The Fatou property of submonoids of $A^* \times B^*$

Luc Boasson and Olivier Carton

IRIF – CNRS and Université Paris Diderot

ANR DeLTA – Bordeaux
Janvier 2017
Outline

The problem

(Well) Known results

What’s new

Nasty examples

Technics

Motivation
Global vs Local constraints

Global constraints of input/output: \( \text{monoid } M \subseteq A^* \times B^* \)

A relation \( R \subseteq M \) realized by some transducer.

Length-preserving function

\[
\{a, b\}^* \rightarrow \{a, b\}^*
\]

\[
a_1a_2a_3 \cdots a_n \mapsto a_2a_3 \cdots a_na_1
\]

Here \( M = \{(u, v) : |u| = |v|\} \)

Question: can \( R \) be realized by a transducer with labels in \( M \)?

This question can be rephrased as the following **Fatou property**:

\[
R \in \text{Rat}(A^* \times B^*) \land R \subseteq M \quad \overset{?}{\Rightarrow} \quad R \in \text{Rat}(M)
\]
Transducers

Transitions $p \xrightarrow{a:v} q$ for $a \in A$, $v \in B^*$. 
Example
Example

\[
\begin{array}{c}
\text{0 : 0} \quad q_0 \quad 1 : 1 \quad q_1 \quad 1 : \varepsilon \\
\end{array}
\]

\[
\begin{array}{cccccccccc}
0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & \ldots\\
\end{array}
\]

\[
\begin{array}{c}
\text{q}_0 \\
\end{array}
\]

\[
\begin{array}{cccccccccc}
0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & \ldots\\
\end{array}
\]
Example
Example

\[
\begin{array}{c}
0: 0 \quad q_0 \\
\rightarrow \\
1: 1 \\
\rightarrow \\
0: 0 \\
\rightarrow \\
1: \varepsilon \\
q_1
\end{array}
\]

\[
\begin{array}{cccccccccc}
\vdots
& 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
\end{array}
\]
Example

\[
\begin{array}{c}
q_0 & \rightarrow & 1:1 \\
0:0 & \rightarrow & q_1 \\
0:0 & \rightarrow & \varepsilon \\
\end{array}
\]

---

\[
\begin{array}{cccccccc}
0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
\end{array}
\]

---

\[
\begin{array}{cccccccc}
0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\
\end{array}
\]
Example

\[
\begin{array}{c}
q_0 \\
0 : 0 \\
0 : 0 \\
\end{array}
\quad \begin{array}{c}
1 : 1 \\
q_1 \\
1 : \varepsilon \\
\end{array}
\]

\[
\begin{array}{ccccccccc}
0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
\end{array}
\]

\[
\begin{array}{c}
q_0 \\
\end{array}
\]

\[
\begin{array}{ccccccccc}
0 & 1 & 0 & 0 & 1 \\
\end{array}
\]
Example
Example

The diagram represents a finite automaton with two states, $q_0$ and $q_1$. The transitions are as follows:
- From $q_0$, on input 0, it stays in $q_0$.
- From $q_0$, on input 1, it goes to $q_1$.
- From $q_1$, on input 0, it stays in $q_1$.
- From $q_1$, on input 1, it goes to $q_1$.

The automaton accepts the language $L_2 = \{0^n1^n \mid n \geq 0\}$.
Example

\[ q_0 \quad \rightarrow \quad 1 : 1 \quad \rightarrow \quad q_1 \]

0 : 0  \quad \rightarrow \quad q_0 \quad  \rightarrow \quad 0 : 0  \quad \rightarrow \quad q_1 \quad  \rightarrow \quad 1 : \varepsilon

\[ \begin{array}{cccccccc}
0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\
\end{array} \]

\[ \begin{array}{cccccccc}
0 & 1 & 0 & 0 & 1 & 0 \quad \cdots
\end{array} \]
Length preserving relations

\[ M = \{(u, v) : |u| = |v|\} \]

Theorem (Elgot & Mezei 65)

The following conditions are equivalent for a rational relation \( R \)

\[ R \subseteq M \]

\[ R \text{ is realized by a transducer such that each transition } p \xrightarrow{u:v} q \text{ satisfies } (u, v) \in M, \text{ that is } |u| = |v|. \]
Length decreasing relations

\[ M = \{(u, v) : |u| \geq |v|\} \]

**Theorem (Leguy 81)**

The following conditions are equivalent for a rational relation \( R \)

\( R \subseteq M \)

\( R \) is realized by a transducer such that each transition \( p \xrightarrow{u:v} q \) satisfies \((u, v) \in M\), that is \( |u| \geq |v| \).

Key ingredients:

- The condition is satisfied for cycles.
- The delay is therefore bounded.
- The synchronized transducer uses two buffers of bounded size.
Stable monoids

A submonoid of $\mathbb{N}^k$ is called stable if one of the two equivalent statements hold.

- $N = C \cap H$ where $C$ closed cone of $\mathbb{N}^k$ and $H$ subgroup of $\mathbb{Z}^k$
- $\forall u \in N, \forall v \in \mathbb{N}^k \quad u + v^\oplus \subseteq N \implies v \in N$

Theorem (Sakarovitch)

If $M = \mu^{-1}(N)$ where $\mu : A^* \times B^* \to \mathbb{N}^k$ and $N$ stable monoid of $\mathbb{N}^k$, then $M$ has the Fatou property.

\[ R \in \text{Rat}(A^* \times B^*) \land R \subseteq M \implies R \in \text{Rat}(M) \]
Negative example

Let $M$ be the monoid

$$M = \{(a, \varepsilon), (\varepsilon, a), (b, b)\}^* = \{(u, v) : |u|_b = |v|_b\}$$

This transducer realizes the relation $R$ given by

$$R = \{(ba^n, a^n b) : n \geq 0\} \subset M.$$ 

This relation cannot be realized by a transducer over $M$. The monoid $M$ does not have the Fatou property.
A monoid $M \subseteq A^* \times B^*$ is quasi-unary if

$$(u, v) \in M \land |v| = |v'| \implies (u, v') \in M.$$
The subword monoid

A word \( v = a_1 \cdots a_n \) is a **subword** of a word \( u \), denoted \( v|u \), if there exist \( n + 1 \) words \( u_0, \ldots, u_n \) such that \( u = u_0 a_1 u_1 \cdots a_n u_n \).

\[ u \]
\[ \begin{array}{ccccccc}
    & u_0 & a_1 & u_1 & a_2 & u_2 & a_3 & \cdots & a_n & u_n \\
\end{array} \]

\[ v \]
\[ \begin{array}{ccccccc}
    & a_1 & a_2 & a_3 & \cdots & a_n \\
\end{array} \]

---

**Theorem**

*The monoid \( M = \{(u, v) : v|u\} \) has the Fatou property.*
Nasty examples

Let $\mu : A^* \to B^*$ be a morphism. The graph of $\mu$

$$M = \{(u, \mu(u)) : u \in A^*\}$$

has the Fatou property. What about finite substitutions?

$$M = \{(a, a), (a, b), (a, \varepsilon), (b, b)\}^*$$

$R = \{(a^{2m+1}ba^{2n+1}, ba^m n b) : m, n \geq 0\}$

The relation $R$ satisfies $R \subset M$ since $m + n \leq \max(2m, 2n)$ but it cannot realized by a transducer with labels in $M$.

$$M = \{(a, a), (a, \varepsilon), (b, b), (b, a)\}^*$$

Open question: does this monoid $M$ have the Fatou property?
Delays

A run $\rho$ of label $(u, v) = (u_1 \cdots u_n, v_1 \cdots v_n)$

$$q_0 \xrightarrow{u_1:v_1} q_1 \xrightarrow{u_2:v_2} q_2 \cdots q_{n-1} \xrightarrow{u_n:v_n} q_n$$

A $M$-factorization $\phi$ of $(u, v)$

$$(u, v) = (a_1, w_1) \cdots (a_m, w_m) \quad \text{where} \quad (a_i, w_i) \in M$$

The delay between $\rho$ and $\phi$ in $q_i$ is

$$d(\rho, \phi, i) = |v_1 \cdots v_i| - |w_1 \cdots w_j| \quad \text{where} \quad u_1 \cdots u_i = a_1 \cdots a_j$$

The delay between $\rho$ and $\phi$ is $d(\rho, \phi) = \max_{1 \leq i \leq n} d(\rho, \phi, i)$.  
The delay of $\rho$ is $d(\rho) = \min_\phi d(\rho, \phi)$

The delay of the transducer is $d(\mathcal{T}) = \max_\rho d(\rho)$
Delays (continued)

Proposition

Any transducer $\mathcal{T}$ of bounded delay $n$ such that $|\mathcal{T}| \subseteq M$ is equivalent to the transducer $\mathcal{T}_n$ over $M$.

For all known example ($M_\approx$, $M_\leq$, subword monoid), it is shown that if transducer $\mathcal{T}$ satisfies $|\mathcal{T}| \subseteq M$, then its delay must be bounded. What about the converse?

Proposition

Let $M$ be a fixed monoid. If there exists a computable function $f$ such that for any transducer $\mathcal{T}$, $|\mathcal{T}| \subseteq M$ implies that the delay of $\mathcal{T}$ is bounded by $f(\mathcal{T})$, then, it is decidable whether $|\mathcal{T}| \subseteq M$. 
Non-finitely generated monoid

For any non-negative real number $\alpha$, let $M_\alpha$

$$M_\alpha = \{(u, v) : |u| \leq \alpha |v|\}$$

Proposition

*For any irrational number $\alpha$, the monoid $M_\alpha$ is not finitely generated but it has the Fatou property.*

A rational non-finitely generated cannot have the Fatou property.
The weak Fatou property

A monoid $M$ has the weak Fatou property if

$$R \in \text{Rat}(A^* \times B^*) \land R \subseteq M \implies R \in \text{Rat}(\text{Rec}(M))$$

Proposition

The monoid $M = \{(\varepsilon, \varepsilon)\} \cup \{(u, v) : |u| \neq 0\}$ has the weak Fatou property.

For any non-negative real number $\alpha$, let $M_\alpha$

$$M'_\alpha = \{(u, v) : |u| < \alpha |v|\}$$

Proposition

For any rational number $\alpha$, the monoid $M'_\alpha$ is not finitely generated and it does not have the weak Fatou property.
Infinite words

\[ R = \{(a^n (b a)^\omega, b a^n b^\omega) : n \geq 0\} = (\varepsilon, b)(a, a)^*(b a, b)^\omega. \]

For each integer \( n \), the output \( \omega \)-word \( b a^n b^\omega \) is a subword of the input \( \omega \)-word \( a^n (b a)^\omega \). However, this relation cannot be realized by a transducer with labels in \( M = \{(u, v) : v|u\}\).
Normality (Borel 1909)

The number of occurrences of a word \( u \) in a word \( w \) is

\[
\text{occ}(w, u) = \#\{i : w[i..i + |u| - 1] = u\}
\]

An infinite word \( x \in A^\omega \) (resp. a real number \( \xi \)) is simply normal (in base \( b \)) if for any \( a \in A \),

\[
\lim_{n \to \infty} \frac{\text{occ}(x[1..n], a)}{n} = \frac{1}{b}.
\]

An infinite word \( x \in A^\omega \) (resp. a real number \( \xi \)) is normal (in base \( b \)) if for any \( u \in A^* \),

\[
\lim_{n \to \infty} \frac{\text{occ}(x[1..n], u)}{n} = \frac{1}{b^{|u|}}.
\]

In base \( b = 2 \), this means

- the frequencies in \( x \) of the 2 digits 0 and 1 are 1/2,
- the frequencies in \( x \) of the 4 words 00, 01, 10, 11 are 1/4,
- the frequencies in \( x \) of the 8 words 000, 001, \ldots, 111 are 1/8,
- \ldots
Selection rules

- If $x = a_1a_2a_3 \cdots$ is a normal infinite word, then so is $x' = a_2a_3a_4 \cdots$ made of symbols at all positions but the first one.
- If $x = a_1a_2a_3 \cdots$ is normal infinite word, then so is $x' = a_2a_4a_6 \cdots$ made of symbols at even positions.
- What about selecting symbols at positions $2^n$?
- What about selecting symbols at prime positions?
- What about selecting symbols following a 1?
- What about selecting symbols followed by a 1?
Prefix selection

Let $L \subseteq A^*$ be a set of finite words and $x = a_1a_2a_3\cdots \in A^\omega$. The prefix selection of $x$ by $L$ is the word $x \upharpoonright L = a_{i_1}a_{i_2}a_{i_3}\cdots$ where $\{i_1 < i_2 < i_3 < \cdots\} = \{i : a_1a_2\cdots a_{i-1} \in L\}$.

Example (Symbols following a 1)

If $L = (0 + 1)^*1$, then $i_1 - 1, i_2 - 1, i_3 - 1$ are the positions of 1 in $x$ and $x \upharpoonright L$ is made of the symbols following a 1.

Theorem (Agafonov 1968)

Prefix selection by a rational set of finite words preserves normality.

The selection can be realized by a transducer.

Example (Selection of symbols following a 1)

\[
\begin{array}{c}
0: \varepsilon \\
\downarrow \\
0 \\
\end{array}
\quad
\begin{array}{c}
1: \varepsilon \\
\end{array}
\quad
\begin{array}{c}
1: 1 \\
0: 0 \\
\end{array}
\quad
\begin{array}{c}
0: 1 \\
\end{array}
\]

\begin{array}{c}
0: \varepsilon \\
\downarrow \\
1 \\
\end{array}
\quad
\begin{array}{c}
1: \varepsilon \\
\end{array}
\quad
\begin{array}{c}
1: 1 \\
0: 0 \\
\end{array}
\quad
\begin{array}{c}
0: 1 \\
\end{array}
Open questions

- Find a characterization of submonoids $M \subseteq A^* \times B^*$ having the Fatou property.
- For each submonoid $M \subseteq A^* \times B^*$ having the Fatou property, this naturally raises the second problem whether the Fatou property implies the decidability of the inclusion $R \subseteq M$. 