COMS W4115
Programming Languages and Translators
Homework Assignment 1

Prof. Stephen A. Edwards
Columbia University

Due February 17th, 2005
at 11:59 PM

CVN students: FAX the solutions to CVN.

Write both your name and your Columbia ID (e.g., se2007) on your solutions.

Do this assignment alone. You may consult the instructor, but not other students.

1. Scanners

   (a) Write a regular expression that accepts only nonnegative even binary numbers, e.g., 100, 110, 10, and 0.

   (b) Draw a DFA that accepts only nonnegative decimal integer numbers divisible by three, e.g., 0, 3, 6, 9, 12, and 561. Hint: sum their digits.

   (c) Using ANTLR-like syntax, write a scanner for C's floating point numbers, as defined by Ritchie.

   > A floating constant consists of an integer part, a decimal point, a fraction part, an `e`, and an optionally signed integer exponent. The integer and fraction parts both consist of a sequence of digits. Either the integer part or the fraction part (not both) may be missing; either the decimal point or the `e` and the exponent (not both) may be missing.

   Hint: make sure your scanner accepts constants such as `1. 0.5e-15 .3e+3 .2 1e5` but not integer constants such as `42`

2. Draw a DFA for a scanner that recognizes and distinguishes the following set of keywords. Draw accepting states with double lines and label them with the name of the keyword they accept. Follow the definition of a DFA given in class.

   ```
   auto case char const continue default
   do double else enum if ifelse
   ```

3. Dragon book, 3.16, p. 149:

   Construct nondeterministic finite automata for the following regular expressions using Algorithm 3.3 (p. 122, shown in class) and show how the NFA simulates the input string *ababbab*. Follow the simulation procedure shown in class.

   (a) $(a\,|\,b)^*$

   (b) $((\varepsilon\,|\,a)b^*)^*$

   (c) $(a\,|\,b)^*abb(a\,|\,b)^*$

4. Dragon book, 4.23, p. 270:

   (a) Using the grammar

   $$S \rightarrow (L)\,|\,a$$
   $$L \rightarrow L,S\,|\,S$$

   construct a rightmost derivation for $(a,(a,a))$ and show the handle of each right-sentinel form.

   (b) Show the steps of a shift-reduce (bottom-up) parser corresponding to this rightmost derivation.

   (c) Show the steps in the bottom-up construction of a parse tree during this shift-reduce parse.

5. Disambiguate and remove left recursion from the following grammar (i.e., show an equivalent grammar):

   $e \rightarrow e$ `>>` $e\,|\,e$ `?` e `:` $e\,|\,e$ `->` $e\,|$ **id**

   Use C's precedence rules, i.e., the precedence of `->` is higher than that of `>>` is higher than that of `?:`. Hint: make sure your grammar can accept an expression such as **id?id:id->id?id:id** with the right precedence rules, e.g., as if it had been written **id?id:**(**id->id**)**?id:id**.