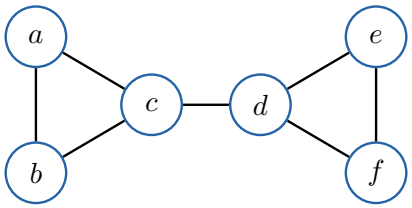
Graph 1

Mingxun Zhou (mingxunz@ust.hk)

COMP2711 Discrete Mathematical Tools

1 Graph models and degrees

Graphs model relationships: people and friendships, junctions and roads, or courses and prerequisites. We first describe graphs using vertices, edges, and degrees; then use counting arguments to prove their properties. Finally, Hall's theorem tells us when a bipartite graph admits a matching covering one part.



MUST UNDERSTAND

A **simple undirected graph** is a pair $G = (V, E)$, where V is a nonempty set of vertices and E is a set of unordered pairs of distinct vertices. Thus it has no loops and no multiple edges. We usually write $|V| = n$ for the number of vertices and $|E| = m$ for the number of edges. In the picture,

$$V = \{a, b, c, d, e, f\}, \qquad E = \{ab, ac, bc, cd, de, df, ef\},$$

where ab abbreviates $\{a, b\}$. Vertices a, b are **adjacent** when $ab \in E$. That edge is **incident** with a and b. The **neighbors** of c are $N(c) = \{a, b, d\}$.

A **subgraph** $H = (V', E')$ of $G = (V, E)$ keeps some vertices and some edges: $V' \subseteq V$, $E' \subseteq E$, and both endpoints of every edge in E' belong to V'. We need not keep every original edge between the retained vertices.

MUST UNDERSTAND

The **degree** $\deg(v)$ of a vertex counts its incident edge-ends. In a simple graph this is the number of neighbors. The pictured degrees, in vertex order, are $2, 2, 3, 3, 2, 2$. A vertex of degree zero is isolated.

Other edge models

Friendship is mutual, but following someone need not be. A **directed graph** associates each edge with an ordered pair (u, v): an arrow from u to v. In-degree $\deg^-(v)$ counts incoming edges; out-degree $\deg^+(v)$ counts outgoing edges. For arcs $(b, a), (c, a), (a, d)$, the in-degrees in order a, b, c, d are $(2, 0, 0, 1)$ and the out-degrees are $(1, 1, 1, 0)$. A directed loop contributes one to each count.

Using Rosen's convention, an undirected **multigraph** allows multiple edges but no loops; a **pseudograph** also allows loops. Distinct edges can have the same endpoints. For edges ab, ab, bc, cc, the degrees are $2, 3, 3$: the loop at c contributes two edge-ends. Multiple transport links and a route returning to its starting point motivate these models.

Two degree identities

If 12 people each have five friends within the group, adding their friend counts gives 60. Each friendship is counted at both endpoints, so there are $60/2 = 30$ friendships.

MUST UNDERSTAND

Handshaking theorem. For every finite undirected graph, including graphs with loops and multiple edges,

$$\sum_{v \in V} \deg(v) = 2m.$$

Each edge contributes two edge-ends to this sum, even when both ends meet the same vertex. In the opening example the sum is $2 + 2 + 3 + 3 + 2 + 2 = 14$, giving seven edges.

For a directed graph, each edge contributes one to the out-degree of its tail and one to the in-degree of its head. Consequently,

$$\sum_{v \in V} \deg^-(v) = m = \sum_{v \in V} \deg^+(v).$$

The earlier three-arc example has both sums equal to three.

2 Important types of graphs

A **complete graph** K_n joins every pair of distinct vertices. It models a group in which everyone is friends with everyone else. Each vertex has degree $n - 1$, and choosing the two endpoints gives

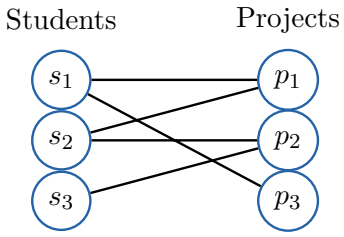
$$m = \binom{n}{2} = \frac{n(n-1)}{2}.$$

A **cycle graph** C_n, $n \geq 3$, models a ring network. Its edges are $v_1v_2, v_2v_3, \ldots, v_{n-1}v_n, v_nv_1$. Every vertex has degree two and there are n edges.

An undirected cycle returns to its starting vertex without repeating other vertices and has at least three edges. A **tree** is a connected undirected graph with no cycles. Here connected means that any two vertices can be joined by a sequence of adjacent vertices. A directed cycle follows arrows back to its start, without repeating other vertices. A **directed acyclic graph** (DAG) has no directed cycle. Course prerequisites form a DAG when there is no circular chain of requirements.

An acyclic simple undirected graph on $n \geq 1$ vertices has at most $n-1$ edges. To see this, a longest path with no repeated vertices has an endpoint of degree at most one; otherwise we could extend the path or create a cycle. Remove that endpoint and apply induction. A path attains $n-1$. A simple DAG has at most $\binom{n}{2}$ edges, since loops and opposite pairs of arrows would create directed cycles. Orient every edge from the smaller to the larger label to attain the bound.

A simple graph is **bipartite** if $V = L \cup R$, $L \cap R = \varnothing$, and every edge joins one vertex in L to one in R. For example, put students on one side, projects on the other, and draw an edge when a student is interested in a project.



Equivalently, vertices can be colored with two colors so adjacent vertices differ: the two color classes are the two parts. Every tree is bipartite: choose a starting vertex and alternate colors along branches. The absence of cycles prevents conflicting color requirements. An even cycle is bipartite by alternating colors. An odd cycle is not: alternating colors around it makes the closing edge join two vertices of the same color.

The **complete bipartite graph** $K_{r,s}$ contains every edge between parts of sizes r and s, with no edges within either part. Thus $n = r + s$ and $m = rs$. Vertices on the r-side have degree s; vertices on the s-side have degree r.

3 Case studies

Must two people have the same number of friends?

Every simple graph on $n \geq 2$ vertices has two vertices with the same degree. Each degree belongs to $\{0, 1, \ldots, n-1\}$. However, degrees zero and $n-1$ cannot both occur: a vertex adjacent to everyone would also be adjacent to the purported isolated vertex. Therefore at most $n-1$ degree values occur among n vertices. The pigeonhole principle gives two equal degrees.

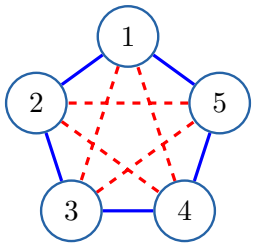
The six-person theorem

Theorem. Among any six people, there are three mutual friends or three mutual strangers. Friendship is mutual, and strangers means not friends.

Color every pair of people blue if they are friends and red if they are strangers. We must find a triangle whose three edges have the same color.

Proof. Choose a person v. Among the five edges from v, at least three have the same color, by the pigeonhole principle. Call their other endpoints a, b, c, and call this color blue (interchange the two colors if necessary). If any of ab, ac, bc is blue, that pair together with v forms a blue triangle. Otherwise all three are red, so a, b, c form a red triangle.

Why six? With five people, color the edges of a five-cycle blue and all remaining pairs red. The red edges also form a five-cycle. Neither color contains a triangle, so five people do not suffice.

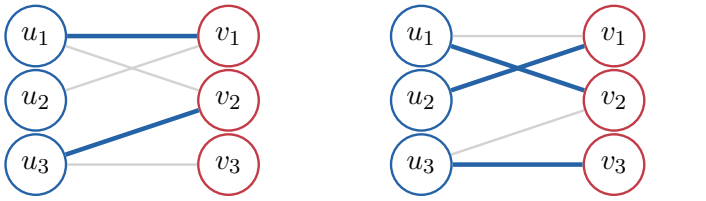


4 Bipartite matching and Hall's theorem

MUST UNDERSTAND

A **matching** is a set of edges with no shared endpoints. A **maximum matching** has the largest possible number of edges. For a bipartite graph with parts L, R, a matching **covers** L if every vertex in L is an endpoint of a matching edge. Vertices in R may remain unmatched.

Example. Suppose the edges are $u_1v_1, u_1v_2, u_2v_1, u_3v_2, u_3v_3$. The two highlighted edges on the left form a matching. The matching on the right has three edges and is maximum because $|L| = 3$.



Matching: 2 edges Maximum matching: 3 edges

Example: Pairing participants. At an event, each man and woman names their desired partners. Let L and R be the two groups, and put an edge between two participants if both name each other. Can everyone be paired with exactly one mutually desired partner? This asks for a **perfect matching**: a matching covering every vertex in $L \cup R$. When $|L| = |R|$, a matching covering L is perfect.

When does a matching cover L?

The inequality $|R| \geq |L|$ alone is not enough. If three vertices in L have only two neighbors between them, at least one of the three must remain unmatched. For $S \subseteq L$, let

$$N(S) = \{v \in R : v \text{ is adjacent to at least one vertex in } S\}.$$

This is the union of the neighbor sets of the vertices in S.

MUST UNDERSTAND

Hall's theorem. Let G be a finite bipartite graph with parts L, R. If

$$|N(S)| \geq |S| \qquad \text{for every } S \subseteq L,$$

then G has a matching covering every vertex in L. Conversely, a matching covering L pairs the vertices of each S with distinct vertices of $N(S)$, so the condition is necessary as well.

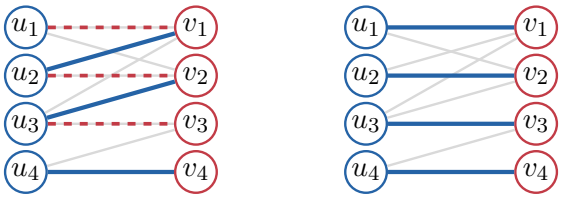
Enlarging a matching. In the lecture's four-by-four example, take

$$M = \{u_2v_1, u_3v_2, u_4v_4\}.$$

The vertices u_1 and v_3 are unmatched. Follow

$$u_1, v_1, u_2, v_2, u_3, v_3.$$

Its edges alternate between edges outside M and edges in M. Switch the choices along this path: remove u_2v_1, u_3v_2, and add u_1v_1, u_2v_2, u_3v_3. Together with u_4v_4, we now have four matching edges.



Before: 3 edges After: 4 edges

Every internal vertex loses one matching edge and gains one; each endpoint gains one. This is why the switch gives a larger matching. An alternating path with distinct vertices and unmatched endpoints is called an *augmenting path*.

Proof of Hall's theorem. Assume Hall's condition. Choose a maximum matching M and suppose $u \in L$ is unmatched. Explore all alternating paths from u: unused edges go from L to R, and matching edges go back. Skip already reached vertices, so the search terminates and each reached vertex has a simple alternating path from u.

If an unmatched vertex in R is reached, switching along its path enlarges M, a contradiction.

Otherwise, let S be the reached vertices in L. All of $N(S)$ is reached: unused edges are explored, and a reached vertex's matching edge leads back to a previously reached vertex. Every vertex in $N(S)$ is matched to a vertex in S, and every vertex in $S \setminus \{u\}$ was reached through its matching edge. Thus the matching pairs $N(S)$ with $S \setminus \{u\}$, giving

$$|N(S)| = |S| - 1 < |S|,$$

again a contradiction. Therefore M covers L.

5 Practice

1. Common neighbors. A simple graph has $n \geq 3$ vertices, each of degree greater than $n/2$. Prove that any two distinct vertices have a common neighbor.

Solution. For distinct u, v, inclusion–exclusion gives

$$|N(u) \cap N(v)| = \deg(u) + \deg(v) - |N(u) \cup N(v)| > n - n = 0.$$

Thus they have a common neighbor.

2. Apply Hall's theorem. A simple bipartite graph has parts L, R. Every vertex has degree $d \geq 1$. Prove that there is a matching covering L.

Solution. For any $S \subseteq L$, count the edges from S to $N(S)$. There are exactly $d|S|$ of them. Each vertex of $N(S)$ is incident with at most d of them, so $d|S| \leq d|N(S)|$. Divide by $d > 0$ and apply Hall's theorem.

3. Extend the six-person theorem. Prove that any simple graph on ten vertices contains three pairwise adjacent vertices or four vertices with no edges between them.

Solution. If there is a triangle, we are done. Otherwise, if a vertex has at least four neighbors, any four of them have no edges between them. If every degree is at most three, choose a vertex v and six of its nonneighbors. By the six-person theorem, three of them have no edges between them. Those three together with v give the desired four vertices.

6 Optional case studies

The core lecture ends with matching and Hall's theorem. The following two extensions can be studied independently: the friendship paradox uses expectation and double counting; the triangle-free bound uses degrees and a bipartite construction.

Case study 3: The friendship paradox

Do you feel like your friends are more popular than you? Return to the opening six-person graph. Person a has two friends, b and c, whose degrees are two and three. Their average is $5/2$, greater than a's degree two. The same holds for b, e, f: four of the six people have fewer friends than their friends do on average. In contrast, c, d each have degree three and their friends' average degree is $(2 + 2 + 3)/3 = 7/3$.

Theorem. Let $G = (V, E)$ be a finite simple undirected graph with no isolated vertices. Define

$$f(v) = \frac{1}{\deg(v)} \sum_{u \in N(v)} \deg(u).$$

If v is sampled uniformly from V, then

$$\mathbb{E}_v[f(v)] \geq \mathbb{E}_v[\deg(v)].$$

Thus the expected average degree of your friends is at least your expected degree. For our example,

$$\mathbb{E}_v[f(v)] = \frac{4(5/2) + 2(7/3)}{6} = \frac{22}{9}, \qquad \mathbb{E}_v[\deg(v)] = \frac{14}{6} = \frac{7}{3}.$$

Proof. For a uniformly sampled vertex, each outcome has probability $1/n$. By the probability-times-value definition of expectation,

$$\mathbb{E}_v[f(v)] = \sum_{v \in V} \frac{1}{n} \cdot f(v) = \frac{1}{n} \sum_{v \in V} f(v),$$
$$\mathbb{E}_v[\deg(v)] = \sum_{v \in V} \frac{1}{n} \cdot \deg(v) = \frac{1}{n} \sum_{v \in V} \deg(v).$$

It therefore suffices to prove $\sum_v f(v) \geq \sum_v \deg(v) = 2m$, where the equality is the handshaking theorem.

First enumerate each vertex v, then its neighbors u. Substituting the definition of $f(v)$ gives

$$\sum_{v \in V} f(v) = \sum_{v \in V} \frac{1}{\deg(v)} \sum_{u \in N(v)} \deg(u)$$
$$= \sum_{v \in V} \sum_{u \in N(v)} \frac{\deg(u)}{\deg(v)}.$$

The factor $1/\deg(v)$ is constant while u runs over the neighbors of v. Now change the enumeration order: consider each undirected edge $\{u, v\}$ once. It contributes two terms in the original sum, one when enumerating v and its neighbor u, and one when enumerating u and its neighbor v. Hence

$$\sum_{v \in V} f(v) = \sum_{\{u,v\} \in E} \left(\frac{\deg(u)}{\deg(v)} + \frac{\deg(v)}{\deg(u)} \right).$$

Recall that $a/b + b/a \geq 2$ for positive integers a, b. Since each edge has endpoints of positive degree,

$$\frac{\deg(u)}{\deg(v)} + \frac{\deg(v)}{\deg(u)} \geq 2.$$

Therefore $\sum_v f(v) \geq 2m = \sum_v \deg(v)$ by the handshaking theorem. Dividing by n gives $\mathbb{E}_v[f(v)] \geq \mathbb{E}_v[\deg(v)]$. Equality holds exactly when the endpoints of every edge have equal degree.

Why exclude isolated vertices? Their degree is zero, so the friends' average $f(v)$ is undefined. If we instead define $f(v) = 0$ at isolated vertices, they contribute zero to both sums and the same expectation inequality remains valid.

This is an expectation inequality; it does not claim that a majority of vertices satisfy $\deg(v) \leq f(v)$.

Further reading: Gupta and Chakrabarti, *Friends of Friends*, gives this per-person averaging argument; see also Friendship paradox.

Case study 4: How many edges without a triangle?

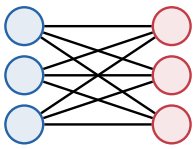
A triangle is a cycle of length three. A graph is *triangle-free* if it contains no triangle.

Problem. Among simple undirected graphs on $n \geq 1$ vertices with no triangle, what is the largest possible number of edges?

Construction. Split the vertices into parts of sizes a and $n - a$, and join every pair across the parts. The resulting complete bipartite graph $K_{a,n-a}$ is triangle-free: among any three vertices, two lie in the same part and are not adjacent. It has

$$a(n-a) = \frac{n^2}{4} - \left(a - \frac{n}{2}\right)^2$$

edges. Choosing the parts as equally as possible gives $\lfloor n^2/4 \rfloor$ edges. For $n = 1$, use the edgeless graph. A triangle-free graph need not be bipartite: C_5 is a counterexample. We therefore need an upper bound that applies to *all* triangle-free graphs.



$K_{3,3}$: nine edges, no triangle

Mantel's theorem. Every triangle-free simple undirected graph on $n \geq 1$ vertices has at most $\lfloor n^2/4 \rfloor$ edges. The balanced complete bipartite construction attains this bound.

Finding the proof. The bipartite construction suggests looking for a set of vertices with no edges inside it. In a triangle-free graph, the neighbors of any vertex form such a set: an edge between two neighbors would complete a triangle. We first use any vertex to obtain an upper bound, then choose the vertex to make that bound useful.

Proof. Pick any vertex v, and set

$$A = N(v), \qquad B = V \setminus A.$$

Then $|A| = \deg(v)$, $|B| = n - \deg(v)$, and there are no edges inside A. Notice that $v \in B$, because a vertex is not its own neighbor. Every edge therefore has at least one endpoint in B, so

$$m \leq \sum_{u \in B} \deg(u).$$

An edge between A and B is counted once; an edge inside B is counted twice. Such double counting is harmless for this upper bound. We do not assume that B has no internal edges.

How should we choose v? Let Δ be the maximum degree of the graph. For any choice of v,

$$\sum_{u \in B} \deg(u) \leq \sum_{u \in B} \Delta = |B|\Delta = (n - \deg(v))\Delta.$$

Now choose v with maximum degree, so $\deg(v) = \Delta$. We obtain

$$m \leq \sum_{u \in B} \deg(u) \leq (n - \Delta)\Delta.$$

To finish, use a simple product bound: if $a + b = n$, then

$$ab = a(n - a) = \frac{n^2}{4} - \left(a - \frac{n}{2}\right)^2 \leq \frac{n^2}{4},$$

with equality exactly when $a = b = n/2$. Taking $a = n - \Delta$ and $b = \Delta$ gives $m \leq n^2/4$. Since m is an integer, $m \leq \lfloor n^2/4 \rfloor$. This also covers the edgeless case $\Delta = 0$.

Optional exercise. A triangle-free graph has six vertices, each of degree three. Prove that it is $K_{3,3}$.

Solution. Choose a vertex v, and set $A = N(v)$ and $B = V \setminus A$. Each part has three vertices. No edges lie inside A, so each vertex of A, having degree three, is adjacent to all of B. Each vertex of B now has its three neighbors in A and has no additional edges. Thus the graph is $K_{3,3}$.

Reference: Rosen, *Discrete Mathematics and Its Applications*, 8th ed., Sections 10.1–10.2 (Hall's theorem: Section 10.2, Theorem 5); the six-person theorem is Example 13 in Section 6.2. Our proof of Hall's theorem uses alternating paths rather than Rosen's induction proof. Mantel's theorem and the friendship paradox are optional instructor-selected extensions. Representations, connectivity, Euler/Hamilton routes, and coloring continue in L19.