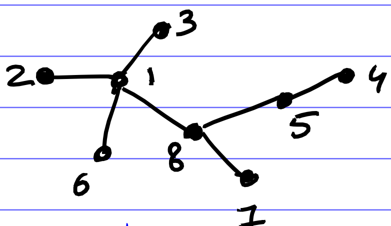
Counting labelled trees

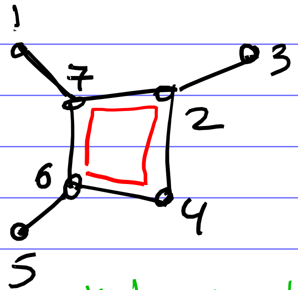
A **labelled tree** T is a simple graph on n vertices labelled by $1, 2, \dots, n$ such that

- T is a connected graph
- T has no cycles (i.e. closed walks)

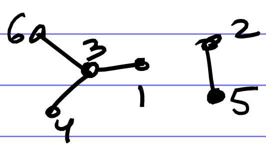
Examples:



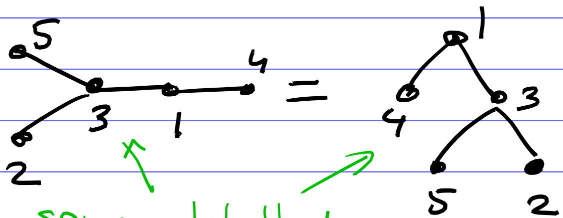
a labelled tree



not a tree
(It has a cycle)

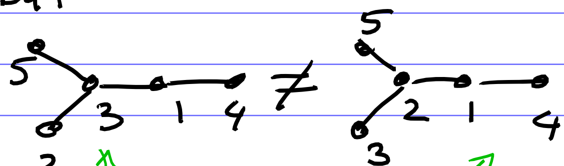


not a tree
(It is disconnected)

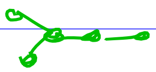


same labelled trees

but



different labelled trees

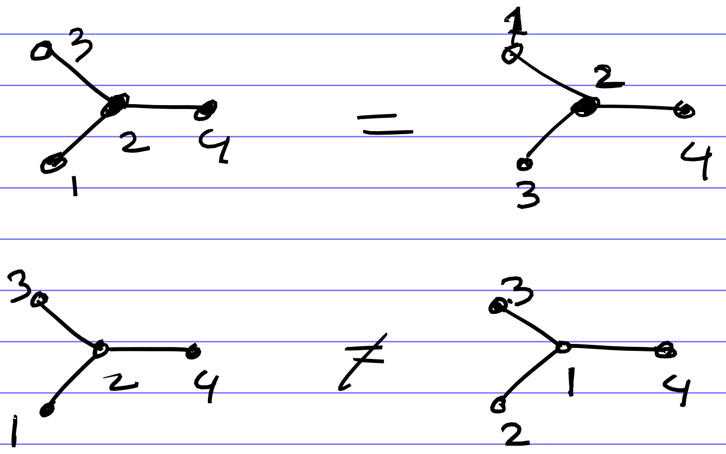
(If we remove the labels we get the same **unlabelled tree** )

Remark. There are many different ways to draw the same labelled tree on the plane. A drawing of a tree is <u>not</u> a part of its structure.

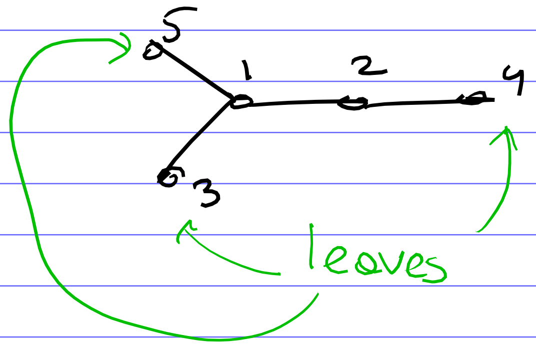
(In contrast, for a <u>binary</u> tree, its drawing is a part of its structure)

Switching two labels in a labelled tree may or may not produce a different labelled tree:

Example.



Def. A <u>leaf</u> in a tree is a vertex of degree 1



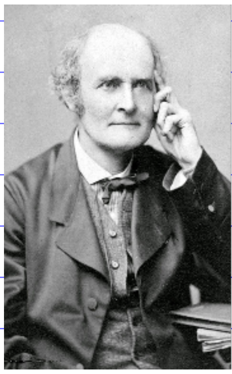
Lemma Any tree with $n \geq 2$ vertices has at least 2 leaves.

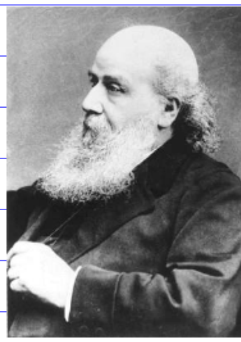
Lemma Any tree with n vertices has exactly $n-1$ edges.

Both lemmas can be easily proved by induction.

Cayley's Formula

Theorem The number of labelled trees on n vertices equals n^{n-2}.





Arthur Cayley
1821 - 1895

Carl Wilhelm Borchard
1817 - 1880

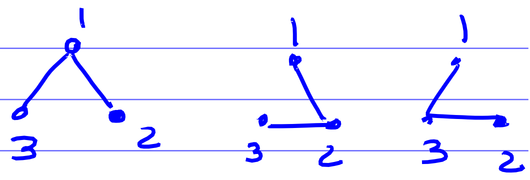
James Joseph Sylvester
1814 - 1897

Remark. This formula is named after A. Cayley 1889. But it was known earlier to Sylvester 1857 and Borchard 1860.

Actually, Cayley's paper refers to the earlier paper of Borchard. Cayley proved an extension of Cayley's formula.

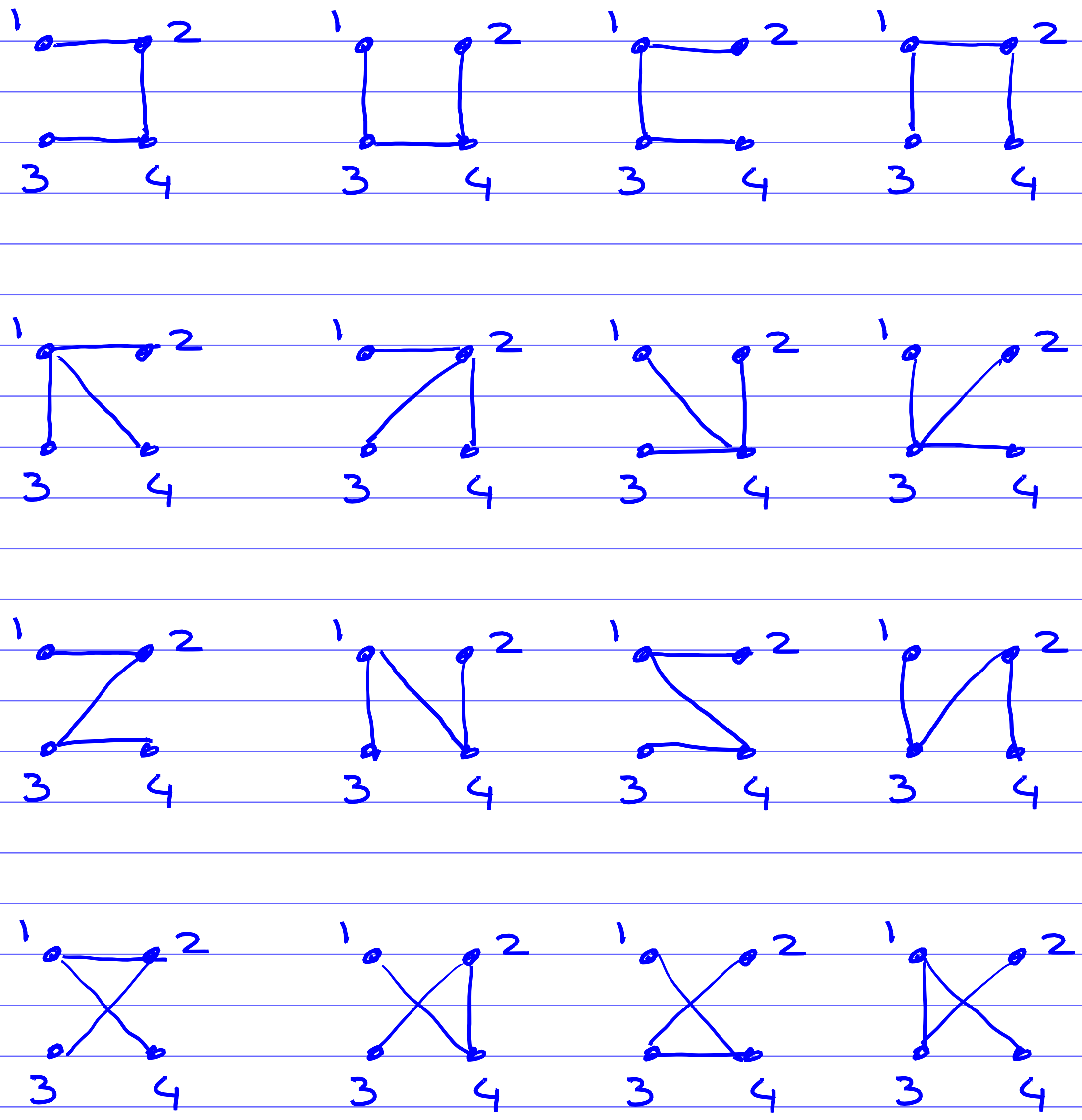
Examples: $n = 3$

$3^1 = 3$ labelled trees:



$n = 4$

$4^2 = 16$ labelled trees:



How to prove Cayley's formula?

There are many known proofs, and we'll discuss some of them...

Let's first try to prove it by induction on n.

But n^{n-2} does not satisfy a simple recurrence.

It is not clear how to prove Cayley's formula (# trees $= n^{n-2}$) by induction.

However we give a simple inductive argument to a generalization of Cayley's formula.

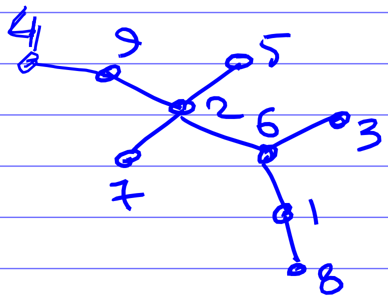
A Multivariate extension of Cayley's formula.

For a labelled tree T on n vertices, define the monomial in $x_1, \dots, x_n$

$$x^T := x_1^{\deg_T(1)-1} x_2^{\deg_T(2)-1} \cdots x_n^{\deg_T(n)-1},$$

where $\deg_T(v)$ is the <u>degree</u> of vertex v in T (i.e. # edges incident to v).

<u>Example.</u> $T =$

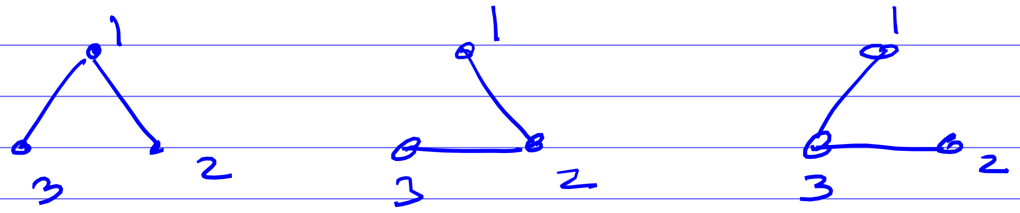


$$x^T = x_1 x_2^3 x_4^0 x_5^0 x_6^2 x_7^0 x_8^0 x_9^1$$

Let

$$F_n(x_1, x_2, \dots, x_n) := \sum_{\substack{T \text{ labelled} \\ \text{tree on} \\ n \text{ vertices}}} x^T$$

<u>Example</u> $F_3 = x_1 + x_2 + x_3$



$x_1^1 x_2^0 x_3^0$ $\quad$ $x_1^0 x_2^1 x_3^0$ $\quad$ $x_1^0 x_2^0 x_3^1$

Theorem $\boxed{F_n(x_1,\dots,x_n) = (x_1+\dots+x_n)^{n-2}}$

for $n \geq 2$.

Clearly, if we set $x_1 = \dots = x_n = 1$, we get Cayley's formula n^{n-2}

Equivalently the theorem can be formulated as follows:

For a given sequence $d_1, \dots, d_n$ $\left(d_i \geq 1, \ \sum_{i=1}^{n} d_i = 2n-2\right)$ the number of labelled trees T on n vertices such that $\deg_T(\sigma) = d_\sigma$ for $\sigma = 1, \dots, n$ equals the multinomial coefficient

$$\binom{n-2}{d_1-1 \ \ d_2-1 \ \dots \ d_n-1}.$$

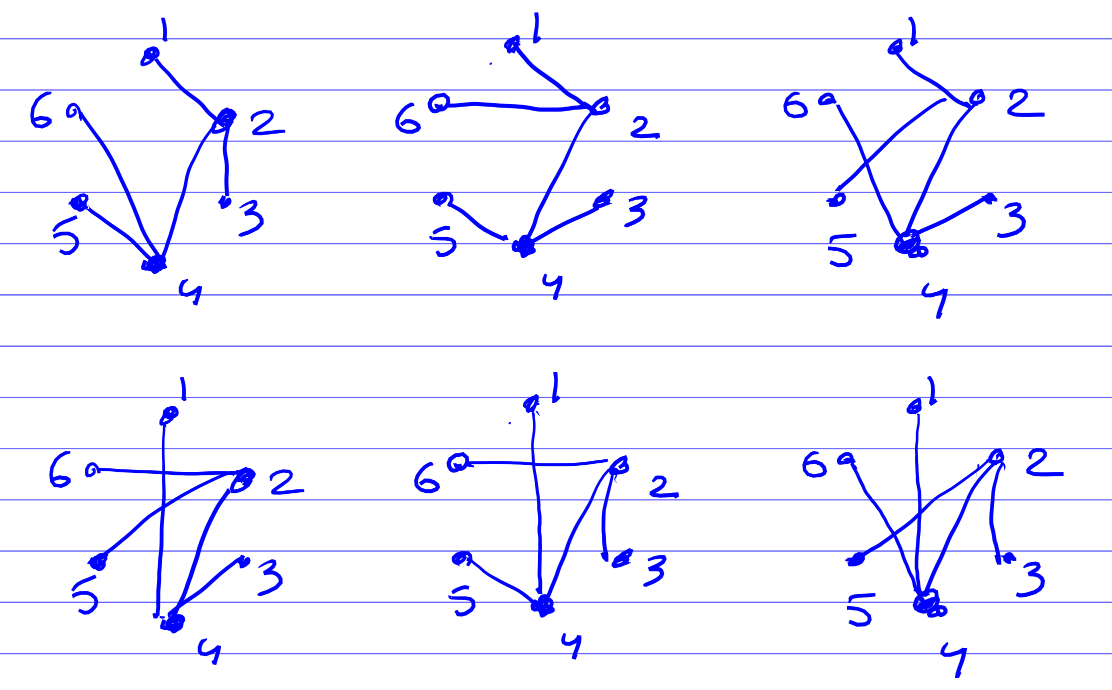
Lemma We have $\sum_{\sigma=1}^{n} \deg_T(\sigma) = 2(n-1)$.

Proof In this sum we count each edge 2 times. $\square$

Example $n = 6$

$(d_1, \dots, d_6) = (1, 3, 1, 3, 1, 1)$

trees with these degrees equals $\binom{4}{0,2,0,2,0,0} = \frac{4!}{2!\,2!} = 6$



6 labelled trees with degree vector $(1,3,1,3,1,1)$.

Proof (Rényi 1967)

Alfréd Rényi
(1921 - 1970)

Let $G_n(x_1, \ldots, x_n)$

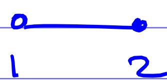
$$:= F_n(x_1, \ldots, x_n) - (x_1 + \ldots + x_n)^{n-2}.$$

Let's prove that

$G_n(x_1, \ldots, x_n) = 0$ by induction on n.

Base: $n = 2$

$$F_2(x_1, x_2) = x_1^0 \, x_2^0 = 1$$



$$G_2(x_1, x_2) = F_2(x_1, x_2) - (x_1 + x_2)^0$$

$$= 1 - 1 = 0.$$ ✓

Induction Step $n \geq 3$

Assume that $G_{n-1}(x_1, \ldots, x_{n-1}) = 0$.

Let's us specialize $x_n = 0$ in $G_n(x_1, \dots, x_n)$:

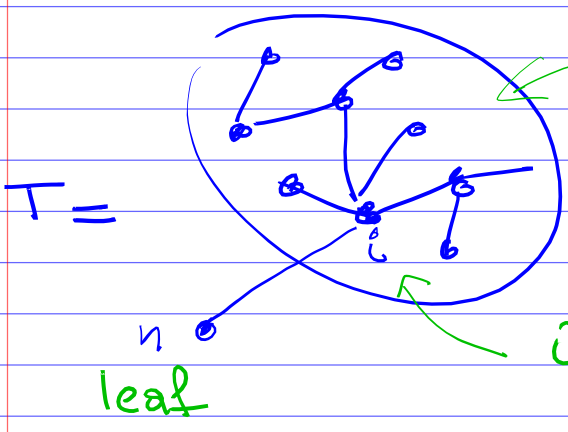
$$G_n(x_1, \dots, x_{n-1}, 0) =$$

$$= F_n(x_1, \dots x_{n-1}, 0) - (x_1 + \dots + x_{n-1})^{n-2}$$

$$= \sum_{\substack{T \text{ labelled} \\ \text{tree on } n \\ \text{vertices}}} x_1^{\deg(1)-1} \cdots \underline{0}^{\deg(n)-1} - (x_1 + \dots + x_{n-1})^{n-2}$$

only trees T such that vertex n is a leaf make a contribution: $0^0 = 1$
$\qquad 0^a = 0 \quad \forall\, a \geq 1$)

$$= \sum_{\substack{T: \\ n \text{ is a leaf}}} x_1^{\deg(1)} \cdots x_{n-1}^{\deg(n-1)} - (x_1 + \dots + x_{n-1})^{n-2}$$



any labelled tree on vertices $1, 2, \dots, n-1$

i is any vertex $1, 2, \dots, n-1$

$$= (x_1 + \dots + x_{n-1}) F_{n-1}(x_1, \dots x_{n-1})$$
$$- (x_1 + \dots + x_{n-1}) \cdot (x_1 + \dots + x_{n-1})^{n-3}$$

$$= (x_1 + \dots + x_{n-1}) \, G_{n-1}(x_1, \dots, x_{n-1})$$

$= 0$ by the inductive hypothesis.

This implies that

$$G_n(x_1, \dots, x_n)\Big|_{x_i = 0} = 0$$

for any $i = 1, \dots, n$, because $G_n(x_1, \dots, x_n)$ is symmetric with respect to permutations of the variables $x_1, \dots, x_n$.

Now we know

- $G_n(x_1, \dots, x_n)$ is a polynomial in $x_1, \dots, x_n$ of degree $\le n-2$.

- $G_n(x_1, \dots, x_n)\big|_{x_i=0} = 0$ for any $i = 1, \dots, n$.

These two facts imply that $G_n(x_1, \dots, x_n) = 0$.

Indeed, if $G_n(x_1, \dots, x_n) \neq 0$ and it contains some monomial $x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$ with a non-zero coefficient. Then $a_1 + \dots + a_n < n$. So by the pigeon hole principle $\exists\, i$ s.t. $a_i = 0$. So this monomial would survive in the specialization $G_n\big|_{x_i=0}$.

But we proved that $G_n\big|_{x_i=0}$ should be 0.

So we deduced that $G_n(x_1, \dots, x_n) = 0$ for any $n \geq 2$. Q.E.D.

The previous proof is nice.
But it looks like a miracle that we obtained the formula $(x_1+\dots+x_n)^{n-2}$ almost out of nothing just by making some trivial observations about polynomial.

Can we give a combinatorial proof of Cayley's formula?

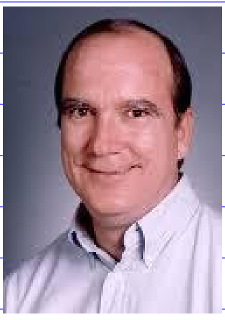
Actually, there are several nice combinatorial proofs:

- Prüfer's code, 1918
- Eğecioğlu – Remmel, 1986; ~ Joyal 1981

Let's start with this proof.

Proof by Eğecioğlu-Remmel, 1986

based on Joyal 1981



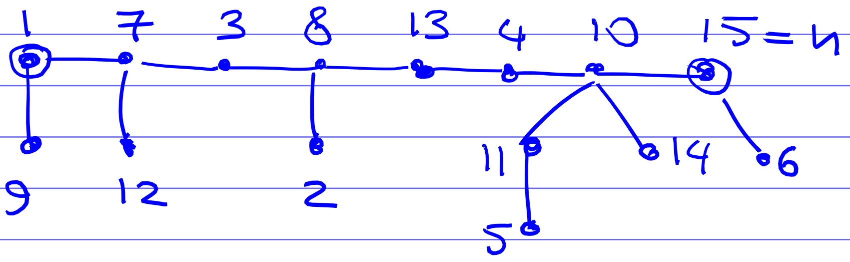
André Joyal   Ömer Eğecioğlu   Jeff Remmel

Clearly n^{n-2} is the number of all maps from an $(n-2)$-element set to an n-element set.

Let's construct a bijection between all labelled trees T on n vertices and all maps $f: \{1,\dots,n\} \to \{1,\dots,n\}$ such that $f(1)=1$, $f(n)=n$.

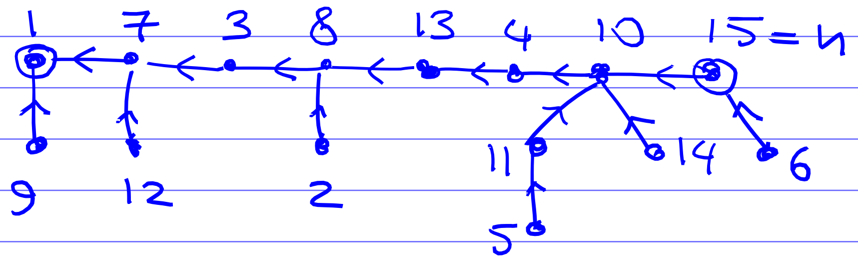
Example $n=15$

T =



Here we've drawn a tree T so that the path P from the maximal vertex n to vertex 1 is on the top; and some branches (small trees) are attached to some vertices on this path.

Let us orient all edges of T towards the vertex 1:

$T =$ 

Consider the path P from n to 1:

$$P = 15 \to 10 \to 4 \to 13 \to 8 \to 3 \to 7 \to 1$$

Find all left-to-right minima in this sequence, i.e. entries which are less than all previous entries.

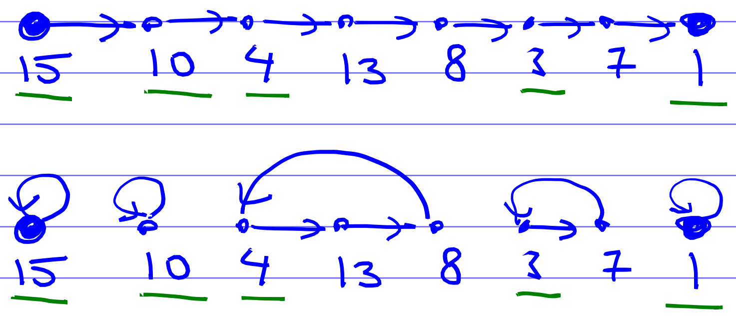
related to records

Replace P by the product of cycles starting at the left-to-right minima.

$\underline{15}, \underline{10}, \underline{4}, 13, 8, \underline{3}, 7, \underline{1}$

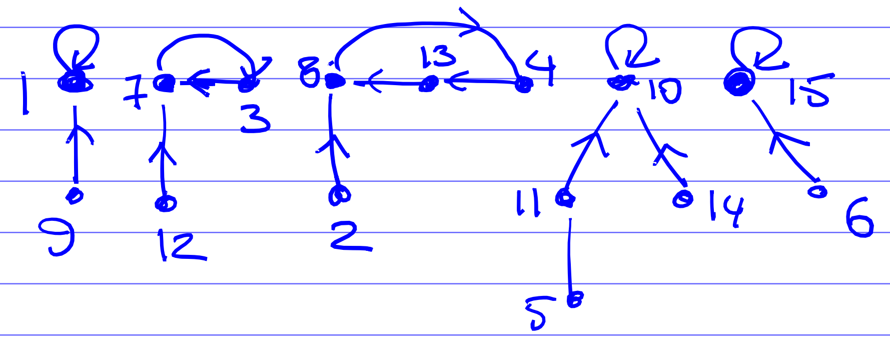
$\rightsquigarrow (\underline{15})\ (\underline{10})\ (\underline{4}, 13, 8)\ (\underline{3}, 7)\ (\underline{1})$

or



Remark. We already used a similar transformation when we discussed statistics on permutations and showed that # cycles in a permutation is equidistributed with # records.

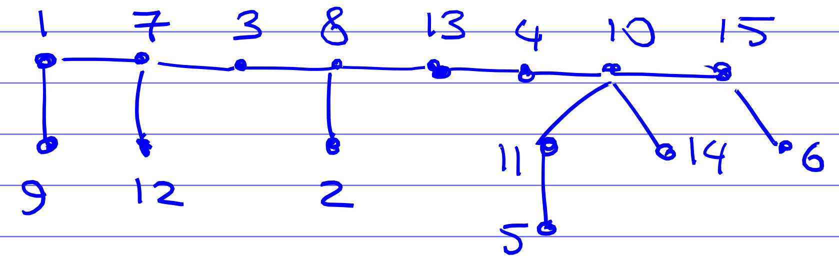
Let's combine this collection of cycles with other parts of T (the "little branches" at vertices of P)

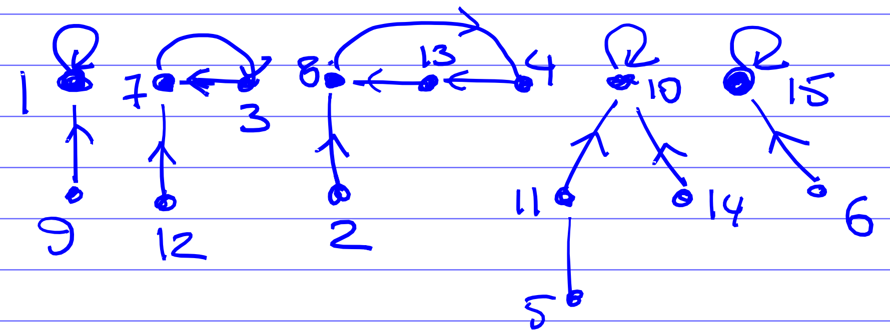


This directed graph represents a map $f : [n] \to [n]$ such that $f(1) = 1$ & $f(n) = n$.

This gives the needed bijection $T \Longleftrightarrow f$.

$T =$





$$f: \begin{array}{ccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 1 & 8 & 7 & 13 & 11 & 15 & 3 & 4 & 1 & 10 & 10 & 7 & 8 & 10 & 15 \end{array}$$

Clearly, this is a bijection.

$$\left\{ \begin{array}{l} \text{all labelled} \\ \text{trees } T \\ \text{on } n \text{ vert.} \end{array} \right\} \xleftrightarrow{\text{bij.}} \left\{ \begin{array}{l} \text{all maps} \\ f : [n] \to [n] \\ \text{s.t. } f(1) = 1 \\ f(n) = n \end{array} \right\}$$

Indeed, the inverse map is given, as follows

- draw a map f as a directed graph.
- consider all directed cycles in this graph.
- Mark the minimal vertices in all cycles
- Arrange the cycles so that their minimal elements go in the decreasing order
- Replace all cycles by one path from n to 1.
- Ignore the directions of the edges (i.e. erase the arrows).
- We obtain a labelled tree.

$\square$