CS 344: Design and Analysis of Computer Algorithms **Rutgers: Fall 2019**

Graphs: Definitions and Preliminaries

October 29, 2019

Instructor: Sepehr Assadi *Scribes: Zach Langley and Harsha Srimath Tirumala*

Undirected Graphs

An undirected graph G is a set of ***vertices*** V together with a set of ***edges*** E; an edge is an (unordered) pair of distinct vertices. We abbreviate the edge $\{u, v\}$ as uv. Graphs are of two kinds: undirected graphs and directed graphs. In undirected graphs, the edge uv is the same as the edge vu, which should be clear from its definition.

We will always use n to denote the number of vertices of a graph and m to denote the number of edges. Note that an algorithm which runs in time $O(n + m)$ is *linear* in the size of the input (which may not be linear in n).

Basic Terminology

- We say that vertices u and v are the ***endpoints*** of the edge uv and that uv ***joins*** u and v.
- A vertex u is ***adjacent to*** a vertex v if u and v are joined by some edge.
- We say an edge e is ***incident*** to a vertex v if v is one of the endpoints of e.
- The ***degree*** of a vertex u, written $\deg(u)$, is the number of edges incident to it.
- The ***neighborhood*** of a vertex v, denoted $N(v)$, is the set of adjacent vertices to v – any vertex $u \in N(v)$ is also called a ***neighbor of*** v.
- A ***walk*** is a sequence P of vertices $v_1, \ldots, v_{k+1}$ such that $v_i v_{i+1}$ is an edge for all $i \in \{1, \ldots, k\}$. A ***path*** is a walk in which all vertices are distinct. We say that P is a v_1-v_k ***path***. The ***length*** of a such a walk (path) is k, the number of edges traversed.
- A ***closed walk*** is a walk that starts and ends at the same vertex. A ***cycle*** is a closed walk of length at least 3 in which all vertices are distinct, except of course at the first and last vertex.
- The ***distance*** between two vertices u and v is the length of the shortest path between them.
- Consider the equivalence relation in which two vertices are related if there is a path between them (you should verify that this is an equivalence relation). The equivalence classes of this relation are called ***components*** of G. That is, the components of G are the maximally connected subgraphs.
- A graph is ***connected*** if there is a path between any two vertices $u \in V$ and $v \in V$. Equivalently, a graph is connected if it contains exactly one component.
- A graph $G' = (V', E')$ is a ***subgraph*** of a graph $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$.
- An (undirected) graph without any cycle is a ***forest***.
- A forest with one component is a ***tree***. Alternatively, a tree is an undirected acyclic connected graph.
- A ***leaf*** of a tree is a vertex of degree 1.

- A ***spanning tree*** of a graph $G = (V, E)$ is a subgraph $T = (V, F)$ which is a tree. Note that spanning trees are on the *same vertex set*.
- A graph is ***bipartite*** if its vertex set can be partitioned into two classes A and B such that all of the edges of the graph go between A and B.
- A graph $G = (V, E)$ is k***-colorable*** if there is a color assignment $c : V \to [k]$ such that $uv \in E \implies c(u) \neq c(v)$. It is easy to verify that a graph is bipartite if and only if it is 2-colorable.
- A graph is d***-regular*** when all of its vertices have degree d.
- A ***clique*** is a set of vertices in a graph which are all connected to each other.

Directed Graphs

A directed graph (sometimes called a digraph) is a set of ***vertices*** V together with a set of ***arcs*** A; an arc is an ***ordered*** pair of distinct vertices. The difference between directed graphs and undirected graphs is precisely that arcs are ordered, whereas edges are not. We abbreviate the arc (u, v) as $\overrightarrow{uv}$. We may abuse the notation and refer to arcs as edges as well whenever it is clear from the context.

Observe that undirected graphs can be viewed as a special case of directed graphs, since for every edge uv in we can create two arcs $\overrightarrow{uv}$ and $\overrightarrow{vu}$.

Basic Terminology

- We say that the arc $a = \overrightarrow{uv}$ ***enters*** v and ***leaves*** u. We also call u the ***tail*** of a and v the ***head*** of a.
- The ***in-degree*** of a vertex is the number of arcs entering it. The ***out-degree*** of a vertex is the number of arcs leaving it.
- A ***source*** is a vertex of in-degree 0, and a ***sink*** is a vertex of out-degree 0.
- The notions of walks, paths, cycles, distance, etc., are the same as in undirected graphs, except that we must follow the orientation of the now directed edges.
- We say that a vertex u is ***reachable*** from a vertex v if there is a directed path from v to u; in that case, we may also say that v can ***reach*** u.
- A directed graph is ***strongly connected*** if for any two vertices u and v, there is a directed path between them.
- A ***directed acyclic graph*** or a ***DAG*** is a directed graph which contains no directed cycle.
- The ***transitive closure*** of a directed graph $G = (V, A)$ is a directed graph on V where there is an arc from u to v whenever G contains a path from u to v.

Properties

- The Handshaking Lemma: For any undirected graph G, the sum of the degrees of its vertices is twice its number of edges, i.e., $\sum_{v \in V} \deg(v) = 2m$. This lemma gets its name from the following observation: if you sum up the number of hands shaken by each person during the course of some event, the result will be an even number.
- For any undirected graph G, the number of vertices of odd degree is even. (This is an immediate corollary of the handshaking lemma.)
- Every directed acyclic graph (DAG) G has at least one source and one sink. Moreover, if there is *exactly* one source in G, then all vertices of G are reachable from this sink; similarly, if there is *exactly* one sink, then all vertices of G can reach this sink.

- A forest with n vertices and k components has exactly $n-k$ edges. An immediate corollary is that a tree on n vertices has $n-1$ edges.
- An undirected graph is bipartite if and only if it does not contain an odd cycle (as a subgraph).
- An undirected graph has between 0 and $\binom{n}{2}$ edges. A directed graph has between 0 and $n \cdot (n-1)$ edges.

Graph Representations

There are two common ways to represent graphs: the adjacency list and the adjacency matrix.

In the adjacency list representation, we maintain n lists, one for each vertex. The list corresponding to vertex u contains all vertices that are adjacent to u in the graph. For a directed graph, the list corresponding to vertex u contains all vertices x such that $\overrightarrow{ux} \in A$ (that is, the arcs leaving u in G).

As the name suggest, the adjacency matrix representation of G is a matrix that encodes the adjacency relationships between all pairs of vertices in the graph G. Concretely, the adjacency matrix A is the $n \times n$ matrix such that $A(i,j) = 1$ if $ij \in E$ and $A(i,j) = 0$ otherwise.

Weighted Graphs

A ***weighted graph*** is a graph $G = (V, E)$ in which every edge has an associated weight. Formally, have a function $w : E \to \mathbb{R}$. We may also speak of weighted directed graphs, which are exactly the same but for directed graphs; in the directed setting, the weight of $\overrightarrow{uv}$ need not be the same as the weight of $\overrightarrow{vu}$.