CIS 160
Recitation Guide - Week 7

**Topics Covered:** Graphs, Probability (Inclusion-Exclusion, Conditional)

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

**Solution:**
We will use the (non-standard) notation $\lambda(T)$ to denote the number of leaves in a tree $T$. Thus, we can rewrite the claim as $\lambda(T) \geq \Delta(T)$.

**Direct Proof:**
Let $v \in V$ have degree $\Delta$ in $T = (V,E)$. Consider the subgraph induced on the vertices $V \setminus \{v\}$. Each neighbor of $v$ is in a distinct component in this graph, because we have destroyed the unique path between any two of $v$’s neighbors in $T$. Thus there are $\Delta$ components, each of which is a tree.

There are two possibilities for each component. If a component is a single node, then this single node is a leaf adjacent to $v$ in $T$. If the 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 component is still a leaf in $T$. In any case, each component contains at least one leaf of $T$ and hence $T$ must have $\Delta$ leaves.

**Maximal Path:**
Let $v \in V$ have degree $\Delta$. 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 $\left(\binom{\Delta}{2}\right)$ such paths, since any pair of starting edges gives a different path. We know that any such path $p_{i,j}$ must terminate in two leaves (call them $w_{i,j}$ and $x_{i,j}$). Lastly, note that since there is a unique path between any two vertices in a tree, every pair of leaves admits at most one maximal path. If there were $\lambda(T) < \Delta$ leaves, we would only have $\binom{\lambda(T)}{2} < \binom{\Delta}{2}$ distinct maximal paths, a contradiction; we must then have $\lambda(T) \geq \Delta$.

**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 $\lambda(T) < \Delta$ leaves. For each $u_i \in N(v)$, let $p_i$ be a maximum length path beginning with $v, \{v, u_i\}, u_i$. Note that there must be $\Delta$ such paths. We know that any such path $p_i$ must terminate in a leaf $\ell_i$.

By the Pigeonhole Principle, where the pigeons are the terminating leaves of each path and the holes are the $\lambda(T)$ leaves available, we know that, since $\lambda(T) < \Delta$, two paths share the same terminating leaf, say $\ell_\omega$.

This is a contradiction, since there can only exist one unique path between $\ell_\omega$ and $v$.

**Induction on the number of vertices:**
Let us prove this by induction on the number of vertices in the graph $n$. 

1
We formulate a proposition $P(n)$ which is: in a tree with $n$ vertices and maximum degree $\Delta$, the number of leaves in the tree is at least $\Delta$.

**Induction Hypothesis:** Assume that $P(k)$ is true, for some $k \in \mathbb{Z}^+, k \geq 3$.

**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:** Consider an arbitrary tree $T = (V, E)$ such that $|V| = k + 1$ and it has maximum degree $\Delta$. 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 $\Delta$ 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 $\ell$ 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 $\Delta$ leaves.

2. There is some vertex in $T'$ that has degree $\Delta$.
   
   By the Induction Hypothesis, we have that $T'$ must have $\Delta$ leaves.

   There are two more cases here:

   (a) $a$ is a leaf in $T'$

   In this case, the addition of $\ell$ does not change the number of leaves, which means we have at least $\Delta$ leaves in $T$, as desired.

   (b) $a$ is not a leaf in $T'$

   In this case, the addition of $\ell$ 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 \( \Delta \), the number of leaves in the tree is at least \( \Delta \).

**Induction Hypothesis:** Assume that \( P(j) \) is true, for all \( j \in \mathbb{Z}, 1 \leq j \leq k \), for some \( k \in \mathbb{Z}^+ \), \( k \geq 2 \).

**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:** Let \( T \) be a tree with \( k + 1 \) edges and with a maximum degree \( \Delta \). Let \( v \) be a vertex with degree \( \Delta \), 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 \( \Delta \)

   From the induction hypothesis, we have that there must be \( \Delta \) 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 \( \Delta \) 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 \( \Delta \) leaves, as required.

2. \( v \) is the only vertex with degree \( \Delta \) 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 \(\Delta\) 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 \(\Delta\) 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 \(A = \{v \in V \mid \text{deg}(v) = \Delta\}\), \(B = \{v \in V \mid 1 < \text{deg}(v) < \Delta\}\), and \(C = \{v \in V \mid \text{deg}(v) = 1\}\). 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.

\[
2n - 2 = \sum_{v \in V} \text{deg}(v) \\
= \sum_{v \in A} \text{deg}(v) + \sum_{v \in B} \text{deg}(v) + \sum_{v \in C} \text{deg}(v) \\
= \Delta \cdot |A| + \sum_{v \in B} \text{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 \quad \text{(since \(|A| \geq 1\))}
\]

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

\[
2n - 2 \geq (\Delta - 2) - |C| + 2n \\
-2 \geq \Delta - 2 - |C| \\
|C| \geq \Delta
\]

Hence we have that the number of leaves is at least \(\Delta\).

Minimal Counterexample:

Consider a minimal counterexample, i.e. a tree \(T\) which violates this property with the minimum possible number of vertices, say \(m\). We know that the case for \(m = 1, 2\) can be handled easily, so we may assume that \(m \geq 3\), i.e. the tree has at least 3 vertices. Now pick an arbitrary leaf \(\ell\) and name its only neighbor in the graph \(v\); remove \(\ell\). Consider the resulting graph \(T'\). Note that \(T'\) has exactly \(m - 1\) vertices. The following cases can occur:

Case 1: \(\Delta(T) = \Delta(T')\).
Note that by removing a single leaf, we can never increase the number of leaves in the graph. It follows that $T'$ has at most as many leaves as $T$, i.e.

$$\lambda(T') \leq \lambda(T) < \Delta(T) = \Delta(T')$$

But this means $\lambda(T') < \Delta(T')$ and $T'$ has $m-1$ vertices. This is a contradiction, as we chose $T$ to be tree with the fewest vertices which violates the claim.

**Case 2:** $v$ is a leaf in $T'$ and $\Delta(T) \neq \Delta(T')$.

The only vertex whose degree can be affected by removing $\ell$ is $v$. Then $v$ must have degree 1 and all other vertices must have degree $\leq 1$. The only trees for which this hold have exactly 1 or 2 vertices; we already know that these cases do not violate the claim. As such, we have a contradiction (it’s impossible for us to end up in this scenario).

**Case 3:** $v$ is not a leaf in $T'$ and $\Delta(T) \neq \Delta(T')$. If the maximum degree changes by removing this leaf, that means that it must decrease by exactly one (we cannot increase degree by removing edges and only removed one edge). In other words, $\Delta(T') = \Delta(T) - 1$. Note that the number of leaves in $T'$ is $\lambda(T) - 1$, since we removed $\ell$ and $v$ is not a leaf. It follows that

$$\lambda(T') = \lambda(T) - 1 < \Delta(T) - 1 = \Delta(T')$$

Again, $\lambda(T') < \Delta(T')$ and we have a contradiction, since $T$ is not minimal.

In every case we have a contradiction - it must be the case that the set of counterexample is empty, i.e. there are no trees which violate the claim, and we are finished.
**Problem 2:**

It is the Fall 2019 semester and Professor Gandhi is now teaching 3 full sections (001, 002, and 003) of CIS 160. On any given day of classes, there is a $\frac{2}{5}$ chance that he misses his first section, a $\frac{2}{5}$ chance that he misses his second section, and a $\frac{1}{5}$ chance that he misses his last section (and missing one does not impact whether he misses the other). Using the inclusion-exclusion formula, calculate the probability that on a given day, Professor Gandhi does not attend all of his classes.

**Solution:**

Let the sample space $\Omega$ be a set of 3-tuples, where any particular outcome can be expressed as $w = (w_1, w_2, w_3)$, such that $w_1, w_2, w_3$ are equal to 1 if Professor Ghandi attended sections 001, 002, and 003 respectively, and 0 otherwise. Note that this sample space has a non-uniform probability distribution.

Let $A$, $B$, $C$ be the events in which Professor Gandhi misses Section 001, 002 and 003, respectively. If he does not attend every class, then he must be missing at least one. Therefore, we are looking for $\Pr[A \cup B \cup C]$.

By the Inclusion-Exclusion Principle,

$$\Pr[A \cup B \cup C] = \Pr[A] + \Pr[B] + \Pr[C] - \Pr[A \cap B] - \Pr[A \cap C] - \Pr[B \cap C] + \Pr[A \cap B \cap C]$$

Note that $A$, $B$, and $C$ are independent events, so:

$$\Pr[A \cap B] = \Pr[A] \cdot \Pr[B] = \frac{2}{5} \cdot \frac{2}{5} = \frac{4}{25}$$
$$\Pr[B \cap C] = \Pr[B] \cdot \Pr[C] = \frac{2}{5} \cdot \frac{1}{5} = \frac{2}{25}$$
$$\Pr[A \cap C] = \Pr[A] \cdot \Pr[C] = \frac{2}{5} \cdot \frac{1}{5} = \frac{2}{25}$$
$$\Pr[A \cap B \cap C] = \Pr[A] \cdot \Pr[B] \cdot \Pr[C] = \frac{2}{5} \cdot \frac{2}{5} \cdot \frac{1}{5} = \frac{4}{125}$$

Plugging it in, we have:

$$\Pr[A \cup B \cup C] = \Pr[A] + \Pr[B] + \Pr[C] - \Pr[A \cap B] - \Pr[A \cap C] - \Pr[B \cap C] + \Pr[A \cap B \cap C]$$
$$= \frac{2}{5} + \frac{2}{5} + \frac{1}{5} - \frac{4}{25} - \frac{2}{25} - \frac{2}{25} + \frac{4}{125}$$
$$= 1 - \frac{8}{25} + \frac{4}{125}$$
$$= \frac{89}{125}$$

A faster way to solve the problem (without using inclusion-exclusion) is to notice the following
equality:

\[
\Pr[A \cup B \cup C] = 1 - \Pr[(A \cup B \cup C)^c] \\
= 1 - \Pr[A^c \cap B^c \cap C^c] \\
= 1 - \Pr[A^c] \cdot \Pr[B^c] \cdot \Pr[C^c]
\]

Note that

\[
\Pr[A^c] = 1 - \Pr[A] = 1 - \frac{2}{5} = \frac{3}{5} \\
\Pr[B^c] = 1 - \Pr[B] = 1 - \frac{2}{5} = \frac{3}{5} \\
\Pr[C^c] = 1 - \Pr[C] = 1 - \frac{1}{5} = \frac{4}{5}
\]

So, we have that:

\[
\Pr[A \cup B \cup C] = 1 - \Pr[A^c] \cdot \Pr[B^c] \cdot \Pr[C^c] \\
= 1 - \frac{3}{5} \cdot \frac{3}{5} \cdot \frac{4}{5} \\
= \frac{89}{125}
\]
Problem 3:
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 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 $\Omega$ 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:

\[
\begin{align*}
\Pr(A) &= \frac{1}{100} \\
\Pr(\overline{A}) &= \frac{99}{100} \\
\Pr(B|A) &= 1 \\
\Pr(B|\overline{A}) &= \left(\frac{1}{2}\right)^7
\end{align*}
\]

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

\[
\begin{align*}
\Pr(A|B) &= \frac{\Pr(A \cap B)}{\Pr(B)} \\
&= \frac{\Pr(A) \times \Pr(B|A)}{\Pr(B)} \\
&= \frac{\Pr(A) \times \Pr(B|A)}{\Pr(A \cap B) + \Pr(\overline{A} \cap B)} \quad \text{(Total Probability Theorem)} \\
&= \frac{\Pr(A) \times \Pr(B|A)}{\Pr(A) \times \Pr(B|A) + \Pr(A) \times \Pr(B|A)} \\
&= \frac{1}{100} \times 1 \\
&= \frac{128}{227} \approx 0.564
\end{align*}
\]