# **BOUNDED MODEL CHECKING**

KV Software Verification WS 18/19



Martina Seidl Institute for Formal Models and Verification



# Example: Verifiction with SMT



### **Expressiveness against Efficiency**

- SAT: efficient, involved encodings
- FO (first-order logic): often too powerful
- satisfiability with respect to some theory is required (non-standard interpretations are not of interest)

**Example:**  $x + y < z \lor \neg (x + 1 \le y \rightarrow x < z)$ 

- theory needs not be first-order axiomatizable
- specialized inference method for each theory



SMT: sweetspot between SAT and FO

- propositional logic + domain specific reasoning
- □ in general more efficient than with general-purpose solvers with incorporated theory axioms

## J⊼N

#### Satisfiability Modulo Theories (SMT)

 $f(x) \neq f(y) \ \land \ x+u = 3 \ \land \ v+y = 3 \ \land \ u = a[z] \ \land \ v = a[w] \ \land \ z = w$ 

#### formulas in first-order logic

- □ usually without quantifiers, variables implicitly existentially quantified
- □ but with sorted / typed symbols and
- □ functions / constants / predicates are interpreted
- □ SMT quantifier reasoning weaker than in first-order theorem proving (FO)
- much richer language compared to propositional logic (SAT)
- many (industrial) applications
  - □ standardized language SMTLIB used in applications and competitions

```
int middle (int x, int y, int z) {
  int m = z;
  if (y < z) {
   if (x < y)
     m = y;
   else if (x < z)
      m = y;
  } else {
   if (x > y)
     m = y;
    else if (x > z)
      m = x;
  }
  return m;
}
```

This program is supposed to return the middle (median) of three numbers.

```
int middle (int x, int y, int z) {
 int m = z;
 if (y < z) {
   if (x < y)
    m = y;
   else if (x < z)
     m = y;
 } else {
   if (x > y)
   m = y;
   else if (x > z)
     m = x;
 }
 return m;
}
```

Some test cases:

middle	(1,	2,	3)	= 2
middle	(1,	З,	2)	= 2
middle	(2,	З,	1)	= 2
middle	(3,	1,	2)	= 2
middle	(3,	2,	1)	= 2
middle	(1,	1,	1)	= 1
middle	(1,	1,	2)	= 1
middle	(1,	2,	1)	= 1
middle	(2,	1,	1)	= 1
middle	(1,	2,	2)	= 2
middle	(2,	1,	2)	= 2
middle	(2,	2,	1)	= 2

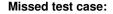
```
int middle (int x, int y, int z) {
  int m = z;
 if (y < z) {
   if (x < y)
    m = y;
   else if (x < z)
     m = y;
  } else {
   if (x > y)
    m = y;
   else if (x > z)
     m = x;
  }
  return m;
}
```

Missed test case:

middle (2, 1, 3) = 1

## J⊼N

```
int middle (int x, int y, int z) {
 int m = z;
 if (y < z) {
   if (x < y)
    m = y;
   else if (x < z)
    m = y;
  } else {
   if (x > y)
    m = y;
    else if (x > z)
     m = x;
  }
  return m;
}
```



middle (2, 1, 3) = 1

**BUG**!

### J⊼∩

 $\rightarrow$ 

#### **Specification for Middle**

Let a be an array of size 3 indexed from 0 to 2.

$$\begin{array}{c} a[i] = x \land a[j] = y \land a[k] = z \\ \land \\ a[0] \le a[1] \land a[1] \le a[2] \\ \land \\ i \ne j \land i \ne k \land j \ne k \\ \rightarrow \\ m = a[1] \end{array}$$

Note: coming up with this specification is a manual process

### **Encoding of Middle in Logic**

### J⊻U

### **Encoding of Middle in Logic**

$$\begin{array}{l} (y < z \land x < y \rightarrow m = y) \\ \land \\ (y < z \land x \ge y \land x < z \rightarrow m = y) \\ \land \\ (y < z \land x \ge y \land x \ge z \rightarrow m = z) \\ \land \\ (y \ge z \land x > y \rightarrow m = y) \\ \land \\ (y \ge z \land x \le y \land x > z \rightarrow m = x) \\ \land \\ (y \ge z \land x \le y \land x \le z \rightarrow m = z) \end{array}$$

### **Encoding of Middle in Logic**

$$\begin{array}{l} (y < z \land x < y \rightarrow m = y) \\ \land \\ (y < z \land x \ge y \land x < z \rightarrow m = y) \\ \land \\ (y < z \land x \ge y \land x \ge z \rightarrow m = z) \\ \land \\ (y \ge z \land x \ge y \land x \ge z \rightarrow m = z) \\ \land \\ (y \ge z \land x \le y \land x > z \rightarrow m = x) \\ \land \\ (y \ge z \land x \le y \land x \le z \rightarrow m = z) \end{array}$$

automatic

### J⊻U

Let P be the encoding of the program, and S of the specification

**\square** program is correct if " $P \rightarrow S$ " is valid

Let P be the encoding of the program, and S of the specification

**\square** program is correct if " $P \rightarrow S$ " is valid

**\blacksquare** program has a bug if " $P \rightarrow S$ " is invalid

Let P be the encoding of the program, and S of the specification

**\square** program is correct if " $P \rightarrow S$ " is valid

**I** program has a bug if " $P \rightarrow S$ " is invalid

**\blacksquare** program has a bug if negation of " $P \rightarrow S$ " is satisfiable

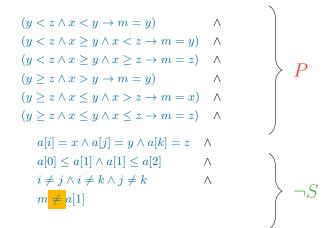
Let P be the encoding of the program, and S of the specification

- **\square** program is correct if " $P \rightarrow S$ " is valid
- **\square** program has a bug if " $P \rightarrow S$ " is invalid
- **\blacksquare** program has a bug if negation of " $P \rightarrow S$ " is satisfiable
- **I** program has a bug if " $P \land \neg S$ " is satisfiable (has a model)

Let *P* be the encoding of the program, and *S* of the specification **\blacksquare** program has a bug if "*P*  $\land \neg S$ " is satisfiable (has a model)

 $\begin{array}{c} (y < z \land x < y \rightarrow m = y) & \land \\ (y < z \land x \ge y \land x < z \rightarrow m = y) & \land \\ (y < z \land x \ge y \land x \ge z \rightarrow m = z) & \land \\ (y \ge z \land x \ge y \land x \ge z \rightarrow m = z) & \land \\ (y \ge z \land x \le y \land x > z \rightarrow m = x) & \land \\ (y \ge z \land x \le y \land x \le z \rightarrow m = z) & \land \end{array} \right\} P$ 

Let *P* be the encoding of the program, and *S* of the specification  $\blacksquare$  program has a bug if "*P*  $\land \neg S$ " is satisfiable (has a model)



```
(set-logic QF AUFLIA)
(declare-fun x () Int) (declare-fun y () Int) (declare-fun z () Int)
(declare-fun i () Int) (declare-fun i () Int) (declare-fun k () Int)
(declare-fun m () Int) (declare-fun a () (Arrav Int Int))
(assert (=> (and (< y z) (< x y) ) (= m y)))
(assert (=> (and (< y z) (>= x y) (< x z)) (= m y)))
(assert (=> (and (< v z) (>= x v) (>= x z)) (= m z)))
(assert (=> (and (>= v z) (> x v) ) (= m v)))
(assert (=> (and (>= y z) (<= x y) (> x z) ) (= m x)))
(assert (=> (and (>= y z) (<= x y) (<= x z)) (= m z)))
(assert (and (<= 0 i) (<= i 2) (<= 0 j) (<= j 2) (<= 0 k) (<= k 2)))
(assert (and (= (select a i) x) (= (select a j) v) (= (select a k) z)))
(assert (<= (select a 0) (select a 1) (select a 2)))
(assert (distinct i j k))
(assert (distinct m (select a 1)))
(check-sat) (get-model) (exit)
```

```
(set-logic QF AUFLIA)
(declare-fun x () Int) (declare-fun y () Int) (declare-fun z () Int)
(declare-fun i () Int) (declare-fun i () Int) (declare-fun k () Int)
(declare-fun m () Int) (declare-fun a () (Arrav Int Int))
(assert (=> (and (< y z) (< x y) ) (= m y)))
(assert (=> (and (< y z) (>= x y) (< x z)) (= m y)))
(assert (=> (and (< v z) (>= x v) (>= x z)) (= m z)))
(assert (=> (and (>= v z) (> x v) ) (= m v)))
(assert (=> (and (>= y z) (<= x y) (> x z) ) (= m x)))
(assert (=> (and (>= y z) (<= x y) (<= x z)) (= m z)))
(assert (and (<= 0 i) (<= i 2) (<= 0 j) (<= j 2) (<= 0 k) (<= k 2)))
(assert (and (= (select a i) x) (= (select a j) v) (= (select a k) z)))
(assert (<= (select a 0) (select a 1) (select a 2)))
(assert (distinct i j k))
(assert (distinct m (select a 1)))
(check-sat) (get-model) (exit)
```

### J⊼N

```
(set-logic QF AUFLIA)
(declare-fun x () Int) (declare-fun y () Int) (declare-fun z () Int)
(declare-fun i () Int) (declare-fun i () Int) (declare-fun k () Int)
(declare-fun m () Int) (declare-fun a () (Arrav Int Int))
(assert (=> (and (< y z) (< x y) ) (= m y)))
(assert (=> (and (< y z) (>= x y) (< x z)) (= m y)))
(assert (=> (and (< v z) (>= x v) (>= x z)) (= m z)))
(assert (=> (and (>= v z) (> x v) ) (= m v)))
(assert (=> (and (>= y z) (<= x y) (> x z) ) (= m x)))
(assert (=> (and (>= y z) (<= x y) (<= x z)) (= m z)))
(assert (and (<= 0 i) (<= i 2) (<= 0 j) (<= j 2) (<= 0 k) (<= k 2)))
(assert (and (= (select a i) x) (= (select a j) v) (= (select a k) z)))
(assert (<= (select a 0) (select a 1) (select a 2)))
(assert (distinct i j k))
(assert (distinct m (select a 1)))
(check-sat) (get-model) (exit)
```

```
(set-logic QF AUFLIA)
(declare-fun x () Int) (declare-fun y () Int) (declare-fun z () Int)
(declare-fun i () Int) (declare-fun i () Int) (declare-fun k () Int)
(declare-fun m () Int) (declare-fun a () (Arrav Int Int))
(assert (=> (and (< y z) (< x y) ) (= m y)))
(assert (=> (and (< y z) (>= x y) (< x z)) (= m y)))
(assert (=> (and (< v z) (>= x v) (>= x z)) (= m z)))
(assert (=> (and (>= v z) (> x v) ) (= m v)))
(assert (=> (and (>= y z) (<= x y) (> x z) ) (= m x)))
(assert (=> (and (>= y z) (<= x y) (<= x z)) (= m z)))
(assert (and (<= 0 i) (<= i 2) (<= 0 j) (<= j 2) (<= 0 k) (<= k 2)))
(assert (and (= (select a i) x) (= (select a j) v) (= (select a k) z)))
(assert (<= (select a 0) (select a 1) (select a 2)))
(assert (distinct i j k))
(assert (distinct m (select a 1)))
(check-sat) (get-model) (exit)
```

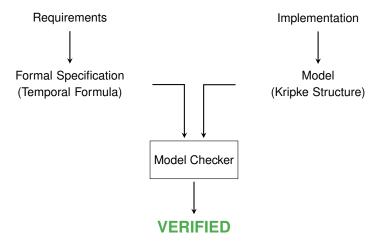
```
(set-logic QF AUFLIA)
(declare-fun x () Int) (declare-fun y () Int) (declare-fun z () Int)
(declare-fun i () Int) (declare-fun i () Int) (declare-fun k () Int)
(declare-fun m () Int) (declare-fun a () (Arrav Int Int))
(assert (=> (and (< y z) (< x y) ) (= m y)))
(assert (=> (and (< y z) (>= x y) (< x z)) (= m y)))
(assert (=> (and (< v z) (>= x v) (>= x z)) (= m z)))
(assert (=> (and (>= v z) (> x v) ) (= m v)))
(assert (=> (and (>= y z) (<= x y) (> x z) ) (= m x)))
(assert (=> (and (>= y z) (<= x y) (<= x z)) (= m z)))
(assert (and (<= 0 i) (<= i 2) (<= 0 j) (<= j 2) (<= 0 k) (<= k 2)))
(assert (and (= (select a i) x) (= (select a j) y) (= (select a k) z)))
(assert (<= (select a 0) (select a 1) (select a 2)))
(assert (distinct i j k))
(assert (distinct m (select a 1)))
(check-sat) (get-model) (exit)
```

```
(set-logic QF AUFLIA)
(declare-fun x () Int) (declare-fun y () Int) (declare-fun z () Int)
(declare-fun i () Int) (declare-fun i () Int) (declare-fun k () Int)
(declare-fun m () Int) (declare-fun a () (Arrav Int Int))
(assert (=> (and (< y z) (< x y) ) (= m y)))
(assert (=> (and (< y z) (>= x y) (< x z)) (= m y)))
(assert (=> (and (< v z) (>= x v) (>= x z)) (= m z)))
(assert (=> (and (>= v z) (> x v) ) (= m v)))
(assert (=> (and (>= y z) (<= x y) (> x z) ) (= m x)))
(assert (=> (and (>= y z) (<= x y) (<= x z)) (= m z)))
(assert (and (<= 0 i) (<= i 2) (<= 0 j) (<= j 2) (<= 0 k) (<= k 2)))
(assert (and (= (select a i) x) (= (select a j) v) (= (select a k) z)))
(assert (<= (select a 0) (select a 1) (select a 2)))
(assert (distinct i j k))
(assert (distinct m (select a 1)))
(check-sat) (get-model) (exit)
```

# **Bounded Model Checking**

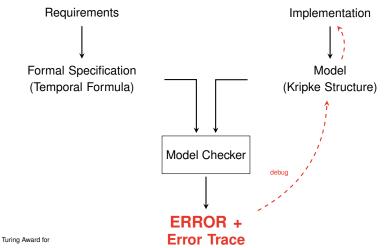


### **Model Checking**





### **Model Checking**



- Queille, J. P.; Sifakis, J. (1982), "Specification and verification of concurrent systems in CESAR", International Symposium on Programming, citations: 1900
- Edmund M. Clarke, E. Allen Emerson: "Design and Synthesis of Synchronization Skeletons Using Branching-Time Temporal Logic". Logic of Programs 1981: 52-71, citations: 3895

#### **Types of Model Checking**

**General question**: Given a system K and a property p, does p hold for K (i.e., for all initial states of K) ?

Explicit state model checking

- $\hfill\square$  enumeration of the state space
- □ state explosion problem

Symbolic model checking

□ representation of model checking problem as logical formula (e.g., in propositional logic (SAT) or QBF)

### J⊻U

#### **Bounded Model Checking**

basic idea: search for a counter-example of bounded length k

- encoding in propositional logic (or extensions)
- use SAT solvers to find such a counter-example: formula is satisfiable iff a bug is found, i.e., an execution of program that violates the claim.
- benefits:
  - □ bit-precise encoding of the real semantics
  - □ powerful SAT solvers
  - $\Box$  difficulty of the problem is controllable (by selection of k)
- $\blacksquare drawback: incomplete for k that is too small$

#### $\Rightarrow$ can be used for debugging

Armin Biere, Alessandro Cimatti, Edmund M. Clarke, Yunshan Zhu (1999) Symbolic Model Checking without BDDs. TACAS 193-207, citations: 2580



### Propositional Satisfiability (SAT)

Given propositional formula  $\phi.$  Is there a satisfying truth assignment for  $\phi?$ 

- SAT solvers are very powerful solving tools
- Using SAT as a "programming language" is very successful in many domains

#### Example

Given: 
$$\phi = (x_1 \vee \neg x_2) \land (\neg x_1 \vee x_3)$$

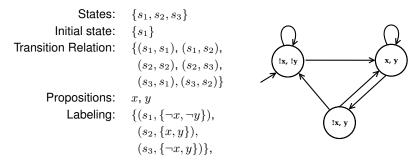
Question: Is  $\phi$  satisfiable?

Yes! For example:  $x_1 = x_3 =$ true,  $x_2 =$ false.

## J⊼N

#### Symbolic System Representation

Kripke Structure: Description of the System



#### Translation to SAT

Initial state: Transition Relation:

$$I((x,y)) = \neg x \land \neg y$$
  

$$T((x,y), (x',y')) = ((x' \Leftrightarrow x \lor y) \land (y' \Leftrightarrow y)) \lor$$
  

$$((x' \Leftrightarrow \neg y) \land (y' \Leftrightarrow x \lor \neg y))$$

#### **Bounded Model Checking (Safety)**

- Given a Kripke structure *K*. Is there a path of length *k* to a **bad state** *s*, i.e., a certain property *p* is violated in *s*?
- In other words: there is a path where Gp does not hold in K
- Observation: if Gp does not hold in K, there is a finite counter-example.
- Idea: consider paths of fixed length k
  - $\Box$  encode problem to propositional formula  $\phi$
  - □ pass problem to SAT solver
  - $\Box \phi$  is true  $\Leftrightarrow$  model of  $\phi$  is counter-example
  - $\hfill\square$  if  $\phi$  is false, then increase k

### J⊻U

#### **Bounded Model Checking (Safety)**

A bounded model checking (BMC) problem for Kripke structure K and safety property Gp is encoded by

$$I(s_0) \wedge \mathcal{T}(s_0, s_1) \wedge \mathcal{T}(s_1, s_2) \wedge \ldots \wedge \mathcal{T}(s_{k-1}, s_k) \wedge B(s_k)$$

where

- $\blacksquare$   $I(s_0)$  is true  $\Leftrightarrow s_0$  is an initial state
- $\blacksquare$   $\mathcal{T}$  is the transition relation of K
- **\blacksquare**  $B(s_k)$  is true  $\Leftrightarrow s_k$  is a bad state, i.e.,  $\neg p$  holds in  $s_k$

### J⊻U

# Bounded Model Checking for Software



#### **Bounded Model Checking of ANSI-C Programs**

#### idea:

- □ unwind program into equation
- □ check equation using SAT/SMT
- benefits:
  - □ completely automated
  - treatment of pointers and dynamic memory is possible
- properties:
  - □ simple assertions
  - □ run time errors (pointers/arrays)
  - □ run time guarantees (WCET)

#### for example implemented in tool CBMC

A tool for checking ANSI-C programs E Clarke, D Kroening, F Lerda Tools and Algorithms for the Construction and Analysis of Systems, 168-176, citations: 1339

#### From C to SAT/SMT

removal of side effects
 example: j=i++ is rewritten to j=i; i=i+1
 control flow is made explicit
 example: continue, break are replaced by goto
 transformation of loops to while (...) ...

■ while (...) ... loops are unwound

- □ all loops must be bounded
  - $\rightarrow$  analysis may become incomplete
- constant loop bounds are found automatically, others must be specified by user
- $\hfill\square$  to ensure sufficient unwinding, "unwinding assertions" are added

### J⊻U

### From C to SAT/SMT: Loop Unwinding

#### original function:

```
void f (...) {
    ...
    while (cond) {
        body;
    }
    rest;
}
```

#### with unwounded loop:

after last iteration an assertion is added:

violated if program runs longer than bound permits

### From C to SAT/SMT: SSA

single static assignment (SSA) form: fresh variable for LHS of each assignment

#### example:

x = x + y; x = x \* 2; a[i] = 100;

#### is translated to

x1 = x0 + y0; x2 = x1 \* 2; a1[i0] = 100;

from which the following SMT formula can be derived

$$(x_1 = x_0 + y_0) \land (x_2 = x_1 * 2) \land (a_1[i_0] = 100)$$

#### From C to SAT/SMT: Conditionals

for each join point, new variables with selectors are addedexample:

original program:rewritten program:if (v)if (v0)x = y;x0 = y0;elsex = z;w = x;x2 = v0 ? x0 : x1;w = x;x2 = v0 ? x0 : x1;

#### From C to SAT/SMT: Example

int main () { int main () { int x, y; int x, y; y = 1; y1 = 1; if (x) if(x0)  $y_2 = y_{1-1};$ y-; else else y3 = y1+1; v++;  $\Rightarrow$  $\Rightarrow$ y4 = x0 ? y2 : y3;assert (y==2 || y==3); assert (y4==2 || y4==3); } }

$$((y_1 = 1) \land (y_2 = y_1 - 1) \land (y_3 = y_1 + 1) \land (y_4 = x_0 ? y_2 : y_3))$$
$$\rightarrow ((y_4 = 2) \lor (y_4 = 3))$$

### J⊻U

#### Arrays

functions "read" and "write": read(a, i), write(a, i, v)
 axioms

array congruence

$$\forall a, i, j \colon i = j \rightarrow \mathsf{read}(a, i) = \mathsf{read}(a, j)$$

read over write 1

$$\forall a, v, i, j \colon i = j \rightarrow \mathsf{read}(\mathsf{write}(a, i, v), j) = v$$

read over write 2

 $\forall a, v, i, j : i \neq j \rightarrow \mathsf{read}(\mathsf{write}(a, i, v), j) = \mathsf{read}(a, j)$ 

used to model memory (HW and SW)

### J⊻U

eagerly reduce arrays to uninterpreted functions:

```
read(write(a, i, v), j) replaced by (i = j ? v : read(a, j))
```

eagerly reduce arrays to uninterpreted functions:

```
\mathsf{read}(\mathsf{write}(a,i,v),j) \quad \mathsf{replaced by} \quad (i=j \ ? \ v:\mathsf{read}(a,j))
```

Example:

 $i \neq j \land u = \operatorname{read}(\operatorname{write}(a, i, v), j) \land v = \operatorname{read}(a, j) \land u \neq v$ 

eagerly reduce arrays to uninterpreted functions:

```
\mathsf{read}(\mathsf{write}(a,i,v),j) \quad \mathsf{replaced by} \quad (i=j \ ? \ v:\mathsf{read}(a,j))
```

Example:

 $i \neq j \land u = \operatorname{read}(\operatorname{write}(a, i, v), j) \land v = \operatorname{read}(a, j) \land u \neq v$ 

 $i \neq j \ \land \ u = (i = j \ ? \ v : \mathsf{read}(a, j)) \ \land \ v = \mathsf{read}(a, j) \ \land \ u \neq v$ 

eagerly reduce arrays to uninterpreted functions:

 $\mathsf{read}(\mathsf{write}(a,i,v),j) \quad \mathsf{replaced by} \quad (i=j \ ? \ v:\mathsf{read}(a,j))$ 

Example:

 $i \neq j \land u = \operatorname{read}(\operatorname{write}(a, i, v), j) \land v = \operatorname{read}(a, j) \land u \neq v$ 

 $i \neq j \ \land \ u = (i = j \ ? \ v : \mathsf{read}(a, j)) \ \land \ v = \mathsf{read}(a, j) \ \land \ u \neq v$ 

 $i \neq j \ \land \ u = \mathsf{read}(a,j) \ \land \ v = \mathsf{read}(a,j) \ \land \ u \neq v$ 

eagerly reduce arrays to uninterpreted functions:

read(write(a, i, v), j) replaced by (i = j ? v : read(a, j))Example:

 $i \neq j \land u = \operatorname{read}(\operatorname{write}(a, i, v), j) \land v = \operatorname{read}(a, j) \land u \neq v$ 

$$i \neq j \ \land \ u = (i = j \ ? \ v : \mathsf{read}(a, j)) \ \land \ v = \mathsf{read}(a, j) \ \land \ u \neq v$$

$$i \neq j \ \land \ u = \mathsf{read}(a,j) \ \land \ v = \mathsf{read}(a,j) \ \land \ u \neq v$$

$$i \neq j \ \land \ u = \mathsf{read}(a, j) = \mathsf{read}(a, j) = v \ \land \ u \neq v$$

eagerly reduce arrays to uninterpreted functions:

read(write(a, i, v), j) replaced by (i = j ? v : read(a, j))Example:

 $i \neq j \land u = \operatorname{read}(\operatorname{write}(a, i, v), j) \land v = \operatorname{read}(a, j) \land u \neq v$ 

$$i \neq j \ \land \ u = (i = j \ ? \ v : \mathsf{read}(a, j)) \ \land \ v = \mathsf{read}(a, j) \ \land \ u \neq v$$

$$i \neq j \land u = \operatorname{read}(a, j) \land v = \operatorname{read}(a, j) \land u \neq v$$

$$i \neq j \land u = \operatorname{read}(a, j) = \operatorname{read}(a, j) = v \land u \neq v$$
  
UNSATISFIABLE

