CS172 Computability & Complexity (Spring'09)

Instructor: Mihai Pătraşcu GSI: Omid Etesami

$\boxed{\text{Midterm Exam}}$ *1.5 hours*

All electronics must be off. You may bring one sheet of paper (double-sided, letter size) containing any notes you care to use. No other materials allowed. You may answer questions in any order.

1. [10 POINTS] Write your name, email, and **student ID** on top of all sheets of paper. Make sure your cell phone (or other device) will not produce sound during the exam.

2. [10 POINTS] Draw the automaton (with ε-transitions) that the Knuth-Morris-Pratt algorithm builds for the needle *tatutatutata*.

3. [20 POINTS] The **?** operator from the C programming language is used in expressions like: $A?B : C$. Here A, B, C are themselves expressions. If A evaluates to non-zero, the entire expression "$A?B : C$" evaluates to the value of B; if A is zero, the entire expression evaluates to the value of C. For instance, $7 - 7?3 + 9 : 1 + 2$ evaluates to 3. The ? operator has lower precedence than $+$ and $-$, so the expression is the same as $(7 - 7)?(3 + 9) : (1 + 2)$.

(a) Write an expression that evaluates to different values if ? : is left-associative versus right-associative.

(b) Write a context free grammar that parses expressions involving $+$, $-$, ? :, parentheses, and single-digit numbers. Assume ? is left-associative. You may ignore the use of minus for negation (ignore expressions like -3).

(c) Same as (b), but ? is right associative.

(d) Give succinct pseudocode for parsing the grammar in (c).

(e) Informally explain why the following grammar is harder to parse. Illustrate with a couple of examples. (Note: this grammar is rather unrelated to the answers you must give in (b) and (c). Please do not get confused by using this as a starting point.)

$$\begin{array}{rcl}
S & \to & T?T:T \ \mid \ U?U\#U \\
T & \to & T+V \ \mid \ T-V \ \mid \ V \\
U & \to & V+U \ \mid \ V-U \ \mid \ V \\
V & \to & 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9
\end{array}$$

4. [10 POINTS] Consider a stream containing $n-1$ distinct numbers from $\{1, \ldots, n\}$. In other words, the stream contains all numbers from 1 to n, except one. Describe an algorithm using $O(\lg n)$ bits of space that outputs the missing number. You may assume n is known in advance.

5. [10 POINTS] Consider a stream of n integers in $\{0, \ldots, n^2\}$. The goal is to determine whether the average of the values is also an element of the stream. For instance, in the stream $[3, 7, 1, 2, 2]$, the average is 3, which does appear in the stream. In the stream $[4, 7, 1, 2, 2]$, the average is 3.2, which does not appear in the stream.

Show that any algorithm answering this question must use $\Omega(n)$ bits of memory. (Your proof will likely invoke the communication lower bound for INDEXING.)

6. [10 POINTS] Prove or disprove:

(a) There exists a bijection between the set of real numbers in $(0, 1)$ and the set of points with real coordinates in $(0, 1)^2$.

(b) Let $\mathcal{F}$ be the set of functions $f : \mathbb{N} \to \mathbb{N}$ satisfying: $(\forall) n \in \mathbb{N}, n \leq f(n) \leq n^2$. The set $\mathcal{F}$ is countable.

6. [10 POINTS] Socrates arrives in a new town with n people, at most $\frac{n}{2} - 1$ of which are liars. His goal is to determine which ones are honest and which not.

To accomplish this goal, he can organize debates between any pair of the n people. After A finished debating B, A tells Socrates what he believes about B, and B tells what he believes about A. An honest person will always identify the other person correctly (as an honest man, or as a liar). But a liar might say anything (including the truth).

How can Socrates identify the honest people?

8. [20 POINTS] A little known fact is that the Incas designed a Turing Machine many centuries ago. Their machine model was identical to the regular Turing Machine, except for the Sacred Input Commendment: "Thou shalt not overwrite a tape location where thy input hast been placed."

(a) Assume the input is represented with one free (unwritten) cell in between each two input cells. That is, if the input is $s_1, s_2, s_3 \ldots$, the beginning of the tape will read s_1, free cell, s_2, free cell, s_3, free cell, etc. Show that the Inca Machine can compute anything that a regular Turing Machine can compute.

(b) Now assume the input is written contiguously at the beginning of the tape (as usual for Turing Machines). Show that the Inca Machine is exactly as powerful as a Read-Only Turing Machine. A Read-Only Turing Machine is a Turing Machine that cannot write to tape at all (it may only move its head around and read cells).

(c) What exactly is the set of languages decidable by a Read-Only Turing Machine?