Auto-similarity in rational base number systems

Victor Marsault*,
joint work with Shigeki Akiyama† and Jacques Sakarovitch*

*CNRS / Telecom-ParisTech, Paris, France
†University of Tsukuba, Japan

WORDS, Turku, Finland
2013-09-17

Outline

Integer base

- base $p \geq 2$
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

Integer base

- base $p \geq 2$
- alphabet $A_p = \{0, 1, \dots, p-1\}$

- evaluation $\pi(a_n \cdots a_1 a_0) = \sum_{i=0}^{n} a_i p^i$
- $\pi(A_p^*) = \mathbb{N}$

Integer base

- base $p \geq 2$
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- evaluation $\pi(a_n \cdots a_1 a_0) = \sum_{i=0}^{n} a_i p^i$
- $\pi(A_p^*) = \mathbb{N}$

- representation $\langle n \rangle_p = \langle n' \rangle_p . a$
  - (n', a) is the Euclidean division of n by p.
- $\langle \mathbb{N} \rangle_p = (A_p \backslash \{0\}) \cdot A_p^*$

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}} . a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.

Example: computation of $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} =$$

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q
- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.

Example: computation of $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} =$$

$$\underset{\substack{\uparrow \\ q}}{\boxed{2}} \times \underset{\substack{\uparrow \\ n}}{\boxed{3}} = \underset{\substack{\uparrow \\ p}}{\boxed{3}} \times N_1 + a_0;$$

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}} . a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.

Example: computation of $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} =$$

$$2 \times 3 = 3 \times N_1 + a_0; \qquad \Rightarrow N_1 = 2 \text{ and } a_0 = 0.$$

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.

Example: computation of $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} = \langle 2 \rangle_{\frac{3}{2}} 0 =$$

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.

Example: computation of $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} = \langle 2 \rangle_{\frac{3}{2}} 0 =$$

$2 \times 2 = 3 \times N_2 + a_1$;

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.

Example: computation of $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} = \langle 2 \rangle_{\frac{3}{2}} 0 =$$

$2 \times 2 = 3 \times N_2 + a_1; \qquad \Rightarrow N_2 = 1$ and $a_1 = 1$.

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.

Example: computation of $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} = \langle 2 \rangle_{\frac{3}{2}} 0 = \langle 1 \rangle_{\frac{3}{2}} 10 =$$

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.

Example: computation of $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} = \langle 2 \rangle_{\frac{3}{2}} 0 = \langle 1 \rangle_{\frac{3}{2}} 10 =$$

$2 \times 1 = 3 \times N_3 + a_2;$

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.

Example: computation of $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} = \langle 2 \rangle_{\frac{3}{2}} 0 = \langle 1 \rangle_{\frac{3}{2}} 10 =$$

$2 \times 1 = 3 \times N_3 + a_2; \qquad \Rightarrow N_3 = 0$ and $a_2 = 2$.

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q
- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}} . a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.

Example: computation of $\langle 3 \rangle_{\frac{3}{2}}$:

$$\langle 3 \rangle_{\frac{3}{2}} = \langle 2 \rangle_{\frac{3}{2}} 0 = \langle 1 \rangle_{\frac{3}{2}} 10 = 210$$

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.
- $L_{\frac{p}{q}} = \langle \mathbb{N} \rangle_{\frac{p}{q}}$

Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.
- $L_{\frac{p}{q}} = \langle \mathbb{N} \rangle_{\frac{p}{q}}$

- evaluation $\pi(a_n \cdots a_1 a_0) = \sum_{i=0}^{n} (\frac{a_i}{q})(\frac{p}{q})^i$

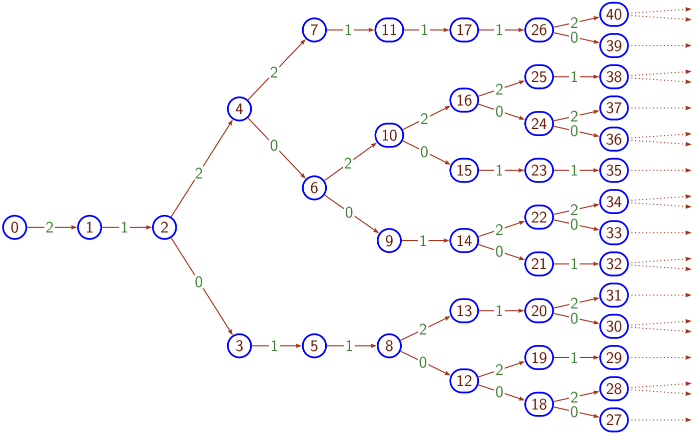
Rational base

- a base $\frac{p}{q}$ with $p > q$ and p coprime with q
- alphabet $A_p = \{0, 1, \ldots, p-1\}$

- representation $\langle n \rangle_{\frac{p}{q}} = \langle n' \rangle_{\frac{p}{q}}.a$:
  - (n', a) is the Euclidean division of $(\mathbf{q} \times n)$ by p.
- $L_{\frac{p}{q}} = \langle \mathbb{N} \rangle_{\frac{p}{q}}$

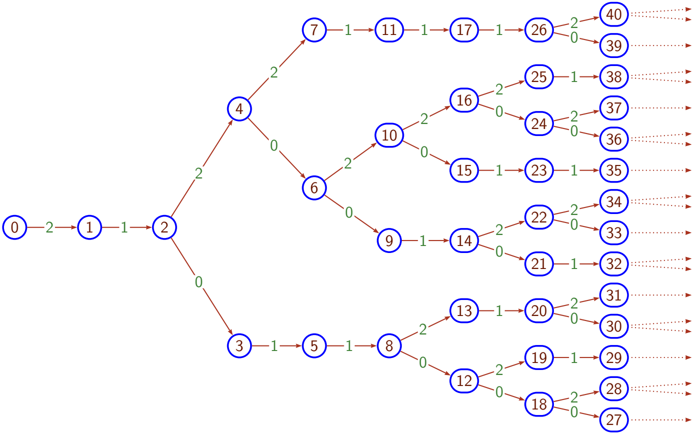
- evaluation $\pi(a_n \cdots a_1 a_0) = \sum_{i=0}^{n} (\frac{a_i}{q})(\frac{p}{q})^i$
  - if $\pi(u) = n \in \mathbb{N}$, u is of the form $0^k \langle n \rangle$
  - $\mathbb{N} \subsetneq \pi(A_p^*) \subsetneq \mathbb{Q}$

- $L_{\frac{p}{q}}$ is prefix-closed and right-extendable.



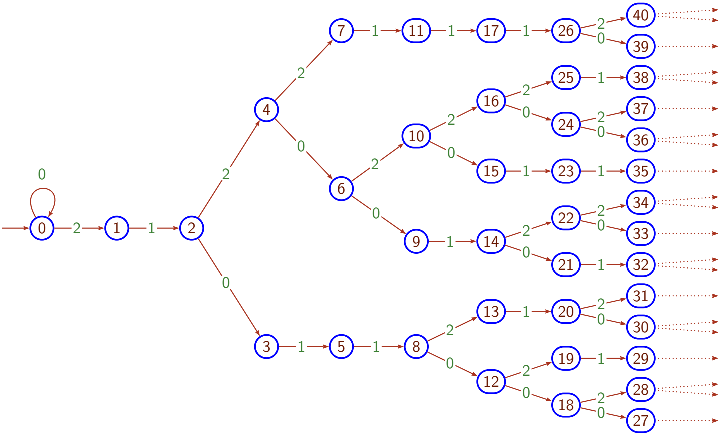
Rational base – the language $L_{\frac{p}{q}}$

- $L_{\frac{p}{q}}$ is prefix-closed and right-extendable.
- $L_{\frac{p}{q}}$ is not rational (not even context-free) [AFS'08].



Rational base – the automaton $T_{\frac{p}{q}}$

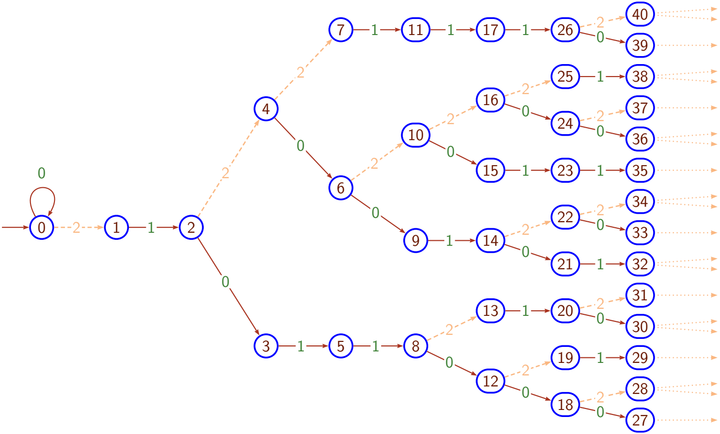
- $T_{\frac{p}{q}}$ accepts $0^* L_{\frac{p}{q}}$ (that is, the words u such that $\pi(u) \in \mathbb{N}$)



Outline

Minimal words – Definition

w_n: the (infinite) word starting from n taking the lowest branch.



Minimal words – Properties

Given an integer n, the minimal word w_n is

- over the alphabet $\{0, \ldots, (q-1)\} = A_q$
- the unique word over A_q readable from n

Minimal words – Properties

Given an integer n, the minimal word w_n is

- over the alphabet $\{0, \ldots, (q-1)\} = A_q$
- the unique word over A_q readable from n

Proposition [AFS'08]

$$w_n \text{ and } w_m \text{ have the same prefix of length } k.$$

$$\Updownarrow$$

$$n \equiv m \; [q^k]$$

Minimal words – Properties

Given an integer n, the minimal word w_n is

- over the alphabet $\{0, \ldots, (q-1)\} = A_q$
- the unique word over A_q readable from n
- different from w_m (for $m \neq n$)
- aperiodic

Proposition [AFS'08]

$$w_n \text{ and } w_m \text{ have the same prefix of length } k. \\ \Updownarrow \\ n \equiv m \ [q^k]$$

Minimal words – Set properties

W: the set of minimal words.

Topological properties

- The topological closure of W is A_q^ω whole.
- The interior of W is *empty*.

Shift operation

- W is stable by shift
- W cannot be finitely generated through shift.

Derived function

$$\begin{array}{rccc} \gamma : & A_q^\omega & \longrightarrow & A_q^\omega \\ & w_n & \longmapsto & w_{n+1} \end{array}$$

Derived function

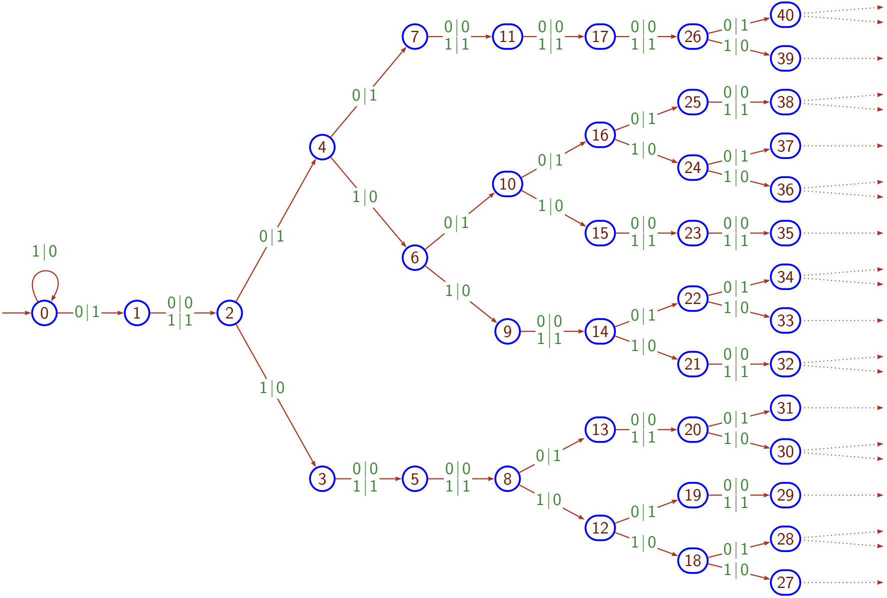
$$\begin{array}{rccc} \gamma: & A_q^\omega & \longrightarrow & A_q^\omega \\ & w_n & \longmapsto & w_{n+1} \end{array}$$

Remark

$w_n = a.w_{n+p}$ for some letter a and integer p.
(Or, equivalently $\gamma^p(w_n)$ is the shifted of w_n.)

Outline

Derived transducer for $\frac{3}{2}$: $D_{\frac{3}{2}}$

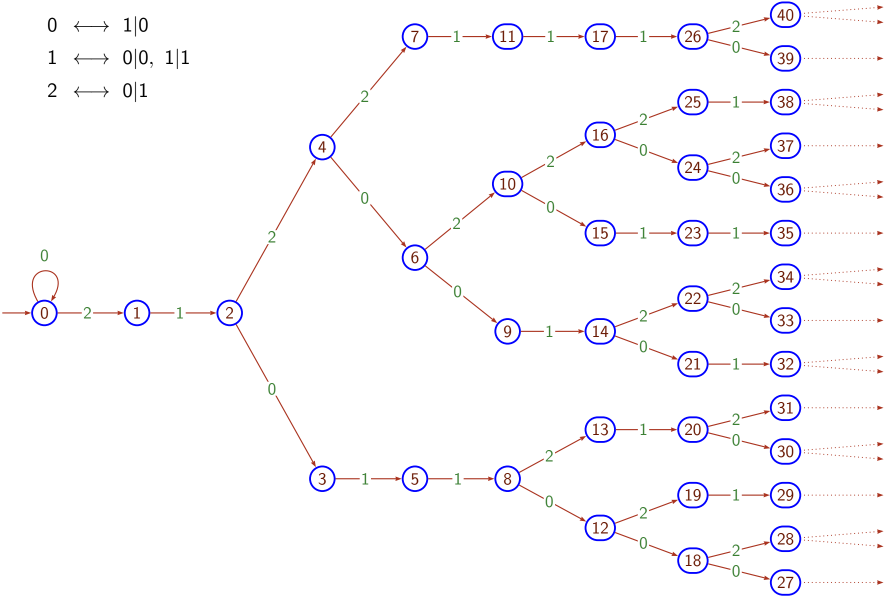


A simple label substitution...

$0 \longleftrightarrow 1|0$

$1 \longleftrightarrow 0|0,\ 1|1$

$2 \longleftrightarrow 0|1$



Proposition

If $\boxed{p = 2q - 1}$,

- the underlying graph of $D_{\frac{p}{q}}$ and $T_{\frac{p}{q}}$ are identical;
- the labels of the transitions of $D_{\frac{p}{q}}$ are obtained by an (injective) substitution from those of $T_{\frac{p}{q}}$.

Auto-similarity

Proposition

If $\boxed{p = 2q - 1}$,

- the underlying graph of $D_{\frac{p}{q}}$ and $T_{\frac{p}{q}}$ are identical;
- the labels of the transitions of $D_{\frac{p}{q}}$ are obtained by an (injective) substitution from those of $T_{\frac{p}{q}}$.

Theorem

The derived transducer $D_{\frac{p}{q}}$ is locally computable from $T_{\frac{p}{q}}$.

Transformation from $T_{\frac{p}{q}}$ to $D_{\frac{p}{q}}$

Step 1: changing the alphabet

$$A_p \longrightarrow B_{p,q} = \{p-(2q-1), \ldots, p-1\}$$

- $B_{p,q}$ always has $(2q-1)$ elements
- The maximal element of A_p and $B_{p,q}$ are the same.

Transformation from $T_{\frac{p}{q}}$ to $D_{\frac{p}{q}}$

Step 1: changing the alphabet

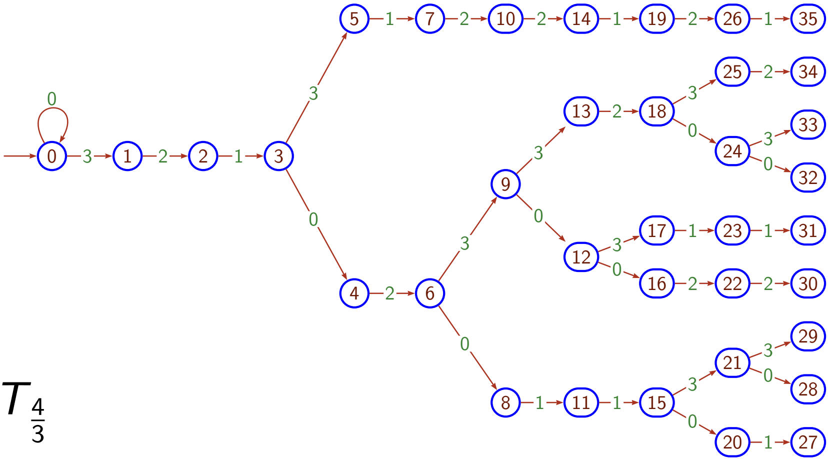
$$A_p \quad \longrightarrow \quad B_{p,q} = \{p-(2q-1), \ldots, p-1\}$$

- $B_{p,q}$ always has $(2q-1)$ elements
- The maximal element of A_p and $B_{p,q}$ are the same.

- if $p = (2q-1)$, $A_p = B_{p,q}$
- if $p < (2q-1)$, $A_p \subseteq B_{p,q}$ (the base $\frac{p}{q}$ is "too small")
- if $p > (2q-1)$, $A_p \supseteq B_{p,q}$ (the base $\frac{p}{q}$ is "too big")

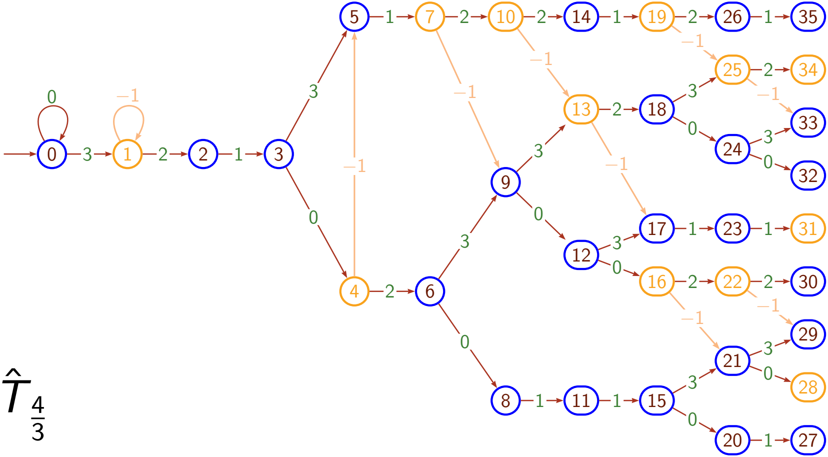
Example of the “small” base $\frac{4}{3}$ (Step 1)

- $A_4 = \{0, 1, 2, 3\} \quad \subseteq \quad \{-1, 0, 1, 2, 3\} = B_{4,3}$



$T_{\frac{4}{3}}$

- $A_4 = \{0, 1, 2, 3\} \subseteq \{-1, 0, 1, 2, 3\} = B_{4,3}$



$\hat{T}_{\frac{4}{3}}$

Transformation from $T_{\frac{p}{q}}$ to $D_{\frac{p}{q}}$

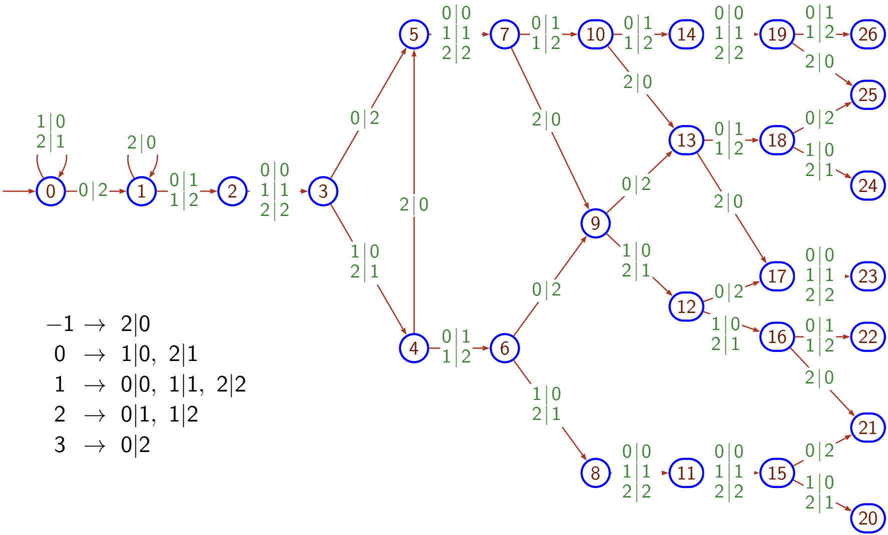
Step 1: changing the alphabet

$$A_p \quad \longrightarrow \quad B_{p,q} = \{p-(2q-1), \ldots, p-1\}$$

Step 2: changing the labels

$$\begin{array}{rccl} \omega : & B_{p,q} & \longrightarrow & \mathbb{P}(A_p \times A_p) \\ & a & \longmapsto & \{(b|c) \mid (b-c) = \underbrace{a-(p-q)}_{\text{distance to the center of B}}\} \end{array}$$

Example of the "small" base $\frac{4}{3}$ (Step 2)



$-1 \to 2|0$

$0 \to 1|0,\ 2|1$

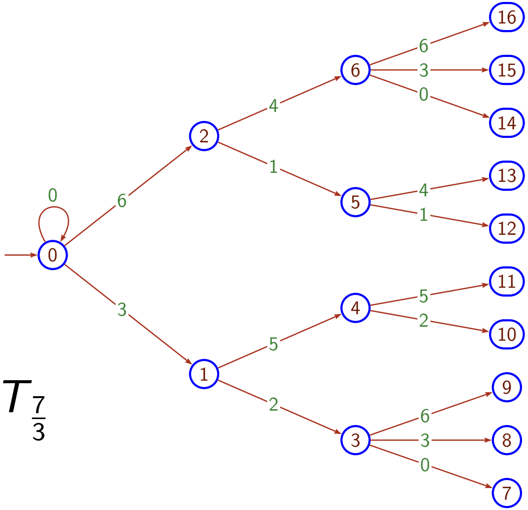
$1 \to 0|0,\ 1|1,\ 2|2$

$2 \to 0|1,\ 1|2$

$3 \to 0|2$

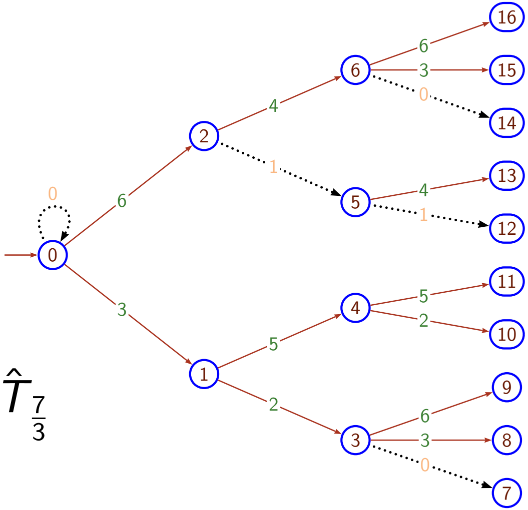
Example of the “big” base $\frac{7}{3}$ (Step 1)

- $A_7 = \{0, 1, 2, 3, 4, 5, 6\} \supseteq \{2, 3, 4, 5, 6\} = B_{7,3}$



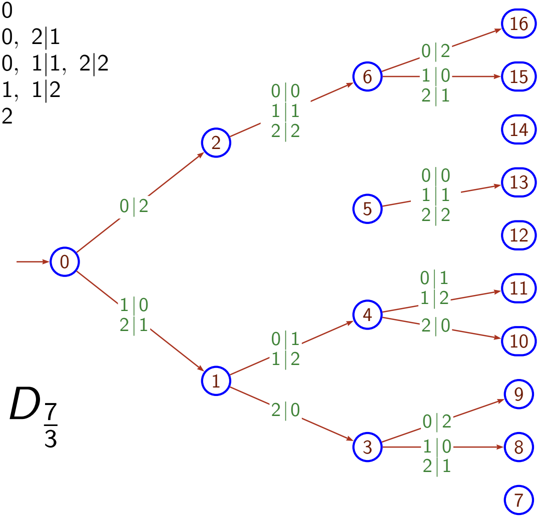
Example of the “big” base $\frac{7}{3}$ (Step 1)

- $A_7 = \{0,1,2,3,4,5,6\} \quad \supseteq \quad \{2,3,4,5,6\} = B_{7,3}$



Example of the "big" base $\frac{7}{3}$ (Step 2)

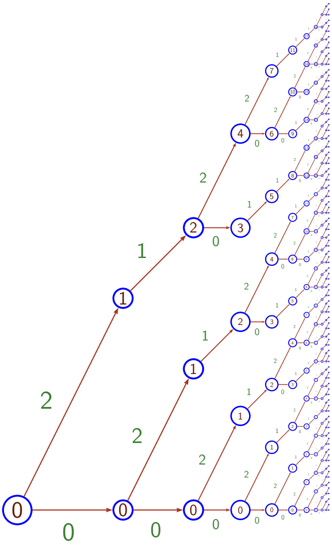
$$\begin{aligned}
2 &\to 2|0 \\
3 &\to 1|0,\ 2|1 \\
4 &\to 0|0,\ 1|1,\ 2|2 \\
5 &\to 0|1,\ 1|2 \\
6 &\to 0|2
\end{aligned}$$



Outline

Real evaluation

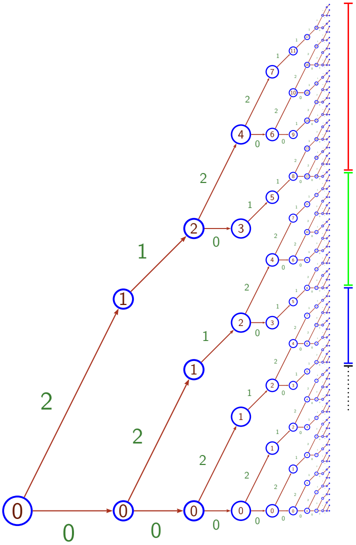
$$\rho(a_1 a_2 \cdots a_n \cdots) \quad = \quad \sum_{i \geq 0} \frac{a_i}{q} \left(\frac{p}{q}\right)^{-i} \quad .$$



Span and renormalised span

Definition – span of the node X

The length of the interval reachable from X in the tree.



$\text{span}(\text{"}2\text{"})\ (=\ \text{span}(\text{"}21\text{"}))$

$\text{span}(\text{"}02\text{"})\ =\ \text{span}(\text{"}2\text{"}) \times \frac{3}{2}$

$\text{span}(\text{"}002\text{"})\ =\ \text{span}(\text{"}2\text{"}) \times \left(\frac{3}{2}\right)^2$

Span and renormalised span

Definition – span of the node X

The length of the interval reachable from X in the tree.

Definition – renormalised span of the node X

the span of X multiplied by $(\frac{p}{q})^k$, where k is the depth of X.

$S_{\frac{p}{q}}$ denotes the set of the renormalised span of every node.

Span words

Definition

- The span of n is represented by the word $(w'_n \ominus w_n)$, where:
  - w'_n is the *maximal* word starting from n;
  - "$\ominus$" denotes the digit-wise subtraction.
    (Example : $321 \ominus 012 = 31(\text{-}1)$)
- It is called the *span-word* of n and is over the alphabet $B_{p,q}$.

Span words

Definition

- The span of n is represented by the word $(w'_n \ominus w_n)$, where:
  - w'_n is the *maximal* word starting from n;
  - "$\ominus$" denotes the digit-wise subtraction.
    (Example : $321 \ominus 012 = 31(\text{-}1)$)
- It is called the *span-word* of n and is over the alphabet $B_{p,q}$.

Proposition

$\hat{T}_{\frac{p}{q}}$ accepts the topological closure of the language of the span-words.

Topological Property of $S_{\frac{p}{q}}$

Theorem

- If $p \leq 2q - 1$, $S_{\frac{p}{q}}$ is dense.
- If $p > 2q - 1$, $S_{\frac{p}{q}}$ is nowhere dense.

Conclusion and future work

- The derived transducer somehow requires the same structure as the original tree.
- The topological properties of the set of spans divides the rational base number systems in two classes.
- The cases $p = 2q - 1$ is remarkable in both constructions.

Next question

For a given integer n,
is there a *finite* transducer realising $w_n \mapsto w_{n+1}$?