Universality of unambiguous register automata

Corentin Barloy - University of Lille
Joint work with Lorenzo Clemente - MIMUW (Warsaw)

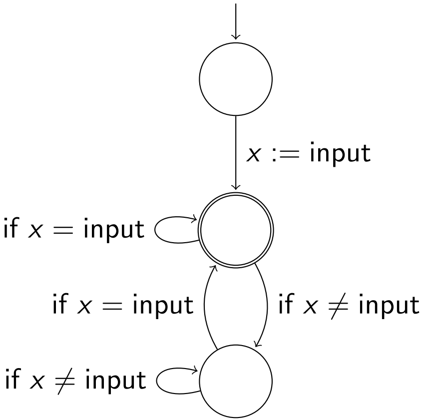
STACS 2021

1. Unambiguous register automata
2. Solving the universality problem

Unambiguous register automata

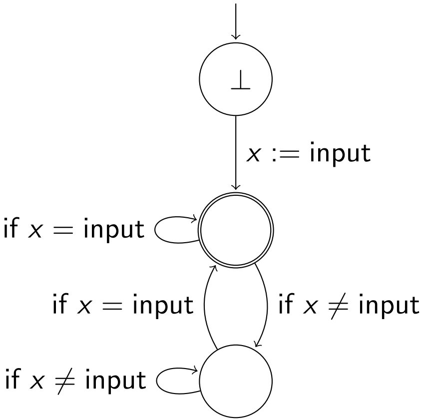
A first (deterministic) register automaton

x is a register that stores values from $\mathbb{N}$



A first (deterministic) register automaton

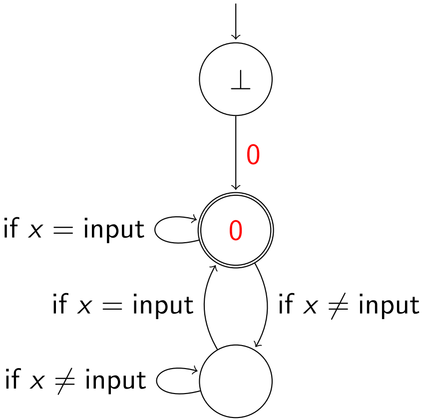
x is a register that stores values from $\mathbb{N}$



$w = 01501$

A first (deterministic) register automaton

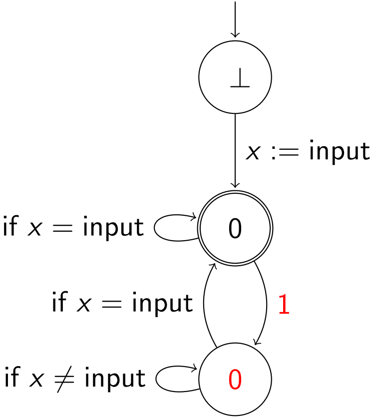
x is a register that stores values from $\mathbb{N}$



$w = 01501$

A first (deterministic) register automaton

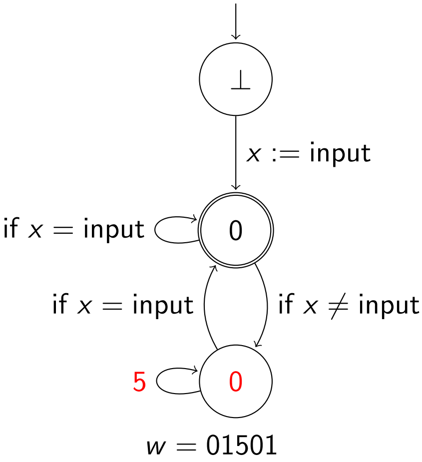
x is a register that stores values from $\mathbb{N}$



$w = 01501$

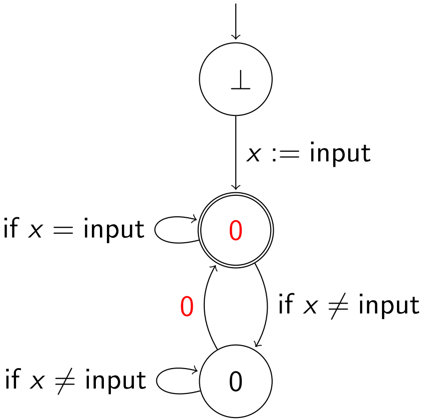
A first (deterministic) register automaton

x is a register that stores values from $\mathbb{N}$



A first (deterministic) register automaton

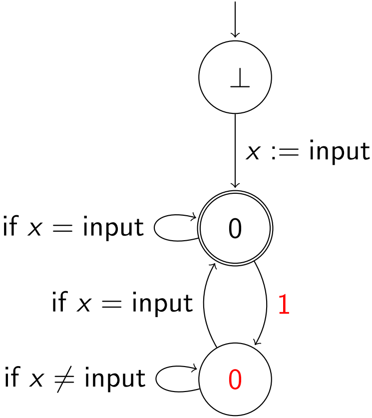
x is a register that stores values from $\mathbb{N}$



$w = 01501$

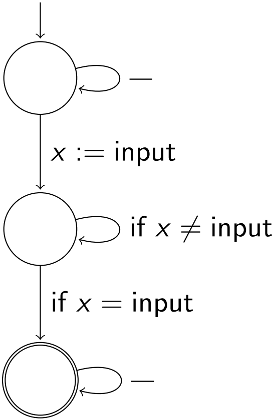
A first (deterministic) register automaton

x is a register that stores values from $\mathbb{N}$



$w = 01501$

A second (non-deterministic) one

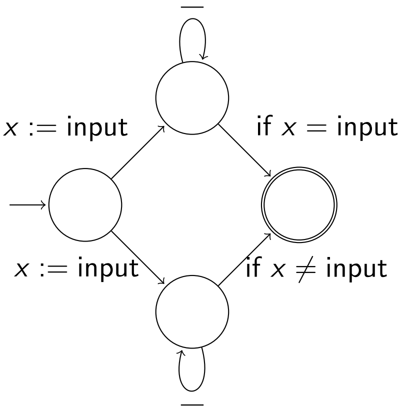


A third (unambiguous) one

Unambiguity: there is a most one accepting run for each word.

A third (unambiguous) one

Unambiguity: there is a most one accepting run for each word.



Decision problems

The universality problem asks whether a given unambiguous register automaton accepts every word.

Decision problems

The universality problem asks whether a given unambiguous register automaton accepts every word.

$$\text{Inclusion} \iff \text{Equivalence} \iff \text{Universality}$$

Decision problems

The universality problem asks whether a given unambiguous register automaton accepts every word.

$$\text{Inclusion} \iff \text{Equivalence} \iff \text{Universality}$$

This is quite surprising: for DCFG, Universality is decidable but not Inclusion!

Main result

Theorem [Mottet, Quaas STACS'19]

The universality problem for unambiguous register automata in decidable in 2EXP-SPACE.

Main result

Theorem [Mottet, Quaas STACS'19]

The universality problem for unambiguous register automata in decidable in 2EXP-SPACE.

Theorem [B, Clemente]

The universality problem for unambiguous register automata in decidable in 2EXP-TIME.

Main result

Theorem [Mottet, Quaas STACS'19]

The universality problem for unambiguous register automata in decidable in 2EXP-SPACE.

Theorem [B, Clemente]

The universality problem for unambiguous register automata in decidable in 2EXP-TIME.

Theorem [Bojańczyk, Klin, Moerman]

The universality problem for unambiguous register automata **with guessing**, and over equality or **order**, is decidable in **EXP-TIME**.

Main result

Theorem [Mottet, Quaas STACS'19]

The universality problem for unambiguous register automata in decidable in 2EXP-SPACE.

Theorem [B, Clemente]

The universality problem for unambiguous register automata in decidable in 2EXP-TIME.

Theorem [Bojańczyk, Klin, Moerman]

The universality problem for unambiguous register automata **with guessing**, and over equality or **order**, is decidable in **EXP-TIME**.

- Undecidable in general.

Main result

Theorem [Mottet, Quaas STACS'19]

The universality problem for unambiguous register automata in decidable in 2EXP-SPACE.

Theorem [B, Clemente]

The universality problem for unambiguous register automata in decidable in 2EXP-TIME.

Theorem [Bojańczyk, Klin, Moerman]

The universality problem for unambiguous register automata **with guessing**, and over equality or **order**, is decidable in **EXP-TIME**.

- Undecidable in general.
- Extend Schutzenberger counting approach.

Main result

Theorem [Mottet, Quaas STACS'19]

The universality problem for unambiguous register automata in decidable in 2EXP-SPACE.

Theorem [B, Clemente]

The universality problem for unambiguous register automata in decidable in 2EXP-TIME.

Theorem [Bojańczyk, Klin, Moerman]

The universality problem for unambiguous register automata **with guessing**, and over equality or **order**, is decidable in **EXP-TIME**.

- Undecidable in general.
- Extend Schutzenberger counting approach.
- There is hope for generalisations to other data domains.

Solving the universality problem

Orbits

For k fixed, $\mathbb{N}^k$ is infinite.

Orbits

For k fixed, $\mathbb{N}^k$ is infinite.
Let φ be a bijection of $\mathbb{N}$. The orbit of a word u is the set:

$$o(u) = \{v \mid \exists \varphi \text{ such that } \varphi(u) = v\} \ .$$

Orbits

For k fixed, $\mathbb{N}^k$ is infinite.
Let φ be a bijection of $\mathbb{N}$. The orbit of a word u is the set:

$$o(u) = \{v \mid \exists \varphi \text{ such that } \varphi(u) = v\} \ .$$

$\mathbb{N}^k$ has B_k many orbits; where B_k is the k^{th} Bell number.

Ryll-Nardzewski function

$s \in$ control locations $\times$ orbit of register values

Ryll-Nardzewski function

$s \in$ control locations $\times$ orbit of register values

$F_s(\quad) = |\{o(r) \mid r \text{ is a run ending in } s \qquad\qquad\qquad\qquad \}|$.

Ryll-Nardzewski function

$s \in$ control locations $\times$ orbit of register values

$F_s(n, \ ) = |\{o(r) \mid r \text{ is a run ending in } s, \text{ of length } n \qquad\qquad \}| \ .$

Ryll-Nardzewski function

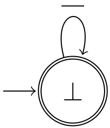
$s \in$ control locations $\times$ orbit of register values

$$F_s(n,k) = |\{o(r) \mid r \text{ is a run ending in } s, \text{ of length } n \text{ and over } k \text{ distinct letters}\}| \ .$$

Ryll-Nardzewski function

$s \in$ control locations $\times$ orbit of register values

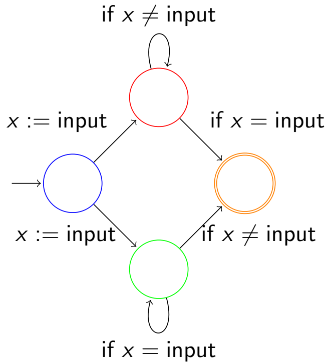
$F_s(n,k) = |\{o(r) \mid r \text{ is a run ending in } s, \text{ of length } n \text{ and over } k \text{ distinct letters}\}|$.



$$F(n+1,k+1) = F(n,k) + (k+1)\cdot F(n,k+1)$$

(Stirling numbers of the second kind)

Counting orbits



n = length of the word, k = number of distinct data values

$$
\begin{cases}
f_{\color{blue}\bullet}(n+1,k+1) = & 0 \\
f_{\color{red}\bullet}(n+1,k+1) = & f_{\color{blue}\bullet}(n,k) + (k+1)\cdot f_{\color{blue}\bullet}(n,k+1) + f_{\color{red}\bullet}(n,k) + k\cdot f_{\color{red}\bullet}(n,k+1) \\
f_{\color{green}\bullet}(n+1,k+1) = & f_{\color{blue}\bullet}(n,k) + (k+1)\cdot f_{\color{blue}\bullet}(n,k+1) + f_{\color{green}\bullet}(n,k+1) \\
f_{\color{orange}\bullet}(n+1,k+1) = & f_{\color{red}\bullet}(n,k+1) + f_{\color{green}\bullet}(n,k) + k\cdot f_{\color{green}\bullet}(n,k+1)
\end{cases}
$$

Zeroness problem

Universality of register automata reduces to the zeroness problem of linrec sequences:
Given a linrec system, is it the case that the function defined is zero everywhere?

Zeroness problem

Universality of register automata reduces to the zeroness problem of linrec sequences:
Given a linrec system, is it the case that the function defined is zero everywhere?

- Compute the number $S(n, k)$ of orbits of all words of length n and width k.

$$S(n+1, k+1) = (k+1) \cdot S(n, k+1) + S(n, k)$$

Zeroness problem

Universality of register automata reduces to the zeroness problem of linrec sequences:
Given a linrec system, is it the case that the function defined is zero everywhere?

- Compute the number $S(n,k)$ of orbits of all words of length n and width k.

$$S(n+1,k+1) = (k+1)\cdot S(n,k+1) + S(n,k)$$

- $\mathcal{A}$ universal $\Leftrightarrow$ $\forall n,k, f_{\bullet}(n,k) = S(n,k) \Leftrightarrow \forall n,k, f_{\bullet}(n,k) - S(n,k) = 0$

Zeroness problem

Universality of register automata reduces to the zeroness problem of linrec sequences:
Given a linrec system, is it the case that the function defined is zero everywhere?

- Compute the number $S(n,k)$ of orbits of all words of length n and width k.

$$S(n+1,k+1) = (k+1) \cdot S(n,k+1) + S(n,k)$$

- $\mathcal{A}$ universal $\Leftrightarrow$ $\forall n, k, f_{\bullet}(n,k) = S(n,k) \Leftrightarrow \forall n, k, f_{\bullet}(n,k) - S(n,k) = 0$
- Linrec is closed by substraction.

Modelling linrec

$$f(n+1,k+1) - f(n+1,k) - f(n,k) - (k+1)\cdot f(n,k+1) = 0$$
$$\downarrow$$
$$\partial_1\partial_2 \cdot f - \partial_1 \cdot f - f - (k+1)\partial_2 \cdot f = 0$$
$$\downarrow$$
$$(\partial_1\partial_2 - \partial_1 - (k+1)\partial_2 - 1)\cdot f = 0$$

Modelling linrec

$$f(n+1,k+1) - f(n+1,k) - f(n,k) - (k+1)\cdot f(n,k+1) = 0$$
$$\downarrow$$
$$\partial_1\partial_2 \cdot f - \partial_1 \cdot f - f - (k+1)\partial_2 \cdot f = 0$$
$$\downarrow$$
$$\underbrace{(\partial_1\partial_2 - \partial_1 - (k+1)\partial_2 - 1)}_{\text{An example of Ore polynomial}} \cdot f = 0$$

Modelling linrec

$$f(n+1,k+1) - f(n+1,k) - f(n,k) - (k+1)\cdot f(n,k+1) = 0$$
$$\downarrow$$
$$\partial_1\partial_2 \cdot f - \partial_1 \cdot f - f - (k+1)\partial_2 \cdot f = 0$$
$$\downarrow$$
$$\underbrace{(\partial_1\partial_2 - \partial_1 - (k+1)\partial_2 - 1)}_{\text{An example of Ore polynomial}} \cdot f = 0$$

Note that those operators do not commute:

$$\partial_1 p(n,k) = p(n+1,k)\partial_1 \neq p(n,k)\partial_1 \quad .$$

Modelling linrec

A system of equations with variables f_i, S and g, with coefficients that are operators.

$$
\left\{
\begin{array}{rrrrr}
(\partial_1\partial_2)\cdot f_{\color{blue}\bullet} & & & & = 0 \\
-\partial_2\cdot f_{\color{blue}\bullet} & +(\partial_1\partial_2-(k+1)\partial_2-1)\cdot f_{\color{red}\bullet} & & & = 0 \\
-\partial_2\cdot f_{\color{blue}\bullet} & & + \quad (\partial_1\partial_2-\partial_2)\cdot f_{\color{green}\bullet} & & = 0 \\
 & -\partial_2\cdot f_{\color{red}\bullet} & + \quad (-(k-1)\partial_2-1)\cdot f_{\color{green}\bullet} & +(\partial_1\partial_2)\cdot f_{\color{orange}\bullet} & = 0 \\
 & (\partial_1\partial_2-(k+1)\partial_2-1)\cdot S & & & = 0 \\
 & g-S+f_{\color{orange}\bullet} & & & = 0
\end{array}
\right.
$$

A system of equations with variables f_i, S and g, with coefficients that are operators.

$$
\begin{cases}
(\partial_1\partial_2)\cdot f_{\color{blue}\bullet} & & & & = 0 \\
-\partial_2 \cdot f_{\color{blue}\bullet} & +(\partial_1\partial_2 - (k+1)\partial_2 - 1)\cdot f_{\color{red}\bullet} & & & = 0 \\
-\partial_2 \cdot f_{\color{blue}\bullet} & & + \quad (\partial_1\partial_2 - \partial_2)\cdot f_{\color{green}\bullet} & & = 0 \\
& -\partial_2 \cdot f_{\color{red}\bullet} & + \quad (-(k-1)\partial_2 - 1)\cdot f_{\color{green}\bullet} & +(\partial_1\partial_2)\cdot f_{\color{orange}\bullet} & = 0 \\
& (\partial_1\partial_2 - (k+1)\partial_2 - 1)\cdot S & & & = 0 \\
& g - S + f_{\color{orange}\bullet} & & & = 0
\end{cases}
$$

Goal: find an operator L such that $L \cdot g = 0$.

Euclidean properties

Let L_1 and L_2 be two operators.

- It is possible to perform an Euclidean division of L_1 by L_2.
- There exists a common left multiple of L_1 and L_2. (Trivial for commutative ring!)

Euclidean properties

Let L_1 and L_2 be two operators.

- It is possible to perform an Euclidean division of L_1 by L_2.
- There exists a common left multiple of L_1 and L_2. (Trivial for commutative ring!)

With $L_1 = \partial_2 - (k+1)$ and $L_2 = \partial_2 - 1$:

$$\begin{aligned}
\mathrm{CLM}(L_1, L_2) &= (k\partial_2 - (k+1)) \cdot L_1 \\
&= (k\partial_2 - (k+1)^2) \cdot L_2 \\
&= k\partial_2^2 - (k^2 + 3k + 1)\partial_2 + (k+1)^2
\end{aligned}$$

The elimination process

$$\left\{\begin{array}{lllll}
(\partial_1\partial_2)\cdot f_{\textcolor{blue}{\bullet}} & & & & = 0 \\
-\partial_2\cdot f_{\textcolor{blue}{\bullet}} & +(\partial_1\partial_2-(k+1)\partial_2-1)\cdot f_{\textcolor{red}{\bullet}} & & & = 0 \\
-\partial_2\cdot f_{\textcolor{blue}{\bullet}} & & +\quad(\partial_1\partial_2-\partial_2)\cdot f_{\textcolor{green}{\bullet}} & & = 0 \\
 & -\partial_2\cdot f_{\textcolor{red}{\bullet}} & +\quad(-(k-1)\partial_2-1)\cdot f_{\textcolor{green}{\bullet}} & +(\partial_1\partial_2)\cdot f_{\textcolor{orange}{\bullet}} & = 0 \\
 & (\partial_1\partial_2-(k+1)\partial_2-1)\cdot S & & & = 0 \\
 & g-S+f_{\textcolor{orange}{\bullet}} & & & = 0
\end{array}\right.$$

$$\Downarrow \text{ multiply lines 2 and 3 by } \partial_1$$

$$\left\{\begin{array}{lllll}
(\partial_1\partial_2)\cdot f_{\textcolor{blue}{\bullet}} & & & & = 0 \\
-\textcolor{red}{\partial_1}\partial_2\cdot f_{\textcolor{blue}{\bullet}} & +(\partial_1^2\partial_2-(k+1)\partial_1\partial_2-\partial_1)\cdot f_{\textcolor{red}{\bullet}} & & & = 0 \\
-\textcolor{red}{\partial_1}\partial_2\cdot f_{\textcolor{blue}{\bullet}} & & +\quad(\partial_1^2\partial_2-\partial_1\partial_2)\cdot f_{\textcolor{green}{\bullet}} & & = 0 \\
 & -\partial_2\cdot f_{\textcolor{red}{\bullet}} & +\quad(-(k-1)\partial_2-1)\cdot f_{\textcolor{green}{\bullet}} & +(\partial_1\partial_2)\cdot f_{\textcolor{orange}{\bullet}} & = 0 \\
 & (\partial_1\partial_2-(k+1)\partial_2-1)\cdot S & & & = 0 \\
 & g-S+f_{\textcolor{orange}{\bullet}} & & & = 0
\end{array}\right.$$

The elimination process

$$\left\{\begin{array}{lclcll}
(\partial_1^2\partial_2 - (k+1)\partial_1\partial_2 - \partial_1)\cdot f_{\color{red}\bullet} & & & & & = 0\\
 & + & (\partial_1^2\partial_2 - \partial_1\partial_2)\cdot f_{\color{green}\bullet} & & & = 0\\
-\partial_2\cdot f_{\color{red}\bullet} & + & (-(k-1)\partial_2 - 1)\cdot f_{\color{green}\bullet} & + & (\partial_1\partial_2)\cdot f_{\color{orange}\bullet} & = 0\\
 & & (\partial_1\partial_2 - (k+1)\partial_2 - 1)\cdot S & & & = 0\\
 & & g - S + f_{\color{orange}\bullet} & & & = 0
\end{array}\right.$$

$\Uparrow$ add first line to second and third lines

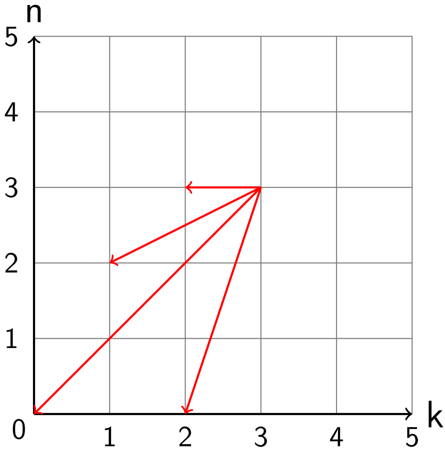
$$\left\{\begin{array}{lclclcll}
(\partial_1\partial_2)\cdot f_{\color{blue}\bullet} & & & & & & & = 0\\
-\partial_1\partial_2\cdot f_{\color{blue}\bullet} & + & (\partial_1^2\partial_2 - (k+1)\partial_1\partial_2 - \partial_1)\cdot f_{\color{red}\bullet} & & & & & = 0\\
-\partial_1\partial_2\cdot f_{\color{blue}\bullet} & & & + & (\partial_1^2\partial_2 - \partial_1\partial_2)\cdot f_{\color{green}\bullet} & & & = 0\\
 & & -\partial_2\cdot f_{\color{red}\bullet} & + & (-(k-1)\partial_2 - 1)\cdot f_{\color{green}\bullet} & + & (\partial_1\partial_2)\cdot f_{\color{orange}\bullet} & = 0\\
 & & (\partial_1\partial_2 - (k+1)\partial_2 - 1)\cdot S & & & & & = 0\\
 & & g - S + f_{\color{orange}\bullet} & & & & & = 0
\end{array}\right.$$

The elimination process

$$((k^2 - 5k + 6)\partial_1^3\partial_2^3 + (2k + 2)\partial_1^3\partial_2^2 - 2\partial_1^2\partial_2 + \partial_2^2 - (3k + 3)) \cdot g = 0$$

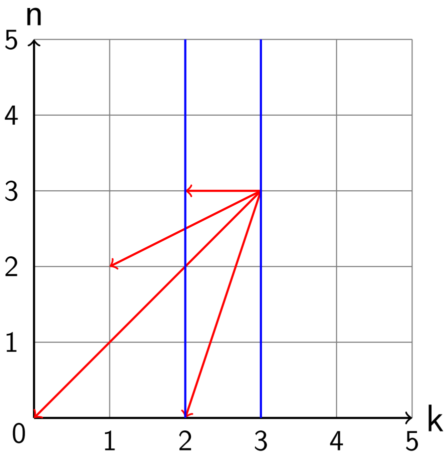
Deciding zeroness of g

$$((k^2 - 5k + 6)\partial_1^3\partial_2^3 + (2k+2)\partial_1^3\partial_2^2 - 2\partial_1^2\partial_2 + \partial_2^2 - (3k+3)) \cdot g = 0$$



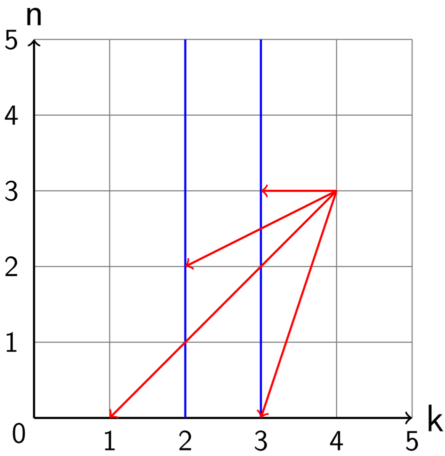
Deciding zeroness of g

$$((k^2 - 5k + 6)\partial_1^3\partial_2^3 + (2k+2)\partial_1^3\partial_2^2 - 2\partial_1^2\partial_2 + \partial_2^2 - (3k+3)) \cdot g = 0$$



Deciding zeroness of g

$$((k^2 - 5k + 6)\partial_1^3\partial_2^3 + (2k+2)\partial_1^3\partial_2^2 - 2\partial_1^2\partial_2 + \partial_2^2 - (3k+3)) \cdot g = 0$$



This yields a 4-EXP-SPACE complexity bound for the universality problem!

$$2^{2^{2^{2^n}}}$$

Complexity

This yields a 4-EXP-SPACE complexity bound for the universality problem!

$$2^{2^{2^{2^{n}}}}$$

- Lagrange bound on zero of polynomials.
- Zeroness of sections (dimension 1).

Hermite forms

Idea: Instead of removing one variable at a time, invert the matrix using (non-commutative) linear algebra.

Hermite forms

Idea: Instead of removing one variable at a time, invert the matrix using (non-commutative) linear algebra.

- Put matrices of operators in a triangular form (Hermite form):

$$\begin{pmatrix} (\partial_1 - 1)\partial_2 & -\partial_2 \\ -k\partial_2 - 1 & \partial_1\partial_2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & (\frac{k}{\partial_1 - 1} - \partial_1)\partial_2 \\ 0 & \partial_2^2 - \frac{1}{\partial_1^2 - \partial_1 - (k+1)}\partial_2 \end{pmatrix}$$

Hermite forms

Idea: Instead of removing one variable at a time, invert the matrix using (non-commutative) linear algebra.

- Put matrices of operators in a triangular form (Hermite form):

$$\begin{pmatrix} (\partial_1 - 1)\partial_2 & -\partial_2 \\ -k\partial_2 - 1 & \partial_1\partial_2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & (\frac{k}{\partial_1 - 1} - \partial_1)\partial_2 \\ 0 & \partial_2^2 - \frac{1}{\partial_1^2 - \partial_1 - (k+1)}\partial_2 \end{pmatrix}$$

- Exponential bound on the coefficient of the Hermite form [Giesbrecht,Kim 11].

Hermite forms

Idea: Instead of removing one variable at a time, invert the matrix using (non-commutative) linear algebra.

- Put matrices of operators in a triangular form (Hermite form):

$$\begin{pmatrix} (\partial_1 - 1)\partial_2 & -\partial_2 \\ -k\partial_2 - 1 & \partial_1\partial_2 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & (\frac{k}{\partial_1 - 1} - \partial_1)\partial_2 \\ 0 & \partial_2^2 - \frac{1}{\partial_1^2 - \partial_1 - (k+1)}\partial_2 \end{pmatrix}$$

- Exponential bound on the coefficient of the Hermite form [Giesbrecht,Kim 11].
- Obtain a bound on the length of a short witness of non-universality.

Main theorems

Theorem [B.,Clemente 20]

The zeroness problem for linrec sequences with univariate polynomial coefficients is decidable in EXP-TIME.

Theorem [B.,Clemente 20]

The universality problem for unambiguous register automata is decidable

- in 2EXP-TIME in general.
- in EXP-TIME (for a fixed number of registers).

Main theorems

Theorem [B.,Clemente 20]

The zeroness problem for linrec sequences with univariate polynomial coefficients is decidable in EXP-TIME.

Theorem [B.,Clemente 20]

The universality problem for unambiguous register automata is decidable

- in 2EXP-TIME in general.
- in EXP-TIME (for a fixed number of registers).

(The complexity becomes EXP-TIME and PTIME if the monicity conjecture holds)

Conclusion

- Improving the complexity. Monicity conjecture: monic cancelling relations suffice:

$$(\cancel{(k^2-5k+6)}\partial_1^3\partial_2^3 + (2k+2)\partial_1^3\partial_2^2 - 2\partial_1^2\partial_2 + \partial_2^2 - (3k+3)) \cdot g = 0$$

This is false for linrec in general, we have to restrict to equations arising from automata. This would give an EXP-TIME bound.

Conclusion

- Improving the complexity. Monicity conjecture: monic cancelling relations suffice:

$$(\cancel{(k^2-5k+6)}\partial_1^3\partial_2^3 + (2k+2)\partial_1^3\partial_2^2 - 2\partial_1^2\partial_2 + \partial_2^2 - (3k+3)) \cdot g = 0$$

  This is false for linrec in general, we have to restrict to equations arising from automata. This would give an EXP-TIME bound.
- Extend to other structures: other atoms, timed automata, pushdown automata...

Conclusion

- Improving the complexity. Monicity conjecture: monic cancelling relations suffice:

$$(\cancel{(k^2-5k+6)}\partial_1^3\partial_2^3 + (2k+2)\partial_1^3\partial_2^2 - 2\partial_1^2\partial_2 + \partial_2^2 - (3k+3)) \cdot g = 0$$

  This is false for linrec in general, we have to restrict to equations arising from automata. This would give an EXP-TIME bound.
- Extend to other structures: other atoms, timed automata, pushdown automata. . .
- Extend to weighted automata.

M. Giesbrecht and M. S. Kim.
Computing the Hermite form of a matrix of Ore polynomials.
Journal of Algebra, 376:341–362, 2013.

A. Mottet and K. Quaas.
The containment problem for unambiguous register automata and unambiguous timed automata.
Theory of Computing Systems, 2020.

O. Ore.
Theory of non-commutative polynomials.
Annals of Mathematics, 34(3):480–508, 1933.

R. Stearns and H. Hunt.
On the equivalence and containment problems for unambiguous regular expressions, grammars, and automata.
In *Proc. of SFCS'81*, pages 74–81, Washington, DC, USA, 1981. IEEE Computer Society.