Normal, Abby Normal, Prefix Normal

Péter Burcsi, Gabriele Fici, Zsuzsanna Lipták, Frank Ruskey, and Joe Sawada

FUN 2014
Lipari, 1-3 July 2014

Prefix normal words

Fici, Lipták (DLT 2011)

Definition

A binary word w is prefix normal (w.r.t. 1) if no substring has more 1s than the prefix of the same length.

Example

$$w = 1011000100110111 0010$$

$$w' = 11101001011001010010$$

Prefix normal words

Fici, Lipták (DLT 2011)

Definition

A binary word w is prefix normal (w.r.t. 1) if no substring has more 1s than the prefix of the same length.

Example

$$w = 10110001001101110010 \quad NO$$

$$w' = 11101001011001010010 \quad YES$$

Prefix normal words

Fici, Lipták (DLT 2011)

Definition

A binary word w is prefix normal (w.r.t. 1) if no substring has more 1s than the prefix of the same length.

Example

$$w = 1011000100110111 0010 \quad \textit{NO}$$

$$w' = 1110100101100101 0010 \quad \textit{YES}$$

all p.n. words of length 4:
$1111, 1110, 1101, 1100, 1010, 1001, 1000, 0000.$

Prefix normal games

Let's play a game.

Alice and Bob are constructing a binary string of length n together.
Alice wins if the word is prefix normal.
Bob wins if it is not.

Prefix normal games

Let's play a game.

Alice and Bob are constructing a binary string of length n together.
Alice wins if the word is prefix normal.
Bob wins if it is not.

Example ($n = 5$)

1. *start* _ _ _ _ _

Prefix normal games

Let's play a game.

Alice and Bob are constructing a binary string of length n together.
Alice wins if the word is prefix normal.
Bob wins if it is not.

Example ($n = 5$)

1.	*start*	_	_	_	_	_
2.	*Alice*	1	_	_	_	_

Prefix normal games

Let's play a game.

Alice and Bob are constructing a binary string of length n together.
Alice wins if the word is prefix normal.
Bob wins if it is not.

Example ($n = 5$)

1.	*start*	_	_	_	_	_
2.	*Alice*	1	_	_	_	_
3.	*Bob*	1	0	_	_	_

Prefix normal games

Let's play a game.

Alice and Bob are constructing a binary string of length n together.
Alice wins if the word is prefix normal.
Bob wins if it is not.

Example ($n = 5$)

1.	*start*	_	_	_	_	_
2.	*Alice*	1	_	_	_	_
3.	*Bob*	1	0	_	_	_
4.	*Alice*	1	0	_	0	_

Prefix normal games

Let's play a game.

Alice and Bob are constructing a binary string of length n together.
Alice wins if the word is prefix normal.
Bob wins if it is not.

Example ($n = 5$)

1.	*start*	_	_	_	_	_
2.	*Alice*	1	_	_	_	_
3.	*Bob*	1	0	_	_	_
4.	*Alice*	1	0	_	0	_

Alice has won: $10000, 10100, 10101$ are all p.n.

Prefix normal games

Exercise

Find n_0 s.t.

- Alice has a winning strategy for all $n < n_0$, and
- Bob has a winning strategy for all $n \geq n_0$.

Some properties of prefix normal words

- w is p.n. $\Rightarrow$ $w0$ is p.n.
- w is p.n. $\Rightarrow$ every prefix of w is p.n.
- for any w of length n, $v = 1^n w$ is p.n.
- ...

Prefix normal forms

- Let $F(w, k)$ = max #1s in a substring of length k

  Ex. $\mathbf{w} = \mathbf{11010010101}$, $F(w, 1) = 1, F(w, 2) = 2, F(w, 3) = 2, \ldots$

Prefix normal forms

- Let $F(w,k)$ = max #1s in a substring of length k

  Ex. $\mathbf{w} = \mathbf{11010010101}$, $F(w,1) = 1, F(w,2) = 2, F(w,3) = 2, \ldots$

  If w p.n. then for all k:

$$\#1\text{s in pref}(w,k) = F(w,k) \quad \text{(by definition)}$$

Prefix normal forms

- Let $F(w,k)$ = max #1s in a substring of length k

  Ex. $\mathbf{w} = \mathbf{11010010101}$, $F(w,1) = 1, F(w,2) = 2, F(w,3) = 2, \ldots$

  If w p.n. then for all k:

  $$\#1\text{s in pref}(w,k) = F(w,k) \quad \text{(by definition)}$$

- **prefix normal form**:
  For every w exists unique p.n. word w' s.t.

  $$\forall k : F(w,k) = F(w',k)$$

  $w' = \mathrm{PNF}_1(w)$. (next slides)

Where do prefix normal words come from?

Binary Jumbled Pattern Matching

Does $\mathbf{w} = \mathbf{10100110110001110010}$ have a substring of length 11 containing exactly 5 ones?

Where do prefix normal words come from?

Binary Jumbled Pattern Matching

Does $\mathbf{w} = \mathbf{10100110110001110010}$ have a substring of length 11 containing exactly 5 ones? (Online: easy, $\mathcal{O}(|w|)$. Indexed: ?)

Where do prefix normal words come from?

Binary Jumbled Pattern Matching

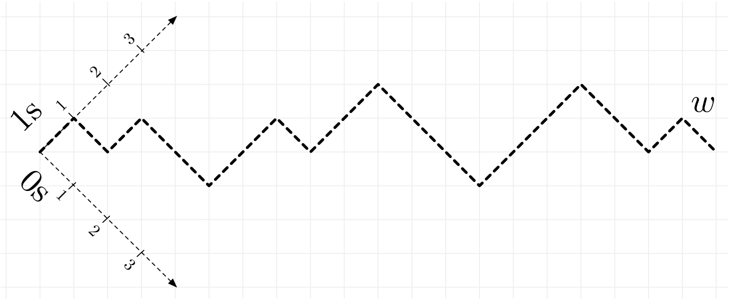
Does $\mathbf{w} = \mathbf{10100110110001110010}$ have a substring of length 11 containing exactly 5 ones? (Online: easy, $\mathcal{O}(|w|)$. Indexed: ?)

Recent papers on this problem:
PSC 2009, FUN 2010, IPL 2010, JDA 2012, ToCS 2012, IJFCS 2012, CPM 2012, SPIRE 2012, IPL 2013, IPL 2013, ESA 2013 x 2, SPIRE 2013, TCS 2014, PhTRS-A 2014, CPM 2014, CPM 2014, ISIT 2014, ICALP 2014, ...

(red ones have an intersection with the authors of this paper)

BJPM with prefix normal forms

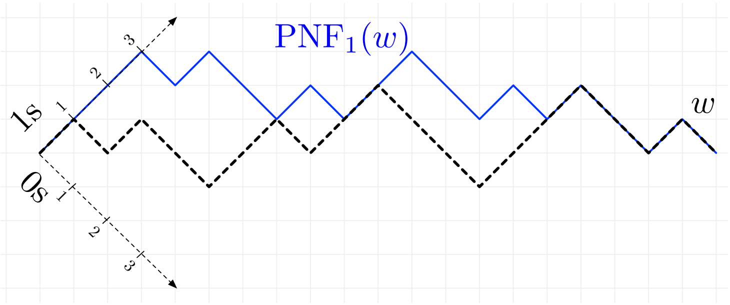
Does $\mathbf{w} = \mathbf{1010011011000111 0010}$ have a substring of length 11 containing exactly 5 ones?



$\diagup = 1$, $\diagdown = 0$,

BJPM with prefix normal forms

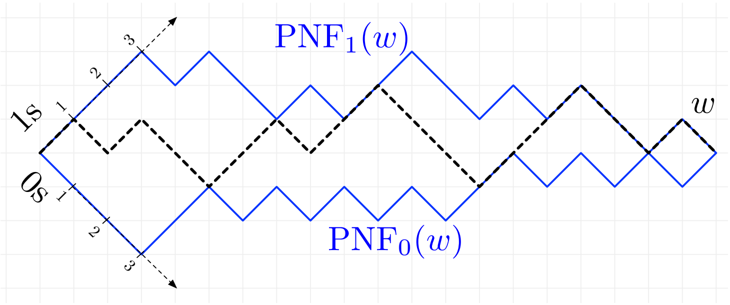
Does **w** = **10100110110001110010** have a substring of length 11 containing exactly 5 ones?



╱ = 1, ╲ = 0,

BJPM with prefix normal forms

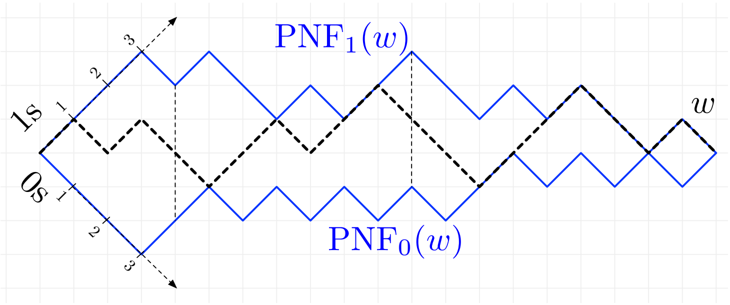
Does **w** = **101001101100011100100** have a substring of length 11 containing exactly 5 ones?



╱ = 1, ╲ = 0, area between blue lines: all substrings of *w*

BJPM with prefix normal forms

Does **w** = **10100110110001110010** have a substring of length 11 containing exactly 5 ones?

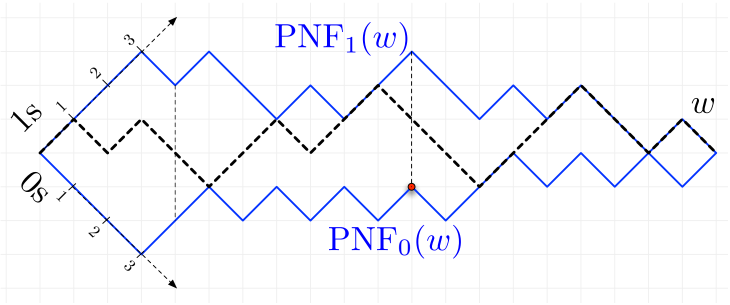


$\diagup = 1$, $\diagdown = 0$, area between blue lines: all substrings of w

verticals: fixed length substrings $k = 4, 11$.

BJPM with prefix normal forms

Does $\mathbf{w} = \mathbf{10100110110001110010}$ have a substring of length 11 containing exactly 5 ones? YES



$\diagup = 1$, $\diagdown = 0$, area between blue lines: all substrings of w
verticals: fixed length substrings $k = 4, 11$.

BJPM with prefix normal forms

Does $\mathbf{w} = \mathbf{10100110110001110010}$ have a substring of length 11 containing exactly 5 ones?

$\#$ 1s in $\text{pref}(\mathrm{PNF}_1(w), 11) = 7$
$\#$ 1s in $\text{pref}(\mathrm{PNF}_0(w), 11) = 5$

$$7 \geq 5 \geq 5 \quad \leadsto \text{YES}$$

Thus, fast computation of PNFs yields fast solution to BJPM.

BJPM with prefix normal forms

Does $\mathbf{w} = \mathbf{1010011011000111001 0}$ have a substring of length 11 containing exactly 5 ones?

$\#$ 1s in pref$(\mathrm{PNF}_1(w), 11) = 7$
$\#$ 1s in pref$(\mathrm{PNF}_0(w), 11) = 5$

$$7 \geq 5 \geq 5 \quad \rightsquigarrow \text{YES}$$

Thus, fast computation of PNFs yields fast solution to BJPM.
Better understanding p.n. words might be useful for BJPM.

BJPM with prefix normal forms

Does $\mathbf{w} = \mathbf{101001101100011100010}$ have a substring of length 11 containing exactly 5 ones?

$\#$ 1s in pref$(\mathrm{PNF}_1(w), 11) = 7$
$\#$ 1s in pref$(\mathrm{PNF}_0(w), 11) = 5$

$$7 \geq 5 \geq 5 \quad \rightsquigarrow \text{YES}$$

Thus, fast computation of PNFs yields fast solution to BJPM.
Better understanding p.n. words might be useful for BJPM.
Or might just be fun.

Some questions about prefix normal words

1. Compute prefix normal forms: Given $w \in \{0,1\}^n$, compute $\mathrm{PNF}_1(w)$ – for Binary Jumbled Pattern Matching. – next

Some questions about prefix normal words

1. Compute prefix normal forms: Given $w \in \{0,1\}^n$, compute $\mathrm{PNF}_1(w)$ – for Binary Jumbled Pattern Matching. – next
2. Testing: Given w, is w prefix normal? – naïve: $O(n^2)$, lit: $O(n^2/polylog\, n)$, this paper: $O(n^2)$ w-c and $O(n)$ expected conjectured

Some questions about prefix normal words

1. Compute prefix normal forms: Given $w \in \{0,1\}^n$, compute $\mathrm{PNF}_1(w)$ – for Binary Jumbled Pattern Matching. – next
2. Testing: Given w, is w prefix normal? – naïve: $O(n^2)$, lit: $O(n^2/polylog\, n)$, this paper: $O(n^2)$ w-c and $O(n)$ expected conjectured
3. Generation (listing): List all prefix normal words of length n. – our paper CPM 2014, based on p.n. words being a bubble language

Some questions about prefix normal words

1. Compute prefix normal forms: Given $w \in \{0,1\}^n$, compute $\mathrm{PNF}_1(w)$ – for Binary Jumbled Pattern Matching. – next
2. Testing: Given w, is w prefix normal? – naïve: $O(n^2)$, lit: $O(n^2/polylog\, n)$, this paper: $O(n^2)$ w-c and $O(n)$ expected conjectured
3. Generation (listing): List all prefix normal words of length n. – our paper CPM 2014, based on p.n. words being a bubble language
4. Enumeration (counting): number of p.n. words of length n? – partial results in this paper, ongoing

Some questions about prefix normal words

1. Compute prefix normal forms: Given $w \in \{0,1\}^n$, compute $\mathrm{PNF}_1(w)$ – for Binary Jumbled Pattern Matching. – next
2. Testing: Given w, is w prefix normal? – naïve: $O(n^2)$, lit: $O(n^2/polylog\, n)$, this paper: $O(n^2)$ w-c and $O(n)$ expected conjectured
3. Generation (listing): List all prefix normal words of length n. – our paper CPM 2014, based on p.n. words being a bubble language
4. Enumeration (counting): number of p.n. words of length n? – partial results in this paper, ongoing
5. In general, properties of p.n. words. – ongoing

Computing the PNF

Our contribution to mechanical algorithm design. Uses a folding ruler and Lipari's Canneto black sandy beach:

Computing the PNF

Our contribution to mechanical algorithm design. Uses a folding ruler and Lipari's Canneto black sandy beach:

Analysis

Uses a quadratic amount of sand but can be faster than other algorithms if implemented by a very fast person.

Enumerating p.n. words

Let $pnw(n) = \#$ prefix normal words of length n.

Easy:

$pnw(n)$ grows exponentially.
Proof: $\forall w : 1^{|w|}w$ is p.n. $\quad \Rightarrow \quad pnw(2n) \geq 2^n$.

Enumerating p.n. words

Let $pnw(n) = \#$ prefix normal words of length n.

Easy:

$pnw(n)$ grows exponentially.
Proof: $\forall w : 1^{|w|}w$ is p.n. $\quad \Rightarrow \quad pnw(2n) \geq 2^n$.

Theorem

Exists $c > 0$*:* $pnw(n) = \Omega(2^{c\sqrt{n \ln n}}) = \Omega((2-\epsilon)^n) \quad \forall \epsilon > 0$.

(Proof uses a variant of the prefix normal game.)

Theorem

$pnw(n) = O\left(\frac{2^n(\ln n)^2}{n}\right) = o(2^n)$.

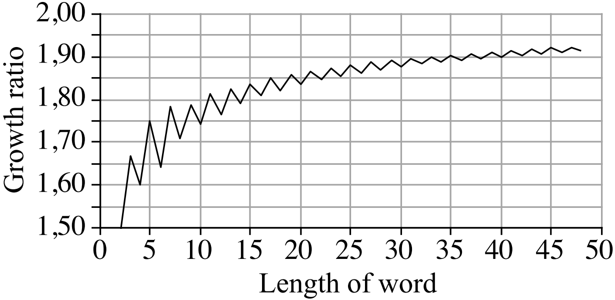
Enumerating p.n. words

Thanks to the new generating algorithm, we can now count p.n. words up to $n = 50$, and also special subsets, e.g. fixed-density, fixed-prefix, fixed-suffix, ... And make intelligent guesses.

Enumerating p.n. words

OEIS sequence no. A194850

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$pnw(n)$	2	3	5	8	14	23	41	70	125	218	395	697	1273	2279	4185	7568



Fixed-density p.n. words

For a word w, its density d is the number of 1s in w,
$pnw(n,d)$ = no. of p.n. words of length n and density d.

Let $f_d(x) = \sum_{n \geq 0} pnw(n,d) x^n$, the generating function of $pnw(n,d)$.

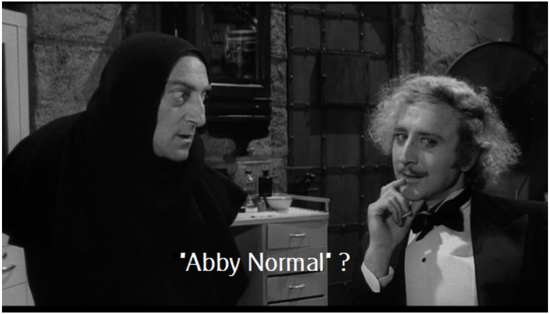
gen.func.	coefficients	closed form
$f_0(x) = \frac{1}{1-x}$	$1,1,1,1,1,1,1,1\ldots$	1
$f_1(x) = \frac{x}{1-x}$	$0,1,1,1,1,1,1,1\ldots$	$[n > 0]$
$f_2(x) = \frac{x^2}{(1-x)^2}$	$0,0,1,2,3,4,5,6\ldots$	$(n-1)$
$f_3(x) = \frac{x^3}{(1-x^2)(1-x)^2}$	$0,0,0,1,2,4,6,9,\ldots$	$pnw(2n) = n(n-1)$, $pnw(2n+1) = (n-1)^2$

...and more (but for larger d becomes less and less manageable) ...

prefix normal not to be confused with abnormal,

prefix normal not to be confused with abnormal, AB normal,

prefix normal not to be confused with abnormal, AB normal, abby normal



prefix normal not to be confused with abnormal, AB normal, abby normal

Mel Brooks: Young Frankenstein (1974)

THANK YOU!