Master Parisien de Recherche en Informatique Année scolaire 2010/2011
École normale supérieure

Cours M.2-6

« Interprétation abstraite: applications à la vérification et à l'analyse statique »

Examen partiel

Patrick Cousot

26 novembre 2010

Course and personal notes are the only allowed documents. It will not be answered to any question during the exam. If a question is ambiguous, imprecise or incorrect, it is part of the question to solve the ambiguity, imprecision or incorrectness by indicating all required hypotheses together with the solution, if any. All questions are independent and can be answered in any order.

Let us recall the following theorem:

Theorem 1 *If $\langle L, \leqslant, \bot\rangle$ is a cpo, $F \in L \to L$ is monotonically increasing, $\langle \overline{L}, \sqsubseteq\rangle$ is a poset, $\alpha \in L \to \overline{L}$ is continuous*[1,2]*, $\overline{F} \in \overline{L} \to \overline{L}$ commutes (resp. semi-commutes) with F that is $\alpha \circ F = \overline{F} \circ \alpha$*[3] *(resp. $\alpha \circ F \sqsubseteq \overline{F} \circ \alpha$) then $\alpha(\mathbf{lfp}^{\leqslant}_{\bot} F) = \mathbf{lfp}^{\sqsubseteq}_{\alpha(\bot)} \overline{F}$ (resp. $\alpha(\mathbf{lfp}^{\leqslant}_{\bot} F) \sqsubseteq \mathbf{lfp}^{\sqsubseteq}_{\alpha(\bot)} \overline{F}$).* ■

which may be useful in some questions.

Question 1

Sintzoff (1972) presents the rule of signs abstraction in the following way:

> "$a \times a + b \times b$ yields always the object "pos" when a and b are the objects "pos" or "neg", and when the valuation is defined as follows :
>
> $$\begin{array}{rcl@{\qquad}rcl}
\text{pos+pos} & = & \text{pos} & \text{pos} \times \text{pos} & = & \text{pos}\\
\text{pos+neg} & = & \text{pos,neg} & \text{pos} \times \text{neg} & = & \text{neg}\\
\text{neg+pos} & = & \text{pos,neg} & \text{neq} \times \text{pos} & = & \text{neg}\\
\text{neg+neg} & = & \text{neg} & \text{neg} \times \text{neg} & = & \text{pos}\\
\text{V(p+q)} & = & \text{V(p)+V(q)} & \text{V(p} \times \text{q)} & = & \text{V(p)} \times \text{V(q)}\\
\text{V(0)} & = & \text{V(1)} \;=\; \ldots \;=\; \text{pos} & & & \\
\text{V(-1)} & = & \text{V(-2)} \;=\; \ldots \;=\; \text{neg} & & &
\end{array}$$
>
> The valuation of $a \times a + b \times b$ yields "pos" by the following computation :

[1] α is *continuous* if and only if it preserves existing lubs of increasing chains.

[2] The continuity hypothesis for α can be restricted to the iterates $F^0 \triangleq \bot$, $F^{n+1} \triangleq F(F^n)$, $F^\omega \triangleq \bigsqcup_{n\geqslant} F^n$ of the least fixpoint of F.

[3] The commutation property $\alpha \circ F(x) = \overline{F} \circ \alpha(x)$ is only required for all $x \in L$ such that $\gamma \circ \alpha(x) \leqslant \mathbf{lfp}^{\leqslant} F$ or even just for the iterates of the least fixpoint of F.

$$\begin{array}{rcl@{\quad}rcl}
\mathrm{V}(a) & = & \text{pos,neg} & \mathrm{V}(b) & = & \text{pos,neg} \\
\mathrm{V}(a \times a) & = & \text{pos} \times \text{pos, neg} \times \text{neg} & \mathrm{V}(b \times b) & = & \text{pos} \times \text{pos, neg} \times \text{neg} \\
& = & \text{pos,pos} \;=\; \text{pos} & & = & \text{pos,pos} \;=\; \text{pos}
\end{array}$$

$$\mathrm{V}(a \times a + b \times b) \;=\; \mathrm{V}(a \times a){+}\mathrm{V}(b \times b) \;=\; \text{pos+pos} \;=\; \text{pos''}$$

What is wrong about it?

Question 2

A multiplication $m \times n = r$ can be checked by summing the digits of integer m modulo 9, summing the digits of n modulo 9, and checking that their product modulo 9 is equal to the sum of the digits of the result r modulo 9. For example,

$$\begin{array}{rr@{\quad}c@{\quad}l@{\quad}c@{\quad}l}
 & 1234 & \rightarrow & 10 \bmod 9 & = & 1 \\
\times & 5678 & \rightarrow & 26 \bmod 9 & = & 8 \\
\hline
= & 7006652 & \rightarrow & 26 \bmod 9 & = & 8
\end{array}$$

succeeds, while

$$\begin{array}{rr@{\quad}c@{\quad}l@{\quad}c@{\quad}l@{\quad}c@{\quad}l}
 & 1234 & \rightarrow & 10 \bmod 9 & = & 1 & & \\
\times & 5678 & \rightarrow & 26 \bmod 9 & = & 8 & & \\
\hline
= & 7006651 & \rightarrow & 27 \bmod 9 & = & 7 & \neq & 1 \times 8 \bmod 9
\end{array}$$

fails.

- Show that this *casting out nines* is an abstraction.
- Is it a proof[4] ?
- Can you cite a sound generalization of the idea used in program analysis?

Question 3

Define the *reflexive transitive closure* $r^\star$ of a relation $r \in \wp(S \times S)$ on a set S as $r^\star \triangleq \bigcup_{n \in \mathbb{N}} r^n$ where the *powers* r^n, $n \in \mathbb{N}$ are defined as $r^0 \triangleq \{\langle s,\ s\rangle \mid s \in S\}$ (which is the identity relation), $r^{n+1} \triangleq r \circ r^n$, and the *composition* of relations is $r \circ r' \triangleq \{\langle s,\ s''\rangle \in S \times S \mid \exists s' \in S : \langle s,\ s'\rangle \in r \wedge \langle s',\ s''\rangle \in r'\}$. Prove that the reflexive transitive closure $r^\star$ of the relation r is an abstraction of the partial trace semantics $\vec{r}$ of this relation r defined as

$$\begin{array}{rcl}
\vec{r}^{\,n} & \triangleq & \{\pi \in S^n \mid \forall i \in [0, n-1] : \langle \pi_i,\ \pi_{i+1}\rangle \in r\}, \quad n > 0 \\
\vec{r} & \triangleq & \displaystyle\bigcup_{n=1}^{+\infty} \vec{r}^{\,n}
\end{array}$$

Question 4

Prove that

Theorem 2 *If $\langle \overline{L},\ \overline{\sqsubseteq},\ \overline{\top}\rangle$ is a dcpo*[5]*, $\overline{F} \in \overline{L} \rightarrow \overline{L}$ is monotonically increasing, $\gamma \in \overline{L} \rightarrow L$ is co-continuous*[6]*, $F \in L \rightarrow L$ commutes with F that is $\gamma \circ \overline{F} = F \circ \gamma$ then $\gamma(\mathbf{gfp}^{\overline{\sqsubseteq}}_{\overline{\top}}\ \overline{F}) = \mathbf{gfp}^{\leqslant}_{\gamma(\overline{\top})}\ F$.* ∎

[4] In French it is called a "proof by 9".

[5] A dual complete partial order (dcpo) has glbs of decreasing chains.

[6] γ is *co-continuous* if and only if it preserves existing glbs of decreasing chains.

Question 5

Let $\langle L,\ \leqslant,\ \bot,\ \neg\rangle$ be a complete Boolean lattice (where $\neg$ is the unique complement). Prove that $\langle L,\ \leqslant\rangle \underset{\neg}{\overset{\neg}{\leftrightarrows}} \langle L,\ \geqslant\rangle$ is a Galois isomorphism.

Question 6

Using Th. 1 and Q. 5, prove the following theorem due to David Park (1969)

Theorem 3 *If $F \in L \to L$ is monotonically increasing on a complete Boolean lattice $\langle L,\ \leqslant,\ \bot,\ \neg\rangle$ then $\neg\ \boldsymbol{lfp}_{\bot}^{\leqslant}\ F = \boldsymbol{gfp}_{\neg\bot}^{\leqslant}\ \neg \circ F \circ \neg$.* ■

Question 7

Prove the following theorem providing a condition for fixpoints of an increasing map on a complete Boolean lattice to be unique.

Theorem 4 (D. Park) *Let $f \in \mathcal{L} \xrightarrow{\nearrow} \mathcal{L}$ be an increasing map on the complete Boolean lattice $\langle L,\ \sqsubseteq,\ \bot,\ \top,\ \sqcup,\ \sqcap,\ \neg\rangle$. Then*

(1) $\boldsymbol{lfp}\,\widetilde{f} \sqcap \boldsymbol{lfp}\,f = \bot$

(2) $(\boldsymbol{lfp}\,\widetilde{f} \sqcup \boldsymbol{lfp}\,f = \top) \quad\Leftrightarrow\quad (\boldsymbol{lfp}\,f = \boldsymbol{gfp}\,f)$

where $\widetilde{f} \triangleq \neg \circ f \circ \neg$. ■

Question 8

The interval analysis of the following program

$$\boldsymbol{P} \quad \triangleq \quad {}^{1}\texttt{x := 100 ; while }{}^{2}\texttt{(x =/= 0) do }{}^{3}\texttt{x := (x - 1); od}^{4}\,.$$

has the following interval equations

$$\left\{\begin{array}{lcl} X_1 & = & [\texttt{min\_int},\ \texttt{max\_int}]\} \\ X_2 & = & [100,\ 100] \sqcup \llparenthesis\, X_3 = \emptyset \mathrel{?} \emptyset \mathrel{:} \texttt{let}\,[a,\ b] = X_3\ \texttt{in} \\ & & \qquad [\max(a-1,\ \texttt{min\_int}), \max(b-1, \texttt{min\_int})]\,\rrparenthesis\} \\ X_3 & = & (X_2 \sqcap [\texttt{min\_int},\ -1]) \sqcup (X_2 \sqcap [1,\ \texttt{max\_int}]) \\ X_4 & = & X_2 \sqcap [0,\ 0] \end{array}\right.$$

The resolution of the equations by iteration with widening/narrowing yields a rather imprecise result.

```
% ocamlc interval.ml intervalWidening.ml intervalNarrowing.ml \
? invariant.ml invariantWidening.ml invariantNarrowing.ml \
? transformerBounded.ml iterator.ml \
? reachability_narrowing_bounded.ml
% time ./a.out
 1:(-1073741824,1073741823) 2:(-1073741824,100) 3:(-1073741824,100) 4:(0,0)
0.000u 0.000s 0:00.00 0.0%      0+0k 0+0io 0pf+0w
%
```

Propose a refinement of the interval widening to improve the precision of the analysis.

Question 9

Consider the syntax of the `repeat` command is

$$
\begin{array}{lcll}
\boldsymbol{C} & \in & \mathbb{C}, & \text{commands} \\
\boldsymbol{C} & ::= & \ldots & \\
& | & \texttt{repeat}\,\boldsymbol{C}\,\texttt{until}\,{}^{\ell}\boldsymbol{B} & \text{where } \ell \notin \mathsf{in}[\![\boldsymbol{C}]\!] \\
& & & \text{and } \mathsf{in}[\![\texttt{repeat}\,\boldsymbol{C}\,\texttt{until}\,{}^{\ell}\boldsymbol{B}]\!] \triangleq \{\ell\} \cup \mathsf{in}[\![\boldsymbol{C}]\!]
\end{array}
$$

Execution of the $\texttt{repeat}\,\boldsymbol{C}\,\texttt{until}\,{}^{\ell}\boldsymbol{B}$ command starts with that of the loop body $\boldsymbol{C}$

$$\mathsf{at}[\![\texttt{repeat}\,\boldsymbol{C}\,\texttt{until}\,{}^{\ell}\boldsymbol{B}]\!] \quad \triangleq \quad \mathsf{at}[\![\boldsymbol{C}]\!]$$

Execution of the loop body $\boldsymbol{C}$ ends at label ℓ just before evaluation of the condition $\boldsymbol{B}$

$$
\begin{array}{lcll}
\boldsymbol{C} & ::= & \ldots & \\
& | & \texttt{repeat}\,\boldsymbol{C}_1\,\texttt{until}\,{}^{\ell}\boldsymbol{B} & \mathsf{aft}[\![\texttt{repeat}\,\boldsymbol{C}\,\texttt{until}\,{}^{\ell}\boldsymbol{B}]\!] \triangleq \mathsf{aft}[\![\boldsymbol{C}]\!] \\
& & & \qquad\qquad\qquad \mathsf{aft}[\![\boldsymbol{C}_1]\!] \triangleq \ell
\end{array}
$$

Define the transitional semantics $\mathsf{T}[\![\texttt{repeat}\,\boldsymbol{C}\,\texttt{until}\,{}^{\ell}\boldsymbol{B}]\!]$ of the `repeat` command so that execution of the loop body $\boldsymbol{C}$ is repeated until the condition $\boldsymbol{B}$ is true.

Question 10

Let U be a universe and $F \in \wp(U) \xrightarrow{\nearrow} \wp(U)$ a $\subseteq$-increasing function on $\wp(U)$ defining $\mathbf{lfp}^{\emptyset}_{\subseteq} F \in \wp(U)$. What is the set R of inference rules such that the formal system $\langle U,\ R\rangle$ defines exactly the same set $\mathbf{lfp}^{\emptyset}_{\subseteq} F$?

Question 11

Let us consider the following program $\boldsymbol{P}$.

```
¹x := ? ;
while ²(1 < x) do
  ³x := x - 2
 od⁴.
```

Formally define the program property that once initialized the variable `x` keeps the same parity.

Question 12

Define the abstraction and concretization for the "bounding abstraction" of the trace semantics which cuts traces at a given depth n, so that the abstract semantics has all its traces of length at most n. This abstraction is left implicit in bounded model-checking.

Question 13

Consider the following *transition abstraction* from sets of traces to a transition relation.

$$
\begin{array}{lcl}
\alpha^{\tau} & \in & \wp(\mathcal{S}^{+}) \mapsto \wp(\mathcal{S} \times \mathcal{S}) \\
\alpha^{\tau}(T) & \triangleq & \{\langle \pi_i,\ \pi_{i+1}\rangle \mid \exists n \geqslant 1 : \pi \in T \cap \mathcal{S}^{n} \wedge 0 \leqslant i < n - 1\}
\end{array}
$$

1. Provide an example proving that this abstraction can loose information on the set of traces.
2. Provide a characterization of those sets of traces for which the abstraction loose no information.

Question 14

Prove that given $h \in \mathfrak{X} \in\mapsto \wp(\mathfrak{Y})$, defining $\alpha^h(P) \triangleq \bigcup\{h(x) \mid x \in P\}$ yields a Galois connection $\langle \wp(\mathfrak{X}),\ \subseteq\rangle \xrightleftharpoons[\alpha^h]{\gamma^h} \langle \wp(\mathfrak{Y}),\ \subseteq\rangle$ (this is the abstraction commonly used in model-checking).

Prove that any Galois connection $\langle \wp(\mathfrak{X}),\ \subseteq\rangle \xrightleftharpoons[\alpha]{\gamma} \langle \wp(\mathfrak{Y}),\ \subseteq\rangle$ can be put in that form for an appropriate choice of h such that $\alpha = \alpha^h$ and $\gamma = \gamma^h$.

Provide an example of abstraction $\langle \wp(\mathfrak{X}),\ \subseteq\rangle \xrightleftharpoons[\alpha]{\gamma} \langle A,\ \sqsubseteq\rangle$ that cannot be put in that form.

Question 15

Show that the program property "to be deterministic" (i.e. to have only one possible execution trace, either finite or infinite) is neither a safety nor a liveness property.

Question 16

Is the infinite union of safety properties a safety property?

Question 17

The abstract best transformer for the interval abstraction has been shown to be

$$[a_1, b_1] \overline{-} [a_2, b_2] \quad = \quad [a_1 - b_2, b_1 - a_2] \tag{1}$$

Assume that variable `x` has a value $x \in [-100, 100]$. For the value of the expression `x - x`, calculate $x \overline{-} x$ as indicated in (1). Can you do better? Why does $\overline{-}$ is so-called the best abstraction of $-$ on powersets?

Question 18

The partial trace semantics of a transition system $\langle \mathcal{S},\ \mathcal{I},\ \mathcal{F},\ \mathsf{T}\rangle$ is $\mathbf{P}^t \triangleq \{\pi \in \mathcal{S}^n \mid n \geqslant 1 \wedge \forall i \in [0, n-2] : \langle \pi_i,\ \pi_{i+1}\rangle \in \mathsf{T}\} = \mathbf{lfp}^{\subseteq}\, \mathsf{F}^t$ where $\mathsf{F}^t(X) \triangleq \mathcal{S}^1 \cup \{\pi\sigma\sigma' \mid \pi\sigma \in X \wedge \langle\sigma,\ \sigma'\rangle \in \mathsf{T}\}$. Ihe prefix trace semantics is defined as the restriction of the prefix trace semantics to traces starting with an initial state. $\mathbf{P}^{ti} \triangleq \{\pi \in \mathcal{S}^n \mid n \geqslant 1 \wedge \pi_0 \in \mathcal{I} \wedge \forall i \in [0, n-2] : \langle \pi_i,\ \pi_{i+1}\rangle \in \mathsf{T}\}$. This prefix trace semantics has the following fixpoint characterization $\mathbf{P}^{ti} = \mathbf{lfp}^{\subseteq}\, \mathsf{F}^{ti}$ where $\mathsf{F}^{ti}(X) \triangleq \{\pi \in \mathcal{S}^1 \mid \pi_0 \in \mathcal{I}\} \cup \{\pi\sigma\sigma' \mid \pi\sigma \in X \wedge \langle\sigma,\ \sigma'\rangle \in \mathsf{T}\}$.

Find an abstraction α such that $\mathbf{P}^{ti} = \alpha(\mathbf{P}^t)$. Then derive the fixpoint characterization of $\mathbf{P}^{ti}$ from that of $\mathbf{P}^t$ using Th. 1.

Question 19

Consider the polyhedral abstraction where the abstract properties P are the conjunction of linear inequalities $\bigwedge_{i=1}^{m} \sum_{j=1}^{n} a_i^j \leqslant b_i$ written

$$\begin{array}{rcl} P & = & Ax \leqslant b \\ & = & \{A_i x + b_i \mid i \in [1, m]\} \\ & = & \{a_i^1 x_1 + \ldots + a_i^j x_j + \ldots + a_i^n x_n \leqslant b_i \mid i \in [1, m]\} \\ & = & \{\displaystyle\sum_{j=1}^{n} a_i^j \leqslant b_i \mid i \in [1, m]\} \end{array}$$

We say that polyhedron P *entails* a constraint $\varphi = \sum_i a_i x_i \leqslant b_i$ when $\bigwedge P \Rightarrow \varphi$ that is $\{x \in \mathbb{Q}^n \mid Ax \leqslant b\} \subseteq \{x \in \mathbb{Q}^n \mid \sum_i a_i x_i \leqslant b_i\}$, written $P \models \varphi$.

The concretization is

$$\gamma^P(P) \quad \triangleq \quad \gamma^P(\langle A,\ b\rangle) \quad \triangleq \quad \{x \in \mathbb{Q}^n \mid Ax \leqslant b\}$$

that is the set of all possible values of the program numerical variables that satisfy all constraints in P.

Prove that $\forall \varphi_1 \in P_1 : \gamma^P(P_1) = \gamma^P((P_1 \setminus \{\varphi_1\}) \cup \{\varphi_2\})$ if and only if $P_1 \models \varphi_2$ and $((P_1 \setminus \{\varphi_1\}) \cup \{\varphi_2\}) \models \varphi_1$.

Question 20

Prove the following statements to be wrong: "*the widening approach to program static analysis is useless since it is always possible to perform an iterative static analysis using a finite abstract domain*"Ãě [7] and "*widenings can always be designed by further abstraction in an abstract domain satisfying the ascending chain condition*" [8].

[7] R.B. Kieburtz and M. Napierala. Abstract semantics. In S. Abramsky and C. Hankin, eds., *Abstract Interpretation of Declarative Languages*, chapter 7, pp. 143–180. Ellis Horwood, Chichester, U.K., 1987.

[8] C. Hankin, S. Hunt: Approximate Fixed Points in Abstract Interpretation. In *Sci. Comput. Program.* 22(3):283–306 (1994)