

Lecture 17

Lecturer: Debmalya Panigrahi

Scribe: Erin Taylor

1 Overview

In this lecture, we introduce directed graphs. We study connectivity in directed graphs. Finally, we introduce acyclic directed graphs, an important type of directed graphs.

2 Directed Graphs

2.1 Definitions

So far, we have worked with simple undirected graphs. Now we will introduce a new type of graph: directed graphs, also called digraphs. As usual, $G = (V, E)$ and V is a finite, non-empty set of vertices. Now, our edges will have directions. So, if $(u, v) \in E$, the edge is directed from u to v . We will assume that our graphs have no self-loops, so edges of the form (v, v) will not be included in E . Thus, $E \subseteq V \times V \setminus \{(v, v) : \forall v \in V\}$. As usual, we let $|E| = m$ and $|V| = n$. See Figure 1 for an example of a directed graph.

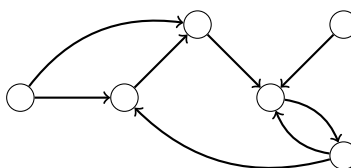


Figure 1: A directed graph.

Definition 1. The in-degree of a vertex, denoted $d_{in}(v)$, is the number of incoming edges incident on a vertex:

$$d_{in}(v) = |\{(u, v) \in E : u \in V\}|.$$

Definition 2. The out-degree of a vertex, denoted $d_{out}(v)$, is the number of outgoing edges incident on a vertex:

$$d_{out}(v) = |\{(v, u) \in E : u \in V\}|.$$

Lemma 1.

$$\sum_{v \in V} d_{in}(v) = \sum_{v \in V} d_{out}(v) = m$$

Proof. Intuitively, an edge (u, v) contributes one to the in-degree of vertex v and the out-degree of vertex u . We will prove this formally by induction on the number of edges.

Base case: If $|E| = 0$, then every vertex has in-degree and out-degree 0 and the Lemma holds.

Inductive hypothesis: Let m be an arbitrary positive integer. For any graph G' with $m' < m$ edges, assume $\sum_{v \in V} d_{in}(v) = \sum_{v \in V} d_{out}(v) = m'$.

Inductive step: Fix an arbitrary edge of G and remove it. Let this edge be $e = (u, v)$. Let $G' = (V, E')$ where $E' = E \setminus \{(u, v)\}$. G' is a graph with $m - 1$ edges. Let $d'(v)$ be the degree of $v \in V$ in G' . By our inductive hypothesis,

$$\sum_{v \in V} d_{in}(v) = \sum_{v \in V} d_{out}(v) = m - 1.$$

For all $v \in V \setminus \{u, v\}$, $d'_{out}(v) = d_{out}(v)$ and $d'_{in} = d_{in}(v)$. Additionally, the in-degree of u did not change, since we removed an outgoing edge from u . The out-degree of u decreased by 1: $d_{out}(u) = 1 + d'_{out}(u)$. Likewise, we removed an incoming edge from v , so $d_{in}(v) = 1 + d'_{in}(v)$. The out-degree of v did not change. The out-degrees:

$$\sum_{v \in V} d_{out}(v) = 1 + \sum_{v \in V} d'_{out}(v) = 1 + m - 1 = m.$$

The in-degrees:

$$\sum_{v \in V} d_{in}(v) = 1 + \sum_{v \in V} d'_{in}(v) = 1 + m - 1 = m.$$

Thus, $\sum_{v \in V} d_{in}(v) = \sum_{v \in V} d_{out}(v) = m$ as desired. Therefore, Lemma 1 holds for any directed graph. \square

Paths and walks are defined identically in undirected and directed graphs. A *walk* in a digraph is a sequence of vertices such that there exists a directed edge from each vertex to the next vertex on the walk. A *path* is a walk that does not repeat vertices. Figure 2 shows an example of a directed path.

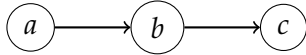


Figure 2: A valid directed path from a to c .

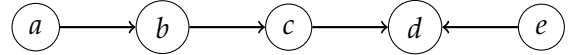


Figure 3: An example where there is no path from a to e .

Observe that paths may only go in one direction in directed graphs (a path may exist from u to v , but not from v to u). In Figure 3, see that there is no path from a to d . Next, we will define connectivity in digraphs.

2.2 Connected Components

In undirected graphs, the relation defined by connectivity was an equivalence relation, but in digraphs we have seen that connectivity is no longer as straightforward. We give two notions of connectivity for digraphs.

Definition 3. Vertices a and b are strongly connected if there is a path from a to b and a path from b to a .

Definition 4. Vertices a and b are weakly connected if there is a path from a to b . (Note that there may or may not be a path from b to a).

Consider the relation defined by weak connectivity. We say that vertices are vacuously connected to themselves, so it is reflexive. The relation is also transitive, but not symmetric. For example, in Figure 3, a, b are weakly connected, but b, a are not weakly connected.

However, the relation defined by strong connectivity is an equivalence relation on the vertices of a directed graph. We verify the three necessary properties.

- *Reflexivity*: For any vertex v , v is (vacuously) strongly connected to itself.
- *Symmetry*: This follows from the symmetry of the definition of strong connectivity. If s and t are strongly connected, then there is a path from s to t and a path from t to s , so t is also strongly connected to s .
- *Transitivity*: Suppose vertex a is strongly connected to vertex b and vertex b is strongly connected to vertex c . This implies there exists a path from a to b and a path from b to c . We can combine these two paths to obtain a walk from a to c . This implies there is a path from a to c . Similarly, there exists a path from c to b and a path from b to a . Combining these paths, we obtain a walk from c to a . This implies there is a path from c to a . Thus, a and c are strongly connected, and the relation is transitive.

Definition 5. Strongly connected components are the equivalence classes of the equivalence relation strong connectivity on the vertices of a directed graph.

Consider a directed cycle, such as the one shown in Figure 4. The vertices in a cycle are strongly connected. Thus, a directed cycle has a single strongly connected component.

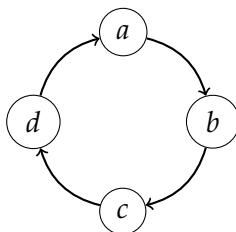


Figure 4: A directed cycle.

Now, consider a directed graph that does not have any cycle. Such a digraph is called a *directed acyclic graph* (DAG). The strongly connected components in a DAG are the singleton vertices. We will prove this with the following claim.

Claim 2. In a DAG, no two vertices can be strongly connected.

Before we prove this Claim, we will prove a useful lemma.



Figure 5: Base Case of Proof of Lemma 3.

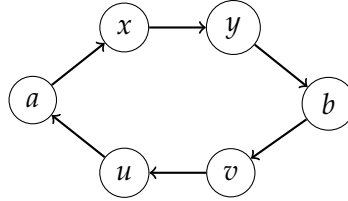


Figure 6: Disjoint paths from a to b and b to a .

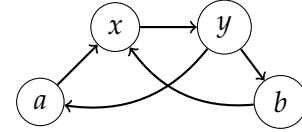


Figure 7: Non-disjoint paths from a to b and b to a .

Lemma 3. Suppose $G = (V, E)$ is a directed graph. If there is a path from vertex a to vertex b and a path from b to a , then there is a cycle containing a in G .

Proof. Suppose we have directed graph $G = (V, E)$ in which there is a path from vertex a to vertex b and a path from b to a . Let P be the path from a to b and Q be the path from b to a . We will prove Lemma 3 by induction on the sum of the lengths of the paths: $\ell = |P| + |Q|$.

Base case: The smallest possible distinct paths from a to b and b to a are two paths consisting of one edge each. In this case, $(a, b), (b, a) \in E$ and these two edges form a cycle.

Inductive hypothesis: Assume the lemma holds for paths P', Q' such that $|P'| + |Q'| < \ell$.

Inductive step: Suppose we have a graph containing paths P, Q where $\ell = |P| + |Q|$. If P and Q are disjoint, then $P + Q$ is a cycle containing a . Otherwise, paths P and Q share a vertex. Let this vertex be x . Consider the subset of path P from a to x , let this be P' . Similarly, let Q' be the subset of Q starting at x and ending at a . $|P'| + |Q'| < \ell$ since $|P'| < |P|$ and $|Q'| < |Q|$. By the inductive hypothesis, there exists a cycle containing a . Thus, in all cases the Lemma holds. \square

Now the proof of Claim 2 is straightforward.

Proof of Claim 2. We will prove the claim by contradiction. Assume two vertices in a DAG are strongly connected, let these be vertices u and v . This implies there is a path from u to v and a path from v to u . By Lemma 3, there is a cycle containing u . This contradicts that the graph is a DAG. Thus, no two vertices are strongly connected in a DAG. \square

2.3 Structure Of Directed Graphs

Strongly connected components and DAGs are useful for describing the structure of any directed graph. Consider ‘contracting’ each strongly connected component of a graph into a single vertex, and add an edge from one component to another component if there is an edge from some vertex in the first component to some vertex in the second. The result of this procedure is another directed graph, and it is acyclic. Intuitively, if there was a cycle containing multiple strongly connected components then they could be merged into a single strongly connected component. See Figure 8 for an example. Thus, any directed graph is a DAG on its strongly connected components. We will state and prove this theorem formally.

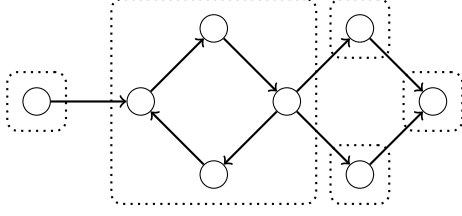


Figure 8: A directed graph and its strongly connected components (each in a dashed box).

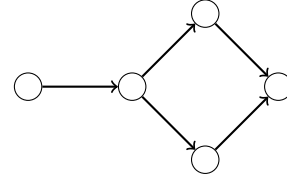


Figure 9: The component graph of the graph in Figure 8.

Definition 6. Suppose a directed graph $G = (V, E)$ has strongly connected components C_1, C_2, \dots, C_k . The component graph of G is $G^{\text{SCC}} = (V^{\text{SCC}}, E^{\text{SCC}})$ where $V^{\text{SCC}} = \{1, \dots, k\}$ and $E^{\text{SCC}} = \{(i, j) : \exists(u, v) \in E \text{ s.t. } u \in C_i, v \in C_j\}$.

Theorem 4. Every component graph is a DAG.

Proof. We will prove this theorem by contradiction. Suppose there exists a component graph containing a cycle. Let $G = (V, E)$ be the underlying directed graph, and G^{SCC} be the component graph of G . Suppose $C = c_1, \dots, c_k$ are vertices in G^{SCC} forming a cycle. Let C_1, \dots, C_k be the corresponding strongly connected components of G .

Since C is a cycle, there must be an edge between components C_i and C_{i+1} in the cycle (for $i = 1, \dots, k-1$). Let such an edge be $(out_i, in_{i+1}) \in E$ for $i \in 1, \dots, k-1$. Intuitively, out_i is the vertex in C_i with an outgoing edge to C_{i+1} . Similarly, in_i is the incoming edge from the previous component in the cycle. Also, there is an edge from C_k to C_1 , so $(out_k, in_1) \in E$. Within any component, there is a path from in_i to out_i , because all vertices in this component are strongly connected. Thus, for any two components in a cycle, there is a path from the outgoing vertex of the first component to the incoming vertex of the second component.

Suppose C_i, C_j are two components in this cycle. Let $u, v \in G$ be vertices such that $u \in C_i$ and $v \in C_j$. u, v are in different strongly connected components in G , so there cannot be paths both from u to v and from v to u in G .

Since c_i and c_j are vertices in a cycle, there exists a path from some the outgoing vertex of C_i to the incoming vertex of C_j by the previous argument. C_i is a strongly connected component, so there must exist a path from u to out_i . Similarly, C_j is a strongly connected component, so there must exist a path from in_j to v . By concatenating these three paths, we have found a path in G from u to v .

We can similarly find a path from v to u . Since c_i and c_j are in a cycle, there exists a path from some vertex $out_j \in C_j$ to some vertex in $in_i \in C_i$ by the earlier argument. Since C_j is a strongly connected component, there exists a path from v to c_j . Similarly, since C_i is a strongly connected component, there exists a path from c_i to u . By concatenating these three paths, we have found a path in G from v to u .

We have showed there is a path from u to v and from v to u in G . However, u and v were assumed to be in different strongly connected components. This is a contradiction. \square

3 Summary

In this lecture, we introduced directed graphs and explored the differences between directed and undirected graphs. We then studied connectivity in digraphs, both weak and strong. Finally, we showed that connectivity allows us to make generalizations about the structure of all directed graphs.