Proving termination through conditional termination

Cristina Borralleras, Marc Brockschmidt, Daniel Larraz, Albert Oliveras, Enric Rodríguez-Carbonell and <u>Albert Rubio</u>

> Universitat de Vic Universitat Politècnica de Catalunya - Barcelona Tech Microsoft Research, Cambridge

supported by MINECO/FEDER project TIN2015-69175-C4-3-R

Bergen, Norway October 2016

1 Introduction

- 2 SMT/Max-SMT solving
- 3 (Conditional) Invariant generation
- 4 Termination analysis
- 5 Compositional termination analysis
- 6 Conclusions and current work

1 Introduction

- 2 SMT/Max-SMT solving
- 3 (Conditional) Invariant generation
- 4 Termination analysis
- 5 Compositional termination analysis
- 6 Conclusions and current work

• Main Goal: Build static analysis tools for programmers.

- Fully automatic.
- Efficient.
- Scalable.

- Main Goal: Build static analysis tools for programmers.
 - Fully automatic.
 - Efficient.
 - Scalable.
- Strategy: Take advantage of powerful arithmetic constraint solvers.
 SMT solvers

Constraint-based Program Analysis techniques

- Main Goal: Build static analysis tools for programmers.
 - Fully automatic.
 - Efficient.
 - Scalable.
- Strategy: Take advantage of powerful arithmetic constraint solvers. Max-SMT solvers

Constraint-based Program Analysis techniques

- Main Goal: Build static analysis tools for programmers.
 - Fully automatic.
 - Efficient.
 - Scalable.
- Strategy: Take advantage of powerful arithmetic constraint solvers. Max-SMT solvers

Constraint-based Program Analysis techniques

Goal: Prove termination of imperative programs automatically

- Main Goal: Build static analysis tools for programmers.
 - Fully automatic.
 - Efficient.
 - Scalable.
- Strategy: Take advantage of powerful arithmetic constraint solvers. Max-SMT solvers

Constraint-based Program Analysis techniques

- Goal: Prove termination of imperative programs automatically
 - Find ranking functions.
 - Find supporting invariants.
 - How to guide the search!.

- Main Goal: Build static analysis tools for programmers.
 - Fully automatic.
 - Efficient.
 - Scalable.
- Strategy: Take advantage of powerful arithmetic constraint solvers. Max-SMT solvers

Constraint-based Program Analysis techniques

- Goal: Prove termination of imperative programs automatically
 - Find ranking functions.
 - Find supporting (conditional) invariants.
 - How to guide the search!.

1 Introduction

2 SMT/Max-SMT solving

- 3 (Conditional) Invariant generation
- 4 Termination analysis
- 5 Compositional termination analysis
- 6 Conclusions and current work

We make extensive use of SMT solvers inside our program analysis tools. Input: Given a boolean formula φ over some theory T. Question: Is there any model that satisfies the formula? Example: T = non-linear (polynomial) integer/real arithmetic. $(x^2 + y^2 > 2 \lor x \cdot z \le y \lor y \cdot z < z^2) \land (x > y \lor 0 < z)$ $\{x = 0, y = 1, z = 1\}$

We make extensive use of SMT solvers inside our program analysis tools. Input: Given a boolean formula φ over some theory T. Question: Is there any model that satisfies the formula? Example: T = non-linear (polynomial) integer/real arithmetic. $(x^2 + y^2 > 2 \lor x \cdot z \le y \lor y \cdot z < z^2) \land (x > y \lor 0 < z)$ $\{x = 0, y = 1, z = 1\}$

We make extensive use of SMT solvers inside our program analysis tools. Input: Given a boolean formula φ over some theory T. Question: Is there any model that satisfies the formula? Example: T = non-linear (polynomial) integer/real arithmetic. $(x^2 + y^2 > 2 \lor x \cdot z \le y \lor y \cdot z < z^2) \land (x > y \lor 0 < z)$

$$\{x = 0, y = 1, z = 1\}$$

Non-linear arithmetic decidability:

- Integers: undecidable (Hilbert's 10th problem).
- <u>Reals</u>: decidable (Tarski) but algorithms have prohibitive complexity.

We make extensive use of SMT solvers inside our program analysis tools. Input: Given a boolean formula φ over some theory T. Question: Is there any model that satisfies the formula? Example: T = non-linear (polynomial) integer/real arithmetic. $(x^2 + y^2 > 2 \lor x \cdot z \le y \lor y \cdot z < z^2) \land (x > y \lor 0 < z)$

$$\{x = 0, y = 1, z = 1\}$$

Non-linear arithmetic decidability:

- Integers: undecidable (Hilbert's 10th problem).
- <u>Reals</u>: decidable (Tarski) but algorithms have prohibitive complexity.

Incomplete solvers focus on either satisfiability or unsatisfiability.

We make extensive use of SMT solvers inside our program analysis tools. **Input:** Given a boolean formula φ over some theory T. **Question:** Is there any <u>model</u> that satisfies the formula? Example: T = non-linear (polynomial) integer/real arithmetic. $(x^2 + y^2 > 2 \lor x \cdot z \le y \lor y \cdot z < z^2) \land (x > y \lor 0 < z)$

$$\{x = 0, y = 1, z = 1\}$$

Non-linear arithmetic decidability:

- Integers: undecidable (Hilbert's 10th problem).
- <u>Reals</u>: decidable (Tarski) but algorithms have prohibitive complexity.

Incomplete solvers focus on either satisfiability or unsatisfiability.

We make extensive use of SMT solvers inside our program analysis tools. Input: Given a boolean formula φ over some theory T. Question: Is there any model that satisfies the formula? Example: T = non-linear (polynomial) integer/real arithmetic. $(x^2 + y^2 > 2 \lor x \cdot z \le y \lor y \cdot z < z^2) \land (x > y \lor 0 < z)$

$${x = 0, y = 1, z = 1}$$

- Need to handle large formulas with non-linear arithmetic and <u>complex</u> boolean structure.
- Barcelogic has shown to be the best SMT-solver proving satisfiability of this kind of problems.

(Weighted) Max-SMT problem

Input: Given an SMT formula $\varphi = C_1 \land \ldots \land C_m$ in CNF, where some of the clauses are hard and the others soft with a weight.

Output: An assignment for the hard clauses that minimizes the sum of the weights of the falsified soft clauses.

$$(x^2 + y^2 > 2 \lor x \cdot z \le y \lor y \cdot z < z^2) \land (x > y \lor 0 < z \lor w(5)) \land \dots$$

1 Introduction

- 2 SMT/Max-SMT solving
- 3 (Conditional) Invariant generation
- 4 Termination analysis
- 5 Compositional termination analysis
- 6 Conclusions and current work

An <u>invariant</u> of a program at a location is an assertion over the program variables that remains true whenever the location is reached.

An <u>invariant</u> of a program at a location is an assertion over the program variables that remains true whenever the location is reached.

Definition

An invariant is said to be inductive at a program location if:

- Initiation condition: It holds the first time the location is reached.
- Consecution condition: It is preserved under every cycle back to the location.

An <u>invariant</u> of a program at a location is an assertion over the program variables that remains true whenever the location is reached.

Definition

An invariant is said to be inductive at a program location if:

- Initiation condition: It holds the first time the location is reached.
- Consecution condition: It is preserved under every cycle back to the location.
- We focus on inductive invariants.

We inspire ourselves with the constraint-based method [CSS'03].

Assume input programs consist of linear expressions.

We inspire ourselves with the constraint-based method [CSS'03]. Assume input programs consist of linear expressions.

Keys:

■ Use a template for candidate invariants.

 $c_1x_1+\ldots+c_nx_n+d\leq 0$

We inspire ourselves with the constraint-based method [CSS'03]. Assume input programs consist of linear expressions.

Keys:

■ Use a template for candidate invariants.

$$c_1x_1+\ldots+c_nx_n+d\leq 0$$

■ Impose <u>initiation</u> and <u>consecution</u> conditions obtaining an ∃∀ problem over <u>non-linear</u> arithmetic.

We inspire ourselves with the constraint-based method [CSS'03]. Assume input programs consist of linear expressions.

Keys:

■ Use a template for candidate invariants.

$$c_1x_1+\ldots+c_nx_n+d\leq 0$$

- Impose <u>initiation</u> and <u>consecution</u> conditions obtaining an ∃∀ problem over <u>non-linear</u> arithmetic.
- Transform it using Farkas' Lemma into an ∃ problem over non-linear arithmetic.

```
int isqrt(int N) {
    int a = 0, s = 1, t = 1;
    // Inv: c_1a + c_2s + c_3t + d \le 0
    while (s \le N) {
        a = a + 1;
        s = s + t + 2;
        t = t + 2;
    }
    return a;
}
```

```
int isqrt(int N) {
   int a = 0, s = 1, t = 1;
   // Inv: c_1 a + c_2 s + c_3 t + d < 0
   while (s < N) {
      a = a + 1:
      s = s + t + 2;
       t = t + 2;
   }
   return a;
}
\exists c_1, c_2, c_3, d \forall a, s, t
true \Longrightarrow c_1 \cdot 0 + c_2 \cdot 1 + c_3 \cdot 1 + d \leq 0 \land Initiation condition
s \leq N \wedge c_1 \cdot a + c_2 \cdot s + c_3 \cdot t + d \leq 0 \Longrightarrow c_1 \cdot (a+1) + c_2 \cdot (s+t+2) + c_3 \cdot (t+2) + d \leq 0
                                        consecution condition
```

```
int isqrt(int N) {
   int a = 0, s = 1, t = 1;
   // Inv: c_1 a + c_2 s + c_3 t + d < 0
   while (s < N) {
      a = a + 1;
      s = s + t + 2;
       t = t + 2;
   }
   return a;
}
\exists c_1, c_2, c_3, d \forall a, s, t
c_2+c_3+d\leq 0 \wedge
                                   Initiation condition
s \leq N \land c_1 \cdot a + c_2 \cdot s + c_3 \cdot t + d \leq 0 \Longrightarrow c_1 \cdot a + c_2 \cdot s + (c_2 + c_3) \cdot t + c_1 + 2c_2 + 2c_3 + d \leq 0
                                        consecution condition
      Albert Rubio (UPC)
                                 Proving termination through conditional term
                                                                                    Bergen 2016
                                                                                                   11 / 34
```

Square root of a natural number N:

```
int isqrt(int N) {
  int a = 0, s = 1, t = 1;
  // Inv: c_1 a + c_2 s + c_3 t + d < 0
  while (s < N) {
    a = a + 1;
    s = s + t + 2:
    t = t + 2;
  }
  return a;
}
Apply Farkas' Lemma to remove \forall a, s, t
```

Use Barcelogic to solve the non-linear SMT problem!

```
int isqrt(int N) {
  int a = 0, s = 1, t = 1;
  // Inv: c_1 a + c_2 s + c_3 t + d \le 0
  while (s \leq N) {
    a = a + 1;
    s = s + t + 2;
    t = t + 2;
  }
  return a;
ን
```

$$\{c_1 = -2, c_2 = 0, c_3 = 1, d = -1\}$$

```
int isqrt(int N) {
  int a = 0, s = 1, t = 1;
  // Inv: -2a + 0s + 1t - 1 < 0
  while (s < N) {
    a = a + 1;
    s = s + t + 2;
    t = t + 2;
  }
  return a;
}
```

$$\{c_1 = -2, c_2 = 0, c_3 = 1, d = -1\}$$

A formula is a conditional (inductive) invariant at a program location if:

Consecution condition holds.

A formula is a conditional (inductive) invariant at a program location if:

- Consecution condition holds.
- but Initiation condition may not hold.

A formula is a conditional (inductive) invariant at a program location if:

- Consecution condition holds. Hard
- but Initiation condition may not hold.

A formula is a conditional (inductive) invariant at a program location if:

- Consecution condition holds. Hard
- but Initiation condition may not hold. Soft

Key: We prefer invariants but we can live with conditional invariants

A formula is a conditional (inductive) invariant at a program location if:

- Consecution condition holds. Hard
- but Initiation condition may not hold. Soft

Key: We prefer invariants but we can live with conditional invariants

Consider that this is an optimization problem rather than a satisfiability problem
Definition

A formula is a conditional (inductive) invariant at a program location if:

- Consecution condition holds. Hard
- but Initiation condition may not hold. Soft

Key: We prefer invariants but we can live with conditional invariants

Consider that this is an optimization problem rather than a satisfiability problem

Encode the problem using Max-SMT,

We use Barcelogic to solve it.

1 Introduction

- 2 SMT/Max-SMT solving
- 3 (Conditional) Invariant generation
- 4 Termination analysis
- 5 Compositional termination analysis
- 6 Conclusions and current work

Basic method: find a single ranking function f: States $\rightarrow \mathbb{Z}$, with $f(S) \ge 0$ and f(S) > f(S') after every iteration.

Ranking functions and Invariants

Basic method: find a single ranking function f: States $\rightarrow \mathbb{Z}$, with $f(S) \ge 0$ and f(S) > f(S') after every iteration.

It does not work in practice in many cases

What is (at least) necessary?

Ranking functions and Invariants

Basic method: find a single ranking function f: States $\rightarrow \mathbb{Z}$, with $f(S) \ge 0$ and f(S) > f(S') after every iteration.

It does not work in practice in many cases

What is (at least) necessary?

- Find supporting (conditional) invariants
- Consider a (lexicographic) combination of ranking functions

Ranking functions and Invariants

Basic method: find a single ranking function f: States $\rightarrow \mathbb{Z}$, with $f(S) \ge 0$ and f(S) > f(S') after every iteration.

It does not work in practice in many cases

What is (at least) necessary?

Find supporting (conditional) invariants

Consider a (lexicographic) combination of ranking functions

int main() { $y > 0 \wedge z < 0$ $\tau_1: \begin{array}{c} \wedge x' = x \\ \wedge y' = y + z \end{array}$ int x, y, z; y < 0 $\begin{array}{c} \wedge y' = y + z \\ \wedge z' = z - 1 \end{array} \quad \begin{array}{c} \wedge x' = x \\ \wedge y' = y \end{array}$ x = nondet();y = nondet(); $\wedge z' = z$ z = nondet();while $(y \ge 0 \&\& z \ne 0)$ { τ_0 : true if $(z < 0) \{ y = y + z;$ ℓ_0 ℓ_1 z = z - 1: } else { x = x - z; z = 0 $\wedge x' = x$ y = y + x; $\tau_4: \quad \stackrel{\frown}{\wedge} y' = y$ $y > 0 \wedge z > 0$ z = z + 1; $\tau_2: \begin{array}{c} \wedge x' = x - z \\ \wedge y' = y + x \end{array}$ $\wedge z' = z$ $\wedge z' = z + 1$

 ℓ_2

In order to discard a transition τ_i we need to find a ranking function f over the integers such that:

1
$$\tau_i \Longrightarrow f(x_1, \dots, x_n) \ge 0$$
 (bounded)
2 $\tau_i \Longrightarrow f(x_1, \dots, x_n) > f(x'_1, \dots, x'_n)$ (strict-decreasing)
3 $\tau_j \Longrightarrow f(x_1, \dots, x_n) \ge f(x'_1, \dots, x'_n)$ for all j (non-increasing)

Use a linear template for the ranking function as well.

In order to prove properties of the ranking function we may need to generate invariants.

Generation of both conditional invariants and ranking functions should be combined in the same optimization problem.

In order to prove properties of the ranking function we may need to generate invariants.

Generation of both conditional invariants and ranking functions should be combined in the same optimization problem.

$$\begin{array}{l} \mathbb{1} \ \mathcal{I} \wedge \tau_i \Longrightarrow f(x_1, \dots, x_n) \ge 0 & (bounded) \\ \mathbb{2} \ \mathcal{I} \wedge \tau_i \Longrightarrow f(x_1, \dots, x_n) > f(x'_1, \dots, x'_n) & (strict-decreasing) \\ \mathbb{3} \ \mathcal{I} \wedge \tau_j \Longrightarrow f(x_1, \dots, x_n) \ge f(x'_1, \dots, x'_n) \text{ for all } j & (non-increasing) \end{array}$$

In order to prove properties of the ranking function we may need to generate invariants.

Generation of both conditional invariants and ranking functions should be combined in the same optimization problem.

$$\begin{array}{l} \mathbf{1} \ \mathcal{I} \wedge \tau_i \Longrightarrow f(x_1, \dots, x_n) \ge 0 & (bounded) \\ \\ \mathbf{2} \ \mathcal{I} \wedge \tau_i \Longrightarrow f(x_1, \dots, x_n) > f(x'_1, \dots, x'_n) & (strict-decreasing) \\ \\ \\ \mathbf{3} \ \mathcal{I} \wedge \tau_j \Longrightarrow f(x_1, \dots, x_n) \ge f(x'_1, \dots, x'_n) \text{ for all } j & (non-increasing) \end{array}$$

Considering conditional invariants give more chances to the solver

In order to prove properties of the ranking function we may need to generate invariants.

Generation of both conditional invariants and ranking functions should be combined in the same optimization problem.

1
$$\mathcal{I} \wedge \tau_i \Longrightarrow f(x_1, \dots, x_n) \ge 0$$
 (bounded)
2 $\mathcal{I} \wedge \tau_i \Longrightarrow f(x_1, \dots, x_n) > f(x'_1, \dots, x'_n)$ (strict-decreasing)
3 $\mathcal{I} \wedge \tau_j \Longrightarrow f(x_1, \dots, x_n) \ge f(x'_1, \dots, x'_n)$ for all j (non-increasing)

Considering conditional invariants give more chances to the solver

But we get a conditional termination proof



$$y \ge 0 \land z < 0$$

$$\land x' = x$$

$$\land y' = y + z$$

$$\land z' = z - 1$$

$$\ell_{0}$$

$$\tau_{0}: true$$

$$\ell_{1}$$

$$y \ge 0 \land z > 0$$

$$\tau_{2}: \land x' = x - z$$

$$\land y' = y + x$$

$$\land z' = z + 1$$

- **z < 0** is a conditional invariant at location ℓ_1
- y is a ranking function
 - **1** au_1 is bounded and strictly decreasing
 - **2** au_2 is disabled

$$y \ge 0 \land z < 0$$

$$\land x' = x$$

$$\land y' = y + z$$

$$\land z' = z - 1$$

$$\ell_0$$

$$\tau_0: true$$

$$\ell_1$$

$$y \ge 0 \land z > 0$$

$$\tau_2: \land x' = x - z$$

$$\land y' = y + x$$

$$\land z' = z + 1$$

We have a conditional proof:

The system terminates if the condition z < 0 holds at l_0 (or τ_0)

$$y \ge 0 \land z < 0$$

$$\land x' = x$$

$$\land y' = y + z$$

$$\land z' = z - 1$$

$$\ell_0$$

$$\tau_0: z < 0$$

$$\ell_1$$

$$\psi \ge 0 \land z > 0$$

$$\tau_2: \land x' = x - z$$

$$\land y' = y + x$$

$$\land z' = z + 1$$

We have a conditional proof:

The system terminates if the condition z < 0 holds at ℓ_0 (or τ_0)

In order to complete the termination proof we have to consider the complementary problem.

<u>Narrow</u> the transitions removing all states that we already now that are terminating.

We can do better than just add the negation of the condition in the entry.



We know more!:

whenever z < 0 holds at ℓ_1 the system terminates



Narrow the transition system according to this:

whenever z < 0 holds at ℓ_1 the system terminates

Assume we have the following transition system:



After sending the problem to our Max-SMT solver we get:



- Conditional invariant I_1 at location l_1 .
- Conditional invariant I_2 at location l_2 .

After sending the problem to our Max-SMT solver we get:



- Conditional invariant I₁ at location l₁.
- Conditional invariant *I*₂ at location *l*₂.
- If I_1 holds in location l_1 then I_2 holds in location l_2 .
- I₂ is preserved in l_2 .
- If I_2 holds in location l_2 then I_1 holds in location l_1 .
- If I_2 holds in l_2 and I_2 holds in l_2 then it terminates.

After sending the problem to our Max-SMT solver we get:



- Conditional invariant I_1 at location l_1 .
- Conditional invariant I_2 at location l_2 .

Therefore

- If I_1 holds in location l_1 we are done.
- If I_2 holds in location l_2 we are done.

After narrowing



Remains to be proved

Therefore

- The entry au_0 is narrowed with $\neg I'_1$
- Transition τ_1 is narrowed with $\neg I_1$ and $\neg I'_2$
- Transition τ_2 is narrowed with $\neg I_2$ and $\neg I'_2$
- **Transition** τ_3 is narrowed with $\neg I_2$ and $\neg I'_1$



Narrow the transition system according to this:

whenever z < 0 holds at ℓ_1 the system terminates

After simplifying the transition system we get:



$$\begin{array}{l} y \geq 0 \wedge z > 0 \\ \tau_2: \quad \wedge x' = x - z \\ \wedge y' = y + x \\ \wedge z' = z + 1 \end{array}$$

After simplifying the transition system we get:



$$y \ge 0 \land z > 0$$

$$\tau_2: \land x' = x - z$$

$$\land y' = y + x$$

$$\land z' = z + 1$$

Conditionally terminates:

- x < 0 is a conditional invariant at location ℓ_1
- y is a ranking function
 - **1** au_2 is bounded and strictly decreasing

Narrowing again with the complement of x < 0 we get:



Narrowing again with the complement of x < 0 we get:



Which terminates with x as a ranking function

Conditional termination provides a natural way of

proving termination by cases

Conditional termination provides a natural way of

proving termination by cases

The Max-SMT solver tries to get a direct proof

Conditional termination provides a natural way of

proving termination by cases

The Max-SMT solver tries to get a direct proof

but if this is not possible (in a given time limit)

it can provide a conditional proof (soft constraints) which give us a progress.

Conditional termination provides a natural way of

proving termination by cases

The Max-SMT solver tries to get a direct proof

but if this is not possible (in a given time limit)

it can provide a conditional proof (soft constraints) which give us a progress.

An additional advantage (key in some case):

If we cannot prove termination of the narrowed transition system

we can use it to try to prove non-termination

as the non-terminating execution (if any) should be there!

1 Introduction

- 2 SMT/Max-SMT solving
- 3 (Conditional) Invariant generation
- 4 Termination analysis
- 5 Compositional termination analysis
- 6 Conclusions and current work

Aim: prove termination in large programs (several consecutive loops). New approach:

- **1** Obtain a conditional termination proof.
- 2 Check (compositionally) the condition as a Safety property.

Simple example:

```
assume(x > y \&\& y \ge 0);
while (y > 0) \{
x = x - 1;
y = y - 1;
}
while (y < 0) \{
y = y + x;
}
```
Scalable Termination Analysis

Aim: prove termination in large programs (several consecutive loops). New approach:

- **1** Obtain a conditional termination proof.
- 2 Check (compositionally) the condition as a Safety property.

Simple example:

```
assume(x > y \&\& y \ge 0);
while (y > 0) \{
x = x - 1;
y = y - 1;
}
assert(x > 0); Rank: -y
while (y < 0) \{
y = y + x;
}
```

Aim: verify termination in large programs (several consecutive loops). **Key ideas:**

- Generate conditional proofs:
 - Find conditional invariants implying termination
- Check the condition as a Safety property of previous loops.

Aim: verify termination in large programs (several consecutive loops). **Key ideas:**

- Generate conditional proofs:
 - Find conditional invariants implying termination
- Check the condition as a Safety property of previous loops.
- In case of failure of the Safety checker Narrow the loop and try again!

Aim: verify termination in large programs (several consecutive loops). **Key ideas:**

Generate conditional proofs:

Find conditional invariants implying termination

- Check the condition as a Safety property of previous loops.
- In case of failure of the Safety checker Narrow the loop and try again!

We can handle every loop (or SCC in general) independently

Our techniques have been implemented in VeryMax(already presented)

These techniques can be highly parallelized (sharing few information).

Compared to last year competitors in TermComp on (335) Integer C programs

Tool	Terminating
AProVE	208(5)
HipTNT+	210(5)
UltimateBuchiAutomizer	207
VeryMax	213

1 Introduction

- 2 SMT/Max-SMT solving
- 3 (Conditional) Invariant generation
- 4 Termination analysis
- 5 Compositional termination analysis
- 6 Conclusions and current work

Conclusions

Two main conclusions:

- Using SMT and Max-SMT, automatic generation of conditional invariants and ranking function becomes feasible.
- In constraint-based program analysis it is often better to consider that we have optimization problems rather than satisfiability problems!

Under development:

- Combine conditional termination and non-termination analysis.
- Use conditional termination to provide witness of termination. For instance, it has applications to check reachability.

Future developments?:

Generate linear upper bounds

Thank you!