Def **Kontextfreie** Grammatik (CFG)

$$G = (V, \Sigma, P, S)$$

Startsymbol
auch: Axiom

- Nichtterminale
  auch: Variablen
- endliche Menge

- Alphabet
  auch: Terminale, Parameter
- endliche Menge
- $\varepsilon \notin \Sigma$, $| \notin \Sigma$

- Produktionen
  auch Ersetzungsregeln
- $P \subseteq V \times (V \cup \Sigma)^*$
- endliche Menge
- $X \to \gamma$ statt $(X, \gamma) \in P$
- $X \to \gamma_1 | \gamma_2 | \dots | \gamma_n$ statt $(X, \gamma_1), \dots, (X, \gamma_n) \in P$

$$V \cap \Sigma = \emptyset$$

Häufig nur Produktionen und Startsymbol explizit:

Bsp:

$$E \to T \mid E+T$$

$$T \to F \mid T \circ F$$

$$F \to a \mid (E)$$

mit E Startsymbol

$\rightsquigarrow$ vollständig: $(\{E,T,F\}, \{a, (,), +, \circ\},$
$\{(E,T), (E, E+T), \dots\},$
$E)$

Def: Ableitung bzgl. CFG $G = (V, \Sigma, P, S)$

- $\gamma \xrightarrow{0}_G \gamma$ für alle $\gamma \in (V \cup \Sigma)^*$
- $\gamma \xrightarrow{1}_G \gamma'$ für alle $\gamma, \gamma' \in (V \cup \Sigma^*)$ mit
  1. $\gamma = \alpha X \omega$ mit $X \in V$
  2. $\gamma' = \alpha \delta \omega$
  3. $(X, \delta) \in P$
- Induktiv: $\gamma \xrightarrow{n+1}_G \gamma'$ für alle $\gamma, \gamma' \in (V \cup \Sigma)^*$ falls es ein $\gamma'' \in (V \cup \Sigma)^*$ mit $\gamma \xrightarrow{1}_G \gamma'' \xrightarrow{n}_G \gamma'$ gibt.
- Damit: $\gamma \xrightarrow{*}_G \gamma'$ falls $\exists n \in \mathbb{N}_0 : \gamma \xrightarrow{n}_G \gamma'$

  $\gamma \xrightarrow{+}_G \gamma'$ falls $\exists n \in \mathbb{N} : \gamma \xrightarrow{n}_G \gamma'$

Bsp: G: $E \rightarrow T \mid E+F$

$T \rightarrow F \mid T \circ F$

$F \rightarrow a \mid (E)$

- $a+T+a \xrightarrow{1} a+T\circ F+a \xrightarrow{1} a+T\circ a+a$

Def: CFG $G = (V, \Sigma, P, S)$

- $L_X(G) := \{ w \in \Sigma^* \mid X \to_G^* w \}$
- $L(G) := L_S(G)$

$\rightsquigarrow$ $L(G)$ von G erzeugte Sprache

- $L \subseteq \Sigma^*$ ist kontextfrei falls es eine CFG G mit $L = L(G)$ gibt.

- CFL $\hat{=}$ context-free language

$$G: \quad E \longrightarrow T \mid E + F$$
$$T \longrightarrow F \mid T \cdot F$$
$$F \longrightarrow a \mid (E)$$

$\leadsto$ $a \in L(G)$, $(a) \in L(G)$, $a + a \in L(G)$, ...

$\boxed{\text{Def}}$: $S \xrightarrow{1}_G \gamma_1 \xrightarrow{1}_G \gamma_2 \rightarrow \dots \xrightarrow{1}_G w \in \Sigma^*$

Linksableitung: ersetze jeweils das linkeste NT

Rechtsableitung: ——— " — rechteste NT

Bsp.: $E \xrightarrow{1} E + T \xrightarrow{1} T + T \xrightarrow{1} F + T \xrightarrow{1} a + T \xrightarrow{1} a + F \xrightarrow{1} a + a$

$E \xrightarrow{1} E + T \xrightarrow{1} E + F \xrightarrow{1} E + a \xrightarrow{1} T + a \xrightarrow{1}_a F + a \xrightarrow{1} a + a$

<u>Bsp.:</u> $L = \{a^n b^n \mid n \in \mathbb{N}_0\}$ CFL.

wird erzeugt von

$$S \to aSb \mid \varepsilon$$

<u>Bsp.:</u> $L = \{ w \in \{a,b\}^* \mid w = w^R \}$

wird erzeugt von

$$S \to aSa \mid bSb \mid a \mid b \mid \varepsilon$$

$\boxed{\underline{\text{Lem}}:}$ $\quad \alpha_1 \alpha_2 \xrightarrow[G]{n} \beta$

$$\Longleftrightarrow$$
gdw.

$\exists \beta_1, \beta_2, n_1, n_2 : n = n_1 + n_2 \wedge \beta = \beta_1 \beta_2 \wedge \alpha_1 \xrightarrow[G]{n_1} \beta_1$
$\wedge\, \alpha_2 \xrightarrow[G]{n_2} \beta_2$

<u>Bsp.</u>: $\quad E + T \xrightarrow[G]{5} a + a$

mit $E+ \xrightarrow[G]{1} T+ \xrightarrow[G]{1} F+ \xrightarrow[G]{1} a+$

und $T \xrightarrow[G]{1} F \xrightarrow[G]{1} a$

$\leadsto$ also $\beta_1 = a = \beta_2$, $n_1 = 3$, $n_2 = 2$

<u>Beweis</u> mittels Induktion nach n. (Übung)

$\boxed{\underline{\text{Def.}}}$: CFG $G = (V, \Sigma, P, S)$ ist

linear falls $\gamma \in (\Sigma \cup \{\varepsilon\})(V \cup \{\varepsilon\})(\Sigma \cup \{\varepsilon\})$

linkslinear falls $\gamma \in (V \cup \{\varepsilon\})(\Sigma \cup \{\varepsilon\})$ $\quad$ $\underline{\forall (X, \gamma) \in P}$

rechtslinear falls $\gamma \in (\Sigma \cup \{\varepsilon\})(V \cup \{\varepsilon\})$

Bsp.:

$S \to aSa \mid bSb \mid a \mid b \mid \varepsilon$

$S \to aSb \mid \varepsilon$ $\quad$ linear

$S \to aS \mid Sb \mid \varepsilon$

$S \to aS \mid a$ $\quad$ rechtslinear

$S \to \varepsilon \mid Sb$ $\quad$ linkslinear

$\boxed{\text{Lem:}}$ L regulär $\iff L = L(G)$ mit G rechtslinear

$\iff L = L(G)$ mit G linkslinear

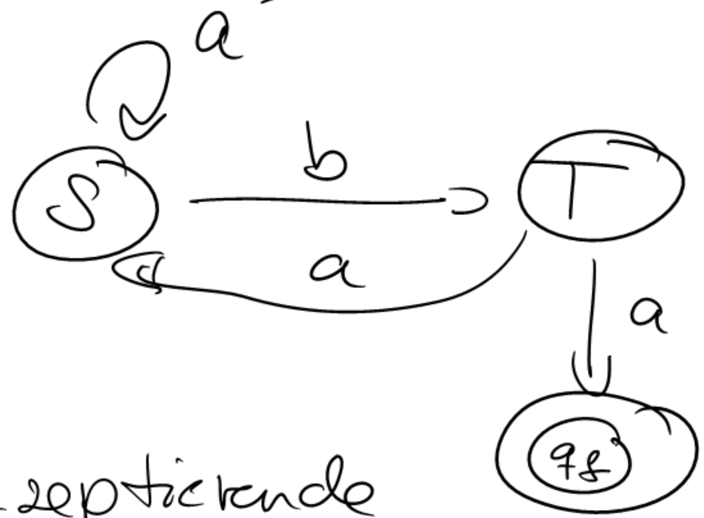
Beweisidee:

rechtslinear

$$S \to aS \mid bT$$
$$T \to a \mid aS$$

$\Longleftrightarrow$

NFA



Ableitung: $\longmapsto$ / $\longleftarrow$ akzeptierende Berechnung

Variablen $\longmapsto$ Zustände + extra q_f

Produktionen $\longmapsto$ Transitionen

Alternative Leseweise von Produktionen

lies $X \to u_1 X_1 u_2 \cdots u_k X_k u_{k+1}$

$(u_1, \dots, u_k \in \Sigma^*, \; X, X_1, \dots, X_k \in V)$

als: Falls $w_1 \in L_{X_1}(G), \dots, w_k \in L_{X_k}(G)$

dann auch $u_1 w_1 u_2 \cdots u_k w_k u_{k+1} \in L_X(G)$

bzw: $u_1 L_{X_1}(G) u_2 \cdots u_k L_{X_k}(G) u_{k+1} \subseteq L_X(G)$

Induktive Definition von L(G)

Bsp.: G: $S \to \varepsilon \mid [S] \mid SS$

$\leadsto$

- Stets $\varepsilon \in L(G)$
- Falls $u \in L(G)$, dann auch $[u] \in L(G)$
- Falls $u, v \in L(G)$, dann auch $uv \in L(G)$

bzw: $\{\varepsilon\} \cup \{[\} L(G) \{]\} \cup L(G)L(G) \subseteq L(G)$

Bsp.: $G: S \to \varepsilon \mid [S] \mid SS$

$\leadsto$
- Stets $\varepsilon \in L(G)$
- Falls $u \in L(G)$, dann auch $[u] \in L(G)$
- Falls $u, v \in L(G)$, dann auch $uv \in L(G)$

Damit: $\varepsilon \in L(G)$

$\underset{\text{impl.}}{\Longrightarrow}$ $[\,] \in L(G)$

$\Longrightarrow$ $[[\,]] \in L(G)$

$\Longrightarrow$ $[[\,]][\,] \in L(G)$

- bottom-up Konstruktion
- immer nur Wörter als Zwischenergebnisse

<u>vs</u> "top-down" Ableitung

$S \to SS \to [S]S \to [[S]]S \to [[\,]]S$

$\to [[\,]][S] \to [[\,]][\,]$

$\rightsquigarrow$ Induktive Definition von $L(G)$
hilfreich, um Eigenschaften von $L(G)$ zu beweisen.

Bsp: G: $S \to \varepsilon \mid [S] \mid SS$

Beh: $\forall w \in L(G)$: $|w|_{[} = |w|_{]}$

Anzahl der Vorkommen von [bzw.] in w

$$|[\,]\,[|_{[} = 2$$

$$|[\,]\,[|_{]} = 1$$

$\rightsquigarrow$ Beweis mittels Induktion nach Produktionsregeln

Bsp: $G: \quad S \to \varepsilon \mid [S] \mid SS$

Beh: $\forall w \in L(G): |w|_{[} = |w|_{]}$

Bew:

- Fall "$S \to \varepsilon$":
  Dann: $w = \varepsilon$ und damit $|\varepsilon|_{[} = 0 = |\varepsilon|_{]}$

- Fall "$S \to [S]$":
  Ann: $w \in L(G)$ beliebig mit $|w|_{[} = |w|_{]}$
  Dann $|[w]|_{[} = 1 + |w|_{[} \overset{\text{Ann}}{=} 1 + |w|_{]} = |[w]|_{]}$

- Fall "$S \to SS$"
  Ann: $w', w \in L(G)$ beliebig mit $|w|_{[} = |w|_{]}$ und $|w'|_{[} = |w'|_{]}$
  Dann: $|ww'|_{[} = |w|_{[} + |w'|_{[} \overset{\text{Ann.}}{=} |w|_{]} + |w'|_{]} = |ww'|_{]}$

$\rightsquigarrow$ Allgemeines Induktionsprinzip:

Um $\forall X \in V \; \forall w \in L_X(G): \; Q_X(w)$

$\rightsquigarrow$ zu zeigen, zeige für jede Regel

$$X \to u_1 X_1 \dots u_k X_k u_{k+1} \in P$$

<u>Falls</u> $w_1 \in L_{X_1}(G) \wedge \dots \wedge w_k \in L_{X_k}(G)$

$\wedge \; Q_{X_1}(w_1) \wedge \dots \wedge Q_{X_k}(w_k)$

<u>dann</u> $Q_X(u_1 w_1 \dots u_k w_k u_{k+1})$

<u>Bsp.:</u> $G: \quad S \to \varepsilon \mid [S] \mid SS$

<u>Beh.:</u> $\forall w \in L(G): \quad |w|_{[} = |w|_{]}$

$\wedge \; \forall u \preceq w: |u|_{[} \geq |u|_{]}$ } "w ist balanciert"

"u Präfix von w"

<u>Bsp.:</u> $[\,]$, $[[\,]][\,]$, ε balanciert

$[$, $][$, $[\,]][$ nicht balanciert

$\leadsto$ müssen nur noch zeigen:

$\forall w \in L(G): \forall u \preceq w: |u|_{[} \geq |u|_{]}$

• **Beweis:** $G: S \to \varepsilon \mid [S] \mid SS$

Fall "$S \to \varepsilon$":

- keine Variablen, damit keine Annahmen
- direkt beweisen

Sei $w = \varepsilon$, dann ist ε der einzige Präfix und es gilt $|\varepsilon|_[= 0 \geq 0 = |\varepsilon|_]$ ✓

• Fall "$S \to [S]$"

Annahme: $w \in L(G)$ (beliebig) und $\forall u \preceq w: \underline{|u|_[\geq |u|_]}$ (*)

Präfixe von $[w]$

$\rightsquigarrow |[|_[= 1 \geq 0 = |[|_]$

$\rightsquigarrow |[u|_[= 1 + |u|_[\overset{(*)}{\geq} 1 + |u|_] > |u|_] = |[u|_]$

$\rightsquigarrow |[w|_[= 1 + |w|_[= 1 + |w|_] > |w|_] = |[w|_]$

$\rightsquigarrow$ ...

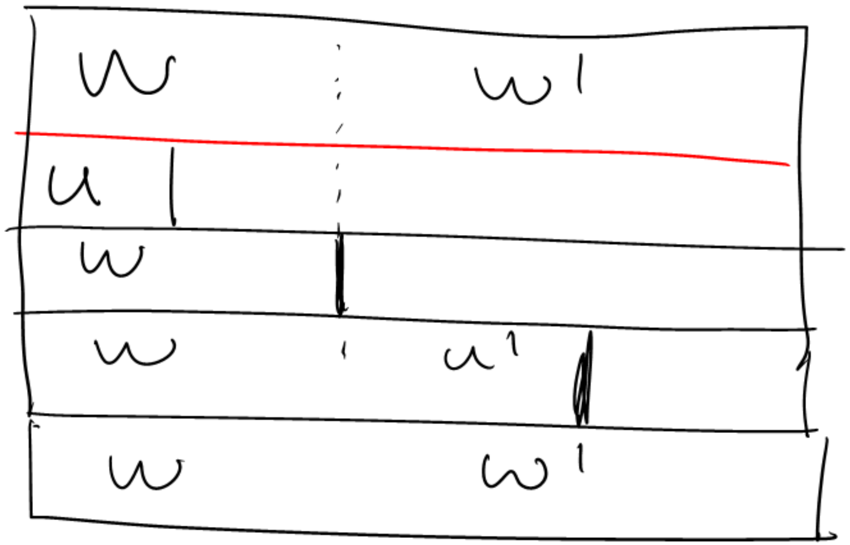
Fall "$S \to SS$":

Annahme $w \in L(G) \wedge w' \in L(G)$

$\wedge \forall u \preceq w : |u|_{[} \geq |u|_{]}$ $*$

$\wedge \forall u' \preceq w' : |u'|_{[} \geq |u'|_{]}$ $*$

Mögliche Präfixe von ww'



$*$

$|w|_{[} = |w|_{]}$

$*$ A $|w|_{[} = |w|_{]}$

$|w|_{[} = |w|_{]} \wedge |w'|_{[} = |w'|_{]}$

Bsp.:

Beh: $w \in \{[,]\}^*$ balanciert $\Longrightarrow w \in L(G)$

$\rightsquigarrow$ Induktion nach Regeln nicht möglich, da $w \in L(G)$ ja zu zeigen

$\rightsquigarrow$ Induktion nach $n = |w|$ hier

$n=0$: also $w = \varepsilon$. Mit $S \rightarrow \varepsilon \in P$ folgt $\varepsilon \in L(G)$ ✓

$n \rightarrow n+1$ ...

$G: S \rightarrow \varepsilon \mid [S] \mid SS$

$n \to n+1$:

Annahme: Für alle balancierten Wörter w mit $|w| \leq n$

gilt $w \in L(G)$

- Sei w balanciert mit $|w| = n+1$, sonst beliebig fixiert.

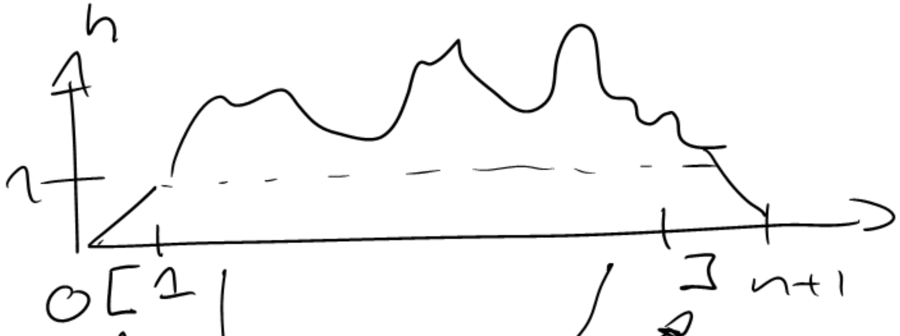
$\Rightarrow w = a_1 a_2 \cdots a_{n+1}$ mit $a_i \in \{[,]\}$

- Setze $h(k) = |a_1 \cdots a_k|_[- |a_1 \cdots a_k|_]$ ("Höhe" bei k)

für $k \in \{0, 1, \ldots, n+1\}$

$\leadsto$
- $h(0) = |\varepsilon|_[- |\varepsilon|_] = 0$
- $h(n+1) = |w|_[- |w|_] = 0$, da w balanciert nach Voraussetzung
- $|h(k+1) - h(k)| = 1$ ⨍
- $h(k) \in \mathbb{N}_0$, da w balanciert

1. Fall h hat nur die Nullstellen $0, n+1$

$\rightsquigarrow$ 

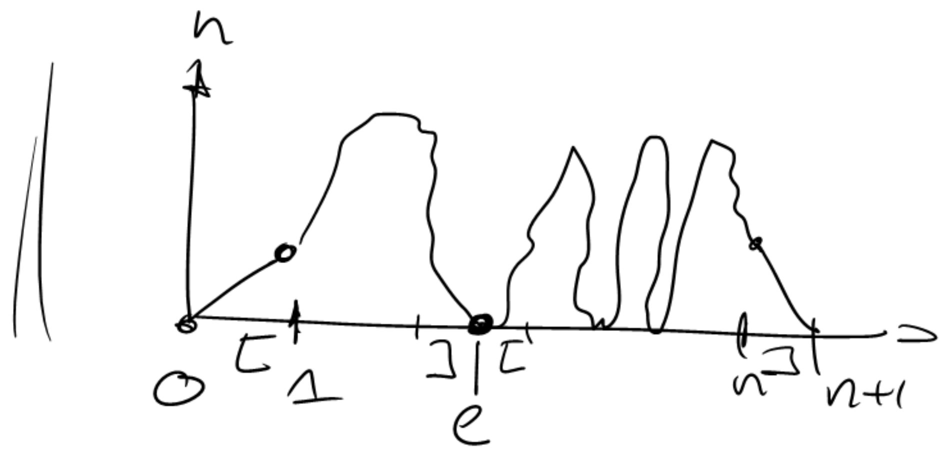
Warum?

$\rightsquigarrow$ $a_2 \dots a_n$ balanciert

$\rightsquigarrow$ $a_2 \dots, a_n \in L(G)$ nach Induktion

$\rightsquigarrow$ $w = [a_2 \dots a_n] \in L(G)$ mittels $S \to [S] \in P$

<u>2. Fall</u> h hat noch mindestens eine Nullstelle in $\{1, \dots, n\}$

$\leadsto$ 

$\leadsto$ $\underbrace{a_1 \dots a_e}_{|\cdot| \le n}$ und $\underbrace{a_{e+1} \dots a_{n+1}}_{|\cdot| \le n}$ balanciert

$\leadsto$ $a_1 \dots a_e \in L(G) \wedge a_{e+1} \dots a_{n+1} \in L(G)$
nach Induktion

$\leadsto$ $a_1 \dots a_e \, a_{e+1} \dots a_{n+1} \in L(G)$ mittels $S \to SS \in P$

$\square$