Sequences of level $1, 2, 3, \ldots, k, \ldots$

G. Sénizergues

LaBRI and Université de Bordeaux I **

Abstract. Sequences of numbers (either natural integers, or integers or rational) of level $k \in \mathbb{N}$ have been defined in [FS06] as the sequences which can be computed by deterministic pushdown automata of level k. We extend this definition to sequences of *words* indexed by *words*. We give characterisations of these sequences in terms of "higher-order" L-systems. In particular sequences of rational numbers of level 3 are characterised by polynomial recurrences (which generalize the P-recurrent sequences studied in [Sta80]). The equality problem for sequences of rational numbers of level 3 is shown decidable.

keywords: Iterated pushdown automata ; recurrent sequences ; equivalence problems.

1 Introduction

The class of pushdown automata of level k (for $k \geq 1$) has been introduced in [Gre70],[Mas74] as a generalisation of the automata and grammars of [Aho68],[Aho69], [Fis68] and has been the object of many further studies: see [Mas76], [ES77], [Dam82], [Eng83], [ES84], [EV86], [DG86], and more recently [Cau02], [KNU02], [CW03],[Fra05].

The class of *integer* sequences computed (in a suitable sense) by such automata was defined in [Fra05],[FS06](we denote it by $\mathbb{S}_k$).

The class $\mathcal{F}(S_k)$ consisting of all the sequences of *rational* numbers which can be decomposed as $\frac{a_n - b_n}{a'_n - b'_n}$ for sequences $a, b, a', b' \in \mathbb{S}_k$ was also introduced. These classes of number sequences fulfil many closure properties and generalize some well-known classes of recurrent sequences (or formal power series). The level 3, for example, contains all the so-called P-recurrent sequences of rational numbers, corresponding also to the D-finite formal power series (see [Sta80] for a survey and [PWZ96] for a thorough study of their algorithmic properties).

We give here several characterisations of the classes $\mathbb{S}_k$ for $k \geq 1$. These characterisations go through generalizations of the above classes $\mathbb{S}_k$ to their analogues

** mailing adress:LaBRI and UFR Math-info, Université Bordeaux1
351 Cours de la libération -33405- Talence Cedex.
email:{ges}@labri.u-bordeaux.fr; fax: 05-56-84-66-69;
URL:http://dept-info.labri.u-bordeaux.fr/∼ges

for sequences of words, formal power series with non-commutative undeterminates and, finally, mappings from words to words. Let us denote by $\mathbb{S}_k(A^*, B^*)$ the class of mappings from A^* to B^* computed by pushdown automata of level k. The elements of $\mathbb{S}_k(A^*, B^*)$ can be characterised by some kind of Lindenmayer-systems of "order k" that we introduce here. As a corollary, S_3 is characterised by polynomial recurrences. The equality problem for two sequences in $\mathcal{F}(S_3)$ can thus be solved by a suitable reduction to polynomial ideal theory (namely to the construction of Gröbner bases). The present text is an extended abstract: it gives the main definitions and states the main results but does not provide any formal proof.

2 Preliminaries

We introduce here some notation and basic definitions which will be used throughout the text.

2.1 Automata

Beside the usual notions of finite automaton and pushdown automaton, we shall consider here the notion of *pushdown automaton of level* k. Such automata are an extension of classical pushdown-automata to a storage structure built iteratively. This storage structure can be described as follows:

Definition 1 (k-iterated pushdown store). *Let Γ be a set. We define inductively the set of* k-iterated pushdown-stores over Γ *by:*

$$0-\mathrm{pds}(\Gamma) = \{\varepsilon\} \quad (k+1)-\mathrm{pds}(\Gamma) = (\Gamma[k-\mathrm{pds}(\Gamma)])^* \quad \mathrm{it}-\mathrm{pds}(\Gamma) = \bigcup_{k \geq 0} k-\mathrm{pds}(\Gamma).$$

The elementary operations that a k-pda can perform are:
- pop of level j (where $1 \leq j \leq k$), which consists of popping the leftmost letter of level j and all the structure which is "above" this letter
- push of level j (where $1 \leq j \leq k$), which consists of pushing a new letter C on the left of the leftmost letter D of level j and copying above this new letter C all the structure which was "above" the letter D.

A transition of the automaton consists, given the word γ made of all the leftmost letters of the k-pushdown (the one of level 1, followed by the one of level 2, ..., followed by the one of level k), the state q and the leftmost letter b (or, possibly, the empty word ε) on the input tape, in performing one of the above elementary operations. More formally,

Definition 2 (k-*pdas*). *Let $k \geq 1$, let $POP = \{\mathrm{pop}_j \,|j \in [k]\}$, $PUSH(\Gamma) = \{\mathrm{push}_j(\gamma)|\gamma \in \Gamma^+, j \in [k]\}$, and $TOPSYMS(\Gamma) = \Gamma^{(k)} - \{\varepsilon\}$.*
A k-iterated pushdown automaton over a terminal alphabet B is a 6-tuple $\mathcal{A} = (Q, B, \Gamma, \delta, q_0, Z_0)$ where

– *Q is a finite set of states, $q_0 \in Q$ denoting the initial state,*
– *Γ is a finite set of pushdown-symbols,$Z_0 \in \Gamma$ denoting the initial symbol,*
– *the transition function δ is a map from $Q \times (B \cup \{\varepsilon\}) \times TOPSYMS(\Gamma)$ into the set of finite subsets of $Q \times (PUSH(\Gamma) \cup POP)$ such that: if $(q, \mathrm{push}_j(\gamma)) \in \delta(p, \bar{b}, \gamma)$ then $j \leq |\gamma| + 1$ and if $(q, \mathrm{pop}_j) \in \delta(p, \bar{b}, \gamma)$ then $j \leq |\gamma|$.*

(see [FS06] for more details). The automaton $\mathcal{A}$ is said *deterministic* iff, for every $q \in Q, \gamma \in \Gamma^{(k)}, b \in B$

$$\mathrm{Card}(\delta(q, \varepsilon, \gamma)) \leq 1 \text{ and } \mathrm{Card}(\delta(q, b, \gamma)) \leq 1, \tag{1}$$

$$\mathrm{Card}(\delta(q, \varepsilon, \gamma)) = 1 \Rightarrow \mathrm{Card}(\delta(q, b, \gamma)) = 0. \tag{2}$$

In order to define a useful notion of map *computed* by a k-pda we introduce the following stronger condition: $\mathcal{A}$ is called *strongly deterministic* iff, for every $q \in Q, \gamma \in \Gamma^{(k)}$

$$\sum_{\bar{b} \in \{\varepsilon\} \cup B} \mathrm{Card}(\delta(q, \bar{b}, \gamma)) \leq 1 \tag{3}$$

In other words, the automaton $\mathcal{A}$ is *strongly deterministic* iff, the leftmost contents γ of the memory and the state q completely determine the transition of $\mathcal{A}$, in particular what letter b (or possibly the empty word) can be read. Therefore, such an automaton $\mathcal{A}$ can accept at most one word w from a given configuration. We say that $\mathcal{A}$ is *level-partitioned* iff Γ is the disjoint union of subsets $\Gamma_1, \Gamma_2, \ldots, \Gamma_k$ such that, in every transition of $\mathcal{A}$, every occurrence of a letter from Γ_i is at level i. It is easy to transform any k-pushdown automaton $\mathcal{A}$ into another one $\mathcal{A}'$ which recognizes the same language and is level-partitioned. Moreover, if $\mathcal{A}$ is strongly deterministic then $\mathcal{A}'$ is strongly deterministic.

Definition 3 (k-computable mapping). *A mapping $f : A^* \mapsto B^*$ is called* k-computable *iff there exists a strongly deterministic k-pda $\mathcal{A}$, over a pushdown-alphabet Γ which is level-partitionned, such that Γ contains $k-1$ symbols $U_1, U_2, \ldots, U_{k-1}$, the alphabet A is a subset of Γ_k and for all $w \in A^*$:*

$$(q_0, f(w), U_1[U_2 \ldots [U_{k-1}[w]] \ldots]) \vdash^*_{\mathcal{A}} (q_0, \varepsilon, \varepsilon).$$

One denotes by $\mathbb{S}_k(A^*, B^*)$ the set of k-computable mappings from A^* to B^*.

The particular case where $\mathrm{Card}(A) = \mathrm{Card}(B) = 1$ was studied in [FS06].

2.2 Number recurrences

Definition 4 ($\mathbb{N}$-rational formal power series). *A mapping $f : A^* \to \mathbb{N}$ is $\mathbb{N}$-rational iff there is an homomorphism $M : A^* \to \mathbb{N}^{d \times d}$ and two vectors L in $\mathbb{B}^{1 \times d}$ and T in $\mathbb{B}^{d \times 1}$ such that, for every $w \in A^*$*

$$f(w) = L \cdot M(w) \cdot T. \tag{4}$$

The map f can also be denoted by $\sum_{w \in A^*} f(w) \cdot w$ which explains our terminology.

Definition 5 (Polynomial recurrent relations). *Given a finite index set $I = [1, n]$ and a family of mappings $f_i : A^* \to \mathbb{N}$ (for $i \in I$), we call* system of polynomial recurrent relations *a system of the form*

$$f_i(aw) = P_{i,a}(f_1(w), f_2(w), \ldots, f_n(w)) \text{ for all } i \in I, a \in A, w \in A^*$$

where $P_i \in \mathbb{N}[X_1, X_2, \ldots, X_n]$.

A sinilar definition can be given for mappings $f_i : A^* \to \mathbb{Z}$ (for $i \in I$) and polynomials $P_i \in \mathbb{Z}[X_1, X_2, \ldots, X_n]$.

2.3 Word recurrences

When considering mappings into *words* instead of integers, one is lead to consider the following kind of recurrent relations.

Definition 6 (catenative recurrent relations). *Given a finite index set $I = [1, n]$ and a family of mappings $f_i : A^* \to B^*$ (for $i \in I$), we call* system of catenative recurrent relations *a system of the form*

$$f_i(aw) = \prod_{j=1}^{\ell(a,i)} f_{\alpha(i,a,j)}(w) \text{ for all } i \in I, a \in A, w \in A^*$$

where $\ell(a, i) \in \mathbb{N}, \alpha(i, a, j) \in I$.

One can check that, in the case where B is reduced to one letter, a mapping $f : A^* \to B^*$ is the first mapping of a family fulfilling a system of catenative recurrent relations iff f is a rational series.

2.4 Recurrences in a monoid

Mezei and Wright developed a general notion of grammar defining languages in general algebras ([MW67]). These ideas lead naturally to the following adaptation to arbitrary monoids of the notion of catenative recurrent relations. Let $(M, \cdot, 1)$ be some monoid.

Definition 7 (recurrent relations in M). *Given a finite index set $I = [1, n]$ and a family of mappings $f_i : A^* \to M$ (for $i \in I$), we call* system of recurrent relations in M *a system of the form*

$$f_i(aw) = \prod_{j=1}^{\ell(a,i)} f_{\alpha(i,a,j)}(w) \text{ for all } i \in I, a \in A, w \in A^*$$

where $\ell(a, i) \in \mathbb{N}, \alpha(i, a, j) \in I$ and the symbol $\prod$ stands for the extension of the binary product in M to an arbitrary finite number of arguments.

The monoid $(Hom(B^*, B^*), \circ, \mathrm{Id})$ will be of particular interest for studying mappings of level $k \geq 3$.

2.5 L-systems

The following notion ([KRS97]) turns out to be crucial for describing all k-computable mappings as compositions of simpler mappings.

Definition 8 (HDT0L sequences). *Let $f : A^* \to B^*$. The mapping f is called a HDT0L sequence iff there exists a finite alphabet C, a homomorphism $H : A^* \to Hom(C^*, C*)$, an homomorhism $h \in Hom(C^*, B^*)$ and a letter $c \in C$ such that, for every $w \in A^*$*

$$f(w) = h(H^w(c)).$$

(here we denote by H^w the image of w by H). The mapping f is called a DT0L when $B = C$ and the homomorphism h is just the identity; f is called a HD0L when A is reduced to one element.

3 Sequences of level 1

Let us mention, just for sake of completeness, the description of level 1 of our hierarchy of mappings.

Theorem 1. *The elements of $\mathbb{S}_1(A^*, B^*)$ are exactly the generalized sequential mappings from A^* to B^*.*

4 Sequences of level 2

N. Marin has shown in her Master thesis ([Mar07]) that $\mathbb{S}_2(A^*, B^*)$ has several nice characterisations.

Theorem 2 ([Mar07]). *Let us consider a mapping $f : A^* \to B^*$. The following properties are equivalent:*
1- $f \in \mathbb{S}_2(A^*, B^*)$
2- There exists a finite family $(f_i)_{i\in[1,n]}$ of mappings $A^* \to B^*$ which fulfils a system of catenative recurrent relations and such that $f = f_1$
3- f is a HDT0L sequence.

This theorem specializes as follows in the particular cases where A or B is reduced to one letter.

Corollary 1. *Let us consider a mapping $f : A^* \to \mathbb{N}$. The following properties are equivalent:*
1- $f \in \mathbb{S}_2(A^*, \mathbb{N})$
2- f is a $\mathbb{N}$-rational power series with non-commutative undeterminates in A

Corollary 2. *Let us consider a mapping* $f : \mathbb{N} \to B^*$*. The following properties are equivalent:*
1- $f \in \mathbb{S}_2(\mathbb{N}, B^*)$
2- There exists a finite family $(f_i)_{i\in[1,n]}$ *of sequences* $\mathbb{N} \to B^*$ *which fulfils a system of catenative recurrent relations and such that* $f = f_1$
3- f *is a HD0L sequence.*

Since J. Honkala has proved that the equivalence for HDT0L sequences is decidable ([Hon00]), theorem 2 implies

Theorem 3. *The equality problem is decidable for mappings in* $\mathbb{S}_2(A^*, B^*)$*.*

5 Sequences of level 3

Theorem 4. *Let us consider a mapping* $f : A^* \to B^*$*. The following properties are equivalent:*
1- $f \in \mathbb{S}_3(A^*, B^*)$
2- There exists a finite family $(H_i)_{i\in[1,n]}$ *of mappings* $A^* \to Hom(C^*, C^*)$ *which fulfils a system of recurrent relations in* $(Hom(C^*, C^*), \circ, \mathrm{Id})$ *, an element* $h \in Hom(C^*, B^*)$ *and a letter* $c \in C$ *such that, for every* $w \in A^*$*:*

$$f(w) = h(H_1(w)(c)).$$

3- f *is a composition of a DT0L sequence* $g : A^* \to C^*$ *by a HDT0L sequence* $h : C^* \to B^*$*.*

This theorem specializes as follows in the particular cases where B is reduced to one letter i.e. when the mapping f is a formal power series.

Corollary 3. *Let us consider a mapping* $f : A^* \to \mathbb{N}$*. The following properties are equivalent:*
1- $f \in \mathbb{S}_3(A^*, \mathbb{N})$
2- There exists a finite family $(H_i)_{i\in[1,n]}$ *of mappings* $A^* \to Hom(C^*, C^*)$ *which fulfils a system of recurrent relations in* $(Hom(C^*, C^*), \circ, \mathrm{Id})$ *, an element* $h \in Hom(C^*, \mathbb{N})$ *and a letter* $c \in C$ *such that, for every* $w \in A^*$*:*

$$f(w) = h(H_1(w)(c)).$$

3- f *is composition of a DT0L sequence* $g : A^* \to C^*$ *by a rational series* $h : C^* \to \mathbb{N}$*.*
4- There exists a finite family $(f_i)_{i\in[1,n]}$ *of mappings* $A^* \to \mathbb{N}$ *fulfilling a system of polynomial recurrent relations and such that* $f = f_1$*.*

Definition 9. *Let* $\mathbb{S}$ *be a set of mappings* $A^* \to B^*$*. We denote by* $\mathcal{D}(\mathbb{S})$ *the set of mappings of the form:*

$$f(w) = g(w) - h(w) \quad \textit{for all } w \in A^*,$$

for some mappings $g, h \in \mathbb{S}$. We denote by $\mathcal{F}(\mathbb{S})$ the set of mappings of the form:

$$f(w) = \frac{g(w) - h(w)}{f'(w) - g'(w)} \quad \text{for all } w \in A^*,$$

for some mappings $f, g, f', g' \in \mathbb{S}$.

Using point 4 of corollary 3 we can prove the following

Theorem 5. *The equality problem is decidable for formal power series in $\mathcal{F}(\mathbb{S}_3(A^*, \mathbb{N}))$.*

The method consists, in a way similar to [Sén99] or [Hon00], in reducing such an equality problem to deciding whether some polynomial belongs to the ideal generated by a finite set of other polynomials.

6 Sequences of level k

Theorem 6. *Let us consider a mapping $f : A^* \to B^*$. The following properties are equivalent:*
1- $f \in \mathbb{S}_k(A^*, B^*)$
2- f is a composition of $k-1$ HDT0L sequences $g_1 : A^* \to C_1^*, \ldots, g_i : C_{i-1}^* \to C_i^*, \ldots, g_{k-1} : C_{k-2}^* \to B^*$. Moreover the $g_1, \ldots, g_{k-2}$ can be chosen to be DT0L's.

From this theorem follows easily the fact that the inclusion $S_k(\mathbb{N}, \mathbb{N}) \subset S_{k+1}(\mathbb{N}, \mathbb{N})$ is strict.

7 Examples and counter-examples

We examine here four examples of mappings $A^* \to \mathbb{N}$ and locate them in the classes $\mathbb{S}_k$ or some related classes.

Example 1. The Fibonacci sequence F_n defined by

$$F_0 = 1, F_1 = 1, F_{n+2} = F_{n+1} + F_n \text{ for all } n \geq 0$$

is clearly in $\mathbb{S}_2$ since $\sum_{n=0}^{\infty} F_n X^n$ is a rational series.

Example 2. Let $G : \{0,1\}^* \to \mathbb{N}$ be defined by

$$G(w) = F_{\nu(w)}$$

where $\nu(w)$ is the natural number expressed by w in base 2. Since $\sum_{n=0}^{\infty} \nu(w)w$ is a rational series, G fulfils point 3 of our characterisation of $\mathbb{S}_3(\{0,1\}^*, \mathbb{N})$.

Example 3. Let us consider the sequence of Catalan numbers: $C_n = \frac{1}{n+1}C_{2n}^n$. This sequence is not residually ultimately periodic ([Ber03]) while we can prove that every sequence in $\bigcup_{k\in\mathbb{N}} \mathbb{S}_k(\mathbb{N},\mathbb{N})$ is residually ultimately periodic. By the same arguments C_n cannot belong to $\mathcal{D}(\mathbb{S}_k)$ for any $k \geq 1$. Nevertheless C_n belongs to $\mathcal{F}(\mathbb{S}_3)$.

Example 4. Let us consider the sequence $D_n = n^n$. This sequence belongs to $\mathbb{S}_4$. It does not belong to $\mathcal{F}(\mathbb{S}_3)$ because, for every $r \geq 1$, the sequences $D_n, D_{n+1}, \ldots, D_{n+r-1}$ are algebraically independent over $\mathbb{Q}$.

References

[Aho68] A.V. Aho. Indexed grammars-an extension of context-free grammars. *J. Assoc for Comput. Mach. 15*, pages 647–671, 1968.

[Aho69] A.V. Aho. Nested stack automata. *J. Assoc for Comput. Mach. 16*, pages 383–406, 1969.

[Ber03] J. Berstel. Personnal communication. 2003.

[Cau02] D. Caucal. On infinite terms having a decidable monadic theory. In *Proceedings MFCS 2002*, pages 165–176. Springer-Verlag, 2002.

[CW03] A. Carayol and S. Wöhrle. The Caucal hierarchy of infinite graphs in terms of logic and higher-order pushdown automata. In *FST TCS 2003: Foundations of software technology and theoretical computer science*, volume 2914 of *Lecture Notes in Comput. Sci.*, pages 112–123. Springer, Berlin, 2003.

[Dam82] W. Damm. The IO- and OI-hierarchies. *TCS 20*, pages 95–207, 1982.

[DG86] W. Damm and A. Goerdt. An automata-theoretical characterization of the OI-hierarchy. *Information and control 71*, pages 1–32, 1986.

[Eng83] J. Engelfriet. Iterated pushdown automata and complexity classes. In *Proc. 15th Annu ACM Sympos*, pages 365–373. Theory Comput, Assoc. Comp Mach., 1983.

[ES77] J. Engelfriet and E.M. Schmidt. IO and OI. I. *J. Comput. System Sci.*, 15(3):328–353, 1977.

[ES84] J. Engelfriet and G. Slutski. Extended macro grammars and stack controlled machines. *Journal of Computer and System sciences 29*, pages 366–408, 1984.

[EV86] J. Engelfriet and H. Vogler. Pushdown machines for the macro tree transducer. *TCS 42*, pages 251–368, 1986.

[Fis68] M.J. Fischer. *Grammars with macro-like productions*. PhD thesis, Harvard University, 1968. See also:Proc. 9th Symp. on Switching and Automata Theory (1968) p. 131-142.

[Fra05] S. Fratani. *Automates à piles de piles ... de piles*. PhD thesis, Bordeaux 1 university, 2005.

[FS06] S. Fratani and G. Sénizergues. Iterated pushdown automata and sequences of rational numbers. *Ann. Pure Appl. Logic*, 141(3):363–411, 2006.

[Gre70] S. Greibach. Full AFL's and nested iterated substitution. *Information and Control*, pages 7–35, 1970.

[Hon00] J. Honkala. A short solution for the HDT0L sequence equivalence problem. *Theoret. Comput. Sci.*, 244(1-2):267–270, 2000.

[KNU02] T. Knapik, D. Niwinski, and P. Urzyczyn. Higher-order pushdown trees are easy. In *FoSSaCs 2002*. LNCS, 2002.

[KRS97] L. Kari, G. Rozenberg, and A. Salomaa. L systems. In *Handbook of formal languages, Vol. 1*, pages 253–328. Springer, Berlin, 1997.

[Mar07] N. Marin. Suites de mots et automates. pages 1–40, 2007. Master thesis. Bordeaux 1 university.

[Mas74] A.N. Maslov. The hierarchiy of indexed languages of an arbitrary level. *Soviet. Math. Dokl. 15*, pages 1170–1174, 1974.

[Mas76] A.N. Maslov. Multilevel stack automata. *Problemy Peredachi Informatsii 12*, pages 38–43, 1976.

[MW67] J. Mezei and J. B. Wright. Algebraic automata and context-free sets. *Information and Control*, 11:3–29, 1967.

[PWZ96] M. Petkovšek, H.S. Wilf, and D. Zeilberger. $A = B$. A K Peters Ltd., Wellesley, MA, 1996.

[Sén99] G. Sénizergues. T(A) = T(B)? In P. van Emde Boas J. Wiedermann and M. Nielsen, editors, *Proceedings ICALP'99*, volume 1644, pages 665–675. Lecture Notes in Computer Science, 1999. Detailed version in technical report 1209-99,Pages 1-61. Can be accessed at URL,http://www.labri.u-bordeaux.fr/∼ges.

[Sta80] R.P. Stanley. Differentiably finite power series. *European Journal of Combinatorics 1*, pages 175–188, 1980.