Lecture 1: Matchings on bipartite graphs

Some good texts on Graph Theory are [3, 12–14].

1 Basic Concepts

An *undirected graph* $G = (V, E)$ consists of a finite set V of *vertices* and a finite multi-set of unordered pairs E of *edges*. A *loop* is an edge of the form (v, v). When E is a proper set (not a multi-set), G is said to be *simple*. When E is an ordered set, the graph is said to be *directed*.

An edge $e = (u, v) \in E(G)$ is said to be *incident* to u and v, while u and v are *adjacent*. The *complement* of a graph G, denoted by $\bar{G}$ is the graph whose vertex set is the same as that of G, and two vertices in $\bar{G}$ are adjacent iff they are not adjacent in G.

A *walk* is a sequence of vertices $v_1, \dots, v_k$ where $v_i v_{i+1} \in E(G)$. A *path* is a walk without repeated vertex in the sequence. A path that starts with u and end with v is called a path from u to v or a (u, v)-path. The length of a path is the number of edges in the path. The *distance* $d(u, v)$ between two vertices u and v is the minimum length of (u, v)-paths. Note that $d(u, v)$ could be infinite. A *cycle* is a walk which starts and ends at the same vertex and all the vertices in the middle do not repeat in the walk. An n*-cycle* or a cycle of length n is a cycle with n edges. The *girth* of a graph G is the minimum length of a cycles.

The *degree* $d_G(v)$ of a vertex v is the number of edges incident to v. A graph is *regular* if all vertices have the same degree. We often use $\Delta(G)$ and $\delta(G)$ to denote the maximum and minimum degree of G, respectively. A graph is k*-regular* if $\Delta(G) = \delta(G) = k$.

A *subgraph* $G' = (V', E')$ of $G = (V, E)$ is a graph such that $V' \subseteq V$ and $E' \subseteq E$. An *induced subgraph* $G' = (V', E')$ of $G = (V, E)$ is a subgraph such that for any $u', v' \in V'$, $u'v' \in E$ implies $u'v' \in E'$. When $V' = V$, G' is said to be a *spanning subgraph* of G.

A graph G is *connected* if there is a path between any two vertices in G. It is *disconnected* otherwise. A *component* of a graph is a maximal connected induced subgraph. A *tree* is a graph with no cycle. A *forest* is a graph all of whose components are trees. A *spanning tree* of a graph G is a spanning subgraph of G which is also a tree.

For any graph G and a subset V' of $V(G)$, we use $G - V'$ to denote the graph obtained by removing all vertices in V' and the edges one of whose end points in is V'. For any subset E' of edges, we use $G - E'$ to denote $(V, E - E')$.

A subset C of vertices is called a *vertex cut* if $G - C$ is not connected. A subset S of edges is said to be an *edge cut* if $G - S$ is disconnected. The *vertex-connectivity* of G, denoted by $\kappa(G)$, is the size of a minimum vertex cut. Similarly, the *edge-connectivity*, denoted by $\kappa'(G)$, is the size of a minimum edge cut. If a vertex cut (edge cut) contains one vertex (edge) only, then the vertex (edge) is called a *cut vertex* (*cut edge* or *bridge*). A graph is k*-connected* if $k \le \kappa(G)$, and k*-edge-connected* if $k \le \kappa'(G)$.

An n*-factor* of a graph G is an n-regular subgraph of G. A matching of G is a 2-factor of G, and it is said to be *perfect* if it contains all vertices of G.

A *bipartite graph* G is a graph whose vertex set $V(G)$ can be partitioned into two non-empty subsets X and Y. This partition is often called the *bipartition* of V. The sets X and Y are often called the *color classes* of G.

A *vertex coloring* of G is a mapping f from V into some finite set C such that $uv \in E$ implies $f(u) \neq f(v)$. Elements of C are called the colors. A graph is k*-colorable* if there is a coloring with at most k colors. The *chromatic number* $\chi(G)$ is the least k such that G is k-colorable. An *edge coloring* of G is a mapping f from E into some finite set C such that two edges of E which are incident to the same vertex have different colors. The concepts of k*-edge-colorable*, and the *chromatic index* $\chi_e(G)$ are defined in the obvious way.

The *genus* of a graph G, denoted by $\gamma(G)$, is the smallest genus of an orientable surface onto which G can be embedded so that no two edges intersect. G is said to be *planar* if $\gamma(G) = 0$.

The *line graph* $L(G)$ of a graph $G = (V, E)$ is the graph whose vertex set is E and whose edge set is E' where $e_1e_2 \in E'$ iff e_1 and e_2 are incident to the same vertex in G.

2 Introduction to matching theory

In this section, we assume all graphs are simple graphs for simplicity, although many results hold for multi-graphs, too. Many parameters of a given graph G are of interest to us. We have seen $\chi(G)$ and $\chi_e(G)$. Let's visit some more.

The size of the largest matching in a graph G, called the *matching number* of G, is denoted by $\nu(G)$. The corresponding matching is called the *maximum matching* of G. For any matching M of G, an M*-alternating path* is a path of G which alternates between edges in M and not in M; an M*-augmenting path* is an M-alternating paths which starts and ends at edges not in M.

Exercise 2.1. Prove that a matching M of a graph G is maximum iff there is no M-augmenting path.

A subset $U \subseteq V(G)$ is called a *vertex cover* of G iff every edge of G is incident to at least one vertex in U. The size of any smallest vertex cover of G is called the *vertex covering number* of G, and is denoted by $\tau(G)$.

An *edge cover* of G is a set of edges whose set of end points is $V(G)$. The size of any smallest edge cover of G is denoted by $\rho(G)$, and is called the *edge covering number* of G.

A set of vertices is *independent* if there's no edge between any two of them. The size of any maximum independent set is called the *independent number* of G, and is denoted by $\alpha(G)$.

This section, beside other things, relates the last four parameters in a very nice way.

2.1 General Results

Let us start with the so-called Gallai identities.

Theorem 2.2 (Gallai Identities, 1959 [7]). *For any graph G, let $n = V(G)$, then*

(i) $\alpha(G) + \tau(G) = n$.

(ii) $\nu(G) + \rho(G) = n$ *if G has no isolated vertex.*

Proof. The basic idea of both proofs is to show that the left side is both $\geq$ and $\leq$ the right hand side.

(i) Let C be a vertex cover of size $\tau(G)$. Then, $V(G) - C$ is an independent set of size $n - \tau(G)$, which implies $\alpha(G) \geq n - \tau(G)$. Conversely, for any independent set I of size $\alpha(G)$, $V(G) - I$ is a vertex cover, implying $n - \alpha(G) \geq \tau(G)$.

(ii) Consider an edge cover L of minimum size $\rho(G)$. Since L is minimal it has to be a union of s stars. The number of vertices in each star is one more than the number of L's edges in the star, and $s \leq \nu(G)$ since taking an edge from each star forms a matching. This means $n = s + |L| \leq \nu(G) + \rho(G)$.

Conversely, consider a maximum matching M of size $\nu(G)$. The set $U = V(G) - V(M)$ is an independent set. For each vertex in U pick an edge incident to it (no isolated vertex). Call the set of edges E'. Clearly E' along with $E(M)$ form an edge cover of G, which means $\rho(G) \leq (n - 2\nu(G)) + \nu(G)$.

□

The following result also gives us some of intuition into seeing the relationship between these graph parameters.

Theorem 2.3. *We have*

(i) A minimal edge cover is minimum iff it contains a maximum matching.

(ii) A maximal matching is maximum iff it is contained in a minimum edge cover.

Proof. (i) We can assume G does not have any isolated vertex, otherwise there is no edge cover.

For necessity, let L be a minimum edge cover, which as we have noticed consists of a set of stars. Since $|L| = \rho(G) = n - \nu(G)$, the number of stars is exactly $\nu(G)$. This means the matching obtained by taking one edge from each star has size $\nu(G)$, i.e. it is a maximum matching.

For sufficiency, let L be a minimal edge cover which contains a maximum matching. The fact that L is minimal implies that L is a collection of stars. So L has exactly $\nu(G)$ stars, i.e. $|L| = n - \nu(G) = \rho(G)$. Hence, L has to be minimum.

(ii) Let M be a maximum matching. Let L be the edge cover obtained by taking an arbitrary edge incident to each vertex of $V(G) - V(M)$ along with all edges of M. Clearly $|L| = \nu(G) + n - 2\nu(G) = \rho(G)$. So L is a minimum edge cover which contains M.

Conversely, let M be any maximal matching which is contained in a minimum edge cover L. The edges of M must come from different starts of L. Moreover, M's maximality implies that each star of L contributes at least an edge for M, so $|M| = n - \rho(G) = \nu(G)$.

□

Exercise 2.4. Show that for any graph G, $\nu(G) \leq \tau(G) \leq 2\nu(G)$.

2.2 Bipartite Graphs

Many practical problems can be formulated in terms of matching problems on bipartite graphs. In this section we restrict our attention to bipartite graphs only.

Theorem 2.5 (König's Minimax Theorem, 1931 [11]). *If G is bipartite, then $\tau(G) = \nu(G)$.*

This theorem is also referred to as the König-Egerváry theorem as Egerváry came up with the same result in [5]. We use $\Gamma_G(X)$ to denote the set of neighbors of X in a graph G. We shall drop the subscript G when there's no confusion.

Theorem 2.6 (P. Hall, 1935 [9]). *Let $G = (A, B; E)$ be a bipartite graph. Then G has a complete matching from A into B if and only if*

$$|\Gamma(X)| \geq |X|, \forall X \subseteq A.$$

Theorem 2.7 (Frobenius, 1917 [6]). *Let $G = (A, B; E)$ be a bipartite graph. Then G has a perfect matching if and only if $|A| = |B|$ and*

$$|\Gamma(X)| \geq |X|, \forall X \subseteq A.$$

Frobenius' Theorem is often called the *Marriage Theorem*. It is interesting to note that all three theorems are equivalent, and the proof of their equivalences isn't so hard to find.

Proof of the equivalence of König's Minimax, Frobenius, and P. Hall's Theorems. We shall show a circular implication.

- **König $\Rightarrow$ Hall.** Necessity is obvious. For sufficiency, assume that for all $X \subseteq A$ we have $|\Gamma(X)| \geq |X|$. Let C be a vertex cover such that $|C| = \tau(G)$ (which is $\leq |A|$ since A is a vertex cover). If $|C| = |A|$, then we are done since that would mean $|A| = \nu(G)$. Assume $|C| < |A|$. Then, since
$$|\Gamma(A - C)| \geq |A - C| = |A| - |A \cap C| > |B \cap C|$$
there is an edge from $A - C$ to $B - C$, contradicting the fact that C is a vertex cover.

- **Hall $\Rightarrow$ Frobenius.** This is immediate. Frobenius' Theorem is clearly a special case of Hall's Theorem.

- **Frobenius $\Rightarrow$ König.** The fact that $\nu(G) \leq \tau(G)$ is obvious since each edge of an matching needs at least one vertex to cover it. We use Frobenius' theorem to show sufficiency. Let C be a vertex cover of G of minimum size $\tau(G)$. To show $|C| \leq \nu(G)$, we only need to find a matching M of G so that $|C| = |M|$. This matching $|M|$ shall be formed by take the union of two matchings M_1 from $A \cap C$ into $B - C$ and M_2 from $B \cap C$ into $A - C$. The constructions of these two sub-matchings are symmetric. Firstly, for every $X \subseteq A \cap C$, $|\Gamma(X)| \geq |X|$, because otherwise we could replace X by $\Gamma(X)$ for a smaller vertex cover. Note that this implies $|A \cap C| \leq |B - C|$. Add dummy vertices to $|A \cap C|$ to make its size equal $|B - C|$, connect all dummy vertices to all vertices in $|B - C|$, and we get Frobenius' conditions satisfied. This implies there is a complete matching from $A \cap C$ into $B - C$, which is our M_1.

□

Given the previous proof, we only need to show one theorem to get the rest. We show one proof (among many) of P. Hall's theorem here, leaving the independent proofs of the other two theorems as exercises.

A Proof P. Hall's Theorem. This is a "proof from the book" [1], courtesy of the great mathematician Paul Erdös. Necessity is obvious. We show sufficiency by induction on $|A|$. When $|A| = 0, 1$, the theorem trivially holds. Suppose $|A| \geq 2$.

Suppose for all $X \subset A$, we have $|\Gamma(X)| > |X|$. Let ab be an edge of G. Let $G' = G - a - b$, then, G' is a bipartite graph with color classes A' and B'. For any $X' \subseteq A'$, we have $|\Gamma_{G'}(X')| \geq |\Gamma_G(X')| - 1 > |X'| - 1$, so that $|\Gamma_{G'}(X')| \geq |X'|$ in G'. Induction hypothesis implies there is a complete matching from A' into B', which along with ab forms a complete matching from A into B.

Now, suppose there is an $X \subset A$ so that $|\Gamma(X)| = |X|$. Let G_1 be the subgraph of G induced by $X \cup \Gamma(X)$, and $G_2 = G - X - \Gamma(X)$. Let A_i, B_i ($i = 1, 2$) be the color classes respectively. It is easy to verify that G_1 and G_2 satisfy the matching conditions, implying there is complete matching of A_i into B_i. These two matchings together form a complete matching from A into B that we are looking for. □

Exercise 2.8. Prove Theorem 2.5 independent of the other two theorems.

Exercise 2.9. Prove Theorem 2.7 independent of other two theorems.

Exercise 2.10. Show that if G is bipartite, then

$$\rho(G) = \alpha(G).$$

Exercise 2.11 (System of Distinct Representatives). Let $\mathcal{S} = \{S_1, \dots, S_n\}$ be a collection of sets. A System of Distinct Representatives (SDR) of $\mathcal{S}$ is a set of n distinct elements $s_1, \dots, s_n$ such that $s_i \in S_i, \forall i$. Show that $\mathcal{S}$ has an SDR iff for every k, $0 \leq k \leq n$, the union of any k of the sets $S_1, \dots, S_n$ has cardinality at least k.

Exercise 2.12. Suppose the elements of $\mathcal{S} = \{S_1, \dots, S_n\}$ all have size $k \geq 1$, and further suppose that no element is contained in more than k sets. Show that there exist k SDR's such that for any i the k representatives of S_i are distinct and thus together form the set S_i.

Exercise 2.13. Let $G = (A, B; E)$ be a bipartite graph. Suppose $S \subseteq A$, $T \subseteq B$, and that there is a matching from S into B and one from T into A. Show that there is then a matching in G covering both S and T.

A *doubly stochastic matrix* is a real, non-negative square matrix whose row and column sums are all 1. A *permutation matrix* is a 01-matrix where there is exactly one 1 in each row and one 1 in each column. A *permutation set* of a permutation matrix A is a subset of n entries of A with no two from the same row or the same column.

Lemma 2.14. *Every doubly stochastic matrix has a permutation set of non-zero entries.*

Proof. Let A be a doubly stochastic matrix. Let $G = (U, V; E)$ be a bipartite graph constructed from A as follows. The sets U and V represent the rows and columns of A, respectively. There is an edge $(u, v) \in E$, where $u \in U, v \in V$, if and only if the entry a_{uv} is not zero. A permutation set of A then corresponds to a perfect matching of G. We shall apply Frobenius theorem here.

Let $X \subseteq U$ be any subset of rows of the matrix A. We want to verify that $|\Gamma(X)| \geq |X|$, namely the number k of *different* columns of A with non-zero entries in X is at least $|X|$. Note that the sum of non-zero entries is X is at most k, and is equal to $|X|$. Hence, $|X| \leq k$ as desired. $\square$

Exercise 2.15 (Birkhoff – von Neumann Theorem). Show that any doubly stochastic matrix A can be written as a convex combination of permutation matrices, namely

$$A = \alpha_1 P_1 + \dots + \alpha_k P_k,$$

where $\alpha_i > 0$ and P_i is a permutation matrix for all i, and $\sum_{i=0}^{k} \alpha_i = 1$. Also show that $k \leq n^2 - n + 1$.

Theorem 2.16 (König's Line Coloring Theorem, 1916 [10]). *For every bipartite graph G, $\chi_e(G) = \Delta(G)$. Here $\chi_e(G)$ is the chromatic index of G, i.e. $\chi_e(G)$ is the minimum integer so that a $\chi_e(G)$-edge-coloring of G exists, and $\Delta(G)$ is the maximum degree of all vertices in G.*

Proof. G can be embedded in a Δ-regular bipartite graph by adding dummy vertices and edges into G (how?). We only need to show that every k-regular bipartite (multi-) graph is k-edge-colorable.

Suppose G is k-regular. It is easy to verify P. Hall's matching condition on G, hence G contains a perfect matching. Color all edges of this matching with one color, then remove them from G we obtain a $(k-1)$-regular bipartite graph. Repeat this process k times and we are done. $\square$

Theorem 2.17 (de Werra (1971, 1975) [2]). *Let k be any positive integer and G be a bipartite graph. Then, G can be written as the union of k edge-disjoint spanning subgraphs $G_1, \dots, G_k$ such that for each $v \in V(G)$:*

$$\left\lfloor \frac{d_G(v)}{k} \right\rfloor \le d_{G_i}(v) \le \left\lceil \frac{d_G(v)}{k} \right\rceil, \forall i, 1 \le i \le k.$$

Proof. Build a bipartite graph G' from G by splitting each vertex $v \in V(G)$ into $\lfloor \frac{d_G(v)}{k} \rfloor$ of degree k, and possibly one more vertex of degree $deg(v) - k\lfloor \frac{d_G(v)}{k} \rfloor$. The graph G' has maximum degree k, and thus by Theorem 2.16 is k-colorable. The sets of same-color edges are matchings $M_1, \dots, M_k$ of G'. Now, "collapsing" G' back to G, and let G_i be the graph formed by edges of M_i after collapsing. It is obvious that for each v, we must have

$$\left\lfloor \frac{d_G(v)}{k} \right\rfloor \le d_{G_i}(v) \le \left\lceil \frac{d_G(v)}{k} \right\rceil, \forall i, 1 \le i \le k.$$

□

Exercise 2.18. Show that König's Line Coloring Theorem is equivalent to de Werra's Theorem.

Exercise 2.19 (Gupta, 1967 [8]). Show that if G is a bipartite graph with minimum degree $\delta = \delta(G)$, then G is the union of δ edge-disjoint edge-covers.

It is even more interesting to know that they are all equivalent to the celebrated Dilworth's theorem. Since Dilworth's theorem requires the language of posets, I won't discuss these results but mention them here for those who have deeper background on posets.

Theorem 2.20 (Dilworth, 1950 [4]). *In any finite poset, the size of any largest antichain equals the size of any smallest chain decomposition.*

Exercise 2.21. Show that Theorem 2.20 is equivalent to Theorem 2.5.

Exercise 2.22. State and prove a "dual" version of Dilworth's theorem.

Exercise 2.23. Two network routers R and S are connected by f fibers. The jth fiber can accommodate up to n_j **different** wavelengths, $1 \le j \le f$.

A set C of connections are routed through (R, S). Each connection in C is to be carried on a pre-assigned wavelength. There are w different wavelengths. In C, there are m_i connections on the ith wavelength, $1 \le i \le w$.

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

Exercise 2.24 (Common System of Distinct Representatives). Let $\mathcal{X} = \{X_1, \dots, X_m\}$ be a collection of sets. A set of distinct elements $X = \{x_1, \dots, x_m\}$ is called a *system of distinct representatives* of $\mathcal{X}$ if there exists a one-to-one mapping $\phi : X \to \mathcal{X}$ such that $x_i \in \phi(x_i), \forall i = 1 \dots m$.

Let $\mathcal{A} = \{A_1, \dots, A_m\}$ and $\mathcal{B} = \{B_1, \dots, B_m\}$ be two collections of subsets of $[n] = \{1, \dots, n\}$, $m \le n$. A common system of distinct representatives (CSDR) is a set $S = \{s_1, \dots, s_m\}$ of m (different) elements such that S represents both $\mathcal{A}$ and $\mathcal{B}$. (Note that the one-to-one mappings from S to $\mathcal{A}$ and $\mathcal{B}$ do not need to be the same.)

Show that $\mathcal{A}$ and $\mathcal{B}$ have a CSDR if and only if

$$\left| \left(\bigcup_{i \in I} A_i \right) \cap \left(\bigcup_{j \in J} B_j \right) \right| \ge |I| + |J| - m, \text{ for all } I, J \subseteq [m].$$

Exercise 2.25. Let m, k be positive integers. Let $G = (A, B; E)$ be a bipartite (multi) graph satisfying the following conditions: (a) all vertices in A have degree m, (b) all vertices in B have degree mk.

Show that we can color the edges of G with m colors such that vertices in A are incident to edges with different colors, and vertices in B are incident to exactly k edges of each color.

References

[1] M. Aigner and G. M. Ziegler, *Proofs from The Book*, Springer-Verlag, Berlin, second ed., 2001. Including illustrations by Karl H. Hofmann.

[2] D. de Werra, *Balanced schedules*, INFOR—Canad. J. Operational Res. and Information Processing, 9 (1971), pp. 230–237.

[3] R. Diestel, *Graph theory*, vol. 173 of Graduate Texts in Mathematics, Springer-Verlag, New York, second ed., 2000.

[4] R. P. Dilworth, *A decomposition theorem for partially ordered sets*, Ann. of Math. (2), 51 (1950), pp. 161–166.

[5] J. Egerváry, *Matrixok kombinatorikus tulajdonságairól*, Mathematikai és Fizikai Lápok, 38 (1931), pp. 19–28.

[6] Frobenius, *Über zerlegbare determinanten*, Sitzungsber. König. Preuss. Akad. Wiss., XVIII (1917), pp. 274–277. Jbuch. 46.144 [xv, 6].

[7] T. Gallai, *Über extreme Punkt- und Kantenmengen*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math., 2 (1959), pp. 133–138.

[8] R. Gupta, *A decomposition theorem for bipartite graphs (results)*, Gordon and Breach, 1967, pp. 135–136. Theory of Graphs (International Symposium, Rome, 1966), Ed: P. Rosenstiehl.

[9] P. Hall, *On representatives of subsets*, J. London Math. Soc., 10 (1936), pp. 26–30.

[10] D. König, *Über graphen und ihre anwendung auf determinantentheorie und mengenlehre*, Math. Ann., 77 (1916), pp. 453–465.

[11] ———, *Graphen und matrizen*, Mathematikai és Fizikai Lápok, 38 (1931), pp. 116–119.

[12] L. Lovász, *Combinatorial problems and exercises*, North-Holland Publishing Co., Amsterdam, second ed., 1993.

[13] L. Lovász and M. D. Plummer, *Matching theory*, North-Holland Publishing Co., Amsterdam, 1986. Annals of Discrete Mathematics, 29.

[14] D. B. West, *Introduction to graph theory*, Prentice Hall Inc., Upper Saddle River, NJ, 1996.