Quantifying CDS Sortability of Permutations Using Strategic Piles

Marisa Gaetz — *Massachusetts Institute of Technology*
Bethany Molokach — *Western Carolina University*
Marion Scheepers — *Boise State University*
Meghan Shanks — *Texas A&M University*

Motivation: Ciliates

Sorting is essential to many disciplines. One biological example is the one-celled organism known as a ciliate, which must sort DNA in its micronuclei in order to create a macronucleus. Biologists hypothesize that ciliates use iterations of the context directed swap (CDS) sorting operation to rearrange their DNA [3]. We investigate the mathematical implications of this hypothesis.

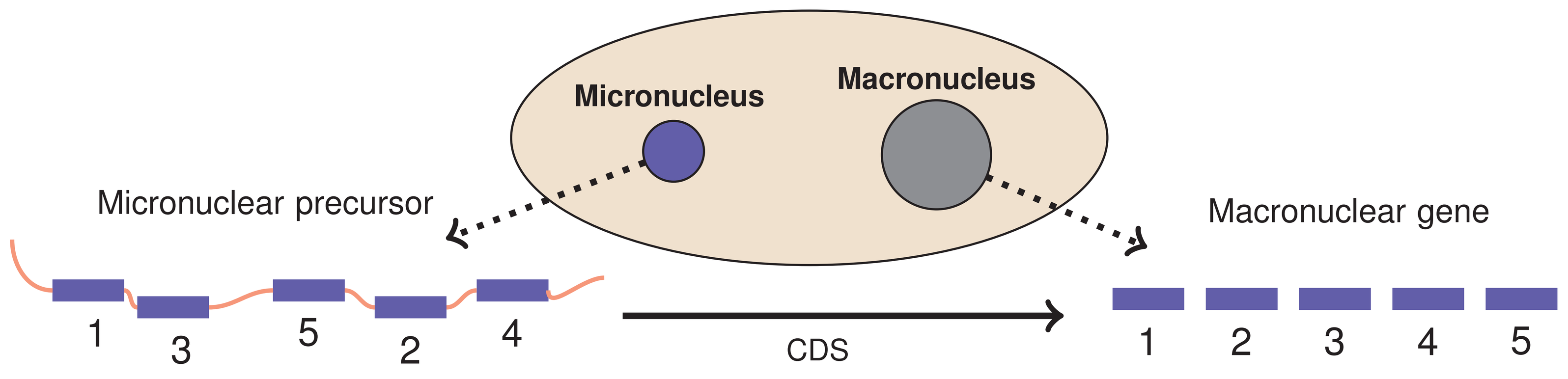

Figure: CDS as used by ciliates

CDS and Strategic Piles

Consider $\pi \in S_n$ where $\pi = [a_1, a_2, \ldots, a_n]$. To each a_i, assign a left pointer, $\langle a_i - 1, a_i\rangle$, and a right pointer, $\langle a_i, a_i + 1\rangle$. The CDS operation uses two pointers p and q in the order $p \ldots q \ldots p \ldots q$ and exchanges the elements flanked by these pointers. [1]

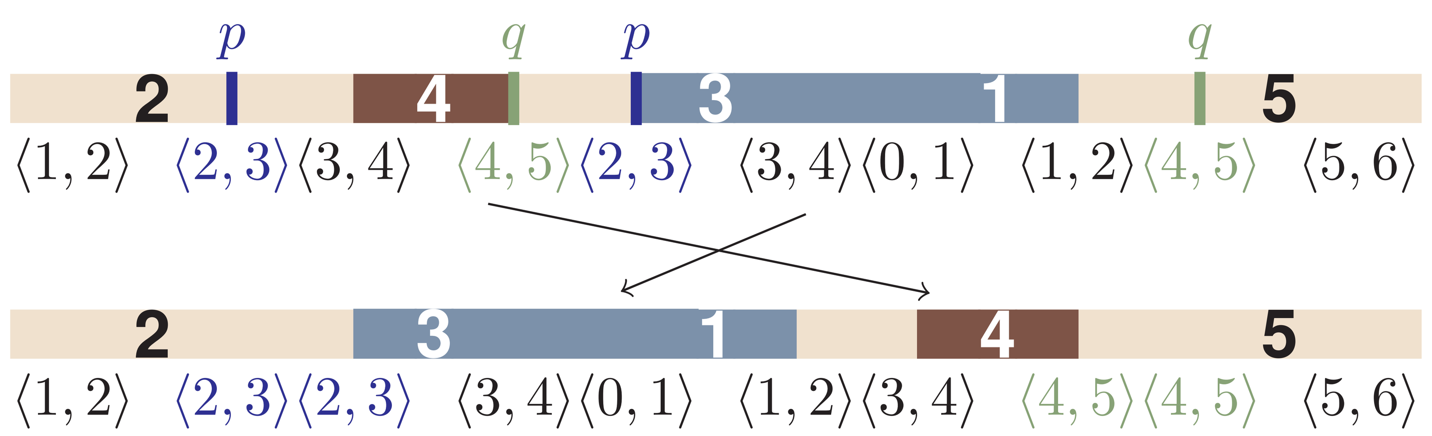

Figure: A valid CDS move with $p = \langle 2, 3\rangle$ and $q = \langle 4, 5\rangle$

Facts and concepts from [1] relating to CDS include:

- CDS cannot sort every permutation.
- **Fixed points** are permutations with no available CDS moves.
- The **cycle permutation of π**, C_π, is defined as $Y_\pi \circ X_n$, where $X_n = (0, 1, \ldots, n)$ and $Y_\pi = (0, a_n, \ldots, a_1)$.
- If $(0, \ldots, n, b_1, \ldots, b_k)$ is in the cycle decomposition of C_π, then the **strategic pile** of π, SP(π), is $\{b_1, b_2, \ldots, b_k\}$. Otherwise, SP($\pi$) $= \emptyset$. There is a one-to-one correspondence between SP(π) and the fixed points of π.

From the above, we see that $|SP(\pi)|$ relates to the degree of unsortability of π. Our goal is to count the number of permutations that, using CDS, can result in k different fixed points. Thus, we quantify permutations with a specific degree of unsortability.

Full Strategic Piles

A permutation $\pi \in S_n$ has a **full strategic pile** if $|\text{SP}(\pi)| = n - 1$. If π has a full strategic pile, then n is even.

> **Theorem**
>
> For an even n, the number of permutations in S_n with a full strategic pile is $\frac{2(n-1)!}{n}$.

Since $X_n = Y_\pi^{-1} \circ C_\pi$, we count the factorizations of X_n into two $(n+1)$-cycles where the second factor has the form $(0, n, \ldots)$:

- X_{n-2} has $\frac{2(n-2)!}{n}$ factorizations into two $(n-1)$-cycles [2].
- There exists a bijection from these factorizations to factorizations of X_n of the form $\gamma \circ (0, n, 1, \ldots)$, where both factors are $(n+1)$-cycles.
- Another $n - 1$ bijections together can map the above set to all factorizations of X_n of the form $\gamma \circ (0, n, \ldots)$.

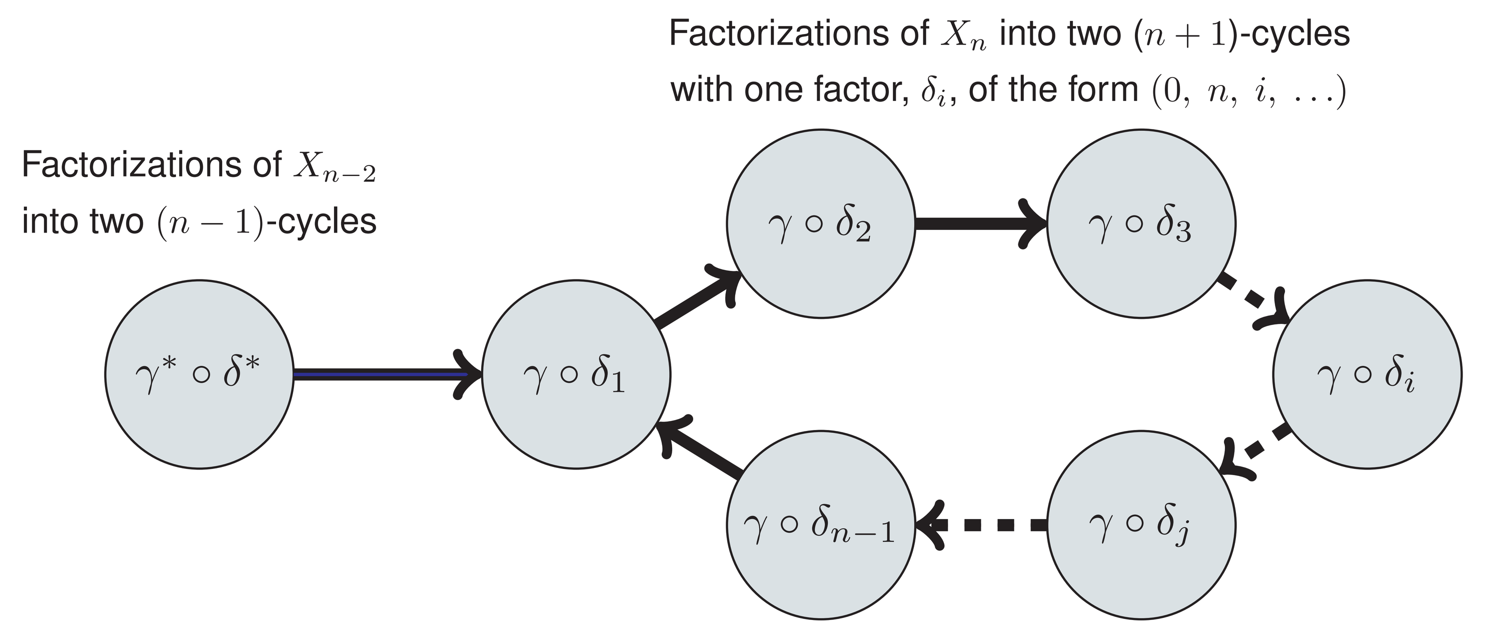

Figure: One bijection that maps $\gamma^* \circ \delta^*$ to $\gamma \circ \delta_1$ and $n - 1$ bijections that map $\gamma \circ \delta_1$ to $\gamma \circ \delta_i$ for a given i.

Therefore, the total number of permutations in S_n with full strategic pile is $\frac{2(n-1)!}{n}$.

Merge Numbers

Permutations with strategic pile $\{b_1, \ldots, b_k\}$ take the form:

$$[b_k + 1, \ldots, b_{x_1}, \underbrace{b_{x_1 - 1} + 1, \ldots, b_{x_2}}_{\text{merge if equal}}, b_{x_2 - 1} + 1, \cdots, \underbrace{b_{x_{k-1}}, b_{x_{k-1} - 1} + 1}_{\text{pair}}, \ldots, b_1].$$

Together, the entries $b_{x_i}, b_{x_i - 1} + 1$ are called a **pair**. A **merge** occurs when $b_{x_i - 1} + 1 = b_{x_{i+1}}$.

It is essential to count the ways the $k - 1$ pairs can be ordered with l merges. We call these quantities **merge numbers**, denoted $c_{k,l}$, and have partially developed an algorithm to find them.

Strategic Piles of Size k

> **Theorem**
>
> The number of permutations $\pi \in S_n$ with $|SP(\pi)| = k$ is
>
> $$(n-k)! \sum_{i=0}^{k-t} c_{k,i} \binom{n - (k+1)}{k - (i+1)}$$
>
> where $t = 2$ for k even and $t = 1$ for k odd.

- There are $(n-k)!$ ways to assign $\{1 \ldots n\}$ to $\{b_1 \ldots b_k\}$ and order the remaining elements.
- The i^{th} term represents arranging pairs using i merges.
  - Each $c_{k,i}$ counts the ways to order pairs using i merges.
  - The binomial coefficients count the ways to place these ordered pairs.

Future Work

- Complete our merge number algorithm, which would require:
  - Determining the integer partitions of n with all parts of size greater than one and an even number of even parts.
  - Establishing the base case for a recursive formula.
- Construct an alternative merge number algorithm that will run in less than exponential time.

References

[1] K.L.M. Adamyk et al., *Sorting Permutations: Games, Genomes and Cycles*, **arXiv:1410.2353**

[2] L. Cangelmi, *Factorization of an n-cycle into two n-cycles*, **European Journal of Combinatorics** 24 (2003) 849–853.

[3] D.M. Prescott, A. Ehrenfeucht and G. Rozenberg, *Template-guided recombination for IES elimination and unscrambling of genes in stichotrichous ciliates*, **Journal of Theoretical Biology** 222 (2003), 323 - 330.

Acknowledgments

Thanks to the National Science Foundation, Grant No. DMS-1359425, and to Boise State University.