Week 6. By Deng Yufan.

Prob 1. ~~Design~~ Design CFG.

a. $\{a^i b^j c^k \mid i \neq j \vee j \neq k\}$.

(start) $S \to aPC \mid QbC \mid AbR \mid ATc$
$A \to aA \mid \varepsilon$
$C \to cC \mid \varepsilon$.
$P \to aP \mid aPb \mid \varepsilon$
$Q \to Qb \mid aQb \mid \varepsilon$.
$R \to bR \mid bRc \mid \varepsilon$.
$T \to Tc \mid bTc \mid \varepsilon$.

b. $\{w \in \{a,b\}^* \mid \forall w' \in \{a,b\}^*, w \neq w'w'\}$.

(start) $S \to A \mid B \mid AB \mid BA$.

$A \to a \mid aAa \mid aAb \mid bAa \mid bAb$

$B \to b \mid aBa \mid aBb \mid bBa \mid bBb$.

c. $\{a^n b^m \mid 3n \geq m \geq 2n\}$.

(start) $S \to aSbb \mid aSbbb \mid \varepsilon$.

d. $\{a^m b^n c^p d^q \mid m+n = p+q\}$.

(start) $S \to aSd \mid P \mid Q$

$P \to bPd \mid R$

$Q \to aQc \mid R$

$R \to bRc \mid \varepsilon$.

e. $\{a^n b^i c^j d^m \mid n+m = i+j\}$.

(start) $S \to PA \mid BQ$

$P \to aPb \mid \varepsilon$. $A \to bAd \mid C$ $B \to aBc \mid D$
$Q \to cQd \mid \varepsilon$. $C \to cCd \mid \varepsilon$ $D \to aDb \mid \varepsilon$.

Prob 2. Consider CFG $G: S \to aSbS \mid bSaS \mid \varepsilon$.

Prove $L(G) = \{w \in \{a,b\}^* \mid \#a = \#b \text{ in } w\}$.

~~Induction with struct of Parse tree to prove~~

We prove $\overset{S(n)=}{\{w \in \{a,b\}^* \mid |w| = 2n,\ \#a = \#b\}} \subseteq L(G)$ for all $n \in N$ by induction n.

As basis. $S \to \varepsilon$, so the statement is true when $n=0$.

As induction. we assume $\le n$ and prove $n+1$. For all $w \in S(n+1)$, if w starts with a, we prove we can always write w as $asbt$, where $\#a(s) = \#b(s)$, $\#a(t) = \#b(t)$.

Prove by contradiction. We know for all prefix x, which $\#a(x) = \#b(x)+1$, the successor of x can't be b. Since w start with a, we can $\boxed{\text{simply induction}}$ to prove for all nonempty prefix x, $\#a(x) > \#b(x)$. But $\#a(w) = \#b(w)$. contradiction.

Simply induction: As basis. w starts with a. so $\#a(w)=1 > \#b(w)=0$. As induction. write the prefix x as yc, where c is a char. If $\#a(y) > \#b(y)+1$, then whatever c is a or b, we have $\#a(x) > \#b(x)$. If $\#a(y) = \#b(y)+1$, then $c = a$, since c can't be b we have known. so $\#a(x) = \#a(y)+1 = \#b(y)+2 = \#b(x)+2$.

So we write w as $asbt$, and $|s| \le 2n$, $|t| \le 2n$. So $s, t \in L(G)$ as the induction assumption. So $S \Rightarrow aSbS \Rightarrow^* asbt = w$. then $S(n+1) \subseteq L(G)$.

So G generates all w that $\#a(w) = \#b(w)$. On the other side, we prove $S \Rightarrow^n_{lm} w \to \#a(w) = \#b(w)$ for all $n \ge 1$ by induction n. this finish the prove.

As basis. $S \overset{\text{only}}{\Rightarrow^1_{lm}} \varepsilon$, and $\#a(\varepsilon) = \#b(\varepsilon) = 0$.

As induction. we assume $\le n$ and prove $n+1$. we have

$$S \Rightarrow^{n+1}_{lm} w \iff aSbS \Rightarrow^n_{lm} w \text{ or } bSaS \Rightarrow^n_{lm} w.$$

$$\iff \exists\, 1 \le i < n,\ S \Rightarrow^i_{lm} s,\ S \Rightarrow^{n-i}_{lm} t.\quad w = asbt \text{ or } bsat.$$

As induction assumption, $\#a(s) = \#b(s)$, $\#a(t) = \#b(t)$. so $\#a(w) = \#b(w)$.

Prob 3. Consider $E \to +EE \mid *EE \mid -EE \mid x \mid y$.

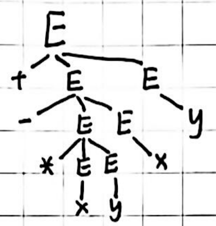
a. Find leftmost, rightmost derivations, and a derivation tree for $+*-xyxy$.

$E \Rightarrow_{lm} +EE \Rightarrow_{lm} +*EEE \Rightarrow_{lm} +*-EEEE$

$\Rightarrow_{lm} +*-xEEE \Rightarrow_{lm} +*-xyEE \Rightarrow_{lm} +*-xyxE \Rightarrow_{lm} +*-xyxy$.

$E \Rightarrow_{rm} +EE \Rightarrow_{rm} +Ey \Rightarrow_{rm} +*EEy \Rightarrow_{rm} +*Exy$

$\Rightarrow_{rm} +*-EExy \Rightarrow_{rm} +*-Eyxy \Rightarrow_{rm} +*-xyxy$.



b. Prove the CFG is unambiguous.

~~We prove for all $w \in L(G)$, there is only one way leftmost derivations for w by induction $|w|$.~~

We first prove for all $n \geq 1$, $E \Rightarrow^{n}_{lm} w \to$ for every ~~nonempty~~ proper prefix x of w, $\#(+*-)(x) \geq \#(xy)(x)$, and $\#(+*-)(w)+1 = \#(xy)(w)$ by induction n.

As basis, when $n=1$, S only generates x and y. Easy to confirm the statement is tr

As induction, we assume $\leq n$ and prove $n+1$.

$E \Rightarrow^{n+1}_{lm} w \iff (+*-)EE \Rightarrow^{n}_{lm} w$ (where () means or).

$\iff E \Rightarrow^{i}_{lm} l,\ E \Rightarrow^{n-i}_{lm} r,\ w = (+*-)lr.\ (1 \leq i < n)$.

We have $\#(+*-)(x)+1 \geq \#(xy)(x)$ for every prefix (no need to be proper) of l by induction assumption. Since w starts with $(+*-)$, so $\#(+*-)(x) \geq \#(xy)(x)$ for all prefix of w up to the end of l.

We also have $\#(+*-)(l)+1 = \#(xy)(l)$, and $\#(+*-)(x) \geq \#(xy)(x)$ for every proper prefix of r. By easy calculation we can have $\#(+*-)(x) \geq \#(xy)(x)$ for all proper prefix of w.

And $\#(+*-)(l)+1 = \#(xy)(l)$, $\#(+*-)(r)+1 = \#(xy)(r)$, w starts with $(+*-)$.
So $\#(+*-)(w)+1 = \#(xy)(w)$.

So the statement is preserved.

Then we prove for all $w \in L(G)$, there is only one leftmost derivation for w by induction on $n = |w| \geq 1$.

As basis, when $n=1$, w can only be x and y. Easy to see there is only one way.

As induction, we assume $\leq n$ and prove $n+1$.

Since $\#(+*-)(x) \geq \#(xy)(x)$, where x is the first char of w, w can only start with $+$, $-$, or $*$. So the first step of derivation is determined to be $E \to +EE \mid *EE \mid -EE$ depending on what w start with.

Then, let w' be w delete the first char. we need to write w' into l and r, that $E \Rightarrow^*_{lm} l$, $E \Rightarrow^*_{lm} r$, to find a derivation for w.

As induction assumption, there is only one leftmost derivation for l and r. So if we proof there is only one way to write w' into lr, the only way of derivation is

$$E \Rightarrow_{lm} (+*-)EE \Rightarrow^*_{lm} (+*-)lE \Rightarrow^*_{lm} (+*-)lr = w.$$

(depend on what w starts with. but only one way)

Since $w \in L(G)$, we can always find at least one way to write w' into lr. If we can write w' into lr and $l'r'$, where $|l| < |l'|$.

As we proved, $\#(+*-)(l)+1 = \#(xy)(l)$. Since l is a proper prefix of l', $\#(+*-)(l) \geq \#(xy)(l)$. Contradiction.

So the induction part is preserved and completes the proof.

Prob 4. Give ambigous and unambiguous CFG for $\{a^m b^n \mid m \geq 2n \geq 0\}$.

Ambiguous: (start) $S \to aSb \mid aSbb \mid \varepsilon$.

Unambiguous: (start) $S \to aSb \mid T$

$T \to aTbb \mid \varepsilon$.