(1.)

On Numeration Systems and Automatic Sequences

Goals:

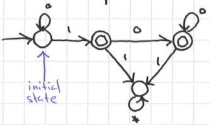
- Introduce k-automatic sequences.
- Decidable properties of them.
- Generalizations and their limits.

(2.)

preliminaries

alphabet A, say $A = \{0,1\}$
$A^* = \{\varepsilon, 0, 1, 00, 10, \dots, 1001, \dots\}$ words over A
language $\mathcal{L} \subseteq A^*$

language $\mathcal{L} \subseteq A^*$ is regular if it is accepted by a DFA (deterministic finite automaton)

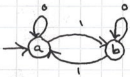


◎ = final
O = non-final

accepts words over $\{0,1\}$ with exactly one 1

k-automatic sequences

k-automatic sequence = sequence obtained by feeding the base-k representations of integers to a DFAO (DFA with output)



[redacted]

$rep_2(n)$ (base-2):	0,	1,	10,	11,	100,	101,	110,	...
x_n:	a,	b,	b,	a,	b,	a,	a,	...

counts if there is even number of 1's in the representation
Thue-Morse word, 2-automatic

- robust class of words of low complexity
- can be studied using automata theory, logic
- Allouche, Shallit: Automatic Sequences

③

Properties of k-automatic sequences

Important decidability result:

Thm (Charlier-Rampersad-Shallit 2012) Let $k \geq 2$. If we can express a property of a k-automatic sequence x using quantifiers, logical operations, integer variables, the operations of addition, subtraction, indexing into x and comparison of integers or elements of x, then this property is decidable.

Reformulation of Büchi's Theorem: first-order theory of $\langle \mathbb{N}, +, V_k \rangle$ is decidable.

$V_k(x)$ = largest power of k dividing x

Application

Detecting overlaps: abaabaa aaaaa

$$(\exists i \geq 0)(\exists \ell \geq 1)(\forall j)(j \leq \ell \rightarrow x[i+j] = x[i+\ell+j])$$

position   length of repetition

Thue (1906): true for the Thue-Morse word

Shallit et al. (2010–): Can we use the decision procedure in practice?

YES! (surprisingly often)

Walnut (Mousavi): Software implementing the decision procedure.

Proves that the Thue-Morse word is overlap-free in ~1s.

(4.)

closure properties

$X \subseteq \mathbb{N}$ is k-recognizable if $\mathrm{rep}_k(X)$ is regular

$X \subseteq \mathbb{N}$ is k-definable if there exists a formula $\varphi(x)$ in $\langle \mathbb{N}, +, V_k \rangle$ such that $X = \{n \in \mathbb{N} : \langle \mathbb{N}, +, V_k \rangle \models \varphi(n)\}$

Thm. (Büchi 1960) Let $k \geq 2$. $X \subseteq \mathbb{N}$ is k-definable iff X is k-recognizable.

core facts: $\mathrm{rep}_k(\{(x, y, x+y) : x, y \in \mathbb{N}\})$ regular, same for $=$, V_k

Thm. $x = x_0 x_1 \cdots \in \mathcal{A}^{\mathbb{N}}$ is k-automatic iff $\mathrm{fiber}_a(x) := \{i \in \mathbb{N} : x_i = a\}$ is k-definable $\forall a \in \mathcal{A}$.

We have the following closure property.

Prop. Let $k \geq 2$, $c > d \geq 0$. If $x = (x_i)_{i \geq 0} \in \mathcal{A}^{\mathbb{N}}$ is k-automatic, then $y = (y_{ci+d})_{i \geq 0}$ is k-automatic.

Proof. $\mathrm{fiber}_a(x)$ is k-definable by $\varphi_{x,a}$ $\forall a$

$\Rightarrow$ $\mathrm{fiber}_a(y)$ —"— $\varphi_{y,a}(i) := \varphi_{x,a}(ci + d)$ $\forall a$

$\uparrow$ $ci = \underbrace{i + \dots + i}_{c \text{ times, constant}}$

$\overset{\text{Thm}}{\Rightarrow}$ y k-automatic

$\square$

5.

questions

- we can represent integers in noncanonical bases, can we generalize the results so far to such cases?

6.

Positional numeration systems

A positional numeration system is an increasing sequence $U = (U_n)_{n\geq 0}$ of integers s.t.

- $U_0 = 1$
- $C_U := \sup_{n\geq 0} \left\lceil \frac{U_{n+1}}{U_n} \right\rceil < \infty$

We represent n greedily as a word in $\{0, 1, \dots, C_U - 1\}^*$, as in the base-$k$ case.

$\text{rep}_U(n)$

Ex. $U = (1, 2, 3, 5, 8, \dots)$ (Fibonacci numbers)

$7 = 5 + 2 = U_3 + U_1 = 1 \cdot U_3 + 0 \cdot U_2 + 1 \cdot U_1 + 0 \cdot U_0$

$\Rightarrow \text{rep}_U(7) = 1010$

also: $7 = 3 + 2 + 2 = 1 \cdot U_2 + 2 \cdot U_1 + 0 \cdot U_0$

$\Rightarrow 120$, not greedy

$\dots 011 \dots \longrightarrow \dots 100 \dots$

not greedy

$\text{rep}_U(\mathbb{N}) = 1\{0, 01\}^* \cup \{0\}$ (regular)

0:	0	4:	101	8:	10000
1:	1	5:	1000	9:	10001
2:	10	6:	1001	10:	10010
3:	100	7:	1010	11:	10100

(count as in base 2, but skip when we have 11)

β-expansions

Let us see next how to associate a "good" numeration system with a real $\beta > 1$.

$T_\beta : [0,1] \to [0,1], \ x \mapsto \beta x \bmod 1$

set $x_i = \lfloor \beta T_\beta^{i-1}(x) \rfloor$ for $i \geq 1$, $x_i \in \{0, \dots, \lceil \beta \rceil - 1\}$

(x_i) codes the orbit of x and $x = \sum_{i=1}^{\infty} \frac{x_i}{\beta^i}$ (also obtained by taking the greedy expansion of this form)

$d_\beta(x) := (x_i)_{i \geq 1}$, β-expansion of x

Ex. $\beta = \frac{1+\sqrt{5}}{2}$, $\beta^2 - \beta - 1 = 0$, $x = 1$

$$\begin{aligned}
x_1 &= \lfloor \beta \rfloor = 1 \\
T_\beta(x) &= \beta - 1 \\
x_2 &= \lfloor \beta(\beta-1) \rfloor = \lfloor \beta^2 - \beta \rfloor = 1 \\
T_\beta^2(x) &= \beta(\beta - 1) - 1 = 0 \\
x_3 &= 0
\end{aligned}$$

$\Rightarrow d_\beta(1) = 1100\dots$

Remark: β integer $\Rightarrow d_\beta(x)$ is expansion of x in base β

Def. If $d_\beta(1)$ is ultimately periodic, then β is a Parry number. If $d_\beta(1) = t_1 \cdots t_m 00\dots$ with $t_m \neq 0$, then we set $d_\beta^*(1) = (t_1 \cdots t_{m-1}(t_m - 1))^\omega$, otherwise $d_\beta^*(1) = d_\beta(1)$.

let β be a Parry number and $d_\beta^*(1) = t_1 t_2 \cdots$

set
$$U_n = t_1 U_{n-1} + \dots + t_n U_0 + 1, \quad n \geq 0$$

Parry numeration system $U := (U_n)$ (associated with β)

8.

Why is a Parry numeration system "good"?

Answer: because then $rep_\beta(\mathbb{N})$ is regular

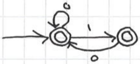
Thm. (Parry 1960) Word w is a greedy expansion of an integer iff $w \leq_{lex} \sigma^k(d_\beta^*(1))$ $\forall k$.

↑ explain

β Parry number $\Rightarrow$ $d_\beta^*(1)$ contains finite information

$\Rightarrow$ checkable by a DFA

Ex. $\beta = \frac{1+\sqrt{5}}{2}$, $d_\beta(1) = 110\ldots$, $d_\beta^*(1) = (10)^\omega$



This is the earlier Fibonacci numeration system (same recurrence).

How to find rep(N)?

A: Take the language accepted by the automaton (without leading 0's) and order in radix order.

Parry-automatic sequence = sequence obtained by feeding the representations of integers in a Parry numeration system to a DFA

use above DFA as DFAO, get the Fibonacci word

abaababaaba ...

(9.)

Pisot numeration systems

Def. An algebraic number is Pisot if its conjugates have modulus <1.

Ex. $\frac{1+\sqrt{5}}{2}$ is Pisot, conjugate $\frac{1-\sqrt{5}}{2} \approx -0{,}61$

Thm. (Schmidt 1980) Pisot numbers are Parry numbers.

$\Rightarrow$ We have a numeration system associated with every Pisot number.

Thm. (Bruyère, Hansel 1997) Büchi's Thm for Pisot numeration systems: $\langle \mathbb{N}, +, V_u \rangle$ decidable.

$V_U(x)$ = largest base element U_n dividing x

$V_U(x) = y \Leftrightarrow rep(x) = a_i \cdots a_j 0^j, \ a_j \neq 0, \ y = U_j$

$\Rightarrow$ Thm of Charlier–Rampersad–Shallit for Pisot-automatic words

$\Rightarrow$ We can study the Fibonacci word with first-order logic.

Works in Walnut!

$\Rightarrow$ We get the same closure properties for Pisot-automatic words.

CONCLUSION: Pisot as "good" as k-automatic.

Extension beyond Pisot

Does this extend for non-Pisot Parry-automatic sequences?

NO (Frougny et al. : addition not necessarily regular)

In a sense: Pisot largest generalization for which we have the "good" properties of k-automatic sequences.

Next: Explicit example that the closure properties break for general Parry-automatic sequences.

(So that audience gets a feeling how such a result could be proved.)

Joint work with A. Massuir and M. Rigo from the University of Liège.

Explicit example

$U := (U_n)$, $U_n = 3U_{n-1} + 2U_{n-2} + 3U_{n-3}$ (4 th degree)
$U_0 = 1$, $U_1 = 4$, $U_2 = 15$, $U_3 = 54$

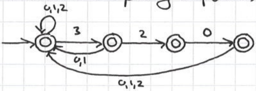
roots of char. polynomial: $\beta \approx 3{,}616$, $\gamma \approx -1{,}097$
$\alpha, \bar{\alpha}$, $|\alpha|, |\bar{\alpha}| < 1$

char. polynomial min. polynomial for β
$\Rightarrow$ β not Pisot

β is Parry number: $d_\beta(1) = 32030^\omega$

$\Rightarrow$ U is a Parry numeration system (recurrence and initial values match with the Parry construction)

automaton accepting $rep_U(\mathbb{N})$:



$d_\beta^*(1) = (3202)^\omega$

0	20
1	21
2	22
3	23
10	30
11	31
12	32
13	100

characteristic sequence of (U_n):

$x = 0100\,10\,00\,00\,00\,00\,01\,00 \ldots$

x is U-automatic (we need to check if the input is of the form $10\ldots0$)

delete every second letter of x:

$y = 00\,10\,00 \ldots$

Claim: y not U-automatic (i.e. closure property breaks)

(12.) y U-automatic $\Rightarrow \mathcal{L} := \{ \mathrm{rep}_U(U_n/2) : U_n \text{ even} \}$ regular

Claim: $\mathrm{rep}_U(U_n/2)$ converges to $d_\beta(\frac{1}{2})$ as $n \to \infty$

Proof: recurrence relation $\Rightarrow U_{n-1} < U_n/2 < U_n$, $n > 1$

$\Rightarrow \mathrm{rep}_U(U_n/2)$ word of length n

say $\mathrm{rep}_U(U_n/2) = d_1 \cdots d_k d_{k+1} \cdots d_n$

extremal values for $d_{k+1} \cdots d_n$: 0^{n-k}, $\mathrm{rep}_U(U_{n-k-1})$

$\Rightarrow \quad 0 \leq \frac{U_n}{2} - d_1 U_{n-1} - \dots - d_k U_{n-k} < U_{n-k} \quad | : U_n$

$\xrightarrow{n \to \infty} \quad 0 \leq \frac{1}{2} - \frac{d_1}{\beta} - \dots - \frac{d_k}{\beta^k} < \frac{1}{\beta^k} \quad$ (recall: $U_n \sim \beta^n$)

$\Rightarrow d_\beta(\frac{1}{2}) = d_1 \cdots d_k \cdots$ $\square$

Claim: $d_\beta(\frac{1}{2})$ ultimately periodic

Proof: $\mathcal{L}$ regular, contains xy^jz $\forall j \geq 0$ (Pumping Lemma)

$\overset{\text{claim}}{\Rightarrow}$ $d_\beta(\frac{1}{2})$ has prefix xy^j $\forall j \geq 0$

$\Rightarrow d_\beta(\frac{1}{2}) = xy^\omega$ $\square$

(13.) write $\underbrace{\frac{1}{2} = \sum_{i=1}^{\infty} \frac{d_i}{\beta^i}}$, (d_i) ultimately periodic

can be written as a polynomial in β
(because (d_i) is ultimately periodic)

$$\Rightarrow \frac{1}{2} = \sum_{i=1}^{\infty} \frac{d_i}{\gamma^i}$$ (since β, γ are conjugates), $\gamma \approx -1{,}09$

$$= \underbrace{\sum_{i=1}^{21} \frac{d_i}{\gamma^i}}_{< -2{,}20} + \underbrace{\sum_{i=22}^{\infty} \frac{d_i}{\gamma^i}}_{\leq 3\left(\frac{1}{\gamma^{22}} + \frac{1}{\gamma^{24}} + \dots\right) < 2{,}33}, \quad d_i \in \{0,1,2,3\}$$

$< 0{,}13$ $\lightning$