Master Parisien de Recherche en Informatique
École normale supérieure

Année scolaire 2009/2010

Cours M.2-6

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

Examen partiel

Patrick Cousot

20 novembre 2009

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.

We describe the syntax of grammars using the following meta-grammar (that is grammar of grammars).

$\mathbb{T}$			terminals T
$\mathbb{N}$			nonterminals N
$\mathbb{V}$	$\triangleq$	$\mathbb{T} \cup \mathbb{N}$	vocabulary ($\mathbb{T} \cap \mathbb{N} = \emptyset$)
G	$::=$	$P\ G \quad \vert \quad P$	grammar
P	$::=$	N '$::=$' ARS	production
ARS	$::=$	RS '$\vert$' $ARS \quad \vert \quad RS$	alternative right sides
RS	$::=$	$S\ RS \quad \vert \quad S$	right sides
S	$::=$	$N \quad \vert \quad T \quad \vert \quad$ 'ε'	symbols

This meta-grammar has the meta-symbols $::=$, $|$, ε, the meta-terminals $\{$'$::=$', '$|$', 'ε'$\} \cup \mathbb{V}$ such that $\{$'$::=$'; '$|$', 'ε'$\} \notin \mathbb{V}$ and the meta-nonterminals $\{G, P, ARS, RS, S, N, T\} \notin \mathbb{V}$. We assume that all productions of the grammar with the same left side nonterminal have their right sides grouped, with the alternative right sides separated by |. For example

$$
\begin{array}{lcl}
X & ::= & Y X \\
 & | & \varepsilon \\
Y & ::= & a \\
 & | & b
\end{array}
$$

The *terminals* of a grammar are

$$
\begin{array}{rcl}
\mathbb{T}[\![P G]\!] & \triangleq & \mathbb{T}[\![P]\!] \cup \mathbb{T}[\![G]\!] \\
\mathbb{T}[\![N \text{ `::=' } ARS]\!] & \triangleq & \mathbb{T}[\![ARS]\!] \\
\mathbb{T}[\![RS \text{ `|' } ARS]\!] & \triangleq & \mathbb{T}[\![RS]\!] \cup \mathbb{T}[\![ARS]\!] \\
\mathbb{T}[\![S\ RS]\!] & \triangleq & \mathbb{T}[\![S]\!] \cup \mathbb{T}[\![RS]\!] \\
\mathbb{T}[\![N]\!] & \triangleq & \emptyset \\
\mathbb{T}[\![T]\!] & \triangleq & \{T\} \\
\mathbb{T}[\![\text{`}\varepsilon\text{'}]\!] & \triangleq & \emptyset
\end{array}
$$

The definition is by structural induction, using the following well-founded relation $\lhd$ (*is a strict syntactic component of*). $P \lhd PG$, $G \lhd PG$, $N \lhd N$ '::=' ARS, $ARS \lhd N$ '::=' ARS, $RS \lhd RS$ '|' ARS, $ARS \lhd RS$ '|' ARS, $S \lhd S\ RS$, $RS \lhd S\ RS$.
The *nonterminals* of a grammar are

$$
\begin{array}{rcl}
\mathbb{N}[\![P G]\!] & \triangleq & \mathbb{N}[\![P]\!] \cup \mathbb{N}[\![G]\!] \\
\mathbb{N}[\![N \text{ `::=' } ARS]\!] & \triangleq & \{N\} \cup \mathbb{N}[\![ARS]\!] \\
\mathbb{N}[\![RS \text{ `|' } ARS]\!] & \triangleq & \mathbb{N}[\![RS]\!] \cup \mathbb{N}[\![ARS]\!] \\
\mathbb{N}[\![S\ RS]\!] & \triangleq & \mathbb{N}[\![S]\!] \cup \mathbb{N}[\![RS]\!] \\
\mathbb{N}[\![N]\!] & \triangleq & \{N\} \\
\mathbb{N}[\![T]\!] & \triangleq & \emptyset \\
\mathbb{N}[\![\text{`}\varepsilon\text{'}]\!] & \triangleq & \emptyset
\end{array}
$$

The *vocabulary* of a grammar is

$$
\mathbb{V}[\![G]\!] \quad \triangleq \quad \mathbb{T}[\![G]\!] \cup \mathbb{N}[\![G]\!]
$$

The *protosentences* $p \in \mathbb{V}^\star$ of the grammar are the possibly empty finite sequences of terminals and nonterminals of the grammar.
The *productions* of a grammar are

$$
\begin{array}{rcl}
\mathbb{P}[\![PG]\!] & \triangleq & \mathbb{P}[\![P]\!] \;\cup\; \mathbb{P}[\![G]\!] \\
\mathbb{P}[\![N \text{ `::=' } ARS]\!] & \triangleq & \{\langle N,\, r\rangle \mid r \in \mathbb{A}[\![ARS]\!]\} \\
\mathbb{A}[\![RS \text{ `|' } ARS]\!] & \triangleq & \mathbb{A}[\![RS]\!] \;\cup\; \{\mathbb{R}[\![ARS]\!]\} \\
\mathbb{A}[\![RS]\!] & \triangleq & \{\mathbb{R}[\![RS]\!]\} \\
\mathbb{R}[\![S\ RS]\!] & \triangleq & \mathbb{R}[\![S]\!] \cdot \mathbb{R}[\![RS]\!] \\
\mathbb{R}[\![N]\!] & \triangleq & N^{1} \\
\mathbb{R}[\![T]\!] & \triangleq & T \\
\mathbb{R}[\![\text{`}\varepsilon\text{'}]\!] & \triangleq & \epsilon
\end{array}
$$

For example

$$
\begin{array}{rl}
 & \mathrm{P}[\![X ::= YX \mid \varepsilon \\
 & \quad Y ::= a \mid b \qquad\qquad]\!] \\
= & \{\langle X,\, YX\rangle, \langle X,\, \epsilon\rangle, \langle Y,\, a\rangle, \langle Y,\, b\rangle\}
\end{array}
$$

The *transition system* of a grammar is a relation between protosentences of the grammar where a transition consists in replacing a nonterminal N in a string pNq by the right member r of a production $\langle N,\, r\rangle$ of the grammar G to obtain prq. The transition system of a grammar is

$$
\tau[\![G]\!] \quad \triangleq \quad \{\langle pNq,\, prq\rangle \mid p, q \in \mathbb{V}[\![G]\!]^{\star} \wedge \langle N,\, r\rangle \in \mathbb{P}[\![G]\!]\} \tag{1}
$$

For example, the transition system of the grammar G:

$$
\begin{array}{rcl}
X & = & YX \mid \varepsilon \\
Y & = & a \mid b
\end{array}
$$

is:

$$
\begin{array}{rcl}
\tau[\![G]\!] & = & \{\langle pXq,\, pYXq\rangle, \\
 & & \;\langle pXq,\, pq\rangle, \\
 & & \;\langle pYq,\, paq\rangle, \\
 & & \;\langle pYq,\, pbq\rangle \mid p, q \in \{X, Y, a, b\}^{\star}\}
\end{array}
$$

An example of derivation is ($\alpha \rightarrow \beta$ means $\langle \alpha,\, \beta\rangle \in \tau[\![G]\!]$):

$$
X \rightarrow YX \rightarrow YYX \rightarrow YbX \rightarrow abX \rightarrow ab
$$

[1] We assimilate a symbol to the string of one element which is this symbol.

Question 1

Provide a structural definition of the transition system of a grammar (by induction on the meta-grammar).

Question 2

Prove that the correctness of the structural definition of the transition system of a grammar (that is the equivalence of the definitions in questions ?? and 1).

Question 3

Let us define the *reflexive transitive closure* $r^\star$ of a relation $r \in \wp(S \times S)$ on a set S as $r^\star \triangleq \bigcup_{n \geqslant 0} r^n$ where the *powers* r^n of r are $r^0 \triangleq \{\langle x,\ x\rangle \mid x \in S\} \triangleq \mathbb{1}_S$ (identity relation), $r^{n+1} = r^n \circ r = r \circ r^n$, and the composition of relations is $r \circ r' \triangleq \{\langle x,\ x''\rangle \mid \exists x' \in S : \langle x,\ x'\rangle \in r \wedge \langle x',\ x''\rangle \in r'\}$. Prove that $r^\star = \mathbf{lfp}^{\subseteq}_{\emptyset}\ \lambda X \bullet \mathbb{1}_S \cup r \circ X = \mathbf{lfp}^{\subseteq}_{\emptyset}\ \lambda X \bullet \mathbb{1}_S \cup X \circ r$ (where $\mathbf{lfp}^{\leqslant}_{a}\ f$ is the $\leqslant$-least fixpoint of f which is $\leqslant$-greater than or equal to a, if any).

Question 4

The derivation semantics of a grammar is the reflexive transitive closure $\tau[\![G]\!]^\star$ of its transition semantics $\tau[\![G]\!]$ defined in questions ?? and 1. Let us define the $\subseteq$-increasing transformer:

$$B[\![ARS]\!] \quad \in \quad \wp(\mathbb{V}^\star \times \mathbb{V}^\star) \xrightarrow{\nearrow} \wp(\mathbb{V}^\star)$$

as follows:

$$\begin{aligned} B[\![RS \mid ARS]\!]r &\triangleq B[\![RS]\!]r \ \cup\ B[\![ARS]\!]r &(2)\\ B[\![SRS]\!]r &\triangleq B[\![S]\!]r \cdot B[\![RS]\!]r &(3)\\ B[\![N]\!]r &\triangleq \{p \mid \langle N,\ p\rangle \in r\} &(4)\\ B[\![T]\!]r &\triangleq \{T\} &(5)\\ B[\![\varepsilon]\!]r &\triangleq \{\varepsilon\} &(6)\end{aligned}$$

where $X \cdot Y = \{pq \mid p \in X \wedge q \in Y\}$ is the concatenation of sets of protosentences and ε is the empty protosentence.

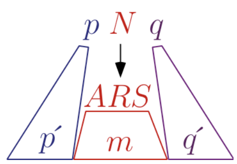
Let us define

$$B[\![G]\!] \quad \in \quad \wp(\mathbb{V}^\star \times \mathbb{V}^\star) \xrightarrow{\nearrow} \wp(\mathbb{V}^\star \times \mathbb{V}^\star)$$

as follows:

$$
\begin{aligned}
B[\![PG]\!]r &\triangleq B[\![P]\!]r \cup B[\![G]\!]r && (7)\\
B[\![N ::= ARS]\!]r &\triangleq \mathsf{l}_{\mathbb{V}^\star} \cup \{\langle pNq,\ p'mq'\rangle \mid \langle p,\ p'\rangle \in r \wedge && (8)\\
&\qquad\qquad m \in B[\![ARS]\!]r \wedge \langle q,\ q'\rangle \in r\}
\end{aligned}
$$

which can be illustrated as follows



Prove that $\mathbf{lfp}^{\subseteq}_{\emptyset}\, B[\![G]\!] = \tau[\![G]\!]^\star$.