Graph 2

Mingxun Zhou (mingxunz@ust.hk)

COMP2711 Discrete Mathematical Tools

1 Graph representations

How do we store a graph in a computer? A picture shows its shape, but graph algorithms need access to vertices and their neighbors. Two common representations are adjacency lists and adjacency matrices. Write $G = (V, E)$, $n = |V|$, and $m = |E|$. Unless stated otherwise, graphs are simple and undirected.

For the graph with vertices a, b, c, d and edges ab, ac, bc, cd, the adjacency lists and adjacency matrix (in that vertex order) are

$$\begin{array}{c|l} a & b, c \\ b & a, c \\ c & a, b, d \\ d & c \end{array} \qquad A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

MUST UNDERSTAND

An **adjacency list** stores the neighbors of each vertex. An **adjacency matrix** has $A_{ij} = 1$ exactly when v_i and v_j are adjacent. For a simple undirected graph, the matrix is symmetric, its diagonal is zero, and each row sum is the corresponding degree.

Operation	Unsorted adjacency lists	Adjacency matrix
Storage	$\Theta(n + m)$	$\Theta(n^2)$
Test adjacency of u, v	$O(\deg(u))$	$O(1)$
Scan neighbors of u	$\Theta(\deg(u))$	$\Theta(n)$

Here scan time counts neighbor entries; there is also constant overhead for an empty list.

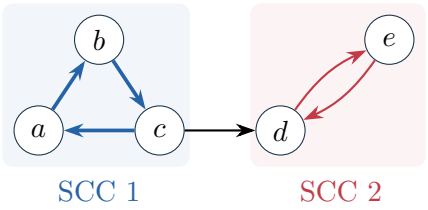
2 Paths and connectivity

A **path** starts at a vertex and moves to a neighbor at each step. Vertices and edges may repeat unless we impose an additional restriction. Its length is the number of steps (edges). For example, take edges ab, bc, cd, de, eb. The path a, b, c, d, e starts at a, ends at e, and has length four. It does not use edge eb. In a directed graph, each step must follow the direction of its edge. For parallel edges, specify which edge is used at each step.

Vertex v is **reachable from** u if there is a path starting at u and ending at v. For example, with edges ab, bc, de, the path a, b, c makes c reachable from a, but d is not reachable from a. In an undirected graph, reversing a path shows that reachability is mutual; in a directed graph it need not be.

An undirected graph is **connected** if any two vertices are reachable from each other (i.e., a path connects them). A connected component is a maximal connected subgraph. For example, the graph with edges ab, bc, ca, cd, de is connected. Removing cd disconnects it into the two components $\{a, b, c\}$ and $\{d, e\}$.

In a directed graph, a **strongly connected component** (SCC) is a maximal group of vertices where every vertex is reachable from every other. For example, take directed edges $a \to b$, $b \to c$, $c \to a$, $c \to d$, $d \to e$, and $e \to d$. The SCCs are $\{a, b, c\}$ and $\{d, e\}$. The edge $c \to d$ lets us move from the first group to the second, but there is no path back, so the two groups do not form one SCC.



Euler and Hamilton paths

An **Euler path** covers every edge exactly once. A **Hamilton path** covers every vertex exactly once. In the graph with edges ab, bc, cd, da, ac, the sequence a, b, c, d, a, c is an Euler path: its five steps use the five edges once. In the same graph, a, b, c, d is a Hamilton path: it visits all four vertices once, leaving edges da and ac unused.

3 Euler paths

MUST UNDERSTAND

An **Euler path** uses every edge exactly once. Vertices may repeat. An **closed Euler path** additionally ends where it starts.

The seven bridges of Königsberg give a multigraph: four land regions are vertices, bridges are edges, and two pairs of regions have parallel edges. Its degrees are $5, 3, 3, 3$.

Euler path existence. A connected undirected graph has an Euler path if and only if it has zero or two odd-degree vertices. Parallel edges are allowed. More precisely:

- It has a closed Euler path if and only if every degree is even.
- It has an open Euler path if and only if exactly two vertices have odd degree. These are its endpoints.

Isolated vertices do not affect an edge-covering route. Two disjoint triangles show why even degrees alone are insufficient.

Exercise: Why is the number of odd-degree vertices even? The handshaking theorem gives $\sum_v \deg(v) = 2m$. The even degrees contribute an even sum. Therefore the odd degrees also have an even sum, which requires an even number of them.

All-even case: strong induction on m. We prove that every connected undirected graph with all degrees even has a closed Euler path. For $m = 0$, the graph consists of one vertex; the length-zero path suffices. Assume the claim holds for every such graph with fewer than m edges, and let $m > 0$.

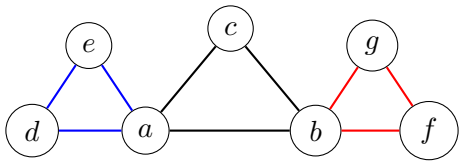
Start at a vertex a and follow unused edges, stopping when we first return to a. We cannot get stuck at another vertex: previous arrivals and departures used pairs of edges, and the latest arrival uses one more; even total degree leaves an unused departure. Since there are only finitely many edges, we must return to a. Call this nonempty closed path P. No edge of P is repeated.

Remove the edges of P, retaining all vertices. At every vertex, P uses an even number of incident edge-ends, so the remaining degrees are even. Each remaining connected component has fewer than m edges. The induction hypothesis gives a closed Euler path in each component, including the trivial path for an isolated vertex.

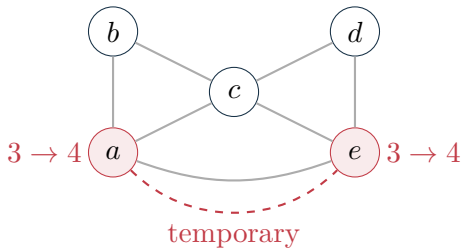
Every component containing edges shares a vertex with P. To see this, take a path in the original connected graph from that component to P. Before its first meeting with P, none of its edges was removed, so that first meeting vertex is still in the same remaining component. Start the component's closed Euler path at this shared vertex, and insert it when P visits that vertex. Repeat for every component containing edges. The paths use disjoint edge sets, so the result is a closed path using every original edge exactly once. This completes the induction.

Example. Take three triangles with edges ab, bc, ca, ad, de, ea, and bf, fg, gb. Choose $P = a, b, c, a$. Removing its edges leaves triangles ade and bfg, and the isolated vertex c. Insert a, d, e, a at a and b, f, g, b at b to obtain

$$a, d, e, a, b, f, g, b, c, a.$$



Two odd vertices. If exactly a, e have odd degree, add a temporary edge ae, allowing a parallel edge if necessary. All degrees become even. Find a closed Euler path; it can start anywhere. Delete the temporary edge, and read the closed path starting just after the cut. This gives an Euler path between a and e.



For example, start at c in the augmented graph:

$$c, d, e \underbrace{\longrightarrow}_{\text{temporary edge}} a, b, c, a, e, c.$$

After cutting, start at a and move the initial segment c, d, e to the end:

$$a, b, c, a, e, c, d, e.$$

This uses every original edge exactly once. In particular, the original edge ae is still used; only the added copy is removed.

Necessity. At each internal visit, an arrival edge pairs with a departure edge. A closed path pairs every incident edge at every vertex. In an open route, only the start and finish have one unpaired edge. Thus there are zero or two odd vertices, respectively. Königsberg has four odd vertices, so neither kind of Euler route exists.

4 Hamilton paths and cycles

MUST UNDERSTAND

A **Hamilton path** visits every vertex exactly once. A **cycle** in a simple graph returns to its start with no other repeated vertex and has length at least three. A **Hamilton cycle** is a cycle visiting every vertex. Edges outside the route need not be used.

In the graph with edges ab, bc, cd, da, ac, the sequence a, b, c, d, a is a Hamilton cycle: it visits all four vertices and returns to the start. The diagonal ac is unused. In contrast, the Euler path a, b, c, d, a, c uses all five edges, repeating vertices.

Degree parity characterizes Euler routes but does not characterize Hamilton routes. A proposed Hamilton cycle is easy to check: verify that every vertex appears exactly once and that consecutive vertices, including the last and first, are adjacent. Deciding whether one exists is NP-complete. This connects the lecture to the earlier distinction between checking and finding.

5 Graph coloring

Suppose exams sharing students cannot occur simultaneously. Make each exam a vertex and join conflicting exams by an edge. A time slot is a color. In the lecture example, Math, CS, and Physics conflict pairwise, so at least three slots are necessary. The assignment below uses three and is therefore optimal:

Slot	Exams
1	Math, History
2	CS, English
3	Physics

The remaining conflicts are CS–History, Physics–English, and History–English; all are separated by this assignment.

MUST UNDERSTAND

A **proper coloring** assigns a color to each vertex so that adjacent vertices have different colors. The **chromatic number** $\chi(G)$ is the fewest colors in a proper coloring. To prove $\chi(G) = k$, show both that k colors suffice (construct a coloring) and that fewer cannot suffice (prove a lower bound).

Examples. For $n \geq 1$, $\chi(K_n) = n$, since every pair is adjacent. A bipartite graph with at least one edge has chromatic number 2: color its two parts differently, and the edge forces two colors. Conversely, a proper two-coloring partitions the vertices into two parts with no edges inside either part. A nonempty edgeless graph needs one color.

For $n \geq 3$,

$$\chi(C_n) = \begin{cases} 2 & n \text{ even}, \\ 3 & n \text{ odd}. \end{cases}$$

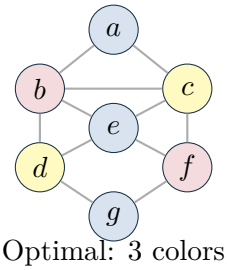
Two colors must alternate along a cycle. They close consistently exactly when the length is even. For an odd cycle, color the first $n - 1$ vertices alternately and use a third color on the last.

Greedy coloring

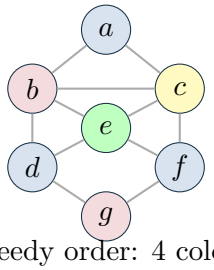
The lecture uses the following order: start anywhere, then repeatedly choose an uncolored vertex adjacent to a colored vertex. Once a component is finished, start another. Assign each chosen vertex the smallest positive color absent from its already colored neighbors.

The coloring is proper: for each edge, its later-colored endpoint avoids the color of its earlier endpoint.

Example. In the graph below, processing a, b, c, d, e, f, g gives colors $1, 2, 3, 1, 4, 1, 2$. When e is processed, its colored neighbors b, c, d already use colors $2, 3, 1$, forcing color 4. Yet the graph needs only three colors: the left drawing supplies a 3-coloring, and triangle abc forces at least three.



Optimal: 3 colors



Greedy order: 4 colors

Theorem. For a nonempty simple graph with maximum degree Δ,

$$\chi(G) \leq \Delta + 1.$$

At most $\deg(v) \leq \Delta$ colors are forbidden when a vertex v is colored. Thus one of $1, \ldots, \Delta + 1$ is available. This bound holds for any vertex order, even though the number of colors used can depend on the order.

Two colors versus three

Exercise. Why does the lecture's greedy order use at most two colors whenever the graph is bipartite?

Proof. Consider one connected component with parts L, R. Give its first vertex color 1 and suppose it lies in L. Inductively, all colored vertices in L have color 1 and those in R have color 2. Every later vertex has a colored neighbor. If it lies in R, all its colored neighbors have color 1, so it receives color 2; if it lies in L, all its colored neighbors have color 2, so it receives color 1. The same argument applies independently to each component.

The choice of order matters. For the path $a - b - c - d$, the arbitrary order a, d, b, c gives colors $1, 1, 2, 3$, even though the graph is bipartite. This order does not follow the lecture's rule: after a, the vertex d has no colored neighbor.

To decide two-colorability, explore each component with a queue or stack. Give an uncolored neighbor the opposite color; if an edge has two equal-colored endpoints, reject. The proof above shows that a bipartite graph cannot cause rejection. With adjacency lists, each vertex and each edge is processed a constant number of times, giving $\Theta(n + m)$ worst-case time.

Deciding whether a graph is 3-colorable is NP-complete. A proposed 3-coloring is easy to check by inspecting every edge, but no polynomial-time algorithm for finding one on arbitrary graphs is known. Computing $\chi(G)$ is an optimization problem; the NP-complete statement concerns the decision question "Is $\chi(G) \leq 3$?"

Why odd cycles obstruct two-coloring

Theorem. A graph is bipartite if and only if it has no odd cycle. If it is bipartite, a cycle alternates between the two parts, so its length is even. Conversely, work in one connected component, choose a root, and form a tree of shortest paths from the root. Color vertices by the parity of their distance from the root. Suppose an edge uv joined equal-parity levels. Adjacent distances differ by at most one, so they must be equal. The two tree paths from their last common vertex to u, v, together with uv, form a cycle of odd length: the two tree paths have equal length, followed by one extra edge. This contradicts the hypothesis. Thus every edge joins different colors. Apply this construction independently in each component.

6 Practice

1. For which $n \geq 3$ does K_n have a closed Euler path? A Hamilton cycle? What is its chromatic number?
2. A connected graph has exactly two odd-degree vertices. Adding an edge between them makes a closed Euler path possible. Why may a parallel edge be necessary?
3. Give a triangle-free graph that needs three colors. Explain why this does not contradict the bipartite criterion.

Solutions. (1) Every degree in K_n is $n-1$, so it has a closed Euler path exactly when n is odd. It always has a Hamilton cycle for $n \geq 3$, obtained by listing all vertices and returning. Its chromatic number is n. (2) The odd vertices may already be adjacent; the temporary edge must be a distinct additional edge, not reuse an existing one. (3) C_5 is triangle-free and needs three colors. It contains an odd cycle of length five.

Further topics

These algorithms are pointers to later courses, not required material.

Problem	Algorithms or methods
Minimum spanning trees	Prim; Kruskal
Shortest paths	Dijkstra; Bellman–Ford
Bipartite matching	Hopcroft–Karp
Network flow	Ford–Fulkerson; Edmonds–Karp
Graph traversal	BFS; DFS
Topological sorting	Kahn; DFS
Strongly connected components	Tarjan; Kosaraju
Graph isomorphism	VF2 (backtracking)
PageRank	Power iteration

Optional reading

Hamilton sufficient conditions. For a simple graph on $n \geq 3$ vertices, Dirac's theorem guarantees a Hamilton cycle if every degree is at least $n/2$. Ore's theorem requires $\deg(u) + \deg(v) \geq n$ for every nonadjacent pair. Neither condition is necessary: C_5 satisfies neither but is Hamiltonian.

Chinese postman. Find a shortest closed route covering every edge, allowing repetition. With nonnegative edge weights and exactly two odd vertices, duplicate a shortest path between them and find a closed Euler path. With more odd vertices, choosing how to pair them is another optimization problem.

Four colors. A planar graph can be drawn without edge crossings except at common endpoints. Every simple planar graph admits a coloring with at most four colors (proof omitted).

Reading. Rosen, *Discrete Mathematics and Its Applications*, 8th ed., Sections 10.3 (representations only), 10.4 (connectivity), 10.5 (Euler and Hamilton routes), and 10.8 (coloring).