

Challenges in Decomposing Encodings of Verification Problems

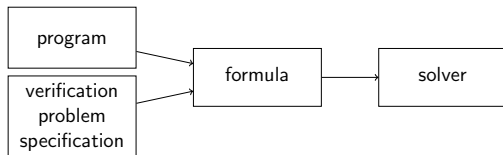
Peter Schrammel



HCVS 2016
Eindhoven, The Netherlands

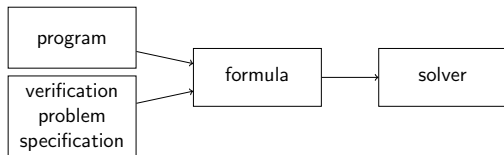
Motivation

Modern software verification tools:



Motivation

Modern software verification tools:

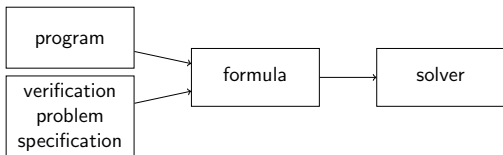


Large programs:

- Problem: Formula too large for existing backend solvers

Motivation

Modern software verification tools:

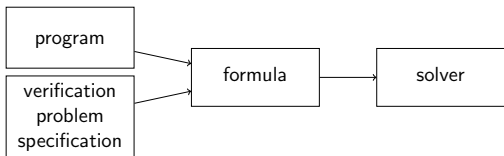


Large programs:

- Problem: Formula too large for existing backend solvers
- Solution: Make formula smaller

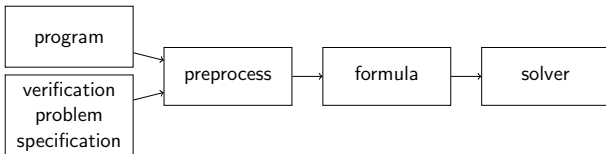
Motivation

Modern software verification tools:



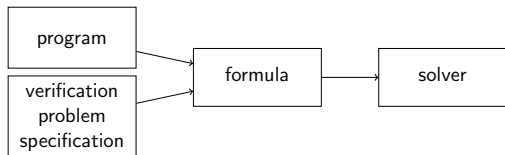
Large programs:

- Problem: Formula too large for existing backend solvers
- Solution: Make formula smaller



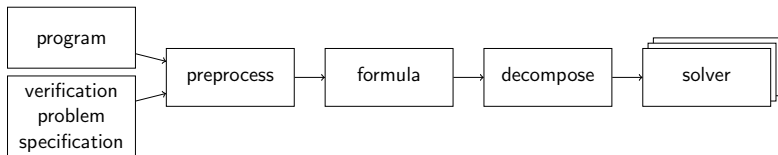
Motivation

Modern software verification tools:



Large programs:

- Problem: Formula too large for existing backend solvers
- Solution: Make formula smaller



Case Studies: Termination Analyses (ASE'15)

Universal termination:

- Result: terminating / potentially non-term. / non-terminating
- Decision problem

Conditional termination:

- Result: sufficient precondition for termination
- Inference problem

Case Studies: Termination Analyses (ASE'15)

Universal termination:

- Result: terminating / potentially non-term. / non-terminating
- Decision problem

Conditional termination:

- Result: sufficient precondition for termination
- Inference problem

Example: Universal Termination Analysis

Encoding of the modular universal termination problem:

$$\begin{aligned}
 & \exists_2 \text{Summary}_{f_1}, \dots, \text{Summary}_{f_n} : \bigwedge_{f \in F} \\
 & \quad \exists_2 \text{Inv}_f, \text{RR}_f : \forall \mathbf{x}^{\text{in}}_f, \mathbf{x}_f, \mathbf{x}'_f, \mathbf{x}^{\text{out}}_f : \\
 & \quad \quad \text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \implies \text{Inv}_f(\mathbf{x}_f) \\
 & \quad \wedge \text{Inv}_f(\mathbf{x}_f) \wedge \text{Trans}_f(\mathbf{x}_f, \mathbf{x}'_f) \wedge \bigwedge_{h_i \in H_f} \text{Summary}_h(\mathbf{x}^{\text{p-in}}_{h_i}, \mathbf{x}^{\text{p-out}}_{h_i}) \\
 & \quad \quad \implies \text{Inv}_f(\mathbf{x}'_f) \wedge \text{RR}_f(\mathbf{x}_f, \mathbf{x}'_f) \\
 & \quad \wedge \text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \wedge \text{Inv}_f(\mathbf{x}'_f) \wedge \text{Out}_f(\mathbf{x}'_f, \mathbf{x}^{\text{out}}_f) \\
 & \quad \quad \implies \text{Summary}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}^{\text{out}}_f)
 \end{aligned}$$

Decomposition:

- Procedural, top-down, context-sensitive

Example: Universal Termination Analysis

Benchmarks:

- Product line benchmarks from SV-COMP (597 benchmarks, 1.6 MLOC)
- Non-trivial procedural structure (on average 67 procedures, 5.5 loops)

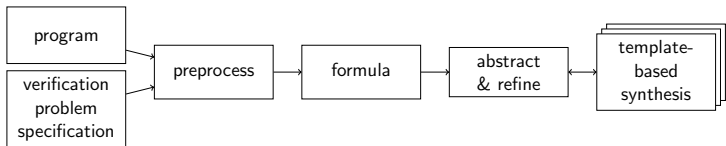
Results:

	expected	2LS IPTA	2LS MTA	TAN	Ultimate
terminating	264	249	26	18	50
non-terminating	333	320	333	3	324
potentially non-terminating	—	14	1	425	0
timed out (0.5h)	—	14	237	150	43
errors	—	0	0	1	180
total run time (h)	—	58.7	119.6	92.8	23.9

2LS for Program Analysis

<http://www.cprover.org/2LS>

- Verification and static analysis on logical formulae
- Approximates solution to 2OL by reduction to FOL



- Bit-precise analysis
(including floating-point arithmetic)
- Template-based synthesis
(using strategy iteration for optimisation)

SV-COMP'16 

Analysis features:

- Interprocedural static analysis, termination analysis
- Incremental BMC, k -induction, $k|k|$ (SAS'15)

Logical Specification of Verification Problems

Safety verification:

$$\begin{aligned} \exists_2 Inv. \quad & \forall \mathbf{x}^{in}, \mathbf{x}, \mathbf{x}'. \\ & (Init(\mathbf{x}^{in}, \mathbf{x}) \implies Inv(\mathbf{x})) \wedge \\ & (Inv(\mathbf{x}) \wedge Trans(\mathbf{x}, \mathbf{x}') \implies Inv(\mathbf{x}')) \wedge \\ & (Inv(\mathbf{x}) \implies \neg Err(\mathbf{x})) \end{aligned}$$

(Blass and Gurevich '87, Grebenshchikov et al '12, David et al '15, ...)

Logical Specification of Verification Problems

Safety verification:

$$\begin{aligned} \exists_2 Inv. \quad & \forall \mathbf{x}^{in}, \mathbf{x}. \\ & (Init(\mathbf{x}^{in}, \mathbf{x}) \implies Inv(\mathbf{x})) \wedge \\ & (Inv(\mathbf{x}) \wedge Trans(\mathbf{x}, \mathbf{x}') \implies Inv(\mathbf{x}')) \wedge \\ & (Inv(\mathbf{x}) \implies \neg Err(\mathbf{x})) \end{aligned}$$

Invariant inference:

$$\begin{aligned} \text{min } Inv. \quad & \forall \mathbf{x}, \mathbf{x}'. \\ & (Init(\mathbf{x}) \implies Inv(\mathbf{x})) \wedge \\ & (Inv(\mathbf{x}) \wedge Trans(\mathbf{x}, \mathbf{x}') \implies Inv(\mathbf{x}')) \end{aligned}$$

(Blass and Gurevich '87, Grebenschikov et al '12, David et al '15, ...)

Logical Specification of Verification Problems

Safety verification:

$$\begin{aligned} \exists_2 \text{Inv. } \forall \mathbf{x}^{in}, \mathbf{x}, \mathbf{x}'. \\ & (\text{Init}(\mathbf{x}^{in}, \mathbf{x}) \implies \text{Inv}(\mathbf{x})) \wedge \\ & (\text{Inv}(\mathbf{x}) \wedge \text{Trans}(\mathbf{x}, \mathbf{x}') \implies \text{Inv}(\mathbf{x}')) \wedge \\ & (\text{Inv}(\mathbf{x}) \implies \neg \text{Err}(\mathbf{x})) \end{aligned}$$

Invariant inference:

$$\begin{aligned} \text{min Inv. } \forall \mathbf{x}, \mathbf{x}'. \\ & (\text{Init}(\mathbf{x}) \implies \text{Inv}(\mathbf{x})) \wedge \\ & (\text{Inv}(\mathbf{x}) \wedge \text{Trans}(\mathbf{x}, \mathbf{x}') \implies \text{Inv}(\mathbf{x}')) \end{aligned}$$

Termination verification:

$$\begin{aligned} \exists_2 \text{RR, Inv. } \forall \mathbf{x}, \mathbf{x}'. \\ & (\text{Init}(\mathbf{x}) \implies \text{Inv}(\mathbf{x})) \wedge \\ & (\text{Inv}(\mathbf{x}) \wedge \text{Trans}(\mathbf{x}, \mathbf{x}') \implies \text{Inv}(\mathbf{x}') \wedge \text{RR}(\mathbf{x}, \mathbf{x}')) \end{aligned}$$

...

(Blass and Gurevich '87, Grebenshchikov et al '12, David et al '15, ...)

Template-Based Synthesis

Reduction to first-order logic via templates, e.g. safety verification:

$$\begin{aligned} \exists_2 Inv. \forall \mathbf{x}, \mathbf{x}' 1. & \quad (Init(\mathbf{x}) \implies Inv(\mathbf{x})) \wedge \\ & \quad (Inv(\mathbf{x}) \wedge Trans(\mathbf{x}, \mathbf{x}') \implies Inv(\mathbf{x}')) \wedge \\ & \quad (Inv(\mathbf{x}) \implies \neg Err(\mathbf{x})) \end{aligned}$$

Template-Based Synthesis

Reduction to first-order logic via templates, e.g. safety verification:

$$\begin{aligned} \exists \mathbf{d}. \quad \forall \mathbf{x}, \mathbf{x}'. \quad & (Init(\mathbf{x}) \implies \mathcal{T}(\mathbf{x}, \mathbf{d})) \wedge \\ & (\mathcal{T}(\mathbf{x}, \mathbf{d}) \wedge Trans(\mathbf{x}, \mathbf{x}') \implies \mathcal{T}(\mathbf{x}', \mathbf{d})) \wedge \\ & (\mathcal{T}(\mathbf{x}, \mathbf{d}) \implies \neg Err(\mathbf{x})) \end{aligned}$$

where \mathbf{d} are template parameters.

(Graf & Saïdi CAV'97, ..., Reps et al, ... Brauer et al, ..., Srivastava et al, ...)

Template-Based Synthesis

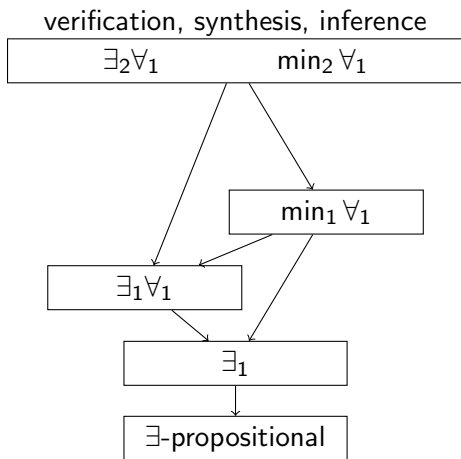
Reduction to first-order logic via templates, e.g. invariant inference:

$$\begin{aligned} \min \mathbf{d}. \quad & \forall \mathbf{x}, \mathbf{x}'. \quad (\text{Init}(\mathbf{x}) \implies \mathcal{T}(\mathbf{x}, \mathbf{d})) \wedge \\ & (\mathcal{T}(\mathbf{x}, \mathbf{d}) \wedge \text{Trans}(\mathbf{x}, \mathbf{x}') \implies \mathcal{T}(\mathbf{x}', \mathbf{d})) \end{aligned}$$

where \mathbf{d} are template parameters.

(Graf & Saïdi CAV'97, ..., Reps et al, ... Brauer et al, ..., Srivastava et al, ...)

Solver Hierarchy



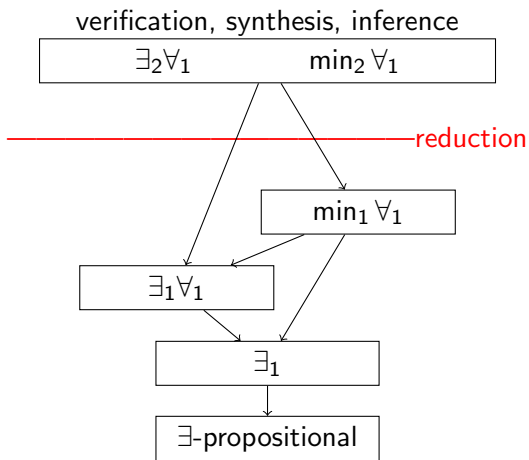
Eldarica, Spacer, ...

Symba, MathSAT-opt, ...

} CVC4, Z3, MathSAT, ...

MiniSAT, Glucose, ...

Solver Hierarchy



Eldarica, Spacer, ...

Symba, MathSAT-opt, ...

} CVC4, Z3, MathSAT, ...

MiniSAT, Glucose, ...

Program Encoding

Non-recursive programs with multiple procedures

Procedure f : ($Init(\mathbf{x}^{in}, \mathbf{x})$, $Trans(\mathbf{x}, \mathbf{x}')$, $Out(\mathbf{x}, \mathbf{x}^{out})$)

Program Encoding

Non-recursive programs with multiple procedures

Procedure f : ($Init(x^{in}, x)$, $Trans(x, x')$, $Out(x, x^{out})$)

```

unsigned f(unsigned z) {
  unsigned w = 0;                 $w_0 = 0$ 
  if(z>0)                         $\wedge g_4 = z > 0$ 
    w = h(z);                     $\wedge h_0(z, r_{h_0}) \wedge w_1 = r_{h_0}$ 
                                 $\wedge w_2^\phi = g_4 ? w_1 : w_0$ 
  return w;                       $\wedge r_h = x_1^\phi$ 
}

unsigned h(unsigned y) {
  unsigned x;                     $g_0 = true$ 
  for (x=0;                        $\wedge x_0 = 0$ 
      x<10;                         $\wedge g_1 = g_0 \wedge x_1^\phi = (ls_3 ? x_3^{lb} : x_0)$ 
      x+=y);                        $\wedge g_2 = (x_1^\phi < 10 \wedge g_1)$ 
  return x;                        $\wedge x_2 = x_1^\phi + y$ 
                                 $\wedge r_h = x_1^\phi$ 
}

```

Program Encoding

Non-recursive programs with multiple procedures

Procedure f : ($Init(x^{in}, x)$, $Trans(x, x')$, $Out(x, x^{out})$)

unsigned f(unsigned z) {	
unsigned w = 0;	$w_0 = 0$
if (z>0)	$\wedge g_4 = z > 0$
w = h(z);	$\wedge h_0(z, r_{h_0}) \wedge w_1 = r_{h_0}$
	$\wedge w_2^\phi = g_4 ? w_1 : w_0$
return w;	$\wedge r_h = x_1^\phi$
}	
unsigned h(unsigned y) {	
unsigned x;	$g_0 = true$
for (x=0;	$\wedge x_0 = 0$
	$\wedge g_1 = g_0 \wedge x_1^\phi = (ls_3 ? x_3^{lb} : x_0)$
x<10;	$\wedge g_2 = (x_1^\phi < 10 \wedge g_1)$
x+=y);	$\wedge x_2 = x_1^\phi + y$
return x;	$\wedge r_h = x_1^\phi$
}	

Example: Modular Universal Termination Problem

$$\begin{aligned}
 & \exists_2 \text{Summary}_{f_1}, \dots, \text{Summary}_{f_n} : \bigwedge_{f \in F} \\
 & \quad \exists_2 \text{Inv}_f, \text{RR}_f : \forall \mathbf{x}^{\text{in}}_f, \mathbf{x}_f, \mathbf{x}'_f, \mathbf{x}^{\text{out}}_f : \\
 & \quad \quad \text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \implies \text{Inv}_f(\mathbf{x}_f) \\
 & \quad \wedge \text{Inv}_f(\mathbf{x}_f) \wedge \text{Trans}_f(\mathbf{x}_f, \mathbf{x}'_f) \wedge \bigwedge_{h_i \in H_f} \text{Summary}_{h_i}(\mathbf{x}^{\text{p-in}}_{h_i}, \mathbf{x}^{\text{p-out}}_{h_i}) \\
 & \quad \quad \implies \text{Inv}_f(\mathbf{x}'_f) \wedge \text{RR}_f(\mathbf{x}_f, \mathbf{x}'_f) \\
 & \quad \wedge \text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \wedge \text{Inv}_f(\mathbf{x}'_f) \wedge \text{Out}_f(\mathbf{x}'_f, \mathbf{x}^{\text{out}}_f) \\
 & \quad \quad \implies \text{Summary}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}^{\text{out}}_f)
 \end{aligned}$$

Decomposition into a sequence of subproblems

- Classical approach: Follow the call graph top-down

Decomposition

Soundness of the decomposition by

- Soundness of the individual subproblems
- Soundness of the combination of the subproblem results
- Induction over the recursive decomposition algorithm

Decomposition

Soundness of the decomposition by

- Soundness of the individual subproblems
(e.g. calling contexts, summaries, ...)
- Soundness of the combination of the subproblem results
(e.g. joins, fixed points, upwards propagation, ...)
- Induction over the recursive decomposition algorithm
(e.g. call graph traversal)

Decomposition

Soundness of the decomposition by

- Soundness of the individual subproblems (e.g. calling contexts, summaries, ...)
- Soundness of the combination of the subproblem results (e.g. joins, fixed points, upwards propagation, ...)
- Induction over the recursive decomposition algorithm (e.g. call graph traversal)

Problem

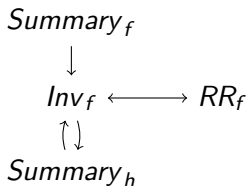
- Cyclically dependent predicates

Example: Universal Termination

$$\begin{aligned}
 & \exists_2 \text{Summary}_{f_1}, \dots, \text{Summary}_{f_n} : \bigwedge_{f \in F} \\
 & \quad \exists_2 \text{Inv}_f, \text{RR}_f : \forall \mathbf{x}^{\text{in}}_f, \mathbf{x}_f, \mathbf{x}'_f, \mathbf{x}^{\text{out}}_f : \\
 & \quad \quad \text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \implies \text{Inv}_f(\mathbf{x}_f) \\
 & \quad \wedge \text{Inv}_f(\mathbf{x}_f) \wedge \text{Trans}_f(\mathbf{x}_f, \mathbf{x}'_f) \wedge \bigwedge_{h_i \in H_f} \text{Summary}_h(\mathbf{x}^{\text{p-in}}_{h_i}, \mathbf{x}^{\text{p-out}}_{h_i}) \\
 & \quad \quad \implies \text{Inv}_f(\mathbf{x}'_f) \wedge \text{RR}_f(\mathbf{x}_f, \mathbf{x}'_f) \\
 & \quad \wedge \text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \wedge \text{Inv}_f(\mathbf{x}'_f) \wedge \text{Out}_f(\mathbf{x}'_f, \mathbf{x}^{\text{out}}_f) \\
 & \quad \quad \implies \text{Summary}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}^{\text{out}}_f)
 \end{aligned}$$

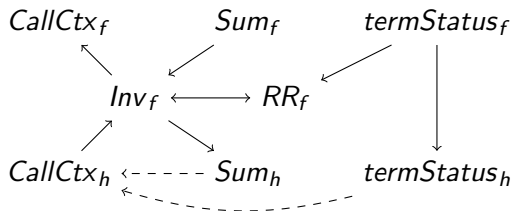
Example: Universal Termination

Cyclically dependent predicates:



Example: Universal Termination

Cyclically dependent predicates:



Example: Universal Termination

$$\begin{aligned}
 & \exists_2 \text{Summary}_{f_1}, \dots, \text{Summary}_{f_n} : \bigwedge_{f \in F} \\
 & \quad \exists_2 \text{Inv}_f, \text{RR}_f : \forall \mathbf{x}^{\text{in}}_f, \mathbf{x}_f, \mathbf{x}'_f, \mathbf{x}^{\text{out}}_f : \\
 & \quad \quad \text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \implies \text{Inv}_f(\mathbf{x}_f) \\
 & \quad \wedge \text{Inv}_f(\mathbf{x}_f) \wedge \text{Trans}_f(\mathbf{x}_f, \mathbf{x}'_f) \wedge \bigwedge_{h_i \in H_f} \text{Summary}_h(\mathbf{x}^{\text{p-in}}_{h_i}, \mathbf{x}^{\text{p-out}}_{h_i}) \\
 & \quad \quad \implies \text{Inv}_f(\mathbf{x}'_f) \wedge \text{RR}_f(\mathbf{x}_f, \mathbf{x}'_f) \\
 & \quad \wedge \text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \wedge \text{Inv}_f(\mathbf{x}'_f) \wedge \text{Out}_f(\mathbf{x}'_f, \mathbf{x}^{\text{out}}_f) \\
 & \quad \quad \implies \text{Summary}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}^{\text{out}}_f)
 \end{aligned}$$

Example: Universal Termination

Summary (callee's perspective):

$\exists_2 \text{Summary}_{f_1}, \dots, \text{Summary}_{f_n} : \bigwedge_{f \in F}$

For each f : $\exists_2 \text{Sum}_f :$

$\exists_2 \text{Inv}_f, \text{RR}_f : \forall \mathbf{x}^{\text{in}}_f, \mathbf{x}_f, \mathbf{x}'_f, \mathbf{x}^{\text{out}}_f :$

$\text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \implies \text{Inv}_f(\mathbf{x}_f)$

$\wedge \text{Inv}_f(\mathbf{x}_f) \wedge \text{Trans}_f(\mathbf{x}_f, \mathbf{x}'_f) \wedge \bigwedge_{h_i \in H_f} \text{Sum}_h(\mathbf{x}^{\text{p-in}}_{h_i}, \mathbf{x}^{\text{p-out}}_{h_i})$
 $\implies \text{Inv}_f(\mathbf{x}'_f) \wedge \text{RR}_f(\mathbf{x}_f, \mathbf{x}'_f)$

$\wedge \text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \wedge \text{Inv}_f(\mathbf{x}'_f) \wedge \text{Out}_f(\mathbf{x}'_f, \mathbf{x}^{\text{out}}_f)$
 $\implies \text{Sum}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}^{\text{out}}_f)$

Example: Universal Termination

Summary (callee's perspective):

$\exists_2 \text{Summary}_{f_1}, \dots, \text{Summary}_{f_n} : \bigwedge_{f \in F}$

For each f : Given $\text{CallCtx}_f : \exists_2 \text{Sum}_f :$

$\exists_2 \text{Inv}_f, \text{RR}_f : \forall \mathbf{x}^{\text{in}}_f, \mathbf{x}_f, \mathbf{x}'_f, \mathbf{x}^{\text{out}}_f :$

$\text{CallCtx}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}^{\text{out}}_f) \wedge \text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \implies \text{Inv}_f(\mathbf{x}_f)$

$\wedge \text{Inv}_f(\mathbf{x}_f) \wedge \text{Trans}_f(\mathbf{x}_f, \mathbf{x}'_f) \wedge \bigwedge_{h_i \in H_f} \text{Sum}_h(\mathbf{x}^{\text{p-in}}_{h_i}, \mathbf{x}^{\text{p-out}}_{h_i})$
 $\implies \text{Inv}_f(\mathbf{x}'_f) \wedge \text{RR}_f(\mathbf{x}_f, \mathbf{x}'_f)$

$\wedge \text{CallCtx}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}^{\text{out}}_f) \wedge \text{Init}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}_f) \wedge \text{Inv}_f(\mathbf{x}'_f) \wedge \text{Out}_f(\mathbf{x}'_f, \mathbf{x}^{\text{out}}_f)$
 $\implies \text{Sum}_f(\mathbf{x}^{\text{in}}_f, \mathbf{x}^{\text{out}}_f)$

Example: Universal Termination

Calling context (caller's perspective):

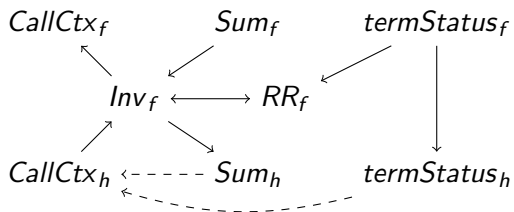
For each f : Given $CallCtx_f$: For each $h_j \in H_f$: $\exists_2 CallCtx_{h_j}$:

$\exists_2 Inv_f : \forall \mathbf{x}^{in}_f, \mathbf{x}_f, \mathbf{x}'_f, \mathbf{x}^{out}_f :$

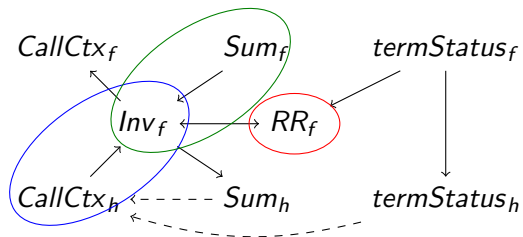
$CallCtx_f(\mathbf{x}^{in}_f, \mathbf{x}^{out}_f) \wedge Init_f(\mathbf{x}^{in}_f, \mathbf{x}_f) \implies Inv_f(\mathbf{x}_f)$

$\wedge Inv_f(\mathbf{x}_f) \wedge Trans_f(\mathbf{x}_f, \mathbf{x}'_f) \wedge \bigwedge_{h_i \in H_f} Sum_h(\mathbf{x}^{p-in}_{h_i}, \mathbf{x}^{p-out}_{h_i})$
 $\implies Inv_f(\mathbf{x}'_f) \wedge CallCtx_{h_j}(\mathbf{x}^{p-in}_{h_j}, \mathbf{x}^{p-out}_{h_j})$

Example: Universal Termination



Example: Universal Termination



- ① Calling contexts of procedure calls h in f
- ② Recurse
- ③ Invariants and summary of procedure f
- ④ Termination argument for procedure f
- ⑤ Determine termination status of f

Example 2: Conditional Termination

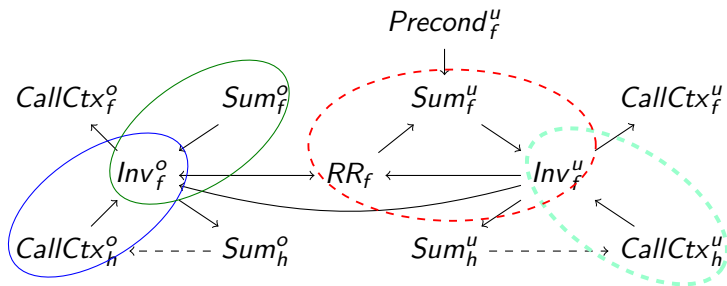
Universal termination:

- Result: terminating / potentially non-term. / non-terminating
- Decision problem

Conditional termination:

- Result: sufficient precondition for termination
- Inference problem

Sufficient Preconditions for Termination



over-approximations

under-approximations

Lessons Learned

2LS for program analysis <http://www.cprover.org/2LS>

- Modular analyses and verification algorithms as decompositions of large formulae
- Predicate inference (abstract interpretation, synthesis, optimisation)
- Should expose solver interface

Observations:

- Decomposition increases scalability (smaller formulae, parallelise)
- Decomposition introduces abstraction
- Decomposition does not eliminate fixed points
- Subproblems of decision problems may be inference problems
- Syntax is a bad criterion for decomposition.

Lessons Learned

Criteria for decomposition?

- Syntax
 - eliminate syntactic bias (control where to cut)
- Abstract domains
 - (e.g. invariants vs ranking functions, lfp vs gfp)
- Predicate interdependencies
 - avoid cutting cycles (still need fixed point, lfp vs gfp)
- Precision / solving capacity-driven (inference problems)
 - As much as possible, as little as necessary
- Property-driven (decision problems)
 - Decomposition = abstraction (start rough and refine)

Lessons Learned

Challenges

- Dynamic decomposition
decomposing upfront is difficult
- Better predicate inference algorithms
(product domains, under-approximations, lfp+gfp, ...)
- Decomposition of cycles?
→ precise handling of recursive functions
- Precision / solving capacity-driven (inference problems)
→ best-effort refinement
- Property-driven (decision problems)
→ modular refinement algorithms