Refining Model Checking by Abstract Interpretation

Patrick COUSOT
DMI, École normale supérieure
45 rue d'Ulm
75230 Paris cedex 05
France
cousot@dmi.ens.fr
http://www.dmi.ens.fr/~cousot

Udine, Italy, Gio. 24 Sep. 1998, 15h00–16h00

Combining model-checking and abstract interpretation

Why?

- Model-checking:
  - finite state space;
  - sound and complete property verification.
- Abstract Interpretation:
  - infinite state space;
  - Sound but uncomplete property determination.

Combining model-checking and abstract interpretation

How?

1. *Model abstraction:*
   - The finite model is an abstraction of the system;

     ⇒ Exact properties of an approximate model.
2. *Abstract symbolic methods:*
   - Use symbolic representations of properties (BDDs, convex polyhedra, ...);
   - One can make approximations (e.g. widenings);

     ⇒ Approximate properties of an exact model.

A new combination... [1]

3. *Parallel combination of model-checking and abstract interpretation:*
   - Model-checking:
     * Exact symbolic representation of properties;
     * The model is an exact representation of the system;

       ⇒ Exact properties of exact model
   - Abstract interpretation:
     * Preliminary/parallel analysis of the model by abstract interpretation;

       ⇒ Limit the state search space.

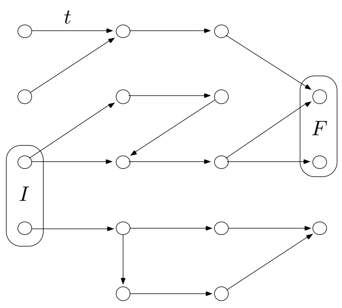
   ⟹ Exact properties of an exact sub-model.

[1] P. Cousot and R. Cousot. Refining Model Checking by Abstract Interpretation. *Automated Software Engineering*, 6(1), 1999, to appear.

Transition systems

Transition system $\langle S,\ t,\ I,\ F\rangle$:

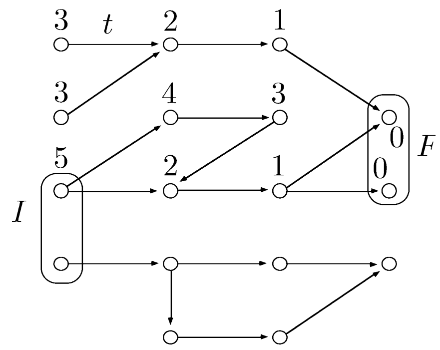
- S : set of states
- $t \subseteq S \times S$: transition relation
- $I \subseteq S$: set of initial states
- $F \subseteq S$: set of final states ($I \cap F = \emptyset$)



○: state, $\longrightarrow$: transition

Example: maximum delay problem [2]

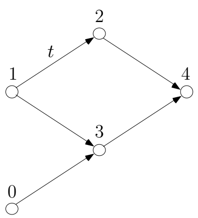
Find the maximum delay to reach a final state starting from some initial state:



[2] Halbwachs, N. Delays analysis in synchronous programs. CAV'93, LNCS 697, 1993, pp. 333–346.

Pre- and post-images

$$
\begin{array}{rcll}
\mathsf{pre}[t]\,P & \stackrel{\text{def}}{=} & \{s \mid \exists s' : \langle s,\ s'\rangle \in t \wedge s' \in P\} & \text{pre-image} \\
\widetilde{\mathsf{pre}}[t]\,P & \stackrel{\text{def}}{=} & \neg\,\mathsf{pre}[t](\neg P) & \text{dual pre-image} \\
& = & \{s \mid \forall s' : \langle s,\ s'\rangle \in t \Longrightarrow s' \in P\} & \\
\mathsf{post}[t]\,P & \stackrel{\text{def}}{=} & \{s' \mid \exists s : s \in P \wedge \langle s,\ s'\rangle \in t\} & \text{post-image} \\
\widetilde{\mathsf{post}}[t]\,P & \stackrel{\text{def}}{=} & \neg\,\mathsf{post}[t](\neg P) & \text{dual post-image} \\
& = & \{s' \mid \forall s : \langle s,\ s'\rangle \in t \Longrightarrow s \in P\} &
\end{array}
$$



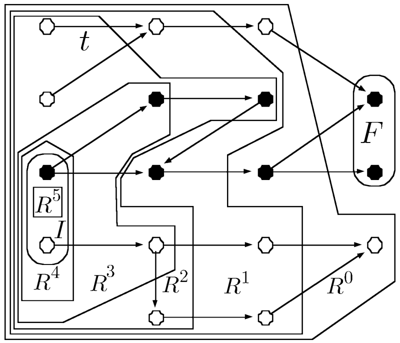
$$
\begin{array}{rcl}
\mathsf{pre}[t]\{3\} & = & \{0, 1\} \\
\widetilde{\mathsf{pre}}[t]\{3\} & = & \{0, 4\} \\
\mathsf{post}[t]\{1\} & = & \{2, 3\} \\
\widetilde{\mathsf{post}}[t]\{1\} & = & \{0, 1, 2\}
\end{array}
$$

Maximum delay algorithm "maximum1" [3]

procedure maximum1 $(I,\ F)$;
$R' := S$;
$n := 0$;
$R := (S - F)$;
while $(R \neq R' \wedge R \cap I \neq \emptyset)$ **do**
  $R' := R$;
  $n := n + 1$;
  $R := \mathsf{pre}[t]\,R' \cap (S - F)$;
od;
return if $(R' = R)$ **then** ∞ **else** n;

[3] Campos, S., Clarke, E., Marrero, W., and Minea, M. Verus: A tool for quantitative analysis of finite-state real-time systems. Proc. ACM SIGPLAN 1995 Workshop on Languages, Compilers & Tools for Real-Time Systems, La Jolla, Calif., jun 21–22, 1995, pp. 75–83.

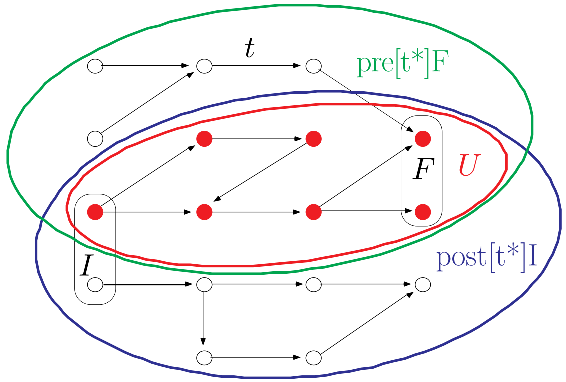
Execution trace of the "maximum1" algorithm



It is useless to explore the states which are not:

- descendants of the initial states;
- ancestors of the initial states.

Descendants of the initial states I which are ancestors of the initial states F

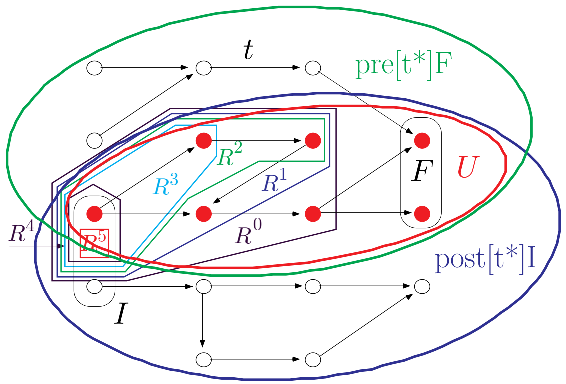


Maximum delay algorithm "maximum2" (with state search space restriction)

procedure maximum2 (I, F);
$R' := S$;
$n := 0$;
$R := \boxed{(U_0 - F)}$;
while $(R \neq R' \wedge R \cap I \neq \emptyset)$ **do**
  $R' := R$;
  $n := n + 1$;
  $R :=$ pre$[t]\ R' \cap \boxed{(U_n - F)}$;
od;
return if $(R' = R)$ **then** ∞ **else** n;

where: $\forall n \geq 0 : U_n \supseteq U \stackrel{\text{def}}{=}$ post$[t^\star]\ I \cap$ pre$[t^\star]\ F$

Execution trace of the "maximum2" algorithm



- Any upper-approximations $U_0, U_1, \ldots, U_n, \ldots$ of U can be used;
- In the worst case $U_n = S$, hence "maximum2" $=$ "maximum1".

Tarski's fixpoint theorem

A monotonic map $\varphi \in L \mapsto L$ on a complete lattice:

$$\langle L,\ \sqsubseteq,\ \bot,\ \top,\ \sqcup,\ \sqcap\rangle$$

has a least fixpoint:

$$\mathsf{lfp}\,\varphi = \sqcap\{x \in L \mid \varphi(x) \sqsubseteq x\}$$

(such that $\varphi(\mathsf{lfp}\,\varphi) = \mathsf{lfp}\,\varphi$ and $\varphi(x) = x$ implies $\mathsf{lfp}\,\varphi \sqsubseteq x$) and, dually, a greatest fixpoint:

$$\mathsf{gfp}\,\varphi = \sqcup\{x \in L \mid x \sqsubseteq \varphi(x)\}$$

Fixpoint characterization of the descendants of the initial states I which are ancestors of the initial states F

$$\begin{aligned}
\mathsf{pre}[t^\star]\,F &= \mathsf{lfp}^{\subseteq}\,\lambda X\boldsymbol{\cdot}\,F \cup \mathsf{pre}[t]\,X &&= \mathsf{lfp}^{\subseteq}_{F}\,\lambda X\boldsymbol{\cdot}\,X \cup \mathsf{pre}[t]\,X,\\
\widetilde{\mathsf{pre}}[t^\star]\,F &= \mathsf{gfp}^{\subseteq}\,\lambda X\boldsymbol{\cdot}\,F \cap \widetilde{\mathsf{pre}}[t]\,X &&= \mathsf{gfp}^{\subseteq}_{F}\,\lambda X\boldsymbol{\cdot}\,X \cap \widetilde{\mathsf{pre}}[t]\,X,\\
\mathsf{post}[t^\star]\,I &= \mathsf{lfp}^{\subseteq}\,\lambda X\boldsymbol{\cdot}\,I \cup \mathsf{post}[t]\,X &&= \mathsf{lfp}^{\subseteq}_{I}\,\lambda X\boldsymbol{\cdot}\,X \cup \mathsf{post}[t]\,X,\\
\widetilde{\mathsf{post}}[t^\star]\,I &= \mathsf{gfp}^{\subseteq}\,\lambda X\boldsymbol{\cdot}\,I \cap \widetilde{\mathsf{post}}[t]\,X &&= \mathsf{gfp}^{\subseteq}_{I}\,\lambda X\boldsymbol{\cdot}\,X \cap \widetilde{\mathsf{post}}[t]\,X\ .
\end{aligned}$$

Kleenian fixpoint theorem [4]

The least fixpoint of an upper-continuous map $\varphi \in L \mapsto L$ on a cpo $\langle L,\ \sqsubseteq,\ \bot,\ \sqcup\rangle$ is:

$$\mathsf{lfp}\,\varphi = \bigsqcup_{n \geq 0} \varphi^n(\bot)$$

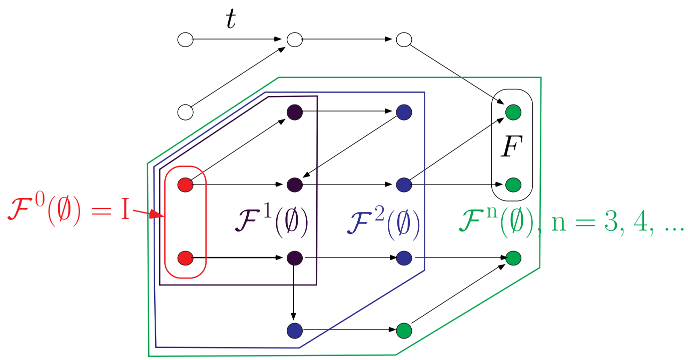
where the iterates $\varphi^n(x)$ of φ from x are:

- $\varphi^0(x) \stackrel{\text{def}}{=} x$;
- $\varphi^{n+1}(x) \stackrel{\text{def}}{=} \varphi(\varphi^n(x))$ for all $x \in L$.

[4] Can be generalized to monotonic non-continuous maps.

Iterative characterization of the descendants of the initial states I

$$\mathsf{post}[t^\star]\,I = \mathsf{lfp}^{\subseteq}\,\mathcal{F} = \bigcup_{n \in \mathbb{N}} \mathcal{F}^n(\emptyset) \qquad \text{where} \quad \mathcal{F}(X) = I \cup \mathsf{post}[t]\,X$$



Analysis of the model by abstract interpretation

- We can compute:

$$U_0 \supseteq U_1 \supseteq \ldots \supseteq U_n \supseteq U \stackrel{\text{def}}{=} \text{post}[t^\star]\, I \cap \text{pre}[t^\star]\, F$$

by abstract interpretation;
- The abstract interpretation can be done in parallel with the model-checking (at almost no supplementary cost);
- The abstract interpretation results are used on the fly for U_n as they become available to restrict the state search space;
- Several restriction operators have been proposed for symbolic model checking (with BDDs (cofactor, constrain, restrict) & convex polyhedra[5]).

[5] Halbwachs, N. and Raymond, P. On the use of approximations in symbolic model checking. Tech. rep. SPECTRE L21 (jan 1996), VERIMAG laboratory, Grenoble, France.

Upper approximation D of $\text{post}[t^\star]\, I = \text{lfp}^{\subseteq}\, \lambda X \cdot I \cup \text{post}[t]\, X$ by abstract interpretation[6]

1. Consider an abstract domain $\langle L, \sqsubseteq\rangle$ approximating sets of states $\langle \wp(S), \subseteq\rangle$;
2. define a correspondence:

$$\langle \wp(S), \subseteq\rangle \underset{\alpha}{\overset{\gamma}{\leftrightarrows}} \langle L, \sqsubseteq\rangle$$

which is a Galois connection:

$$\forall P \in \wp(S) : \forall Q \in L : \alpha(P) \sqsubseteq Q \Longleftrightarrow P \subseteq \gamma(Q) .$$

The abstract value $\alpha(P)$ is the approximation of $P \subseteq S$: $P \subseteq \gamma(\alpha(P))$.

[6] Cousot, P. and Cousot, R. Abstract interpretation: a unified lattice model for static analysis of programs by construction or approximation of fixpoints. 4th POPL, Los Angeles, 1977, pp. 238–252.

3. Define an abstract post-image transformer $\mathcal{F} \in L \stackrel{\text{m}}{\longmapsto} L$:

$$\forall Q \in L : \alpha \circ (\lambda X \cdot I \cup \text{post}[t]\, X) \circ \gamma(Q) \sqsubseteq \mathcal{F}(Q)$$

4. Define a *widening operator* $\triangledown \in L \times L \mapsto L$:
   - it is an upper approximation:

     $\forall x, y \in L : x \sqsubseteq x \triangledown y$ and $\forall x, y \in L : y \sqsubseteq x \triangledown y$.
   - it enforces finite convergence of $\mathcal{F}$-upward iterates:

     for all increasing chains $x^0 \sqsubseteq x^1 \sqsubseteq \ldots \sqsubseteq x^i \sqsubseteq \ldots$ the increasing chain defined by $y^0 = x^0, \ldots, y^{i+1} = y^i \triangledown x^{i+1}, \ldots$ is not strictly increasing.

5. The *upward forward iteration sequence with widening*:
   - $\hat{\mathcal{F}}^0 \stackrel{\text{def}}{=} \alpha(\emptyset)$,
   - $\hat{\mathcal{F}}^{i+1} \stackrel{\text{def}}{=} \hat{\mathcal{F}}^i$ if $\mathcal{F}(\hat{\mathcal{F}}^i) \sqsubseteq \hat{\mathcal{F}}^i$
   - $\hat{\mathcal{F}}^{i+1} \stackrel{\text{def}}{=} \hat{\mathcal{F}}^i \triangledown \mathcal{F}(\hat{\mathcal{F}}^i)$ otherwise

   is ultimately stationary; Its limit $\hat{\mathcal{F}}$ is a sound upper approximation of $\text{post}[t^\star]\, I$ in that:

$$\text{post}[t^\star]\, I \subseteq \gamma(\text{lfp}^{\sqsubseteq}\, \mathcal{F}) \subseteq \gamma(\hat{\mathcal{F}}) .$$

6. Define a *narrowing operator* $\triangle \in L \times L \mapsto L$ such that:
   - it is an upper approximation
     $\forall x, y \in L : x \sqsubseteq y \Longrightarrow x \sqsubseteq x \triangle y \sqsubseteq y.$
   - it enforces finite convergence of $\mathcal{F}$-downward iterates:
     For all decreasing chains $x^0 \sqsupseteq x^1 \sqsupseteq \ldots$ the decreasing chain defined by $y^0 = x^0, \ldots, y^{i+1} = y^i \triangle x^{i+1}, \ldots$ is not strictly decreasing.

7. the *downward forward iteration sequence with narrowing*:
   - $\check{\mathcal{F}}^0 \stackrel{\text{def}}{=} \hat{\mathcal{F}}$,
   - $\check{\mathcal{F}}^{i+1} \stackrel{\text{def}}{=} \check{\mathcal{F}}^i$ if $\mathcal{F}(\check{\mathcal{F}}^i) = \check{\mathcal{F}}^i$
   - $\check{\mathcal{F}}^{i+1} \stackrel{\text{def}}{=} \check{\mathcal{F}}^i \triangle \mathcal{F}(\check{\mathcal{F}}^i)$ otherwise

   is ultimately stationary;
   its limit $\check{\mathcal{F}}$ is a better sound upper approximation $\mathsf{post}[t^\star]\, I$ in that:

$$\mathsf{post}[t^\star]\, I \subseteq \gamma(\mathsf{lfp}^{\sqsubseteq}\, \mathcal{F}) \subseteq \gamma(\check{\mathcal{F}}) \subseteq \gamma(\hat{\mathcal{F}})\ .$$

Abstract interpretation design

- The design of:
  - the abstract algebra $\langle L, \sqsubseteq, \bot, \top, \sqcup, \sqcap, \triangledown, \triangle, f_1, \ldots, f_n\rangle$
  - the transformer $\mathcal{F}$ (usually composed out of the primitives f_1, …, f_n)

  are problem dependent;
- Natural choices in the model-checking context are:
  - BDDs (discrete systems),
  - Convex polyhedra (hybrid systems);

  for which widening operators have been defined [7, 8].

[7] Mauborgne, L. Abstract interpretation using typed decision graphs. Sci. Comput. Prog., 31(1):91–112, 1998.
[8] Cousot, P. and Halbwachs, N. Automatic discovery of linear restraints among variables of a program. In 5th POPL, Tucson, 1978, pp. 84–97.

Upper approximation A of $\mathsf{pre}[t^\star]\, F = \mathsf{lfp}^{\subseteq}\, \lambda X \cdot F \cup \mathsf{pre}[t]\, X$ by abstract interpretation [9]

Use the same abstract algebra $\langle L, \sqsubseteq, \bot, \top, \sqcup, \sqcap, \triangledown, \triangle, f_1, \ldots, f_n\rangle$:

8. Define an abstract pre-image transformer $\mathcal{B} \in L \stackrel{\mathrm{m}}{\longmapsto} L$:

$$\forall Q \in L : \alpha \circ (\lambda X \cdot F \cup \mathsf{pre}[t]\, X) \circ \gamma(Q) \sqsubseteq \mathcal{B}(Q)$$

9. First use an *upward backward iteration sequence with widening* finitely converging to $\hat{\mathcal{B}}$;
10. Improve by a *downward iteration sequence with narrowing* finitely converging to $\check{\mathcal{B}}$ such that:

$$\mathsf{pre}[t^\star]\, F = \mathsf{lfp}^{\subseteq}\, \lambda X \cdot F \cup \mathsf{pre}[t]\, X \subseteq \gamma(\mathsf{lfp}^{\sqsubseteq}\, \mathcal{B}) \subseteq \gamma(\check{\mathcal{B}}) \subseteq \gamma(\hat{\mathcal{B}})$$

[9] Cousot, P. and Cousot, R. Systematic design of program analysis frameworks. In 6th POPL, San Antonio, 1979, pp. 269–282.

Sequence of upper approximations $U_0, U_1, \ldots, U_n, \ldots$ of $U = \textbf{post}[t^\star]\, I \cap \textbf{pre}[t^\star]\, F$ by abstract interpretation[10, 11]

- $U_0 = S$, all states;
- U_1 is the γ-concretization of the limit of the upward forward iteration sequence with widening for $\mathcal{F}$;
- U_2 is the γ-concretization of the limit of the corresponding downward forward iteration sequence with narrowing for $\mathcal{F}$ starting from U_0;
- $\cdots$

[10] Cousot, P. Méthodes itératives de construction et d'approximation de points fixes d'opérateurs monotones sur un treillis, analyse sémantique de programmes. Ph. D. thesis, Université scientifique et médicale de Grenoble, 1978.

[11] Cousot, P. and Cousot, R. Abstract interpretation and application to logic programs. J. Logic Prog. 13, 2–3, 103–179. (The editor of JLP has mistakenly published the unreadable galley proof. For a correct version of this paper, see `http://www.ens.fr/~cousot`.)

- $\cdots$
- U^{4n+3} is the γ-concretization of the limit of the upward backward iteration sequence with widening for $\lambda X \cdot (U^{4n+2} \sqcap \mathcal{B}(X))$;
- U^{4n+4} is the γ-concretization of the limit of the corresponding downward backward iteration sequence with narrowing for $\lambda X \cdot (U^{4n+2} \sqcap \mathcal{B}(X))$ starting from U^{4n+3};
- U^{4n+5} is the γ-concretization of the limit of the upward forward iteration sequence with widening for $\lambda X \cdot (U^{4n+4} \sqcap \mathcal{F}(X))$;
- U^{4n+6} is the γ-concretization of the limit of the corresponding downward forward iteration sequence with narrowing for $\lambda X \cdot (U^{4n+4} \sqcap \mathcal{F}(X))$ starting from U^{4n+5};
- $\cdots$

Correctness

- The sequence $U_0, U_1, U_2, \ldots, U^{4n+3}, U^{4n+4}, U^{4n+5}, U^{4n+6}, \ldots$ is a descending chain;

$\Rightarrow$ The restriction is more and more precise as the model-checking goes on;

- All elements U_k is the sequence are sound:

$$\text{post}[t^\star]\, I \cap \text{pre}[t^\star]\, F \subseteq U_k$$

- Stop the abstract interpretation computation with a narrowing or when the parallel model-checking terminates;

Parallel programming

$\boxed{[\![\ \underline{\text{analysis}}\ ()\ \|\ \textbf{return}\ \underline{\text{maximum2}}\ (I,\ F)\ ]\!]}$

Communication:

- "**send**(V);" and "**receive**(U);" : asynchronous one-place buffered communication where the buffer is initialized to the supremum $\top = \alpha(S)$
- "**send**(V);" replaces the current value of the buffer with V
- "**receive**(U);" assigns to U the current value of the buffer which is left unchanged.

Maximum delay

function maximum2 (I, F);
$R' := S$;
$n := 0$;
receive(U);
$R := \boxed{(U - F)}$;
while $(R \neq R' \wedge R \cap I \neq \emptyset)$ **do**
  $R' := R$;
  $n := n + 1$;
  receive(U);
  $R :=$ pre$[t]\ R' \cap \boxed{(U - F)}$;
od;
return if $(R' = R)$ **then** ∞ **else** n.

Analysis by abstract interpretation

function fixapp (φ, P); $X := \bot$; **loop** $Y := X$; $X := \varphi(Y) \sqcap P$; **exit if** $X \sqsubseteq Y$ $X := (Y \nabla X) \sqcap P$; **forever**; **send**$(\gamma(Y))$; **while** $X \neq Y$ **do** $Y := X$; $X := \varphi(Y) \sqcap P$; $X := (Y \triangle X) \sqcap P$; **od**; **return** X.	**function** analysis (); $D := \top$; $A := \top$; **repeat** $D :=$ fixapp $(\mathcal{F}, A)$; **send**$(\gamma(D))$; $A :=$ fixapp $(\mathcal{B}, D)$; **send**$(\gamma(A))$; **until** $A = D$; **return** $\gamma(A)$.

Problematic termination

- The abstract interpretation always terminate;
- The abstract interpretation is approximate so the state-space restriction may not be finite;

$\Rightarrow$ The parallel combination of abstract interpretation and model-checking is incomplete since it may not terminate;

- In case of nontermination the information gathered by abstract interpretation is reusable for verification by:
  - abstract symbolic methods,
  - model abstraction;

  which are also incomplete but guarantee termination.

Conclusion

- We have proposed a method for the parallel combination of model-analysis by abstract interpretation and verification by model-checking where the verification:
  - makes no approximation on states and transitions,
  - explores an (hopefully finite) subgraph;
- Semi-algorithm since there is no guarantee that the explored subgraph will be finite:
  - classical model-checking would have failed anyway,
  - case by case experimentation is needed;
- The method should be used before resorting to model-checking of a more abstract model (the information gathered about the exact model being reusable).