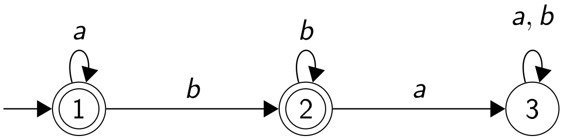
State complexity of the multiples of the Thue-Morse set

Adeline Massuir
Joint work with Émilie Charlier and Célia Cisternino

École Jeunes Chercheurs en Informatique Mathématique

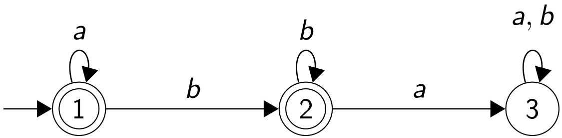
March 7th, 2019

Deterministic finite automaton (DFA) : $\mathscr{A} = (Q, q_0, F, \Sigma, \delta)$



Automata

Deterministic finite automaton (DFA) : $\mathscr{A} = (Q, q_0, F, \Sigma, \delta)$



Language – Regular language

Minimal automaton

Theorem

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

Minimal automaton

Theorem

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

Theorem

An automaton is minimal if and only if it is accessible and reduced.

Minimal automaton

Theorem

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

Theorem

An automaton is minimal if and only if it is accessible and reduced.

One algorithm :

Minimal automaton

Theorem

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

Theorem

An automaton is minimal if and only if it is accessible and reduced.

One algorithm :

1. Eject non accessible states

Minimal automaton

Theorem

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

Theorem

An automaton is minimal if and only if it is accessible and reduced.

One algorithm :

1. Eject non accessible states
2. Look for undistinguished states

State complexity

Definition

The *state complexity* of a regular language is the number of states of its minimal automaton.

What do we want to do?

Definition

Let $b \in \mathbb{N}_{\geq 2}$. A subset X of $\mathbb{N}$ is *b-recognizable* if $\text{rep}_b(X)$ is regular.

What do we want to do ?

Definition

Let $b \in \mathbb{N}_{\geq 2}$. A subset X of $\mathbb{N}$ is *b-recognizable* if $\text{rep}_b(X)$ is regular.

It is equivalent to work with $0^* \text{rep}_b(X)$.

What do we want to do ?

Definition

Let $b \in \mathbb{N}_{\geq 2}$. A subset X of $\mathbb{N}$ is *b-recognizable* if $\mathrm{rep}_b(X)$ is regular.

It is equivalent to work with $0^* \mathrm{rep}_b(X)$.

Theorem

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 language $0^* \operatorname{rep}_b(m\,\mathbb{N})$ is

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

Thue-Morse

0

Thue-Morse

0

1

Thue-Morse

01

1

Thue-Morse

01

10

Thue-Morse

0110

10

Thue-Morse

0110

1001

Thue-Morse

01101001

1001

Thue-Morse

01101001

10010110

Thue-Morse

0110100110010110

10010110

Thue-Morse

$$0110100110010110\ldots$$

Thue-Morse

$$0110100110010110\ldots$$

Definition

The Thue-Morse set is the set

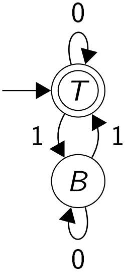
$$\mathscr{T} = \{n \in \mathbb{N} : |\operatorname{rep}_2(n)|_1 \in 2\,\mathbb{N}\}\,.$$

$$0110100110010110\ldots$$

Definition

The Thue-Morse set is the set

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



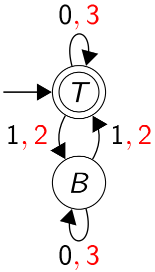
Thue-Morse

0110100110010110...

Definition

The Thue-Morse set is the set

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



Definition

Let $p, q \in \mathbb{N}_{\geq 2}$. We say that p and q are *multiplicatively independant* if

$$p^a = q^b \Rightarrow a = b = 0.$$

They are said *multiplicativement dependant* otherwise.

Definition

Let $p, q \in \mathbb{N}_{\geq 2}$. We say that p and q are *multiplicatively independant* if

$$p^a = q^b \Rightarrow a = b = 0.$$

They are said *multiplicativement dependant* otherwise.

Theorem [Cobham, 1969]

- Let b, b' two multiplicatively independant 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 dependant bases. A subset of $\mathbb{N}$ is b-recognizable iff it is b'-recognizable.

The result

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

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.

The method

Automaton	Language accepted
$\mathscr{A}_{\mathscr{T},2^p}$	$(0,0)^* \operatorname{rep}_{2^p}(\mathscr{T} \times \mathbb{N})$

The method

Automaton	Language accepted
$\mathscr{A}_{\mathscr{T},2^p}$	$(0,0)^* \operatorname{rep}_{2^p}(\mathscr{T} \times \mathbb{N})$
$\mathscr{A}_{m,2^p}$	$(0,0)^* \operatorname{rep}_{2^p}(\{(n, mn) : n \in \mathbb{N}\})$

The method

Automaton	Language accepted
$\mathscr{A}_{\mathscr{T},2^p}$	$(0,0)^* \operatorname{rep}_{2^p}(\mathscr{T} \times \mathbb{N})$
$\mathscr{A}_{m,2^p}$	$(0,0)^* \operatorname{rep}_{2^p}(\{(n, mn) : n \in \mathbb{N}\})$
$\mathscr{A}_{\mathscr{T},2^p} \times \mathscr{A}_{m,2^p}$	$(0,0)^* \operatorname{rep}_{2^p}(\{(t, mt) : t \in \mathscr{T}\})$

The method

Automaton	Language accepted
$\mathscr{A}_{\mathscr{T},2^p}$	$(0,0)^* \operatorname{rep}_{2^p}(\mathscr{T} \times \mathbb{N})$
$\mathscr{A}_{m,2^p}$	$(0,0)^* \operatorname{rep}_{2^p}(\{(n, mn) : n \in \mathbb{N}\})$
$\mathscr{A}_{\mathscr{T},2^p} \times \mathscr{A}_{m,2^p}$	$(0,0)^* \operatorname{rep}_{2^p}(\{(t, mt) : t \in \mathscr{T}\})$
$\pi\left(\mathscr{A}_{\mathscr{T},2^p} \times \mathscr{A}_{m,2^p}\right)$	$0^* \operatorname{rep}_{2^p}(m\mathscr{T})$

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

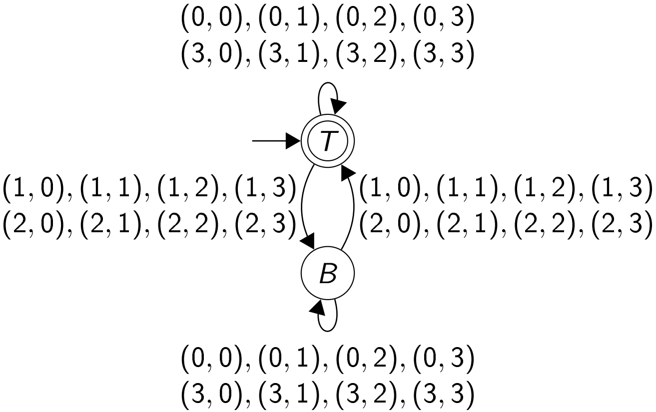
$$(0,0)^* \{\mathrm{rep}_{2^p}(t,n) : t \in \mathscr{T}, n \in \mathbb{N}\}$$

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

$(0,0)^* \{\mathrm{rep}_{2^p}(t,n) : t \in \mathscr{T}, n \in \mathbb{N}\}$

States	T, B
Initial state	T
Final states	T
Alphabet	$\{0,\ldots,2^p-1\}^2$
Transitions	$\delta_{\mathscr{T},2^p}(X,(a,b)) = \begin{cases} X & \text{if } a \in \mathscr{T} \\ \overline{X} & \text{else.} \end{cases}$

The automaton $\mathscr{A}_{\mathscr{T},4}$



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

$(0,0)^* \{\mathrm{rep}_b(n, mn) : n \in \mathbb{N}\}$

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

$$(0,0)^* \{\mathrm{rep}_b(n, mn) : n \in \mathbb{N}\}$$

States	$0, \ldots, m-1$
Initial state	0
Final states	0
Alphabet	$\{0, \ldots, b-1\}^2$
Transitions	$\delta_{m,b}(i, (d, e)) = j \Leftrightarrow bi + e = md + j$

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

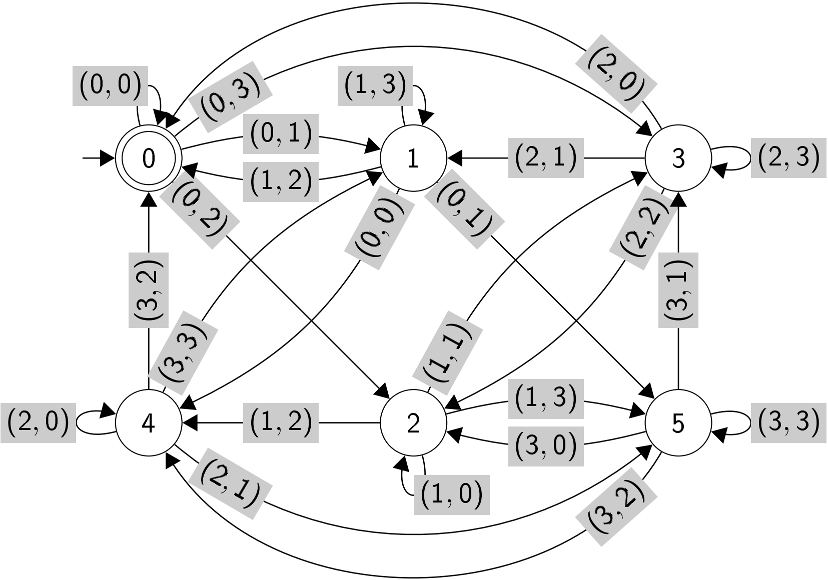
$(0,0)^* \{\text{rep}_b(n, mn) : n \in \mathbb{N}\}$

States	$0, \ldots, m-1$
Initial state	0
Final states	0
Alphabet	$\{0, \ldots, b-1\}^2$
Transitions	$\delta_{m,b}(i,(d,e)) = j \Leftrightarrow bi + e = md + j$

For all $i, j \in \{0, \ldots, m-1\}$, for all $u, v \in \{0, \ldots, b-1\}^*$,

$$\delta_{m,b}(i,(u,v)) = j \Leftrightarrow b^{|(u,v)|} i + \text{val}_b(v) = m\,\text{val}_b(u) + j.$$

The automaton $\mathscr{A}_{6,4}$



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

$(0,0)^* \{\mathrm{rep}_{2^p}(t, mt) : t \in \mathscr{T}\}$

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

$(0,0)^* \{\mathrm{rep}_{2^p}(t, mt) : t \in \mathscr{T}\}$

States	$(0, T), \dots, (m-1, T), (0, B), \dots, (m-1, B)$
Initial state	$(0, T)$
Final states	$(0, T)$
Alphabet	$\{0, \dots, 2^p - 1\}^2$
Transitions	$\delta_{\mathscr{T},2^p}((i, X), (u, v)) = (j, Y)$ $\Leftrightarrow 2^{p\lvert(u,v)\rvert} i + \mathrm{val}_{2^p}(v) = m\,\mathrm{val}_{2^p}(u) + j$ and $Y = \begin{cases} X \text{ if } \mathrm{val}_{2^p}(u) \in \mathscr{T} \\ \overline{X} \text{ else.} \end{cases}$

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

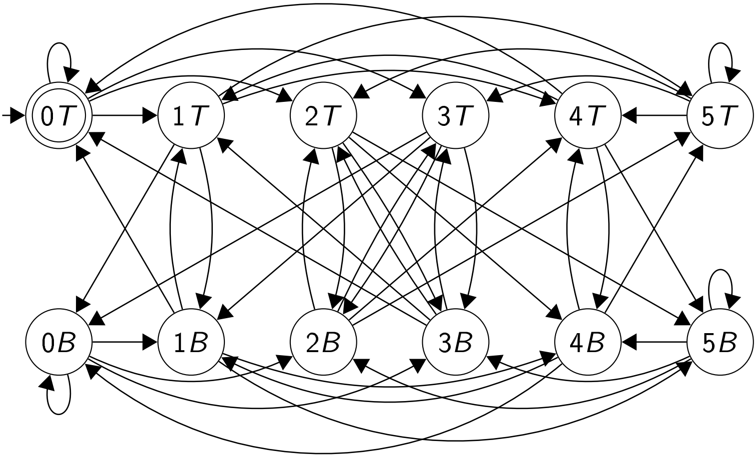
$$(0,0)^* \{\mathrm{rep}_{2^p}(t, mt) : t \in \mathscr{T}\}$$

States	$(0,T),\ldots,(m-1,T),(0,B),\ldots,(m-1,B)$
Initial state	$(0,T)$
Final states	$(0,T)$
Alphabet	$\{0,\ldots,2^p-1\}^2$
Transitions	$\delta_{\mathscr{T},2^p}((i,X),(u,v)) = (j,Y)$ $\Leftrightarrow 2^{p\lvert(u,v)\rvert} i + \mathrm{val}_{2^p}(v) = m\,\mathrm{val}_{2^p}(u) + j$ and $Y = \begin{cases} X & \text{if } \mathrm{val}_{2^p}(u) \in \mathscr{T} \\ \overline{X} & \text{else.} \end{cases}$

Remark

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

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

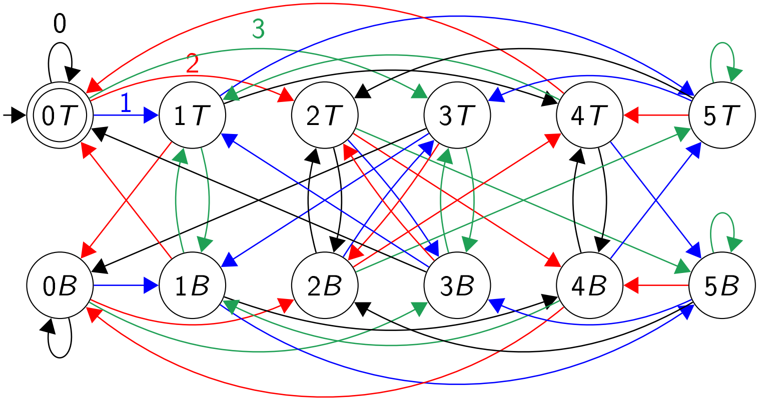


The projected automaton $\pi\left(\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}\right)$

$0^* \operatorname{rep}_{2^p}\left(m\mathscr{T}\right) = 0^* \left\{\operatorname{rep}_{2^p}(mt) : t \in \mathscr{T}\right\}$

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

$$0^* \operatorname{rep}_{2^p}(m\mathscr{T}) = 0^* \{\operatorname{rep}_{2^p}(mt) : t \in \mathscr{T}\}$$



Proposition

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

- deterministic,
- accessible,
- coaccessible.

Proposition

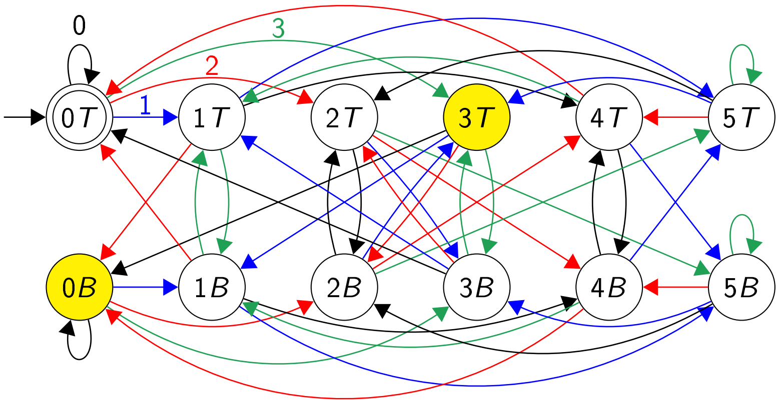
The automaton $\pi\left(\mathscr{A}_{m,2^p} \times \mathscr{A}_{\mathscr{T},2^p}\right)$ is

- deterministic,
- accessible,
- coaccessible.

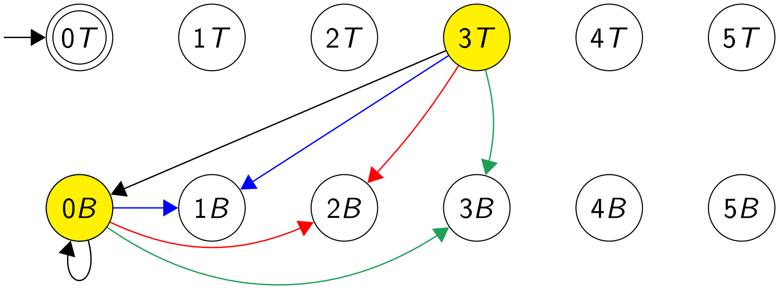
Proposition

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 disjoined for all $i \in \{0, \dots, m-1\}$.

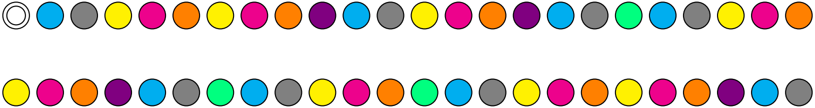
The automaton $\pi(\mathscr{A}_{6,4} \times \mathscr{A}_{\mathscr{T},4})$



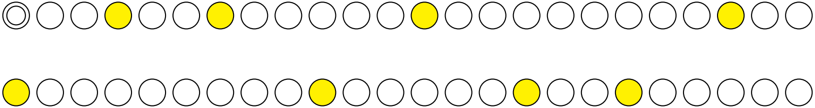
The automaton $\pi(\mathscr{A}_{6,4} \times \mathscr{A}_{\mathscr{T},4})$



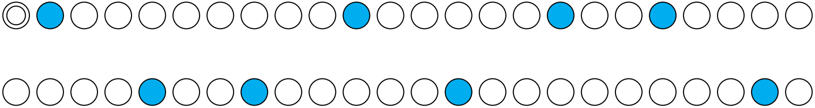
The automaton $\pi(\mathscr{A}_{24,4} \times \mathscr{A}_{\mathscr{T},4})$



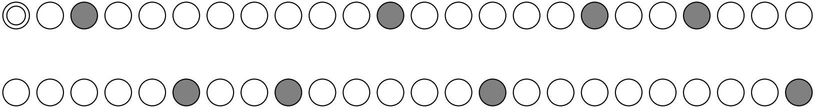
The automaton $\pi(\mathscr{A}_{24,4} \times \mathscr{A}_{\mathscr{T},4})$



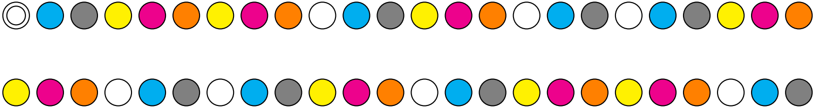
The automaton $\pi\left(\mathscr{A}_{24,4} \times \mathscr{A}_{\mathscr{T},4}\right)$



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



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



Recall

For $m \in \mathbb{N}, p \in \mathbb{N}_{\geq 1}$, we write $m = k2^z$ with z odd.

Recall

For $m \in \mathbb{N}, p \in \mathbb{N}_{\geq 1}$, we write $m = k2^z$ with z odd.

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

$$T_n := \begin{cases} T \text{ if } n \in \mathscr{T} \\ B \text{ else.} \end{cases}$$

Recall

For $m \in \mathbb{N}, p \in \mathbb{N}_{\geq 1}$, we write $m = k2^z$ with z odd.

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

$$T_n := \begin{cases} T \text{ if } n \in \mathscr{T} \\ B \text{ else.} \end{cases}$$

Definition

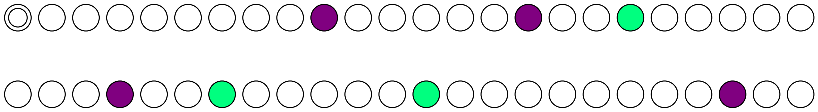
For all $j \in \{1, \dots, k-1\}$, we set

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

We also set

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

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



Definition

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

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

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

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

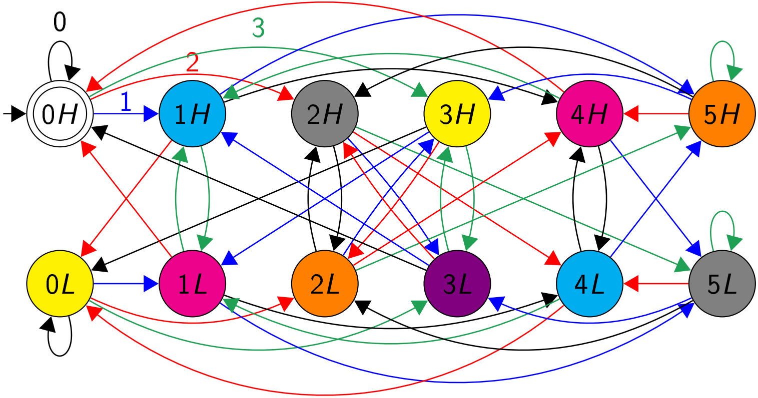
We can build a new automaton

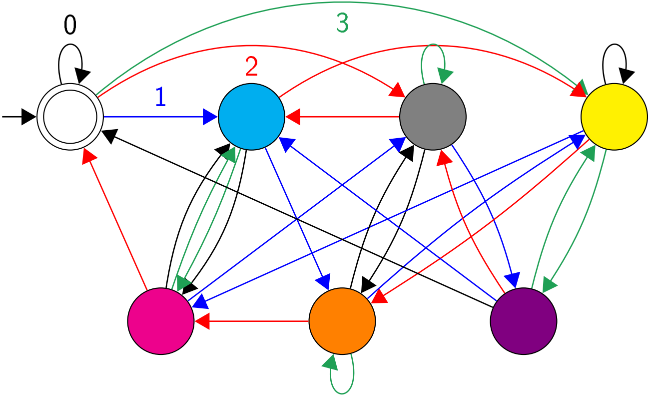
We can build a new automaton which is

- accessible

We can build a new automaton which is

- accessible
- reduced.


0
3
2
1
0H
1H
2H
3H
4H
5H
0L
1L
2L
3L
4L
5L


0
1
2
3

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 equal to

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

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 equal to

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

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

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

To go further

What about the language

$$0^* \operatorname{rep}_{2^p}(m\mathscr{T} + r)$$

where $r \in \{0, \ldots, m-1\}$?