CMSC330 Spring 2011 Midterm #1 - SOLUTION

Name ________________________________

Discussion Time (circle one):  9am  10am  11am  12pm  1pm  2pm

Do not start this exam until you are told to do so!

Instructions

• You have 75 minutes to take this midterm.
• This is a closed-book exam. No notes or other aids are allowed.
• If you have a question, please raise your hand and wait for the instructor.
• Answer essay questions concisely using 2-3 sentences. Longer answers are not necessary and a penalty may be applied.
• To be eligible for partial credit, show your work and clearly indicate your answers.
• Write neatly. Credit cannot be given for illegible answers.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 True or False</td>
<td>/10</td>
</tr>
<tr>
<td>2 Regular expressions &amp; finite automata</td>
<td>/20</td>
</tr>
<tr>
<td>3 RE to NFA to DFA</td>
<td>/20</td>
</tr>
<tr>
<td>4 CFGs</td>
<td>/20</td>
</tr>
<tr>
<td>5 Ruby Programming</td>
<td>/30</td>
</tr>
<tr>
<td>Total</td>
<td>/100</td>
</tr>
</tbody>
</table>
1. (10 pts) Indicate whether each of the following statements is true (“T”) or false (“F”)

a. _T_ Ruby is a scripting language.

b. _T_ In Ruby classes are also objects.

c. _F_ Ruby uses static type checking.

d. _F_ The backend of a compiler handles program parsing.

e. _F_ There is no such thing as an interpreted object-oriented language.

f. _T_ A language is regular if and only if there is a DFA that accepts it.

g. _T_ Every DFA is also an NFA.

h. _F_ Every NFA is also a DFA.

i. _F_ If a string is accepted by a CFG then there is a unique parse tree for that string.

j. _T_ Every string accepted by a CFG is also a sentential form of the CFG.
2. **Regular expressions and finite automata**

a. (5 pts) Give a regular expression over the alphabet \{0,1\} for binary numbers with no leading 0s. For example, the language contains 0, 10, 11, etc., but not \(\varepsilon, 00, 01, \text{etc.}\)

\[0 | 1(0|1)^*\]

b. (5 pts) Give a DFA for the language in part a.

(closed state $S_3$ can be omitted.)

![DFA Diagram](image)


c. (5 points) Suppose that $s$ is a string in Ruby, that $R$ is a regular expression, and that the Ruby expression $s \sim /R/$ evaluates to 0. Does this mean that $s$ is in the language of $R$? Explain.

No it does not. If $s \sim /R/$ evaluates to 0 then this means that a substring of $s$ beginning at position 0 in $s$ has been found that is in the language of $R$. However, only this prefix of $s$ is guaranteed to be in \([R]\); it does not mean that all of $s$ is in \([R]\).

d. (5 pts) A class of languages is closed with respect to union if whenever two languages $L_1$, $L_2$ are in the class, then so is $L_1 \cup L_2$. Is the class of regular languages closed with respect to union? (That is, if two languages are regular, is their union also regular?) Briefly explain your answer.

The class of regular languages is indeed closed with respect to union. Suppose that $L_1$ and $L_2$ are regular languages. This means there are regular expressions $r_1$ and $r_2$ such that $L_1 = [r_1]$ and $L_2 = [r_2]$. By definition of $\cup$, then, $r_1 | r_2$ is a regular expression with $[r_1 | r_2] = L_1 \cup L_2$, and thus $L_1 \cup L_2$ is regular.
3. (20 pts) **RE to NFA to DFA**
   
a. (10 pts) Create a NFA for the regular expression $0(0|1)^*0$ *using the method described in lecture*. In particular, be sure to show the regular expressions associated with each state, and that the transitions and final-state labelings match what the method would produce.

   (Leading $\varepsilon$ can be omitted.)

   ![NFA Diagram](image1)

   ![NFA Diagram](image2)

   ![DFA Diagram](image3)

   ![DFA Diagram](image4)

   b. (10 pts) Apply the subset construction algorithm discussed in class to convert the following NFA to a DFA. Show the NFA states associated with each state in your DFA.
4. (20 pts) CFGs
   a. (8 points) Give a context-free grammar for the language \( L = \{ w \in \{a,b\}^* \mid w \text{ is a palindrome} \} \). A palindrome is a string that reads the same from left to right and right to left. So \( \varepsilon \), aba, and abba are palindromes, while ab and baba are not.

   \[
   S \rightarrow \varepsilon | a | b | aSa | bSb
   \]

   b. (8 points) Consider the grammar \( G \) given as follows.

   \[
   E \rightarrow E \circ E \mid E \Box E \mid (E) \mid 0 \mid 1
   \]

   Give two different parse trees for the string \( 0 \circ 1 \Box 0 \).

   ![Parse Trees]

   c. (4 points) Consider the following grammar, which is unambiguous and accepts the same language as the grammar in part b.

   \[
   E \rightarrow E \circ C \mid C
   C \rightarrow C \Box T \mid T
   T \rightarrow (E) \mid 0 \mid 1
   \]

   Which operator has higher precedence, according to this grammar: \( \circ \) or \( \Box \)? (Hint: construct the parse tree for \( 1 \Box 0 \circ 1 \Box 0 \). What does this tell you?)

   \( \Box \) has higher precedence.

   ![Parse Tree]
5. (30 pts) **Ruby programming**

In this problem, you will write a series of Ruby methods that, combined, work with airport weather data. The initial input you are given is a file containing lines of the form

```
dd-mm-yyyy, hh:mm, apt, temp
```

where

- **dd** is a two-digit day, ranging from 01 to 31
- **mm** is a two-digit month, ranging from 01 to 12
- **yyyy** is a four-digit year (any four digits are allowed)
- **apt** is a three-letter airport code that must be one of: DCA, BOS, ORD, ATL, SFO
- **temp** is a temperature value, which may be any (positive, zero or negative) integer

The following methods on arrays / strings may be helpful.

```
a = s.split(“x”)  // returns array of substrings of s delimited by x
s.to_i            // returns integer value for string s
a = h.keys        // returns keys in hash h as an array a
a.sort!           // sorts elements of array a in place
a.size            // number of elements in the array
a.each { ... }    // apply code block to each element in array
a.push / a.pop     // treat array as stack
```

a. (10 pts) Write a Ruby regular expression that exactly matches the line format for the input file. That is, if `R` is your regular expression and `l` is an input line from the file, then `l =~ R` should return true if and only if the line is in the correct format. (You do not need to check whether the day is valid for the given month, e.g., for purposes of this problem, 31-02-2011 is a valid date.)

```
mm = "0[1-9]|1[0]|11|12"
yyyy = "\d{4}"  # extra backslash to avoid escape in string
hh = "[0-1][0-9]|2[0-3]"
min = "[0-5][0-9]"
apt = "DCA|BOS|ORD|ATL|SFO"
temp = "-?\d+"   # yes, -0 is allowed
r = Regexp.new("(#{dd})-(#{mm})-(#{yyyy}),(#{hh}):(#{min}),(#{apt}),(#{temp})")
```

(Newline issues ignored.)
b. (10 pts) Write a Ruby method `parse(l)` that, given a line `l` in the format above, returns an array of the form `[dd, mm, yyyy, hh, mm, apt, temp]`, where `dd, mm, yyyy, hh, and temp` are integers and `apt` is a string. You can assume the line is valid according to the regexp in part (a). We’ll call the kind of array returned by `parse()` an *entry*.

```ruby
def parse(l)
  r = l.chomp.split(/[,:]/)  
  r = r[0].split('-') + r[1..-1]
  [0,1,2,3,4,6].each { |i| r[i] = r[i].to_i }
  r
end
```

c. (10 pts) Write a Ruby method `avg(data, apt, dd, mm, yyyy)` that takes an array of entries (i.e., an array of the arrays returned by `parse`), a three-letter airport code as a string, and integers corresponding to a day, month and year, and returns the average temperature recorded across all entries for that airport on that date. For this problem, it doesn’t matter what time of day those temperatures were recorded—just average them all.

```ruby
def avg(data, apt, dd, mm, yyyy)
  n = sum = 0
  data.each { |e|
      n += 1
      sum += e[6]
    end
  }
  sum/n
end
```