Abschlusseigenschaften (kfS)

Substitution von Sprachen: → Verallgemeinerung von Homomorphismen
- Abbildung σ von Wörtern in Sprachen

$\sigma: \Sigma^* \to \mathcal{L}$ ist Substitution, wenn $\sigma(v_1 \ldots v_n) = \sigma(v_1) \circ \ldots \circ \sigma(v_n)$ für alle $v_i \in \Sigma$

$\sigma(L) = \bigcup\{\sigma(w) \mid w \in L\}$ ist das Abbild der Wörter von L unter σ

- Sehr ausdrucksstark:

- $L_1 \cup L_2 = \sigma(\{1,2\})$ für $\sigma(1) = L_1$ und $\sigma(2) = L_2$
- $L_1 \circ L_2 = \sigma(\{12\})$ für $\sigma(1) = L_1$ und $\sigma(2) = L_2$
- $L^* = \sigma(\{1\}^*)$ für $\sigma(1) = L$

Verwende Abgeschlossenheit unter Substitutionen

- **L_1, L_2 kontextfrei $\Rightarrow$ $L_1 \cup L_2$ kontextfrei**
  - Sei $\sigma(1)=L_1$ und $\sigma(2)=L_2$
  - Dann ist $\sigma:\{1,2\}\to\mathcal{L}_2$ Substitution und $L_1 \cup L_2 = \sigma(\{1,2\}) \in \mathcal{L}_2$
- **L_1, L_2 kontextfrei $\Rightarrow$ $L_1 \circ L_2$ kontextfrei**
  - Sei $\sigma(1)=L_1$ und $\sigma(2)=L_2$
  - Dann ist $\sigma:\{1,2\}\to\mathcal{L}_2$ Substitution und $L_1 \circ L_2 = \sigma(\{12\}) \in \mathcal{L}_2$
- **L kontextfrei $\Rightarrow$ L^* kontextfrei**
  - Für $\sigma(1)=L$ ist $\sigma:\{1\}\to\mathcal{L}_2$ Substitution und $L^* = \sigma(\{1\}^*) \in \mathcal{L}_2$
- **L kontextfrei $\Rightarrow$ L^+ kontextfrei**
  - Für $\sigma(1)=L$ ist $\sigma:\{1\}\to\mathcal{L}_2$ Substitution und $L^+ = \sigma(\{1\}^+) \in \mathcal{L}_2$
- **$L \in \mathcal{L}_2$, h Homomorphismus $\Rightarrow$ $h(L)$ kontextfrei**
  - Für $\sigma(a)=\{h(a)\}$ ist $\sigma:T\to\mathcal{L}_2$ Substitution und $h(L) = \sigma(L) \in \mathcal{L}_2$

Normalformen:

↳ um leichte Analyse zu realisieren, Systematische Form ⇒ Chomsky-Normalform:

Für Grammatik $G = (V, T, P, S)$ gilt entweder $A \to BC$ oder $A \to a$
$(A, B, C \in V, a \in T)$

↳ kontextsensitiv

- **Jede kontextfreie Grammatik G mit $\epsilon \notin L(G)$ ist in Chomsky-Normalform transformierbar**
  1. Eliminierung von ϵ-Produktionen $A \to \epsilon$
  2. Eliminierung von Einheitsproduktionen $A \to B$
  3. Eliminierung unnützer Symbole
  4. Separieren von Terminalsymbolen und Variablen in Produktionen
  5. Aufspalten von Produktionen $A \to \alpha$ mit $|\alpha|>2$

Aufblähung/Transformationszeit meist quadratisch relativ zur Größe von G

Grammatiken mit $\epsilon \in L(G)$ benötigen Sonderregelung (Folie 18)

1. ε - Produktion eliminieren: $\exists A \to \varepsilon$

$P = \{ S \to ABa,\ A \to aAA \mid C,\ B \to bBB \mid \varepsilon,\ C \to \varepsilon \}$

$S \to ABa, Aa, Ba, a$
$A \to aAA, aA, C, \varepsilon$
$B \to bBB, bB, b, \varepsilon$
$C \to \varepsilon$

$\to$

$S \to ABa, Aa, Ba, a$
$A \to aAA, aA, C,$
$B \to bBB, bB, b$

2. Einheitsproduktionen eliminieren

(Einheitspaar (A,B) $A \xrightarrow{*} B$) Ist (A,B) und (B,C) dann auch (A,C) Einheitspaar)

$$P' = \{\ E \to T \mid E{+}T,\ T \to F \mid T{*}F,\ F \to I \mid (E)\ \ I \to a \mid b \mid c \mid Ia \mid Ib \mid Ic \mid I0 \mid I1\ \}$$

- **Bestimme alle Einheitspaare (A,B) mit $A \xRightarrow{*} B$**
  1.: (E,E), (T,T), (F,F) und (I,I) sind Einheitspaare
  2.: (E,T), (T,F) und (F,I) sind ebenfalls Einheitspaare
  3.: (E,F) und (T,I) sind ebenfalls Einheitspaare
  4.: (E,I) ist ebenfalls Einheitspaar
  5.: Keine weiteren Einheitspaare möglich
- **Erzeuge Grammatik ohne Einheitsproduktionen**
  – Einheitspaare mit E: $\{E \to E{+}T \mid T{*}F \mid (E) \mid a \mid b \mid c \mid Ia \mid Ib \mid Ic \mid I0 \mid I1\}$
  – Einheitspaare mit T: $\{T \to T{*}F \mid (E) \mid a \mid b \mid c \mid Ia \mid Ib \mid Ic \mid I0 \mid I1\}$
  – Einheitspaare mit F: $\{F \to (E) \mid a \mid b \mid c \mid Ia \mid Ib \mid Ic \mid I0 \mid I1\}$
  – Einheitspaare mit I: $\{I \to a \mid b \mid c \mid Ia \mid Ib \mid Ic \mid I0 \mid I1\}$

3. Unnütze Symbole entfernen:

- X nützlich falls $S \xrightarrow{*} \alpha X \beta \xrightarrow{*} w \in T^*$
  ↳ Erzeugend ($X \xrightarrow{*} v \in T^*$) und erreichbar ($S \xrightarrow{*} \alpha X \beta$)

- **Beispiel: $P = \{\ S \to AB \mid a,\ A \to b\ \}$**
  · Erreichbar: S, A, B, a, und b erzeugend: S, A, a, und b
  – Nach Elimination von B: $\{\ S \to a,\ A \to b\ \}$
  · Erreichbar: S und a erzeugend: S, A, a, und b
  – Nach Elimination von A: $\{\ S \to a\ \}$
  · Erreichbar: S und a erzeugend: S und a

  Erzeugte Produktionenmenge ist äquivalent zu P

4. Berechnung Erzeugender / Erreichbarer Symbole

- **Beispiel: $P = \{\ S \to AB \mid a,\ A \to b\ \}$**
  – Erzeugende Symbole: 1.: a und b sind erzeugend
  2.: S und A sind ebenfalls erzeugend
  3.: Keine weiteren Symbole sind erzeugend
  – Erreichbare Symbole: 1.: S ist erreichbar
  2.: A, B und a sind ebenfalls erreichbar
  3.: b ist ebenfalls erreichbar

5. Erzeugung der Chomsky-Normalform:

- **Separiere Terminalsymbole von Variablen**
  - Für jedes Terminalsymbol $a \in T$ erzeuge neue Variable X_a
  - Ersetze Produktionen $A \to \alpha$ mit $|\alpha| \geq 2$ durch $A \to \alpha_X$ (a $\in T$ ersetzt durch X_a)
  - Ergänze Produktionen $X_a \to a$ für alle $a \in T$
- **Spalte Produktionen $A \to \alpha$ mit $|\alpha| > 2$**
  - Ersetze jede Produktion $A \to X_1..X_k$ durch $k-1$ Produktionen
    $A \to X_1 Y_1,\ Y_1 \to X_2 Y_2,\ \ldots Y_{k-2} \to X_{k-1} X_k$, wobei alle Y_i neue Variablen

$$\begin{aligned}
\boldsymbol{P} = \{ & \boldsymbol{E \to E+T \mid T*F \mid (E) \mid a \mid b \mid c \mid Ia \mid Ib \mid Ic \mid I0 \mid I1} \\
& \boldsymbol{T \to T*F \mid (E) \mid a \mid b \mid c \mid Ia \mid Ib \mid Ic \mid I0 \mid I1} \\
& \boldsymbol{F \to (E) \mid a \mid b \mid c \mid Ia \mid Ib \mid Ic \mid I0 \mid I1} \\
& \boldsymbol{I \to a \mid b \mid c \mid Ia \mid Ib \mid Ic \mid I0 \mid I1}\ \}
\end{aligned}$$

- **Separiere Terminalsymbole von Variablen**

$$\begin{aligned}
P' = \{ & E \to EX_+T \mid TX_*F \mid X_(EX_) \mid a \mid b \mid c \mid IX_a \mid IX_b \mid IX_c \mid IX_0 \mid IX_1 \\
& T \to TX_*F \mid X_(EX_) \mid a \mid b \mid c \mid IX_a \mid IX_b \mid IX_c \mid IX_0 \mid IX_1 \\
& F \to X_(EX_) \mid a \mid b \mid c \mid IX_a \mid IX_b \mid IX_c \mid IX_0 \mid IX_1 \\
& I \to a \mid b \mid c \mid IX_a \mid IX_b \mid IX_c \mid IX_0 \mid IX_1 \\
& X_a \to a,\ X_b \to b,\ X_c \to c,\ X_0 \to 0,\ X_1 \to 1,\ X_+ \to +,\ X_* \to *,\ X_(\to (,\ X_) \to)\ \}
\end{aligned}$$

- **Spalte Produktionen $A \to \alpha$ mit $|\alpha| > 2$**

$$\begin{aligned}
P' = \{ & E \to EY_1 \mid TY_2 \mid X_(Y_3 \mid a \mid b \mid c \mid IX_a \mid IX_b \mid IX_c \mid IX_0 \mid IX_1 \\
& Y_1 \to X_+T,\ Y_2 \to X_*F,\ Y_3 \to EX_) \\
& T \to TY_2 \mid X_(Y_3 \mid a \mid b \mid c \mid IX_a \mid IX_b \mid IX_c \mid IX_0 \mid IX_1 \\
& F \to X_(Y_3 \mid a \mid b \mid c \mid IX_a \mid IX_b \mid IX_c \mid IX_0 \mid IX_1 \\
& I \to a \mid b \mid c \mid IX_a \mid IX_b \mid IX_c \mid IX_0 \mid IX_1 \\
& X_a \to a,\ X_b \to b,\ X_c \to c,\ X_0 \to 0,\ X_1 \to 1,\ X_+ \to +,\ X_* \to *,\ X_(\to (,\ X_) \to)\}
\end{aligned}$$