1) Unione di grammatiche. Supponiamo di avere $G_i = (V_i, \Sigma_i, R_i, S_i)$ tutte CFG $\forall i \in [k]$. $k \in \mathbb{N}$.

Vogliamo CFG $G = (V, \Sigma, R, S)$ t.c. $L(G) = \bigcup_i L(G_i)$.

Idea naturale:

- $V = \bigcup_i V_i \cup \{S\}$. Wlog, possiamo assumere $V_i \cap V_j = \emptyset$ $\forall i, j \in [k]$ t.c. $i \neq j$.

- S nuova variabile iniziale

- $\Sigma = \bigcup_i \Sigma_i$

- $R = \bigcup_i R_i \cup \{ S \to S_1 \mid \dots \mid S_k \}$

Correttezza: $\bigcup_i L(G_i) = L(G)$.

Dimostro prima $\bigcup_i L(G_i) \subseteq L(G)$.

Sia $w \in \bigcup_i L(G_i)$: $\exists\, j \in [k]$ t.c. $w \in L(G_j)$. Ovvero:

$[k] = \{1, 2, \dots, k\}$

$S_j \overset{*}{\Longrightarrow}_{G_j} w$

Ma allora per definizione:

$$S \Rightarrow S_i \overset{*}{\Rightarrow}_{G_i} w$$

ovvero $w \in L(G)$.

Altra direzione: $L(G) \subseteq \bigcup_i L(G_i)$.

Sia $w \in L(G)$, ovvero $S \overset{*}{\Rightarrow}_G w$.

Per definizione $\exists\, i \in [k]$ t.c.

$$S \Rightarrow S_i \overset{*}{\Rightarrow}_G w$$

Siccome V_i è disgiunto da tutte le altre variabili:

$$S \Rightarrow \underbrace{S_i \overset{*}{\Rightarrow}_{G_i} w}_{w \in L(G_i) \subseteq \bigcup_j L(G_j)}$$

Quindi

$\blacksquare$

2) Da DFA a CFG. Dato DFA $D = (Q, \Sigma, \delta, q_0, F)$ voglio definire $G = (V, \Sigma, R, S)$ t.c.

$$L(G) = L(D).$$

Idea naturale:

- $V = \{ V_q : q \in Q \}$
- $S = V_{q_0}$
- Aggiungo le regole $V_q \to a V_p$

per ogni $p, q \in Q$, $a \in \Sigma$ t.c.

$$\delta(q, a) = p.$$

- $\forall\, q \in F$, aggiungo $V_q \to \varepsilon$.

Esercizio: Dimostrare la correttezza in generale.

(Esempio:

$$q_0 \xrightarrow{a} q_1 \xrightarrow{b} q_2$$

$w = ab \in L(A)$)

G pure produce $w = ab$;

$$S = V_{q_0} \xrightarrow[R_1]{} a\, V_{q_1} \quad ; \quad V_{q_1} \xrightarrow[R_2]{} b\, V_{q_2}$$

$$; \quad V_{q_2} \xrightarrow[R_3]{} \varepsilon$$

$$S \underset{R_1}{\Rightarrow} a V_{q_1} \underset{R_2}{\Rightarrow} ab\, V_{q_2} \underset{R_3}{\Rightarrow} ab \quad ,)$$

3) È spesso utile usare la ricorsione

$$R \to 0R1$$

Regole di questo tipo consentono di ricordare "informazione limitata".

ESERCIZI: Progettare CFG per i linguaggi

$L_1 = \{ w \in \{0,1\}^* : w$ contiene almeno 3 '1' $\}$

$$L_2 = \{ w \in \{0,1\}^* : w = w^R , |w| \text{ è pari} \}$$

$$L_3 = \{ a^i b^j c^k : i + j = k, \; i, j, k \geq 0 \}$$

Vediamo ora che le CFG hanno una forma CANONICA.

DEF (FORMA NORMALE DI CHOMSKY)

Una CFG è in forma normale se ogni regola è del tipo:

$$A \to BC$$

$$A \to a$$

con $A, B, C \in V$ e $a \in \Sigma$.

$B, C \neq S$

Inoltre è ammessa la regola $S \to \varepsilon$.

TEO Ogni CFG ammette una CFG equivalente in forma normale.

DIM. In primis, aggiungo var. iniziale S_0 insieme alla regola $S_0 \to S$.

Poi elimino le ε-regole, $B \to \varepsilon$.

Per ogni occorrenza di B a DX di una regola, aggiungo nuove regole con occorrenza cancellata. Es:

$R \to uBv$ aggiungo $R \to uv$

$R \to u B v B w$ aggiungo

$R \to u v B w$; $R \to u B v w$

; $R \to u v w$

Poi, regole unitarie $A \to B$. Le elimino e poi per ogni occorrenza di regola $B \to u$ aggiungo $A \to u$ (a meno che questa regola non era già stata eliminata).

Infine trasformo regole restanti

$$A \to u_1 u_2 \dots u_k \qquad k \geq 3$$

$$A \to u_1 A_1 \; ; \; A_1 \to u_2 A_2$$

$$\dots \; A_{k-2} \to u_{k-1} u_k$$

dove A_i nuove variabili. Se u_i è terminale lo sostituisco con U_i e aggiungo $U_i \to u_i$

Running Example:

$$S \to ASA \mid aB \qquad S_0 \to S$$

$$A \to B \mid S$$

$$B \to b \mid \varepsilon$$

$$S_0 \to S$$

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

(Remove $B \to \varepsilon$)

$$A \to B \mid S \mid \varepsilon$$

$$B \to b \mid \cancel{\varepsilon}$$

$S_0 \rightarrow S$

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

(Rimuovo $A \rightarrow \varepsilon$)

$A \rightarrow B \mid S \mid \cancel{\varepsilon}$

$B \rightarrow b$

$S_0 \rightarrow S$

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

$A \rightarrow B \mid S$

$B \rightarrow b$

Rimuovo $S \to S$ e $S_0 \to S$

$S_0 \to ASA \mid aB \mid a \mid SA \mid AS$

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

$A \to B \mid S$

$B \to b$

Rimuovo $A \to B$

$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$

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

$A \rightarrow b \mid S$

$B \rightarrow b$

Rimuovo $A \rightarrow S$

$S_0 \rightarrow ASA \mid aB \mid a \mid SA \mid AS$

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

$A \to b \mid ASA \mid eB \mid e \mid SA \mid AS$

$B \to b$

Completo le sostituzioni:

$S_0 \to AA_1 \mid UB \mid e \mid SA \mid AS$

$S \to AA_1 \mid UB \mid e \mid SA \mid AS$

$A \to b \mid AA_1 \mid UB \mid e \mid SA \mid AS$

$A_1 \to SA$

$U \to e$; $B \to b$