State complexity of the multiples of the Thue-Morse set

Célia Cisternino
Joint work with Émilie Charlier and Adeline Massuir

July 8, 2019

Numeration and Substitution,
Vienna (Austria)
Definition

A DFA is *minimal* iff it is *reduced* and *accessible*.

- Trim minimal

Theorem

For any regular language $L$, there exists a unique (up to isomorphism) minimal automaton accepting $L$.

Definition

The *state complexity* of a regular language is equal to the number of states of its minimal automaton.
Definition
A DFA has **disjoint states** if, for distinct states \( p \) and \( q \), we have \( L(p) \cap L(q) = \emptyset \).

Proposition
Any coaccessible DFA having disjoint states is reduced.
**Definition**

For a base $b$, a subset $X$ of $\mathbb{N}$ is said to be $b$-recognizable if the language $0^*\text{rep}_b(X)$ is regular.

**Proposition**

Let $b \in \mathbb{N}_{\geq 2}$ and $m \in \mathbb{N}$. If $X$ is $b$-recognizable, then so is $mX$. 
**Multiplicatively independent integers:**

\[(p^a = q^b) \Rightarrow (a = b = 0)\]

**Theorem (Cobham, 1969)**

- Let \(b, b'\) be two multiplicatively independent bases. Then a subset of \(\mathbb{N}\) is both \(b\)-recognizable and \(b'\)-reconnaissable if and only if it is a finite union of arithmetic progressions.

- Let \(b, b'\) be two multiplicatively dependent bases. Then a subset of \(\mathbb{N}\) is \(b\)-recognizable if and only if it is \(b'\)-recognizable.
The **Thue-Morse set**:

\[ T = \{ n \in \mathbb{N}: |\text{rep}_2(n)|_1 \in 2\mathbb{N} \} \]

Characteristic sequence: 1001011001101001 \cdots
The set $T$ is $2^p$-recognizable for all $p \in \mathbb{N}_{\geq 1}$ and is not $b$-recognizable for any other base $b$.

$A_4 \cap T = \{0, 3\}$

$A_4 \cap (\mathbb{N} \setminus T) = \{1, 2\}$

$X_a = \begin{cases} X & \text{if } a \in T \\ \overline{X} & \text{otherwise} \end{cases}$
Main Theorem

Lemma

For any $m \in \mathbb{N}$ and $p \in \mathbb{N}_{\geq 1}$, the set $mT$ is $2^p$-recognizable.

Theorem

Let $m \in \mathbb{N}$ and $p \in \mathbb{N}_{\geq 1}$. Then the state complexity of the language $0^*\text{rep}_{2^p}(mT)$ is equal to

$$2k + \left\lceil \frac{z}{p} \right\rceil$$

if $m = k2^z$ with $k$ odd.
## Method

<table>
<thead>
<tr>
<th>Automaton</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{A}_T, 2^p$</td>
<td>$(0, 0)^* { \text{rep}_{2^p}(t, n) : t \in T, n \in \mathbb{N} }$</td>
</tr>
<tr>
<td>$\mathcal{A}_m, 2^p$</td>
<td>$(0, 0)^* { \text{rep}_{2^p}(n, mn) : n \in \mathbb{N} }$</td>
</tr>
<tr>
<td>$\mathcal{A}_m, 2^p \times \mathcal{A}_T, 2^p$</td>
<td>$(0, 0)^* { \text{rep}_{2^p}(t, mt) : t \in T }$</td>
</tr>
<tr>
<td>$\Pi(\mathcal{A}_m, 2^p \times \mathcal{A}_T, 2^p)$</td>
<td>$0^* { \text{rep}_{2^p}(mt) : t \in T }$</td>
</tr>
</tbody>
</table>
The automaton $A_{T,2^p}$

Edges:

$$(0, 0), (0, 1), (0, 2), (0, 3)$$
$$(3, 0), (3, 1), (3, 2), (3, 3)$$

$$(1, 0), (1, 1), (1, 2), (1, 3)$$
$$(3, 0), (3, 1), (3, 2), (3, 3)$$

$$(2, 0), (2, 1), (2, 2), (2, 3)$$
$$(3, 0), (3, 1), (3, 2), (3, 3)$$

$$(1, 0), (1, 1), (1, 2), (1, 3)$$
$$(2, 0), (2, 1), (2, 2), (2, 3)$$

$$(d, e)$$

$Y = X_d$
Proposition

The automaton $A_{\mathcal{T},2^p}$

- accepts $(0, 0)^* \{\text{rep}_{2^p}(t, n) : t \in \mathcal{T}, n \in \mathbb{N}\}$
- is accessible
- is coaccessible
- has disjoint states
- is trim minimal
- is complete
The automaton $A_{m,b}$

Edges:
$$(d, e)$$

$bi + e = md + j$
Proposition

The automaton $A_{m,b}$
- accepts $(0, 0)^* \{ \text{rep}_b(n, mn): n \in \mathbb{N} \}$
- is accessible
- is coaccessible
- has disjoint states
- is trim minimal

Remark: The automaton $A_{m,b}$ is not complete.
The product automaton $A_{m,2p} \times A_{\mathcal{T},2p}$
**Proposition**

The automaton $A_{m,2^p} \times A_{\mathcal{T},2^p}$

- accepts $(0, 0)^* \{\text{rep}_{2^p}(t, mt) : t \in \mathcal{T}\}$
- is accessible
- is coaccessible
- has disjoint states
- is trim minimal

**Remark:** The automaton $A_{m,2^p} \times A_{\mathcal{T},2^p}$ is not complete.
Projection of $\mathcal{A}_{m,2^p} \times \mathcal{A}_{T,2^p}$
Proposition

The automaton $\Pi(A_{m,2^p} \times A_{\mathcal{T},2^p})$

- accepts $0^* \{\text{rep}_{2^p}(mt) : t \in \mathcal{T}\}$
- is deterministic
- is accessible
- is coaccessible
- has disjoint states if $m$ is odd
- is trim minimal if $m$ is odd

Corollary

The state complexity of $m\mathcal{T}$ in base $2^p$ is $2m$ if $m$ is odd.
Minimisation of $\prod(\mathcal{A}_{m,2^p} \times \mathcal{A}_{T,2^p})$
Let \( m = k \cdot 2^z \) where \( k, z \in \mathbb{N} \) with \( k \) odd.

For \( (j, X) \in (\{1, \ldots, k - 1\} \times \{T, B\}) \cup \{(0, B)\} \), the class of \( (j, X) \) is

\[
[(j, X)] = \{(j + k\ell, X_\ell) : 0 \leq \ell \leq 2^z - 1\}.
\]

Moreover, the class of \( (0, T) \) is

\[
[(0, T)] = \{(0, T)\}.
\]
For $\alpha \in \{0, \ldots, z - 1\}$, we define a pre-classe $C_\alpha$ of size $2^\alpha$:

$$C_\alpha = \left\{ \left( \frac{m}{2^{\alpha+1}} + \frac{m}{2^\alpha} \ell, B_\ell \right) : 0 \leq \ell \leq 2^\alpha - 1 \right\}$$

For all $\beta \in \{0, \ldots, \left\lceil \frac{z}{p} \right\rceil - 2\}$, we define

$$\Gamma_\beta = \bigcup_{\alpha \in \{\beta p, \ldots, (\beta + 1)p - 1\}} C_\alpha$$

and

$$\Gamma_{\left\lceil \frac{z}{p} \right\rceil - 1} = \bigcup_{\alpha \in \{(\left\lceil \frac{z}{p} \right\rceil - 1)p, \ldots, z - 1\}} C_\alpha$$
In this example $m = 3 \cdot 2^3$ and $b = 4$. So, $k = 3$, $z = 3$, $p = 2$, and $\left\lceil \frac{z}{p} \right\rceil = 2$. We obtain

\[
C_0 = \{(12, B)\} \\
C_1 = \{(6, B), (18, T)\} \\
C_2 = \{(3, B), (9, T), (15, T), (21, B)\}
\]

and

\[
\Gamma_1 = C_0 \cup C_1 = \{(6, B), (12, B), (18, T)\} \\
\Gamma_2 = C_2 = \{(3, B), (9, T), (15, T), (21, B)\}
\]
# Counting and Conclusion

<table>
<thead>
<tr>
<th>Classes</th>
<th>Number of such classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[(j, X)]$ for $(j, X) \in ({1, \ldots, k - 1} \times {H, B})$</td>
<td>$2(k - 1)$</td>
</tr>
<tr>
<td>$[(0, B)]$</td>
<td>$1$</td>
</tr>
<tr>
<td>$[(0, H)]$</td>
<td>$1$</td>
</tr>
<tr>
<td>$\Gamma_\beta$ for $\beta \in {0, \ldots, \left\lfloor \frac{z}{p} \right\rfloor - 2}$</td>
<td>$\left\lfloor \frac{z}{p} \right\rfloor - 1$</td>
</tr>
<tr>
<td>$\Gamma_{\left\lfloor \frac{z}{p} \right\rfloor - 1}$</td>
<td>$1$</td>
</tr>
</tbody>
</table>

**Total** = $2k + \left\lfloor \frac{z}{p} \right\rfloor$
Theorem

Let $m \in \mathbb{N}$ and $p \in \mathbb{N}_{\geq 1}$. Then the state complexity of the language $0^*\text{rep}_{2^p}(mT)$ is equal to

$$2k + \left\lfloor \frac{z}{p} \right\rfloor$$

if $m = k2^z$ with $k$ odd.
The state complexity of the Thue-Morse set \( T \) in base 4 is equal to 
\[
\frac{3}{2} + \left\lceil \frac{1}{2} \right\rceil.
\]
The state complexity of $6T$ in base 4 is equal to

$$2.3 + \left\lceil \frac{1}{2} \right\rceil.$$
Thank you!