Coalescing particles on an interval and Mallows(q) permutations

Nathaniel B. Alloway*, Hugo Panzo, Ph.D., Evan Socher

Saint Louis University

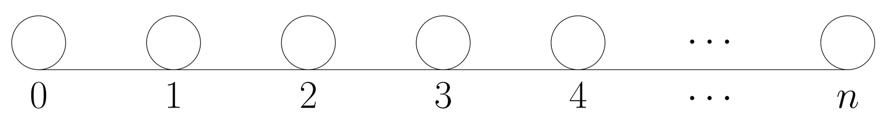
Background

Dynamical systems like particle models and cellular automota are used to model physical phenomena and study computer algorithms. They are also rich sources of mathematical inquiry.

This poster presents a particle model based on the paper [3]. The original model studied in that paper has its origins in understanding self-organizing binary search trees.

The particle model

Fix $n \in \mathbb{Z}_+$. Take a collection of $n+1$ identical particles and place them on a number line, at positions $0, 1, \ldots, n$.



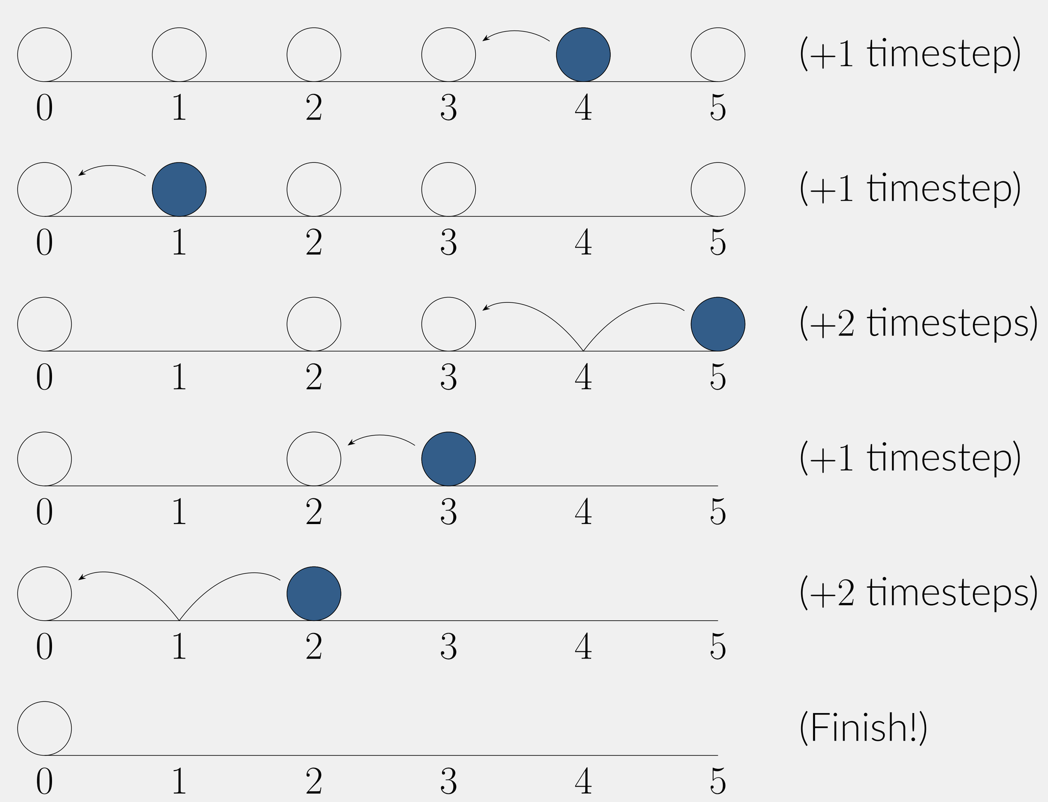
We will evolve this system over the course of n iterations, moving particles one at a time until they all end up at position 0.

(1) Choose a particle at a positive position $1, 2, \ldots, n$ uniformly at random.
(2) Move the particle left at a rate of 1 unit per timestep.
(3) When two particles coincide, the two particles coalesce into a single particle, and movement stops.
(4) Repeat steps (1)–(3) until all particles have coalesced with the particle at position 0.

Note that the amount of timesteps that occur in one iteration depends on the order in which particles are chosen.

Example evolution

Set $n = 5$. Here is one possible evolution:



Define the coalescence time of the system to be the total number of timesteps that it takes for the evolution to finish. The coalescence time is a random variable, which we denote by T_n.

Notice that the system evolution is completely determined by the order in which particles are chosen. Suppose the particle starting at position $j \in [n]$ is chosen during iteration $\sigma_j \in [n]$. Then the permutation

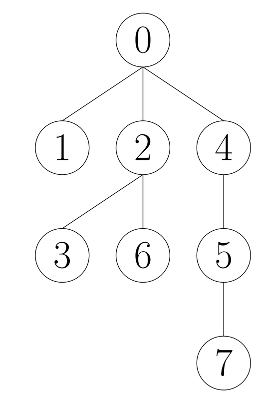
$$\sigma = \sigma_1\sigma_2 \ldots \sigma_n \in \mathcal{S}_n$$

gives the selection order permutation for the system.

Example. In the previous example, we have $\sigma = 25413 \in \mathcal{S}_5$ with $T_n = 1+1+2+1+2 = 7$.

A bijection with recursive rooted trees

The book [5] describes a standard bijection between $\mathcal{S}_n$ and recursive rooted trees on $n+1$ nodes. After a slight modification, this bijection sends the coalescence time of a selection order permutation to the total path length of the resulting tree, which is the sum of the root-to-node distances in the tree.



Example. In the tree at right, the root-to-node distance for the node 3 is 2 while the root-to-node distance for the node 7 is 3. The total path length is 12.

Example of the modified bijection

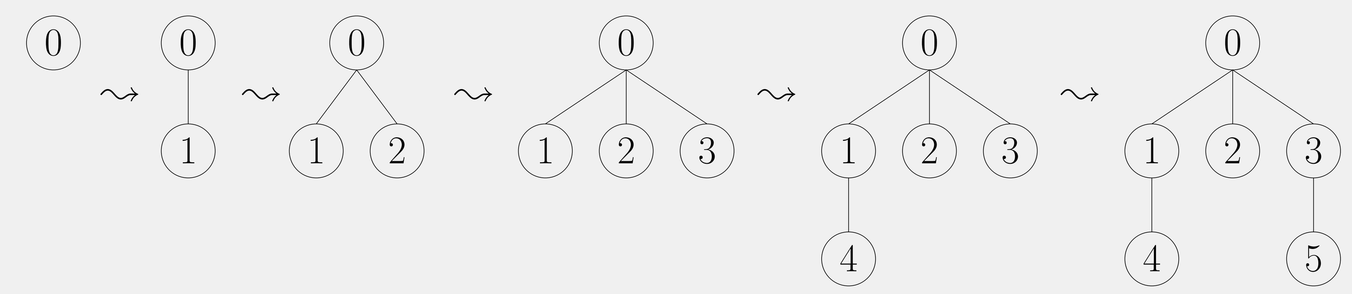
Start by reversing σ and inverting the values of each of its entries:

$$\sigma = 25413 \mapsto 31452 \mapsto 35214$$

Prepend a 0 to the front of the resulting string. Re-interpret the string as a sequence of substrings, where the j-th substring in the sequence contains only the numbers $0, 1, \ldots, j$:

$$0 \rightsquigarrow 01 \rightsquigarrow 021 \rightsquigarrow 0321 \rightsquigarrow 03214 \rightsquigarrow 035214$$

Recursively build a tree as follows: Start with a root node labeled 0. On step j, look at the j-th substring, find the number to the immediate left of j in the substring, and append a node labeled j to the node with that number in the tree.



The tree constructed at the very end is the result of the bijection.

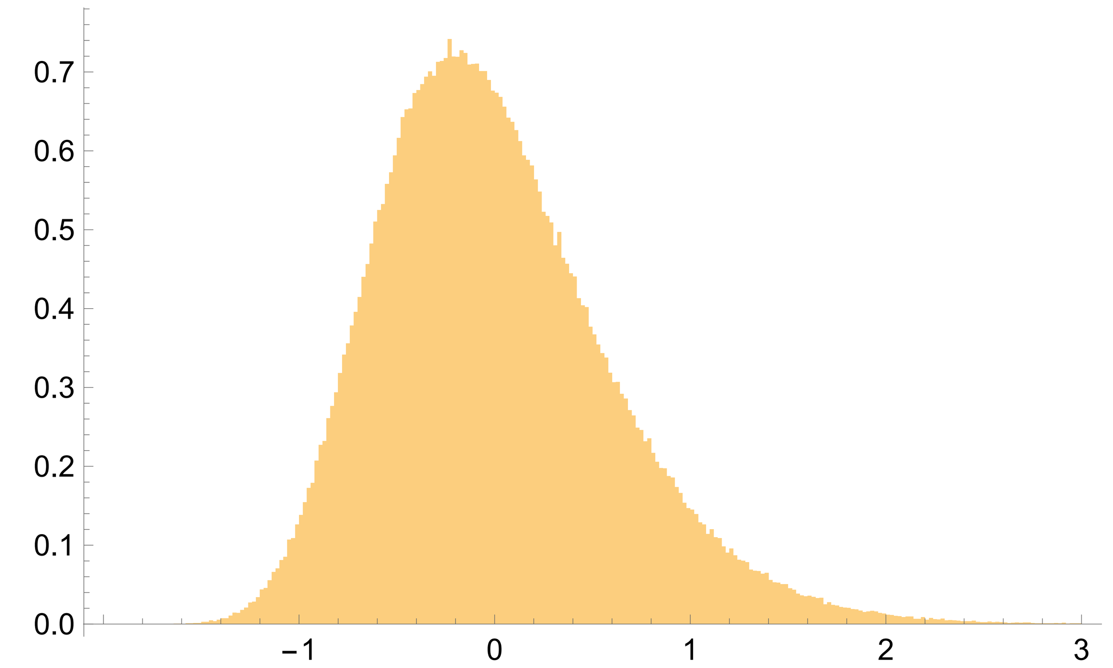
The scaling limit of T_n

Give the set of recursive rooted trees on $n+1$ nodes the uniform distribution. Under the bijection, T_n is equal in distribution to the total path length of a recursive rooted tree.

Theorem (Dobrow and Fill, [2]). *Let $U \sim \mathrm{Unif}([n])$, and let T_j, T_j' be copies of the total path length random variable for $j \in \{0, 1, \ldots, n-1\}$, all variables mutually independent. Then*

$$T_n \overset{d}{=} T_{U-1} + T_{n-U} + U.$$

The scaling limit of T_n (i.e. the distribution you get by normalizing and sending $n \to +\infty$) was first studied in [4]. In [2], the recursion is used to show that the scaling limit is not normal!



Above is a histogram of 5×10^5 simulated values of T_n after normalizing, for $n = 10^5$.

q-analogues

A q-analogue of an object is a parametrized version of the object, depending on a value q, such that sending $q \to 1$ in the limit recovers the original object. In several branches of mathematics, q-analogues appear as natural generalizations of objects. We highlight some examples.

Let $q > 0$ and $n \in \mathbb{Z}_+$.

- q-integers. Define $n_q := 1 + q + q^2 + \cdots + q^{n-1}$.
- q-factorials. Define $n!_q := 1_q \cdot 2_q \cdots n_q$.
- q-harmonic numbers. Define $H_n(q) := \sum_{j=1}^{n}(1/j_q)$.
- Discrete q-uniform distribution. We say that $U_n(q)$ has the discrete q-uniform distribution if it has support $[n]$ and satisfies
$$\mathrm{Prob}(U_n(q) = j) = q^{n-j}/n_q, \quad j \in [n].$$

The inversion count $\mathrm{inv}\,\sigma$ of $\sigma = \sigma_1\sigma_2\ldots\sigma_n \in \mathcal{S}_n$ is the number of pairs $i < j$ with $\sigma_i > \sigma_j$.

Example. For $\sigma = 25413$, we have $\mathrm{inv}\,\sigma = 6$.

We say that a *permutation-valued* random variable Π with support $\mathcal{S}_n$ has the Mallows(q) distribution if it satisfies

$$\mathrm{Prob}(\Pi = \sigma) = q^{\mathrm{inv}\,\sigma}/n!_q, \quad \sigma \in \mathcal{S}_n.$$

This is a q-analogue of the uniform distribution on $\mathcal{S}_n$.

Results

Theorem. *If the selection order permutation is Mallows(q), then $\mathbb{E}[T_n] = \sum_{j=1}^{n} H_j(q)$.*

Theorem. *Suppose the selection order permutation is Mallows(q). Let $U_n(q)$ be discrete q-uniform, and let T_j, T_j' be copies of the coalescence time variable for $j \in \{0, 1, \ldots, n-1\}$, all variables mutually independent. Then*

$$T_n \overset{d}{=} T_{U_n(q)-1} + T_{n-U_n(q)} + U_n(q).$$

Discussion

- The q-analogue $\sum_{j=1}^{n} H_j(q)$ is not the usual q-analogue of the "hyper-harmonic number" that appears in the literature.
- "Mallows(q)-distributed binary trees" have been studied before in [1]. Under the bijection with recursive rooted trees, we have "Mallows(q)-distributed recursive rooted trees" instead. As far we can tell, these have not yet been studied!
- Ongoing work is to determine the asymptotic growth rate for the variance of T_n when $q < 1$ and $q > 1$ and to determine the limit of the distribution for T_n when sending $n \to +\infty$ before or after sending $q \to 1$.

Acknowledgements

The presenting author would like to thank Dr. Özlem Uğurlu for introducing to them the concept of q-analogues.

References

[1] Louigi Addario-Berry. Benoît Corsini. The height of Mallows trees. *The Annals of Probability*, 49(5):2220–2271, 2021.

[2] Robert P. Dobrow and James Allen Fill. Total path length for recursive trees. *Combinatorics, Probability, and Computing*, 8(4):317–333, 1999.

[3] Michael Larsen and Russell Lyons. Coalescing particles on an interval. *Journal of Theoretical Probability*, 12(1):201–205, 1999.

[4] Hosam M. Mahmoud. Limiting distributions for path lengths in recursive trees. *Probability in the Engineering and Informational Sciences*, 5 (1):53–59, 1991.

[5] Richard P. Stanley. *Enumerative Combinatorics*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2nd edition, 2012.