Satisfaction of Polynomial Constraints over Finite Domains

Stefania Monica

Dipartimento di Scienze Matematiche, Fisiche e Informatiche
Università degli Studi di Parma
stefania.monica@unipr.it

February 15\textsuperscript{th}, 2018
The work is motivated by practical considerations:

- *Polynomial constraints* are ubiquitous;
- Polynomial constraints often involve variables that take values from *finite domains*;
- Typical finite-domain constraint solvers are *incomplete*.

**Example 1**

The following is a CLP(FD) constraint:

\[
X \text{ in } -10..10, \ Y \text{ in } -10..10, \ X \times Y \ #>= \ 21.
\]

SWI-Prolog (version 6.6.6) rewrites it to:

\[
X \text{ in } -10..-1\//1..10, \ Y \text{ in } -10..-1\//1..10, \\
Z \text{ in } 21..100, \ X \times Y \ #= \ Z.
\]
The Constraint Language

The signature $\Sigma$ of the considered constraint language is

$$\Sigma = \langle \mathcal{V}, \mathcal{F}, \mathcal{P} \rangle$$

where

- $\mathcal{V}$ is a denumerable set of variable symbols;
- $\mathcal{F} = O \cup Z$ is the set of constant symbols and function symbols with $O = \{+, \ast\}$ and $Z = \{0, 1, -1, 2, -2, \ldots\}$;
- $\mathcal{P} = \{=, \neq, <, \leq, >, \geq\}$ is the finite set of constraint predicate symbols.

Signature $\Sigma$ can express equalities, inequalities, and disequalities among multi-variate polynomials with integer coefficients, where

- A *primitive constraint* is any atomic predicate built using the symbols from signature $\Sigma$;
- A *(non-primitive) constraint* is a conjunction of primitive constraints.
Polynomial Constraints in Canonical Form

Polynomial constraints with variables that take values from subsets of \( \mathbb{Z} \) can be written in a *canonical form*

\[
p(x) \geq 0,
\]

where \( x = (x_1, \ldots, x_n) \) are the \( n \in \mathbb{N} \) variables of polynomials.

**Proposition 1**

The following co-implications hold for every polynomial \( p(x) \) with integer coefficients and integer variables \( x \):

\[
\begin{align*}
p(x) \leq 0 & \iff -p(x) \geq 0 \quad (1) \\
p(x) > 0 & \iff p(x) \geq 1 \iff p(x) - 1 \geq 0 \quad (2) \\
p(x) < 0 & \iff p(x) \leq -1 \iff -p(x) - 1 \geq 0 \quad (3) \\
p(x) = 0 & \iff p^2(x) \leq 0 \iff -p^2(x) \geq 0 \quad (4) \\
p(x) \neq 0 & \iff p^2(x) > 0 \iff p^2(x) - 1 \geq 0. \quad (5)
\end{align*}
\]
All variables $\{x_i\}_{i=1}^n$ are initially restricted over *known intervals* 

$$x_i \in [\underline{x}_i^{(0)} .. \overline{x}_i^{(0)}] \quad i \in [1..n].$$

Starting from the initial box $B^{(0)} = [\underline{x}_1^{(0)} .. \overline{x}_1^{(0)}] \times \cdots \times [\underline{x}_n^{(0)} .. \overline{x}_n^{(0)}]$: 

- The algorithm tests if each constraint is consistent in the current box;  
- If no consistency can be certified, a variable is selected and the current box is split into two disjoint boxes; 
- The obtained disjoint boxes are processed recursively.

Currently, variables are selected using a fixed lexicographic order, and the generic interval $[\underline{x}_i .. \overline{x}_i]$ (with $\underline{x}_i < \overline{x}_i$) of the selected variable $x_i$ is split into two disjoint intervals $[\underline{x}_i .. s_i]$ and $[(s_i + 1) .. \overline{x}_i]$ where

$$s_i = \left\lfloor \frac{(\underline{x}_i + \overline{x}_i)}{2} \right\rfloor.$$
The consistency of single constraints is verified in a box as follows:

- For each primitive constraint $p(x) \geq 0$, a lower bound $l$ and an upper bound $u$ for $p(x)$ in the current box are computed;
- If $u < 0$ the constraint is not satisfiable in the current box;
- If $l \geq 0$ the constraint is consistent over the current box;
- Otherwise, the current box should be split and analysed recursively.

Only the signs of suitable lower and upper bounds is needed. Such signs can be computed using either:

- The *modified Bernstein form* of polynomials; or
- The values of polynomial functions for specific assignments of variables.
Given two multi-indices \( I \in \mathbb{N}^n \) and \( J \in \mathbb{N}^n \), with 
\( I = (i_1, i_2, \ldots, i_n) \) and \( J = (j_1, j_2, \ldots, j_n) \),

\[
|I| = \sum_{j=1}^{n} i_j \quad \text{(order of I)} \tag{6}
\]

\[
I! = \prod_{j=1}^{n} i_j! \quad \text{(factorial of I)} \tag{7}
\]

\[
I + J = (i_1 + j_1, i_2 + j_2, \ldots, i_n + j_n) \tag{8}
\]

\[
\sum_{l \leq J} (\cdot) = \sum_{i_1=0}^{j_1} \sum_{i_2=0}^{j_2} \cdots \sum_{i_n=0}^{j_n} (\cdot) \tag{9}
\]

\[
\sum_{|I| \leq k} (\cdot) = \sum_{i_1=0}^{k} \sum_{i_2=0}^{k} \cdots \sum_{i_n=0}^{k} (\cdot). \tag{10}
\]
The box $B$ built using $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ and $\mathbf{v} = (\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n) \in \mathbb{R}^n$ is denoted as

$$B = [\mathbf{v}, \mathbf{v}] = [v_1, \overline{v}_1] \times [v_2, \overline{v}_2] \times \cdots \times [v_n, \overline{v}_n],$$

and it is the empty set if and only if for some $0 \leq i \leq n$, $v_i > \overline{v}_i$.

We are now interested in studying the sign of a generic multivariate polynomials functions in $n \in \mathbb{N}$ variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$

$$p(\mathbf{x}) = \sum_{I \leq L} a_I \mathbf{x}^I,$$

where $L \in \mathbb{N}^n$ is its multi-degree, and $\{a_I\}_{I \leq L} \subset \mathbb{R}$ are its coefficients, over a known box $B = [\mathbf{v}, \mathbf{v}]$. 

Satisfaction of polynomial constraints over finite domains
We consider the affine change of variable between $\mathbf{x}$ and a new variable $\mathbf{t}$ defined over the unit box $[0, 1]^n$

$$x_i = x_i + (\bar{x}_i - x_i)t_i$$

This leads to

$$p(\mathbf{x}) = \sum_{l \leq L} a_l \mathbf{x}^l = \sum_{l \leq L} c_l \mathbf{t}^l$$

where

$$c_l = \sum_{J=1}^{L} a_l \binom{J}{l} \mathbf{x}^{J-l}(\bar{x} - \mathbf{x})^l \quad \binom{J}{l} = \prod_{k=0}^{n} \binom{j_k}{i_k}$$
The Bernstein basis is defined as

\[ B^L_I(x) = \prod_{j=0}^{n} B^L_{ij}(x_j) \]

with

\[ B^L_{ij}(x_j) = \binom{I}{j} \frac{(x_j - x_j)^{i} (x_j - x_j)^{j-i}}{(\bar{x}_j - x_j)^{l}} \]

The polynomial \( p(x) \) can then be written in Bernstein form as

\[ p(x) = \sum_{I \leq L} b_I B^L_I(x) \]

where

\[ b_I = \sum_{J \leq l} \binom{I}{J} c_J \quad I \leq L \]
We derive the modified Bernstein form

\[ p(x) = \sum_{I \leq L} \tilde{b}_I \tilde{B}_I^L(x), \]

where

\[ \tilde{b}_I = \sum_{J \leq I} \binom{L}{I} \binom{J}{I} c_J = \sum_{J \leq I} \binom{L - J}{I - J} c_J \quad I \leq L \]

and

\[ \tilde{B}_I^L(x) = \frac{B_I^L(x)}{\binom{L}{I}}. \]
Properties of the Modified Bernstein Form

It is worth noting that:

- \( \text{range}(p) \subseteq [\min_{I \leq L} \{b_I\}, \max_{I \leq L} \{b_I\}] \),
- \( \{b_I\}_{I \leq L} \) are not always integer even if \( \{a_I\}_{I \leq L} \) are integer.

We are interested in the signs of lower and upper bounds for a polynomial, which can be effectively computed using \( \{\tilde{b}_I\}_{I \leq L} \).

**Proposition 2**

The following properties hold:

- If \( \min_{I \leq L} \{b_I\} \odot 0 \) then for some multi-index \( J \), \( \tilde{b}_J \odot 0 \), with \( \odot \in \{=, \neq, <, \leq, \geq, >\} \);
- If \( \max_{I \leq L} \{b_I\} \odot 0 \) then for some multi-index \( J \), \( \tilde{b}_J \odot 0 \), with \( \odot \in \{=, \neq, <, \leq, \geq, >\} \);
- If \( \{a_I\}_{I \leq L} \) are integer, then \( \{\tilde{b}_I\}_{I \leq L} \) are integer.
Example: $x^2 - 9 \geq 0, \ x \in [-10..10]$
Let \( p : \mathbb{R} \rightarrow \mathbb{R} \) be a polynomial function, and \( B = [b, \overline{b}] \subset \mathbb{R} \) be a closed interval.

The following corollary of the Taylor’s theorem holds.

**Proposition 3**

For all \( x \in B \), and for all \( k \in \mathbb{N} \):

\[
p(x) = \sum_{i=0}^{k} \frac{D^{(i)} p(b)}{i!} (x - b)^i + R_{b,k}(x - b),
\]

where the remainder \( R_{b,k}(h) \) is given in Lagrange form by

\[
R_{b,k}(h) = \frac{D^{(k+1)} p(b + ch)}{(k + 1)!} h^{(k+1)}
\]

for some \( c \in (0, 1) \).
In particular, the following proposition holds (see also Rivlin, 1970) using a truncation for $k = 2$.

**Proposition 4**

Let $p : \mathbb{R} \rightarrow \mathbb{R}$ be a polynomial function of degree $l \in \mathbb{N}$ with coefficients $\{a_i\}_{i=0}^{l} \subset \mathbb{R}$, and let $U = [0, 1] \subset \mathbb{R}$. If

$$m = \min_{x \in U} p(x) \quad \text{and} \quad \overline{m} = \max_{x \in U} p(x),$$

then

$$\min_{x \in \{0,1\}} p(x) - \delta \leq m \leq \max_{x \in \{0,1\}} p(x) + \delta$$

where

$$\delta = \frac{1}{8} \sum_{i=2}^{l} (i - 1)i|a_i|.$$
The following corollary of the Taylor’s theorem holds.

**Proposition 5**

Let \( p : \mathbb{R}^n \rightarrow \mathbb{R} \) be a polynomial function in \( n \in \mathbb{N} \) variables, and \( B = [\underline{b}, \overline{b}] \subset \mathbb{R}^n \) be a box. Then, for all \( x \in B \), and for all \( k \in \mathbb{N} \)

\[
p(x) = \sum_{|I| \leq k} \frac{\partial^I p(b)}{I!} (x - b)^I + R_{b,k}(x - b)
\]

(18)

where \( \partial^I f = \partial^{i_1} \partial^{i_2} \cdots \partial^{i_n} f \) and the remainder \( R_{b,k}(h) \) is given in Lagrange form by

\[
R_{b,k}(h) = \sum_{|I| = k+1} \frac{\partial^I p(b + ch)}{I!} h^I
\]

(19)

for some \( c \in (0, 1) \).
Previous proposition with a truncation to $k = 2$ is sufficient to prove the following proposition.

**Proposition 6**

Let $U = [0, 1]^n \subset \mathbb{R}^n$, with

$$m = \min_{x \in U} p(x), \quad M = \max_{x \in U} p(x). \quad (20)$$

Then

$$\min_{x \in C} p(x) - \delta \leq m \leq \max_{x \in C} p(x) + \delta, \quad (21)$$

where $C = \{0, 1\}^n$ is the set of the corners of $U$ and

$$\delta = \frac{1}{8} \sum_{I \leq L} |I||I|(|I| - 1)|a_I|. \quad (22)$$
Example: $x^2 + y^2 + 100 \leq 0, \ x \in [1..25], \ y \in [1..25]$
Benchmarks

<table>
<thead>
<tr>
<th>D</th>
<th>PolyFD [ms]</th>
<th>Gecode [ms]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^1$</td>
<td>28</td>
<td>17</td>
</tr>
<tr>
<td>$10^2$</td>
<td>80</td>
<td>41</td>
</tr>
<tr>
<td>$10^3$</td>
<td>128</td>
<td>1022</td>
</tr>
<tr>
<td>$10^4$</td>
<td>191</td>
<td>10564</td>
</tr>
<tr>
<td>$10^5$</td>
<td>247</td>
<td>28121</td>
</tr>
</tbody>
</table>

**Table:** $3x^4 + 6x^2 + 5y^4 + 20y^2 + 33 \leq 0$, $x \in [-D..D]$, $y \in [-D..D]$

<table>
<thead>
<tr>
<th>D</th>
<th>PolyFD [ms]</th>
<th>Gecode [ms]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^1$</td>
<td>426</td>
<td>13</td>
</tr>
<tr>
<td>$10^2$</td>
<td>1430</td>
<td>7574</td>
</tr>
<tr>
<td>$10^3$</td>
<td>2886</td>
<td>&gt; 5 min</td>
</tr>
<tr>
<td>$10^4$</td>
<td>4363</td>
<td>&gt; 5 min</td>
</tr>
<tr>
<td>$10^5$</td>
<td>5877</td>
<td>&gt; 5 min</td>
</tr>
</tbody>
</table>

**Table:** $3x^4 + 6x^2 + 5y^4 + 10y^2 + 7z^4 + 14z^2 + 15 \leq 0$, $x \in [-D..D]$, $y \in [-D..D]$, $z \in [-D..D]$
Conclusions

Preliminary results on constraint solving for polynomial constraints over finite domains were presented:

- A canonical form of constraints;
- A subdivision algorithm that uses only the signs of polynomials functions over boxes;
- Algorithms to compute the signs of needed lower and upper bounds using only integer arithmetic.

Ongoing work targets an implementation to support further investigation on possible improvements:

- A drop-in replacement of the FD solver that ships with SWI-Prolog;
- A reusable library in Java, available online;
- A reusable library in C++, not yet available.


Thank you for your attention

Joint work with Federico Bergenti

Progetto GNCS Giovani Ricercatori 2016

stefania.monica@unipr.it