Discrepancy and nested perfect necklaces

Verónica $\operatorname{Becher}^{1^*}$ <u>Olivier Carton</u>^{2*}

 $^1 \mathrm{Universidad}$ de Buenos Aires

 $^2\mathrm{IRIF},$ Université Paris Cité & CNRS

 $^* \rm Supported$ by LIA SINFIN

LaBRI - November 2022

Outline

Normality

Discrepancy

Nested perfect necklaces

Affine necklaces

| | | |

Normal sequences

A sequence of symbols is **normal** if all finite words of the same length occur in it with the same frequency.

If $x \in A^{\mathbb{N}}$ and $w \in A^*$, the frequency of w in x is defined by

$$\operatorname{freq}(x,w) = \lim_{N \to \infty} \frac{|x[1:N]|_w}{N}.$$

where $|z|_w$ denotes the number of occurrences of w in z.

A sequence $x \in A^{\mathbb{N}}$ is normal if for each $w \in A^*$:

$$freq(x,w) = \frac{1}{(\#A)^{|w|}}$$

where ► #A is the cardinality of the alphabet A |w| is the length of w.

Normal sequences (continued)

Theorem (Borel, 1909)

The decimal expansion of almost every real number in [0, 1) is a normal sequence of symbols in $\{0, 1, \ldots, 9\}$.

Nevertheless, not so many examples have been proved normal. Some of them are:

• Champernowne 1933 (natural numbers):

 $12345678910111213141516171819202122232425\cdots$

► Besicovitch 1935 (squares):

 $149162536496481100121144169196225256289324\cdots$

► Copeland and Erdős 1946 (primes):

 $235711131719232931374143475359616771737983\cdots$

Equidistribution

A sequence $(x_n)_{n \ge 1}$ of real numbers in [0; 1) is equidistributed if for any real numbers α, β such that $0 \le \alpha \le \beta \le 1$,

$$\lim_{N \to \infty} \frac{\#\{1 \le n \le N : x_n \in [\alpha; \beta]\}}{N} = \beta - \alpha$$

Example

The sequence $(\{n\alpha\})_{n\geq 1}$ is equidistributed for each irrational real number α (where $\{x\} = x - \lfloor x \rfloor$ is the fractional part of x).

Proposition

The base b expansion of the real number α is normal if and only if the sequence $(\{b^n \alpha\})_{n \ge 1}$ is equidistributed.

Equidistribution

A sequence $(x_n)_{n \ge 1}$ of real numbers in [0; 1) is equidistributed if for any real numbers α, β such that $0 \le \alpha \le \beta \le 1$,

$$\lim_{N \to \infty} \frac{\#\{1 \le n \le N : x_n \in [\alpha; \beta]\}}{N} = \beta - \alpha$$

Example

The sequence $(\{n\alpha\})_{n\geq 1}$ is equidistributed for each irrational real number α (where $\{x\} = x - \lfloor x \rfloor$ is the fractional part of x).

Proposition

The base b expansion of the real number α is normal if and only if the sequence $(\{b^n \alpha\})_{n \ge 1}$ is equidistributed.

The block w of length k occurs in α at position n + 1 iff $\{b^n \alpha\}$ lies in the interval $[0.w; 0.w + 2^{-k}]$.

Discrepancy

The discrepancy D_N of a sequence $(x_n)_{n \ge 1}$ of real numbers in [0; 1) is defined by

$$D_N = \sup_{0 \le \alpha \le \beta \le 1} \left| \frac{\#\{1 \le n \le N : x_n \in [\alpha; \beta]\}}{N} - (\beta - \alpha) \right|.$$

Proposition

The sequence
$$(x_n)_{n \ge 1}$$
 is equidistributed iff $\lim_{N \to \infty} D_N = 0$.

| | | |

Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.



Theorem (Schmidt 72)

There exists a constant C > 0 such that for any sequence of real numbers, its discrepancy satisfies $D_N \ge C \log N/N$ for infinitely many $N \ge 1$.

This lower bound is achieved by the van der Corput sequence.



For a sequence of the form $(\{b^n \alpha\})_{n \ge 1}$, the lowest discrepancy known so far is $(\log N)^2/N$. It is achieved by Levin's sequence.

A necklace is a cyclically considered word (written on a circle).



A necklace is a cyclically considered word (written on a circle).



A necklace is a cyclically considered word (written on a circle).



A necklace is a cyclically considered word (written on a circle).



A necklace is a cyclically considered word (written on a circle).



A necklace is a cyclically considered word (written on a circle).



A necklace is a cyclically considered word (written on a circle).



A necklace is a cyclically considered word (written on a circle).



A necklace is a cyclically considered word (written on a circle).



De Bruijn necklaces as Eulerian circuits

A k-de Bruijn necklace is Eulerian circuit in the local graph whose vertices are words of length k - 1 and edges words of length k.



De Bruijn necklaces as Eulerian circuits

A k-de Bruijn necklace is Eulerian circuit in the local graph whose vertices are words of length k - 1 and edges words of length k.

The 3-de Bruijn necklace 00010111 corresponds to the following Eulerian circuit.



BEST theorem

Theorem (BEST 1951)

The number ec(G) of Eulerian circuits in a connected graph G = (V, E) is

$$ec(G) = t_w(G) \prod_{v \in V} (deg(v) - 1)!$$

where $t_w(G)$ is the number of spanning trees directed to w. This number is equal to any cofactor of the Laplacian matrix.

Proposition

The number of k-de Bruijn necklaces over an alphabet of size n is $\frac{(n!)^{n^{k-1}}}{n^k}.$

For n = 2, $\blacktriangleright k = 1$: 01 (and 10) $\blacktriangleright k = 2$: 0011 (0110, 1100 and 1001)

• k = 3 : 00010111 and 11101000

A (k, m)-perfect necklace is a necklace in which each word of length k has m occurrences at positions which are different modulo m. Its length must be $m(\#A)^k$.

A (k, m)-perfect necklace is a necklace in which each word of length k has m occurrences at positions which are different modulo m. Its length must be $m(\#A)^k$.

1	0	0	word	mod		word	\mod
1		0	00	1	\checkmark	10	1
0		1	00	0		10	0
0		1	01	1		11	1
0	1	T	01	0		11	0

A (k, m)-perfect necklace is a necklace in which each word of length k has m occurrences at positions which are different modulo m. Its length must be $m(\#A)^k$.

1	0	0	word	mod		word	\mod
1		0	00	1	\checkmark	10	1
0		1	00	0		10	0
0		1	01	1		11	1
U	1	T	01	0	\checkmark	11	0

A (k, m)-perfect necklace is a necklace in which each word of length k has m occurrences at positions which are different modulo m. Its length must be $m(\#A)^k$.

$\begin{array}{c} 0\\ 1\end{array}$	0	word	mod		word	mod		
		0	00	1	\checkmark	10	1	
0		1	00	0		10	0	
0 1		1	01	1		11	1	\checkmark
	1	1	01	0	\checkmark	11	0	

A (k, m)-perfect necklace is a necklace in which each word of length k has m occurrences at positions which are different modulo m. Its length must be $m(\#A)^k$.

1	0	0	word	mod		word	mod	
T		0	00	1	\checkmark	10	1	
0		1	00	0		10	0	
0		1	01	1		11	1	\checkmark
	1	1	01	0	\checkmark	11	0	√

A (k, m)-perfect necklace is a necklace in which each word of length k has m occurrences at positions which are different modulo m. Its length must be $m(\#A)^k$. The perfect perfect perfect perfect perfect perfect perfect.

1	0	0	word	mod		word	mod	
1		0	00	1	\checkmark	10	1	\checkmark
0		1	00	0		10	0	
$egin{array}{c} 0 & & \\ & 1 & \end{array}$		1	01	1		11	1	\checkmark
	1	1	01	0	\checkmark	11	0	\checkmark

A (k, m)-perfect necklace is a necklace in which each word of length k has m occurrences at positions which are different modulo m. Its length must be $m(\#A)^k$.

1	0	0	word	mod		word	mod	
1	0	00	1	\checkmark	10	1	\checkmark	
0		1	00	0	\checkmark	10	0	
0		1	01	1		11	1	\checkmark
0	1	1	01	0	\checkmark	11	0	\checkmark

A (k, m)-perfect necklace is a necklace in which each word of length k has m occurrences at positions which are different modulo m. Its length must be $m(\#A)^k$.

1	0	0	word	mod		word	mod	
1	0	00	1	\checkmark	10	1	\checkmark	
0		1	00	0	\checkmark	10	0	
0		1	01	1	\checkmark	11	1	\checkmark
0	1	T	01	0	\checkmark	11	0	\checkmark

A (k, m)-perfect necklace is a necklace in which each word of length k has m occurrences at positions which are different modulo m. Its length must be $m(\#A)^k$.

1	0	0	word	mod		word	mod	
1	0	00	1	\checkmark	10	1	\checkmark	
0		1	00	0	\checkmark	10	0	\checkmark
$egin{array}{c} 0 & & \\ & 1 & \end{array}$		1	01	1	\checkmark	11	1	\checkmark
	1	T	01	0	\checkmark	11	0	\checkmark

A (k, m)-perfect necklace is a necklace in which each word of length k has m occurrences at positions which are different modulo m. Its length must be $m(\#A)^k$. The necklace 00010111 is not a (2, 2)-perfect necklace because

the two occurrences of 01 have the same position modulo 2. The necklace 00111001 is a (2, 2)-perfect necklace.

1	0	0	word	mod		word	mod	
T		0	00	1	\checkmark	10	1	\checkmark
0		1	00	0	\checkmark	10	0	\checkmark
$egin{array}{c} 0 & & \\ & 1 & \end{array}$		1	01	1	\checkmark	11	1	\checkmark
	1	T	01	0	\checkmark	11	0	\checkmark

More generally, the concatenation in lexicographic order of all words of length k gives a (k, k)-perfect necklace. Here is k = 3:

000001010011100101110111.

Perfect necklaces as Eulerian circuits

A (k, m)-perfect is an Eulerian circuit in the graph whose vertex set is $A^{k-1} \times \{0, \dots, m-1\}$.



Perfect necklaces as Eulerian circuits

A (k, m)-perfect is an Eulerian circuit in the graph whose vertex set is $A^{k-1} \times \{0, \ldots, m-1\}$. The (2, 2)-perfect necklace 00111001 corresponds to the following Eulerian circuit.



Nested perfect necklaces

A (k, m)-nested perfect necklace is a necklace such that for each $1 \leq \ell \leq k$, each block of length $m(\#A)^{\ell}$ starting at a position multiple of $m(\#A)^{\ell}$ is a (ℓ, m) -perfect necklace. Its length is $m(\#A)^k$.

Example

The necklace 00111001 is a (2, 2)-nested perfect necklace since

- ▶ 00111001 is a (2, 2)-perfect necklace
- ▶ 0011 and 1001 are (1, 2)-perfect necklaces

Nested perfect necklaces

A (k, m)-nested perfect necklace is a necklace such that for each $1 \leq \ell \leq k$, each block of length $m(\#A)^{\ell}$ starting at a position multiple of $m(\#A)^{\ell}$ is a (ℓ, m) -perfect necklace. Its length is $m(\#A)^k$.

Example

- is a (4, 4)-nested perfect necklace since
 - it is a (4, 4)-perfect necklace
 - made of two (3, 4)-perfect necklaces

• made of four (2, 4)-perfect necklaces

Proposition

Levin's sequence of discrepancy $(\log N)^2/N$ is a concatenation $u_0u_1u_2\cdots$ where each u_n is a $(2^n, 2^n)$ -nested perfect necklace of length 2^{n+2^n} :



Levin's construction

Let $n \ge 0$. Let \hat{M}_n be the $2^n \times 2^n$ -matrix defined by $M_{1,i} = M_{i,2^n} = 1$ for $1 \le i \le 2^n$ and $M_{i,j} = M_{i-1,j} + M_{i,j+1}$. $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 4 & 3 & 2 & 1 \end{pmatrix}$

$$\hat{M}_2 = \begin{pmatrix} 4 & 3 & 2 & 1 \\ 10 & 6 & 3 & 1 \\ 20 & 10 & 4 & 1 \end{pmatrix}$$

Let M_n be the matrix over \mathbb{F}_2 defined by $M_n \equiv \hat{M}_n \mod 2$. The matrices M_n can also be defined recursively by

$$M_0 = (1)$$
 and $M_{n+1} = \begin{pmatrix} M_n & M_n \\ 0 & M_n \end{pmatrix}$.

The matrices M_1 and M_2 are then

$$M_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$
 and $M_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

| | | | |

Levin's construction (continued)

Words over $\{0, 1\}$ are identified with column vectors over \mathbb{F}_2 . Let $w_1, \ldots, w_{2^{2^n}}$ be the enumeration in lexicographic order of all words of length 2^n over $\{0, 1\}$. Let $w'_i = M_n w_i$ for each $1 \leq i \leq 2^{2^n}$.

Theorem

The concatenation $w'_1 \cdots w'_{2^{2^n}}$ is a $(2^n, 2^n)$ -nested perfect necklace.

Examples

For $n = 1$,		$i \mid$	1	2	3	4	
	_	w_i	00	01	10	11	
		w_i'	00	11	10	01	
For $n = 2$,	i	1		2		16	j
	w_i	000)0	0001		111	
	w'_i	000	00	1111		000)

More matrices

Let k_1, \ldots, k_m be integers such that $k_i - 1 \leq k_{i+1} \leq k_i$ for $1 \leq i < 2^n$ and $k_{2^n} = 0$ Define

$$M_n^{k_1,...,k_{2^n}} = (\sigma^{k_1}(C_1),\ldots,\sigma^{k_m}(C_m)).$$

where $M_n = (C_1, \ldots, C_m)$. For n = 2, the eight matrices:



Affine necklaces

Let M be one matrix $M_n^{k_1,\ldots,k_{2^n}}$ and z be a fixed vector of length 2^n . Let $w'_i = Mw_i + z$ for each $1 \le i \le 2^{2^n}$. An affine necklace is the concatenation $w'_1 \cdots w'_{2^n}$.

Theorem

A necklace is a nested perfect necklace if and only if it is an affine necklace. There are $2^{2^{n+1}-1}$ $(2^n, 2^n)$ -nested perfect necklaces.

Theorem

Any sequence which is the concatenation of $(2^n, 2^n)$ -nested perfect necklaces for n = 0, 1, 2, ... has discrepancy $(\log N)^2/N$ (as the one defined by Levin).

Open questions

- What is the number of (k, m)-nested perfect necklaces over some alphabet A? We only know when #A = 2 and m is a power of 2.
- ▶ What is the least possible discrepancy for a sequence of the form $(\{b^n \alpha\})_{n \ge 1}$ for some real number α ?
- ▶ 2-dimensional objects : nested perfect toroidal arrays



Merci