Automatic sequences in rational base numeration systems (and even more)
Joint work Michel Rigo (ULiège)

Manon Stipulanti (ULiège) FNRS Postdoc Researcher

ULiège Discrete Math. Seminar
Liège (Belgium)
March 31, 2021
How it started

Eric Rowland presented our paper


at the *One World Combinatorics on Words Seminar.* He (notably) talked about the word

$$w_{3/2} = 0011021001120011031001130011021001140011031 \cdots ,$$

which is the lexicographically least word avoiding $3/2$-powers on the alphabet of non-negative integers. Then Michel asked:

Is this word $3/2$-regular?

So we started to look at automaticity and regularity in rational bases.
The plan

Define:

- Abstract numeration systems
- Labeled trees and periodic labeled signatures
- Rational base numeration systems
- Automatic sequences

Two main results:

- A version of Cobham’s theorem
- “Factor complexity” in trees
Abstract numeration systems

Introduced by Pierre Lecomte and Michel Rigo in 2001, an abstract numeration system (ANS for short) is a triple $S = (L, A, <)$ where

- $(A, <)$ is a totally ordered (finite) alphabet
- $L$ is an infinite language over $A$.

We say that $L$ is the numeration language.

The map $\text{rep}_S : \mathbb{N} \to L$ is the one-to-one correspondence mapping $n \in \mathbb{N}$ onto the $(n+1)$st word in the radix ordered language $L$. This word is called the $S$-representation of $n$. (The $S$-representation of 0 is the first word in $L$.)

The inverse map is denoted by $\text{val}_S : L \to \mathbb{N}$. For any word $w$ in $L$, $\text{val}_S(w)$ is its $S$-(numerical) value.

Remark: Nothing is assumed on the language $L$, i.e., it might well be neither regular nor prefix-closed.
Examples of ANS

• Integer base numeration systems:
  Let $k \geq 2$ be an integer. We let $A_k = \{0, 1, \ldots, k - 1\}$ and

  $L_k = \{\varepsilon\} \cup \{1, \ldots, k - 1\}\{0, \ldots, k - 1\}^*$.

  Then $S = (L_k, A_k, <)$ with $0 < 1 < \cdots < k - 1$ is an ANS.

• Fibonacci (or Zeckendorff) numeration system:
  Let $S = (L, \{0, 1\}, <)$ with $0 < 1$ and $L$ containing words avoiding 11.
  In radix order: $L = \{\varepsilon, 1, 10, 100, 101, 1000, 1001, 1010, \ldots\}$

<table>
<thead>
<tr>
<th>$S$-representation</th>
<th>$S$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>101</td>
<td>4</td>
</tr>
<tr>
<td>1000</td>
<td>5</td>
</tr>
</tbody>
</table>
A “non-standard” ANS

$S = (a^*b^*, \{a, b\}, <)$ with $a < b$

$a^*b^* = \{\varepsilon, a, b, aa, ab, bb, aaa, aab, abb, \ldots\}$ in radix order

<table>
<thead>
<tr>
<th>$S$-representation</th>
<th>$S$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$</td>
<td>0</td>
</tr>
<tr>
<td>$a$</td>
<td>1</td>
</tr>
<tr>
<td>$b$</td>
<td>2</td>
</tr>
<tr>
<td>$aa$</td>
<td>3</td>
</tr>
<tr>
<td>$ab$</td>
<td>4</td>
</tr>
<tr>
<td>$bb$</td>
<td>5</td>
</tr>
<tr>
<td>$aaa$</td>
<td>6</td>
</tr>
<tr>
<td>$aab$</td>
<td>7</td>
</tr>
</tbody>
</table>

Position of a word in $a^*b^*$: $\text{val}_S(a^pb^q) = \frac{(p+q)(p+q+1)}{2} + q$

For instance, $\text{val}_S(ab) = \frac{2 \cdot 3}{2} + 1 = 4$. 

Automaticity in rational bases

Manon Stipulanti (ULiège)
Prefix-closed languages define labeled trees.

Let $S = (L, A, <)$ be an ANS where $L$ is prefix-closed. We define the tree $T(L)$ as follows.

- The set of nodes of $T(L)$ is $L$.
- If $w$ and $wd$ are words in $L$ with $d \in A$, then there is an edge from $w$ to $wd$ with label $d$ in $T(L)$.

The children of a node are ordered by the labels of the letters in the ordered alphabet $A$.

Nodes are enumerated by breadth-first traversal.

Example: $S = (a^*b^*, \{a, b\}, <)$ with $a < b$

$a^*b^* = \{\varepsilon, a, b, aa, ab, bb, aaa, \ldots\}$
Let $T$ be a labeled tree.

The **signature** of $T$ is the sequence of the degrees of the nodes visited by the breadth-first traversal of $T$.

The **labeling** of $T$ is the sequence of the labels of the edges visited by the breadth-first traversal of $T$.

**Example:**

![Diagram of a labeled tree]

**Signature:**

2, 2, 1, 2, 1, 1, 2, 1, 1, 2, …

**Labeling:**

$a, b, a, b, b, a, b, b, b, a, b, …$

**Remark:** Sometimes it is convenient to consider **i-trees**: the root is assumed to be a child of itself.
Let \( p \) and \( q \) be two relatively prime integers with \( p > q > 1 \).

For a word \( w = w_\ell w_{\ell-1} \cdots w_0 \in A_p^* \), the value of \( w \) in base \( \frac{p}{q} \) is the rational number

\[
\text{val}_{\frac{p}{q}}(w) = \sum_{i=0}^{\ell} \frac{w_i}{q} \left( \frac{p}{q} \right)^i.
\]

Example: \( p = 3, \ q = 2 \) \( A_3 = \{0, 1, 2\} \)

<table>
<thead>
<tr>
<th>( w \in {0, 1, 2}^* )</th>
<th>( \text{val}_{\frac{3}{2}}(w) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon, 0 )</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>( \frac{1}{2} \cdot \left( \frac{3}{2} \right)^0 = \frac{1}{2} )</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{2}{2} \cdot \left( \frac{3}{2} \right)^0 = 1 )</td>
</tr>
<tr>
<td>10</td>
<td>( \frac{1}{2} \cdot \left( \frac{3}{2} \right)^1 + \frac{0}{2} \cdot \left( \frac{3}{2} \right)^0 = \frac{3}{4} )</td>
</tr>
<tr>
<td>21</td>
<td>( \frac{2}{2} \cdot \left( \frac{3}{2} \right)^1 + \frac{1}{2} \cdot \left( \frac{3}{2} \right)^0 = 2 )</td>
</tr>
</tbody>
</table>

Remark: \( \text{val}_{\frac{p}{q}}(w) \) is not always an integer.
Definition: A word $w \in A_p^*$ is a representation of an integer $n \geq 0$ in base $\frac{p}{q}$ if $\text{val}_{\frac{p}{q}}(w) = n$.

**Theorem** (Akiyama, Frougny, Sakarovitch, 2008)
Representations in rational bases are unique up to leading zeroes.

In base $\frac{p}{q}$:

$\text{rep}_{\frac{p}{q}}(n)$ denotes the representation of $n$ that does not start with 0.

By convention, the representation of 0 is the empty word $\varepsilon$.

The numeration language is the set

$$L_{\frac{p}{q}} = \left\{ \text{rep}_{\frac{p}{q}}(n) \mid n \geq 0 \right\}.$$

Example: In base $\frac{3}{2}$:

$$L_{\frac{3}{2}} = \{ \varepsilon, 2, 21, 210, 212, 2101, 2120, 2122, \ldots \}$$
Properties:

- For all $u, v \in A_p^*$, $\text{val}_{\frac{p}{q}}(uv) = \text{val}_{\frac{p}{q}}(u) \left( \frac{p}{q} \right)^{|v|} + \text{val}_{\frac{p}{q}}(v)$.
- $m < n$ if and only if $\text{rep}_{\frac{p}{a}}(m) < \text{rep}_{\frac{p}{a}}(n)$ for the radix order.
- $L_{\frac{p}{q}} \subseteq A_p^*$ is not regular.
- $L_{\frac{p}{q}} \subseteq A_p^*$ is prefix-closed.

Example: base $\frac{3}{2}$

Signature: $2, 1, 2, 1, 2, 1, \ldots$

(i-tree: if we add an edge of label 0 onto the root)

Labeling: $0, 2, 1, 0, 2, 1, 0, 2, 1, \ldots$

$n$th node: $\text{rep}_{\frac{3}{2}}(n)$

(breadth-first traversal)

Edges: $n \xrightarrow{a \in A_3} m \iff m = \frac{3}{2} \cdot n + \frac{a}{2}$. 
ANS given by periodic labeled signatures

A labeled signature is an infinite sequence \((w_n)_{n \geq 0}\) of finite words providing

- a signature \(\{|w_n|\}_{n \geq 0}\) and
- a consistent labeling (made of the sequence of letters of \((w_n)_{n \geq 0}\)) of a tree.

The canonical breadth-first traversal of this tree produces an ANS.

Example:

Labeled signature: \((02, 1)\omega\)

Signature: \(|02|, |1|, |02|, |1|, \ldots\)

\[=2 \quad =1 \quad =2 \quad =1\]

Labeling: 0, 2, 1, 0, 2, 1, 0, 2, 1, \ldots

Base \(\frac{3}{2}\)
Example: Rational bases

<table>
<thead>
<tr>
<th>$p/q$</th>
<th>corresp. labeled sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/2</td>
<td>$(02, 1)^\omega$</td>
</tr>
<tr>
<td>5/2</td>
<td>$(024, 13)^\omega$</td>
</tr>
<tr>
<td>7/3</td>
<td>$(036, 25, 14)^\omega$</td>
</tr>
<tr>
<td>11/4</td>
<td>$(048, 159, 26(10), 37)^\omega$</td>
</tr>
</tbody>
</table>

Example: i-tree associated with the labeled signature $(023, 14, 5)^\omega$
Automatic sequences

**Definition**: Let $S = (L, A, <)$ be an ANS and let $B$ be a finite alphabet. An infinite word $x = x_0x_1x_2\cdots \in B^\mathbb{N}$ is $S$-automatic if there exists a deterministic finite automaton with output (DFAO for short) $A = (Q, q_0, A, \delta, \mu : Q \rightarrow B)$ such that

$$x_n = \mu(\delta(q_0, \text{rep}_S(n))) \quad \forall n \geq 0.$$ 

We read most significant digits first (not a restriction).

**Remark**: We talk about...

- $k$-automatic seq. in the base-$k$ numeration system $(L_k, A_k, <)$
- $\frac{p}{q}$-automatic seq. in the base-$\frac{p}{q}$ numeration system $(L_{\frac{p}{q}}, A_{\frac{p}{q}}, <)$.
Toy example: the sum-of-digits in base $\frac{3}{2}$

<table>
<thead>
<tr>
<th>$n$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{rep}_{\frac{3}{2}}(n)$</td>
<td>$\varepsilon$</td>
<td>2</td>
<td>21</td>
<td>210</td>
<td>212</td>
<td>2101</td>
<td>2120</td>
<td>2122</td>
</tr>
<tr>
<td>$s(n)$</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>$t(n) = s(n) \mod 2$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The sequence $t$ is $\frac{3}{2}$-automatic as it is generated by the following DFAO when reading base-$\frac{3}{2}$ representations:

![DFAO diagram]
Periodic labeled signature \((023, 14, 5)^\omega\) producing the i-tree

The sum-of-digits in \(S\) modulo 2

<table>
<thead>
<tr>
<th>(n)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{rep}_S(n))</td>
<td>(\varepsilon)</td>
<td>2</td>
<td>3</td>
<td>21</td>
<td>24</td>
<td>35</td>
<td>210</td>
<td>212</td>
</tr>
<tr>
<td>(\text{s.o.d})</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>(\text{mod}_2)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

is \(S\)-automatic since it is generated by the DFAO

The first few words in the corresponding ANS \(S\):

\(\varepsilon, 2, 3, 21, 24, 35, 210, 212, 213, 241, 244, 355, \ldots\)
Cobham’s theorem

**Theorem** (Cobham, 1972)

A sequence is $k$-automatic if and only if it is the image under a coding of a fixed point of a $k$-uniform morphism.

Many generalizations exist.

**Goal:** Generalization to our context of $S$-automatic sequences for ANS built on tree languages with a purely periodic labeled signatures.

**Needed:** alternate fixed points and block substitutions.
Alternating morphisms

The Kolakoski–Oldenburger word is the unique word $k$ over $\{1, 2\}$ starting with 2 and satisfying $\Delta(k) = k$ where $\Delta$ is the run-length encoding map

$$k = 2211212212211 \cdots.$$  [OEIS, A000002]

How to build it?
“Write what you read”: each term of $k$ generates a run of one or two future terms.

2 starting point
22 the first letter 2 generates a run of “22”
2211 the second letter 2 generates a run of “11”
22112 the first letter 1 generates a run of “2”
221121 the second letter 1 generates a run of “1”
22112122 and so on and so forth...

It is a well-known and challenging object of study in CoW. Conjecture: The density of 1 in $k$ is $\frac{1}{2}$. (Still open.)
Alternate fixed point

Alternative definition (Culik, Karhumäki, Lepistö, 1992): \( k = k_0k_1k_2 \cdots \)
can be obtained by periodically iterating two morphisms

\[
k = h_0(k_0)h_1(k_1) \cdots h_0(k_{2n})h_1(k_{2n+1}) \cdots
\]

where

\[
h_0 : \begin{cases} 
1 & \mapsto 2 \\
2 & \mapsto 22
\end{cases}
\quad \text{and} \quad
h_1 : \begin{cases} 
1 & \mapsto 1 \\
2 & \mapsto 11
\end{cases}
\]

\[
\begin{array}{|c|c|c|c|c|c|c|c|c|}
\hline
k & 2 & 2 & 1 & 1 & 2 & 1 & 2 & 1 \\
\hline
(h_0, h_1) & h_0 & h_1 & h_0 & h_1 & h_0 & h_1 & h_0 & h_1 \hline
(h_0, h_1)(k) = k & 22 & 11 & 2 & 1 & 22 & 1 & 22 & 11 \hline
\end{array}
\]

We say that \( k \) is an alternate fixed point of \((h_0, h_1)\).

Definition

Let \( r \geq 1 \) be an integer and let \( f_0, \ldots, f_{r-1} \) be morphisms over an alphabet \( A \).

An word \( w = w_0w_1 \cdots \in A^\omega \) is an alternate fixed point of \((f_0, \ldots, f_{r-1})\)
if \( w = f_0(w_0)f_1(w_1) \cdots f_{r-1}(w_{r-1})f_0(w_r) \cdots f_{i \mod r}(w_i) \cdots \).
Alternative definition (Dekking, 1980): Since

\[
\begin{align*}
    k &= h_0(2)h_1(2) = g(22) \\
    &= h_0(1)h_1(1) = g(11) \\
    &= h_0(2)h_1(1) = g(21) \\
    &= h_0(2)h_1(2) = g(22) \\
    &= h_0(1)h_1(2) = g(12) \\
\end{align*}
\]

\[= 2211 \ 21 \ 221 \ 2211 \ 211 \ldots\]

\(k\) is also a fixed point of the \(2\)-block substitution

\[
g : \begin{cases} 
11 & \mapsto h_0(1)h_1(1) = 21 \\
12 & \mapsto h_0(1)h_1(2) = 211 \\
21 & \mapsto h_0(2)h_1(1) = 221 \\
22 & \mapsto h_0(2)h_1(2) = 2211. 
\end{cases}
\]

Remarked: Lengths of images under \(g\) are not all equal.
**Definition**

An *r*-block substitution \( g : A^r \rightarrow A^* \) maps a word \( w_0 \cdots w_{rn-1} \in A^* \) to

\[
g(w_0 \cdots w_{r-1})g(w_r \cdots w_{2r-1}) \cdots g(w_{r(n-1)} \cdots w_{rn-1}).
\]

If the length of the word is not a multiple of \( r \), then the suffix of the word is ignored under the action of \( g \).

An infinite word \( w = w_0w_1 \cdots \in A^\omega \) is a **fixed point of the *r*-block substitution** \( g : A^r \rightarrow A^* \) if

\[
w = g(w_0 \cdots w_{r-1})g(w_r \cdots w_{2r-1}) \cdots.
\]

**Proposition**

If an infinite word over \( A \) is an alternate fixed point of \((f_0, \ldots, f_{r-1})\), then it is a fixed point of an *r*-block substitution.

**Proof:** For every of length-*r* word \( a_0 \cdots a_{r-1} \), define the *r*-block substitution \( g : A^r \rightarrow A^* \) by \( g(a_0 \cdots a_{r-1}) = f_0(a_0) \cdots f_{r-1}(a_{r-1}) \).
Goal: From a DFAO, build alternating morphisms.

ANS built on the purely periodic labeled signature 
\((w_0, w_1, \ldots, w_{r-1})^\omega\)

\(x\) automatic produced by a DFA \(A = (Q, q_0, A, \delta)\)

\[f_i : Q \rightarrow Q_{|w_i|} : q \mapsto \delta(q, w_i, 0) \cdot \delta(q, w_i, |w_i| - 1)\]
\(\forall i \in \{0, \ldots, r-1\}\)

\(x\) is the alternate fixed point of \((f_0, \ldots, f_{r-1})\)
starting with \(q_0\)

Base \(\frac{3}{2}\) \((02, 1)^\omega\)

Sum-of-digits mod 2
\(t = 0011101111101 \cdots\) \(\frac{3}{2}\)-automatic

\[f_0 : \begin{cases} 0 \mapsto \delta(0, 0) = 00 \\ 1 \mapsto \delta(1, 0) = 11 \end{cases}\]

\[f_1 : \begin{cases} 0 \mapsto \delta(0, 1) = 1 \\ 1 \mapsto \delta(1, 1) = 0 \end{cases}\]

\[f_0(0) = 00 \\
 f_0(0)f_1(0) = 00 1 \\
 f_0(0)f_1(0)f_0(1) = 00 1 11 \\
 f_0(0)f_1(0)f_0(1)f_1(1)f_0(1) = 00 1 11 0 11 1\]

\(t\) alternate fixed point of \((f_0, f_1)\)
Proposition (Rigo, S., 2021)

Let \( r \geq 1 \) be an integer and let \( A \) be a finite alphabet of digits. Let \( w_0, \ldots, w_{r-1} \) be \( r \) non-empty words in \( \text{inc}(A^*) \).

Consider the language \( L(s) \) of the i-tree generated by the purely periodic signature \( s = (w_0, w_1, \ldots, w_{r-1})^\omega \) and the corresponding ANS \( S = (L(s), A, <) \).

Let \( A = (Q, q_0, A, \delta) \) be a DFA.

For \( i \in \{0, \ldots, r-1\} \), we define the \( r \) morphisms from \( Q^* \) to itself by

\[
f_i : Q \rightarrow Q^{|w_i|}, \quad q \mapsto \delta(q, w_i, 0) \cdots \delta(q, w_i, |w_i| - 1),
\]

where \( w_{i,j} \) denotes the \( j \)th letter of \( w_i \).

The alternate fixed point \( x = x_0x_1 \cdots \) of \((f_0, \ldots, f_{r-1})\) starting with \( q_0 \) is the sequence of states reached in \( A \) when reading the words of \( L(s) \) in increasing radix order, i.e., for all \( n \geq 0 \), \( x_n = \delta(q_0, \text{rep}_S(n)) \).
Alternate fixed points to automatic sequences

Goal: From alternating morphisms, build a DFAO.

\( (f_0, \ldots, f_{r-1}) \) \( f_i : A^* \rightarrow A^* \) \( \ell_i \)-unif.

\( f_0 \) is prolongable on \( a \)

ANS \( S \) built on the purely periodic labeled signature \( (w_0, \ldots, w_{r-1})^\omega \)

\[
\begin{align*}
w_0 &= 0 \cdots (\ell_0 - 1) \\
w_1 &= \ell_0(\ell_0 + 1) \cdots (\ell_0 + \ell_1 - 1) \\
\vdots \\
w_{r-1} &= (\sum_{j<r} \ell_j) \cdots (\sum_{j<r} \ell_j - 1)
\end{align*}
\]

\( \mathcal{A} = (A, a, B, \delta) \) with

\[
\begin{align*}
B &= \{0, \ldots, \sum_{j<r} \ell_j - 1\} \\
\delta(b, i) &= [f_{j_i}(b)]_{i_j} \text{ for } b \in A, i \in B \text{ such that } \quad i = \sum_{k \leq j - 1} \ell_k + t \text{ with } j \geq 0, 0 \leq t < \ell_j
\end{align*}
\]

The alternate fixed point of \( (f_0, \ldots, f_{r-1}) \) starting with \( a \) is \( S \)-automatic (produced by \( \mathcal{A} \))

\[
\begin{align*}
\mathcal{A} &= (\{0, 1\}, 0, \{0, 1, 2\}, \delta) \\
\ell_0 &= 2 \\
\ell_1 &= 1
\end{align*}
\]

\[
\begin{align*}
f_0 : \begin{cases} 
0 & \mapsto 00 \\
1 & \mapsto 11
\end{cases} \\
f_1 : \begin{cases} 
0 & \mapsto 1 \\
1 & \mapsto 0
\end{cases}
\end{align*}
\]

The alternate fixed point of \( (f_0, f_1) \) is automatic. Up to a coding of the ANS, it is equal to \( t \) (\( \frac{3}{2} \)-automatic).
Proposition (Rigo, S., 2021)

Let $r \geq 1$ be an integer and let $A$ be a finite alphabet. For all $i \in \{0, \ldots, r - 1\}$, let $f_i : A^* \to A^*$ be a $\ell_i$-uniform morphism such that $f_0$ is prolongable on some $a \in A$.

Consider the language $L(s)$ of the i-tree generated by the purely periodic labeled signature

$$s = \left(0 \cdots (\ell_0 - 1), \ell_0(\ell_0 + 1) \cdots (\ell_0 + \ell_1 - 1), \ldots, \left(\sum_{j < r - 1} \ell_j\right) \cdots \left(\sum_{j < r - 1} \ell_j - 1\right)\right)$$

and the corresponding ANS $S = (L(s), B, <)$.

Let $A = (A, a, B, \delta)$ be the DFA where $B = \{0, \ldots, \sum_{j < r} \ell_j - 1\}$ and its transition function $\delta : A \times B \to A$ is defined as follows: for all $i \in B$, $\exists! j_i \geq 0$ and $\exists! t_i \geq 0$ such that $i = \sum_{k \leq j_i - 1} \ell_k + t_i$ with $t_i < \ell_{j_i}$, so we set $\delta(b, i) = \left[f_{j_i}(b)\right]_{t_i}$ $\forall b \in A$.

Then the alternate fixed point $x = x_0 x_1 \cdots$ of $(f_0, \ldots, f_{r-1})$ starting with $a$ is the sequence of the states reached in $A$ when reading the words of $L(s)$ by increasing radix order, i.e., for all $n \geq 0$, $x_n = \delta(a, \text{rep}_S(n))$. 
**Theorem (Rigo, S., 2021)**

Let $A, B$ be two finite alphabets. An infinite word over $B$ is the image under a coding $g : A \to B$ of an alternate fixed point of uniform morphisms (not necessarily of the same length) over $A$ if and only if it is $S$-automatic for an ANS built on a tree language with a purely periodic labeled signature.

**Proof:** It follows from the previous two propositions.
Corollary (Rigo, S., 2021)

If a sequence is $\frac{p}{q}$-automatic, then it is the image under a coding of a fixed point of a $q$-block substitution whose images all have length $p$.

Example: The sum-of-digits mod 2 $t = 0011101111101 \cdots$ is $\frac{3}{2}$-automatic.

$$f_0 : \begin{cases} 0 \mapsto 00 \\ 1 \mapsto 11 \end{cases} \quad f_1 : \begin{cases} 0 \mapsto 1 \\ 1 \mapsto 0 \end{cases}$$

Then $t$ is also a fixed point of the 2-block substitution

$$g : \begin{cases} 00 \mapsto f_0(0)f_1(0) = 001 \\ 01 \mapsto f_0(0)f_1(1) = 000 \\ 10 \mapsto f_0(1)f_1(0) = 111 \\ 11 \mapsto f_0(1)f_1(1) = 110. \end{cases}$$

Observe that images under $g$ all have length 3.
Non-morphic sequences

**Theorem (Cobham, 1972) rewritten**

An automatic sequence in an integer base is a morphic word, i.e., the image, under a coding, of a fixed point of a prolongable morphism.

Consider the morphisms $g_0 : 0 \mapsto 01, 1 \mapsto 00$ and $g_1 : 0 \mapsto 1, 1 \mapsto 0$ yielding the 2-block substitution

$$h_2 : 00 \mapsto 011, \ 01 \mapsto 010, \ 10 \mapsto 001, \ 11 \mapsto 000$$

producing the word $F_2 = 01011000011\cdots$. It is $\frac{3}{2}$-automatic as it is generated by the following DFAO (built thanks to our proposition):

$$\delta(0, 0) = [g_0(0)]_0 = 0 \quad \delta(1, 0) = [g_0(1)]_0 = 0$$
$$\delta(0, 1) = [g_0(0)]_1 = 1 \quad \delta(1, 1) = [g_0(1)]_1 = 0$$
$$\delta(0, 2) = [g_1(0)]_0 = 1 \quad \delta(1, 2) = [g_1(1)]_0 = 0$$

**Property (Lepistö, 1993)**

The word $F_2$ is neither purely morphic nor morphic.
Decorated trees (not only for Christmas)

ANS $S = (L, A,<)$ where $L = \{w_0 < w_1 < \cdots\}$ is prefix-closed producing a tree $T(L)$.

We now add an extra information on every node.

**Definition:** Let $x = x_0x_1\cdots \in B^\mathbb{N}$ where $B$ is a finite alphabet. A decoration of $T(L)$ by $x$ is the map $L \rightarrow B : w_n \mapsto x_n$.

**Example:** (0 means black, 1 means red)

**Base 2**

Decoration: Thue-Morse

0110100110010110100110010110\cdots

**Base $\frac{3}{2}$**

Decoration: Sum-of-digits modulo 2

0011101111101101110110111011011\cdots
**Factors of trees (subtrees)**

**Definition:** Let $T$ be a labeled tree. The **domain** $\text{dom}(T)$ of $T$ is the set of labels of paths from the root to its nodes.

**Example:**
The domain of the tree on the right is \{\varepsilon, 21, 210, 212, 2101, 2120, 2122\}.

**Definition:** Let $w \in L$ and let $h \geq 0$. We let $T[w, h]$ denote the **factor of height** $h$ rooted at $w$ of $T(L)$. The **prefix of height** $h$ of $T(L)$ is the factor $T[\varepsilon, h]$.

**Example:** The previous tree is a factor of height 4 of $T(L_{\frac{3}{2}})$. It is its prefix of height 4.
Definition: Two factors \( T[w, h] \) and \( T[w’, h] \) of the same height are equal if they have the same domain and the same decorations.

We let \( F_h = \{ T[w, h] \mid w \in L \} \) denote the set of factors of height \( h \) occurring in \( T(L) \).

Base 2
Decoration:
Thue-Morse
0110100110010110

Base \( \frac{3}{2} \)
Decoration:
Sum-of-digits modulo 2
0011101111011011
Theorem (Rigo, S., 2021)

A sequence $x$ is $k$-automatic if and only if, in the labeled tree $T(L_k)$ decorated by $x$, there exists a height $h \geq 0$ such that $\#F_h = \#F_{h+1}$.

Example: Base 2, decoration being Thue-Morse 0110100110010110 ···

Each factor $T[w, h]$ (not the prefix) is determined by $\delta(q_0, w)$: it is a full binary tree, and the decorations are determined by $\tau(\delta(q_0, wu))$ with $u \in A_{2}^{\leq h}$. Therefore $\#F_h \leq \#Q + 1$.

\[
\begin{array}{c}
\begin{array}{c}
0 \quad 1 \\
1 \\
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
1 \\
0 \\
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
0 \\
1 \\
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
0 \\
1 \\
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
0 \\
1 \\
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
0 \\
1 \\
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
1 \\
0 \\
\end{array}
\end{array}
\begin{array}{c}
\begin{array}{c}
1 \\
0 \\
\end{array}
\end{array}
\end{array}
\]

Theorem (Rigo, S., 2021)

Let $S = (L, A, <)$ be an ANS built on a prefix-closed regular language $L$. A sequence $x$ is $S$-automatic if and only if, in the labeled tree $T(L)$ decorated by $x$, there exists a height $h \geq 0$ such that $\#F_h = \#F_{h+1}$.
The rational bases case

Several extensions: in base $\frac{p}{q}$, except for the height-$h$ prefix, each factor of height $h$ is extended in exactly $q$ ways to a factor of height $h + 1$. To the first (leftmost) leaf of a factor of height $h$ are attached children corresponding to one of the $q$ words of the periodic signature.

**Example:** Base $\frac{3}{2}$ periodic labeled signature $(02, 1)\omega$
Factors appearing infinitely often

Definition: Let $T$ be a labeled decorated tree and let $h \geq 0$. We let $F_h^\infty \subseteq F_h$ denote the set of factors of height $h$ occurring infinitely often in $T$.

For any suitable letter $a$ in the signature of $T$, we let $F_{h,a}^\infty \subseteq F_h^\infty$ denote the set of factors of height $h$ occurring infinitely often in $T$ such that the label of the edge between the first node on level $h - 1$ and its first child is $a$. (Otherwise stated, the first word of length $h$ in the domain of the factor ends with $a$.)

Example: Base $\frac{3}{2}$

$F_{2,0}^\infty$: $F_{2,1}^\infty$: 

\begin{align*}
F_{2,0}^\infty: & \quad 0 \quad 2 \\
 & \quad 0 \quad 2 \quad 1
\end{align*}

\begin{align*}
F_{2,1}^\infty: & \quad 1 \\
 & \quad 1 \quad 0 \quad 2
\end{align*}

Automaticity in rational bases

Manon Stipulanti (ULiège)
Example: Base $\frac{3}{2}$
Assumptions: Factors of length 1 in $T(L_{\frac{3}{2}})$ can be extended as follows:

Assumptions:
- The first tree gives the prefix and occurs only once.
- The last eight trees of height 2 occur infinitely often.

Build an NFA:
- States: $\{T[\epsilon, 1]\} \cup F_1^\infty$
- Initial state: root of $T[\epsilon, 1]$
- Final states: nodes of $T[\epsilon, 1]$ and leaves in $F_1^\infty$
- Transitions...
Example for transitions on state 7.
The corresponding tree has two extensions:

Extension 1

The tree hanging to the child 0 (resp. 2) of the root corresponds to state 5 (resp. 7).
Transitions: $7 \rightarrow 5$ and $7 \rightarrow 7$

Extension 2

The tree hanging to the child 0 (resp. 2) of the root corresponds to state 7 (resp. 5).
Transitions: $7 \rightarrow 7$ and $7 \rightarrow 5$
Runs in the NFA for $w = 210 \in L_{\frac{3}{2}}$:

- $q_0 \xrightarrow{2} q_1 \times$
- $q_0 \xrightarrow{2} 0 \xrightarrow{1} 1 \times$
- $q_0 \xrightarrow{2} 0 \xrightarrow{1} 5 \times$
- $q_0 \xrightarrow{2} 0 \xrightarrow{1} 7 \xrightarrow{0} 5 \times$
- $q_0 \xrightarrow{2} 0 \xrightarrow{1} 7 \xrightarrow{0} 7 \times$
- $q_0 \xrightarrow{2} 0 \xrightarrow{1} 7 \xrightarrow{0} 8 \checkmark$

Determinize the NFA (using the usual subset construction).

No output is set for state 2.

This DFA produces the sum-of-digits modulo 2 in base $\frac{3}{2}$ ($\frac{3}{2}$-automatic).

Automaticity in rational bases
Theorem (Rigo, S., 2021)

Let the tree $T(L_{p/q})$ be decorated by a sequence $x$.
Suppose there exists some $h \geq 0$ such that $\#F_{h+1,a}^{\infty} \leq \#F_{h}^{\infty}$ for all $0 \leq j \leq q-1$ and all suitable letters $a \in A_p$.
Then $x$ is $\frac{p}{q}$-automatic.
Example: Base $\frac{3}{2}$
Decoration: the sum-of-digits mod 2 $t = 001110111110 \cdots$, $\frac{3}{2}$-automatic

Theorem (Rigo, S., 2021)

Let $x$ be a $\frac{p}{q}$-automatic sequence generated by a DFAO $\mathcal{A} = (Q, q_0, A_p, \delta, \tau : A_p \rightarrow B)$ with the following property:

$\exists h \geq 0$ s.t. $\forall q \neq q' \in Q$ and $\forall w \in L_p^\frac{q}{q}$, $\exists u \in w^{-1}L_p^\frac{q}{q}$ with $|u| \leq h$ s.t. $\tau(\delta(q, u)) \neq \tau(\delta(q', u))$.

Then in the tree $T(L_p^\frac{q}{q})$ decorated by $x$, we have $\#F_{h+1,a} \leq \#F_h^\infty$ for all $0 \leq j \leq q - 1$ and all suitable letters $a \in A_p$. 
References

References

○ V. Marsault, On $\frac{p}{q}$-recognisable sets, arXiv:1801.08707.