Automata
and
FO[$\mathbb{R}; +, <, 1]$-definable set of reals

Arthur MILCHIOR

LACL, UPEC, Créteil, France

6 september 2016
Journée Montoise d’Informatique Théorique
The main results

Theorem

*It is decidable in linear time whether $R \subseteq \mathbb{R}^{\geq 0}$, accepted by a minimal automaton in base $b \geq 2$ is $\text{FO}[\mathbb{R}; +, <, 1]$-definable.*

Theorem

*Let $\mathcal{A}$ be a real automaton which accepts a $\text{FO}[\mathbb{R}; +, <, 1]$-definable. There exists an existential $\text{FO}[\mathbb{R}; <, +\mathbb{Q}]$-formula of length $O\left(n^3 \log(n)b\right)$ which defines $\mathcal{A}^\mathbb{R}$."

Outline:

Introduction

Methods for similar problems

Description of automata
Characterization of $\text{FO}[\mathbb{R}; +, <, 1]$-definable sets

**Definition (Simple set)**

A simple set $R \subseteq \mathbb{R}^{\geq 0}$ is a finite union of intervals with rational bounds.

Example:

$$\left[3, \frac{20}{3}\right) \cup \left(\frac{22}{3}, +\infty\right).$$

**Theorem (Weispfenning 99)**

A set is $\text{FO}[\mathbb{R}; +, <, 1]$-definable iff it is simple.

**Theorem (Ferrante, Rackoff, 75)**

$\text{FO}[\mathbb{R}; +, <, 1]$ admits quantifier elimination.
Characterization of $\text{FO}[\mathbb{R}, \mathbb{Z}; +, <, 1]$-definable sets

Definition (Periodically simple set)
A periodically simple set $R \subseteq \mathbb{R}^\geq_0$ is a set of the form $A \cup (m\mathbb{N} + B)$ with $A$ and $B$ simple and $m$ rational.

Example:
$$\{3\} \cup \left(4\mathbb{N} + \left(2, \frac{5}{2}\right)\right)$$

Theorem (Weispfenning 99)
A set is $\text{FO}[\mathbb{R}, \mathbb{Z}; +, <, 1]$-definable iff it is periodically simple.

Theorem (Weispfenning 99)
$\text{FO}[\mathbb{R}, \mathbb{Z}; +, <, 1, \lfloor \rfloor, \text{mod}]$ admits quantifier elimination.
Representation of reals.

<table>
<thead>
<tr>
<th>( \frac{2}{3} )</th>
<th>Base 2-expansion</th>
<th>Base 3-expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bullet (10)^\omega )</td>
<td>( \bullet 2(0)^\omega )</td>
<td></td>
</tr>
<tr>
<td>( 1 \bullet 01(0)^\omega )</td>
<td>( \bullet 1(2)^\omega )</td>
<td></td>
</tr>
<tr>
<td>( 1 \bullet 00(1)^\omega )</td>
<td>( 1 \bullet (02)^\omega )</td>
<td></td>
</tr>
</tbody>
</table>

Figure: \( \{\frac{2}{3}, \frac{5}{4}\} \) in base 2
Representation of reals.

<table>
<thead>
<tr>
<th>Base 2-expansion</th>
<th>Base 3-expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{2}{3} )</td>
<td>( \bullet (10)^\omega )</td>
</tr>
<tr>
<td>( \frac{5}{4} )</td>
<td>( 1 \bullet 01(0)^\omega )</td>
</tr>
<tr>
<td></td>
<td>( 1 \bullet 00(1)^\omega )</td>
</tr>
</tbody>
</table>

Figure: \( \{ \frac{2}{3}, \frac{5}{4} \} \) in base 2
Representation of reals.

<table>
<thead>
<tr>
<th>Base 2-expansion</th>
<th>Base 3-expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bullet(10)^\omega )</td>
<td>( \bullet 2(0)^\omega )</td>
</tr>
<tr>
<td>( \frac{2}{3} )</td>
<td>( \bullet 1(2)^\omega )</td>
</tr>
<tr>
<td>( \frac{5}{4} )</td>
<td>( 1 \bullet (02)^\omega )</td>
</tr>
<tr>
<td>( 1 \bullet 01(0)^\omega )</td>
<td></td>
</tr>
<tr>
<td>( 1 \bullet 00(1)^\omega )</td>
<td></td>
</tr>
</tbody>
</table>

Figure: \( \{ \frac{2}{3}, \frac{5}{4} \} \) in base 2
Representation of reals.

<table>
<thead>
<tr>
<th></th>
<th>Base 2-expansion</th>
<th>Base 3-expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\frac{2}{3})</td>
<td>(\bullet (10)^\omega)</td>
<td>(\bullet 2(0)^\omega)</td>
</tr>
</tbody>
</table>
| \(\frac{5}{4}\) | \(1 \bullet 01(0)^\omega\)  
\(1 \bullet 00(1)^\omega\) | \(\bullet 1(2)^\omega\)  
\(1 \bullet (02)^\omega\) |

Figure: \(\left\{ \frac{2}{3}, \frac{5}{4} \right\}\) in base 2
### Representation of reals.

<table>
<thead>
<tr>
<th></th>
<th>Base 2-expansion</th>
<th>Base 3-expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{2}{3} )</td>
<td>( \bullet (10)^\omega )</td>
<td>( \bullet 2(0)^\omega )</td>
</tr>
<tr>
<td>( \frac{5}{4} )</td>
<td>( 1 \bullet 01(0)^\omega ) ( 1 \bullet 00(1)^\omega )</td>
<td>( \bullet 1(2)^\omega ) ( 1 \bullet (02)^\omega )</td>
</tr>
</tbody>
</table>

**Figure:** \( \left\{ \frac{2}{3}, \frac{5}{4} \right\} \) in base 2
Expressivity of base $b$-automata

Theorem (Boigelot, Rassart, Wolper, 98)

A set $R \subseteq (\mathbb{R}^{\geq 0})^d$ is accepted by an automaton in base $b$ iff it is $\text{FO}[\mathbb{R}, \mathbb{Z}; +, <, X_b, 1]$-definable.

$X_b(x, u, k)$ holds iff $u = 2^z$ for $z \in \mathbb{Z}$ and the $z$-th bit of $x$ is $k$. 
Example: Digit with infinitely many 0 and 1

$\forall r > 0. \exists u < n.X_2(x, u, 1) \land \exists u < n.X_2(x, u, 0)$
Sets accepted in multiple basis.

Theorem (Boigelot, Brusten, Leroux, 09)

A set $R \subseteq (\mathbb{R}_{\geq 0})^d$ is accepted by automata in all basis iff it is $\text{FO}[\mathbb{R}, \mathbb{Z}; +, <, 1]$-definable.

Figure: $\left\{ \frac{2}{3}, \frac{5}{4} \right\}$ in base 3
Weak Büchi Automata

**Definition**
An automaton is weak if all or none of the states of a strongly connected components are accepting.

**Theorem (Löding - 2001)**
*Weak Büchi automata admits minimization in time $O(n \log(n))$.***
Properties of $\text{FO}[\mathbb{R}, \mathbb{Z}; +, <, 1]$ 

Theorem (Boigelot, Brusten, Bruyère, 08)

$\text{FO}[\mathbb{R}, \mathbb{Z}; +, <, 1]$-definable sets are accepted by weak-Büchi automata.

Example

A weak Büchi automaton accepting a non-$\text{FO}[\mathbb{R}, \mathbb{Z}; +, <, 1]$-definable set.

![Diagram](image)

**Figure:** Reals without 1 in its base-3 representation
Similar problem for FO[+] and integers

Theorem (Recall: main result)

It is decidable in linear time whether \( R \subseteq \mathbb{R}^{\geq 0} \), accepted by a minimal automaton in base \( b \geq 2 \) is FO[\( \mathbb{R}; +, <, 1 \)]-definable.

Theorem

Whether \( R \subseteq \mathbb{N}^d \) accepted by a minimal automaton is FO[\( \mathbb{N}; + \)]-definable is decidable.

<table>
<thead>
<tr>
<th>dimension ( d )</th>
<th>time complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(Honkala 86)</td>
</tr>
<tr>
<td>any</td>
<td>3EXP</td>
</tr>
<tr>
<td>any</td>
<td>polynomial</td>
</tr>
<tr>
<td>1</td>
<td>quasi-linear</td>
</tr>
</tbody>
</table>
Honkala’s algorithm

Theorem (Honkala 86)
Whether $R \subseteq \mathbb{N}$ accepted by an automaton is $\text{FO}[\mathbb{N}; +]$-definable is decidable.

Lemma
The threshold $t$ and the period $p$ is at most $b^n$ where $n$ is the number of states of the automaton $A$.

Algorithm

(1) Runs over all sets $R \subseteq \mathbb{N}$ with threshold and period less than $b^n$. 
(2) Generates the minimal automaton $A_R$ in base $b$ accepting $R$. 
(3) Accepts if $A_R = A$. 

Example of application of Honkala’s method

\[
\begin{align*}
q_0 & \rightarrow (0, 1), (1, 0) \\
& \rightarrow (1, 1) \\
& \rightarrow (\cdot, 0) \\
& \rightarrow (\cdot, \cdot) \\
q_1 & \rightarrow (1, 1) \\
& \rightarrow (\cdot, 0) \\
& \rightarrow (\cdot, \cdot) \\
q_2 & \rightarrow (0, 0) \\
\end{align*}
\]

If this automaton accepts an ultimately periodic set, the threshold and period are at most \(2^3 = 8\).

Proposition

This method allows to recognize any class \(C\) of languages such that:

1. there exists a function \(s\) from language to integers,
2. for all language \(L \in C\), the minimal automaton accepting \(L\) has at least \(s(L)\) states,
3. for each \(c \in \mathbb{N}\), the set \(\{L \mid s(L) = c\}\) is finite and computable.

Application

For \(C\) the periodically simple sets, take \(s(R)\) as the length of the greatest rational.
Muchnik’s algorithm

Theorem (Muchnik 91)

There exists $\phi_d \in FO[\mathbb{N}; +, R]$ which states "$R \subseteq \mathbb{N}^d$ is $FO[\mathbb{N}; +]$ – definable”.

Corollary (Muchnik 91)

Whether $R \subseteq \mathbb{N}^d$ accepted by an automaton $\mathcal{A}$ is $FO[\mathbb{N}; +]$-definable is decidable in $4\text{EXP}$-time.

Algorithm

(1) Transform $\phi_d$ into an automaton $\mathcal{A}'$ where $R$ is encoded by $\mathcal{A}$.
(2) Accepts if $\mathcal{A}'$ accepts a non-empty language.
Example of application of Muchnik’s example.

This method works for any class $C$ of set of tuple of numbers such that there exists a $\phi \in \text{FO}[\mathbb{R}, \mathbb{Z}; +, X_b, R]$ which holds if and only if $R \in C$.

Example

Whether an automaton accepts a subsemigroup of $(\mathbb{R}^d, +)$ is decidable.

$$\forall \ x_1, \ldots, x_d \ y_1, \ldots, y_d \left\{ \begin{array}{l} (x_1, \ldots, x_d) \in R \land \ (y_1, \ldots, y_d) \in R \\ (x_1 + y_1, \ldots, x_d + y_d) \in R \end{array} \right\} \rightarrow (x_1 + y_1, \ldots, x_d + y_d) \in R$$

asserts that $R$ is a subsemigroup of $(\mathbb{R}^d, +)$. 
Marsault-Sakarovitch’s algorithm

Theorem (Marsault-Sakarovitch 13)
Whether \( R \subseteq \mathbb{N} \) accepted by a minimal automaton \( A \) is
\( \text{FO[}\mathbb{N}; +\text{]} \)-definable is decidable in linear time.

Proposition
Let \( \mathbb{L} \) be a class of language and \( \mathbb{A} \) a class of automata such that:

1. each language of \( \mathbb{L} \) is accepted by an automaton of \( \mathbb{A} \),
2. all automata of \( \mathbb{A} \) accepts a language belonging to \( \mathbb{L} \),
3. \( \mathbb{A} \) is closed under quotient and
4. it is decidable in time \( t(n) \) whether an automaton belongs to \( \mathbb{A} \).

It is decidable in time \( t(n) \) whether a minimal automaton accepts a
language of \( \mathbb{L} \).
Form of an automaton accepting a $\text{FO}^R[+; +, <, 1]$-definable subset of $[0, 1]$
Example \( \left[ \frac{1}{4}, \frac{1}{3} \right) \cup \left\{ \frac{11}{24}, \frac{2}{3} \right\} \)

Lemma

*Cyclic states accepts a set which is a finite union of elements of \( \left\{ \left[ 0, r \right), \{ r \}, (r, 1] \right\} \) for some rational \( r \).*
Describing $\mathbb{A}$.

Using Marsault-Sakarovitch’s method, $\mathbb{A}$ is the set of automata satisfying:

1. At most two leaves, accepting $\emptyset$ and $[0, 1]$,
2. Let $C$ be a strongly connected component which is not a leaf. Let $q \in C$, $a \in [b - 1]$. If $\delta(q, a)$ does not belong to $C$ then it belongs to a leaf.
3. $C$ is a cycle.
4. ...
Using Marsault-Sakarovitch’s method, $\mathcal{A}$ is the set of automata satisfying:

1. At most two leaves, accepting $\emptyset$ and $[0, 1]$,
2. Let $C$ be a strongly connected component which is not a leaf. Let $q \in C$, $a \in [b - 1]$. If $\delta(q, a)$ does not belong to $C$ then it belongs to a leaf.
3. For each $q$ in $C$, there exists a single digit $d_q \in [b - 1]$ such that $\delta(q, d_q) \in C$.
4. For each $q, q' \in C$, for each $a < d_q$, $a' < d_{q'}$, $\delta(q, a) = \delta(q', a')$. 
Future research

Solving the similar problem for:

\[ R \subseteq [0, 1]^d, \]

\[ R \subseteq \mathbb{N}^d \] and the logic FO[<, mod].