University of Virginia - cs3102: Theory of Computation	Spring 2010
Exam 1	**2 March 2010**

Honor Policy. For this exam, you must **work alone**. You may consult the single page of notes you brought, but may not look at any other materials. You may not aid or accept aid from other students.

Directions. Answer all 6 questions, including all sub-parts. You may use the backs of pages for your scratch work, but we will only grade answers that are written in the answer spaces, or that are found following clearly marked arrows from these boxes. The space provided for each answer is designed to be big enough to easily fit a full credit, correct answer. If you feel like you need more space to write your answer, then either you have extremely large handwriting, or your answer is incorrect, inelegant, or you are providing more detail than is needed for full credit. In general, the questions are organized by topic and the sub-parts of each question are intended to get progressively more challenging (so question 5a may be easier than question 1d).

Your Name:

UVa Email Id:

Problem 1: Definitions. For each question, provide a correct, clear, precise, and concise answer from the perspective of a theoretical computer scientist.

a. (5) What is a *language*?

b. (5) How do we measure the *power* of a type of machine such as a DFA or NPDA?

c. (5) What does it mean for a machine model to be *nondeterministic*?

Problem 2: Language Classification.

For each of these questions, provide a convincing argument supporting the proposition. You may use any technique you wish to make your argument, and may assume any of the properties we have proved in class or in the book.

a. (5, semi-trick question) Show that the language produced by the context-free grammar below is a regular language:

$$S \rightarrow 0S0 \mid 1S1$$

b. (5) Show that the language produced by the context-free grammar below is a regular language:

$$S \rightarrow 0S \mid S1 \mid \epsilon$$

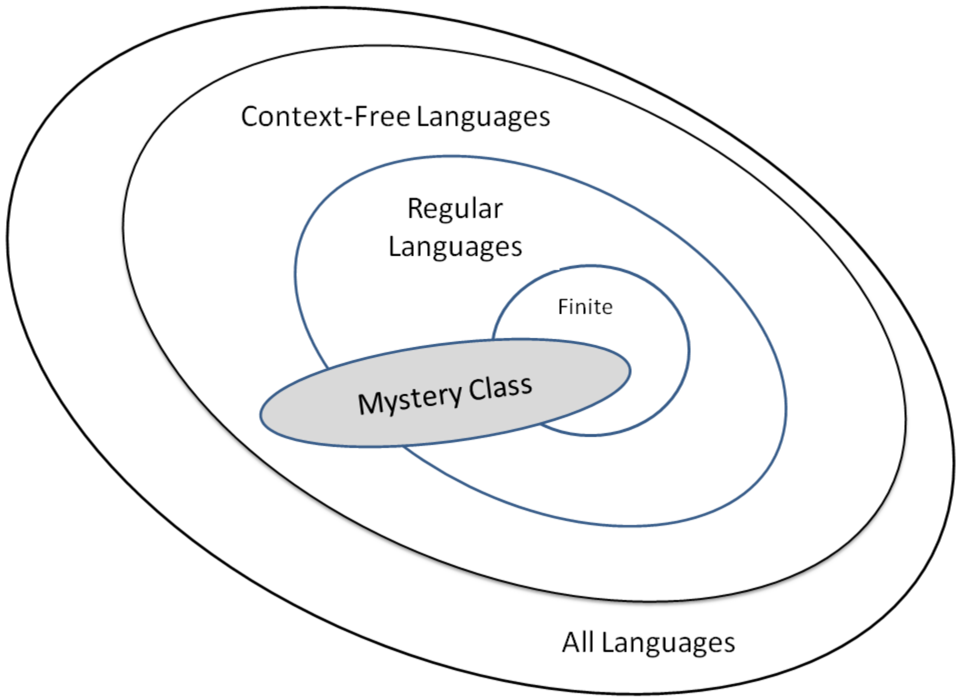
c. (5) Show that the language $NOTREPEATING$ defined below is **not** a regular language:

$$NOTREPEATING = \{w | w \in \{0,1\}^* \wedge \text{there is no } x \text{ such that } w = xx\}$$

d. (5 + 5) A *squarefree* sequence is a sequence that contains no adjacent repeating subsequences of any non-zero length. For example, ab and $abcba$ are squarefree, but aa is not since it contains a^2 and $abcbacbab$ is not since it contains $(cba)^2$. Show the language consisting of all squarefree strings in $\{a,b,c\}^*$ is not context-free.

Problem 3: Language Classes.

(15) Define a type of machine that recognizes the set of languages in the *Mystery Class* ellipse depicted below:



Your class should include infinitely many non-regular languages, but there should be infinitely many finite languages that are not included in your class. For full credit, your answer should include a convincing argument why your class includes infinitely many non-regular languages and excludes infinitely many finite languages.

Problem 4: Broken Proofs.

Each of these "proofs" claims to prove a conjecture that is false. For each proof, identify the *first step* that is wrong, and briefly and clearly explain why. The explanation is more important than the step you identify.

a. (5) **False Conjecture:** The intersection of two context free languages is context free.

Claimed Proof.

1. We define the intersection of two languages

$$A \cap B = \{w | w \in A \land w \in B\}.$$

2. The language $X = A \cap B$ is a subset of the language A: $X \subseteq A$. This is true since every string in X must also be in A.

3. Since X is a subset of A, and A is context free, X is context free.

b. (7) **False Conjecture:** All whole numbers are even.

Claimed Proof. We use induction on the numbers to prove the conjecture.

1. A number n is *even* if $2m = n$ for some integer m.
2. *Basis:* 0 is even. Select $m = 0$, then $2m = 0$.
3. *Induction:* Assume the conjecture holds for all $i < n$. We show that it holds for n.
4. $n = k + 2$ for some integer $k < n$.
5. By the induction hypothesis, $k = 2m$ for some integer m.
6. $n = k + 2 = (2m) + 2 = 2(m + 1)$. Thus, n is even.

c.(8, *tricky*) **False Conjecture:** The language $TRIPLES$ is not regular.

$$TRIPLES = \left\{1^{3n} | n \geq 0\right\}$$

Claimed Proof. We use the pumping lemma for regular languages to obtain a contradiction.

1. Assume $TRIPLES$ is regular. Then, there exists some DFA M with pumping length p that recognizes $TRIPLES$.
2. Choose $s = 1^p$.
3. The pumping lemma requires that there is a way to divide s into $s = xyz$ where $|y| \geq 1$ and $|xy| \leq p$ and $xy^iz \in TRIPLES$ for all $i \geq 0$.
4. Since s is all 1s, we know y can contain only 1s.
5. Choose $y = 1$.
6. Choose $i = 2$.
7. Since xy^2z now has $p + 1$ ones, it is not in the language $TRIPLES$.
8. Thus, we have a contradiction of the pumping lemma and $TRIPLES$ must not be regular.

d. (*Bonus*; don't work on this one until you finish the rest of the exam)

In the 2010 annual Latkes (fried potato cakes) v. Hamentash (triangular cookies) debate at MIT, Michael Sipser presented the proof described below (taken from *Faculty fling fake facts in food fight*, The Tech, 26 February 2010, with numbers added):

> Sipser wrapped things up for Team Hamentash with the HamenTheorem, which proves by contradiction that the hamentash is better than the latke. (1) First, the proof assumes latkes are best. (2) Then by obviousness, he claimed that hamentashen are better than nothing, and (3) by first assumption, claimed that nothing is better than latkes. (4) Therefore, Sipser argued that the HamenTheorem proved that hamentashen are better than latkes.

Since everyone knows latkes are better, the proof must be flawed. Explain the flaw in Sipser's proof.

Problem 5: Leaky PDAs.

Consider a new machine model known as a *LeakyPDA*. A LeakyPDA is similar to a deterministic PDA except that the stack has a limited depth. A LeakyPDA is specified by the 7-tuple $(Q, \Sigma, D, \Gamma, \delta, q_0, F)$ where $D \in \mathcal{N}$ is a natural number giving the maximum stack depth and the other elements have the same meaning as for a deterministic PDA. When the stack contains D elements, if a transition pushes another element on the stack, the bottom element of the stack is removed. That is, the result of following the transition rule,

$$\delta(q_i, a, \epsilon) \rightarrow (q_t, h_p)$$

starting with a stack $\gamma_0\gamma_1, \ldots, \gamma_{D-2}\gamma_{D-1}$ is the stack $h_p\gamma_0\gamma_1, \ldots, \gamma_{D-2}$.

(10) Precisely describe the computing power of a LeakyPDA. For full credit, your answer should include a convincing proof supporting your answer.

Problem 6: Language Class Differences.

a. (10) If A and B are regular languages, is $A - B$ always a regular language? Recall that the difference of two sets is defined as

$$A - B = \{s : s \in A \land x \notin B\}$$

(For full credit, your answer should include a convincing proof supporting your answer.)

b. (10) If A and B are context-free languages, is $A - B$ always a context-free language? (For full credit, your answer should include a convincing proof supporting your answer.)

1 (15)	2 (20)	3 (15)	4 (20)	5 (10)	6 (20)	Total