A strong version of Cobham's theorem

and other thoughts about decidability in expansions of Presburger arithmetic

Philipp Hieronymi

ICALP, Trends in Arithmetic Theories, July 2022

Universität Bonn
Mathematisches Institut

This is (essentially) a follow-up talk about two topics Émilie mentioned in her beautiful talk: Cobhman's theorem and using decidability to prove theorems in combinatorics on words.

This is (essentially) a follow-up talk about two topics Émilie mentioned in her beautiful talk: Cobhman's theorem and using decidability to prove theorems in combinatorics on words.

A logician's point of view.

This is (essentially) a follow-up talk about two topics Émilie mentioned in her beautiful talk: Cobhman's theorem and using decidability to prove theorems in combinatorics on words.

A logician's point of view.

Presburger (1929). The theory of $(\mathbb{N}, +)$ is decidable.

This is (essentially) a follow-up talk about two topics Émilie mentioned in her beautiful talk: Cobhman's theorem and using decidability to prove theorems in combinatorics on words.

A logician's point of view.

Presburger (1929). The theory of $(\mathbb{N}, +)$ is decidable.

Gödel (1931). The theory of $(\mathbb{N}, +, \cdot)$ is undecidable.

This is (essentially) a follow-up talk about two topics Émilie mentioned in her beautiful talk: Cobhman's theorem and using decidability to prove theorems in combinatorics on words.

A logician's point of view.

Presburger (1929). The theory of $(\mathbb{N}, +)$ is decidable.

Gödel (1931). The theory of $(\mathbb{N}, +, \cdot)$ is undecidable.

A classical research program in logic. Study expansions of $(\mathbb{N}, +)$ by fragments of multiplication.

This is (essentially) a follow-up talk about two topics Émilie mentioned in her beautiful talk: Cobhman's theorem and using decidability to prove theorems in combinatorics on words.

A logician's point of view.

Presburger (1929). The theory of $(\mathbb{N}, +)$ is decidable.

Gödel (1931). The theory of $(\mathbb{N}, +, \cdot)$ is undecidable.

A classical research program in logic. Study expansions of $(\mathbb{N}, +)$ by fragments of multiplication.

Putman (1958). The theory of $(\mathbb{N}, +, \{n^2 \ : \ n \in \mathbb{N}\})$ is undecidable.

This is (essentially) a follow-up talk about two topics Émilie mentioned in her beautiful talk: Cobhman's theorem and using decidability to prove theorems in combinatorics on words.

A logician's point of view.

Presburger (1929). The theory of $(\mathbb{N}, +)$ is decidable.

Gödel (1931). The theory of $(\mathbb{N}, +, \cdot)$ is undecidable.

A classical research program in logic. Study expansions of $(\mathbb{N}, +)$ by fragments of multiplication.

Putman (1958). The theory of $(\mathbb{N}, +, \{n^2 : n \in \mathbb{N}\})$ is undecidable.

A closely connected program. Study expansions of $(\mathbb{R}, <, +, \mathbb{N})$ by fragments of multiplication.

A strong version of Cobham's theorem.

Cobham's theorem (1969). Let $k, \ell \geq 2$ be two multiplicatively independent integers. A set $X \subseteq \mathbb{N}$ is both k-recognizable and ℓ-recognizable if and only if it is ultimately periodic.

Cobham's theorem (1969). Let $k, \ell \geq 2$ be two multiplicatively independent integers. A set $X \subseteq \mathbb{N}$ is both k-recognizable and ℓ-recognizable if and only if it is ultimately periodic.

A set $X \subseteq \mathbb{N}$ is k-**recognizable** if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton.

Cobham's theorem (1969). Let $k, \ell \geq 2$ be two multiplicatively independent integers. A set $X \subseteq \mathbb{N}$ is both k-recognizable and ℓ-recognizable if and only if it is ultimately periodic.

A set $X \subseteq \mathbb{N}$ is k-**recognizable** if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton.

Examples.

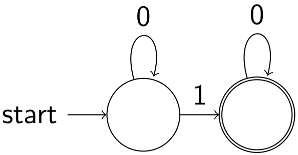
- $k^{\mathbb{N}}$ is k-recognizable,

Cobham's theorem (1969). Let $k, \ell \geq 2$ be two multiplicatively independent integers. A set $X \subseteq \mathbb{N}$ is both k-recognizable and ℓ-recognizable if and only if it is ultimately periodic.

A set $X \subseteq \mathbb{N}$ is k-**recognizable** if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton.

Examples.

- $k^{\mathbb{N}}$ is k-recognizable,

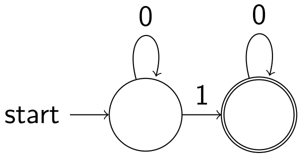


Cobham's theorem (1969). Let $k, \ell \geq 2$ be two multiplicatively independent integers. A set $X \subseteq \mathbb{N}$ is both k-recognizable and ℓ-recognizable if and only if it is ultimately periodic.

A set $X \subseteq \mathbb{N}$ is k-**recognizable** if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton.

Examples.

- $k^{\mathbb{N}}$ is k-recognizable,
- $\{n \in \mathbb{N} \ : \ s_2(n) \text{ is even } \}$ is 2-recognizable, where $s_2(n)$ is the binary digit sum - **Thue-Morse set**.

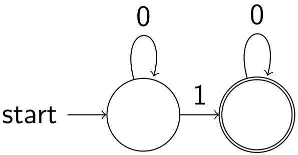


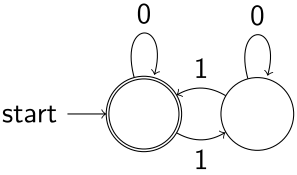
Cobham's theorem (1969). Let $k, \ell \geq 2$ be two multiplicatively independent integers. A set $X \subseteq \mathbb{N}$ is both k-recognizable and ℓ-recognizable if and only if it is ultimately periodic.

A set $X \subseteq \mathbb{N}$ is k-**recognizable** if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton.

Examples.

- $k^{\mathbb{N}}$ is k-recognizable,
- $\{n \in \mathbb{N} \ : \ s_2(n) \text{ is even }\}$ is 2-recognizable, where $s_2(n)$ is the binary digit sum - **Thue-Morse set**.





A set $X \subseteq \mathbb{N}^n$ is k-**recognizable** if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton.

A set $X \subseteq \mathbb{N}^n$ is k-**recognizable** if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton.

Semenov (1977). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. A set $X \subseteq \mathbb{N}^n$ is both k-recognizable and ℓ-recognizable if and only if it is semilinear.

A set $X \subseteq \mathbb{N}^n$ is k-**recognizable** if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton.

Semenov (1977). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. A set $X \subseteq \mathbb{N}^n$ is both k-recognizable and ℓ-recognizable if and only if it is semilinear.

Ginsburg-Spanier (1966). A subset of $\mathbb{N}^n$ is semilinear if and only if it is definable in $(\mathbb{N}, +)$.

A set $X \subseteq \mathbb{N}^n$ is k-**recognizable** if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton.

Semenov (1977). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. A set $X \subseteq \mathbb{N}^n$ is both k-recognizable and ℓ-recognizable if and only if it is semilinear.

Ginsburg-Spanier (1966). A subset of $\mathbb{N}^n$ is semilinear if and only if it is definable in $(\mathbb{N}, +)$.

Büchi arithmetic. For $k \in \mathbb{N}_{\geq 2}$, let $V_k(x) : \mathbb{N} \to k^{\mathbb{N}}$ be the function that maps x to the largest power of k dividing x.

A set $X \subseteq \mathbb{N}^n$ is k-**recognizable** if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton.

Semenov (1977). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. A set $X \subseteq \mathbb{N}^n$ is both k-recognizable and ℓ-recognizable if and only if it is semilinear.

Ginsburg-Spanier (1966). A subset of $\mathbb{N}^n$ is semilinear if and only if it is definable in $(\mathbb{N}, +)$.

Büchi arithmetic. For $k \in \mathbb{N}_{\geq 2}$, let $V_k(x) : \mathbb{N} \to k^{\mathbb{N}}$ be the function that maps x to the largest power of k dividing x.

Büchi(1960)-Bruyère(1985). Let $k \in \mathbb{N}_{\geq 2}$ and $X \subseteq \mathbb{N}^n$. Then X is k-recognizable if and only if X is definable in $(\mathbb{N}, +, V_k)$. Thus, the theory of $(\mathbb{N}, +, V_k)$ is decidable. In particular, for each k-recognizable $X \subseteq \mathbb{N}^n$, the theory of $(\mathbb{N}, +, X)$ is decidable.

A set $X \subseteq \mathbb{N}^n$ is k-**recognizable** if the language consisting of the base-k representations of the elements of X is accepted by a finite automaton.

Semenov (1977). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. A set $X \subseteq \mathbb{N}^n$ is both k-recognizable and ℓ-recognizable if and only if it is semilinear.

Ginsburg-Spanier (1966). A subset of $\mathbb{N}^n$ is semilinear if and only if it is definable in $(\mathbb{N}, +)$.

Büchi arithmetic. For $k \in \mathbb{N}_{\geq 2}$, let $V_k(x) : \mathbb{N} \to k^{\mathbb{N}}$ be the function that maps x to the largest power of k dividing x.

Büchi(1960)-Bruyère(1985). Let $k \in \mathbb{N}_{\geq 2}$ and $X \subseteq \mathbb{N}^n$. Then X is k-recognizable if and only if X is definable in $(\mathbb{N}, +, V_k)$. Thus, the theory of $(\mathbb{N}, +, V_k)$ is decidable. In particular, for each k-recognizable $X \subseteq \mathbb{N}^n$, the theory of $(\mathbb{N}, +, X)$ is decidable.

Cobham-Semenov restated. Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. A set $X \subseteq \mathbb{N}^n$ is definable in both $(\mathbb{N}, +, V_k)$ and $(\mathbb{N}, +, V_\ell)$ if and only if it is definable in $(\mathbb{N}, +)$.

H.-Schulz (2022). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let $X \subseteq \mathbb{N}^m$ and $Y \subseteq \mathbb{N}^n$ be such that

- X is k-recognizable, but not semilinear,
- Y is ℓ-recognizable, but not semilinear.

Then the theory of $(\mathbb{N}, +, X, Y)$ is undecidable.

H.-Schulz (2022). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let $X \subseteq \mathbb{N}^m$ and $Y \subseteq \mathbb{N}^n$ be such that

- X is k-recognizable, but not semilinear,
- Y is ℓ-recognizable, but not semilinear.

Then the theory of $(\mathbb{N}, +, X, Y)$ is undecidable.

H.-Schulz restated. Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let $X \subseteq \mathbb{N}^m$ and $Y \subseteq \mathbb{N}^n$ be such that

- X is definable in $(\mathbb{N}, +, V_k)$, but not in $(\mathbb{N}, +)$,
- Y is definable in $(\mathbb{N}, +, V_\ell)$, but not in $(\mathbb{N}, +)$.

Then the theory of $(\mathbb{N}, +, X, Y)$ is undecidable.

H.-Schulz (2022). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let $X \subseteq \mathbb{N}^m$ and $Y \subseteq \mathbb{N}^n$ be such that

- X is k-recognizable, but not semilinear,
- Y is ℓ-recognizable, but not semilinear.

Then the theory of $(\mathbb{N}, +, X, Y)$ is undecidable.

H.-Schulz restated. Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let $X \subseteq \mathbb{N}^m$ and $Y \subseteq \mathbb{N}^n$ be such that

- X is definable in $(\mathbb{N}, +, V_k)$, but not in $(\mathbb{N}, +)$,
- Y is definable in $(\mathbb{N}, +, V_\ell)$, but not in $(\mathbb{N}, +)$.

Then the theory of $(\mathbb{N}, +, X, Y)$ is undecidable.

Proof of Cobham-Semenov.

Suppose $X \subseteq \mathbb{N}^n$ is definable in both $(\mathbb{N}, +, V_k)$ and $(\mathbb{N}, +, V_\ell)$, but not in $(\mathbb{N}, +)$. Then the theory of $(\mathbb{N}, +, X, X)$ is undecidable. However, then the theory of $(\mathbb{N}, +, X)$ is undecidable.

Villemaire (1992). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. The theory of $(\mathbb{N}, +, V_k, V_\ell)$ is undecidable.

Villemaire (1992). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. The theory of $(\mathbb{N}, +, V_k, V_\ell)$ is undecidable.

Bès (1996). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let Y be definable in $(\mathbb{N}, +, V_\ell)$, but not in $(\mathbb{N}, +)$. Then the theory of $(\mathbb{N}, +, V_k, Y)$ is undecidable.

Villemaire (1992). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. The theory of $(\mathbb{N}, +, V_k, V_\ell)$ is undecidable.

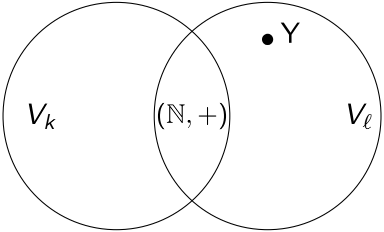
Bès (1996). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let Y be definable in $(\mathbb{N}, +, V_\ell)$, but not in $(\mathbb{N}, +)$. Then the theory of $(\mathbb{N}, +, V_k, Y)$ is undecidable.

In both cases $(\mathbb{N}, +, V_k, V_\ell)$ and $(\mathbb{N}, +, V_k, Y)$ define multiplication. Hence undecidability follows from Gödel's theorem that the theory of $(\mathbb{N}, +, \cdot)$ is undecidable.

Villemaire (1992). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. The theory of $(\mathbb{N}, +, V_k, V_\ell)$ is undecidable.

Bès (1996). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let Y be definable in $(\mathbb{N}, +, V_\ell)$, but not in $(\mathbb{N}, +)$. Then the theory of $(\mathbb{N}, +, V_k, Y)$ is undecidable.

In both cases $(\mathbb{N}, +, V_k, V_\ell)$ and $(\mathbb{N}, +, V_k, Y)$ define multiplication. Hence undecidability follows from Gödel's theorem that the theory of $(\mathbb{N}, +, \cdot)$ is undecidable.



H.-Schulz (2022). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let $X \subseteq \mathbb{N}^m$ and $Y \subseteq \mathbb{N}^n$ be such that

- X is definable in $(\mathbb{N}, +, V_k)$, but not in $(\mathbb{N}, +)$
- Y is definable in $(\mathbb{N}, +, V_\ell)$, but not in $(\mathbb{N}, +)$.

Then the theory of $(\mathbb{N}, +, X, Y)$ is undecidable.

H.-Schulz (2022). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let $X \subseteq \mathbb{N}^m$ and $Y \subseteq \mathbb{N}^n$ be such that

- X is definable in $(\mathbb{N}, +, V_k)$, but not in $(\mathbb{N}, +)$
- Y is definable in $(\mathbb{N}, +, V_\ell)$, but not in $(\mathbb{N}, +)$.

Then the theory of $(\mathbb{N}, +, X, Y)$ is undecidable.

Bès (1996). Let $k \in \mathbb{N}_{\geq 2}$, and let $X \subseteq \mathbb{N}^n$ be definable in $(\mathbb{N}, +, V_k)$, but not in $(\mathbb{N}, +)$. Then $(\mathbb{N}, +, X)$ defines $k^{\mathbb{N}}$.

H.-Schulz (2022). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let $X \subseteq \mathbb{N}^m$ and $Y \subseteq \mathbb{N}^n$ be such that

- X is definable in $(\mathbb{N}, +, V_k)$, but not in $(\mathbb{N}, +)$
- Y is definable in $(\mathbb{N}, +, V_\ell)$, but not in $(\mathbb{N}, +)$.

Then the theory of $(\mathbb{N}, +, X, Y)$ is undecidable.

Bès (1996). Let $k \in \mathbb{N}_{\geq 2}$, and let $X \subseteq \mathbb{N}^n$ be definable in $(\mathbb{N}, +, V_k)$, but not in $(\mathbb{N}, +)$. Then $(\mathbb{N}, +, X)$ defines $k^{\mathbb{N}}$.

Left to show:

Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. Then the theory of $(\mathbb{N}, +, k^{\mathbb{N}}, \ell^{\mathbb{N}})$ is undecidable.

H.-Schulz (2022). Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent, and let $X \subseteq \mathbb{N}^m$ and $Y \subseteq \mathbb{N}^n$ be such that

- X is definable in $(\mathbb{N}, +, V_k)$, but not in $(\mathbb{N}, +)$
- Y is definable in $(\mathbb{N}, +, V_\ell)$, but not in $(\mathbb{N}, +)$.

Then the theory of $(\mathbb{N}, +, X, Y)$ is undecidable.

Bès (1996). Let $k \in \mathbb{N}_{\geq 2}$, and let $X \subseteq \mathbb{N}^n$ be definable in $(\mathbb{N}, +, V_k)$, but not in $(\mathbb{N}, +)$. Then $(\mathbb{N}, +, X)$ defines $k^{\mathbb{N}}$.

Left to show:

Let $k, \ell \in \mathbb{N}_{\geq 2}$ be multiplicatively independent. Then the theory of $(\mathbb{N}, +, k^{\mathbb{N}}, \ell^{\mathbb{N}})$ is undecidable.

This question is an old question. Bruyère, Cherlin and van den Dries asked this question as early as 1985, and it has been restated in the literature many times.

This is kind of unfortunate. Even the theory of $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$ includes many non-trivial number-theoretic statements about 2 and 3.

This is kind of unfortunate. Even the theory of $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$ includes many non-trivial number-theoretic statements about 2 and 3.

Corollary of Baker's theorem on linear forms. For every $m \in \mathbb{N}$, there exists $C(m)$ such that if $n_1, n_2 \in \mathbb{N}$ with $2^{n_1} - 3^{n_2} = m$, then $n_1, n_2 \leq C$.

This is kind of unfortunate. Even the theory of $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$ includes many non-trivial number-theoretic statements about 2 and 3.

Corollary of Baker's theorem on linear forms. For every $m \in \mathbb{N}$, there exists $C(m)$ such that if $n_1, n_2 \in \mathbb{N}$ with $2^{n_1} - 3^{n_2} = m$, then $n_1, n_2 \leq C$.

In $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$:

$$\forall u \exists v \forall x \in k^{\mathbb{N}} \forall y \in \ell^{\mathbb{N}} \ (x \geq v \ \vee \ y \geq v) \rightarrow |x - y| > u.$$

This is kind of unfortunate. Even the theory of $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$ includes many non-trivial number-theoretic statements about 2 and 3.

Corollary of Baker's theorem on linear forms. For every $m \in \mathbb{N}$, there exists $C(m)$ such that if $n_1, n_2 \in \mathbb{N}$ with $2^{n_1} - 3^{n_2} = m$, then $n_1, n_2 \leq C$.

In $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$:

$$\forall u \exists v \forall x \in k^{\mathbb{N}} \forall y \in \ell^{\mathbb{N}} \ (x \geq v \ \vee \ y \geq v) \to |x - y| > u.$$

What does that mean?

You probably can't automatically prove theorems worth a Fields medal.

This is kind of unfortunate. Even the theory of $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$ includes many non-trivial number-theoretic statements about 2 and 3.

Corollary of Baker's theorem on linear forms. For every $m \in \mathbb{N}$, there exists $C(m)$ such that if $n_1, n_2 \in \mathbb{N}$ with $2^{n_1} - 3^{n_2} = m$, then $n_1, n_2 \leq C$.

In $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$:

$$\forall u \exists v \forall x \in k^{\mathbb{N}} \forall y \in \ell^{\mathbb{N}} \ (x \geq v \ \vee \ y \geq v) \to |x - y| > u.$$

What does that mean?

You probably can't automatically prove theorems worth a Fields medal.

Open question. What fragments of the theory of $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$ are decidable?

For simplicity, let $k = 2$ and $\ell = 3$. So consider $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$.

For simplicity, let $k = 2$ and $\ell = 3$. So consider $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$.

Let $\lambda : \mathbb{N} \to 2^{\mathbb{N}}$ map x to the unique element $2^m \in 2^{\mathbb{N}}$ with $2^m \leq x < 2^{m+1}$.

For simplicity, let $k = 2$ and $\ell = 3$. So consider $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$.

Let $\lambda : \mathbb{N} \to 2^{\mathbb{N}}$ map x to the unique element $2^m \in 2^{\mathbb{N}}$ with $2^m \leq x < 2^{m+1}$.

For $y \in 2^{\mathbb{N}}$, define $S(y)$ to be the set of all $x \in 3^{\mathbb{N}}$ such that $\lambda(x - \lambda(x)) = y$.

In words: $S(y)$ is the set of all powers of 3 for which y is the second largest power of 2 that appears in the binary representation of x.

For simplicity, let $k = 2$ and $\ell = 3$. So consider $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$.

Let $\lambda : \mathbb{N} \to 2^{\mathbb{N}}$ map x to the unique element $2^m \in 2^{\mathbb{N}}$ with $2^m \leq x < 2^{m+1}$.

For $y \in 2^{\mathbb{N}}$, define $S(y)$ to be the set of all $x \in 3^{\mathbb{N}}$ such that $\lambda(x - \lambda(x)) = y$.

In words: $S(y)$ is the set of all powers of 3 for which y is the second largest power of 2 that appears in the binary representation of x.

For example: $27 = 16 + 8 + 2 + 1$. So $27 \in S(8)$.

For simplicity, let $k = 2$ and $\ell = 3$. So consider $(\mathbb{N}, +, 2^{\mathbb{N}}, 3^{\mathbb{N}})$.

Let $\lambda : \mathbb{N} \to 2^{\mathbb{N}}$ map x to the unique element $2^m \in 2^{\mathbb{N}}$ with $2^m \leq x < 2^{m+1}$.

For $y \in 2^{\mathbb{N}}$, define $S(y)$ to be the set of all $x \in 3^{\mathbb{N}}$ such that $\lambda(x - \lambda(x)) = y$.

In words: $S(y)$ is the set of all powers of 3 for which y is the second largest power of 2 that appears in the binary representation of x.

For example: $27 = 16 + 8 + 2 + 1$. So $27 \in S(8)$.

Fact. For all $y \in 2^{\mathbb{N}}$, $S(y)$ is finite. However, for all $m, n \in \mathbb{N}$ there is $y \in 2^{\mathbb{N}}$ such that $y > m$ and $|S(y)| > n$.

Main Lemma. Let $m, n \in \mathbb{N}$, let $Z_1, \ldots, Z_m$ be a partition of $\{1, \ldots, n\}$. Then there are $s \in \mathbb{N}_{>0}$ and $t_1 < \cdots < t_n$ such that for $i = 1, \ldots, m$

$$S(2^{s+i}) \cap [3^{t_1}, 3^{t_n}] = \{3^{t_j} \; : j \in Z_i\}.$$

Main Lemma. Let $m, n \in \mathbb{N}$, let $Z_1, \ldots, Z_m$ be a partition of $\{1, \ldots, n\}$. Then there are $s \in \mathbb{N}_{>0}$ and $t_1 < \cdots < t_n$ such that for $i = 1, \ldots, m$

$$S(2^{s+i}) \cap [3^{t_1}, 3^{t_n}] = \{3^{t_j} \ : j \in Z_i\}.$$

$Z_1 = 1, 3, 5,$

$Z_2 = 2, 6, 7,$

$Z_3 = 4$

	3^{t_1}	3^{t_2}	3^{t_3}	3^{t_4}	3^{t_5}	3^{t_6}	3^{t_7}
2^{s+1}							
2^{s+2}							
2^{s+3}							

Main Lemma. Let $m, n \in \mathbb{N}$, let $Z_1, \dots, Z_m$ be a partition of $\{1, \dots, n\}$. Then there are $s \in \mathbb{N}_{>0}$ and $t_1 < \cdots < t_n$ such that for $i = 1, \dots, m$

$$S(2^{s+i}) \cap [3^{t_1}, 3^{t_n}] = \{3^{t_j} \ : j \in Z_i\}.$$

$Z_1 = 1, 3, 5,$

$Z_2 = 2, 6, 7,$

$Z_3 = 4$

	3^{t_1}	3^{t_2}	3^{t_3}	3^{t_4}	3^{t_5}	3^{t_6}	3^{t_7}
2^{s+1}							
2^{s+2}							
2^{s+3}							

- Proof of Main Lemma just uses density of $2^{-\mathbb{N}}3^{\mathbb{N}}$ in $\mathbb{R}_{>0}$.

Main Lemma. Let $m, n \in \mathbb{N}$, let $Z_1, \ldots, Z_m$ be a partition of $\{1, \ldots, n\}$. Then there are $s \in \mathbb{N}_{>0}$ and $t_1 < \cdots < t_n$ such that for $i = 1, \ldots, m$

$$S(2^{s+i}) \cap [3^{t_1}, 3^{t_n}] = \{3^{t_j} \ : j \in Z_i\}.$$

$Z_1 = 1, 3, 5,$

$Z_2 = 2, 6, 7,$

$Z_3 = 4$

	3^{t_1}	3^{t_2}	3^{t_3}	3^{t_4}	3^{t_5}	3^{t_6}	3^{t_7}
2^{s+1}							
2^{s+2}							
2^{s+3}							

- Proof of Main Lemma just uses density of $2^{-\mathbb{N}}3^{\mathbb{N}}$ in $\mathbb{R}_{>0}$.
- This allows us to code/interpret arbitrary large finite subsets of $\mathbb{N}^2$.

Main Lemma. Let $m, n \in \mathbb{N}$, let $Z_1, \dots, Z_m$ be a partition of $\{1, \dots, n\}$. Then there are $s \in \mathbb{N}_{>0}$ and $t_1 < \dots < t_n$ such that for $i = 1, \dots, m$

$$S(2^{s+i}) \cap [3^{t_1}, 3^{t_n}] = \{3^{t_j} \ : j \in Z_i\}.$$

$Z_1 = 1, 3, 5,$

$Z_2 = 2, 6, 7,$

$Z_3 = 4$

	3^{t_1}	3^{t_2}	3^{t_3}	3^{t_4}	3^{t_5}	3^{t_6}	3^{t_7}
2^{s+1}							
2^{s+2}							
2^{s+3}							

- Proof of Main Lemma just uses density of $2^{-\mathbb{N}}3^{\mathbb{N}}$ in $\mathbb{R}_{>0}$.
- This allows us to code/interpret arbitrary large finite subsets of $\mathbb{N}^2$.
- Such theories are known to be undecidable, as the halting problem or the tiling problem can be encoded in such theories.

Schulz (2022). $(\mathbb{N}, +, k^{\mathbb{N}}, \ell^{\mathbb{N}})$ does *not* define multiplication.

Schulz (2022). $(\mathbb{N}, +, k^{\mathbb{N}}, \ell^{\mathbb{N}})$ does *not* define multiplication.

Definition. A set $X \subseteq \mathbb{Z}$ has **zero natural density** if

$$\limsup_{n\to\infty} \frac{|X \cap \{-n, -n+1, \ldots, n-1, n\}|}{2n+1} = 0.$$

Schulz (2022). $(\mathbb{N}, +, k^{\mathbb{N}}, \ell^{\mathbb{N}})$ does *not* define multiplication.

Definition. A set $X \subseteq \mathbb{Z}$ has **zero natural density** if

$$\limsup_{n\to\infty} \frac{|X \cap \{-n, -n+1, \ldots, n-1, n\}|}{2n+1} = 0.$$

The set X is **piecewise syndetic** if there is $b \in \mathbb{N}$ such that there are arbitrarily long intervals of $\mathbb{Z}$ where the gaps in X are bounded by b.

Schulz (2022). $(\mathbb{N}, +, k^{\mathbb{N}}, \ell^{\mathbb{N}})$ does *not* define multiplication.

Definition. A set $X \subseteq \mathbb{Z}$ has **zero natural density** if

$$\limsup_{n\to\infty} \frac{|X \cap \{-n, -n+1, \ldots, n-1, n\}|}{2n+1} = 0.$$

The set X is **piecewise syndetic** if there is $b \in \mathbb{N}$ such that there are arbitrarily long intervals of $\mathbb{Z}$ where the gaps in X are bounded by b.

Schulz (2022). ('Friedman-Miller for Presburger') Let $E \subseteq \mathbb{Z}$ be such that for every $m \in N$ and every $h : \mathbb{Z}^m \to \mathbb{Z}$ definable in $(\mathbb{Z}, <, +)$, the image $h(E^m)$ has zero natural density. Then every subset of $\mathbb{Z}$ definable in $(\mathbb{Z}, <, +, E)^{\#}$ is either piecewise syndetic or has zero natural density.

Schulz (2022). $(\mathbb{N}, +, k^{\mathbb{N}}, \ell^{\mathbb{N}})$ does *not* define multiplication.

Definition. A set $X \subseteq \mathbb{Z}$ has **zero natural density** if

$$\limsup_{n\to\infty} \frac{|X \cap \{-n, -n+1, \ldots, n-1, n\}|}{2n+1} = 0.$$

The set X is **piecewise syndetic** if there is $b \in \mathbb{N}$ such that there are arbitrarily long intervals of $\mathbb{Z}$ where the gaps in X are bounded by b.

Schulz (2022). ('Friedman-Miller for Presburger') Let $E \subseteq \mathbb{Z}$ be such that for every $m \in N$ and every $h : \mathbb{Z}^m \to \mathbb{Z}$ definable in $(\mathbb{Z}, <, +)$, the image $h(E^m)$ has zero natural density. Then every subset of $\mathbb{Z}$ definable in $(\mathbb{Z}, <, +, E)^{\#}$ is either piecewise syndetic or has zero natural density.

Schulz (2022). Let E be the union $2^{\mathbb{N}} \cup 3^{\mathbb{N}}$, and let $f : \mathbb{Z}^m \to \mathbb{Z}$ be definable in $(\mathbb{Z}, <, +)$. Then the image $f(E^n)$ has zero natural density.

Schulz (2022). $(\mathbb{N}, +, k^{\mathbb{N}}, \ell^{\mathbb{N}})$ does *not* define multiplication.

Definition. A set $X \subseteq \mathbb{Z}$ has **zero natural density** if

$$\limsup_{n\to\infty} \frac{|X \cap \{-n, -n+1, \dots, n-1, n\}|}{2n+1} = 0.$$

The set X is **piecewise syndetic** if there is $b \in \mathbb{N}$ such that there are arbitrarily long intervals of $\mathbb{Z}$ where the gaps in X are bounded by b.

Schulz (2022). ('Friedman-Miller for Presburger') Let $E \subseteq \mathbb{Z}$ be such that for every $m \in N$ and every $h : \mathbb{Z}^m \to \mathbb{Z}$ definable in $(\mathbb{Z}, <, +)$, the image $h(E^m)$ has zero natural density. Then every subset of $\mathbb{Z}$ definable in $(\mathbb{Z}, <, +, E)^{\#}$ is either piecewise syndetic or has zero natural density.

Schulz (2022). Let E be the union $2^{\mathbb{N}} \cup 3^{\mathbb{N}}$, and let $f : \mathbb{Z}^m \to \mathbb{Z}$ be definable in $(\mathbb{Z}, <, +)$. Then the image $f(E^n)$ has zero natural density.

Folklore. The set of square-free integers is not piecewise syndetic and does not have zero natural density.

Combinatorics on words.

We observed that $\{n \in \mathbb{N} \ : \ s_2(n) \text{ is even } \}$ is 2-recognizable, where $s_2(n)$ is the binary digit sum. Thus the function $f : \mathbb{N} \to \{0, 1\}$ given by

$$n \mapsto \begin{cases} 0 & \text{if } s_2(n) \text{ is even} \\ 1 & \text{otherwise.} \end{cases}$$

is definable in $(\mathbb{N}, +, V_2)$.

We observed that $\{n \in \mathbb{N} \ : \ s_2(n) \text{ is even } \}$ is 2-recognizable, where $s_2(n)$ is the binary digit sum. Thus the function $f : \mathbb{N} \to \{0,1\}$ given by

$$n \mapsto \begin{cases} 0 & \text{if } s_2(n) \text{ is even} \\ 1 & \text{otherwise.} \end{cases}$$

is definable in $(\mathbb{N}, +, V_2)$.

The word $f(0)f(1)f(2)\dots$ is the **Thue-Morse sequence**.

We observed that $\{n \in \mathbb{N} \ : \ s_2(n) \text{ is even } \}$ is 2-recognizable, where $s_2(n)$ is the binary digit sum. Thus the function $f : \mathbb{N} \to \{0, 1\}$ given by

$$n \mapsto \begin{cases} 0 & \text{if } s_2(n) \text{ is even} \\ 1 & \text{otherwise.} \end{cases}$$

is definable in $(\mathbb{N}, +, V_2)$.

The word $f(0)f(1)f(2)\dots$ is the **Thue-Morse sequence**.

Jeff Shallit's idea. Use decision procedure for $(\mathbb{N}, +, V_2)$ to decide statements about the Thue-Morse sequence.

We observed that $\{n \in \mathbb{N} \;:\; s_2(n) \text{ is even }\}$ is 2-recognizable, where $s_2(n)$ is the binary digit sum. Thus the function $f : \mathbb{N} \to \{0, 1\}$ given by

$$n \mapsto \begin{cases} 0 & \text{if } s_2(n) \text{ is even} \\ 1 & \text{otherwise.} \end{cases}$$

is definable in $(\mathbb{N}, +, V_2)$.

The word $f(0)f(1)f(2)\dots$ is the **Thue-Morse sequence**.

Jeff Shallit's idea. Use decision procedure for $(\mathbb{N}, +, V_2)$ to decide statements about the Thue-Morse sequence.

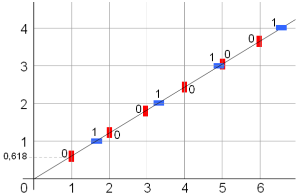
Example. To check that the Thue-Morse sequence in not eventually periodic, we have to decide

$$(\mathbb{N}, +, V_2) \models \forall p \; (p > 0) \to \Big(\forall i \; \exists j \; j > i \;\wedge\; f(j) \neq f(j+p)\Big)$$

Different words.

The **characteristic Sturmian word with slope** a is the infinite $\{0,1\}$-word $\mathbf{c_a} = c_a(0)c_a(1)c_a(2)\dots$ such that for all $n \in \mathbb{N}$

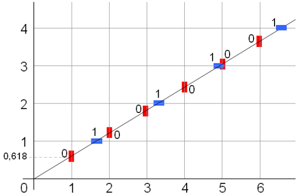
$$c_a(n) = \lfloor a(n+1) \rfloor - \lfloor an \rfloor - \lfloor a \rfloor.$$



Different words.

The **characteristic Sturmian word with slope** a is the infinite $\{0,1\}$-word $\mathbf{c_a} = c_a(0)c_a(1)c_a(2)\ldots$ such that for all $n \in \mathbb{N}$

$$c_a(n) = \lfloor a(n+1) \rfloor - \lfloor an \rfloor - \lfloor a \rfloor.$$

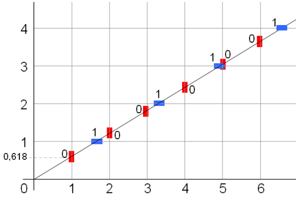


Example. Let $a = 1/\varphi$, where φ is the golden ration. Then $\mathbf{c_a}$ starts with 0100101001.

Different words.

The **characteristic Sturmian word with slope** a is the infinite $\{0,1\}$-word $\mathbf{c_a} = c_a(0)c_a(1)c_a(2)\dots$ such that for all $n \in \mathbb{N}$

$$c_a(n) = \lfloor a(n+1)\rfloor - \lfloor an\rfloor - \lfloor a\rfloor.$$



Example. Let $a = 1/\varphi$, where φ is the golden ration. Then $\mathbf{c_a}$ starts with 0100101001.This word is also called the **Fibonacci word**:

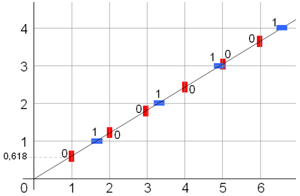
$$S_0 := 0, S_1 := 01, S_2 := 010, S_3 := 01001, S_4 := 01001010, S_5 := 0100101001001\dots$$

Goal. Use decidability to prove theorems about Sturmian words.

Different words.

The **characteristic Sturmian word with slope** a is the infinite $\{0,1\}$-word $\mathbf{c_a} = c_a(0)c_a(1)c_a(2)\dots$ such that for all $n \in \mathbb{N}$

$$c_a(n) = \lfloor a(n+1) \rfloor - \lfloor an \rfloor - \lfloor a \rfloor.$$



Example. Let $a = 1/\varphi$, where φ is the golden ration. Then $\mathbf{c_a}$ starts with 0100101001.This word is also called the **Fibonacci word**:

$$S_0 := 0, S_1 := 01, S_2 := 010, S_3 := 01001, S_4 := 01001010, S_5 := 0100101001001\dots$$

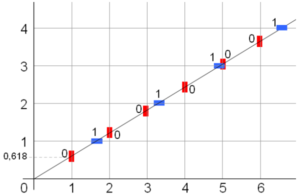
Goal. Use decidability to prove theorems about Sturmian words.

Problem. No Sturmian word with irrational slope is definable in $(\mathbb{N}, +, V_k)$ for any k.

Different words.

The **characteristic Sturmian word with slope** a is the infinite $\{0,1\}$-word $\mathbf{c_a} = c_a(0)c_a(1)c_a(2)\dots$ such that for all $n \in \mathbb{N}$

$$c_a(n) = \lfloor a(n+1) \rfloor - \lfloor an \rfloor - \lfloor a \rfloor.$$



Example. Let $a = 1/\varphi$, where φ is the golden ration. Then $\mathbf{c_a}$ starts with 0100101001.This word is also called the **Fibonacci word**:

$$S_0 := 0, S_1 := 01, S_2 := 010, S_3 := 01001, S_4 := 01001010, S_5 := 0100101001001\dots$$

Goal. Use decidability to prove theorems about Sturmian words.

Problem. No Sturmian word with irrational slope is definable in $(\mathbb{N}, +, V_k)$ for any k.

Solution. Replace k-ary representations by different non-standard representations.

A **continued fraction expansion** $[a_0; a_1, \ldots, a_k, \ldots]$ is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}$$

Let a be a real number with continued fraction expansion $[a_0; a_1, \ldots, a_k, \ldots]$.

A **continued fraction expansion** $[a_0; a_1, \ldots, a_k, \ldots]$ is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}$$

Let a be a real number with continued fraction expansion $[a_0; a_1, \ldots, a_k, \ldots]$.

Set $q_{-1} := 0$ and $q_0 := 1$, and for $k \geq 0$,

$$q_{k+1} := a_{k+1} \cdot q_k + q_{k-1}.$$

A **continued fraction expansion** $[a_0; a_1, \ldots, a_k, \ldots]$ is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}$$

Let a be a real number with continued fraction expansion $[a_0; a_1, \ldots, a_k, \ldots]$.

Set $q_{-1} := 0$ and $q_0 := 1$, and for $k \geq 0$,

$$q_{k+1} := a_{k+1} \cdot q_k + q_{k-1}.$$

Example. $1/\varphi = [0; 1, 1, \ldots]$ and $\sqrt{2} - 1 = [0; 2, 2, \ldots]$ and $\sqrt{3} - 1 = [0; 1, 2, 1, 2, \ldots]$.

A **continued fraction expansion** $[a_0; a_1, \ldots, a_k, \ldots]$ is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}$$

Let a be a real number with continued fraction expansion $[a_0; a_1, \ldots, a_k, \ldots]$.

Set $q_{-1} := 0$ and $q_0 := 1$, and for $k \geq 0$,

$$q_{k+1} := a_{k+1} \cdot q_k + q_{k-1}.$$

Example. $1/\varphi = [0; 1, 1, \ldots]$ and $\sqrt{2} - 1 = [0; 2, 2, \ldots]$ and $\sqrt{3} - 1 = [0; 1, 2, 1, 2, \ldots]$.
So for $a = 1/\varphi$, we get that $q_k = F_k$, where F_k is the k-th Fibonacci number.

A **continued fraction expansion** $[a_0; a_1, \ldots, a_k, \ldots]$ is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}$$

Let a be a real number with continued fraction expansion $[a_0; a_1, \ldots, a_k, \ldots]$.

Set $q_{-1} := 0$ and $q_0 := 1$, and for $k \geq 0$,

$$q_{k+1} := a_{k+1} \cdot q_k + q_{k-1}.$$

Example. $1/\varphi = [0; 1, 1, \ldots]$ and $\sqrt{2} - 1 = [0; 2, 2, \ldots]$ and $\sqrt{3} - 1 = [0; 1, 2, 1, 2, \ldots]$.
So for $a = 1/\varphi$, we get that $q_k = F_k$, where F_k is the k-th Fibonacci number.

Zeckendorf representation (1972). Let $N \in \mathbb{N}$ and F_k be the k-th Fibonacci number. Then N can be written uniquely as

$$N = \sum_{k=1}^{n} b_{k+1} F_k,$$

where $b_k \in \{0, 1\}$ and if $b_{k+1} = 1$, then $b_k = 0$.

Zeckendorf representation (1972). Let $N \in \mathbb{N}$ and F_k be the k-th Fibonacci number. Then N can be written uniquely as

$$N = \sum_{k=1}^{n} b_{k+1} F_k,$$

where $b_k \in \{0, 1\}$ and if $b_{k+1} = 1$, then $b_k = 0$.

Example.

Zeckendorf representation (1972). Let $N \in \mathbb{N}$ and F_k be the k-th Fibonacci number. Then N can be written uniquely as

$$N = \sum_{k=1}^{n} b_{k+1} F_k,$$

where $b_k \in \{0, 1\}$ and if $b_{k+1} = 1$, then $b_k = 0$.

Example. $F_{-1} = 0, F_0 = 1, F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5,\ F_5 = 8, \ldots$

Zeckendorf representation (1972). Let $N \in \mathbb{N}$ and F_k be the k-th Fibonacci number. Then N can be written uniquely as

$$N = \sum_{k=1}^{n} b_{k+1} F_k,$$

where $b_k \in \{0, 1\}$ and if $b_{k+1} = 1$, then $b_k = 0$.

Example. $F_{-1} = 0, F_0 = 1, F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5,\ F_5 = 8,$...

Zeckendorf representation of 9 is $8 + 1$ and not $5 + 3 + 1$.

Zeckendorf representation (1972). Let $N \in \mathbb{N}$ and F_k be the k-th Fibonacci number. Then N can be written uniquely as

$$N = \sum_{k=1}^{n} b_{k+1} F_k,$$

where $b_k \in \{0, 1\}$ and if $b_{k+1} = 1$, then $b_k = 0$.

Example. $F_{-1} = 0, F_0 = 1, F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5,\ F_5 = 8,\ldots$

Zeckendorf representation of 9 is $8 + 1$ and not $5 + 3 + 1$.

Ostrowski (1918). Every natural number N can be written uniquely as

$$N = \sum_{k=0}^{n} b_{k+1} q_k,$$

where $b_k \in \mathbb{N}$ such that $b_1 < a_1$, $b_k \leq a_k$ and, if $b_k = a_k$, $b_{k-1} = 0$.

Let $V_a : \mathbb{N} \to \mathbb{N}$ be the function that maps $x \geq 1$ with Ostrowski representation $b_n \dots b_1$ to the least q_k with $b_{k+1} \neq 0$, and 0 to 1.

Let $V_a : \mathbb{N} \to \mathbb{N}$ be the function that maps $x \geq 1$ with Ostrowski representation $b_n \dots b_1$ to the least q_k with $b_{k+1} \neq 0$, and 0 to 1.

H.-Terry (2016). Let a be quadratic. The first-order logical theory $\mathrm{FO}(\mathbb{N}, +, V_a)$ is decidable.

Let $V_a : \mathbb{N} \to \mathbb{N}$ be the function that maps $x \geq 1$ with Ostrowski representation $b_n \dots b_1$ to the least q_k with $b_{k+1} \neq 0$, and 0 to 1.

H.-Terry (2016). Let a be quadratic. The first-order logical theory $\mathrm{FO}(\mathbb{N}, +, V_a)$ is decidable.

H.-Ma-Oei-Schaeffer-Schulz-Shallit (2021). The first-order logical theory of

$$\{(\mathbb{N}, +, V_a) \ : \ a \in (0,1) \setminus \mathbb{Q}\}$$

is decidable. That is: it is decidable whether statements hold in all structures of the form $(\mathbb{N}, +, V_a)$.

Let $V_a : \mathbb{N} \to \mathbb{N}$ be the function that maps $x \geq 1$ with Ostrowski representation $b_n \dots b_1$ to the least q_k with $b_{k+1} \neq 0$, and 0 to 1.

H.-Terry (2016). Let a be quadratic. The first-order logical theory $\mathrm{FO}(\mathbb{N}, +, V_a)$ is decidable.

H.-Ma-Oei-Schaeffer-Schulz-Shallit (2021). The first-order logical theory of

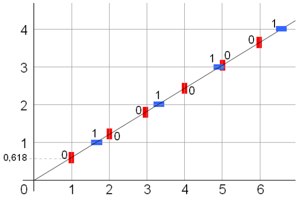
$$\{(\mathbb{N}, +, V_a) \ : \ a \in (0,1) \setminus \mathbb{Q}\}$$

is decidable. That is: it is decidable whether statements hold in all structures of the form $(\mathbb{N}, +, V_a)$.

Proof Strategy. Show that these structures are uniformly ω-automatic. Uses general adder in Ostrowski numeration systems due to Baranwal, Schaeffer and Shallit.

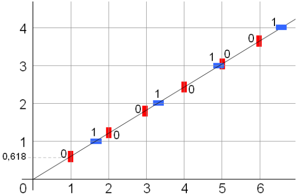
The **characteristic Sturmian word with slope** a is the infinite $\{0,1\}$-word $\mathbf{c_a} = c_a(0)c_a(1)c_a(2)\dots$ such that for all $n \in \mathbb{N}$

$$c_a(n) = \lfloor a(n+1) \rfloor - \lfloor an \rfloor - \lfloor a \rfloor.$$



The **characteristic Sturmian word with slope** a is the infinite $\{0,1\}$-word $\mathbf{c_a} = c_a(0)c_a(1)c_a(2)\dots$ such that for all $n \in \mathbb{N}$

$$c_a(n) = \lfloor a(n+1) \rfloor - \lfloor an \rfloor - \lfloor a \rfloor.$$

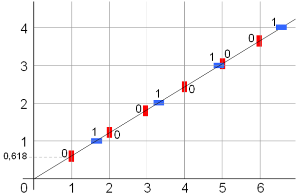


Fact. Let $n \in \mathbb{N}_{\geq 1}$. Then the following are equivalent:

- the n-th digit of the characteristic Sturmian word with slope a is 1.
- the a-Ostrowski representation of n ends with an odd number of 0's.

The **characteristic Sturmian word with slope** a is the infinite $\{0,1\}$-word $\mathbf{c_a} = c_a(0)c_a(1)c_a(2)\ldots$ such that for all $n \in \mathbb{N}$

$$c_a(n) = \lfloor a(n+1) \rfloor - \lfloor an \rfloor - \lfloor a \rfloor.$$



Fact. Let $n \in \mathbb{N}_{\geq 1}$. Then the following are equivalent:

- the n-th digit of the characteristic Sturmian word with slope a is 1.
- the a-Ostrowski representation of n ends with an odd number of 0's.

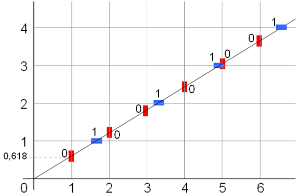
Corollary. The set

$$\{n \in \mathbb{N} \ : \ c_a(n) = 1\}$$

is definable in $(\mathbb{N}, +, V_a)$.

The **characteristic Sturmian word with slope** a is the infinite $\{0,1\}$-word $\mathbf{c_a} = c_a(0)c_a(1)c_a(2)\dots$ such that for all $n \in \mathbb{N}$

$$c_a(n) = \lfloor a(n+1) \rfloor - \lfloor an \rfloor - \lfloor a \rfloor.$$



Fact. Let $n \in \mathbb{N}_{\geq 1}$. Then the following are equivalent:

- the n-th digit of the characteristic Sturmian word with slope a is 1.
- the a-Ostrowski representation of n ends with an odd number of 0's.

Corollary. The set

$$\{n \in \mathbb{N} \ : \ c_a(n) = 1\}$$

is definable in $(\mathbb{N}, +, V_a)$.

H.-Ma-Oei-Schaeffer-Schulz-Shallit (2021). The theory T_{Sturmian} of

$$\{(\mathbb{N}, +, 0, 1, n \mapsto c_a(n)) \ : \ a \in (0,1) \setminus \mathbb{Q} \ \}$$

is decidable.

Let $\mathcal{L}_c$ be the language of $\mathcal{N}_a := (\mathbb{N}, +, 0, 1, n \mapsto c_a(n))$.

Let $\mathcal{L}_c$ be the language of $\mathcal{N}_a := (\mathbb{N}, +, 0, 1, n \mapsto c_a(n))$. Consider the $\mathcal{L}_c$-sentence φ

$$\forall p \; (p > 0) \to \Big(\forall i \; \exists j \; j > i \; \wedge \; c(j) \neq c(j+p)\Big)$$

Let $\mathcal{L}_c$ be the language of $\mathcal{N}_a := (\mathbb{N}, +, 0, 1, n \mapsto c_a(n))$. Consider the $\mathcal{L}_c$-sentence φ

$$\forall p\ (p > 0) \to \Big(\forall i\ \exists j\ j > i\ \wedge\ c(j) \neq c(j + p)\Big)$$

We observe that

$$\mathcal{N}_a \models \varphi \text{ if and only if } \mathbf{c_a} \text{ is not eventually periodic.}$$

Let $\mathcal{L}_c$ be the language of $\mathcal{N}_a := (\mathbb{N}, +, 0, 1, n \mapsto c_a(n))$. Consider the $\mathcal{L}_c$-sentence φ

$$\forall p \; (p > 0) \to \Big(\forall i \; \exists j \; j > i \; \wedge \; c(j) \neq c(j+p)\Big)$$

We observe that

$$\mathcal{N}_a \models \varphi \text{ if and only if } \mathbf{c_a} \text{ is not eventually periodic.}$$

Thus

$$T_{\text{Sturmian}} \models \varphi \text{ if and only if all Sturmian words are not evenutually periodic.}$$

Let $\mathcal{L}_c$ be the language of $\mathcal{N}_a := (\mathbb{N}, +, 0, 1, n \mapsto c_a(n))$. Consider the $\mathcal{L}_c$-sentence φ

$$\forall p \ (p > 0) \to \Big(\forall i \ \exists j \ j > i \ \wedge \ c(j) \neq c(j+p)\Big)$$

We observe that

$$\mathcal{N}_a \models \varphi \text{ if and only if } \mathbf{c_a} \text{ is not eventually periodic.}$$

Thus

$$T_{\text{Sturmian}} \models \varphi \text{ if and only if all Sturmian words are not evenututally periodic.}$$

The decision procedure for T_{Sturmian} allows us to check that no Sturmian word is eventually periodic.

Given a word $w \in \{0,1\}^*$, let w^R be the reverse of w. We say w is a **palindrome** if $w = w^R$.

Given a word $w \in \{0,1\}^*$, let w^R be the reverse of w. We say w is a **palindrome** if $w = w^R$.

Consider the following $\mathcal{L}_c$-formula $\psi(i,n)$:

$$\forall j(j \leq n \to c(i+j) = c(i+n-j)).$$

Given a word $w \in \{0,1\}^*$, let w^R be the reverse of w. We say w is a **palindrome** if $w = w^R$.

Consider the following $\mathcal{L}_c$-formula $\psi(i,n)$:

$$\forall j(j \leq n \rightarrow c(i+j) = c(i+n-j)).$$

Then for $i,n \in \mathbb{N}$

$$\mathcal{N}_a \models \psi(i,n) \text{ if and only if } c_a(i)\ldots c_a(i+n) \text{ is a palindrome.}$$

Given a word $w \in \{0,1\}^*$, let w^R be the reverse of w. We say w is a **palindrome** if $w = w^R$.

Consider the following $\mathcal{L}_c$-formula $\psi(i,n)$:

$$\forall j(j \leq n \to c(i+j) = c(i+n-j)).$$

Then for $i, n \in \mathbb{N}$

$$\mathcal{N}_a \models \psi(i,n) \text{ if and only if } c_a(i)\dots c_a(i+n) \text{ is a palindrome.}$$

Thus

$$T_{\text{Sturmian}} \models \forall n > 0 \exists i \ \psi(i,n)$$

if and only if

every characteristic Sturmian word contains palindromes of every length.

For $w \in \{0,1\}^*$, let $\overline{w}$ be the $\{0,1\}$-word obtained by replacing each 1 in w by 0 and each 0 in w by 1. We say v is an **antisquare** if $v = w\overline{w}$.

For $w \in \{0,1\}^*$, let $\overline{w}$ be the $\{0,1\}$-word obtained by replacing each 1 in w by 0 and each 0 in w by 1. We say v is an **antisquare** if $v = w\overline{w}$.

Consider the following $\mathcal{L}_c$-formula $\chi(i,n)$:

$$\forall j (j \leq n \rightarrow c(i+j) \neq c(i+n+1+j)).$$

For $w \in \{0,1\}^*$, let $\overline{w}$ be the $\{0,1\}$-word obtained by replacing each 1 in w by 0 and each 0 in w by 1. We say v is an **antisquare** if $v = w\overline{w}$.

Consider the following $\mathcal{L}_c$-formula $\chi(i, n)$:

$$\forall j (j \leq n \rightarrow c(i+j) \neq c(i+n+1+j)).$$

Then for $i, n \in \mathbb{N}$

$$\mathcal{N}_a \models \chi(i, n) \text{ if and only if } c_a(i) \ldots c_a(i+2n) \text{ is an antisquare.}$$

For $w \in \{0,1\}^*$, let $\overline{w}$ be the $\{0,1\}$-word obtained by replacing each 1 in w by 0 and each 0 in w by 1. We say v is an **antisquare** if $v = w\overline{w}$.

Consider the following $\mathcal{L}_c$-formula $\chi(i,n)$:

$$\forall j (j \leq n \to c(i+j) \neq c(i+n+1+j)).$$

Then for $i, n \in \mathbb{N}$

$$\mathcal{N}_a \models \chi(i,n) \text{ if and only if } c_a(i) \dots c_a(i+2n) \text{ is an antisquare.}$$

Thus

$$T_{\text{Sturmian}} \models \exists m \forall i \forall n \, (\chi(i,n) \to n \leq m)$$

if and only if

every characteristic Sturmian word contains finitely many antisquares.

An implementation: **Pecan**

- Try Pecan at `http://reedoei.com/pecan`
- Git: `https://github.com/ReedOei/Pecan`

Pecan improves on **Walnut** by Mousavi, another automated theorem prover for deciding combinatorial properties of automatic words, by using Büchi automata instead of finite automata.

This difference enables Pecan to handle uncountable families of sequences, allowing us quantify over all Sturmian words.

We used Pecan to re-prove theorems from papers about Sturmian words (papers as recent as 2020) and have established first new results. Currently working to prove a conjecture of Jason Bell.

				Max		Final	
Name	Quant.	At.	Runtime	States	Edges	States	Edges
Mirror invariant	$\exists$	1	8.1	1440	16840	1129	9666
Unbordered	$\exists^3$	2	0.5	275	1156	92	410
Cube	$\exists$	4	0.7	936	5956	126	561
Least period	$\forall$	4	2605.2	352577	6098198	577	4161
Max unb. subf.	$\forall$	4	26.4	25200	196575	585	4345
Palindrome	$\exists^2$	4	5.1	1934	12337	922	6274
Period	$\exists^2$	5	64.1	5853	103886	1660	17570
Recurrent	$\forall\exists$	5	272.6	61713	960207	34	212
Special factor	$\exists^3\forall$	8	1361.8	17738	103274	4594	25349
Factor Lt (idx)	$\exists\forall^2$	11	702.7	1057221	22348882	2204	25026
Ev. periodic	$\exists^2\forall\exists^2$	12	216.6	78338	1001075	1	0
Reverse factor	$\exists\forall^2$	12	842.0	1408050	22780414	1440	16840
Antipalindrome	$\exists^2\forall^3$	13	242.2	78396	1668960	200	834
Antisquare	$\forall^3$	13	1844.3	2542937	31570114	136	539
Square	$\forall^3$	13	2138.0	1908657	23683717	155	747
$(01)^*\|(10)^*$	$\forall$	16	77.9	5409	72739	103	456

- Of course, Pecan can also be used to produce part of a proof rather than just proving first-order statements.

- Of course, Pecan can also be used to produce part of a proof rather than just proving first-order statements.
- Decision procedure is more general than what is presented here. For example, we can check whether a given statement holds for a dense subset of Sturmian words.

- Of course, Pecan can also be used to produce part of a proof rather than just proving first-order statements.
- Decision procedure is more general than what is presented here. For example, we can check whether a given statement holds for a dense subset of Sturmian words.
- We can also check whether statements hold for all elements in certain subsets of irrational numbers:

- Of course, Pecan can also be used to produce part of a proof rather than just proving first-order statements.
- Decision procedure is more general than what is presented here. For example, we can check whether a given statement holds for a dense subset of Sturmian words.
- We can also check whether statements hold for all elements in certain subsets of irrational numbers:
    1. the set of all irrational a such that the terms in the continued fraction expansion of a are powers of 2,

- Of course, Pecan can also be used to produce part of a proof rather than just proving first-order statements.
- Decision procedure is more general than what is presented here. For example, we can check whether a given statement holds for a dense subset of Sturmian words.
- We can also check whether statements hold for all elements in certain subsets of irrational numbers:
    1. the set of all irrational a such that the terms in the continued fraction expansion of a are powers of 2,
    2. the set of all irrational a such that the terms in the continued fraction expansion of a are not in some fixed finite set, and

- Of course, Pecan can also be used to produce part of a proof rather than just proving first-order statements.
- Decision procedure is more general than what is presented here. For example, we can check whether a given statement holds for a dense subset of Sturmian words.
- We can also check whether statements hold for all elements in certain subsets of irrational numbers:
    1. the set of all irrational a such that the terms in the continued fraction expansion of a are powers of 2,
    2. the set of all irrational a such that the terms in the continued fraction expansion of a are not in some fixed finite set, and
    3. the set of all irrational a such that all even terms in their continued fraction expansion are 1.

- Of course, Pecan can also be used to produce part of a proof rather than just proving first-order statements.
- Decision procedure is more general than what is presented here. For example, we can check whether a given statement holds for a dense subset of Sturmian words.
- We can also check whether statements hold for all elements in certain subsets of irrational numbers:
  1. the set of all irrational a such that the terms in the continued fraction expansion of a are powers of 2,
  2. the set of all irrational a such that the terms in the continued fraction expansion of a are not in some fixed finite set, and
  3. the set of all irrational a such that all even terms in their continued fraction expansion are 1.

What does that mean?

While you probably can't automatically prove theorems worth a Fields medal, there are interesting non-trivial mathematical theorems Pecan can handle for you.

- Of course, Pecan can also be used to produce part of a proof rather than just proving first-order statements.
- Decision procedure is more general than what is presented here. For example, we can check whether a given statement holds for a dense subset of Sturmian words.
- We can also check whether statements hold for all elements in certain subsets of irrational numbers:
  1. the set of all irrational a such that the terms in the continued fraction expansion of a are powers of 2,
  2. the set of all irrational a such that the terms in the continued fraction expansion of a are not in some fixed finite set, and
  3. the set of all irrational a such that all even terms in their continued fraction expansion are 1.

What does that mean?

While you probably can't automatically prove theorems worth a Fields medal, there are interesting non-trivial mathematical theorems Pecan can handle for you.