Abelian powers in automatic sequences are not always automatic

Luke Schaeffer

University of Waterloo

June 11, 2013

Powers in Infinite Words

Definition

Let $d \geq 2$ be an integer. A *dth power* is a finite word $x \in \Sigma^*$ such that x is of the form

$$y^d = \overbrace{yy \cdots y}^{d \text{ times}}$$

for some word $y \in \Sigma^*$.

The English word $\texttt{tartar} = (\texttt{tar})^2$ is an example of a square.

Definition

Let $d \geq 2$ be an integer. An *abelian dth power* is a finite word $x \in \Sigma^*$ where $x = x_1 \cdots x_d$ where each x_i is a permutation of any other x_j.

The word `reappear` is an example of an abelian square.

Notation

Definition

Suppose $n \geq 0$ and $k \geq 2$ are integers. Let $[n]_k \in \{0, 1, \ldots, k-1\}^*$ denote the base-k representation of the integer n.

Definition

A subset X of $\mathbb{N}^m$ is *k-automatic* if there exists an automaton T such that $(i_1, \ldots, i_m) \in X$ if and only if T accepts words $[i_1]_k, \ldots, [i_m]_k$ in parallel.

Theorem

A set $X \subseteq \mathbb{N}^m$ is k-automatic if and only if the predicate $P(i_1, \ldots, i_m) := (i_1, \ldots, i_m) \in X$ is in the theory $\langle \mathbb{N}, 0, 1, <, +, V_k \rangle$.

Powers with Predicates

Theorem

Fix an integer $d \geq 2$. Suppose w is a k-automatic sequence. The set

$$\{(i, p) : w[i..i + pd - 1] \text{ is a } d\text{th power}\} \subseteq \mathbb{N}^2$$

is k-automatic.

Proof.

The word $w[i..i + pd - 1]$ is a dth power if and only if each consecutive pair of blocks, $w[i + (j - 1)p..i + (j + 1)d - 1]$, is a square. Hence, it suffices to prove the case $d = 2$.

The subword $w[i..i + 2p - 1]$ is a square if the two halves, $w[i..i + p - 1]$ and $w[i + p..i + 2p - 1]$, are equal. □

Convention for Occurrences of Subwords

- Position/Period

$$w[i..i + pd - 1] \to (i, p)$$

- Position/Length

$$w[i..i + n - 1] \to (i, n)$$

- Endpoints

$$w[i..j] \to (i, j)$$

- Endpoints and midpoint

$$w[i..j] \to (i, (i + j)/2, j)$$

Theorem

If a set of occurrences is k-automatic in any of the above conventions, then it is k-automatic for every convention.

Outline

Question

We can express the query "is the subword $w[i..j]$ a dth power?" as a predicate, and the set of all such subwords is k-automatic.

Q: Can we do the same for abelian powers?
A: No, the set of occurrences of abelian squares in a k-automatic word is not always k-automatic.

- Three examples where we *can* describe abelian powers:
  - When the set of powers is trivial
  - When we have a letter frequency automaton
  - An ad hoc approach
- One example where we *cannot* describe abelian powers.

Trivial Set of Abelian Powers

Consider the word

$$\mathbf{q} = 0112122312232330\cdots$$

where $\mathbf{q}[i]$ is the number of ones in $[i]_2$ modulo 4.

Theorem

The infinite word $\mathbf{q} \in \{0, 1, 2, 3\}^\omega$ contains no abelian cubes.

Corollary

The set of occurrences of abelian cubes in $\mathbf{q}$,

$$\{(i, n) \in \mathbb{N} : \mathbf{q}[i..i+n-1] \textit{ is an abelian cube}\},$$

is empty, and therefore k-automatic for all $k \geq 2$.

Abelian Squares in Thue-Morse

Recall the Thue-Morse sequence,

$$\mathbf{t} = 0110100110010110100101100110 1001 \cdots$$

where $\mathbf{t}[i]$ is the number of ones in $[i]_2$ modulo 2.

It is not hard to see that $\mathbf{t}$ is composed of 01 blocks and 10 blocks.

- Even length prefix $\implies$ equal number of zeros and ones
- Odd length prefix $\implies$ extra zero or one, depending on last symbol

Theorem

Let $|x|_y$ denote the number of occurrences of $y \in \Sigma^*$ as a subword of $x \in \Sigma^*$. There is an automaton M that shows the set

$$\{(n, |\mathbf{t}[0..n-1]|_0) : n \in \mathbb{N}\},$$

is 2-automatic.

We can use M to build an automaton for abelian squares.

Abelian Squares in Thue-Morse

Theorem

The following are equivalent:

- $\mathbf{t}[i..i+2n-1]$ *is an abelian square.*
- $\mathbf{t}[i..i+n-1]$ *and* $\mathbf{t}[i+n..i+2n-1]$ *contain the same number of zeros.*
- $|\mathbf{t}[0..i-1]|_0$, $|\mathbf{t}[0..i+n-1]|_0$, $|\mathbf{t}[0..i+2n-1]|_0$ *is an arithmetic progression.*
- $|\mathbf{t}[0..i-1]|_0 + |\mathbf{t}[0..i+2n-1]|_0 = 2\,|\mathbf{t}[0..i+n-1]|_0$.

We can decide whether $\mathbf{t}[i..i+2n-1]$ is an abelian square using M and the following predicate.

$$(\exists a, b, c\ \ M(i,a) \wedge M(i+n,b) \wedge M(i+2n,c) \wedge (a+c=2b))$$

Characteristic Word of $\{2^k - 1\}_{k=0}^{\infty}$

Define an infinite word $\mathbf{c} \in \{0,1\}^\omega$,

$$\mathbf{c} = 11010001000000010000\cdots$$

where $\mathbf{c}[i] = 1$ if and only if i is of the form $2^k - 1$.
The set

$$\{(n, |\mathbf{c}[0..n-1]|_1) : n \in \mathbb{N}\}$$

cannot be k-automatic because of the following theorem:

Theorem

Let $f : \mathbb{N} \to \mathbb{N}$ be a non-decreasing function. If $\{(n, f(n)) : n \in \mathbb{N}\}$ is k-automatic, then $f(n)$ is $\Theta(1)$ or $\Theta(n)$.

Abelian squares in $\mathbf{c}$

Suppose $x = \mathbf{c}[i..i+2n-1]$ is an abelian square.

- There is a single 1 in the range $\mathbf{c}[j..2j]$ for any j.
- The factor $\mathbf{c}[i+n..i+2n-1]$ contains at most one 1, and therefore so does $\mathbf{c}[i..i+n-1]$.
- If $\mathbf{c}[i+n..i+2n-1]$ does not contain a 1, then $x = 0^{2n}$. There is an automaton to check if $\mathbf{c}[i..i+2n-1] = 0^{2n}$ using predicates.

  $$(\forall j\ (j < i) \vee (i+n-1 < j) \vee (\mathbf{c}[j] = 0))$$

- Otherwise, each half of x, $\mathbf{c}[i..i+n-1]$ and $\mathbf{c}[i+n..i+2n-1]$, contains exactly one 1. This can also be expressed with predicates.

  $$\begin{aligned}(\exists k\ (\mathbf{c}[k] = 1)\wedge\\ &(i \leq k) \wedge (k \leq i+n-1)\wedge\\ &(\forall j\ (j < i) \vee (i+n-1 < j)\vee\\ &\quad(\mathbf{c}[j] = 0) \vee (j = k)))\end{aligned}$$

- The set of occurrences of abelian squares in $\mathbf{c}$ is 2-automatic.

Paperfolding Word

The paperfolding word,

$$\mathbf{p} := 110110011100100\cdots$$

is the result of iterating the map $f(w) = w1\overline{w}^R$, where $\overline{w}^R$ denotes the *reverse complement* of w.

$$f(\varepsilon) = 1$$
$$f(1) = 110$$
$$f(110) = 1101100$$
$$f(1101100) = 110110011100100$$
$$\vdots$$

Goal

Show that the occurrences of abelian squares in $\mathbf{p}$ are not 2-automatic.

Symbol Freq. in Paperfolding Word

Definition

Define a function $\Delta : \mathbb{N} \to \mathbb{Z}$ where $\Delta(i) = |\mathbf{p}[1..i]|_1 - |\mathbf{p}[1..i]|_0$.

The value $\Delta(i)$ turns out to be more convenient than $|\mathbf{p}[1..i]|_0$ or $|\mathbf{p}[1..i]|_1$, but we can still use it to check for abelian squares.

Theorem

A subword $\mathbf{p}[i+1..i+2p]$ is an abelian square if and only if $\Delta(i) + \Delta(i+2p) = 2\Delta(i+p)$.

Useful result about Δ

Theorem

Let i be a natural number. Suppose u is a binary representation for i with at least one leading zero. Then

$$\Delta(i) = |u|_{01} + |u|_{10}\,.$$

We cannot construct an automaton for the set

$$\{(n, \Delta(n)) : n \in \mathbb{N}\}$$

because when $[n]_2 = (10)^k$ the theorem gives $\Delta(n) = 2k = \Theta(\log n)$.

Main Result

Theorem

The set of occurrences of abelian squares in $\mathbf{p}$ *is not 2-automatic.*

Proof.

Suppose for a contradiction that the set

$$\{(i, j, k) : \mathbf{p}[i..k] \text{ is an abelian square and } i + k = 2j\}.$$

is 2-automatic. Let M be the automaton accepting the set in binary.

Proof Cont'd

Proof.

Consider (i, j, k) such that

$$[i]_2 = (0000)^m(0010)^n$$
$$[j]_2 = (0100)^m(0110)^n$$
$$[k]_2 = (1000)^m(1010)^n$$

for $m, n \in \mathbb{N}$. Clearly $i + k = 2j$ for all $m, n \geq 0$ and

$$\Delta(i) = 2n$$
$$\Delta(j) = 2m + 2n$$
$$\Delta(k) = 2m + 4n.$$

We can show that (i, j, k) corresponds to an abelian square if and only if

$$2m + 6n = \Delta(i) + \Delta(k) = 2\Delta(j) = 4m + 4n$$

Proof Cont'd

Proof.

Therefore, the automaton should accept (i, j, k) if and only if $m = n$. The automaton has finitely many states, so there exist m, m' such that M is in the same state after reading

$$\begin{pmatrix} 0000 \\ 0100 \\ 1000 \end{pmatrix}^m \text{ or } \begin{pmatrix} 0000 \\ 0100 \\ 1000 \end{pmatrix}^{m'}.$$

From that state, the automaton accepts on input

$$\begin{pmatrix} 0010 \\ 0110 \\ 1010 \end{pmatrix}^n$$

if and only if $n = m$, but also if and only if $n = m'$. Contradiction. □

Logical Expressibility

Recall the following theorem.

Theorem

A set $X \subseteq \mathbb{N}^m$ is k-automatic if and only if the predicate $P(i_1, \ldots, i_m) := (i_1, \ldots, i_m) \in X$ is in the theory $\langle \mathbb{N}, 0, 1, <, +, V_k \rangle$.

Since the set of occurrences of abelian squares in **p** is not 2-automatic, we cannot express it in the theory $\langle \mathbb{N}, 0, 1, <, +, V_2 \rangle$.

Cobham's Theorem

Theorem

Let k, ℓ be multiplicatively dependent integers. Suppose $w \in \Sigma^\omega$ is a k-automatic sequence. Then w is ℓ-automatic.

Theorem (Cobham)

Let $w \in \Sigma^\omega$ be an aperiodic sequence. If w is both k-automatic and ℓ-automatic, then k and ℓ are multiplicatively dependent.

Corollary

The set of occurrences of abelian squares in $\mathbf{p}$ *is not k-automatic for any $k \geq 2$.*

Proof of the Corollary

Proof.

Suppose the occurrences of abelian squares in $\mathbf{p}$ is k-automatic for some $k \geq 2$.

- On the one hand, if k is multiplicatively dependent with 2 then the set of occurrences is 2-automatic, contradicting our main result. Therefore k and ℓ are not multiplicatively dependent.
- On the other hand, the set of occurrences of abelian squares of length 2 in $\mathbf{p}$ is k-automatic.
  - There is an abelian square of length 2 starting at position i if and only if $\mathbf{p}[i] = \mathbf{p}[i+1]$, or equivalently, if $\mathbf{p}[i+1] - \mathbf{p}[i] \equiv 0 \pmod 2$.
  - We can recover $\mathbf{p}$ from the first differences, $\mathbf{p}[i+1] - \mathbf{p}[i] \equiv 0 \pmod 2$, and it is k-automatic.
  - Apply Cobham's theorem to $\mathbf{p}$, therefore k and 2 are multiplicatively dependent.

□

The End

Thank You!