

SAT and SMT Solving

Sarah Winkler

SS 2018

Department of Computer Science
University of Innsbruck

Outline

- Summary of Last Week
- Bit Vectors

1

Summary of Last Week

Input to Satisfiability Problem for Equality Logic

conjunction φ of equality logic literals over set of variables V

Definitions

- ▶ $\varphi_ =$ is set of positive literals (equality literals) in φ
- ▶ φ_{\neq} is set of negative literals (inequality literals) in φ
- ▶ equality graph is undirected graph $G_=(\varphi) = (V, \varphi_ =, \varphi_{\neq})$

Definitions

equality graph $G_=(\varphi) = (V, \varphi_ =, \varphi_{\neq})$

- ▶ contradictory cycle is cycle with exactly one φ_{\neq} edge
- ▶ contradictory cycle is simple if it contains no node twice

Lemma

φ is satisfiable iff $G_=(\varphi)$ contains no simple contradictory cycles

2

Idea (Branch and Bound)

- ▶ given \mathbb{R}^2 solution α , add constraints to exclude α but preserve \mathbb{Z}^2 solutions:
if $a < \alpha(x) < a_1$, use Simplex on problems $C \wedge x \leq a$ and $C \wedge x \geq a + 1$
- ▶ might not terminate if solution space is unbounded

Algorithm BranchAndBound(φ)

Input: LIA constraint φ
Output: unsatisfiable, or satisfying assignment

let res be result of deciding φ over \mathbb{R} ▷ e.g. by Simplex
if res is unsatisfiable **then**
 return unsatisfiable
else if res is solution over \mathbb{Z} **then**
 return res
else
 let x be variable assigned non-integer value q in res
 $res = \text{BranchAndBound}(\varphi \wedge x \leq \lfloor q \rfloor)$
 return $res \neq \text{unsatisfiable} ? res : \text{BranchAndBound}(\varphi \wedge x \geq \lceil q \rceil)$

3

Definition (Cut)

given solution α to problem over \mathbb{R}^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

Gomory Cuts: Assumptions

- ▶ DPLL(T) Simplex returned solution α to

$$A\vec{x}_N = \vec{x}_B \tag{1}$$

$$-\infty \leq l_i \leq x_i \leq u_i \leq +\infty \tag{2}$$
- ▶ for some $i \in B$ have $\alpha(x_i) \notin \mathbb{Z}$ and for all $j \in N$ value $\alpha(x_j)$ is l_j or u_j

Notation

- ▶ write $c = \alpha(x_i) - \lfloor \alpha(x_i) \rfloor$
- ▶ by assumption all nonbasic variables are assigned bounds, so can split

$$L^+ = \{j \in L \mid \alpha(x_j) = l_j \text{ and } A_{ij} \geq 0\} \quad U^+ = \{j \in U \mid \alpha(x_j) = u_j \text{ and } A_{ij} \geq 0\}$$

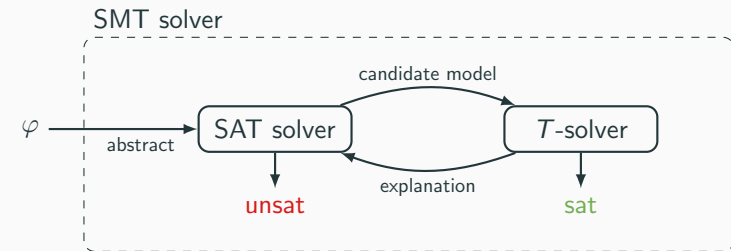
$$L^- = \{j \in L \mid \alpha(x_j) = l_j \text{ and } A_{ij} < 0\} \quad U^- = \{j \in U \mid \alpha(x_j) = u_j \text{ and } A_{ij} < 0\}$$

Lemma (Gomory Cut)

$$\sum_{j \in L^+} \frac{A_{ij}}{1-c} (x_j - l_j) - \sum_{j \in U^-} \frac{A_{ij}}{1-c} (u_j - x_j) - \sum_{j \in L^-} \frac{A_{ij}}{c} (x_j - l_j) + \sum_{j \in U^+} \frac{A_{ij}}{c} (u_j - x_j) \geq 1 \tag{4}$$

Bit Vectors

How to Be Lazy



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 ✓
- DPLL(T) Simplex ✓
- DPLL(T) Simplex + cuts ✓
- bit-blasting

5

Definition (Bit Vector Theory)

- ▶ **variable** x_k is list of length k of propositional variables $x_{k-1} \dots x_2 x_1 x_0$
- ▶ **constant** n_k is bit list of length k
- ▶ formulas built according to grammar

$formula := (formula \vee formula) \mid (formula \wedge formula) \mid (\neg formula) \mid atom$

$atom := term \ rel \ term \mid true \mid false$

$rel := = \mid \neq \mid \geq_u \mid \geq_s \mid >_u \mid >_s$

$term := (term \ binop \ term) \mid (unop \ term) \mid var \mid constant \mid term[i:j] \mid (formula \ ? \ term : term)$

$binop := + \mid - \mid \times \mid \div_u \mid \div_s \mid \%_u \mid \%_s \mid \ll \mid \gg_u \mid \gg_s \mid \& \mid \mid \mid \mid \wedge \mid ::$

$unop := \sim \mid -$

- ▶ **axioms** are equality axioms plus rules for arithmetic/comparison/bitwise operations on bit vectors of length k
- ▶ **solution** assigns bit list of length k to variables x_k

6

Examples

- ▶ $x_4 + y_4 = 7_4$
satisfiable: $v(x_4) = 4_4$ and $v(y_4) = 3_4$

- ▶ $x_4 + 2_4 <_u x_4$
satisfiable: $v(x_4) = 15_4$

overflow semantics!

- ▶ $(x_4 \times y_4 = 6_4) \wedge (x_4 \& y_4 = 2_4)$
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 $v(y_4) = 15_4$

unsigned \gg_u shifts in 0s
signed \gg_s shifts in sign bits

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

7

Notation for Constants

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

Example

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

More examples

- ▶ $-a_4 = a_4$
satisfiable: $v(a_4) = -8_4 = x8_4$

negation uses two's complement

- ▶ $a_8 \div_u b_8 = a_8 \gg_u 1_8$
satisfiable: $v(a_8) = 8_8$ and $v(b_8) = 2_8$

satisfied for powers of 2 (and 0)

- ▶ $a_8 \& (a_8 - 1_8) = 0_8$
satisfiable: $v(a_8) = 8_8$ or $x0_8, x1_8, x2_8, x4_8, x8_8, x10_8, x20_8, x40_8, x80_8$

8

Remarks

- ▶ theory is decidable because carrier is finite
- ▶ common decision procedures use translation to SAT (**bit blasting**)
 - ▶ eager: no DPPL(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)

$$x_1 \neq 0_1 \wedge (y_3 :: x_1) \%_u 2_4 = 0_4 \rightarrow x_1 = 1_1 \wedge (y_3 :: x_1) \%_u 2_4 = 0_4 \\ \rightarrow (y_3 :: 1_1) \%_u 2_4 = 0_4 \rightarrow F$$

Definition (Bit Blasting: Formulas)

bit blasting transformation B transforms BV formula into propositional formula:

$$B(\varphi \vee \psi) = B(\varphi) \vee B(\psi)$$

$$B(\varphi \wedge \psi) = B(\varphi) \wedge B(\psi)$$

$$B(\neg \varphi) = \neg B(\varphi)$$

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

bit blasting B_t for term t
returns (result u , side condition φ)

B_r transforms atom into propositional formula

9

Definition (Bit Blasting: Atoms)

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

► equality

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

► inequality

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

► 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

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

10

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

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

► bitwise negation

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

11

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

12

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:
 \mathbf{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

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

► subtraction

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

13

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

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

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

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

14

Example (SMT-LIB 2 for BV)

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



BV in SMT-LIB 2

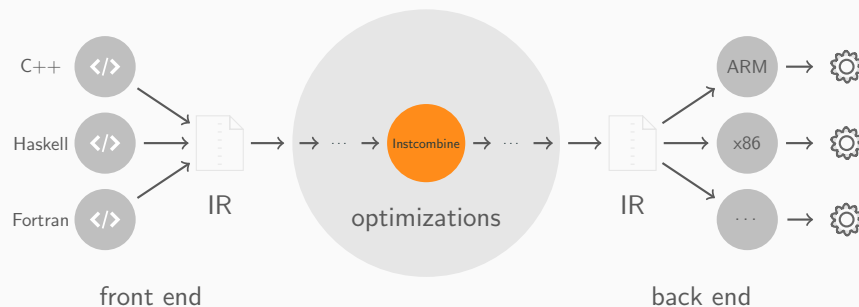
- ▶ `(_ BitVec k)` is sort of bitvectors of length k
- ▶ `#xa` is constant in hexadecimal
- ▶ `#b1000` is constant in binary
- ▶ `bvadd`, `bvsub`, `bvmul` are arithmetic operations, `bvudiv` and `bvssdiv` are unsigned and signed division
- ▶ `bvult` and `bvule` are unsigned, `bvslt` and `bvslle` are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

15

Application: Verifying Compiler Optimizations (1)

LLVM

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



16

Application: Verifying Compiler Optimizations (2)

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

```
int foo(int z) {
  int x = 4 * (z | 1001);
  return -256 & x;
}

define i32 @foo(i32) #0 {
  %2 = or i32 %0, 1001
  %3 = mul nsw i32 4, %2
  %4 = xor i32 -256, %3
  ret i32 %4
}

define i32 @foo(i32) #0 {
  %2 = shl i32 %0, 2
  %3 = or i32 %2, 4004
  %4 = xor i32 %3, -256
  ret i32 %4
}
```

17

Application: Verifying Compiler Optimizations (3)


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 (bvsub #x00000000 (bvdiv x c))))
(assert (= after (bvdiv x (bvneg c))))
(assert (not (= before after)))
(assert (not (= c #x00000000)))
(check-sat)
```



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

18

Bit Vectors 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
```

19

Application: Detecting Nontermination

```
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

20

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.

21