CIS260-201/204–Spring 2008

Trees

Friday, March 28

Proposition: Let G be a forest with n vertices and c components. Then there are $n - c$ edges in G.
Proof: In recitation we directly derived the result. Here we will explore another way of proving this statement by induction.

Since there are two independent variables in the formula, namely, n and c, we need to do induction on both variables. The good news is we can do this one at a time, but we need to make sure that we do induction for each variable. The idea is to fix the other variable and vary the one we want to induct on. If you have a calculus background, this is like taking a partial derivative, when you treat other variables as constants and differentiate the function with respect to the variable in interest.

We will do induction on n first. First of all, note that the number of components cannot exceed the number of vertices. Otherwise we would have an "empty" component, which is not allowed.

Base Case: $(n = c)$ [Note that we treat c as a constant. Since c is the least possible integer n can take on, it becomes our base case.] In this case, there is exactly one vertex in each component. Therefore, there is no edges at all, but $|E| = n - c = c - c = 0$, which is correct.

Inductive Step: Assume that the statement is true for $n = k$. We need to show that it still holds true for $n = k + 1$, i.e., the number of edges in a forest with $k + 1$ vertices and c components is $k + 1 - c$.

By the inductive hypothesis, a forest with k vertices and c components has $k - c$ edges. If we want $k + 1$ vertices and c components, we can only add a vertex to one of the existing components (so that the number of components stays the same). Adding a vertex to one of the components requires us to attach an edge from the new vertex to one of the vertices in that component. Note that adding a new vertex does not create a cycle, so the whole graph is still a forest. Now we have a forest with $k + 1$ vertices and c components, and the number of edges is $k - c + 1 = (k + 1) - c$, which is what we are looking for.

That's it for the induction on n. Now we need to do induction on c, leaving n fixed.

Base Case: $(c = 1)$ If the forest with n vertices has one component, then the forest is actually a tree. By a theorem proven in class, this forest (tree) has $n - 1$ edges. But $n - 1 = n - c$, so the formula holds for $c = 1$.

Inductive Step: Assume that the statement is true for $c = k$. We need to show that it still holds true for $c = k + 1$, i.e., the number of edges in a forest with n vertices and $k + 1$ components is $n - (k + 1)$.

Consider any forest with n vertices and $k + 1$ components. Adding an edge between any two components does not create a cycle, so those two components become one, i.e., a tree. But now the number of components is $(k + 1) - 1 = k$. By the inductive hypothesis, this forest with n vertices and k components has $n - k$ edges. This forest has one extra edge from the original forest. Hence, the number of edges in the original forest is $n - k - 1 = n - (k + 1)$, as desired.

This completes the proof of the proposition. □

Since we have some space left, the following is not directly related to our course, but it might be helpful in other courses such as CIS121.

Lemma: Let q be a constant. Then

$$\sum_{i=0}^{n} q^i = \frac{q^{n+1} - 1}{q - 1}.$$

This sum is called a geometric series.
Proof: Let

$$A = \sum_{i=0}^{n} q^i = 1 + q + q^2 + \cdots + q^n. \tag{1}$$

Multiplying both sides by q, we obtain

$$qA = \sum_{i=1}^{n+1} q^i = q + q^2 + \cdots + q^n + q^{n+1}. \tag{2}$$

Subtracting (1) from (2), we obtain

$$(q - 1)A = q^{n+1} - 1.$$

Therefore,

$$A = \sum_{i=0}^{n} q^i = \frac{q^{n+1} - 1}{q - 1},$$

as required. □

Exercise: Prove the above lemma by induction.