A Generalised Twinning Property for Minimisation of Cost Register Automata

Laure Daviaud
LIP, ENS Lyon

Joint work with P-A.Reynier and J-M.Talbot
LIF, Aix-Marseille Université

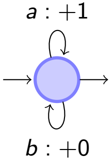
LABRI, February 2016

A first example

Given a word $w \in \{a, b\}^*$, compute $\{|w|_a, |w|_b\}$.

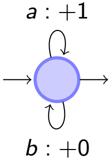
A first example

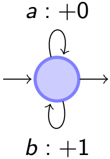
Given a word $w \in \{a, b\}^*$, compute $\{|w|_a, |w|_b\}$.



A first example

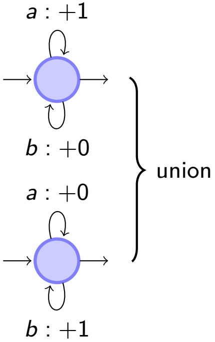
Given a word $w \in \{a, b\}^*$, compute $\{|w|_a, |w|_b\}$.





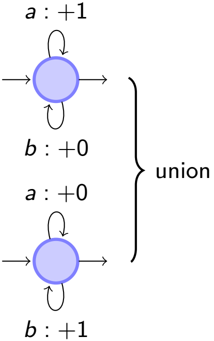
A first example

Given a word $w \in \{a, b\}^*$, compute $\{|w|_a, |w|_b\}$.

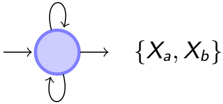


A first example

Given a word $w \in \{a, b\}^*$, compute $\{|w|_a, |w|_b\}$.



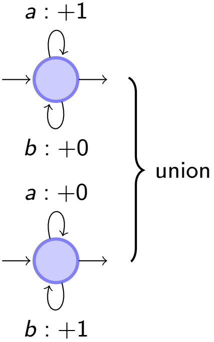
$$a : \begin{cases} X_a := X_a + 1 \\ X_b := X_b \end{cases}$$

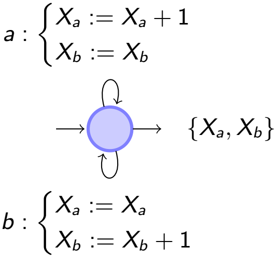


$$b : \begin{cases} X_a := X_a \\ X_b := X_b + 1 \end{cases}$$

A first example

Given a word $w \in \{a, b\}^*$, compute $\{|w|_a, |w|_b\}$.





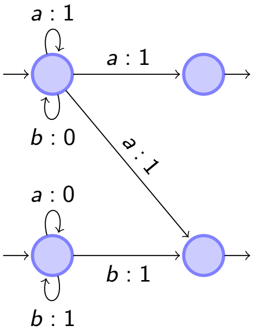
Question: How many values do we need to keep in memory?

Weighted automata (in a restricted case)

Non deterministic finite automaton whose transitions are weighted by elements of a semiring $\longrightarrow$ here $(\mathcal{P}_f(G), \cup, \cdot)$, with $(G, \cdot)$ a group

Weight of a run ρ: $\omega(\rho) =$ product of the weights of the transitions

Function: $w \mapsto \{\omega(\rho) \quad | \ \rho$ accepting run labelled by $w \quad \}$



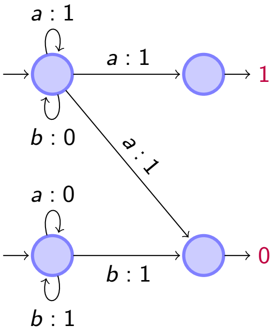
$(\mathbb{Z}, +)$

Weighted automata (in a restricted case)

Non deterministic finite automaton whose transitions are weighted by elements of a semiring $\longrightarrow$ here $(\mathcal{P}_f(G), \cup, \cdot)$, with $(G, \cdot)$ a group + an function t from the final states to G.

Weight of a run ρ: $\omega(\rho) =$ product of the weights of the transitions

Function: $w \mapsto \{\omega(\rho)t(q) \mid \rho$ accepting run labelled by w ending in $q\}$



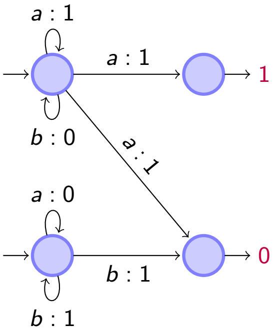
$(\mathbb{Z}, +)$

Weighted automata (in a restricted case)

Non deterministic finite automaton whose transitions are weighted by elements of a semiring $\longrightarrow$ here $(\mathcal{P}_f(G), \cup, \cdot)$, with $(G, \cdot)$ a group
$+$ an function t from the final states to G.

Weight of a run ρ: $\omega(\rho) =$ product of the weights of the transitions

Function: $w \mapsto \{\omega(\rho)t(q) \mid \rho$ accepting run labelled by w ending in $q\}$



$(\mathbb{Z}, +)$

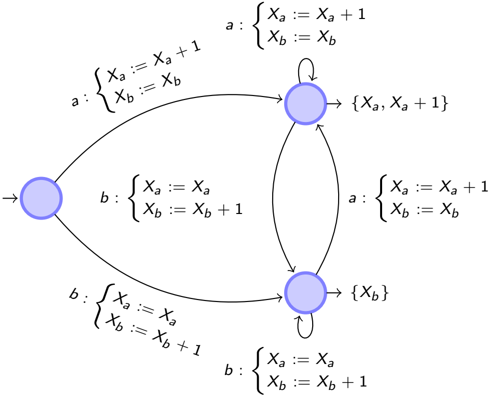
$[\![\mathcal{W}]\!](w) = \{|w|_a, |w|_a + 1\}$
if w ends with an a

$[\![\mathcal{W}]\!](w) = \{|w|_b\}$
if w ends with a b

Cost register automata (in a restricted case too)

Deterministic finite state machine with registers + an output function

Register updates: $X := Y\alpha$ with $\alpha \in G$.



Infinitary group

Definition

A group G is said to be **infinitary** if for all $\alpha, \beta, \gamma \in G$ such that $\alpha\beta\gamma \neq \beta$, the set $\{\alpha^n \beta \gamma^n \mid n \in \mathbb{N}\}$ is infinite.

- $(\mathbb{Z}, +)$, $(\mathbb{R}, \times)$ are infinitary.
- The free group generated by a finite set is infinitary.

Infinitary group

Definition

A group G is said to be **infinitary** if for all $\alpha, \beta, \gamma \in G$ such that $\alpha\beta\gamma \neq \beta$, the set $\{\alpha^n \beta \gamma^n \mid n \in \mathbb{N}\}$ is infinite.

- $(\mathbb{Z}, +)$, $(\mathbb{R}, \times)$ are infinitary.
- The free group generated by a finite set is infinitary.

Why infinitary group ?... See later !

Several notions...

Valuedness (of a function)

Ambiguity (of a weighted automaton)

Register complexity (of a function)

Several notions...

Valuedness (of a function)

- ℓ-valued : for all words w, $|f(w)| \leqslant \ell$
- we only consider finite-valued functions

Ambiguity (of a weighted automaton)

Register complexity (of a function)

Several notions...

Valuedness (of a function)

- ℓ-valued : for all words w, $|f(w)| \leqslant \ell$
- we only consider finite-valued functions

Ambiguity (of a weighted automaton)

- number of accepting paths labelled by a word in a weighted automaton
- ℓ-valued $=$ ℓ-ambiguous [Filiot, Gentilini, Raskin]

Register complexity (of a function)

Several notions...

Valuedness (of a function)

- ℓ-valued : for all words w, $|f(w)| \leqslant \ell$
- we only consider finite-valued functions

Ambiguity (of a weighted automaton)

- number of accepting paths labelled by a word in a weighted automaton
- ℓ-valued $=$ ℓ-ambiguous [Filiot, Gentilini, Raskin]

Register complexity (of a function)

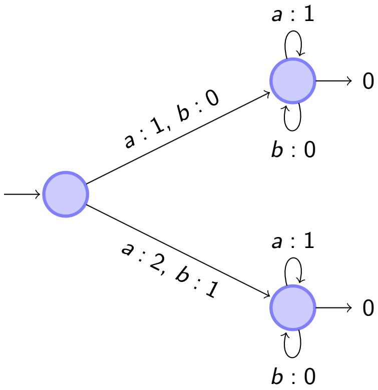
- minimal number of registers needed to compute a given function by a cost register automaton
- PSPACE-complete for one-valued additive register cost functions (unambiguous WA over $(\mathbb{Z}, +)$) [Alur, Raghothaman]

The question

Over an infinitary group,
characterise (effectively) the register complexity of a function computed by a finite-valued weighted automaton.

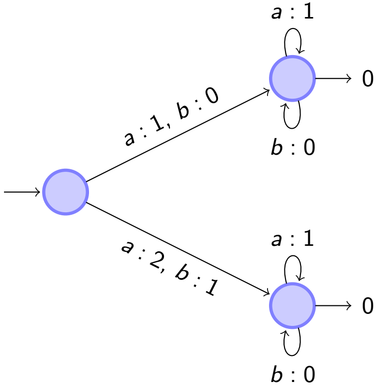
A very very simple example

$$w \mapsto \{|w|_a, |w|_a + 1\} \text{ in } (\mathbb{Z}, +)$$



A very very simple example

$$w \mapsto \{|w|_a, |w|_a + 1\} \text{ in } (\mathbb{Z}, +)$$



Definition

Given $\alpha,\ \beta \in G$, the **delay** between α and β is $\alpha^{-1}\beta$. It is denoted by $delay(\alpha, \beta)$.

Twinning property [Choffrut]

Definition

A weighted automaton satisfies the **twinning property** if for all initial states p, p' and co-accessible states q, q', for all words u, v such that:

$$p \xrightarrow{u:\alpha} q \xrightarrow{v:\beta} q$$

$$p' \xrightarrow{u:\alpha'} q' \xrightarrow{v:\beta'} q'$$

then $\mathit{delay}(\alpha, \alpha') = \mathit{delay}(\alpha\beta, \alpha'\beta')$

One register (cost register automata)
$=$ Deterministic (weighted automata)
$=$ TP (weighted automata)

Generalisation

Theorem

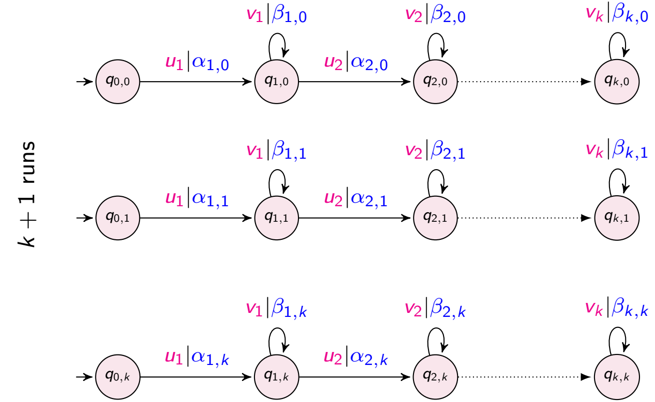
Let $\mathcal{W}$ be a finite-valued weighted automaton over an infinitary group, and k be a positive integer.
The following assertions are equivalent:

- $\mathcal{W}$ satisfies the twinning property of order k,
- $[\![\mathcal{W}]\!]$ has register complexity k,
- $[\![\mathcal{W}]\!]$ satisfies the k-bounded variation property,

And everything is effective...

Twinning Property of order k

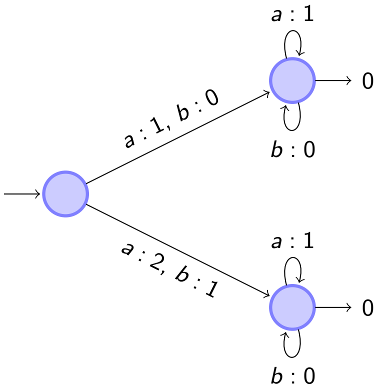
The weighted automaton satisfies the **twinning property of order** k if for all $q_{0,j}$ initial and $q_{k,j}$ co-accessible such that:



there are $j \neq j'$ such that for all $i \in \{1, \ldots, k\}$,

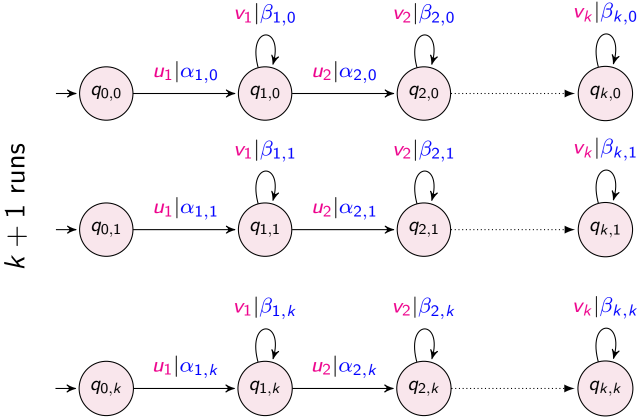
$$delay(\alpha_{1,j} \cdots \alpha_{i,j}, \alpha_{1,j'} \cdots \alpha_{i,j'}) = delay(\alpha_{1,j} \cdots \alpha_{i,j}\beta_{i,j}, \alpha_{1,j'} \cdots \alpha_{i,j'}\beta_{i,j'})$$

Twinning property of order k



- Commutative case Vs non commutative case
- Decidability
- Infinitary here !!!

Twinning property of order k



If the twinning property is not satisfied → construction of a sequence of words that have $k+1$ diverging behaviours [infinitarity + pumping the loops the right number of times]

Main result

Theorem

Let $\mathcal{W}$ be a finite-valued weighted automaton over an infinitary group, and k be a positive integer.
The following assertions are equivalent:

- $\mathcal{W}$ satisfies the twinning property of order k,
- $[\![\mathcal{W}]\!]$ has register complexity k,
- $[\![\mathcal{W}]\!]$ satisfies the k-bounded variation property,

And everything is effective...

Bounded variation property

Notions of distance:

- On words: $dist(u, v) = |u| + |v| - 2 * |lcp(u, v)|$ where $lcp(u, v)$ is the longest common prefix of the two words u and v.
- On a finitely generated group G with a finite set of generators Γ, $d(\alpha, \beta)$ is the minimal length of a path linking α and β in the undirected right Cayley graph of (G, Γ).

Bounded variation property

Notions of distance:

- On words: $dist(u, v) = |u| + |v| - 2 * |lcp(u, v)|$ where $lcp(u, v)$ is the longest common prefix of the two words u and v.
- On a finitely generated group G with a finite set of generators Γ, $d(\alpha, \beta)$ is the minimal length of a path linking α and β in the undirected right Cayley graph of (G, Γ).

Definition

A function $f : A^* \to \mathcal{P}_f(G)$ satisfies the k-**bounded variation property** if for all n, there is N such that for all words $w_0, \ldots, w_k \in A^*$ and all $\alpha_0 \in f(w_0), \ldots, \alpha_k \in f(w_k)$,
if for all $0 \leqslant i, j \leqslant k$, $dist(w_i, w_j) \leqslant n$ then there are $0 \leqslant i < j \leqslant k$ such that $d(\alpha_i, \alpha_j) \leqslant N$.

Bounded variation property

Notions of distance:

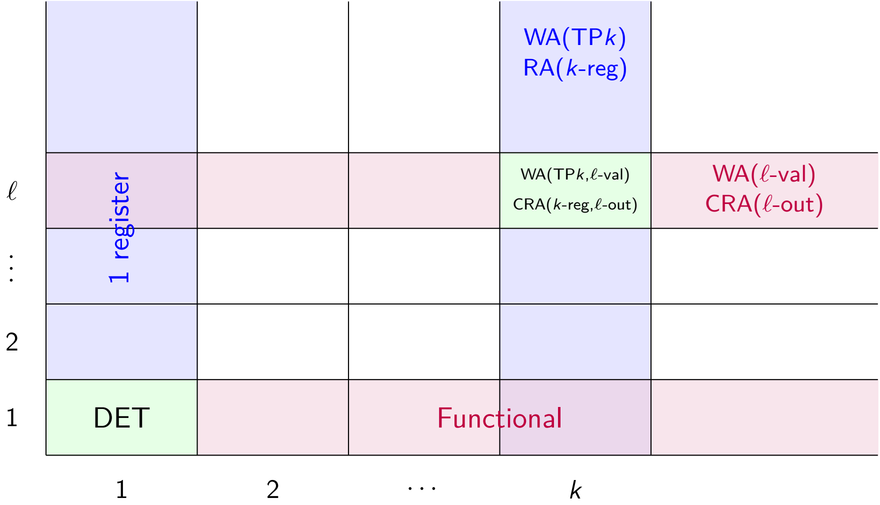
- On words: $dist(u, v) = |u| + |v| - 2 * |lcp(u, v)|$ where $lcp(u, v)$ is the longest common prefix of the two words u and v.
- On a finitely generated group G with a finite set of generators Γ, $d(\alpha, \beta)$ is the minimal length of a path linking α and β in the undirected right Cayley graph of (G, Γ).

Definition

A function $f : A^* \to \mathcal{P}_f(G)$ satisfies the k-**bounded variation property** if for all n, there is N such that for all words $w_0, \ldots, w_k \in A^*$ and all $\alpha_0 \in f(w_0), \ldots, \alpha_k \in f(w_k)$,
if for all $0 \leqslant i, j \leqslant k, dist(w_i, w_j) \leqslant n$ then there are $0 \leqslant i < j \leqslant k$ such that $d(\alpha_i, \alpha_j) \leqslant N$.

- Generalisation of the bounded variation property of Choffrut.
- A machine-independent characterisation of TPk.
- Finitely generated case Vs non finitely generated case.

Hierarchy


WA(TPk)
RA(k-reg)
WA(TPk,ℓ-val)
CRA(k-reg,ℓ-out)
WA(ℓ-val)
CRA(ℓ-out)
1 register
DET
Functional
ℓ
$\vdots$
2
1
1
2
$\cdots$
k

Transducers

Theorem

Let $\mathcal{T}$ be an ℓ-valued transducer from A^* to B^*, and k be a positive integer. The following assertions are equivalent:

- $\mathcal{T}$ satisfies the twinning property of order k,
- $[\![\mathcal{T}]\!]$ satisfies the k-bounded variation property,
- $[\![\mathcal{T}]\!]$ is computed by a cost register automaton over B^* with k registers and ℓ outputs.

Sketch of the proof

$\mathcal{W}$ finite-valued weighted automaton over an infinitary group.

First step: $\mathcal{W}$ satisfies the twinning property of order k
$\Leftrightarrow$ $[\![\mathcal{W}]\!]$ satisfies the k-bounded variation property

Second step: $\mathcal{W}$ satisfies the twinning property of order k
$\Leftrightarrow$ $[\![\mathcal{W}]\!]$ has register complexity k

- Register complexity k $\implies$ Twinning property of order k
- Twinning property of order k $\implies$ Register complexity k

Conclusion and open questions

- TPk $\Leftrightarrow$ BVk $\Leftrightarrow$ Register complexity k
  for infinitary groups and transducers... (at least)
- Minimisation of cost register automata

- Generalise the case of transducers
- Generalisation to visibly pushdown