



# Pushdown automata of order 2 and applications to recognizability

J. Ferté, N. Marin, G. Sénizergues

Université de Bordeaux, LaBRI,

06 Décembre 2011

Projet FREC, 5-6 Décembre 2011

# contents

- 1 The results
- 2 Preliminaries
- 3 Characterization theorem
- 4 Skimming theorem
- 5 Rational sets of endomorphisms
- 6 Quadratic word equations

# THE RESULTS

# Context

[Fratani-Sénizergues APAL'06 Integer sequences of level  $k$ ]

[Sénizergues CSR'07 Word sequences of level  $k$ ]

## Theorem

*A map  $f : X^* \rightarrow Y^*$  has level  $k \geq 2$  iff it is a  $(k - 1)$ -fold composition of maps of level 2.*

## Theorem

Let  $X, Y$  be two finite alphabets,  $f : X^* \rightarrow Y^*$  a mapping. The following conditions are equivalent :

- 1-  $f$  belongs to  $\mathbb{S}_2(X^*, Y^*)$  i.e. is computed by some strongly deterministic *pushdown of pushdown automaton*
- 2-  $f$  is a *HDTOL*
- 3-  $f$  fulfills a system of *catenative recurrent* equations

A1: **skimming** theorem for rational series [Schutzenberger Cong. Int. Math, 1970].

New proof + extension to HDTOL's.

A2: **rational subsets** of  $\text{HOM}(A^*, A^*)$  are not closed under intersection

A3: the set of solutions of a **quadratic equation**, over words, is an index language.

# PRELIMINARIES

Formalism: [Damm-Goerdts Information and Control 86]

The set of **pushdown stores** of level 2 over  $\Gamma$  is:

$$2 - \text{pds}(\Gamma) := (\Gamma[\Gamma^*])^*$$

Function:

$$\text{topsyms} : 2 - \text{pds}(\Gamma) \rightarrow \Gamma^{\leq 2}$$

Operations:

$$\text{pop}_1, \text{pop}_2, \text{push}_1, \text{push}_2$$

Examples:

$$\text{pop}_1(A_0[xyy]A_1[xyy]A_1[\varepsilon]) = A_1[xyy]A_1[\varepsilon]$$

$$\text{pop}_2(A_0[xyy]A_1[xyy]A_1[\varepsilon]) = A_0[yy]A_1[xyy]A_1[\varepsilon].$$

$$\text{push}_1(A_1)(A_0[xyy]A_1[xyy]A_1[\varepsilon]) = A_1[xyy]A_0[xyy]A_1[xyy]A_1[\varepsilon])$$

$$\text{push}_2(y)(A_0[xyy]A_1[xyy]A_1[\varepsilon]) = A_0[yyy]A_1[xyy]A_1[\varepsilon])$$

$$\text{push}_2(xy)(A_0[xyy]A_1[xyy]A_1[\varepsilon]) = A_0[xyyy]A_1[xyy]A_1[\varepsilon])$$



# strong determinism

Let  $\mathbb{A} = (Q, X, \Gamma, \delta, q_0, A_0)$  be a 2-pda.

$\mathbb{A}$  will be said **strongly** deterministic if :

$\forall q \in Q, \forall \gamma \in \text{TOPSYMB}(\Gamma),$

$$\sum_{\bar{a} \in X^{\leq 1}} \text{Card}(\delta(q, \bar{a}, \gamma)) \leq 1$$

Note: very close to a pushdown transducer of level 2 from strings to strings; just see  $\Gamma$  as the input alphabet,  $X$  as the output alphabet.

## Definition (2-computable sequences)

A mapping  $f : X^* \rightarrow Y^*$  is called a **2-computable mapping** (or sequence) iff there exists a *2-sdpda*  $\mathbb{A}$ , over a pushdown-alphabet  $\Gamma \supseteq X$ , with terminal alphabet  $Y$ , such that, for every  $u \in X^*$ :

$$(q_0, f(u), A_0[u]) \vdash_{\mathbb{A}}^* (q_0, \varepsilon, \varepsilon).$$

We denote by  $\mathbb{S}_2(X^*, Y^*)$  the set of all 2-computable sequences mappings from  $X^*$  to  $Y^*$ .

## Definition (catenative recurrent relations)

Given a finite set  $I$  and a family of mappings indexed by  $I$ ,  $f_i : X^* \rightarrow Y^*$  (for  $i \in I$ ), we call **system of catenative recurrent relations** over the family  $(f_i)_{i \in I}$  a system of the form

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

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

### Definition (HDTOL, [Karhumaki and alii Handbook, 1997])

Let  $f : X^* \rightarrow Y^*$ . The mapping  $f$  is called a **HDTOL** mapping (or sequence) iff there exists a finite alphabet  $A$ , a homomorphism  $H : X^* \rightarrow \text{Hom}(A^*, A^*)$ , an homomorphism  $h \in \text{Hom}(A^*, Y^*)$  and a letter  $a \in A$  such that, for every  $w \in X^*$

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

# CHARACTERIZATION THEOREM

# Main theorem

## Theorem

Let  $X, Y$  be two finite alphabets,  $f : X^* \rightarrow Y^*$  a mapping. The following conditions are equivalent :

- 1-  $f$  belongs to  $\mathbb{S}_2(X^*, Y^*)$  i.e. is computed by some strongly deterministic *pushdown of pushdown automaton*
- 2-  $f$  is a *HDTOL*
- 3-  $f$  fulfills a system of *catenative recurrent equations*

## Example: automaton

$$\mathbb{A} := (Q, \{a, b\}, \Gamma, \delta, q_0, A_0):$$

$$Q := \{q_0, q_1, q_{0,x}, q_{1,x}, q_{0,y}, q_{1,y}\}$$

$$\Gamma := \{A_0, A_1, x, y, a, b\}$$

$$\delta(q_0, \varepsilon, A_i z) := (q_{i,z}, \text{pop}_2) \quad \text{for } i \in \{0, 1\}, z \in \{x, y\}$$

$$\delta(q_{0,x}, \varepsilon, A_0 \bar{a}) := (q_0, \text{push}_1(A_0 A_1)) \quad \text{for } \bar{a} \in \{x, y, \varepsilon\}$$

$$\delta(q_{0,y}, \varepsilon, A_0 \bar{a}) := (q_0, \text{push}_1(A_1)) \quad \text{for } \bar{a} \in \{x, y, \varepsilon\}$$

$$\delta(q_{1,x}, \varepsilon, A_1 \bar{a}) := (q_0, \text{push}_1(A_1 A_0)) \quad \text{for } \bar{a} \in \{x, y, \varepsilon\}$$

$$\delta(q_{1,y}, \varepsilon, A_1 \bar{a}) := (q_0, \text{push}_1(A_1 A_0)) \quad \text{for } \bar{a} \in \{x, y, \varepsilon\}$$

$$\delta(q_0, a, A_0) := (q_0, \text{pop}_1)$$

$$\delta(q_0, b, A_1) := (q_0, \text{pop}_1)$$

# Example: derivations

Basic derivations:

$$(q_0, A_0[xu], q_0) \rightarrow_{\mathbb{A}}^* (q_0, A_0[u], q_0)(q_0, A_1[u], q_0)$$

$$(q_0, A_0[yu], q_0) \rightarrow_{\mathbb{A}}^* (q_0, A_1[u], q_0)$$

$$(q_0, A_1[xu], q_0) \rightarrow_{\mathbb{A}}^* (q_0, A_1[u], q_0)(q_0, A_0[u], q_0)$$

$$(q_0, A_1[yu], q_0) \rightarrow_{\mathbb{A}}^* (q_0, A_1[u], q_0)(q_0, A_0[u], q_0)$$

$$(q_0, A_0[\varepsilon], q_0) \rightarrow_{\mathbb{A}}^* a$$

$$(q_0, A_1[\varepsilon], q_0) \rightarrow_{\mathbb{A}}^* b.$$



Example:  $f_0(x^3)$ 

$$(q_0, A_0[xu], q_0) \rightarrow_{\mathbb{A}}^* (q_0, A_0[u], q_0)(q_0, A_1[u], q_0)$$

$$(q_0, A_0[yu], q_0) \rightarrow_{\mathbb{A}}^* (q_0, A_1[u], q_0)$$

$$(q_0, A_1[xu], q_0) \rightarrow_{\mathbb{A}}^* (q_0, A_1[u], q_0)(q_0, A_0[u], q_0)$$

$$(q_0, A_1[yu], q_0) \rightarrow_{\mathbb{A}}^* (q_0, A_1[u], q_0)(q_0, A_0[u], q_0)$$

Therefore we get:

$$(q_0, A_0[xxx], q_0) \rightarrow_{\mathbb{A}}^* (q_0, A_0[xx], q_0)(q_0, A_1[xx], q_0)$$

$$\rightarrow_{\mathbb{A}}^* (q_0, A_0[x], q_0)(q_0, A_1[x], q_0)(q_0, A_1[x], q_0)(q_0, A_0[x], q_0)$$

$$\rightarrow_{\mathbb{A}}^* (q_0, A_0[\varepsilon], q_0)(q_0, A_1[\varepsilon], q_0)^2(q_0, A_0[\varepsilon], q_0) \cdot$$

$$(q_0, A_1[\varepsilon], q_0)(q_0, A_0[\varepsilon], q_0)^2(q_0, A_1[\varepsilon], q_0)$$

$$\rightarrow_{\mathbb{A}}^* \text{abbabaab.}$$

## Example: morphisms

The corresponding **HDTOL**:

$X := \{x, y\}$ ,  $Y := \{a, b\}$ ,  $\Phi : X^* \rightarrow \text{HOM}(Y^*, Y^*)$  is defined by

$$\Phi^x(a) := ab, \quad \Phi^x(b) := ba, \quad \Phi^y(a) := b, \quad \Phi^y(b) := ba.$$

We obtain the family  $(f_i)_{i \in \{0,1\}}$  of HDTOL maps:

$$f_0(u) := \Phi^u(a), \quad f_1(u) := \Phi^u(b).$$

(Note that  $n \mapsto f_0(x^n)$  is the Thue-Morse sequence and  $n \mapsto f_0(y^n)$  is the Fibonacci sequence).

## Example: catenative recurrence

These maps fulfill the system of **catenative relations**:

$$\left\{ \begin{array}{l} f_0(xu) = f_0(u)f_1(u) \\ f_0(yu) = f_1(u) \\ f_1(xu) = f_1(u)f_0(u) \\ f_1(yu) = f_1(u)f_0(u) \end{array} \right.$$

# SKIMMING THEOREM

# rational series:recall

## Definition

Let  $X$  be a finite alphabet and  $f : X^* \rightarrow \mathbb{N}$  be some map. The map  $f$  is a **rational series** iff,  $f$  belongs to some **finitely generated**  $\mathbb{N}$ -submodule  $M$  of  $\mathbb{N} \langle\langle X \rangle\rangle$ , which is **closed under left-residual**.

## Theorem

Let  $X$  be a finite alphabet,  $f : Y^* \rightarrow \mathbb{N}$  be some map and  $f' : Y^* \rightarrow \mathbb{N}$  be defined by:

$$f'(u) := \max\{0, f(u) - 1\}$$

If  $f$  is a  $\mathbb{N}$ -rational series, then  $f'$  is  $\mathbb{N}$ -rational too.

First proof: [Schutzenberger Cong. Int. Math, 1970]

Other proof: [Sakarovitch-De Souza STACS 2008]

## Definition

Let  $X, Y$  be finite alphabets and  $f : X^* \rightarrow Y^*$  be some map. The map  $f$  is a **HDTOL** iff,  $f$  belongs to some **finitely generated** submonoid  $M$  of  $\langle \text{MAP}(X^*, Y^*), \cdot, w \mapsto \varepsilon \rangle$ , which is **closed under left-residual** by pairs  $(x, \varepsilon)$  (for  $x \in X$ ).

## Theorem

Let  $X, Y$  be finite alphabets. The set of HDTOL's from  $X^*$  to  $Y^*$  is closed under left-residual by *pairs*  $(\varepsilon, y)$  (for  $y \in Y$ )



## Theorem

Let  $X, Y, Z, T$  be four finite alphabets and  $f : X^* \rightarrow Y^*, g : Y^* \rightarrow Z^*, h : Z^* \rightarrow T^*$ , be some maps. If  $f, h$  are *rational maps* and  $g$  is an *HDTOL*, then  $f \circ g \circ h$  is an *HDTOL*.

**Sketch of proof:**

Case 1:  $f$  is a gsm

use the (standard) definition of [HDTOL's](#)

Case 2:  $h$  is a r-to-l gsm

use [catenative recurrences](#)

Case 3:  $h$  is a l-to-r gsm

use [sdpda](#) of level 2.

Principle: direct product of the spda with the gsm.  $\diamond$

# RATIONAL SETS of ENDOMORPHISMS

Question: Is the set of **rational subsets** of  $\langle \text{HOM}(A^*, A^*), \circ, \text{Id} \rangle$  closed under **intersection** ?

Notation:  $w \diamond f$  means  $f(w)$ .

## Lemma

For every rational subset  $R$  of  $\text{HOM}(A^*, A^*)$  and every  $w \in A^*$ , the set  $\{w \diamond f \mid f \in R\}$  is an *indexed-language*.

### Sketch of proof:

There exists some finite alphabet  $X$ , a monoid homomorphism  $H : X^* \rightarrow \text{HOM}(A^*, A^*)$  and a rational subset  $R_X \subseteq X^*$  such that

$$R = \{H^u \mid u \in R_X\}.$$

The language  $L$  can be rewritten as

$$L = \{w \diamond H^u \mid u \in R_X\}.$$

$L$  is recognized by an adaptation of the 2-dpda computing  $u \mapsto (w \diamond H^u)$ .  $\diamond$

$$A := \{x, y, z, t\}$$

$$h_1 : \quad x \mapsto xy, \quad y \mapsto y, \quad z \mapsto z, \quad t \mapsto t,$$

$$h_2 : \quad x \mapsto x, \quad y \mapsto yx, \quad z \mapsto z, \quad t \mapsto t,$$

$$h_3 : \quad x \mapsto x, \quad y \mapsto y, \quad z \mapsto zt, \quad t \mapsto t,$$

Let

$$R := (h_1 \circ h_3)^* \circ h_2^*, \quad S := h_1^* \circ (h_2 \circ h_3)^*$$

$$\begin{aligned}h_1 : & \quad x \mapsto xy, & y \mapsto y, & z \mapsto z, & t \mapsto t, \\h_2 : & \quad x \mapsto x, & y \mapsto yx, & z \mapsto z, & t \mapsto t, \\h_3 : & \quad x \mapsto x, & y \mapsto y, & z \mapsto zt, & t \mapsto t,\end{aligned}$$

We observe that  $h_1 \circ h_3 = h_3 \circ h_1$  and  $h_2 \circ h_3 = h_3 \circ h_2$ .

$$R \cap S = \{h_1^n h_2^n h_3^n \mid n \geq 0\}.$$

It follows that

$$\{x \diamond h \mid h \in R \cap S\} = \{x(yx^n)^n \mid n \geq 0\}.$$

But this last language has been shown to do **not be an indexed-language** in:

[Hayashi 73 Th. 5.3]

[Gilman 96 Corollary 4].

By Lemma 6:  $R \cap S$  is **not rational**.

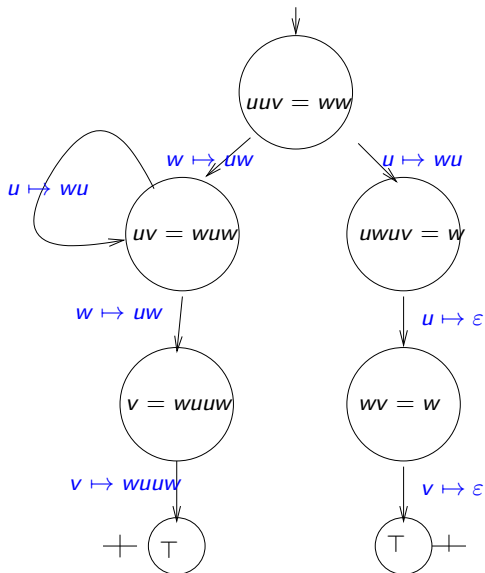


# QUADRATIC WORD EQUATIONS

Equation:

$$S : uv = vw$$

## The automaton:



$$\text{SOL}_h(\mathcal{S}) = (f + g \circ h^* \circ k) \circ \text{HOM}(U^*, U^*).$$

where

$$\begin{array}{lll} f : & u \mapsto w, & v \mapsto \varepsilon, & w \mapsto w, \\ g : & u \mapsto u, & v \mapsto v, & w \mapsto uw, \\ h : & u \mapsto wu, & v \mapsto v, & w \mapsto w, \\ k : & u \mapsto u, & v \mapsto wuuv, & w \mapsto uw. \end{array}$$

$$\begin{aligned}
 \text{SOL}_w(\mathcal{S}) &:= \{\varphi(u)\#\varphi(v)\#\varphi(w) \mid \varphi \in \text{SOL}_h\} \\
 \text{SOL}_w(\mathcal{S}) &= \{\bar{w}\#\varepsilon\#\bar{w} \mid \bar{w} \in \mathcal{U}^*\} \cup \\
 &\quad \{(\bar{u}\bar{w})^n\bar{u}\#\bar{w}\bar{u}^2\bar{w}\#(\bar{u}\bar{w})^n\bar{u}(\bar{u}\bar{w}) \mid \bar{u}, \bar{v}, \bar{w} \in \mathcal{U}^*\}.
 \end{aligned}$$

## Theorem

*For every finite quadratic system of word equations  $\mathcal{S}$ , the set of hom-solutions  $\text{SOL}_h(\mathcal{S})$  is *rational*.*

(almost true)

follows from [Diekert Lothaire chap. 12]

## Theorem

*For every finite quadratic system of word equations  $\mathcal{S}$ , the set of word-solutions  $\text{SOL}_w(\mathcal{S})$  is an *indexed-language*.*

(perfectly true)