Counting Trees. Prüfer.

from notes of Jerry Lodder; modified by R Montgomery.

On Friday Jan 8, Nathan sketched an induction argument worked out by his group that every tree on n vertices has $n-1$ edges. Their proof relied crucially on the assertion that every finite tree has at least one "leaf". A "leaf" is a vertex of degree 1. Prove this assertion by completing:

Exercise 0.1. Suppose that G is a finite graph all of whose vertices has degree two or greater. Prove that a cycle passes through each vertex. Conclude that G cannot be a tree.

Why does the exercise show every tree has at least one leaf? Here is the Nathan's group inductive argument, sketched a bit more carefully.

Exercise 0.2. *To prove,* via induction, that every tree on n vertices has exactly $n-1$ edges.

BASE CASE: Explain why the result holds for $n = 2$.

INDUCTIVE HYPOTHESIS: If T is a tree on n vertices, then T has $n-1$ edges.

SKETCH PROOF: Let S be a tree on $n+1$ vertices. Delete a leaf (vertex of degree one !) from S and the edge connected to the leaf. Is the graph formed by these deletions still a tree? Why? Carefully apply the inductive hypothesis to finish the argument.

Exercise 0.3. Draw several connected graphs which have exactly one vertex of degree 1. Are any of them trees.

Exercise 0.4. Prove: every tree has at least TWO leaves. Do so using the degree sum formula $\Sigma_{v \in V} deg(v) = 2\#\text{edges}$.

Let us back up and go through two basic characterizations of trees:

Exercise 0.5. Let G be a connected graph. We say that G is minimally connected if the removal of any edge of G (without deleting any vertices) results in a disconnected graph.

(a) Show that a connected, minimally connected graph has no cycles.

(b) Show that a connected graph with no cycles is minimally connected.

COMMENTARY: The exercise shows that the two conditions are equivalent. A graph satisfying either condition is a tree.

READ Prüfer's paper. Here is a synopsis in exercises.

Exercise 0.6. Let T be a tree on n ordered vertices and let b_1 be a leaf of T (the first leaf, if necessary). Let a_1 be the vertex to which b_1 is connected by an edge b_1a_1, and let T' be the graph constructed from T by deleting the vertex b_1 and deleting the edge b_1a_1 (do not delete the vertex a_1). Prove that T' is a tree, by showing that either

(a) T' is a connected graph that contains no cycles; or

(b) T' is a connected, minimally connected graph.

Which argument, (a) or (b), does Prüfer's paper suggest? How many vertices does T' contain?

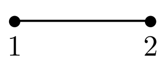
Exercise 0.7. Let T be a labelled tree on n vertices, whose vertices are completely ordered from 'smallest' to 'greatest'. Let a_1 be smallest leaf. Let T' be the tree on $n-1$ vertices constructed by deleting a_1, as per Exercise 0.6. Define $\mathcal{S}$ recursively by

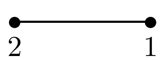
$$\mathcal{S}(T) = (a_1,\ \mathcal{S}(T')),$$

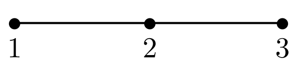
for $n > 3$. Does this construction match Prüfer's description? Prüfer says: "If $n-1 > 2$ also, then one determines the town a_2 with which the first endpoint b_2 of the new net is directly connected. We take a_2 as the next element of the symbol. Then we strike out the town b_2 and the segment $b_2\,a_2$. We obtain a net with $n-3$ segments and the same properties. We continue this procedure until we finally obtain a net with only one segment joining 2 towns. Then nothing more is included in the symbol."

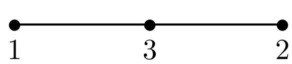
Which construction do you find easier to understand, the recursive definition or Prüfer's description? Which would be easier to implement? Why?

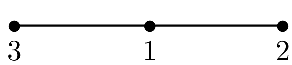
Exercise 0.8. Find the Prüfer symbols of the following labeled trees. Explain your solutions.

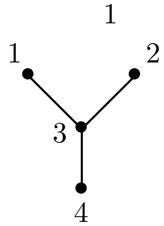
(a) 

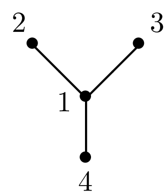
(b) 

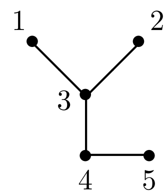
(c) 

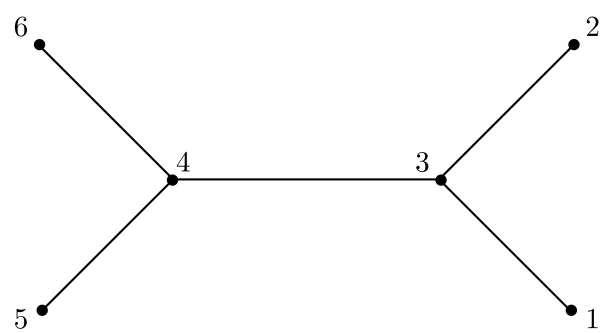
(d) 

(e) 

(f) 

(g) 

(h) 

(i) 

Exercise 0.9. Draw the labeled trees which have the following Prüfer symbols σ. Be sure to explain your solutions.

(a) The symbol $\sigma = (1)$, where the labels of the vertices are 1, 2, 3.

(b) The symbol $\sigma = (1,\ 1)$, where the labels of the vertices are 1, 2, 3, 4.

(c) The symbol $\sigma = (1,\ 1,\ 4)$, where the labels of the vertices are 1, 2, 3, 4, 5.

(d) The symbol $\sigma = (1,\ 1,\ 2)$, where the labels of the vertices are 1, 2, 3, 4, 5.

Exercise 0.10. Let $V = \{v_1,\ v_2,\ \ldots,\ v_n\}$ be a set of vertices. How many sequences or "words" of length $n-2$ are there in the characters $v_1,\ v_2,\ \ldots,\ v_n$? Characters may be repeated? Two sequences $(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_{n-2})$ and $(\beta_1,\ \beta_2,\ \ldots,\ \beta_{n-2})$ are counted as the same if and only if $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2$, $\ldots$, $\alpha_{n-2} = \beta_{n-2}$. Explain your answer.

1 Counting Unlabelled trees.

Background. Cayley's formula asserts that #(labelled trees on n vertices) $= n^{n-2}$. Is there a nice formula for counting the number $T(n)$ of unlabeled trees on n vertices? Starting from $n = 2$ we have $T(n) : T(2) = 1, T(3) = 1, T(4) = 2, T(5) = 3$.

After class on Friday, S (a student) made a beautiful stab at a conjectured formula for $T(n)$: GUESS: $T(n) = F_{n-1}$ where F_n denotes the nth Fibonacci number. His conjecture was based in part on the fact that $2 = 1 + 1$ and $3 = 1 + 2$. Recall the Fibonacci series: $F_n : 1, 1, 2, 3, 5, 8, 11, \ldots$ with $F_{n+1} = F_n + F_{n-1}$.

Exercise 1.1. Show that $T(6) = 6$, that is, the number of distinct unlabeled trees on 6 vertices, is 6.

Sorry, S. your guess is wrong! It is an open problem to find a closed form formula for $T(n)$!

Exercise 1.2. A. Show that $n^{n-2}/n! < T(n)$ by looking at how the symmetric group S_n acts on labelled trees. Use $|S_n| = n!$

B. Show that $T(n) < n^{n-2} - (n! - 2)$ by looking at how badly we over count for the linear graph (how many labeling there are for the linear graph and the graph with a branch at at the next-to-last vertex.

C. Use Stirling's approximation as described in wiki and Part A to show that $\frac{1}{\sqrt{2\pi}} \frac{e^n}{\sqrt{n}} < T(n)$.

A New Proof of a Theorem about Permutations.

by Heinz Prüfer from Berlin.

In the Berlin Mathematical Society, Herr Dziobek has announced a theorem ... . His proof ... is not particularly simple, and it is perhaps of interest to look at another proof which depends entirely on combinatorial considerations. I shall express it in an intuitive geometrical garb, as posed by Herr Professor Schur in a problem to the University of Berlin's mathematical seminar:

Consider a country with n towns. These towns must be connected by a railway network of $n-1$ single segments (the smallest possible number) in such a way that one can travel from each town to every other town. There are n^{n-2} different railway networks of this kind.

By a single segment is meant a stretch of railway that connects only two towns. The theorem can be proved by assigning to each railway network, in a unique way, a symbol $\{a_1,\ a_2,\ \ldots,\ a_{n-2}\}$, whose $n-2$ elements can be selected independently[1] from any of the numbers 1, 2, ..., n. There are n^{n-2} such symbols, and this fact, together with the one-to-one correspondence between networks and symbols, will complete the proof.

In the case $n=2$, the empty symbol corresponds to the only possible network, consisting of just one single segment that connects both towns. If $n>2$, we denote the towns by the numbers 1, 2, ..., n and specify them in a fixed sequence. The towns at which only one segment terminates we call the endpoints. [Every network has endpoints] for otherwise there would be at least two segments terminating at each town, and there would be at least $\dfrac{2n}{2}=n$ segments.

In order to define the symbol belonging to a given net for $n>2$, we proceed as follows.

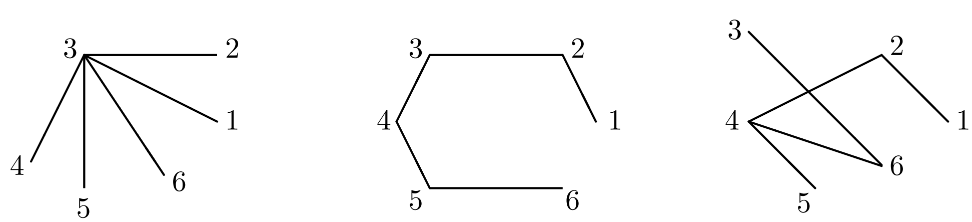
Let b_1 be the first town which is an endpoint of the net, and a_1 the town which is directly joined to b_1. Then a_1 is the first element of the symbol. We now strike out the town b_1 and the segment $b_1\,a_1$. There remains a net containing $n-2$ segments that connects $n-1$ towns in such a way that one can travel from each town to any other.

If $n-1>2$ also, then one determines the town a_2 with which the first endpoint b_2 of the new net is directly connected. We take a_2 as the next element of the symbol. Then we strike out the town b_2 and the segment $b_2\,a_2$. We obtain a net with $n-3$ segments and the same properties.

We continue this procedure until we finally obtain a net with only one segment joining 2 towns. Then nothing more is included in the symbol.

Examples:

Nets:



Symbols: $\{3,\ 3,\ 3,\ 3\}$ $\{2,\ 3,\ 4,\ 5\}$ $\{2,\ 4,\ 6,\ 4\}$

Each town at which m segments terminate occurs exactly $m-1$ times in the symbol. For, in the formation of the symbol by successively removing segments, a town appears in the symbol only when

[1] The entries (elements) of a symbol may be repeated.

one of its incident edges is removed, except in the case that this edge is the last one having that town as endpoint.

Conversely, if we are given a particular symbol $\{a_1,\ a_2,\ \ldots,\ a_{n-2}\}$, other than the empty symbol, then we write down the numbers 1, 2, ..., n, and find the first number that does not appear in the symbol. Let this be b_1. Then we connect the towns b_1 and a_1 by a segment. We now strike out the first element of the symbol and the number b_1.

If $\{a_2,\ a_3,\ \ldots,\ a_{n-2}\}$ is also not the empty symbol, then we find b_2, the first of the $n-1$ remaining numbers that does not appear in the symbol. Connect the towns b_2 and a_2. Then strike out the number b_2 and the element a_2 in the symbol.

In this way we eventually obtain the empty symbol. When that happens, we join the last two towns not yet crossed out.

That the system of segments obtained by this construction actually is a net, and that this net and no other actually gives rise to the given symbol, follows from an induction argument. For, if a net is represented by a symbol, then the towns which do not appear in the symbol are just the endpoints of the net. As the segment $b_1\,a_1$ is the only line ending at b_1, it [segment $b_1\,a_1$] must appear in the net. But we may assume that we have proved that the symbol $\{a_2,\ a_3,\ \ldots, a_{n-2}\}$ corresponds to just one net connecting all the towns except b_1, and that this net was obtained by the construction, so that the truth of the proposition follows for the symbol $\{a_1,\ a_2,\ \ldots,\ a_{n-2}\}$.