**Todd Wenker**

**CSE 340 Fall 2015**

**HOMEWORK 1**

Assigned 8/31/2015

Due 9/9/2015 by 11:59:59 pm on blackboard.

**All submissions must be PDF and should be typed. Exception can only be made for drawing parse trees, which can be handwritten and scanned in the submitted document.**

**Note.** Some of these problems are taken from the Dragon book.

**Problem 1**. Consider the following regular expressions (we omit the dot operator)

R0 = 1|2|3|4|5|6|7|8|9

R1 = 0|1|2|3|4|5|6|7|8|9

R2 = (0|1)\* R0 (0|1)

R3 = 00 R0\*(0|1)\*

R4 = R3\* R2\*000

Assume that the longest prefix-matching rule is used. Assume that ties are broken in favor of the regular expression listed first in the list.

1. Give an example of input for which getToken() returns R0­
   1. getToken() 🡪 1. This would return R­0  as it is a single digit (R2-4 are sequences of digits) and because R0  comes before R1 in the list of regular expressions, the tie is broken in its favor.
2. Give an example of input for which getToken() returns R1
   1. getToken() 🡪 0. R1 is a single digit from 0-9. However, because R0 is a single digit from1-9 and is listed in the list of regular expressions first, the only way getToken() will return R1 is when the input is just 0.
3. Give an example of input for which getToken() returns R2
   1. getToken()🡪 1001081. The regular expression R2 is split into three parts. The first part, (0|1)\*, means a sequence of 1’s and 0’s need to start the input. In this input, 10010 satisfy that condition. The second part, R0, means that after the first segment, the input must have a single digit between 1 and 9. The third part, (0|1) means that the sequence must end with either a 0 or a 1. This input has a 1 at the end so it satisfies this condition.
4. Give an example of input for which getToken() returns R3
   1. getToken() 🡪003426401001. The regular expression R3 is split into three parts. The first is that it starts with 00
5. Give an example of input for which getToken() returns R4
   1. getToken()🡪 00551 101010191 000. The segment of the input 00551 satisfies the R3\* portion of the expression. The segment 101010191 satisfies the R2\* portion of the expression. 000 is always the ending three digits of this expression.
6. If getToken() if called repeatedly on the following input, what is the sequence of tokens returned?

99001101678100010101030123457000010

9🡪R0

9🡪R0

001101🡪R3

6🡪 R0

7🡪 R0

8🡪 R0

100010101030 🡪 R2

1🡪 R0

2🡪 R0

3🡪 R0

4🡪 R0

5🡪 R0

7🡪R0

000010🡪R2

**Explain your answers**

**Problem 2.** Consider the grammar

S → AB

A → BaA | bB

B → aSB | AS | ε

1. Show that this grammar is ambiguous by constructing two different leftmost derivations for the sentence abab. Answer below.
2. Show that this grammar is ambiguous by constructing two different parse trees for the string abab. Answer below.

**Problem 3.** Compute FIRST sets for the following grammar.

S → aAB | CD

A → CD | SE | ε

B → aSB | AS

C → cC | ε

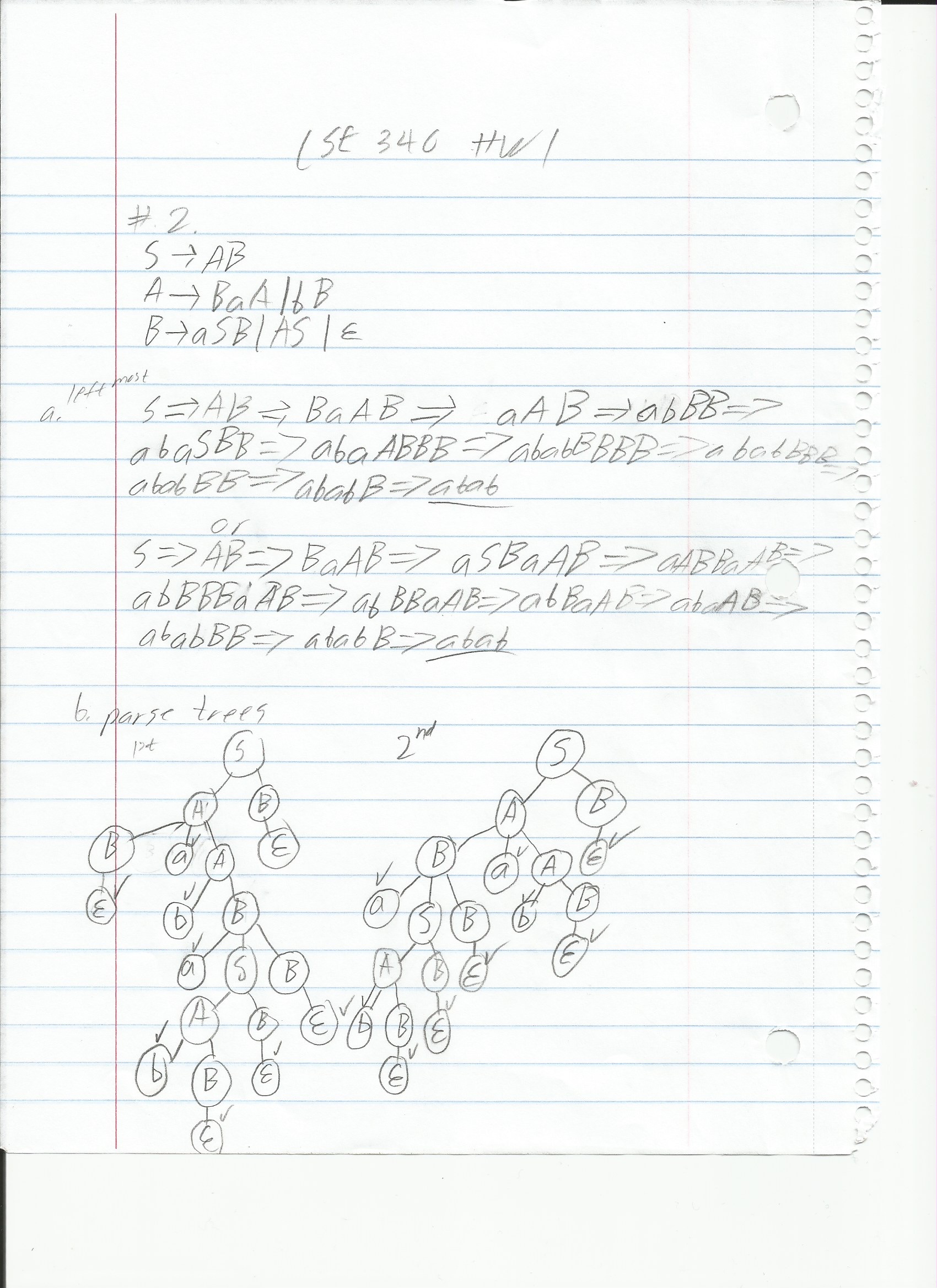
D → CDd | ε

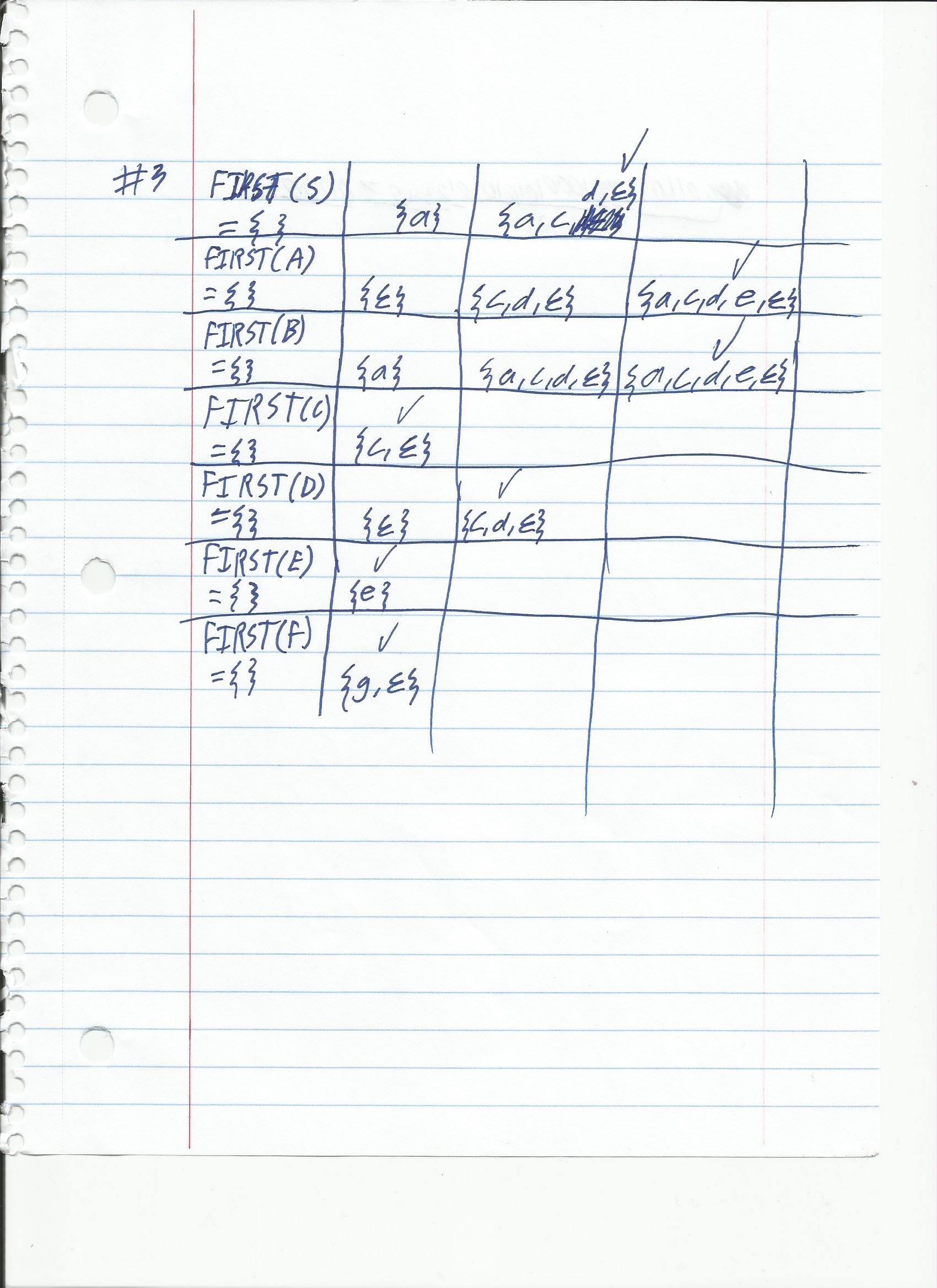
E → eFg

F → Fg | ε

Answer below.

**Show your work. An answer by itself does not count.**



****