On the Maximum Length of Multiset Cycle Packings

Humayun Hassan

Euler Circle
altmanhumayun@gmail.com

The Trade-off: Speed vs. Order

Autonomous robots need to know where they are on a track marked with colors.

- **Cameras** read the colors in order, but cost time, power, and money.
- **Photodiodes** are cheap and fast, but only report *how many* of each color are in view — not the order.

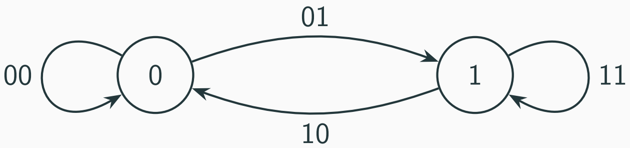
$$\text{R} \quad \text{B} \quad \text{R} \quad \underbrace{\text{B} \quad \text{R} \quad \text{R}}_{\{\{2\ \text{Red},\ 1\ \text{Blue}\}\}}$$

Two different arrangements of colors can produce the exact same reading. That's the problem this talk is about.

Background: De Bruijn Sequences

When order *is* available, this is a solved problem.

- Fix an alphabet of k colors and a window of size m.
- A **de Bruijn sequence** is a cyclic word of length k^m in which every ordered m-tuple appears exactly once.
- Equivalently: an Eulerian circuit on the de Bruijn graph, whose edges are m-tuples and whose vertices are $(m-1)$-tuples.



$k = 2$, $m = 2$: the circuit $0 \to 0 \to 1 \to 1 \to 0$ reads out 0011.

Forgetting Order

A photodiode sensor doesn't give us the tuple $(x_1, \ldots, x_m)$ — it gives us the observation function

$$f(x_1, \ldots, x_m) = \{\{x_1, \ldots, x_m\}\}.$$

- **Ordered domain:** k^m tuples.
- **Unordered codomain:** $\binom{k+m-1}{m}$ multisets.

f is highly non-injective: many distinct tuples land on the same multiset. Any de Bruijn-style construction built on ordered tuples is now the wrong tool.

When Can We Track Perfectly?

Definition

An **M-cycle** is a cyclic sequence of colors in which every size-m multiset over the alphabet appears exactly once as a window.

Theorem (Necessary condition, Hurlbert–Johnson–Zahl)

If an M-cycle exists, then

$$k \;\Big|\; \binom{k+m-1}{m}.$$

$k = 3,\ \ m = 3$

$\binom{5}{3} = 10$, and $3 \nmid 10$. No M-cycle exists for this pair.

Contribution 1: First Attempt

If we can't hit all 10 multisets, the obvious fix is to drop one and aim for length 9, since $3 \mid 9$.

This doesn't work. Removing an arbitrary multiset unbalances how many times each color appears across all the windows in the cycle. The remaining states can no longer be strung into a closed, consistent sequence.

So the real question isn't just *how many* states to drop — it's *which ones*, and how many is actually forced.

An Upper Bound on $M_3(k)$

Track, for each color, its total multiplicity across every window in the cycle. For the cycle to close consistently, this total must be divisible by m.

Theorem (Upper bound, $m = 3$, $3 \mid k$)

$$M_3(k) \;\le\; \binom{k+2}{3} - \frac{k}{3}.$$

At least $k/3$ multisets must be left out of any valid packing.

What Gets Dropped

The $k/3$ omitted states aren't arbitrary — they're forced into a specific shape.

Theorem (Rainbow partition)

When the bound above is achieved, the omitted multisets partition the alphabet $\mathbb{Z}_k$ exactly: each is a rainbow triple of three distinct colors, and together they use every color exactly once.

This turns "find a valid packing" into a much more structured search problem.

Finding a Construction

Searching for a valid cycle directly, via integer linear programming over individual edges, doesn't scale: the number of edge variables grows like $O(k^{2m})$.

Fix: exploit symmetry.

- Let $\sigma(x) = x + 1 \pmod{k}$ act on colors.
- States group into orbits under $\langle \sigma \rangle$.
- Solve the ILP once per orbit instead of once per edge — a factor-of-k reduction that makes the problem tractable.

The Bound Is Tight

The orbit-reduced ILP produces explicit constructions matching the bound exactly:

k	Upper bound	Constructed length
6	54	54
9	162	162
12	360	360

For each of these, the theoretical ceiling is actually achievable.

What's Still Open

Conjecture (1D Multiset Packing)

For each fixed m, there is a threshold k_0 such that for all $k \geq k_0$,

$$M_m(k) \;=\; k \left\lfloor \frac{1}{k}\binom{k+m-1}{m} \right\rfloor .$$

- Verified computationally for $k = 6, 9, 12$.
- The case $k \equiv 0 \pmod{9}$ is a genuinely open subcase, even for $m = 3$ — the orbit argument doesn't yet settle it.

Contribution 2: Just Sum the Colors?

A simpler idea, worth ruling out before building anything more complicated: give each color a numeric weight and have the sensor report the sum.

$$f(M) = \sum_{c \in M} c.$$

Could a single integer serve as a reliable position readout?

Why Scalar Sums Fail

Theorem (Inevitable collisions)

For $k \geq 3$, $m \geq 2$, no assignment of bounded, sequentially-ordered integer weights makes the scalar-sum map injective.

- **Domain** (multisets): grows like $O(k^m)$.
- **Codomain** (sums): confined to an interval of width $O(mk)$.

Polynomial growth outruns linear growth. Past a certain k, distinct multisets are guaranteed to share a sum.

Takeaways

- Dropping order breaks the classical de Bruijn construction — the observation map is far from injective.
- Optimal multiset packings are governed by a color-balance constraint, and the states that must be dropped form a rainbow partition of the alphabet.
- Scalar-sum encodings can't rescue injectivity: polynomial domain growth always outpaces linear codomain growth.

Thank you — questions?