

SAT and SMT Solving

Sarah Winkler

KRDB
Department of Computer Science
Free University of Bozen-Bolzano

lecture 10
WS 2022

Satisfiability in Linear Integer Arithmetic

Fact

for every LIA problem can compute bounds to get equisatisfiable bounded problem,
so BranchAndBound terminates

Definition (Cut)

given solution α to problem over \mathbb{Q}^n , cut is inequality $a_1x_1 + \dots + a_nx_n \leq b$
which is not satisfied by α but by every \mathbb{Z}^n -solution

Outline

- Summary of Last Week
- Bit Vectors

1

Gomory Cuts

Gomory Cuts: Assumptions

- ▶ DPLL(T) Simplex returned solution α and final tableau A such that

$$A\bar{x}_I = \bar{x}_D \quad l_i \leq x_i \leq u_i$$

- ▶ for some $x_i \in D$ have $\alpha(x_i) \notin \mathbb{Z}$ and for all $x_j \in I$ value $\alpha(x_j)$ is l_j or u_j

Notation

- ▶ write $c = \alpha(x_i) - \lfloor \alpha(x_i) \rfloor$
- ▶ split independent variables I into $L = \{x_j \mid \alpha(x_j) = l_j\}$ and $U = \{x_j \mid \alpha(x_j) = u_j\}$
- ▶ $L^+ = \{x_j \in L \mid \alpha(x_j) = l_j \text{ and } A_{ij} \geq 0\}$ $U^+ = \{x_j \in U \mid \alpha(x_j) = u_j \text{ and } A_{ij} \geq 0\}$
 $L^- = \{x_j \in L \mid \alpha(x_j) = l_j \text{ and } A_{ij} < 0\}$ $U^- = \{x_j \in U \mid \alpha(x_j) = u_j \text{ and } A_{ij} < 0\}$

Lemma (Gomory Cut)

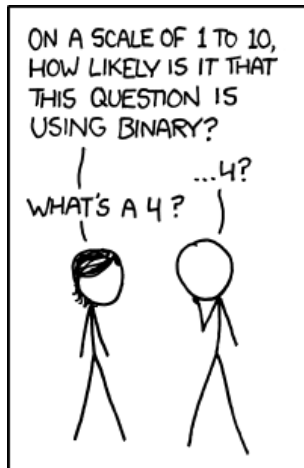
$$\sum_{x_j \in L^+} \frac{A_{ij}}{1-c} (x_j - l_j) - \sum_{x_j \in U^-} \frac{A_{ij}}{1-c} (u_j - x_j) - \sum_{x_j \in L^-} \frac{A_{ij}}{c} (x_j - l_j) + \sum_{x_j \in U^+} \frac{A_{ij}}{c} (u_j - x_j) \geq 1$$

2

3

Outline

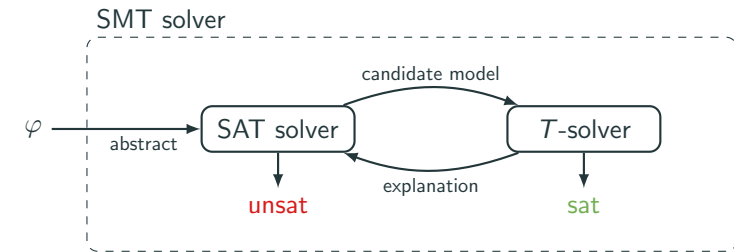
- Summary of Last Week
- Bit Vectors



Disclaimer

rest of lecture assumes brains in binary mode

Theories in SMT Solving



Theory T

- ▶ equality logic
- ▶ equality + uninterpreted functions (EUF)
- ▶ linear real arithmetic (LRA)
- ▶ linear integer arithmetic (LIA)
- ▶ bitvectors (BV)
- ▶ arrays (A)

T -solving method

- equality graphs ✓
- congruence closure ✓
- Simplex ✓
- Simplex + cuts + bounds ✓
- bit-blasting

4

5

Flashback to the Binary World

Binary representation

- ▶ k -bit representation of non-negative number n

$$\boxed{n \bmod 2^{k-1}} \boxed{n \bmod 2^{k-2}} \dots \boxed{n \bmod 2^1} \boxed{n \bmod 2^0}$$

denoted n_k

- ▶ ... of negative number $-n$ is $(\sim n_k) + 1_k$

2-complement

Operations on binary numbers

(for fixed bitwidth)

- ▶ $\&$, $|$, and \sim are bitwise and, or, and negation
- ▶ $+$, $-$, \times are addition, subtraction, and multiplication (with overflow)

Example

▶ 5_4	<table><tr><td>0</td><td>1</td><td>0</td><td>1</td></tr></table>	0	1	0	1	▶ 13_4	<table><tr><td>1</td><td>1</td><td>0</td><td>1</td></tr></table>	1	1	0	1						
0	1	0	1														
1	1	0	1														
▶ $5_4 \& 13_4$	<table><tr><td>0</td><td>1</td><td>0</td><td>1</td></tr></table>	0	1	0	1	▶ $5_4 \mid 13_4$	<table><tr><td>1</td><td>1</td><td>0</td><td>1</td></tr></table>	1	1	0	1	▶ $\sim 5_4$	<table><tr><td>1</td><td>0</td><td>1</td><td>0</td></tr></table>	1	0	1	0
0	1	0	1														
1	1	0	1														
1	0	1	0														
▶ $5_4 + 1_4$	<table><tr><td>0</td><td>1</td><td>1</td><td>0</td></tr></table>	0	1	1	0	▶ $5_4 + 13_4$	<table><tr><td>0</td><td>0</td><td>1</td><td>0</td></tr></table>	0	0	1	0						
0	1	1	0														
0	0	1	0														
▶ -5_4	<table><tr><td>1</td><td>0</td><td>1</td><td>1</td></tr></table>	1	0	1	1	▶ -1_4	<table><tr><td>1</td><td>1</td><td>1</td><td>1</td></tr></table>	1	1	1	1						
1	0	1	1														
1	1	1	1														

6

7

Comparison operators

- ▶ $<_u, \leq_u, \dots$ considers operands as **unsigned** numbers
- ▶ $<_s, \leq_s, \dots$ considers operands as **signed** numbers

Example

- ▶ 3_4

0	0	1	1
---	---	---	---

 $<_u$

1	1	0	1
---	---	---	---

 13_4
- ▶ 3_4

0	0	1	1
---	---	---	---

 $\not<_s$

1	1	0	1
---	---	---	---

 -3_4

Binary operations and sign

- ▶ $+$, $-$, \times work independently of whether operands are considered signed
- ▶ division and modulo depend on signedness: distinguish \div_u , $\%_u$ and \div_s , $\%_s$

Example

$$\begin{array}{rcl}
 3_4 & \begin{array}{|c|c|c|c|} \hline 0 & 0 & 1 & 1 \\ \hline \end{array} & 3_4 \\
 10_4 + & \begin{array}{|c|c|c|c|} \hline 1 & 0 & 1 & 0 \\ \hline \end{array} & -6_4 \\
 \hline
 13_4 & \begin{array}{|c|c|c|c|} \hline 1 & 1 & 0 & 1 \\ \hline \end{array} & -3_4
 \end{array}
 \quad
 \begin{array}{rcl}
 & \begin{array}{|c|c|c|c|} \hline 1 & 0 & 1 & 0 \\ \hline \end{array} & 10_4 \\
 \div_u & \begin{array}{|c|c|c|c|} \hline 0 & 0 & 1 & 1 \\ \hline \end{array} & 3_4 \\
 \hline
 & \begin{array}{|c|c|c|c|} \hline 0 & 0 & 1 & 1 \\ \hline \end{array} & 3_4
 \end{array}
 \quad
 \begin{array}{rcl}
 & \begin{array}{|c|c|c|c|} \hline 1 & 0 & 1 & 0 \\ \hline \end{array} & -6_4 \\
 \div_s & \begin{array}{|c|c|c|c|} \hline 0 & 0 & 1 & 1 \\ \hline \end{array} & 3_4 \\
 \hline
 & \begin{array}{|c|c|c|c|} \hline 1 & 1 & 1 & 0 \\ \hline \end{array} & -2_4
 \end{array}$$

8

Examples

- ▶ $x_4 + y_4 = 7_4$
satisfiable: $v(x_4) = 4_4$ and $v(y_4) = 3_4$
- ▶ $x_4 + 2_4 <_u x_4$ overflow semantics!
satisfiable: $v(x_4) = 15_4$
- ▶ $(x_4 \times y_4 = 6_4) \wedge (x_4 \& y_4 = 2_4)$
satisfiable: $v(x_4) = 3_4$, $v(y_4) = 2_4$
- ▶ $(x_4 \geq_u y_4) \wedge \neg(x_4 \geq_s y_4)$
satisfiable: $v(x_4) = 8_4$, $v(y_4) = 0_4$
- ▶ $(x_4 \ll 2_4 = 12_4) \wedge (x_4 + 1_4 = 12_4)$
satisfiable: $v(x_4) = 11_4$
- ▶ $(8_4 \gg_u 2_4 = 2_4) \wedge (8_4 \gg_s 2_4 = 14_4)$
holds
- ▶ $(x_4[1:0] :: x_4[3:2] = 2_4) \wedge (y_4[2:0] = 7_3)$
satisfiable: $v(x_4) = 8_4$ and $v(y_4) = 15_4$

\gg_u shifts in 0s,
 \gg_s shifts in sign bits

$x[i:j]$ denotes $x_i \dots x_j$
and $::$ is concatenation

10

Definition (Bit Vector Theory)

for bitwidth k , theory BV_k is given by

- ▶ signature
 - ▶ constants n_k for all $n < 2^k$
 - ▶ binary function symbols $+$, $-$, \times , \div_u , \div_s , $\%_u$, $\%_s$, \ll , \gg_u , \gg_s , $\&$, $|$, \wedge
 - ▶ unary function symbols $-$ and \sim
 - ▶ predicates $=$, \neq , \geq_u , \geq_s , $>_u$, and $>_s$
- ▶ **axioms** are equality axioms plus all correct arithmetic, comparison, and bit operations on binary numbers with k bits

Remark

- ▶ theories $BV_{k_1}, \dots, BV_{k_m}$ of different bit widths can be combined
- ▶ can also use binary $::$ for concatenation and unary $(\cdot)[i:j]$ to extract bits

Definitions

- ▶ **variable** x_k is list of length k of propositional variables $x_{k-1} \dots x_2 x_1 x_0$
- ▶ **valuation** v assigns element in $\{T, F\}^k$ to variable x_k ,
(usually written as binary number with k bits)

9

Notation for Constants

- ▶ n_k is binary representation of n in k bits
- ▶ xn_k is binary representation of hexadecimal n in k bits

Example

- ▶ $0_1, 3_2, 10_4, 1024_{32}, \dots$
- ▶ $x0_4, xa_4, xb0_8, x11cf_{16}, xffffff_{32}, \dots$

More examples

- ▶ $-a_4 = a_4$
satisfiable: $v(a_4) = -8_4 = x8_4$
- ▶ $a_8 \div_u b_8 = a_8 \gg_u 1_8$
satisfiable: $v(a_8) = 8_8$ and $v(b_8) = 2_8$
- ▶ $a_8 \& (a_8 - 1_8) = 0_8$
satisfiable: $v(a_8) = 8_8$ or $x0_8, x1_8, x2_8, x4_8, x8_8, x10_8, x20_8, x40_8, x80_8$

negation uses two's complement

satisfied by powers of 2 (and 0)

11

Remarks

- ▶ theory is decidable because carrier is finite
- ▶ common decision procedures use translation to SAT (**bit blasting**)
 - ▶ eager: no DPLL(T), bit-blast entire formula to SAT problem
 - ▶ lazy: second SAT solver as BV theory solver, bit-blast only BV atoms
- ▶ solvers heavily rely on **preprocessing via rewriting**

Example (Preprocessing)

$$\mathbf{x}_1 \neq \mathbf{0}_1 \wedge (\mathbf{y}_3 :: \mathbf{x}_1) \%_u \mathbf{2}_4 = \mathbf{0}_4 \rightarrow \mathbf{x}_1 = \mathbf{1}_1 \wedge (\mathbf{y}_3 :: \mathbf{x}_1) \%_u \mathbf{2}_4 = \mathbf{0}_4 \\ \rightarrow (\mathbf{y}_3 :: \mathbf{1}_1) \%_u \mathbf{2}_4 = \mathbf{0}_4 \rightarrow \mathbf{F}$$

Definition (Bit Blasting: Formulas)

bit blasting transformation \mathbf{B} transforms BV formula into propositional formula:

$$\mathbf{B}(\varphi \vee \psi) = \mathbf{B}(\varphi) \vee \mathbf{B}(\psi)$$

$$\mathbf{B}(\varphi \wedge \psi) = \mathbf{B}(\varphi) \wedge \mathbf{B}(\psi)$$

$$\mathbf{B}(\neg\varphi) = \neg\mathbf{B}(\varphi)$$

$$\mathbf{B}(t_1 \text{ rel } t_2) = \mathbf{B}_r(u_1 \text{ rel } u_2) \wedge \varphi_1 \wedge \varphi_2 \quad \text{if } \mathbf{B}_t(t_1) = (u_1, \varphi_1) \text{ and } \mathbf{B}_t(t_2) = (u_2, \varphi_2)$$

bit blasting \mathbf{B}_t for term t
returns (result u , side condition φ)

\mathbf{B}_r transforms atom into propositional formula

12

Definition (Bit Blasting: Bitwise Operations)

for bit vectors \mathbf{x}_k and \mathbf{y}_k use fresh variable \mathbf{z}_k and set

- ▶ **bitwise and**

$$\mathbf{B}_t(\mathbf{x}_k \& \mathbf{y}_k) = (\mathbf{z}_k, \varphi) \quad \varphi = \bigwedge_{i=0}^{k-1} z_i \leftrightarrow (x_i \wedge y_i)$$

- ▶ **bitwise or**

$$\mathbf{B}_t(\mathbf{x}_k | \mathbf{y}_k) = (\mathbf{z}_k, \varphi) \quad \varphi = \bigwedge_{i=0}^{k-1} z_i \leftrightarrow (x_i \vee y_i)$$

- ▶ **bitwise exclusive or**

$$\mathbf{B}_t(\mathbf{x}_k \wedge \mathbf{y}_k) = (\mathbf{z}_k, \varphi) \quad \varphi = \bigwedge_{i=0}^{k-1} z_i \leftrightarrow (x_i \oplus y_i)$$

- ▶ **bitwise negation**

$$\mathbf{B}_t(-\mathbf{x}_k) = (\mathbf{z}_k, \varphi) \quad \varphi = \bigwedge_{i=0}^{k-1} z_i \leftrightarrow \neg x_i$$

14

Definition (Bit Blasting: Atoms)

for bit vectors \mathbf{x}_k and \mathbf{y}_k set

- ▶ **equality**

$$\mathbf{B}_r(\mathbf{x}_{k+1} = \mathbf{y}_{k+1}) = (x_k \leftrightarrow y_k) \wedge \dots \wedge (x_1 \leftrightarrow y_1) \wedge (x_0 \leftrightarrow y_0)$$

- ▶ **inequality**

$$\mathbf{B}_r(\mathbf{x}_{k+1} \neq \mathbf{y}_{k+1}) = (x_k \oplus y_k) \vee \dots \vee (x_1 \oplus y_1) \vee (x_0 \oplus y_0)$$

- ▶ **unsigned greater-than or equal**

$$\mathbf{B}_r(\mathbf{x}_1 \geq_u \mathbf{y}_1) = y_0 \rightarrow x_0$$

$$\mathbf{B}_r(\mathbf{x}_{k+1} \geq_u \mathbf{y}_{k+1}) = (x_k \wedge \neg y_k) \vee ((x_k \leftrightarrow y_k) \wedge \mathbf{B}(\mathbf{x}[k-1:0] \geq \mathbf{y}[k-1:0]))$$

- ▶ **unsigned greater-than**

$$\mathbf{B}(\mathbf{x}_k >_u \mathbf{y}_k) = \mathbf{B}(\mathbf{x}_k \geq \mathbf{y}_k) \wedge \mathbf{B}(\mathbf{x}_k \neq \mathbf{y}_k)$$

13

Definition (Bit Blasting: Concatenation, Extraction, If)

- ▶ **concatenation**

$$\mathbf{B}_t(\mathbf{x}_k :: \mathbf{y}_m) = (\mathbf{x}_k \mathbf{y}_m, \top)$$

for bit vectors \mathbf{x}_k and \mathbf{y}_m

- ▶ **extraction**

$$\mathbf{B}_t(\mathbf{x}[n:m]) = (\mathbf{z}_{n-m+1}, \varphi) \quad \varphi = \bigwedge_{i=0}^{n-m} z_i \leftrightarrow x_{i+m}$$

for bit vector \mathbf{x}_k , $k > n \geq m \geq 0$ and fresh variable \mathbf{z}_{n-m+1}

- ▶ **if-then-else**

$$\mathbf{B}_t(p ? \mathbf{x}_k : \mathbf{y}_k) = (\mathbf{z}_k, \varphi) \quad \varphi = \bigwedge_{i=0}^{k-1} (p \rightarrow (z_i \leftrightarrow x_i)) \wedge (\neg p \rightarrow (z_i \leftrightarrow y_i))$$

for formula p and bit vectors \mathbf{x}_k and \mathbf{y}_k

15

Definition (Bit Blasting: Addition and Subtraction)

► addition

$$\mathbf{B}_t(\mathbf{x}_k + \mathbf{y}_k) = (\mathbf{s}_k, \varphi)$$

where

$$\varphi = (c_0 \leftrightarrow x_0 \wedge y_0) \wedge (s_0 \leftrightarrow x_0 \oplus y_0) \wedge \bigwedge_{i=1}^{k-1} (c_i \leftrightarrow \min2(x_i, y_i, c_{i-1})) \wedge (s_i \leftrightarrow x_i \oplus y_i \oplus c_{i-1})$$

ripple-carry adder:
 c_k are carry bits

for fresh variables \mathbf{s}_k and \mathbf{c}_k and $\min2(a, b, d) = (a \wedge b) \vee (a \wedge d) \vee (b \wedge d)$

► unary minus

$$\mathbf{B}_t(-\mathbf{x}_k) = \mathbf{B}_t(\sim \mathbf{x}_k + \mathbf{1}_k)$$

► subtraction

$$\mathbf{B}_t(\mathbf{x}_k - \mathbf{y}_k) = \mathbf{B}_t(\mathbf{x}_k + (-\mathbf{y}_k))$$

16

Example (SMT-LIB 2 for BV)

$(\mathbf{a}_4 + \mathbf{b}_4 <_u \mathbf{b}_4) \wedge (\mathbf{a}_4 \neq \mathbf{10}_4) \wedge (\mathbf{a}_4 \& \mathbf{b}_4 = \mathbf{8}_4)$ is expressed as

```
(declare-const a (_ BitVec 4))
(declare-const b (_ BitVec 4))
(assert (bvult (bvadd a b) b))
(assert (not (= a #xa)))
(assert (= (bvand a b) #b1000))
(check-sat)
```



Bit vectors in SMT-LIB 2

- `(_ BitVec k)` is sort of bitvectors of length k
- `#xa` is constant in hexadecimal
- `#b1000` is constant in binary
- `bvadd`, `bvsub`, `bvmul` are arithmetic operations, `bvudiv` and `bvsdiv` are unsigned and signed division
- `bvult` and `bvule` are unsigned, `bvslt` and `bvsle` are signed $<$ and \leq
- `bvshl`, `bvlshr`, `bvashr` are shifts
- `bvand`, `bvor` are bitwise logical operations

18

Definition (Bit Blasting: Multiplication and Division)

for bit vectors \mathbf{x}_k and \mathbf{y}_k set

► multiplication

$$\mathbf{B}_t(\mathbf{x}_k \times \mathbf{y}_k) = \mathbf{B}_t(\text{mul}(\mathbf{x}_k, \mathbf{y}_k, 0))$$

where `mul` is defined by recursion on last argument:

$$\text{mul}(\mathbf{x}_k, \mathbf{y}_k, k) = \mathbf{0}_k$$

$$\text{mul}(\mathbf{x}_k, \mathbf{y}_k, i) = \text{mul}(\mathbf{x}_k \ll \mathbf{1}_k, \mathbf{y}_k, i+1) + (y_i ? \mathbf{x}_k : \mathbf{0}_k) \quad \text{if } i < k$$

shift-and-add

► unsigned division

$$\mathbf{B}_t(\mathbf{x}_k \div_u \mathbf{y}_k) = (\mathbf{q}_k, \varphi)$$

$$\varphi = \mathbf{B}(\mathbf{y}_k \neq \mathbf{0}_k \rightarrow (\mathbf{q}_k \times \mathbf{y}_k + \mathbf{r}_k = \mathbf{x}_k \wedge \mathbf{r}_k < \mathbf{y}_k \wedge \mathbf{q}_k < \mathbf{x}_k))$$

for fresh variables \mathbf{q}_k and \mathbf{r}_k

17

Bit Vectors in python/z3

```
from z3 import *
x = BitVec("x", 8)
y = BitVec("y", 8)
zero = BitVecVal(0, 8)
one = BitVecVal(1, 8)
r = y ^ ((x ^ y) & (zero -(If(x < y, one, zero))))
m = If(x < y, x, y)
solve(r != m) # shorthand for checking single formula
```

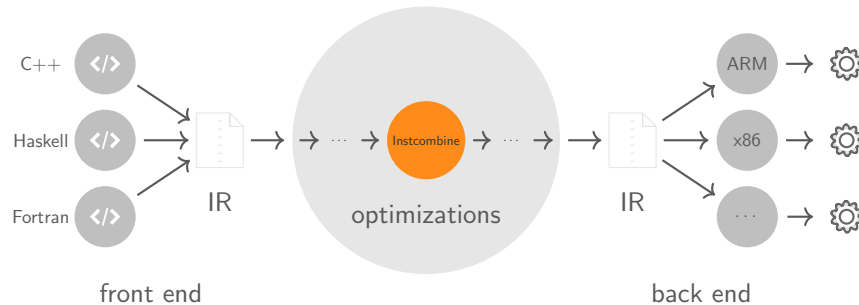
- `BitVec(name, k)` creates variable with k bits
- `BitVecVal(s, k)` is constant c in k bits
- `+`, `-`, `*` are arithmetic operations
- `&`, `|`, `~`, `^` are bitwise operations
- comparisons `<`, `<=`, `>`, `>=` are signed, use `ULT`, `ULE`, `UGT`, `UGE` for unsigned
- `<<` is left shift, `>>` is `>>_s`, `LShR` is `>>_u`
- division `/` and modulo `%` is signed, use `UDiv` and `URem` for unsigned
- for valuations, solver returns integers by default

19

Application 1: Verifying Compiler Optimizations

LLVM

- ▶ open-source umbrella project: set of reusable toolchain components: libraries, assemblers, compilers, debuggers, ...
- ▶ compilation toolchain includes peephole optimizations in **Instcombine** pass



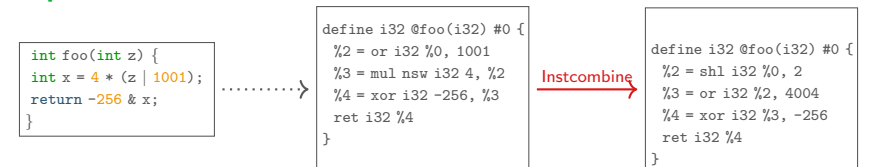
20

Application 1: Verifying Compiler Optimizations

Instcombine Pass

- ▶ over 1000 **algebraic simplifications of expressions**
 - ▶ transform multiplies with constant power-of-two argument into shifts
 - ▶ bitwise operators with constant operands are always grouped so that shifts are performed first, then ors, then ands, then xors
 - ▶ changing bitwidth of variables
 - ▶ ...
- ▶ code is community maintained
- ▶ sometimes optimizations have **errors**—and compiler bugs are critical

Example



21

Application 1: Verifying Compiler Optimizations

Alive Project

- ▶ represent Instcombine optimizations in domain-specific language, e.g.

```
Name: PR20186
%a = sdiv %X, C
%r = sub 0, %a
=>
%r = sdiv %X, -C
```

- ▶ check correctness by means of SMT encoding

```
(declare-const x (_ BitVec 32))
(declare-const c (_ BitVec 32))
(declare-const before (_ BitVec 32))
(declare-const after (_ BitVec 32))
(assert (= before (bvsdiv #x00000000 (bvsdiv x c))))
(assert (= after (bvsdiv x (bvneg c))))
(assert (not (= before after)))
(assert (not (= c #x00000000)))
(check-sat)
```



- ▶ **wrong** for `c = x = #x80000000`

22

Same in python/z3

```
from z3 import *
x = BitVec('x', 32) # create variable named x with 32 bits
c = BitVec('c', 32)
```

```
before = BitVecVal(0, 32) - (x / c)
after = x / - c
```

```
solver = Solver()
solver.add(c != BitVecVal(0, 32)) # exclude case where c=0
solver.add(after != before)
```

```
result = solver.check()
if result == z3.sat:
    m = solver.model()
    print m[x], m[c] # 2147483648 2147483648
    print m.eval(before), m.eval(after) # 4294967295 1
```

23

Application 2: Detecting Nontermination in Programs

```
int bsearch(int a[], int k, unsigned int lo, unsigned int hi) {
    unsigned int mid;
    while (lo < hi) {
        mid = (lo + hi)/2;
        if (a[mid] < k)
            lo = mid + 1;
        else if (a[mid] > k)
            hi = mid - 1;
        else
            return mid;
    }
    return -1;
}
```

- ▶ (former) implementation of binary search in Java library
- ▶ **loops** for inputs `lo=1` and `hi=UINT_MAX` if `a[0] < k`.
- ▶ SMT encoding can find values such that parameters stay the same in recursive call

24

Bibliography



Daniel Kroening and Ofer Strichman

Bit Vectors

Chapter 6 of Decision Procedures — An Algorithmic Point of View
Springer, 2008



Nuno Lopes, David Menendez, Sarantosh Nagarakatte, and John Regehr.

Provably Correct Peephole Optimizations with Alive.

Proc. 36th PLDI, pp. 22–32, 2013.

25