

Verification of Imperative Programs through Transformation of Constraint Logic Programs

Emanuele De Angelis¹, Fabio Fioravanti¹,
Alberto Pettorossi², and Maurizio Proietti³

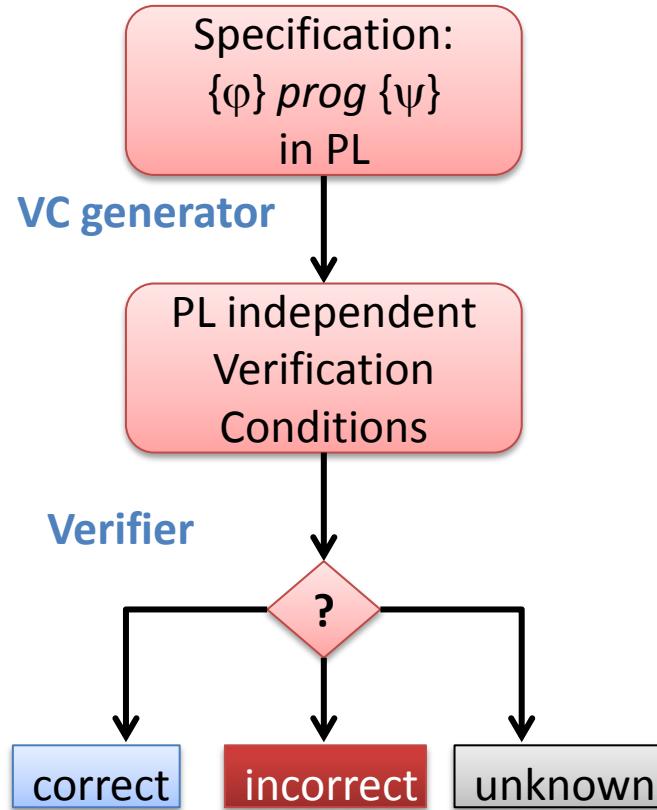
¹University of Chieti-Pescara ‘G. d’Annunzio’, Italy

²University of Rome ‘Tor Vergata’, Italy

³CNR - Istituto di Analisi dei Sistemi ed Informatica, Rome, Italy

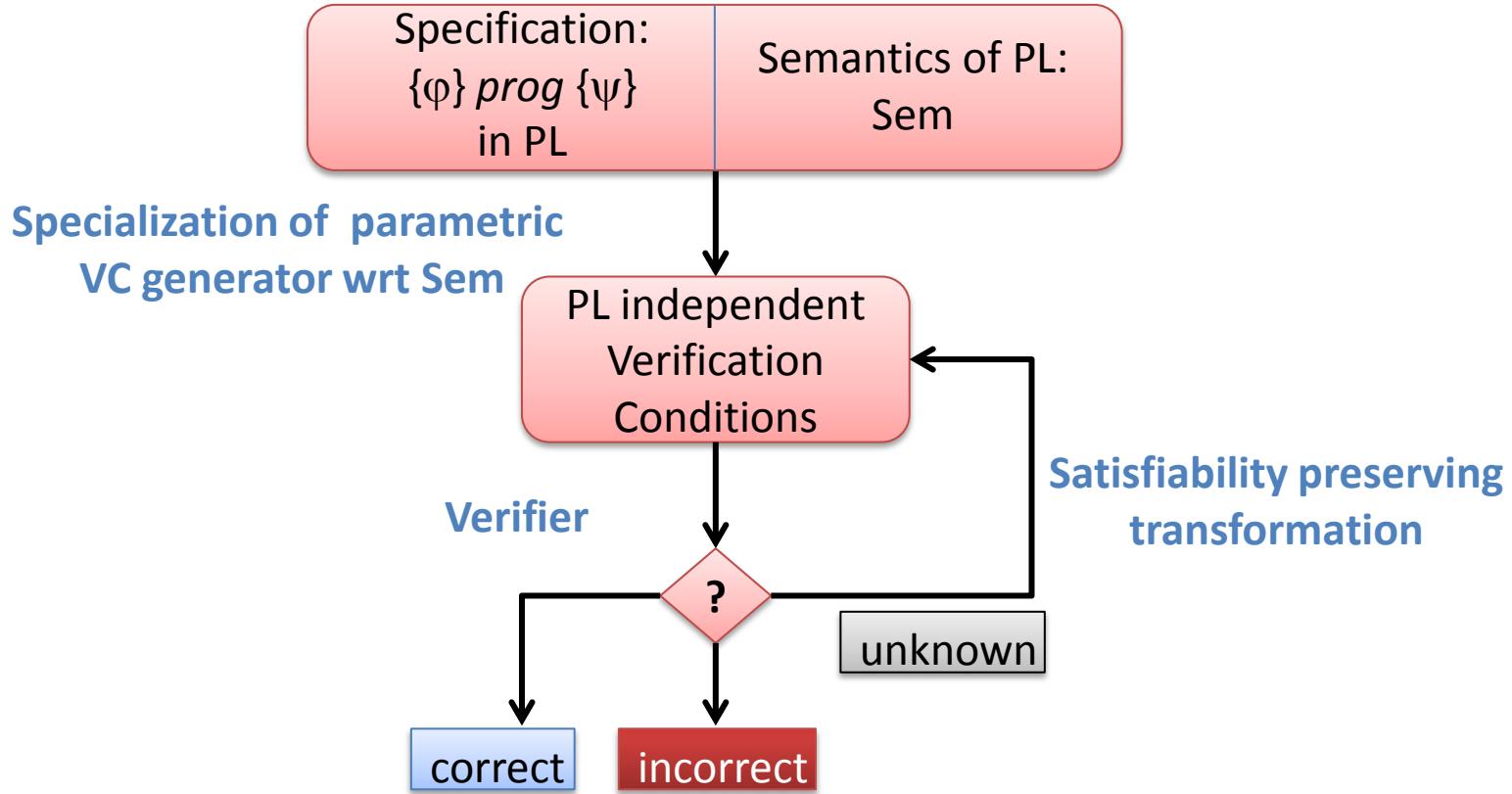
Milano, 25 Settembre 2014

Program Verification



- The VC generator has to be rewritten when PL is changed
- The work done is lost if the answer is “unknown”

Transformation-based Verification



- The VC generator is parametric wrt semantics of PL
- The work done is kept even if the answer is “unknown”

Parametric verification via CLP transformations

- Constraint Logic Programming (CLP) as a *metalanguage* for representing the semantics of PL
- The verification method based on CLP transformation is *parametric* with respect to:
 - programming language and its operational semantics
 - properties and proof rules
 - theory of data structures.
- The input and the output of transformation are semantically equivalent CLP programs. This allows:
 - *composition* of verification tasks
 - *iteration* for refining verification
 - *interoperation* with other verifiers that use CLP (Horn clause) format.

Outline of the Talk

- CLP representation of
 - the imperative program
 - the semantics of the imperative language (*interpreter*)
 - the property to be verified
- Verification method based on CLP program transformation
 - *Transformation rules and strategies* that preserve the semantics of CLP
 - *VC generation* by specialization of the interpreter
 - *VC transformation* by propagation of the property to be verified
- Improving precision via *iterated* VC transformation
- Experimental evaluation: The VeriMAP system

CLP as a metalanguage for imperative programs

CLP with integer constraints

- A CLP *clause* is an implication $c \wedge G \rightarrow H$, written as:

$H :- c, G.$

where H is an atom, c is a constraint, and G is a conjunction of atoms

- A *constraint* is a conjunction of linear equalities/inequalities over integers ($p_1 = p_2$, $p_1 \geq p_2$, $p_1 > p_2$)
- A CLP *program* is a set of CLP clauses
- Semantics: *least model* of the program with the fixed interpretation of constraints.

Imperative Programs over Integers

- We consider an imperative language with integer variables, assignment, if-else, while-loop, and goto.
- Program *increase*:

```
while(x < n){  
    x=x+1;  
    y=x+y;  
}
```

- Partial Correctness Specification

$$\{x = 0 \wedge y = 0\} \text{ increase } \{x \leq y\}$$

Encoding of an Imperative Program into CLP

A program is represented as a set of atoms $\text{at}(label, command)$.

Program *increase*:

```
 $\ell_0$  : while(x < n){  
 $\ell_1$  :     x=x+1;  
 $\ell_2$  :     y=x+y;  
 $\ell_3$  : }
```

CLP encoding of *increase*:

```
at( $\ell_0$ , ite(less(int(x), int(n)),  $\ell_1$ ,  $\ell_h$ )).  
at( $\ell_1$ , asgn(int(x), plus(int(x), int(1)))).  
at( $\ell_2$ , asgn(int(y), plus(int(x), int(y)))).  
at( $\ell_3$ , goto( $\ell_0$ )).  
at( $\ell_h$ , halt).
```

A *transition semantics* is defined by:

- a set of *configurations*, i.e., a CLP term: $\text{cf}(C, S)$

where:

- C is a labeled *command*
- S is a *store*,
i.e., a list of [variable identifier, value] pairs:

$[[\text{int}(x), 2], [\text{int}(y), 3]]$

- a *transition relation*: $\text{tr}(\text{cf}(C, S), \text{cf}(C_1, S_1))$

L: $\text{Id} = \text{Expr}$	$\text{tr}(\text{cf}(\text{cmd}(L, \text{asgn}(\text{Id}, \text{Expr})), S), \text{cf}(\text{cmd}(L_1, C_1), S_1)) :-$ $\quad \text{aeval}(\text{Expr}, S, V), \quad \text{evaluate expression}$ $\quad \text{update}(\text{Id}, V, S, S_1), \quad \text{update store}$ $\quad \text{nextlabel}(L, L_1), \quad \text{next label}$ $\quad \text{at}(L_1, C_1). \quad \text{next command}$
L: $\text{if } (\text{Expr}) \{$ $\quad \text{goto } L_1;$ $\} \text{ else}$ $\quad \text{goto } L_2$ $\}$	$\text{tr}(\text{cf}(\text{cmd}(L, \text{ite}(\text{Expr}, L_1, L_2)), S), \text{cf}(C, S)) :-$ $\quad \text{beval}(\text{Expr}, S), \quad \text{expression is true}$ $\quad \text{at}(L_1, C). \quad \text{next command}$ $\text{tr}(\text{cf}(\text{cmd}(L, \text{ite}(\text{Expr}, L_1, L_2)), S), \text{cf}(C, S)) :-$ $\quad \text{beval}(\text{not}(\text{Expr}), S), \quad \text{expression is false}$ $\quad \text{at}(L_2, C). \quad \text{next command}$
L: $\text{goto } L_1$	$\text{tr}(\text{cf}(\text{cmd}(L, \text{goto}(L_1)), S), \text{cf}(C, S)) :- \text{at}(L_1, C).$ $\quad \text{at}(L_1, C). \quad \text{next command}$

CLP encoding of (in)correctness

Given the specification $\{\varphi_{init}\} \text{ prog } \{\psi\}$ define $\varphi_{error} \equiv \neg\psi$

Definition (Program Incorrectness)

A program *prog* is *incorrect* w.r.t. φ_{init} and φ_{error} if from an initial configuration satisfying φ_{init} it is possible to reach a final configuration satisfying φ_{error} .

Otherwise, program *prog* is *correct*.

Definition (CLP encoding of incorrectness: The interpreter *Int*)

incorrect :- initConf(X), reach(X).

reach(X) :- tr(X,Y), reach(Y).

reach(X) :- errorConf(X).

initConf(X) \equiv X is a configuration satisfying φ_{init}

errorConf(X) \equiv X is a configuration satisfying φ_{error}

reachability

Theorem (Correctness of Encoding)

prog is correct iff *incorrect* $\notin M(\text{Int})$ (the least model of *Int*)

Partial Correctness Specification

$\{x = 0 \wedge y = 0\}$	φ_{init}
<i>increase</i>	
$\{x \leq y\}$	ψ
$\{x > y\}$	$\varphi_{error} \equiv \neg \psi$

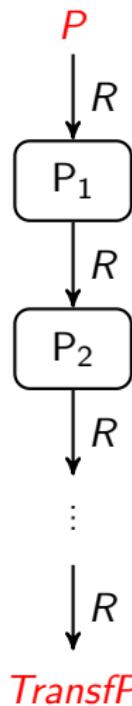
Initial and Error Configurations

```
initConf(cf(cmd(0,ite(...)), [[int(x),X],[int(y),Y],[int(n),N]]))  
      :- X=0, Y=0.           $\varphi_{init}$   
errorConf(cf(cmd(h,halt), [[int(x),X],[int(y),Y],[int(n),N]]))  
      :- X>Y.             $\varphi_{error}$ 
```

Transformation of CLP

Unfold/Fold Program Transformation

[Burstall-Darlington 77, Tamaki-Sato 84, Etalle-Gabrielli 96]



- transformation *rules*:
 $R \in \{ \text{Definition, Unfolding, Folding, Clause Removal} \}$
- the transformation rules *preserve the least model*:
$$\text{incorrect} \in M(P) \text{ iff } \text{incorrect} \in M(TransfP)$$
- the rules must be guided by a *strategy*.

Rules for Transforming CLP Programs

R1. **Definition.** Introducing a new predicate

```
newp(X) :- c, A
```

Rules for Transforming CLP Programs

R1. **Definition.** Introducing a new predicate

`newp(X) :- c, A`

R2. **Unfolding.** A symbolic evaluation step (resolution)

given $H :- c, \underline{A}, G$

$\underline{A} :- d_1, G_1, \dots, \underline{A} :- d_m, G_m$

derive $H :- c, d_1, G_1, G, \dots, H :- c, d_m, G_m, G$

Rules for Transforming CLP Programs

R1. **Definition.** Introducing a new predicate

$\text{newp}(X) :- c, A$

R2. **Unfolding.** A symbolic evaluation step (resolution)

given $H :- c, \underline{A}, G$
 $\underline{A} :- d_1, G_1, \dots, \underline{A} :- d_m, G_m$

derive $H :- c, d_1, G_1, G, \dots, H :- c, d_m, G_m, G$

R3. **Folding.** Matching the body of a predicate definition

given $H :- d, \underline{A}, G$
 $\text{newp}(X) :- c, \underline{A}$ and $d \rightarrow c$

derive $H :- d, \text{newp}(X), G$

Rules for Transforming CLP Programs

R1. Definition. Introducing a new predicate

`newp(X) :- c, A`

R2. Unfolding. A symbolic evaluation step (resolution)

given $H :- c, \underline{A}, G$
 $\underline{A} :- d_1, G_1, \dots, \underline{A} :- d_m, G_m$

derive $H :- c, d_1, G_1, G, \dots, H :- c, d_m, G_m, G$

R3. Folding. Matching the body of a predicate definition

given $H :- d, \underline{A}, G$
 `newp(X) :- c, \underline{A}` and $d \rightarrow c$

derive $H :- d, \textcolor{red}{\underline{\text{newp}(X)}}, G$

R4. Clause Removal. Removal of clauses with unsatisfiable constraint or subsumed by others

An Example of Transformation

Initial program P :

```
incorrect :- X=1, reach(N,X).  
reach(N,X) :- X<N, X'=X+1, reach(N,X').  
reach(N,X) :- X<0, X>=N.
```

An Example of Transformation

Initial program P :

```
incorrect :- X=1, reach(N,X).  
reach(N,X) :- X<N, X'=X+1, reach(N,X').  
reach(N,X) :- X<0, X>=N.
```

definition:

```
newp(N,X) :- X>0, reach(N,X).
```

An Example of Transformation

Initial program P :

```
incorrect :- X=1, reach(N,X).  
reach(N,X) :- X<N, X'=X+1, reach(N,X').  
reach(N,X) :- X<0, X>=N.
```

definition:

```
newp(N,X) :- X>0, reach(N,X).
```

folding: % $X=1 \rightarrow X>0$

```
incorrect :- X=1, newp(N,X).
```

An Example of Transformation

Initial program P :

```
incorrect :- X=1, reach(N,X).  
reach(N,X) :- X<N, X'=X+1, reach(N,X').  
reach(N,X) :- X<0, X>=N.
```

definition:

```
newp(N,X) :- X>0, reach(N,X).
```

folding: % $X=1 \rightarrow X>0$

```
incorrect :- X=1, newp(N,X).
```

unfolding:

```
newp(N,X) :- X>0, reach(N,X).
```

An Example of Transformation

Initial program P :

```
incorrect :- X=1, reach(N,X).  
reach(N,X) :- X<N, X'=X+1, reach(N,X').  
reach(N,X) :- X<0, X>=N.
```

definition:

```
newp(N,X) :- X>0, reach(N,X).
```

folding: % $X=1 \rightarrow X>0$

```
incorrect :- X=1, newp(N,X).
```

unfolding:

```
newp(N,X) :- X>0, X<N, X'=X+1, reach(N,X').
```

An Example of Transformation

Initial program P :

```
incorrect :- X=1, reach(N,X).  
reach(N,X) :- X<N, X'=X+1, reach(N,X').  
reach(N,X) :- X<0, X>=N.
```

definition:

```
newp(N,X) :- X>0, reach(N,X).
```

folding: % $X=1 \rightarrow X>0$

```
incorrect :- X=1, newp(N,X).
```

unfolding:

```
newp(N,X) :- X>0, X<N, X'=X+1, reach(N,X').
```

```
newp(N,X) :- X>0, reach(N,X).
```

An Example of Transformation

Initial program P :

```
incorrect :- X=1, reach(N,X).  
reach(N,X) :- X<N, X'=X+1, reach(N,X').  
reach(N,X) :- X<0, X>=N.
```

definition:

```
newp(N,X) :- X>0, reach(N,X).
```

folding: % $X=1 \rightarrow X>0$

```
incorrect :- X=1, newp(N,X).
```

unfolding:

```
newp(N,X) :- X>0, X<N, X'=X+1, reach(N,X').
```

```
newp(N,X) :- X>0, reach(N,X).
```

An Example of Transformation

Initial program P :

```
incorrect :- X=1, reach(N,X).  
reach(N,X) :- X<N, X'=X+1, reach(N,X').  
reach(N,X) :- X<0, X>=N.
```

definition:

```
newp(N,X) :- X>0, reach(N,X).
```

folding: % $X=1 \rightarrow X>0$

```
incorrect :- X=1, newp(N,X).
```

unfolding:

```
newp(N,X) :- X>0, X<N, X'=X+1, reach(N,X').
```

```
newp(N,X) :- X>0, X<0, X>=N.
```

An Example of Transformation

Initial program P :

```
incorrect :- X=1, reach(N,X).  
reach(N,X) :- X<N, X'=X+1, reach(N,X').  
reach(N,X) :- X<0, X>=N.
```

definition:

```
newp(N,X) :- X>0, reach(N,X).
```

folding: % $X=1 \rightarrow X>0$

```
incorrect :- X=1, newp(N,X).
```

unfolding:

```
newp(N,X) :- X>0, X<N, X'=X+1, reach(N,X').
```

```
newp(N,X) :- X>0, X<0, X>=N. clause removal
```

An Example of Transformation

Initial program P :

```
incorrect :- X=1, reach(N,X).  
reach(N,X) :- X<N, X'=X+1, reach(N,X').  
reach(N,X) :- X<0, X>=N.
```

definition:

```
newp(N,X) :- X>0, reach(N,X).
```

folding: % $X=1 \rightarrow X>0$

```
incorrect :- X=1, newp(N,X).
```

unfolding:

```
newp(N,X) :- X>0, X<N, X'=X+1, reach(N,X').
```

```
newp(N,X) :- X>0, X<0, X>=N. clause removal
```

foldng:

```
newp(N,X) :- X>0, X<N, X'=X+1, newp(N,X').
```

An Example of Transformation

Initial program P :

```
incorrect :- X=1, reach(N,X).  
reach(N,X) :- X<N, X'=X+1, reach(N,X').  
reach(N,X) :- X<0, X>=N.
```

definition:

```
newp(N,X) :- X>0, reach(N,X).
```

folding: % $X=1 \rightarrow X>0$

```
incorrect :- X=1, newp(N,X).
```

unfolding:

```
newp(N,X) :- X>0, X<N, X'=X+1, reach(N,X').
```

```
newp(N,X) :- X>0, X<0, X>=N. clause removal
```

foldng:

```
newp(N,X) :- X>0, X<N, X'=X+1, newp(N,X').
```

Transformed program $TransfP$:

```
incorrect :- X=1, newp(N,X).
```

```
newp(N,X) :- X<N, X'=X+1, X>0, newp(N,X').
```

% No facts. $M(P) = M(TransfP) = \emptyset$. Thus, $incorrect \notin M(P)$.

The Transformation Strategy

$\text{Transform}(P)$

```
TransfP = ∅;  
Defs = {incorrect :- initConf(X), reach(X)};  
while ∃cl ∈ Defs do  
    Cls = Unfold(cl);  
    Cls = RemoveClauses(Cls);  
    Defs = (Defs – {cl}) ∪ Define(Cls);  
    TransfP = TransfP ∪ Fold(Cls, Defs);  
od
```

Theorem (Termination and Correctness of the Transformation Strategy)

- For a *suitable* Define, $\text{Transform}(P)$ terminates for all P ;
- $\text{incorrect} \in M(P)$ iff $\text{incorrect} \in M(\text{Transf}P)$

Generalization Strategies

- The most critical transformation step in the Transform strategy is the *introduction of new predicate definitions* (Define) to be used for folding.

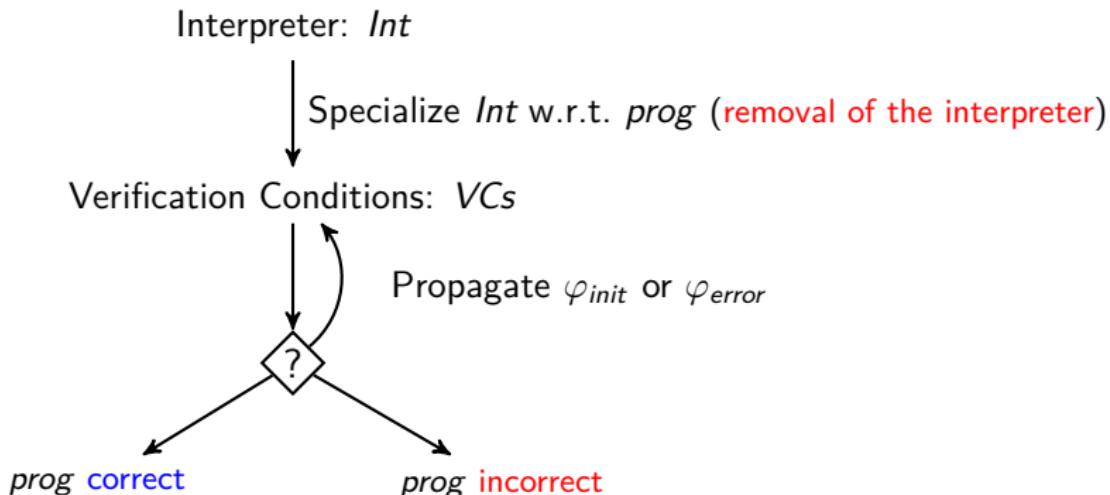
- Given $p(X) :- c(X, Y), \underline{q(Y)}.$

Introduce $\text{newp}(Y) :- d(Y), \underline{q(Y)}.$

where $c(X, Y) \rightarrow d(Y)$ ($d(Y)$ is a *generalization* of $c(X, Y)$)

and *fold*: $p(X) :- c(X, Y), \text{newp}(Y).$

The Transformation-based Verification Method



- prog correct if no constrained facts appear in the VCs.
- prog incorrect if the fact `incorrect.` appears in the VCs.

Generation of Verification Conditions

Generating Verification Conditions via Specialization

Removal of the interpreter: An application of *Transform* that *specializes Int* w.r.t. *prog* by unfolding all occurrences of:

- *tr* (i.e., the operational semantics of the imperative language)
- *at* (i.e., the encoding of *prog*)

New predicate definitions correspond to a subset of the *program points*:

```
new1(X,Y,N) :- reach(cf(cmd(0,ite(...))),  
                      [[int(x),X],[int(y),Y],[int(n),N]])).
```

VC: The Specialized Interpreter for *increase* (Verification Conditions)

```
incorrect :- X=0, Y=0, new1(X,Y,N).
```

```
new1(X,Y,N) :- X<N, X1=X+1, Y1=X1+Y, new1(X1,Y1,N).
```

```
new1(X,Y,N) :- X≥N, X>Y.
```

The fact *incorrect*. is not in VC: we cannot infer that *increase* is *incorrect*.
A constrained fact is in VC: we cannot infer that *increase* is *correct*.

Propagation of Initial and Error Conditions

Propagation of φ_{init}

Propagation of the initial configuration φ_{init} : An application of *Transform* where new predicate definitions are introduced by *widening* [Cousot-Cousot 77].

Propagation of φ_{init}

Propagation of the initial configuration φ_{init} : An application of *Transform* where new predicate definitions are introduced by *widening* [Cousot-Cousot 77].

Definition 1.

```
incorrect :- X=0, Y=0, new1(X,Y,N) .
```

Propagation of φ_{init}

Propagation of the initial configuration φ_{init} : An application of *Transform* where new predicate definitions are introduced by *widening* [Cousot-Cousot 77].

Definition 1.

```
incorrect :- X=0, Y=0, new1(X,Y,N).
```

Definition 2.

```
new2(X,Y,N) :- X=1, Y=1, N>0, new1(X,Y,N).
```

Propagation of φ_{init}

Propagation of the initial configuration φ_{init} : An application of *Transform* where new predicate definitions are introduced by *widening* [Cousot-Cousot 77].

Definition 1.

```
incorrect :- X=0, Y=0, new1(X,Y,N).
```

Definition 2.

```
new2(X,Y,N) :- X=1, Y=1, N>0, new1(X,Y,N).
```

Candidate new definition:

```
new3(Xr,Yr,Nr) :- Xr=1, Yr=1, X=2, Y=3, N>1, new1(X,Y,N).
```

Potential nontermination of *Transform* is detected!

Propagation of φ_{init}

Propagation of the initial configuration φ_{init} : An application of *Transform* where new predicate definitions are introduced by *widening* [Cousot-Cousot 77].

Definition 1.

```
incorrect :- X=0, Y=0, new1(X,Y,N).
```

Definition 2.

```
new2(X,Y,N) :- X=1, Y=1, N>0, new1(X,Y,N).
```

Candidate new definition:

```
new3(Xr,Yr,Nr) :- Xr=1, Yr=1, X=2, Y=3, N>1, new1(X,Y,N).
```

Potential nontermination of *Transform* is detected!

Definition 3. % Generalization based on widening

```
new3(X,Y,N) :- X $\geq$ 1, Y $\geq$ 1, N>0, new1(X,Y,N).
```

More complex generalization strategies based on combinations of widening and *convex hull* [Cousot-Halbwachs 78]

VC₁: Transformed Verification Conditions for *increase*

... propagating the constraint X=0, Y=0.

```
incorrect :- N>0, X1=1, Y1=1, new2(X1, Y1, N).  
new2(X, Y, N) :- X=1, Y=1, N>1, X1=2, Y1=3, new3(X1, Y1, N).  
new3(X, Y, N) :- X1≥1, Y1≥X1, X<N, X1=X+1, Y1=X1+Y, new3(X1, Y1, N).  
new3(X, Y, N) :- Y≥1, N>0, X≥N, X>Y.
```

The fact incorrect. is not in VC₁: we cannot infer that *increase* is incorrect.

A constrained fact is in VC₁: we cannot infer that *increase* is correct.

Program Reversal

P:

```
incorrect :- a(X), p(X).  
p(X) :- c(X,Y), p(Y).  
p(X) :- b(X).
```



RevP:

```
incorrect :- b(X), p(X).  
p(Y) :- c(X,Y), p(X).  
p(X) :- a(X).
```

$\text{incorrect} \in M(P)$ iff $\text{incorrect} \in M(\text{RevP})$

Propagation of φ_{error}

VC₁: Transformed Verification Conditions for *increase*

incorrect :- N>0, X1=1, Y1=1, new2(X1, Y1, N).

new2(X, Y, N) :- X=1, Y=1, N>1, X1=2, Y1=3, new3(X1, Y1, N).

new3(X, Y, N) :- X1 \geq 1, Y1 \geq X1, X<N, X1=X+1, Y1=X1+Y, new3(X1, Y1, N).

new3(X, Y, N) :- Y \geq 1, N>0, X \geq N, X>Y.

Propagation of φ_{error}

VC₁: Transformed Verification Conditions for *increase*

```
incorrect :- N>0, X1=1, Y1=1, new2(X1, Y1, N).  
new2(X, Y, N) :- X=1, Y=1, N>1, X1=2, Y1=3, new3(X1, Y1, N).  
new3(X, Y, N) :- X1≥1, Y1≥X1, X<N, X1=X+1, Y1=X1+Y, new3(X1, Y1, N).  
new3(X, Y, N) :- Y≥1, N>0, X≥N, X>Y.
```

VC₂: Reversed VC₁

```
incorrect :- Y≥1, N>0, X≥N, X>Y, new3(X, Y, N).  
new3(X1, Y1, N) :- X1≥1, Y1≥X1, X<N, X1=X+1, Y1=X1+Y, new3(X, Y, N).  
new3(X1, Y1, N) :- X=1, Y=1, N>1, X1=2, Y1=3, new2(X, Y, N).  
new2(X1, Y1, N) :- N>0, X1=1, Y1=1.
```

Propagation of φ_{error}

VC₁: Transformed Verification Conditions for *increase*

```
incorrect :- N>0, X1=1, Y1=1, new2(X1, Y1, N).  
new2(X, Y, N) :- X=1, Y=1, N>1, X1=2, Y1=3, new3(X1, Y1, N).  
new3(X, Y, N) :- X1≥1, Y1≥X1, X<N, X1=X+1, Y1=X1+Y, new3(X1, Y1, N).  
new3(X, Y, N) :- Y≥1, N>0, X≥N, X>Y.
```

VC₂: Reversed VC₁

```
incorrect :- Y≥1, N>0, X≥N, X>Y, new3(X, Y, N).  
new3(X1, Y1, N) :- X1≥1, Y1≥X1, X<N, X1=X+1, Y1=X1+Y, new3(X, Y, N).  
new3(X1, Y1, N) :- X=1, Y=1, N>1, X1=2, Y1=3, new2(X, Y, N).  
new2(X1, Y1, N) :- N>0, X1=1, Y1=1.
```

VC₃: Transformed VC₂

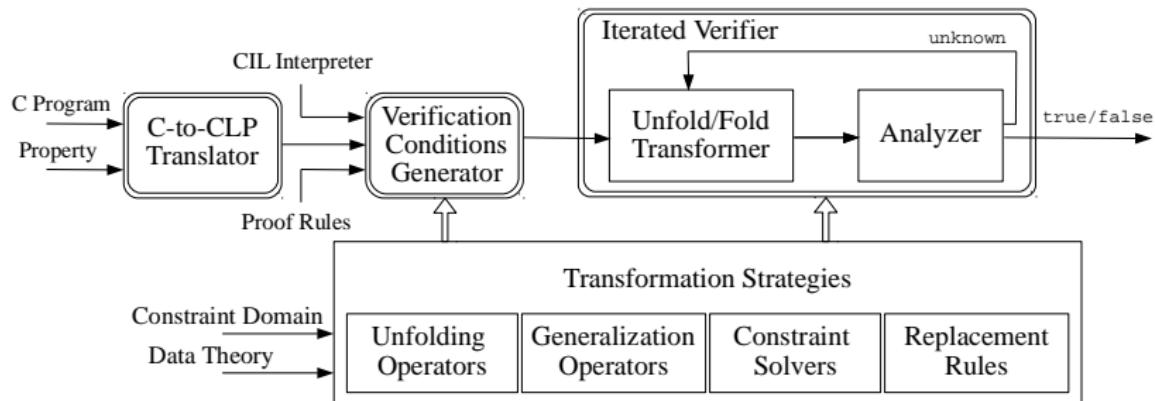
by propagating the constraint $Y \geq 1, N > 0, X \geq N, X > Y$.

```
incorrect :- Y≥1, N>0, X≥N, X>Y, new4(X, Y, N).
```

No constrained facts: *increase* is **correct**.

Experimental Results

The VeriMAP tool <http://map.uniroma2.it/VeriMAP>
[DFPP PEPM 2013, VMCAI 2014, TACAS 2014]



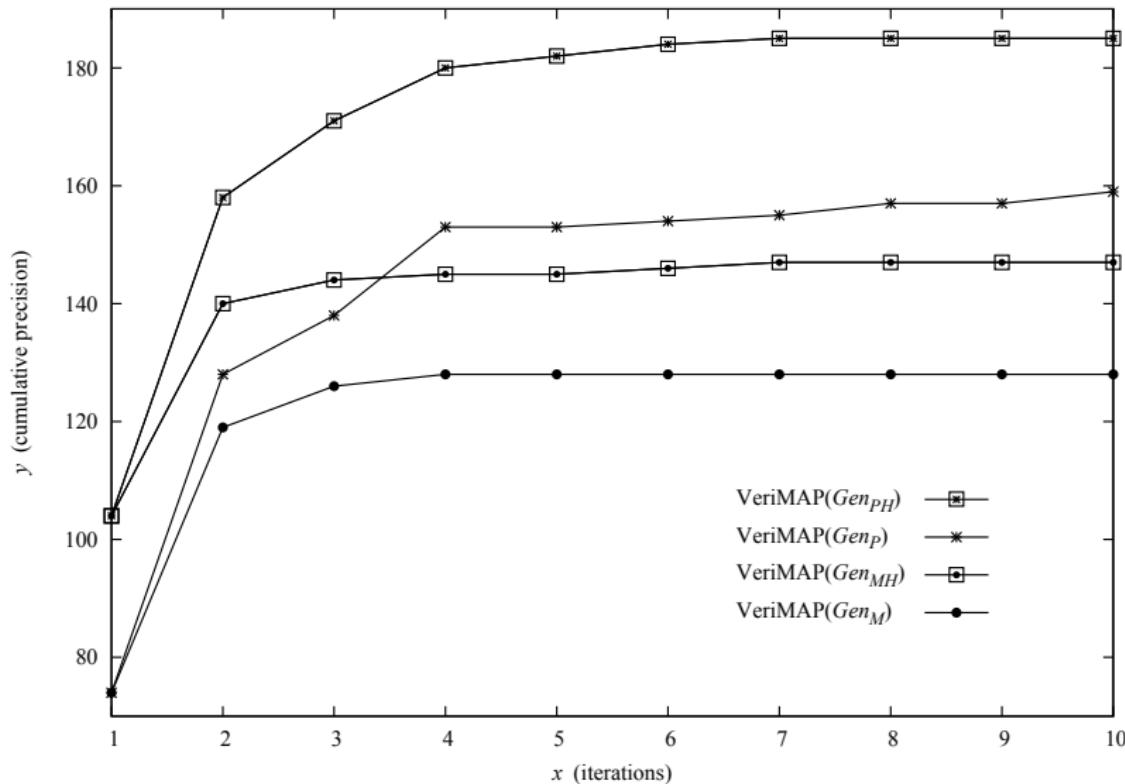
Experimental Evaluation

216 examples taken from: DAGGER, TRACER, InvGen, and TACAS 2013 Software Verification Competition.

		VeriMAP	ARMC	HSF(C)	TRACER
1	<i>correct answers</i>	185	138	160	103
2	safe problems	154	112	138	85
3	unsafe problems	31	26	22	18
4	<i>incorrect answers</i>	0	9	4	14
5	false alarms	0	8	3	14
6	missed bugs	0	1	1	0
7	<i>errors</i>	0	18	0	22
8	<i>timed-out problems</i>	31	51	52	77
9	<i>total score</i>	339 (0)	210 (-40)	278 (-20)	132 (-56)
10	<i>total time</i>	10717.34	15788.21	15770.33	23259.19
11	<i>average time</i>	57.93	114.41	98.56	225.82

- ARMC [Podelski, Rybalchenko PADL 2007]
- HSF(C) [Grebenshchikov et al. TACAS 2012]
- TRACER [Jaffar, Murali, Navas, Santosa CAV 2012]

Improving Precision by Iteration



Future Work

Exploit parametricity of the transformation-based approach and extend to:

- Recursive functions
- More data structure theories (lists, heaps, etc.)
- Other programming languages and properties

Thanks for your attention!

Try the VeriMAP tool <http://map.uniroma2.it/VeriMAP>