Extra lecture notes : Day 13

The extra stuff today explores the connection between matchings in bipartite graphs and flows in networks.

We begin with the following corollary of the FF algorithm used to prove the Max-flow min-cut (MFMC) theorem :

Corollary to MFMC *Let G be a network in which the capacity of every edge is an integer. Then there exists a maximum flow for which the flow along every edge is also an integer.*

PROOF : In each step of the FF algorithm, when we augment a given flow along an augmenting $s - t$ path, we do so by an integer value.

The link to matching problems is most clearly seen, I think, via

König's theorem *Let G be a bipartite graph. The maximum size of a matching in G equals the minimum number of vertices needed to cover all the edges of G.*

FIRST PROOF (USING HALL'S THEOREM) : Let m be the maximum size of a matching in G and n the minimum number of vertices needed to cover all the edges.

Easy part : $m \leq n$.

If S is a covering set of vertices of size n and M any set of edges, then each edge in M must cover a vertex in S. If M is a matching then each vertex in S is covered at most once, and so M cannot possess more edges than there are vertices in S.

Harder part : $m \geq n$.

Again let S be a covering set of vertices of size n. We'll show that there must exist a matching in G of size n. Divide the vertices of G into the following four subsets :

$$A := X \backslash S,$$

$$
\begin{aligned}
B &:= X \cap S, \\
C &:= Y \cap S, \\
D &:= Y \backslash S.
\end{aligned}
$$

Then $|S| = |B| + |C| = n$. That S covers all edges in G means that there are no edges between A and D.

Claim : There is a perfect matching for B in the bipartite graph (B, D).

Suppose otherwise. Then by Hall's theorem, there must exist a subset E of B with fewer than $|E|$ neighbours in D. Denote this set of neighbours by $J_D(E)$. But then the set of vertices

$$S^* := C \cup (B \backslash E) \cup J_D(E)$$

also covers all edges in G and is smaller than S, contradicting minimality of the latter. This establishes our claim.

By a similar argument there exists a perfect matching for C in the bipartite graph (A, C). Putting these two matchings together, we get a matching in G consisting of $|B| + |C|$ edges, v.s.v.

Second Proof (using MFMC Theorem) : Let G^* be the following network :

(i) every edge in G^* has capacity 1,

(ii) G^* contains a copy of G, with all edges now directed from X to Y,

(iii) G^* contains two extra vertices s and t (placed to the left of X and to the right of Y respectively). There is an edge directed from s to each vertex of X, and an edge directed to t from each vertex of Y.

By the MFMC theorem, our proof is then complete once we establish the following two facts :

Fact 1 : The maximum strength of a flow in G^* equals the maximum size of a matching in G.

Fact 2 : There is a cut (S^*, T^*) of minimum capacity such that G is covered by a set of $c(S^*, T^*)$ vertices.

More precsiely, Facts 1 and 2 establish, in the notation of our first proof, the inequality $m \geq n$. But as we've seen, the reverse inequality is trivial.

Proof of Fact 1 : Let f be a flow of maximal strength constructed by applying the FF algorithm. By the above Corollary to MFMC, f is integer-valued. In the case of G^* this means that the flow along every edge is either 0 or 1. The edges in the bipartite G which have positive flow then form a matching M in G whose size is just $|f|$. Conversely, given any matching M_0 in G we can construct a flow in G^* of strength $|M_0|$ just by saturating all the edges in M_0 and all adjacent edges through s and t. It follows that the matching M is maximal, which establishes Fact 1.

Proof of Fact 2 : Let (S_0, T_0) be any cut in G^*. We divide the vertices of G into the following four subsets :

$$A := S_0 \cap X, \tag{1}$$
$$B := T_0 \cap X, \tag{2}$$
$$C := S_0 \cap Y, \tag{3}$$
$$D := S_0 \cap Y. \tag{4}$$

Since there is an edge from s to each vertex of B, and similarly one from each vertex of C to t, we have that

$$c(S_0, T_0) = |B| + |C| + \text{no. of edges } A \to D. \tag{5}$$

Suppose there is at least one edge in G from A to D, say the edge (v, w). Let now (S_1, T_1) be the cut of G^* given by

$$S_1 := S_0 \backslash \{v\}, \quad T_1 := T_0 \cup \{v\}.$$

I claim that $c(S_1, T_1) \leq c(S_0, T_0)$. For on the one hand, the capacity goes up by at most 1 because the edge (s, v) now goes across the cut, but on the other hand it goes down by at least 1 since the edge (v, w) no longer goes across the cut.

Now if we start with a cut (S_0, T_0) of minimal capacity then, by repeated application of the above idea, we may replace it with a cut (S^*, T^*), also of minimal capacity, such that, in the notation of (1)-(4), there are no edges in G between A and D. But this just means that the vertices in $B \cup C$ cover all edges in G. But, by (5), we also have in this case that $c(S^*, T^*) = |B| + |C|$. Hence, we have indeed established the statement of Fact 2.