The interval analysis of multilinear expressions

C. Laneve, T.A. Lascu, V. Sordoni

Università di Bologna
TAPAS 2010

Interval analysis

Input:

- a polynomial arithmetic expression E
- the range of every variable $x_i \in E$

Output:

- the range of expression E

Context:

- *static analysis* of source code

Example instance

Example

Compute the range of:

$$E = x * (y - z) + z$$

knowing that:

$$\begin{cases} x \in X = [\,\underline{X},\ \overline{X}\,] = & [0, 1] \\ y \in Y = [\,\underline{Y},\ \overline{Y}\,] = & [0, 2000] \\ z \in Z = [\,\underline{Z},\ \overline{Z}\,] = & [0, 2000] \end{cases}$$

Historic technique

Standard Interval Arithmetics (SIA) - [Moore '66]

Example

$$
\begin{aligned}
X * (Y - Z) + Z &= [0,1] * ([0,2000] - [0,2000]) + [0,2000] = \\
&= [0,1] * [0 - 2000, 2000 - 0] + [0,2000] = \\
&= [0,1] * [-2000, 2000] + [0,2000] = \\
&= [-2000, 2000] + [0,2000] = \\
&= [-2000, 4000]
\end{aligned}
$$

Drawback: expressions with *multiple occurrences of the same variable*

Example

$$range_{SIA}(E) = [-2000, 4000] \supsetneq [0, 2000] = range(E)$$

Dependency Problem: different occurrences of the same variable are abstracted by independent intervals

Multilinear expressions

Expressions that are linear w.r.t. every variable.

Example

$$E = x * (y - z) + z$$

Meaningful class because all expressions in it share an interesting property ...

Property of multilinear expressions

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a multilinear function. If f has a local minimum or a local maximum then f is constant.

Corollary

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a multilinear function. The lower and upper bounds of f in the hypercube

$$H = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

occur at the vertices of H.

Vertex Evaluation Technique (VE)

VertexEvaluation:

1. find the hypercube's vertices
2. evaluate the expression in each of these points
3. keep the minimum and maximum values: $min(E)$ and $max(E)$
4. set $range_{VE}(E) := [min(E), max(E)]$

Computational complexity: $O(n \cdot 2^{2^n})$

Gaganov, *Computational complexity of the range of a polynomial in several variables*, ['85] shows that this problem is *NP-hard*.

Vertex Evaluation Technique (VE)

VertexEvaluation:

1. find the hypercube's vertices
2. evaluate the expression in each of these points
3. keep the minimum and maximum values: $min(E)$ and $max(E)$
4. set $range_{VE}(E) := [min(E), max(E)]$

Computational complexity: $O(n \cdot 2^{2^n})$

Gaganov, *Computational complexity of the range of a polynomial in several variables*, ['85] shows that this problem is *NP-hard*.

Generalization

Two steps:

1. reduce a generic expression to a multilinear one
2. apply VertexEvaluation

Example

$$E = x^5 - x^3 z + xy - xz + z$$

1^{st} step: u replaces x^3 ($u \in U = X^3$)

$$E' = ux^2 - uz + xy - xz + z$$

2^{nd} step: v replaces x^2 ($v \in V = X^2$)

$$E'' = uv - uz + xy - xz + z$$

N.B. reduction strategy is not unique!

Precision

Proposition

In the general case:

$$range(E) \subseteq range_{VE}(E)$$

because we loose some information on dependence.

Example

$$ux^2 - uz + xy - xz + z \quad \rightsquigarrow \quad uv - uz + xy - xz + z$$

in the right-hand side there is no dependence between x and v!

Is it good?

Computational complexity: $O(n(1 + d/2) \cdot 2^{2n(1+d/2)})$

Pros:

- multilinear case $\rightarrow$ exact range
- we pay an exponential cost at compile time

Cons:

- no guarantee that it returns a sharper range

Comparison with Miné's symbolic methods - 1

Two possible sources of imprecision:

- over-approximating the range of exponential terms as x^n
- poor handling of dependence

Tradeoff between these two aspects.

Example expression:

$$E = x^3 - x^2 y + y$$

Miné's method gives:

$$x^3 - x^2y + y \rightsquigarrow [a,b]^3 + (-[a,b]^2 + [1,1])y$$

our technique gives:

$$x^3 - x^2y + y \rightsquigarrow ux - uy + y \quad \text{with } u \in U = X^2$$

Instance	Miné	Our technique	Exact range
$x \in [-2,4]$	$[-23,64]$	$[-47,64]$	$[-11,64]$
$x \in [-2,2]$	$[-1,7]$	$[-1,7]$	$[1,7]$
$x \in [0,2]$	$[-3,9]$	$[-4,4]$	$[0,4]$

Comparison with Miné's symbolic methods - 2

Miné's method gives:

$$x^3 - x^2y + y \rightsquigarrow [a, b]^3 + (-[a, b]^2 + [1, 1])y$$

our technique gives:

$$x^3 - x^2y + y \rightsquigarrow ux - uy + y \quad \text{with } u \in U = X^2$$

Instance	Miné	Our technique	Exact range
$x \in [-2, 4]$	$[-23, 64]$	$[-47, 64]$	$[-11, 64]$
$x \in [-2, 2]$	$[-1, 7]$	$[-1, 7]$	$[1, 7]$
$x \in [0, 2]$	$[-3, 9]$	$[-4, 4]$	$[0, 4]$

Comparison with Miné's symbolic methods - 3

Conjecture

In the case of positive intervals our technique attains a sharper range.

. . . still to be proven . . . for the moment we can show that the following result holds:

Proposition

When all variables' ranges are positive:

$$range_{VE}(E) \subseteq range_{Miné}(E)$$

Conclusions

Contribution:

- an alternative technique:
  - multilinear case $\rightarrow$ exact range
  - general case $\rightarrow$ over-approximation
  - soundness proof
- tool:
  - to be used @ Magneti Marelli
  - static analysis of C code

For the future:

1. impact of different reduction strategies
2. estimate of the error introduced by the chosen strategy
3. classification of expressions w.r.t. precision

Conclusions

Contribution:

- an alternative technique:
  - multilinear case $\rightarrow$ exact range
  - general case $\rightarrow$ over-approximation
  - soundness proof
- tool:
  - to be used @ Magneti Marelli
  - static analysis of C code

For the future:

1. impact of different reduction strategies
2. estimate of the error introduced by the chosen strategy
3. classification of expressions w.r.t. precision