State complexity of the multiples of the Thue-Morse set

Adeline Massuir
Joint work with Émilie Charlier and Célia Cisternino

17^{th} Mons Theorical Computer Science Days – Bordeaux

September 13^{th} 2018

Basics

- Alphabet, letter, word, language, automaton, regular language

Basics

- Alphabet, letter, word, language, automaton, regular language
- $|w|$, $|w|_a$, ε

Basics

- Alphabet, letter, word, language, automaton, regular language
- $|w|$, $|w|_a$, ε
- State complexity

Basics

- Alphabet, letter, word, language, automaton, regular language
- $|w|$, $|w|_a$, ε
- State complexity
- Reduced, accessible, minimal automaton

Basics

- Alphabet, letter, word, language, automaton, regular language
- $|w|$, $|w|_a$, ε
- State complexity
- Reduced, accessible, minimal automaton
- Trim minimal automaton

Basics

- Alphabet, letter, word, language, automaton, regular language
- $|w|$, $|w|_a$, ε
- State complexity
- Reduced, accessible, minimal automaton
- Trim minimal automaton
- Disjoint states : $L(q) \cap L(q') = \emptyset$

Basics

- Alphabet, letter, word, language, automaton, regular language
- $|w|$, $|w|_a$, ε
- State complexity
- Reduced, accessible, minimal automaton
- Trim minimal automaton
- Disjoint states : $L(q) \cap L(q') = \emptyset$
- Coaccessible state, automaton

Let $b \in \mathbb{N}_{\geq 2}$

- Let $n = \sum_{j=0}^{\ell-1} c_j b^j \in \mathbb{N}$ with $c_j \in A_b := \{0, \ldots, b-1\}$ and $c_{\ell-1} \neq 0$

Let $b \in \mathbb{N}_{\geq 2}$

- Let $n = \sum_{j=0}^{\ell-1} c_j b^j \in \mathbb{N}$ with $c_j \in A_b := \{0, \dots, b-1\}$ and $c_{\ell-1} \neq 0$
  - $\text{rep}_b(n) = c_{\ell-1} \dots c_0$, $\text{rep}_b(0) = \varepsilon$

Let $b \in \mathbb{N}_{\geq 2}$

- Let $n = \sum_{j=0}^{\ell-1} c_j b^j \in \mathbb{N}$ with $c_j \in A_b := \{0, \ldots, b-1\}$ and $c_{\ell-1} \neq 0$
  - $\text{rep}_b(n) = c_{\ell-1} \ldots c_0$, $\text{rep}_b(0) = \varepsilon$
  - $\text{val}_b(c_{\ell-1} \ldots c_0) = n$, $\text{val}_b(\varepsilon) = 0$

Let $b \in \mathbb{N}_{\geq 2}$

- Let $n = \sum_{j=0}^{\ell-1} c_j b^j \in \mathbb{N}$ with $c_j \in A_b := \{0, \ldots, b-1\}$ and $c_{\ell-1} \neq 0$
  - $\text{rep}_b(n) = c_{\ell-1} \ldots c_0$, $\text{rep}_b(0) = \varepsilon$
  - $\text{val}_b(c_{\ell-1} \ldots c_0) = n$, $\text{val}_b(\varepsilon) = 0$
- b-recognizable set

Let $b \in \mathbb{N}_{\geq 2}$

- Let $n = \sum_{j=0}^{\ell-1} c_j b^j \in \mathbb{N}$ with $c_j \in A_b := \{0, \ldots, b-1\}$ and $c_{\ell-1} \neq 0$
  - $\text{rep}_b(n) = c_{\ell-1} \ldots c_0$, $\text{rep}_b(0) = \varepsilon$
  - $\text{val}_b(c_{\ell-1} \ldots c_0) = n$, $\text{val}_b(\varepsilon) = 0$
- b-recognizable set
- $u = u_1 \ldots u_n \in A^*, v = v_1 \ldots v_n \in B^*$,

$$(u, v) = (u_1, v_1) \ldots (u_n, v_n) \in (A \times B)^*$$

Let $b \in \mathbb{N}_{\geq 2}$

- Let $n = \sum_{j=0}^{\ell-1} c_j b^j \in \mathbb{N}$ with $c_j \in A_b := \{0, \ldots, b-1\}$ and $c_{\ell-1} \neq 0$
  - $\text{rep}_b(n) = c_{\ell-1} \ldots c_0$, $\text{rep}_b(0) = \varepsilon$
  - $\text{val}_b(c_{\ell-1} \ldots c_0) = n$, $\text{val}_b(\varepsilon) = 0$

- b-recognizable set

- $u = u_1 \ldots u_n \in A^*, v = v_1 \ldots v_n \in B^*$,

$$(u, v) = (u_1, v_1) \ldots (u_n, v_n) \in (A \times B)^*$$

- $\text{rep}_b(n_1, n_2) = \left(0^{\ell - |\text{rep}_b(n_1)|}\,\text{rep}_b(n_1), 0^{\ell - |\text{rep}_b(n_2)|}\,\text{rep}_b(n_2)\right)$
  where $\ell = \max\{|\text{rep}_b(n_1)|, |\text{rep}_b(n_2)|\}$

Goal and method

Multiplicatively independent integers ($p^a = q^b \Rightarrow a = b = 0$)

Goal and method

Multiplicatively independent integers ($p^a = q^b \Rightarrow a = b = 0$)

Theorem [Cobham, 1969]

- Let b, b' two multiplicatively independent bases.
  A subset of $\mathbb{N}$ is both b-recognizable and b'-recognizable iff it is a finite union of arithmetic progressions.
- Let b, b' two multiplicatively dependent bases.
  A subset of $\mathbb{N}$ is b-recognizable iff it is b'-recognizable.

Goal and method

Multiplicatively independent integers ($p^a = q^b \Rightarrow a = b = 0$)

Theorem [Cobham, 1969]

- Let b, b' two multiplicatively independent bases. A subset of $\mathbb{N}$ is both b-recognizable and b'-recognizable iff it is a finite union of arithmetic progressions.
- Let b, b' two multiplicatively dependent bases. A subset of $\mathbb{N}$ is b-recognizable iff it is b'-recognizable.

Proposition

Let $b \in \mathbb{N}_{\geq 2}$ and $m \in \mathbb{N}$. If $X \subseteq \mathbb{N}$ is b-recognizable, so is mX.

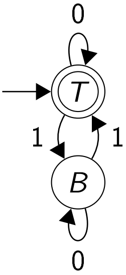
Theorem [Alexeev, 2004]

The state complexity of the base-b representations of $m\mathbb{N}$ is

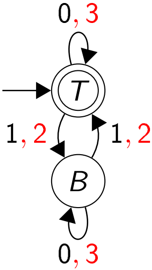
$$\min_{N\geq 0}\left\{\frac{m}{\gcd\left(m,b^N\right)}+\sum_{n=0}^{N-1}\frac{b^n}{\gcd\left(b^n,m\right)}\right\}$$

$$\mathscr{T} = \{n \in \mathbb{N} : |\mathrm{rep}_2(n)|_1 \in 2\,\mathbb{N}\}$$

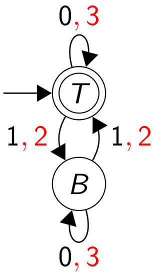
$$\mathscr{T} = \{n \in \mathbb{N} : |\mathsf{rep}_2(n)|_1 \in 2\,\mathbb{N}\}$$



$$\mathscr{T} = \{n \in \mathbb{N} : |\operatorname{rep}_2(n)|_1 \in 2\,\mathbb{N}\}$$



$$\mathscr{T} = \{n \in \mathbb{N} : |\operatorname{rep}_2(n)|_1 \in 2\mathbb{N}\}$$



Theorem

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

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

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

$$\mathscr{T} = \{n \in \mathbb{N} : |\operatorname{rep}_2(n)|_1 \in 2\,\mathbb{N}\}$$

Automaton	Set recognized
$\mathscr{A}_{\mathscr{T},2^p}$	$\mathscr{T} \times \mathbb{N}$

$$\mathscr{T} = \{n \in \mathbb{N} : |\operatorname{rep}_2(n)|_1 \in 2\,\mathbb{N}\}$$

Automaton	Set recognized
$\mathscr{A}_{\mathscr{T},2^p}$	$\mathscr{T} \times \mathbb{N}$
$\mathscr{A}_{m,2^p}$	$\{(n, mn) : n \in \mathbb{N}\}$

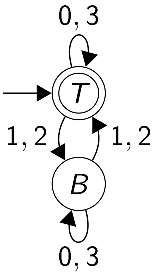
$$\mathscr{T} = \{n \in \mathbb{N} : |\operatorname{rep}_2(n)|_1 \in 2\,\mathbb{N}\}$$

Automaton	Set recognized
$\mathscr{A}_{\mathscr{T},2^p}$	$\mathscr{T} \times \mathbb{N}$
$\mathscr{A}_{m,2^p}$	$\{(n, mn) : n \in \mathbb{N}\}$
$\mathscr{A}_{\mathscr{T},2^p} \times \mathscr{A}_{m,2^p}$	$\{(t, mt) : t \in \mathscr{T}\}$

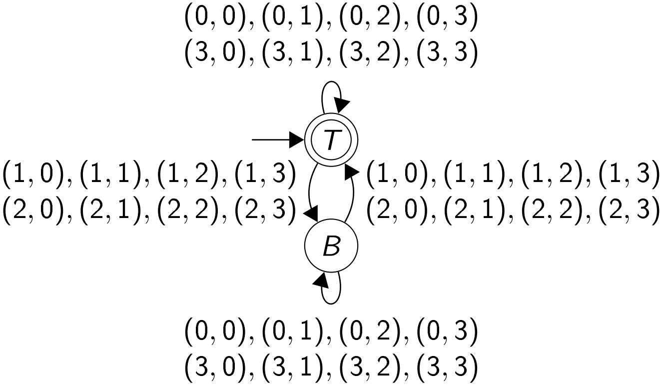
$$\mathscr{T} = \{n \in \mathbb{N} : |\operatorname{rep}_2(n)|_1 \in 2\mathbb{N}\}$$

Automaton	Set recognized
$\mathscr{A}_{\mathscr{T},2^p}$	$\mathscr{T} \times \mathbb{N}$
$\mathscr{A}_{m,2^p}$	$\{(n, mn) : n \in \mathbb{N}\}$
$\mathscr{A}_{\mathscr{T},2^p} \times \mathscr{A}_{m,2^p}$	$\{(t, mt) : t \in \mathscr{T}\}$
$\pi\left(\mathscr{A}_{\mathscr{T},2^p} \times \mathscr{A}_{m,2^p}\right)$	$m\mathscr{T}$

The automaton $\mathscr{A}_{\mathscr{T},2^p}$: $\mathscr{T} \times \mathbb{N}$ in base 2^p



The automaton $\mathscr{A}_{\mathscr{T},2^p}$: $\mathscr{T} \times \mathbb{N}$ in base 2^p



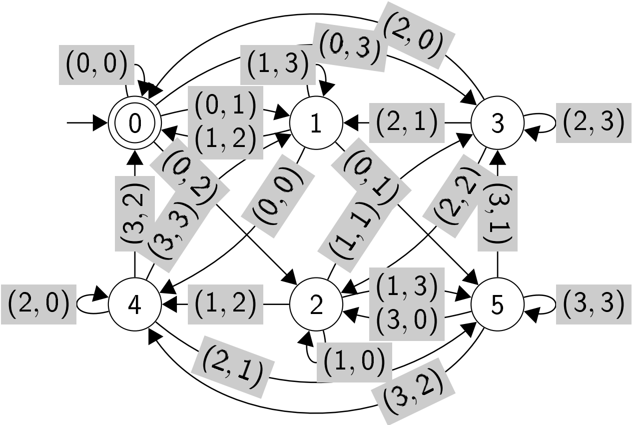
$$\delta_{\mathscr{T},2^p}(X,(d,e)) = \begin{cases} X \text{ if } d \in \mathscr{T} \\ \overline{X} \text{ otherwise} \end{cases}$$

Lemma

The automaton $\mathscr{A}_{\mathscr{T},2^p}$

- recognizes $\mathscr{T} \times \mathbb{N}$ in base 2^p
- is minimal
- is complete
- is trim
- has disjoint states

The automaton $\mathscr{A}_{m,b}$: $\{(n, mn) : n \in \mathbb{N}\}$ in base b



$$\delta_{m,b}(i,(d,e)) = j \qquad \Leftrightarrow \qquad bi + e = md + j$$

Proposition

The automaton $\mathscr{A}_{m,b}$

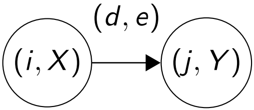
- is accessible
- is coaccessible
- has disjoint states
- is minimal
- is trim

The product $\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}$: $\{(t, mt) : t \in \mathscr{T}\}$ in base 2^p

$$(0, T), \ldots, (m-1, T) \qquad (0, B), \ldots (m-1, B)$$

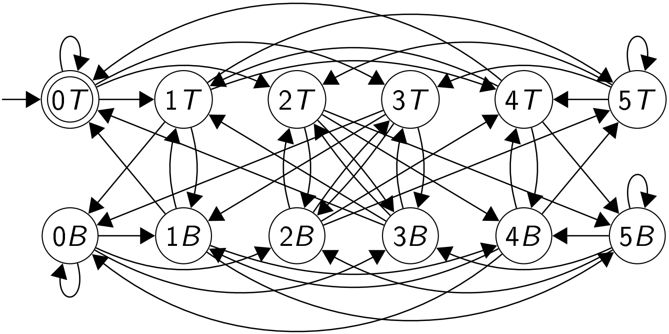
The product $\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}$: $\{(t, mt) : t \in \mathscr{T}\}$ in base 2^p

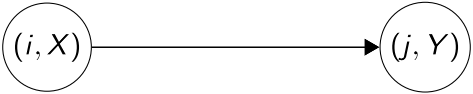
$(0, T), \ldots, (m-1, T)$ $\quad$ $(0, B), \ldots (m-1, B)$



$$2^p i + e = md + j$$

$$Y = \begin{cases} X \text{ if } d \in \mathscr{T} \\ \overline{X} \text{ otherwise} \end{cases}$$





Reading (d, e) :

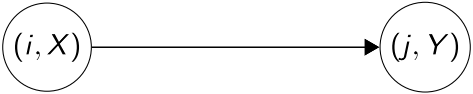
$2^p i + e = md + j$

$$Y = \begin{cases} X \text{ if } d \in \mathscr{T} \\ \overline{X} \text{ otherwise} \end{cases}$$

Reading (u, v) :

$2^{p|(u,v)|} i + \mathrm{val}_{2^p}(v) = m\,\mathrm{val}_{2^p}(u) + j$

$$Y = \begin{cases} X \text{ if } \mathrm{val}_{2^p}(u) \in \mathscr{T} \\ \overline{X} \text{ otherwise} \end{cases}$$



Reading (d, e) :

$2^p i + e = md + j$

$$Y = \begin{cases} X \text{ if } d \in \mathscr{T} \\ \overline{X} \text{ otherwise} \end{cases}$$

Reading (u, v) :

$2^{p|(u,v)|} i + \mathrm{val}_{2^p}(v) = m\,\mathrm{val}_{2^p}(u) + j$

$$Y = \begin{cases} X \text{ if } \mathrm{val}_{2^p}(u) \in \mathscr{T} \\ \overline{X} \text{ otherwise} \end{cases}$$

Remark

Given i, X, v, there exist unique j, Y, u such that we have a transition labeled by (u, v) from (i, X) to (j, Y).

Proposition

The automaton $\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}$

- is accessible
- is coaccessible
- has disjoint states
- is minimal
- is trim

Proposition

The automaton $\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}$

- is accessible
- is coaccessible
- has disjoint states
- is minimal
- is trim

$$(0, T) \to (i, T) : \quad \mathrm{rep}_{2^p}(0, i)$$

Proposition

The automaton $\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}$

- is accessible
- is coaccessible
- has disjoint states
- is minimal
- is trim

$$
\begin{aligned}
(0, T) \to (i, T) &: \quad \text{rep}_{2^p}(0, i) \\
(0, T) \to (i, B) &: \quad \text{rep}_{2^p}(1, m + i)
\end{aligned}
$$

Proposition

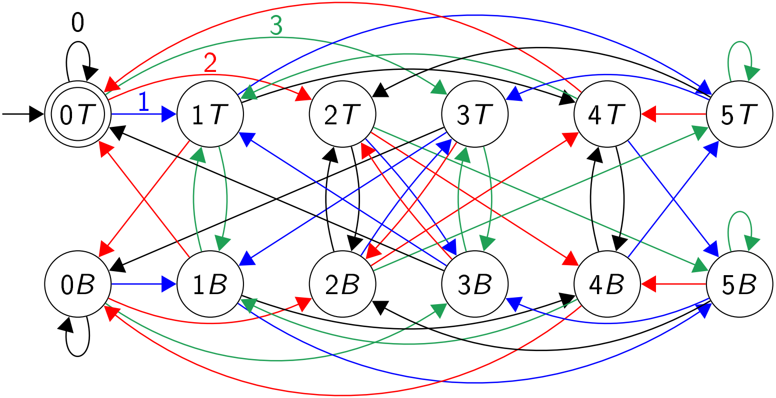
The automaton $\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}$

- is accessible
- is coaccessible
- has disjoint states
- is minimal
- is trim

$$(0,T) \to (i,T): \quad \mathrm{rep}_{2^p}(0,i)$$
$$(0,T) \to (i,B): \quad \mathrm{rep}_{2^p}(1,m+i)$$

$$(0,B) \to (0,T): \quad \mathrm{rep}_{2^p}(1,m) \text{ if } p|z$$
$$\mathrm{rep}_{2^p}\left(2^{p-r}, k2^{q+1}\right) \text{ if } z = qp + r \text{ with } q \in \mathbb{N} \text{ and } r \in \{1,\dots,p-1\}$$

The automaton $\pi\left(\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}\right)$: $m\mathscr{T}$ in base 2^p



Proposition

The automaton $\pi\left(\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}\right)$ is

- deterministic
- accessible
- coaccessible

Proposition

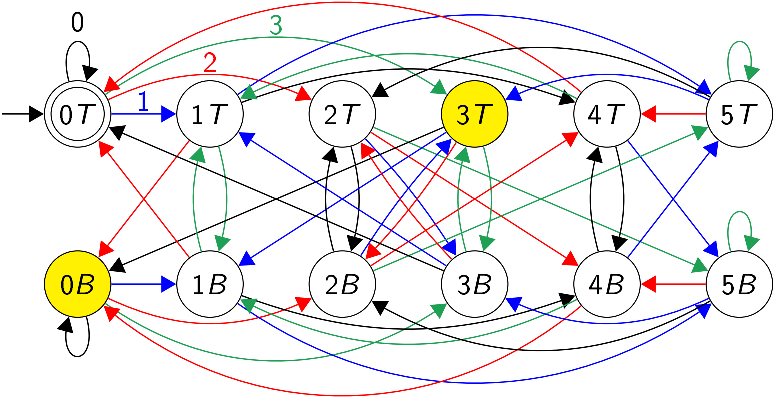
The automaton $\pi\left(\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}\right)$ is

- deterministic
- accessible
- coaccessible

Proposition

In the automaton $\pi\left(\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}\right)$, the states (i, T) and (i, B) are disjoint for all $i \in \{0, \ldots, m-1\}$.

The automaton $\pi(\mathscr{A}_{6,4} \times \mathscr{A}_{\mathscr{T},4})$



For all $n \in \mathbb{N}$, we set

$$X_n := \begin{cases} T \text{ if } n \in \mathscr{T} \\ B \text{ otherwise} \end{cases}$$

For all $n \in \mathbb{N}$, we set

$$X_n := \begin{cases} T \text{ if } n \in \mathscr{T} \\ B \text{ otherwise} \end{cases}$$

Definition

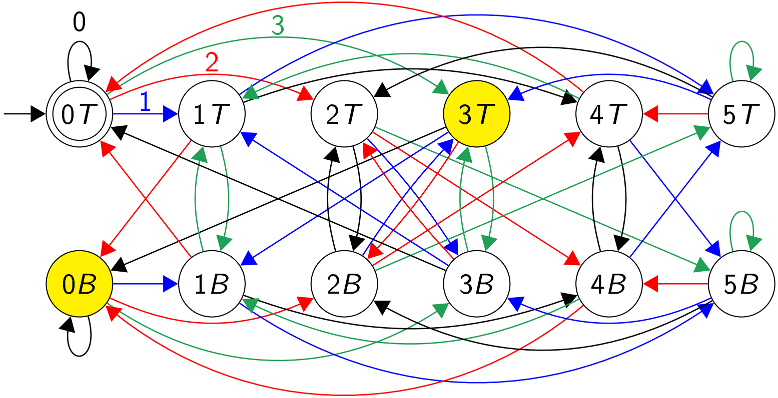
For all $j \in \{1, \ldots, k-1\}$, we set

$$[(j,T)] := \{(j + k\ell, X_\ell) : 0 \leq \ell \leq 2^z - 1\}$$
$$[(j,B)] := \left\{\left(j + k\ell, \overline{X_\ell}\right) : 0 \leq \ell \leq 2^z - 1\right\}$$

We also set

$$[(0,T)] := \{(0,T)\}$$
$$[(0,B)] := \left\{\left(k\ell, \overline{X_\ell}\right) : 0 \leq \ell \leq 2^z - 1\right\}$$

$p = 2, m = 6 = 3 \times 2^1 \Rightarrow k = 3, z = 1$



$$[(0, B)] = \left\{ \left(3\ell, \overline{X_\ell}\right) : 0 \le \ell \le 2 - 1 \right\}$$
$$= \left\{ \left(0, \overline{X_0}\right), \left(3, \overline{X_1}\right) \right\} = \{(0, B), (3, T)\}$$

Definition

For all $\alpha \in \{0, \ldots, z-1\}$, we set

$$C_\alpha := \left\{ \left(k2^{z-\alpha-1} + k2^{z-\alpha}\ell, \overline{X_\ell} \right) : 0 \leq \ell \leq 2^\alpha - 1 \right\}$$

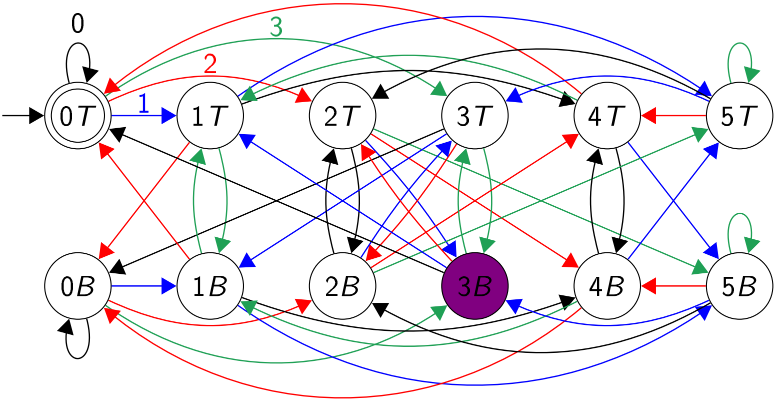
For all $\beta \in \left\{0, \ldots, \left\lceil \frac{z}{p} \right\rceil - 2\right\}$, we set

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

We also set

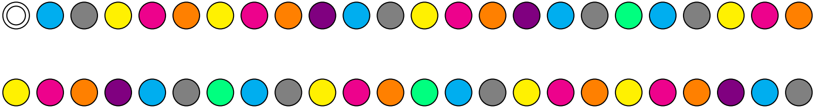
$$\Gamma_{\left\lceil \frac{z}{p} \right\rceil - 1} := \bigcup_{\alpha \in \left\{ \left(\left\lceil \frac{z}{p} \right\rceil - 1 \right) p, \ldots, z-1 \right\}} C_\alpha$$

$p = 2, m = 6 = 3 \times 2^1 \Rightarrow k = 3, z = 1$



$$\Gamma_{\lceil \frac{1}{2} \rceil - 1} = \Gamma_0 = C_0 = \left\{ \left(3 \times 2^{1-0-1} + 3 \times 2^{1-0}\ell, \overline{X_\ell} \right) : 0 \leq \ell \leq 2^0 - 1 \right\}$$
$$= \{(3, B)\}$$

The automaton $\pi\left(\mathscr{A}_{24,4} \times \mathscr{A}_{\mathscr{T},4}\right)$



We can build a new automaton

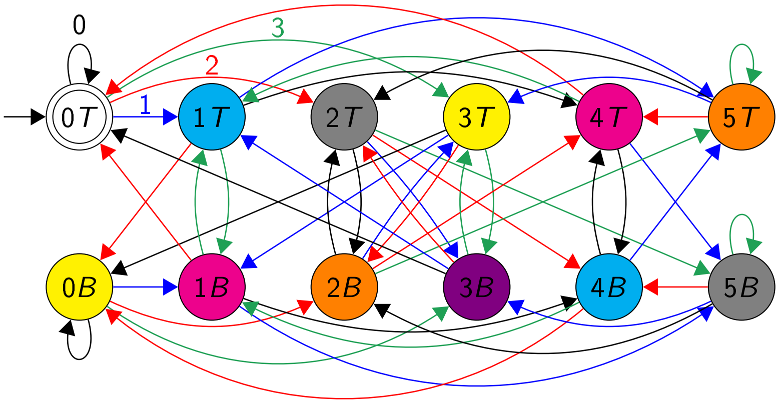
We can build a new automaton which is

- accessible

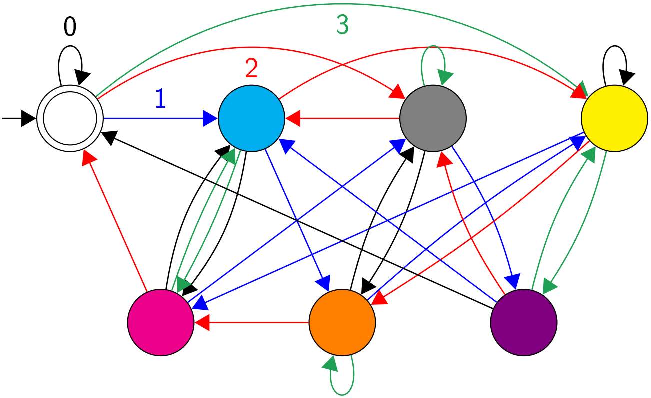
We can build a new automaton which is

- accessible
- reduced

Automaton recognizing $6\mathscr{T}$ in base 4



Automaton recognizing $6\mathscr{T}$ in base 4



Theorem

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

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

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

Theorem

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

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

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

$2 \times 3 + \left\lceil \frac{1}{2} \right\rceil = 7$

The automaton $\pi(\mathscr{A}_{24,4} \times \mathscr{A}_{\mathscr{T},4})$