# Homework #7

#1. True/ False

a. The Pumping Lemma for CFL's can be used to show a language is context-free
True False

b. The string $z = a^k b^{k+1} c^k$ can be used to show $\{a^n b^n c^n\}$ is not context free
True False

c. The string $z = a^k a^k a^k$ can be used to show $\{w\, w^R w \mid w\ \varepsilon\ \{a,b\}^*\}$ is not context-free
True False

d. Given a CFG, G, and a string w, it is decidable whether w ε L(G) True False

e. The intersection of a context-free language and a regular language is context-free
True False

#2. Why must we remove the recursive Start to convert to Chomsky Normal form?

If S occurs on the right hand side, and $S \rightarrow \lambda$ is a production, then S can occur in the middle of a derivation and essentially shorten a sentential form. We don't want this because we want a relationship between the length of the derivation and the length of the string generated.

#3. What is the relationship between the length of a string and the length of its derivation if the grammar is in Greibach normal Form?

If there are no lambda productions, each production produces exactly 1 terminal, so the derivation is the same as the string length in this case.

#4. Convert the following grammar to Chomsky Normal Form. Show work clearly.

Solution

Remove Recursive Start:

$$S' \to S$$
$$S \to aSb \mid BB \mid BCD \mid ab \mid BC$$
$$A \to B \mid DD \mid BCB \mid D \mid \lambda$$
$$B \to AB \mid C$$
$$C \to Cc \mid c$$
$$D \to Scc \mid cc$$

Remove λ Productions:

A is nullable, but the others are not (no production has an "A" on its rhs).

New grammar:

S' → S
S → aSb | BB | BCD| ab | BC
A → B | DD | BCB | D
B → AB | B |C
C → Cc | c
D → Scc |cc

Remove unit rules:

S' → aSb | BB | BCD| ab | BC
S → aSb | BB | BCD| ab | BC

A → AB | Cc | c | DD | BCB | Scc |cc
B → AB | Cc | c
C → Cc | c
D → Scc |cc

Remove useless:

Term = {D, C, A, S, S'} so all useful so far

Reach = {S', S, B, C, D, A} so all useful:

S' → aSb | BB | BCD| ab | BC
S → aSb | BB | BCD| ab | BC
A → AB | Cc | c | DD | BCB | Scc |cc
B → AB | Cc | c
C → Cc | c
D → Scc |cc

Convert to CNF (lots of ways to do this)

$S' \rightarrow AT_1 \mid BB \mid B\ T_3 \mid A'\ T_2 \mid BC$
$A' \rightarrow a$
$T_1 \rightarrow S\ T_2$
$T_2 \rightarrow b$
$T_3 \rightarrow CD$
$A \rightarrow AB \mid CT4 \mid c \mid DD \mid BT_6 \mid Scc \mid cc$
$S \rightarrow AT_1 \mid BB \mid B\ T_3 \mid A\ T_2 \mid BC$
$B \rightarrow AB \mid CT_4 \mid c$
$T_4 \rightarrow c$
$D \rightarrow S\ T_5 \mid T_4\ T_4$
$T_5 \rightarrow T_4\ T_4$
$T_6 \rightarrow CB$

#5. Show the following languages are or are not context-free:

a) $\{a^i b^j c^m \mid i \neq j \neq m\}$

Many of you told me this was not context-free. I think you got caught in what $i \neq j \neq m$ means. Note that when $i = 2, j = 3, m = 2, i \neq j \neq m$. ($\neq$ is not transitive).

Here is a grammar:
In what follows:

A generates 0 or more *a*'s
B generates 0 or more *b*'s
C generates 0 or more *c*'s

E first generates an equal number of *b*'s and *c*'s then 1 or more *b*'s (via B) or 1 or more *c*'s (via *cC*).
So E ➔ b*c* (#b ≠ #c)
Similarly D ➔ a*b* (#a ≠ #b)

It is context-free:

$S \rightarrow A\ E \mid D\ C$
$A \rightarrow a\ A \mid \lambda$
$B \rightarrow b\ B \mid b$
$C \rightarrow c\ C \mid \lambda$
$D \rightarrow a\ D\ b \mid B \mid a\ A$
$E \rightarrow b\ E\ c \mid B \mid cC$

S ➔ A E ➔ a A E ➔ a a A E ➔ a a E ➔ a a b E c ➔ a a b b E c c ➔ aabbbcc
etc.

b) $\{a^i b^j c^m \mid i < j < m;\ i,j,k \geq 0\}$

Not context free:

- Let n be the integer from the CFL pumping lemma
- Choose $x = a^n b^{n+1} c^{n+2}$
- Consider all strings u, v, w, y, z such that
  - x = uvwyz
  - |vwy| <= n
  - |vy| >= 1
- There are five possible cases for strings v and y
  - v and y contain at least 1 a and 0 b's and c's
  - v and y contain at least 1 a and 1 b and 0 c's
  - v and y contain at least 1 b and 0 a's and c's
  - v and y contain at least 1 b and 1 c and 0 a's
  - v and y contain at least 1 c and 0 a's and b's
    - In summary, v and y cannot contain both a's and c's.
      - This follows from the fact that |vwy| <= n
    - v and y must contain at least one character
      - This follows from the fact that |vy| >= 1
- We will show for each case that there exists a k such that $uv^kwy^kz$ is not in L
- Case 1: v and y contain at least 1 a and 0 b's and c's
  - Choose k=2
  - $uv^2wy^2z$ is not in L
    - $uv^2wy^2z$ contains exactly n+1 b's
      - This follows from the fact that uvwyz contains exactly n+1 b's and
      - v and y contain 0 b's
    - $uv^2wy^2z$ contains at least n+1 a's
      - This follows from uvwyz contains exactly n a's and
      - v and y contain at least 1 a
- Case 2: v and y contain at least 1 a and b and 0 c's
  - Choose k=2
  - $uv^2wy^2z$ is not in L
    - $uv^2wy^2z$ contains exactly n+2 c's
      - This follows from the fact that uvwyz contains exactly n+2 c's and
      - v and y contain 0 c's
    - $uv^2wy^2z$ contains at least n+2 b's
      - This follows from uvwyz contains exactly n+1 b's and
      - v and y contain at least 1 b
- Case 3: v and y contain 0 a's and c's and at least 1 b
  - Choose k=2
    - Could choose k=0 as well

            - $uv^2wy^2z$ is not in L
                - $uv^2wy^2z$ contains exactly n+2 c's
                    - This follows from the fact that uvwyz contains exactly n+2 c's and
                    - v and y contain 0 c's
                - $uv^2wy^2z$ contains at least n+2 b's
                    - This follows from uvwyz contains exactly n+1 b's and
                    - v and y contain at least 1 b
    - Case 4: v and y contain 0 a's and at least 1 b and 1 c
        - Choose k=0
        - uwz is not in L
            - uwz contains exactly n a's
                - This follows from the fact that uvwyz contains exactly n a's and
                - v and y contain 0 a's
            - uwz contains at most n b's
                - This follows from uvwyz contains exactly n+1 b's and
                - v and y contain at least one b
    - Case 5: v and y contain 0 a's and b's and at least 1 c
        - Choose k=0
        - uwz is not in L
            - uwz contains exactly n+1 b's
                - This follows from the fact that uvwyz contains exactly n+1 b's and
                - v and y contain 0 b's
            - uwz contains at most n+1 c's
                - This follows from uvwyz contains exactly n+2 c's and
                - v and y contain at least 1 c
    - In all cases, we have shown that there exists a choice of k such that $uv^kwy^kz$ is not in L.
    - Thus L does not satisfy the pumping condition.
    - Thus L is not a context-free language.