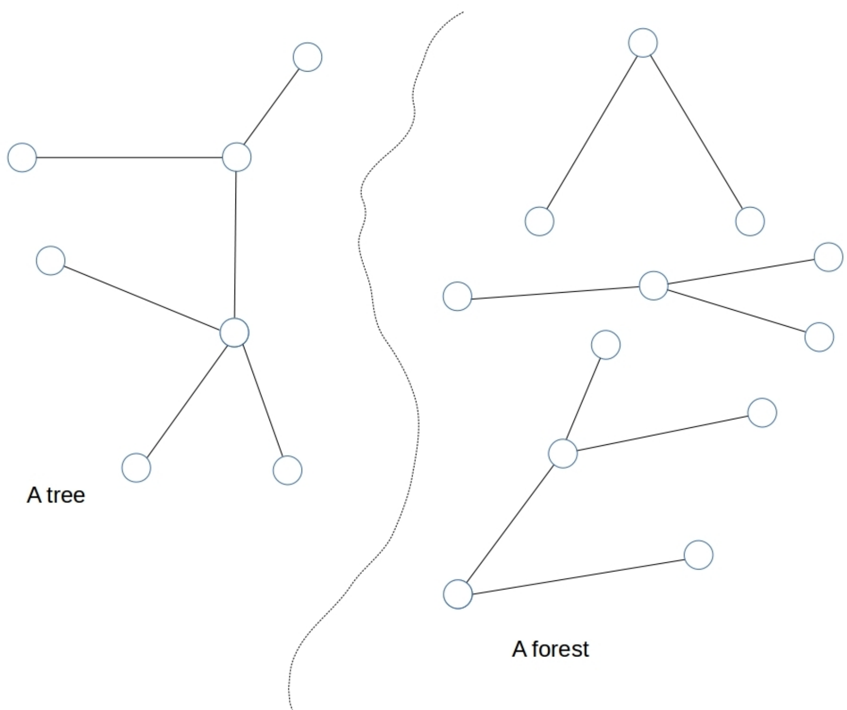
Trees

Recall that a cycle is a path that begins and ends at the same vertex (and contains no vertex more than once).

A connected graph that has no cycles is called a **tree.**

A graph that has no cycles is called a **forest.** Each connected component of a forest is a tree.



There are a number of different ways to define trees:

> A tree is a graph in which every pair of vertices is joined by exactly one path.
>
> A tree is a connected graph in which every edge is a cut-edge.
>
> A tree is a connected graph in which $|E| = |V| - 1$

We can prove that each of these is equivalent to the original definition ("a tree is a connected graph without cycles")

Leaves

A vertex of degree 1 in a tree is called a **leaf** of the tree.

Theorem: Let T be a tree with ≥ 2 vertices. Then T contains at least two leaves.

Proof: Let P be the longest path in T, and let vertices x and y be the first and last vertices in P.

Suppose x is not a leaf. Then x has degree ≥ 2, which means that x has a neighbour that is not in P (if all neighbours of x are in P, then the graph contains a cycle). Let z be this neighbour of x. Then since z is not in P, we can create a new path P' by adding the edge {xz} to P. But P' is longer than P, which is impossible since we chose P to be the longest path in T. Therefore x does **not** have any neighbour other than the one in P. Therefore x has degree 1, so it is a leaf. Now we can repeat the same argument for vertex y, so y is also a leaf.

In future courses we will spend a lot of time looking at **spanning trees of a graph** ... which are defined to be spanning subgraphs that are also trees.

Here's an important result about trees:

Theorem: Let T be a tree on ≥ 2 vertices, with leaf vertex x. Then T – x is a tree.

Proof: We need to show that T – x is connected and has no cycles. To show that T – x is connected, let a and b be any two vertices in T – x. Since they are in T – x, they are also in T, which means T contains an ab-path P. Since x has degree 1, x cannot be in P, so P is completely in T – x. Thus all pairs of vertices in T – x are connected, so T – x is connected. Showing that T – x has no cycles is trivial: T has no cycles since it is a tree, and deleting x cannot create a cycle ... so T – x has no cycles. Thus T – x is a connected graph with no cycles ... which means it is a tree.

This is important because it lets us use induction to prove many properties of trees. For example ...

Theorem: Let T be a tree on n vertices. Then |E| = |T| - 1

Proof: We will use induction on the number of vertices in the tree.

Base case : The only tree with one vertex is K_1 ... which has one vertex and zero edges, so the claim is true when the number of vertices in the tree is one.

Inductive Assumption: Assume the claim is true for all trees on $\leq$ k vertices.

Let T be a tree on k+1 vertices. Since k+1 $\geq$ 2 we know T has a leaf x, and *we know T – x is a tree!* And we know T – x has exactly k vertices!! So we know that the number of edges in T – x is exactly k – 1 (by the Inductive Assumption)!!! So T has k+1 vertices and k edges!!!! So the claim is true for T !!!!!

Therefore the claim is true for all trees !!!!!!

And now, like our distant ancestors millions of years ago, it is time to leave the safety of the trees and venture out into the wide plains ... where we will need a map.