### **A Compositional and Monotone Approach to Termination** Towards a more dependable and scalable termination analysis

Shaowei Zhu **Advisor: Zachary Kincaid Princeton University** May 2020

## Background

## program correctness = partial correctness + termination proof

It does the right thing if it terminates

It does terminate



### **Programs as transition systems**

- A program statement defines a transition on the state space S
- Transition system  $\langle S, R \rangle$ 
  - S: state space
  - $R \subseteq S \times S$ : transition relation

### **Ranking functions**

- Loop terminates iff it has a ranking function
- Ranking function r for a loop with state space S and body R
  - $r: S \rightarrow B$  is a mapping from state space S into a set B equipped with a well-founded relation >
  - $(s_1, s_2) \in R \Rightarrow r(s_1) > r(s_2)$
- Examples: linear ranking functions, lexicographic linear ranking functions

### // ranking function: x while (x > 0) { $\mathbf{x} = \mathbf{x} - \mathbf{1};$

// ranking function: < i, j > while (i > 0) { if (j > 0) { j--; } else { = 100;





- Complete procedure for lexicographic ranking function synthesis for a single loop (Gonnord et al., 2015)
- Deciding termination of linear transition systems (Tiwari, 2004)
- Weakest terminating pre-conditions for octagonal relations (Bozga et al., 2012)

- Repeatedly sample a lasso-shaped trace and synthesize ranking function for it (Cook et al., 2006)
- Extend single-loop ranking function synthesis to whole program termination by analyzing paths between cut points (Gonnord et al., 2015)
- Guess ranking function candidates, and then synthesize conditions under which they become real ranking functions (Cook et al., 2008)





### Notivation

#### "Because Infer is *compositional*, it can operate incrementally, quickly producing results on a code diff even if the diff is part of a codebase in the millions of lines"

– Peter O'Hearn

Want: a compositional analysis of termination analyze each loop in isolation and combine the analysis results It brings us: parallelism and scalability, incremental analysis, etc.

### Want: a compositional analysis of termination

# Requires: an analysis that generates sufficient conditions for termination



- Consider a compositional analysis for termination that looks at each loop in isolation
- When analyzing the while loop, we don't know what values could x take
- The analysis need to figure out a sufficient condition for termination and checks if the context implies that condition

if (\_\_VERIFIER\_nondet\_int()  $\neq$  0) { x = 1;} else { x = -1;

// conditions: x > 0 or x < 0while (y < 100 & z < 100) { y = y + x;z = z - x;



### "Tools can understand our programs, but we cannot understand our tools"

– Leino and Moskal

# 

#### widely used in modern tools, but lead to unpredictability in the analysis



### Ultimate Automizer

- #1 tool in Software Verification
   Competition termination category
- Supports linear, lexicographic linear, multiphase, piecewise ranking functions, etc.
- Can prove this program terminates within a few seconds, but ...

```
int main() {
    int x, y, z, a, b;
    int n;
      while (z \ge 0 \& WERIFIER_nondet_int()) 
          z = z - 1;
      }
      if (a = b) {
          while (x \ge 0 \& y \ge 0) {
              y = y - z;
              while (z > 0) {
                  z = z - 1;
                  y = y + 2 * z - x;
              }
              x = x + a - b - 1;
    return 0;
```

### Ultimate Automizer

- #1 tool in Software Verification Competition termination category
- Supports linear, lexicographic linear, multiphase, piecewise ranking functions, etc.
- Can prove this program terminates within a few seconds, but it will not terminate after 30 minutes on this slightly modified program which ought to be "easier" to reason about

15



# Want: a predictable analysis of termination produces a more precise result when given more precise information







This work presents a termination analysis that is:

#### Predictable

- precise analysis results
- Compositional
  - Analyzing on a per-loop basis

Adding information to program leads to more

A monotone and compositional

# analysis for conditional termination

## How we achieve compositionality and predictability?

- <u>composing the results</u>
  - Extend Tarjan's framework to termination analysis
- program
  - Exploit the compositional framework
    - Reduce termination of programs to termination of single loops
    - Access to a summary for the body of each loop  $\bullet$
  - - Abstract transition formulas into linear models that can be handled by Tiwari's method

#### Compositionality: compute termination conditions for a program by analyzing its components and

#### <u>Predictability: give a monotone procedure that computes a sufficient terminating precondition for a</u>

Utilize decision procedure for linear loops to generate terminating pre-conditions for summaries

19



#### An algebraic framework for program analysis

- (Tarjan, 1981) views the control flow graph of the program as a labelled graph where labels are transitions relations
  - Can efficiently compute regular expressions that recognize all paths between two program locations
  - Certain program analysis can be defined using interpretation of regular expressions that represent sets of paths

#### $a \in \Sigma$ $e \in \text{RegExp}(\Sigma) ::= a \mid 0 \mid 1 \mid e_1 + e_2 \mid e_1e_2 \mid e^*$



### Interpreting path expressions as transition formulas

- Transition formula  $\mathbf{TF}$ 
  - A formula in linear integer arithmetic over program variables and primed program variables, and loop counter
  - Describes a transition relation
- Compositional recurrence analysis (Kincaid et al., 2015) interprets path expressions as transition formulas

 $x \in Var$  $n \in \mathbb{Z}$  $y \in \mathsf{BoundVar}$  $t \in \text{Term} ::= x \mid x' \mid y \mid n \mid n \cdot t \mid t_1 + t_2$  $T \in \mathsf{TF} ::= t_1 < t_2 \mid t_1 = t_2 \mid T_1 \land T_2 \mid T_1 \lor T_2 \mid \exists y . T$ 

 $T_1 + {}^{\mathcal{T}} T_2 \triangleq T_1 \lor T_2$  $0^{\mathcal{T}} \triangleq \text{False}$  $1^{\mathscr{T}} \triangleq \bigwedge x' = x$ x∈Var  $T_1 \cdot \mathcal{T}_2 \triangleq \exists \operatorname{Var}'' \cdot T_1[\operatorname{Var}' \mapsto \operatorname{Var}''] \wedge T_2[\operatorname{Var} \mapsto \operatorname{Var}'']$  $T^{*^{\mathscr{T}}} \triangleq$  $\mathcal{T}[[(u, x := e, v)]] \triangleq x' = e \land \qquad \bigwedge \qquad y' = y$ y≠x∈Var  $\mathcal{T}[(u, \operatorname{assume}(c), v)]] \triangleq c \land \bigwedge x' = x$ x∈Var



## An algebraic framework for compositional termination analys

We augment Tarjan's framework with  $\omega$ -regular expressions

- State formula SF
  - formula in LIA over program variables that describes set of program states
- Termination analysis defined as interpretation of  $\omega$ -path expressions
  - atc:  $TF \rightarrow TF$ , approximate transitive closure opera provides the summary of loop body (Kincaid et al., 20)
  - swf:  $TF \rightarrow SF,$  gives a set of initial states from whic the transition is guaranteed to terminate
  - Let p be the set of  $\omega$ -paths starting at program entry, if a state from which there is an non-terminating execution, then  $s \models \mathcal{T}(p)$

22

$$a \in \Sigma$$

$$e \in \operatorname{RegExp}(\Sigma) ::= a \mid 0 \mid 1 \mid e_{1} + e_{2} \mid e_{1}e_{2} \mid e^{*}$$

$$f \in \omega \operatorname{RegExp}(\Sigma) ::= e^{\omega} \mid e; f \mid f_{1} \oplus f_{2}$$

$$S^{\mathcal{T}} \triangleq \operatorname{TF}$$

$$T_{1} + ^{\mathcal{T}}T_{2} \triangleq T_{1} \vee T_{2}$$

$$0^{\mathcal{T}} \triangleq \operatorname{False}$$

$$1^{\mathcal{T}} \triangleq \bigwedge_{x \in \operatorname{Var}} x' = x$$

$$T_{1} \cdot ^{\mathcal{T}}T_{2} \triangleq \exists \operatorname{Var}'' \cdot T_{1} |\operatorname{Var}' \mapsto \operatorname{Var}''| \wedge T_{2} |\operatorname{Var}$$

$$T^{*\mathcal{T}} \triangleq \operatorname{ato}(T)$$

$$S^{\mathcal{T}}[[(u, x := e, v)]] \triangleq x' = e \wedge \bigwedge_{y \neq x \in \operatorname{Var}} y' = y$$

$$S^{\mathcal{T}}[(u, \operatorname{assume}(c), v)]] \triangleq c \wedge \bigwedge_{x \in \operatorname{Var}} x' = x$$

$$L^{\mathcal{T}} \triangleq \operatorname{SF}$$

$$T^{\omega^{\mathcal{T}}} \triangleq \exists \operatorname{Var}' \cdot T \wedge S |\operatorname{Var} \mapsto \operatorname{Var}'|$$

$$S_{1} \oplus ^{\mathcal{T}}S_{2} \triangleq S_{1} \vee S_{2}$$



## How we achieve compositionality and predictability?

- composing the results
  - Extend Tarjan's framework to termination analysis
- <u>program</u>
  - Exploit the compositional framework
    - Reduce termination of programs to termination of single loops (swf\_perator)
    - Access to a summary for the body of each loop (atc operator)
  - - Abstract transition formulas into linear models that can be handled by Tiwari's method

#### Compositionality: compute termination conditions for a program by analyzing its components and



#### Predictability: give a monotone procedure that computes a sufficient terminating precondition for a

Utilize decision procedure for linear loops to generate terminating pre-conditions for summaries



#### Termination analysis of the whole program

Define swf() operator that finds terminating preconditions for a transition formula

Reduce

in a monotone manner

Terminating preconditions for linear transition systems with rational eigenvalues based on (Tiwari, 2004) Compositional Recurrence analysis (Kincaid et al., 2015)

# Tiwari has a way to synthesize conditions for termination, but ...

### **Long-term Dynamics of Linear Transition Systems**

 Suppose the loop to the right does not terminate after k iterations, then what are the values of x, y, yand z?



#### while (x > 0) { X = X + Y $\underline{Y} = \underline{Y} + \underline{Z};$



#### Long-term Dynamics of Linear Transition Systems Suppose the loop to the right does not terminate after k iterations, then what are the values of x, y, y

and z?

• 
$$z^{(k)} = z^{(0)}$$

• 
$$y^{(k)} = y^{(0)} + kz^{(0)}$$

• 
$$x^{(k)} = x^{(0)} + ky^{(0)} + \frac{k(k-1)}{2}z^{(0)}$$

Obviously the value of x affects termination, so we look at

• 
$$x^{(k)} = x^{(0)} + k(y^{(0)} - \frac{1}{2}z^{(0)}) + k^{2}$$

#### while (x > 0) { X = X + V;Y = Y + Z;





### Long-term Dynamics of Linear Transition Systems • $x^{(k)} = x^{(0)} + k(y^{(0)} - \frac{1}{2}z^{(0)}) + k^2\frac{z^{(0)}}{2}$

• Consider when k gets large enough  $k^2 \gg k \gg 1$ 



while (x > 0) { X = X + Y;y = y + z;

28



### Long-term Dynamics of Linear Transition Systems

If the loop never terminates, k can actually get large enough and  $x^{(k)} > 0$  still holds • If  $\frac{z^{(0)}}{2} \neq 0$  then  $x^{(k)} \approx k^2 \frac{z^{(0)}}{2} > 0 \Rightarrow \frac{z^{(0)}}{2} > 0$ • If  $\frac{z^{(0)}}{2} = 0$  but  $y^{(0)} - \frac{1}{2}z^{(0)} \neq 0$ , then  $x^{(k)} \approx k(y^{(0)} - \frac{1}{2}z^{(0)}) > 0 \Rightarrow y^{(0)} - \frac{1}{2}z^{(0)} > 0$ If  $\frac{z^{(0)}}{2} = 0$ ,  $y^{(0)} - \frac{1}{2}z^{(0)} = 0$ , then  $x^{(k)} \approx x^{(0)} > 0$ 

- while (x > 0) { X = X + V;Y = Y + Z;



### Long-term Dynamics of Linear Transition Systems

Thus non-termination implies the following:

• 
$$\frac{z^{(0)}}{2} > 0$$
, or  
•  $\frac{z^{(0)}}{2} = 0 \land y^{(0)} - \frac{1}{2}z^{(0)} > 0$ , or  
•  $\frac{z^{(0)}}{2} = 0 \land y^{(0)} - \frac{1}{2}z^{(0)} = 0 \land x^{(0)}$ 

z < 0 $\vee (z = 0 \land y < 0)$  $\lor (z = 0 \land y = 0 \land x \leq 0)$ while (x > 0) { X = X + V;y = y + z;

#### The negation of the above condition implies termination! (Tiwari, 2004)



### **Dominant Term Analysis (DTA) exploits long-time dynamics**

- Using Tiwari's analysis to generate terminating conditions requires
  - 1. Finding loop guards expressed in linear terms of program variables (polyhedral guards)
  - 2. Linear loop update as matrix multiplication  $\mathbf{x}' = A\mathbf{x}$ 
    - Exists a linear ordering on the eigenvalues of A





### **Dominant Term Analysis (DTA) exploits long-time dynamics**

- Using Tiwari's analysis to generate terminating conditions requires
  - 1. Finding loop guards expressed in linear terms of program variables (polyhedral guards)
  - 2. Linear loop update as matrix multiplication  $\mathbf{x}' = A\mathbf{x}$ 
    - Exists a linear ordering on the eigenvalues of A

Q: When is this possible? What if we have to compare  $e^{\frac{k\pi}{3}i}$  and  $e^{\frac{3k\pi}{4}i}$ ? A: That is hard to do. Want to consider only matrices with real/rational eigenvalues



# Tiwari has a way to synthesize conditions for termination, but ...

Tiwari has a way to synthesize conditions for termination, but only applies to *linear transition systems with polyhedral guards and rational spectra* 

#### Termination analysis of the whole program

Define swf() operator that finds terminating preconditions for a transition formula

Reduce using theory of best abstractions in a monotone manner

> Terminating preconditions for linear transition systems with rational eigenvalues based on (Tiwari, 2004)

Compositional Recurrence analysis (Kincaid et al., 2015)

#### Simulation as abstraction

- Transition system  $A = \langle S_A, R_A \rangle$
- Transition system  $\widetilde{A} = \langle \widetilde{S}_A, \widetilde{R}_A \rangle$
- $p: S_A \to \widetilde{S_A}$  is a simulation from transition system A to transition system A if for all A -transitions  $(a, a') \in R_A$ ,  $(p(a), p(a')) \in R_A$
- Existence of simulation p from A to A has the consequence that termination of Aimplies termination of A
  - If have a terminating condition C for Athen we can get a sufficient terminating condition for A as  $p^{-1}(C)$



36

#### Simulation as abstraction

- Transition system  $A = \langle S_A, R_A \rangle$
- Transition system  $\widetilde{A} = \langle \widetilde{S_A}, \widetilde{R_A} \rangle$
- $p: S_A \to \widetilde{S_A}$  is a simulation from transition system A to transition system A if for all A -transitions  $(a, a') \in R_A$ ,  $(p(a), p(a')) \in R_A$
- Existence of simulation p from A to  $\overline{A}$  has the consequence that termination of Aimplies termination of A
  - If have a terminating condition C for Athen we can get a sufficient terminating condition for A as  $p^{-1}(C)$

while  $(x + y \ge 0)$  { if (\_\_VERIFIER\_nondet\_int()) { x = x - 1;} else { y = y - 1;

Transition formula of A:  $x + y \ge 0 \land ((x' = x - 1 \land y' = y) \lor (y' = y - 1 \land x' = x))$ Simulation  $s: \begin{bmatrix} x \\ y \end{bmatrix} \rightarrow x + y$ Transition formula of  $\widetilde{A}$ :  $t \ge 0 \land t' = t - 1$ 









#### Theory of best abstractions

• A is a best abstraction of Awithin class & with respect to a certain class of simulations, e.g., linear simulations



Class  $\mathscr{C}$  (over-approximating transition systems)



### **Best abstraction and monotonicity**

- We want to get sufficient terminating conditions for transition system A
- Suppose we have a procedure that computes swf for transition systems in a class  $\mathscr{C}$ 
  - Using the best abstraction A and simulation  $p: A \rightarrow A$ , we get sufficient terminating conditions  $p^{-1}(swf(A))$
  - Using some other abstraction *B* and simulation  $q: A \rightarrow B$ , we get sufficient terminating conditions  $q^{-1}(swf(B))$
- Under reasonable conditions, using the best abstraction yields the weakest terminating preconditions, compared to using any other abstraction in  $\mathscr{C}$



#### **Best abstraction** ! simulation $\overline{q}$ Other abstraction B

Class  $\mathscr{C}$  (over-approximating transition systems)





#### Termination analysis of the whole program

Define swf() operator that finds terminating preconditions for a transition formula

Reduce using theory of best abstractions in a monotone manner

> Terminating preconditions for linear transition systems with rational eigenvalues based on (Tiwari, 2004)

Compositional Recurrence analysis (Kincaid et al., 2015)

#### Termination analysis of the whole program



Compositional Recurrence analysis (Kincaid et al., 2015)

Example: Transition formula  $x' = x \land x' = 0$  has abstractions x' = x, x' = 2x, x' = 3x, etc. but does not have a best abstraction

### **Theorem: transition formulas in LIA have best** abstractions in DATS w.r.t. linear simulations

- DATS: deterministic affine transition system
  - State space  $\mathbb{Q}^n$
  - $(\mathbf{u}, \mathbf{v}) \in R \Leftrightarrow \mathbf{v} = T\mathbf{u} + \mathbf{c} \wedge D\mathbf{u} = d$
- cf. TDATS (total deterministic affine transition system)
  - $(\mathbf{u}, \mathbf{v}) \in R \Leftrightarrow \mathbf{v} = P\mathbf{u} + \mathbf{c}$
- Note: transition formulas in LIA does not necessarily have best abstractions in TDATS w.r.t. linear simulations

### Simplifying DATS

- DATS: deterministic affine transition system
  - $(\mathbf{u}, \mathbf{v}) \in R \Leftrightarrow \mathbf{v} = T\mathbf{u} + \mathbf{c} \wedge D\mathbf{u} = d$
- Simplification 1: easy to homogenize this in a DLTS feldeterministic linear transition system)
  - $(\mathbf{u}',\mathbf{v}') \in R \Leftrightarrow \mathbf{v}' = T'\mathbf{u}' \wedge D'\mathbf{u}' = 0$
- study the exponentiated version of  $T_0$  for the termination of DLTS

Define the *domain* of a transition relation as  $dom(R) \triangleq \{x \in S : \exists y . xRy\}$ 

- Define the *invariant domain* of transition relation R as dom<sup>\*</sup>(R)  $\triangleq$   $\int$  dom( $R^n$ )
  - Intuition: initial states outside of dom<sup>\*</sup> are certainly terminating pre-states, so only cares about states inside the invariant domain

Simplification 2: construct a representation matrix  $T_0$  that behaves exactly the same as T on the domain we care about and maps everything else to 0, then

#### Termination analysis of the whole program



Compositional Recurrence analysis (Kincaid et al.)

#### Termination analysis of the whole program



**Compositional Recurrence** analysis (Kincaid et al.)  $(\mathbf{u}, \mathbf{v}) \in R \Leftrightarrow A\mathbf{v} = B\mathbf{u}$  $(\mathbf{u}, \mathbf{v}) \in R \Leftrightarrow \mathbf{v} = T\mathbf{u} \wedge D\mathbf{u} = 0$ Next: spectral theory of DLTS

### Spectral theory of DLTS

- Spectrum of total deterministic transition systems (TDLTS) v = Tu
  - spec(*T*)  $\triangleq \{\lambda \in \mathbb{C} : \exists v, v \neq 0.Tv = \lambda v\}$
- Spectrum of deterministic transition systems (DLTS)  $v = Tu \wedge Du = 0$ 
  - spec $(T, D) \triangleq \{\lambda \in \mathbb{C} : \exists v \in \text{dom}^*(T, D), v \neq 0.Tv = \lambda v\}$
- QDLTS: DLTS with rational spectrum
- Theorem: the representation  $T_0$  for a QDLTS  $v = Tu \wedge Du = 0$  has rational eigenvalues

### Best QDLTS abstraction

- Theorem: any DLTS has a best abstraction as a QDLTS w.r.t. linear simulations
- Proof is constructive and we give an algorithm to compute the best QDLTS abstraction
- Algorithm idea: repeatedly construct DLTS with lower dimensions but remains an over-approximation of the original DLTS, until we obtain a QDLTS or we run out of dimensions





#### Best QDLTS abstraction restricted to invariant domain is what Tiwari needs

- Properties of a QDLTS:
  - Every iteration is a linear update restricted to a linear space  $\mathbf{x}' = T\mathbf{x} \wedge D\mathbf{x} = 0$
  - Spectrum of the above DLTS is all rational

- Another look:
  - Starting from x within the invariant domain, the updates are characterized by  $\mathbf{x}' = T_0 \mathbf{x}$
  - Eigenvalues of  $T_0$  are all rational
- initial states outside of the invariant domain are certainly terminating pre-states



#### Best QDLTS abstraction restricted to invariant domain is what Tiwari needs

- Tiwari needs:
  - Polyhedral guards

Procedure to compute the best polyhedral guards given a transition formula and simulation

Linear updates in the loop body  $\mathbf{x}' = A\mathbf{x}$ 



Update matrix A has rational eigenvalues

- Another look:
  - Starting from x within the invariant domain, the updates are characterized by  $\mathbf{x}' = T_0 \mathbf{x}$
  - Eigenvalues of  $T_0$  are all rational
- initial states outside of the invariant domain are certainly terminating pre-states



#### Termination analysis of the whole program



**Compositional Recurrence** analysis (Kincaid et al.)

Define swf() operator that finds terminating preconditions for a transition formula

> Terminating preconditions for linear transition systems with rational eigenvalues (Tiwari)

Transition formula F

Best DLTS abstraction G

Best QDLTS abstraction Q

Linear simulation S

#### Sufficient preconditions for *F* to terminate

#### Sufficient preconditions for *G* to terminate

#### Linear simulation *T* Monotone

Monotone

### Sufficient preconditions for Q to terminate





## How we achieve compositionality and predictability?

- composing the results
  - Extend Tarjan's framework to termination analysis
- <u>program</u>
  - Exploit the compositional framework
    - Reduce termination of programs to termination of single loops (swf\_perator)
    - Access to a summary for the body of each loop (atc operator)
  - Utilize decision procedure for linear loops to generate terminating pre-conditions for summaries
    - Abstract transition formulas into linear models that can be handled by Tiwari's method

#### Compositionality: compute termination conditions for a program by analyzing its components and



#### Predictability: give a monotone procedure that computes a sufficient terminating precondition for a



#### **Preliminary Experimental Results**

- **SV-COMP** termination benchmarks
  - Not geared towards conditional 0 termination
  - Subset containing only terminating programs
- Conclusions
  - Monotonicity and compositionally 0 come at a cost



![](_page_52_Picture_9.jpeg)

![](_page_52_Picture_10.jpeg)

![](_page_52_Picture_11.jpeg)

### Future work

- Create more conditional termination benchmarks and evaluate the tool on those benchmarks
- Use termination arguments to enhance loop invariant generation
- Monotone conditional termination analysis for other ranking function templates
- Explore how mathematical theory of dynamical systems could further help with program analysis

### References

- [1] B. Cook, A. Podelski, and A. Rybalchenko, "Abstraction Refinement for Termination.," SAS, 2005.
- [2] A. Tiwari, "Termination of Linear Programs.," CAV, 2004.
- POPL, 2009.
- 1981.
- [5] T. W. Reps, S. Sagiv, and G. Yorsh, "Symbolic Implementation of the Best Transformer.," VMCAI, 2004.
- [6] R. Jhala and R. Majumdar, "Software model checking.," ACM Comput. Surv., 2009.
- [7] L. Gonnord, D. Monniaux, and G. Radanne, "Synthesis of Ranking Functions using Extremal Counterexamples," {ACM} SIGPLAN Notices, vol. 50, no. 6, pp. 608–618, Jun. 2015.
- 1995.
- [9] A. Farzan and Z. Kincaid, "Compositional Recurrence Analysis.," FMCAD, 2015.
- [10] K. L. McMillan, "Lazy Abstraction with Interpolants.," CAV, 2006.
- [11] M. Bozga, R. Iosif, and F. Konecný, "Deciding Conditional Termination.," TACAS, 2012.

• [3] C. Calcagno, D. Distefano, P. W. O'Hearn, and H. Yang, "Compositional shape analysis by means of bi-abduction.,"

• [4] R. E. Tarjan, "A Unified Approach to Path Problems," Journal of the ACM (JACM), vol. 28, no. 3, pp. 577–593, Jul.

• [8] T. W. Reps, S. Horwitz, and S. Sagiv, "Precise Interprocedural Dataflow Analysis via Graph Reachability.," POPL,

### Thanks for your attention!