Formal Methods in Software Developement

The basic ideas of the virtual substitution and the cylindrical algebraic decomposition for solving real arithmetic problems

Mădălina Erașcu

West University of Timișoara Faculty of Mathematics and Informatics

Based on slides of the lecture Satisfiability Checking (Erika Ábrahám), RTWH Aachen

December 1, 2018

Domain	+	+, ·
Reals R	linear real arithmetic decidable Fourier-Motzkin Simplex	non-linear real arithmetic decidable Interval constraint propagation Virtual substitution Cylindrical algebraic decomposition
Integers Z	linear integer arithmetic decidable Branch-and-bound Omega test	non-linear integer arithmetic undecidable Bit-blasting Branch-and-bound

Reminder: Non-linear real arithmetic (NRA)

Real arithmetic: first-order theory ($\mathbb{R},+,\cdot,0,1,<$) over the reals with addition and multiplication.

Syntax of real arithmetic

Polynomials: $t := 0 \mid 1 \mid x \mid t+t \mid t \cdot t$

Constraints: c := t < t

Formulas: $\varphi ::= c \mid \neg \varphi \mid \varphi \land \varphi \mid \exists x. \varphi$

where x is a variable.

- Syntactic sugar for constraints: $t_1 \le t_2$, $t_1 = t_2$, $t_1 \ne t_2$.
- Notation: $D[x_1, \ldots, x_n]$ $(n \ge 1)$ is the set of all polynomials in variables x_1, \ldots, x_n with coefficients from some domain D.
- E.g., $\mathbb{Z}[x_1, \dots, x_n]$ is the set of all polynomials over variables x_1, \dots, x_n with coefficients from \mathbb{Z} .
- Naming in math: real algebra (instead of real arithmetic).

We consider the satisfiability problem for the quantifier-free fragment QFNRA of real arithmetic (equivalently, we consider the existential fragment, i.e., no universal quantifiers and no negation of expressions containing quantifiers).

We consider the satisfiability problem for the quantifier-free fragment QFNRA of real arithmetic (equivalently, we consider the existential fragment, i.e., no universal quantifiers and no negation of expressions containing quantifiers).

Given a QFNRA formula φ containing polynomials from $\mathbb{Z}[x_1,\ldots,x_n]$, the QFNRA satisfiability problem is to decide whether there are real values $v=(v_i,\ldots,v_n)\in\mathbb{R}^n$ such that substituting v_i for x_i for each $i=1,\ldots n$ in φ (notation: $\varphi[\vec{v}/\vec{x}]$) evaluates the formula to true.

- We consider the satisfiability problem for the quantifier-free fragment QFNRA of real arithmetic (equivalently, we consider the existential fragment, i.e., no universal quantifiers and no negation of expressions containing quantifiers).
 - Given a QFNRA formula φ containing polynomials from $\mathbb{Z}[x_1,\ldots,x_n]$, the QFNRA satisfiability problem is to decide whether there are real values $v=(v_i,\ldots,v_n)\in\mathbb{R}^n$ such that substituting v_i for x_i for each $i=1,\ldots n$ in φ (notation: $\varphi[\vec{v}/\vec{x}]$) evaluates the formula to true.
- QFLRA (quantifier-free linear real arithmetic) example:

$$\exists x. \exists y. \ x + 2y > 10 \land x \ge y \land (x < 0 \lor 2y > x)$$

- We consider the satisfiability problem for the quantifier-free fragment QFNRA of real arithmetic (equivalently, we consider the existential fragment, i.e., no universal quantifiers and no negation of expressions containing quantifiers).
 - Given a QFNRA formula φ containing polynomials from $\mathbb{Z}[x_1,\ldots,x_n]$, the QFNRA satisfiability problem is to decide whether there are real values $v=(v_i,\ldots,v_n)\in\mathbb{R}^n$ such that substituting v_i for x_i for each $i=1,\ldots n$ in φ (notation: $\varphi[\vec{v}/\vec{x}]$) evaluates the formula to true.
- QFLRA (quantifier-free linear real arithmetic) example:

$$\exists x. \exists y. \ x + 2y > 10 \land x \ge y \land (x < 0 \lor 2y > x)$$

QFNRA (quantifier-free non-linear real arithmetic) example:

$$\exists x. \ \exists y. \ (x^2 - 4x^3y^2 > 0 \land x - y = 1)$$

Notations

- Assume: variables $x_1, ..., x_n$, coefficient domain $D = \mathbb{Z}$
- Monomial: product of variables (the empty product represents the constant 1). Examples: xy^2 , u^3vz^2
- Term: product of an integer coefficient and a monomial.
 - Examples: $2xy^2$, $3u^3vz^2$
- Polynomial $p \in \mathbb{Z}[x_1, ..., x_n]$: sum of terms
 - Example: $2xy^2 + 3u^3vz^2$
- Polynomial constraint in canonical form: $p \sim 0$, $\sim \in \{<, \leq, =, \geq, >\}$.
 - Example: $2xy^2 + 3u^3vz^2 5 < 0$
- A polynomial in one variable is called univariate.
 - A polynomial in more than one variables is called multivariate.
 - Multivariate polynomials can be seen as univariate polynomials with polynomial coefficients (notation: $p \in \mathbb{Z}[x_1, \dots, x_{n-1}][x_n]$).

Real arithmetic: On the border of decidability

Theorem (Alfred Tarski 1948)

The FO theory of $(\mathbb{R}, +, \cdot, 0, 1, <)$ is decidable.

- Tarski's proof was constructive, i.e., it defined a decision procedure.
- However, its time-complexity in the number of variables was non-elementary ("greater than all finite towers of powers of 2").

Real arithmetic: Some historical facts

- 1637 Descartes' rule of signs
- 1835 Jaques Charles François Sturm's theorem
- 1948 Alfred Tarski's "A decision method for elementary algebra and geometry"
- 1975 Cylindrical algebraic decomposition (CAD) method by George E. Collins
- 1979-80 First implementation of the CAD method by Dennis S. Arnon
 - 1988 Virtual substitution by Volker Weispfenning
 - 1990 First implementation of virtual substitution (Klaus-Dieter Burhenne)
 - 1993 Gröbner bases approach by P. Pedersen, M.-F. Roy, A. Szpirglas, later extended by V. Weispfenning
 - 1994 Implementation of the Gröbner bases approach (Andreas Dolzmann)

Real arithmetic: Implementations

Interval arithmetic

Ariadne, iSAT, SMT-RAT, ...

Virtual substitution (VS)

■ Computer algebra system Redlog, SMT-RAT, ...

Cylindrical algebraic decomposition (CAD)

QEPCAD, Redlog, SMT-RAT, ...

Real arithmetic: Implementations

Interval arithmetic

Ariadne, iSAT, SMT-RAT, ...

Virtual substitution (VS)

■ Computer algebra system Redlog, SMT-RAT, ...

Cylindrical algebraic decomposition (CAD)

QEPCAD, Redlog, SMT-RAT, ...

The virtual substitution (VS) and the cylindrical algebraic decomposition (CAD) are quantifier elimination methods.

The idea of quantifier elimination

Given: FO sentence φ over $(\mathbb{R},+,\cdot,0,1,<)$ containing n quantifiers

1 Transform φ into prenex normal form:

$$\varphi \equiv Q_1 x_1 \dots Q_n x_n \varphi_n(x_1, \dots, x_n)$$

where φ_n is a quantifier-free NRA formula with variables x_1, \dots, x_n .

2 Eliminate iteratively the quantifiers $Q_n \dots Q_1$ and thus the quantified variables:

$$\varphi \equiv Q_1 x_1 \dots Q_{n-1} x_{n-1} Q_n x_n \quad \varphi_n(x_1, \dots, x_n)
\equiv Q_1 x_1 \dots Q_{n-1} x_{n-1} \quad \varphi_{n-1}(x_1, \dots, x_{n-1})
\dots
\equiv Q_1 x_1 \quad \varphi_1(x_1)
\equiv \varphi_0()$$

Removing universal quantification

Is it sufficient to eliminate existential quantifiers?

Removing universal quantification

Is it sufficient to eliminate existential quantifiers?

$$\exists x_1. \exists x_2. \quad \forall x_3. \quad \exists x_4. \quad \forall x_5. \quad \forall x_6. \quad \exists x_7. \exists x_8. \quad \varphi'$$

$$\equiv \exists x_1. \exists x_2. \ \neg(\exists x_3. \ \neg(\exists x_4. \ \neg(\exists x_5. \ \neg(\neg(\exists x_6. \ \neg(\exists x_7. \ \exists x_8. \quad \varphi' \quad)))))))$$

$$\equiv \exists x_1. \exists x_2. \ \neg(\exists x_3. \ \neg(\exists x_4. \ \neg(\exists x_5. \quad \exists x_6. \ \neg(\exists x_7. \ \exists x_8. \quad \varphi' \quad)))))$$

Removing universal quantification

Is it sufficient to eliminate existential quantifiers?

$$\exists x_1. \ \exists x_2. \quad \forall x_3. \quad \exists x_4. \quad \forall x_5. \quad \forall x_6. \quad \exists x_7. \ \exists x_8. \quad \varphi'$$

$$\equiv \exists x_1. \ \exists x_2. \ \neg(\exists x_3. \ \neg(\exists x_4. \ \neg(\exists x_5. \ \neg(\neg(\exists x_6. \ \neg(\exists x_7. \ \exists x_8. \quad \varphi' \))))))$$

$$\equiv \exists x_1. \ \exists x_2. \ \neg(\exists x_3. \ \neg(\exists x_4. \ \neg(\exists x_5. \quad \exists x_6. \ \neg(\exists x_7. \ \exists x_8. \quad \varphi' \))))$$

But: increased complexity

$$p \ge 0 \qquad \equiv \\
p \le 0 \qquad \equiv \\
p > 0 \qquad \equiv \\
p < 0 \qquad \equiv \\
p \ne 0 \qquad \equiv$$

$$p \ge 0$$
 \equiv $\exists \epsilon. \ p - \epsilon^2 = 0$
 $p \le 0$ \equiv
 $p > 0$ \equiv
 $p < 0$ \equiv
 $p \ne 0$ \equiv

$$p \ge 0$$
 \equiv $\exists \epsilon. \ p - \epsilon^2 = 0$
 $p \le 0$ \equiv $\exists \epsilon. \ p + \epsilon^2 = 0$
 $p > 0$ \equiv
 $p < 0$ \equiv
 $p \ne 0$ \equiv

$$\begin{array}{lll} p \geq 0 & \equiv & \exists \epsilon. \ p - \epsilon^2 = 0 \\ p \leq 0 & \equiv & \exists \epsilon. \ p + \epsilon^2 = 0 \\ p > 0 & \equiv & \exists \epsilon. \ 1 - p \cdot \epsilon^2 = 0 \\ p < 0 & \equiv & \\ p \neq 0 & \equiv & \end{array}$$

$$p \ge 0 \qquad \equiv \qquad \exists \epsilon. \ p - \epsilon^2 = 0$$

$$p \le 0 \qquad \equiv \qquad \exists \epsilon. \ p + \epsilon^2 = 0$$

$$p > 0 \qquad \equiv \qquad \exists \epsilon. \ 1 - p \cdot \epsilon^2 = 0$$

$$p < 0 \qquad \equiv \qquad \exists \epsilon. \ 1 + p \cdot \epsilon^2 = 0$$

$$p \ne 0 \qquad \equiv \qquad \exists \epsilon. \ 1 + p \cdot \epsilon^2 = 0$$

$$p \ge 0 \qquad \equiv \qquad \exists \epsilon. \ p - \epsilon^2 = 0$$

$$p \le 0 \qquad \equiv \qquad \exists \epsilon. \ p + \epsilon^2 = 0$$

$$p > 0 \qquad \equiv \qquad \exists \epsilon. \ 1 - p \cdot \epsilon^2 = 0$$

$$p < 0 \qquad \equiv \qquad \exists \epsilon. \ 1 + p \cdot \epsilon^2 = 0$$

$$p \ne 0 \qquad \equiv \qquad \neg (p = 0)$$

Is it sufficient to handle equations?

$$p \ge 0 \qquad \equiv \qquad \exists \epsilon. \ p - \epsilon^2 = 0$$

$$p \le 0 \qquad \equiv \qquad \exists \epsilon. \ p + \epsilon^2 = 0$$

$$p > 0 \qquad \equiv \qquad \exists \epsilon. \ 1 - p \cdot \epsilon^2 = 0$$

$$p < 0 \qquad \equiv \qquad \exists \epsilon. \ 1 + p \cdot \epsilon^2 = 0$$

$$p \ne 0 \qquad \equiv \qquad \neg (p = 0)$$

But: increased complexity



Quantifier elimination with VS and CAD: Finite abstraction

- The degree of a polynomial is the highest degree of its monomials, when expressed in canonical form. The degree of a monomial is the sum of the exponents of the variables that appear in it. The word degree is now standard, but in some older books, the word order may be used instead.
- Each univariate polynomial p(x) of degree d has d complex roots.
- Each univariate polynomial p(x) of degree d has at most d real roots.

Quantifier elimination with VS and CAD: Finite abstraction

- The degree of a polynomial is the highest degree of its monomials, when expressed in canonical form. The degree of a monomial is the sum of the exponents of the variables that appear in it. The word degree is now standard, but in some older books, the word order may be used instead.
- Each univariate polynomial p(x) of degree d has d complex roots.
- Each univariate polynomial p(x) of degree d has at most d real roots.
- The sign of *p* is invariant between each two successive real roots.

Quantifier elimination with VS and CAD: Finite abstraction

- The degree of a polynomial is the highest degree of its monomials, when expressed in canonical form. The degree of a monomial is the sum of the exponents of the variables that appear in it. The word degree is now standard, but in some older books, the word order may be used instead.
- Each univariate polynomial p(x) of degree d has d complex roots.
- Each univariate polynomial p(x) of degree d has at most d real roots.
- The sign of p is invariant between each two successive real roots. This implies that, if we know all roots, we can partition \mathbb{R} into at most 2d + 1 sign invariant regions for p.
- Similar facts hold also for formulas: for each QFNRA formula there is a finite partitioning of the state space such that the formula's truth value is invariant in each partition.

• Given: $\varphi = \exists x_1....\exists x_n. \varphi_n$, where φ_n is a quantifier-free FO sentence over $(\mathbb{R}, +, \cdot, 0, 1, <)$

- Given: $\varphi = \exists x_1....\exists x_n. \varphi_n$, where φ_n is a quantifier-free FO sentence over $(\mathbb{R}, +, \cdot, 0, 1, <)$
- Problem: \mathbb{R} is uncountably infinite.

- Given: $\varphi = \exists x_1....\exists x_n. \varphi_n$, where φ_n is a quantifier-free FO sentence over $(\mathbb{R}, +, \cdot, 0, 1, <)$
- Problem: \mathbb{R} is uncountably infinite.
- Idea: Find a finite set $T \subset \mathbb{R}$ with

$$\exists x_1. \ldots \exists x_n. \varphi_n \quad \Leftrightarrow \quad \exists x_1. \ldots \exists x_{n-1}. \bigvee_{t \in T} \varphi_n[t/x_n]$$

- Given: $\varphi = \exists x_1....\exists x_n. \varphi_n$, where φ_n is a quantifier-free FO sentence over $(\mathbb{R}, +, \cdot, 0, 1, <)$
- Problem: \mathbb{R} is uncountably infinite.
- Idea: Find a finite set $T \subset \mathbb{R}$ with

$$\exists x_1. \ldots \exists x_n. \varphi_n \Leftrightarrow \exists x_1. \ldots \exists x_{n-1}. \bigvee_{t \in T} \varphi_n[t/x_n]$$

T consists of a test (sample) points from all sign-invariant regions that might contain solutions.

- Given: $\varphi = \exists x_1. \ldots \exists x_n. \varphi_n$, where φ_n is a quantifier-free FO sentence over $(\mathbb{R}, +, \cdot, 0, 1, <)$
- Problem: \mathbb{R} is uncountably infinite.
- Idea: Find a finite set $T \subset \mathbb{R}$ with

$$\exists x_1. \ldots \exists x_n. \varphi_n \Leftrightarrow \exists x_1. \ldots \exists x_{n-1}. \bigvee_{t \in T} \varphi_n[t/x_n]$$

T consists of a test (sample) points from all sign-invariant regions that might contain solutions.

■ What remains: Determine the real roots of polynomials.

Real roots of univariate polynomials

What are the degrees and the real roots of these polynomials?

Polynomial	Degree	Real roots
X		
2x - 5		
x^2		
$x^2 - 1$		
$x^2 + 1$		
$x^2 - 2$		
$2x^6 - 5x^4 + 3x^2 - 6$		

Real roots of univariate polynomials

What are the degrees and the real roots of these polynomials?

Polynomial	Degree	Real roots
X	1	
2x - 5	1	
x^2	2	
$x^2 - 1$	2	
$x^2 + 1$	2	
$x^2 - 2$	2	
$2x^6 - 5x^4 + 3x^2 - 6$	6	

Polynomial	Degree	Real roots
X	1	0
2x - 5	1	
x^2	2	
$x^2 - 1$	2	
$x^2 + 1$	2	
$x^2 - 2$	2	
$2x^6 - 5x^4 + 3x^2 - 6$	6	

Polynomial	Degree	Real roots
X	1	0
2x - 5	1	2.5
x^2	2	
$x^2 - 1$	2	
$x^2 + 1$	2	
$x^2 - 2$	2	
$2x^6 - 5x^4 + 3x^2 - 6$	6	

Polynomial	Degree	Real roots
X	1	0
2x - 5	1	2.5
x^2	2	0
$x^2 - 1$	2	
$x^2 + 1$	2	
$x^2 - 2$	2	
$2x^6 - 5x^4 + 3x^2 - 6$	6	

Polynomial	Degree	Real roots
X	1	0
2x - 5	1	2.5
x^2	2	0
$x^2 - 1$	2	1, -1
$x^2 + 1$	2	
$x^2 - 2$	2	
$2x^6 - 5x^4 + 3x^2 - 6$	6	

Polynomial	Degree	Real roots
X	1	0
2x - 5	1	2.5
x^2	2	0
$x^2 - 1$	2	1, -1
$x^2 + 1$	2	-
$x^2 - 2$	2	
$2x^6 - 5x^4 + 3x^2 - 6$	6	

Polynomial	Degree	Real roots
X	1	0
2x - 5	1	2.5
x^2	2	0
$x^2 - 1$	2	1, -1
$x^2 + 1$	2	-
$x^2 - 2$	2	$\sqrt{2}$, $-\sqrt{2}$
$2x^6 - 5x^4 + 3x^2 - 6$	6	

Polynomial	Degree	Real roots
X	1	0
2x - 5	1	2.5
x^2	2	0
$x^2 - 1$	2	1, -1
$x^2 + 1$	2	-
$x^2 - 2$	2	$\sqrt{2}$, $-\sqrt{2}$
$2x^6 - 5x^4 + 3x^2 - 6$	6	???

Real roots of univariate quadratic polynomials $ax^2 + bx + c$ ($a, b, c \in \mathbb{Z}$):

Real roots of univariate quadratic polynomials

$$ax^2 + bx + c \ (a, b, c \in \mathbb{Z})$$
:

$$\begin{array}{ll} \frac{-b\pm\sqrt{b^2-4ac}}{2a} & \text{if } a\neq 0,\ b^2-4ac\geq 0 \\ x=-\frac{c}{b} & \text{if } a=0,\ b\neq 0 \\ \mathbb{R} & \text{if } a=0,\ b=0,\ c=0 \\ \text{none} & \text{else.} \end{array}$$

Real roots of univariate quadratic polynomials $ax^2 + bx + c$ ($a, b, c \in \mathbb{Z}$):

$$\begin{array}{ll} \frac{-b\pm\sqrt{b^2-4ac}}{2a} & \text{if } a\neq 0,\ b^2-4ac\geq 0\\ x=-\frac{c}{b} & \text{if } a=0,\ b\neq 0\\ \mathbb{R} & \text{if } a=0,\ b=0,\ c=0\\ \text{none} & \text{else}. \end{array}$$

Real roots of multivariate quadratic polynomials

$$p_a x^2 + p_b x + p_c \ (p_a, p_b, p_c \in \mathbb{Z}[\vec{y}])$$
:

Real roots of univariate quadratic polynomials

$$ax^2 + bx + c \ (a, b, c \in \mathbb{Z})$$
:

$$\frac{-b\pm\sqrt{b^2-4ac}}{2a} \quad \text{if } a \neq 0, \ b^2-4ac \geq 0$$

$$x = -\frac{c}{b} \quad \text{if } a = 0, \ b \neq 0$$

$$\mathbb{R} \quad \text{if } a = 0, \ b = 0, \ c = 0$$

$$\text{none} \quad \text{else.}$$

Real roots of multivariate quadratic polynomials

$$p_a x^2 + p_b x + p_c \ (p_a, p_b, p_c \in \mathbb{Z}[\vec{y}])$$
:

$$\begin{array}{ll} \frac{-p_b\pm\sqrt{p_b^2-4p_ap_c}}{2p_a} & \text{if } p_a\neq 0,\ p_b^2-4p_ap_c\geq 0\\ x=-\frac{p_c}{p_b} & \text{if } p_a=0,\ p_b\neq 0\\ \mathbb{R} & \text{if } p_a=0,\ p_b=0,\ p_c=0\\ \text{none} & \text{else.} \end{array}$$

Real roots of univariate quadratic polynomials

$$ax^2 + bx + c \ (a, b, c \in \mathbb{Z})$$
:

$$\frac{-b\pm\sqrt{b^2-4ac}}{2a} \quad \text{if } a \neq 0, \ b^2-4ac \geq 0$$

$$x = -\frac{c}{b} \quad \text{if } a = 0, \ b \neq 0$$

$$\mathbb{R} \quad \text{if } a = 0, \ b = 0, \ c = 0$$
none else.

Real roots of multivariate quadratic polynomials

$$p_{a}x^{2} + p_{b}x + p_{c} \ (p_{a}, p_{b}, p_{c} \in \mathbb{Z}[\vec{y}])$$
:
$$\frac{-p_{b}\pm\sqrt{p_{b}^{2}-4p_{a}p_{c}}}{2p_{a}} \quad \text{if } p_{a} \neq 0, \ p_{b}^{2}-4p_{a}p_{c} \geq 0$$

$$x = -\frac{p_{c}}{p_{b}} \qquad \text{if } p_{a} = 0, \ p_{b} \neq 0$$

$$\mathbb{R} \qquad \qquad \text{if } p_{a} = 0, \ p_{b} = 0, \ p_{c} = 0$$
 none else.

Problem: expressions not in QFNRA. Solution: virtual substitution.

CAD: Real root isolation

CAD: Real root isolation

- For polynomials of degree 5 or higher, no solution equations exist.
- Instead of computing the zeros, we will isolate them: for each zero we define an interval in which this single zero is included.
- This is the so-called interval representation of zeros: (p, I) with univariate polynomial p and real interval I, such that p has exactly one real root in I.
- We need to be able to compute with this representation, e.g., substitute such a real root for a variable in a univariate polynomial contstraint and check its validity.
- We will see later (for the cylindrical algebraic decomposition) how it works.