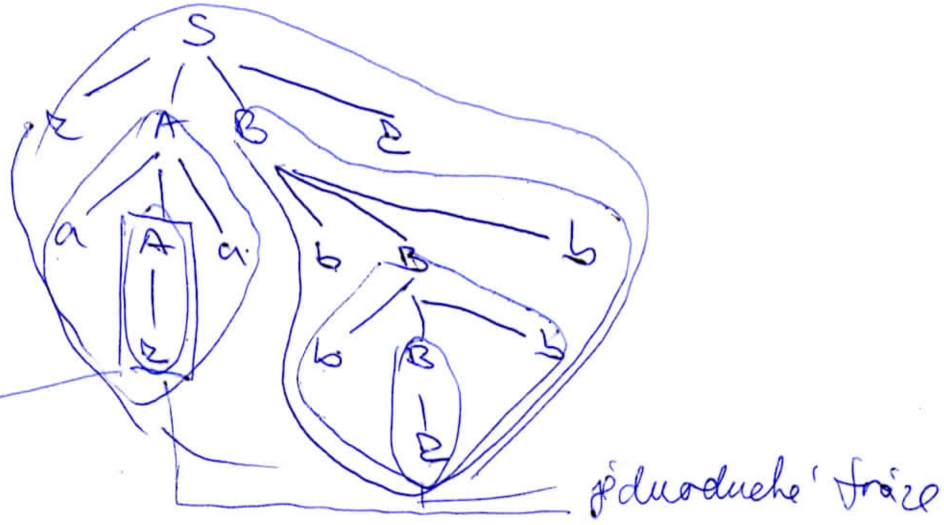
BEZKONTEXTOVÉ JAZYKY

- sestavte bezk. gramatiku, která generuje jazyk
  $L = \{ z a^n z a^n b^m z b^m z \mid m, n \geq 0 \}$.
  Uvažujte pak větu zazabbzbbz, sestavte pro ni levou derivaci, derivační strom a znázorněte v něm fráze, jednoduché fráze a l-fráze.

- $P: S \longrightarrow zABz \quad A \rightarrow aAa \mid z \quad B \rightarrow bBb \mid z$
  $G = (\{S, A, B\}, \{a, b, z\}, P, S)$
- $S \Rightarrow z\underline{A}Bz \Rightarrow za\underline{A}aBz \Rightarrow zaza\underline{B}z \Rightarrow$
  $\Rightarrow zazab\underline{B}bz \Rightarrow zazabb\underline{B}bbz \Rightarrow zazabbzbbz$

- der. strom:
  s vyzn. frázemi



Odstranění nedostupných a zbytečných symbolů

$X \in N \cup \Sigma$ je nedostupný $\overset{def}{\Leftrightarrow} \neg \exists \alpha, \beta \in (N \cup \Sigma)^*: S \overset{*}{\Rightarrow} \alpha X \beta$

$X \in N \cup \Sigma$ je zbytečný $\overset{def}{\Leftrightarrow} \neg \exists \alpha, \beta \in (N \cup \Sigma)^* \exists w \in \Sigma^*: S \overset{*}{\Rightarrow} \alpha X \beta \overset{*}{\Rightarrow} w$

- odstranění zbytečných symbolů

1. Zkonstruuje se mn-a N_t neterm. gen. term. řetězce

   $N_t = \{ A \in N \mid \exists w \in \Sigma^*: A \overset{*}{\Rightarrow} w \}$

   N_t se počítá iteračně:

   - $N_t^0 = \emptyset$, $i = 0$
   - $N_t^{i+1} = \{ A \in N \mid \exists (A \to \alpha) \in P: \alpha \in (\Sigma \cup N_t^i)^* \} \cup N_t^i$

   if $N_t^{i+1} \neq N_t^i$, i++

   if $N_t^{i+1} = N_t^i = N_t$

2. Odstraníme všechny neterminály, které nejsou v N_t (mimo S) + pravidla založená na těchto symbolech

3. Zkonstruuje se množina dostupných symbolů V:

   $V_0 = \{S\}$, $i = 0$

   $V_{i+1} = \{ X \in N \cup X \mid \exists (A \to \alpha X \beta) \in P: A \in V_i, \alpha, \beta \in (N \cup \Sigma)^* \} \cup V_i$

4. Odstraňte nedosažitelné symboly a pravidla na nich založená.

- Mějme gramatiku $G = (\{S, A, B\}, \{a, b\}, P, S)$, kde
  $P: S \to a \mid A, A \to AB, B \to b$
  Odstraňte zbytečné symboly (algoritmicky):
  1. $N_t^0 = \emptyset$
     $N_t^1 = \{S, B\}, N_t^2 = \{S, B\} = N_t^1 = N_t$
  2. $G' = (\{S, B\}, \{a, b\}, \{S \to a, B \to b\}, S)$
  3. $V_0 = \{S\}, V_1 = \{a, S\}, V_2 = \{a, S\} = V_1 = V$
  4. $G'' = (\{S\}, \{a\}, \{S \to a\}, S)$.

Odstranění ε-pravidel

$N_\varepsilon = \{A \in N \mid A \overset{*}{\Rightarrow} \varepsilon\}$

- zapište alg., který vypočte N_ε.

$\emptyset^* = \{\varepsilon\}$

Vstup: bezk. gr. $G = (N, \Sigma, P, S)$
Výstup: $N_\varepsilon = \{A \in N \mid A \overset{*}{\Rightarrow} \varepsilon\}$
Metoda: 1. $N_\varepsilon^0 = \emptyset, i = 0$
2. $N_\varepsilon^{i+1} = \{A \in N \mid \exists (A \to \alpha) \in P : \alpha \in (N_\varepsilon^i)^*\}$

3. je-li $N_\varepsilon^{i+1} = N_\varepsilon^i$, pak konec a $N_\varepsilon = N_\varepsilon^i$,
jinak i++ a předchod na bod 2.

- odstranění ε-pravidel:

$A \to \alpha_0 B_1 \alpha_1 B_2 \dots \alpha_n B_n \alpha_{n+1}$ kde $B_1, B_2, \dots, B_n \in N_\varepsilon$

nahradíme za 2^n pravidel tak, že uvažíme všechny kombinace $B_i \begin{cases} B_i \\ \varepsilon \end{cases}$

- Algoritmicky odstraňte ε pravidla z gramatiky s pravidly:

$S \to BC \mid d$

$B \to D \mid bCdS$

$C \to f \mid \varepsilon$

$D \to C \mid aB$

- $N_\varepsilon^0 = \emptyset$, $N_\varepsilon^1 = \{C\}$, $N_\varepsilon^2 = \{C, D\}$, $N_\varepsilon^3 = \{C, D, B\}$,
$N_\varepsilon^4 = \{C, D, B, S\} = N_\varepsilon^5 = N_\varepsilon$

- nová přepisovací pravidla

$S' \to S \mid \varepsilon$ (kde S' je nový st. symbol)

$S \to BC \mid C \mid B \mid \cancel{\varepsilon} \mid d$

$B \to D \mid \cancel{\varepsilon} \mid bCdS \mid bdS \mid bCd \mid bd$

$C \to f \mid \cancel{\varepsilon}$

$D \to C \mid \cancel{X} \mid aB \mid a$

Odstranění jednoduchých pravidel ($A \to B$, $A, B \in N$)

- $\forall A \in N$ spočítej $N_A = \{B \in N \mid A \overset{*}{\Rightarrow} B\}$
- následně $\forall B \to \alpha$, kde $\alpha \notin N$, najdeme všechna $A \in N$ taková, že $B \in N_A$, doplníme $A \to \alpha$ (následně vyhodíme jedn. prav.)
- Výpočet N_A:
  - $N_A^0 = \{A\}$, $i = 0$
  - $N_A^{i+1} = \{B \in N \mid \exists (C \to \alpha B \beta) \in P : C \in N_A^i, \alpha, \beta \in N_\varepsilon^*\} \cup N_A^i$

    $N_A^{i+1} \neq N_A^i$, $i{+}{+}$

    $N_A^{i+1} = N_A^i = N_A$
- Odstraníme jednoduchá pravidla z gramatiky, která vznikla v předchozím příkladě a má tedy pravidla:

  $S' \to S \mid \varepsilon$

  $S \to BC \mid B \mid C \mid d$

  $B \to D \mid bCdS \mid bdS \mid fCd \mid bd$

  $C \to f$

  $D \to C \mid aB \mid a$
- $N_\varepsilon^0 = \emptyset$, $N_\varepsilon^1 = \{S'\} = N_\varepsilon^2 = N_\varepsilon$

  $N_{S'}^0 = \{S'\}$, $N_{S'}^1 = \{S', S\}$, $N_{S'}^2 = \{S', S, B, C\}$, $N_{S'}^3 = \{S', S, B, C, D\} = N_{S'}^4 = N_{S'}$

$N_S^0 = \{S\}, N_S^1 = \{S, B, C\}, N_S^2 = \{S, B, C, D\} = N_S^3 = N_S$

$N_B^0 = \{B\}, N_B^1 = \{B, D\}, N_B^2 = \{B, D, C\} = N_B^3 = N_B$

$N_C^0 = \{C\} = N_C^1 = N_C$

$N_D^0 = \{D\}, N_D^1 = \{D, C\} = N_D^2 = N_D$

- $S' \to BC \mid d \mid bCdS \mid bdS \mid fCd \mid fd \mid f \mid aB \mid a \mid \varepsilon$

$S \to BC \mid d \mid bCdS \mid bdS \mid fCd \mid fd \mid f \mid aB \mid a$

$B \to bCdS \mid fdS \mid fCd \mid fd \mid f \mid aB \mid a$

$C \to f$

$D \to aB \mid a \mid f$

výsl. mn-a pravidel

- Zapište algoritmus testu, zda daná gramatika obsahuje cyklus.

vstup: $G = (N, \Sigma, P, S)$ bezk. gramatika

Výstup: ANO, pokud G obsahuje cyklus ($\exists A \in N: A \xRightarrow{+} A$)

NE jinak

Metoda: 1. Vypočteme N_ε.

2. Zavedeme relaci $\varrho \subseteq N \times N$ tak, že

$\forall A, B \in N: A \varrho B \overset{def}{\iff} \exists (A \to \alpha B \beta) \in P: \alpha, \beta \in N_\varepsilon^*$.

3. Vypočteme ϱ^+ např. Warshallovým alg.

4. ANO, je-li $A \in N$ takové, že $A \varrho^+ A$.
NE, jinak.

Odstranění levé rekurze ($A \overset{+}{\Rightarrow} A\alpha$)

- odstranění přímé levé rekurze

$$A \to A\alpha_1 | \dots | A\alpha_m | \beta_1 | \dots | \beta_n.$$

$$A \Rightarrow A\alpha_{i_1} \Rightarrow A\alpha_{i_2}\alpha_{i_1} \Rightarrow \dots \Rightarrow A\alpha_{i_n}\alpha_{i_{n-1}} \dots \alpha_{i_1} \Rightarrow \beta_j\alpha_{i_n} \dots \alpha_{i_1}$$

$$A \to \beta_1 | \dots | \beta_n | \beta_1 A' | \dots | \beta_n A'$$
$$A' \to \alpha_1 | \dots | \alpha_m | \alpha_1 A' | \dots | \alpha_m A'$$

- nepřímá levá rek.: převede se na přímou dosazováním pravidel do sebe

---

- Mějme gramatiku s pravidly $S \to S\overset{\alpha_1}{a} | \overset{\beta_1}{Ac} | \overset{\beta_2}{b}$, $A \to Sd | f$

Odstraňte levou rekurzi (algoritmicky)

- $S \to Ac | b | AcS' | bS'$
$S' \to a | aS'$
$A \to Sd | f$

} odstranění přímé levé rek. na symbolu S

- převod nepřímé levé rek. na přímou:

$S \to Ac | b | AcS' | bS'$

$S' \to a \mid aS'$

$A \to A\underbrace{cd}_{\alpha_1} \mid \underbrace{bd}_{\beta_1} \mid A\underbrace{cS'd}_{\alpha_2} \mid \underbrace{bS'd}_{\beta_2} \mid \underbrace{f}_{\beta_3}$

- odstranění přímé levé rekurze na symbolu A

$S \to Ac \mid b \mid AcS' \mid bS'$, $S' \to a \mid aS'$

$A \to bd \mid bS'd \mid f \mid bdA' \mid bS'dA' \mid fA'$

$A' \to cd \mid cS'd \mid cdA' \mid cS'dA'$ výsledná pravidla

- Chomského NF: $A \to BC \mid a$ (příp. $S \to \varepsilon$)
- Greibachové NF: $A \to a\alpha$, $\alpha \in N^*$
- Mějme gramatiku s pravidly

$A \to BcC \mid dDaDaD$

$B \to aB \mid Cd \mid f$

$C \to bB \mid a$

$D \to dD \mid c$

Převeďte do GNF (algoritmicky).

---

1. Odstranění ε-pravidel a levé rekurze
   - v tomto případě nedojde ke změně
2. Sestavíme relaci $< \subseteq N \times N$ takovou, že

$X < Y \overset{def}{\Leftrightarrow} \exists (X \to Y\beta) \in P$

- V našem případě: $A < B$ (z $A \to BcC$), $B < C$ (z $B \to Cd$)

- Relaci $<$ zúplníme lib. způsobem, např. $A < B < C < D$

3. Dosadíme pravidla do sebe od nejv. symbolu, pokud na levé straně pravé str. přep. pr. je neterm.:

$A \to BcC \mid dDaDaD$
$B \to aB \mid bBd \mid ad \mid f$
$C \to bB \mid a$
$D \to dD \mid c$

(dosazení C pravidel)

$A \to aBcC \mid bBdcC \mid adcC \mid fcC \mid dDaDaD$
$B \to aB \mid bBd \mid ad \mid f$
$C \to bB \mid a$
$D \to dD \mid c$

(dosazení B pravidel)

4. Odstraníme terminály, které nejsou na levé pozici pravé strany:

$A \to aBC'C \mid bBD'C'C \mid aD'C'C \mid fC'C \mid dDA'DA'D$
$C' \to c \quad D' \to d \quad A' \to a$
$B \to aB \mid bBD' \mid aD' \mid f$

$C \to bB \mid a$

$D \to dD \mid c$

- Mějme $G = (\{S\}, \{a,b\}, \{S \to aSb \mid \varepsilon\}, S)$ a $L = \{a^n b^n \mid n \geq 0\}$. Dokažte, že $L(G) = L$.

Důkaz:

a) $L(G) \subseteq L$: Indukcí k délce i slova. Vzhledem k tomu, že pravidla $S \to aSb \mid \varepsilon$ mohou generovat slova sudé délky, stačí indukce přes sudá čísla.

- bázový případ: $i = 0$
  Jediné slovo délky 0 je ε, ale víme, že $\varepsilon \in L$, protože $\varepsilon = a^0 b^0 \in L$ (pro $n = 0$)
- indukční krok:
  - Předpokládejme, že tvrzení platí pro libovolné slovo $w \in L(G)$ sudé délky menší než i, $i \geq 0$.
  - Ukážeme, že tvrzení platí i pro slova délky $i+2$
  - Slovo délky $i+2$ obsahuje alespoň dva symboly a tedy vypadá derivace $S \Rightarrow aSb \overset{*}{\Rightarrow} awb$, kde $|w| = i$
  - Na základě ind. předpokladu víme, že $w \in L$ (víme totiž že $S \overset{*}{\Rightarrow} w$ a $|w| = i \leq i$).

- z $w \in L$ víme, že $w = a^n b^n$ pro $n \geq 0$.
Ovšem $a a^n b^n b = a^{n+1} b^{n+1} \in L$.

b) $L \subseteq L(G)$ Indukcí na délkou i slova z L, přičemž z def. pravidly L plyne, že řetězce $\in L$ mají délku sudou (značně $L = \{a^n b^n \mid n \geq 0\}$, tedy $|w| = 2n$, $n \geq 0$, pro libovolné $w \in L$) a tedy postačí indukce přes sudá čísla.

- bázový případ ($i = 0$): Platí, protože jediný řetězec délky 0 je ε a $(S \to \varepsilon) \in P$ a tedy $\varepsilon \in L(G)$.
- indukční krok:
  - Předpokládejme, že tvrzení platí pro slova $w \in L$ taková, že $|w| \leq i$ pro nějaké sudé $i \geq 0$.
  - Ukážeme, že pak platí i pro $w \in L$, kde $|w| = i + 2$.
  - Řetězec $w \in L$ takový, že $|w| = i + 2$, má podobu $a^n b^n$ pro nějaké $n \geq 0$.
  - z toho, že $i + 2 \geq 2$, víme, že $n \geq 1$, tedy lze w rozepsat jako $a a^{n-1} b^{n-1} b$

- Ovšem $|a^{n-1}b^{n-1}| = i$ a tedy na $a^{n-1}b^{n-1}$ se vztahuje ind. předpoklad a tedy $a^{n-1}b^{n-1} \in L(G)$, (neboli $S \overset{*}{\Rightarrow} a^{n-1}b^{n-1}$)
- Pak ale můžu zkonstruovat derivaci:

$$S \Rightarrow aSb \overset{*}{\Rightarrow} aa^{n-1}b^{n-1}b = a^n b^n$$

a tedy $a^n b^n \in L(G)$. $\square$