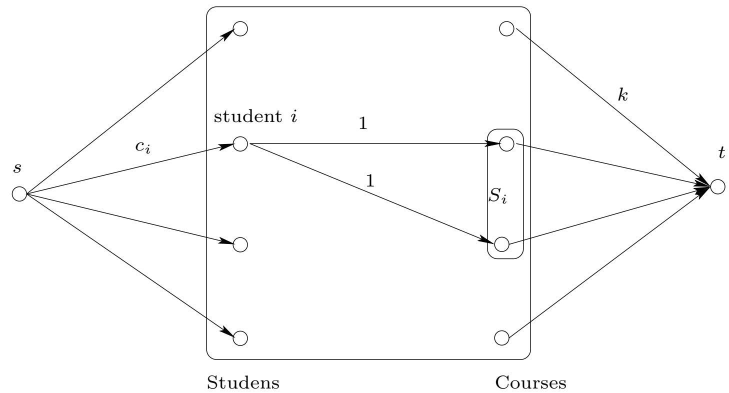


Figure 8: Flow network for question 4(a).

Claim. the students can register so that no course will be canceled if and only if the above flow network has max flow value nk, where n is the number of courses.

The proof of this is pretty standard, so I will not type the details here. Please do this on your own and seek help from me or the TAs if you have trouble expressing this kind of proof.

(b) This time, the flow network is a little bit more involved, see Figure 9. For each course v, create l "buffer vertices $v_1, \ldots, v_l$, one per ethnic group. Instead of connected directly student i to course $v \in S_i$, we connect i to the buffer vertex v_j corresponding to the ethnic group that student i belongs. The edges (v_j, v) all have capacities $k-1$. Thus, in order to get a total flow of k per course, there cannot be only students from one ethnic group j for any course because at the most this group can send $k-1$ units of flow to course v.

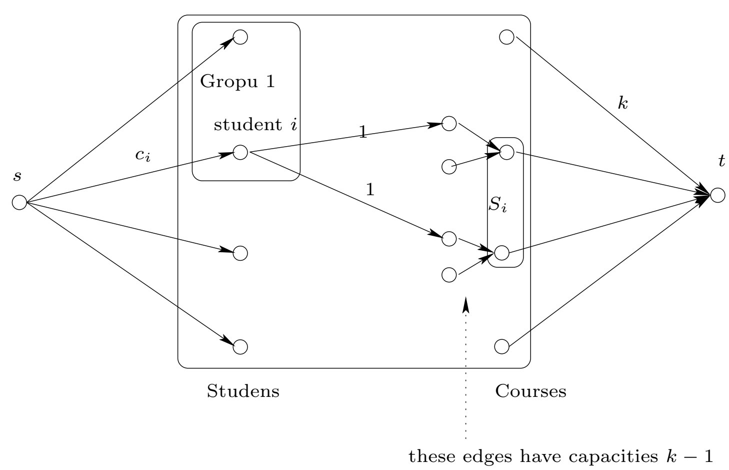


Figure 9: Flow network for question 4(b).

□

Problem 5. Let $G = (V, E)$ be a directed graph, with source $s \in V$, sink $t \in V$, and nonnegative edge capacities $\{c_e\}$. Give a polynomial-time algorithm to decide whether G has a *unique* minimum s,t-cut (i.e., an s,t-cut of capacity strictly less that that of all other s,t-cut).

Solution Sketch. Compute a minimum cut (A, B) with capacity C. Define

$$\begin{aligned} X &= \{s\} \cup \{u \mid (u,v) \in [A,B] \text{for some } v\} \\ Y &= \{t\} \cup \{v \mid (u,v) \in [A,B] \text{for some } u\} \end{aligned}$$

In English, X consists of s and all tails of edges crossing the cut, and Y consists of t and all heads of edges crossing the cut.

Increase the capacity of some edge (u, v) crossing the cut by $C+1$. In the new flow network, if there is another cut with capacity C, then that cut is different from (A, B) and it is also minimum for the original flow network. Repeat the procedure for all edges (u, v) crossing (A, B). If we found no cut with capacity C, we are now certain of the following: if there was another minimum cut other than (A, B), this cut must contain all edges crossing (A, B).

We next devise another test to check if there's another minimum cut (A', B') containing all edges crossing $[A, B]$. Note that $X \subset A'$ an $Y \subset B'$. Hence, (A', B') can be obtained from (A, B) by moving some vertices in $A - X$ into B and/or some vertices in $B - Y$ into A.

If (A', B') exists, and B' contains some vertices in $A - X$, we can safely assume that A' does not contain any vertices in $B - Y$. If A' does, we can move them back to B' while keeping the cut capacity the same. The converse is symmetric, namely A' may contain some vertices in $B - Y$ and B' does not contain any vertex from $A - X$.

Thus, we only need to consider the case when $B' = B \cup T$, where $T \subset A - X$, and $A' = A - X - T$.

How must T be? Since the capacity of (A', B') is exactly C, the capacity of the cut $(A - X - T, T)$ must be zero! Hence, to find out the existence of such T, we can go through each vertex $d \in A - X$ and compute the minimum s,d-cut capacity of the sub-network induced by A. If we found one d for which there is a cut (S, T) $(s \in S, d \in T)$ with zero capacity, then we can move T to B to obtain (A', B'). If no such d is found, (A, B) is the unique minimum cut in the original flow network.

□

Problem 6. Two optical switches R and S are connected by f optical fibers. There are totally w different wavelengths $\lambda_1, \ldots, \lambda_w$. However, due to various physical limitations, the jth fiber can only accommodate up to n_j different wavelengths (any subset of at most n_j wavelengths is OK), for $1 \leq j \leq f$.

A set C of connections are to be routed from switch R to switch S. Each connection in C is to be carried on a pre-assigned wavelength. In the set C, there are m_i connections which were pre-assigned with wavelength λ_i, $1 \leq i \leq w$.

We are to route the connections in C through (R, S), namely each connection in C is assigned to one of the f fibers such that no two connections with the same wavelength are assigned on the same fiber, and that the jth fiber does not get assigned to more than n_j connections.

Suppose $m_1 \geq \cdots \geq m_w$, and $n_1 \leq \cdots \leq n_f$. Show that the routing can be done if and only if, for all k, and l, where $0 \leq k \leq w$, $0 \leq l \leq f$, it holds that $k(f-l) + \sum_{j=1}^{l} n_j \geq \sum_{i=1}^{k} m_i$.

Solution Sketch. We first set up a flow network in the usual manner. The network D consists of vertices $\{s\} \cup \Lambda \cup \mathcal{F} \cup \{t\}$, where $\Lambda = \{1, \ldots, w\}$ represents the set of wavelengths, $F = \{1, \ldots, f\}$ represents the set of fibers. The edges are as follows: (i) there's an edge (s, i) for each $i \in \Lambda$ with capacity m_i, (ii) there's an edge (i, j) for each $i \in \Lambda$ and $j \in \mathcal{F}$ with capacity 1, and (iii) there's an edge (j, t) from each $j \in \mathcal{F}$ to the sink t with capacity n_j.

Claim. The routing can be done if and only if the above flow network has max-flow capacity $\sum_{i \in \Lambda} m_i$.

Proof. Please prove this "if and only if" on your own using the integrality theorem.

Now, D does have a cut with capacity $\sum_{i \in \Lambda} m_i$ (the cut separating s from the rest). Hence, the max-flow value would be $\sum_{i \in \Lambda} m_i$ iff all cuts in D have capacity at least $\sum_{i \in \Lambda} m_i$.

Consider a generic cut $(\{s\} \cup X \cup Y, \{t\} \cup X' \cup Y')$, where $X \subset \Lambda$ has size k and $Y \subset \mathcal{F}$ has size l, where $0 \leq k \leq w$ and $0 \leq l \leq f$. This cut's capacity is exactly

$$\sum_{i \in X'} m_i + k(f - l) + \sum_{j \in Y} n_j.$$

Hence, the cut's capacity is at least $\sum_{i \in \Lambda} m_i$ if and only if

$$\sum_{i \in X'} m_i + k(f - l) + \sum_{j \in Y} n_j \geq \sum_{i \in \Lambda} m_i,$$

which is equivalent to

$$k(f - l) + \sum_{j \in Y} n_j \geq \sum_{i \in X} m_i.$$

Consequently, the routing can be done if and only if, for any $0 \leq k \leq w$ and $0 \leq l \leq f$, and for any subset X of Λ of size k and any subset Y of $\mathcal{F}$ of size l, it holds that

$$k(f - l) + \sum_{j \in Y} n_j \geq \sum_{i \in X} m_i.$$

This relation holds iff it holds for the smallest l values of n_j and the largest k values of m_i. This is exactly what we're trying to prove. $\square$