MA238: Discrete Mathematics

10 Graph Theory

10.1 Graphs & Graph Models

Definition. *A graph (or undirected graph) $G=(V,E)$ where V is a nonempty set of elements called vertices and where E is a set of unordered pairs of elements of V called edges.*

Example. $G=(V,E)$ where $V=\{1,2,3,4\}$ and $E=\{\{1,1\},\{1,2\},\{2,3\},\{3,4\},\{4,1\},\{1,3\}\}$.

We can represent a graph by making a point for each vertex and a curve joining a and b for each edge, $\{a,b\}$.

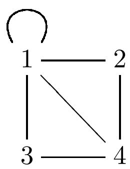

Figure 1. Graph $G=(V,E)$

Definition. *For an edge $e=\{a,b\}$, a and b are called the endpoints of e, and e is said to connect or join its endpoints a and b.*

Example. Gene interactions.

Definition. *A directed graph or digraph, $G=(V,E)$ consists of a nonempty set, V, of elements called vertices and a set, E, of ordered pairs of elements of V, called [directed] edges or arcs.*

Example. $G=(V,E)$ where $V=\{1,2,3,4\}$ and $E=\{(1,1),(1,2),(2,3),(3,2),(3,4),(4,1),(1,3)\}$

We represent a digraph by making a point for each vertex and a curve joining a and b for each edge, (a,b) with an arrow pointing from a to b.

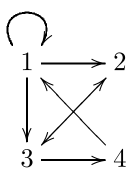

Figure 2. Graph of $G=(V,E)$

Definition. *For a directed edge, $e=(a,b)$, a is called the tail of e (start or initial vertex of e), denoted $t(e)$, and b is called the head of e (end or terminal vertex of e), denoted $h(e)$.*

Vertices a and b are called the endpoints or ends of e.

Remark. "Vertex" is singular. Its plural is "vertices". **There is no such thng as one vertice**.

Definition. *A loop in a graph or an undirected graph is an edge with both endpoints the same.*

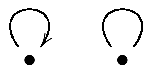

Figure 3. *Directed and Undireced Loops*

Definition. *In a graph, multiple edges, are distinct edges with the same endpoints. We sometimes speak of the multiplicity of an edge with endpoints a and b which means the number of edges with these endpoints.*

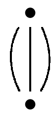

Figure 4. *These points have multiplicity 3*

Definition. *A graph without loops or multiple edges is called a simple graph.*

Definition. *In a digraph, multiple edges are edges with the same head and tail.*

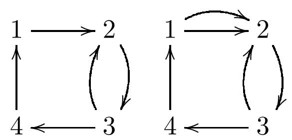

Figure 5. *First graph does not have multiple edges, second one does*

Definition. *The multiplicity of an edge, with tail a and head b, is the number of edges with tail a and head b. A digraph without loops or multiple edges is called a simple directed graph.*

Remark. *A mixed graph is one with both directed and undirected edges. Ex. A roadmap*

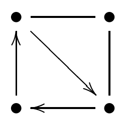

Figure 6. *Example of a mixed graph where undirected edges are two way streets*

Example. A digraph. The web can be modelled as a digraph where each webpage corresponds as a vertex and there is an edge from webpage A, to webpage B if there is a direct link on A to B. (Some people prefer to think of the web graph as an undirected graph).

Example. Another example of a digraph, a precedence digraph. The vertices correspond to activities. There is an edge direced from the vertex representing activity a, to the vertex representing activity b, exactly when activity a must be completed before activity b can begin.

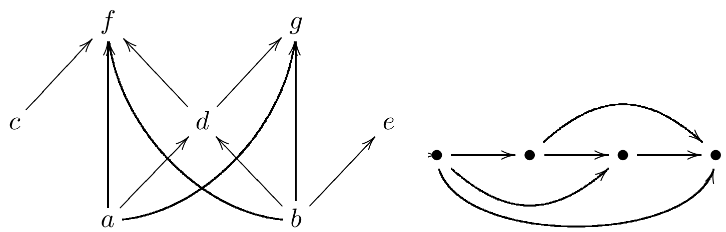

Figure 7. Examples Precedence Digraph

Suppose each vertex corresponds to a job that takes 1 day to do. How can we schedule the jobs to days to minimize the number of days required and which is feasible (which means when job j is stared, all predecessing jobs are actually already completed).

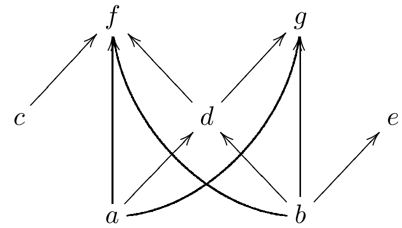

Figure 8. Initial job schedule

We can do jobs c, a, b on day one, the remaing job graph is

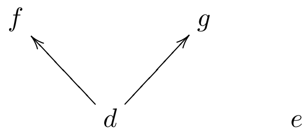

Figure 9. job schedule for day 2

We can now do d, e, then do f, g on the next day

$$f \qquad\qquad g$$

Figure 10. jub schedule for day 3

Suppose the foreman wants you to finish the job in 2 days. You can show him the directed path

$$a \longrightarrow d \longrightarrow g$$

Figure 11. This will obviously take 3 days

10.2 Graph Terminology & Special Types of Graphs

Definition. *If $e = \{u, v\}$ is an edge of graph G, then vertices u and v are said to be adjacent and edge e is said to be incident to its endpoints u and v. Vertices u and v are called neighbours. The set $N(v)$ of all neighbours of vertex v, is called the neighbourhood of v or neighbour set of v.*

Definition. *The degree of a vertex v in an indirected graph, G, is denoted $\deg(v)$ or $d(v)$, and this is the number of non-loop edges indcident to v plus 2 times the number of loop edges incident to v. Think of the $\deg(v)$ as the number of edge ends.*

Example. **insert graph**

A vertex is isolated if it degree is 0 and is a pendant vertex if its degree is 1.

Example. **insert graph**

Theorem. ***The Handshaking Theorem***

Let $G=(V,E)$ *be a undirected graph. Then* $2|E|=\sum_{v\in V}\deg(v)$.

Acknowledgments. *To be sure we understand the terminology, we will check for above.*

insert graph

$2|E|=2\cdot 6=12$ *and* $\sum_{v\in V}\deg(v)=\sum\{\deg(v): v\in V\}=3+4+3+2=12$

Proof. *Let* $G=(V,E)$ *be an undirected graph.*

Consider the sum: $\sum_{v\in V}\deg(v)$. *A non-loop edge* $e=\{u,v\}$ *contributes 1 to the degree of* u *and 1 to the degree of* v*, so altogether 2 to the sum.*

A loop edge $e=\{u,u\}$ *contributes 2 to the degree of* u *and thus, 2 to the sum.*

$\therefore$ *Each edge of* G *contributes 2 to the sum.*

$$\therefore 2|E|=\sum_{v\in V}\deg(v)$$

□

Theorem. ***Corollary of Prev. Theorem***

An undirected graph has even number of vertices of odd degree (possibly 0, but not part of theorem).

Proof. *Let* $G=(V,E)$ *be an undirected graph. By the* ***Handshaking Theorem****,*

$$2|E|=\sum_{v\in V}\deg(v)=\sum_{v\in V,\deg(v)\text{ is even}}\deg(v)+\sum_{v\in V,\deg(v)\text{ is odd}}\deg(v)$$

$2|E|$ *and* $\sum_{v\in V}\deg(v)$ *are even. Thus,* $\sum_{v\in V,\deg(v)\text{ is odd}}\deg(v)$ *is even. The only way that a sum of odd numbers can be even, is if there are an even number of odd terms in the sum.*

$\therefore$ *The number of vertices of odd degree is even.*

□

Definition. *The* <u>*degree sequence*</u> *of a graph is a sequence of the degrees of the vertices in non-increasing order.*

Definition. *A sequence* $d_1,d_2,...,d_n$ *is called* <u>*graphic*</u> *if it is a degree sequence of a simple graph.*

Question. *Let* $d_1,d_2,...,d_n$ *be a non-increasing sequence of non-negative integers.*

1. *Is* $d_1,...,d_n$ *the degree sequence of a graph?*
2. *Is this the degree sequence of a simple graph?*

Definition. *A graph is called* <u>*regular*</u> *if every vertex has the same degree.*

Definition. *In a digraph, $G = (V, E)$, the indegree of a vertex is the number of edges whose head is v and denoted $\deg^{-}(v) = \deg^{\text{in}}(v)$.*

The outdegree of v is the number of number of edges whose tail is v, and is denoted $\deg^{+}(v) = \deg^{\text{out}}(v)$.

Theorem. *Let $G = (V, E)$ be a directed graph. Then:*

$$\sum \{\deg^{-}(v) : v \in V\} = \sum \{\deg^{+}(v) : v \in V\} = |E|$$

Proof. *Let $G = (V, E)$ be a digraph. For each edge, $e \in E$;*

- *if e is a loop, say $e = (w, w)$, then e cotributes 1 to the $\deg^{-}(w)$ and contributes 1 to the $\deg^{+}(w)$ and so, 1 to each sum in the theorem*

- *if e is not a loop, say $e = (u, v)$, where $u \neq v$, then e contributes 1 to $\deg^{+}(u)$ and 1 to $\deg^{-}(v)$ and so, 1 to each sum in the theorem* □

Definition. *The underlying undirected graph of a digraph, G, is obtained from G by removing the directions of the edges.*

Some Special Graphs

1. **Complete Graph** The complete graph on n vertices, denoted K_n, is the simplest graph on n vertices that contains exactly one one edge between every distinct pair of vertices $(n \geq 1)$.

2. **Cycles** The cycle, C_n, $n \geq 3$, consists of n vertices $v_1, v_2, ..., v_n$ and edges $\{v_1, v_2\}, \{v_2, v_3\}, ...,$ $\{v_n, v_1\}$.

3. **Wheels** The wheel, W_n, is obtained from C_n by adding another vertex, v_{n+1}, and joining it to each of the vertices of C_n.

4. **N-Cubes (hypercubes)** The n-dimensional hypercube or n-cube, denoted N_n, is the graph whose vertices correspond to the certain number of 0-1 strings of length n; two vertices are adjacent if the strings correspond to differ in exactly one digit $(n \geq 1)$.

5. **Bipartite Graph** A graph, $G = (V, E)$, is called bipartite if its vertex-set can be divided into two, disjoint sets, V_1 and V_2, such that every edge has one end in V_1 and the other in V_2.

6. **Complete Bipartite Graph** The complete bipartite graph, $K_{m,n}$ ($m \geq 1$ and $n \geq 1$), is a graph whose vertices are partitioned two sets, V_1 with m vertices and V_2 with n vertices and there is an edge between two vertices exactly when one is in V_1 and the other is in V_2.

7. **Paths** The path, P_n, $n \geq 1$, consists of n vertices $v_1, ..., v_n$ and edges $\{v_1, v_2\}, \{v_2, v_3\}, ...,$ $\{v_{n-1}, v_n\}$.

Definition. *The complementary graph, or complement, $\bar{G}$ of a simple graph, G, has the same vertices as G and two vertices are adjacent in $\bar{G}$ exactly when they are not adjacent in G.*

Remark. How many edges does K_n have?

For each of the n vertices, it is adjacent to $n-1$ others. Thus, $n(n-1)$ counts all the edges twice. Thus, the number of edges is $\frac{n(n-1)}{2} = \binom{n}{2}$.

Or, apply the *Handshaking Theorem*: The degree of each vertex in K_n is $n-1$. The theorem says,

$$2|E| = \sum_{v \in V} \deg(v)$$

$=(n-1)+...+(n-1)$ (n times)

$\therefore |E| = \frac{n(n-1)}{2}$

Remark. How many edges for $K_{m,n}$ have?

It has $m \cdot n$ edges. For each of the m vertices of V_1, it is joined to the n vertices of V_2.

Or, apply the *Handshaking Theorem*.

If the degree of vertex, v, in G is k and G has n vertices, what is the degree of v in $\bar{G}$?

The $\deg(v)$ in G and $\deg(v)$ in $\bar{G} = \deg(v)$ in $K_n = n-1$.

Definition. *A matching, M, in a graph $G=(V,E)$ is a set of edges, no two which are incident to the same vertex.*

A matching is called a maximum (largest), if it has the largest number of any matching edges in the graph.

Definition. *A subgraph of a graph $G=(V,E)$ is a graph, $H=(W,F)$ where $W \subseteq V$ and $F \subseteq E$.*

A subgraph induced by subset W of V is the graph wth vertex-set W, and edge-set all edges with both ends in W.

Note. *There are many subgraphs with vertex-set, W. There is only 1 subgraph induced by W.*

Definition. *The union of two graphs $G_1=(V_1,E_1)$ and $G_2=(V_2,E_2)$ is the graph with vertex-set $V_1 \cup V_2$ and edge-set $E_1 \cup E_2$. The union of G_1 and G_2 is denoted $G_1 \cup G_2$.*

If G_1 and G_2 are simple graphs and $G_1 \cup G_2$ is also a simple graph.

10.3 Representing Graphs & Graph Isomorphism

Representing Graphs

1. Adjacency Lists - Specifies the vertices adjacent to a given vertex

   Definition. *The adjacency matrix A of a simple graph, G, with vertices v_1, ..., v_n is an $n \times n$ 0-1 matrix whose $(i,j)^{\text{th}}$ entry is 1 when v_i and v_j are adjacent and 0 when they are not adjacent.*

   If the graph is not simple, then the $(i,j)^{\text{th}}$ entry is the number of edges pointing from v_i to v_j. In particular, the $(i,i)^{\text{th}}$ entry of the adjacency matrix is the number of loops at v_i.

   If $G=(V,E)$ is a digraph, then the $(i,j)^{\text{th}}$ if the number of edges (v_i,v_j), that is, the number of edges with tail, v_i and head v_j.

   Definition. *The incidence matrix, M or M_G, of a loop-free graph, G, with n vertices $v_1,...v_n$ and m edges $e_1,...,e_m$ is an $n \times m$ matrix whose $(i,j)^{\text{th}}$ entry is 1 when e_j is incident to v_i and 0 otherwise.*

Graph Isomorphism

Definition. *Two simple graphs, $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are isomorphic if there is a 1-1 correspondence (a bijection), f, from V_1 to V_2 with the property that vertices a and b are adjacent in G_1 iff vertices $f(a)$ and $f(b)$ are adjacent in G_2.*

Remark. There is no efficient (polynomial time) algorithm known for: *Are these two graphs efficient?* Although, the problem is known to be NP-complete.

To demonstrate that two graphs, G_1 and G_2 are ismorphic, give a bijection, f, between their vertices and check:

- For each edge, $\{a, b\}$ of G_1, chekc that $\{f(a), f(b)\}$ is an edge of G_2
- For each pair of vertices a, b such that $\{a, b\}$ is not an edge of G_1; check that $\{f(a), f(b)\}$ is not an edge of G_2

Simplification. For each edge of $\{a, b\}$ of G_1 check that $\{f(a), f(b)\}$ is an edge of G_2. Next, we can check that G_1 and G_2 have the same number of edges.

Example. Are these two graphs isomorphic?

Definition. *A property preserved by isomorphism of two graphs is called a graph invariant.*

Ex. Num. of vertices, num. of edges, degree sequence, cycle of q vertices, path of q vertices, Hamilton Cycle (cycle through all vertices), matching of size q

Remark. If G_1 has a certain graph invariant, and G_2 doesn't, then they are not isomorphic. But, we don't know a long enough list of invariants so that if all graphs agree on invariants, they are isomorphic.

The only algorithm we know that is always guarenteed to correctly determine whether or not two graphs are ismorphic is to check all $n!$ bijections from V_1 to V_2.

10.4 Connectivity

Definition. *A walk in a graph is a sequence of vertices v_i and edges v_j of the form:*

$$v_0, \{v_0, v_1\}, v_1, \{v_1, v_2\}, v_2, ..., v_{k-1}, \{v_{k-1}, v_k\}, v_k$$

In a walk, vertices and edges may be repeated. Sometimes it is referred to as a walk from v_0 to v_k. The length of a walk is the number of edges, not distinct edges.

A walk is called a circuit if it starts and ends at the same vertex. The circuit is said to pass through $v_0, ..., v_k$ and traverse the edges $\{v_0, v_1\}, ..., \{v_{k-1}, v_k\}$.

A walk is called a trail its edges are distinct.

A walk is called a path its vertices are distinct.

A circuit is called a cycle if $v_0, ..., v_{k-1}$ are distinct and $v_0 = v_k$.

In simple graphs, a walk can be denoted just by listing its vertices $v_0, ..., v_k$.

Definition. *A undirected graph is called connected for every distinct pair of vertices, a and b, there is a walk from a to b.*

Theorem. *In every connected graph, for every pair of distinct vertices, a and b, there is a path from a to b.*

Proof. *Assume G, is a connected graph that a and b are distinct vertices of G. By the definition of connected, there is a way, W, from a to b.*

Let W^{+} be a walk from a to b of shortest length.

If W^{*} is not a path, thn there is a repeated vertex, say $v_l = v_m$.

Ex. $W^{*} = v_0 = a, \{v_0, v_1\}, v_1, ...v_l, \{v_l, v_m\}, v_{l+1}, \cancel{...}, \{v_{m-1}, v_m\}, v_m, ..., v_k = b$

Then $W^{**} = v_0 = a, \{v_0, v_1\}, v_1, ..., v_l, \{v_m, v_{m+1}\}, ..., v_k = b$

This contradicts our choice of W^{*} as the shortest walk from a to b. So it is not true that W^{*} is not a path - that W^{*} is a path. □

Definition. *A connected (component) of a graph, G, is a connected subgraph of G which is not a proper subgraph of G; that is, a connected component of G is a maximal connected subgraph of G.*

Convention. What is the difference between *maximal* and *maximum*?

Answer. *Maximum* always refers to size and means largest. *Maximal* means that "with respect to inclusion." A set is maximal with a certain property if you can't add to it and still have the property.

Example. Maximal or maximum complete subgraphs of a graph.

Question. *Does a connected component need to be an induced subgraph?*

Answer. *YES. If a and b were in the "component" and $\{a, b\}$ is an edge of G, but not in the component, we could add $\{a, b\}$ and still be connected which means that what we started with wasn't maximal $\Rightarrow\Leftarrow$ since components are **maximal** connected subgraphs.*

Definition. *A walk in a directed graph, G, is a walk in the underlying undirectted graph of G. Circuits, cycles, trails and paths are defined similarly.*

A directed walk in a digraph, G, is a sequence of the form $v_0, (v_0, v_1), v_1, ..., v_{k-1}, (v_{k-1}, v_k), v_k$.

Directed circuits, cycles, trails and paths are defined similarly.

Definition. *A digraph is (weakly) connected if its underlying undirected graph is connected.*

Definition. *A digraph is called strongly connected if for every pair of distinct vertices, a and b, there is a directed path from a to b and there is a directed path from b to a.*

Definition. *The strongly connected components (or strong components) of a directed graph, G, are the maximal strongly connected subgraphs of G; that is subgraphs which are strongly connected but which are not contained in some other strongly connected subgraph.*

Definition. *A cut-vertex in a graph, G, is a vertex, v, such that if we remove it, the remaining graph (called $G - v$) has more components than G.*

Definition. *A <u>cut-edge</u> (<u>bridge</u>, <u>isthmus</u>), e, in a graph, G, is an edge such that if we remove it, the remaining graph, $G-e$, has more components than G.*

Theorem. *Let G, be a graph with adjacency matrix, A, where the i^{th} row of A corresponds to vertex v_i of G.*

The number of different walks of length r from vertex v to vertex v_j is the $(i, j)^{\text{th}}$ entry of the matrix A^r, for $r = 1, 2, \ldots$

Proof. *By simple induction on r.*

Base case: $r = 1$. A walk of length 1 corresponds to an edge. The $(i, j)^{\text{th}}$ entry of $A^1 = A$ is, by definition, the number of edges joining v_i and v_j. ($S(r)$).

IH: Assume that the theorem holds for r. (That is, assume that the statement $S(r)$:

which is: the number of vertices of different walks of length r from vertex v_i to vertex v_j is the $(i, j)^{\text{th}}$ entry of A^r is true.

Prove that the theorem holds for $r+1$ (ex. Prove that $S(r+1)$ is true). $A^{r+1} = A^r A$

The (i, j) entry of A^{r+1} is:

$$b_{i1}a_{1j} + b_{i2}a_{2j} + \ldots + b_{ik}a_{kj} + \ldots + b_{in}a_{nj}$$

where b_{ik} is the number of walks from v_i to v_k of length r and where a_{kj} is the number of edges joining v_k and v_j (by def. of adjacency matrix).

A walk of length $r+1$ from v_i to v_j consists of a walk of length r from v_i to some vertex v_k and then an edge from v_k to v_j.

By IH, the number of walks of length r from v_i to v_k is b_{ik} (the $(i, k)^{\text{th}}$ entry of A^r). So the number of walks of length $r+1$ from v_i to v_j whose second-last vertex is v_k from $b_{ik}a_{kj}$, since a_{kj} is the number of edges from v_k to v_j in G.

Summing over all possible second last vertices, v_k, we get $b_{ii}a_{ij} + \ldots + b_{in}a_{nj}$ which is the $(i, j)^{\text{th}}$ entry of A^r. □

10.5 Euler Trails/Circuits & Hamilton Paths/Cycles

Definition. *An <u>Euler trail</u> in a graph G, is a trail which traverses every edge of G exactly once (can repeat vertices).*

An <u>Euler circuit</u> is an Euler trail that begins and ends at the same vertex.

Definition. *A <u>Hamilton path</u> is a path which goes through every vertex exactly once.*

A <u>Hamilton</u> cycle which goes through every vertex and then back to the starting vertex.

Theorem. *A graph, G, with at least one edge has an Euler circuit if and only if G is connected and every vertex has even degree.*

Proof. *($\Rightarrow$)*

Assume G is a graph with at least one edge and one Euler circuit. To see that G is connected;

let a, b, $c \in V(G)$ s.t there is a trail from a to b namely, the part of the Euler circuit from an occurance of a to an occurance of b. $\therefore G$ is connected.

To see that every vertex has even degree; let s be the start vertex of the Euler circuit. For every other vertex, v, each time the circuit comes to v it also leaves v. $\therefore$ the ciruit traverses 2 new edges incident to v or a new edge and a new half loop incident to v, or two new half loops.

So for each visit to v, a count of 2 is used from $\deg(v)$. $\therefore$ Every vertex except s has even degree.

For s, the first edge of the circuit is distinct from the last - these use 2 from $\deg(s)$. Every other visit to s uses a count of 2 from $\deg(s)$.

$\therefore \deg(s)$ *is also even.*

($\Leftarrow$)

Lemma. *If G is a connected graph with at least one edge and every vertex has even degree, then G contains a loop, a digon or a cycle.*

Proof. *Do it yourself.* □

Assume G is a connected graph with at least one edge and every vertex has even degree.

Proof. *By induction on the $|E(G)|$.*

Base case: $|E(G)| = 1$ or 2. All of these graphs have an Euler circuit.

IH: Assume that for all graphs with fewer edges than G with fewer edges than G, the theorem holds; that is, assume that H is a connected graph, with fewer edges than G and H has all vertex degrees even, then H has an Euler circuit.

By the lemma, G contains a loop, digon or cycle, call it C and let s be a vertex on C. Remove the edges of C to get a new graph, G'. Each component of G' has even degree because if it lies on C, then its degree in G' is 2 smaller than than its degree in G and otherwise its degree is the same in G and in G'.

By the IH, each component of G' with at least one edge has an Euler circuit. To form an Euler circuit of G, start at s on C, travel along C until you reach a vertex, say s_1, of a component of G' with at least one edge, traverse the Euler circuit of that component ending up back at s_1, continue along C until you reach another vertex of a component of G' with at least one edge and repeat, you will end up with an Euler circuit of G. □ □

Theorem. *A graph, G, has an Euler trail but no Euler circuit iff G is connected and has two vertices of odd degree.*

Proof. *($\Rightarrow$) Obvious*

($\Leftarrow$)

Assume G is connected and has exactly two vertices, a and b, of odd degree.

Let G' be the graph obtained from G by adding an edge $\{a, b\}$. Then G' has at least one edge, is connected and all vertex degrees are equal. By the previous **Theorem**, *G' has an Euler circuit, C.*

Then $C - \{a, b\}$ is an Euler trail of G with endpoints a and b. □

Remark. *It may be easiest to see this by reordering the edges of the Euler trail, making the added edge $\{a, b\}$ the last edge of C.*

Theorem. ***Ore's Theorem***

If G is a simple graph with n vertices, where $n \geq 3$ and for every pair of non-adjacent vertices, u and v, $\deg_G(u) + \deg_G(v) \geq n$, then G has a Hamilton cylce.

Theorem. ***Dirac's Theorem*** *(Corollary of Ore's Theorem)*

If G is a simple graph with n vertices, where $n \geq 3$ such that the degree of every vertex is at least $\frac{n}{2}$ then G has a Hamilton cycle.

Proof. *Of Dirac's Theorem assuming Ore's Theorem is true.*

Let G be a simple graph with n vertices where $n \geq 3$. Assume condition of Dirac's theorem holds; ex. for every vertex v of G, $\deg(v) \geq \frac{n}{2}$.

Let u and v be any two non-adjacent vertices of G. By condition of Dirac's theorem,

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

$\therefore$ *holds.*

$\therefore$ *By Ore's theorem, G has a Hamilton cycle.* □

Proof. *Of* ***Ore's Theorem***

Assume G is a simple graph, with n vertices where $n \geq 3$ and the condition holds. Assume G does not have a Hamilton cylce.

Add edges to G to get a simple graph, H, and no Hamilton cycle but s.t if we add any edge between non-adjacent vertices of H, the new graph has a Hamilton cycle. Note that H is not K_n, since K_n has a Hamilton cycle and H doesn't. $\therefore$ H has at least two non-adjacent vertices. Say a and b are non-adjacent in H.

If for some i, $\{a, v_{i-1}\}$ and $\{b, v_i\}$ were both edges of H, $(4 \leq i \leq n)$, then H would have a Hamilton cycle: $0, v_{i-1}, v_{i-2}, ..., b, v_i, v_{i+1}, ..., v_n, a$.

H does not have a Hamilton cycle, so $\{a, v_{i-1}\}$ and $\{b, v_i\}$ can't both be in H $(4 \leq i \leq n)$. ***

Define $S = \{i \colon \{b, v_i\} \in E(H)\} \subseteq \{3, 4, ..., n\}$ and $|S| = \deg_H(b)$

Define $T = \{i \colon \{a, v_{i-1}\} \in E(H)\} \subseteq \{4, ..., n+1\}$ and $|T| = \deg_H(a)$

*** *Tells us that there is no i in both S and T; $S \cap T = \emptyset$. We have $S \cup T \subseteq \{3, 4, ..., n\}$.*

$$\deg_G(a) + \deg_G(b) \leq \deg_H(a) + \deg_H(b) = |T| + |S| = |S \cup T| \,(\text{since disjoint}) \leq n - 1$$

and a and b are not adjacent in H, $\therefore$ not adjacent in G. Thus, a contradiction has arisen from the ashes. Our assumption that G had no Hamilton cycle, was incorrect and G does have a Hamilton cycle. □

Theorem. *A digraph, G, with at least 1 edge has a* *directed Euler circuit* *$\Leftrightarrow G$ is connected and* $\deg^-(v) = \deg^+(v)$ *for all vertices v.*

Proof. *Adapt the undirected proof for undirected graphs.* □

Definition. *A* *tournament* *is a directed a loop-free graph s.t for every distinct pair of vertices u and v, there is exactly one edge (either (u, v) or (v, u)).*

A tournament of n vertices is somtimes denoted K_n^* but there can be non-isomorphic K_n^*'s.

Definition. *A* $\underline{\textit{directed Hamilton} \begin{cases} \textit{path} \\ \textit{cycle} \end{cases}}$ *is a directed* $\begin{cases} \textit{path} \\ \textit{cycle} \end{cases}$ *that*

$\begin{cases} \textit{contains each vertex exactly once} \\ \textit{contains each vertex exactly once but then goes back to the starting vertex} \end{cases}.$

Round-Robin Tournament

Each team plays each other exactly once (no ties). Represent it by a directed graph as follows:

Vertices correspond to team and edge, $\{a, b\}$ means team a beat team b.

Theorem. ***Redei's Theorem***

Every tournament has a directed Hamilton path. (You can always rank a round robin tournament from first ... last.

Proof. *Let P_m be a directed path in a tournament K_n^* $(n \geq 2)$, $m-1$ edges (m vertices), $m \geq 2$.*

P_m always exists since a single edge will work for $m=2$. Say P_m is:

$$v_1, (v_1, v_2), v_2, (v_2, v_3), ..., (v_{m-1}, v_m), v_m$$

If $m=n$, then P_m is a directed Hamilton path and we are done.

Otherwise there is some vertex, v not on P_m. We'll show how to get a longer directed path.

Because we have a tournament, K_n^*, v is joined to every other vertex.

If (v, v_1) is am edge of K_n^*, then $v, (v, v_1), P_m$ is a larger directed path. If (v_m, v) is an edge of K_n^*, then $P_m, (v_m, v), v$ is a longer directed path.

So assume these two edges are not in K_n^*; then (v, v) and (v, v_m) are edges of K_n^*. Thus, they must be the same v_i on P_m s.t (v_i, v) and (v, v_{mi}). Then $v_1, ..., v_i, v_{i+1}, ..., v_m$ is a longer directed path.

We've shown that if P_m is not a Hamilton directed path then we can get a longer directed path, so repeating this several times we will find a Hamilton directed path. $\square$

Definition. *A graph is planar if it can be drawn in the plan with edges intersecting only at vertices. "Don't cross the streams".*

A drawing of a graph G, in the plane without crossings is called a planar representation.

Definition. *Given a planar representation of a graph, G, imagine cutting along the edges of G. The pieces that are left are called regions.*

Theorem. ***Euler's Formula***

Let $G=(V, E)$ be a connected planar graph and let R be the set of regions in a planar representation of G. Then:

$$|V| - |E| + |R| = 2$$

Corollary. *Let $G=(V, E)$ be a loop-free, connected, simple, planar graph with $|E| > 2$ and let R be the set of regions in a planar representation of G. Then:*

i. $3|R| \leq 2|E|$

ii. $|E| \leq 3|V| - 6$

Proof. *Each region is bounded by 3 edges, at least. Each edge is in the boundary of two regions.*

$3|R| \leq \sum \{\text{number of edges of region } F \colon F \in R\} = 2|E|$ $\square$

Proof.

$$\begin{aligned} 3|V|-3|E|+3|R| &= 6 \\ +2|E|-3|R| &\geq 0 \\ 3|V|-|E| &\geq 6 \\ \Rightarrow |E| \leq 3|V|-6 & \end{aligned}$$

□

Corollary. *Let $G=(V,\,E)$ be a loop-free, connected, simple, planar graph with $|E|>2$ and no cycles of length 3 (bounded by at least 4 edges), and let R be the set of regions in a planar representation of G. Then:*

$$|E| \leq 2|V|-4$$

Definition. *LEt $G=(V,E)$ be a loop-free graph with at least one edge. An elementary subdivision of G is obtained by removing edge $\{u,w\}$ and adding a new vertex v and two new edge $\{u,v\}$ and $\{v,w\}$.*

Definition. *Graph G_1 and G_2 are homeomorphic if they can be obtained from the same graph by a sequence of elementary subdivisions.*

Theorem. ***Kuratowski's Theorem***

Graph G is nonplanar iff G contains a subgraph homeomorphic to $K_{3,3}$ or K_5.

10.8 Graph Colouring

Definition. *Given a planar representation $G=(V,E)$, the dual G^d is obtained as follows:*

1. *Place one vertex in each region, including the out, infinite, region.*
2. *For each edge of G, join the vertices of G^d corresponding to the regions of G on either side of the edge.*

Proposition. *If G is a connected, planar representation, then $(G^d)^d$ is isomorphic to G. May not hold if G is not connected.*

A colouring of the regions of a planar representation G, corresponds to a colouring of the vertices of G^d (which is also a planar representation) where we want adjcent vertices to get different colours.

Definition. *A colouring (or proper vertex colouring), of a simple graph is an assignment of colours to the vertices, so that adjacent vertices have different colours. (Doesn't make sense if the graph has loops).*

The Original 4-Colour Problem. Can every planar representation have its regions coloured with at most 4 colours so that regions that share an edge (boundary), get different colours? Is equivalent to: Can every simple, planar graph [have its vertices] coloured with at most 4 colours?

Definition. *The minimum number of colours needed to colour a graph G is called its chromatic number and is denoted $\chi(G)$.*

The Four Colour problem is:

Does every simple, planar reprensentation G have $\chi(G) \leq 4$?

What is $\chi(G)$ for some of our special classes of graph.

Exam timetable problem:

Finding $\chi(G)$ would tell us the minimum number of exam slots required (not counting extra constraints like no one has 3 exams in 24 hours).

Recall. If $G = (V, E)$ is a simple, connected, planar representation, then:

$$|E| \leq 3|V| = 6$$

Lemma. *If $G = (V, E)$ is a simple, connected and planar representation, then there is a vertex with degree at most 5.*

Proof. *Suppose $G = (V, E)$ is a simple, connected and planar representation , then*

$$|E| \leq 3|V| - 6.$$

Suppose further that every vertex of G has $\deg(v) \geq 6$.

$$2|E| = \sum_{v \in V} \deg(v) \geq \sum_{v \in V} 6 = 6|V|$$

So, $|E| \geq 3|V|$. Thus,

$$3|V| \leq |E| \leq 3|V| - 6$$

but $3|V|$ cannot be less than or equal to $3|V| - 6$. So our assumption is false.

$\therefore$ *There must be some vertex x with $\deg(x) < 6$. Ex. $\deg(x) \leq 5$.* □

Theorem. ***The 6-Colour Theorem***

Let G be a simple, planar representation. Then $\chi(G) \leq 6$.

Proof. *If a graph is not connected, we can colour each component, and this gives of a colouring of the whole graph with the max number of colours used in any component.*

So when talking about colouring, we only need to consider connected graphs. □

Proof. 6-Colour Theorem *(real proof)*

Base case: The theorem is true for any simple graph with at most 6 vertices. (Give each vertex its own colour).

IH: Every simple, connected planar graph with k vertices can be coloured with at most 6 colours.

Prove the theorem is true for $k + 1$ vertices.

Let $G = (V, E)$ be a simple, connected planar graph with $k + 1$ vertices. G has a vertex x with $\deg(x) \leq 5$. $G - x$ is a simple planar graph with k vertices so, by the IH it can be coloured with ≤ 6 colours.

In this colouring, the neighbours of x get a most 5 different colours. There is an unused colour that can be used for x and we get a colouring of G. □

Theorem. ***The 5-Colour Theorem***

Any simple, planar graph with at most 5 colours. Thus for any simple planar graph, G, $\chi(G) \leq 5$.

Proof. *Similar to proof of the 6-Colour Theorem. Induction on number of vertices.*

Base case: The result is obviously true for simple planar graps with ≤ 5 vertices. Give each vertex their own colour.

IH: Every simple planar graph with k vertices can be coloured with ≤ 5 colours.

Prove true for simple, planar graphs with $k+1$ vertices. Let G be a simple planar graph with $k+1$ vertices.

G contains a vertex with degree at most 5.

By the IH, $G - x$ can be coloured with ≤ 5 colours. If $\deg(x) \leq 4$, then there is some colour not used on the neighbours of x, so there is a colour for x and we are done.

Now assume $\deg(x) = 5$. Now consider a planar representation $\tilde{G}$ of G and let $v_1, \ldots, v_5$ be the neighbours of x in clockwise order in the planar representation.

If in the colouring of $G - x$, ≤ 4 colours are used for $v_1, \ldots, v_5$ then there is a colour for x and we are done. So assume v_i has colour c_i $(1 \leq i \leq 5)$.

Define H_{ij} to be a subgraph of $\tilde{G}$ induced by the vertices of colours c_i and c_j $(1 \leq i \leq j \leq 5)$.

There are two possibilities:

i. *v_1 and v_3 are not in the same component of H_{13}. Interchange the colours c_1 and c_3 on the component of H_{13} containing v_1. Now v_1 and v_3 both have colour c_3 and we have a new colouring of $G - x$. Thus we can use c_1 for x.*

ii. *v_1 and v_3 are in the same component of H_{13}. Now consider H_{24}, the subgraph of G induced by vertices of colour c_2 and c_4. There can not be a path from v_2 to v_4 in H_{24} because such a path would start inside the cycle formed by the $v_1 - v_3$ path in the H_{13}, together with x and end outside that cycle, which is not possible in our planar representation $\tilde{G}$.* □

11 Trees

11.1 Introduction to Trees

Definition. *A simple undirected graph is called a <u>tree</u> if it is connected and has no cycles.*

A <u>forest</u> is a graph whose connected components is a tree.

Theorem. *Let a and b be vertices of tree, $T = (V, E)$. Then there is a unique path in T connecting a and b.*

Lemma. *If $G = (V, E)$ is a tree and $e = \{a, b\} \in E$, then $G - e$ has exactly two components and each is a tree.*

Proof. *Let $G = (V, E)$ be a tree and $e = \{a, b\}$ be an edge of G.*

$a, e = \{a, b\}, b$ is the unique path in G connecting a and b. So, $G - e$ has no path connecting a and b, so a and b are in different components of $G - e$.

Each component of $G - e$ is a tree (since deleting an edge can't create a cycle).

Let V_1 be the set of vertices, v of G, s.t. the unique path connecting a and v in G does use $\{a, b\}$. V_1 induces a connected subgraph of $G - e$.

Let V_2 be the set of vertices, w of G, s.t. the unique path connecting a and w in G does not contain $\{a, b\}$.

V_2 also induces a connected subgraph of $G - e$ (think about why).

$$V_1 \cup V_2 = V \text{ and } V_1 \cap V_2 \neq \emptyset,$$

so $G - e$ has exactly two components. □

Theorem. *In any tree, $T = (V, E)$: $|E| = |V| - 1$.*

Proof. *By complete induction on $|E|$.*

Base case: $|E| = 0$, T must have one vertex, $\therefore |V| = 1$. If $|E| = 1$, T has two vertices and one edge between, $\therefore |V| = 2$

IH: If $T = (V, E)$ is a tree with $<k$ edges, then $|E| = |V| - 1$.

Prove for trees with k edges. Let $T = (V, E)$ be a tree with k edges and $e = \{a, b\} \in E$.

By the lemma, $T - e$ is a forest consisting of two trees. $T_1 = (V_1, E_1)$ and $T_2 = (V_2, E_2)$ and each has fewer than k edges.

So IH holds:

$$\begin{aligned}
|E_1| &= |V_1| - 1 \\
|E_2| &= |V_2| - 1 \\
|E| &= |E_1| + |E_2| + 1 \\
&= (|V_1| - 1) + (|V_2| - 1) + 1 \\
&= |V_1| + |V_2| - 1 \\
&= |V| - 1
\end{aligned}$$

□

Applications of Trees: Saturated Hydrocarbons

A hydrocarbon consists of C_5 and H_5 and is saturated iff it only has single bonds.

If a saturated hydrocarbon has n carbons and k hydrogens, how are n and k related? Think trees, where C's have degree 4 and H's have degree 1.

$$\begin{aligned}
\sum_{v \in V} \deg(v) &= 4n + 1k \\
2|E| &= 2(|V| - 1) = 2((n + k) - 1) \\
4n + k &= 2n + 2k - 2 \\
2n + 2 &= k
\end{aligned}$$

Theorem. *Every tree with at least two vertices has at least two vertices of degree 1.*

Proof. *Let $T = (V, E)$ be a tree with $|E| \geq 1$, so $|V| \geq 2$. Say, $|V| = n$. Assume that T has at most one vertex of degree 1. $\therefore n - 1$ vertices of T have degree ≥ 2.*

$$2|E| = \sum_{v \in V} \deg(v) \geq 1 + 2(n - 1) = 2n - 1$$

(at most one vertex of degree 1, all the rest have deg ≥ 2)

$$\begin{aligned} 2|E| &= 2(|V|-1) = 2n-2 \\ 2n-2 &= 2|E| \geq 2n-1 \end{aligned}$$

Thus, a contradiction. $\therefore$ Our assumption was wrong and G at least 2 vertices degree 1. □

Theorem. *Let $G = (V, E)$ be a simple graph. The following are equivalent:*

a) G is a tree

b) G is connected. Removal of any edge disconnects G into two trees

c) G contains no cycles and $|V| = |E| + 1$

d) G is connected and $|V| = |E| + 1$

e) G contains no cyles and if $\{a, b\}$ is not an edge of G, then the graph G' obtained by adding $\{a, b\}$ to G contains exactly one cycle

Proof.

a⇒b:

Assume a. By definition, G is connected. We've already proved that removal of any edge disconnects G into 2 trees. $\therefore$ b holds.

b⇒c:

Assume b. $\therefore$ G is connected. If G contains a cycle and e is an edge of the cycle, then $G - e$ is still connected. So it is true that G consists of two trees (which contradicts b). $\therefore$ G does not contain any cycles.

Since G is connected and has no cycles, G is a tree and by prev. Thm, $|V| = |E| + 1$ □

c⇒d:n

Assume c. Assume G is not connected. Since G has no cycles, G is a forest with $k \geq 2$ components. Let the component of G be $T_i = (V_i, E_i)$, where $1 \leq i \leq k$.

$$|E| = \sum_{i=1}^{k} |E_i| = \sum_{i=1}^{k} (|V_i| + 1) = \sum_{i=1}^{k} |V_i| + \sum_{i=1}^{k} 1 = |V| + k$$

$|E| = |V| + k$ which contradicts c. $\therefore$ G is connected. □

d⇒e:

Assume d. Suppose G contains a cycle C. For C, the number of vertices equal the number of edges to reach a vertex from C, we need one new edge. It follows that for G, the number of vertices $\leq$ the number of edges. ex. $|E| = |V|$. But this contradicts the assumption the assuption, $\therefore$ G does not contain any cycles.

Since G is connected, if we add an edge $\{a, b\}$ between an exisiting a and b, we will get a cycle.

So assume, when we add an edge $e = \{a, b\}$ between existing vertices, a and b, we get two cycles, C_1 and C_2. Then $C_1 \cup C_2 - \{e\}$ is a circuit and contains a cycle C which is in G. This contradicts the end of d⇒e. $\therefore$ When $\{a, b\}$ is added, one unique cycle is created. □

e⇒a:

Assume e. Assume G is not connected. If we choose a vertex a in one component and a vertex b is another component, then adding edge $\{a,b\}$ does not create a cycle, and this contradicts e.

$\therefore$ G is connected. $\square$

Theorem. *A full $n-ary$ tree with i internal vertices has $n = mi+1$ vertices.*

Proof. *Each vertex except the root is the child of some vertex (and only one vertex).*

$\therefore$ The number of vertices that are children = (# number of interval vertices)(# children our interval vertex has) $= mi$

 # vertices $= mi+1$ $\square$

Theorem. *A full $m-ary$ tree with:*

i. *n vertices has $i = \frac{n-1}{m}$ interval vertices and $l = \frac{(m-1)n+1}{m}$ leaves*

ii. *i interval vertices has $n = m_i + 1$ vertices and $l = (m-1)i+1$ leaves*

iii. *l leaves has $n = \frac{l-1}{m-1}$ vertices and $i = \frac{l-1}{m-1}$ internal vertices*

Proof.

i.

Suppose T is a full $m-ary$ tree with n vertices. By prev. Thm. $n = mi+1$

So $\frac{n-1}{m} = i$:

$$l = n - i = n - \frac{n-1}{m} = \frac{nm - n + 1}{m} = \frac{n(m-1)+1}{m}$$ $\square$

ii.

Suppose T is a full $m-ary$ tree with i internal vertices. By prev. Thm. $n = mi+1$:

$$l = n - i = (mi+1) - i = mi - i + 1 = (m-1)i + 1$$

iii.

Suppose T is a full $m-ary$ tree with l leaves

$$\begin{aligned} & n = i + l \\ \rightarrow\ & n = mi + 1 \\ \rightarrow\ & i = n - l \\ \rightarrow\ & i = \frac{n-1}{m} \end{aligned}$$

...

Definition. *The level of a vertex v in a rooted tree is the length of the unique directed path from the root to v. The level of the root is 0. The height of the tree is the maximum level: the length of the longest-directed path from the root to any vertex.*

Definition. *A root $m-ary$ tree is called balanced if all the leaves are at level equal to the height h or equal level equal to $h-1$.*

Theorem. *There are at most m^h leaves in an $m-$ ary tree of height h.*

Proof. *(By induction on the height of h)*

Base case:

$$\begin{aligned} h=0 \quad & T \text{ is} \cdot \quad \#\text{leaves} = 1 = m^0 \\ h=1 \quad & .... \quad \#\text{leaves} \le m = m^1 \end{aligned}$$

IH: Assume the statement is true, the height is h. Prove the statement for height $h+1$.

Any $m-$ ary tree, T', of height $h+1$ is obtained from an $m-$ ary tree, T, of height h by adding $\le m$ children to each leaf of T.

T had $\le m^h$ leaves.

$$\therefore T' \text{ has} \le m^h \cdot m = m^{h+1}$$

□

Corollary. *of prev **Thm.** If an $m-$ ary tree of height h has l leaves, then $h \ge \lceil \log_m(l) \rceil$. If the $m-$ ary tree is full and balanced, then $h = \lceil \log_m(l) \rceil$.*

Proof. *(by prev. Thm.)*

For an $m-$ ary tree of height h, $l \ge m^h$.

$\therefore$ By the definition of a logarithm, $h \ge \log_m(l)$. Then, since h is an integer, $h \ge \lceil \log_m(l) \rceil$ (do it yourself). □

11.2 Minimum (or Maximum) Weight Spanning Trees

Definition. *A spanning tree of a graph $G=(V,E)$ is a subgraph of G which is a tree and contains all vertices of G.*

A spanning forest of a graph G consists of a spanning tree of each component of G.

Given a graph with weights on its edges, the weight of a spanning tree is the sums of weights of the edges of the tree.

Definition. *For a graph $G=(V,E)$ and a subset S of V, $\Delta(S)$ is the set of edges with one end in S and the other end not in S.*

Algorithm

Kruskal's Algorithm
Input: A loop-free graph, $G=(V,E)$ with n vertices and weight w_j on the edges j.
Output: A minimum weight spanning forest

To start, $T=\emptyset$.
In general, add to T an edge of minimum weight, as long as it does not create a cycle with T.
When this can't be done anymore:

- If T has $n-1$ edges, T is a spanning tree (G is connected)
- If T has fewer than $n-1$ edges, T is a spanning forest. It can be shown that G is connected

Algorithm

Prim's Algorithm

Input: A loop-free graph, $G=(V,E)$ with n vertices and weight w_j on the edges j.
Output: A minimum weight spanning tree or a non-empty set S of vertices, $S \neq V$ with $\Delta(S)=\emptyset$
Choose some vertex r to be the starting vertex. In general, we have a set of tree vertices S. Choose an edge e in $\Delta(S)$ of minimum weight. Add e to the tree and add the end of e not already in S to S. If $\Delta(S)=\emptyset$:

- Then if $S=V$, then the chosen edges form a minimum weight spanning tree of G
- If $S \neq V$, then S is a non-empty set of vertices with $\Delta(S)=\emptyset$, so G is not connected. The chosen edges are a minimum weight spanning tree of the component of G containing r.

Proof. It works!
Let F be the set of chosen edges. F together with the vertice F, meets is clearly connected. We can see that F has no cycles as follows:
Suppose it contains a cycle, C. Let f be the last edge of C chosen. Then when f was chosen both of the ends were already in S. Contradiction! $\therefore$ F contains no cycle. So F together with S is a tree. The rest of the proof follows from the following.

Claim. Let $G=(V,E)$ be a connected graph with weights w_j on the edges j. Let E' be a connected subset of edges of some minimum weight spanning tree, T' of G. Let e^* be a minimum weight edge of $\Delta(V')$, where V' is the set of vertices met by E'. Then $E' \cup \{e^*\}$ is a subset of a minimum weight spanning tree. $\square$