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).

\[
\begin{align*}
T & \quad \text{terminals } T \\
N & \quad \text{nonterminals } N \\
\mathcal{V} & \triangleq T \cup N \quad \text{vocabulary } (T \cap N = \emptyset) \\
G & \ ::= \ P\, G \mid P \quad \text{grammar} \\
P & \ ::= \ N \ ::= \ ARS \quad \text{production} \\
ARS & \ ::= \ RS \mid RS \mid RS \quad \text{alternative right sides} \\
RS & \ ::= \ S\, RS \mid S \quad \text{right sides} \\
S & \ ::= \ N \mid T \mid \varepsilon \quad \text{symbols}
\end{align*}
\]

This meta-grammar has the meta-symbols ::=, |, ε, the meta-terminals {::=, |, ε} \cup \mathcal{V} such that {::=; |; ε} \notin \mathcal{V} and the meta-nonterminals \{G, P, ARS, RS, S, N, T\} \notin \mathcal{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

\[ X ::= YX \]

\[ Y ::= a \]

The terminals of a grammar are

\[ T[P] \triangleq T + T[G] \]
\[ T[N ::= ARS] \triangleq T[ARS] \]
\[ T[S RS] \triangleq T[S] + T[RS] \]
\[ T[N] \triangleq \emptyset \]
\[ T[T] \triangleq \{T\} \]
\[ T[\varepsilon] \triangleq \emptyset \]

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

The nonterminals of a grammar are

\[ N[N ::= ARS] \triangleq \{N\} + N[ARS] \]
\[ N[S RS] \triangleq N[S] + N[RS] \]
\[ N[N] \triangleq \{N\} \]
\[ N[T] \triangleq \emptyset \]
\[ N[\varepsilon] \triangleq \emptyset \]

The vocabulary of a grammar is


The protosentences \(p \in V^*\) of the grammar are the possibly empty finite sequences of terminals and nonterminals of the grammar.

The productions of a grammar are
\[ P[PG] \triangleq P[P] \cup P[G] \\
\[ P[N ::= \text{ARS}] \triangleq \{\langle N, r \rangle | r \in A[\text{ARS}]\} \\
\[ A[RS ::= \text{ARS}] \triangleq A[RS] \cup \{R[\text{ARS}]\} \\
\[ A[RS] \triangleq \{R[RS]\} \\
\[ R[RS] \triangleq R[S] \cdot R[RS] \\
\[ R[N] \triangleq N^+ \\
\[ R[T] \triangleq T \\
\[ R[\varepsilon] \triangleq \varepsilon \\
\]

For example

\[
\begin{align*}
P[X ::= YX | \varepsilon] \\
Y ::= a | b
\end{align*}
\]

\[
= \{\langle X, YX\rangle, \langle X, \varepsilon\rangle, \langle Y, a\rangle, \langle Y, b\rangle\}
\]

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

\[
\tau[G] \triangleq \{\langle pNq, prq \rangle | p, q \in V[G]^* \land \langle N, r \rangle \in P[G]\} 
\]

(1)

For example, the transition system of the grammar \( G \):

\[
X = YX | \varepsilon \\
Y = a | b
\]

is:

\[
\tau[G] = \{\langle pXq, pYXq \rangle, \langle pXq, pq \rangle, \langle pYq, paq \rangle, \langle pYq, pbq \rangle | p, q \in \{X, Y, a, b\}^*\}
\]

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

\[
X \rightarrow YYX \rightarrow YbX \rightarrow abX \rightarrow ab
\]

\(^1\text{We assimilate a symbol to the string of one element which is this symbol.}\)
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 $r^*$ of a relation $r \in \wp(S \times S)$ on a set $S$ as $r^* \triangleq \bigcup_{n \geq 0} r^n$ where the powers $r^n$ of $r$ are $r^0 \triangleq \{\langle x, x \rangle \mid x \in S\} \triangleq I_S$ (identity relation), $r^{n+1} = r^n \circ r = r \circ r^n$, and the composition of relations is $r \circ r' \triangleq \{\langle x, x'' \rangle \mid \exists x' \in S : \langle x, x' \rangle \in r \land \langle x', x'' \rangle \in r'\}$. Prove that $r^* = \text{lfp}_a \lambda X \cdot I_S \cup r \circ X = \text{lfp}_a \lambda X \cdot I_S \cup X \circ r$ (where $\text{lfp}_a f$ is the $\leq$-least fixpoint of $f$ which is $\leq$-greater than or equal to $a$, if any).

Question 4

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

$$B[ARS] \in \wp(V^\times V^\times) \xrightarrow{\lambda} \wp(V^\times)$$

as follows:

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

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

Let us define

$$B[G] \in \wp(V^\times V^\times) \xrightarrow{\lambda} \wp(V^\times V^\times)$$
as follows:

\[
B[PG]_r \triangleq B[P]_r \cup B[G]_r \\
B[N \equiv ARS]_r \triangleq 1_\gamma \cup \{\langle pNq, p'mq' \rangle \mid \langle p, p' \rangle \in r \land m \in B[ARS]_r \land \langle q, q' \rangle \in r\}
\]  

which can be illustrated as follows

\[
\begin{array}{c}
p \\
\downarrow \quad ARS \\
q \\
\end{array}
\begin{array}{c}
p' \\
\downarrow m \\
q' \\
\end{array}
\]