Multiset Combinatorial Gray Codes and Universal Cycles

Yuan-Hsun Lo

Department of Applied Mathematics
National Pingtung University

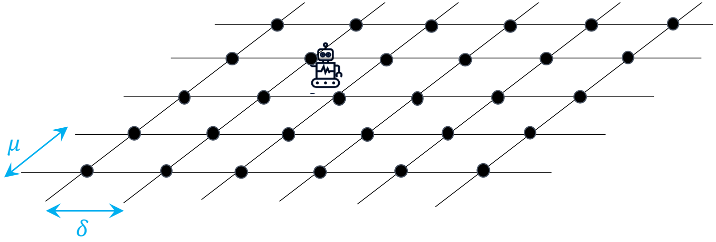
2025 TMS Annual Meeting

Joint with C. S. Chen (Nokia Bell Labs), W. S. Wong (CUHK), and T.-L. Wong (NSYSU)

Outline

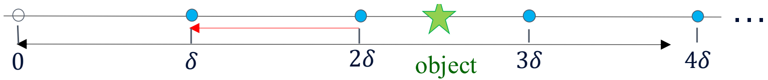
Object Tracking Problem

- Localization of an object in a *proximity sensor network*.
- Each sensor node is assigned a unique identification number (ID).
- The monitored region is a (flat or cyclic) 2D map.
- The localization accuracy $\delta \times \mu$ is pre-assigned.
- How to minimize the number of required IDs when considering the detection range of each sensor?

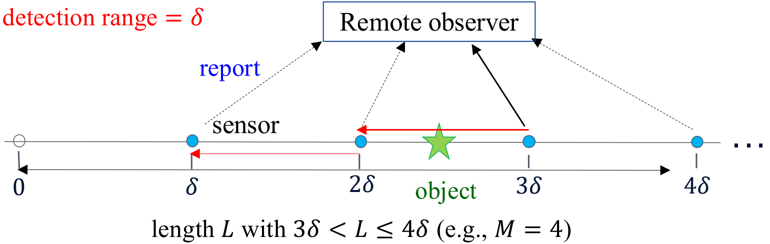


Object Tracking Problem (1D)

- An object randomly appears on a line of length L.
- Time is divided into descrete time slots of duration h.
- At the beginning of each time slot, we want to determine the position of the object, with upside precision δ.
  - the system says it is located at 2δ, it means: in $[2\delta, 3\delta)$
  - δ: tracking accuracy



Object Tracking Problem (1D)

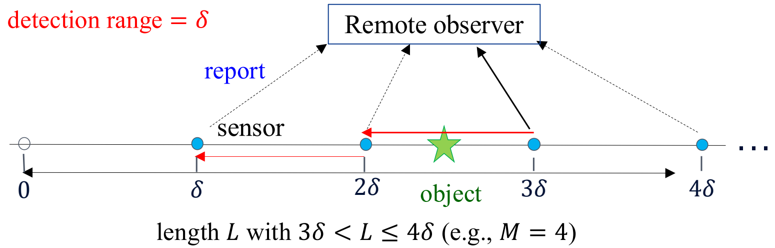


A straightforward solution:

- put one sensor at each position located $\delta, 2\delta, \ldots, \lceil L/\delta \rceil \delta$
- each sensor detects objects to its left with detection range δ
- each sensor reports the location of the detected object to a remote observer

$M := \lceil L/\delta \rceil$ unique IDs are necessary for the M sensors, each of which requires $\log_2 M$ bits to represent.

Object Tracking Problem (1D)

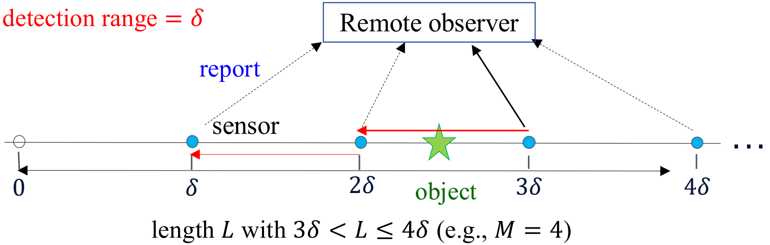


A straightforward solution:

- put one sensor at each position located $\delta, 2\delta, \ldots, \lceil L/\delta \rceil \delta$
- each sensor detects objects to its left with detection range δ
- each sensor reports the location of the detected object to a remote observer

$M := \lceil L/\delta \rceil$ unique IDs are necessary for the M sensors, each of which requires $\log_2 M$ bits to represent.

Object Tracking Problem (1D)



Straightforward Protocol

$M := \lceil L/\delta \rceil$ unique IDs are necessary for the M sensors, each of which requires $\log_2 M$ bits to represent.

- Time is divided into descrete time slots of duration h.
- Assume the communication channel is with date rate R (bits/time-unit).
- The straightforward protocol works whenever

$$\log_2 \lceil L/\delta \rceil = \log_2 M \leq Rh \qquad (*)$$

Object Tracking Problem (1D)

Straightforward Protocol

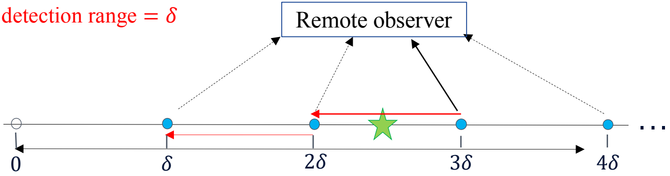
$M := \lceil L/\delta \rceil$ unique IDs are necessary for the M sensors, each of which requires $\log_2 M$ bits to represent.

- Time is divided into descrete time slots of duration h.
- Assume the communication channel is with date rate R (bits/time-unit).
- The straightforward protocol works whenever

$$\log_2 \lceil L/\delta \rceil = \log_2 M \leq Rh \tag{$*$}$$

- $(*)$ does not hold if $L > \delta 2^{Rh}$
- If the detection range $> \delta$, the straightforward protocol seems a bit lavish.

Object Tracking Problem (1D)



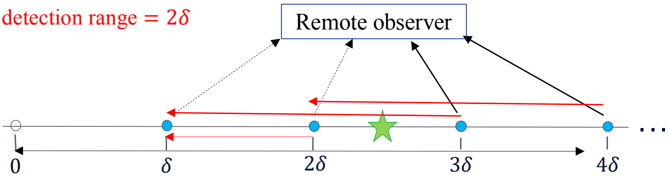
- Some IDs may be reused to reduce the number of IDs.

Multiset Color Coding Problem

Assume the detection range is $m\delta$, for $m \in \mathbb{N}$, where δ is the tracking accuracy. Find a set of $k \leq M$ IDs and select one ID for each sensor so that, the combinations of m consecutive IDs for all intervals $[(j-1)\delta, j\delta]$ are all distinct (distinguishable).

Objective: For given M and m, minimize the value k.

Object Tracking Problem (1D)



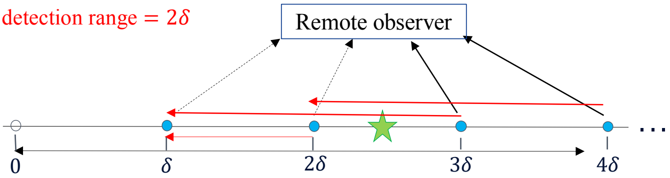
- Some IDs may be reused to reduce the number of IDs.

Multiset Color Coding Problem

Assume the detection range is $m\delta$, for $m \in \mathbb{N}$, where δ is the tracking accuracy. Find a set of $k \leq M$ IDs and select one ID for each sensor so that, the combinations of m consecutive IDs for all intervals $[(j-1)\delta, j\delta]$ are all distinct (distinguishable).

Objective: For given M and m, minimize the value k.

Object Tracking Problem (1D)



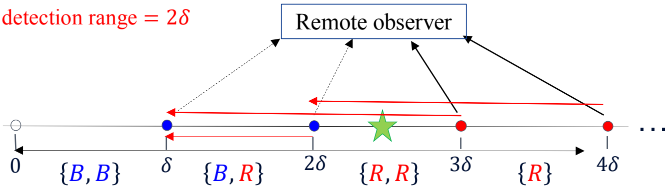
- Some IDs may be reused to reduce the number of IDs.

Multiset Color Coding Problem

Assume the detection range is $m\delta$, for $m \in \mathbb{N}$, where δ is the tracking accuracy. Find a set of $k \leq M$ IDs and select one ID for each sensor so that, the combinations of m consecutive IDs for all intervals $[(j-1)\delta, j\delta]$ are all distinct (distinguishable).

Objective: For given M and m, minimize the value k.

Object Tracking Problem (1D)



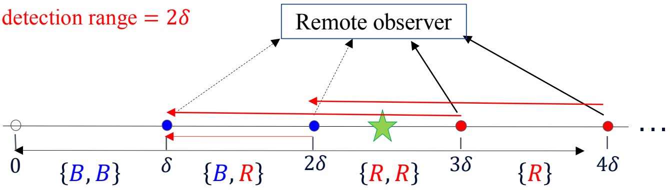
- Some IDs may be reused to reduce the number of IDs.

Multiset Color Coding Problem

Assume the detection range is $m\delta$, for $m \in \mathbb{N}$, where δ is the tracking accuracy. Find a set of $k \leq M$ IDs and select one ID for each sensor so that, the combinations of m consecutive IDs for all intervals $[(j-1)\delta, j\delta]$ are all distinct (distinguishable).

Objective: For given M and m, minimize the value k.

Object Tracking Problem (1D)



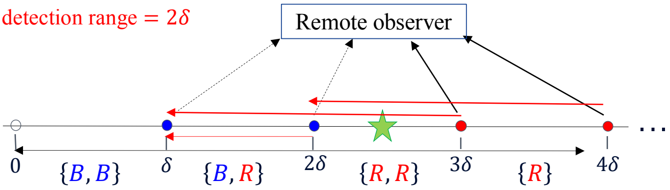
- Some IDs may be reused to reduce the number of IDs.

Multiset Color Coding Problem

Assume the detection range is $m\delta$, for $m \in \mathbb{N}$, where δ is the tracking accuracy. Find a set of $k \leq M$ IDs and select one ID for each sensor so that, the combinations of m consecutive IDs for all intervals $[(j-1)\delta, j\delta]$ are all distinct (distinguishable).

Objective: For given M and m, minimize the value k.

Object Tracking Problem (1D)



- Some IDs may be reused to reduce the number of IDs.

Multiset Color Coding Problem

Assume the detection range is $m\delta$, for $m \in \mathbb{N}$, where δ is the tracking accuracy. Find a set of $k \leq M$ IDs and select one ID for each sensor so that, the combinations of m consecutive IDs for all intervals $[(j-1)\delta, j\delta]$ are all distinct (distinguishable).

Objective: For given M and m, minimize the value k.

Multiset Combinatorial Gray Codes: Distinguishable Sequences

Notation and Definitions

- Let $[k] := \{1, 2, \ldots, k\}$, a set of k colors (IDs).
- Consider a sequence $S = s_0 s_1 \cdots s_{M-1}$ in which $s_i \in [k]$.
- Let $S_t(m)$ denote the **multiset** $\{s_t, s_{t+1}, \ldots, s_{t+m-1}\}$.

Definition (Multiset Combinatorial Gray Codes)

- m-distinguishable: multisets $S_t(m)$, $t \in \{0, 1, \ldots, M-m\}$, are all distinct
- cyclic m-distinguishable: multisets $S_t(m)$ are all distinct for $t \in \mathbb{Z}_M$

⋆ Code symbols of successive m-blocks can differ by at most two elements, multiplicity counting.

Example.

- 21133212 is 3-distinguishable, but not cyclic 3-distinguishable
- 111222333 is both 3-distinguishable and cyclic 3-distinguishable

C. Savage, "A survey of combinatorial Gray codes," *SIAM Review*, vol. 39, no. 4, pp. 605–629, 1997.

Notation and Definitions

- Let $[k] := \{1, 2, \ldots, k\}$, a set of k colors (IDs).
- Consider a sequence $S = s_0 s_1 \cdots s_{M-1}$ in which $s_i \in [k]$.
- Let $S_t(m)$ denote the **multiset** $\{s_t, s_{t+1}, \ldots, s_{t+m-1}\}$.

Definition (Multiset Combinatorial Gray Codes)

- m-distinguishable: multisets $S_t(m)$, $t \in \{0, 1, \ldots, M-m\}$, are all distinct
- cyclic m-distinguishable: multisets $S_t(m)$ are all distinct for $t \in \mathbb{Z}_M$

⋆ Code symbols of successive m-blocks can differ by at most two elements, multiplicity counting.

Example.

- 21133212 is 3-distinguishable, but not cyclic 3-distinguishable
- 111222333 is both 3-distinguishable and cyclic 3-distinguishable

C. Savage, "A survey of combinatorial Gray codes," *SIAM Review*, vol. 39, no. 4, pp. 605–629, 1997.

An Equivalent Problem

For given m and sequence length M, minimize the number of colors k.

$$\Updownarrow$$

For given m and color size k, maximize the sequence length M. The maximum length is denoted by $M_m^c(k)$ or $M_m(k)$

Proposition (Natural Upper Bounds)

For given positive integers m and k, one has

$$M_m^c(k) \leq \binom{k+m-1}{m} \quad \text{and} \quad M_m(k) \leq \binom{k+m-1}{m} + m - 1.$$

Proof (for cyclic case).

- $S_t(m) = \{1^{e_1}, \ldots, k^{e_k}\}$, where e_s is the multiplicity of the element s
- $e_1 + e_2 + \cdots + e_k = m$ with each $e_s \geq 0$
- $\sharp$ of non-negative integral solutions for above is $H_m^k = \binom{k+m-1}{m}$

An Equivalent Problem

For given m and sequence length M, minimize the number of colors k.

$$\Updownarrow$$

For given m and color size k, maximize the sequence length M. The maximum length is denoted by $M_m^c(k)$ or $M_m(k)$

Proposition (Natural Upper Bounds)

For given positive integers m and k, one has

$$M_m^c(k) \leq \binom{k+m-1}{m} \quad \text{and} \quad M_m(k) \leq \binom{k+m-1}{m} + m - 1.$$

Proof (for cyclic case).

- $S_t(m) = \{1^{e_1}, \ldots, k^{e_k}\}$, where e_s is the multiplicity of the element s
- $e_1 + e_2 + \cdots + e_k = m$ with each $e_s \geq 0$
- $\sharp$ of non-negative integral solutions for above is $H_m^k = \binom{k+m-1}{m}$

Main Results

Related Works

- de Bruijn sequence: a cyclic sequence of length k^m in which every m-tuples in $[k]^m$ occurs exactly once as a substring.
  - $m = 4, k = 2$: 1111211221212222

Definition (Universal cycles for multisets, Mcycles)

An (m, k)-Mcycle is a cyclic m-distinguishable sequence S in which

- every m-multiset of $[k]$ appears exactly once as $S_t(m)$ for some t.

Example. $m = 3, k = 5$:

1112335 2223441 3334552 4445113 5551224

N. G. de Bruijn, "A combinatorial problem," *Porc. Koninklijke Nederlandse Akademie V. Wetenschappen*, vol 49, pp. 758–764, 1946.

G. Hurlbert, T. Johnson, and J. Zahl, "On universal cycles for multisets," *Discrete Math.*, vol. 309, pp. 5321–5327, 2009.

Related Works

- de Bruijn sequence: a cyclic sequence of length k^m in which every m-tuples in $[k]^m$ occurs exactly once as a substring.
  - $m = 4, k = 2$: 1111211221212222

Definition (Universal cycles for multisets, Mcycles)

An (m, k)-Mcycle is a cyclic m-distinguishable sequence S in which

- every m-multiset of $[k]$ appears exactly once as $S_t(m)$ for some t.

Example. $m = 3, k = 5$:

1112335 2223441 3334552 4445113 5551224

N. G. de Bruijn, "A combinatorial problem," *Porc. Koninklijke Nederlandse Akademie V. Wetenschappen*, vol 49, pp. 758–764, 1946.

G. Hurlbert, T. Johnson, and J. Zahl, "On universal cycles for multisets," *Discrete Math.*, vol. 309, pp. 5321–5327, 2009.

Mcycles

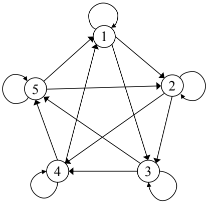
Definition (Universal cycles for multisets, Mcycles)

An (m, k)-Mcycle is a cyclic m-distinguishable sequence S in which

- every m-multiset of $[k]$ appears exactly once as $S_t(m)$ for some t.

Some facts:

- An (m, k)-Mcycle is of length $\binom{k+m-1}{m}$.
- An (m, k)-Mcycle exists, then $k \,\big|\, \binom{k+m-1}{m}$
- A $(2, k)$-Mcycle is an Eulerian circuit of K_k^ℓ (with loops)
- A $(3, k)$-Mcycle exists for $k \,\big|\, \binom{k+2}{3}$, i.e., $k \equiv 1, 2 \pmod 3$



1 1 2 2 3 3 4 4 5 5 1 3 5 2 4

Mcycles

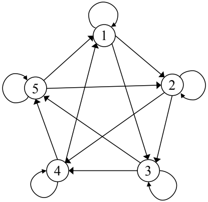
Definition (Universal cycles for multisets, Mcycles)

An (m,k)-Mcycle is a cyclic m-distinguishable sequence S in which

- every m-multiset of $[k]$ appears exactly once as $S_t(m)$ for some t.

Some facts:

- An (m,k)-Mcycle is of length $\binom{k+m-1}{m}$.
- An (m,k)-Mcycle exists, then $k \mid \binom{k+m-1}{m}$
- A $(\mathbf{2},k)$-Mcycle is an Eulerian circuit of K_k^{ℓ} (with loops)
- A $(\mathbf{3},k)$-Mcycle exists for $k \mid \binom{k+2}{3}$, i.e., $k \equiv 1, 2 \pmod 3$



Mcycles

Definition (Universal cycles for multisets, Mcycles)

An (m,k)-Mcycle is a cyclic m-distinguishable sequence S in which

- every m-multiset of $[k]$ appears exactly once as $S_t(m)$ for some t.

Some facts:

- An (m,k)-Mcycle is of length $\binom{k+m-1}{m}$.
- An (m,k)-Mcycle exists, then $k \mid \binom{k+m-1}{m}$
- A $(\mathbf{2},k)$-Mcycle is an Eulerian circuit of K_k^{ℓ} (with loops)
- A $(\mathbf{3},k)$-Mcycle exists for $k \mid \binom{k+2}{3}$, i.e., $k \equiv 1,2 \pmod 3$

Corollary (Hurlbert–Johnson–Zahl, 2009)

It follows from the results of Mcycles that

$$M_m^c(k) = \begin{cases} \binom{k+1}{2} & \text{if } m = 2 \text{ and } k \equiv 1 \pmod 2, \\ \binom{k+2}{3} & \text{if } m = 3 \text{ and } k \equiv 1,2 \pmod 3. \end{cases}$$

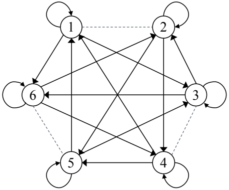
Main Results I

Theorem (–, 2024)

For the missing cases, we have

$$M_m^c(k) \geq \begin{cases} \binom{k+1}{2} - \frac{k}{2} & \text{if } m = 2 \text{ and } k \equiv 0 \pmod 2, \\ \binom{k+2}{3} - \frac{k}{3} & \text{if } m = 3 \text{ and } k \equiv 0 \pmod 3. \end{cases}$$

($m = 2$)



1 1 3 3 2 2 4 4 5

6 4 3 5 2 6 6 1 5

C. S. Chen, Y.-H. Lo, W. S. Wong, and Y. Zhang, "Object tracking using multiset color coding", in *International Symposium on Information Theory and Its Applications*, Taipei, Taiwan, Nov. 2024, pp. 378–383.

Main Results I

Theorem (–, 2024)

For the missing cases, we have

$$M_m^c(k) \geq \begin{cases} \binom{k+1}{2} - \frac{k}{2} & \text{if } m = 2 \text{ and } k \equiv 0 \pmod 2, \\ \binom{k+2}{3} - \frac{k}{3} & \text{if } m = 3 \text{ and } k \equiv 0 \pmod 3. \end{cases}$$

($m = 3$)

- By induction on k.
  - $k = 3$: 111222333
  - $k = 6$: 11122 23331 16631 55224 53532 44336 21414 62625 14365 55444 6665
- Assume on $[k-3]$, it is of the form $S'T'$
- For $[k]$, the obtained sequence is of the form $WXYZ$, where
  - $W = S'T'$
  - X is obtained from T' by $k-5 \mapsto k-2$, $k-4 \mapsto k-1$, and $k-3 \mapsto k$.
  - Y and Z are two special patterns, according to the parity of k.

C. S. Chen, Y.-H. Lo, W. S. Wong, and Y. Zhang, "Object tracking using multiset color coding", in *International Symposium on Information Theory and Its Applications*, Taipei, Taiwan, Nov. 2024, pp. 378–383.

Main Results I

Theorem (–, 2025+)

Suppose p is a prime and k is divided by p. Then

$$M_p^c(k) \le \binom{k+p-1}{p} - \frac{k}{p}.$$

Corollary (–, 2025+)

For the missing cases, we have

$$M_m^c(k) = \begin{cases} \binom{k+1}{2} - \frac{k}{2} & \text{if } m = 2 \text{ and } k \equiv 0 \pmod 2, \\ \binom{k+2}{3} - \frac{k}{3} & \text{if } m = 3 \text{ and } k \equiv 0 \pmod 3. \end{cases}$$

C. S. Chen, Y.-H. Lo, W. S. Wong, and Y. Zhang, "Object tracking using multiset color coding", in *International Symposium on Information Theory and Its Applications*, Taipei, Taiwan, Nov. 2024, pp. 378–383.

C. S. Chen, W. S. Wong, Y.-H. Lo, and T.-L. Wong, "Multiset combinatorial Gray codes with application to proximity sensor networks," *arXiv:2410.15428*

Main Results I

Theorem (–, 2025+)

Suppose p is a prime and k is divided by p. Then

$$M_p^c(k) \le \binom{k+p-1}{p} - \frac{k}{p}.$$

Proof Sketch.

- Fix an element $a \in [k]$.
- For any multiset A, let $\varphi_a(A) = \sharp$ of a's in A

1. S is a longest cyclic p-distinguishable sequence, $|S| = M$.
   - $\sum_{t=0}^{M-1} \varphi_a(S_p(t)) = p \cdot (\sharp \text{ of } a\text{'s in } S)$
2. $\mathcal{B}_{a,i}$ = collection of all p-multisets containing exactly $(p-i)$ a's
   - $\sum_{A \in \mathcal{B}_{a,i}} \varphi_a(A) = (p-i)H_i^{k-1}$, a multiple of p except when $i = 1$
   - $p \nmid \sum_{i=0}^{p-1} \sum_{A \in \mathcal{B}_{a,i}} \varphi_a(A)$

$\Rightarrow$ S can not include all p-multisets of $[k]$ containing at least one a

Main Results II: Synthetic Construction

Theorem (–, 2025+)

Let S be a cyclic m_1-distinguishable sequence on k_1 colors of length M_1 and T a cyclic m_2-distinguishable sequence on k_2 colors of length M_2. If $m_i | M_i$ and $\gcd(d, M_i/m_i d) = 1$ for $i = 1, 2$, where $d = \gcd(M_1/m_1, M_2/m_2) \geq 2$, then there exists a cyclic $(m_1 + m_2)$-distinguishable sequence with $k_1 + k_2$ colors of length $\frac{M_1 M_2}{d}(\frac{1}{m_1} + \frac{1}{m_2})$.

Example.

- $S = \underbrace{11}_{\alpha_0}\underbrace{33}_{\alpha_1}\underbrace{52}_{\alpha_2}\underbrace{41}_{\alpha_3}\underbrace{23}_{\alpha_4}\underbrace{45}_{\alpha_5}$
- $T = \underbrace{666}_{\beta_0}\underbrace{777}_{\beta_1}\underbrace{888}_{\beta_2}\underbrace{999}_{\beta_3}\underbrace{000}_{\beta_4}\underbrace{668}_{\beta_5}\underbrace{800}_{\beta_6}\underbrace{796}_{\beta_7}\underbrace{807}_{\beta_8}\underbrace{799}_{\beta_9}.$

$$\begin{aligned} S \times T = {} & \alpha_0\beta_0\alpha_1\beta_1\alpha_2\beta_2\alpha_3\beta_3\alpha_4\beta_4\alpha_5\beta_5 \quad \alpha_0\beta_6\alpha_1\beta_7\alpha_2\beta_8\alpha_3\beta_9\alpha_4\beta_0\alpha_5\beta_1 \\ & \alpha_0\beta_2\alpha_1\beta_3\alpha_2\beta_4\alpha_3\beta_5\alpha_4\beta_6\alpha_5\beta_7 \quad \alpha_0\beta_8\alpha_1\beta_9\alpha_2\beta_0\alpha_3\beta_1\alpha_4\beta_2\alpha_5\beta_3 \\ & \alpha_0\beta_4\alpha_1\beta_5\alpha_2\beta_6\alpha_3\beta_7\alpha_4\beta_8\alpha_5\beta_9. \end{aligned}$$

Main Results II: Synthetic Construction

Theorem (–, 2025+)

Let S be a cyclic m_1-distinguishable sequence on k_1 colors of length M_1 and T a cyclic m_2-distinguishable sequence on k_2 colors of length M_2. If $m_i|M_i$ and $\gcd(d, M_i/m_i d) = 1$ for $i = 1, 2$, where $d = \gcd(M_1/m_1, M_2/m_2) \geq 2$, then there exists a cyclic $(m_1 + m_2)$-distinguishable sequence with $k_1 + k_2$ colors of length $\frac{M_1 M_2}{d}(\frac{1}{m_1} + \frac{1}{m_2})$.

Theorem (–, 2025+)

For any m and k, by viewing $M_m^c(k)$ and $M_m(k)$ as functions of k, one has

$$M_m^c(k) = \theta(k^m) \quad \text{and} \quad M_m(k) = \theta(k^m).$$

2D Multiset Combinatorial Gray Codes

- 2D color mapping $\Phi : \mathbb{Z}_M \times \mathbb{Z}_N \to [k]$
- coding window size: $m \times n$
- For $(x_0, y_0) \in \mathbb{Z}_M \times \mathbb{Z}_N$, define a multiset

$$S_{m,n}(x_0, y_0) := \{\Phi(x_0 + i, y_0 + j) : 0 \leq i < m, 0 \leq j < n\}.$$

Definition

The color mapping Φ is called

- (m, n)-distinguishable if the multisets $S_{m,n}(x_0, y_0)$ are all distinct for $0 \leq x_0 < M - m$ and $0 \leq y_0 < N - n$
- cyclic (m, n)-distinguishable if the multisets $S_{m,n}(x_0, y_0)$ are all distinct for $(x_0, y_0) \in \mathbb{Z}_M \times \mathbb{Z}_N$.

Objective: given M, N, m, n, minimize k $\longleftarrow$ $K_{M,N}(m, n)$ or $K^c_{M,N}(m, n)$

2D Multiset Combinatorial Gray Codes

- 2D color mapping $\Phi : \mathbb{Z}_M \times \mathbb{Z}_N \to [k]$
- coding window size: $m \times n$
- For $(x_0, y_0) \in \mathbb{Z}_M \times \mathbb{Z}_N$, define a multiset

$$S_{m,n}(x_0, y_0) := \{\Phi(x_0 + i, y_0 + j) : 0 \leq i < m, 0 \leq j < n\}.$$

Definition

The color mapping Φ is called

- (m, n)-distinguishable if the multisets $S_{m,n}(x_0, y_0)$ are all distinct for $0 \leq x_0 < M - m$ and $0 \leq y_0 < N - n$
- cyclic (m, n)-distinguishable if the multisets $S_{m,n}(x_0, y_0)$ are all distinct for $(x_0, y_0) \in \mathbb{Z}_M \times \mathbb{Z}_N$.

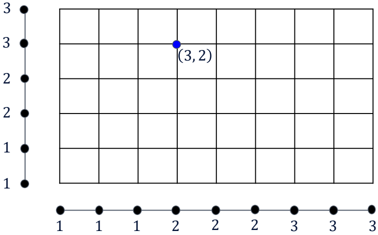
Objective: given M, N, m, n, minimize k $\longleftarrow$ $K_{M,N}(m, n)$ or $K^c_{M,N}(m, n)$

Product Code

- $S = s_0 s_1 \cdots s_{M-1}$ in which $s_i \in [k]$
- $T = t_0 t_1 \cdots t_{N-1}$ in which $t_i \in [h]$
- Define $\Phi(x, y) = (s_x, t_y), \forall (x, y) \in \mathbb{Z}_M \times \mathbb{Z}_N$.

Proposition

- The color mapping Φ is (cyclic) (m, n)-distinguishable if S is (cyclic) m-distinguishable and T is (cyclic) n-distinguishable.
- $K_{M,N}(m, n) \leq K_M(m) \cdot K_N(n)$ and $K^c_{M,N}(m, n) \leq K^c_M(m) \cdot K^c_N(n)$



Main Result III: color-coding gain

The number of bits to represent an ID :

- in "straightforward protocol", it is $\log_2 MN$
- in "multiset combinatorial Gray coding protocol", it is $\log_2 K_{M,N}(m,n)$

Define the **color-coding gain** as

$$\mathcal{R}_{M,N}(m,n) := \lim_{M,N\to\infty} \frac{\log_2 K_{M,N}(m,n)}{\log_2 MN}.$$

Theorem (–, 2025+)

It holds that

$$\frac{1}{mn} \le \mathcal{R}_{M,N}(m,n) \le \frac{\log_2 k + \log_2 h}{m\log_2 k + n\log_2 h},$$

where $k = K_M(m)$ and $h = K_N(n)$. In particular, when $m = n$, one has

$$\frac{1}{m^2} \le \mathcal{R}_{M,N}(m,m) \le \frac{1}{m}.$$

Main Result III: color-coding gain

The number of bits to represent an ID :

- in "straightforward protocol", it is $\log_2 MN$
- in "multiset combinatorial Gray coding protocol", it is $\log_2 K_{M,N}(m,n)$

Define the **color-coding gain** as

$$\mathcal{R}_{M,N}(m,n) := \lim_{M,N\to\infty} \frac{\log_2 K_{M,N}(m,n)}{\log_2 MN}.$$

Theorem (–, 2025+)

It holds that

$$\frac{1}{mn} \leq \mathcal{R}_{M,N}(m,n) \leq \frac{\log_2 k + \log_2 h}{m \log_2 k + n \log_2 h},$$

where $k = K_M(m)$ and $h = K_N(n)$. In particular, when $m = n$, one has

$$\frac{1}{m^2} \leq \mathcal{R}_{M,N}(m,m) \leq \frac{1}{m}.$$

Main Result III: color-coding gain

Let $k = K_M(m)$ and $h = K_N(n)$. Then,

$$\frac{1}{mn} \leq \mathcal{R}_{M,N}(m,n) \leq \frac{\log_2 k + \log_2 h}{m \log_2 k + n \log_2 h}.$$

Proof.

1. - $(M-m+1)(N-n+1) \leq H^k_{mn} = \binom{k+mn-1}{mn} \Rightarrow MN \leq \frac{k^{mn}}{(mn)!}$
   - Apply *Stirling's approximation formula*.

2. - Product Code: $K_{M,N}(m,n) \leq K_M(m) \cdot K_N(n)$
   - $\dfrac{\log_2 K_{M,N}(m,n)}{\log_2 MN} \leq \dfrac{\log_2 K_M(m) + \log_2 K_N(n)}{\log_2 M + \log_2 N}$
   - $M_m(k) = \theta(k^m) \Rightarrow M \approx k^m$ and $N \approx h^n$

Concluding Remarks

Remark 1: Universal Cycles

Definition (Universal cycles, Ucycles)

An (m,k)-Ucycle is a cyclic m-distinguishable sequence S on $[k]$ in which

1. there is no repeated elements in any $S_t(m)$, and
2. every m-subset of $[k]$ appears exactly once as $S_t(m)$ for some t.

Some facts:

- A (k,m)-Ucycle is of length $\binom{k}{m}$.
- A (k,m)-Ucycle exists, then $m \big| \binom{k-1}{m-1}$

Conjecture

For every $m \in \mathbb{N}$, there exists $n_o \in \mathbb{N}$ such that for all $k \geq n_o$, there exists an (m,k)-Ucycle whenever m divides $\binom{k-1}{m-1}$.

F. Chung, P. Diaconis, and R. Graham, "Universal cycles for combinatorial structures," *Discrete Math.*, vol. 110, pp. 43–59, 1992.

Remark 1: Universal Cycles

Theorem (Glock–Joos–Kühn–Osthus, 2020)

For every $m \in \mathbb{N}$, there exists $n_o \in \mathbb{N}$ such that for all $k \geq n_o$, there exists an (m,k)-Ucycle whenever m divides $\binom{k-1}{m-1}$.

- This theorem is proved by a probabilistic method.
- Some constructive results can be found in:
  - Chung–Diaconis–Graham (1992): $m = 2$
  - Jackson (1993): $m = 3$, partial $m = 4, 5$
  - Hurlbert (1994): $m = 3, 4, 6$ when $\gcd(m,k) = 1$.

S. Glock, F. Joos, D. Kühn, and D. Osthus, "Euler tours in hypergraphs," *Combinatorica*, vol. 40, pp. 679–690, 2020.

F. Chung, P. Diaconis, and R. Graham, "Universal cycles for combinatorial structures," *Discrete Math.*, vol. 110, pp. 43–59, 1992.

B. W. Jackson, "Universal cycles of k-subsets and k-permutations," *Discrete Math.*, vol. 117, pp. 141–150, 1993.

G. Hurlbert, "On universal cycles for k-subsets of an n-set," *SIAM J. Discrete Math.*, vol. 7, 598–604, 1994.

Remark 2: Higher Dimensional Extension

This concept was extended to d-dimension for any $d \geq 2$:

- $\Phi : \mathbb{Z}_M^d \to [k]$
- the window size: m^d

Theorem (Chen–Keevash–Kennedy–de Panafieu–Vetta (2024))

Fix a dimension $d \geq 2$ and the number of colors k of the form $bd+1$ for some $b \geq 1$. For any window size m multiple of $2(k-1)$, there exists a cyclic $(m)^d$-distinguishable integer lattice $\mathbb{Z}_M^d$ with

$$M \sim C_k^{1/d} \cdot m^{k-1} \quad \textit{where} \quad C_k = \left(\frac{2}{k-1}\right)^{k-1}.$$

- When $d = 2 \Rightarrow M \sim \left(\sqrt{\frac{2}{k-1}}m\right)^{k-1}$
- Product Code: $M \sim (\sqrt{k})^m > \left(\frac{1}{2q} \cdot m\right)^{(k-1)q}$, when set $m = 2(k-1)q$

C. S. Chen, P. Keevash, S. Kennedy, É. de Panafieu, and A. Vetta, "Robot positioning using torus packing for multisets," *51st International Colloquium on Automata, Languages, and Programming (ICALP)*, 2024.

Conclusion

- Formulate the problem of 1D and 2D multiset combinatorial Gray codes
- General upper bounds on the grid size
- For 1D multiset combinatorial Gray codes
  - $M_p^c(k) = \binom{k+p-1}{p} - \frac{k}{p}$ for any prime p and integer k with $p|k$
  - Exact values of $M_2^c(k)$ and $M_3^c(k)$
  - $M_m^c(k) = \theta(k^m)$
- For 2D multiset combinatorial Gray codes
  - Product Code
  - The color-coding gain has $\dfrac{1}{m^2} \leq \mathcal{R}_{M,N}(m,m) \leq \dfrac{1}{m}$

Future and Ongoing Works

1. For 1D case
   - Optimal constructions for (cyclic) m-distinguishable sequences, for $m \geq 4$.
   - For any any prime p and integer k with $p|k$, is $M_p^c(k) = \binom{k+p-1}{p} - \frac{k}{p}$?
   - Any other constructive methods for universal cycles?
   - The sufficient condition of the existence of an Mcycle is necessary?
2. For 2D case
   - Find a multiset combinatorial Gray code such that $\mathcal{R}_{M,N}(m,m) = \dfrac{1}{m^2}$
   - Any optimal construction?
3. Extend our results on 1D and 2D grids to higher dimensional cases.

Thank you for your listening