Context-free languages

- CNF
- Membership
- Emptineess

Closure properties

CFLs are closed under union, concatenation and $*$.

$$L = \{a^n b^n c^n \mid n \in \mathbb{N}\}$$

$$L_1 = \{a^n b^n c^m \mid m, n \in \mathbb{N}\}$$
$$L_2 = \{a^m b^n c^n \mid m, n \in \mathbb{N}\}$$

CFLs

CFLs are not closed under intersection or complement.

Chomsky Normal Form

- Find all nullable variables
  $(X \Rightarrow^* \epsilon)$
- Eliminate all rules of the form
  $Y \rightarrow \epsilon$
- Find all pairs (X, Y) s.t.
  $X \Rightarrow^* Y$
- Using these pairs, eliminate all unit rules $(A \rightarrow B)$
- Rewrite rules so that RHS has only 2 var.

$A \rightarrow BC$

$A \rightarrow a$

$S \to ASA \mid aB$

$A \to B \mid S$

$B \to b \mid \epsilon$

Nullable : $\{B, A\}$.

Removing the rules of the form $X \to \epsilon$

$S \to ASA \mid SA \mid AS \mid \cancel{S} \mid aB \mid a$

$A \to B \mid S$

$B \to b$

(A,B)

(A,S)

$S \to ASA \mid SA \mid AS \mid aB \mid a$

$A \to b \mid ASA \mid SA \mid AS \mid aB \mid a$

$B \to b$

Removing unit rules

Copying from previous slide

$S \rightarrow ASA \mid SA \mid AS \mid aB \mid a$

$A \rightarrow b \mid ASA \mid SA \mid AS \mid aB \mid a$

$B \rightarrow b$

Getting rid of rules $X \rightarrow w$ where w has letters and variables

$S \rightarrow ASA \mid SA \mid AS \mid X_aB \mid a$

$A \rightarrow b \mid ASA \mid SA \mid AS \mid X_aB \mid a$

$B \rightarrow b$

$X_a \rightarrow a$

Making RHS length 2 or less.

CNF

$S \rightarrow X_{AS}A \mid SA \mid AS \mid X_aB \mid a$

$A \rightarrow b \mid X_{AS}A \mid SA \mid AS \mid X_aB \mid a$

$B \rightarrow b$

$X_{AS} \rightarrow AS \quad ; \quad X_a \rightarrow a$

Effectiveness of CNF

Let $G = (V, \Sigma, R, S)$ be a CFG in CNF.

$$S \Rightarrow w_1 \Rightarrow w_2 \Rightarrow w_3 \Rightarrow \dots w_k = w \in \Sigma^*.$$

Let $|w| = n$

Then each w_i : $|w_i| \leq w$.

But notice there are no unit rules.

The number of "$A \rightarrow a$" rules used in the derivation is n, and number of "$A \rightarrow BC$" rules is at most $n-1$

If G is in CNF,

and $S \Rightarrow w_1 \ldots w_k \Rightarrow w$

then $k \leq 2n-1$

Hence membership for CFGs is solvable:

- Given CFG G and $w \in \Sigma^*$ does $w \in L(G)$?
  - Convert CFG to CNF
  - Check all derivations of size $2|w|-1$
  - If any of them generate w, say YES else say NO

Emptiness.

Given a CFG G, is $L(G) = \phi$?

$$S \to AB$$
$$A \to B \mid BB$$
$$B \to cD \mid c$$

$S \to A$
$A \to AB$
$B \to AA$
$C \to d$

B can generate something.
A can generate something
S can generate something.

C can generate something.

Algorithm

$Y := \{ X \mid X \to \alpha,\ \alpha \overset{\in \Sigma^*}{\text{made of } \Sigma \text{ only}} \}$

while (Y does not stabilize) {

$\quad Y := Y \cup \{ X \mid X \to w$ and every nonterminal/var in w is in $Y \}$;

}

$L(G) \neq \emptyset$ iff $S \in Y$ (start var. belongs to Y).

$Y := \{ X \mid X \to \alpha,\ \alpha \in \Sigma^* \}$

$Y_{old} := \emptyset$

while ($Y \neq Y_{old}$) {

$\quad Y_{old} := Y$

$\quad Y = Y \cup \{ \ldots \}$

}

Closure properties contd.

Closure under homomorphism.

$G \quad A \rightarrow aAb \mid \epsilon$

$h(a) = 00 \quad h(b) = 1$

$G' : \quad A \rightarrow 00A1 \mid \epsilon$

Claim. $h(L(G)) = L(G')$

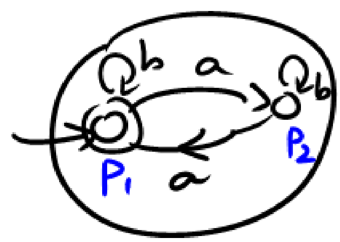
Hence CFLs are closed under homomorphisms.

Given a CFG & NFA

L L'

Is $L \cap L'$ a CFG?

$A \to aAb \mid \epsilon$



$$S \to A_{p_1,p_2} \mid \epsilon$$

$$A_{p_1,p_2} \to a \cancel{A_{p_2,p_1}} b \quad a A_{p_2,p_2} b$$

$$A_{p_2,p_2} \to a A_{p_1,p_2} b \mid \epsilon$$

In general, let

$G = (V, \Sigma, R, S)$ be in CNF

and $A = (Q, \Sigma, \delta, q_0, F)$ be an DFA.

$$G' = (\{X_{q,q'} \mid q, q' \in Q,\ X \in V\} \cup \{S_0\},$$
$$\Sigma,\ R',\ S_0)$$

$$S_0 \to S_{q_0, q_f} \quad \forall\, q_f \in F.$$

$$X_{q,q'} \to Y_{q,q''} Z_{q'',q'} \qquad X \to YZ$$

$$X_{q,q'} \to a \quad \text{iff } q \xrightarrow{a} q' \qquad X \to a$$