Lecture 18 : Digraphs

Instructor: Dieter van Melkebeek Scribe: Dalibor Zelený

DRAFT

Last time we discussed relations and used graphs and digraphs to represent them. Afterwards, we started focusing on directed graphs, and we will finish that discussion today. After that we will start looking at graphs whose edges don't have a direction.

18.1 Directed Graphs

Recall that a digraph G is a pair of sets V and E, where V is the set of vertices and E is a multiset of edges. The edges are elements of the Cartesian product $V \times V$. If the edge multiset E is actually a set, we say G is *simple*, and in this case G represents some relation.

18.1.1 Transitive Closure of a Directed Graph

Last time we defined what we mean by a path and a cycle in a directed graph. The notion of a path also induces a relation R on V. In particular, uRv if there is a path from u to v in G. This relation is called the *transitive closure* of G.

Transitive closure is reflexive. This is because a path from $v \in V$ to itself is one where we start at v and stay put. This path has length zero and doesn't require any edges, which means it's present in every graph, even in one without any edges.

Transitive closure is not symmetric. The simplest counterexample is the graph with 2 vertices u and v and one edge (u, v). There is a path from u to v, but there is no path from v to u. We will discuss graphs for which the transitive closure is symmetric in the second part of this lecture.

Transitive closure is, as the name suggests, transitive. Indeed, let $u, v, w \in V$, and suppose there is a path $e_1, e_2, \ldots, e_r$ from u to v, and a path $f_1, f_2, \ldots, f_s$ from v to w. Then the sequence of edges $e_1, e_1, \ldots, e_r, f_1, f_2, \ldots, f_r$ is a path from u to w.

18.1.2 Directed Acyclic Graphs

A directed graph is *acyclic* if it contains no cycles. Se call such a graph a *directed acyclic graph*, or DAG for short.

Directed acyclic graphs describe many real life situations. For example, consider registering for classes. Some courses have other courses as prerequisites. We can model this relation on the set of courses using a directed graph. Each course has a corresponding vertex in the graph, and there is an edge from the vertex representing course c_1 to a vertex representing course c_2 if c_1 is a prerequisite for c_2. If there is a cycle in this graph, we won't be able to take any course whose corresponding vertex is on that cycle, or any course whose corresponding vertex lies on a path from some vertex in that cycle. Thus, the graph representing the "prerequisite" relation better be a directed acyclic graph.

Other examples of relations represented by directed acyclic graphs are dependencies between different pieces of software on a computer. Again, if there are circular dependencies, we would not be able to install some software.

The same thing goes for programming parts of a bigger project. We first need to code up the small parts before we can get the bigger parts to work. If there is a circular dependency, we may not be able to finish our program, or, at the very least, we will have to code up a significant portion of it before we can test its functionality (this often suggests that we designed the program poorly).

Example 18.1: Let's consider a more mundane problem: getting dressed. The graph of the "put this on before you put that on" relation is in Figure 18.1. Piece of clothing a is related to piece of clothing b if you need to put on a before putting on b. For example, you first need to put on your right sock before you put on your right shoe. This is indicated by an edge going from the vertex labeled "right sock" to the vertex labeled "right shoe" in Figure 18.1. Observe that the graph is acyclic. If it were cyclic, we would not be able to get dressed. ⊠

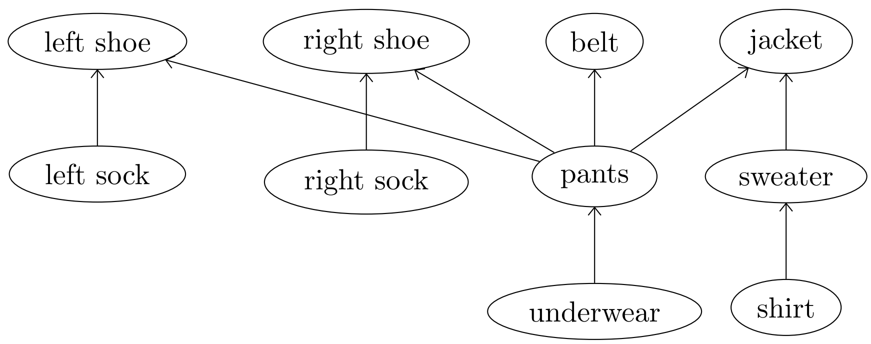


Figure 18.1: Graph of the clothing dependency relation.

A directed acyclic graph has no edges of the form (v, v) because such an edge would form a cycle of length 1. We now show that we can take any directed acyclic graph and arrange all its vertices on one line so that if $u, v \in V$ and $(u, v) \in E$, then the vertex u is "to the left" of the vertex v. In other words, all arrows in a picture representing that graph point from the left to the right. More formally, we show that every directed acyclic graph has a topological ordering.

Definition 18.1. *A* topological ordering *of a graph $G = (V, E)$ is a total order $\leq$ on V such that if $(u, v) \in E$ for two different vertices $u, v \in V$, then $u \leq v$.*

Theorem 18.2. *Every directed acyclic graph has a topological ordering.*

Before we prove Theorem 18.2, let's see an example involving the graph from Figure 18.1.

Example 18.2: Consider the rearrangement of the nodes of the graph from Figure 18.1 shown in Figure 18.2. Notice that all edges go from left to right, so we have a topological ordering of the graph from Figure 18.1.

More formally, the relation $\leq$ on the set of vertices defined by $u \leq v$ if u is to the left of v in Figure 18.2 satisfies $(u, v) \in E \Rightarrow u \leq v$ for all $u, v \in V$ such that $u \neq v$, and is a total order.

Notice that if we go through the vertices in the graph in Figure 18.2 from left to right and put on a piece of clothing when we reach its corresponding vertex, we successfully get dressed. ⊠

To prove Theorem 18.2, we need to define the following properties of vertices.

Definition 18.3. *Let $G = (V, E)$ be a graph, and let $v \in V$. The* in-degree *of v is the number of edges in E that have v as the endpoint. The* out-degree *of v is the number of edges in E that have v as the starting point.*

Example 18.3: Consider the graph in Figure 18.2. The vertex "right sock" has in-degree zero because there are no edges coming in, and has out-degree 1 because there is one edge leaving it

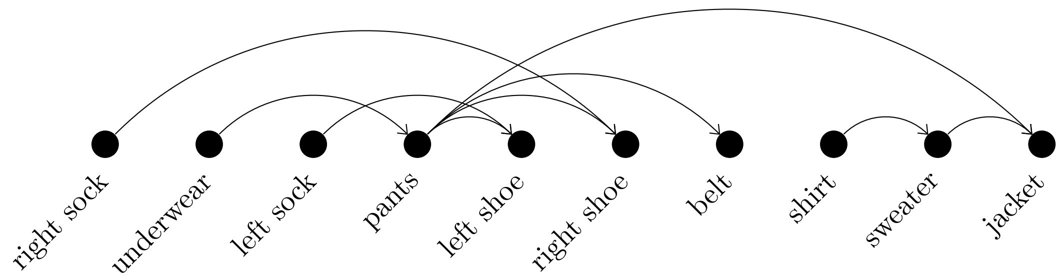


Figure 18.2: A topological ordering of the graph of clothing dependencies from Figure 18.1.

(the one that goes to vertex "right shoe"). The vertex "pants" has in-degree 1 and out-degree 4. The vertex "right shoe" has in-degree 2 and out-degree 0. ⊠

Four vertices in the directed acyclic graph of Figure 18.1 have in-degree zero: right sock, underwear, left sock, and shirt. This observation is a key step towards proving Theorem 18.2.

Claim 18.4. *Every finite nonempty directed acyclic graph has a vertex of in-degree* 0*.*

We will prove Claim 18.4 in a moment. First let's see how it helps us prove Theorem 18.2.

Proof of Theorem 18.2. We prove by induction that every directed acyclic graph on n vertices has a topological ordering.

In the base case, we have a graph with one vertex. Its topological ordering $\leq$ is the empty relation. There are no edges in this graph, so $\leq$ satisfies $(\forall u, v \in V)\ \ (u, v) \in E \Rightarrow u \leq v$ trivially.

Now suppose that all directed acyclic graphs on n vertices have a topological ordering, and consider a directed acyclic graph G on $n+1$ vertices. By Claim 18.4, G has a vertex u of in-degree zero. Thus, for all vertices $v \neq u$ we define $u \leq v$ which guarantees that the implication $(u, v) \in E \Rightarrow u \leq v$ is satisfied by all vertices $v \in V$.

Now consider the subgraph G' of G obtained by removing u and all edges going from u. The subgraph G' is a directed acyclic graph because we obtained it from a directed acyclic graph by removing one vertex and some edges, which cannot introduce any cycles. Furthermore, G' has n vertices, so it has a topological ordering $\preceq$ by the induction hypothesis. Then for two vertices v, w different from u in G, we define $v \leq w \iff v \preceq w$.

We claim that $\leq$ is a topological ordering of G.

Consider any two vertices $v, w \in V$ such that $v \leq w$ and $w \leq v$. If both v and w are vertices of G', we also have $v \preceq w$ and $w \preceq v$ by definition of $\leq$. But $\preceq$ is antisymmetric by the induction hypothesis, so $v = w$ in that case. For the case $u = v$ (or $u = w$), recall that we defined $u \leq v$ for all vertices $v \neq u$, and did not set $v \leq u$ for any $v \in V$. Thus, we see that $\leq$ is antisymmetric.

Next, pick any three vertices $v, w, x \in V$ where $v \leq w$ and $w \leq x$. If all three are vertices of G', we also have $v \leq x$ because $\leq$ is the same as $\preceq$ on G', and $\preceq$ is transitive. Now suppose one of v, w, x is u. Then $v = u$ because no vertex $v \in V$ satisfies $v \leq u$, and $w, x \neq u$ for the same reason. Now $u \leq x$ for any vertex $x \neq u$, so the implication $u \leq w \wedge w \leq x \Rightarrow u \leq x$ is true regardless of what w is. Hence, $\leq$ is transitive.

We showed that $\leq$ is antisymmetric and transitive, so $\leq$ is an order relation. Also notice that for any pair $v, w \in V$ with $v \neq w$, we either have $v \leq w$ or $w \leq v$. If both v and w are vertices of G', this follows because $\preceq$ is a total order, and if one of them is u, it follows because $u \leq v$ for any vertex $v \neq u$ by definition. Thus, $\leq$ is a total order. □

Now let's argue Claim 18.4. We give a proof using invariants.

Proof of Claim 18.4. We argue by contradiction. Suppose there are no vertices with zero in-degree in the graph $G = (V, E)$, and let $|V| = n$.

We construct a cycle in G in stages. At the beginning of stage i for $1 \leq i \leq n$, we have the following invariant: There are distinct vertices $v_1, \dots, v_i \in V$ such that there is a path $(v_i, v_{i-1}), \dots, (v_3, v_2), (v_2, v_1)$ from v_i to v_1.

When $i = 1$, we can pick any vertex as v_1. The path from v_1 to v_1 is the empty path. Thus, the invariant holds at the beginning of the first stage.

Now suppose the invariant holds at the beginning of stage $i < n$. The vertex v_i has in-degree at least 1 by assumption, so there is some vertex $u \in V$ such that $(u, v_i) \in E$. Now if $u = v_j$ for some $j \in \{1, \dots, i\}$, this edge completes the cycle $(v_i, v_{i-1}), \dots, (v_{j+1}, v_j), (v_j, v_i)$, and we get a contradiction with the fact that the graph is acyclic.

So assume $u \neq v_j$ for any $j \in \{1, \dots, i\}$ (this is possible since $i < n$). Then define $v_{i+1} = u$ and move to stage $i+1$. Observe that the vertices $v_1, \dots, v_{i+1}$ are distinct because $v_1, \dots, v_i$ are distinct by the induction hypothesis and v_{i+1} is different from all of $v_1, \dots, v_i$ by construction. Also, the path $(v_{i+1}, v_i), \dots, (v_2, v_1)$ is present in the graph because the path $(v_i, v_{i-1}), \dots, (v_2, v_1)$ is present by the induction hypothesis, and the edge (v_{i+1}, v_i) is present by construction. Thus, the invariant is maintained at the beginning of stage $i + 1$.

Now consider the situation at the beginning of stage n. There is a path $(v_n, v_{n-1}), \dots, (v_2, v_1)$, and this path goes through all vertices of G. However, v_n has in-degree 1, so there is a vertex v_j such that $(v_j, v_n) \in E$. The edge (v_j, v_n) completes the cycle $(v_n, v_{n-1}), \dots, (v_{j+1}, v_j), (v_j, v_n)$, so, again, we get a contradiction with the fact that G is acyclic.

Since all cases lead to a contradiction, we must reject the assumption that all vertices in G have nonzero in-degree. It follows that at least one vertex in G has in-degree zero. □

Now let's return to a fact we mentioned in the previous lecture. Consider the digraph $G = (V, E)$ that represents an order relation R. Recall that R is antisymmetric and transitive. Remove self loops (edges of the form (v, v)) from G to get a graph $G' = (V, E')$ that represents a relation R', and observe that R' is antireflexive in addition to being antisymmetric and transitive. Now let $u, v \in V$ be two different vertices. Because R is antisymmetric, at most one of (u, v), (v, u) is in E'. Also, there are no cycles in G' because we obtained G' from G by removing all self-loops, and there were no other cycles in G because G is the graph that corresponds to an order relation (a cycle $(v_1, v_2), \dots, (v_r, v_1)$ in G would imply $v_1 R v_j$ and $v_j R v_1$ for some $v_j \neq v_1$ by transitivity, which would mean R is not antisymmetric). Then G' is a directed acyclic graph, and has a topological ordering $\leq$.

Now consider the following relation $\tilde{R}$. We have $(u, v) \in \tilde{R}$ if either $u \leq v$ in a topological ordering of G', or if $u = v$ and uRu. Thus, if uRu, also $u\tilde{R}u$. Also, if $u \neq v$ and uRv, we have $u \leq v$ in the topological ordering of G', and, therefore, $u\tilde{R}v$. Therefore, $\tilde{R}$ is an extension of R. It is also a total order because a topological ordering is a total order. Hence, we just proved the following consequence of Theorem 18.2.

Corollary 18.5. *Every order relation R can be extended to some total order $\tilde{R}$.*

18.2 Graphs

Now let's shift our attention to graphs where edges do not have direction. Sometimes people refer to such graphs as undirected graphs. We only use the term graphs, and drop the adjective "undirected".

A graph G consists of a vertex set V and an edge multiset E. We can view such a graph as a digraph where (u, v) is an edge if and only if (v, u) is an edge. We collapse those two edges into one edge $\{u, v\} \in E$ (where the curly braces indicate that order does not matter, which is consistent with the fact that the edge doesn't have a direction). If there is a self loop, we represent the edge as $\{v\}$. With this representation, E is a multiset consisting of one- and two-element subsets of V. If an edge e has a vertex v as one of its endpoints, we say e and v are *incident*.

18.2.1 Graphs Representing Relations

The fact that edges no longer have direction means that we cannot represent every relation using a graph with only one vertex for each element of the domain. An edge $\{u, v\}$ would not tell us whether uRv, vRu, or both. However, we can use a graph to represent any symmetric relation R since in that case we can interpret the edge $\{u, v\}$ unambiguously as both uRv and vRu.

It is still possible to represent an ordinary relation using a graph. We just need more vertices. Suppose R is a relation on A. We make disjoint vertex sets V_{left} and V_{right}, each labeled with elements of A. For each $a \in A$, we have vertices $a_{\text{left}} \in V_{\text{left}}$ and $a_{\text{right}} \in V_{\text{right}}$. The vertex set of G is then $V = V_{\text{left}} \cup V_{\text{right}}$. If aRb, there is an edge between a_{left} and b_{right}, and there are no other edges. The edge $\{a_{\text{left}}, b_{\text{right}}\}$ only indicates aRb, and not bRa.

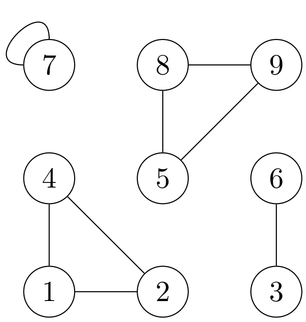
We call any graph whose vertex set can be partitioned into two sets V_{left} and V_{right} such that no edge is a subset of V_{left} or V_{right} a *bipartite graph*.

18.2.2 Transitive Closure of a Graph

Observe that in a graph, there is a path from u to v if and only if there is a path from v to u (just traverse the edges on the path from u to v in reverse order). Thus, the transitive closure of a graph is symmetric. Note that transitive closure remains reflexive and transitive, so the transitive closure of a graph is an equivalence relation.

Definition 18.6. *Let $G = (V, E)$ be a graph and consider its transitive closure. Let $V_1, V_2, \ldots, V_r$ be the equivalence classes in the partition of V induced by the transitive closure. For each equivalence class V_i, the subgraph of G with vertex set V_i and edges from G that involve vertices in V_i is called a* connected component. *A graph is called* connected *if it has at most one connected component.*

Example 18.4: Consider the graph in Figure 18.3a. It has four connected components. We list them alongside the graph in Figure 18.3b. ⊠



(a) a graph with four connected components.

No.	**Vertices**	**Edges**
1	1, 2, 4	$\{1, 2\}$, $\{1, 4\}$, $\{2, 4\}$
2	3, 6	$\{3, 6\}$
3	5, 8, 9	$\{5, 8\}$, $\{5, 9\}$, $\{8, 9\}$
4	7	$\{7\}$

(b) List of connected components of the graph in part (a).

Figure 18.3: An example illustrating the definition of a connected component.

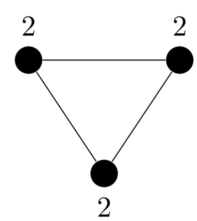
Since there is no direction to edges, we don't talk about in-degree and out-degree of a vertex in a graph. Instead, we only talk about the degree.

Definition 18.7. *Let $G = (V, E)$ be a graph, and let $v \in V$. The* degree *of v is the number of edges containing v, counting a self-loop (an edge $\{v\}$) twice. We use the notation* $\deg(v) = k$ *to denote that vertex v has degree k.*

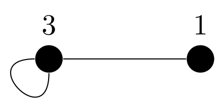
Example 18.5: The vertex 1 in the graph in Figure 18.3a has degree 2 because it belongs to two edges, namely $\{1, 2\}$ and $\{1, 4\}$. The vertex 7 also has degree 2 because it belongs to the self-loop $\{7\}$ which contributes 2 towards its degree. The only vertices of degree other than 2 are vertices 3 and 6, both of which have degree 1. ⊠

There is also a relationship between the sum of the degrees of all vertices of a graph and the number of that graph's edges. Before we state it, let's look at a few examples in Figure 18.4. The vertex labels in Figure 18.4 indicate the vertices' degrees. For each graph, we find the sum of the degrees of its vertices and the number of edges.

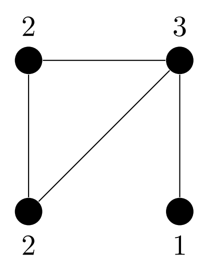
(a) $\sum_{v \in V} \deg(v) = 2$, $|E| = 1$



(b) $\sum_{v \in V} \deg(v) = 6$, $|E| = 3$



(c) $\sum_{v \in V} \deg(v) = 4$, $|E| = 2$. Recall that a self loop contributes 2 towards the degree of a vertex, which is why the degree of the vertex on the left is 3.



(d) $\sum_{v \in V} \deg(v) = 8$, $|E| = 4$

Figure 18.4: Finding the relationship between the sum of the degrees of vertices in a graph and the number of edges in that graph. In the pictures above, the vertex labels indicate the vertices' degrees.

Theorem 18.8. *In every graph $G = (V, E)$,*

$$\sum_{v \in V} \deg(v) = 2|E|. \tag{18.1}$$

Proof. Consider the contribution of edge $e \in E$ to each side of (18.1).

First consider the contribution of e to the left-hand side of (18.1). If e is a self-loop $\{v\}$ for some $v \in V$, it contributes 2 to the degree of v, so its contribution to the left-hand side of (18.1) is

2. If e is not a self-loop, it has the form $\{u, v\}$ for some $u, v \in V$, and it contributes 1 to the degree of v and 1 to the degree of u for a total contribution of 2 to the left-hand side of (18.1).

Now consider e's contribution to the right-hand side of (18.1). Because e contributes 1 to $|E|$ and the right-hand side of (18.1) is $2|E|$, e contributes 2 to the right-hand side of (18.1).

We have shown that every edge gives the same contribution to both sides of (18.1), which means that (18.1) holds. $\square$

Theorem 18.8 has the following consequence.

Corollary 18.9. *The number of vertices of odd degree in a graph is even.*

Proof. Consider the sets $V_1 = \{v \in V \mid v \text{ has odd degree}\}$ and $V_2 = \{v \in V \mid v \text{ has even degree}\}$. These sets form a partition of V, so we can write $\sum_{v \in V} \deg(v) = \sum_{v \in V_1} \deg(v) + \sum_{v \in V_2} \deg(v)$. Substituting that into (18.1) yields $\sum_{v \in V_1} \deg(v) + \sum_{v \in V_2} \deg(v) = 2|E|$, so

$$\sum_{v \in V_1} \deg(v) = 2|E| - \sum_{v \in V_2} \deg(v). \tag{18.2}$$

The right-hand side of (18.2) is even because $2|E|$ is even and $\sum_{v \in V_2} \deg(v)$ is a sum of even numbers, which is even. Thus, the left-hand side of (18.2), $\sum_{v \in V_1} \deg(v)$, is also even. Moreover, the latter sum is a sum of odd numbers. Only an even number of odd numbers can add up to an even number, so the left-hand side of (18.2) is a sum of an even number of terms, which implies that $|V_1|$ is even. $\square$