Section III.A.4. Leaves

Our trees and forests are <u>finite</u> graphs.

Definition. A **leaf** is a vertex of degree 1.

Observation. An isolated vertex is not a leaf.

Observation. A trivial tree (a single isolated vertex) has no leaves.

Theorem. Every nontrivial tree has at least two leaves.

Proof 1. Assume there are counterexamples. Let T be a counterexample with the smallest number of edges. This number of edges is at least one since T is nontrivial.

If T has no leaves, then every vertex has degree at least 2. From earlier results, T must have a cycle. This is not allowed, so this case fails.

It follows that T has at least one leaf ℓ. (typeset this as \ell) If we delete the edge e incident with ℓ, then $T - e$ has two components. One of those components consists only of the vertex ℓ. Hence, $T - \ell$ is a tree. But then either $T - \ell$ is trivial, in which case T is isomorphic to K_2 and has two leaves, or $T - \ell$ is nontrivial. In this case, the graph $T - \ell$ is a nontrivial tree with fewer edges than T. Hence, $T - \ell$ is not a counterexample to the theorem and so $T - \ell$ has at least two leaves, x and y. Reinserting edge e and vertex ℓ has the effect of increasing the degree of one vertex by 1. At most one of vertices x or y would be affected by adding edge e and vertex ℓ in this manner. The other one is still a leaf and so T has at least two leaves.

Proof 2. Let T be a tree with $n \geq 2$ vertices. By previous results, T has $n - 1$ edges. For integers $k \geq 0$, let n_k be the number of vertices of degree k. Quick observation: $n_0 = 0$ because there are no isolated vertices. We want to show $n_1 \geq 2$. Two quick formulas:

$$\sum_{k=0}^{\infty} n_k = n_0 + n_1 + n_2 + n_3 + \cdots = n$$

$$\sum_{k=0}^{\infty} kn_k = 0n_0 + 1n_1 + 2n_2 + 3n_3 + \cdots = \text{total degree} = \text{twice edges} = 2(n-1) = 2n - 2$$

We now observe

$$\sum_{k=0}^{\infty} kn_k = \left(\sum_{k=0}^{\infty} 2n_k\right) - 2.$$

Rearranging:

$$2 = \sum_{k=0}^{\infty} 2n_k - \sum_{k=0}^{\infty} kn_k = \sum_{k=0}^{\infty} (2-k)n_k.$$

Expanding out a few terms on the right:

$$
\begin{aligned}
2 &= 2n_0 + 1n_1 + \sum_{k=2}^{\infty}(2-k)n_k \\
&= n_1 + \sum_{k=2}^{\infty}(2-k)n_k \\
&\leq n_1
\end{aligned}
$$

since every term of $\sum_{k=2}^{\infty}(2-k)n_k$ is nonpositive. Hence, $n_1 \geq 2$ and so T has at least two leaves.

This formula says a bit more: For any tree,

$$2 = n_1 - n_3 - 2n_4 - 3n_5 - 4n_6 - \cdots$$

Notice that n_2 does not appear; the only restriction on vertices of degree 2 in a tree is that the quantity $n_2 \geq 0$.

Proof 3. Let T be a nontrivial tree and let

$$P: v_0, e_1, v_1, \dots, v_{k-1}, e_k, v_k$$

be a longest path in T. Since T is nontrivial, we must have $k \geq 1$ and $v_0 \neq v_k$.

We claim v_0 and v_k are both leaves in T. Suppose v_0 is not a leaf. This would imply that there exists $v_{-1} \neq v_1$ such that v_{-1} is adjacent to v_0. If $v_{-1} = v_j$, a vertex on P, then

$$Q: v_{-1}, e_0, v_0, e_1, v_1, \dots, e_j, v_j = v_{-1}$$

is a cycle, contradicting T being a tree. If v_{-1} is not on P, then

$$Q: v_{-1}, e_0, v_0, e_1, v_1, \dots, e_k, v_k$$

is a longer path than P, contradicting P being a longest path. Thus v_{-1} cannot exist and so v_0 is a leaf.

We can apply similar reasoning to v_k, adjusting indices as needed to show it, too, must be a leaf.

Moral of the story: In any tree, the endpoints of a longest path must be leaves.

Next time, I will briefly discuss “spanning trees”; these are subgraphs of a connected graph that are trees and that contain all vertices of the graph.