

## Problem Set 1

---

- For each of the following languages, either write a regular expression that recognizes the language, or briefly explain why no such regular expression exists.
  - The language of valid identifiers in C: all strings of letters, numbers and underscores which do not begin with a number.
  - The language of strings over  $\{a, b\}$  which contain exactly two  $as$ .
  - The language of strings over  $\{a, b\}$  with more  $bs$  than  $as$ .
  - The language of strings over  $\{0, 1\}$  such that two zeros are separated by a string of length  $4i$  for any  $i$  (*i.e.*, there exists a pair of such zeros in the string).
- Write a CFG in Backus-Naur Form (not EBNF) for the following. Make sure your CFG is unambiguous.
  - A Java while loop (you may assume  $\langle \text{stmt} \rangle$  and  $\langle \text{boolean-expr} \rangle$  are defined).
  - Valid e-mail addresses. (For simplicity, assume the user name part and each part of the domain name contain only lower-case letters.)
  - The language  $L = \{a^i b^j c^k \mid k = i + j\}$ . ( $a^i$  means a string of  $as$  where there are exactly  $i$   $as$ .)
- The Java<sup>TM</sup> Tutorial describes the `switch` statement at <http://java.sun.com/docs/books/tutorial/java/nutsandbolts/switch.html>.
  - Write a CFG in BNF for the Java `switch` statement. You may assume that the following non-terminals are already defined:  

$\langle \text{stmt} \rangle$	Statement (including loops, conditional, method calls, etc, but <b>not break</b> ;) )
$\langle \text{int-expr} \rangle$	Integer expression
  - Draw a parse tree for the following code, using your grammar. You may assume that `x`, `1`, and `2` can be derived directly from  $\langle \text{int-expr} \rangle$ , and that `stmt1`, `stmt2`, and `stmt3` can be derived directly from  $\langle \text{stmt} \rangle$ .

```
switch (x) {
    case 1: stmt1;
    case 2: stmt2; break;
    default: stmt3;
}
```
  - Rewrite your grammar concisely using EBNF.

4. What languages are generated by the following grammars? Try to describe the language concisely in English—*e.g.*, “All strings of zeros and ones which contain at least 5 zeros”. You may also use simple mathematical notation, such as  $a^n b^n$ , or  $w_i$  to refer to the  $i$ th position in string  $w$ .

(a)  $\langle S \rangle \rightarrow a \langle S \rangle a \mid b \langle S \rangle b \mid c \langle S \rangle c \mid a \mid b \mid c \mid \langle \text{empty} \rangle$

(b)  $\langle S \rangle \rightarrow \langle S \rangle a \langle S \rangle b \langle S \rangle \mid \langle \text{empty} \rangle$

(c)  $\langle S \rangle \rightarrow a \langle S \rangle b \langle C \rangle \mid \langle \text{empty} \rangle$

$\langle C \rangle \rightarrow \langle \text{empty} \rangle \mid c \mid c \langle C \rangle$

(d)  $\langle S \rangle \rightarrow \langle A \rangle \langle S \rangle \langle B \rangle \mid \langle \text{empty} \rangle$

$\langle A \rangle \rightarrow ac \mid a$

$\langle B \rangle \rightarrow cb \mid b$

5. For each of the grammars above, indicate whether it is ambiguous or not. If it is ambiguous, prove it.

## Due Date

---

This assignment is due Monday, 23 September at noon.

## Silent Policy

---

A silent policy will take effect 24 hours before this assignment is due. This means that no question asked after noon on Sunday, 22 September will be answered, whether it is asked on the newsgroup, by e-mail or in person.

## Handing It In

---

You can hand in this assignment on paper in the drop box (Sandford Flemming, 2nd floor, by the overpass to Pratt) or give it to your TA **before** the beginning of the day section tutorial. You may also hand it in electronically on CDF using the following command:

```
submit -c csc324h -a PS1 ps1.txt
```

We will accept your submission in plain ASCII (in which case your submission file must be called `ps1.txt`), in PostScript (`ps1.ps`), or in PDF (`ps1.pdf`). We recommend ASCII—the electronic submission option is meant to be convenient and easy, not to create additional work for you, so don’t waste your time on a fancy presentation. You may draw your parse trees using any reasonable representation, as long as it is clear.

Type “`man submit`” on CDF for more information about the submit command.