

Syntaks og semantik

Lektion 7

6 marts 2007

CFG

PDA

Lukningsegenskaber

Fra sidst

- 1 Kontekstfrie grammatikker
- 2 Pushdown-automater
- 3 Lukningsegenskaber

Definition 2.2: En kontekstfri grammatik (CFG) er en 4-tupel

$G = (V, \Sigma, R, S)$, hvor delene er

- 1 V : en endelig mængde af **variable**
- 2 Σ : en endelig mængde af **terminaler**, med $V \cap \Sigma = \emptyset$
- 3 $R : V \rightarrow \mathcal{P}((V \cup \Sigma)^*)$: **produktioner / regler**
- 4 $S \in V$: **startvariablen**

– produktioner skrives $A \rightarrow w$ i stedet for $w \in R(A)$

- Hvis $u, v, w \in (V \cup \Sigma)^*$ er ord og $A \rightarrow w$ er en produktion, siges uAv at **frembringe** uwv : $uAv \Rightarrow uwv$.
- Hvis $u, v \in (V \cup \Sigma)^*$ er ord, siges u at **derivere** v : $u \xRightarrow{*} v$, hvis $u = v$ (!) eller der findes en følge u_1, u_2, \dots, u_k af ord således at $u \Rightarrow u_1 \Rightarrow u_2 \Rightarrow \dots \Rightarrow u_k \Rightarrow v$.
- **Sproget** som G genererer er $\llbracket G \rrbracket = \{w \in \Sigma^* \mid S \xRightarrow{*} w\}$.

– dvs. et ord $w \in \Sigma^*$ genereres af G hvis og kun hvis der findes en **derivation** $S \Rightarrow w_1 \Rightarrow w_2 \Rightarrow \dots \Rightarrow w_k \Rightarrow w$, hvor alle $w_i \in (V \cup \Sigma)^*$.

3/14

Eksempel: Opgave 2.6 d (ca.)

$$\begin{aligned} S &\rightarrow A\#T\#A \\ T &\rightarrow aTa \mid bTb \mid \#A\# \\ A &\rightarrow aA \mid bA \mid \varepsilon \mid A\#A \end{aligned}$$

Genererer sproget

$$\{x_1\#x_2\#\dots\#x_k \mid k \geq 5, \text{ alle } x_i \in \{a, b\}^*, \\ \text{og } x_i = x_j^R \text{ for to indices } i \neq j\}$$

Definition 2.13: En **pushdown-automat (PDA)** er en 6-tupel $M = (Q, \Sigma, \Gamma, \delta, q_0, F)$, hvor delene er

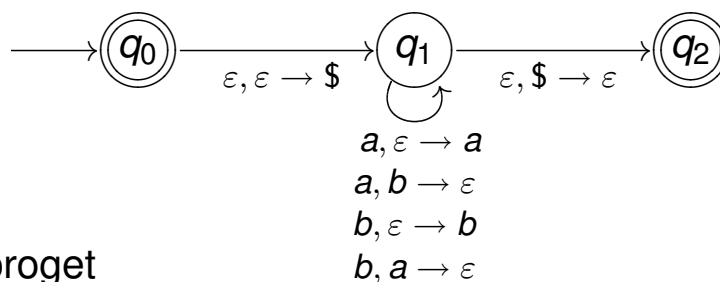
- 1 Q : en endelig mængde af tilstande
- 2 Σ : input-alfabetet
- 3 Γ : stack-alfabetet
- 4 $\delta : Q \times \Sigma_\epsilon \times \Gamma_\epsilon \rightarrow \mathcal{P}(Q \times \Gamma_\epsilon)$: transitionsfunktionen
- 5 $q_0 \in Q$: starttilstanden
- 6 $F \subseteq Q$: mængden af accepttilstande

M siges at **acceptere** et ord $w \in \Sigma^*$ hvis der findes $m \in \mathbb{N}$ og $w_1, w_2, \dots, w_m \in \Sigma_\epsilon$, $r_0, r_1, \dots, r_m \in Q$ og $s_0, s_1, \dots, s_m \in \Gamma^*$ således at $w = w_1 w_2 \dots w_m$ og

- 1 $r_0 = q_0$ og $s_0 = \epsilon$,
- 2 for alle $i = 0, 1, \dots, m-1$ findes $a, b \in \Gamma_\epsilon$ og $t \in \Gamma^*$ som opfylder $s_i = at$, $s_{i+1} = bt$ og $(r_{i+1}, b) \in \delta(r_i, w_{i+1}, a)$, og
- 3 $r_m \in F$.

5/14

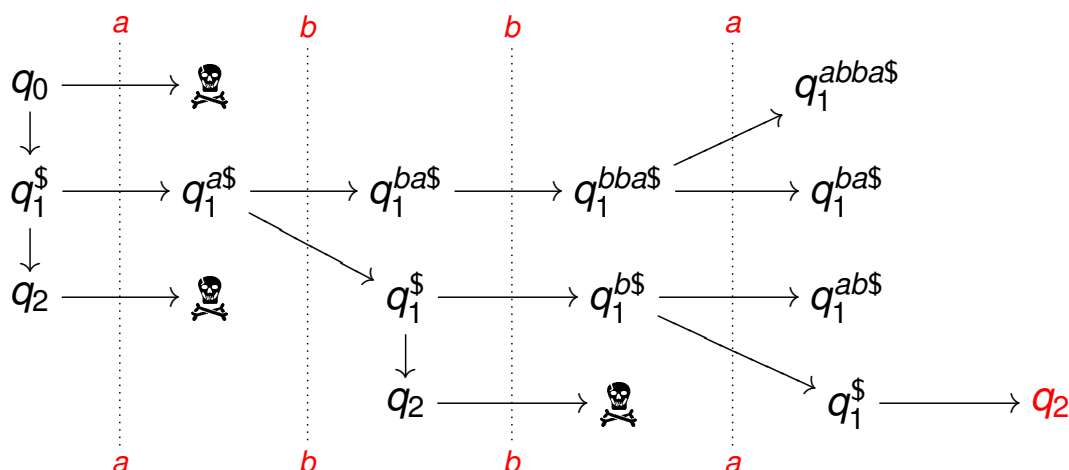
Eksempel:



Genkender sproget

$$\{w \in \{a, b\}^* \mid \text{antallet af } a \text{ i } w = \text{antallet af } b \text{ i } w\}$$

At læse strengen *abba*:



6/14

Definition: Et sprog siges at være **kontekstfrit** hvis der findes en CFG der genererer det.

Sætning 2.20: Et sprog er kontekstfrit hvis og kun hvis der findes en PDA der genkender det.

Bevis lige om lidt.

7/14

Sætning: Klassen af kontekstfrie sprog er lukket under \cup , \circ og $*$.

Bevis: (Opgave 2.8) Lad A_1 og A_2 være kontekstfrie sprog over et fælles alfabet Σ .

- \cup : Lad $G_1 = (V_1, \Sigma, R_1, S_1)$, $G_2 = (V_2, \Sigma, R_2, S_2)$ være CFGs med $\llbracket G_1 \rrbracket = A_1$ og $\llbracket G_2 \rrbracket = A_2$. Konstruér en ny CFG $G = (V, \Sigma, R, S)$ ved $V = V_1 \cup V_2 \cup \{S\}$ og $R = R_1 \cup R_2 \cup \{S \rightarrow S_1 \mid S_2\}$. Da er $\llbracket G \rrbracket = A_1 \cup A_2$.
- \circ : Lad $M_1 = (Q_1, \Sigma, \Gamma_1, \delta_1, q_1, F_1)$, $M_2 = (Q_2, \Sigma, \Gamma_2, \delta_2, q_2, F_2)$ være PDAs med $\llbracket M_1 \rrbracket = A_1$ og $\llbracket M_2 \rrbracket = A_2$. Antag at $\Gamma_1 \cap \Gamma_2 = \emptyset$. Konstruér en ny PDA $M = (Q, \Sigma, \Gamma, \delta, q_1, F_2)$ ved $Q = Q_1 \cup Q_2$, $\Gamma = \Gamma_1 \cup \Gamma_2$ og $\delta = \delta_1 \cup \delta_2 \cup \{(q_f, \varepsilon, \varepsilon) \rightarrow (q_2, \varepsilon) \mid q_f \in F_1\}$. Da er $\llbracket M \rrbracket = A_1 \circ A_2$.
- $*$: Lad $G_1 = (V_1, \Sigma, R_1, S_1)$ være en CFG med $\llbracket G_1 \rrbracket = A_1$. Konstruér en ny CFG $G = (V, \Sigma, R, S)$ ved $V = V_1 \cup \{S\}$ og $R = \{S \rightarrow \varepsilon \mid SS \mid S_1\}$. Da er $\llbracket G \rrbracket = A_1^*$.

8/14

Kontekstfrie grammatikker og pushdown-automater

- 4 Ethvert kontekstfrit sprog genkendes af en PDA
- 5 Ethvert sprog genkendt af en PDA er kontekstfrit

9/14

Lemma 2.21: Lad Σ være et alfabet og $A \subseteq \Sigma^*$ et kontekstfrit sprog. Da findes en PDA P med $\llbracket P \rrbracket = A$.

Bevis: Lad $G = (V, \Sigma, R, S)$ være en CFG med $\llbracket G \rrbracket = A$.
Idéen er at PDAen, givet en inputstreng s , nondeterministisk forsøger at finde en derivation for s i G :

- 1 Push S på stacken
- 2 Hvis topsymbalet på stacken er en variabel A : Pop A og push højresiden w af en produktion $A \rightarrow w$ i R . (Dø hvis der ikke er nogen produktion $A \rightarrow w$ i R .)
- 3 Hvis topsymbalet på stacken er en terminal a : Sammenlign med næste inputsymbol. Hvis de er ens, pop a . Hvis de ikke er ens, dø.
- 4 Gentag step 2 og 3 indtil stacken er tom.

Lemma 2.21: Lad Σ være et alfabet og $A \subseteq \Sigma^*$ et kontekstfrit sprog. Da findes en PDA P med $\llbracket P \rrbracket = A$.

Bevis: Lad $G = (V, \Sigma, R, S)$ være en CFG med $\llbracket G \rrbracket = A$.

Vi konstruerer først en “generaliseret PDA”

$P = (Q, \Sigma, \Gamma, \delta, q_s, F)$, der kan pushe strenge i stedet for bare symboler. Lad $Q = \{q_s, q_\ell, q_f\}$, $F = \{q_a\}$ og $\Gamma = V \cup \Sigma \cup \{\$ \}$.

Lad

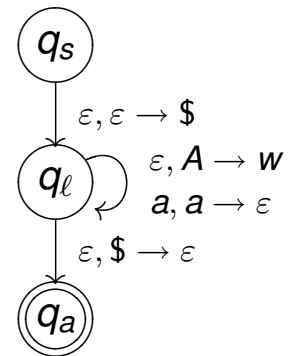
$$\delta(q_s, \varepsilon, \varepsilon) = \{(q_\ell, S\$)\}$$

$$\delta(q_\ell, \varepsilon, A) = \{(q_\ell, w) \mid w \in R(A)\} \quad \text{for alle } A \in V$$

$$\delta(q_\ell, a, a) = \{(q_\ell, \varepsilon)\} \quad \text{for alle } a \in \Sigma$$

$$\delta(q_\ell, \varepsilon, \$) = \{(q_a, \varepsilon)\}$$

$$\delta(q, a, b) = \emptyset \quad \text{for alle andre}$$



Lav til sidst P om til en “almindelig” PDA ved at erstatte enhver transition $q \xrightarrow{a, b \rightarrow s_1 s_2 \dots s_n} r$ med (nye tilstande og) en følge

$$q \xrightarrow{a, b \rightarrow s_n} q_1 \xrightarrow{\varepsilon, \varepsilon \rightarrow s_{n-1}} q_2 \longrightarrow \dots \longrightarrow q_{n-1} \xrightarrow{\varepsilon, \varepsilon \rightarrow s_1} r.$$

11 / 14

Lemma 2.27: Lad Σ være et alfabet og P en PDA over Σ . Da findes en CFG G over Σ med $\llbracket G \rrbracket = \llbracket P \rrbracket$.

Bevis: Lad $P = (Q, \Sigma, \Gamma, \delta, q_0, F)$. Vi konstruerer $G = (V, \Sigma, R, S)$:

- ① Sørg for at P kun har én accepttilstand q_a og at stacken tømmes før P går i q_a .
- ② Lad $V = \{A_{pq} \mid p, q \in Q\}$, og sørg for at A_{pq} deriverer præcist de strenge som bringer P fra p med tom stack til q med tom stack.
- ③ Lad $S = A_{q_0 q_a}$. Voilà!

Lemma 2.27: Lad Σ være et alfabet og P en PDA over Σ . Da findes en CFG G over Σ med $\llbracket G \rrbracket = \llbracket P \rrbracket$.

Bevis: Lad $P = (Q, \Sigma, \Gamma, \delta, q_0, F)$. Vi konstruerer $G = (V, \Sigma, R, S)$:

- 1 Sørg for at P kun har én accepttilstand q_a og at stacken tømmes før P går i q_a .

Nyt stacksymbol $\$$. Tre nye tilstande: q_s , q_e og q_a . Nye

transitioner: $q_s \xrightarrow{\varepsilon, \varepsilon \rightarrow \$} q_0$, $q \xrightarrow{\varepsilon, \varepsilon \rightarrow \varepsilon} q_e$ for alle $q \in F$,

$q_e \xrightarrow{\varepsilon, a \rightarrow \varepsilon} q_e$ for alle $a \in \Sigma$, og $q_e \xrightarrow{\varepsilon, \$ \rightarrow \varepsilon} q_a$.

Sørg for at enhver transition *enten* pusher *eller* popper.

- Erstat enhver transition $q \xrightarrow{a, b \rightarrow c} r$ med
 $q \xrightarrow{a, b \rightarrow \varepsilon} q_1 \xrightarrow{\varepsilon, \varepsilon \rightarrow c} r$
- Erstat enhver transition $q \xrightarrow{a, \varepsilon \rightarrow \varepsilon} r$ med
 $q \xrightarrow{a, \varepsilon \rightarrow x} q_1 \xrightarrow{\varepsilon, x \rightarrow \varepsilon} r$ for et eller andet symbol $x \in \Gamma$.

13 / 14

Lemma 2.27: Lad Σ være et alfabet og P en PDA over Σ . Da findes en CFG G over Σ med $\llbracket G \rrbracket = \llbracket P \rrbracket$.

Bevis: Lad $P = (Q, \Sigma, \Gamma, \delta, q_0, F)$. Vi konstruerer $G = (V, \Sigma, R, S)$:

- 2 Lad $V = \{A_{pq} \mid p, q \in Q\}$, og sørg for at A_{pq} deriverer præcist de strenge som bringer P fra p med tom stack til q med tom stack.

- Lav en produktion $A_{pp} \rightarrow \varepsilon$ for alle $p \in Q$ (terminering)
- Lav en produktion $A_{pq} \rightarrow A_{pr}A_{rq}$ for alle $p, q, r \in Q$ (rekursion)
- For alle $p, q, r, s \in Q$: Hvis $p \xrightarrow{a, \varepsilon \rightarrow t} r$ og $s \xrightarrow{b, t \rightarrow \varepsilon} q$ for nogle $a, b \in \Sigma_\varepsilon$ og et $t \in \Gamma$: Lav en produktion
 $A_{pq} \rightarrow aA_{rs}b$. (produktion)

– der skal *argumenteres* for at dette giver det rigtige resultat!

14 / 14