#### 49th IEEE Conference on Decision and Control Atlanta, GA

Pre-Conference Workshop on Verification of Control Systems

December 14, 2010

# Verification of Control Systems by Abstract Interpretation

# Patrick Cousot

#### **CIMS-NYU & ENS**

pcousot@cs.nyu.edu cousot@ens.fr cs.nyu.edu/~pcousot di.ens.fr/~cousot

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

# Motivation

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

2

© P. Cousot

### All computer scientists have experienced bugs







Ariane 5.01 failure (overflow error)

Patriot failure (float rounding error)

Mars orbiter loss (unit error)

Russian Proton-M rocket carrying 3 Glonass-M satellites (unknown programming error)

- Checking the presence of bugs is great
- Proving their absence is even better!

# Abstract interpretation

Patrick Cousot & Radhia Cousot. Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints. In *Conference Record of the Fourth Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*, pages 238–252, Los Angeles, California, 1977. ACM Press, New York, NY, USA

Patrick Cousot & Radhia Cousot. Systematic design of program analysis frameworks. In *Conference Record of the Sixth Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages* pages 269–282, San Antonio, Texas, 1979. ACM Press, New York, U.S.A.

Patrick Cousot & Radhia Cousot. Abstract interpretation frameworks. *Journal of Logic and Computation*, 2(4):511-547, August 1992.

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

#### Abstract interpretation

- Started in the 70's
- Statically and automatically inferring properties of the behavior of programs/computer systems for program analysis (proof, verification, optimization, transformation, etc.)
- Based on the idea that undecidability and complexity of automated static program analysis can be fought by approximation
- Applications of abstract interpretation do scale up!

### Application to static analysis and verification

- Static analysis consists in automatically answering questions about the runtime executions of programs
- Static means « at compile time », by examining the program text only, without executions on computers
- Automatic means by a computer, without human intervention during the analysis



# Fighting undecidability and complexity in program verification

- Any automatic program verification method will definitely fail on infinitely many programs (Gödel)
- Solutions:
  - Ask for human help (theorem-prover based deductive methods)
  - Consider (small enough) finite systems (modelchecking)
  - Do complete abstractions or else sound approximations (abstract interpretation)



# An informal introduction to abstract interpretation

P. Cousot & R. Cousot. A gentle introduction to formal verification of computer systems by abstract interpretation. In *Logics and Languages for Reliability and Security*, J. Esparza, O. Grumberg, & M. Broy (Eds), NATO Science Series III: Computer and Systems Sciences, © IOS Press, 2010, Pages 1–29.

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

# I) Define the programming language semantics Formalize the concrete **execution** of programs (e.g. transition system) (x,y) t=0 X t=2 t = lTrajectory Space/time trajectory in state space











# Unsound validation: testing Try a few cases Forbidden zone Error !!! Test of a few trajectories 15

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010



#### Unsound validation: static analysis

Many static analysis tools are **unsound** (e.g. Coverity, etc.) so inconclusive



#### Incompleteness

When abstract proofs may fail while concrete proofs would succeed



#### True error

The abstract alarm may correspond to a concrete error



#### False alarm

The abstract alarm may correspond to no concrete error (false negative)



49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

## What to do about false alarms?

- Automatic refinement: inefficient and may not terminate (Gödel)
- Domain-specific abstraction:
  - Adapt the abstraction to the programming paradigms typically used in given domain-specific applications
  - e.g. synchronous control/command: no recursion, no dynamic memory allocation, maximum execution time, etc.

# Just a bit less informally ...

Patrick Cousot & Radhia Cousot. Basic Concepts of Abstract Interpretation. In *Building the Information Society*, René Jacquard (Ed.), Kluwer Academic Publishers, pp. 359–366, 2004. (IFIP WCC 2004 Toulouse, Topical Day on Abstract Interpretation, Tuesday 24 August 2004).

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

22

© P. Cousot

## Semantics and specification

- Small-step operational semantics (of a program)  $\langle \Sigma, \tau, \ln it \rangle$
- Reachable states

 $\mathsf{post}[\tau^*]$ Init

 $au^*$ 

transitive closure

 $\mathsf{post}\llbracket r\rrbracket X \triangleq \{s' \mid \exists s \in X : r(s, s')\}$ 

reachable states

post-image

fixpoint characterization

• Absence of run-time errors

 $\begin{array}{l} \mathsf{Bad} \\ \mathsf{post}[\![\tau^*]\!] \mathsf{Init} \subseteq \Sigma \setminus \mathsf{Bad} \\ \mathsf{no} \ \mathsf{erroneous} \ \mathsf{state} \ \mathsf{is} \ \mathsf{reachable} \ \mathsf{from} \ \mathsf{initial} \ \mathsf{states} \end{array}$ 

 $= \mathsf{lfp}^{\subseteq} \lambda X \cdot \mathsf{lnit} \cup \mathsf{post}[\tau] X$ 

## Fixpoint abstraction and approximation

**Theorem** If  $\langle L, \leq, \perp \rangle$  is a complete lattice or a cpo,  $F \in L \to L$  is increasing,  $\langle \overline{L}, \subseteq \rangle$  is a poset,  $\alpha \in L \to \overline{L}$  is continuous<sup>(6),(7)</sup>,  $\overline{F} \in \overline{L} \to \overline{L}$  commutes (resp. semicommutes) with F that is  $\alpha \circ F = \overline{F} \circ \alpha$  (resp.  $\alpha \circ F \subseteq \overline{F} \circ \alpha$ ) then  $\alpha(\mathsf{lfp}_{\perp}^{\leq} F) = \mathsf{lfp}_{\alpha(\perp)}^{\subseteq} \overline{F}$  (resp.  $\alpha(\mathsf{lfp}_{\perp}^{\leq} F) \subseteq \mathsf{lfp}_{\alpha(\perp)}^{\subseteq} \overline{F})$ .

#### Example

 $1\Sigma$ 

identity

 $au^* = \mathsf{lfp}^{\subseteq} \lambda R ullet \mathbb{1}_{\Sigma} \cup R \circ au$ 

transitive closure

 $\lambda R \cdot \mathsf{post}[\![R]\!]\mathsf{Init} \circ \lambda R \cdot \mathbb{1}_{\Sigma} \cup R \circ \tau] = \lambda R \cdot \mathsf{Init} \cup \mathsf{post}[\![\tau]\!](\mathsf{post}[\![R]\!]\mathsf{Init})$ 

commutativity

 $\Rightarrow \operatorname{post}[\tau^*] \operatorname{Init} = \operatorname{lfp}^{\subseteq} \lambda X \cdot \operatorname{Init} \cup \operatorname{post}[\tau] X$ 

reachable states

<sup>&</sup>lt;sup>(6)</sup>  $\alpha$  is *continuous* if and only if it preserves existing lubs of increasing chains.

<sup>&</sup>lt;sup>(7)</sup> The continuity hypothesis for  $\alpha$  can be restricted to the iterates of the least fixpoint of F.



# **Convergence** acceleration • Fixpoints can be computed iteratively $\operatorname{lfp}^{\sqsubseteq}_{\perp} F = \bigsqcup_{n \ge 0} F^n(\bot)$ May not finitely converge • Accelerate convergence by a widening $\nabla$ $F^{\uparrow 0} \triangleq \bot, \quad \dots, \quad F^{\uparrow n+1} \triangleq F^{\uparrow n} \nabla F(F^{\uparrow n}) \quad \text{until} \quad F(F^{\uparrow \ell}) \sqsubseteq F^{\uparrow \ell}$ • followed by a narrowing $\Delta$ $F^{\downarrow 0} \triangleq F^{\uparrow \ell}, \quad \dots, \quad F^{\downarrow n+1} \triangleq F^{\downarrow n} \bigtriangleup F(F^{\downarrow n}) \quad \text{until} \quad F(F^{\downarrow \ell}) = F^{\downarrow \ell}$ • to get an over-approximation $\mathsf{lfp}_{\bot}^{\sqsubseteq} F \sqsubseteq F^{\downarrow \ell} \sqsubseteq F^{\uparrow \ell}$



Soundness and (In)completeness • Soundness: We effectively compute  $F^{\downarrow \ell}$  such that  $\alpha(\mathsf{post}[\tau^*]\mathsf{Init}) = \alpha(\mathsf{lfp}^{\subseteq} \lambda X \cdot \mathsf{Init} \cup \mathsf{post}[\tau]X) \sqsubseteq \mathsf{lfp}^{\sqsubseteq}F \sqsubseteq F^{\downarrow \ell} \sqsubseteq F^{\uparrow \ell}$ and check that  $F^{\downarrow \ell} \cap \mathsf{Bad} = \emptyset$  in the abstract, **proving**  $\text{post}[\tau^*]$ Init = Ifp<sup> $\subseteq \ \lambda X$  · Init  $\cup$  post $[\tau] X \subseteq \Sigma \setminus$  Bad</sup> • [In]completeness: may produce false alarms (e.g. take  $Bad = \neg post[\tau^*]$  Init so that no abstraction is possible) • In practice, can be made complete for domainspecific programs (such as synchronous control/

command) with adequate abstractions.

## In practice ...

- Proceed by structural induction in the syntax of programs
- Use chaotic/asynchronous iterations within loops
- Use many abstractions combined in an approximate reduced product with partial iterative reduction
  - Example:
    - $x \in [1, 13]$
    - even(x)

reduction:

•  $x \in [2, 12]$ 

# Applications of abstract interpretation to the static analysis of aerospace control systems

Julien Bertrane, Patrick Cousot, Radhia Cousot, Jérôme Feret, Laurent Mauborgne, Antoine Miné, Xavier Rival. Static Analysis and Verification of Aerospace Software by Abstract Interpretation. AIAA Infotech@Aerospace 2010, AIAA 2010-3385, 20–22 April 2010, Atlanta, GA, http://pdf.aiaa.org/preview/2010/CDReadyMIAA10\_2358/PV2010\_3385.pdf

Patrick Cousot. Integrating Physical Systems in the Static Analysis of Embedded Control Software. The Third Asian Symposium on Programming Languages and Systems (APLAS'05), Tsukuba, Japan, November 3–5, 2005. Lecture Notes in Computer Science, volume 3780, © Springer, Berlin, pp. 135–138.

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

# ASTRÉE

Patrick Cousot, Radhia Cousot, Jérôme Feret, Antoine Miné, David Monniaux, Laurent Mauborgne, Xavier Rival. The ASTRÉE Analyzer. ESOP 2005: The European Symposium on Programming, Edinburgh, Scotland, April 2–10, 2005. Lecture Notes in Computer Science 3444, © Springer, Berlin, pp. 21–30.

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

31

© P. Cousot

# Target language and applications

- C programming language
  - Without recursion, longjump, dynamic memory allocation, conflicting side effects, backward jumps, system calls (stubs)
  - With all its horrors (union, pointer arithmetics, etc)
  - Reasonably extending the standard (e.g. size & endianess of integers, IEEE 754-1985 floats, etc)
- Synchronous control/command
  - e.g. generated from Scade

# The semantics of C implementations is very hard to define

What is the effect of out-of-bounds array indexing?

```
% cat unpredictable.c
#include <stdio.h>
int main () { int n, T[1];
    n = 2147483647;
    printf("n = %i, T[n] = %i\n", n, T[n]);
}
```

Yields different results on different machines:

n = 2147483647, T[n] = 2147483647 Macintosh PPC n = 2147483647, T[n] = -1208492044 Macintosh Intel n = 2147483647, T[n] = -135294988 PC Intel 32 bits Bus error PC Intel 64 bits

# Implicit specification

- Absence of runtime errors: overflows, division by zero, buffer overflow, null & dangling pointers, alignment errors, ...
- Semantics of runtime errors:
  - Terminating execution: stop (e.g. floating-point exceptions when traps are activated)
  - Predictable outcome: go on with worst case (e.g. signed integer overflows result in some integer, some options: e.g. modulo arithmetics)
  - Unpredictable outcome: stop (e.g. memory corruption)

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010



#### Example of general purpose abstraction: octagons

- Invariants of the form  $\pm x \pm y \leq c$ , with  $\mathcal{O}(\mathbb{N}^2)$  memory and  $\mathcal{O}(\mathbb{N}^3)$  time cost.
- Example:

- At ★, the interval domain gives
   L ≤ max(max A, (max Z)+(max V)).
- In fact, we have  $L \leq A$ .
- To discover this, we must know at  $\bigstar$  that R = A-Z and R > V.
- Here, R = A-Z cannot be discovered, but we get  $L-Z \leq \max R$  which is sufficient.
- We use many octagons on small packs of variables instead of a large one using all variables to cut costs.

© P. Cousot

## Example of general purpose abstraction: decision trees

/\* boolean.c \*/

typedef enum {F=0,T=1} BOOL; BOOL B; void main () { unsigned int X, Y; while (1) { B = (X == 0);if (!B) { Y = 1 / X;



The boolean relation abstract domain is parameterized by the height of the decision tree (an analyzer option) and the abstract domain at the leaves

#### Example of domain-specific abstraction: ellipses typedef enum {FALSE = 0, TRUE = 1} BOOLEAN; BOOLEAN INIT; float P, X; void filter () { static float E[2], S[2]; if (INIT) { S[0] = X; P = X; E[0] = X;} else { P = (((((0.5 \* X) - (E[0] \* 0.7)) + (E[1] \* 0.4)))+ $(S[0] * 1.5)) - (S[1] * 0.7)); \}$ E[1] = E[0]; E[0] = X; S[1] = S[0]; S[0] = P;/\* S[0], S[1] in [-1327.02698354, 1327.02698354] \*/ } void main () { X = 0.2 \* X + 5; INIT = TRUE; while (1) { X = 0.9 \* X + 35; /\* simulated filter input filter (); INIT = FALSE; } }

#### Example of domain-specific abstraction (I): exponentials

```
% cat count.c
typedef enum {FALSE = 0, TRUE = 1} BOOLEAN;
volatile BOOLEAN I; int R; BOOLEAN T;
void main() {
 R = 0;
  while (TRUE) {
    __ASTREE_log_vars((R));
                                          \leftarrow potential overflow!
    if (I) { R = R + 1; }
    else { R = 0; }
    T = (R \ge 100);
    __ASTREE_wait_for_clock(());
  }}
% cat count.config
__ASTREE_volatile_input((I [0,1]));
__ASTREE_max_clock((3600000));
% astree -exec-fn main -config-sem count.config count.c|grep '|R|'
|R| <= 0. + clock *1. <= 3600001.
```

39

© P. Cousot

#### Example of domain-specific abstraction (II): exponentials

```
% cat retro.c
typedef enum {FALSE=0, TRUE=1} BOOL;
BOOL FIRST;
volatile BOOL SWITCH;
volatile float E;
float P, X, A, B;
void dev( )
{ X=E;
  if (FIRST) { P = X; }
  else
    \{ P = (P - ((((2.0 * P) - A) - B)) \}
            * 4.491048e-03)); };
  B = A;
  if (SWITCH) \{A = P;\}
  else {A = X;}
}
Cumulated rounding error
```

```
void main()
{ FIRST = TRUE:
  while (TRUE) {
    dev():
    FIRST = FALSE;
    ASTREE wait for clock(()):
  }}
% cat retro.config
__ASTREE_volatile_input((E [-15.0, 15.0]));
__ASTREE_volatile_input((SWITCH [0,1]));
__ASTREE_max_clock((3600000));
|P| <= (15. + 5.87747175411e-39
/ 1.19209290217e-07) * (1 +
1.19209290217e-07) clock - 5.87747175411e-39
/ 1.19209290217e-07 <= 23.0393526881
                                   \overrightarrow{x}
```

40

© P. Cousot

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

#### Example of domain-specific abstraction (III): scaling

```
% gcc scale.c
 % cat -n scale.c
                                        % ./a.out
  1 int main () \{
  2 float x; x = 0.7000001; x = 0.69999988079071
  3 while (1) {
  4 = x - 3.0;
  5 \quad x = x * 3.0;
  6 __ASTREE_log_vars((x));
  7 __ASTREE_wait_for_clock(());
  8 }
  9 }
 % cat scale.config
  __ASTREE_max_clock((100000000));
 % astree -exec-fn main -config-sem scale.config -unroll Q, scale.
   |& grep "x in" | tail -1
 direct = <float-interval: x in [0.69999986887, 0.7000000-084]
 %
49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010
```

© P. Cousor



#### An erroneous common belief on static analyzers

"The properties that can be proved by static analyzers are often simple" [2]

Like in mathematics:

- May be simple to state (no overflow)
- But harder to discover (S[0], S[1] in [-1327.02698354, 1327.02698354])
- And difficult to prove (since it requires finding a non trivial non-linear invariant for second order filters with complex roots [Fer04], which can hardly be found by exhaustive enumeration)

Reference

[2] Vijay D'Silva, Daniel Kroening, and Georg Weissenbacher. A Survey of Automated Techniques for Formal Software Verification. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Vol. 27, No. 7, July 2008.

# Industrial applications

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

# **Examples of applications**

- Verification of the absence of runtime-errors in
  - Fly-by-wire flight control systems





• ATV docking system



 Flight warning system (on-going work)



## Industrialization

#### • 8 years of research (CNRS/ENS/INRIA):

#### www.astree.ens.fr



#### • Industrialization by AbsInt (since Jan. 2010):

www.absint.com/astree/Absint<br/>Argewandte Informatik

# On-going work

Julien Bertrane, Patrick Cousot, Radhia Cousot, Jérôme Feret, Laurent Mauborgne, Antoine Miné, Xavier Rival. Static Analysis and Verification of Aerospace Software by Abstract Interpretation. AIAA Infotech@Aerospace 2010, AIAA 2010-3385, 20–22 April 2010, Atlanta, GA, http://pdf.aiaa.org/preview/2010/ CDReadyMIAA10\_2358/PV2010\_3385.pdf

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

# Verification of target programs

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

# Verification of compiled programs

- The valid source may be proved correct while the certified compiler is incorrect so the target program may go wrong
- Possible approaches:
  - Verification at the target level
  - Source to target proof translation and proof check on the target
  - Translation validation (local verification of equivalence of run-time error free source and target)
  - Formally certified compilers

# Verification of imperfectly clocked synchronous systems

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

© P. Cousot

# Imperfect synchrony

• Example of (buggy) communicating synchronous systems:



## Semantics and abstractions

- Continuous semantics (value s(t) of signals s at any time t)
- Clock ticks and serial communications do happen in known time intervals  $[l, h], l \leq h$
- Examples of abstractions:
  - $\forall t \in [a; b] : s(t) = x.$
  - $\exists t \in [a; b] : s(t) = x.$
  - change counting (≤ k, a ► <b) and (≥ k, a ► <b)</li>
     (signal changes less (more) than k times in time interval [a, b])

## Example of static analysis



# THÉSÉE:Verification of embedded real-time parallel C programs

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

© P. Cousot

# Parallel programs

- Bounded number of processes with shared memory, events, semaphores, message queues, blackboards,...
- Processes created at initialization only
- Real time operating system (ARINC 653) with fixed priorities (highest priority runs first)
- Scheduled on a single processor

### Verified properties

- Absence of runtime errors
- Absence of unprotected data races

## Semantics

#### • No memory consistency model for C

 Optimizing compilers consider sequential processes out of their execution context

init: $flag1 = flag2 = 0$		
process 1:	process 2:	
flag1 = 1;	flag2 = 1;	write to flag1/2 and
if (!flag2)	if (!flag1)	read of flag2/1 are
{	{	independent so can be
<pre>/* critical section */</pre>	<pre>/* critical section */</pre>	reordered → error!

#### • We assume:

- sequential consistency in absence of data race
- for data races, values are limited by possible interleavings between synchronization points

## Abstractions

- Based on Astrée for the sequential processes
- Takes scheduling into account
- OS entry points (semaphores, logbooks, sampling and queuing ports, buffers, blackboards, ...) are all stubbed (using Astrée stubbing directives)
- Interference between processes: flow-insensitive abstraction of the writes to shared memory and inter-process communications

# Example of application: FWS



- Degraded mode (5 processes, 100 000 LOCS)
  - Ih40 on 64-bit 2.66 GHz Intel server
  - 98 alarms
- Full mode (15 processes, 1 600 000 LOCS)
  - 50 h
  - 12 000 alarms in May, 7000 in December 2010
  - <u>III more work to be done III</u> (e.g. analysis of complex data structures, interference, logs, etc)

# Conclusion

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

© P. Cousot

## **Cost-effective verification**

- The rumor has it that:
  - Manuel validation (testing) is costly, unsafe, not a verification!
  - Formal proofs by theorem provers are extremely laborious hence costly
  - Model-checkers do not scale up
- Why not try abstract interpretation?
  - Domain-specific static analysis scales and can deliver no false alarm

# The End

Cited references are available online at URL http://www.di.ens.fr/~cousot/COUSOTpapers.shtml

49th iEEE Conference on Decision and Control, Atlanta GA, Pre-Conference Workshop on Verification of Control Systems, Dec. 14, 2010

© P. Cousot