COL 352 Lecture 17 Feb 18

Language class	Generation	Recognition	Properties	Limitations
Regular	reg. expr.	DFA/NFA	P.L., Closure properties, decision algo	Can't handle $a^i b^i$ etc.
Context Free				

Set of rules

1. $S \to ab$ Production rules
2. $S \to aSb$
3. $S \to \epsilon$

$$S \xrightarrow{2} aSb \xrightarrow{2} aaSbb \xrightarrow{2} aaaSbbb$$

$$\xrightarrow{1} aaaabbbb$$

$$S \xrightarrow{*} w \mid w = a^i b^i, \; i \geq 0$$

apply repeatedly one of the rules

Any sequence of substitution must begin with the special variable S

$$G: \left\{ \begin{array}{ll} V = S & s = S \\ T = \{a, b, \epsilon\} & P = 1, 2, 3 \end{array} \right\}$$

Convention: Capital letters for variables
small case for terminals

$$AB \longrightarrow ABB$$

Context Free Grammar (CFG), the prodn rules have exactly one symbol on the LHS

Grammar $G = (V, T, P, S)$

V: set of variables that appear of LHS

T: Set of alphabet Terminals

P: Set of rules or productions

S: $\in V$ start symbol

L_{eq}: Equal number of a's and b's

$a^i b^i$ ababab, aababb,

Is L_{eq} regular?

Context Free Language (CFL): all languages that can be generated using CFG

Is L_{eq} CFL?

$S, \{a, b, \varepsilon\}, S,$

$$S \to \varepsilon \mid ba \mid ab \mid baS \mid abS$$

$V = \{S, A, B\}$ $\quad T = \{a, b\}$

$S \to aB \mid bA \mid \varepsilon$

$A \to a \mid bAA \mid aS$

$B \to b \mid aBB \mid bS$

aababb

$S \to aB \to aaBB \to aabB \to aabaBB \xrightarrow{*} aababb$

Claim 1. $S \xrightarrow{*} w$ iff w has equal # a's and b's

for $|w| \geq 1$ 2. $A \xrightarrow{*} w$ iff w has one more a than b

3. $B \xrightarrow{*} w$ iff w has one more b than a

Proof by induction on $|w|$

Base case $|w| = 1$

S: no strings of length 1, so true

A: $A \to a$ only strings of length 1

B: $B \to b$

Suppose true for all $|w| \leq k-1$

Consider any string $|w| = k$

S

$w \to aw_1$, $w \to bw_1$

If w has equal # a's and b's then $S \xrightarrow{*} w$

$|w_1| = k-1$

$S \to aB \xrightarrow{*} aw_1$

since B generates all string $\leq$ length $k-1$ with one extra b

So $B \xrightarrow{*} w_1$

Conversely if $S \xrightarrow{*} w$ then $w = a w_1$ (extra b)

or $w = b w_2$ (extra a)

Let $S \to aB$ So $B \xrightarrow{*} w_1$

from I.H. w_1 has an extra b

Prove it for all the three assertions

A, B, S and their converse

$A \to a \mid bAA \mid aS$

If $A \xrightarrow{*} w$ $|w| = k$ then w has one more a than b

Let $w = b w_1$ $A \to bAA \xrightarrow{*} b w_1' A \xrightarrow{*} b w_1' w_2''$

Suppose $A \to w_1'$ $A \to w_1''$ $|w_1'| \leq k-1$

$|w_1''|$

From I.H. w_1', w_1'' will have one more a than b

So overall one more a than b

If w has one more a than b then $A \xrightarrow{*} w$

$|w| = k$ $w = a w_1$ $w_1 \neq \varepsilon$ eq # a's and b's

$A \to aS \xrightarrow{*} a w_1$ where $S \xrightarrow{*} w_1$

$w = b\,w_1$ $\qquad$ w_1 has 2 extra a's than b's

$w_1 = \underbrace{x_1\, x_2\, x_3 \ldots}_{A}\ \underbrace{x_{k-1}}_{A}$ $\qquad$ $x_i \in \{a, b\}$

difference between #a's and b's for each position of the string w_1

0 -1 0 1 0 1 2

b a a b a a

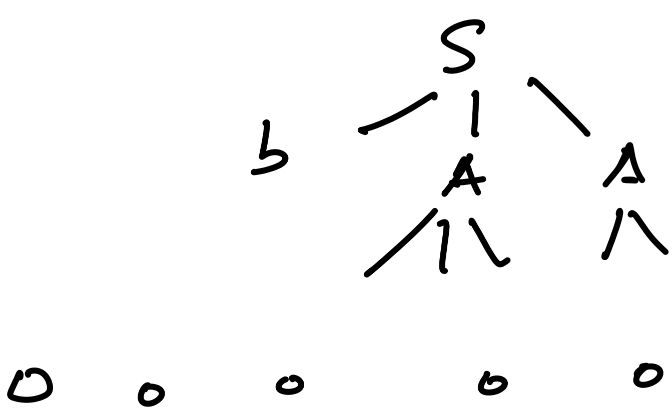
H.W. Problem. Design a CFG for strings over a, b s.t.

#a's = 2x #b's

Different ways of writing CFG

Membership problem

Given a CFG $G = (V, T, S, P)$ and a string $w \in T^*$ does $S \xrightarrow{*} w$



Derivation Tree

← Terminals

Canonical forms of CFG

Chomsky Normal Form	Greibach Normal Form
$A \to BC$	$A \to a\,BCD$
$A \to a$	$A \to a$

Claim: Any given CFG can be transformed into an equivalent CNF or GNF