COMP3630/6363 Assignment 2, 2009

Due date: beginning of lecture, Wednesday April 29

Hand-written answers are acceptable if written neatly.

Correct answers might be given less than full credit if they are unnecessarily complicated.

Late penalty: 10% cumulative per day, ignoring weekends.
Not accepted more than 14 days late without special permission.

Question 1

Write a context-free grammar for each of the following languages.

(a) Real numbers that consist of one or more decimal digits, followed optionally by a decimal point and one or more decimal digits. The part before the decimal point (or the whole string in case there is no decimal point) must not have a leading 0 unless it is the string "0". Examples: `29` `23.034` `0.00`

(b) Properly formed non-empty arithmetic expressions using variables (whose names are a single lowercase letter), binary + and −, and parentheses. Examples:
`x+y` `(x+(a-b+y+(z))-x)` `((((x))))` `z`

In the case of (a), give an unambiguous grammar. In the case of (b), give an ambiguous grammar and prove using an example that it is ambiguous. **Unnecessarily complicated or messy answers will lose marks even if technically correct.**

Question 2

Convert the following grammar to Chomsky Normal Form **using the algorithm given in the textbook and lectures**. Show the grammar after each step.

$$
\begin{aligned}
S &\to SSC \mid SaX \mid b \\
A &\to bb \\
X &\to \epsilon \mid Xb \mid XBY \\
Y &\to X \\
B &\to Ba \mid CS \\
C &\to B
\end{aligned}
$$

Question 3

Prove that the following language is context-free by describing a push-down automaton that accepts it: Strings over $\{a, b, c\}$ such that for any prefix the number of 'b's is at least equal to the number of 'a's. You don't need to give all the transitions in detail, but you must describe the PDA with sufficient care that the correctness is obvious to your lecturer even if he is tired and grumpy.

Question 4

One of following languages is context-free and the other is not. Provide a grammar for the one that is and a proof using the pumping lemma for the one that isn't.

(i) $\{a^i b^j c^{i+j+1} \mid i, j \geq 0\}$

(ii) $\{a^i b^j c^{ij+1} \mid i, j \geq 0\}$

Question 5 (COMP6363 and PhDs only)

Two strings can be *interleaved* in various ways. For example the possible results of interleaving *abbc* and *def* include *adebbfc*, *dabebcf*, *deabbcf*, and lots of others. Note that the order of the characters of the two original strings is preserved during the interleaving. (If this is not clear, please ask.)

Given languages L_1, L_2 over alphabet Σ, the *interleaving* of L_1 and L_2 is the language over Σ defined by

$$L_1 ⋓ L_2 = \{\, x \mid x \text{ is formed by interleaving a string from } L_1 \text{ and a string from } L_2 \,\}.$$

In each of the following cases, either prove that the claim is true or prove that the claim is false.

(a) If L_1 and L_2 are regular, then $L_1 ⋓ L_2$ is regular.

(b) If L_1 and L_2 are context-free, then $L_1 ⋓ L_2$ is context-free.

Question 6 (Optional: good answers get bonus credit even if incomplete)

The following or may not be true. Either prove it or disprove it.

Conjectured Theorem. *Let Σ be an alphabet and consider four infinite sequences of strings over Σ:*

$$\begin{array}{l} \alpha_1,\ \alpha_2,\ \alpha_3,\ \ldots \\ \beta_1,\ \beta_2,\ \beta_3,\ \ldots \\ \gamma_1,\ \gamma_2,\ \gamma_3,\ \ldots \\ \delta_1,\ \delta_2,\ \delta_3,\ \ldots \end{array}$$

Suppose L is a language over Σ such that, for all $i, j, k, \ell \geq 1$, $\alpha_i \beta_j \gamma_k \delta_\ell \in L$ if and only if $i = k$ and $j = \ell$. Then L is not context-free.