CS 536

Midterm Exam

Friday, November 9, 2012

11:00 AM — 1:00 PM

Instructions

Answer question 1 and any three other questions. (If you answer more, only the first four will count.) Point values are as indicated. Please try to make your answers neat and coherent. Remember, if we can't read it, it's wrong. Partial credit will be given, so try to put something down for each question (a blank answer always gets 0 points!).

1. (1 point)
   In what year were the 2012 Olympic Games held?

2. Write regular expression definitions for the following token classes:

   (a) (11 points)
   Unsigned integer literals that contain no unnecessary leading zeroes. That is, 0, 10, 123, and 1000001 are allowed, but 01, 007 and 000 are not allowed.

   (b) (11 points)
   A CSX-style single line comment that does not contain the character pair `##` anywhere within its text. Thus

   ```
   // We're #1!!
   ```

   is OK, but

   ```
   // ## This is really a comment!
   ```

   is not allowed.

   (c) (11 points)
   A CSX-style string literal that contains exactly one newline in its text (and thus spans exactly two lines). For example

   ```
   write("Hello
   World!");
   ```

3. (33 points)
   Write a JLex token definition that specifies a C-like comment whose body is delimited by `/*` and `*/`. Individual `*`'s and `/`'s may appear in the comment body, but the pair `*/` may not.

4. (33 points)
   Assume we add a conditional compilation facility to CSX. The symbol `#if` will be followed by an integer literal. If the integer literal is non-zero, all the characters up to a `#end` symbol will be scanned as usual. The `#if`, the integer literal and the `#end` will be deleted (just like white space or comments). If the integer literal following the `#if` is zero, all the characters up to the `#end` will be skipped, as will the `#if`, the integer literal and the `#end`. Thus if we have

```
#if 1
a = b+c;
#end
```

the assignment will be scanned, but if we have

```
#if 0
a = b+c;
#end
```

the assignment will be skipped by the scanner.

Outline how you'd add conditional compilation to your CSX scanner. You may assume that `#if` and `#end` pairs are not allowed to nest.

5. (33 points)

Let s be a string. Define delete1(s) to be a function that deletes a single character from s. If s is n characters long, delete1 returns a set of n (or fewer) strings since there are n characters that might be deleted in a string of length n.

For example, delete1(abc) = { bc, ac, ab }. delete1 applied to a set of strings is the union of delete1 applied to members of the set. Hence delete1(ab, de) = { b, a ,e, d }.

Let R be any regular set (my choice) that does not include λ. Show that delete1(R) is a regular set.

6. (a) (16 points)

Let DFA be any deterministic finite automaton (my choice). Assume you know DFA contains exactly n states and that it accepts at least one string of length n or greater. Show that DFA must also accept at least one string of length 2n or greater.

(b) (17 points)

Let F be any deterministic finite automation (my choice). Explain how to test whether F accepts all possible strings (that is, no string at all is rejected).