Observe that directed graphs and undirected graphs are different (but similar) objects mathematically. Certain notions (such as path) are defined for both, but other notions (such as connectivity) are only defined for one.

In a digraph, the number of edges coming out of a vertex is called the *out-degree* of that vertex, and the number of edges coming in is called the *in-degree*. In an undirected graph we just talk about the *degree* of a vertex, as the number of edges which are *incident* on this vertex. By the *degree* of a graph, we usually mean the maximum degree of its vertices.

In a directed graph, each edge contributes 1 to the in-degree of a vertex and contributes one to the out-degree of each vertex, and thus we have

Observation: For a digraph $G = (V, E)$,

$$\sum_{v \in V} \textit{in-deg}(v) = \sum_{v \in V} \textit{out-deg}(v) = |E|.$$

($|E|$ means the cardinality of the set E, i.e. the number of edges).

In an undirected graph each edge contributes once to the outdegree of two different edges and thus we have

Observation: For an undirected graph $G = (V, E)$,

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

Lecture 21: More on Graphs

(Tuesday, April 14, 1998)

Read: Sections 5.4, 5.5.

Graphs: Last time we introduced the notion of a graph (undirected) and a digraph (directed). We defined vertices, edges, and the notion of degrees of vertices. Today we continue this discussion. Recall that graphs and digraphs both consist of two objects, a set of vertices and a set of edges. For graphs edges are undirected and for graphs they are directed.

Paths and Cycles: Let's concentrate on directed graphs for the moment. A *path* in a directed graph is a sequence of vertices $\langle v_0, v_1, \ldots, v_k \rangle$ such that (v_{i-1}, v_i) is an edge for $i = 1, 2, \ldots, k$. The *length* of the path is the number of edges, k. We say that w is *reachable* from u if there is a path from u to w. Note that every vertex is reachable from itself by a path that uses zero edges. A path is *simple* if all vertices (except possibly the first and last) are distinct.

A *cycle* in a digraph is a path containing at least one edge and for which $v_0 = v_k$. A cycle is *simple* if, in addition, $v_1, \ldots, v_k$ are distinct. (Note: A self-loop counts as a simple cycle of length 1).

In undirected graphs we define path and cycle exactly the same, but for a *simple cycle* we add the requirement that the cycle visit at least three distinct vertices. This is to rule out the degenerate cycle $\langle u, w, u \rangle$, which simply jumps back and forth along a single edge.

There are two interesting classes cycles. A *Hamiltonian cycle* is a cycle that visits every vertex in a graph exactly once. A *Eulerian cycle* is a cycle (not necessarily simple) that visits every edge of a graph exactly once. (By the way, this is pronounced "Oiler-ian" and not "Yooler-ian".) There are also "path" versions in which you need not return to the starting vertex.

One of the early problems which motivated interest in graph theory was the *Königsberg Bridge Problem*. This city sits on the Pregel River as is joined by 7 bridges. The question is whether it is possible to cross all 7 bridges without visiting any bridge twice. Leonard Euler showed that it is not possible, by showing that this question could be posed as a problem of whether the multi-graph shown below has an Eulerian path, and then proving necessary and sufficient conditions for a graph to have such a path.

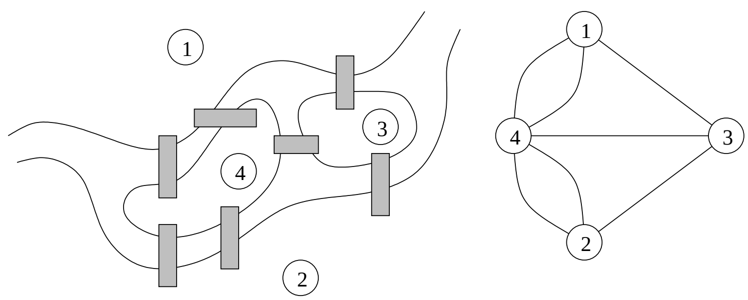

Figure 19: Bridge's at Königsberg Problem.

Euler proved that for a graph to have an Eulerian path, all but at most two of the vertices must have even degree. In this graph, all 4 of the vertices have odd degree.

Connectivity and acyclic graphs: A graph is said to be *acyclic* if it contains no simple cycles. A graph is *connected* if every vertex can reach every other vertex. An acyclic connected graph is called a *free tree* or simply *tree* for short. The term "free" is intended to emphasize the fact that the tree has no root, in contrast to a *rooted tree*, as is usually seen in data structures.

Observe that a free tree is a minimally connected graph in the sense that the removal of any causes the resulting graph to be disconnected. Furthermore, there is a unique path between any two vertices in a free tree. The addition of any edge to a free tree results in the creation of a single cycle.

The "reachability" relation is an equivalence relation on vertices, that is, it is reflexive (a vertex is reachable from itself), symmetric (if u is reachable from v, then v is reachable from u), and transitive (if u is reachable from v and v is reachable from w, then u is reachable from w). This implies that the reachability relation partitions the vertices of the graph in equivalence classes. These are called *connected components*.

A connected graph has a single connected component. An acyclic graph (which is not necessarily connected) consists of many free trees, and is called (what else?) a *forest*.

A digraph is *strongly connected* if for any two vertices u and v, u can reach v and v can reach u. (There is another type of connectivity in digraphs called *weak connectivity*, but we will not consider it.) As with connected components in graphs, strongly connectivity defines an equivalence partition on the vertices. These are called the *strongly connected components* of the digraph.

A directed graph that is acyclic is called a *DAG*, for *directed acyclic graph*. Note that it is different from a directed tree.

Isomorphism: Two graphs $G = (V, E)$ and $G' = (V', E')$ are said to be *isomorphic* if there is a bijection (that is, a 1–1 and onto) function $f : V \to V'$, such that $(u, v) \in E$ if and only if $(f(u), f(v)) \in E'$. Isomorphic graphs are essentially "equal" except that their vertices have been given different names.

Determining whether graphs are isomorphic is not as easy as it might seem at first. For example, consider the graphs in the figure. Clearly (a) and (b) seem to appear more similar to each other than to (c), but in fact looks are deceiving. Observe that in all three cases all the vertices have degree 3, so that is not much of a help. Observe there are simple cycles of length 4 in (a), but the smallest simple cycles in (b) and (c) are of length 5. This implies that (a) cannot be isomorphic to either (b) or (c). It turns

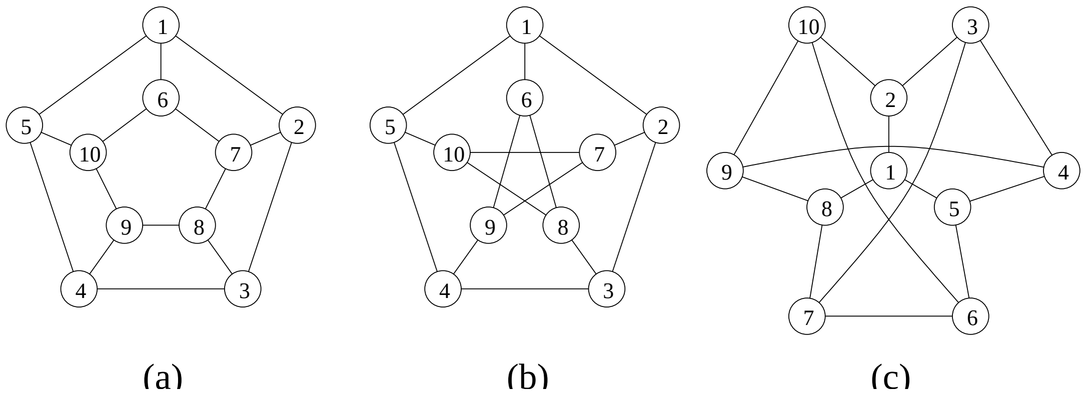

Figure 20: Graph isomorphism.

out that (b) and (c) are isomorphic. One possible isomorphism mapping is given below. The notation $(u \to v)$ means that vertex u from graph (b) is mapped to vertex v in graph (c). Check that each edge from (b) is mapped to an edge of (c).

$$\{(1 \to 1), (2 \to 2), (3 \to 3), (4 \to 7), (5 \to 8), (6 \to 5), (7 \to 10), (8 \to 4), (9 \to 6), (10 \to 9)\}.$$

Subgraphs and special graphs: A graph $G' = (V', E')$ is a *subgraph* of $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$. Given a subset $V' \subseteq V$, the subgraph *induced* by V' is the graph $G' = (V', E')$ where

$$E' = \{(u, v) \in E \mid u, v \in V'\}.$$

In other words, take all the edges of G that join pairs of vertices in V'.

An undirected graph that has the maximum possible number of edges is called a *complete graph*. Complete graphs are often denoted with the letter K. For example, K_5 is the complete graph on 5 vertices. Given a graph G, a subset of vertices $V' \subseteq V$ is said to form a *clique* if the subgraph induced by V' is complete. In other words, all the vertices of V' are adjacent to one another. A subset of vertices V' forms an *independent set* if the subgraph induced by V' has no edges. For example, in the figure below (a), the subset $\{1, 2, 4, 6\}$ is a clique, and $\{3, 4, 7, 8\}$ is an independent set.

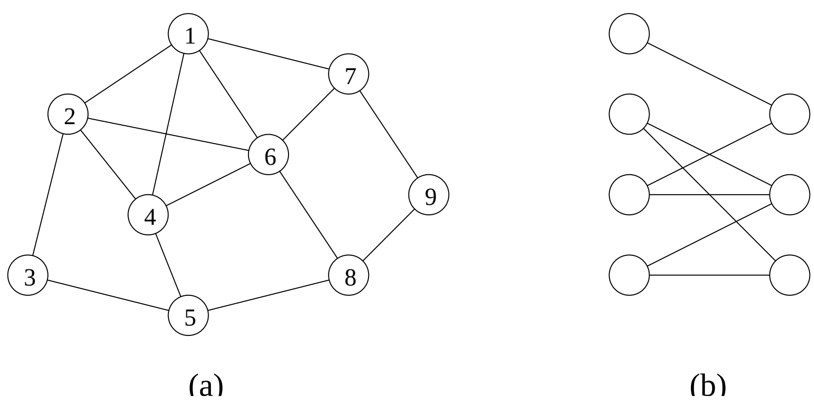

Figure 21: Cliques, Independent set, Bipartite graphs.

A *bipartite graph* is an undirected graph in which the vertices can be partitioned into two sets V_1 and V_2 such that all the edges go between a vertex in V_1 and V_2 (never within the same group). For example, the graph shown in the figure (b), is bipartite.

The *complement* of a graph $G = (V, E)$, often denoted $\bar{G}$ is a graph on the same vertex set, but in which the edge set has been complemented. The *reversal* of a directed graph, often denoted G^R is a graph on the same vertex set in which all the edge directions have been reversed. This may also be called the *transpose* and denoted G^T.

A graph is *planar* if it can be drawn on the plane such that no two edges cross over one another. Planar graphs are important special cases of graphs, since they arise in applications of geographic information systems (as subdivisions of region into smaller subregions), circuits (where wires cannot cross), solid modeling (for modeling complex surfaces as collections of small triangles). In general there may be many different ways to draw a planar graph in the plane. For example, the figure below shows two essentially different drawings of the same graph. Such a drawing is called a *planar embedding*. The *neighbors* of a vertex are the vertices that it is adjacent to. An embedding is determined by the counterclockwise cyclic ordering of the neighbors about all the vertices. For example, in the embedding on the left, the neighbors of vertex 1 in counterclockwise order are $\langle 2, 3, 4, 5\rangle$, but on the right the order is $\langle 2, 5, 4, 3\rangle$. Thus the two embeddings are different.

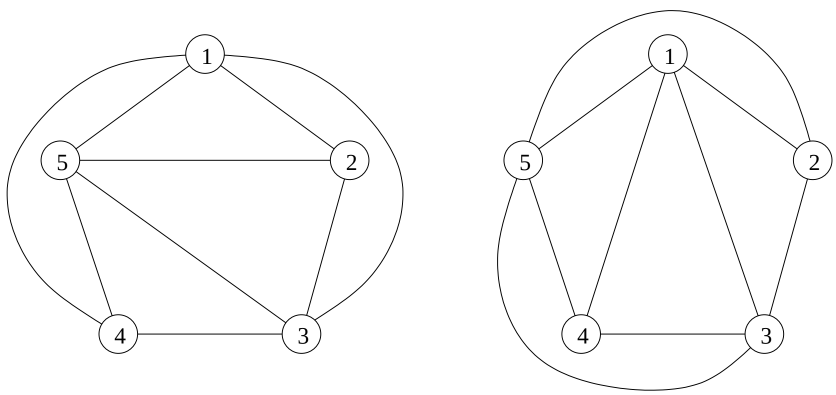

Figure 22: Planar Embeddings.

An important fact about planar embeddings of graphs is that they naturally subdivide the plane into regions, called *faces*. For example, in the figure on the left, the triangular region bounded by vertices $\langle 1, 2, 5\rangle$ is a face. There is always one face, called the *unbounded face* that surrounds the whole graph. This embedding has 6 faces (including the unbounded face). Notice that the other embedding also has 6 faces. Is it possible that two different embeddings have different numbers of faces? The answer is no. The reason stems from an important observation called *Euler's formula*, which relates the numbers of vertices, edges, and faces in a planar graph.

Euler's Formula: A connected planar embedding of a graph with V vertices, E edges, and F faces, satisfies:

$$V - E + F = 2.$$

In the examples above, both graphs have 5 vertices, and 9 edges, and so by Euler's formula they have $F = 2 - V + E = 2 - 5 + 9 = 6$ faces.

Size Issues: When referring to graphs and digraphs we will always let $n = |V|$ and $e = |E|$. (Our textbook usually uses the abuse of notation $V = |V|$ and $E = |E|$. Beware, the sometimes V is a set, and sometimes it is a number. Some authors use $m = |E|$.)

Because the running time of an algorithm will depend on the size of the graph, it is important to know how n and e relate to one another.