CS143 Spring 2025 – Written Assignment 2

Due Monday, April 28, 2025 11:59 PM PDT

This assignment covers context free grammars and parsing. You may discuss this assignment with other students and work on the problems together. However, your write-up should be your own individual work, and you should indicate in your submission who you worked with, if applicable. Assignments can be submitted electronically through Gradescope as a PDF by 11:59 PM PDT. Please review the the course policies for more information: `https://web.stanford.edu/class/cs143/policies/`. A LaTeX template for writing your solutions is available on the course website. If you need to draw parse trees in LaTeX, you may use the `forest` package: `https://ctan.org/pkg/forest`.

1. Give a context-free grammar (CFG) for each of the following languages. Any grammar is acceptable—including ambiguous grammars—as long as it has the correct language. The start symbol should be S.

   (a) The set of all strings over the alphabet $\{a, b, c\}$ such that the number of a's plus the number of b's is divisible by 3. Example Strings in the Language:

   ε abbc   aaacc   bbbacabccc

   Strings not in the Language:

   a   bb   acbc   abbbccc

   Solution:

   (b) The set of all strings over the alphabet $\{x, (,), ;\}$ representing nested tuples of x's where each tuple has an even length.

   Example Strings in the Language:

   $()$ $(x; ())$ $((); x; ((); x); x)$

   Strings not in the Language:

   ε x $((); x; x)$ $(x; (); (x; (); x); x)$

   Solution:

(c) The set of all strings over the alphabet $\{0, 1\}$ where no consecutive 0's appear (no substring "00"). Example Strings in the Language:

$$\varepsilon \qquad 1 \qquad 0 \qquad 01 \qquad 10 \qquad 101 \qquad 010 \qquad 1010101$$

Strings not in the Language:

$$00 \qquad 100 \qquad 001 \qquad 1001 \qquad 10010$$

Solution:

(d) The set of all strings over the alphabet $\{0, 1\}$ in the language $L : \{0^i 1^j 0^k \mid j = i + k\}$. Example Strings in the Language:

$$\varepsilon \qquad 10 \qquad 01 \qquad 0110 \qquad 000111$$

Strings not in the Language:

$$0 \qquad 00 \qquad 001 \qquad 0010 \qquad 01110$$

Solution:

2. Consider the following grammar for binary strings that involves the alphabet $\{a, b\}$:

$$E \to Ea \mid Eb \mid aE \mid bE \mid T$$
$$T \to a \mid b \mid \varepsilon$$

Is this grammar ambiguous or not? If yes, give an example of an expression with two different parse trees, draw the parse trees, and make the grammar unambiguous. If not, explain why it is unambiguous.

Solution:

3. (a) Eliminate left recursion from the following grammar:

$$
\begin{aligned}
S &\to S(T) \mid Sa \mid [T] \mid Tb \\
T &\to T(S) \mid Tc \mid d
\end{aligned}
$$

Solution:

(b) Left factor the following grammar:

$$
\begin{aligned}
S &\to (T + T) \mid (T) \\
T &\to U * T \mid U{*}? \mid [U] \\
U &\to U0 \mid U1 \mid \varepsilon
\end{aligned}
$$

Solution:

4. Consider the following CFG, where the set of terminals is $\{0, 1, (,), ;\}$:

$$
\begin{aligned}
S &\to (T \\
T &\to CA \mid) \\
A &\to ; B \mid) \\
B &\to CA \mid) \\
C &\to 0 \mid 1 \mid S
\end{aligned}
$$

(a) Construct the FIRST sets for each of the nonterminals.
Solution:

- S:
- T:
- A:
- B:
- C:

(b) Construct the FOLLOW sets for each of the nonterminals.
Solution:

- S:
- T:
- A:
- B:
- C:

(c) Construct the LL(1) parsing table for the grammar.
Solution:

Nonterminal	(	)	;	0	1	$
S						
T						
A						
B						
C						

(d) Show the sequence of stack, input and action configurations that occur during an LL(1) parse of the string "(( ) ; 0)". At the beginning of the parse, the stack should contain a single S.
Solution:

Stack	Input	Action

5. Consider the following grammar G over the alphabet $\{x, =\}$:

$$
\begin{aligned}
S' &\to S \\
S &\to L = R \\
S &\to R \\
L &\to x \\
R &\to L
\end{aligned}
$$

You want to implement G using an SLR(1) parser. Note that we have already added the $S' \to S$ production for you.

(a) Construct the DFA of the LR(0) machine, and identify all conflicting states and conflicts that prevent the grammar from being LR(0).

Solution:

(b) Now, for each conflicting state in the DFA that prevents it from being LR(0), identify the FOLLOW sets of the left-hand nonterminals. Is the grammar SLR(1)? Explain. Your explanation must reference at least one of the identified FOLLOW sets from each conflicting DFA state.

Solution: