On the Unsolvability of Loop Analysis

Anton Varonka

Joint work with Laura Kovács

TU Wien Informatics

AVM 2022

The Collatz Problem

$x := \text{input}()$
while $x \neq 1$ **do**
  $x \bmod 2 = 0 \rightarrow x := \frac{1}{2}x$
  $x \bmod 2 = 1 \rightarrow x := 3x + 1$

any input from the domain $\mathbb{D}$: does program terminate on all of them?

The Collatz Problem

$x := \text{input}()$
while $x \neq 1$ **do**
  $x \bmod 2 = 0 \rightarrow x := \frac{1}{2}x$
  $x \bmod 2 = 1 \rightarrow x := 3x + 1$

any input from the domain $\mathbb{D}$: does program terminate on all of them?

This talk is not about the Collatz Conjecture.

Collatz as Decision Problem

$x := c$
while $x \neq 1$ **do**
$\quad x \bmod N = 0 \rightarrow x := a_0 x + b_0$
$\quad \vdots$
$\quad x \bmod N = N-1 \rightarrow x := a_{N-1} x + b_{N-1}$

Collatz as Decision Problem

$$
\begin{array}{l}
x := c \\
\textbf{while } x \neq 1 \textbf{ do} \\
\quad x \bmod N = 0 \rightarrow x := a_0 x + b_0 \\
\quad \vdots \\
\quad x \bmod N = N-1 \rightarrow x := a_{N-1}x + b_{N-1}
\end{array}
$$

Termination is in fact

undecidable (Conway + Kurtz & Simon).

Collatz as Decision Problem

$$
\begin{aligned}
&x := c \\
&\textbf{while } x \neq 1 \textbf{ do} \\
&\quad x \bmod N = 0 \to x := a_0 x + b_0 \\
&\quad \vdots \\
&\quad x \bmod N = N-1 \to x := a_{N-1}x + b_{N-1}
\end{aligned}
$$

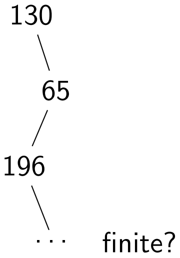
Halting: input c *fixed*
vs.
Termination: *any* input c

Termination is in fact

undecidable (Conway + Kurtz & Simon).

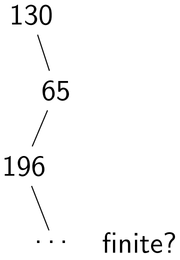
Tweaking the undecidability

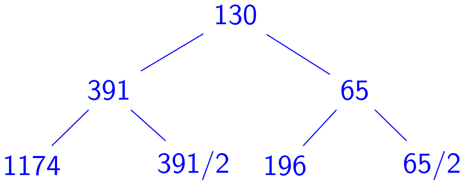
1. Determinism: input given → single path;



Tweaking the undecidability

1. Determinism: input given → single path; allow multiple?





Tweaking the undecidability

1. Determinism: input given → single path; allow multiple?

2. Dimension: one variable or many?

$$(x, y) := (2x - 3, x + 2y)$$

Tweaking the undecidability

1. Determinism: input given → single path; allow multiple?

finite?

2. Dimension: one variable or many?

$$(x, y) := (2x - 3, x + 2y)$$

3. Loop condition: stop if reach the point or interval/. . . ?

$$\text{while } x \geq 0$$

Programs we consider

Non-deterministic.

Sample loop

while **do**

$\boldsymbol{\sigma_1}$:

or

$\boldsymbol{\sigma_2}$:

or

$\boldsymbol{\sigma_3}$:

Programs we consider

Non-deterministic. Arbitrary dimension.

Sample loop (in 3D)

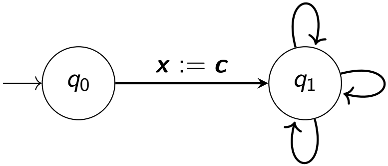
$(x, y, z) := (-1, -1, 2)$
while **do**
$\boldsymbol{\sigma_1}$: $(x, y, z) := (x - y + 1, y - 2z, 2z - x - 1)$
or
$\boldsymbol{\sigma_2}$: $(x, y, z) := (-\frac{3}{2}, x + y + \frac{1}{2}, -x - y + 1)$
or
$\boldsymbol{\sigma_3}$: $(x, y, z) := (2x, y + z, -x)$

Programs we consider

Non-deterministic. Arbitrary dimension. Linear inequality conditions.

Sample loop (in 3D)

$(x, y, z) := (-1, -1, 2)$
while $x + 2y + 3z > 0 \wedge x \leq 10$ **do**
$\boldsymbol{\sigma_1}$: $(x, y, z) := (x - y + 1, y - 2z, 2z - x - 1)$
or
$\boldsymbol{\sigma_2}$: $(x, y, z) := (-\frac{3}{2}, x + y + \frac{1}{2}, -x - y + 1)$
or
$\boldsymbol{\sigma_3}$: $(x, y, z) := (2x, y + z, -x)$



Program correctness:

- Termination on all branches
- Finding good invariants

Termination Result

Theorem 1.
Termination of multi-path affine loops with linear inequality conditions **is undecidable**.

Proof by reduction from the Post's Correspondence Problem (its complement). A loop terminates iff an instance of PCP has no solution.

Remains undecidable with:

- just 4 variables, or
- just 2 linear updates.

Invariants

$(x, y, z) := (-1, -1, 2)$
while true **do**
$(x, y, z) := (x - y + 1, y - z, 2z + y - 1)$
or
$(x, y, z) := (-\frac{3}{2}, x + y + \frac{1}{2}, z + 1)$
or
$(x, y, z) := (2x, y + z, -x)$

Inductive invariant is a relation between variables of a loop $\mathcal{L}$ which is *preserved under any update of* $\mathcal{L}$.

$$f(x, y, z) = 0$$

Invariants

$$
\begin{aligned}
&(x, y, z) := (-1, -1, 2)\\
&\textbf{while } \text{true } \textbf{do}\\
&\quad (x, y, z) := (x - y + 1, y - z, 2z + y - 1)\\
&\quad \text{or}\\
&\quad (x, y, z) := (-\tfrac{3}{2}, x + y + \tfrac{1}{2}, z + 1)\\
&\quad \text{or}\\
&\quad (x, y, z) := (2x, y + z, -x)
\end{aligned}
$$

Inductive invariant is a relation between variables of a loop $\mathcal{L}$ which is *preserved under any update of* $\mathcal{L}$.

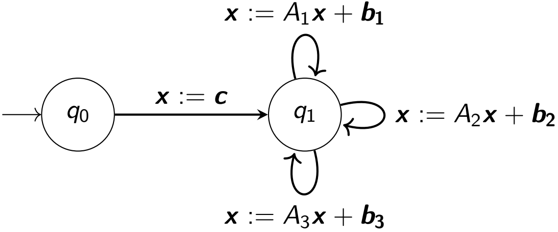
$$x + y + z = 0$$

Strongest Algebraic Invariants

Algebraic invariants are those defined by polynomial equations. There exists the smallest algebraic invariant.

$$x^2 - y^3 = 0 \land y - 2z + 1 = 0$$

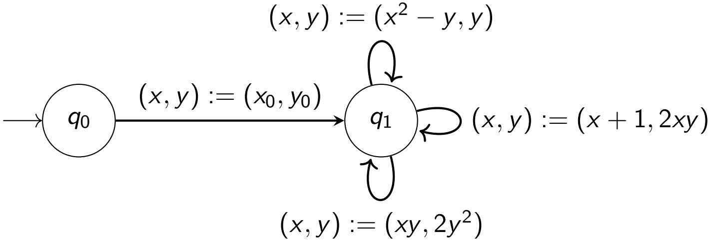
The so-called "Zariski closure" of a set of points in $\mathbb{Q}^d$.
For non-deterministic affine updates: **algorithm** by Hrushovski, Ouaknine, Pouly, Worrell.



Invariant Result

Theorem 2.

Finding the strongest algebraic invariant of a multi-path loop with update degrees ≤ 2 is algorithmically **unsolvable**.



Open questions

1. Halting for multi-path affine loops

Open questions

1. Halting for multi-path affine loops
2. Invariants for deterministic loops with non-affine updates

$$\textit{while}\ \ \text{true}\ \ \textit{do}\ \ (x, y) := (x^2 - y, xy)$$

Open questions

1. Halting for multi-path affine loops
2. Invariants for deterministic loops with non-affine updates

$$\textit{while}\ \ \text{true}\ \ \textit{do}\ \ (x, y) := (x^2 - y, xy)$$

3. Termination of linear-constraint loops

$$\textit{while}\ \ B\boldsymbol{x} \geq \boldsymbol{b}\ \ \textit{do}\ \ A\begin{bmatrix}\boldsymbol{x} & \boldsymbol{x'}\end{bmatrix}^T \leq \boldsymbol{c}$$

Thank you!

undecidable termination

non-determinism without control

+

affine updates and conditions

unsolvable invariants

non-determinism without control

+

quadratic updates