State complexity of the multiples of the Thue-Morse set

Emilie Charlier, Célia Cisternino, Adeline Massuir
University of Liège, Departement of Mathematics, Liège, Belgium
{echarlier, ccisternino, a.massuir}@uliege.be

Let $\mathcal{T}$ be the set of nonnegative integers having an even number of occurrences of the digit 1 when written in base 2:

$$\mathcal{T} = \{n \in \mathbb{N} \colon |\text{rep}_2(n)|_1 \in 2\mathbb{N}\}.$$

This set $\mathcal{T}$ is called the Thue-Morse set of integers. The first elements of $\mathcal{T}$ are $0, 3, 5, 6, 9, \ldots$. It is 2-recognizable, which means that there is a finite automaton accepting exactly the binary representations of the integers in $\mathcal{T}$.

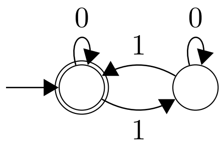

Figure 1: The Thue-Morse set is 2-recognizable.

Moreover, the Thue-Morse set $\mathcal{T}$ is 2^p-recognizable for all $p \in \mathbb{N}_{\geq 1}$ and is not b-recognizable for any other base b. This is a consequence of the famous theorem of Cobham [3]. The automaton accepting $0^*\text{rep}_{2^p}(\mathcal{T})$ can be easily deduced from the one shown on Figure 1.

Knowing that the property of being b-recognizable is stable under multiplication by a constant and under translation, we are interested in caracterizing the state complexity of the sets $m\mathcal{T} + r$ in base 2^p depending on the multiple m and the remainder r, which means that we are trying to compute the number of states of the trim minimal automaton accepting the languages $0^*\text{rep}_{2^p}(m\mathcal{T} + r)$.

The main result of this work is the following one.

Theorem 1. *Let $m \in \mathbb{N}$, $r \in \{0, \ldots m-1\}$ and $p \in \mathbb{N}_{\geq 1}$. The state complexity of the language $0^*\text{rep}_{2^p}(m\mathcal{T} + r)$ is equal to*

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

if $m = k2^z$ where k is an odd integer.

In order to prove this theorem, we construct consecutive automata and the minimization of the last automaton obtained is the key of the reasoning.
This work is closely related to the article [1], which concerns the state complexity of the set $m\mathbb{N}$ in base b for all $m, b \in \mathbb{N}_0$. However, the way to prove our results is different. An extended question has also been considered in [2] by studing the state complexity of the greedy representations of $m\mathbb{N}$, $m \in \mathbb{N}_{\geq 2}$ for a wide class of linear numeration systems.

References

[1] Boris Alexeev, *Minimal DFA for testing divisibility*, Journal of Computer and System Sciences 69 (2004), 2, pp. 235–243.

[2] É. Charlier, N. Rampersad, M. Rigo and L. Waxweiler. *The minimal automaton recognizing $m\mathbb{N}$ in a linear numeration system.* Integers 11B, A4, 24 pp, 2011.

[3] A. Cobham, *On the base-dependence of sets of numbers recognizable by finite automata.* Math. Systems Theory, 3:186–192, 1969.