Structural Properties of Bounded Languages with Respect to Multiplication by a Constant

Emilie Charlier, Michel Rigo

Department of Mathematics
University of Liège

Journées Montoises d'Informatique Théorique
à Rennes, 2006

1) Abstract numeration systems

2) Main question

3) First results about S-recognizability

4) Bounded languages

5) B_ℓ-representation of an integer

6) Multiplication by $\lambda = \beta^\ell$

1) Abstract numeration systems

Definition. [P. Lecomte, M. Rigo, 2001] An *abstract numeration system* is a triple $S = (L, \Sigma, <)$ where L is a regular language over a totally ordered alphabet $(\Sigma, <)$.

Enumerating the words of L with respect to the genealogical ordering induced by $<$ gives a one-to-one correspondence

$$\mathrm{rep}_S : \mathbb{N} \to L \qquad \mathrm{val}_S = \mathrm{rep}_S^{-1} : L \to \mathbb{N}.$$

Examples

1) a^*

n	0	1	2	3	4	$\cdots$
rep(n)	ε	a	aa	aaa	$aaaa$	$\cdots$

2) $\{a,b\}^*,\ a < b$

n	0	1	2	3	4	5	6	7	$\cdots$
rep(n)	ε	a	b	aa	ab	ba	bb	aaa	$\cdots$

3) $a^*b^*,\ a < b$

n	0	1	2	3	4	5	6	$\cdots$
rep(n)	ε	a	b	aa	ab	bb	aaa	$\cdots$

Remark. This generalizes "classical" Pisot systems like integer base systems or Fibonacci system.

$$L = \{\varepsilon\} \cup \{1, \ldots, k-1\}\{0, \ldots, k-1\}^* \text{ or } L = \{\varepsilon\} \cup 1\{0, 01\}^*$$

Definition. A set $X \subseteq \mathbb{N}$ is S-*recognizable if* $\mathrm{rep}_S(X) \subseteq \Sigma^*$ *is a regular language (accepted by a DFA).*

2) Main question

If $S = (L, \Sigma, <)$ is an abstract numeration system, *can we find some necessary and sufficient condition on* $\lambda \in \mathbb{N}$ *such that for any* S*-recognizable set* X*, the set* λX *is still* S*-recognizable ?*

$$X \ S\text{-rec} \quad \xrightarrow{?} \quad \lambda X \ S\text{-rec}$$

3) First results about S-recognizability

Theorem 1. Let $S = (L, \Sigma, <)$ be an abstract numeration system. Any arithmetic progression is S-recognizable.

Definition. We denote by $\mathbf{u}_L(n)$ the number of words of length n belonging to L.

Theorem 2. [Polynomial case] Let $L \subseteq \Sigma^*$ be a regular language such that $\mathbf{u}_L(n) \in \Theta(n^k)$, $k \in \mathbb{N}$ and $S = (L, \Sigma, <)$. Preservation of the S-recognizability after multiplication by λ holds only if $\lambda = \beta^{k+1}$ for some $\beta \in \mathbb{N}$.

Definition. A language L is *slender* if $\mathbf{u}_L(n) \in O(1)$.

Theorem 3. [Slender case] Let $L \subset \Sigma^*$ be a slender regular language and $S = (L, \Sigma, <)$. A set $X \subseteq \mathbb{N}$ is S-recognizable if and only if X is a finite union of arithmetic progressions.

Corollary. Let S be a numeration system built on a slender language. If $X \subseteq \mathbb{N}$ is S-recognizable then λX is S-recognizable for all $\lambda \in \mathbb{N}$.

Theorem 4. Let $\beta > 0$. For the abstract numeration system

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

the multiplication by β^2 preserves S-recognizability if and only if β is an odd integer.

4) Bounded languages, notation

We denote by $\mathcal{B}_\ell = a_1^* \cdots a_\ell^*$ the bounded language over the totally ordered alphabet $\Sigma_\ell = \{a_1 < \ldots < a_\ell\}$ of size $\ell \geq 1$.

We consider abstract numeration systems of the form $(\mathcal{B}_\ell, \Sigma_\ell)$ and we denote by rep_ℓ and val_ℓ the corresponding bijections.

A set $X \subseteq \mathbb{N}$ is said to be $\mathcal{B}_\ell$-*recognizable* if $\mathrm{rep}_\ell(X)$ is a regular language over the alphabet Σ_ℓ.

In this context, multiplication by a constant λ can be viewed as a transformation

$$f_\lambda : \mathcal{B}_\ell \to \mathcal{B}_\ell.$$

The question becomes then :

Can we determine some necessary and sufficient condition under which this transformation preserves regular subsets of $\mathcal{B}_\ell$?

Example

Let $\ell = 2$, $\Sigma_2 = \{a, b\}$ and $\lambda = 25$.

$$\begin{array}{rcl} 8 & \xrightarrow{\times 25} & 200 \\ \mathsf{rep}_2 \downarrow & & \downarrow \mathsf{rep}_2 \\ a\,b^2 & \xrightarrow{f_{25}} & a^9\,b^{10} \end{array} \qquad \qquad \begin{array}{rcl} \mathbb{N} & \xrightarrow{\times \lambda} & \mathbb{N} \\ \mathsf{rep}_\ell \downarrow & & \downarrow \mathsf{rep}_\ell \\ \mathcal{B}_\ell & \xrightarrow{f_\lambda} & \mathcal{B}_\ell \end{array}$$

Thus multiplication by $\lambda = 25$ induces a mapping f_λ onto $\mathcal{B}_2$ such that for $w, w' \in \mathcal{B}_2$, $f_\lambda(w) = w'$ if and only if $\mathsf{val}_2(w') = 25\,\mathsf{val}_2(w)$.

5) B_ℓ-representation of an integer

We set

$$\mathbf{u}_\ell(n) := \mathbf{u}_{\mathcal{B}_\ell}(n) = \#(\mathcal{B}_\ell \cap \Sigma_\ell^n)$$

and

$$\mathbf{v}_\ell(n) := \#(\mathcal{B}_\ell \cap \Sigma_\ell^{\leq n}) = \sum_{i=0}^{n} \mathbf{u}_\ell(i).$$

Lemma 1. For all $\ell \geq 1$ and $n \geq 0$, we have

$$\mathbf{u}_{\ell+1}(n) = \mathbf{v}_\ell(n) \tag{1}$$

and

$$\mathbf{u}_\ell(n) = \binom{n+\ell-1}{\ell-1}. \tag{2}$$

Lemma 2. Let $S = (a_1^* \cdots a_\ell^*, \{a_1 < \cdots < a_\ell\})$. We have

$$\mathsf{val}_\ell(a_1^{n_1} \cdots a_\ell^{n_\ell}) = \sum_{i=1}^{\ell} \binom{n_i + \cdots + n_\ell + \ell - i}{\ell - i + 1}.$$

Corollary. [Katona, 1966] Let $\ell \in \mathbb{N} \setminus \{0\}$. Any integer n can be uniquely written as

$$n = \binom{z_\ell}{\ell} + \binom{z_{\ell-1}}{\ell - 1} + \cdots + \binom{z_1}{1} \tag{3}$$

with $z_\ell > z_{\ell-1} > \cdots > z_1 \geq 0$.

Example

Consider the words of length 3 in the language $a^*b^*c^*$,

$$aaa < aab < aac < abb < abc < acc < bbb < bbc < bcc < ccc.$$

We have $\mathrm{val}_3(aaa) = \binom{5}{3} = 10$ and $\mathrm{val}_3(acc) = 15$.

If we apply the erasing morphism $\varphi : \{a,b,c\} \to \{a,b,c\}^*$ defined by

$$\varphi(a) = \varepsilon, \varphi(b) = b, \varphi(c) = c$$

on the words of length 3, we get

$$\varepsilon < b < c < bb < bc < cc < bbb < bbc < bcc < ccc.$$

So we have $\mathrm{val}_3(acc) = \mathrm{val}_3(aaa) + \mathrm{val}_2(cc)$ where val_2 is considered as a map defined on the language b^*c^*.

Algorithm computing rep$_\ell(n)$.

Let `n` be an integer and `l` be a positive integer.

For i=l,l-1,...,1 do

if n>0,

find t such that $\binom{\mathtt{t}}{\mathtt{i}} \leq \mathtt{n} < \binom{\mathtt{t+1}}{\mathtt{i}}$

z(i)$\longleftarrow$t

n$\longleftarrow$n-$\binom{\mathtt{t}}{\mathtt{i}}$

otherwise, z(i)$\longleftarrow$i-1

Consider now the triangular system having $\alpha_1, \ldots, \alpha_\ell$ as unknowns

$$\alpha_i + \cdots + \alpha_\ell = \mathtt{z}(\ell - i + 1) - \ell + i, \quad i = 1, \ldots, \ell.$$

One has rep$_\ell(\mathtt{n}) = a_1^{\alpha_1} \cdots a_\ell^{\alpha_\ell}$.

Example

For $\ell = 3$, one gets for instance

$$12345678901234567890 = \binom{4199737}{3} + \binom{3803913}{2} + \binom{1580642}{1}$$

and solving the system

$$\left\{\begin{array}{rcl} n_1 + n_2 + n_3 & = & 4199737 - 2 \\ n_2 + n_3 & = & 3803913 - 1 \\ n_3 & = & 1580642 \end{array}\right.$$

$$\Leftrightarrow (n_1, n_2, n_3) = (395823, 2223270, 1580642),$$

we have

$$\text{rep}_3(12345678901234567890) = a^{395823} b^{2223270} c^{1580642}.$$

Remark. In particular, we have $\mathbf{u}_{\mathcal{B}_\ell}(n) \in \Theta(n^{\ell-1})$.

So we have to focus only on multiplicators of the kind

$$\lambda = \beta^\ell.$$

6) Multiplication by $\lambda = \beta^\ell$

Theorem. For the abstract numeration system

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

if $\beta \in \mathbb{N} \setminus \{0,1\}$ is such that $\beta \not\equiv \pm 1 \pmod 6$ then the multiplication by β^3 does not preserve the S-recognizability.

For instance, if $\beta \equiv 2 \pmod 6$, for n large enough, we have

$$\mathrm{rep}_3\left[(6k+2)^3\,\mathrm{val}_3(a^n)\right] = a^r\, b^{s+(3k+1)n}\, c^{t+(3k+1)n}$$

where the constants r, s, t are given by

$$r = 4k+6k^2, \quad s = 5k+11k^2+24k^3+18k^4, \quad t = -3k-17k^2-24k^3-18k^4.$$

Conjecture. Multiplication by β^ℓ preserves S-recognizability for the abstract numeration system

$$S = (a_1^* \cdots a_\ell^*, \{a_1 < \cdots < a_\ell\})$$

built on the bounded language $\mathcal{B}_\ell$ over ℓ letters if and only if

$$\beta = \prod_{i=1}^{k} p_i^{\theta_i}$$

where $p_1, \ldots, p_k$ are prime numbers strictly greater than ℓ.

Lemma 1. For $n \in \mathbb{N}$ large enough, we have

$$|\mathsf{rep}_\ell(\beta^\ell n)| = \beta\,|\mathsf{rep}_\ell(n)| + \frac{(\beta-1)(\ell-1)}{2} + i$$

with $i \in \{-1, 0, \ldots, \beta-1\}$.

Definition. For all $i \in \{-1, 0, \ldots, \beta-1\}$ and $k \in \mathbb{N}$ large enough, we define

$$\mathcal{R}_{i,k} := \left\{ n \in \mathbb{N} : |\mathsf{rep}_\ell(n)| = k \text{ and } |\mathsf{rep}_\ell(\beta^\ell n)| = \beta\,k + \frac{(\beta-1)(\ell-1)}{2} + i \right\}.$$

We assume that β satisfies the condition of the Conjecture.

Proposition. Let $i \in \{0, \dots, \beta - 1\}$. There exists a constant $\mathbf{L} \geq 0$ (depending only on ℓ and β) such that for all $k \geq \mathbf{L}$, if $m = \min \mathcal{R}_{i,k}$ and $n = \min \mathcal{R}_{i,k+\beta^{\ell-1}}$ then

$$\forall t \in \{2, \dots, \ell\} : |\mathrm{rep}_\ell(\beta^\ell m)|_{a_t} = |\mathrm{rep}_\ell(\beta^\ell n)|_{a_t}.$$

Furthermore, $|\mathrm{rep}_\ell(\beta^\ell m)|_{a_1} + \beta^\ell = |\mathrm{rep}_\ell(\beta^\ell n)|_{a_1}$.