CS1231 Chapter 12

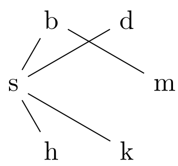
Trees

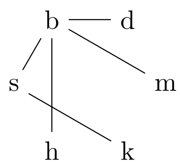
⊛ Every node is grown out from a root
⊛ This growing may involve branching
⊛ Different branches stay separate

12.1 Characterizations

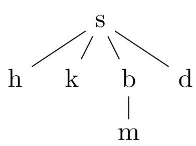
Definition 12.1.1. A *tree* is a connected acyclic undirected graph.

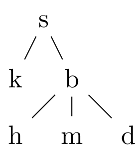
Example 12.1.2. The following are drawings of trees.





One can redraw these respectively as follows.





Proposition 12.1.3. An undirected graph with no loop is a tree if and only if between any two vertices there is exactly one path in the graph.

Proof. Being connected is by definition equivalent to having at least one path between any two vertices. For an undirected graph with no loop, being acyclic is by Theorem 11.2.5 equivalent to having at most one path between any two vertices. The proposition then follows from the definition of trees. □

Theorem 12.1.4. A connected undirected graph G is a tree if and only if removing any edge disconnects G, i.e., for every $e \in \mathrm{E}(G)$, the graph $(\mathrm{V}(G), \mathrm{E}(G) \setminus \{e\})$ is not connected.

Proof. ($\Rightarrow$) Suppose G is a tree. Take any $uv \in \mathrm{E}(G)$. Being acyclic, we know G has no loop. So $u \neq v$, and thus uv is a path between u and v in G. In view of Proposition 12.1.3, this is the only path between u and v in G. Hence in $(\mathrm{V}(G), \mathrm{E}(G) \setminus \{uv\})$ there is no path between u and v. This shows $(\mathrm{V}(G), \mathrm{E}(G) \setminus \{uv\})$ is not connected.

($\Leftarrow$) We prove this by contraposition. Suppose G is not a tree. As G is connected, we deduce that G must be cyclic. If G has a loop, say $uu \in \mathrm{E}(G)$, then $(\mathrm{V}(G), \mathrm{E}(G) \setminus \{uu\})$

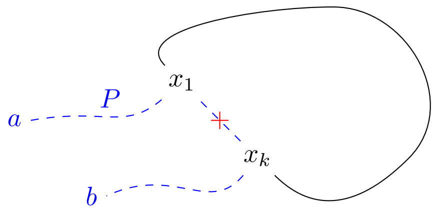


Figure 12.1: Removing an edge from a cycle

remains connected because by definition no path can have a loop. So suppose G has a cycle, say $x_1x_2\dots x_kx_1$. We show that $(\mathrm{V}(G), \mathrm{E}(G) \setminus \{x_1x_k\})$ remains connected; this will imply what we want.

Take any $a, b \in \mathrm{V}(G)$. We want a path between a and b in $(\mathrm{V}(G), \mathrm{E}(G) \setminus \{x_1x_k\})$. Use the connectedness of G to find a path $P = y_0y_1\dots y_\ell$ in G where $y_0 = a$ and $y_\ell = b$.

Case 1: suppose $x_1x_k \notin \mathrm{E}(P)$. Then P is a path between a and b in $(\mathrm{V}(G), \mathrm{E}(G) \setminus \{x_1x_k\})$.

Case 2: suppose $x_1x_k \in \mathrm{E}(P)$. Swapping a and b if needed, say $x_1 = y_r$ and $x_k = y_{r+1}$. Now in $(\mathrm{V}(G), \mathrm{E}(G) \setminus \{x_1x_k\})$,

- between a and y_r there is a path $y_0y_1\dots y_r$;
- between y_r and y_{r+1} there is a path $x_1x_2\dots x_k$;
- between y_{r+1} and b there is a path $y_{r+1}y_{r+2}\dots y_\ell$.

So two applications of Lemma 11.1.10 give us a path between a and b in $(\mathrm{V}(G), \mathrm{E}(G) \setminus \{x_1x_k\})$. □

Exercise 12.1.5. Which proof technique(s) is/are used in our proof of Theorem 12.1.4? ✎ 12a

Theorem 12.1.6. Let G be a finite tree with at least one vertex. Then $|\mathrm{E}(G)| = |\mathrm{V}(G)| - 1$.

Proof. We proceed by strong induction on the number of vertices in G.

(Base step) Let G be a tree with exactly one vertex. Note that G cannot have a loop because this would make G cyclic. So G cannot have any edge. Thus $|\mathrm{E}(G)| = 0 = 1 - 1 = |\mathrm{V}(G)| - 1$.

(Induction step) Let $k \in \mathbb{Z}^+$ such that the equation holds for all trees with at most k vertices. Consider a tree G with exactly $k+1$ vertices. As $k \geqslant 1$, we know G has at least $1 + 1 = 2$ vertices. Being a tree, the graph G must be connected. So G must have an edge, say uv, where $u \neq v$. Define $G_{uv} = (\mathrm{V}(G), \mathrm{E}(G) \setminus \{uv\})$. Then G_{uv} is not connected by Theorem 12.1.4.

Apply Proposition 11.3.6 to find connected components H_u, H_v of G_{uv} with $u \in \mathrm{V}(H_u)$ and $v \in \mathrm{V}(H_v)$. These two connected components are distinct, and they are the only connected components of G_{uv}.

On the one hand, by the definition of connected components, we know H_u, H_v are connected. On the other hand, both H_u, H_v are subgraphs of G_{uv} and thus of G. As G is acyclic, both H_u and H_v must also be acyclic. Putting these together, we conclude that H_u and H_v are both trees. Since $H_u \neq H_v$, we know $v \notin \mathrm{V}(H_u)$ and $u \notin \mathrm{V}(H_v)$ by Tutorial Exercise 11.4(a). Thus, by the Difference Rule,

$$|\mathrm{V}(H_u)| \leqslant |\mathrm{V}(G) \setminus \{v\}| = |\mathrm{V}(G)| - |\{v\}| = (k+1) - 1 = k,$$

and similarly $|\mathrm{V}(H_v)| \leqslant k$. Hence the induction hypothesis tells us $|\mathrm{E}(H_u)| = |\mathrm{V}(H_u)| - 1$ and $|\mathrm{E}(H_v)| = |\mathrm{V}(H_v)| - 1$. Therefore, by the Addition Rule,

$$\begin{aligned} |\mathrm{E}(G)| &= |\mathrm{E}(H_u)| + \overset{\text{removed edge}}{|\{uv\}|} + |\mathrm{E}(H_v)| \\ &= \big(|\mathrm{V}(H_u)| - 1\big) + 1 + \big(|\mathrm{V}(H_v)| - 1\big) = |\mathrm{V}(G)| - 1. \end{aligned}$$

This completes the induction. □

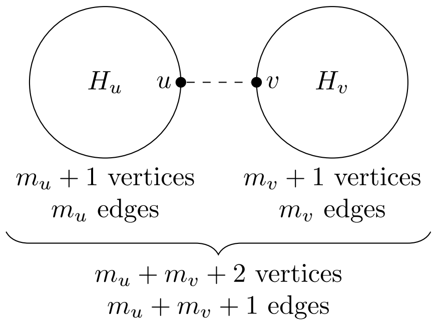


Figure 12.2: Induction step in the proof of Theorem 12.1.6

Exercise 12.1.7. Which proof technique(s) is/are used in our proof of Theorem 12.1.6? ✎ 12b

Explanation of why G has an edge uv with $u \neq v$ in the proof above (extra material). We know G has at least two vertices. Between these two distinct vertices, there is a path by connectedness. To connect two distinct vertices, this path must have at least one edge uv where $u \neq v$. ■

Explanation of why $H_u \neq H_v$ in the proof above (extra material). As $u \neq v$, we know uv is a path between u and v in G. In view of Proposition 12.1.3, this is the only path between u and v in G because G is a tree. Hence in G_{uv} there is no path between u and v. So $H_u \neq H_v$ by Theorem 11.3.7. ■

Explanation of why H_u, H_v are the only connected components of G_{uv} in the proof above (extra material). Take any connected component H of G_{uv}. As G_{uv} has at least two vertices, namely u, v, we know from Tutorial Exercise 11.4(a) that $\mathrm{V}(H) \neq \varnothing$. Pick $a \in \mathrm{V}(H)$. Use the connectedness of G to find a path $P = x_0 x_1 \ldots x_\ell$ in G where $x_0 = a$ and $x_\ell = v$.

Case 1: suppose $uv \notin \mathrm{E}(P)$. Then P is a path between a and v in G_{uv}. So there is some connected component, say $H_{a,v}$, with both a and v in it by Theorem 11.3.7. In view of Tutorial Exercise 11.4(a), we must have $H = H_{a,v} = H_v$ because $a \in \mathrm{V}(H) \cap \mathrm{V}(H_{a,v})$ and $v \in \mathrm{V}(H_{a,v}) \cap \mathrm{V}(H_v)$.

Case 2: suppose $uv \in \mathrm{E}(P)$. Then $u = x_{\ell-1}$ because $v = x_\ell$. So $x_0 x_1 \ldots x_{\ell-1}$ is a path between a and u in G_{uv}. As in the previous case, Theorem 11.3.7 then gives some connected component, say $H_{a,u}$, with both a and u in it. In view of Tutorial Exercise 11.4(a), we must have $H = H_{a,u} = H_u$ because $a \in \mathrm{V}(H) \cap \mathrm{V}(H_{a,u})$ and $u \in \mathrm{V}(H_{a,u}) \cap \mathrm{V}(H_u)$.

Thus $H = H_u$ or $H = H_v$ in all cases. ■

Theorem 12.1.8. Let G be a connected cyclic finite undirected graph. Then $|\mathrm{E}(G)| \geqslant |\mathrm{V}(G)|$.

connected cyclic → |E(G)| > |V(G)| - 1
connected acyclic → |E(G)| = |V(G)| - 1
connected → |E(G)| ≥ |V(G)| - 1

Proof. Let G be a connected cyclic finite undirected graph. Then the definition of cyclicity tells us that G contains either a loop or a cycle. Run the following procedure.

1. If G has a loop, then set $H_0 = (\{u\}, \{uu\})$ where $uu \in \mathrm{E}(G)$; else set H_0 to be any cycle in G.

   // Note that $|\mathrm{V}(H_0)| = |\mathrm{E}(H_0)| \geqslant 1$ in either case.

2. Initialize $i = 0$.

   while there are still vertices not in the subgraph

3. While $\mathrm{V}(G) \setminus \mathrm{V}(H_i) \neq \varnothing$ do:

   3.1. Use the connectedness of G to find a path $P_i = x_0^i x_1^i \ldots x_{\ell_i}^i$ in G where $x_0^i \notin \mathrm{V}(H_i)$ and $x_{\ell_i}^i \in \mathrm{V}(H_i)$.

   // Note that this implies $\ell_i \geqslant 1$.

   3.2. Set r_i to be the smallest element of $\{1, 2, \ldots, \ell_i\}$ such that $x_{r_i}^i \in \mathrm{V}(H_i)$.

   // Such r_i exists by the Well-Ordering Principle because $\ell_i \in \{1, 2, \ldots, \ell_i\}$ and $x_{\ell_i}^i \in \mathrm{V}(H_i)$.

   3.3. Set $H_{i+1} = \big(\mathrm{V}(H_i) \cup \{x_{r_i-1}^i\}, \mathrm{E}(H_i) \cup \{x_{r_i-1}^i x_{r_i}^i\}\big)$.

   // Note that $x_{r_i-1}^i \notin \mathrm{V}(H_i)$ by the smallestness of r_i. So $|\mathrm{V}(H_{i+1})| = |\mathrm{V}(H_i)| + 1 = |\mathrm{E}(H_i)| + 1 = |\mathrm{E}(H_{i+1})|$.

   3.4. Increment i to $i + 1$.

This procedure must stop after finitely many steps because G is finite. Then a run results in $H_0, H_1, \ldots, H_k$ where $k \in \mathbb{N}$. Notice $\mathrm{V}(H_k) = \mathrm{V}(G)$ as the stopping condition is reached. So, as each H_i is a subgraph of G,

$$|\mathrm{E}(G)| \geqslant |\mathrm{E}(H_k)| = |\mathrm{V}(H_k)| = |\mathrm{V}(G)|. \qquad \square$$

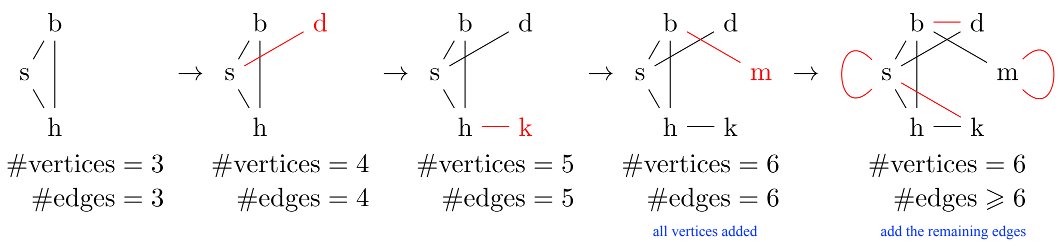


Figure 12.3: A run of the procedure in the proof of Theorem 12.1.8 from a cycle

Exercise 12.1.9. Which proof technique(s) is/are used in our proof of Theorem 12.1.8? ✎ 12c

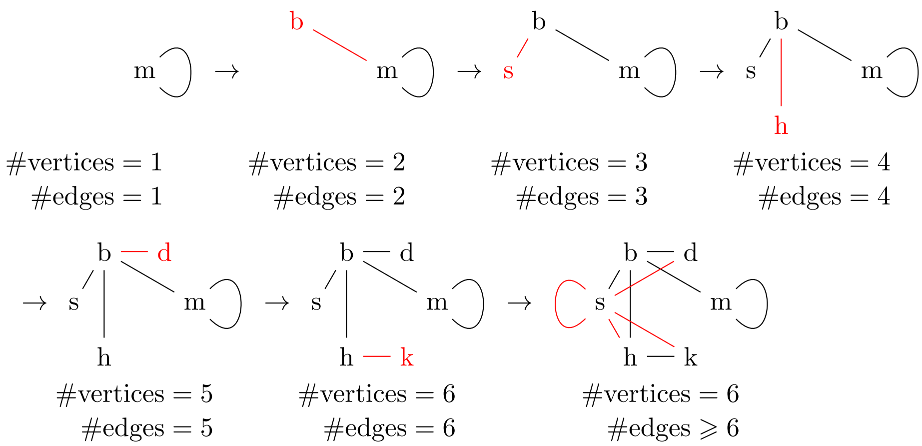


Figure 12.4: A run of the procedure in the proof of Theorem 12.1.8 from a loop

Corollary 12.1.10. The following are equivalent for all finite undirected graphs G with at least one vertex.

(i) G is a tree, i.e., it is connected and acyclic.

(ii) G has no loop, and between any two vertices there is exactly one path in G.

(iii) G is connected, and removing any edge disconnects G.

(iv) G is connected and $|\mathrm{E}(G)| = |\mathrm{V}(G)| - 1$.

(v) G is acyclic, and adding any new edge makes G cyclic.

(vi) G is acyclic and $|\mathrm{E}(G)| = |\mathrm{V}(G)| - 1$.

Proof. The equivalence of (i) and (ii) is Proposition 12.1.3. The equivalence of (i) and (iii) is Theorem 12.1.4. The equivalence of (i) and (iv) follows from Theorem 12.1.6 and Theorem 12.1.8. In a similar way, the equivalence of (i) and (vi) follows from Theorem 12.1.6 and Tutorial Exercise 12.4. The equivalence of (i) and (v) is Tutorial Exercise 12.3. □

12.2 Roots

Definition 12.2.1. A *rooted tree* is a tree with a distinguished vertex called the *root*. In a finite rooted tree,

(1) the *height* is the length of a longest path between the root and some vertex;

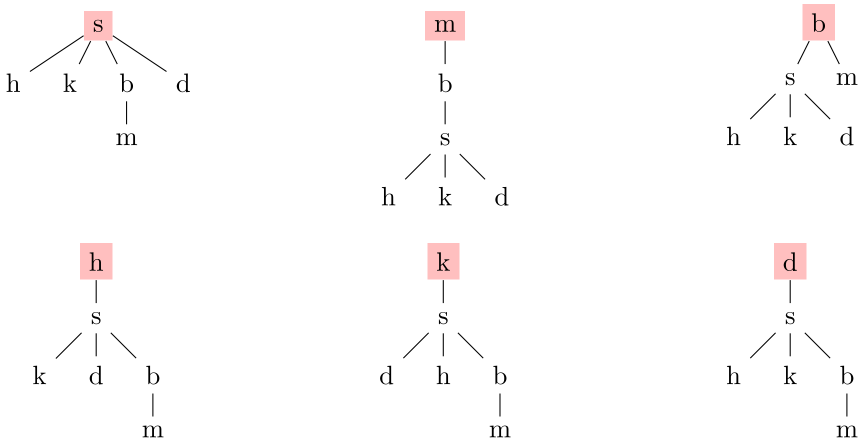
(2) every vertex y except the root has a *parent*, which is defined to be the vertex x such that the edge xy is in the path (unique by Proposition 12.1.3) between y and the root;

(3) a vertex y is said to be a *child* of a vertex x if x is the parent of y;

(4) a *terminal vertex* or a *leaf* is a vertex with no child;

(5) an *internal vertex* or a *parent* is a vertex that is not terminal.

Example 12.2.2. Here are drawings of the left tree in Example 12.1.2 with different choices of roots drawn at the top.



In the first tree where s is the root,

(1) s is a parent of h, k, b, d but not of m;

(2) h, k, b, d are children of s, but m is not.

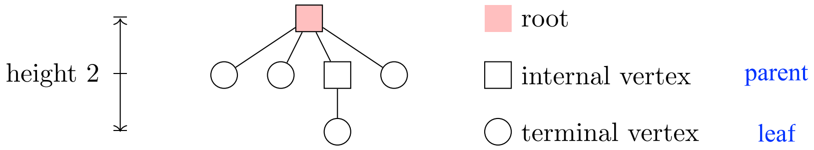


Figure 12.5: Illustrations of some terms for rooted trees

Proposition 12.2.3. Let T be a finite rooted tree of height h in which every vertex has at most two children. Then T has at most 2^h terminal vertices.

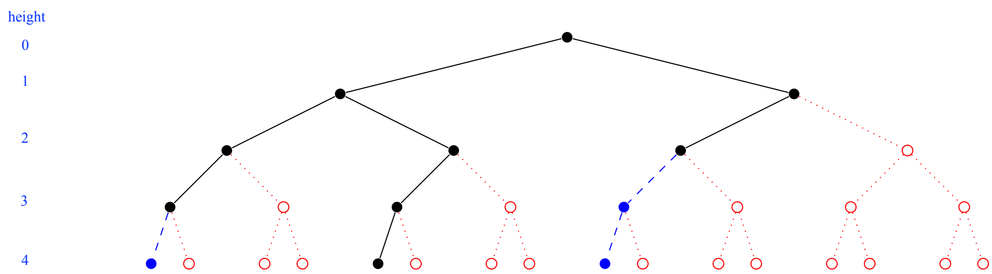


Figure 12.6: Counting terminal vertices in a rooted tree

Proof. As there is a unique path between the root and any terminal vertex in view of Proposition 12.1.3, we may equivalently count the paths $x_0x_1\dots x_\ell$ where x_0 is the root and x_ℓ is a terminal vertex. By the definition of heights, such $\ell \leqslant h$. For each $i \in \{1, 2, \dots, h\}$ such that $x_1, x_2, \dots, x_{i-1}$ are chosen,

- if x_{i-1} has no child, then x_{i-1} is already a terminal vertex, and thus there is no x_i;
- if x_{i-1} has a child, then x_{i-1} has one or two children by assumption, and x_i must be one of these children.

Here, if x_{i-1} has a child, then there must be an x_i, because otherwise $x_\ell = x_{i-1}$, which is not a terminal vertex. Therefore, by Corollary 10.2.3, there are at most

$$\underbrace{2 \times 2 \times \cdots \times 2}_{h\text{-many 2's}} = 2^h$$

such paths. ☐

Exercise 12.2.4. Which proof technique(s) is/are used in our proof of Proposition 12.2.3? 12d

Theorem 12.2.5. Let T be a finite rooted tree in which every internal vertex has exactly two children. If T has exactly t terminal vertices, then T has exactly $t-1$ internal vertices and thus exactly $2t-1$ vertices in total.

t players = terminal vertices
t-1 matches = internal vertices

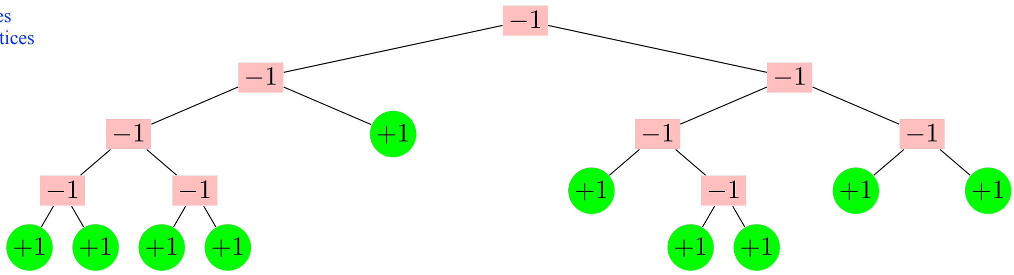


Figure 12.7: Internal versus terminal vertices (knockout tournament)

Proof. Denote by i the number of internal vertices in T. By definition, every vertex is either internal or terminal, but not both. So T has exactly $i+t$ vertices by the Addition Rule. We count in two different ways the number of elements in

$$\{(x, y) \in \mathrm{V}(G)^2 : y \text{ is a child of } x \text{ in } T\}.$$

- The variable x can only take one of i internal vertices, and then the variable y can only take one of the two children the internal vertex chosen. So there are exactly $2i$ ways to substitute objects into (x, y) by the General Multiplication Rule.
- The variable y can only take one of $i+t$ vertices in T, except the root, and then the variable x can only take the parent of this vertex. So there are exactly $i+t-1$ ways to substitute objects into (x, y) by the General Multiplication Rule.

It follows that $2i = i + t - 1$. Hence $i = t - 1$. ☐

Exercise 12.2.6. Which proof technique(s) is/are used in our proof of Theorem 12.2.5? 12e

Tutorial exercises

An asterisk (*) indicates a more challenging exercise.

12.1. Consider the graph G where

$$\mathrm{V}(G) = \{1, 2, 3, 4, 5, 6, 7, 8, 9\};$$
$$\mathrm{E}(G) = \{12, 23, 45, 67, 68, 78, 89\}.$$

(a) Draw G.

(b) Draw all subgraphs of G that is a tree and contains the vertex 7. Put the vertex 7 at the top of your drawings.

(c) Calculate the number of graphs that have the same vertices as G and are isomorphic to G. Briefly explain your calculations.

12.2. Recall the definition of isomorphism for graphs from Tutorial Exercise 11.3. Let T_1, T_2 be trees with roots r_1, r_2 respectively. These two rooted tree are *isomorphic* if there is an isomorphism f from T_1 to T_2 such that $f(r_1) = r_2$.

Let $n \in \{1, 2, 3, 4\}$.

(a) How many trees are there whose vertices are precisely $1, 2, \ldots, n$?

(b) How many trees are there with n vertices if we count isomorphic trees as one?

(c) How many rooted trees are there with n vertices if we count isomorphic rooted trees as one?

Justify your answers.

12.3. The aim of this exercise is to show that trees are precisely the acyclic graphs with a maximal set of edges.

Prove that an acyclic undirected graph G is a tree if and only if adding any new edge makes G cyclic, i.e., for all $u, v \in \mathrm{V}(G)$, if $uv \notin \mathrm{E}(G)$, then $(\mathrm{V}(G), \mathrm{E}(G) \cup \{uv\})$ is cyclic.

12.4. We proved Theorem 12.1.8 that $|\mathrm{E}(G)| > |\mathrm{V}(G)| - 1$ for all connected cyclic finite undirected graphs G. In this exercise, we prove that the opposite inequality holds for all unconnected acyclic finite undirected graphs.

Let G be an unconnected acyclic finite undirected graph. Prove that

$$|\mathrm{E}(G)| < |\mathrm{V}(G)| - 1.$$

12.5. Is it true that, for all finite undirected graphs G, if $|\mathrm{E}(G)| \geqslant |\mathrm{V}(G)| - 1$, then G is connected? Prove that your answer is correct.

12.6. Is it true that, for all finite undirected graphs G, if $|\mathrm{E}(G)| \leqslant |\mathrm{V}(G)| - 1$, then G is acyclic? Prove that your answer is correct.

12.7. (Induction corner) Using induction, prove Theorem 12.2.5, i.e.,

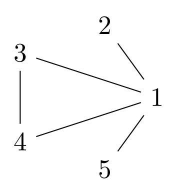
> Let T be a finite rooted tree in which every internal vertex has exactly two children. If T has exactly t terminal vertices, then T has exactly $t-1$ internal vertices.

Extra exercises

12.8. In this exercise, we look into the notion of spanning trees.

Definition. Let G be an undirected graph. A *spanning tree* of G is a subgraph T of G such that T is a tree and $\mathrm{V}(T) = \mathrm{V}(G)$.

(a) Draw all spanning trees of the graph drawn below.



(b) Let G be a finite tree. Prove that G has exactly one spanning tree.

(c)* Let G be a finite undirected graph with no loop. Prove that, if G has exactly one spanning tree, then G is a tree.

12.9. Let T be a finite tree with at least two vertices. Prove that T has at least two vertices that are each in exactly one edge.

(Hint: Consider the end-points of a path of maximum length in T. Why does such a path exist?)