

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an **undirected** graph, so that for an edge $(i, j) \in \mathcal{E}$, we have $(i, j) = (j, i)$. We say that an edge, $(i, j) \in \mathcal{G}$, is **adjacent** to a node, k , if $k = i$ or $k = j$.

We make the following assumptions about graphs:

1. There are no edges from a node to itself, i.e., no edges of the form (i, i) .
2. There is not more than one edge between two nodes, i.e., an edge (i, j) does not appear twice.
3. We assume that each node has at least one edge

Definition 1 (path) A **path** is an ordered sequence of nodes, $i_1, i_2, \dots, i_k \in \mathcal{V}$ where no node repeats and there are edges $(i_1, i_2), (i_2, i_3), \dots, (i_{k-1}, i_k) \in \mathcal{E}$. We will generally denote paths by listing the sequence of nodes in parentheses i.e., (i_1, i_2, \dots, i_k) . Note that every arc, $(i, j) \in \mathcal{E}$ is also a two node path.

Definition 2 (cycle) A **cycle** is an ordered sequence of nodes, $(i_1, i_2, \dots, i_k, i_{k+1})$, so that (i_1, i_2, \dots, i_k) is a path, $i_1 = i_{k+1}$, and the edge (i_k, i_1) is in the graph.

Definition 3 (connected) A graph is said to be **connected** if there is a path between every pair of nodes.

Definition 4 (tour) A **tour** of a graph is a cycle that visits every node in the graph and returns to the first node listed. Note that every node is visited exactly once except the first node, which is visited twice.

Definition 5 (spanning tree) The **spanning tree** of a graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a graph $T = (\mathcal{V}, \mathcal{E}_T)$ where $\mathcal{E}_T \subseteq \mathcal{E}$ satisfying the following two properties:

- (1) T is connected; and
- (2) T has no cycles.

We note that a tree is any graph that satisfies properties (1) and (2). A tree is a spanning tree of \mathcal{G} if the tree has the same nodes as \mathcal{G} .

In Figure 1, we have a connected graph as there are paths between every pair of nodes. Also, $(1, 3, 5)$ and $(1, 2, 4, 5)$ are both paths, between nodes 1 and 5, but $(1, 3, 2, 4, 5)$ is not as $(3, 2)$ is not an edge. In Figure 1, $(1, 2, 4, 6, 3, 1)$ is a cycle but $(1, 4, 5, 3, 4)$ is not. However, $(4, 5, 3, 4)$ is a cycle. Edge sets that comprise a spanning tree include $\{(1, 2), (2, 4), (4, 6), (5, 6), (3, 5)\}$ and $\{(2, 4), (1, 4), (3, 4), (5, 4), (6, 4)\}$.

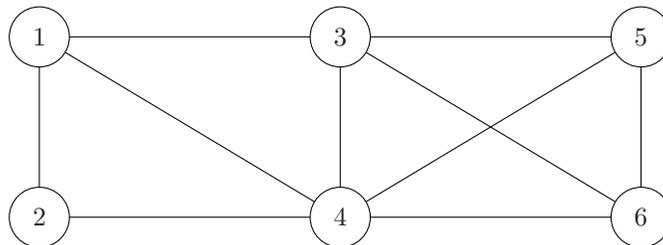


Figure 1: A graph example

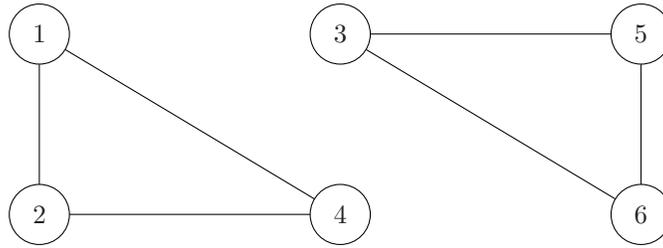


Figure 2: A second graph example with no spanning tree

In Figure 2, we have a graph that is not connected as there is no path between several nodes, including nodes 1 and 3.

Properties of trees Here are a few elementary properties of trees.

Property 1 *Every tree has at least one node with exactly one arc adjacent to it. Such a node is called a leaf.*

Proof: Suppose we are given a tree. For a contradiction, assume the tree has no leaves. If so, start at any node and color it red. Choose any arc adjacent to the red node, and color it red. Then, we repeat the following: we consider the node on the other end of the red arc. As the node is not a leaf, there are at least two arcs adjacent to it. Color one of these arcs red, and consider the node at the other end of the red arc. Color this new node red. As the new node is also not a leaf, there must be at least one arc that is not colored red. Color this arc red and repeat the following. We consider the node at the end of the arc just colored. If the node is not red, we color it red as before, and choose an arc as before. If the node is red, we have repeated a node, so the red nodes can be used to find a cycle. The process of coloring nodes red can only be done as many times as there are nodes in the graph. As there are finitely many nodes in the graph, we eventually find a cycle, a contradiction, as the graph is a tree. ■

Property 2 *For a given tree $\mathcal{T} = (\mathcal{V}, \mathcal{E}_T)$, the number of edges is one less than the number of nodes, i.e., $|\mathcal{E}_T| = |\mathcal{V}| - 1$.*

Proof: We prove the property using induction on the number of nodes in the graph. As a base case, note that the unique tree on two nodes has one edge, as it is connected and has no cycles. For the inductive case, let $n \geq 2$ be given, and assume that all trees with n nodes have exactly $n - 1$ edges. Now let a tree with $n + 1$ nodes be given. By Property 1, the tree has a leaf. Consider the subgraph that results when the leaf node and arc adjacent to it is removed from the tree. The subgraph is still connected as any path with the leaf arc ends at the leaf, and so is not on any path between non-leaf nodes. Moreover, removing the leaf node and arc cannot create any cycles. Thus, the subgraph is connected and has no cycles, i.e., it is a tree. Note that the number of nodes in the subgraph is one less than the original tree, i.e., the subgraph has n nodes. By the inductive hypothesis, the subgraph has $n - 1$ edges. But then, as we removed exactly one edge from the tree to obtain the subgraph, the tree must have n edges. ■