CIS 160

Recitation Guide - Week 7

Topics Covered: Graphs, Trees, Conditional Probability

Problem 1: Let T be a tree where the maximum degree is Δ. Prove that T has at least Δ leaves.

Solution:

Direct Proof:

Let $v \in V$ have degree Δ in $T = (V, E)$. Consider the subgraph induced on the vertices $V \setminus \{v\}$. Each neighbor of v is in a distinct connected component in this graph, because we have destroyed the unique path between any two of v's neighbors in T. Thus there are Δ connected components, each of which is a tree (because each connected component is also acyclic - a cycle in one of these components would also be a cycle in T, which is impossible).

There are two possibilities for each connected component. If a connected component is a single node, then this single node is a leaf adjacent to v in T. If the connected component has at least 2 nodes, then it has at least 2 leaves. One of the leaves may be adjacent to v and not a leaf in T. But the other leaf in this connected component is still a leaf in T. In any case, each connected component contains at least one leaf of T and hence T must have Δ leaves.

Maximal Path:

We first prove a lemma.

Lemma: All maximal paths in a tree must start and end with leaves. We say a path is maximal if it cannot be extended.

Proof: Suppose for the sake of contradiction that at least one of the endpoints is not a leaf. Let v be this vertex. Notice that $deg(v) \geq 2$, since it is not a leaf. Since v is at the end of a maximal path, there must exist some vertex u that is a neighbor of v, but not the neighbor of v in the maximal path. However, this creates a cycle, as u must lie on the maximal path (or the path could be extended), which is a contradiction since trees do not have cycles. ■

Let $v \in V$ have degree Δ. For each $u_i, u_j \in N(v)$, let $P_{i,j}$ be a maximal path including $u_i - v - u_j$. Note that there must be at least $\binom{\Delta}{2}$ such distinct maximal paths of this form, since each pair of starting edges gives a different path. We know from the above lemma that any maximal path in T must terminate in two leaves. Let λ denote the number of leaves in T. Thus, there are at most $\binom{\lambda}{2}$ maximal paths in the tree. Thus, $\binom{\lambda}{2} \geq \binom{\Delta}{2}$, so we must then have $\lambda \geq \Delta$. Thus, T has at least Δ leaves.

Contradiction:

Assume that $\Delta \geq 2$, since the cases of $\Delta = 0$ and $\Delta = 1$ are clearly true. Suppose for the sake of contradiction that there are at most $\psi < \Delta$ leaves. Let $v \in V$ have degree Δ. Consider $S = \{u \in V \mid \{u, v\} \in E\}$. Note that S is the set of v's neighbors, and $|S| = \Delta$.

For all $u_i \in S$, there exists at least one path that starts with $\{v, u_i\}$ that ends with a leaf. We pick any such leaf for each edge $\{v, u_i\}$ and call the leaf l_i. Note there is a unique l_i corresponding to each u_i, as trees are acyclic, so we have Δ l_i's in total. Hence, by the Pigeonhole Principle, where the pigeons are the terminating leaves l_i of each path and the holes are the ψ leaves available, we know that $\lceil \frac{\Delta}{\psi} \rceil \geq \lceil \frac{\Delta}{\Delta-1} \rceil = \lceil 1 + \frac{1}{\Delta-1} \rceil$ (since $\Delta \geq 2$) $= 2$ paths share the same terminating leaf, say ℓ_ω.

This is a contradiction, since the path between ℓ_ω and v are unique in a tree.

Induction on the number of vertices:

Let us prove this by induction on the number of vertices in the graph n.

We formulate a proposition $P(n)$ which is: in a tree with n vertices and maximum degree Δ, the number of leaves in the tree is at least Δ.

Base Case (n= 1, 2 and 3): The case of $n = 1$ is trivial - a graph of just 1 node has maximum degree 0 and at least 0 leaves. There is only one possible tree when $n = 2$: $T = (V, E)$, $V = \{u, v\}$, $E = \{\{u, v\}\}$. Here $\Delta = 1$, and we have 2 leaves, so it checks out as required.

There is only one possible tree when $n = 3$: $T = (V, E)$, $V = \{u, v, w\}$, $E = \{\{u, v\}, \{v, w\}\}$. Here $\Delta = 2$, and we have 2 leaves, so it checks out as required.

We choose to show three base cases here to avoid a slightly unfortunate edge case in the Induction Step.

Induction Step: Assume that (IH) $P(k)$ is true, for some $k \in \mathbb{Z}^+, k \geq 3$. Consider an arbitrary tree $T = (V, E)$ such that $|V| = k+1$ and it has maximum degree Δ. Let $\ell \in V$ be an arbitrary leaf in T who has some neighbor a. Consider $T' = (V', E')$ where $V' = V \setminus \ell$ and $E' = E \setminus \{a, \ell\}$.

We know that $|V'| = k$ and is a tree (since removal of a leaf can never disconnect a tree), so we can apply the Induction Hypothesis on T'.

Note that there are two cases here:

1. a was the only vertex of degree Δ in T.

   It must be the case then that a has degree $\Delta - 1$ in T' and is of maximum degree. The Induction Hypothesis gives us that T' must have at least $\Delta - 1$ leaves.

   Further note if a is a leaf in T', then it must be the case that $n = 3$ (convince yourself of this), and that is already shown to be true by the base case. Hence, going forward we will operate under the assumption that a is not a leaf.

   Adding ℓ back to T' to reconstruct T increases the number of leaves by one (since a is not a leaf), so we have that T has at least Δ leaves.

2. There is some vertex in T' that has degree Δ.

   By the Induction Hypothesis, we have that T' must have Δ leaves.

   There are two more cases here:

   (a) a is a leaf in T'

In this case, the addition of ℓ does not change the number of leaves, which means we have at least Δ leaves in T, as desired.

(b) a is not a leaf in T'

In this case, the addition of ℓ increases the number of leaves by 1, which means we have at least $\Delta + 1$ leaves in T, which proves our claim.

Induction on the number of edges:

You can do a similar procedure to the induction on the number of vertices in order to perform induction on the number of edges. Note that in this case you would consider the subgraph induced by the vertices other than the leaf.

Strong Induction on the number of edges:

Let us prove this by induction on the number of edges in the graph m.

We formulate a proposition $P(m)$ which is: in a tree with m edges and maximum degree Δ, the number of leaves in the tree is at least Δ.

Base Case (m=0, 1, 2): There is only one possible tree (of one vertex) when $m = 0$. Here $\Delta = 0$, and we have at least 0 leaves.

There is only one possible tree when $m = 1$: $T = (V, E)$, $V = \{u, v\}$, $E = \{\{u, v\}\}$. Here $\Delta = 1$, and we have 2 leaves, so it checks out as required.

There is only one possible tree when $m = 2$: $T = (V, E)$, $V = \{u, v, w\}$, $E = \{\{u, v\}, \{v, w\}\}$. Here $\Delta = 2$, and we have 2 leaves, so it checks out as required.

We choose to show two base cases here to avoid a slightly unfortunate edge case in the Induction Step.

Induction Step: Assume that (IH) $P(j)$ is true, for all $j \in \mathbb{Z}, 1 \leq j \leq k$, for some $k \in \mathbb{Z}^+$, $k \geq 2$. Let T be a tree with $k+1$ edges and with a maximum degree Δ. Let v be a vertex with degree Δ, and u be an arbitrary neighbor of v. Let us consider $G' = (V', E')$, where $V' = V$, $E' = E \setminus \{\{u, v\}\}$. Note that G' must have had two connected components C_1 and C_2, which are both trees when a subgraph is induced on each of them. Let C_1 be the component with v, and let C_2 be the component with u.

There are two cases here:

1. There is another vertex in C_1 that has degree Δ

   From the induction hypothesis, we have that there must be Δ leaves in C_1. Let us reconstruct T from G'.

   There are two cases here:

   (a) $|C_2| = 1$

   In this case, if v is a leaf in G', then the addition of $\{u, v\}$ will not change the number of leaves. Therefore we have that T must have at least Δ leaves. If v is not a leaf, then the addition of $\{u, v\}$ will add an additional leaf, so we have that T must have at least $\Delta + 1$ leaves.

(b) $|C_2| \geq 2$

In this case, C_2 must have two leaves. Hence there are at least $\Delta + 2$ leaves in G'. Notice that the addition of the edge $\{u, v\}$ can decrease the number of leaves by up to 2 (if u and v were both leaves in G'). Hence we have that T has at least Δ leaves, as required.

2. v is the only vertex with degree Δ in T.

   Hence, $\Delta(C_1) = \Delta - 1$. From the induction hypothesis, we know that C_1 must have $\Delta - 1$ leaves. We further note that if v is a leaf in G', it must be that $m = 2$ (convince yourself of this), and we have already shown the validity of this in the base case. We will therefore operate now under the assumption that v is not a leaf.

   There are two cases here:

   (a) $|C_2| = 1$

   Since v is not a leaf, then the addition of $\{u, v\}$ will add an additional leaf, so we have that T must have at least Δ leaves.

   (b) $|C_2| \geq 2$

   In this case, C_2 must have two leaves. Hence there are at least $\Delta + 1$ leaves in G'. Notice that the addition of the edge $\{u, v\}$ can decrease the number of leaves by up to 1 (if u is a leaf in G'). Hence we have that T has at least Δ leaves, as required.

Using inequalities:

We know that a tree with n vertices must have $n - 1$ edges. Since the sum of the degrees of all the vertices in a graph must be twice the number of edges, we know that the total of all degrees in the tree must be $2n - 2$.

Let us consider the following partitioning of the vertices in V. Let $A = \{v \in V \mid deg(v) = \Delta\}$, $B = \{v \in V \mid 1 < deg(v) < \Delta\}$, and $C = \{v \in V \mid deg(v) = 1\}$. Note that $V = A \cup B \cup C$ and $A \cap B = \emptyset, A \cap C = \emptyset, B \cap C = \emptyset$. Note that C is the set of leaves.

$$
\begin{aligned}
2n - 2 &= \sum_{v \in V} deg(v) \\
&= \sum_{v \in A} deg(v) + \sum_{v \in B} deg(v) + \sum_{v \in C} deg(v) \\
&= \Delta \cdot |A| + \sum_{v \in B} deg(v) + |C| \\
&\geq \Delta \cdot |A| + |C| + 2 \cdot |B| \\
&= \Delta \cdot |A| + |C| + 2 \cdot (n - |A| - |C|) \\
&= (\Delta - 2) \cdot |A| - |C| + 2n \\
&\geq (\Delta - 2) - |C| + 2n
\end{aligned}
$$

Hence we have established that $2n - 2 \geq (\Delta - 2) - |C| + 2n$. Further, we have that:

$$\begin{aligned} 2n - 2 &\geq (\Delta - 2) - |C| + 2n \\ -2 &\geq \Delta - 2 - |C| \\ |C| &\geq \Delta \end{aligned}$$

Hence we have that the number of leaves is at least Δ.

Problem 2:

You run into a town with 100 robots. You know that 99 of these robots tell the truth half the time and lie the other half. You also know that there is exactly 1 truthful robot in town, who always tells the truth. You take a robot at random and ask a question seven times, and the robot tells the truth every time. What is the probability that this is the truthful robot?

Solution:

We can define each outcome in the sample space Ω as an ordered pair, where the first element represents whether or not the robot is truthful, and the second element is the robot's answers to the 7 questions, i.e. a sequence of length 7 of elements from $\{T,F\}$.

Let A be the event: selected robot is truthful. Let B be the event: selected robot tells the truth seven times.

Note that from parsing the question, we know the following probabilities:

$$\Pr[A] = \frac{1}{100} \qquad \Pr[\overline{A}] = \frac{99}{100}$$

$$\Pr[B|A] = 1 \qquad \Pr[B|\overline{A}] = \left(\frac{1}{2}\right)^7$$

We want $\Pr[A|B]$ (make sure you can justify each step in this simplification):

$$\begin{aligned}
\Pr[A|B] &= \frac{\Pr[A \cap B]}{\Pr[B]} && \text{(Conditional Probability)} \\
&= \frac{\Pr[A] \times \Pr[B|A]}{\Pr[B]} && \text{(Conditional Probability)} \\
&= \frac{\Pr[A] \times \Pr[B|A]}{\Pr[A \cap B] + \Pr[\overline{A} \cap B]} && \text{(Total Probability Theorem)} \\
&= \frac{\Pr[A] \times \Pr[B|A]}{\Pr[A] \times \Pr[B|A] + \Pr[\overline{A}] \times \Pr[B|\overline{A}]} \\
&= \frac{\frac{1}{100} \times 1}{\frac{1}{100} \times 1 + \frac{99}{100} \times (\frac{1}{2})^7} \\
&= \frac{128}{227} \approx \boxed{0.564}
\end{aligned}$$

Problem 3:
Prove that G or the complement of G is connected. Note that the complement of a graph $G = (V, E)$ is $G^c = (V, E')$ and $\forall u, v \in V$, $\{u, v\} \in E' \iff \{u, v\} \notin E$.

Solution:

If G is connected we are done.

If G is not connected then G is composed of multiple connected components. We want to prove that given two arbitrary vertices in G there must be a path between them in G^c. Let these two arbitrary vertices be u and v.

Case 1: u and v do not share an edge in G

This means they must share an edge in G^c and thus there is a path from u to v in G^c.

Case 2: u and v share an edge in G

This means they were part of the same connected component in G. Take an arbitrary vertex x in a different connected component in G. Edges $u - x$ and $v - x$ must both exist in G^c. Thus, there is a path $u - x - v$ between vertices u and v.

Thus, we have shown that there exists a path between any two arbitrary vertices in G^c. By definition G^c must be connected. The claim is proved.