Statistics of parking functions and labeled forests

${\sf Mei}\;{\sf Yin}^1$

Department of Mathematics, University of Denver

June 21, 2024

¹Joint work with Stephan Wagner. Presentation supported by NSF-AWM. $= -9 \circ 0$

Parking functions: An introduction

Parking functions were introduced by Konheim and Weiss (1966) under the name of "parking disciplines," in their study of the hash storage structure, and have since found many applications in combinatorics, probability, algebra, and computer science.

Consider a parking lot with *n* parking spots placed sequentially along a one-way street. A line of $m \le n$ cars enters the lot, one by one. The *i*th car drives to its preferred spot π_i and parks there if possible; if the spot is already occupied then the car parks in the first available spot after that. The list of preferences $\pi = (\pi_1, \ldots, \pi_m)$ is called a parking function if all cars successfully park. (This generalizes the term classical parking function where m = n.)



▲ロ ▶ ▲ □ ▶ ▲ □ ▶ ▲ □ ▶ ▲ □ ▶ ● ○ ○ ○

n = 1: 1

n = 2: 11, 12, 21

n = 3: 111, 112, 121, 211, 113, 131, 311, 122, 212, 221, 123, 132, 213, 231, 312, 321 We denote by PF(m, n) the set of parking functions π with m cars and n parking spots, and by $\mathcal{F}(m, n)$ the set of rooted forests Fwith n + 1 vertices and m edges (equivalently, n - m + 1 distinct tree components) such that a specified set of n - m + 1 vertices are the roots of the different trees. We label the roots of F by $\{01, 02, \ldots, 0(n - m + 1)\}$ and the non-root vertices by $\{1, 2, \ldots, m\}$. We further denote by $\mathcal{T}(n)$ the set of rooted trees \mathcal{T} on the vertex set $\{0, 1, \ldots, n\}$ with root 0.

Note that $|PF(n, n)| = |\mathcal{T}(n)|$ and more generally $|PF(m, n)| = |\mathcal{F}(m, n)|$.

Many bijections between the two combinatorial objects exist; we will explore one later!

Pigeonhole principle

It is well-known and easy to see that π is a parking function if and only if

$$\#\{k: \pi_k \leq i\} \geq m - n + i$$
, for $i = n - m + 1, \dots, n$.

Equivalently, if $\lambda_1 \leq \cdots \leq \lambda_m$ is the (weakly) increasing rearrangement of π_1, \ldots, π_m , then π is a parking function if and only if $\lambda_i \leq n - m + i$ for $1 \leq i \leq m$.

ション ふゆ アメリア メリア しょうくしゃ

Two immediate observations:

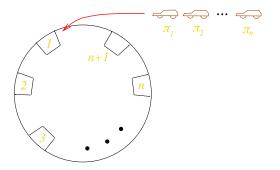
- Parking functions are invariant under the action of the symmetric group \mathfrak{S}_m permuting the *m* cars, that is, permuting the list of preferences π .
- When some π_i takes values in the set {1, 2, ..., n m + 1}, changing π_i to any other value in the set {1, 2, ..., n m + 1} has no effect on π being a parking function.

One less immediate observation:

Every parking function π in PF(m, n) can be uniquely decomposed into an arbitrary function π_a : A → [n - m + 1] on a set A ⊆ [m] of cardinality s and a function that is equivalent to a parking function π_p in PF(m - s, m - 1).

Circular symmetry

Pollak's ingenious circle argument: Assign $m \le n$ cars on a circle with n + 1 spots. Those car assignments where spot n + 1 is left empty after circular rotation give valid parking functions.



▲ロ ▶ ▲ □ ▶ ▲ □ ▶ ▲ □ ▶ ▲ □ ▶ ● ○ ○ ○

Significance of our approach

Unlike Pollak's original argument where the parking statistics are studied after all cars have parked, we investigate the individual parking statistics for each car the moment it is parked on the circle. This "seemingly small step ahead" provides a lot more useful information about the parking scenario.

Main Theorem (parking functions)

$$\sum_{\pi \in \operatorname{PF}(m,n)} x^{\operatorname{slev}(\pi)} y^{\operatorname{lel}(\pi)}$$

= $(n - m + 1)xy [(m - 1)((n - m + 1)x + y + m - 1)^{m-2} + (xy + (n - m)x + 1)(xy + (n - m)x + m)^{m-2}].$

- slev(π) (size of level set): total number of cars whose desired spot is in the range {1, 2, ..., n - m + 1}. (new statistic!)
- ones(π) (1's): total number of cars whose desired spot is spot
 1. (when m = n, slev(π) = ones(π).)
- lel(π) (leading elements): total number of cars whose desired spot is the same as that of the first car. (new statistic!)

An equivalent formulation of Main Theorem (parking functions)

Let $s, t \ge 1$. We have

$$\#\{\pi \in PF(m, n): \text{ slev}(\pi) = s \text{ and } \text{lel}(\pi) = t\}$$

= $\binom{m-2}{s-1, t-1, m-s-t} (n-m+1)^s (m-1)^{m-s-t+1}$
+ $\binom{m-1}{t-1, s-t, m-s} s(n-m+1)(n-m)^{s-t} m^{m-s-1}.$

・ロト・西ト・ヨト・ヨー うらぐ

Some immediate corollaries

Using standard probability tools, some asymptotic analysis of the above parking statistics readily follows, approximated by normal or Poisson distributions.

Take m = cn for some 0 < c < 1 as $n \to \infty$. Consider the parking preference $\pi \in PF(m, n)$ chosen uniformly at random. Then we have

$$lel(\pi) - 1 \xrightarrow{d} Poisson(c).$$

$$\frac{\operatorname{slev}(\pi) - c(1-c)n}{\sqrt{c^2(1-c)n}} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1).$$

・ロト・西ト・モン・モー ひゃぐ

Some immediate corollaries

Using standard probability tools, some asymptotic analysis of the above parking statistics readily follows, approximated by normal or Poisson distributions.

Take m = cn for some 0 < c < 1 as $n \to \infty$. Consider the parking preference $\pi \in PF(m, n)$ chosen uniformly at random. Then we have

а

$$egin{aligned} & ext{lel}(\pi)-1 \stackrel{ ext{d}}{ o} ext{Poisson}(c). \ & ext{slev}(\pi)-c(1-c)n \ & ext{d} \ & ext{$$

Outline of proof

We set $slev(\pi) = s$ and classify into two situations: $\pi_1 \in \{1, 2, \dots, n - m + 1\}$ and $\pi_1 \notin \{1, 2, \dots, n - m + 1\}$. By Pollak's argument,

$$\#\{\pi \in \mathrm{PF}(m,n): \pi_1 = 1\} = (n-m+2)(n+1)^{m-2},$$

which implies that

$$\#\{\pi \in \mathrm{PF}(m,n) : \pi_1 \in \{1,2,\ldots,n-m+1\}\} \\ = (n-m+1)(n-m+2)(n+1)^{m-2}$$

▲ロ ▶ ▲周 ▶ ▲ ヨ ▶ ▲ ヨ ▶ → 目 → の Q @

Generating function when $\pi_1 \in \{1, 2, \dots, n-m+1\}$:

$$\sum_{s=1}^{m} \binom{m-1}{s-1} y(n-m+1)(n-m+y)^{s-1} sm^{m-s-1} x^{s}.$$

- x^s : $slev(\pi) = s$.
- $\sum_{s=1}^{m}$: s can be any value from 1 (only the first car is in the level set) to m (all m cars are in the level set).
- $\binom{m-1}{s-1}$: besides the first car, we choose s-1 cars out of the remaining m-1 cars to constitute the level set.
- $y(n-m+1)(n-m+y)^{s-1}$: n-m+1 choices for the spot of any car in the level set. Each of the s-1 later cars independently has the same probability $\frac{1}{n-m+1}$ of being mapped to the same element as car 1, which combined contributes $(n-m+1)^s y\left(\frac{n-m}{n-m+1}+\frac{y}{n-m+1}\right)^{s-1}$ to $lel(\pi)$.
- sm^{m-s-1} : cars that are not in the level set constitute a parking function in PF(m-s, m-1).

Applying identities from the binomial distribution:

$$\sum_{s=1}^{m} {m-1 \choose s-1} y(n-m+1)(n-m+y)^{s-1} sm^{m-s-1} x^s$$

= $(n-m+1)xy \sum_{s=1}^{m} {m-1 \choose s-1} s(xy+(n-m)x)^{s-1} m^{m-s-1}$
= $(n-m+1)xy(xy+(n-m)x+m)^{m-2}(xy+(n-m)x+1).$

・ロト・日本・ヨト・ヨト・ヨー つへぐ

Generating function when $\pi_1 \notin \{1, 2, \dots, n-m+1\}$:

$$\sum_{s=0}^{m-1} \binom{m-1}{s} (n-m+1)^s ys(m-1+y)^{m-s-1} x^s$$

- x^s : $slev(\pi) = s$.
- $\sum_{s=0}^{m-1}$: s can be any value from 0 (no car is in the level set) to m-1 (all remaining m-1 cars are in the level set).
- $\binom{m-1}{s}$: we choose *s* cars out of the remaining m-1 cars to constitute the level set.
- (n − m + 1)^s: n − m + 1 choices for the spot of any car in the level set.
- ys(m-1+y)^{m-s-1}: cars that are not in the level set constitute a parking function in PF(m s, m 1). Each of the m-s-1 later cars independently has the same probability ¹/_m of being mapped to the same element as car 1, which combined contributes sm^{m-s-1}y (^{m-1}/_m + ^y/_m)^{m-s-1} to lel(π).

Applying identities from the binomial distribution:

$$\sum_{s=0}^{m-1} \binom{m-1}{s} (n-m+1)^s ys(m-1+y)^{m-s-1} x^s$$

= $(n-m+1)xy \sum_{s=0}^{m-1} \binom{m-1}{s} s((n-m+1)x)^{s-1} (m-1+y)^{m-s-1}$
= $(n-m+1)xy(m-1)((n-m+1)x+y+m-1)^{m-2}.$

・ロト・母ト・ヨト・ヨト・ヨー つへぐ

The mysterious pair $(ones(\pi), lel(\pi))$

For classical parking functions, the level set statistic $slev(\pi)$ reduces to the 1's statistic $ones(\pi)$, and is equidistributed with the leading elements statistic $lel(\pi)$. This feature of parking functions is quite mysterious as these two parking function statistics seem unrelated and are not of the same nature.

While the leading elements statistic is invariant under circular rotation, it does not satisfy permutation symmetry as permuting the entries might change the first element. On the other hand, though the 1's statistic is invariant under permuting all the entries, it does not exhibit circular rotation invariance. Indeed, only 1 out of n + 1 rotations of an assignment of n cars on a circle with n + 1 spots gives a valid parking function. It is thus intriguing what is hidden behind the pair of statistics $(ones(\pi), lel(\pi))$.

The less mysterious pair $(\deg_T(0), \deg_T(p))$

Under the bijective correspondence induced by breadth first search (BFS), the seemingly unrelated leading elements statistic and 1's statistic for classical parking functions both become degree statistics for certain vertices in the tree: one records the degree of a movable root (parent of a fixed vertex), while the other records the degree of the fixed root.

 $(ones(\pi), lel(\pi)) \leftrightarrow (\deg_{\mathcal{T}}(0), \deg_{\mathcal{T}}(p))$, where $\deg_{\mathcal{T}}(0)$ is the degree of the root vertex 0 and $\deg_{\mathcal{T}}(p)$ is the degree of the parent of vertex 1 in the tree \mathcal{T} .

The BFS construction between parking functions and rooted trees goes back to Foata and Riordan (1974). See Yan (2015) and also Chassaing and Marckert (2001). There are many more interesting statistics of rooted trees: descents in Eğecioğlu and Remmel (1986), inversions in De Oliveira and Vergnas (2010), and runs in Lackner and Panholzer (2020).

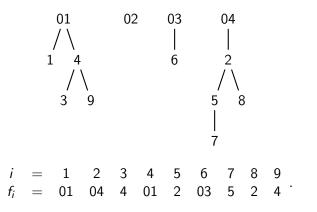
Main Theorem (BFS)

The breadth first search algorithm connecting parking functions and rooted forests has some interesting implications:

- The number of times π_i appears in a parking function π ∈ PF(m, n) equals the degree of the parent of vertex i in the corresponding forest F ∈ F(m, n).
- The number of times 1, 2, ..., n − m + 1 appears in a parking function π ∈ PF(m, n) respectively equals the degree of the root vertex 01, 02, ..., 0(n − m + 1) in the corresponding forest F ∈ 𝔅(m, n).

Outline of proof

A forest $F \in \mathcal{F}(m, n)$ may be represented by an acyclic function f, where for a non-root vertex i, $f_i = j$ indicates that vertex j is the parent of vertex i in a tree component of the forest.



▲ロ ▶ ▲ □ ▶ ▲ □ ▶ ▲ □ ▶ ▲ □ ▶ ● ○ ○ ○

We read the vertices of the forest in breadth first search (BFS) order. That is, read root vertices in order first, then all vertices at level 1 (children of a root), then those at level 2 (distance 2 from a root), and so on, where vertices at a given level are naturally ordered in order of increasing predecessor, and, if they have the same predecessor, increasing order. Applying BFS to the forest F, we have

 $v_{01}, \ldots, v_{04}, v_5, \ldots, v_{13} = 01, 02, 03, 04, 1, 4, 6, 2, 3, 9, 5, 8, 7.$

We let σ_f^{-1} be the vertex ordering once we remove the root vertices and σ_f be the inverse order permutation of σ_f^{-1} .

$$i = 1 2 3 4 5 6 7 8 9$$

 $\sigma_f^{-1}(i) = 1 4 6 2 3 9 5 8 7$
 $\sigma_f(i) = 1 4 5 2 7 3 9 8 6$

We further let $t(f) = (r_1, \ldots, r_{12})$ with r_i recording the degree of v_i , starting with v_{01} and ending with v_{12} (ignoring the final vertex v_{13}), that is,

$$t(f) = (2, 0, 1, 1, 0, 2, 0, 2, 0, 0, 1, 0).$$

ション ふぼう メリン メリン しょうくしゃ

The sequence t(f) is referred to as the forest specification of F.

For a parking function $\pi \in PF(m, n)$, the associated specification is $s(\pi) = (r_1, \ldots, r_n)$, where $r_k = \#\{i : \pi_i = k\}$ records the number of cars whose parking preference is spot k. The order **permutation** $\tau_{\pi} \in \mathfrak{S}_m$, on the other hand, is defined by $\tau_{\pi}(i) = \#\{j : \pi_j < \pi_i, \text{ or } \pi_j = \pi_i \text{ and } j \le i\}$, and so is the permutation that orders the list, without switching elements that are the same. In words, $\tau_{\pi}(i)$ is the position of the entry π_i in the non-decreasing rearrangement of π .

Example: for $\pi = (3, 1, 3, 1)$, $\tau_{\pi}(1) = 3$, $\tau_{\pi}(2) = 1$, $\tau_{\pi}(3) = 4$, and $\tau_{\pi}(4) = 2$.

We can easily recover a parking function π by replacing *i* in τ_{π} with the *i*th smallest term in the sequence $1^{r_1} \dots n^{r_n}$.

For a parking function $\pi \in PF(m, n)$, the associated specification is $s(\pi) = (r_1, \ldots, r_n)$, where $r_k = \#\{i : \pi_i = k\}$ records the number of cars whose parking preference is spot k. The order **permutation** $\tau_{\pi} \in \mathfrak{S}_m$, on the other hand, is defined by $\tau_{\pi}(i) = \#\{j : \pi_j < \pi_i, \text{ or } \pi_j = \pi_i \text{ and } j \le i\}$, and so is the permutation that orders the list, without switching elements that are the same. In words, $\tau_{\pi}(i)$ is the position of the entry π_i in the non-decreasing rearrangement of π .

Example: for $\pi = (3, 1, 3, 1)$, $\tau_{\pi}(1) = 3$, $\tau_{\pi}(2) = 1$, $\tau_{\pi}(3) = 4$, and $\tau_{\pi}(4) = 2$.

We can easily recover a parking function π by replacing *i* in τ_{π} with the *i*th smallest term in the sequence $1^{r_1} \dots n^{r_n}$.

The breadth first search algorithm bijectively connects parking functions and rooted forests, where $(t(f), \sigma_f) = (s(\pi), \tau_{\pi})$. Continuing with our earlier example, we have

$$s(\pi) = (2, 0, 1, 1, 0, 2, 0, 2, 0, 0, 1, 0),$$

and

ション ふゆ アメリア メリア しょうくしゃ

We form the non-decreasing rearrangement sequence associated with $s(\pi)$:

$$1^2, 3^1, 4^1, 6^2, 8^2, 11^1 = 1, 1, 3, 4, 6, 6, 8, 8, 11.$$

Replacing *i* in τ_{π} with the *i*th smallest term in this sequence yields the corresponding parking function $\pi \in PF(9, 12)$ given below:

$$i = 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ \pi_i = 1 \ 4 \ 6 \ 1 \ 8 \ 3 \ 11 \ 8 \ 6 \ .$$

The bijective formulas

From a forest F to a parking function π , we have

$$\pi_i = \begin{cases} j & f_i = 0j \text{ for some } j = 1, 2, \dots, n - m + 1, \\ (n - m + 1) + \sigma_f(f_i) & \text{otherwise.} \end{cases}$$

Conversely, from a parking function π to a forest ${\it F}$, we have

$$f_i = \begin{cases} 0j & \pi_i = j \text{ for some } j = 1, 2, \dots, n - m + 1, \\ \tau_\pi^{-1}(\pi_i - (n - m + 1)) & \text{otherwise.} \end{cases}$$

▲□▶ ▲□▶ ▲目▶ ▲目▶ - 目 - のへで

Main Theorem (rooted forests)

$$\sum_{F \in \mathcal{F}(m,n)} x^{\deg_F(0)} y^{\deg_F(p)}$$

= $(n - m + 1)xy [(m - 1)((n - m + 1)x + y + m - 1)^{m-2} + (xy + (n - m)x + 1)(xy + (n - m)x + m)^{m-2}].$

deg_F(0): total degree of all root vertices 01,...,0(n − m + 1).

• $\deg_F(p)$: degree of the parent of vertex 1. (new statistic!) By degree, we generally mean more precisely the number of children of a vertex in a rooted tree, which is 1 less than the degree in the graph-theoretical sense for non-root vertices.

An equivalent formulation of Main Theorem (rooted forests)

Let $s, t \ge 1$. We have

$$\#\{F \in \mathcal{F}(m,n): \deg_F(0) = s \text{ and } \deg_F(p) = t\} \\ = \binom{m-2}{s-1, t-1, m-s-t} (n-m+1)^s (m-1)^{m-s-t+1} \\ + \binom{m-1}{t-1, s-t, m-s} s(n-m+1)(n-m)^{s-t} m^{m-s-1}.$$

◆□ ▶ ◆□ ▶ ◆ 臣 ▶ ◆ 臣 ▶ ○ 臣 ○ のへで

Some immediate corollaries

Using standard probability tools, some asymptotic analysis of the above forest statistics readily follows, approximated by normal or Poisson distributions.

Take m = cn for some 0 < c < 1 as $n \to \infty$. Consider the labeled forest $F \in \mathcal{F}(m, n)$ chosen uniformly at random. Then we have

$$\deg_F(p) - 1 \stackrel{\mathrm{d}}{\rightarrow} \operatorname{Poisson}(c).$$

$$\frac{\deg_F(0) - c(1-c)n}{\sqrt{c^2(1-c)n}} \stackrel{\mathrm{d}}{\to} \mathcal{N}(0,1).$$

▲□▶ ▲□▶ ▲三▶ ▲三▶ 三三 - のへぐ

Some immediate corollaries

Using standard probability tools, some asymptotic analysis of the above forest statistics readily follows, approximated by normal or Poisson distributions.

Take m = cn for some 0 < c < 1 as $n \to \infty$. Consider the labeled forest $F \in \mathcal{F}(m, n)$ chosen uniformly at random. Then we have

$$\deg_F(p) - 1 \stackrel{\mathrm{d}}{\rightarrow} \operatorname{Poisson}(c).$$

$$\frac{\deg_F(0) - c(1-c)n}{\sqrt{c^2(1-c)n}} \stackrel{\mathrm{d}}{\to} \mathcal{N}(0,1).$$

Outline of proof

We use the well-known result that the number of rooted forests with vertex set [a] and b components whose root labels are given is ba^{a-b-1} (Cayley's formula). Thus the number of such rooted forests with one distinguished vertex (possibly one of the roots) is ba^{a-b} , and by symmetry the number of such rooted forests where a vertex in the first component is distinguished must be a^{a-b} .

Key ingredient is a vertex splitting argument. Vertex p (parent of vertex 1) and vertex 1 are first merged together to form a distinguished vertex. This vertex later splits, and all former children of p become children of 1.

There are

$$\binom{m-2}{s-1, t-1, m-s-t}(n-m+1)^s(m-1)^{m-s-t+1}$$

possible forests where parent p of vertex 1 is not one of the roots.

- Choose a label $r \in [m] \setminus \{1\}$ (m-1 possibilities).
- Choose two disjoint sets of labels $\{x_1, x_2, \ldots, x_{s-1}\}$ and $\{y_1, y_2, \ldots, y_{t-1}\}$ from $[m] \setminus \{1, r\} \left(\binom{m-2}{s-1, t-1, m-s-t}\right)$ possibilities).
- Choose a rooted forest on [m] \ {1} with root labels
 r, x₁, x₂,..., x_{s-1}, y₁, y₂,..., y_{t-1} and a distinguished vertex p
 in the first component ((m − 1)^{m-s-t} possibilities).
- Split the distinguished vertex *p* into two vertices, labeled *p* and 1 respectively, where *p* is the parent and 1 is the child, and all former children of *p* now become children of 1.
- Add roots 01,02,..., 0(n m + 1), and connect each of the vertices r, x₁, x₂,..., x_{s-1} with one of these roots by an edge ((n m + 1)^s possibilities).
- Add edges between vertex p and $y_1, y_2, \ldots, y_{t_{\overline{a}}1, \ldots, y_{t_{\overline{a}}1, \ldots, y_{\overline{a}}}}$

There are

$$\binom{m-1}{t-1,s-t,m-s}s(n-m+1)(n-m)^{s-t}m^{m-s-1}$$

possible forests where parent p of vertex 1 is one of the roots.

- Select a set of labels {x₁, x₂,..., x_{s-1}} from [m] \ {1} ((^{m-1}_{s-1}) possibilities).
- Construct a rooted forest with vertex set [m] and root labels $1, x_1, x_2, \ldots, x_{s-1}$ (sm^{m-s-1} possibilities).
- Among the labels x₁, x₂,..., x_{s-1}, choose the siblings of vertex 1 (^{s-1}_{t-1}) possibilities).
- Pick one of the n m + 1 roots 01, 02, ..., 0(n m + 1) as the parent p of vertex 1, and connect it and all the siblings chosen in the previous step to it by an edge (n m + 1 possibilities).
- Pick one of the other n m roots as parent for each of the remaining vertices with label in the set {x₁, x₂, ..., x_{s-1}} ((n m)^{s-t} possibilities).

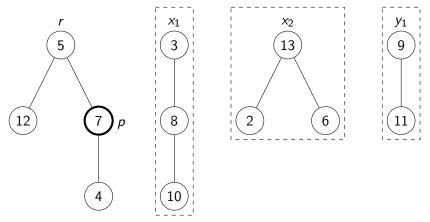
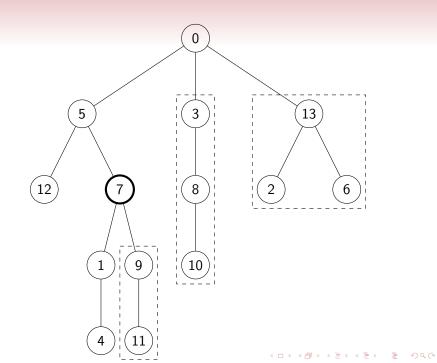


Illustration of the procedure: the rooted forest with the distinguished vertex p indicated by a thick node, and the final tree (next page).

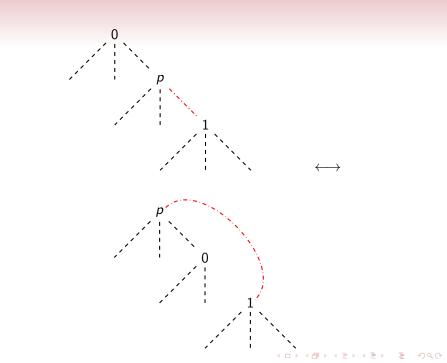


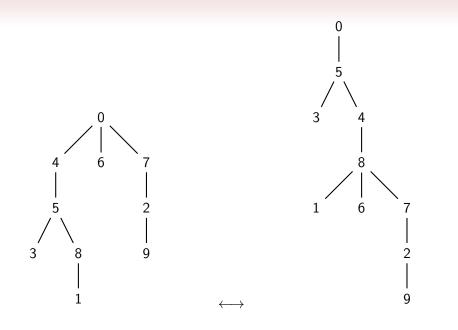
An explicit bijection $\deg_{\mathcal{T}}(0) \leftrightarrow \deg_{\mathcal{T}}(p) \text{ and } \operatorname{ones}(\pi) \leftrightarrow \operatorname{lel}(\pi)$

To construct an explicit bijection in the tree T, we proceed in steps.

- Remove the edge connecting vertices 1 and *p*.
- Connect vertices 0 and 1 by an edge.
- Interchange vertices 0 and p.

This map has the extra benefit of being an involution. Moreover, the degrees of all vertices except 0 and p are preserved.





◆□▶ ◆□▶ ◆ 臣▶ ◆ 臣▶ ○臣 ○ のへぐ

Some nice features are hence introduced in the corresponding parking function bijection under BFS, where

$$\pi = (8, 4, 5, 1, 2, 1, 1, 5, 6) \leftrightarrow \pi' = (5, 8, 2, 2, 1, 5, 5, 4, 9).$$

We see that $ones(\pi)$ and $lel(\pi)$ are switched, but the frequencies of the non-1 and non-leading elements are preserved up to permutation.

ション ふゆ アメリア メリア しょうくしゃ

There are other cool results that I'd love to share with you, but let's stop here for now!

Thank You! Questions?

▲□▶ ▲□▶ ▲目▶ ▲目▶ 三日 - のへの