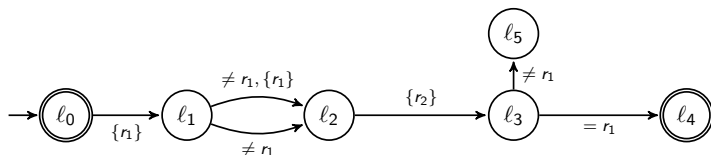


The Containment Problem for Unambiguous Register Automata

Antoine Mottet, Karin Quaas

QuantLA Workshop 2018

- Extension of finite automata to infinite alphabets ($\Sigma \times \mathbb{N}$)

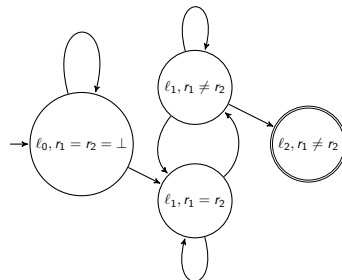
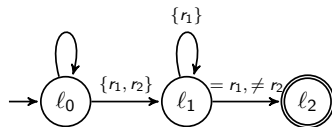


$$\begin{array}{ccccccc}
 (\ell_0, \perp, \perp) & \xrightarrow{\begin{pmatrix} a \\ 0 \end{pmatrix}} & (\ell_0, 0, \perp) & \xrightarrow{\begin{pmatrix} a \\ 1 \end{pmatrix}} & (\ell_2, 0, \perp) & \xrightarrow{\begin{pmatrix} a \\ 3 \end{pmatrix}} & (\ell_3, 0, 3) & \xrightarrow{\begin{pmatrix} a \\ 0 \end{pmatrix}} & (\ell_4, 0, 3) \\
 & & & \longrightarrow & (\ell_2, 1, \perp) & \longrightarrow & (\ell_3, 1, 3) & \longrightarrow & (\ell_5, 1, 3)
 \end{array}$$

- Recognizers of **orbits**:

$$\begin{pmatrix} a \\ 0 \end{pmatrix} \begin{pmatrix} a \\ 1 \end{pmatrix} \begin{pmatrix} a \\ 3 \end{pmatrix} \begin{pmatrix} a \\ 0 \end{pmatrix} \sim \begin{pmatrix} a \\ 4 \end{pmatrix} \begin{pmatrix} a \\ 3 \end{pmatrix} \begin{pmatrix} a \\ 1 \end{pmatrix} \begin{pmatrix} a \\ 4 \end{pmatrix}.$$

- ▶ Projection of $L \subseteq (\Sigma \times \mathbb{N})^*$ onto Σ^* : set of words $w \in \Sigma^*$ such that $(w_1, d_1) \dots (w_n, d_n) \in L$ for some $d_1, \dots, d_n \in \mathbb{N}$.
- ▶ Projection of recognizable L is regular (rec. by **orbit automaton**).
- ▶ Emptiness of L is decidable: $L = \emptyset \Leftrightarrow$ its projection is empty.

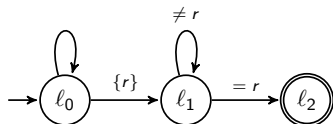


Definition

An automaton is unambiguous if every word has at most **1** accepting run.

- ▶ Deterministic \subseteq Unambiguous \subseteq Non-deterministic,
- ▶ Ambiguity as a resource (STAA?),
- ▶ Collapses and non-collapses depending on model of computation,
- ▶ Succinctness,
- ▶ Important problems related to unambiguity (parity games in $UP \setminus P?$).

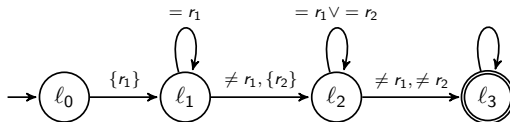
- $L = \{d_1 \dots d_n \in \mathbb{N}^* \mid \exists i \in \{1, \dots, n-1\} : d_i = d_n\}$



- \bar{L} not recognizable (even by nondeterministic RA):
- say \mathcal{A} is a k -register RA recognizing \bar{L} ,
 - $(0)(1) \dots (k)(k+1)(k+2) \in \bar{L}$, so \exists accepting run of \mathcal{A} .
 - Let $d \in \{1, \dots, k+1\}$ be one of the forgotten values.
 - $(0)(1) \dots (k)(k+1)(d) \notin \bar{L}$ has the same accepting run.
- In particular L not recognizable by deterministic RA.

$$\{d_1 \cdots d_n \in \mathbb{N}^* \mid \#\{d_1, \dots, d_n\} \geq 3\}$$

- Recognizable by deterministic RA:



- Needs 2 registers.
- Exists a 1-register unambiguous RA.

- Universality: Given \mathcal{B} , determine if $L(\mathcal{B}) = (\Sigma \times \mathbb{N})^*$.

\mathcal{B}	DRA	URA	NRA
1 register	NL-complete	?	Ackermann-complete
≥ 2 registers	NL-complete	?	Undecidable
*	PSPACE-complete	?	Undecidable

- Containment: Given \mathcal{A}, \mathcal{B} , determine if $L(\mathcal{A}) \subseteq L(\mathcal{B})$.

\mathcal{B}	DRA	URA	NRA
1 register	PSPACE-complete	?	Ackermann-complete
*	PSPACE-complete	?	Undecidable

- $L(\mathcal{A}) \subseteq L(\mathcal{B}) \Leftrightarrow L(\mathcal{A}) \cap \overline{L(\mathcal{B})} = \emptyset$
 \rightsquigarrow “on-the-fly” complementation.

- Configuration C of n -register \mathcal{B} : set of tuples $(\ell^{\mathcal{B}}, d_1, \dots, d_n)$.
- Synchronized configuration of \mathcal{A} and \mathcal{B} : $((\ell^{\mathcal{A}}, d_1, \dots, d_m), C)$.
- (Infinitely branching) transition system $(\mathbb{S}, \rightarrow)$ on synchronized configurations:

$$((\ell^{\mathcal{A}}, d_1, \dots, d_m), C) \rightarrow ((\ell'^{\mathcal{A}}, e_1, \dots, e_m), C')$$

if $(\ell^{\mathcal{A}}, d_1, \dots, d_m) \xrightarrow{\begin{pmatrix} \sigma \\ d \end{pmatrix}} (\ell'^{\mathcal{A}}, e_1, \dots, e_m)$ and $C \xrightarrow{\begin{pmatrix} \sigma \\ d \end{pmatrix}} C'$ for some $(\sigma, d) \in \Sigma \times \mathbb{N}$.

- $((\ell^{\mathcal{A}}, d_1, \dots, d_m), C)$ **bad** if $\ell^{\mathcal{A}}$ accepting and C not accepting.
- $L(\mathcal{A}) \not\subseteq L(\mathcal{B}) \Leftrightarrow \exists \text{bad reachable configuration in } (\mathbb{S}, \rightarrow)$.

The **non-deterministic** case:

- ▶ Infinite branching: only consider “essentially different” successors.
- ▶ Infinite depth: ...
 - ▶ \preceq is a **well-quasi-order** if for every infinite sequence S_0, S_1, \dots , there exist $i < j$ such that $S_i \preceq S_j$.
 - ▶ For 1 register: define a WQO $S \preceq S'$ on synchronized configurations such that if S' reaches a bad configuration in k steps, then S reaches bad in k steps.
 - ▶ For ≥ 2 registers: no such WQO exists (because of undecidability).

The **unambiguous** case: try to bound size of configurations.

Definition

An ***n*-type** is a satisfiable conjunction $\varphi(x_1, \dots, x_n)$ of $=$ and \neq that is maximal (any formula containing φ is equivalent to φ or unsatisfiable).

- ▶ $x_1 = x_2$ and $x_1 \neq x_2$ are the only 2-types,
- ▶ $x_1 = x_2 \wedge x_2 \neq x_3$ and $x_1 \neq x_2 \wedge x_1 = x_3$ are 3-types (there are 5 in total),
- ▶ In general, there are at most $n^n = O(2^{n^2})$ types with n variables (**Bell numbers**).
- ▶ Every tuple $(d_1, \dots, d_n) \in \mathbb{N}^n$ has a type $\text{tp}(d_1, \dots, d_n)$.
- ▶ $\text{tp}(d_1, \dots, d_n) = \text{tp}(e_1, \dots, e_n) \Leftrightarrow \exists \text{ permutation } \alpha \text{ of } \mathbb{N} \text{ s.t. } \alpha(d_i) = e_i$.

Consider

$$C = \{(\ell, 1, 2), (\ell'', 1, 2), (\ell', 3, 4), (\ell', 2, 5), (\ell'', 4, 5), (\ell, 1, 3), (\ell'', 1, 3)\}$$

- ▶ Pick $\varphi(x_1, x_2, x_3, x_4)$ a 4-type.
- ▶ For (d_1, d_2) , compute

$$L_\varphi(d_1, d_2) := \{\ell \mid \exists e_1, e_2 : (\ell, e_1, e_2) \in C \text{ and } \mathbb{N} \models \varphi(d_1, d_2, e_1, e_2)\}.$$

- | | |
|---|--|
| ▶ $\varphi := (x_1 = x_3 \neq x_2 = x_4)$ | ▶ $\psi := \{x_2, x_3\}, \{x_1\}, \{x_4\}$ |
| ▶ $L_\varphi(1, 2) = \{\ell, \ell''\}$ | ▶ $L_\psi(1, 2) = \{\ell'\}$ |
| ▶ $L_\varphi(2, 5) = \{\ell'\}$, | ▶ $L_\psi(2, 5) = \emptyset$, |
| ▶ $L_\varphi(3, 4) = \{\ell'\}$, | ▶ $L_\psi(3, 4) = \{\ell''\}$, |
| ▶ $L_\varphi(1, 3) = \{\ell, \ell''\}$. | ▶ $L_\psi(1, 3) = \{\ell'\}$. |

- ▶ $\bar{d} \equiv_C \bar{e}$ if for every $2n$ -type φ , $L_\varphi(\bar{d}) = L_\varphi(\bar{e})$.
- ▶ $(1, 2) \equiv_C (1, 3)$.
- ▶ Generalize \equiv_C to synchronized configurations.

- ▶ C **coverable** if $\exists C' \supseteq C$ reachable.

Proposition (M-Quaas '18)

C **coverable configuration**. \bar{a}, \bar{b} such that $\bar{a} \equiv_C \bar{b}$.

Let $C' = C \setminus \{(\ell, \bar{b}) \in C\}$.

C reaches a bad configuration in k steps iff

C' reaches a bad configuration in k steps.

- ▶ If C is coverable, C' is coverable.
- ▶ Given C coverable, one can decide in exponential time if the proposition applies.
- ▶ If **collapsing** does not apply to C , number of data in $C \leq 2^{|\mathcal{B}| \times B_{2n+m}} \leq 2^{|\mathcal{B}| \times 2^{(2n+m)^2}}$.
- ▶ \rightsquigarrow number of collapsed configurations $\leq 2^{2^{\text{poly}(|\mathcal{A}|, |\mathcal{B}|)}}$.

Decide $L(\mathcal{A}) \subseteq L(\mathcal{B})$:

- ▶ Start exploring reachable synchronized configurations, starting from $((\ell_{\text{in}}^{\mathcal{A}}, \perp), \{(\ell_{\text{in}}^{\mathcal{B}}, \perp)\})$.
- ▶ When reaching S and S can be collapsed to S' , pretend we reached S' . If S' is bad, reject.
- ▶ When everything has been reached, accept.
- ▶ At most $2^{2^{\text{poly}(|\mathcal{A}|, |\mathcal{B}|)}}$ collapsed configurations.
 \rightsquigarrow 2-EXPSPACE algorithm.

\mathcal{B}	DRA	URA	NRA
1 register *	PSPACE-comp. PSPACE-comp.	EXPSPACE 2-EXPSPACE	Ackermann-comp. Undecidable

- ▶ For register automata:
 - ▶ Lower bounds,
 - ▶ Length of shortest witnesses for $L(\mathcal{A}) \not\subseteq L(\mathcal{B})$?
 - ▶ Minimal number of data in witness for $L(\mathcal{A}) \not\subseteq L(\mathcal{B})$?
 - ▶ Bounded amount of ambiguity?
- ▶ For RAs over **ordered domain**: decidability for ≥ 2 registers?
- ▶ **Timed** automata: decidability for ≥ 2 clocks?