TWO PROOFS OF CAYLEY'S THEOREM

TITU ANDREESCU AND COSMIN POHOATA

Abstract. We present two proofs of the celebrated Cayley theorem that the number of spanning trees of a complete graph on n vertices is n^{n-2}.

In this expository note we present two proofs of Cayley's theorem that are not as popular as they deserve to be. To set up the story we revisit first some terminology. By a graph G we mean a pair $(V(G), E(G))$, where $V(G)$ is a set of points (or *vertices*) and $E(G)$ is a subset of $V(G) \times V(G)$. Visually, each element from the set $E(G)$ can be represented as a curve connecting the two corresponding points, thus we will call these the *edges* of graph G. Note that with the above definition, graphs cannot have multiple edges between two given vertices u and v, but they can have loops (i.e. edges in $E(G)$ of the form (w, w) for $w \in V(G)$). However, in this paper we won't allow such pairs.

Such simple graphs can be *directed* or *undirected* depending on whether the pairs in $E(G) \subset V(G) \times V(G)$ are ordered or not. A *path* is a graph P with vertex set $V(P) = \{v_1, \ldots, v_n\}$ and edge set $E(P) = \{(v_i, v_{i+1}) : 1 \leq i \leq n-1\}$. The number n will be refered to as the length of the path. A *cycle* (of length n) is a graph C of the form $C = P \cup (v_n, v_1)$. Or in other words, a cycle of length n is a graph obtained from a path of length n by adding an edge between its endpoints.

Furthermore, all the graphs in this note will be *connected*. A graph G is *connected* if between any two vertices $u, v \in V(G)$, there is a path $P \subset G$ with endpoints u and v. A graph H is called *disconnected* if it is not connected. A *tree* of size n is a connected graph T with n vertices that does not contain any cycle as a subgraph (i.e. there is no cycle C such that $V(C) \subset V(T)$ and $E(C) \subset E(T)$). We leave the following two observations as exercises for the reader:

Fact 1. For any tree T, $|E(T)| = |V(T)| - 1$.

Fact 2. If G is a connected graph, then there is some tree T such that $V(T) \subset V(G)$ and $E(T) \subset E(G)$ such that $|V(T)| = n$. In other words, every connected graph G has a *spanning tree*.

It turns out spanning trees are extremely beautiful objects that combinatorialists and computer scientists appreciate very much for their structure. There are incredibly many properties and problems that one can think of, and some end up being quite challenging. One such famous puzzle is even older than graph theory itself.

Cayley's Theorem. The number of spanning trees of a complete graph on n vertices is n^{n-2}.

Here, by a *complete graph on n vertices* we mean a graph K_n with n vertices where $E(G)$ is the set of all possible pairs $V(K_n) \times V(K_n)$. In particular, note that $|E(G)| = \binom{n}{2}$, since we are only considering *simple graphs* that do not have loops or multiple edges.

There are many remarkable proofs of Cayley's result from 1889. Martin Aigner and Günter Ziegler even dedicated a chapter in their famous *Proofs from the Book*, where they survey some motivating small cases and four different proofs. Here, we begin by revisiting the fourth one due to Jim Pitman, which uses double counting, since it is the most natural one in our opinion, it is from The Book, and can be slightly simplified (especially notation-wise). Afterwards, we present a second more high-tech proof using branching processes that the second author found out about from [3]. The latter is quite unknown in literature and it is rarely brought up into the discussion because it is significantly more sophisticated than other proofs. Nonetheless, we hope to shed some light on it, since it's quite beautiful!

Double Counting Proof. We begin by saying that a directed tree is a rooted graph that has a simple path from the root to every vertex in the graph.

The quantity we will compute in two different ways is the number τ of different sequences of directed edges that can be added to an empty graph on n vertices to yield a rooted tree.

First way: we start with a spanning tree on the empty graph (T_n choices). Pick a root for the tree (n choices), and note that given the root, the direction of every edge is fully determined. By Fact 1, there are $n-1$ directed edges to insert in any order in our graph ($(n-1)!$ ways to order them). In total, there are $T_n \cdot n \cdot (n-1)!$ different sequences of directed edges to add in a graph so as to form a directed rooted tree; so

$$\tau = T_n \cdot n \cdot (n-1)!$$

Second way: we start with an empty graph on n vertices and we add the $n-1$ directed edges one by one until we construct a rooted tree spanning the n vertices. At every step $k = 1, \dots, n-1$, let n_k be the number of possible directed edges from which to choose the edge to add. Note that an edge is completely defined when its tail and head are picked. Hence, the number of possible directed edges at every step is the product of the number of ways to choose a tail and the number of ways to choose a head. Initially, we have a forest of n empty rooted trees, hence $n_1 = n(n-1)$ (since we have n ways to choose the tail and $n-1$ ways to choose the head. Next, note that the graph is a collection of $n-1$ trees (initially there were n trivial trees, but now we joined two vertices together). When we choose the k-th edge, we can still pick any of the n vertices as the tail, but now we only have $n-k$ options for the head (since the edge has to be pointed to the root of a tree that is different from the one where the tail belongs). Consequently, $n_k = n(n-k)$ for all $k = 1, \dots, n-1$. It thus follows that there are

$$\tau = \prod_{k=1}^{n} n_k = n^{n-1}(n-1)!$$

ways to add the edges.

Putting the two counts together, we get that $T_n = n^{n-2}$, as desired. □

We now move on to the second proof, for which we need a little bit of background. We will cover everything below, however we will assume a little familarity with basic probability theory. We follow closely the material from [3].

Branching Processes Overture. A *branching process* is the simplest possible model for a population evolving in time. Suppose each organism independently gives birth to a number of children with the same distribituion. We denote the offspring distribution by $\{p_i\}_{i=0}^{\infty}$, where p_i denotes the probability that an individual has i children. We denote by

Z_n the number of individuals in the nth generation, where by convention we let $Z_0 = 1$. Then Z_n satisfies the recursion relation

$$Z_n = \sum_{i=1}^{Z_{n-1}} X_{n,i},$$

where $\{X_{n,i}\}_{n,i\geq 1}$ is a doubly infinite array of i.i.d random variables. We will often write X for the offspring distribution, so that $\{X_{n,i}\}_{n,i\geq 1}$ is a doubly infinite array of independent and identically distribution (briefly *i.i.d.*) random variables with $X_{n,i} \sim X$ for all n, i.

One of the major results of branching processes is that when $E[X] \leq 1$, the population dies out with probability one (unless $X_{1,1} = 1$ with probability one), while if $E[X] > 1$, there is a nonzero probability that the population will not become extinct.

For our proof of Cayley's theorem, we will only need branching processes with Poisson offspring distributions. More precisely, we will assume that the $X_{n,i} \sim X$ are i.i.d. Poisson random variables with mean λ. Even, more precisely, the random variables $X_{n,i}$ are all independent and they all have the same probability mass function

$$f_{X_{n,i}}(k) = P(X_{n,i} = k) = \frac{\lambda^k}{k!}e^{-\lambda},$$

for all $k \in \{0, 1, 2, \ldots\}$.

The remarkable thing about Poisson branching processes is that the total progreny $T^* = \sum_{n=0}^{\infty} Z_n$ (the number of all individuals since the beginning) has a very nice distribution.

Fact 3. For a branching process with i.i.d. offspring X, where X has a Poisson distribution with mean λ,

$$P(T^* = n) = \frac{(\lambda n)^{n-1}}{n!}e^{-\lambda n}.$$

The proof is easy but it requires a certain level of comfort with probability theory. We refer to [3] for the full details. We are now ready for the second proof of Cayley's theorem.

Branching Process Proof. Note first that Cayley's result is equivalent with the fact that the number of labeled trees of size n is equal to n^{n-2}, and so it suffices to prove this version. Furthermore, note that it is convenient to label the vertices of a tree in terms of *words*. These words arise inductively as follows. The root is the word $\emptyset$. The children of the root are the words $1, \ldots, d_\emptyset$, where for a word w, we let d_w denote the number of children of w. The children of 1 are $11, 12, \ldots, 1d_1$, etc. A tree is then uniquely represented by its set of words. For example, the word 1123 represents the third child of the second child of the first child of the first child of the root.

Two trees are the same if and only if they are represented by the same set of words. We obtain a branching process when the variables $(d_w)_w$ are equal to a collection of i.i.d. random variables. For a word w, we let $|w|$ be its length, where $|0| = 0$. The length of a word w is the number of stepts the word is away from the root, and equals its generation.

Let $\mathcal{T}$ denote the family tree of a branching process with Poisson offspring distribution with parameter $\lambda = 1$. By the first paragraph, note that the probability of obtaining a given tree t is given by

$$P(\mathcal{T} = t) = \prod_{w \in t} P(\xi = d_w),$$

where ξ is a Poisson random variable with parameter 1, and d_w is the number of children of the word w in the tree t. Taking $|V(t)| = n$, for a Poisson branching process we thus get

$$P(\xi = d_w) = \frac{e^{-1}}{d_w!},$$

hence

$$P(\mathcal{T} = t) = \frac{e^{-n}}{\prod_{w \in t} d_w!}.$$

We note that the above probability is the same for each tree with the same number of vertices of degree k for each k.

Conditionally on having total progeny $T^* = n$, we introduce a labeling as follows. We give the root label 1, and give all other vertices a label from the set $\{2, \ldots, n\}$, giving a labeled tree on n vertices. Given $\mathcal{T}$, there are precisely

$$\prod_{w \in \mathcal{T}} d_w!$$

possible ways to put down the labels that give rise to the same labeled tree, since permuting the children of any vertex does not change the labeled tree. Also, the probability that w receives label i_w with $i_\emptyset = 1$ is precisely equal to $1/(n-1)!$, where $n = |\mathcal{T}|$. For a labeled tree ℓ, let t_ℓ be any tree, i.e. a collection of words, from which ℓ can be obtained by labelling the vertices. Then, the probability of obtaining a given labeled tree ℓ of arbitrary size equals

$$P(\mathcal{L} = \ell) = P(\mathcal{T} = t_\ell) \cdot \frac{\prod_{w \in t_\ell} d_w!}{(|\ell| - 1)!} = \frac{e^{-|\ell|}}{\prod_{w \in t_\ell} d_w!} \cdot \frac{\prod_{w \in t_\ell} d_w!}{(|\ell| - 1)!} = \frac{e^{-|\ell|}}{(|\ell| - 1)!}.$$

Therefore, conditionally on $T^* = n$, the probability of a given labeled tree $\mathcal{L}$ of size n equals

$$P(\mathcal{L} | |\mathcal{L}| = n) = \frac{P(\mathcal{L} = \ell)}{P(|\mathcal{L}| = n)}.$$

By Fact 3 however,

$$P(|\mathcal{L}| = n) = P(T^* = n) = \frac{e^{-n} n^{n-2}}{(n-1)!}.$$

Thus, it follows that

$$P(\mathcal{L} | |\mathcal{L}| = n) = \frac{P(\mathcal{L} = \ell)}{P(|\mathcal{L}| = n)} = \frac{e^{-n}}{(n-1)!} \cdot \frac{e^{-n} n^{n-2}}{(n-1)!} = \frac{1}{n^{n-2}}.$$

The obtained probability is uniform over all labeled trees. Therefore, the number of labeled trees equals

$$P(\mathcal{L} | |\mathcal{L}| = n)^{-1} = n^{n-2}.$$

This not only proves Cayley's theorem, but also gives an explicit construction of a uniform labeled tree from a Poisson branching process.

References

[1] A. Cayley, "A theorem on trees", Quart. J. Math 23: 376378, 1889.

[2] M. Aigner, G. Ziegler, "Proofs from THE BOOK" (4th ed.), Berlin, New York: Springer-Verlag.

[3] R. v. d. Hofstad, "Random Graphs and Complex Networks" (Lecture Notes from Eindhoven University of Technology), October 20, 2014.