9-3

$E = \{ 0^i 1^j \mid i > j \}$

Given: $p \in N$

Choose: $s \in E$ s.t. $|s| \geq p$

$s = 0^{2p} 1^p$

Given: $s = xyz \quad |y| > 0 \quad |xy| \leq p$

$xyz = 0^{2p} 1^p \quad x = 0^a \quad y = 0^b \quad z = 0^c 1^p$

$a + b + c = 2p \quad a + b \leq p \quad b > 0$

Choose: $i \in N$ s.t. $xy^iz \notin E$

$xy^iz = 0^a 0^{bi} 0^c 1^p$

$a + bi + c > p$

$2p + (i-1)b > p$

$(i-1)b > -p \longrightarrow i-1 > -p/b \Rightarrow i > -p/b + 1$

$-(i-1)b < p$ $\quad p > 0 \quad b > 0$

$b - bi < p$

$s = 0^{p+1} 1^p \quad xyz = 0^{p+1} 1^p \quad x = 0^a \quad y = 0^b \quad z = 0^c 1^p$

$a + b + c = p + 1$

$xy^iz \Rightarrow a + bi + c > p$

$i = 0 \Rightarrow a + c > p \quad b \neq 0 \quad a + c = p \quad p \not> p$

$PLUS = \{ 0^n 1 0^m 1 0^{n+m} \mid n, m \in N \}$

"$n + m = \underline{n+m}$"

Given: p

Choose: $s = 0^p 1 0 1 0^{p+1}$ "$p + 1 = p + 1$"

$x = 0^a \quad y = 0^b \quad z = 0^c 1 0 1 0^{p+1} \quad a + b + c = p$

$xy^iz = 0^{a+bi+c} 1 0 1 0^{p+1}$ $\in PLUS$ iff

$a + bi + c + 1 = p + 1$

$a + bi + c = p = a + b + c$

$bi = b$

$i = 1$

10-1/ Acceptors (DFAs) $\rightarrow$ Language Denot (REX)

$\Downarrow$ (str $\rightarrow$ t or f) $\qquad\qquad$ $\Downarrow$ (rex $\rightarrow$ Set(str))

Context-Free Languages (CFL)

denotation: Context-Free Grammar (CFG)

acceptor: Push-Down Automata (PDA)

CFL like REG

CFG like REX

PDA like DFA

Example CFG:

$A \rightarrow 0A1$ ①

$A \rightarrow \varepsilon$ ②

rule, or production or a substitution rule

lhs ∈ Variable

rhs $\in (\text{Vars} \cup \Sigma)^*$

terminals

non-terminals

one variable is the start variable (A) variable on the lhs of the first rule

A "derivation" of a string of Grammar G:

$A \xrightarrow{①} 0A1 \xrightarrow{①} 00A11 \xrightarrow{①} 000A111 \xrightarrow{②} 000111$

"000111" $\mapsto$ 1112

$A \rightarrow AB$ ① $\quad$ $A \rightarrow 1$ ② $\quad$ $B \rightarrow 0$ ③

$A \xrightarrow{①} AB \xrightarrow{②} 1B$

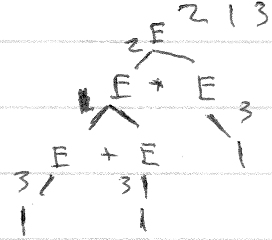
$AB \xrightarrow{③} A0$

A parse tree is the sequence of rules

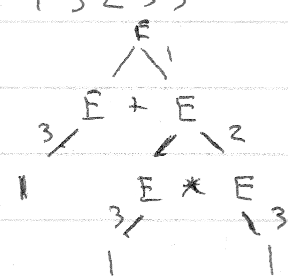
1. $E \rightarrow E + E$
2. $E \rightarrow E * E$
3. $E \rightarrow 1$

What derivation produces $1 + 1 * 1$?

2 1 3 3 3



1 3 2 3 3



ambiguous grammar

A CFG $G = \langle V, \Sigma, R, S \rangle$

V is a finite set — variable / non-terminals

Σ is a finite set — terminals

$V \cap \Sigma = \emptyset$

$S \in V$ — start variable

R is an indexed set of $\underbrace{V}_{\text{lhs}} \times \underbrace{(V \cup \Sigma)^*}_{\text{rhs}}$

every element has a natural label

$d \in$ string of indices

$L(G) = \{ w \mid w \in \Sigma^* \text{ and } S \overset{d}{\Rightarrow}^* w \}$

$u \overset{d}{\Rightarrow}^* v$ (u derives v with d) $u, v \in (V \cup \Sigma)^*$

$d \in N^*$

$u \overset{\varepsilon}{\Rightarrow}^* u$

$u \overset{r}{\Rightarrow} x$ and $x \overset{d}{\Rightarrow}^* V \implies u \overset{rd}{\Rightarrow}^* V$

$u \overset{r}{\Rightarrow} v$ (u yields v with r) $u, v \in (V \cup \Sigma)^*$ $r \in N$

$u \in \Sigma^*$ $A \in V$ $w, v \in (V \cup \Sigma)^*$ $r \in N$

$uAv \overset{r}{\Rightarrow} uwv$ iff $R(r) = (A, w)$

$S \Rightarrow aSb \mid SS \mid \varepsilon$

$V = \{S\}$ $\Sigma = \{a, b\}$ $S = S$

$R = 0 \mapsto (S, aSb)$

$1 \mapsto (S, SS)$

$2 \mapsto (S, \varepsilon)$

$S \overset{d}{\Rightarrow}^* aabaabbb$? $d = 0102002$

$S \overset{0}{\Rightarrow} aSb \overset{1}{\Rightarrow} aSSb \overset{0}{\Rightarrow} aaSbSb \overset{2}{\Rightarrow} aabSb \overset{0}{\Rightarrow} aabaSbb \overset{0}{\Rightarrow}$

$aabaaSbbb \overset{2}{\Rightarrow} aabaabbb$

10-3/ Equal numbers of 0 and 1

$$S \to \varepsilon \mid 0S1 \mid 1S0 \mid SS$$

CFGs are closed under union:

$G_1 = (V_1, \Sigma, R_1, S_1)$ $\quad$ $G_2 = (V_2, \Sigma, R_2, S_2)$

$G_3 = (V_3, \Sigma, R_3, S_3)$

$S_3 \to S_1 \mid S_2$ $\quad$ $V_3 = (1 \times V_1) \cup (V_2 \times 1) \cup \{S_3\}$

CFGs are closed under star:

$S^* \Rightarrow \varepsilon \mid S S^*$

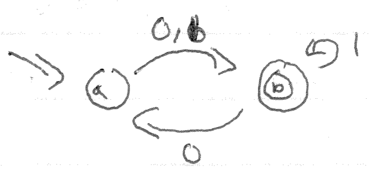
DFA $\subseteq$ CFG

DFA $d = \langle Q, \Sigma, \delta, q_0, F \rangle$

CFG $G = \langle Q, \Sigma, R, q_0 \rangle$

$q_i \to a q_j$ iff $\delta(q_i, a) = q_j$

$q_i \to \varepsilon$ iff $q_i \in F$



$a \to 0b \mid 1b$

$b \to \varepsilon \mid 1b \mid 0a$

$$a \overset{0}{\Rightarrow} 0b \overset{4}{\Rightarrow} 01b \overset{5}{\Rightarrow} 010a \overset{2}{\Rightarrow} 0101b \overset{0}{\Rightarrow} 0101$$

REG are CFGs where rule $(R_x \to \Sigma^* R_y \mid \varepsilon)$

$(R_x \to (V \cup \Sigma)^*)$

$\Sigma^* R_x \Rightarrow \Sigma^* R_y$

Noam Chomsky

10-4

Chomsky Normal Form:

Every rule in R is either:

$A \to BC$ $\quad A \in V,\ B \in V-S,\ C \in V-S$

or $A \to a$ $\quad a \in \Sigma$

or $S \to \varepsilon$

Algorithm:
1. Add a new start state
2. Collapse $A \to \varepsilon$ rules up
3. Remove unit rules $(A \to B)$
4. Add intermediate variables

Example: $S \to 0S1 \mid \varepsilon$

(1) $S_{new} \to S$
$S \to 0S1 \mid \varepsilon$

(2) $S_{new} \to S \mid \varepsilon$
$S \to 0S1 \mid 01$

(3) $S_{new} \to 0S1 \mid 01 \mid \varepsilon$
$S \to 0S1 \mid 01$

(4) $S_{new} \to AC \mid AB \mid \varepsilon$
$S \to AC \mid AB$
$A \to 0$ $\quad C \to SB$
$B \to 1$

$S \to \varepsilon$
$S \to a$
$S \to AB$
$AB \to aB$
$AB \to CDB$

If input string has n-chars the derivation is ..
tree is at most $n+1$ high
when $|s| = 2^n+1$