

Lecture 12: Introduction to Graph Theory

*Instructor: Goutam Paul**Scribe: Anirban Bhattacharyya*

Contents

1	Introduction	12-1
2	Some Basic Terminologies	12-1
3	Graph Coloring	12-6
4	Some Problems	12-7

1 Introduction

Graphs are discrete structures consisting of vertices and edges that connect these vertices. There are different kinds of graphs, depending on whether edges have directions, whether multiple edges can connect the same pair of vertices, and whether loops are allowed. Problems in almost every conceivable discipline can be solved using graph models. In this scribe we'll discuss about some terminologies related to graphs, will have a brief discussion about graph coloring.

2 Some Basic Terminologies

We begin with the definition of an undirected graph.

Definition 2.1 (Undirected Graph). *An undirected graph $G = (V, E)$, where V is the set of vertices or nodes and E or the set of edges contains some of the 2-element subsets of V as its members.*

Example 1. *Let us say $G = (V, E)$ is an undirected graph and V is the set of vertices. $V = \{A, B, C, D, E\}$ and edge set $E = \{\{A, B\}, \{B, C\}, \{D, E\}\}$ G is an undirected graph.*

Thus G can be represented in the way below.

Remark. *Here one should note that for an undirected graph edges have no orientation. The edge $\{A, B\}$ is identical to the edge $\{B, A\}$. That is edges of an undirected graph are not ordered pairs but unordered pairs. The maximum number of edges in an undirected graph without a loop is $n(n - 1)/2$, where n is the number of nodes in the graph.*

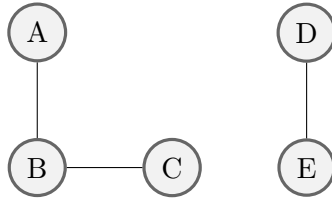


Figure 1: Undirected Graph G

Definition 2.2 (Directed Graph). A *directed graph* or *dgraph* is a graph $G = (V, E)$ where V is the set of vertices or nodes and the set of edges E is a subset of $V \times V$. Thus E is a relation from V to V .

The edges are called *directed edge* and each directed edge is associated with an ordered pair of vertices. Each such edge associated with the ordered pair (X, Y) is said to start at X and end at Y .

Example 2. We assume a graph $G_1 = (V_1, E_1)$. For this graph we have the set of vertices $V_1 = \{A, B, C, D, E\}$ and the edge set is defined as below:

$$E_1 = \{\{A, B\}, \{C, B\}, \{D, E\}, \{E, D\}\}$$

Thus $E_1 \subseteq V_1 \times V_1$ and we can draw the graph in the following way.

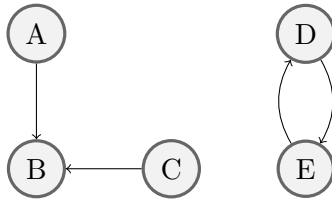


Figure 2: A directed Graph G_1

Remark. The set of vertices V of a graph G may be infinite. A graph with an infinite vertex set or an infinite number of edges is called an **infinite graph**, and in comparison, a graph with a finite vertex set and a finite edge set is called a **finite graph**.

The graphs we discussed about till now, don't have any **multiple edges** or any **self loop**. A **multiple edge** is more than one edges that connect the same two vertices. In other words, the edge set is a **multiset**. Such graph are called **multigraph**. When there are m different edges associated to the same unordered pair of vertices $\{X, Y\}$, we also say that $\{X, Y\}$ is an edge of multiplicity m . That is, we can think of this set of edges as m different copies of an edge $\{X, Y\}$.

A **self loop** is an edge that connects a vertex to itself. Mathematically, if $G = (V, E)$ is a graph then $\forall x \in V, (x, x) \notin E$.

For a directed graph things become a bit more complicated. Similar to an undirected graph, a directed graph may contain loops and it may contain multiple directed edges that start and end at the same vertices. A directed graph may also contain directed edges that

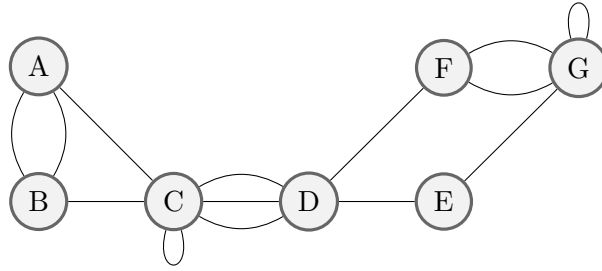


Figure 3: An example of an undirected graph with self loop and multiples edges

connect vertices X and Y in both directions; that is, when a dgraph contains an edge from X to Y , it may also contain one or more edges from Y to X .

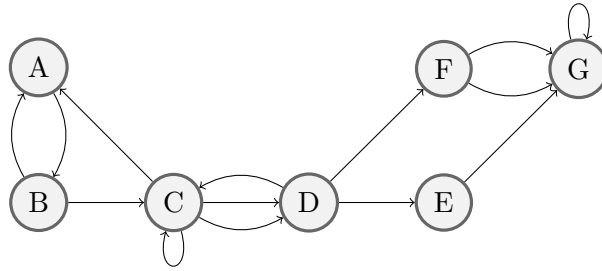


Figure 4: An example of a directed graph with self loop and multiples edges

The above discussion leads to the following definition.

Definition 2.3 (Simple Graph). *A graph $G = (V, E)$ is called simple if and only if it doesn't contain any multiple edges and self loop.*

Remark. *Unless stated otherwise, the graphs we deal with are all simple graphs.*

Definition 2.4 (Subgraph and Induced Subgraph). *A graph $G' = (V', E')$ is called a **subgraph** of $G = (V, E)$ if the vertex set and the edge set are subsets of those of G (i.e. $V' \subseteq V$ and $E' \subseteq E$).*

*If G' is a subgraph of G , then G is said to be a **supergraph** of G' .*

*Let $G = (V, E)$ be any graph, and let $V' \subseteq V$. Then the **induced subgraph** G' is the graph whose vertex set is V' and whose edge set consists of all of the edges in E that have both endpoints in V' . The same definition works for both undirected graphs and directed graphs.*

Remark. *To avoid ambiguity, we say a subgraph can have less edges between the same vertices than the original one. But an induced subgraph can be constructed by deleting vertices (and with them all the incident edges), but no more edges. If additional edges are deleted, the subgraph is not induced.*

Also, a graph is a subgraph of itself.

Definition 2.5 (Complete Graph). *An undirected graph $G = (V, E)$ is called **complete** if E contains all the $\binom{n}{2}$ many 2 element subsets of V where n is the cardinality of V .*

A complete graph on n vertices is denoted by K_n and it contains exactly one edge between each pair of distinct vertices. A simple graph for which there is at least one pair of distinct vertex not connected by an edge is called **noncomplete**. The graphs K_n for $n = 1, 2, 3, 4, 5, 6$ are displayed below.



Figure 5: K_1, K_2

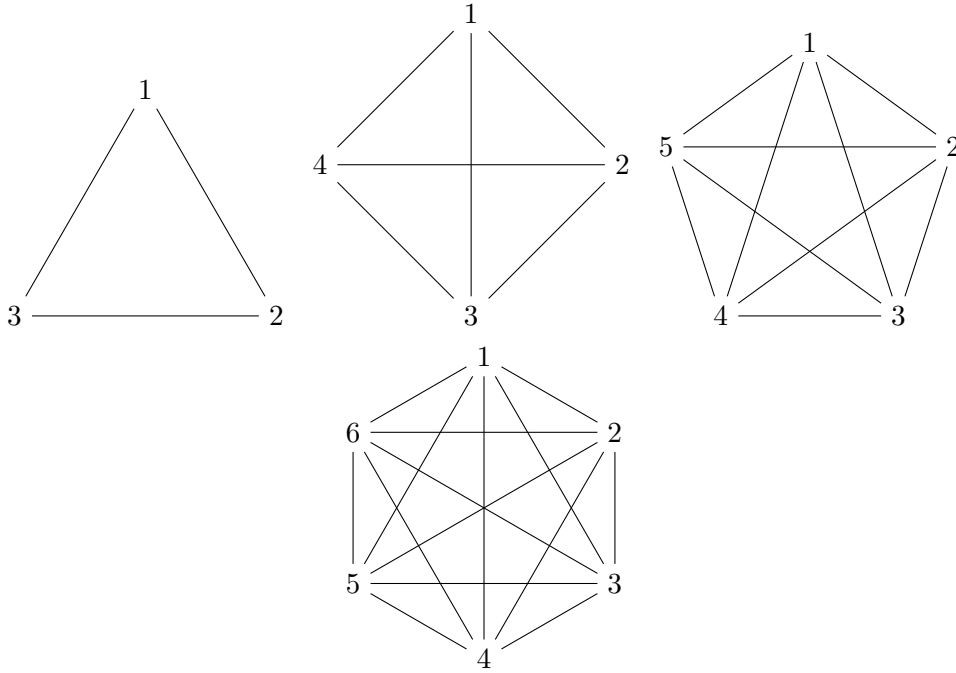


Figure 6: K_3, K_4, K_5, K_6

Definition 2.6 (Clique). A **clique** C , in an undirected graph $G = (V, E)$ is a subset of the vertices, $C \subseteq V$, such that every two distinct vertices are adjacent. This is equivalent to the condition that the induced subgraph of G induced by C is a complete graph.

A **maximal clique** is a clique that cannot be extended by including one more adjacent vertex, that is, a clique which does not exist exclusively within the vertex set of a larger clique.

The clique of largest possible size is called a **maximum clique**. The **clique number** $\omega(G)$ of a graph G is the number of vertices in a maximum clique in G .

Thus we can conclude that maximum cliques are maximal cliques but not necessarily vice versa. In other words, a maximal clique can't be a subset of a larger clique. But maximum means that there is no larger clique in the graph.

As an example, the graph G in Figure 9, has two cliques namely K_3 and K_4 in it. The size of the largest clique in the graph is 4 therefore K_4 is a maximum clique. The K_3 is maximal, meaning that if we add any other vertex in G to the K_3 we no longer have a clique. But it is not maximum since the size of it is $3 < 4$.

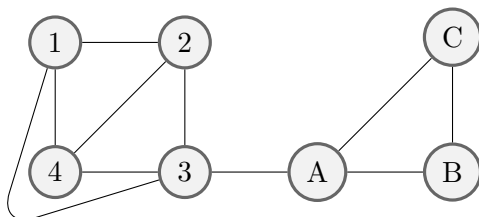


Figure 7

Definition 2.7 (Independent Set). Given $G = (V, E)$, an **independent set** is a set of vertices, no two of which are adjacent. That is, it is a set $I \subseteq V$ of vertices such that for every two vertices in I , there is no edge connecting the two.

The size of an independent set is the number of vertices it contains.

A **maximal independent set** is either an independent set such that adding any other vertex to the set forces the set to contain an edge or the set of all vertices of the empty graph.

A **maximum independent set** is an independent set of largest possible size for the given graph G . This size is called the independence number of G , and denoted by $\alpha(G)$.

Clearly, any subset of an independent set is an independent set. Also as before, we can't increase size of a maximal independent set by adding one more vertex.

In the figure 10, the graph has 5 nodes, namely, $\{A, B, C, D, E\}$. $I = \{A, B, E\}$ is an independent set. Also it is maximal as adding any vertex makes it dependent. As an example, if we add "D" to I, then "B" and "D" will be adjacent.

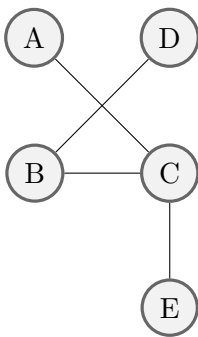


Figure 8

Definition 2.8 (Complement of a Graph). Given a graph $G = (V, E)$, we say $\bar{G} = (V, \bar{E})$ complement of G if $\forall X, \forall Y \in V, \{X, Y\} \in \bar{E}$ but $\{X, Y\} \notin E$.

In the Figure 11, we see a graph G on the left, and complement of it, \bar{G} next to it. Point to note, G is not K_4 here.

To keep track of how many edges are incident to a vertex, we make the following definition.

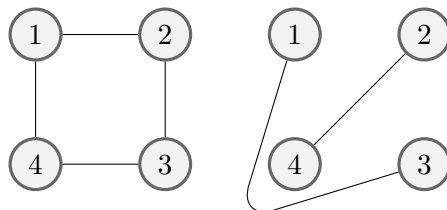


Figure 9: An example of G and \bar{G}

Definition 2.9 (Degree of a Vertex). *The **degree of a vertex** in an undirected graph is the number of edges incident with it, except that a loop at a vertex contributes twice to the degree of that vertex. The degree of the vertex v is denoted by $\text{deg}(v)$.*

A vertex of degree zero is called **isolated**. It follows that an isolated vertex is not adjacent to any vertex. A vertex is **pendant** if and only if it has degree one. Consequently, a pendant vertex is adjacent to exactly one other vertex.

We will calculate the degrees and the neighborhoods of the vertices (the edges adjacent to a vertex) in the graph G depicted below.

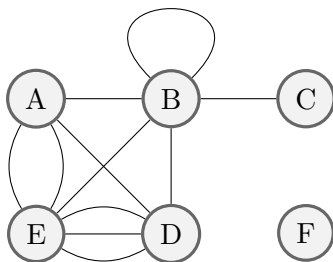


Figure 10: G

In G , $\text{deg}(A) = 4$, $\text{deg}(B)=6$ (as there is a loop, it is counted twice), $\text{deg}(E) = 6$, $\text{deg}(C) = 1$, $\text{deg}(D) = 5$, $\text{deg}(F)=0$. The neighborhoods of these vertices are $N(A) = \{B,D,E\}$, $N(B) = \{A, B, C, D, E\}$, $N(C) = \{B\}$, $N(D) = \{A, B, E\}$, $N(E) = \{A, B, D\}$ and $N(F)=\emptyset$. Thus vertex F is an isolated vertex and vertex C is a pendant vertex.

3 Graph Coloring

Graph coloring is a special case of graph labeling; it is an assignment of labels traditionally called “colors” to elements of a graph subject to certain constraints. In its simplest form, it is a way of coloring the vertices of a graph such that no two adjacent vertices are of the same color; this is called a vertex coloring. Similarly, an edge coloring assigns a color to each edge so that no two adjacent edges are of the same color, and a face coloring of a planar graph assigns a color to each face or region so that no two faces that share a boundary have the same color.

Definition 3.1 (Vertex Coloring). *Given a graph $G = (V, E)$, a **k -coloring** of G is a function $\psi : V \rightarrow \{1, 2, \dots, k\}$, such that if $\{X, Y\} \in E$, then $\psi(X) \neq \psi(Y)$. In other*

words, adjacent vertices get different colors.
If such a coloring exists, then G is called **k -colorable**.

Since a vertex with a loop (i.e., a connection directly back to itself) could never be properly colored, it is understood that graphs in this context are loopless. Also we can state a trivial result here that, K_n is c -colorable iff $c \geq n$.

Definition 3.2 (Edge Coloring). Given a graph $G = (V, E)$, an **edge coloring** of G with k colors is a function $\psi : E \rightarrow \{1, 2, \dots, k\}$. Such a coloring is a **valid edge-coloring** if no two edges sharing a vertex are assigned the same color.

4 Some Problems

Problem 1. Six people meet in a party. Show that either there are at least three who have mutually shaken hands before or there are at least three, no two of whom have shaken hands before. Show also that the number ‘six’ in the statement can’t be replaced by the number ‘five’ or less.

Solution. For convenience we shall refer to any two people who have shaken hands before by the term “friends” or “acquaintances” and any two who have never done so, by the term “strangers”. So three persons who, two by two, are friends would be called “mutual friends”; and in the same manner, three people who, two by two, are strangers would be called “mutual strangers”. So the problem requires us to show that in any party of six people either three of them are mutual strangers or three of them are mutual acquaintances. The proof of this statement requires nothing but a three-step logic.

In order to help the understanding, it is convenient to convert the problem to graph theoretic setting. Suppose we have a complete graph K_6 . Now it has $\binom{6}{2} = 15$ edges. Let the 6 vertices of K_6 stand for the $\binom{6}{2}$ people in our party. Let the edges colored red or green according as the two people represented by the vertices connected by the edge are mutual strangers or acquaintances. The problem now asserts: In whatever way you may color the 15 edges of a K_6 red or green, you can never avoid either a red triangle - that is, a triangle all of whose three sides are red - or, in the alternative, a green triangle. The interesting proof goes as follows.

Consider K_6 in the figure 14. If we choose any one vertex, say A . There are five lines going forth from A . They are colored green or red - some red, some green. We do not know how many of them are red and how many of them are green. It could be all five red; four red and one green; three red and two green; two red and three green; one red and four green; or all five green. The beauty here is the relevance of the PHP. Since there are only two colors and we have five lines which fall into either one of them, the PHP says there are at least three of the same color. (The above listing of the different possibilities certainly confirms this conclusion to arrive at which, however, the listing was not necessary).

Suppose C, D, E are the other ends of these three lines, all of the same color, say, red. If either one of CD, DE, CE is red then that line with the two edges from P meeting it at its ends would give a red triangle. If none of CD, CE, DE is red, then all three are green and we have a green triangle CDE . This concludes the proof. Finally, we note that the number

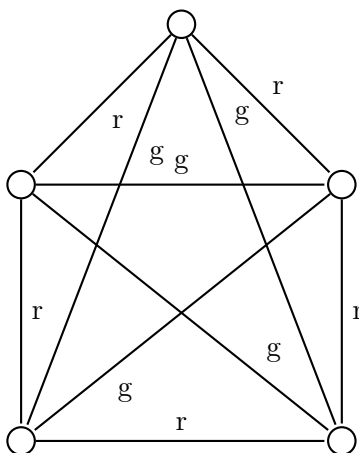


Figure 11

'six' can't be replaced by 'five'. For, if we had a K_5 , and we colored its 10 edges red and green, it is not always true that there exists either a red triangle or green triangle; for, we have to look at edge coloring of K_5 in figure 13 where the letters r and g indicated on the edges stand for red and green respectively. We note that, for this particular coloring, there is neither a red triangle nor a green triangle. This single case tells us that the existence of a monochromatic triangle is not universally true in the case of K_5 . Certainly it can't be so in the case of K_4 or K_3 . Thus 6 is the least possible value of n for which K_n has this property. Note that when K_6 has the property, all K_n , $n > 6$ has also the property.

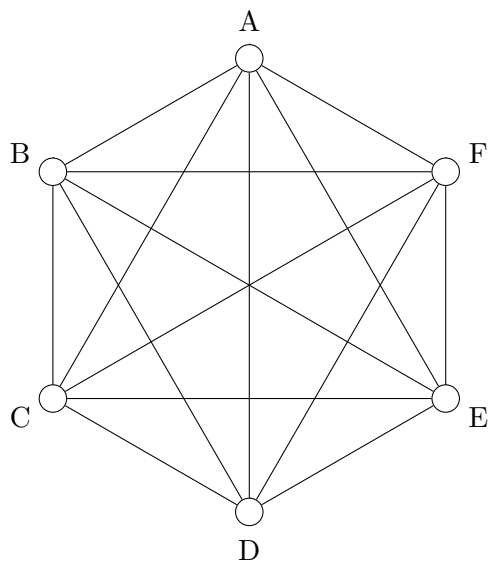


Figure 12