Identifying Minimal Changes in the Zone Abstract Domain

Kenny Ballou
Elena Sherman

Boise State University
Boise, Idaho
United States of America

July 2023

Outline

1. Background and Motivation
   - Zones Domain
   - Exploiting DFA Features
2. Algorithms and Approach
   - Spurious Connections
   - Connected Components
   - Node Neighbors
   - Minimal Neighbors
3. Experimental Results
   - Application
4. Conclusions

Static analysis computes invariants

Static analysis computes invariants

Unit difference, two-variables per inequality

$$x - Z_0 = 0$$
$$w - x \leq 2$$

Static analysis computes invariants

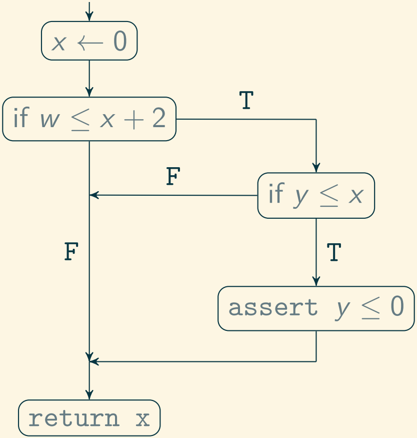
Inequalities as invariants for a simple program

```
int example(int w, int y) {
    int x = 0;
    if (w <= x + 2) {
        if (y <= x) {
            assert y <= 0;
        }
    }
    return x;
}
```

Static analysis computes invariants

Inequalities as invariants for a simple program

```
int example(int w, int y) {
    int x = 0;
    if (w <= x + 2) {
        if (y <= x) {
            assert y <= 0;
        }
    }
    return x;
}
```



$x \leftarrow 0$

if $w \leq x + 2$

T

F

if $y \leq x$

F

T

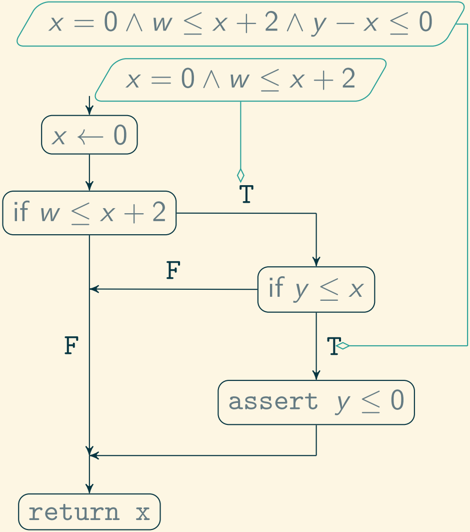
assert $y \leq 0$

return x

Static analysis computes invariants

Inequalities as invariants for a simple program

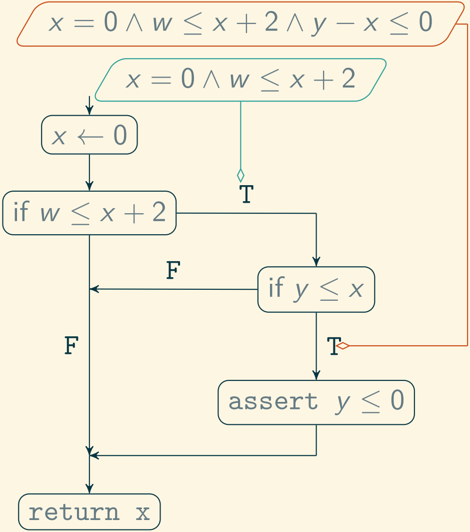
```
int example(int w, int y) {
    int x = 0;
    if (w <= x + 2) {
        if (y <= x) {
            assert y <= 0;
        }
    }
    return x;
}
```



Static analysis computes invariants

Inequalities as invariants for a simple program

```
int example(int w, int y) {
    int x = 0;
    if (w <= x + 2) {
        if (y <= x) {
            assert y <= 0;
        }
    }
    return x;
}
```



$x = 0 \land w \leq x + 2 \land y - x \leq 0$

$x = 0 \land w \leq x + 2$

$x \leftarrow 0$

if $w \leq x + 2$

T

F

if $y \leq x$

F

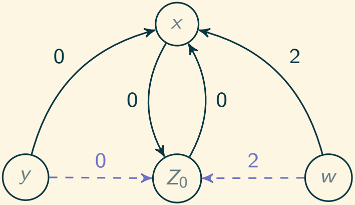
T

assert $y \leq 0$

return x

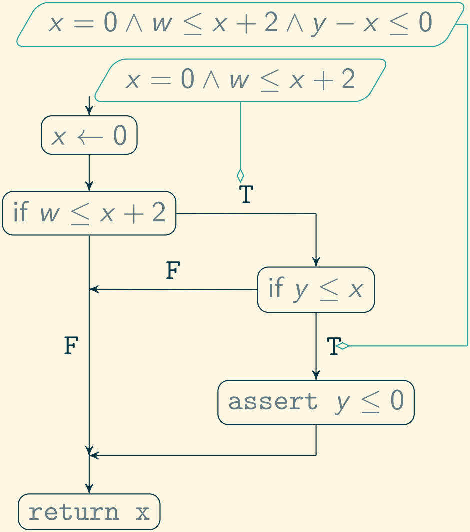
Zone Domain

$$\begin{array}{c} x - Z_0 \leq 0 \\ Z_0 - x \leq 0 \\ w - x \leq 2 \\ y - x \leq 0 \\ \hline y \leq 0 \\ w \leq 2 \end{array}$$

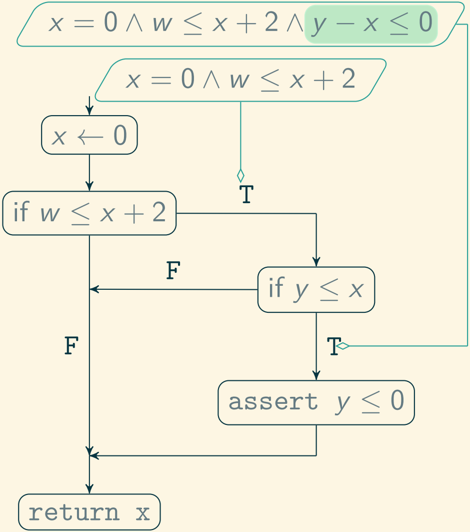


Zonal state representation of data-flow analysis invariant

Data-flow analysis incrementally updates variables



Data-flow analysis incrementally updates variables



Finding Affected Inequalities

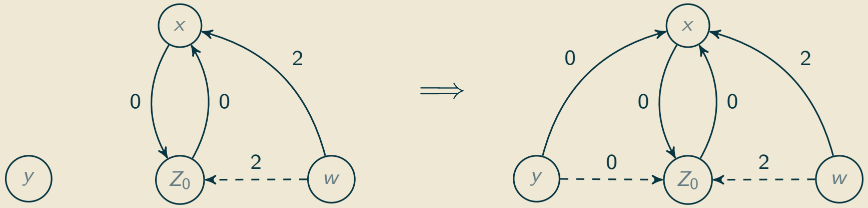
Problem Definition

$$\frac{\begin{array}{r} x = 0 \\ w - x \leq 2 \\ y - x \leq 0 \end{array}}{w \leq 2} \implies \frac{\begin{array}{r} x = 0 \\ w - x \leq 2 \\ y - x \leq 0 \end{array}}{\begin{array}{c} y \leq 0 \\ w \leq 2 \end{array}}$$

What are the changed set of inequalities?

Finding Affected Inequalities

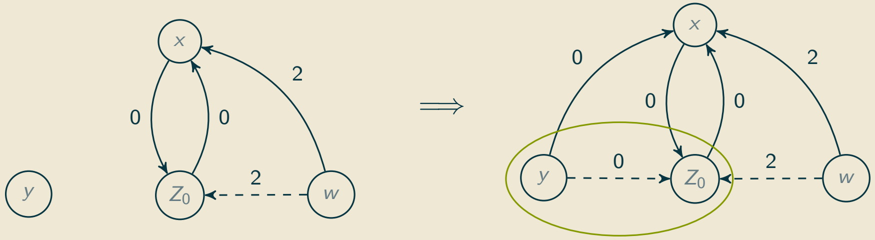
Problem Definition



What are the changed set of inequalities?

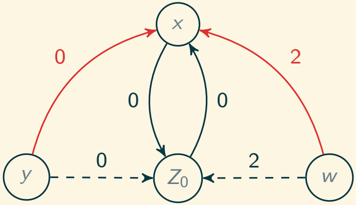
Finding Affected Inequalities

Problem Definition



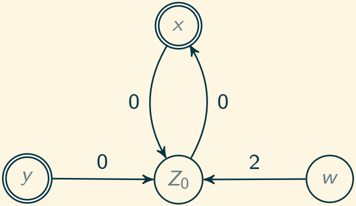
What are the changed set of inequalities?

Spurious Connected Variables[1]



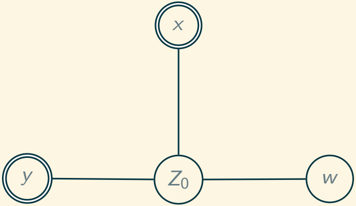
[1] Larsen et al., “Efficient Verification of Real-Time Systems: Compact Data Structure and State-Space Reduction”.

Connected Components



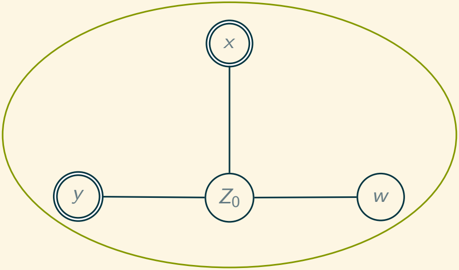
Connected Components

Variable Relation Projection



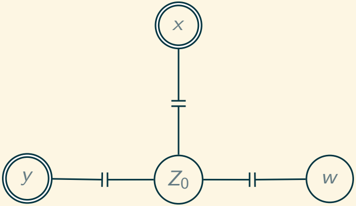
Connected Components

Variable Relation Projection



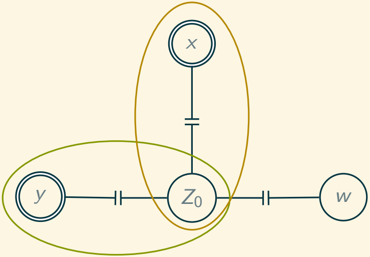
Connected Components

Variable Relation Projection with impassable Z_0



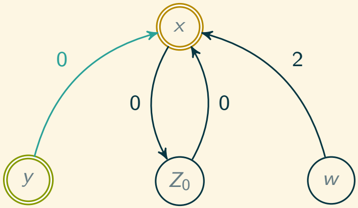
Connected Components

Variable Relation Projection with impassable Z_0



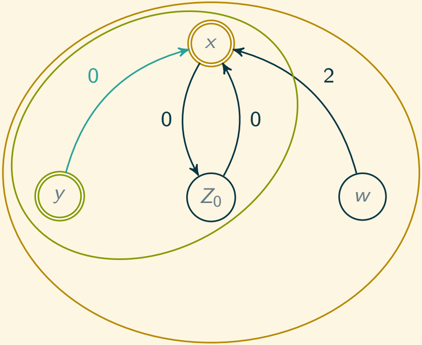
Node Neighbors

Reconsider the out-going state without closed edges



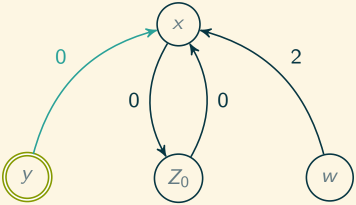
Node Neighbors

Reconsider the out-going state without closed edges



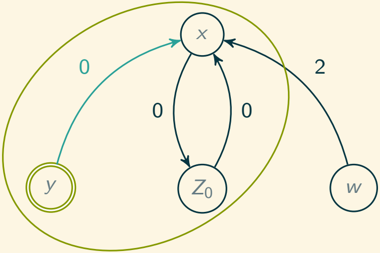
Minimal Neighbors

Again, reconsider the out-going state without closed edges.



Minimal Neighbors

Again, reconsider the out-going state without closed edges.



Logically comparing different abstract domains

Research Questions

RQ1 Do the minimization algorithms reduce the size of a Zone state and improve runtime of domain comparisons?

RQ2 Do the minimization algorithms affect categorization of domain comparison results?

Experiment Setup

- Benchmarks: 127 Java methods
  - Ranging from 4 to 412 Jimple instructions
- Compared Zones to Intervals and Zones to Predicates
- Compared Total Runtime of Z3 to perform logical entailment of every combination, averaging over 5 executions

Experimental results show significant reduction in required number of inequalities for comparison

Average percentage changes in V and E between each technique

State Type	vs.	$\downarrow \Delta$ % V	$\downarrow \Delta$ % E
DFA Subject Programs			
CC	FS	70.37	29.47
NN	CC	0.02	0.01
MN	NN	0.10	0.05
EQBench Subject Programs			
CC	FS	43.0	2.1
NN	CC	0.0	0.0
MN	NN	0.13	0.13

Experimental results show significantly reduced time to solver queries

State Type	$\sim$ Inter, sec.	$\sim$ Pred, sec.
DFA Subject Programs		
FS	4.03	265.91
CC	1.41	4.09
NN	1.41	4.04
MN	1.35	4.05
EQBench Subject Programs		
FS	0.79	5.56
CC	0.63	0.87
NN	0.58	0.9
MN	0.58	0.9

Experimental results show significant improvement in comparison granularity

State	$\succ$ Intervals	$=$ Intervals
DFA Subject Programs		
FS	2898	1002
CC	1194	2706
NN	1191	2709
MN	1164	2736
EQBench Subject Programs		
FS	374	255
CC	131	498
NN	131	498
MN	131	498

Experimental results show significant improvement in comparison granularity

State	$\succ$ **Predicates**	$=$ **Predicates**	$\prec$ **Predicates**	$\prec\succ$ **Predicates**
DFA Subject Programs				
FS	1464	237	167	2032
CC	1324	1930	473	173
NN	1322	1933	473	172
MN	1305	1960	473	162
EQBench Subject Programs				
FS	307	135	46	141
CC	217	322	72	18
NNy	217	322	72	18
MN	217	322	72	18

Conclusion

Experimental Results

- Minimization leads to reduced overall execution time when determining domain categorization.
- Minimization leads to improved granularity when evaluating domain precision.

Conclusion

Experimental Results

- Minimization leads to reduced overall execution time when determining domain categorization.
- Minimization leads to improved granularity when evaluating domain precision.

Algorithms and Approaches

- Spurious Connections → Reduce variable clustering
- Connected Components → Extract subsets using relational projection
- Node Neighbors → Extract subsets based on reachable neighborhoods
- Minimal Neighbors → Extract subsets leveraging semantic information

Future Work

- Extend to other Weakly-Relational Domains, e.g., Octagons
- Extend for comparison between relational domains

Thank you

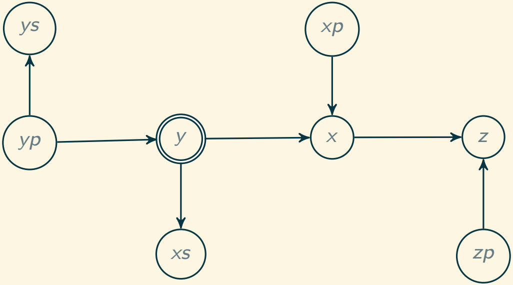
Questions?

The work reported here was supported by the U.S. National Science Foundation under award CCF-19-42044.

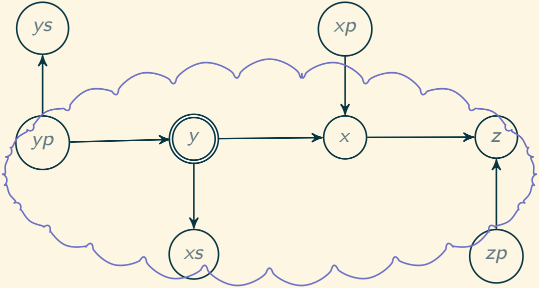
References I

[1] K.G. Larsen et al. “Efficient Verification of Real-Time Systems: Compact Data Structure and State-Space Reduction”. In: *Proceedings Real-Time Systems Symposium*. IEEE Comput. Soc, 1997, pp. 14–24. ISBN: 081868268X. DOI: 10.1109/real.1997.641265.

Extended Examples of the Minimal Neighbors Algorithm



Extended Examples of the Minimal Neighbors Algorithm


ys
xp
yp
y
x
z
xs
zp