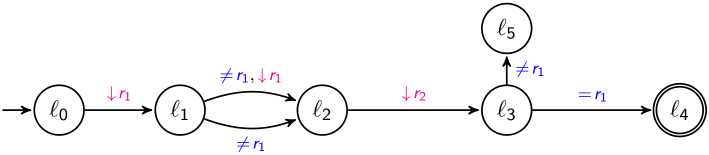
On the Containment Problem for Unambiguous Register Automata

joint work with Antoine Mottet

Karin Quaas
Department für Informatik
Carl von Ossietzky Universität Oldenburg

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

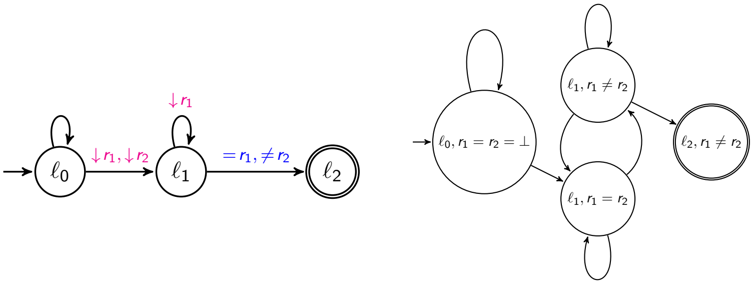


- A run on the input word 0 1 3 0:

$$(\ell_0, \bot, \bot) \xrightarrow{0} (\ell_1, 0, \bot) \xrightarrow{1} (\ell_2, 0, \bot) \xrightarrow{3} (\ell_3, 0, 3) \xrightarrow{0} (\ell_4, 0, 3)$$

- Recognizers of orbits: $0\,1\,3\,0 \sim 4\,3\,1\,4$.

- Emptiness: Given $\mathcal{A}$, determine if $L(\mathcal{A}) = \emptyset$
- Emptiness of RA is decidable (PSPACE-complete): $L(\mathcal{A}) = \emptyset \Leftrightarrow$ language of orbit automaton is empty

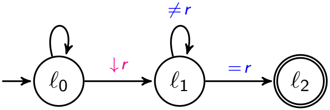


Definition

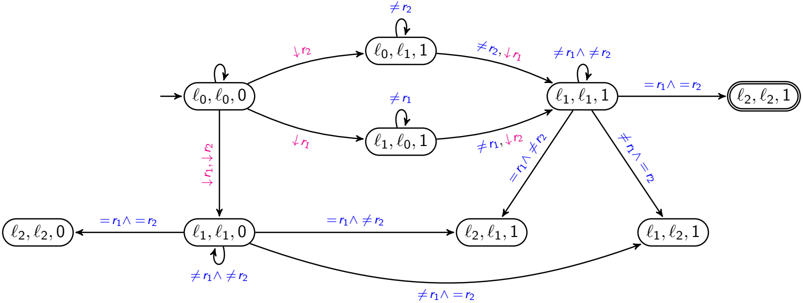
An automaton is unambiguous if every input word has at most one accepting run

- Deterministic $\subseteq$ Unambiguous $\subseteq$ Non-deterministic,
- Collapses and non-collapses depending on model of computation,
- Succinctness,
- Important problems related to unambiguity (open problem: are parity games in $UP \setminus P$).

Given a RA $\mathcal{A}$, decide if $\mathcal{A}$ is unambiguous:

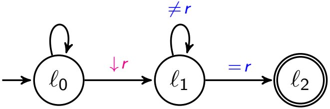


Decidable by a simple extension of the product construction on $\mathcal{A}$



$\mathcal{A}$ is unambiguous iff the language of the product automaton is empty

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

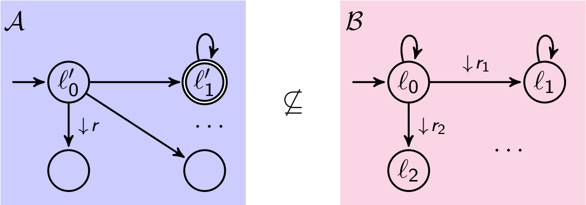


- $\overline{L} = \{d_1 \dots d_n \in \mathbb{N}^* \mid \forall i \in \{1, \dots, n-1\} : d_i \neq d_n\}$
- $\overline{L}$ not recognizable (even by nondeterministic RA):
- In particular L not recognizable by deterministic RA.

- 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
>1 register	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



$$\langle(\ell'_0,0),\{(\ell_0,0,1)\}\rangle$$
$$\Downarrow 2$$
$$\langle(\ell'_1,0),\{(\ell_0,0,1),(\ell_1,2,1),(\ell_2,0,2)\}\rangle \quad \text{witness for } \not\subseteq$$
$$\Downarrow 3$$
$$\langle(\ell'_1,0),\{(\ell_0,0,1),(\ell_1,3,1),(\ell_1,2,1),(\ell_2,0,3)\}\rangle$$
$$\Downarrow 7$$
$$\langle(\ell'_1,0),\{(\ell_0,0,1),(\ell_1,7,1),(\ell_1,3,1),(\ell_1,2,1),(\ell_2,0,7)\}\rangle$$
$$\Downarrow 4$$
$$\ldots$$

- $\mathcal{B}$-set C of n-register $\mathcal{B}$: set of states $(\ell, d_1, \ldots, d_n)$.
- Synchronized state of $\mathcal{A}$ and $\mathcal{B}$: $\langle(\ell_{\mathcal{A}}, d_1, \ldots, d_m), C\rangle$.
- (Infinite) transition system $(\mathbb{S}, \Rightarrow)$ on synchronized states

  $$\langle(\ell_{\mathcal{A}}, d_1, \ldots, d_m), C\rangle \Rightarrow \langle(\ell'_{\mathcal{A}}, e_1, \ldots, e_m), C'\rangle$$

  if $(\ell_{\mathcal{A}}, d_1, \ldots, d_m) \xrightarrow{d} (\ell'_{\mathcal{A}}, e_1, \ldots, e_m)$ and $C \xrightarrow{d} C'$ for some $d \in \mathbb{N}$.
- $\langle(\ell_{\mathcal{A}}, d_1, \ldots, d_m), C\rangle$ bad if $\ell_{\mathcal{A}}$ accepting and C not accepting.
- $L(\mathcal{A}) \not\subseteq L(\mathcal{B})$ iff $\exists$ bad reachable synchronized state in $(\mathbb{S}, \Rightarrow)$.

The non-deterministic case:

- $(\mathbb{S}, \rightarrow)$ is infinite transition system
- Infinite branching: only consider "essentially different" successors
- Infinite depth:
  - $\preceq$ is a well-quasi-order if for every infinite sequence $S_0, S_1, \ldots$, there exist $i < j$ such that $S_i \preceq S_j$
  - For 1 register: define a WQO $S \preceq S'$ on synchronized states such that if S' reaches a bad synchronized state in k steps, then S reaches bad synchronizd state in k steps
  - For ≥ 2 registers: no such WQO exists (because of undecidability)

The unambiguous case: try to bound size of $\mathcal{B}$-sets

- Let

$$\langle(\ell_1', 0), \{(\ell_0, 0, 1), (\ell_1, 7, 1), (\ell_1, 3, 1), (\ell_1, 2, 1), (\ell_2, 0, 7)\}\rangle$$

be a synchronized state in $(\mathbb{S}, \Rightarrow)$

- Do we need to keep both $(\ell_1, 3, 1)$ and $(\ell_1, 2, 1)$, or can we collapse into $\langle(\ell_1', 0), \{(\ell_0, 0, 1), (\ell_1, 7, 1), (\ell_1, 3, 1), (\ell_2, 0, 7)\}\rangle$?
- Goal: find some (decidable) criteria for collapsing such that:

Collapsing Proposition

Let $\langle(\ell, \bar{a}), C\rangle$ be reachable, $C' \subsetneq C$ be the collapse of C.
$\langle(\ell, \bar{a}), C\rangle$ reaches a witness for $\not\subseteq$ in k steps
if, and only if,
$\langle(\ell, \bar{a}), C'\rangle$ reaches a witness for $\not\subseteq$ in k steps.

Definition

An *n*-type is a satisfiable conjunction $\varphi(x_1, \ldots, x_n)$ 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 $\{x_1, x_3\}, \{x_2\}$ 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 $(d_1, \ldots, d_n) \in \mathbb{N}^n$ has a type $\mathrm{tp}(d_1, \ldots, d_n)$.

Consider

$$\langle(\ell_1', 0), \{(\ell_0, 0, 1), (\ell_1, 7, 1), (\ell_1, 3, 1), (\ell_1, 2, 1), (\ell_2, 0, 7)\}\rangle$$

- Pick $\varphi(x_0, x_1, x_2, x_3, x_4)$ a 5-type.
- For (d_1, d_2), compute

$$L_\varphi(d_1, d_2) := \{\ell \in \mathcal{L} \mid \exists e_1, e_2 : (\ell, e_1, e_2) \in C \;\&\; \mathbb{N} \models \varphi(c, d_1, d_2, e_1, e_2)\}$$

- $\varphi := \{x_0\}, \{x_1\}, \{x_2, x_4\}, \{x_3\}$
- $L_\varphi(0, 1) = \emptyset$
- $L_\varphi(7, 1) = \{\ell_1\}$
- $L_\varphi(3, 1) = \{\ell_1\}$
- $L_\varphi(2, 1) = \{\ell_1\}$
- $L_\varphi(0, 7) = \emptyset$

- $\psi := \{x_0, x_3\}, \{x_2\}, \{x_1, x_4\}$
- $L_\psi(0, 1) = \emptyset$
- $L_\psi(7, 1) = \{\ell_2\}$
- $L_\psi(3, 1) = \emptyset$,
- $L_\psi(2, 1) = \emptyset$,
- $L_\psi(0, 7) = \emptyset$.

- $(\ell, d_1, d_2) \equiv_C (\ell', d_1', d_2')$ if $\ell = \ell'$ and for every 5-type φ, $L_\varphi(d_1, d_2) = L_\varphi(d_1', d_2')$
- $(\ell_1, 1, 2) \equiv_C (\ell_1, 1, 3)$

Collapsing Proposition (Mottet & Q, 2018)

Let $\langle(\ell, \bar{a}), C\rangle$ be reachable, and let $(\ell, \bar{d}), (\ell, \bar{e}) \in C$ such that $(\ell, \bar{d}) \equiv_C (\ell, \bar{e})$. Define $C' = C \setminus \{(\ell, \bar{e})\}$, called the collapse of C. Then the following holds:

$\langle(\ell, \bar{a}), C\rangle$ reaches a witness for $\not\subseteq$ in k steps

if, and only if,

$\langle(\ell, \bar{a}), C'\rangle$ reaches a witness for $\not\subseteq$ in k steps.

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

- Start exploring reachable synchronized configurations, starting from $\langle(\ell_{\text{in}}^{\mathcal{A}}, \bot), \{(\ell_{\text{in}}^{\mathcal{B}}, \bot)\}\rangle$.
- 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 $2^{2^{2^{poly(|\mathcal{A}|,|\mathcal{B}|)}}}$ collapsed configurations.
  $\rightsquigarrow$ 2-EXPSPACE algorithm.

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

Open problems

- 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?