

## CS 335 - Homework 1

### Deadline:

IF turned in on-paper: 11:59 am on Friday, February 4

IF submitted electronically: 11:59 pm on Friday, February 4

### How to submit:

Because of the nature of this homework, you have a choice of submitting it on-paper or electronically.

You can hand-write your answers, hand-write and scan them, type them, or whatever *legible* combination you choose.

#### ***IF you choose to submit it on-paper:***

- turn in your homework paper in my mailbox in the department office (BSS 320) by the on-paper deadline shown above. Note that if the department office is closed when you go to turn it in, you can place it in the department drop-box to the left of the department office door.

#### ***IF you choose to submit it electronically:***

- save your file in either .pdf or .txt format, with file name 335hw01.pdf or 335hw01.txt
- submit your file using `~st10/335submit` on nrs-labs, with a homework number of 1, by the electronic-submission deadline shown above

#### **Instructions for using the tool `~st10/335submit`:**

- If it is not already on nrs-labs, transfer your file to be submitted to a directory on nrs-labs.
  - If you are in a campus lab, you can do so by copying it to a folder on the U: drive
  - If you are not in a campus lab, you can do so by using `sftp` (secure file transfer) and connecting to `nrs-labs.humboldt.edu`, and then transferring it
- Once your file to be submitted is in a directory on nrs-labs, then use `ssh` (or Putty in an Academic Computing lab) to connect to `nrs-labs.humboldt.edu`.
- use `cd` to change to the directory containing the file to be submitted -- for example,

```
cd 335hw01
```

- type the command:

```
~st10/335submit
```

...and give the number of the homework being submitted (or whatever number you have been asked to do for lab-related files) when asked, and answer that, `y`, you do want to submit all of the files with `of-interest-to-335` suffixes in the current directory. (Note that I don't mind if a few extraneous files get submitted as well -- I'd rather receive too many files than too few, and typing in all of the file

names for some assignment is just too error-prone...)

- you are expected to carefully check the list of files that the tool believes have been submitted, and make sure all of the files you hoped to submit were indeed submitted! (The most common error is to try to run `~st10/335submit` while in a different directory than where your files are...)

## Purpose:

To practice with regular expressions (RE's), context-free grammars (CFG's), Backus-Naur Form (BNF), derivations, and derivation trees/parse trees

## The Problems:

### ***Problem 1 - 18 points***

For each of the following regular expressions (RE's):

- assume that the alphabet is  $\{0, 1\}$
- describe the language denoted by that RE
- give at least 4 different example strings in the language denoted by that RE
- **either** give at least two examples of strings **not** in that language (but drawn from the same alphabet), **or** explain why you cannot
- give an example of a string from those string or strings in the language whose length is the shortest possible for that language

#### **1 part a**

$(0 + 1)^*1(0 + 1)(0 + 1)(0 + 1)$

#### **1 part b**

$(0 + 1)^*0$

#### **1 part c**

$(0 + 1)^*110^*$

### ***Problem 2 - 20 points***

For each of the following, give an equivalent regular expression. Assume that the alphabet for all of the languages below is  $\{0, 1\}$ .

#### **2 part a**

$\{w \mid w \text{ begins with a } 1 \text{ and ends with a } 0\}$

#### **2 part b**

$\{w \mid w \text{ contains the substring } 1101\}$

**2 part c** $\{w \mid w \text{ has the length } 5\}$ **2 part d** $\{w \mid w \text{ begins with } 00 \text{ or } 111\}$ **Problem 3 - 22 points**

Consider the following context-free grammar  $G$ .

 $A \rightarrow DAD \mid B$  $B \rightarrow mCh \mid hCm$  $C \rightarrow DCD \mid D \mid \epsilon$  $D \rightarrow m \mid h \mid p$ **3 part a**

What are the variables/nonterminals of  $G$ ?

**3 part b**

What are the terminals of  $G$ ?

**3 part c**

Give a **derivation** for the string **hmhm**

**3 part d**

Give a **derivation** for the string **mphmh**

**3 part e**

Give a **derivation tree/parse tree** for the string **hmhm**

**3 part f**

Give a **derivation tree/parse tree** for the string **mphmh**

**Problem 4 - 21 points**

Consider the BNF for an Algol-60-style number posted along with this homework handout (it is adapted from the MacLennan course text, page 152, but you should use the posted version for this question, because of those adaptations). Use this in answering the following questions.

**4 part a**

List all the **terminals** in this BNF.

**4 part b**

List all the **nonterminals/syntactic categories** within this BNF.

**4 part c**

We talked about derivations of strings from a CFG, and how, from a BNF, you can derive examples of legal "chunks" of code from BNF describing a language, or parts of a language.

Consider the following such derivation, showing that 3.87 is indeed a "legal" Algol-60 number: (note that this is not the only possible derivation...)

```

<number> => <unsigned number>
          => <decimal number>
          => <unsigned integer><decimal fraction>
          => <digit><decimal fraction>
          => 3<decimal fraction>
          => 3.<unsigned integer>
          => 3.<unsigned integer><digit>
          => 3.<unsigned integer>7
          => 3.<digit>7
          => 3.87

```

Remembering to only substitute for **one** variable per step in such derivations:

Is  $1E-7$  an Algol-60 number, according to the this BNF? If so, give a **derivation** or a **derivation tree** (your choice) for it; if not, explain why not (according to the above BNF)

**4 part d**

Is  $2E0.6$  an Algol-60 number, according to the this BNF? If so, give a **derivation** or a **derivation tree** (your choice) for it; if not, explain why not (according to the above BNF)

**4 part e**

Is  $E-5$  an Algol-60 number, according to the this BNF? If so, give a **derivation** or a **derivation tree** (your choice) for it; if not, explain why not (according to the above BNF)

**4 part f**

Is  $.8$  an Algol-60 number, according to the this BNF? If so, give a **derivation** or a **derivation tree** (your choice) for it; if not, explain why not (according to the above BNF)

**4 part g**

Is  $4.$  an Algol-60 number, according to the this BNF? If so, give a **derivation** or a **derivation tree** (your choice) for it; if not, explain why not (according to the above BNF)

**Problem 5 - 19 points**

(adapted from MacLennan, Exercise 4-7, page 152)

**5 part a**

Write a BNF description of identifiers that have the following syntax:

A letter is either a, b, c, or d. A digit is either 2, 4, or 6. An identifier is a string of one or more letters, digits, and underscores (`_`), subject to the following restrictions:

- an identifier must start with a letter
- consecutive underscores are not permitted
- an identifier may not end with an underscore

Hint: start by writing out some examples of legal and illegal identifiers so that you can identify the pattern.

**5 part b**

Write a derivation or a derivation tree (your choice), using your BNF, that shows that the string `ab_b4d_c` is a legal identifier.