Practice CS103 Midterm Exam II

This practice exam is closed-book and closed-computer but open-note. You may have a double-sided, 8.5” × 11” sheet of notes with you when you take this exam. Please hand-write all of your solutions on this physical copy of the exam.

On the actual exam, there'd be space here for you to write your name and sign a statement saying you abide by the Honor Code. We're not collecting or grading this exam (though you're welcome to step outside and chat with us about it when you're done!) and this exam doesn't provide any extra credit, so we've opted to skip that boilerplate.

You have three hours to complete this practice midterm. There are 24 total points. This practice midterm is purely optional and will not directly impact your grade in CS103, but we hope that you find it to be a useful way to prepare for the exam. You may find it useful to read through all the questions to get a sense of what this practice midterm contains before you begin.

Question		**Points**	**Grader**
(1) The Pigeonhole Principle	(6)	/ 6	
(2) Functions and Relations	(6)	/ 6	
(3) Regular Languages	(6)	/ 6	
(4) Nonregular Languages	(6)	/ 6	
	(24)	**/ 24**	

Best of luck on the exam!

Problem One: The Pigeonhole Principle (6 Points)

Suppose that there are six flavors of jelly beans and that you have eleven jelly beans of each flavor. You distribute those jelly beans across five jars. Prove that no matter how you distribute them, there will always be a jar with at least three jelly beans of one flavor and at least three jelly beans of a different flavor.

(Giving credit where credit is due: this excellent pigeonhole principle problem comes from a problem set given at MIT. I just thought it was such a good problem that I couldn't pass up on it. ☺)

Problem Two: Functions and Relations (6 Points)

Let A and B be arbitrary sets and let $\leq_B$ be an arbitrary partial order over B. Suppose that we pick an injective function $f : A \to B$. We can then define a relation $\leq_A$ over A as follows: for any $x, y \in A$, we say that $x \leq_A y$ if $f(x) \leq_B f(y)$.

i. **(4 Points)** Prove that $\leq_A$ is a partial order over A.

Let R be a partial order relation over a set A. We say that R is a ***total order*** if it satisfies the following requirement:

$$\forall a \in A.\ \forall b \in A.\ (aRb \lor bRa)$$

This question explores the interplay between total orders and equivalene relations.

ii. **(2 Points)** Prove or disprove: there is a binary relation R over the set $\mathbb{N}$ such that R is both an equivalence relation and a total order.

Problem Three: Regular Languages (6 Points)

The number of characters in a regular expression is defined to be the total number of symbols used to write out the regular expression. For example, **(a|b)*** is a six-character regular expression, and **ab** is a two-character regular expression.

Let Σ = {**a**, **b**}. Find examples of all of the following:

- A regular language over Σ with a one-state NFA but no one-state DFA.
- A regular language over Σ with a one-state DFA but no one-character regular expression.
- A regular language over Σ with a one-character regular expression but no one-state NFA.

Prove that all of your examples have the required properties.

Problem Four: Nonregular Languages (6 Points)

This question has two parts, each of which focuses on a different aspect of the nonregular languages. Don't worry if the two parts seem entirely unrelated – that's intentional.

Let Σ = {**a**, **b**} and consider the language **a***. This language is regular, since it's the language of a regular expression. However, below is an ***incorrect*** proof that **a*** is not a regular language:

> ***Proof:*** We will prove that there are infinitely many strings distinguishable relative to **a***. By the Myhill-Nerode theorem, this shows **a*** is not a regular language.
>
> Consider any string of the form $\mathbf{a}^n$ for some $n \geq 1$ and any string of the form $\mathbf{b}^m$ for some $m \geq 1$. We can append the string a to both of these strings, forming the strings $\mathbf{a}^{n+1}$ and $\mathbf{b}^m\mathbf{a}$. This first string ($\mathbf{a}^{n+1}$) is in **a*** because it consists of some number of copies of **a**. However, the second string ($\mathbf{b}^m\mathbf{a}$) is not in **a*** because it begins with the character **b**. Therefore, these strings are distinguishable relative to **a***. Since there are infinitely many strings of each type, we see that there are infinitely many strings that are distinguishable relative to **a***. Therefore, by the Myhill-Nerode Theorem, we can conclude that **a*** is not a regular language. ■

Something must be wrong with this proof, but what is it?

i. **(3 Points)** Explain what is wrong with this proof. Be specific.

A ***palindrome*** is a string that's the same when read forwards and backwards. For example, the strings **abba**, ε, **aa**, **babaabab**, **a**, **abbabbbabba**, and **b** are all palindromes.

Let Σ = {**a**, **b**} and consider the following language *L* over Σ:

$L = \{ w \in \Sigma^* \mid w$ is a palindrome and the number of **a**'s in w is a multiple of four $\}$.

Note that zero counts as a multiple of four, so $\varepsilon \in L$.

ii. **(3 Points)** Write a context-free grammar for *L*.