Closed Forms for Numerical Loops

Zachary Kincaid[1] Jason Breck[2] John Cyphert[2] Thomas Reps[2,3]

[1]Princeton University [2]University of Wisconsin-Madison [3]GrammaTech, Inc

January 16, 2019

Loop summarization

The problem: given a loop, compute a formula that represents its behavior.

```
while(i < n):
  i := i + 2
  j := j + 1
```

Loop summarization

The problem: given a loop, compute a formula that represents its behavior.

```
while(i < n):
  i := i + 2
  j := j + 1
```

$$\exists k \in \mathbb{N}. \left(\begin{array}{ll} & i' = i + 2k \\ \wedge & j' = j + k \\ \wedge & n' = n \\ \wedge & i' \geq n \wedge (k \geq 1 \Rightarrow i' \leq n + 1) \end{array} \right)$$

Before exec

After exec

Loop counter

Loop summarization

The problem: given a loop, compute a formula that represents its behavior.

```
while(i < n):
  i := i + 2
  j := j + 1
```

$\Rightarrow$

$$i = j = 0 \land n > 0 \land$$

$$\exists k \in \mathbb{N}. \left(\begin{array}{ll} & i' = i + 2k \\ \land & j' = j + k \\ \land & n' = n \\ \land & i' \geq n \land (k \geq 1 \Rightarrow i' \leq n + 1) \end{array} \right)$$

$$\land \neg(2j' = i')$$

Summary can be used to answer questions about program behavior

- Is $\{i = j = 0 \land n > 0\} loop \{2j = i\}$ valid?

Today: Linear loops

non-deterministic

while (*):

$\mathbf{x} := A\mathbf{x}$

$A \in \mathbb{Q}^{n \times n}$

Today: Linear loops

non-deterministic

```
while ( * ):
```

$$\mathbf{x} := A\mathbf{x}$$

$A \in \mathbb{Q}^{n\times n}$

- In the paper: affine & solvable polynomial loops [Rodríguez-Carbonell & Kapur, ISAAC 2004].

Why linear loops?

- Natural problem

Why linear loops?

- Natural problem
- Practical applications
    - Any loop can be *approximated* by a linear loop [KBCR POPL'18]
    - Summary for the approximation gives invariants for the loop

Approximating general loops [KBCR POPL'18]

```
binary-search(A,target):
   lo = 1, hi = size(A), ticks = 0
   while (lo <= hi):
      ticks++;
      mid = lo + (hi-lo)/2
      if A[mid] == target:
         return mid
      else if A[mid] < target:
         lo = mid+1
      else :
         hi = mid-1
```

Not a linear transformation

Approximating general loops [KBCR POPL'18]

```
binary-search(A,target):
   lo = 1, hi = size(A), ticks = 0
   while (lo <= hi):
      ticks++;
      mid = lo + (hi-lo)/2
      if A[mid] == target:
         return mid
      else if A[mid] < target:
         lo = mid+1
      else :
         hi = mid-1
```

while (*):

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} := \begin{bmatrix} 1 & 0 & 1 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Approximating general loops [KBCR POPL'18]

$$\begin{bmatrix} ticks \\ lo \\ hi \\ mid \\ target \\ A \end{bmatrix} \sim \begin{bmatrix} x \\ y \\ z \end{bmatrix} \iff x = ticks \wedge hi - lo \leq y \wedge z = 1$$

```
binary
  lo
  whi
    ticks++;
    mid = lo + (hi-lo)/2
    if A[mid] == target:
      return mid
    else if A[mid] < target:
      lo = mid+1
    else :
      hi = mid-1
```

while (*):

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} := \begin{bmatrix} 1 & 0 & 1 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Approximating general loops [KBCR POPL'18]

```
binary-search(A,target)
   lo = 1, hi = size(A)
   while (lo <= hi):
      ticks++;
      mid = lo + (hi-lo)
      if A[mid] == target
         return mid
      else if A[mid] < target:
         lo = mid+1
      else :
         hi = mid-1
```

$$\begin{tikzcd} s \arrow[r, dotted, no head, "\sim"] \arrow[d] & \mathbf{v} \arrow[d] \\ s' \arrow[r, dotted, no head, "\sim"'] & \mathbf{v}' \end{tikzcd}$$

$$\begin{bmatrix} z \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 1 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Approximating general loops [KBCR POPL'18]

```
binary-search(A,target):
   lo = 1, hi = size(A), ticks = 0
   while (lo <= hi):
      ticks++;
      mid = lo + (hi-lo)/2
      if A[mid] == target:
         return mid
      else if A[mid] < target:
         lo = mid+1
      else :
         hi = mid-1
```

$$\boxed{\exists k \in \mathbb{N}. \begin{pmatrix} x' = x + kz \\ \wedge y' = (1/2)^k y \\ \wedge z' = z \end{pmatrix}}$$

Approximating general loops [KBCR POPL'18]

```
binary-search(A,target):
  lo = 1, hi = size(A), ticks = 0
```

$$\exists k \in \mathbb{N}. \begin{pmatrix} ticks' = ticks + k \\ \wedge hi' - lo' \leq (1/2)^k (hi - lo) \end{pmatrix}$$

```
    mid = lo + (hi-lo)/2
    if A[mid] == target:
      return mid
    else if A[mid] < target:
      lo = mid+1
    else :
      hi = mid-1
```

$$\exists k \in \mathbb{N}. \begin{pmatrix} x' = x + kz \\ \wedge y' = (1/2)^k y \\ \wedge z' = z \end{pmatrix}$$

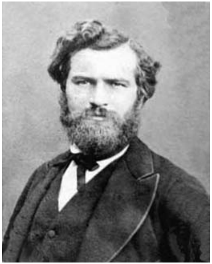
Hasn't this problem already been solved?

Given a square matrix $A \in \mathbb{Q}^{n \times n}$, can compute A^k symbolically

Entries of A^k are exponential polynomials:

$$a_1 \lambda_1^k k^{d_1} + \cdots + a_n \lambda_n^k k^{d_n}$$

Algebraic numbers



Camille Jordan

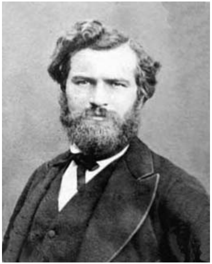
Hasn't this problem already been solved?

Given a square matrix $A \in \mathbb{Q}^{n \times n}$, can compute A^k symbolically

Entries of A^k are exponential polynomials:

$$a_1 \lambda_1^k k^{d_1} + \cdots + a_n \lambda_n^k k^{d_n}$$

Algebraic numbers



Camille Jordan

while $(*)$:
$\quad \mathbf{x} := A\mathbf{x}$

$$\Longrightarrow \quad \exists k \in \mathbb{N}. \mathbf{x}' = A^k \mathbf{x}$$

No.

Skolem's problem (variant):

> *Given an exponential-polynomial f over the algebraic numbers, does there exists some $n \in \mathbb{N}$ such that $f(k) = 0$?*

Decidability of Skolem's problem is unknown!

Thoraf Skolem

No.

Skolem's problem (variant):

> *Given an exponential-polynomial f over the algebraic numbers, does there exists some $n \in \mathbb{N}$ such that $f(k) = 0$?*

Thoraf Skolem

Decidability of Skolem's problem is unknown!

Essential problem: algebraic numbers.

Outline

Starting point of this work: *avoid algebraic numbers*

1. *Periodic rational* matrices have closed forms over $\mathbb{Q}$.
   - Computable in polytime

Outline

Starting point of this work: *avoid algebraic numbers*

1. *Periodic rational* matrices have closed forms over $\mathbb{Q}$.
   - Computable in polytime
2. All matrices have best periodic-rational approximations.

Outline

Starting point of this work: *avoid algebraic numbers*

1. *Periodic rational* matrices have closed forms over $\mathbb{Q}$.
   - Computable in polytime
2. All matrices have best periodic-rational approximations.
3. Exponential-polynomial arithmetic over $\mathbb{Q}$ is decidable.

Closed forms for linear loops

Known:

- Eigenvalues of A are rational $\Rightarrow A^k$ can be expressed in exponential-polynomial arithmetic over $\mathbb{Q}$.

Known:

- Eigenvalues of A are rational $\Rightarrow A^k$ can be expressed in exponential-polynomial arithmetic over $\mathbb{Q}$.
- [Boigelot PhD thesis '99]: A generates a finite monoid $\Rightarrow A^k$ can be expressed in Presburger arithmetic.

Known:

- Eigenvalues of A are rational $\Rightarrow$ A^k can be expressed in exponential-polynomial arithmetic over $\mathbb{Q}$.
- [Boigelot PhD thesis '99]: A generates a finite monoid $\Rightarrow$ A^k can be expressed in Presburger arithmetic.

Common generalization: A matrix A is periodic rational if there is some power p such that A^p has rational eigenvalues.

Known:

- Eigenvalues of A are rational $\Rightarrow$ A^k can be expressed in exponential-polynomial arithmetic over $\mathbb{Q}$.
- [Boigelot PhD thesis '99]: A generates a finite monoid $\Rightarrow$ A^k can be expressed in Presburger arithmetic.

Common generalization: A matrix A is periodic rational if there is some power p such that A^p has rational eigenvalues.

- A periodic rational $\Rightarrow$ can express closed form as

$$\left(\exists k \in \mathbb{N}. \mathbf{x}' = A^k \mathbf{x}\right) \equiv \left(\exists k \in \mathbb{N}. \bigvee_{i=0}^{p-1} k \equiv i \bmod p \land \mathbf{x}' = \underbrace{(A^p)^{\lfloor k/p \rfloor}}_{\text{Rational eigenvalues}} A^i \mathbf{x}\right)$$

- Problem: Rational period of a matrix might be exponential in its size
  - Expressing closed form takes exponential space!

- Problem: Rational period of a matrix might be exponential in its size
  - Expressing closed form takes exponential space!
- Solution: periodic rational spectral decomposition

Periodic rational spectral decomposition (PRSD)

Let $A \in \mathbb{Q}^{n \times n}$ be a square rational matrix. A **periodic rational spectral decomposition** of A is a set of triples

$$\{\langle p_1, \lambda_1, \mathbf{v}_1 \rangle, ..., \langle p_m, \lambda_m, \mathbf{v}_m \rangle\} \subset \mathbb{N} \times \mathbb{Q} \times \mathbb{Q}^n$$

such that

- for each i, $\mathbf{v}_i$ is a generalized eigenvector of A^{p_i}, with eigenvalue λ_i.

Periodic rational spectral decomposition (PRSD)

Let $A \in \mathbb{Q}^{n \times n}$ be a square rational matrix. A **periodic rational spectral decomposition** of A is a set of triples

$$\{\langle p_1, \lambda_1, \mathbf{v}_1 \rangle, ..., \langle p_m, \lambda_m, \mathbf{v}_m \rangle\} \subset \mathbb{N} \times \mathbb{Q} \times \mathbb{Q}^n$$

such that

- for each i, $\mathbf{v}_i$ is a generalized eigenvector of A^{p_i}, with eigenvalue λ_i.
- $\{\mathbf{v}_1, ..., \mathbf{v}_m\}$ is linearly independent

Periodic rational spectral decomposition (PRSD)

Let $A \in \mathbb{Q}^{n\times n}$ be a square rational matrix. A **periodic rational spectral decomposition** of A is a set of triples

$$\{\langle p_1, \lambda_1, \mathbf{v}_1\rangle, ..., \langle p_m, \lambda_m, \mathbf{v}_m\rangle\} \subset \mathbb{N} \times \mathbb{Q} \times \mathbb{Q}^n$$

such that

- for each i, $\mathbf{v}_i$ is a generalized eigenvector of A^{p_i}, with eigenvalue λ_i.
- $\{\mathbf{v}_1, ..., \mathbf{v}_m\}$ is linearly independent
- Informally: $\{\mathbf{v}_1, ..., \mathbf{v}_m\}$ is maximal

Let A be a matrix with PRSD $\{\langle p_1, \lambda_1, \mathbf{v}_1\rangle, ..., \langle p_m, \lambda_m, \mathbf{v}_m\rangle\}$.

- $\left(\mathbf{x}' = A^k\mathbf{x}\right)$ takes exponential space, *but*

Let A be a matrix with PRSD $\{\langle p_1, \lambda_1, \mathbf{v}_1 \rangle, ..., \langle p_m, \lambda_m, \mathbf{v}_m \rangle\}$.

- $\left(\mathbf{x}' = A^k \mathbf{x}\right)$ takes exponential space, *but*
- for any i, $\left(\mathbf{v}_i^T \mathbf{x}' = \mathbf{v}_i^T A^k \mathbf{x}\right)$ can be computed in polytime
  - Intuition: break up period.
    Each $\mathbf{v}_i$ is an easy-to-compute projection

Let A be a matrix with PRSD $\{\langle p_1, \lambda_1, \mathbf{v}_1\rangle, ..., \langle p_m, \lambda_m, \mathbf{v}_m\rangle\}$.

- $\left(\mathbf{x}' = A^k\mathbf{x}\right)$ takes exponential space, *but*
- for any i, $\left(\mathbf{v}_i^T\mathbf{x}' = \mathbf{v}_i^T A^k\mathbf{x}\right)$ can be computed in polytime
  - Intuition: break up period.
    Each $\mathbf{v}_i$ is an easy-to-compute projection

A is periodic rational

$\Longleftrightarrow$

State-space can be recovered from projections

$$\left(\mathbf{x}' = A^k\mathbf{x}\right) \equiv \left(\bigwedge_{i=1}^{m} \mathbf{v}_i^T\mathbf{x}' = \mathbf{v}_i^T A^k\mathbf{x}\right)$$

Approximating linear loops

Let A be a matrix with PRSD $\{\langle p_1, \lambda_1, \mathbf{v}_1\rangle, ..., \langle p_m, \lambda_m, \mathbf{v}_m\rangle\}$.

- Set $V = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_m\end{bmatrix}^T$.

Let A be a matrix with PRSD $\{\langle p_1, \lambda_1, \mathbf{v}_1 \rangle, ..., \langle p_m, \lambda_m, \mathbf{v}_m \rangle\}$.

- Set $V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_m \end{bmatrix}^T$.
- There exists a unique $B \in \mathbb{Q}^{m \times m}$ with $VA = BV$.
    - B is periodic rational

Let A be a matrix with PRSD $\{\langle p_1, \lambda_1, \mathbf{v}_1\rangle, ..., \langle p_m, \lambda_m, \mathbf{v}_m\rangle\}$.

- Set $V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_m \end{bmatrix}^T$.
- There exists a unique $B \in \mathbb{Q}^{m\times m}$ with $VA = BV$.
    - B is periodic rational
    - B simulates A, and V is a simulation:

$$\begin{tikzcd}
\mathbf{a} \arrow[r, dotted, "V"] \arrow[d, "A"'] & \mathbf{b} \arrow[d, "B"] \\
\mathbf{a}' \arrow[r, dotted, "V"'] & \mathbf{b}'
\end{tikzcd}$$

Let A be a matrix with PRSD $\{\langle p_1, \lambda_1, \mathbf{v}_1\rangle, \ldots, \langle p_m, \lambda_m, \mathbf{v}_m\rangle\}$.

- Set $V = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_m\end{bmatrix}^T$.
- There exists a unique $B \in \mathbb{Q}^{m\times m}$ with $VA = BV$.
  - B is periodic rational
  - B simulates A, and V is a simulation:

$$\begin{tikzcd}
\mathbf{a} \arrow[r, dotted, "V"] \arrow[d, "A"'] & \mathbf{b} \arrow[d, "B"] \\
\mathbf{a}' \arrow[r, dotted, "V"'] & \mathbf{b}'
\end{tikzcd}$$

B is the best periodic-rational approximation of A

Invariant generation pipeline

General loop $\xrightarrow{\text{[KBCR POPL'18]}}$ Linear loop

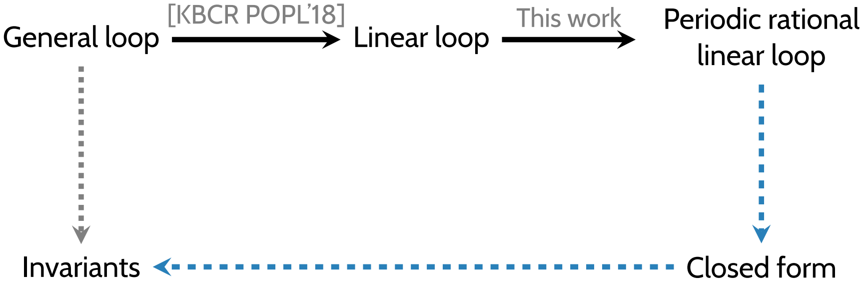
Invariant generation pipeline

General loop $\xrightarrow{\text{[KBCR POPL'18]}}$ Linear loop $\xrightarrow{\text{This work}}$ Periodic rational linear loop

Invariant generation pipeline

$$\begin{tikzcd}
\text{General loop} \arrow[r, "\text{[KBCR POPL'18]}"] & \text{Linear loop} \arrow[r, "\text{This work}"] & \text{Periodic rational linear loop} \arrow[d, dotted] \\
 & & \text{Closed form}
\end{tikzcd}$$

Invariant generation pipeline



Reasoning about non-linear arithmetic

Exponential-polynomial arithmetic is decidable

Two steps:

1. Eliminate all symbols except the loop counter (i.e., program variables)
   - Key idea: terms are linear *over the ring of exponential-polynomials*.
     - $(2^k k^3 - 3^k k^2 + 140 \cdot 3^k)x + (4^k k)y + (2^k)z$
   - Eliminate symbols using linear q.e. [Loos & Weispfennning '93]

Exponential-polynomial arithmetic is decidable

Two steps:

1. Eliminate all symbols except the loop counter (i.e., program variables)
   - Key idea: terms are linear *over the ring of exponential-polynomials*.
     - $(2^k k^3 - 3^k k^2 + 140 \cdot 3^k)x + (4^k k)y + (2^k)z$
   - Eliminate symbols using linear q.e. [Loos & Weispfennning '93]
2. Find a bound for the loop counter
   - Key idea: exponential-polynomials are eventually dominated by the term with largest base (and largest degree)
     - E.g., $2^k k^3 - 3^k k^2 + 140 \cdot 3^k$ is eventully negative

Consequences

Suppose A is periodic rational. The following problems are decidable:

- Is $\{P\}\{\textbf{while}(*) : \mathbf{x} := A\mathbf{x}\}\{Q\}$ valid?

Linear rational arithmetic

Consequences

Suppose A is periodic rational. The following problems are decidable:

- Is $\{P\}\{\textbf{while}(*) : \mathbf{x} := A\mathbf{x}\}\{Q\}$ valid?

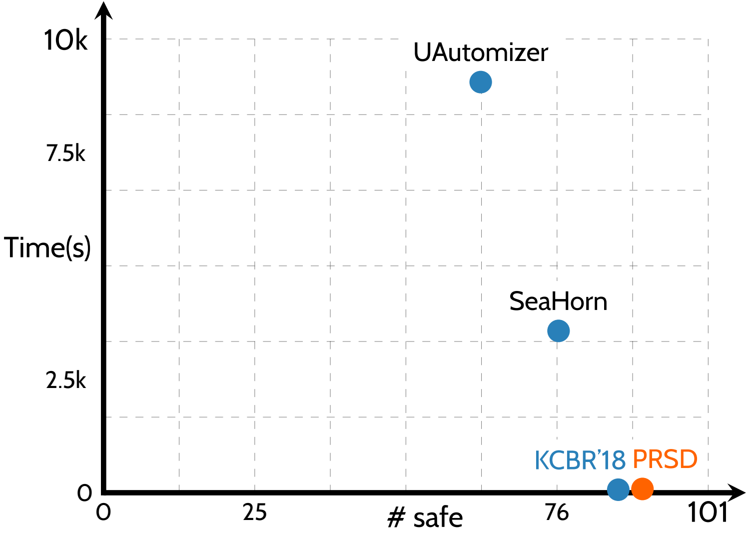
Linear rational arithmetic

- Does $(\mathbf{x} := \mathbf{v}; \textbf{while}(\mathsf{C})\ \textbf{do}\ \mathbf{x} := A\mathbf{x})$ terminate?

Constant vector

Experiments

Suite of 101 microbenchmarks from C4B, HOLA, and literature:



Contributions:

1. Periodic rational linear loops have closed forms over $\mathbb{Q}$.
   - Polytime computation of the summary

Contributions:

1. Periodic rational linear loops have closed forms over $\mathbb{Q}$.
   - Polytime computation of the summary
2. Every matrix has a best periodic-rational approximation.

Contributions:

1. Periodic rational linear loops have closed forms over $\mathbb{Q}$.
    - Polytime computation of the summary
2. Every matrix has a best periodic-rational approximation.
3. Exponential-polynomial arithmetic over $\mathbb{Q}$ is decidable.