Proving Dejean's Conjecture

Narad Rampersad

Department of Mathematics and Statistics
University of Winnipeg

Words

- A word (string) is a sequence of letters (symbols).
- Can be finite or infinite
- Important properties of words: periodicity and repetitions
- Other patterns in words: palindromes, etc.

Repetitions in words

- $w = a_1 a_2 \cdots a_\ell$ has period p if $a_i = a_{i+p}$ for each i.
- If w has length ℓ and period p, then w is a k-power, where $k = \ell/p$.
- k is the exponent of w.

Example

- 0010010 has periods 3 and 6 and is a $7/3$-power.
- 10011001 has periods 4 and 7 and is a 2-power.

Avoiding repetitions

- w avoids k-powers if no factor (substring) is a k-power.
- w avoids k^+-powers if for every $r > k$ no factor is an r-power.

Example

- 012021012102 avoids 2-powers.
- 0110100110010110 avoids 2^+-powers.

Infinite words avoiding repetitions

Avoiding 2-powers with 3 letters (Thue 1906)

Iterate the morphism $0 \to 012;\ 1 \to 02;\ 2 \to 1$:

$$0 \to 012 \to 012021 \to 012021012102 \to \cdots$$

Avoiding 2^+-powers with 2 letters (Thue 1912)

Iterate $0 \to 01;\ 1 \to 10$:

$$0 \to 01 \to 0110 \to 01101001 \to 0110100110010110 \to \cdots$$

Generalizing Thue

- Françoise Dejean (1972) introduced k-powers for non-integral k.
- Thue (1912): 2-powers are avoidable with 3 letters.
- Dejean (1972): $(7/4)^+$-powers are avoidable with 3 letters.
- Both constructions: by iteration of a morphism
- $7/4$ best possible with 3 letters

Dejean's Conjecture

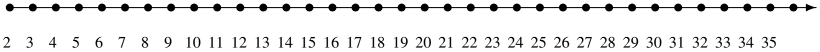
repetition threshold:

$$RT(n) = \inf\{k : \text{some infinite word over an } n\text{-letter alphabet avoids } k\text{-powers}\}$$

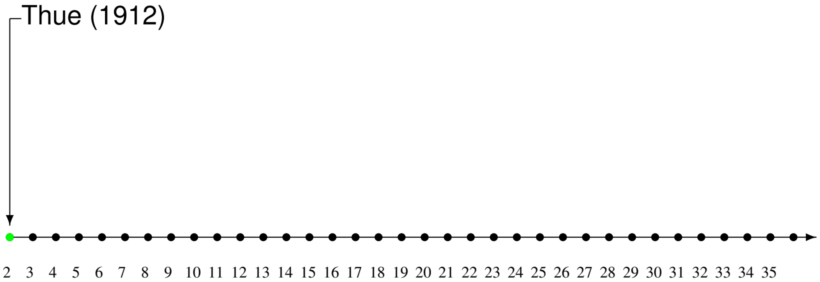
Dejean's Conjecture (1972)

$$RT(n) = \begin{cases} 7/4, & n = 3 \\ 7/5, & n = 4 \\ n/(n-1), & n = 2 \text{ or } n \geq 5. \end{cases}$$

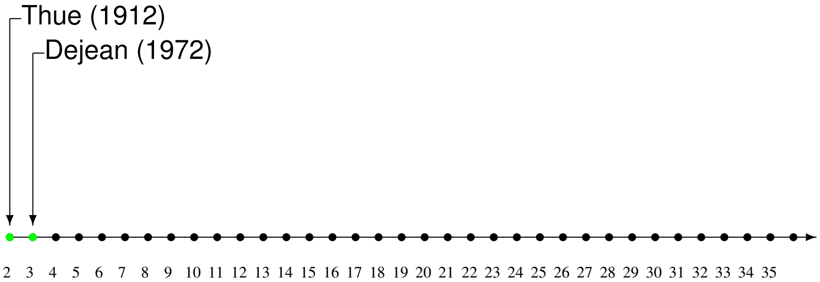
History of the conjecture



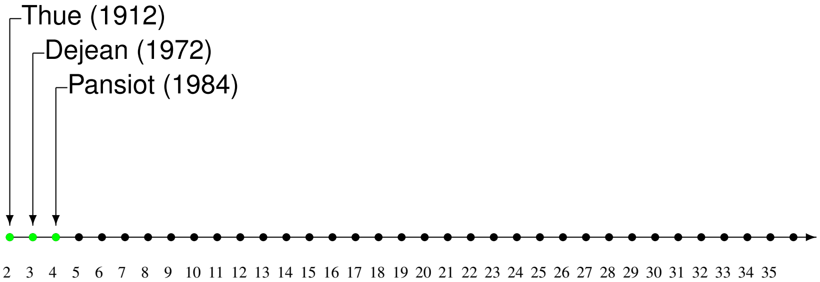
History of the conjecture



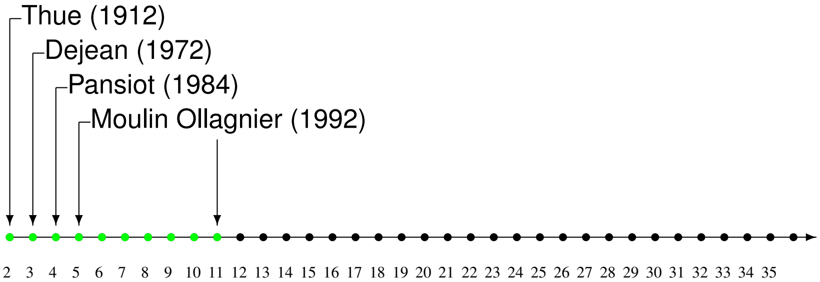
History of the conjecture



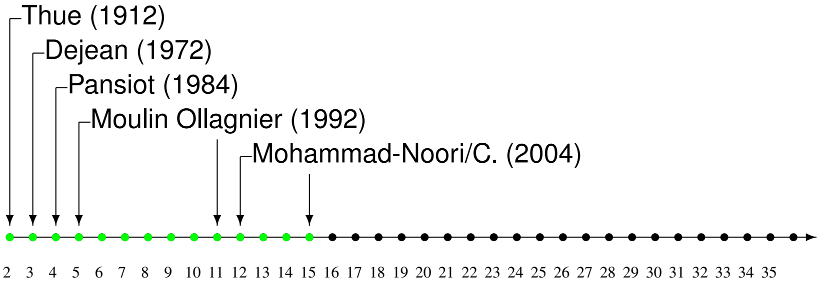
History of the conjecture



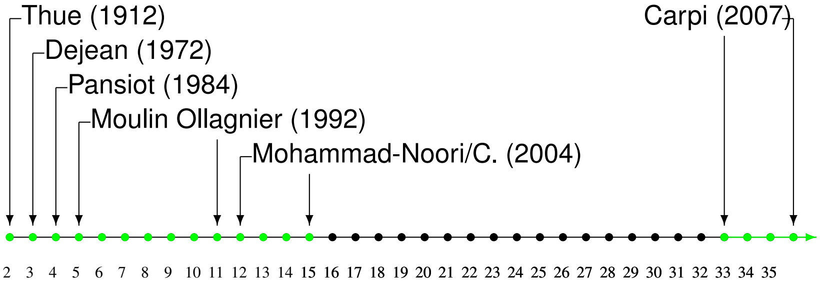
History of the conjecture



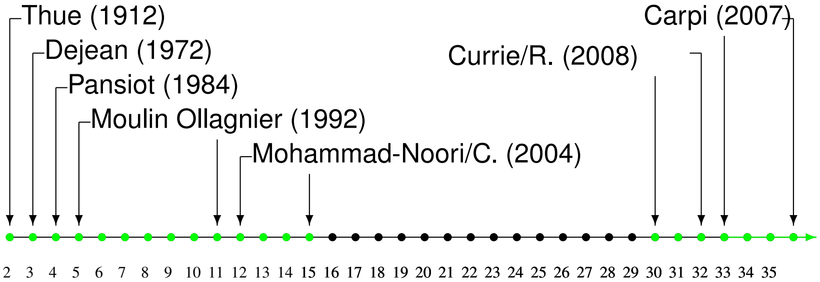
History of the conjecture



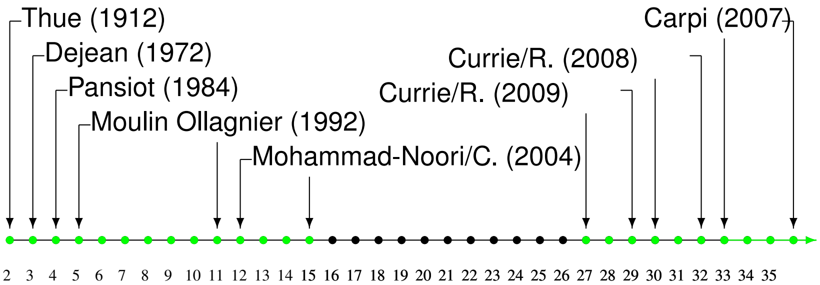
History of the conjecture



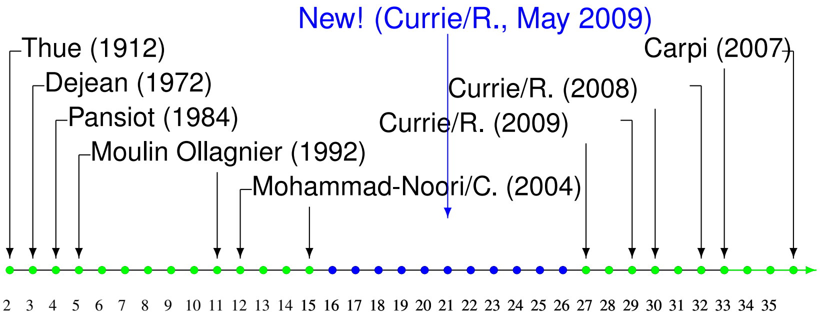
History of the conjecture



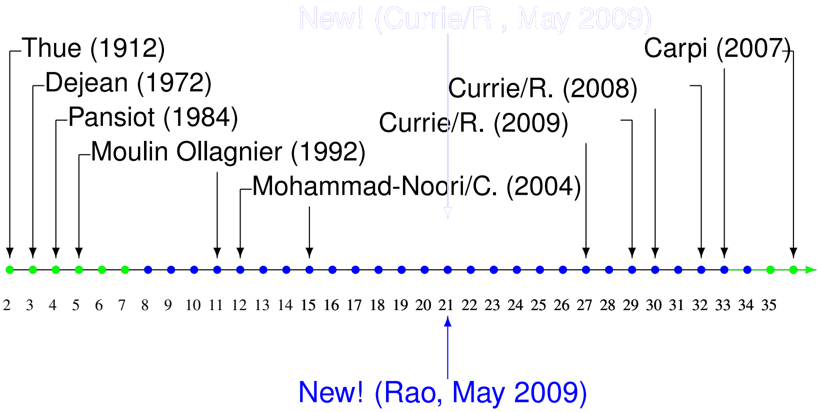
History of the conjecture



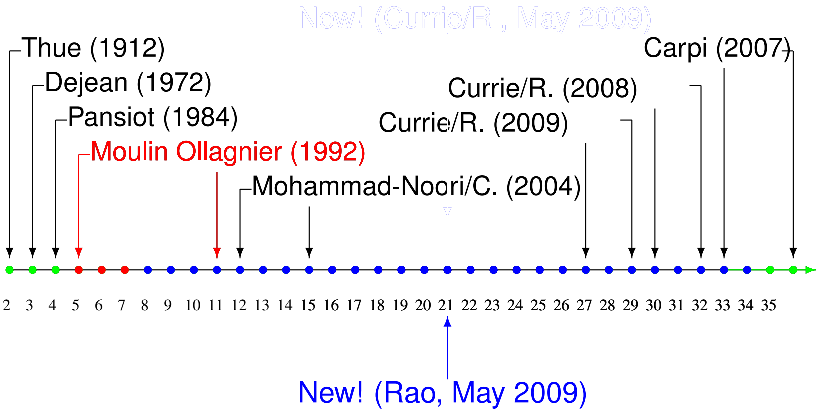
History of the conjecture



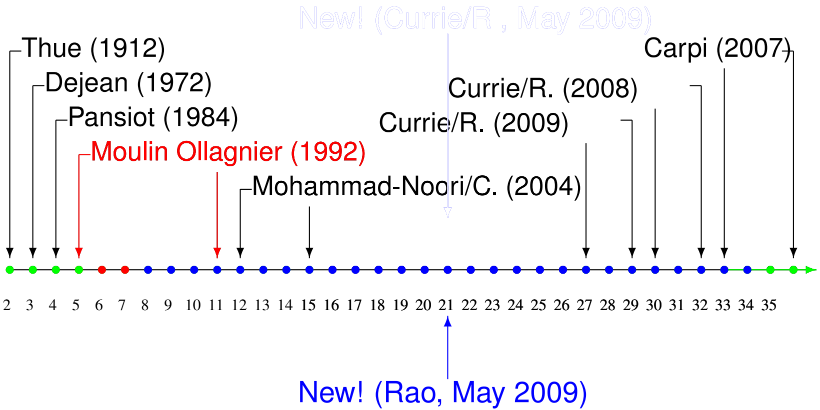
History of the conjecture



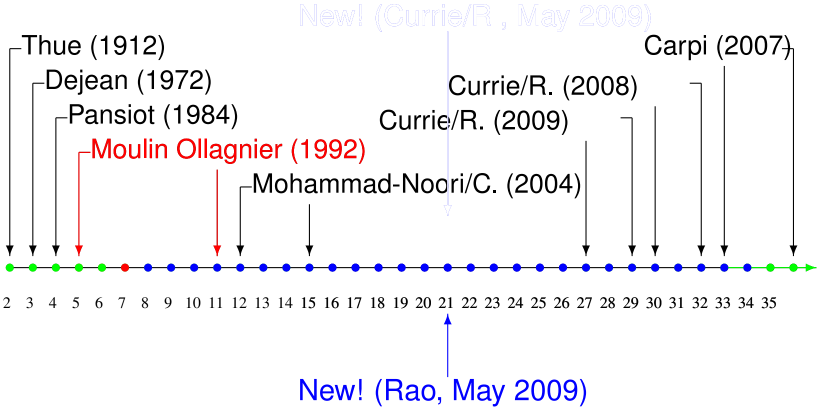
History of the conjecture



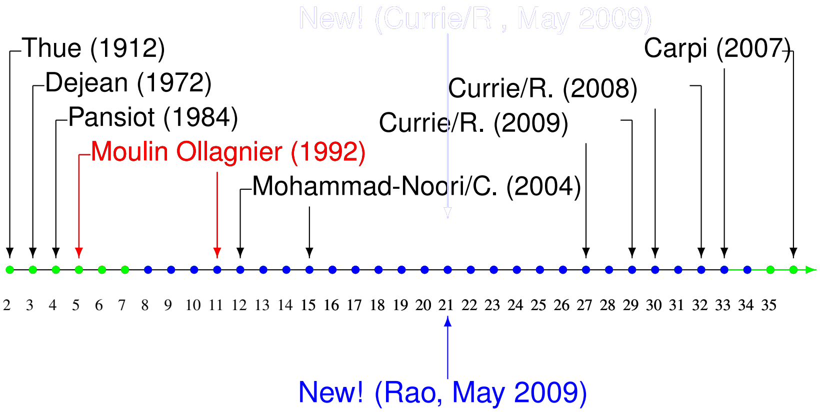
History of the conjecture



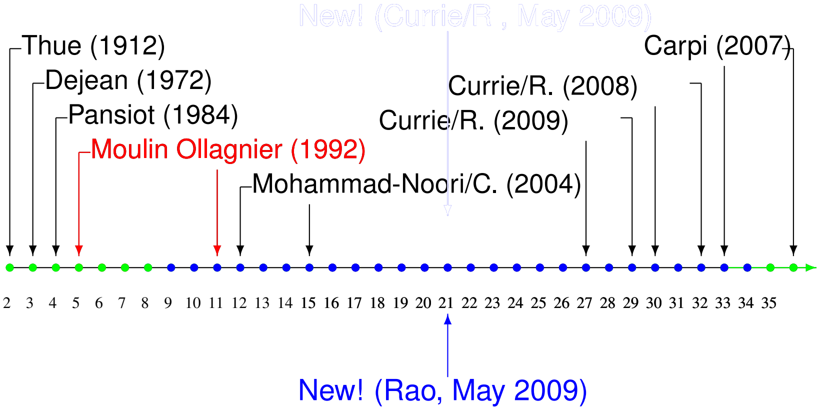
History of the conjecture



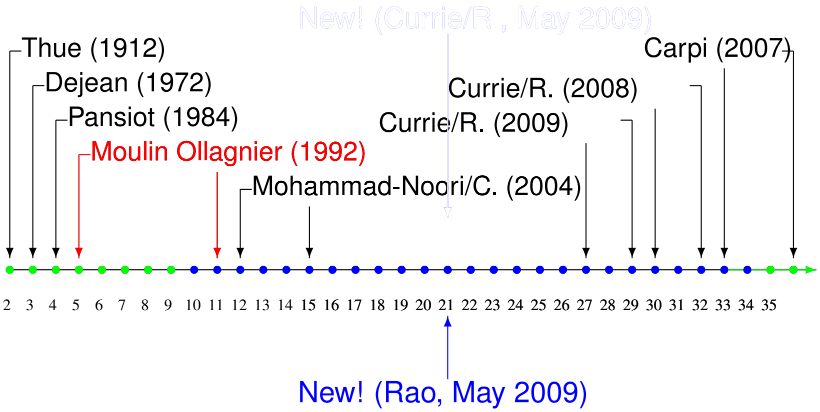
History of the conjecture



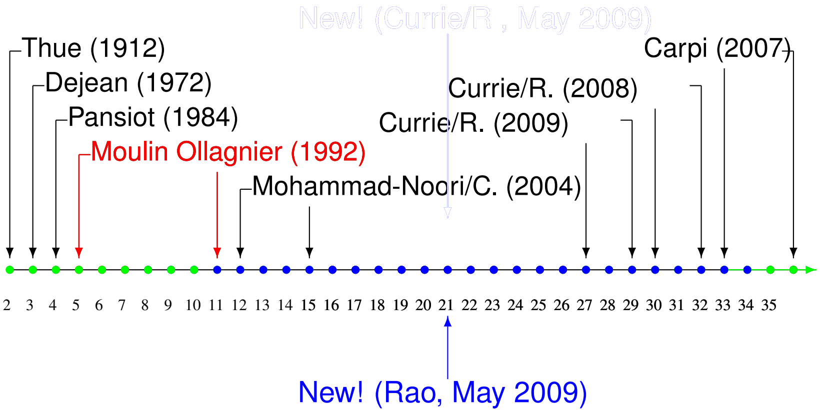
History of the conjecture



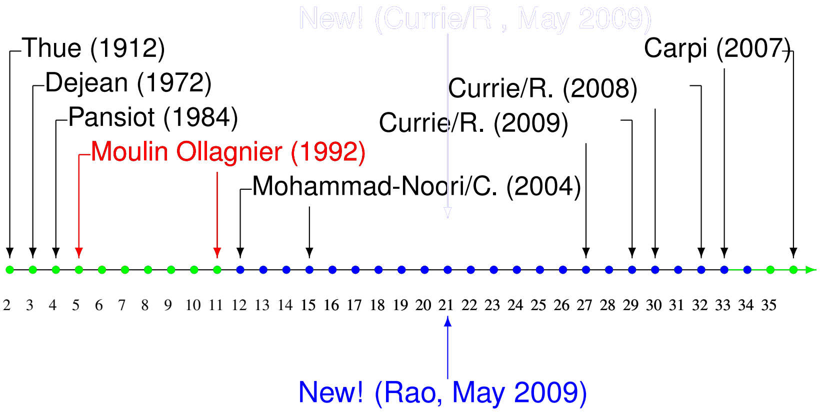
History of the conjecture



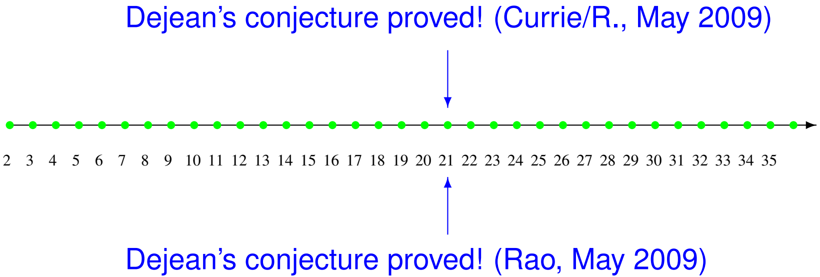
History of the conjecture



History of the conjecture



Dejean's conjecture proved



Pansiot's approach

- Alphabet size n
- A word of length at least $n+2$ must contain a factor with exponent at least $n/(n-1)$.
- If a word avoids $(n/(n-1))^+$-powers, every block of length $n-1$ consists of $n-1$ different letters.

The Pansiot encoding

- The letter following a block y of length $n - 1$ is either
  - the first letter of y; or
  - the unique letter that does not occur in y.
- Pansiot encoding: encode first case with a 0; second case with a 1.
- Can uniquely reconstruct the original word from the Pansiot encoding.

The Pansiot encoding

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101.$$

We reconstruct the original word from the prefix 12345 and the code 0101101.

The Pansiot encoding

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101.$$

We reconstruct the original word from the prefix 12345 and the code 0101101.

The Pansiot encoding

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101.$$

We reconstruct the original word from the prefix 12345 and the code 0101101.

The Pansiot encoding

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101.$$

We reconstruct the original word from the prefix 12345 and the code 0101101.

The Pansiot encoding

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101.$$

We reconstruct the original word from the prefix 12345 and the code 0101101.

The Pansiot encoding

Example (n=6)

Word:

$$1234\textcolor{red}{51632}\textcolor{blue}{4}15$$

Pansiot encoding:

$$0101\textcolor{blue}{1}01.$$

We reconstruct the original word from the prefix 12345 and the code 0101101.

The Pansiot encoding

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101.$$

We reconstruct the original word from the prefix 12345 and the code 0101101.

The Pansiot encoding

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101.$$

We reconstruct the original word from the prefix 12345 and the code 0101101.

The Pansiot encoding

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101.$$

We reconstruct the original word from the prefix 12345 and the code 0101101.

Constructing the Pansiot encoding

- Proving Dejean's conjecture for $n = 4$: need an infinite $(7/5)^+$-power-free word $\mathbf{w}$
- Instead, find the binary Pansiot encoding of $\mathbf{w}$
- Binary encoding: iterate $0 \to 101101$; $1 \to 10$:

$$1 \to 10 \to 10101101 \to 10101101101011011010110110 \to \cdots$$

- Decode:

$$\mathbf{w} = 123421432413423143 21 \cdots$$

A map into the symmetric group

- Moulin Ollagnier proved the conjecture for $5 \leq n \leq 11$.
- His observation: a word $w = a_1 a_2 \cdots a_{n-1}$ containing no repeated letter can be associated with a permutation:

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ a_1 & a_2 & a_3 & \cdots & a_{n-1} & b \end{pmatrix}$$

- b is the unique letter that does not occur in w.

A map into the symmetric group

- Moving from one $(n-1)$-letter block to the next $(n-1)$-letter block by a "0" in the Pansiot encoding corresponds to multiplication on the right by

$$\sigma_0 = \begin{pmatrix} 1 & 2 & 3 & \cdots & k-1 & k \\ 2 & 3 & 4 & \cdots & 1 & k \end{pmatrix}.$$

- Moving from one block to the next by a "1" corresponds to multiplication on the right by

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & \cdots & k-1 & k \\ 2 & 3 & 4 & \cdots & k & 1 \end{pmatrix}.$$

A map into the symmetric group

Example (n=6)

Word:

$$\color{red}{12345}\color{black}{1632415}$$

Pansiot encoding:

$$0101101$$

Permutation:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

A map into the symmetric group

Example (n=6)

Word:

$$1\textcolor{red}{23451}632415$$

Pansiot encoding:

$$0101101$$

Permutation:

$$\textcolor{red}{\sigma_0} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 1 & 6 \end{pmatrix}$$

A map into the symmetric group

Example (n=6)

Word:

$$12345163 2415$$

Pansiot encoding:

$$0101101$$

Permutation:

$$\sigma_0\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 1 & 6 & 2 \end{pmatrix}$$

A map into the symmetric group

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101$$

Permutation:

$$\sigma_0\sigma_1\sigma_0 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 1 & 6 & 3 & 2 \end{pmatrix}$$

A map into the symmetric group

Example (n=6)

Word:

$$1234\color{red}{51632}\color{black}{415}$$

Pansiot encoding:

$$0101101$$

Permutation:

$$\sigma_0\sigma_1\sigma_0\color{red}{\sigma_1}\color{black}{} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 3 & 2 & 4 \end{pmatrix}$$

A map into the symmetric group

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101$$

Permutation:

$$\sigma_0\sigma_1\sigma_0\sigma_1\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 6 & 3 & 2 & 4 & 5 \end{pmatrix}$$

A map into the symmetric group

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101$$

Permutation:

$$\sigma_0\sigma_1\sigma_0\sigma_1\sigma_1\sigma_0 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 3 & 2 & 4 & 1 & 5 \end{pmatrix}$$

A map into the symmetric group

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101$$

Permutation:

$$\sigma_0\sigma_1\sigma_0\sigma_1\sigma_1\sigma_0\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 4 & 1 & 5 & 6 \end{pmatrix}$$

A map into the symmetric group

Example (n=6)

Word:

$$123451632415$$

Pansiot encoding:

$$0101101$$

Permutation:

$$\sigma_0\sigma_1\sigma_0\sigma_1\sigma_1\sigma_0\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 4 & 1 & 5 & 6 \end{pmatrix}$$

A map into the symmetric group

- Define map ψ from the binary Pansiot codewords to the symmetric group S_n by

$$\begin{array}{ccc} 0 & \rightarrow & \sigma_0 \\ 1 & \rightarrow & \sigma_1, \end{array}$$

and if $y = y_0 y_1 \cdots y_\ell$ is a word over $\{0, 1\}$, then

$$y \rightarrow \sigma_{y_0} \sigma_{y_1} \cdots \sigma_{y_\ell}.$$

Kernel repetitions

- Alphabet size n
- w a word over an n-letter alphabet
- x the binary Pansiot encoding of w
- Write $x = pe$ with e also a prefix of x; p non-empty.
- Call p the period and e the excess.
- If $|e| \geq n - 1$ and $\psi(p)$ is the identity permutation, x is a kernel repetition.
- w then has exponent $(|pe| + n - 1)/|p|$.

Kernel repetitions

Example (n=4)

$$\begin{array}{rrcl} \text{Word:} & w & = & 12341341 23413 \\ \text{Pansiot encoding:} & x & = & \underbrace{1100011}_{p}\underbrace{110}_{e} \\ \text{Permutation:} & \psi(p) & = & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix} \end{array}$$

x is a kernel repetition; w has exponent

$$(|pe| + n - 1)/|p| = (10 + 4 - 1)/7 = 13/7.$$

Moulin Ollagnier's approach

- Generate an infinite Pansiot encoding $\mathbf{x}$ by iterating a binary morphism f.
- $\mathbf{x}$ encodes a word $\mathbf{w}$ over an n-letter alphabet.
- $\mathbf{x}$ must not contain a kernel repetition $x = pe$ with $(|pe| + n - 1)/|p| > RT(n)$.

The algebraic condition

- f maps $0 \to f(0)$; $1 \to f(1)$.
- algebraic condition for f: for some permutation τ,

$$\psi(f(0)) = \tau^{-1} \cdot \psi(0) \cdot \tau, \quad \psi(f(1)) = \tau^{-1} \cdot \psi(1) \cdot \tau.$$

- Ensures that f maps kernel repetitions to kernel repetitions
- Long kernel repetitions are the images under f of shorter kernel repetitions (more or less).

Checking the candidate word

- Check finitely many kernel repetitions in $\mathbf{x}$: verify none have $(|pe| + n - 1)/|p| > RT(n)$.
- Check that $\mathbf{w}$ does not contain other forbidden repetitions that do not arise from kernel repetitions in $\mathbf{x}$.
- These have length at most $(n-1)^2$—only finite many to check.

Searching by computer

- Moulin Ollagnier found by computer search binary morphisms to generate $\mathbf{x}$ for $5 \leq n \leq 11$.
- For $n = 5$:

$$
\begin{aligned}
0 &\rightarrow 0101011011010101110110 \\
1 &\rightarrow 1010101011011011101101.
\end{aligned}
$$

The final resolution of the conjecture

- Major breakthrough: Carpi's proof of the conjecture for $n \geq 33$
- We strengthened one part of Carpi's construction, improving this to $n \geq 27$.
- We resolved the remaining open cases by extending Moulin Ollagnier's computer calculations to find suitable morphisms.
- To verify our constructions we checked that they satisfy Moulin Ollagnier's criteria.

The final resolution of the conjecture

- Rao independently resolved the last open cases by a different method.
- He found morphisms, which, when applied to the Thue–Morse word, gave the desired Pansiot encoding.

Our calculations

- Search for candidate morphisms f
- Look for uniform morphisms ($f(0)$ and $f(1)$ have the same length)
- "Guess" $f(0)$ and $f(1)$ have length $4n - 4$ or $4n$

Finding the candidate morphisms

- Backtracking search: find candidates of length $4n-4$ for $f(0)$ and $f(1)$ (length $4n$ for $n = 21$)
- Generate all binary words of length $4n-4$ that are Pansiot encodings of a word avoiding $(n/(n-1))^+$-powers.
- Empirically: number of such words $\approx 1.24^{(4n-4)}$
- For large n (> 15) too many to fit in RAM

Checking the candidates

- Check all pairs of words as candidates for $f(0), f(1)$.
- Candidates should satisfy the algebraic condition.
- Check if candidate morphism avoids forbidden repetitions (finite check).

The computer calculations

- $n = 15$:

$$0 \rightarrow 01101101011011011011010101101010110110110110110101101101$$

$$1 \rightarrow 10101011011011010110101101101011010110110110101101010101$$

- $n = 26$: computation took approx. 6 hrs.

The End