CS172 Computability & Complexity (Spring'09)

Instructor: Mihai Pătraşcu GSI: Omid Etesami

Sample Midterm *Real midterm: Monday March 16*

Midterm policies:

- Bring paper to write on.
- You may bring one (1) sheet of paper (double-sided, letter size) containing any notes you care to use. No other materials allowed.
- All electronics must be off.
- The midterm lasts for 1.5 hours. Usual room, usual time.
- Omid will discuss problems from the sample midterm during this week's section. Of course, you may also ask questions in office hours or by email.

The following are types of questions that you are likely to see on the exam.

1. Educational objective: Design a finite automaton for a given language.
Sample question: Design a deterministic finite automaton for the language: $L =$ strings over $\{a, b, c\}$ that contain both *abac* and *bacac* as substrings.

2. Educational objective: Reason about finite automata.
Sample question: Prove that the intersection of two regular languages is also regular. Use this to conclude that the language L from problem 1 is regular.

Assume we replaced *abac* and *bacac* by two arbitrary strings of length n. How many states (with O-notation) would an automaton like the one you build in question 1 have? How many states would the automaton from question 2 have?

3. Educational objective: Show that a language is not regular using the pumping lemma.
Sample Question: Show that the following language is not regular: $L =$ strings over $\{0, 1\}$ that contain the same number of 0's as of 1's.

4. Educational objective: Demonstrate understanding of the Knuth-Morris-Pratt algorithm.
Sample question: Design the automaton (with ε-transitions) that the KMP algorithm builds for the needle *ananas*. While searching in the haystack *banabananas*, how many ε transitions are taken?

5. Educational objective: Design simple streaming or communication algorithms.
Sample question: You are given a stream of n numbers in $\{0, \ldots, n^2\}$; assume n is known in advance. The goal is to determine whether the third largest value in the stream appears an odd number of times. Design an algorithm using $O(\lg n)$ bits of memory to answer this question.

Sample question: Describe a communication protocol for finding the median with $O(\lg n)$ messages of $O(\lg n)$ bits each.

6. Educational objective: Prove simple streaming/communication lower bounds.

Sample question: Consider a stream of n numbers in $\{0, \ldots, n^2\}$. The goal is to determine whether there exist two numbers in the stream that differ by exactly 13. 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.)

7. Educational objective: Design context free grammars. Reason about ambiguity and left/right associativity.

Sample question: Consider the following grammar for a fragment of natural language:

$$
\begin{array}{rcl}
\mathit{Sentence} & \rightarrow & \mathit{Subject}\ \ \mathit{Verb} \\
\mathit{Subject} & \rightarrow & \mathit{Noun} \\
\mathit{Noun} & \rightarrow & \text{the boy} \mid \text{the teacher} \\
\mathit{Verb} & \rightarrow & \mathit{Intransitive} \mid \mathit{Transitive}\ \mathit{Complement} \\
\mathit{Intransitive} & \rightarrow & \text{ate} \mid \text{slept} \\
\mathit{Transitive} & \rightarrow & \text{saw} \mid \text{introduced} \\
\mathit{Complement} & \rightarrow & \mathit{Noun} \mid \mathit{Pronoun} \\
\mathit{Pronoun} & \rightarrow & \text{him}_{\text{BOY}} \mid \text{him}_{\text{TEACHER}} \mid \text{himself}_{\text{BOY}} \mid \text{himself}_{\text{TEACHER}}
\end{array}
$$

The fragment generates sentences like "the boy saw the teacher" or "the teacher slept." The subscript of the pronoun indicates whom the pronoun refers to (NB: each version of the pronoun is a distinct literal).

In English, a sentence like "the boy saw him" cannot be used if "him" refers to the boy. On the other hand, in "the boy saw himself", "himself" may only refer to the boy. Unfortunately, our grammar generates incorrect sentences like "the boy saw him_{BOY}" or "the teacher introduced $\text{himself}_{\text{BOY}}$."

Fix this shortcoming of the grammar.

A note for your linguistic curiosity: In theoretical linguistics, it is said that a node N of a tree *c-commands* all nodes in the subtree of N's sibling(s). For instance, the subject c-commands the complement in the parse tree of our grammar. The c-command relation seems inherent in human processing of speech. For instance, we can explain the rules of English as follows: (1) reflexive pronouns must be c-commanded by something that refers to the same entity; (2) non-reflexive pronouns may not be c-commanded by anything referring to the same entity. In other languages, other unrelated grammatical rules have to do with whether some part of speech is c-commanded by another.

8. Educational objective: Understand parsing of simple grammars.

Sample question: Write pseudocode for parsing mathematical expressions involving $\times$, /, and $\wedge$ (exponentiation).

9. Educational objective: Reason about computational models.

SAMPLE QUESTION: Professor McDoofus has figured out how to obtain a machine that is more powerful than the Turing Machine and the Random Access Machine. Unlike these machines, which have a linear (potentially infinite) memory, the professor invented a machine with 2-dimensional memory that can stretch infinitely in the entire plane. At every step, the McDoofus Machine may move its head east, west, north, or south.

Prove to the professor that his machine is not any more powerful than the Turing Machine.

10. EDUCATIONAL OBJECTIVE: Reason about Turing-recognizable and Turing-decidable languages.

SAMPLE QUESTION: Show that the class of Turing-recognizable languages is closed under union.

11. EDUCATIONAL OBJECTIVE: Reason about simple diagonalization and countability.

SAMPLE QUESTION: Show that the class of points in $[0,1]^2$ with rational coordinates is countable.