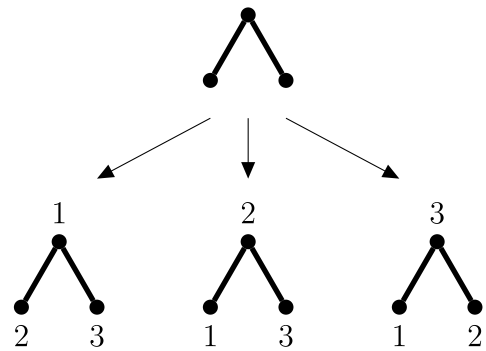

Figure 1: The three labelled trees that arise from a path of length 2.

A *labelled tree* T is a graph G that is a tree together with the *labelling* of its vertices, i.e., a bijection from $V(G)$ to $\{1, 2, \ldots, n\}$. See an example of labelled trees in Figure 1. The purpose of this note is to determine the total number of labelled trees on n vertices, which we denote by t_n.

Theorem 1 (Cayley's formula)**.** *For every $n \in \mathbb{N}$, the number of labelled trees on n vertices t_n is equal to n^{n-2}.*

In order to determine the formula for t_n, we consider a closely related problem of counting the number of so-called *rooted labelled trees.* A rooted labelled tree is a labelled tree T together with one marked vertex $v \in V(T)$, which we call the *root.* We denote a labelled tree T with the root v as T_v. Let t'_n be the number of labelled rooted trees on n vertices. Since for every labelled tree on n vertices there are exactly n choices for the root, we conclude that $t'_n = t_n \cdot n$.

Actually, it would be convenient to compute the number of rooted labelled trees on n vertices in a slightly twisted way, where all the edges are *oriented* – one of the endpoints of an edge will become the *source*, and the other will be the *target*; see Figure 2. Formally, an orientation of the edges of a labelled tree T is a map $o : E(T) \to V(T)$ such that $o(e)$ is the target of e for every $e \in E(T)$.

However, we will be only interested in very specific orientations of the edges of T_v, namely in orientations where every vertex $x \neq v$ is the target of exactly one edge, and the root v is the source of all of its incident edges. Such orientations will be called *outroot* orientations. As it turns out, each rooted tree T_v has exactly one outroot orientation.

Figure 2: An orientation of an edge.

Lemma 1. *Let T_v be a labelled tree with the root v. There exists a unique outroot orientation of the edges of T_v.*

Proof. Let n be the number of vertices of T_v, i.e., $n = |V(T_v)|$. If T_v has only one vertex ($n = 1$), then the vertex must be the root and the statement immediately follows.

For the rest of the proof, we will assume $n \geq 2$. Since every tree with at least two vertices has at least two leaves, we know that T_v contains at least one leaf that is different from v. Let $w \in V(T_v) \setminus \{v\}$ be one such a leaf, and let $z \in V(T_v)$ be the unique vertex adjacent to w in T_v. Note that z might be equal to v.

First, let us show that there exists an outroot orientation of T_v. We proceed by induction on n. We already know that the base case $n = 1$ holds, so we move to the induction step. Let $T'_v := T_v - w$ be the $(n-1)$-vertex tree with the root v. By the induction hypothesis, it has an outroot orientation o' such that every vertex of T'_v except v is the target of exactly one edge. Let o be the following orientation of the edges $e \in E(T_v)$:

- if $e \neq \{w, z\}$, then orient e according to its orientation in o',
- if $e = \{w, z\}$, then orient e so that z is the source and w is the target.

It holds that every vertex of T_v except v is the target of exactly one edge, which is what we wanted to prove.

It remains to show that there exists only one such orientation of T_v. Suppose there are two outroot orientations of T_v, say o_1 and o_2. Our aim is to show that $o_1 = o_2$. Again, we proceed by induction on n and, again, let $T'_v := T_v - w$. Since both o_1 and o_2 are outroot orientations, the leaf w is the target of $\{w, z\}$ in both o_1 and o_2. Therefore, o_1 induces an outroot orientation of T'_v, which we denote by o'_1. Analogously, o_2 induces an outroot orientation of T'_v, which will be denoted by o'_2. By the induction hypothesis, $o'_1 = o'_2$. Since $\{w, z\}$ is oriented in both o_1 and o_2 so that its target is w, it holds that $o_1 = o_2$. $\square$

A rooted labelled tree T_v together with its unique outroot orientation will be called *outroot-oriented labelled tree*, and denoted by $\overrightarrow{T_v}$. In the rest of the note, we focus on counting outroot-oriented labelled trees. By the previous lemma, the number of all such trees on n vertices is equal to t'_n.

All right, before we (finally) move to the step where we perform the promised counting, we alter our problem once more. For any outroot-oriented labelled tree $\overrightarrow{T_v}$ on n vertices, we fix an arbitrary numbering of its edges, i.e., a bijection $b : E\left(\overrightarrow{T_v}\right) \to \{1, \ldots, n-1\}$.

Let t''_n be the number of outroot-oriented labelled trees on n vertices with a fixed numbering of its edges. Since for every tree $\overrightarrow{T_v}$ there are exactly $(n-1)!$ such numberings, it immediately follows that $t''_n = t'_n \cdot (n-1)!$ which in turn means that $t''_n = t_n \cdot n!$. So in order to prove Theorem 1, it is enough to establish the following lemma.

Lemma 2. $t''_n = n^{n-1} \cdot (n-1)!$.

Proof. Fix a number of vertices n. We design a simple procedure that will start with the empty graph with the vertex set $\{1, 2, \ldots, n\}$ and build up a root-oriented labelled tree by adding oriented edges in $n-1$ steps (one edge per step). The order in which the edges are added then determines their numbering, and the root is uniquely determined by being the unique vertex that is not the target of any edge. Since different runs of the procedure lead to different outroot-oriented labelled trees, and every tree can be constructed in this way, it follows that the total number of choices in this procedure is exactly t''_n.

During the whole procedure, we maintain the two key properties that are clearly necessary for succeeding:

- the so-far added edges form an acyclic graph,
- every vertex is the target of at most one edge.

Let us first look on how many choices we have for placing the first edge. The source of the edge can be chosen in n different ways, and once the source s is chosen, we have in total $n-1$ choices for the target (it can be any vertex different from s).

Suppose already k edges have been placed and we are in the $(k+1)$-th step of the procedure, i.e., we are about to place the $(k+1)$-th edge. Since the current graph is acyclic and contains k edges, it has exactly $n-k$ connected components $V_1, V_2, \ldots, V_{n-k}$. Let v_i be the number of vertices in V_i for all $i \in \{1, \ldots, n-k\}$. The edges inside each connected component have to span a tree, i.e., each connected component V_i contains exactly $v_i - 1$ edges. Every edge inside V_i has exactly one target, but no vertex in V_i can be the target for two (or more) edges. Therefore, each connected component V_i contains $v_i - 1$ vertices that are being the target of some edge, which in other words means the component contains exactly one vertex that is still allowed to become a target.

The observation above makes the counting of the possible placements of the $(k+1)$-th edge very easy: first, choose the source s of the edge – there are n ways to do so. Let V_s be the connected component where s belongs to. For the target of the edge, we cannot choose any vertex from V_s, because that would create a cycle. On the other hand, in every connected component other than V_s, we are allowed to choose exactly one vertex as the target (otherwise we would have a vertex that is the target of two edges). That means, there are exactly $n - k - 1 = n - (k+1)$ choices for the target of the $(k+1)$-th edge.

By putting the numbers together, we can see that the number of choices of the source and the target of the k-th edge is equal to $n(n-k)$ for all $k = \{1, 2, \ldots, n-1\}$. Therefore,

$$t_n \cdot n! = t''_n = \prod_{k=1}^{n-1} n(n-k) = n^{n-1} \cdot \prod_{k=1}^{n-1} (n-k) = n^{n-1} \cdot (n-1)!.$$

□

We conclude this note by relating t_n, i.e., the number of n-vertex labelled trees, to the number of n-vertex unlabelled trees, which we denote by u_n.

Corollary 1. *The number of unlabelled trees on n vertices u_n is at least $e^{n-1}/n^{5/2}$.*

Proof. Recall each n-vertex labelled tree can be represented as a pair (G, ℓ), where G is an n-vertex unlabelled tree, and $\ell : V(G) \to \{1, \ldots, n\}$ is a bijection. Every unlabelled tree can lead to at most $n!$ different labelled trees, which implies the bound

$$n^{n-2} = t_n \le u_n \cdot n!.$$

By Stirling's formula $n! \approx \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n$. Since $\sqrt{2\pi} < e$, we conclude that $n! < n^{n+1/2}/e^{n-1}$. Therefore,

$$u_n \ge \frac{n^{n-2}}{n!} > \frac{e^{n-1}}{n^{5/2}}.$$

□

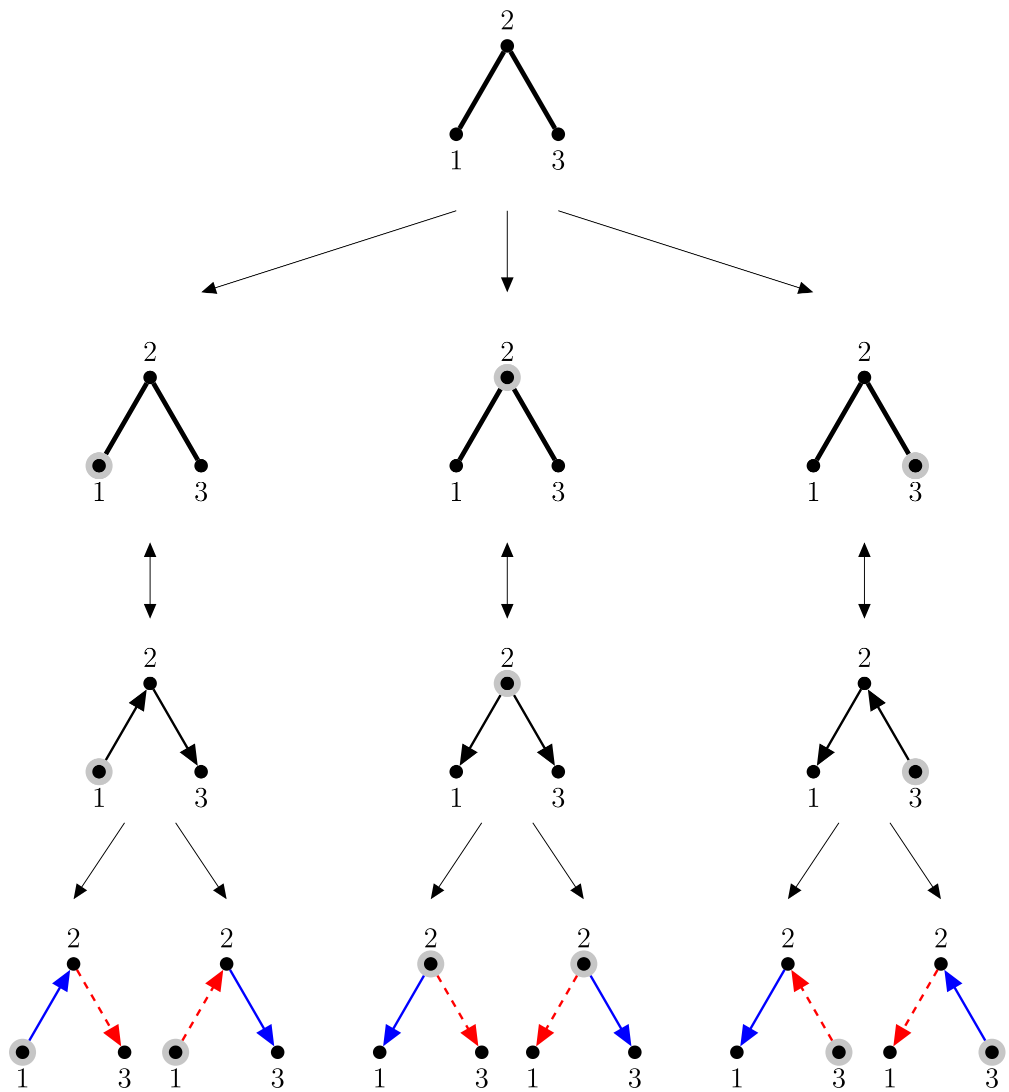

Figure 3: Illustration of the double-counting steps in the proof of Theorem 1. The solid blue edges represents the first edge in the numbering, and the dashed red edges represents the second edge.