

Generating expansions in uncountably many Cantor real bases via a single transducer

Émilie Charlier

joint work with Pierre Popoli and Michel Rigo

Département de mathématiques, ULiège

Arithmétique en Plat Pays
Calais, France, September 21

Cantor real bases and alternate bases

A **Cantor real base** is a sequence $\mathbf{B} = (\beta_n)_{n \geq 0}$ of real numbers such that

- ▶ $\beta_n > 1$ for all n
- ▶ $\prod_{n=0}^{\infty} \beta_n = \infty$.

A **B-representation** of a real number x is an infinite sequence $a = (a_n)_{n \geq 0}$ of integers such that

$$x = \sum_{n=0}^{\infty} \frac{a_n}{\beta_0 \beta_1 \cdots \beta_n}.$$

An **alternate base** is a periodic Cantor base. In this case, we simply write $\mathbf{B} = (\beta_0, \dots, \beta_{p-1})$.

Greedy expansions

For $x \in [0, 1]$, a distinguished $\mathbf{B}$ -representation

$$d_{\mathbf{B}}(x) = (a_n)_{n \geq 0},$$

called the **greedy $\mathbf{B}$ -expansion** of x , is obtained from the greedy algorithm: if the digits $a_0, \dots, a_{N-1}$ have already been defined, then the next digit a_N is the largest possible integer such that

$$\sum_{n=0}^N \frac{a_n}{\beta_0 \beta_1 \cdots \beta_n} \leq x.$$

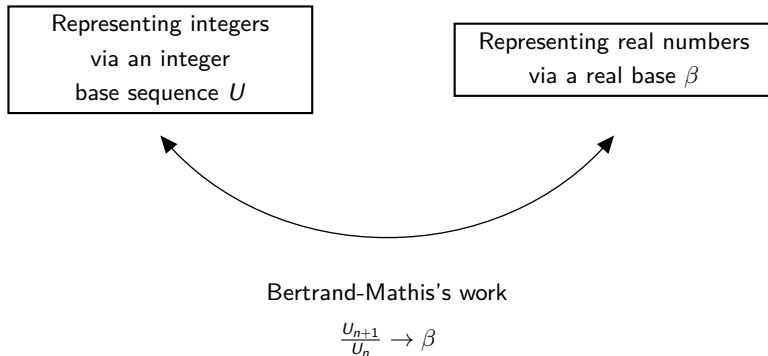
The greedy digits can also be obtained by using the so-called **β -transformation**

$$T_{\beta}: [0, 1] \mapsto [0, 1), \quad x \mapsto \beta x - \lfloor \beta x \rfloor.$$

We then have

$$a_n = \lfloor \beta_n T_{\beta_{n-1}} \cdots T_{\beta_1} T_{\beta_0}(x) \rfloor.$$

Motivation



In the general case, there is a similar relationship with representations of real numbers via some alternate base $\mathbf{B} = (\beta_0, \dots, \beta_{p-1})$.

A linear system with no dominant root

Consider $U = (U_n)_{n \geq 0}$ starting with 1, 3, 4, 9 and such that $U_n = 3U_{n-2} + U_{n-4}$ for $n \geq 4$.

We get $U = (1, 3, 4, 9, 13, 30, 43, 99 \dots)$.

The quotient $\frac{U_{n+1}}{U_n}$ does not have a limit, but alternates between two values

$$\alpha := \lim_{n \rightarrow +\infty} \frac{U_{2n+1}}{U_{2n}} = \frac{1 + \sqrt{13}}{2} \quad \text{and} \quad \beta := \lim_{n \rightarrow +\infty} \frac{U_{2n+2}}{U_{2n+1}} = \frac{5 + \sqrt{13}}{6}.$$

Take a look at the largest representations of each length:

43	30	13	9	4	3	1	$U_n - 1$
							0
						2	2
					1	0	3
				2	0	0	8
			1	0	1	0	12
		2	0	0	1	0	29
	1	0	1	0	1	0	42
2	0	0	1	0	1	0	98

Continuing this way, we will see the two ultimately periodic infinite words $(10)^\omega$ and $200(10)^\omega$.

These infinite words are not random. To understand them, we have to represent 1 using the alternate base $\mathbf{B} = (\alpha, \beta) = (\alpha, \beta, \alpha, \beta, \dots)$.

We can compute $d_{\mathbf{B}}(1) = 2010^\omega$. In particular, we have $1 = \frac{2}{\alpha} + \frac{1}{\alpha^2\beta}$.

Now, considering the shifted alternate base $\sigma(\mathbf{B}) = (\beta, \alpha) = (\beta, \alpha, \beta, \alpha, \dots)$.

Then $d_{\sigma(\mathbf{B})}(1) = 110^\omega$. In particular, we have $1 = \frac{1}{\beta} + \frac{1}{\beta\alpha}$.

We find the largest representations of each length in the numeration system U by using the **quasi-greedy expansions** of 1 both in the alternate bases $\mathbf{B}$ and $\sigma(\mathbf{B})$:

$$d_{\mathbf{B}}^*(1) = 200d_{\sigma(\mathbf{B})}^*(1) = 20010d_{\sigma(\mathbf{B})}^*(1) = 2001010d_{\sigma(\mathbf{B})}^*(1) = \dots = 200(10)^\omega$$

and

$$d_{\sigma(\mathbf{B})}^*(1) = 10d_{\sigma(\mathbf{B})}^*(1) = 1010d_{\sigma(\mathbf{B})}^*(1) = \dots = (10)^\omega.$$

Quasi-greedy expansions

For a Cantor real base $\mathbf{B} = (\beta_n)_{n \geq 0}$, another distinguished $\mathbf{B}$ -representation of $x \in (0, 1]$ is the **quasi-greedy $\mathbf{B}$ -expansion**. It is obtained recursively:

$$d_{\mathbf{B}}^*(x) = \begin{cases} d_{\mathbf{B}}(x), & \text{if } d_{\mathbf{B}}(x) \text{ does not end in } 0^\omega; \\ a_0 a_1 \cdots a_{\ell-2} (a_{\ell-1} - 1) d_{\sigma^\ell(\mathbf{B})}^*(1), & \text{if } d_{\mathbf{B}}(x) = a_0 a_1 \cdots a_{\ell-1} 0^\omega, \ a_{\ell-1} \neq 0, \end{cases}$$

where σ is the shift operator on sequences, so that $\sigma^\ell(\mathbf{B}) = (\beta_n)_{n \geq \ell}$.

Starting point of this work

Let now $\mathbf{B} = (\beta_n)_{n \geq 0} = (\alpha, \beta, \beta, \alpha, \beta, \alpha, \alpha, \beta, \dots)$ be the Thue-Morse sequence over $\{\alpha, \beta\}$:

$$\beta_n = \begin{cases} \alpha & \text{if the binary expansion of } n \text{ has an even number of 1's,} \\ \beta & \text{otherwise.} \end{cases}$$

We compute $d_{\mathbf{B}}(1) = 20010110^\omega$.

$r_0 = 1$	$a_0 = \lfloor \alpha r_0 \rfloor = 2$
$r_1 = T_\alpha(r_0) = \alpha r_0 - a_0 = \frac{-3+\sqrt{13}}{2}$	$a_1 = \lfloor \beta r_1 \rfloor = 0$
$r_2 = T_\beta(r_1) = \beta r_1 - a_1 = \frac{-1+\sqrt{13}}{6}$	$a_2 = \lfloor \beta r_2 \rfloor = 0$
$r_3 = T_\beta(r_2) = \beta r_2 - a_2 = \frac{2+\sqrt{13}}{9}$	$a_3 = \lfloor \alpha r_3 \rfloor = 1$
$r_4 = T_\alpha(r_3) = \alpha r_3 - a_3 = \frac{-1+\sqrt{13}}{6}$	$a_4 = \lfloor \beta r_4 \rfloor = 0$
$r_5 = T_\beta(r_4) = \beta r_4 - a_4 = \frac{2+\sqrt{13}}{9}$	$a_5 = \lfloor \alpha r_5 \rfloor = 1$
$r_6 = T_\alpha(r_5) = \alpha r_5 - a_5 = \frac{-1+\sqrt{13}}{6}$	$a_6 = \lfloor \alpha r_6 \rfloor = 1$
$r_7 = T_\alpha(r_6) = \alpha r_6 - a_6 = 0$	$a_7 = \lfloor \beta r_7 \rfloor = 0$

Some questions we had in mind

- ▶ Do structural properties of the sequence $\mathbf{B}$ yield some structural properties of the $\mathbf{B}$ -expansions?
- ▶ What is the dependence to the involved bases $\beta_0, \beta_1, \beta_2 \dots$?
- ▶ Can we decide some properties of $\mathbf{B}$ -expansions, when $\mathbf{B}$ enjoys some suitable combinatorial structure?
- ▶ Can we decide admissibility with respect to some aperiodic sequence $\mathbf{B}$?

One transducer for uncountably many Cantor real bases

We make use of the notation $T_\beta: [0, 1] \rightarrow [0, 1)$, $r \mapsto \beta r - \lfloor \beta r \rfloor$.

Definition (Greedy transducer)

Let $E \subset \mathbb{R}_{>1}$ be a finite alphabet. We define an infinite transducer denoted by $\mathcal{T}_E$.

- ▶ The set of states is the interval $[0, 1]$;
- ▶ The input alphabet is E ;
- ▶ The output alphabet is $\mathbb{N}$;
- ▶ For all states $r \in [0, 1]$ and all letters $\beta \in E$, there is a transition

$$r \xrightarrow{\beta \mid \lfloor \beta r \rfloor} T_\beta(r).$$

Proposition

When reading $\mathbf{B} \in E^\mathbb{N}$ from a state $r \in [0, 1]$, the transducer $\mathcal{T}_E$ outputs the greedy $\mathbf{B}$ -representation of r .

We make use of the notation $T_{\beta}^*: (0, 1] \rightarrow (0, 1]$, $r \mapsto \beta r - \lceil \beta r - 1 \rceil$.

Definition (Quasi-greedy transducer)

Analogously, we define another infinite transducer denoted by $\mathcal{T}_E^*$, but with set of states given by $(0, 1]$ and with transitions

$$r \xrightarrow{\beta \mid \lceil \beta r - 1 \rceil} T_{\beta}^*(r).$$

Proposition

When reading $\mathbf{B} \in E^{\mathbb{N}}$ from a state $r \in (0, 1]$, the transducer $\mathcal{T}_E^*$ outputs the quasi-greedy $\mathbf{B}$ -expansion of r .

Definition

Considering a given $r \in [0, 1]$ as initial state, we let $\mathcal{T}_{E,r}$ denote the part of $\mathcal{T}_E$ that can be reached from r . We define $\mathcal{T}_{E,r}^*$ analogously for $r \in (0, 1]$.

Proposition

If the transducer $\mathcal{T}_{E,r}^*$ is finite, then so is the transducer $\mathcal{T}_{E,r}$.

Conditions ensuring finiteness

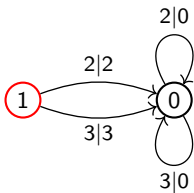
A **Pisot number** is a real algebraic integer which is greater than 1 and whose Galois conjugates all lie in the open unit disc.

Theorem (C-Popoli-Rigo)

Let δ be an algebraic integer of degree d and let $E \subset \mathbb{Z}[\delta]$ be a finite set of Pisot numbers of degree d . For all $r \in \mathbb{Q}(\delta) \cap (0, 1]$, the transducers $\mathcal{T}_{E,r}$ and $\mathcal{T}_{E,r}^*$ are finite.

Examples

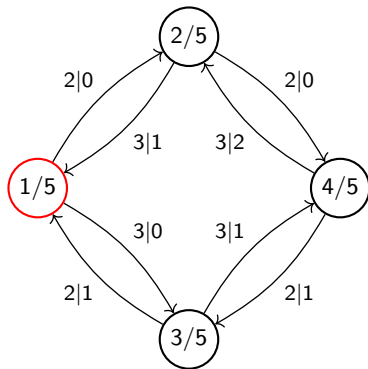
- ▶ For any finite set E of integer bases and any $r \in \mathbb{Q}$, the transducers $\mathcal{T}_{E,r}$ and $\mathcal{T}_{E,r}^*$ are finite.
- ▶ For $E = \{2, 3\}$, we get



$\mathcal{T}_{E,1}$

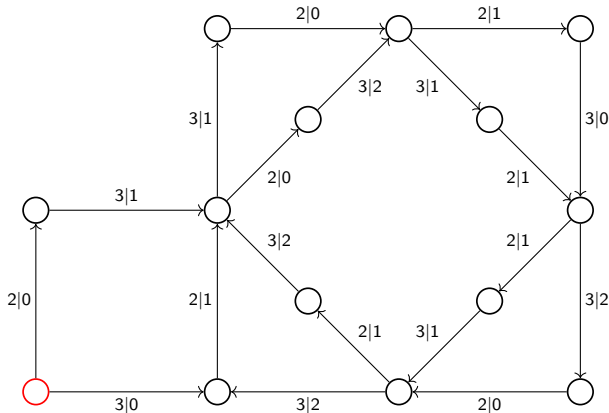


$\mathcal{T}_{E,1}^*$



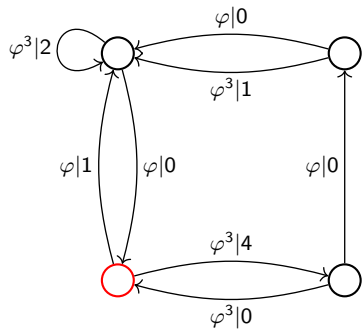
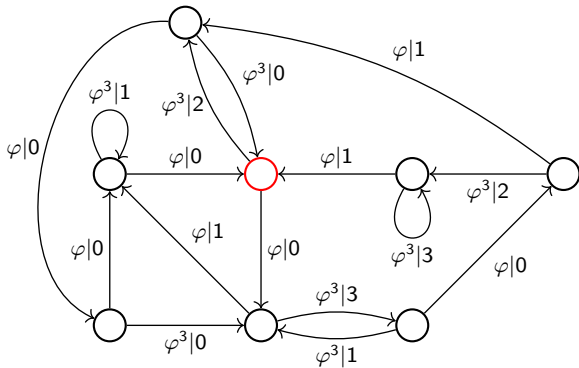
$\mathcal{T}_{E,1/5} = \mathcal{T}_{E,1/5}^*$

- ▶ For $r = \frac{932}{3885}$, we get a transducer $\mathcal{T}_{E,r}$ with 180 states.
- ▶ If we only consider inputs in $\{23, 32\}^\omega$, we obtain a transducer with only 14 states.



- ▶ This is interesting if, for instance, we focus on the Thue-Morse sequence $\mathbf{B} = 23323223 \cdots$ over $\{2, 3\}$.

- ▶ Let $\varphi = \frac{1+\sqrt{5}}{2}$ be the golden ratio and let $E = \{\varphi, \varphi^3\}$.


$$\mathcal{T}_{E,1}^*$$

$$\mathcal{T}_{E, \frac{1}{2}}^*$$

First applications: combinatorial properties preserved through finite transducers

Corollary

- ▶ If $\mathbf{B} \in E^{\mathbb{N}}$ is an ultimately periodic sequence, then for all $r \in \mathbb{Q}(\delta) \cap [0, 1]$, the sequences $d_{\mathbf{B}}(r)$ and $d_{\mathbf{B}}^*(r)$ are also ultimately periodic.
- ▶ If $\mathbf{B} \in E^{\mathbb{N}}$ is a b -automatic sequence for some integer base $b \geq 2$, then for all $r \in \mathbb{Q}(\delta) \cap [0, 1]$, the sequences $d_{\mathbf{B}}(r)$ and $d_{\mathbf{B}}^*(r)$ are also b -automatic.
- ▶ If $\mathbf{B} \in E^{\mathbb{N}}$ is a morphic sequence, then for all $r \in \mathbb{Q}(\delta) \cap [0, 1]$, the sequences $d_{\mathbf{B}}(r)$ and $d_{\mathbf{B}}^*(r)$ are also morphic.

An infinite transducer $\mathcal{T}_{E,1}$

Let φ be the golden ratio and let $E = \{\varphi, 2\}$.

The set of states of the transducer that can be reached from 1 is infinite.

We can see this by reading the Cantor real base $\mathbf{B} = \varphi 2^\omega$ as input.

- ▶ After the first step using base φ , we reach the state $\varphi - 1$.
- ▶ Thus, one has to represent $\varphi - 1$ using only base 2.
- ▶ Since this number is irrational, its 2-expansion is aperiodic and we never encounter the same remainder twice in the greedy algorithm.

An infinite transducer $\mathcal{T}_{E,1}$

Let φ be the golden ratio and let $E = \{\varphi, 2\}$.

The set of states of the transducer that can be reached from 1 is infinite.

We can see this by reading the Cantor real base $\mathbf{B} = \varphi 2^\omega$ as input.

- ▶ After the first step using base φ , we reach the state $\varphi - 1$.
- ▶ Thus, one has to represent $\varphi - 1$ using only base 2.
- ▶ Since this number is irrational, its 2-expansion is aperiodic and we never encounter the same remainder twice in the greedy algorithm.

The assumptions of our result are not fully satisfied: can you see why?

An infinite transducer $\mathcal{T}_{E,1}$

Let φ be the golden ratio and let $E = \{\varphi, 2\}$.

The set of states of the transducer that can be reached from 1 is infinite.

We can see this by reading the Cantor real base $\mathbf{B} = \varphi 2^\omega$ as input.

- ▶ After the first step using base φ , we reach the state $\varphi - 1$.
- ▶ Thus, one has to represent $\varphi - 1$ using only base 2.
- ▶ Since this number is irrational, its 2-expansion is aperiodic and we never encounter the same remainder twice in the greedy algorithm.

The assumptions of our result are not fully satisfied: can you see why?

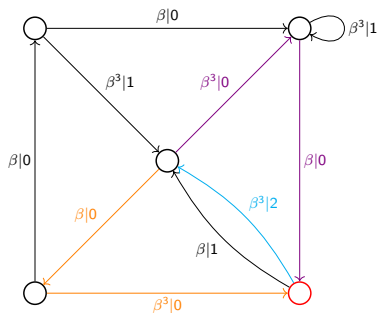
Indeed, φ has algebraic degree 2 whereas 2 has algebraic degree 1.

Cantor real bases giving rise to a periodic expansion of a given real number

Two-walk property

We say that $\mathcal{T}_{E,r}^*$ has the **two-walk property** if one can find two closed walks starting from the same state, with distinct inputs and the same output.

Let β be the smallest Pisot number, i.e., the real root of $X^3 - X - 1$. Consider the set $E = \{\beta, \beta^3\}$.

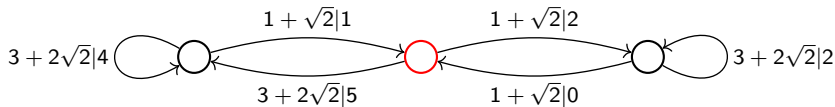


Let us write $u = \beta^3\beta\beta^3$ and $v = \beta^3\beta^3\beta$. Any Cantor real base $\mathbf{B} \in \{u, v\}^\omega$ leads to $d_{\mathbf{B}}^*(1) = (200)^\omega$.

Theorem (C-Popoli-Rigo)

The Cantor real bases $\mathbf{B} \in E^{\mathbb{N}}$ such that $d_{\mathbf{B}}^*(r)$ is ultimately periodic are exactly the ultimately periodic ones if and only if the transducer $\mathcal{T}_{E,r}^*$ does not have the two-walk property.

Consider the set $E = \{\beta, \beta^2\}$, where β is the Pisot number $1 + \sqrt{2}$ and its square β^2 is $3 + 2\sqrt{2}$. The transducer $\mathcal{T}_{E,1}^*$ does not satisfy the two-walk property.



Proposition

The two-walk property is decidable in $\mathcal{O}((\#Q)^3)$ steps, where Q is the set of states.

Deciding admissibility in base B

Main question: Given an infinite sequence $\mathbf{a}$, can we decide whether it is B -admissible, that is, whether there exists some $x \in [0, 1)$ such that $\mathbf{a} = d_B(x)$?

Theorem (Parry 1960)

Let $\beta > 1$ be a real base. A sequence $\mathbf{a}$ over $\mathbb{N}$ is β -admissible if and only if $\sigma^n(\mathbf{a}) <_{\text{lex}} d_\beta^*(1)$ for all $n \in \mathbb{N}$.

Theorem (Caalim-Demegillo 2020, C-Cisternino 2021)

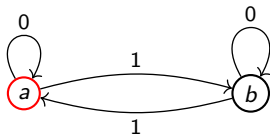
Let B be a Cantor real base. A sequence $\mathbf{a}$ over $\mathbb{N}$ is B -admissible if and only if $\sigma^n(\mathbf{a}) <_{\text{lex}} d_{\sigma^n(B)}^*(1)$ for all $n \in \mathbb{N}$.

- ▶ In the Cantor real base case, these lexicographic conditions are far from being trivial to test, even in a simple situation where the Cantor base is the Thue–Morse word $\mathbf{T} = (2, 3, 3, 2, 3, 2, 2, 3, \dots)$ over $\{2, 3\}$.
- ▶ This is due to the fact that the set $\{\sigma^n(\mathbf{T}) : n \geq 0\}$ is infinite since $\mathbf{T}$ is aperiodic.
- ▶ If $\mathbf{a}$ and B have no specific structure, then there is little chance that we can decide anything.

Automatic sequences

Let $b \geq 2$ be an integer. A sequence $(x_n)_{n \geq 0}$ is **b -automatic** if there exists a DFAO that produces x_n as output when the corresponding input is the b -ary expansion of n .

- For example, the TM sequence is 2-automatic.



n	0	1	2	3	4	5	6	7	8	...
$\text{rep}_2(n)$	ε	1	10	11	100	101	110	111	1000	...
x_n	a	b	b	a	b	a	a	b	b	...

Automatic sequences and logic

Theorem (Büchi 1960, Bruyère 1985)

A sequence is b -automatic if and only if it is definable by a first-order formula of $\langle \mathbb{N}, +, V_b \rangle$, where $V_b(0) = 0$ and for $n \neq 0$, $V_b(n) = b^i$ where b^i is the highest power of b dividing n .

Corollary

The first-order theory of the structure $\langle \mathbb{N}, +, V_b \rangle$ is decidable.

- ▶ This means that we can decide many combinatorial properties of automatic sequences.
- ▶ In our case, two sequences are involved: the base **B** and the given sequence **a** of which we want to decide the admissibility.
- ▶ The lexicographic conditions to test involve the shifts of both these sequences.

A technical decidability result

Theorem (C-Popoli-Rigo)

Let $E \subset \mathbb{R}_{>1}$ be a finite alphabet of real bases, let $\mathbf{B} \in E^{\mathbb{N}}$, and let $b \geq 2$ be an integer base.

Assume that, for each digit a that can occur in a $\mathbf{B}$ -expansion, we are given a first-order formula $\varphi_a(j, n)$ of $\langle \mathbb{N}, +, V_b \rangle$ such that

$$\varphi_a(j, n) \text{ is true} \iff \text{the digit } a \text{ occurs at position } j \text{ in } d_{\sigma^n(\mathbf{B})}^*(1).$$

Then it is decidable whether any given b -automatic sequence over $\mathbb{N}$ is $\mathbf{B}$ -admissible.

The alternate base case

Corollary

Let $b \geq 2$ be an integer, and let $\mathbf{B}$ be an alternate base such that $d_{\sigma^i(\mathbf{B})}^*(1)$ is b -automatic for all i . Then it is decidable whether any given b -automatic sequence over $\mathbb{N}$ is $\mathbf{B}$ -admissible.

A **Parry alternate base** is an alternate base $\mathbf{B}$ such that $d_{\sigma^i(\mathbf{B})}^*(1)$ is ultimately periodic for all i .

Corollary

Let $b \geq 2$ be an integer, and let $\mathbf{B}$ be a Parry alternate base. Then it is decidable whether any given b -automatic sequence over $\mathbb{N}$ is $\mathbf{B}$ -admissible.

Since the finiteness of the transducer $\mathcal{T}_{E,1}^*$ implies that any alternate base $\mathbf{B}$ over E is Parry, we also get:

Corollary

Let $b \geq 2$ be an integer. Suppose that $\mathcal{T}_{E,1}^*$ is finite and let $\mathbf{B}$ be any alternate base over E . Then it is decidable whether any given b -automatic sequence over $\mathbb{N}$ is $\mathbf{B}$ -admissible.

A non-trivial example of decidability with an aperiodic Cantor real base

Proposition

Let $E = \{\beta, \beta^3\}$ and let $\mathbf{B} = (\beta^3, \beta, \beta^3, \beta^3, \beta^3, \beta, \beta^3, \beta^3, \beta, \beta^3, \beta, \beta^3, \beta^3, \beta^3, \beta, \beta^3, \beta, \beta^3, \dots)$ be the Cantor real base obtained as the Thue-Morse sequence over the blocks $(\beta^3, \beta, \beta^3)$ and $(\beta^3, \beta^3, \beta)$.

It is decidable whether any given 2-automatic sequence over $\mathbb{N}$ is $\mathbf{B}$ -admissible.

We used combined tools for proving this result:

- ▶ We took advantage of the 2-automaticity of $\mathbf{B}$ combined with structural properties of the transducer $\mathcal{T}_{E,1}^*$.
- ▶ The two-walk property was used, but it not sufficient. The choice of the blocks $(\beta^3, \beta, \beta^3)$ and $(\beta^3, \beta^3, \beta)$ is also important!
- ▶ This result generalizes to a family of sets E such that the transducer is finite and enjoys some specific (testable) feature.

Thank you!