Abstract-Interpretation-based Static Analysis of Safety-Critical Embedded Software

Patrick Cousot

Patrick.Cousot@ens.fr www.di.ens.fr/~cousot

Workshop Airbus/partners on formal verification tools strategy

Airbus St Martin, M86, Auditorium Dubaï, Dec. 4–5, 2008



— 1 —



Abstract

Static software analysis has known brilliant successes in the small, by proving complex program properties of programs of a few dozen or hundreds of lines, either by systematic exploration of the state space or by interactive deductive methods. To scale up is a definite problem. Very few static analyzers are able to scale up to millions of lines without sacrificing automation and/or soundness and/or precision. Unsound static analysis may be useful for bug finding but is less useless in safety critical applications where the absence of bugs, at least of some categories of common bugs, should be formally verified.

After recalling the basic principles of abstract interpretation including the notions of abstraction, approximation, soundness, completeness, false alarm, etc., we introduce the domain-specific static analyzer ASTRÉE (www.astree.ens.fr) for proving the absence of runtime errors in safety critical real time embedded synchronous software in the large.

The talk emphasizes soundness (no runtime error is ever omitted), parametrization (the ability to refine abstractions by options and analysis directives), extensibility (the easy incorporation of new abstractions to refine the approximation), precision (few or no false alarms for programs in the considered application domain) and scalability (the analyzer scales to millions of lines).

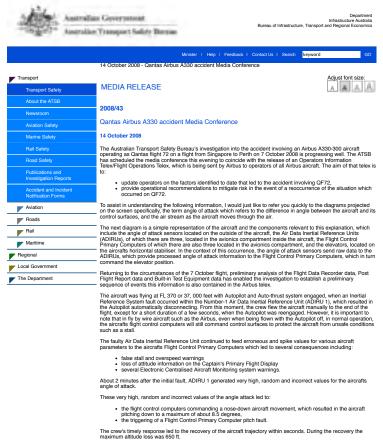
In conclusion, present-day software engineering methodology, which is based on the control of the design, coding and testing processes should evolve in the near future, to incorporate a systematic control of final software product thanks to domain-specific analyzers that scale up.



1. Motivation: bugs are everywhere



Example of bug report



The Digital Flight Data Recorder data show that ADIRU 1 continued to generate random spikes and a second nose-down aircraft movement was encountered later on, but with less significant values in terms of aircraft's trajectory.

At this stage of the investigation, the analysis of available data indicates that the ADIRU 1 abnormal behaviour is

likely as the origin of the event.
The aircraft contains very sophisticated and highly reliable systems. As far as we can understand, this appears to be a unique event and Airbus has advised that it is not aware of any similar event over the many years of operation of the Airbus.
Altous has this eventing, Australian time, issued an Operation Information Teles reflecting the above information. The televa sito forestations the issue of Operational Engineering Diadlens and provides information relating to operational recommendations to operations of A330 and A340 aircraft fitted with the type of ADRIU fitted to the accordina aircraft. These recommended provides are aunded an airminiary final the turn titley event of a similar occurrence. That includes guidance and checklists for crew response in the event of an inertial Reference System Talare.
Meanwhile, the ATSB's investigation is ongoing and will include:
 Devincial of data from the aircraft's three ADIRUIs and detailed seamination and analysis of that data Arrangements are currently browing made for the units to be sent to the comportent manufacturer is faulties in the US as soon as possible and for ATSB investigators to attend and help with that testing, along with the US as soon as possible and for ATSB investigators to attend or the pair with that testing, along with the US as a soon as possible and the ATSB investigators and the US as a total testing and the US as a databased or the US as a sound to the attend of the attend to the attend to the attend to the attend in addition, investigators have been conducting a databate review of the aircraft's maintenance history, including duration frame attendes or the review Attendes of the attend to the attend to the attend to the attendes of the attendes of the Viol's is also ongoing to progress interviews, which will include with injured passengers to understand what occurred in the aircraft dath. The ATSB plants to databate the attendes of the
There has been close and frequent communication between the ATSB, Quartas, Arbus, the EEA, and CASA. That close communication will continue as the investigation progresses to ensure that any additional safety action can be instigated as soon as possible should critical safety factors be identified. The ATSB expects to publish a Preliminary Facular aport in abuot 30 days from the date of the accident.
Media Contact: David Hope 1800 020 616
Related Documents: I Audio file of media conference, 14 October 2008 (18 MB) I
Print Last Updated: 14 October, 2008
Privacy I Copyright I Disclaimer I Linking to the ATSB website I Sitemap

"The Australian Transport Safety Bureau (ATSB) found that the main probable cause of this incident was a *latent software error* which allowed the ADIRU to use data from a failed accelerometer"

http://www.atsb.gov.au/newsroom/2008/release/2008_43.aspx, http://en.wikipedia.org/wiki/Qantas_Flight_72

Airbus, 12/04/2008

— 4 —



2. Varieties of Static Analyses



Static Analysis

- In general static analysis means "the fully automatic verification of properties of program executions using the program text only" (excluding running programs)
- But for trivial cases, it is undecidable
- Alternatives to impossible total verification:
 - under-verification (testing, bounded model-checking, bug pattern mining, etc): bug finding, misses bugs, never ends
 - over-verification (typing, dataflow analysis,etc): no bug missed but false alarms
- Challenge: total verification for a given category of properties and a given family of programs (no bug missed, no false alarm but not for all possible properties of all programs)



Example: Sparrow versus Astrée



```
x[i] = datal;
x[i+2] = datar;
                                       // buffer overrun: LOOP2
     }
}
int atkbd(int *dev, unsigned int type)
{
      const short period[32] =
                 [ 33, 37, 42, 46, 50, 54, 58, 63, 67, 75, 83, 92, 100, 109, 116, 125, 133, 149, 167, 182, 200, 217, 232, 250, 270, 303, 333, 370, 400, 435, 470, 500 ]
      const short delay[4] =
                { 250, 500, 750, 1000 };
      unsigned char param[2];
     int i, j;
      switch (type) {
           case 0x09: return 0;
           case 0x11:
          case 0x14:
          i = j = 0;
           while (i < 32 && period[i] < *dev) i++;</pre>
           while (j < 4 && delay[j] < *dev) j++;
            *dev = period[i];
                                        // buffer overrun: AFTERLOOP2
           *dev = delay[j];
                                        // buffer overrun: AFTERLOOP3
           param[0] = i | (j << 5);
          return 0;
      return -1;
}
 void bo_test()
 {
      atkbd((int*)any(), any());
      library_calls ();
      loop_test();
      param_check(any());
}
 int main()
{
      bo_test();
      return 1;
}
```



http://www.spa-arrow.com/



3. Abstract Interpretation



Example of static analysis

Example after invariant abstraction:

$$\begin{array}{ll} \{y \ge 0\} & \leftarrow & \text{hypothesis} \\ \text{x} := \text{y} \\ \{I(x, y)\} & \leftarrow & \text{loop invariant} \\ \text{while} (\text{x} > 0) \text{ do} \\ \text{x} := \text{x} - 1; \\ \text{od} \end{array}$$

Abstract fixpoint equation:

$$I(x,y) \ = \ x \geqslant 0 \land (x = y \lor I(x+1,y))$$
 (i.e. $I = \mathbf{F}^{\sharp}(I)^{(1)})$

Equivalent Floyd-Naur-Hoare verification conditions:

 $egin{aligned} &(y \geqslant 0 \wedge x = y) \Longrightarrow I(x,y) & ext{initialisation} \ &(I(x,y) \wedge x > 0 \wedge x' = x-1) \Longrightarrow I(x',y) & ext{iteration} \end{aligned}$

(1) We look for the most precise invariant I, implying all others, that is $Ifp^{\Rightarrow} F^{\sharp}$.



$$\begin{array}{l} \text{Accelerated Iterates } I = \bigsqcup_{n \to \infty} \mathbf{F}^{\sharp n}(\text{false}) \stackrel{y}{} \\ I^0(x,y) = \text{false} \qquad n \to \infty \end{array}$$

$$\begin{array}{l} I^1(x,y) = x \geqslant 0 \land (x = y \lor I^0(x+1,y)) \qquad y \\ = 0 \leqslant x = y \end{aligned}$$

$$\begin{array}{l} I^2(x,y) = x \geqslant 0 \land (x = y \lor I^1(x+1,y)) \qquad y \\ = 0 \leqslant x \leqslant y \leqslant x+1 \end{aligned}$$

$$\begin{array}{l} I^3(x,y) = x \geqslant 0 \land (x = y \lor I^2(x+1,y)) \qquad 1 \\ = 0 \leqslant x \leqslant y \leqslant x+2 \end{aligned}$$

$$\begin{array}{l} I^4(x,y) = I^2(x,y) \lor I^3(x,y) \leftarrow \text{widening} \qquad 2 \\ = 0 \leqslant x \leqslant y \end{aligned}$$

$$\begin{array}{l} I^5(x,y) = x \geqslant 0 \land (x = y \lor I^4(x+1,y)) \\ = I^4(x,y) \qquad \text{fixed point!} \end{aligned}$$
The invariants are computer representable

- x

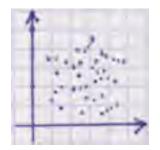
- x

-x

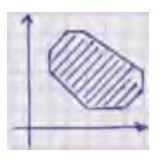
with octagons!



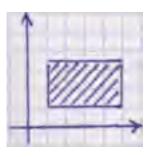
Examples of abstractions used by ASTRÉE



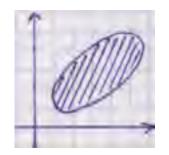
semantics set of points



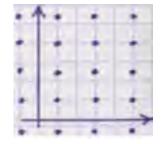
octagons



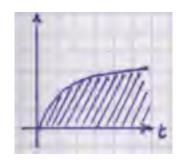
intervals $x \in [a,b]$



ellipsoids $\pm x \pm y \leqslant a \qquad ax^2 + by^2 + cxy \leqslant d$



simple congruences $x \equiv a[b]$



exponentials $x(t)\leqslant a^{bt}$



4. Scaling up



The difficulty of scaling up

- The abstraction must be coarse enough to be effectively computable with reasonable resources
- The abstraction must be precise enough to avoid false alarms
- Abstractions to infinite domains with widenings are more expressive than abstractions to finite domains (when considering the analysis of a programming language) [CC92]
- Abstractions are ultimately incomplete (even intrinsically for some semantics and specifications [CC00])







Abstraction/refinement by tuning the cost/precision ratio in ASTRÉE

- Approximate reduced product of a choice of coarsenable/refinable abstractions
- Tune their precision/cost ratio by
 - Globally by parametrization
 - Locally by (automatic) analysis directives
 - so that the overall abstraction is \underline{not} uniform.



Example of abstract domain choice in ASTRÉE

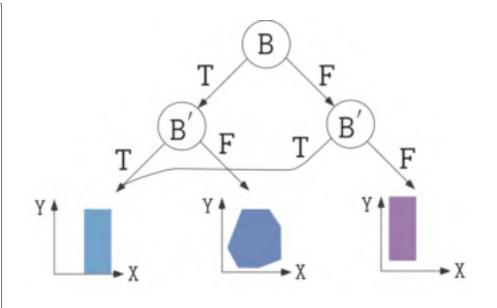
/* Launching the forward abstract interpreter */
/* Domains: Guard domain, and Boolean packs (based on Absolute
value equality relations, and Symbolic constant propagation
(max_depth=20), and Linearization, and Integer intervals, and
congruences, and bitfields, and finite integer sets, and Float
intervals), and Octagons, and High_passband_domain(10), and
Second_order_filter_domain (with real roots)(10), and
Second_order_filter_domain (with complex roots)(10), and
Arithmetico-geometric series, and new clock, and Dependencies
(static), and Equality relations, and Modulo relations, and
Symbolic constant propagation (max_depth=20), and Linearization,
and Integer intervals, and congruences, and bitfields, and
finite integer sets, and Float intervals. */



Example of abstract domain functor in ASTRÉE: decision trees

- Code Sample:

```
/* 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 leafs

P. Cousot 💁

 (\mathbf{C})

Reduction [CC79, CCF⁺08]

Example: reduction of intervals by simple congruences

```
% cat -n congruence.c
     1 /* congruence.c */
     2 int main()
     3 \{ int X; \}
     4 \quad X = 0;
     5 while (X <= 128)
     6 \quad \{ X = X + 4; \};
     7 __ASTREE_log_vars((X));
     8 }
% astree congruence.c -no-relational -exec-fn main |& egrep "(WARN)|(X in)"
direct = <integers (intv+cong+bitfield+set): X in {132} >
Intervals : X \in [129, 132] + \text{congruences} : X = 0 \mod 4 \Longrightarrow
X \in \{132\}.
```

(c) P. Cousot

Parameterized abstractions

- Parameterize the cost / precision ratio of abstractions in the static analyzer
- Examples:
 - array smashing: --smash-threshold n (400 by default) \rightarrow smash elements of arrays of size > n, otherwise individualize array elements (each handled as a simple variable).
 - packing in octogons: (to determine which groups of variables are related by octagons and where)
 - --fewer-oct: no packs at the function level,
 - \cdot --max-array-size-in-octagons n: unsmashed array elements of size > n don't go to octagons packs

Parameterized widenings

- Parameterize the rate and level of precision of widenings in the static analyzer
- Examples:
 - delayed widenings: --forced-union-iterations-at-beginning n (2 by default)
 - thresholds for widening (e.g. for integers):

```
let widening_sequence =
```

```
[ of_int 0; of_int 1; of_int 2; of_int 3; of_int 4; of_int 5;
of_int 32767; of_int 32768; of_int 65535; of_int 65536;
of_string "2147483647"; of_string "2147483648";
of_string "4294967295" ]
```



Analysis directives

- Require a local refinement of an abstract domain
- Example:

```
% cat repeat1.c
typedef enum {FALSE=0,TRUE=1} BOOL;
int main () {
  int x = 100; BOOL b = TRUE;
  while (b) {
   x = x - 1;
    b = (x > 0);
  }
}
% astree -exec-fn main repeat1.c |& egrep "WARN"
repeat1.c:5.8-13::[call#main@2:loop@4>=4:]: WARN: signed int arithmetic
range [-2147483649, 2147483646] not included in [-2147483648, 2147483647]
%
```



Example of directive (cont'd)

```
% cat repeat2.c
typedef enum {FALSE=0,TRUE=1} BOOL;
int main () {
    int x = 100; BOOL b = TRUE;
    __ASTREE_boolean_pack((b,x));
    while (b) {
        x = x - 1;
        b = (x > 0);
    }
}
% astree -exec-fn main repeat2.c |& egrep "WARN"
%
```

The insertion of this directive could be automated in ASTRÉE (if the considered family of programs has "repeat" loops).



Automatic analysis directives

- The directives can be inserted automatically by static analysis

- Example:

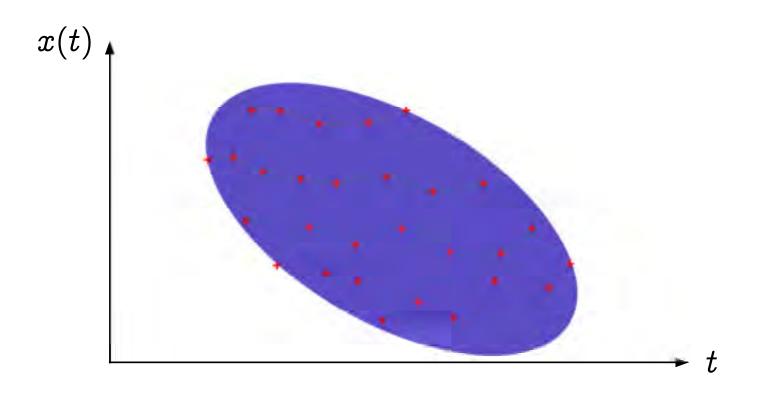
```
% cat p.c
                                              % astree -exec-fn main p.c -dump-partition
int clip(int x, int max, int min) {
                                              . .
 if (max >= min) {
                                              int (clip)(int x, int max, int min)
  if (x \le max) {
                                              {
   max = x;
                                                if ((max >= min))
  }
                                                { __ASTREE_partition_control((0))
  if (x < min) {
                                                  if ((x \le max))
   \max = \min;
                                                   {
  }
                                                     max = x;
 }
                                                   }
 return max;
                                                  if ((x < min))
void main() {
                                                     \max = \min;
 int m = 0; int M = 512; int x, y;
                                                   }
 y = clip(x, M, m);
                                                  __ASTREE_partition_merge_last(());
  __ASTREE_assert(((m<=y) && (y<=M)));
                                                }
}
                                                return max;
% astree -exec-fn main p.c |& grep WARN
                                              }
%
                                              . . .
                                              %
 Airbus, 12/04/2008
                                  — 22 —
                                                                P. Cousot
```

Adding new abstract domains

- The weakest invariant to prove the specification may not be expressible with the current refined abstractions \Rightarrow false alarms cannot be solved
- No solution, but adding a new abstract domain:
 - representation of the abstract properties
 - abstract property transformers for language primitives
 - widening
 - reduction with other abstractions
- Examples : ellipsoids for filters, exponentials for accumulation of small rounding errors, quaternions, ...



Abstraction by ellipsoid for filters



Ellipsoids $(x-a)^2+(y-b)^2\leq c$

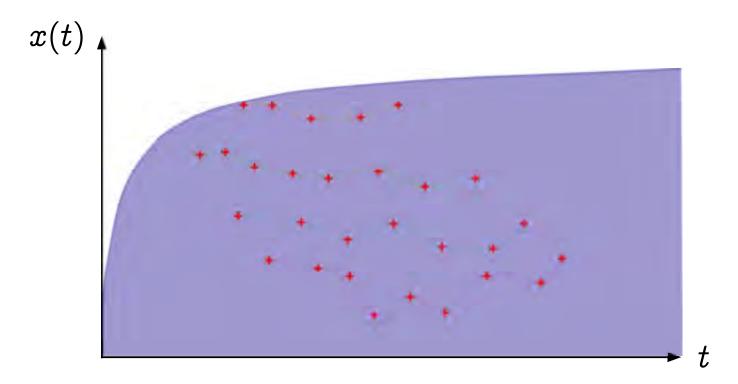


Example of analysis by ASTRÉE

```
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; }
}
```



Abstraction by exponentials for accumulation of small rounding errors



Exponentials $a^x \leq y$





Example of analysis by ASTRÉE

```
% 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)) \}
                  * 5.0e-03)); };
  B = A;
  if (SWITCH) \{A = P;\}
  else {A = X;}
}
```

```
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));
astree -exec-fn main -config-sem retro.config
retro.c |& grep "|P|" | tail -n 1
|P| <=1.000002*((15. +
5.8774718e-39/(1.0000002-1))*(1.0000002)ĉlock -
5.8774718e-39/(1.0000002-1)) + 5.8774718e-39 <=
23.039353
```

6. Industrial Application of ASTRÉE



Industrial results obtained with ASTRÉE

- Automatic proofs of absence of runtime errors in Electric Flight Control Soft-

ware:

- A340/600: 132.000 lines of C, 40mn on a PC 2.8 GHz, 300 Mb (Nov. 2003)
- A380: 1.000.000 lines of C, 34h, 8 Gb (Nov. 2005) no false alarm, World premières !
- Automatic proofs of absence of runtime errors in the ATV software⁽²⁾:
 - C version of the automatic docking software: 102.000 lines of C, 23s on a Quad-Core AMD Opteron[™] processor, 16 Gb (Apr. 2008)



⁽²⁾ the Jules Vernes Automated Transfer Vehicle (ATV) enabling ESA to transport payloads to the International Space Station.

7. Other Prototypes Developped by the ABSTRACTION project/team



Lcertify

- ASTRÉE proves the absence of runtimes errors on C code (with semantics tailored for a given Intel/PPC 32/64 bits machine)
- In absence of runtime errors, LCERTIFY proves the semantic equivalence of the C and machine code
- \Rightarrow the certification is on the flying code.

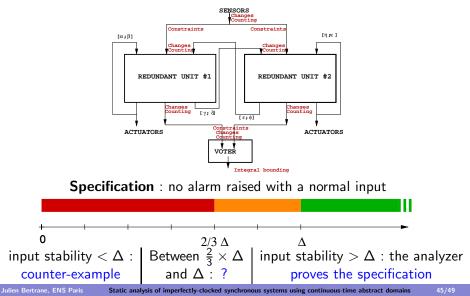


Macro generation

- Most of the control/command code (75%?) is automatically generated from SCADE/SAO, but for the primitives (handcoded by macros)
- Developed a formal specification language to specify the primitive semantics (close to the gascon specification)
- Developped tools
 - to automatically generate C macro code from the specification
 - to verify manual/automatic code with respect to the specification



Static analysis of communicating imperfectly clocked redundant units







Static Analysis of Parallel Code

- Some applications make use of parallel code



- Extremely hard to verify, whichever method is chosen!
- Work on static analysis of parallel code is in progress.



Conclusion

- Vision: to understand the numerical world, different levels of abstraction must be considered
- Theory: abstract interpretation ensures the coherence between abstractions and offers effective approximation techniques to cope with infinite systems
- Applications: the choice of effective abstraction which are coarse enough to be *computable* and precise enough to be *avoid false alarms* is central to master undecidability and complexity in model and program verification
- Software engineering : Manual validation by control of the software design process will ultimately be complemented by the verification of the final product



THE END

Thank you for your attention



9. Bibliography



Short bibliography

- [BCC⁺03] B. Blanchet, P. Cousot, R. Cousot, J. Feret, L. Mauborgne, A. Miné, D. Monniaux, and X. Rival. A static analyzer for large safety-critical software. In *Proceedings of the ACM SIGPLAN '2003 Conference on Programming Language Design and Implementation (PLDI)*, pages 196–207, San Diego, California, United States, 7–14 June 2003. ACM Press, New York, New York, United States.
- [CC77] P. Cousot and R. 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, New York, United States.
- [CC79] P. Cousot and R. 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, New York, United States.
- [CC92] P. Cousot and R. Cousot. Comparing the Galois connection and widening/narrowing approaches to abstract interpretation, invited paper. In M. Bruynooghe and M. Wirsing, editors, Proceedings of the Fourth International Symposium on Programming Language Implementation and Logic Programming, PLILP '92, Leuven, Belgium, 26-28 August 1992, Lecture Notes in Computer Science 631, pages 269-295. Springer, Berlin, Germany, 1992.



- [CC00] P. Cousot and R. Cousot. Temporal abstract interpretation. In Conference Record of the Twentyseventh Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, pages 12–25, Boston, Massachusetts, United States, January 2000. ACM Press, New York, New York, United States.
- [CCF⁺07] P. Cousot, R. Cousot, J. Feret, L. Mauborgne, A. Miné, D. Monniaux, and X. Rival. Varieties of static analyzers: A comparison with ASTRÉE, invited paper. In M. Hinchey, He Jifeng, and J. Sanders, editors, Proceedings of the First IEEE & IFIP International Symposium on Theoretical Aspects of Software Engineering, TASE '07, pages 3–17, Shanghai, China, 6–8 June 2007. IEEE Computer Society Press, Los Alamitos, California, United States.
- [CCF⁺08] P. Cousot, R. Cousot, J. Feret, L. Mauborgne, A. Miné, D. Monniaux, and X. Rival. Combination of abstractions in the ASTRÉE static analyzer. In M. Okada and I. Satoh, editors, *Eleventh Annual Asian Computing Science Conference, ASIAN06*, pages 272–300, Tokyo, Japan, 6–8 December 2006, 2008. Lecture Notes in Computer Science 4435, Springer, Berlin, Germany.
- [DS07] D. Delmas and J. Souyris. ASTRÉE: from research to industry. In G. Filé and H. Riis-Nielson, editors, Proceedings of the Fourteenth International Symposium on Static Analysis, SAS '07, Kongens Lyngby, Denmark, Lecture Notes in Computer Science 4634, pages 437-451. Springer, Berlin, Germany, 22-24 August 2007.

