Generating Loop Invariants via Polynomial Interpolation

Marc Moreno Maza
Joint work with
Rong Xiao

University of Western Ontario, Canada

East China Normal University
May 24, 2012
Plan

1 Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout’s Theorem

2 Invariant ideal of $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3 Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
Plan

1. Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout’s Theorem

2. Invariant ideal of $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3. Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
Loop model under study

While \( C_0 \) do
  if \( C_1 \) then
    \( X := A_1(X) \);
  elif \( C_2 \) then
    \( X := A_2(X) \);
  \ldots
  elif \( C_m \) then
    \( X := A_m(X) \);
  end if
end while

1. Loop variables: \( X = x_1, \ldots, x_s \), rational value scalar
2. Conditions: each \( C_i \) is a quantifier free formula in \( X \) over \( \mathbb{Q} \).
3. Assignments: \( A_i \in \mathbb{Q}[X] \) inducing a polynomial map \( M_i : \mathbb{R}^s \to \mathbb{R}^s \)
4. Initial condition: \( X \)-values defined by a semi-algebraic system.
Basic notions

- $x, y, a, b$ are loop variables since they are updated in the loop or used to update other loop variables.
- The set of the initial values of the loop is
  \[
  \{(x, y, a, b) \mid x = a, y = b, (a, b) \in \mathbb{R}^2\}.
  \]
- The loop trajectory of the above loop starting at $(x, y, a, b) = (1, 0, 1, 0)$ is the sequence:
  \[
  (1, 0, 1, 0), (1, 1, 1, 0), (2, 2, 1, 0), (34, 3, 1, 0).
  \]
- The reachable set $R(L)$ of a loop $L$ consists of all tuples of all trajectories of $L$.
- If $x_1, \ldots, x_s$ are the loop variables of $L$, then a polynomial $P \in \mathbb{Q}[x_1, \ldots, x_s]$ is a (plain) loop invariant of $L$ whenever $R(L) \subseteq V(P)$ holds.
More notions

- The inductive reachable set $R_{\text{ind}}(L)$ of a loop $L$ is the reachable set of the loop obtained from $L$ by replacing the guard condition with true.
- The absolute reachable set $R_{\text{abs}}(L)$ of a loop $L$ is the reachable set of the loop obtained from $L$ by replacing the guard condition with true, ignoring the branch conditions and, at each iteration executing a branch action selected randomly.
- We clearly have
  $$R(L) \subseteq R_{\text{ind}} \subseteq R_{\text{abs}}$$
- If $x_1, \ldots, x_s$ are the loop variables of $L$, then a polynomial $P \in \mathbb{Q}[x_1, \ldots, x_s]$ is an inductive (resp. absolute) loop invariant of $L$ whenever $R_{\text{ind}}(L) \subseteq V(P)$ (resp. $R_{\text{abs}}(L) \subseteq V(P)$) holds.
- We denote by $\mathcal{I}(L)$ (resp. $\mathcal{I}_{\text{ind}}(L), \mathcal{I}_{\text{abs}}(L)$) the set of the polynomials that are plain (resp. inductive, absolute) loop invariants of $L$.
- These are radical ideals such that
  $$\mathcal{I}_{\text{abs}}(L) \subseteq \mathcal{I}_{\text{ind}}(L) \subseteq \mathcal{I}(L)$$
Absolute invariants might be trivial

Consider \( y_1 x_2 + y_2 + y_3 = x_1 \) \((E)\).

If \( x_1 = 0 \) then the equation \((E)\) holds initially and the loop is not entered.

If \( x_1 \neq 0 \) and \( x_2 = 1 \) then \((E)\) and \( y_2 + 1 = x_2 \) hold before each iteration.

If \( x_1 \neq 0 \) and \( x_2 \neq 1 \) then the second action preserves \((E)\).

Therefore \( y_1 x_2 + y_2 + y_3 - x_1 \in \mathcal{I}(L) \) and \( y_1 x_2 + y_2 + y_3 - x_1 \in \mathcal{I}_{\text{ind}}(L) \) both hold.

```
y_1 := 0;
y_2 := 0;
y_3 := x_1;
while y_3 \neq 0 do
    if y_2 + 1 = x_2
        then
            y_1 := y_1 + 1;
y_2 := 0;
y_3 := y_3 - 1;
        else
            y_2 := y_2 + 1;
y_3 := y_3 - 1;
        end if
    end if
end do
```
Absolute invariants might be trivial

Consider \( y_1x_2 + y_2 + y_3 = x_1 \) \((E)\). 
If \( x_1 = 0 \) then the equation \((E)\) holds initially and the loop is not entered. 
If \( x_1 \neq 0 \) and \( x_2 = 1 \) then \((E)\) and 
\( y_2 + 1 = x_2 \) hold before each iteration. 
If \( x_1 \neq 0 \) and \( x_2 \neq 1 \) then the second action preserves \((E)\). 
Therefore \( y_1x_2 + y_2 + y_3 - x_1 \in I(L) \) and 
\( y_1x_2 + y_2 + y_3 - x_1 \in I_{\text{ind}}(L) \) both hold. 
If conditions are ignored, \((x_1, x_2) = (0, 1)\) and execute the first branch once, then we obtain 
\( y_1x_2 = 1 \) and \( y_2 + y_3 = x_1 \). 
Then \((E)\) is violated and we have 
\( I_{\text{abs}}(L) = \langle 0 \rangle \). 

\[
\begin{align*}
y_1 &:= 0; \\
y_2 &:= 0; \\
y_3 &:= x_1; \\
\text{while } y_3 \neq 0 \text{ do} \\
&\quad \text{if } y_2 + 1 = x_2 \\
&\quad \quad \text{then} \\
&\quad \quad \quad y_1 := y_1 + 1; \\
&\quad \quad \quad y_2 := 0; \\
&\quad \quad \quad y_3 := y_3 - 1; \\
&\quad \quad \text{else} \\
&\quad \quad \quad y_2 := y_2 + 1; \\
&\quad \quad \quad y_3 := y_3 - 1; \\
&\quad \text{end if} \\
&\text{end do}
\end{align*}
\]
Inductive invariants might not be plain invariants

\begin{verbatim}
x := 1;
while x \neq 1 do
   x := x + 1;
end do
\end{verbatim}

- $x - 1 = 0$ is an invariant but not an inductive of the following loop.
- Thus $\mathcal{I}_{\text{ind}}(L)$ is strictly smaller than $\mathcal{I}(L)$
Computing inductive invariants via elimination ideals

- Solving for \((x, y)\) as a 2-variable recurrence
  \[
  x(n + 1) = y(n),
  y(n + 1) =
  x(n) + y(n), \text{ with } x(0) = 0, y(0) = 1.
  \]

- We obtain
  \[
  x(n) = \frac{(\sqrt{5}+1)^n}{\sqrt{5}} - \frac{(-\sqrt{5}+1)^n}{\sqrt{5}},
  y(n) = \frac{\sqrt{5}+1}{2} \left(\frac{\sqrt{5}+1}{2}\right)^n - \frac{\sqrt{5}+1}{2} \left(\frac{-\sqrt{5}+1}{2}\right)^n.
  \]

- Let \(u = \left(\frac{\sqrt{5}+1}{2}\right)^n, v = \left(\frac{-\sqrt{5}+1}{2}\right)^n, a = \sqrt{5}\)

- Taking the dependencies \(u^2v^2 = 1, a^2 = 5\) into account, we want
  \[
  \langle x - \frac{au}{5} + \frac{av}{5}, y - a \frac{a+1}{2} \frac{u}{5} + a \frac{-a+1}{2} \frac{v}{5}, a^2 - 5, u^2v^2 - 1 \rangle \cap \mathbb{Q}[x, y],
  \]

  - which is
    \[
    \langle 1 - y^4 + 2xy^3 + x^2y^2 - 2x^3y - x^4 \rangle.
    \]
A natural criterion

\begin{verbatim}
while $C_0$ do
    if $C_1$ then
        $X := A_1(X)$;
    elif $C_2$ then
        $X := A_2(X)$;
    \ldots
    elif $C_m$ then
        $X := A_m(X)$;
    end if
end while
\end{verbatim}

- Let $f \in \mathbb{Q}[X]$ vanishing at each initial condition.
- Assume that for all $i = 1 \cdots m$ we have
  \[ Z_R(A_i(Z_R(f) \cap Z_R(C_i))) \subseteq Z_R(f) \]
- Then we have
  \[ f \in \mathcal{I}_{\text{ind}}(L). \]
- This can be tested with the commands of
  \[ \text{RegularChains:-SemiAlgebraicSetTools} \]
  based on the
  \[ \text{RegularChains:-RealTriangularize} \]
Summary and notes

- Computing $\mathcal{I}_{\text{ind}}(L)$ is a better approximation of $\mathcal{I}(L)$ than $\mathcal{I}_{\text{abs}}(L)$.
- The loop invariant generation methods of (E. Rodriguez-Carbonell & D. Kapur, ISSAC04) and (L. Kovács, TACAS08) focus on $\mathcal{I}_{\text{abs}}(L)$.
Summary and notes

- Computing $\mathcal{I}_{\text{ind}}(L)$ is a better approximation of $\mathcal{I}(L)$ than $\mathcal{I}_{\text{abs}}(L)$.
- The loop invariant generation methods of (E. Rodriguez-Carbonell & D. Kapur, ISSAC04) and (L. Kovács, TACAS08) focus on $\mathcal{I}_{\text{abs}}(L)$.
- In this talk, we target $\mathcal{I}_{\text{ind}}(L)$ (easier to compute than $\mathcal{I}(L)$) and call it the Invariant Ideal of the loop $L$. Same goal as in the preprint (Bin Wu, Liyong Shen, Min Wu, Zhengfeng Yang & Zhenbing Zeng, 2011).
Computing $\mathcal{I}_{\text{ind}}(L)$ is a better approximation of $\mathcal{I}(L)$ than $\mathcal{I}_{\text{abs}}(L)$.

The loop invariant generation methods of (E. Rodriguez-Carbonell & D. Kapur, ISSAC04) and (L. Kovács, TACAS08) focus on $\mathcal{I}_{\text{abs}}(L)$.

In this talk, we target $\mathcal{I}_{\text{ind}}(L)$ (easier to compute than $\mathcal{I}(L)$) and call it the Invariant Ideal of the loop $L$. Same goal as in the preprint (Bin Wu, Liyong Shen, Min Wu, Zhengfeng Yang & Zhenbing Zeng, 2011).

We also want to avoid computing closed forms of loop variables, while

- not making any assumptions on the shape of the polynomial invariants,
- and avoiding an intensive use of expensive algebraic computations other than linear algebra, for which costs are predictable.
Summary and notes

- Computing $I_{\text{ind}}(L)$ is a better approximation of $I(L)$ than $I_{\text{abs}}(L)$.
- The loop invariant generation methods of (E. Rodriguez-Carbonell & D. Kapur, ISSAC04) and (L. Kovács, TACAS08) focus on $I_{\text{abs}}(L)$.
- In this talk, we target $I_{\text{ind}}(L)$ (easier to compute than $I(L)$) and call it the Invariant Ideal of the loop $L$. Same goal as in the preprint (Bin Wu, Liyong Shen, Min Wu, Zhengfeng Yang & Zhenbing Zeng, 2011).
- We also want to avoid computing closed forms of loop variables, while
  - not making any assumptions on the shape of the polynomial invariants,
  - and avoiding an intensive use of expensive algebraic computations other than linear algebra, for which costs are predictable.
- In (Sankaranarayanan, Sipma & Manna, SIGPLAN 2004) (Y. Chen, B. Xia, L. Yang, & N. Zhan, FMHRTS 2007) (D. Kapur Deduction and Applications 2005) template polynomials are used. Moreover, the latter two use real QE.
Computing $\mathcal{I}_{\text{ind}}(L)$ is a better approximation of $\mathcal{I}(L)$ than $\mathcal{I}_{\text{abs}}(L)$. The loop invariant generation methods of (E. Rodriguez-Carbonell & D. Kapur, ISSAC04) and (L. Kovács, TACAS08) focus on $\mathcal{I}_{\text{abs}}(L)$. In this talk, we target $\mathcal{I}_{\text{ind}}(L)$ (easier to compute than $\mathcal{I}(L)$) and call it the Invariant Ideal of the loop $L$. Same goal as in the preprint (Bin Wu, Liyong Shen, Min Wu, Zhengfeng Yang & Zhenbing Zeng, 2011). We also want to avoid computing closed forms of loop variables, while
- not making any assumptions on the shape of the polynomial invariants, and avoiding an intensive use of expensive algebraic computations other than linear algebra, for which costs are predictable.

In (Sankaranarayanan, Sipma & Manna, SIGPLAN 2004) (Y. Chen, B. Xia, L. Yang, & N. Zhan, FMHRTS 2007) (D. Kapur Deduction and Applications 2005) template polynomials are used. Moreover, the latter two use real QE.

Plan

1 Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout's Theorem

2 Invariant ideal of $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3 Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
Poly-geometrical expression

Notations

Let $\alpha_1, \ldots, \alpha_k$ be $k$ elements of $\overline{\mathbb{Q}}^* \setminus \{1\}$. Let $n$ be a variable taking non-negative integer values. We regard $n, \alpha_1^n, \ldots, \alpha_k^n$ as independent variables and we call $\alpha_1^n, \ldots, \alpha_k^n$ $n$-exponential variables.

Definition

Any $f \in \overline{\mathbb{Q}}[n, \alpha_1^n, \ldots, \alpha_k^n]$ is called a poly-geometrical expression in $n$ over $\overline{\mathbb{Q}}$ w.r.t. $\alpha_1, \ldots, \alpha_k$. For such an $f$, we denote by $f|_{n=i}$ the evaluation of $f$ at $i$. For such $f, g$ we write $f = g$ whenever $f|_{n=i} = g|_{n=i}$ holds for all $i$. 
Canonical form of a poly-geometrical expression

Definition

We say that \( f \in \overline{Q}[n, \alpha_1^n, \ldots, \alpha_k^n] \) is in canonical form if there exist

1. \( c_1, \ldots, c_m \in \overline{Q}^* \), and
2. pairwise different couples \( (\beta_1, e_1), \ldots, (\beta_m, e_m) \) all in \( (\overline{Q}^* \setminus \{1\}) \times \mathbb{Z}_{\geq 0} \), and
3. a polynomial \( c_0(n) \in \overline{Q}[n] \), such that
4. each \( \beta_1, \ldots, \beta_m \) is a product of some of the \( \alpha_1, \ldots, \alpha_k \) and such that
5. \( f(n) \) and \( \sum_{i=1}^{m} c_i \beta_i^n n^{e_i} + c_0(n) \) are equal.

When this holds, the polynomial \( c_0(n) \) is the exponential-free part of \( f(n) \).

Proposition

Let \( f \) a poly-geometrical expression in \( n \) over \( \overline{Q} \) w.r.t. \( \alpha_1, \ldots, \alpha_k \). There exists a unique poly-geometrical expression \( c \) in \( n \) over \( \overline{Q} \) w.r.t. \( \alpha_1, \ldots, \alpha_k \) such that \( c \) is in canonical form and such that \( f \) and \( c \) are equal. We call \( c \) the canonical form of \( f \).
Examples of poly-geometrical expressions

Example

The closed form $f := \frac{(n+1)^2 n^2}{4}$ of $\sum_{i=0}^{n} i^3$ is a poly-geometrical expression in $n$ over $\overline{\mathbb{Q}}$ without $n$-exponential variables.

Example

The expression $g := n^2 2^{(n+1)} - n 2^n 3^{n/2}$ is a poly-geometrical in $n$ over $\overline{\mathbb{Q}}$ w.r.t. $2, \sqrt{3}$.

Example

The sum $\sum_{i=1}^{n-1} i^k$ has $n - 1$ terms while its closed form below

$$\sum_{i=1}^{k} \{ \binom{k}{i} \frac{n^{i+1}}{i+1} ,$$

where $\binom{k}{i}$ the number of ways to partition $k$ into $i$ non-zero summands, has a fixed number of terms and thus is poly-geometrical in $n$ over $\overline{\mathbb{Q}}$. 
**Multiplicative relation ideal**

**Definition**

Let \( A := (\alpha_1, \ldots, \alpha_k) \) be a sequence of \( k \) non-zero elements of \( \overline{\mathbb{Q}} \). Let \( e := (e_1, \ldots, e_k) \) be a sequence of \( k \) integers. We say that \( e \) is a multiplicative relation on \( A \) if \( \prod_{i=1}^{k} \alpha_i^{e_i} = 1 \) holds. Such a relation is said non-trivial if there exists \( i \in \{1, \ldots, n\} \) s.t. \( e_i \neq 0 \) holds. If there exists a non-trivial multiplicative relation on \( A \), we say that \( A \) is multiplicatively dependent; otherwise, we say that \( A \) is multiplicatively independent. All multiplicative relations of \( A \) form the multiplicative relation lattice on \( A \).

**Definition**

Let \( A := (\alpha_1, \ldots, \alpha_k) \) be a sequence of \( k \) elements of \( \overline{\mathbb{Q}} \). Assume w.l.o.g. that for some \( \ell \), with \( 1 \leq \ell \leq k \), we have \( \alpha_1 \neq 0, \ldots, \alpha_\ell \neq 0, \alpha_{\ell+1} = \cdots \alpha_k = 0 \). We associate each \( \alpha_i \) with a “new” variable \( y_i \). The binomial ideal \( \text{MRI}(A; y_1, \ldots, y_k) \) of \( \mathbb{Q}[y_1, y_2, \ldots, y_k] \) generated by

\[
\left\{ \prod_{j \in \{1, \ldots, \ell\}, v_j > 0} y_j^{v_j} - \prod_{i \in \{1, \ldots, \ell\}, v_i < 0} y_i^{-v_i} \mid (v_1, \ldots, v_\ell) \in \mathbb{Z} \right\},
\]

and \( \{y_{\ell+1}, \ldots, y_k\} \), where \( \mathbb{Z} \) is the multiplicative relation lattice.
Multiplicative relation ideal: example

Definition

Let \( A := (\alpha_1, \ldots, \alpha_k) \) be a sequence of \( k \) elements of \( \overline{\mathbb{Q}} \). Assume w.l.o.g. that for some \( \ell \), with \( 1 \leq \ell \leq k \), we have \( \alpha_1 \neq 0, \ldots, \alpha_\ell \neq 0 \), \( \alpha_{\ell+1} = \cdots \alpha_k = 0 \). We associate each \( \alpha_i \) with a “new” variable \( y_i \). The binomial ideal \( \text{MRI}(A; y_1, \ldots, y_k) \) of \( \mathbb{Q}[y_1, y_2, \ldots, y_k] \) generated by

\[
\{ \prod_{j \in \{1, \ldots, \ell\}, v_j > 0} y_j^{v_j} - \prod_{i \in \{1, \ldots, \ell\}, v_i < 0} y_i^{-v_i} \mid (v_1, \ldots, v_\ell) \in \mathbb{Z} \},
\]

and \( \{y_{\ell+1}, \ldots, y_k\} \), where \( \mathbb{Z} \) is the multiplicative relation lattice.

Example

Consider \( A = (1/2, 1/3, -1/6, 0) \). The multiplicative relation lattice of \( (1/2, 1/3, -1/6) \) is generated by \( (2, 2, -2) \). Thus the MRI of \( A \) associated with \( y_1, y_2, y_3, y_4 \) is

\[
\langle y_1^2y_2^2 - y_3^2, y_4 \rangle.
\]
Weak multiplicative independence

**Definition**

Let \( A := (\alpha_1, \ldots, \alpha_k) \) be a sequence of \( k \) non-zero algebraic numbers over \( \mathbb{Q} \) and let \( \beta \in \overline{\mathbb{Q}} \). We say \( \beta \) is weakly multiplicatively independent w.r.t. \( A \), if there exist no non-negative integers \( e_1, e_2, \ldots, e_k \) such that
\[
\beta = \prod_{i=1}^{k} \alpha_i^{e_i}
\]
holds.

Furthermore, we say that \( A \) is weakly multiplicatively independent if

(i) \( \alpha_1 \neq 1 \) holds, and

(ii) \( \alpha_i \) is weakly multiplicatively independent w.r.t. \( \{\alpha_1, \ldots, \alpha_{i-1}, 1\} \), for all \( i = 2, \ldots, s \).
Degree estimates for $x$ satisfying $x(n + 1) = \lambda x(n) + h(n)$

**Lemma**

Let $\alpha_1, \ldots, \alpha_k \in \overline{\mathbb{Q}} \setminus \{0, 1\}$. Let $\lambda \in \overline{\mathbb{Q}} \setminus \{0\}$. Let $h(n) \in \overline{\mathbb{Q}}[n, \alpha_1^n, \ldots, \alpha_k^n]$.

Consider the following single-variable recurrence relation $R$:

$$x(n + 1) = \lambda x(n) + h(n).$$

Then, there exists $s(n) \in \overline{\mathbb{Q}}[n, \alpha_1^n, \ldots, \alpha_k^n]$ such that we have

$$\deg(s(n), \alpha_i^n) \leq \deg(h(n), \alpha_i^n) \quad \text{and} \quad \deg(s(n), n) \leq \deg(h(n), n) + 1,$$

and such that

- if $\lambda = 1$ holds, then $s(n)$ solves $R$,
- if $\lambda \neq 1$ holds, then there exists a constant $c$ depending on $x(0)$ (that is, the initial value of $x$) such that $c \lambda^n + s(n)$ solves $R$.

Moreover, in both cases, if the exponential-free part of the canonical form of $(\frac{1}{\lambda})^n h(n)$ is 0, then $\deg(s(n), n) \leq \deg(h(n), n)$ can be required.

This latter hypothesis holds as soon as $\lambda$ is weakly multiplicatively independent w.r.t. $\alpha_1, \ldots, \alpha_k$. 
Plan

1. Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout’s Theorem

2. Invariant ideal of $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - $P$-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3. Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
Degree of an algebraic variety

Notations
Let $\mathbb{K}$ be an algebraically closed field. Let $F \subset \mathbb{K}[x_1, x_2, \ldots, x_s]$. We denote by $V_{\mathbb{K}^s}(F)$ (or simply by $V(F)$ when no confusion is possible) the zero set in $\mathbb{K}^s$ of $F$.

Definition
Let $V \subset \mathbb{K}^s$ be an $r$-dimensional equidimensional algebraic variety. The number of points of intersection of $V$ with an $(n - r)$-dimensional generic linear subspace $L \subset \mathbb{K}^s$ is called the degree of $V$, denoted by $\deg(V)$.

The degree of a non-equidimensional variety is defined to be the sum of the degrees of its equidimensional components.

The degree of an ideal $I \subset \mathbb{K}[x_1, x_2, \ldots, x_s]$ is defined to be the degree of the variety of $I$ in $\mathbb{K}^s$. 
A few well-known properties

Lemma

Let $V \subset \mathbb{K}^s$ be an $r$-dimensional equidimensional algebraic variety of degree $\delta$. Let $L$ be an $(n-r)$-dimensional linear subspace. Then, $L \cap V$ is either of positive dimensional or consists of no more than $\delta$ points.

Lemma

Let $V \subset \mathbb{K}^s$ be an algebraic variety. Let $L$ be a linear map from $\mathbb{K}^s$ to $\mathbb{K}^k$. Then we have $\deg(L(V)) \leq \deg(V)$.


Let $I \subset \mathbb{Q}[x_1, x_2, \ldots, x_s]$ be a radical ideal of degree $\delta$. Then there exist finitely many polynomials in $\mathbb{Q}[x_1, x_2, \ldots, x_s]$ generating $I$ and such that each of this polynomial has total degree less than or equal to $\delta$.

Lemma

Let $V, W, V_1, \ldots, V_e \subset \mathbb{K}^s$ be algebraic varieties such that $V := W \cap \bigcap_{i=1}^e V_i$ holds with $\dim(W) = r$. Then we have

$$\deg(V) \leq \deg(W) \max\{\deg(V_i) \mid i = 1 \cdots e\)^r.$$
A variation on Bezout’s Theorem

Proposition

Let $X = x_1, x_2, \ldots, x_s$ and $Y = y_1, y_2, \ldots, y_t$ be pairwise different $s + t$ variables.

Let $M$ be an ideal in $\mathbb{Q}[Y]$ of degree $d_M$ and dimension $r$.

Let $f_1, f_2, \cdots, f_s$ be $s$ polynomials in $\mathbb{Q}[Y]$, with maximum total degree $d_f$.

Denote by $I$ the ideal $\langle x_1 - f_1, x_2 - f_2, \ldots, x_s - f_s \rangle$.

Then, we have

$$\deg(I + M) \leq d_M d_f^r.$$ 

Remark

Since $I + M$ is an ideal of $\mathbb{Q}[X, Y]$, a direct application of one of the previous lemmas gives

$$\deg(I + M) \leq d_M d_f^{s+r}.$$
A variation on Bezout’s Theorem

This bound is tight

Example

Consider the polynomials of $\mathbb{Q}[x, y, n, m]

\[ g_1 := x - n^2 - n - m \] and
\[ g_2 := y - n^3 - 3n + 1 \]

and the ideals

\[ M := \langle n^2 - m^3 \rangle \] and \[ J := M + \langle g_1, g_2 \rangle \]

With the notations of the proposition we have

\[ d_M := 3, \ r := 1 \] and \[ d_f := 3 \]

Thus the estimated degree is $3 \times 3$. Meanwhile, the true degree of $J$ is indeed 9, which is computed as the (linear space) dimension of

\[ \mathbb{Q}(a, b, c, d, e)[x, y, m, n]/(J + \langle a x + b y + c n + d m + e \rangle), \]

where $a, b, c, d, e$ are indeterminates.
Plan

1. Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout’s Theorem

2. Invariant ideal of $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3. Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
Plan

1. Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout's Theorem

2. Invariant ideal of $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3. Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
The univariate case: recall

Definition

Given a recurrence \( R : x(n + 1) = \lambda x(n) + h(n) \) in \( \mathbb{Q} \), if \( h(n) \) is a poly-geometrical expression in \( n \) over \( \mathbb{Q} \), then \( R \) is called a univariate \( P \)-solvable recurrence.
The multivariate case: setting

Let \( n_1, \ldots, n_k \) be positive integers and define \( s := n_1 + \cdots + n_k \). Let \( M \) be a block-diagonal square matrix over \( \mathbb{Q} \) of order \( s \), with shape:

\[
M := \begin{pmatrix}
M_{n_1 \times n_1} & 0_{n_1 \times n_2} & \cdots & 0_{n_1 \times n_k} \\
0_{n_2 \times n_1} & M_{n_2 \times n_2} & \cdots & 0_{n_2 \times n_k} \\
\vdots & \vdots & \ddots & \vdots \\
0_{n_k \times n_1} & 0_{n_k \times n_2} & \cdots & M_{n_k \times n_k}
\end{pmatrix}
\]

Consider an \( s \)-variable recurrence relation \( R \) in \( x_1, x_2, \ldots, x_s \), with shape:

\[
\begin{pmatrix}
x_1(n + 1) \\
x_2(n + 1) \\
\vdots \\
x_s(n + 1)
\end{pmatrix} = M \times \begin{pmatrix}
x_1(n) \\
x_2(n) \\
\vdots \\
x_s(n)
\end{pmatrix} + \begin{pmatrix}
f_{1n_1 \times 1} \\
f_{2n_2 \times 1} \\
\vdots \\
f_{kn_k \times 1}
\end{pmatrix},
\]

where \( f_1 \) is a vector of length \( n_1 \) with coordinates in \( \mathbb{Q} \) and where \( f_i \) is a tuple of length \( n_i \) with coordinates in the polynomial ring \( \mathbb{Q}[x_1, \ldots, x_{n_1+\cdots+n_{i-1}}] \), for \( i = 2, \ldots, k \).
The multivariate case: definition

Setting (recall)

\[
\begin{pmatrix}
  x_1(n + 1) \\
  x_2(n + 1) \\
  \vdots \\
  x_s(n + 1)
\end{pmatrix}
= M \times
\begin{pmatrix}
  x_1(n) \\
  x_2(n) \\
  \vdots \\
  x_s(n)
\end{pmatrix}
+ \begin{pmatrix}
  f_1 n_1 \times 1 \\
  f_2 n_2 \times 1 \\
  \vdots \\
  f_k n_k \times 1
\end{pmatrix},
\]

where \( f_1 \) is a vector over \( \mathbb{Q} \) of length \( n_1 \) and where \( f_i \) is a tuple of length \( n_i \) with coordinates in \( \mathbb{Q}[x_1, \ldots, x_{n_1} + \ldots + n_{i-1}] \), for \( i = 2, \ldots, k \).

Definition

Then, the recurrence relation \( R \) is called \textit{P-solvable} over \( \mathbb{Q} \) and the matrix \( M \) is called the coefficient matrix of \( R \).

The notion of \textit{P-solvable} recurrence is equivalent to that of \textit{solvable mapping} in (E. Rodriguez-Carbonell & D. Kapur, ISSAC04) or that of \textit{solvable loop} in (L. Kovocs TACAS08) the respective contexts.
Plan

1 Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout's Theorem

2 Invariant ideal of $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - $P$-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3 Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis

Invariant ideal of $P$-solvable recurrences
Degree estimates for solutions of $P$-solvable recurrences
Degree estimates for solutions of \( P \)-solvable recurrences: theorem

Assume \( M \) is in a Jordan normal form. Assume the eigenvalues \( \lambda_1, \ldots, \lambda_s \) of \( M \) (counted with multiplicities) are different from 0, 1, with \( \lambda_i \) being the \( i \)-th diagonal element of \( M \). Assume for each block \( j \) the total degree of any polynomial in \( f_j \) (for \( i = 2 \cdots k \) ) is upper bounded by \( d_j \). For each \( i \), we denote by \( b(i) \) the block number of the index \( i \), that is,

\[
\sum_{j=1}^{b(i)-1} n_j < i \leq \sum_{j=1}^{b(i)} n_j.
\]

Let \( D_1 := n_1 \) and for all \( j \in \{2, \ldots, k\} \) let \( D_j := d_j D_{j-1} + n_j \). Then, there exists a solution \((y_1, y_2, \ldots, y_s)\) for \( R \) of the following form:

\[
y_i := c_i \lambda_i^n + g_i, \quad i = 1 \cdots s \quad \text{where}
\]

(a) \( c_i \) is a constant depending only on the initial value of the recurrence;

(b) \( g_i \) is a poly-geometrical expression in \( n \) w.r.t. \( \lambda_1, \ldots, \lambda_{i-1} \), such that

\[
\deg(g_i) \leq D_{b(i)}.
\]

Moreover, if \( \{\lambda_1, \ldots, \lambda_s\} \) is weakly multiplicatively independent, then, for all \( i = 1, \ldots, k \), we can further choose \( y_i \) such that we have

\[
\deg(g_i, n) = 0 \quad \text{and} \quad \deg(g_i) \leq \prod_{2 \leq t \leq b(i)} \max(d_t, 1).
\]
Degree estimates for solutions of $P$-solvable recurrences: example

Consider the recurrence:

$$
\begin{pmatrix}
  x(n+1) \\
  y(n+1) \\
  z(n+1)
\end{pmatrix} :=
\begin{pmatrix}
  2 & 0 & 0 \\
  0 & 3 & 0 \\
  0 & 0 & 3
\end{pmatrix} \times
\begin{pmatrix}
  x(n) \\
  y(n) \\
  z(n)
\end{pmatrix} +
\begin{pmatrix}
  0 \\
  x(n)^2 \\
  x(n)^3
\end{pmatrix}
$$

Viewing the recurrence as two blocks $(x)$ and $(y, z)$, the degree upper bounds are

$$D_1 := n_1 = 1 \quad \text{and} \quad D_2 := d_2 D_1 + n_2 = 3 \times 1 + 2.$$

If we decouple the $(y, z)$ block to the following two recurrences

$$y(n + 1) = 3y(n) + x(n)^2 \quad \text{and} \quad z(n + 1) = 3z(n) + x(n)^3,$$

then we deduce that the degree of the poly-geometrical expression for $y$ and $z$ are upper bounded by 2 and 3 respectively.
Degree estimates: reduction to the Jordan normal form case

Let $Q$ be a non-singular matrix such that $J := Q M Q^{-1}$ is a Jordan form of $M$. Let the original recurrence $R$ be

$$X(n + 1) = M X(n) + F.$$ 

Consider the following recurrence $R_Q$

$$Y(n + 1) = J Y(n) + QF.$$ 

It is easy to check that if

$$(y_1(n), y_2(n), \ldots, y_s(n))$$

solves $R_Q$, then

$$Q^{-1} (y_1(n), y_2(n), \ldots, y_s(n))$$

solves $R$. Note that an invertible matrix over $\mathbb{Q}$ maps a tuple of poly-geometrical expressions to another tuple of poly-geometrical expressions; moreover it preserves the highest degree among the expressions in the tuple.
## Plan

1. **Preliminaries**
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout's Theorem

2. **Invariant ideal of $P$-solvable recurrences**
   - Degree estimates for solutions of $P$-solvable recurrences
   - $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3. **Loop invariant generation via polynomial interpolation**
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
Invariant ideal of $P$-solvable recurrences

Degree estimates for the invariant ideal: theorem

- Let $R$ be a $P$-solvable recurrence relation with variables $(x_1, x_2, \ldots, x_s)$.
- Let $\mathcal{I} \subset \mathbb{Q}[x_1, x_2, \ldots, x_s]$ be the invariant ideal of $R$.
- Let $A = \alpha_1, \alpha_2, \ldots, \alpha_s$ be the eigenvalues (counted with multiplicities) of the coefficient matrix of $R$.
- Let $\mathcal{M}$ be the multiplicative relation ideal of $A$ associated with variables $y_1, \ldots, y_k$. Let $r$ be the dimension of $\mathcal{M}$.
- Let $f_1(n, \alpha_1^n, \ldots, \alpha_k^n), \ldots, f_s(n, \alpha_1^n, \ldots, \alpha_k^n)$ be $s$ poly-geometrical expressions in $n$ w.r.t. $\alpha_1, \alpha_2, \ldots, \alpha_s$ solving $R$.
- Suppose $R$ has a $k$-block configuration as $(n_1, 1), \ldots, (n_k, d_k)$.
- Let $D_1 := n_1$; and for all $j \in \{2, \ldots, k\}$, let $D_j := d_j D_{j-1} + n_j$.

Then, we have

$$\deg(\mathcal{I}) \leq \deg(\mathcal{M}) D_k^r + 1.$$ 

Moreover, if the degrees of $n$ in $f_i$, for $i = 1 \cdots s$, are all 0, then we have

$$\deg(\mathcal{I}) \leq \deg(\mathcal{M}) D_k^r.$$
Consider again solving for \((x, y)\) as a 2-variable recurrence

\[ x(n + 1) = y(n), \quad y(n + 1) = x(n) + y(n), \quad \text{with } x(0) = 0, y(0) = 1. \]

Recall that we obtained

\[
\begin{align*}
  x(n) &= \frac{(\sqrt{5}+1)^n}{\sqrt{5}} - \frac{(-\sqrt{5}+1)^n}{\sqrt{5}}, \\
  y(n) &= \frac{\sqrt{5}+1}{2} \left( \frac{(\sqrt{5}+1)}{2} \right)^n - \frac{-\sqrt{5}+1}{2} \left( \frac{-\sqrt{5}+1}{2} \right)^n.
\end{align*}
\]

Observe that \(A := \frac{-\sqrt{5}+1}{2}, \frac{\sqrt{5}+1}{2}\) is weakly multiplicatively independent. The multiplicative relation ideal of \(A\) associated with variables \(u, v\) is generated by \(u^2v^2 - 1\) and thus has degree 4 and dimension 1 in \(\mathbb{Q}[u, v]\). Therefore, the previous theorem implies that the degree of invariant ideal bounded by \(4 \times 1^1\). This is sharp since this ideal is

\[
\langle 1 - y^4 + 2xy^3 + x^2y^2 - 2x^3y - x^4 \rangle.
\]
Plan

1. Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout's Theorem

2. Invariant ideal of \( P \)-solvable recurrences
   - Degree estimates for solutions of \( P \)-solvable recurrences
   - \( P \)-solvable recurrences
   - Degree estimates for solutions of \( P \)-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3. Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
Dimension estimates for the invariant ideal: theorem

**Theorem**

Using the same notations as in the definition of $P$-solvable recurrences.

- Let $\lambda_1, \lambda_2, \ldots, \lambda_s$ be the eigenvalues of $M$ counted with multiplicities.
- Let $\mathcal{M}$ be the multiplicative relation ideal of $\lambda_1, \lambda_2, \ldots, \lambda_s$.
- Let $r$ be the dimension of $\mathcal{M}$. Let $\mathcal{I}$ be the invariant ideal of $R$.

Then, we have

$$ \dim(\mathcal{I}) \leq r + 1. $$

Moreover, for generic initial values,

1. we have $r \leq \dim(\mathcal{I})$,
2. if 0 is not an eigenvalue of $M$ and $\lambda_1, \lambda_2, \ldots, \lambda_s$ is weakly multiplicatively independent, then we have $r = \dim(\mathcal{I})$.

**Corollaries**

1. If $r + 1 < s$ holds, then $\mathcal{I}$ is not the zero ideal in $\mathbb{Q}[x_1, x_2, \ldots, x_s]$.
2. Assume that $x_1(0) := a_1, \ldots, x_s(0) := a_s$ are independent indeterminates. If the eigenvalues of $R$ are multiplicatively independent, then the inductive invariant ideal of the loop is the zero ideal in $\mathbb{Q}[a_1, \ldots, a_s, x_1, x_2, \ldots, x_s]$. 
Consider the recurrence:

$$(x(n + 1), y(n + 1)) := (3x(n) + y(n), 2y(n))$$

with $x(0) = a, y(0) = b$.

On one hand, the two eigenvalues are 2 and 3 which are multiplicatively independent.

Therefore, using the previous corollary, the invariant ideal of the corresponding loop is trivial.

On the other hand, for loop variables $(a, b, x, y)$, the reachable set of the loop is

$$\mathcal{R} := \{(a, b, (a+b)3^i - b2^i, b2^i) \mid (a, b) \in \mathbb{Q}^2, \ i \text{ is a non-negative integer}\}.$$ 

Therefore, any polynomial vanishes on all points of $\mathcal{R}$ must be 0.
Consider the linear recurrence

\[ x(n + 1) = 3x(n) - y(n), \quad y(n + 1) = 2y(n) \]

with \((x(0), y(0)) = (a, b)\).

The eigenvalues of the coefficient matrix are 2, 3, which are multiplicatively independent.

One can check that, when \(a = b\), the invariant ideal is generated by \(x - y\).

However, generically, that is when \(a \neq b\) holds, the invariant ideal is the zero ideal.
Plan

1. Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout’s Theorem

2. Invariant ideal of \( P \)-solvable recurrences
   - Degree estimates for solutions of \( P \)-solvable recurrences
   - \( P \)-solvable recurrences
   - Degree estimates for solutions of \( P \)-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3. Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
Loop invariant generation via polynomial interpolation

A direct approach

Loop model under study: recall

while $C_0$ do
  if $C_1$ then
    \[ X := A_1(X); \]
  elif $C_2$ then
    \[ X := A_2(X); \]
  \ldots
  elif $C_m$ then
    \[ X := A_m(X); \]
  end if
end while

1. Loop variables: $X = x_1, \ldots, x_s$, rational value scalar
2. Conditions: each $C_i$ is a quantifier free formula in $X$ over $\mathbb{Q}$.
3. Assignments: $A_i \in \mathbb{Q}[X]$ inducing a polynomial map $M_i : \mathbb{R}^s \mapsto \mathbb{R}^s$
4. Initial condition: $X$-values defined by a semi-algebraic system.
A direct approach

Input

(i) $M := m_1, m_2, \ldots, m_c$ is a sequence of monomials in the loop variables $X$,
(ii) $S := s_1, s_2, \ldots, s_r$ is a set of $r$ points on the inductive trajectory of the loop,
(iii) $E$ is a polynomial system defining the loop initial values,
(iv) $B$ is the transitions $(C_1, A_1), \ldots, (C_m, A_m)$ of the loop.

Algorithm

1. $L := \text{BuildLinSys}(M, S)$
2. $N := \text{LinSolve}(L)$ is full row rank and generates the null space of $L$.
3. $F := \emptyset$
4. For each row vector $v \in N$ do
   
   $F := F \cup \{\text{GenPoly}(M, v)\}$
5. If $Z(E) \not\subseteq Z(F)$ then return FAIL
6. For each branch $(C_i, A_i) \in B$ do
   
   if $A_i(Z(F) \cap Z(C_i)) \not\subseteq Z(F)$ then return FAIL
7. Return $F$, a list of polynomial equation invariants for the target loop.
Plan

1. Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout’s Theorem

2. Invariant ideal of $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3. Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
A small-prime approach: algorithm

Algorithm

1. \( p := \text{MaxMachinePrime}(); \) \( L_p := \text{BuildLinSysModp}(M, S, p); \)
2. \( N_p := \text{LinSolveModp}(L_p, p) \)
3. \( d := \dim(N_p); \) \( N := (N_p); \) \( P := (p); \)
4. While \( p > 2 \) do
   1. If \( d = 0 \) then return FAIL
   2. \( N := \text{RatRecon}(N, P) \)
   3. If \( N \neq \text{FAIL} \) then break;
   4. \( p := \text{PrevPrime}(p); \) \( L_p := \text{BuildLinSysModp}(M, S, p); \)
   5. \( N_p := \text{LinSolveModp}(L_p, p) \)
   6. If \( d > \dim(N_p) \) then \( d := \dim(N_p); \) \( N := (N_p); \) \( P := (p) \)
   7. else \( N := \text{Append}(N, N_p); \) \( P := \text{Append}(P, p) \)
5. If \( p = 2 \) then return FAIL
6. \( F := \emptyset \)
7. For each row vector \( v \in N \) do
   \[ F := F \cup \{\text{GenPoly}(M, v)\} \]
8. If \( Z(E) \not\subseteq Z(F) \) then return FAIL
9. For each branch \( (C_i, A_i) \in B \) do
   if \( A_i(Z(F) \cap Z(C_i)) \not\subseteq Z(F) \) then return FAIL
10. Return \( F \), a list of polynomial equation invariants for the target loop.
A small-prime approach: complexity result

Proposition

Both algorithms run in singly exponential time w.r.t. number of loop variables.

Indeed

- the number of monomials of $M$ is singly exponential w.r.t. number of loop variables.
- applying our criterion to certify the result can be reduced to an ideal membership problem, which is singly exponential w.r.t. number of loop variables.
A small-prime approach: example

Consider the following recurrence relation on \((x, y, z)\):

\[
\begin{pmatrix}
x(n + 1) \\
y(n + 1) \\
z(n + 1)
\end{pmatrix} =
\begin{pmatrix}
0 & 0 & 1 \\
1 & 0 & -3 \\
0 & 1 & 3
\end{pmatrix}
\begin{pmatrix}
x(n) \\
y(n) \\
z(n)
\end{pmatrix}
\]

with initial value \((x(0), y(0), z(0)) = (1, 2, 3)\).

- Note that the characteristic polynomial of the coefficient matrix has 1 as a triple root and the mult. rel. ideal of the eigenvalues is 0-dimensional.
- So the invariant ideal of this recurrence has dimension either 0 or 1.
- On the other hand, we can show that for all \(k \in \mathbb{N}\), we have \(M^k \neq M\); so there are infinitely many points in the set \(\{(x(k), y(k), z(k)) \mid k \in \mathbb{N}\}\), whenever \((x(0), y(0), z(0)) \neq (0, 0, 0)\).
- With our method, we compute the following invariant polynomials

\[
x + y + z - 6, y^2 + 4yz + 4z^2 - 6y - 24z + 20,
\]

which generate a prime ideal of dimension 1, thus the invariant ideal of this.
Plan

1 Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout’s Theorem

2 Invariant ideal of $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3 Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
Implementation of the small-prime approach

- In Maple using LinearAlgebra and RegularChains.
- The interpolation part is done naively: the template set $M$ consists of all monomials up to the target degree.
- A sparse interpolation scheme is work in progress.
- We handle semi-algebraic conditions thanks to RegularChains:-SemiAlgebraicSetTools
- We have applied our code to all example programs used in (E. Rodriguez-Carbonell & D. Kapur, 2007):
  - We are able to find the loop invariants by trying total degree up to 4 for most loops within 60 seconds.
  - In each case, we return a system of generators of the invariant ideal, though we do not have a proof for that fact.
Benchmarks procedure

- “# vars” is the number of loop variables,
- “deg” is the total degree tried for the methods which use a degree bound,
- “PI” is the timing of the our method,
- “AI” (Abstract Interpretation) is the timing of the method described in (E. Rodriguez-Carbonell & D. Kapur, TCS 2007)
- “FP” (ideal fix point, direct use of Gröbner basis techniques) is the timing of the method described in (E. Rodriguez-Carbonell & D. Kapur, JSC 2007)
- “SE” (solving and elimination, direct use of Gröbner basis techniques) is the timing of the method described in (L. Kovocs TACAS08) and implementated in the software ALIGATOR.

- The time unit is the second;
- the “NA” symbol in a time field means that the related method does support the input program;
- the “FAIL” symbol in a time field means that the output is not “correct”.
- All the tests were done using an Intel Core 2 Quad CPU 2.40GHz with 8.0GB memory.
- Computations of multiplicative relation lattice were done (not needed for “PI”) on the same machine with GAP 4.4.12 + Alnuth 2.3.1 + KASH 2.5.
### Timings

<table>
<thead>
<tr>
<th>prog.¹</th>
<th># vars</th>
<th>deg</th>
<th>PI</th>
<th>AI</th>
<th>FP</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>cohencu</td>
<td>4</td>
<td>3</td>
<td>0.6</td>
<td>0.93</td>
<td>0.28</td>
<td>0.13</td>
</tr>
<tr>
<td>cohencu</td>
<td>4</td>
<td>2</td>
<td>0.06</td>
<td>0.76</td>
<td>0.28</td>
<td>0.13</td>
</tr>
<tr>
<td>fermat</td>
<td>5</td>
<td>4</td>
<td>3.74</td>
<td>0.79</td>
<td>0.37</td>
<td>0.1</td>
</tr>
<tr>
<td>prodbin</td>
<td>5</td>
<td>3</td>
<td>1.4</td>
<td>0.74</td>
<td>0.36</td>
<td>0.13</td>
</tr>
<tr>
<td>rk07</td>
<td>6</td>
<td>3</td>
<td>3.1</td>
<td>2.23</td>
<td>NA</td>
<td>0.35</td>
</tr>
<tr>
<td>kov08</td>
<td>3</td>
<td>3</td>
<td>0.2</td>
<td>0.57</td>
<td>0.22</td>
<td>0.01</td>
</tr>
<tr>
<td>sum5</td>
<td>4</td>
<td>5</td>
<td>12</td>
<td>1.60</td>
<td>2.25</td>
<td>0.16²</td>
</tr>
<tr>
<td>wensley2</td>
<td>3</td>
<td>3</td>
<td>0.4</td>
<td>0.84</td>
<td>0.39</td>
<td>0.21</td>
</tr>
<tr>
<td>int-factor</td>
<td>6</td>
<td>3</td>
<td>60.9</td>
<td>1.28</td>
<td>160.7</td>
<td>0.9</td>
</tr>
<tr>
<td>fib(coupled)</td>
<td>4</td>
<td>4</td>
<td>2.4</td>
<td>0.71</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>fib(decoupled)</td>
<td>6</td>
<td>4</td>
<td>4.3</td>
<td>1.28</td>
<td>160.7</td>
<td>FAIL</td>
</tr>
<tr>
<td>non-inv2*</td>
<td>4</td>
<td>3</td>
<td>1.2</td>
<td>3.83</td>
<td>NA</td>
<td>FAIL</td>
</tr>
<tr>
<td>coupled-5-1*</td>
<td>4</td>
<td>4</td>
<td>1.1</td>
<td>9.58</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>coupled-5-2*</td>
<td>5</td>
<td>4</td>
<td>5.38</td>
<td>15.8</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>mannadiv</td>
<td>3</td>
<td>3</td>
<td>0.1</td>
<td>0.83</td>
<td>NA</td>
<td>0.04</td>
</tr>
</tbody>
</table>
Plan

1 Preliminaries
   - Notions on loop invariants
   - Poly-geometric summations
   - A variation on Bezout's Theorem

2 Invariant ideal of $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - $P$-solvable recurrences
   - Degree estimates for solutions of $P$-solvable recurrences
   - Degree estimates for their invariant ideal
   - Dimension estimates for their invariant ideal

3 Loop invariant generation via polynomial interpolation
   - A direct approach
   - A modular method
   - Experimentation
   - Maple Package: ProgramAnalysis
ProgramAnalysis: package architecture

Maple Interface (Commands and Applet)

Parser Utilities (Implemented in C)

Polynomial System Solving Engine (RegularChains library)
Maple session: the input program in a file

\begin{verbatim}
wensley2 := proc(P, Q, E)
local a, b, d, y;
    a := 0;
    b := 1/2 * Q;
    d := 1;
    y := 0;
    #PRE: Q > P and P \geq 0 and E > 0
    while E \leq d do
        if P < a + b then
            b := 1/2 * b;
            d := 1/2 * d
        else
            a := a + b;
            y := y + 1/2 * d;
            b := 1/2 * b;
            d := 1/2 * d
        end if
    end do;
    #POST: P/Q \geq y and y > P/Q - E
    return y
end proc
\end{verbatim}
Loop invariant generation via polynomial interpolation

Maple Package: ProgramAnalysis

Maple session: the sample points

\[
\begin{bmatrix}
0, \frac{5}{2}, 1, 0 \\
\frac{5}{2}, \frac{5}{4}, \frac{1}{2}, \frac{1}{2} \\
\frac{5}{2}, \frac{5}{8}, \frac{1}{4}, \frac{1}{2} \\
\frac{5}{2}, \frac{5}{16}, \frac{1}{8}, \frac{1}{2} \\
\frac{45}{16}, \frac{5}{32}, \frac{1}{16}, \frac{9}{16} \\
\frac{95}{32}, \frac{5}{64}, \\
\frac{1}{32}, \frac{19}{32} \\
\frac{95}{32}, \frac{5}{128}, \frac{1}{64}, \frac{19}{32} \\
\frac{95}{32}, \frac{5}{256}, \frac{1}{128}, \frac{19}{32} \\
\frac{765}{256}, \frac{5}{512}, \frac{1}{256}, \frac{153}{256} \\
\frac{1535}{512}, \\
\frac{5}{1024}, \frac{1}{512}, \frac{307}{512} \\
\frac{1535}{512}, \frac{5}{2048}, \frac{1}{1024}, \frac{307}{512} \\
\frac{1535}{512}, \frac{5}{4096}, \frac{1}{2048}, \frac{307}{512} \\
\frac{12285}{4096}, \frac{5}{8192}, \\
\frac{1}{4096}, \frac{2457}{4096} \\
\frac{24575}{8192}, \frac{5}{16384}, \frac{1}{8192}, \frac{4915}{8192} \\
\frac{24575}{8192}, \frac{5}{32768}, \frac{1}{16384}, \frac{4915}{8192} \\
\frac{24575}{8192}, \\
\frac{5}{65536}, \frac{1}{32768}, \frac{4915}{8192} \\
\frac{196605}{65536}, \frac{5}{131072}, \frac{1}{65536}, \frac{39321}{65536} \\
\frac{393215}{131072}, \frac{5}{262144}, \frac{1}{131072}, \frac{78643}{131072} \\
\frac{78643}{131072}, \frac{5}{262144}, \frac{1}{131072}, \frac{393215}{131072} \\
\frac{393215}{131072}, \frac{5}{524288}, \frac{1}{262144}, \frac{78643}{131072} \\
\frac{393215}{131072}, \frac{5}{1048576}, \frac{1}{524288}, \frac{78643}{131072} \\
\frac{3145725}{1048576}, \frac{5}{2097152}, \frac{1}{1048576}, \frac{629145}{1048576} \\
\frac{6291455}{2097152}, \frac{5}{4194304}, \frac{1}{2097152}, \frac{1258291}{2097152}
\end{bmatrix}
\]
Loop invariant generation via polynomial interpolation

Maple Package: ProgramAnalysis

Maple session: verifying the program

```maple
> mplfile := cat(getenv("MXHOME"), "/mx-2012/programs/wensley2.mpl");
precond := [[Q>P, P>=0, E>0]];
postcond := [[P >= Q*y , Q*y > P - Q*E ]];
guard := [[E<=d]];
ineq_invs := [ P - Q*d < Q*y, Q*y <= P, y>=0];

precond := [[P < Q, 0 <= P, 0 < E]]
postcond := [[Q*y <= P, P - Q*E < Q*y]]
guard := [[E <= d]]
ineq_invs := [ -d Q + P < Q y, Q y <= P, 0 <= y]

> st := time():
eq_invs := LoopEqInv(mplfile); # compute equation invariants
time()-st;
eq_invs := [y Q - a, d Q - 2 b, -2 b y + a d]
0.210

> # verify the specification of the program
st:=time():
LoopVerify(precond, guard, [[op(eq_invs), op(ineq_invs)]], postcond);
time()-st;
true
1.380
```
Xie Xie!