Université Libre de Bruxelles
Faculté des Sciences

INFO-F-403 – Introduction to language theory and compiling

First session examination

January, 17th, 2014

Instructions

— This is a closed book test. You are not allowed to use any kind of reference.
— You can anwser in french or in english.
— Write your first and last names on each sheet that you hand in.
— Write clearly: you can use a pencil or a ballpen or even a quill as long as your answers are readable !
— Always provide full and rigourous justifications along with your answers.
— This test is worth 12 points out of 20. The weight of each questions is given as a reference.
— In your answers (diagrams representing automata, grammars,. . .), you can always use the conventions adopted in the course, without recalling them explicitely. If you deviate from these conventions, be sure to make it clear.

Question 1 — 4 points

Define *formally* the following notions:

1. regular language
2. regular expression (RE)
3. deterministic finite automaton (DFA)
4. non-deterministic finite automaton with ε-transitions (ε-NFA)

and prove *formally* that the classes of REs, DFAs and ε-NFAs recognise *exactly* the class of regular languages. That is, prove that all regular languages can be defined by an RE (resp. DFA, ε-NFA); and that the language defined by an RE (DFA, ε-NFA) is always regular.

Question 2 — 2 points

Define the notion of *ambiguous context-free grammar*. Is the following grammar ambiguous (where the set of terminals is $\{+, \times, id, (,)\}$) ? Prove your answer. If the grammar is ambiguous, give a non-ambiguous context-free grammar that defines the same language.

$$
\begin{array}{lllll}
(1) & E & \to & E + E \\
(2) & & & T \\
(3) & T & \to & T \times T \\
(4) & & & F \\
(5) & F & \to & id \\
(6) & & & (E)
\end{array}
$$

Finally, explain why ambiguous grammars are an issue when building a compiler (you can use the above grammar as an example to support your answer).

Question 3 — 2 points

Explain the role of the *scanner* and the *parser* during compiling. Explain their respective place in the compiling chain, and how they interact.

Question 4 — 1 point

Give the diagram of a *deterministic* pushdown automaton, on the alphabet $\Sigma = \{a, b, c, d, e\}$, that accepts the language $L = \{(ab)^n c(de)^n \mid n \geq 0\}$ using the empty stack acceptance condition.

Question 5 — 3 points

Give the LR(0) canonical finite state machine (CFSM) of the following grammar (where the set of variables is $\{S', S, A, B, C\}$ and the set of terminals is $\{a, b, c, d\}$):

(1)	S'	$\rightarrow$	$S\$$
(2)	S	$\rightarrow$	aAb
(3)			aB
(4)			c
(5)	A	$\rightarrow$	dA
(6)			a
(7)	B	$\rightarrow$	bC
(8)	C	$\rightarrow$	ε