6. Trees

Definition 6.0.1 — Tree. A graph G is called a *tree* if it is connected and has no cycles. That is, a tree is a connected acyclic (circuitless) graph.

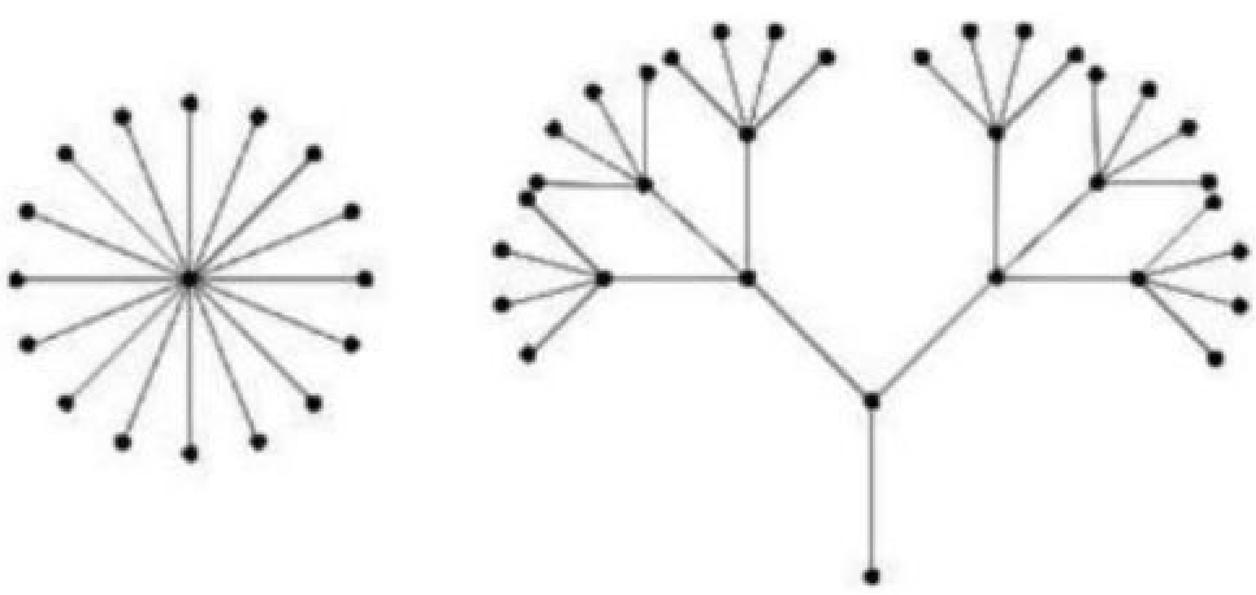

Figure 6.1: Examples of trees

Definition 6.0.2 — Tree. An acyclic graph may possibly be a disconnected graph whose components are trees. Such graphs are called **forests**.

6.1 Properties of Trees

Theorem 6.1.1 A graph is a tree if and only if there is exactly one path between every pair of its vertices.

Proof. Let G be a graph and let there be exactly one path between every pair of vertices in G. So G is connected. If G contains a cycle, say between vertices u and v, then there are

CamScanner

two distinct paths between u and v, which is a contradiction to the hypothesis. Hence, G is connected and is without cycles, therefore it is a tree.

Conversely, let G be a tree. Since G is connected, there is at least one path between every pair of vertices in G. Let there be two distinct paths, say P and P' between two vertices u and v of G. Then, the union of $P \cup P'$ contains a cycle which contradicts the fact that G is a tree. Hence, there is exactly one path between every pair of vertices of a tree. ■

Then, by Definition 3.1.11, we have the following result:

Theorem 6.1.2 All trees are geodetic graphs.

Theorem 6.1.3 A tree with n vertices has $n-1$ edges.

Proof. We prove the result by using mathematical induction on n, the number of vertices. The result is obviously true for $n = 1, 2, 3$. See illustrations in Figure 6.2.

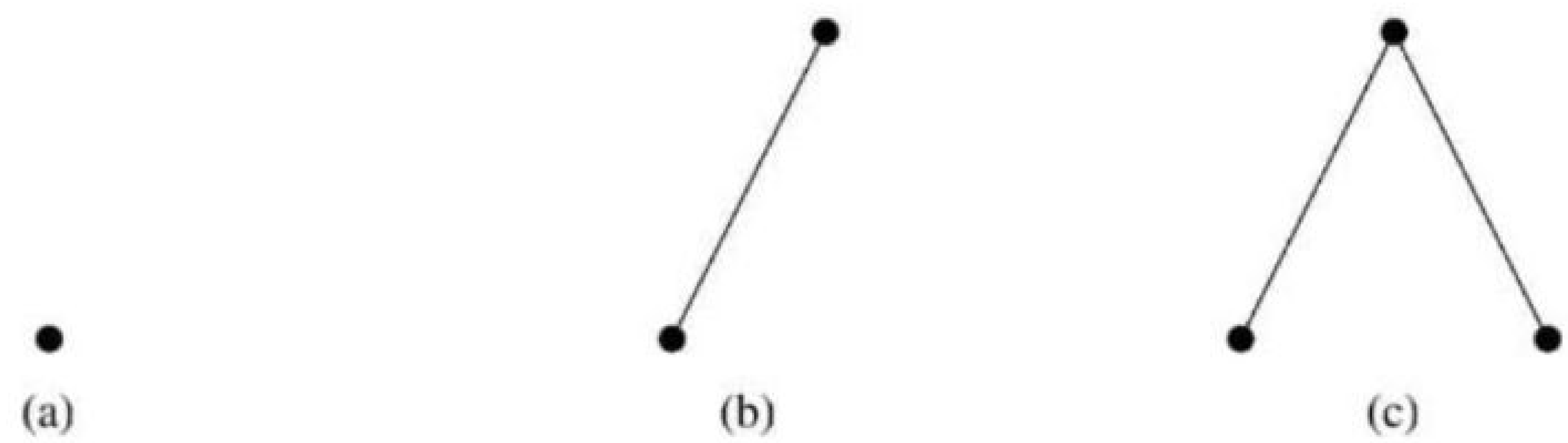

Figure 6.2: Trees with $n = 1, 2, 3$.

Let the result be true for all trees with fewer than n vertices. Let T be a tree with n vertices and let e be an edge with end vertices u and v. So, the only path between u and v is e. Therefore, deletion of e from T disconnects T.

Now, $T - e$ consists of exactly two components T_1 and T_2 say, and as there were no cycles to begin with, each component is a tree. Let n_1 and n_2 be the number of vertices in T_1 and T_2 respectively. Then, note that $n_1 + n_2 = n$. Also, $n_1 < n$ and $n_2 < n$. Thus, by induction hypothesis, the number of edges in T_1 and T_2 are respectively $n_1 - 1$ and $n_2 - 1$. Hence, the number of edges in T is $n_1 - 1 + n_2 - 1 + 1 = n_1 + n_2 - 1 = n - 1$. ■

Theorem 6.1.4 Any connected graph with n vertices and $n-1$ edges is a tree.

Proof. Let G be a connected graph with n vertices and $n-1$ edges. We show that G contains no cycles. Assume to the contrary that G contains cycles. Remove an edge from a cycle so that the resulting graph is again connected. Continue this process of removing one edge from one cycle at a time till the resulting graph H is a tree. As H has n vertices, so the number of edges in H is $n-1$. Now, the number of edges in G is greater than the number of edges in H. That is, $n-1 > n-1$, which is not possible. Hence, G has no cycles and therefore is a tree. ■

CamScanner

Theorem 6.1.5 Every edge of a tree is a cut-edge of G.

Proof. Since a tree T is an acyclic graph, no edge of T is contained in a cycle. Therefore, by Theorem 3.3.1, every edge of T is a cut-edge. ■

A graph is said to be *minimally connected* if removal of any one edge from it disconnects the graph. Clearly, a minimally connected graph has no cycles.

The following theorem is another characterization of trees.

Theorem 6.1.6 A graph is a tree if and only if it is minimally connected.

Proof. Let the graph G be minimally connected. Then, G has no cycles and therefore is a tree. Conversely, let G be a tree. Then, G contains no cycles and deletion of any edge from G disconnects the graph. Hence, G is minimally connected. ■

Theorem 6.1.7 A graph G with n vertices, $n-1$ edges and no cycles is connected.

Proof. Let G be a graph without cycles with n vertices and $n-1$ edges. We have to prove that G is connected. Assume that G is disconnected. So G consists of two or more components and each component is also without cycles. We assume without loss of generality that G has two components, say G_1 and G_2. Add an edge e between a vertex u in G_1 and a vertex v in G_2. Since there is no path between u and v in G, adding e did not create a cycle. Thus $G \cup \{e\}$ is a connected graph (tree) of n vertices, having n edges and no cycles. This contradicts the fact that a tree with n vertices has $n-1$ edges. Hence, G is connected. ■

Theorem 6.1.8 Any tree with at least two vertices has at least two pendant vertices.

Proof. Let the number of vertices in a given tree T be n, where $(n > 1)$. So the number of edges in T is $n-1$. Therefore, the degree sum of the tree is $2(n-1)$ (by the first theorem of graph theory). This degree sum is to be divided among the n vertices. Since a tree is connected it cannot have a vertex of zero degree. Each vertex contributes at least 1 to the above sum. Thus, there must be at least two vertices of degree exactly 1. That is, every tree must have at least two pendant vertices. ■

Theorem 6.1.9 Let G be a graph on n vertices. Then, the following statements are equivalent:

(i) G is a tree.
(ii) G is connected and has $n-1$ edges.
(iii) G is acyclic (circuitless) and has $n-1$ edges.
(iv) There exists exactly one path between every pair of vertices in G.
(v) G is a minimally connected graph.

CamScanner

Proof. The equivalence of these conditions can be established using the results $(i) \implies (ii), (ii) \implies (iii), (iii) \implies (iv), (iv) \implies (v)$ and $(v) \implies (i)$.

Part-$(i) \implies (ii)$: This part states that if G is a tree on n vertices, then G is connected and has $n-1$ edges. Since G is a tree, clearly, by definition of a tree it is connected. The remaining part follows from the result that every tree on n vertices has $n-1$ vertices.

Part-$(ii) \implies (ii)$: This part states that if G is connected and has $n-1$ edges, then G is acyclic and has $n-1$ edges. Clearly, This result follows from the result that a connected graph on n vertices and $n-1$ edges is acyclic.

Part-$(iii) \implies (iv)$: This part states that if G is an acyclic graph on n vertices and has $n-1$ edges, then there exists exactly one path between every pair of vertices in G. By a previous theorem, we have an acyclic graph G on n vertices and $n-1$ edges is connected. Therefore, G is a tree. Hence, by our first theorem, there exists exactly one path between every pair of vertices in G.

Part-$(iv) \implies (v)$: This part states that if there exists exactly one path between every pair of vertices in G, then G is minimally connected. Assume that every pair of vertices in G is connected by a unique path.

Let u and v be any two vertices in G and P be the unique (u,v)-path in G. Let e be any edge in this path P. If we remove the edge from P, then there will be no (u,v)-path in $G-e$. That is, $G-e$ is disconnected. Therefore, G is minimally connected.

Part-$(v) \implies (i)$: This part states that if G is minimally connected, then G is a tree. Clearly, G is connected as it is minimally connected. Since G is minimally connected, removal of any edge makes G disconnected. That is, every edge of G is a cut edge of G. That is, no edge of G is contained in a cycle in G. Therefore, G is acyclic and hence is a tree. ■

Theorem 6.1.10 A vertex v in a tree is a cut-vertex of T if and only if $d(v) \geq 2$.

Proof. Let v be a cut-vertex of a tree T. Since, no pendant vertex of a graph can be its cut-vertex, clearly we have $d(v) \geq 2$.

Let v be a vertex of a tree T such that $d(v) \geq 2$. Then v is called an *internal vertex* (or *intermediate vertex*) of T. Since $d(v) \geq 2$, there are two at least two neighbours for v in T. Let u and w be two neighbours of v. Then, $u-v-w$ is a $(u-w)$-path in G. By Theorem-1, we have the path $u-v-w$ is the unique $(u-w)$-path in G. Therefore, $T-v$ is disconnected and u and w are in different components of T. Therefore, v is a cut-vertex of T. This completes the proof. ■

6.2 Distances in Trees

Definition 6.2.1 — Metric. A *metric* on a set A is a function $d : A \times A \to [0, \infty)$, where $[0, \infty)$ is the set of non-negative real numbers and for all $x, y, z \in A$, the following conditions are satisfied:

1. $d(x,y) \geq 0$ (non-negativity or separation axiom);
2. $d(x,y) = 0 \Leftrightarrow x = y$ (identity of indiscernibles);

CS CamScanner

3. $d(x,y) = d(y,x)$ (symmetry);
4. $d(x,z) \leq d(x,y) + d(y,z)$ (sub-additivity or triangle inequality).

Conditions 1 and 2, are together called a *positive-definite function*.

A metric is sometimes called the *distance function*.

In view of the definition of a metric, we have

Theorem 6.2.1 The distance between vertices of a connected graph is a metric.

Definition 6.2.2 — Center of a graph. A vertex in a graph G with minimum eccentricity is called the *center* of G.

Theorem 6.2.2 Every tree has either one or two centers.

Proof. The maximum distance, $\max d(v, v_i)$ from a given vertex v to any other vertex occurs only when v_i is a pendant vertex. With this observation, let T be a tree having more than two vertices. Tree T has two or more pendant vertices.

Deleting all the pendant vertices from T, the resulting graph T' is again a tree. The removal of all pendant vertices from T uniformly reduces the eccentricities of the remaining vertices (vertices in T') by one. Therefore, the centers of T are also the centers of T'. From T', we remove all pendant vertices and get another tree T''. Continuing this process, we either get a vertex, which is a center of T, or an edge whose end vertices are the two centers of T. ■

6.4 On Counting Trees

A *labelled graph* is a graph, each of whose vertices (or edges) is assigned a unique name ($v_1, v_2, v_3, \ldots$ or $A, B, C, \ldots$) or labels $(1, 2, 3, \ldots)$.

The distinct vertex labelled trees on 4 vertices are given in Figure 6.3.

The distinct unlabelled trees on 4 vertices are given in Figure 6.4.

6.5 Spanning Trees

Definition 6.5.1 — Spanning Tree. A *spanning tree* of a connected graph G is a tree containing all the vertices of G. A *spanning tree* of a graph is a maximal tree subgraph of that graph. A spanning tree of a graph G is sometimes called the *skeleton* or the *scaffold graph*.

Theorem 6.5.1 Every connected graph G has a spanning tree.

Proof. Let G be a connected graph on n vertices. Pick an arbitrary edge of G and name it e_1. If e_1 belongs to a cycle of G, then delete it from G. (Else, leave it unchanged and pick it another one). Let $G_1 = G - e_1$. Now, choose an edge e_2 of G_1. If e_2 belongs to a cycle of G_1, then remove e_2 from G_1. Proceed this step until all cycles in G are removed iteratively. Since G is a finite graph the procedure terminates after a finite number of times. At this stage, we get a subgraph T of G, none of whose edges belong to cycles. Therefore, T is a connected acyclic subgraph of G on n vertices and hence is a spanning tree of G, completing the proof. ■

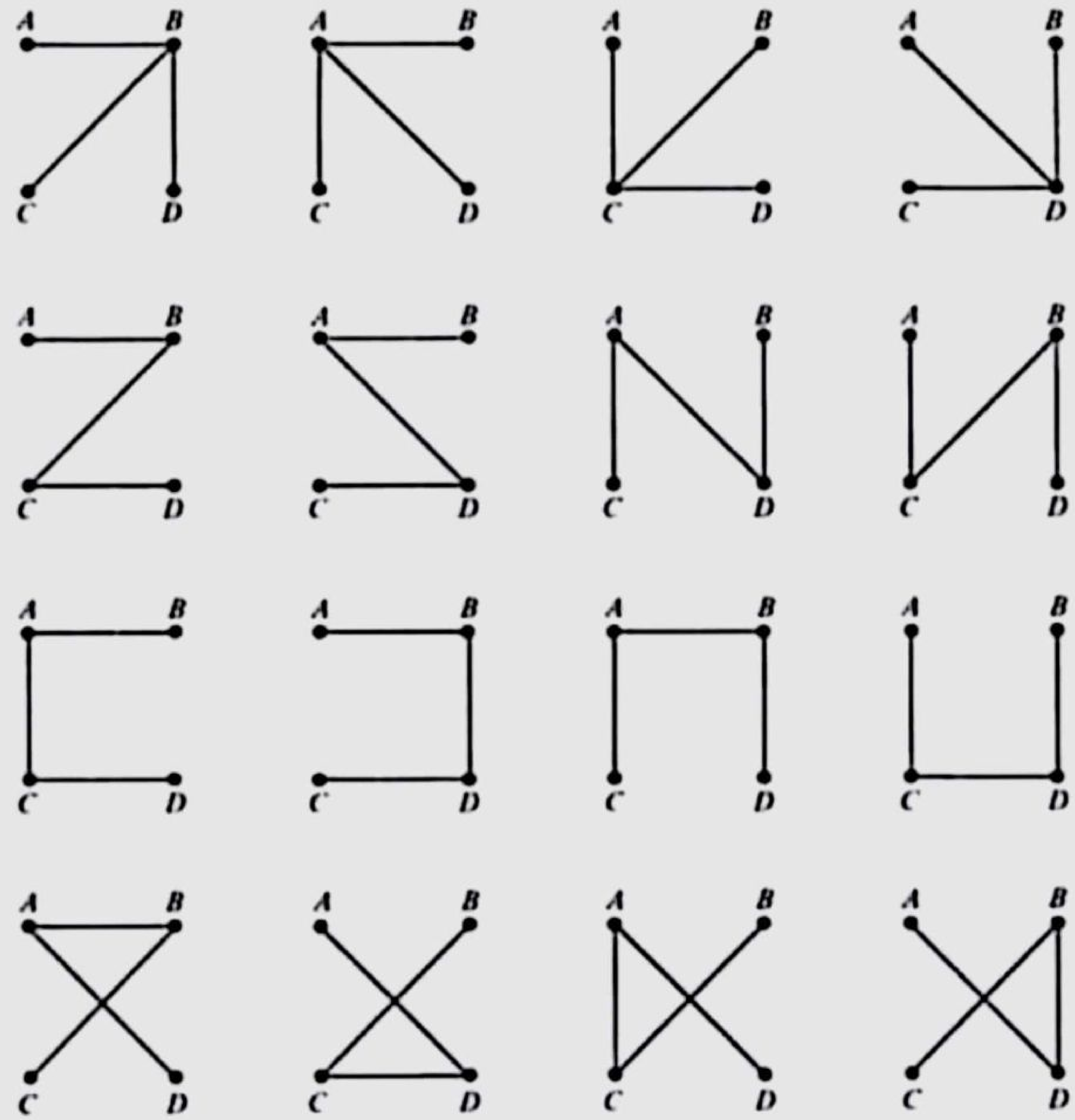

Figure 6.3: Distinct labelled trees on 4 vertices

Figure 6.4: Distinct unlabelled trees on 4 vertices