ĐỒ THỊ - GRAPH

DATA STRUCTURES & ALGORITHMS

ThS Nguyễn Thị Ngọc Diễm
diemntn@uit.edu.vn

NỘI DUNG

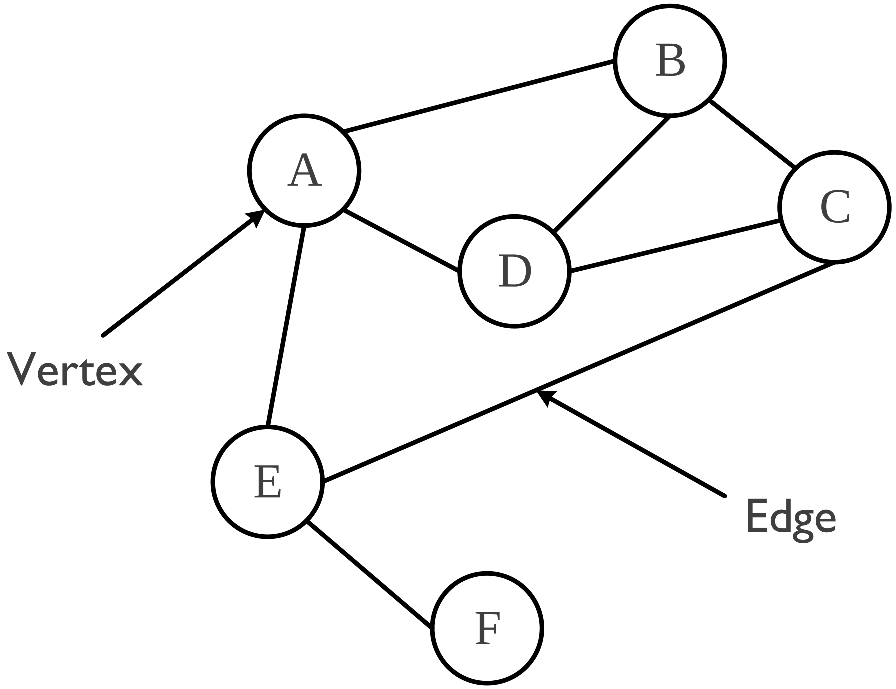
- Các khái niệm trên đồ thị
  - Định nghĩa
  - Các loại đồ thị
  - Đường đi, chu trình, liên thông
- Biểu diễn đồ thị trên máy tính
- Các thuật toán duyệt đồ thị: BFS - DFS
- Một số ứng dụng của tìm kiếm trên đồ thị
  - Bài toán đường đi
  - Bài toán liên thông
  - Bài toán tô màu
  - Bài toán bao đóng

NỘI DUNG

- **Các khái niệm trên đồ thị**
  - Định nghĩa
  - Các loại đồ thị
  - Đường đi, chu trình, liên thông
- Biểu diễn đồ thị trên máy tính
- Các thuật toán duyệt đồ thị: BFS - DFS
- Một số ứng dụng của tìm kiếm trên đồ thị
  - Bài toán đường đi
  - Bài toán liên thông
  - Bài toán tô màu
  - Bài toán bao đóng

ĐỊNH NGHĨA

- Đồ thị (Graph) bao gồm:
  - Tập đỉnh (Vertices, also called nodes or points)
  - Tập cạnh (Edges, also called link or line)



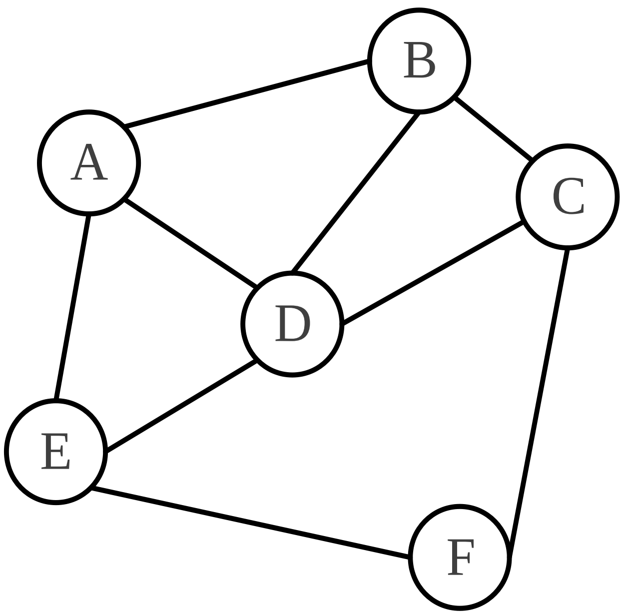
Vertices can be considered "sites" or locations.

Edges represent connections.

Graphs are the basic subject studied by graph theory. The word "graph" was first used in this sense by James Joseph Sylvester in 1878.

ĐỒ THỊ VÔ HƯỚNG

- Một đồ thị vô hướng (undirected graph) $\mathrm{G} = (\mathrm{V}, E)$ được định nghĩa bởi:
  - Tập hợp V được gọi là tập các đỉnh của đồ thị
  - Tập hợp E là tập các cạnh của đồ thị
  - Mỗi cạnh $e \in E$ được liên kết với một cặp đỉnh $\{i, j\} \in \mathrm{V}^2$, không phân biệt thứ tự (unordered pairs).



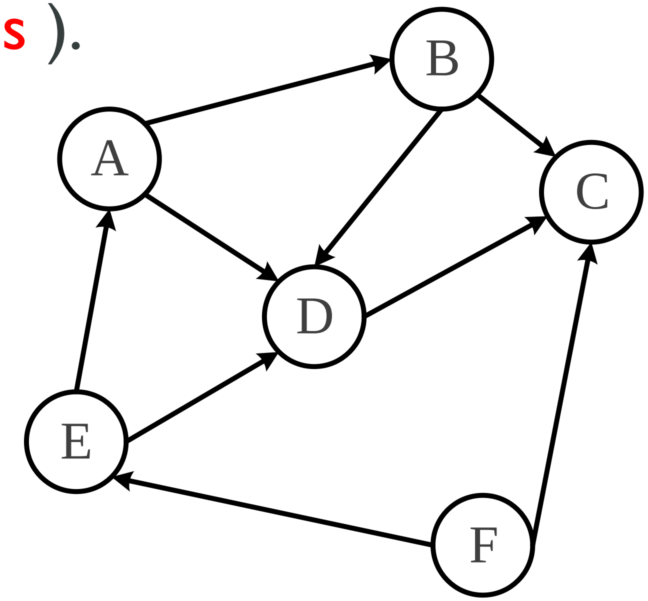
Ví dụ:

V = {A, B, C, D, E, F, G}

E = {(A, B), (A, D), (A, E),
(B, C), (B, D)
(C, D), (C, F),
(E, D), (E, F)}

ĐỒ THỊ CÓ HƯỚNG

- Một đồ thị có hướng (directed graph or digraph) $G = (V, U)$ được định nghĩa bởi:
  - Tập hợp V được gọi là tập các đỉnh của đồ thị
  - Tập hợp U là tập các cạnh của đồ thị;
  - Mỗi cạnh $u \in U$ được liên kết với một cặp đỉnh $(i, j) \in V^2$, có phân biệt thứ tự và không trùng nhau(ordered pairs of distinct vertices).



Ví dụ:

V = {A, B, C, D, E, F}

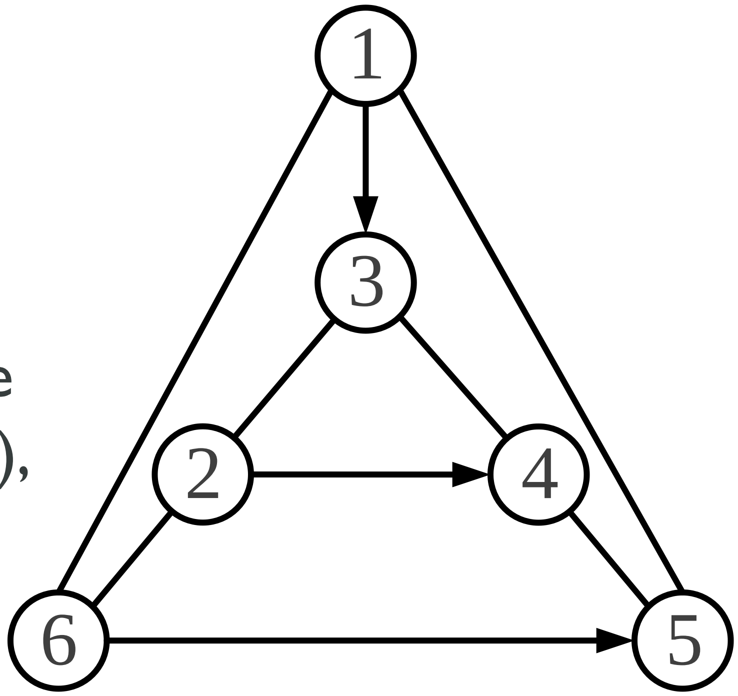
U = {(A, B), (A, D),
(B, C), (B, D)
(D, C),
(E, A), (E, D),
(F, C) (F, E)}

ĐỒ THỊ HỖN HỢP

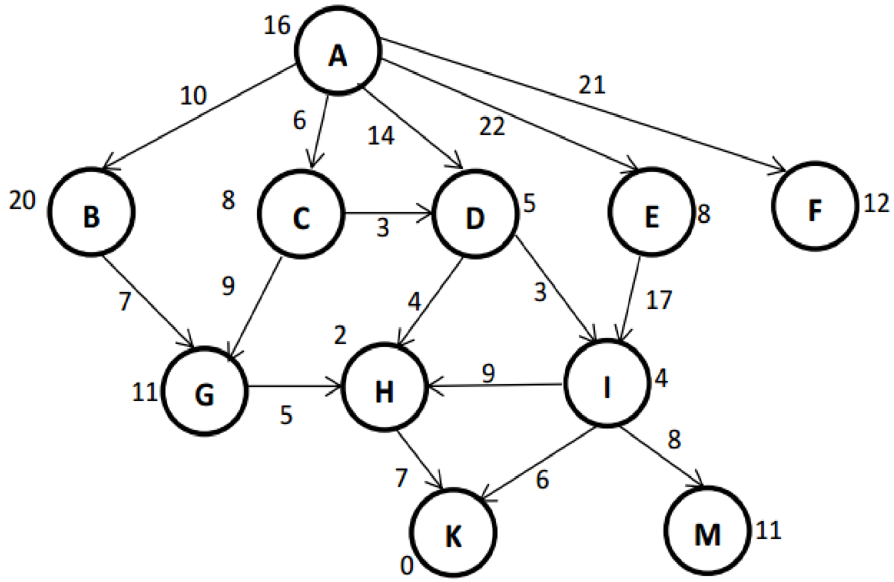
- A mixed graph (Mixed graph) is a graph in which some edges may be directed and some may be undirected.

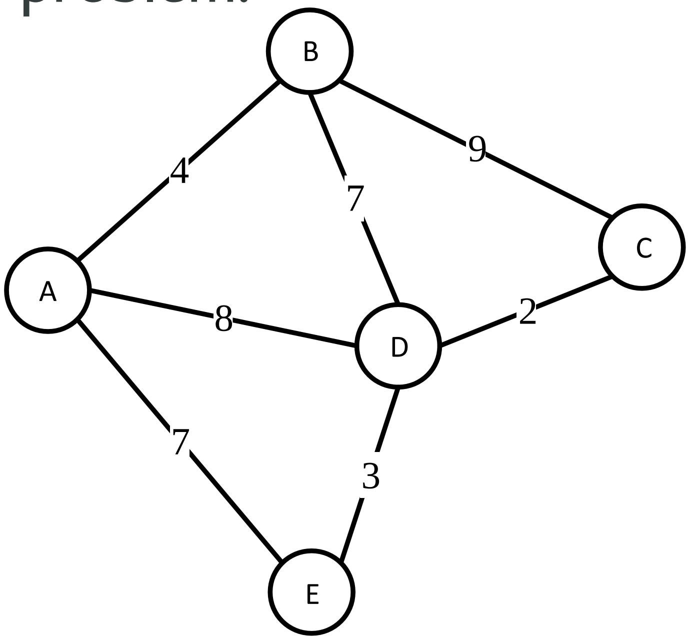
- It is an ordered triple G = (V, E, A) for a mixed simple graph and G = (V, E, A, φE, φA) for a mixed multigraph with V, E (the undirected edges), A (the directed edges), φE and φA defined as above.

- Directed and undirected graphs are special cases of Mixed graph.



- A *weighted graph* is a graph in which a number (the weight) is assigned to each edge.
- Such weights might represent for example costs, lengths or capacities, depending on the problem at hand.
- Such graphs arise in many contexts, for example in shortest path problems such as the traveling salesman problem.

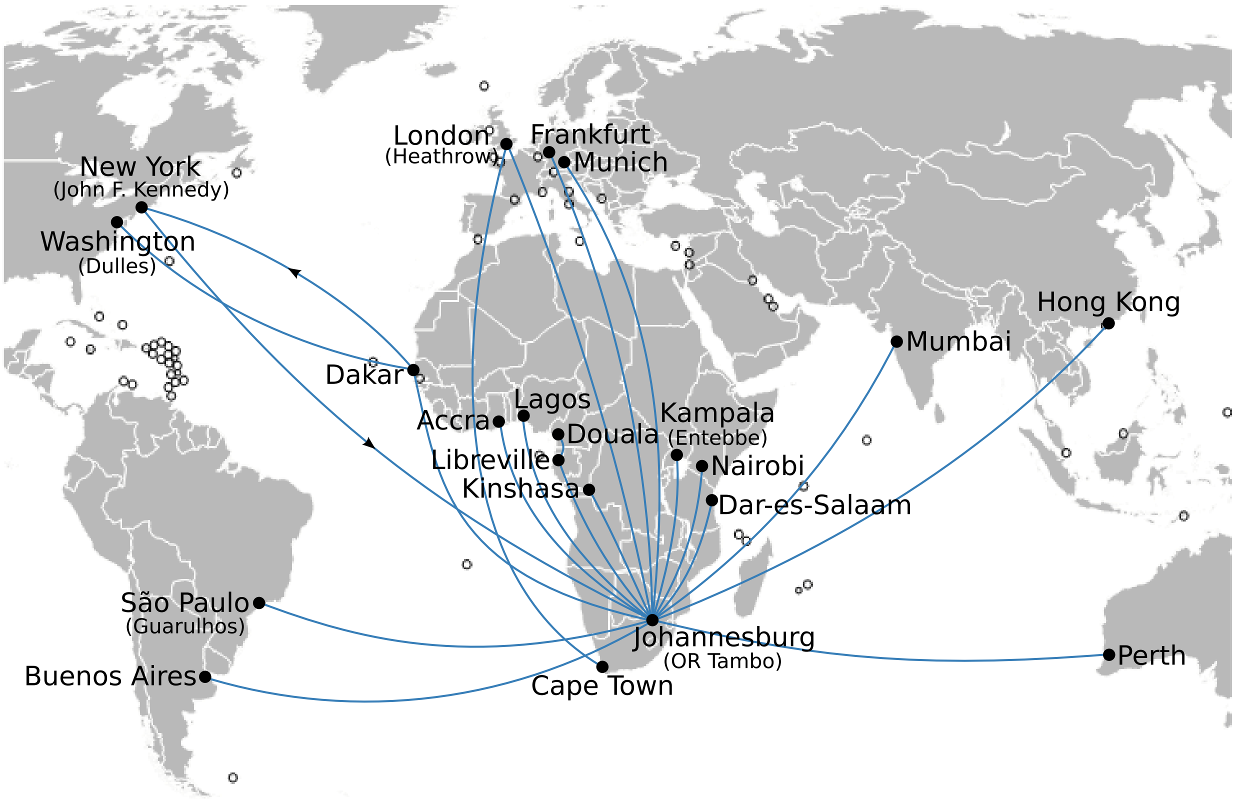




Application: Air flight system

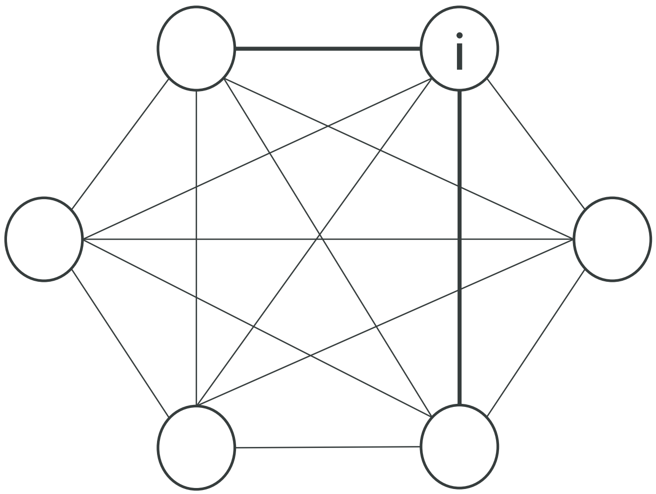
Air flight system:

- Mỗi đỉnh biểu diễn cho một thành phố.
- Mỗi cạnh biểu diễn cho đường bay trực tiếp giữa hai thành phố.
- Một truy vấn trên đường bay trở thành một truy vấn trên cạnh.
- Một truy vấn về cách di chuyển từ thành phố A đến thành phố B trở thành truy vấn "có tồn tại đường đi từ đỉnh A đến đỉnh B".



- Có thể thêm vào chi phí di chuyển cho các cạnh (biểu đồ có trọng số- **weighted graphs**), sau đó truy vấn đường bay nào là ngắn nhất từ A đến B.

Application: Wireless communication



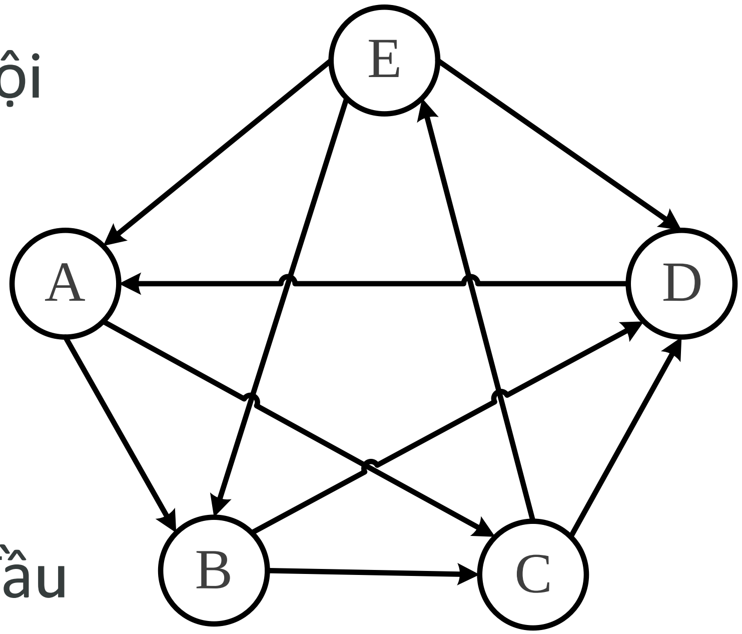
Wireless communication:

- Can be represented by a weighted complete graph
- Each edge represents the Euclidean distance dij between two stations.

- Each station uses a certain power i to transmit messages. Given this power i, only a few nodes can be reached (bold edges). A station reachable by i then use its own power to relay the message to other stations not reachable by i.
- A typical wireless communication problem is: how to broadcast between all stations such that they are all connected and the power consumption is minimized.

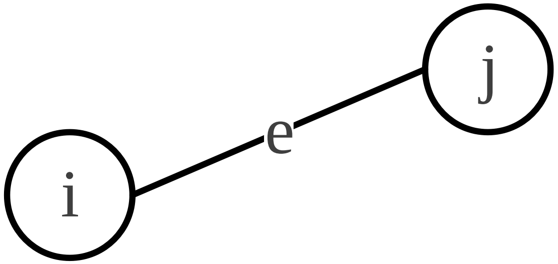
Application: Mô hình thi đấu vòng tròn

- Mô hình đồ thị thi đấu vòng tròn: Là mỗi đội đều thi đấu với tất cả các đội còn lại.
- Giả sử có 5 đội A, B, C, D, E thi đấu theo vòng tròn. Ta có thể biểu diễn kết quả thi đấu như hình bên. Trong đó: Đội ở gốc mũi tên thắng đội ở đầu mũi tên.



ĐỈNH KỀ

- Trên đồ thị vô hướng (undirected graph), xét cạnh e được liên kết với cặp đỉnh (i, j):

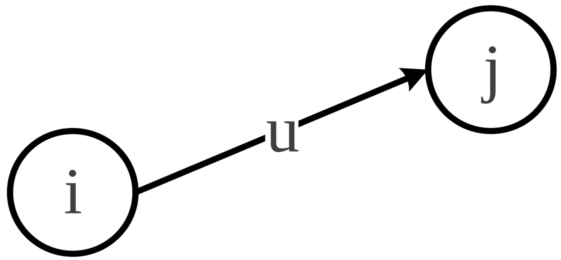


- Cạnh e **kề** (adjacent to) với đỉnh i và đỉnh j (hay đỉnh i và đỉnh j **kề** với cạnh e); có thể viết tắt $e = (i, j)$.

- Đỉnh i và đỉnh j được gọi là 2 đỉnh kề nhau (hay đỉnh i kề với đỉnh j và ngược lại, đỉnh j kề với đỉnh i)

ĐỈNH KỀ

- Trên đồ thị có hướng, xét cạnh u được liên kết với cặp đỉnh (i, j):

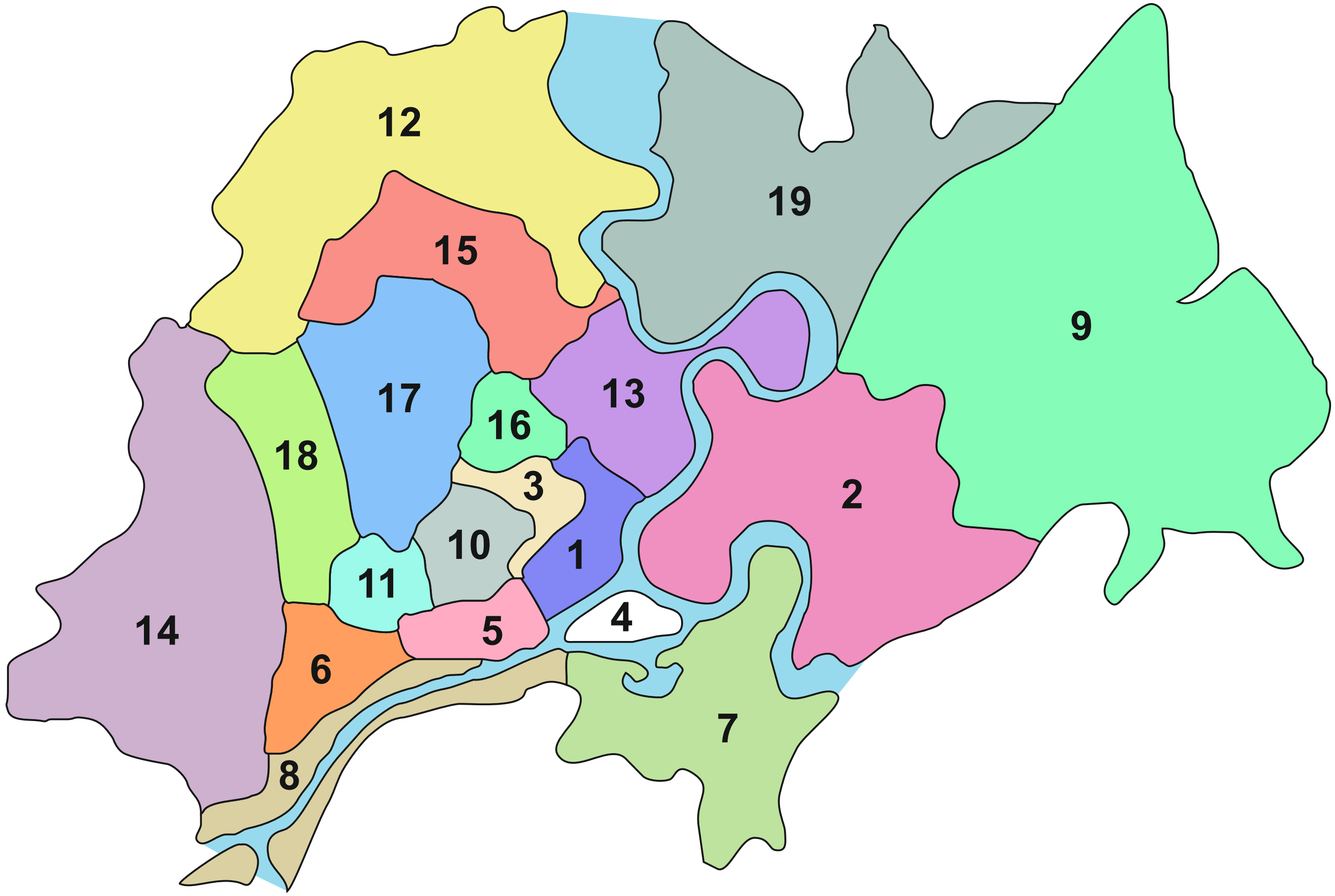


- Cạnh u **kề** (adjacent) với đỉnh i và đỉnh j (hay đỉnh i và đỉnh j **kề** với cạnh u); có thể viết tắt $u = (i, j)$. Cạnh u đi ra khỏi đỉnh i và đi vào đỉnh j

- Đỉnh j được gọi là đỉnh kề (adjacency vertex) của đỉnh i.

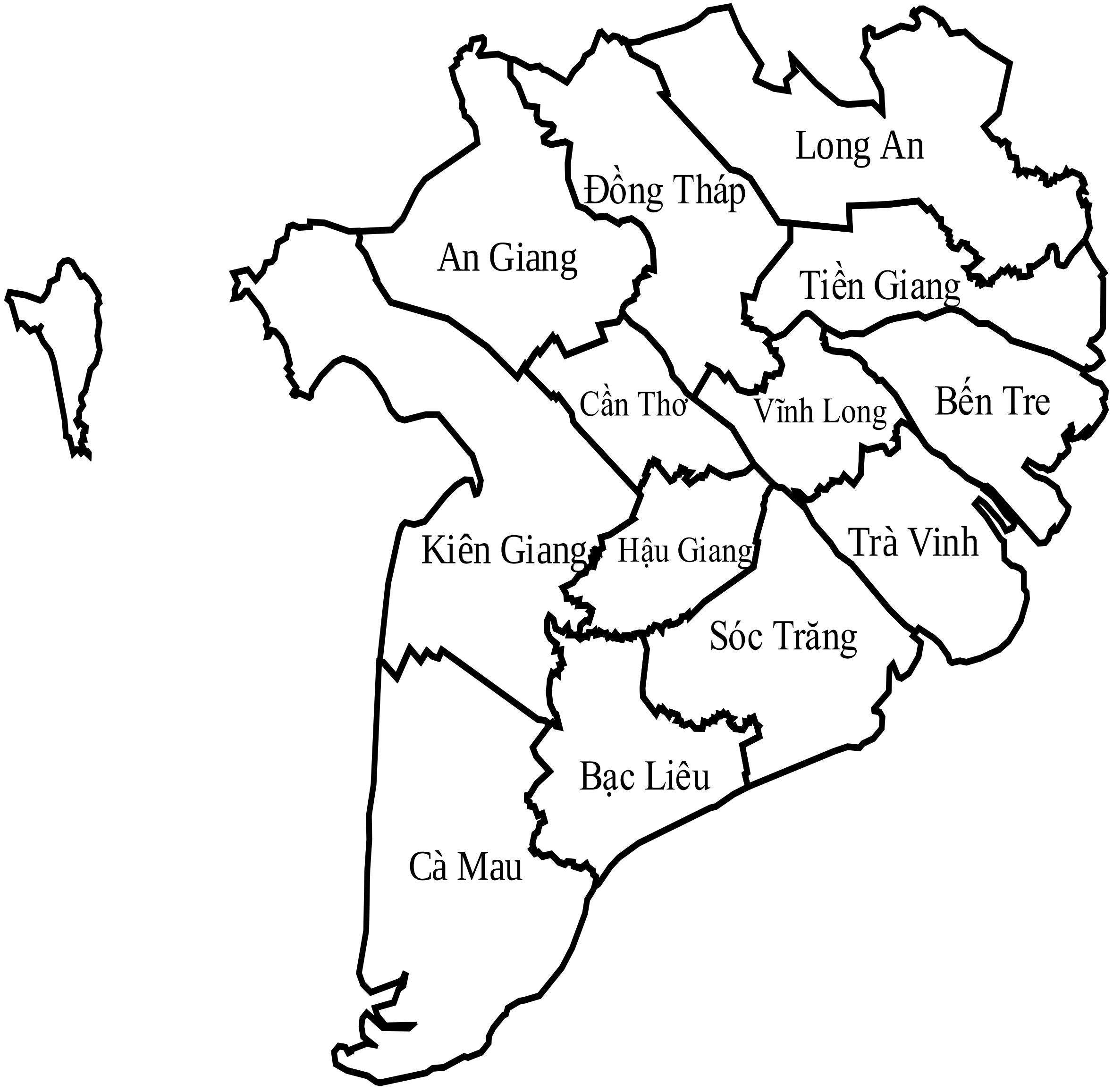
Bài tập

Bài 1: Dựa vào bản đồ bên cạnh, hãy vẽ đồ thị vô hướng biểu diễn đường đi giữa các quận trong thành phố Hồ Chí Minh.

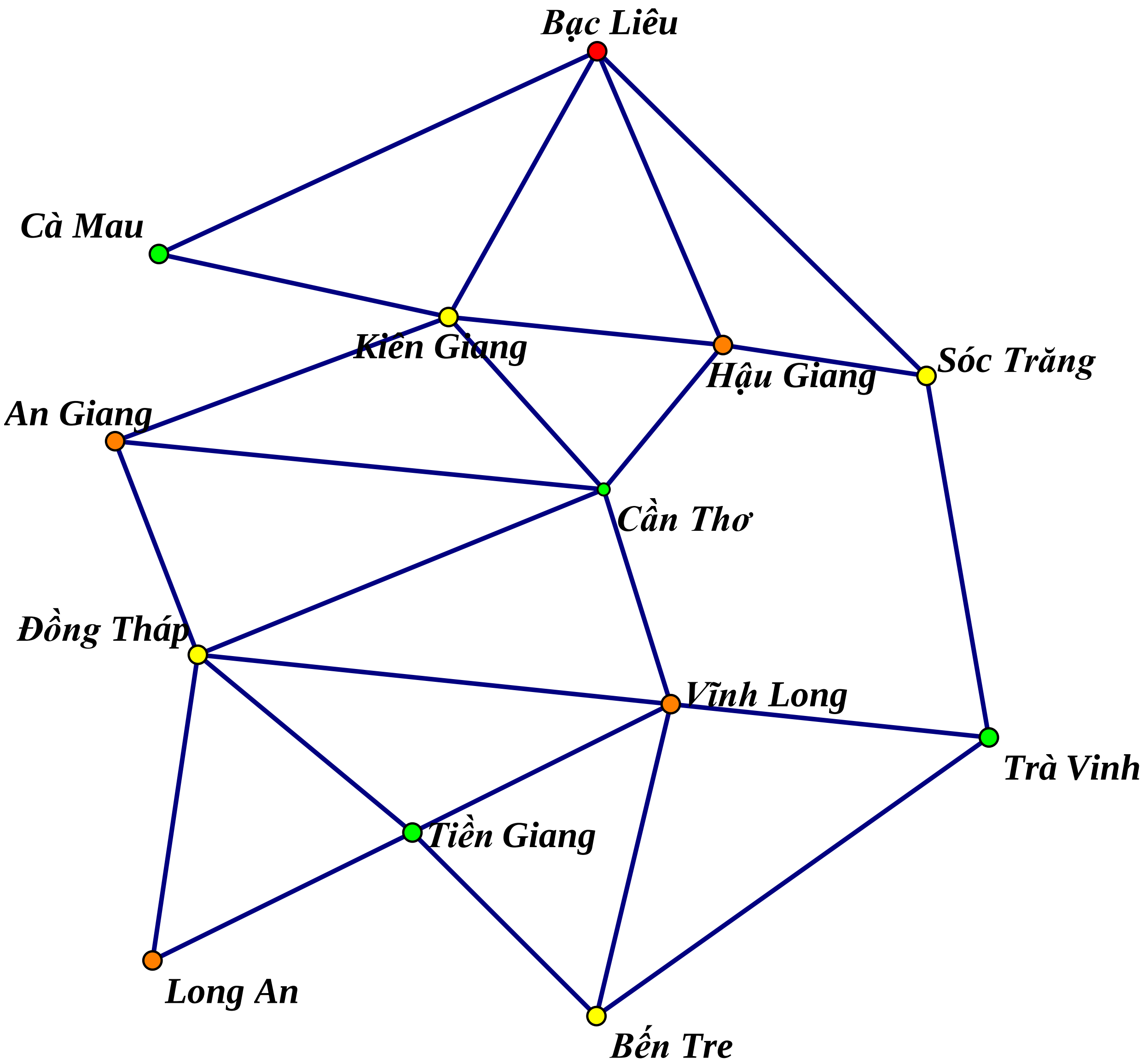


Administrative divisions of HCMC's urban districts.
1-12: Quận 1-12, 13: Bình Thạnh, 14: Bình Tân, 15: Gò Vấp, 16: Phú Nhuận, 17: Tân Bình, 18: Tân Phú, 19: Thủ Đức.

Bài tập



Bài 2: Dựa vào bản đồ bên cạnh, hãy vẽ đồ thị vô hướng biểu diễn đường đi giữa các tỉnh Đồng bằng sông Cửu Long.


Bạc Liêu
Cà Mau
Kiên Giang
Hậu Giang
Sóc Trăng
An Giang
Cần Thơ
Đồng Tháp
Vĩnh Long
Trà Vinh
Tiền Giang
Long An
Bến Tre

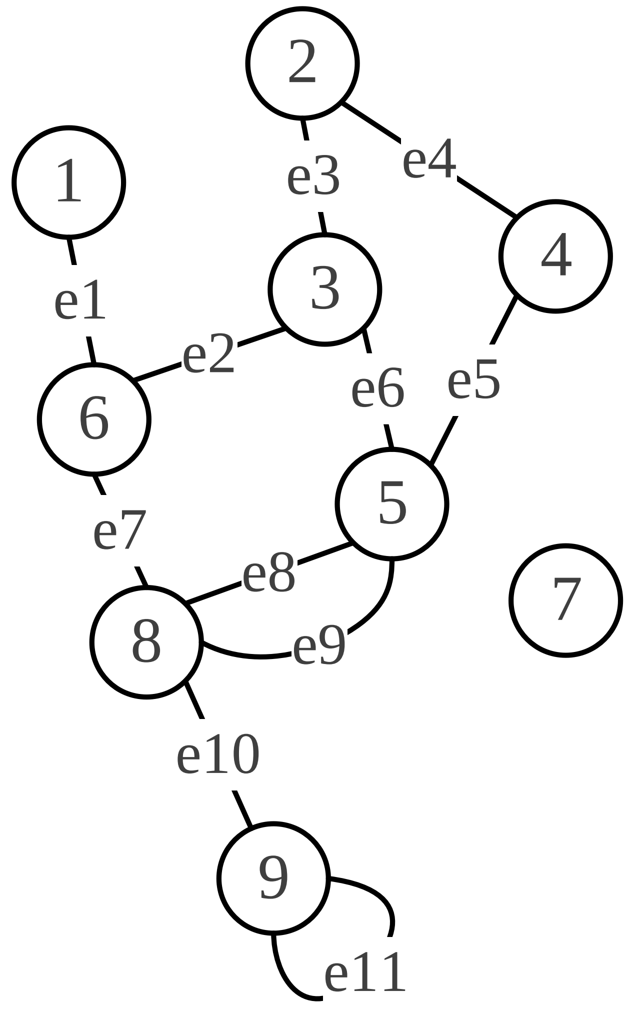
Bài tập

Bài 3: Hãy vẽ đồ thị có trọng số biểu diễn đường đi và khoảng cách giữa 6 thành phố: London, Cambridge, Oxford, Birmingham, Nottingham và Exeter. Khoảng cách tính bằng km giữa các thành phố được cho trong bảng bên dưới.

	L	C	O	B	N	E
London (L)	-	80	56	120	131	200
Cambridge (C)	80	-	100	98	87	250
Oxford (O)	56	100	-	68	103	154
Birmingham (B)	120	98	68	-	54	161
Nottingham (N)	131	87	103	54	-	209
Exeter (E)	200	250	154	161	209	-

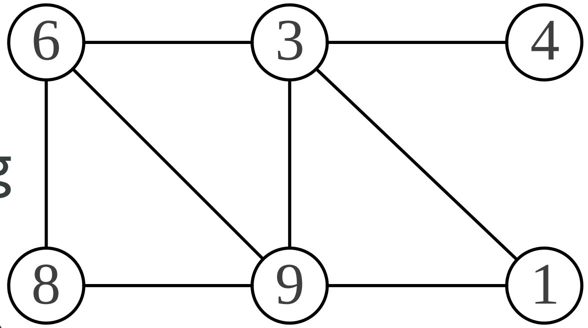
KHUYÊN, ĐỈNH TREO, ĐỈNH CÔ LẬP

- Khuyên (loop): Cạnh có hai đầu trùng nhau (cùng một đỉnh).
- Một *đỉnh cô lập* (isolated vertex) hoặc *đỉnh độc lập* là đỉnh không liên thuộc với một cạnh nào, cũng có nghĩa là đỉnh có bậc 0.
- Đỉnh có bậc 1 là một đỉnh treo hay lá (leaf vertex, pendant vertex).
- Cạnh song song (parallel edge): tồn tại nhiều hơn 1 cạnh giữa 2 đỉnh.
- Ví dụ (hình bên):
  - Khuyên: Đỉnh 9
  - Đỉnh treo: Đỉnh 1
  - Đỉnh cô lập: Đỉnh 7
  - Cạnh song song: e8, e9

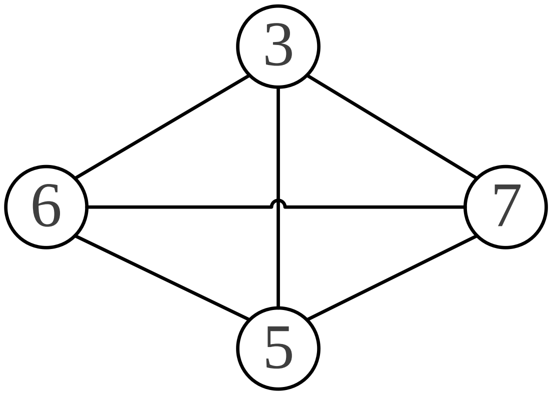


CÁC DẠNG ĐỒ THỊ

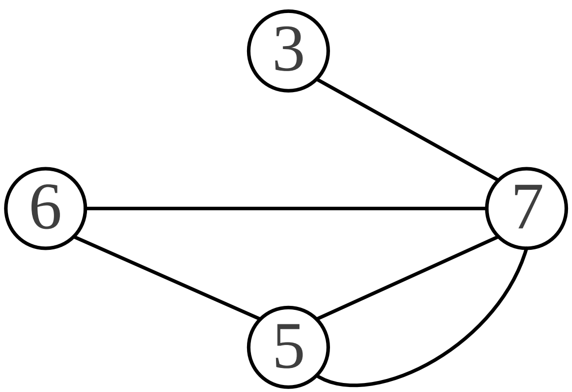
- Đồ thị RỖNG: tập cạnh là tập rỗng
- Đồ thị ĐƠN: không có khuyên và cạnh song song (đa cạnh)



- Đồ thị ĐẦY ĐỦ (complete graph): đồ thị đơn, vô hướng, giữa hai đỉnh bất kỳ đều có đúng một cạnh.
  - Đồ thị đủ N đỉnh ký hiệu là K_N.
  - K_N có $N(N-1)/2$ cạnh.



- ĐA ĐỒ THỊ (multigraph hay pseudograph) là một đồ thị được phép có nhiều cạnh song song. (Tùy định nghĩa mà Đa đồ thị có khuyên hay không có khuyên).



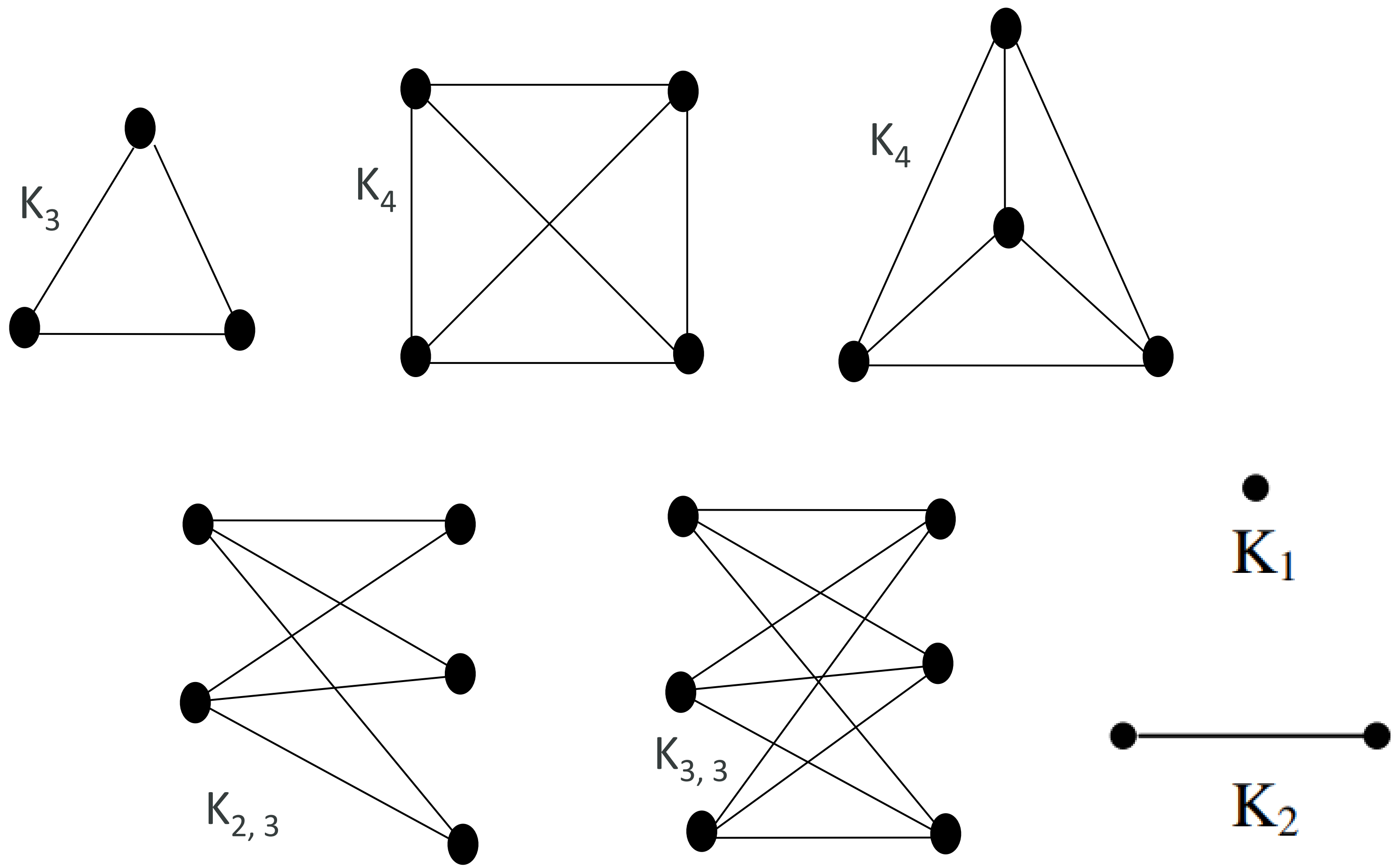
- Cây là một đồ thị liên thông không có chu trình.

CÁC DẠNG ĐỒ THỊ

- Đồ thị LƯỠNG PHÂN (Đồ thị hai phần, đồ thị hai phía, bipartite graph): đồ thị $G = (X, E)$ được gọi là đồ thị lưỡng phân nếu tập X được chia thành hai tập X_1 và X_2 thỏa:
  - X_1 và X_2 phân hoạch;
  - Cạnh chỉ nối giữa X_1 và X_2.

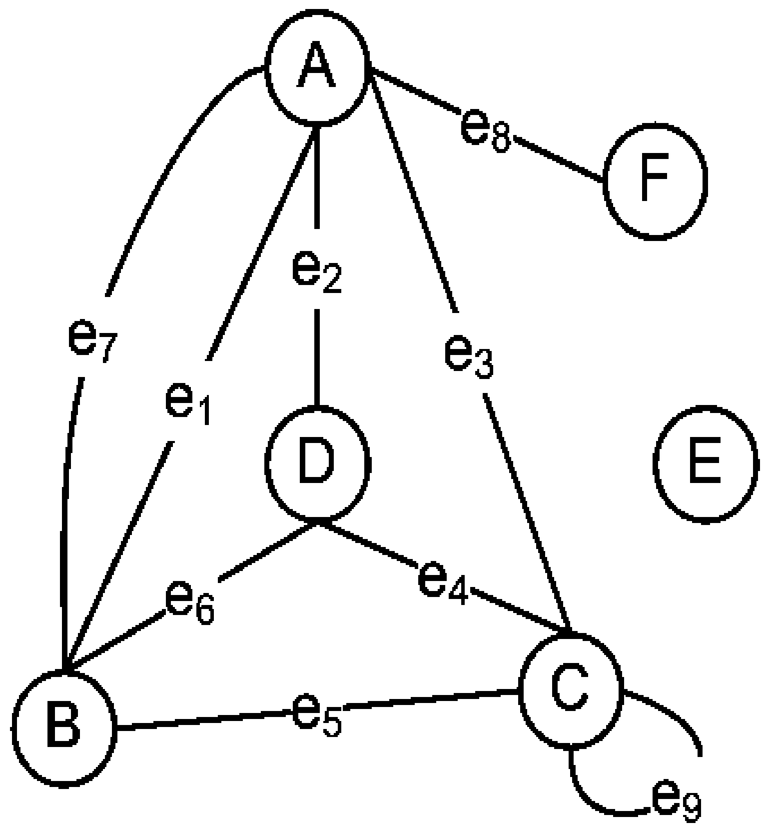
- Đồ thị LƯỠNG PHÂN ĐỦ: là đồ thị lưỡng phân đơn, vô hướng thỏa với $\forall(i, j)$: $i \in X_1$ và $j \in X_2$ có đúng một cạnh i và j.

VÍ DỤ ĐỒ THỊ ĐẦY ĐỦ VÀ ĐỒ THỊ 2 PHÍA



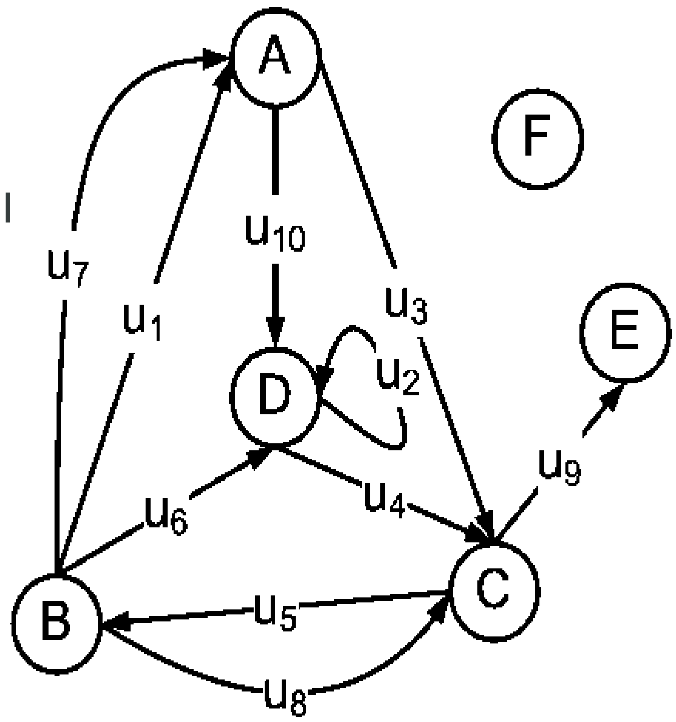
BẬC CỦA ĐỈNH

- Xét đồ thị vô hướng G
  - Bậc của đỉnh x trong đồ thị G là số các cạnh kề với đỉnh x, mỗi khuyên được tính hai lần
  - Ký hiệu: $dG(x)$ (hay $d(x)$) nếu đang xét một đồ thị nào đó.



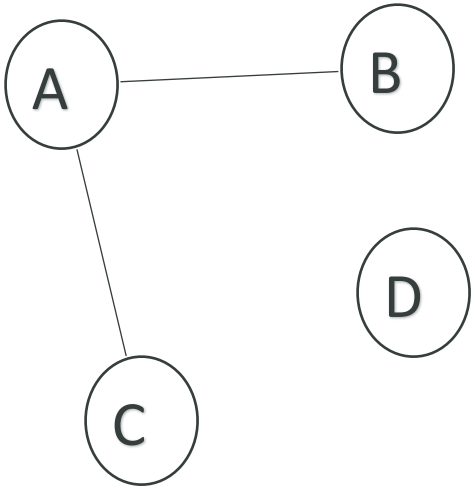
BẬC CỦA ĐỒ THỊ

- Xét đồ thị có hướng G
  - **Nửa bậc ngoài** của đỉnh x là số các cạnh đi ra khỏi đỉnh x, ký hiệu $\mathrm{d}^{+}(x)$.
  - **Nửa bậc trong** của đỉnh x là số các cạ hiệu $\mathrm{d}^{-}(x)$.
  - **Bậc** của đỉnh x: $\mathrm{d}(\mathrm{x}) = \mathrm{d}^{+}(x) + \mathrm{d}^{-}(x)$



BẬC CỦA ĐỈNH

- Đỉnh TREO là đỉnh có bậc bằng 1.
- Đỉnh CÔ LẬP là đỉnh có bậc bằng 0.



MỐI LIÊN HỆ BẬC – SỐ CẠNH

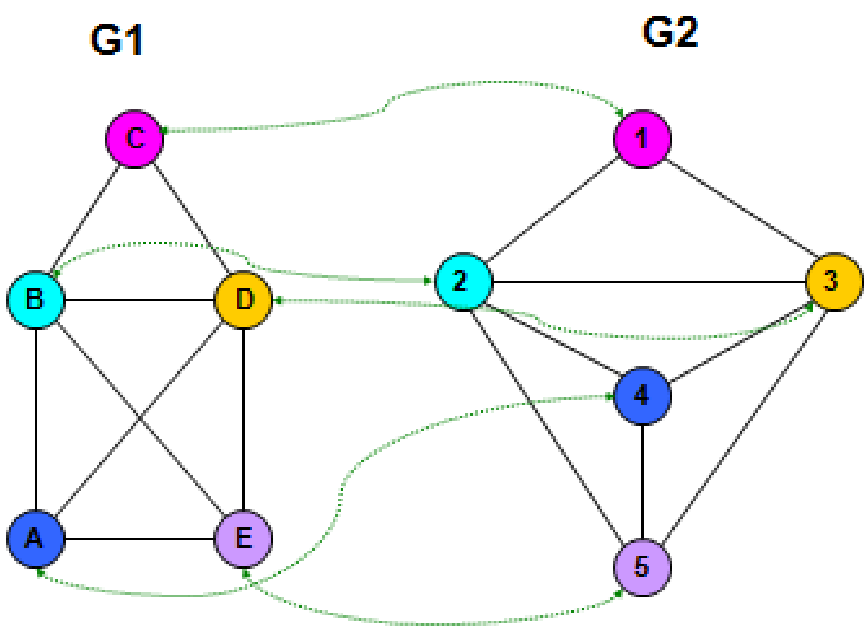
- Định lý:
  - Xét đồ thị có hướng $G = (X, U)$. Ta có:

$$\sum_{x\in X} \mathrm{d}^+(x) = \sum_{x\in X} \mathrm{d}^-(x) \text{ và } \sum_{x\in X} \mathrm{d}(x) = 2|U|$$

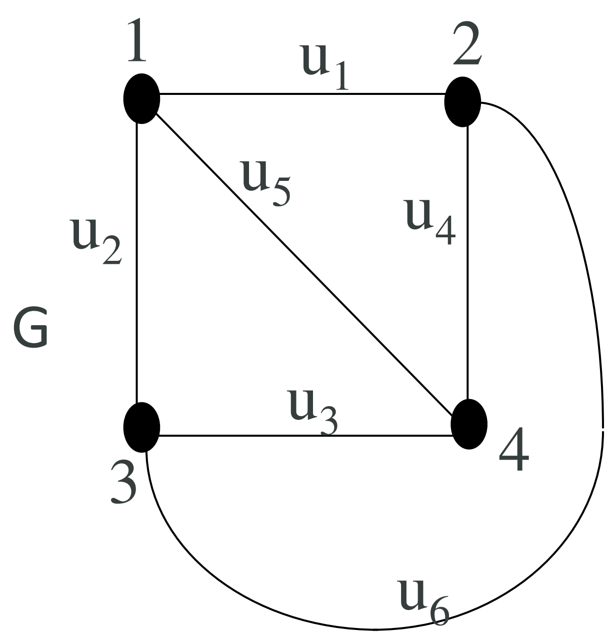
  - Xét đồ thị vô hướng $G = (X, E)$. Ta có: $\sum_{x\in X} \mathrm{d}(x) = 2|E|$

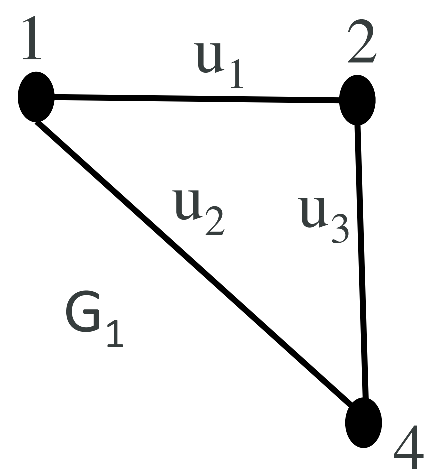
ĐẲNG CẤU ĐỒ THỊ

- Định nghĩa: Hai đồ thị $G_1 = (X_1, U_1)$ và $G_2 = (X_2, U_2)$ được gọi là đẳng cấu với nhau (ký hiệu: $G1 \sim G2$) nếu tồn tại một song ánh f: X_1 -> X_2 sao cho các đỉnh x, y là kề nhau trong G1 khi và chỉ khi f(x) và f(y) là kề nhau trong G2.
- Các tính chất của 2 đồ thị đẳng cấu:
  - Có cùng số đỉnh.
  - Có cùng số đỉnh bậc k, mọi k nguyên dương ≥ 0.
  - Cùng số cạnh.
  - Cùng số thành phần.



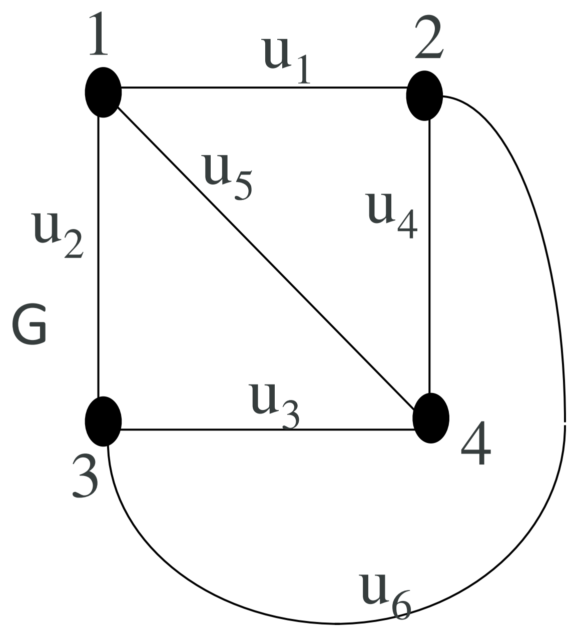
- Xét hai đồ thị $G = (X, U)$ và $G_1 = (X_1, U_1)$. G_1 được gọi là đồ thị con của G và ký hiệu $G_1 \in G$ nếu:
  - $X_1 \in X; U_1 \in U$
  - $u = (i, j) \in U$ của G, nếu $u \in U_1$ thì $i, j \in X_1$

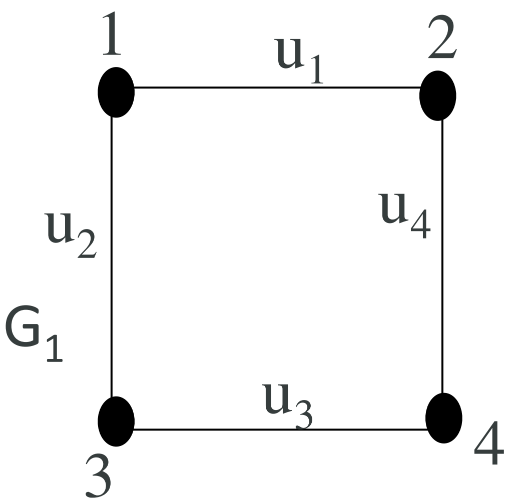




ĐỒ THỊ BỘ PHẬN

- Đồ thị con $G_1 = (X_1, U_1)$ của đồ thị $G = (X, U)$ được gọi là đồ thị bộ phận của G nếu $\mathrm{X} = X_1$.





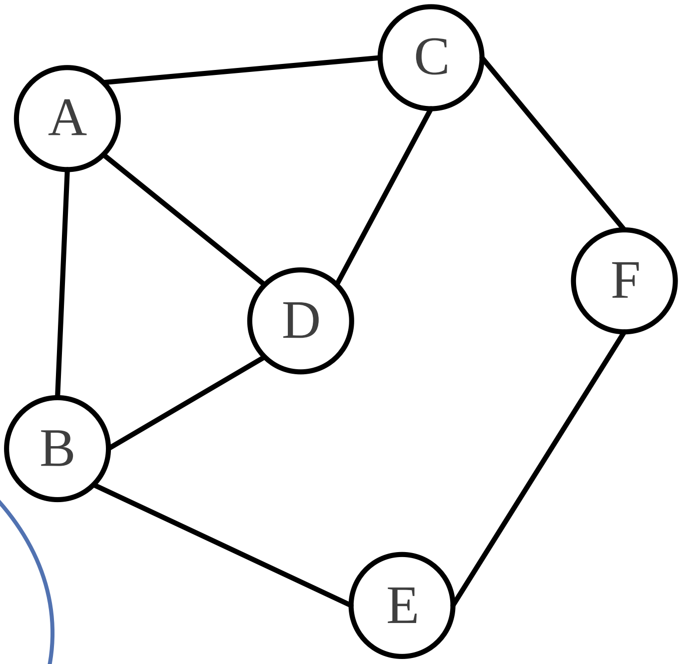
Path between vertices

- A path is a sequence of vertices (v_0, v_1, v_2,… v_k) such that:
  - For $0 \le i < k$, $\{v_i, v_{i+1}\}$ is an edge
  - For $0 \le i < k-1$, $v_i \ne v_{i+2}$ That is, the edge $\{v_i, v_{i+1}\} \ne \{v_{i+1}, v_{i+2}\}$ Note: a path is allowed to go through the same vertex or the same edge any number of times!
- The length of a path is the number of edges on the path

Types of paths

- A path is simple if and only if it does not contain a vertex more than once.
- A path is a cycle if and only if v0= vk
- The beginning and end are the same vertex!
- A path contains a cycle if some vertex appears twice or more

Examples



Are these paths?

Any cycles?

What is the path's length?

1. {a, c, f, e}
2. {a, b, d, c, f, e}
3. {a, c, d, b, d, c, f, e}
4. {a, c, d, b, a}
5. {a, c, f, e, b, d, c, a}

ĐƯỜNG ĐI, CHU TRÌNH

- Một đường đi trong $G = (V, U)$ là một đồ thị con $C = (V, E)$ của G với:
  - $V = \{x_1, x_2, \dots, x_M\}$
  - $E = \{u_1, u_2, \dots, u_{M-1}\}$ với $u_1 = x_1x_2, u_2 = x_2x_3, \dots, u_{\mathrm{M}-1} = x_{\mathrm{M}-1}x_M$; liên kết x_ix_{i+1} không phân biệt thứ tự.
- Khi đó, x_1 và x_M được nối với nhau bằng đường đi C với: $C.x_1$ là đỉnh đầu và $\mathrm{C}.x_M$ là đỉnh cuối.
- **Note: a path is allowed to go through the same vertex or the same edge any number of times!**
- Số cạnh của C được gọi là độ dài của C.
- Khi các cạnh hoàn toàn xác định bởi cặp đỉnh kề, đường có thể viết gọn $(x_1, x_2, \dots, x_M)$

DÂY CHUYỀN, CHU TRÌNH

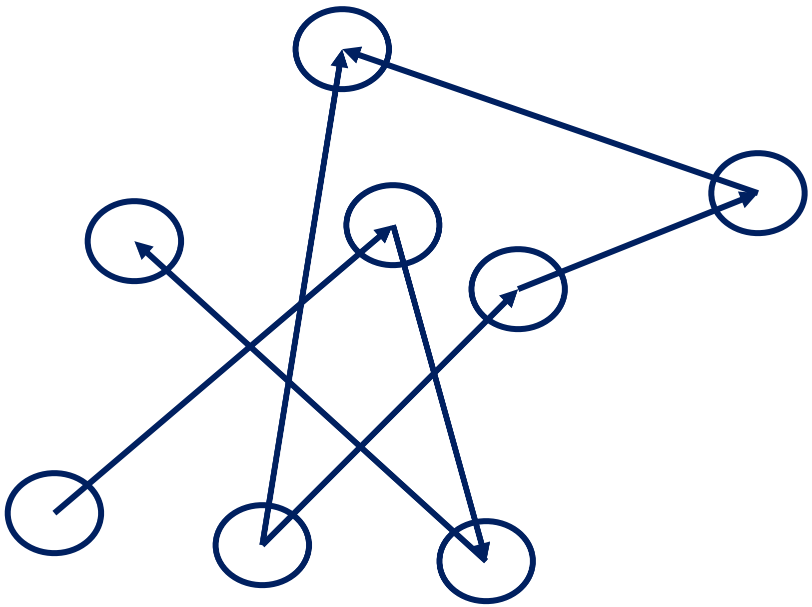
- Dây chuyền đơn là dây chuyền không đi qua
- CHU TRÌNH (cycle): là một dây chuyền có đỉnh đầu và đỉnh cuối trùng nhau.
- Chu trình đơn: là chu trình không đi qua một cạnh nào quá một lần.
- Chu trình SƠ CẤP: là chu trình không chứa cùng một đỉnh quá một lần (trừ đỉnh đầu và đỉnh cuối).
- Chu trình Euler: là chu trình qua tất cả các cạnh, mỗi cạnh đúng một lần.
- Chu trình bao trùm: là cách gọi khác của chu trình Hamilton: là dây chuyền Hamilton xuất phát từ một đỉnh, đi qua tất cả các đỉnh khác của đồ thị, mỗi đỉnh đúng một lần và quay trở về nơi xuất phát.

THÀNH PHẦN LIÊN THÔNG

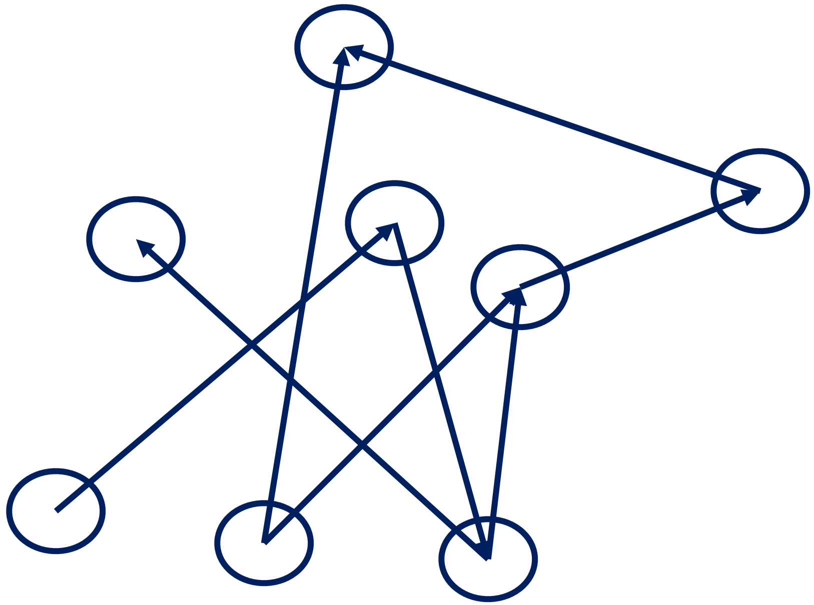
- Một **thành phần liên thông** của một đồ thị vô hướng là một đồ thị con trong đó:
  - Giữa bất kỳ hai đỉnh nào đều có đường đi đến nhau, và
  - Không thể nhận thêm bất kỳ một đỉnh nào mà vẫn duy trì tính chất trên.
  - Một đồ thị liên thông có đúng một thành phần liên thông, chính là toàn bộ đồ thị.

THÀNH PHẦN LIÊN THÔNG

- G gồm 2 thành phần liên thông, H là đồ thị liên thông.



G



H

THÀNH PHẦN LIÊN THÔNG

Thuật toán xác định các thành phần liên thông:

Input: đồ thị $G = (X, E)$, tập X gồm N đỉnh $1, 2, \dots, N$

Output: các đỉnh của G được gán nhãn là số hiệu của thành phần liên thông tương ứng

1. Khởi tạo biến $label = 0$ và gắn nhãn 0 cho tất cả các đỉnh
2. Duyệt qua tất cả các đỉnh $i \in X$

   Nếu nhãn của i là 0

   1. $label = label + 1$
   2. Gán nhãn cho tất cả các đỉnh cùng thuộc thành phần liên thông với i là $label$

THÀNH PHẦN LIÊN THÔNG

Thuật toán gán nhãn các đỉnh cùng thuộc thành phần liên thông với đỉnh i –Visit(i, $label$):

Input: đồ thị $G = (X, E)$, đỉnh i, nhãn $label$

Output: các đỉnh cùng thuộc thành phần liên thông với i được gắn nhãn $label$

1. Gắn nhãn $label$ cho đỉnh i

2. Duyệt qua tất cả các đỉnh $j \in X$ và có cạnh nối với i

    Nếu nhãn của j là 0

        Visit(j , $label$)

Chúc các em học tốt!