Formale Grundlagen der Informatik

12

Eigenschaften kontextfreier Sprachen

Chomsky-Normalform

Pumping-Lemma

Hierarchiebeziehungen

- Die folgenden drei Aussagen sind äquivalent:
  1. L ist eine kontextfreie Sprache.
  2. $L = L(M_1)$ für einen PDA M_1.
  3. $L = N(M_2)$ für einen PDA M_2.

  Sprachfamilie $\mathcal{L}(\mathrm{CF})$
- Jeder NEA ist ein spezieller PDA (*der seinen Keller ignoriert*).
  - ➢ $\mathcal{L}(\mathrm{REG}) \subseteq \mathcal{L}(\mathrm{CF})$
- $\{\ a^n b^n \mid n \geq 0\ \} \in \mathcal{L}(\mathrm{CF}) \setminus \mathcal{L}(\mathrm{REG})$
- Das Wortproblem für kontextfreie Grammatiken ist entscheidbar.

➢ $\mathcal{L}(\mathrm{REG}) \subset \mathcal{L}(\mathrm{CF}) \subseteq \mathcal{L}(\mathrm{REC})$

Motivation

- Gibt es entscheidbare Sprachen, die nicht kontextfrei sind?
- *Kandidat*: Sprache de „Nicht-Quadrate“:

$$L_{nq} = \{\, uxvu'yv' \in \{a,b\}^* \mid x,y \in \{a,b\}, x \neq y, |u| = |u'|, |v| = |v'| \,\}$$

- *Behauptung*: $L_{nq} \notin \mathcal{L}(\mathrm{CF})$
- *Begründung*:
  - ein PDA müsste sich die Längen von u und v im Keller merken …
  - … und in der gleichen Reihenfolge aus dem Keller lesen
  - ➢ nicht möglich
- ***Stimmt das?***

Vereinfachung kontextfreier Grammatiken

- Um wichtige Eigenschaften nachzuweisen:
  möglichst einfache Form kontextfreier Grammatiken!
- **Chomsky-Normalform**
  - benannt nach Noam Chomsky
  - linke Regelseite ist schon einfach: nichtterminales Symbol
  - rechte Regelseite: entweder ein Terminal oder genau zwei Nichtterminale
- **Definition:** Eine kfG $G = (N, T, P, S)$ ist in **Chomsky-Normalform (CNF)**, falls
  $P \subseteq (N \times NN) \cup (N \times T)$ gilt,
  also jede Regel die Form $A \to BC$ oder $A \to a$ hat, $A, B, C \in N$, $a \in T$.
- Jede ε-freie kontextfreie Sprache wird von einer kfG in CNF erzeugt.

ε-Regeln

Lemma 12.1. Zu jeder kfG $G = (N, T, P, S)$ kann eine kfG $G' = (N, T, P', S)$ mit $L(G') = L(G) \setminus \{\varepsilon\}$ konstruiert werden, sodass P' **keine ε-Regeln** (Regeln der Form $\mathrm{A} \rightarrow \varepsilon$) enthält.

Konstruktion:

1. Bestimmen die Menge $\boldsymbol{N_\varepsilon} = \{\, A \in N \mid A \overset{*}{\Longrightarrow} \varepsilon \,\}$ der **eliminierbaren Symbole** iterativ wie folgt:
   1. Füge alle $A \in N$ zu N_ε hinzu, für die $A \rightarrow \varepsilon \in P$.
   2. Wiederhole solange, bis keine neuen Symbole zu N_ε hinzugefügt werden: Füge alle $A \in N$ zu N_ε hinzu, für die $A \rightarrow v \in P$ mit $v \in N_\varepsilon^*$ existiert.

   ➢ Diese Iteration ist endlich, da N endlich ist.

1. Schritt – Beispiel

$S \rightarrow ABC \mid AB$
$A \rightarrow b \mid CD$
$B \rightarrow Bb \mid \varepsilon$
$C \rightarrow BB$
$D \rightarrow aD$

zu N_ε: B
C

$\rightarrow N_\varepsilon = \{B, C\}$

ε-Regeln (Fortsetzung der Konstruktion)

2. Ersetzen jede Regel $A \to X_1 X_2 \dots X_n \in P$, $X_i \in N \cup T$, durch die Menge aller Regeln der Form
$$A \to \alpha_1 \alpha_2 \dots \alpha_n,$$
wobei
   1. $\alpha_i = X_i$, falls $X_i \notin N_\varepsilon$;
   2. $\alpha_i \in \{X_i, \varepsilon\}$, falls $X_i \in N_\varepsilon$;
   3. $\alpha_1 \alpha_2 \dots \alpha_n \neq \varepsilon$.

Die Menge der so erhaltenen Regeln ist die Regelmenge P' von $G' = (N, T, P', S)$.

2. Schritt – Beispiel

$\cancel{S \rightarrow ABC \mid AB}$ $\rightsquigarrow$ $S \rightarrow ABC \mid AC \mid AB \mid A$

$\cancel{A \rightarrow b \mid CD}$ $\rightsquigarrow$ $A \rightarrow b \mid CD \mid D$

$\cancel{B \rightarrow Bb \mid \varepsilon}$ $\rightsquigarrow$ $B \rightarrow Bb \mid b$

$\cancel{C \rightarrow BB}$ $\rightsquigarrow$ $C \rightarrow BB \mid B$

$\cancel{D \rightarrow aD}$ $\rightsquigarrow$ $D \rightarrow aD$

zu N_ε: B

C

$\rightarrow N_\varepsilon = \{B, C\}$

Korrektheit der Konstruktion

Behauptung: Für alle $A \in N$ und alle $w \in T^*$ gilt $A \overset{*}{\underset{G'}{\Longrightarrow}} w$ **gdw.** $w \neq \varepsilon$ und $A \overset{*}{\underset{G}{\Longrightarrow}} w$.
(Dann gilt $w \in L(G')$ gdw. $w \in L(G) \setminus \{\varepsilon\}$.)

Beweis. **1)** Es gelte $A \overset{n}{\underset{G}{\Longrightarrow}} w$ und $w \neq \varepsilon$. Zeigen durch VI über n, dass dann $A \overset{*}{\underset{G'}{\Longrightarrow}} w$.

I.A. ($n = 1$): Wenn $A \underset{G}{\Longrightarrow} w$, dann $A \to w \in P$. Da $w \neq \varepsilon$ ist $A \to w \in P'$, also $A \underset{G'}{\Longrightarrow} w$.

I.S. ($n \to n + 1$): Wenn $A \overset{n+1}{\underset{G}{\Longrightarrow}} w$, dann $A \underset{G}{\Longrightarrow} X_1 X_2 \dots X_k \overset{n}{\underset{G}{\Longrightarrow}} w_1 w_2 \dots w_k$ mit $X_i \in N \cup T$, $w = w_1 w_2 \dots w_k$ und $X_i \overset{*}{\underset{G}{\Longrightarrow}} w_i$ in weniger als n Schritten, $1 \leq i \leq k$.

Für alle i, $1 \leq i \leq k$: Falls $w_i = \varepsilon$, dann $X_i \in N_\varepsilon$.

Dann gilt $A \overset{n}{\underset{G}{\Longrightarrow}} X_{i_1} X_{i_2} \dots X_{i_r}$ $(1 \leq i_j \leq k)$ mit $X_{i_1}, X_{i_2}, \dots, X_{i_r} \notin N_\varepsilon$, also $w_{i_j} \neq \varepsilon$, $1 \leq j \leq r$.

Nach I.V. gilt dann $X_{i_j} \overset{*}{\underset{G'}{\Longrightarrow}} w_{i_j}$, $1 \leq j \leq r$, also $A \overset{*}{\underset{G'}{\Longrightarrow}} w_{i_1} w_{i_2} \dots w_{i_r}$ und $w_{i_1} w_{i_2} \dots w_{i_r} = w$.

Korrektheit der Konstruktion

2) Es gelte $A \overset{n}{\underset{G'}{\Longrightarrow}} w$. Es gilt $w \neq \varepsilon$, da G' keine ε-Regeln hat. VI über n, dass $A \overset{*}{\underset{G}{\Longrightarrow}} w$.

I.A. ($n = 1$): Wenn $A \underset{G'}{\Longrightarrow} w$, dann $A \to w \in P'$. Dann gibt es eine Regel der Form $A \to w_0 B_1 w_1 \dots w_{k-1} B_k w_k \in P$ mit $w_0 w_1 \dots w_k = w$ und $B_i \in N_\varepsilon$ für $1 \leq i \leq k$.
Somit gilt $A \underset{G}{\Longrightarrow} w_0 B_1 w_1 \dots w_{k-1} B_k w_k \overset{*}{\underset{G}{\Longrightarrow}} w$.

I.S. ($n \to n + 1$): Wenn $A \underset{G'}{\Longrightarrow} X_1 X_2 \dots X_k \overset{n}{\underset{G'}{\Longrightarrow}} w$, dann gibt es $A \to \alpha \in P$, und die Regel $A \to X_1 X_2 \dots X_k$ entsteht in P' durch Streichen von Symbolen aus N_ε in α.
Also gilt $A \underset{G}{\Longrightarrow} \alpha \overset{*}{\underset{G}{\Longrightarrow}} X_1 X_2 \dots X_k$.
Außerdem gibt es eine Zerlegung $w = w_1 w_2 \dots w_k$ mit $X_i \overset{*}{\underset{G'}{\Longrightarrow}} w_i$, $1 \leq i \leq k$,
und nach I.V. $X_i \overset{*}{\underset{G}{\Longrightarrow}} w_i$, $1 \leq i \leq k$, also $A \overset{*}{\underset{G}{\Longrightarrow}} w$. $\blacksquare$

Nutzlose Symbole

- Symbol in einer kfG **nutzlos**, wenn es in keiner Satzform vorkommt oder kein terminales Wort daraus ableitbar ist.
- **Lemma 12.2.** Zu jeder kfG $G = (N, T, P, S)$ kann eine kfG $G' = (N', T', P', S)$ mit $L(G) = L(G')$ konstruiert werden, sodass N' **keine nutzlosen Symbole** enthält, das heißt, dass für jedes $A \in (N' \cup T')$ gilt:
  1. es gibt ein $v \in T'^*$, sodass $A \stackrel{*}{\Longrightarrow} v$ (A ist **erzeugend**) und
  2. es gibt ein $\alpha \in (N' \cup T')^*$ mit $|\alpha|_A > 0$, sodass $S \stackrel{*}{\Longrightarrow} \alpha$ (A ist **erreichbar**).

Nutzlose Symbole

Beweis: 1. Konstruieren $G'' = (N'', T, P'', S)$ mit **Eigenschaft 1**.

- Fügen iterativ Nichtterminale $A \in N$ zu N'' hinzu, falls $A \to v \in P$ und $v \in (T \cup N'')^*$.
- Sei P'' die Menge aller Regeln $A \to v \in P$ mit $A \in N''$ und $v \in (T \cup N'')^*$.

➢ Jede Ableitung in G'' ist eine Ableitung in G ist, gilt $L(G'') \subseteq L(G)$.

- Für $L(G) \subseteq L(G'')$ ist zu zeigen:
  Wenn $A \overset{*}{\underset{G}{\Longrightarrow}} v$ für ein $v \in T^*, n \geq 1$, dann gilt $A \in N''$.
- *Induktionsbeweis ähnlich zu Lemma 12.1.*

Nutzlose Symbole

2. Konstruieren **aus $\boldsymbol{G''}$** die Grammatik $G' = (N', T', P', S)$ **mit Eigenschaft 2**.

- Setzen initial $N' = \{S\}, T' = \emptyset$.
- Falls $A \in N'$ und $A \to v \in P''$, dann füge
  - alle Nichtterminale in v der Menge N'
  - und alle Terminale in v der Menge T' hinzu.
- $P' = \{\, A \to v \mid A \to v \in P'', A \in N' \,\}$
- Dann enthalten N' und T' nur noch Symbole, die in Satzformen von G'' vorkommen und es gilt $L(G') = L(G'')$.

$\blacksquare$

Fortsetzung des Beispiels

$\cancel{S \to ABC \mid AB}$ $\quad\rightsquigarrow\quad S \to ABC \mid AC \mid AB \mid A$

$\cancel{A \to b \mid CD}$ $\quad\rightsquigarrow\quad A \to b \cancel{\mid CD \mid D}$

$\cancel{B \to Bb \mid \varepsilon}$ $\quad\rightsquigarrow\quad B \to Bb \mid b$

$\cancel{C \to BB}$ $\quad\rightsquigarrow\quad C \to BB \mid B$

$\cancel{D \to aD}$ $\quad\rightsquigarrow\quad \cancel{D \to aD}$

$N'' = \{A, B, S, C\}$

$N' = \{S, A, B, C\}$

$T' = \{b\}$

- *Die Beseitigung nutzloser Symbole führt keine ε-Regeln ein.*
- *Führt man die Konstruktionen zuerst nach 12.1. und danach nach 12.2. aus, dann erhält man eine kfG ohne nutzlose Symbole und ohne ε-Regeln.*

Kettenregeln

Lemma 12.3. Zu jeder kfG $G = (N, T, P, S)$ ohne ε-Regeln kann eine kfG $G' = (N, T, P', S)$ mit $L(G) = L(G')$ konstruiert werden, sodass P' **keine Kettenregeln** (Regeln der Form A $\rightarrow B$ mit $B \in N$) enthält.

Beweis: Konstruieren $N_c = \{\ (A, B) \mid A, B \in N, A \overset{*}{\Longrightarrow} B\ \}$.
Da P keine ε-Regeln enthält, genügt folgende Prozedur beginnend mit $N_c = \emptyset$:

1. Für alle $A \in N$ füge (A, A) hinzu.
2. Iteriere: Wenn $(A, C) \in N_c$ und $C \rightarrow B \in P$, dann füge (A, B) hinzu.

Nun entsteht P' aus P wie folgt:

1. Streiche alle Kettenregeln.
2. Für jedes $(A, B) \in N_c$ und $B \rightarrow \beta \in P$ mit $\beta \notin N$, füge $A \rightarrow \beta$ hinzu.

Kettenregeln

- $A \overset{*}{\underset{G}{\Longrightarrow}} w$ ohne Kettenregeln gdw. $A \overset{*}{\underset{G'}{\Longrightarrow}} w$, $w \in T^*$.
- Wenn $A = \alpha_0 \underset{G}{\Longrightarrow}_L \alpha_1 \cdots \underset{G}{\Longrightarrow}_L \alpha_n = w$ und für $0 \leq i < j < n$ gilt, dass $|\alpha_i| = |\alpha_j|$ und $\alpha_j \underset{G}{\Longrightarrow} \alpha_{j+1}$ durch Anwendung von $A \to v$ mit $|v| > 1$, dann $\alpha_i \underset{G'}{\Longrightarrow}_L \alpha_{j+1}$ aufgrund der Konstruktion von P'. Also $A \overset{*}{\underset{G'}{\Longrightarrow}} w$.
- Ebenso kann jede Anwendung einer Regel aus $A \to v \in P' \setminus P$ durch eine Ableitung in G simuliert werden, in der zunächst Kettenregeln auf A und angewendet werden, so dass $A \overset{*}{\underset{G}{\Longrightarrow}} B$ und dann die Regel $B \to v$ verwendet wird.
- Somit gilt $L(G) = L(G')$. ■

Fortsetzung des Beispiels

$S \to ABC \mid AB \mid AC \mid \cancel{A}$ $\rightsquigarrow$ $S \to ABC \mid AC \mid AB \mid b$

$A \to b$ $\rightsquigarrow$ $A \to b \mid CD \mid D$

$B \to Bb \mid b$ $\rightsquigarrow$ $B \to Bb \mid b$

$C \to BB \mid \cancel{B}$ $\rightsquigarrow$ $C \to BB \mid B \mid Bb \mid b$

$$N_c = \{(A,A),(B,B),(C,C),(S,S),(S,A),(C,B)\}$$

- *Die Beseitigung der Kettenregeln führt weder ε-Regeln noch nutzlose Symbole ein.*
- *Führt man die Konstruktionen in der Reihenfolge nach 12.1., 12.2, 12.3. aus, dann erhält man eine kfG ohne nutzlose Symbole, ε- und Kettenregeln.*

Chomsky-Normalform (CNF)

Satz 12.4: Zu jeder kfG $G = (N, T, P, S)$ kann eine kfG $G' = (N, T, P', S)$ **in Chomsky-Normalform** mit $L(G') = L(G) \setminus \{\varepsilon\}$ konstruiert werden.

Konstruktion:

1. Beseitigung der ε-Regeln (nach 12.1.)
2. Beseitigung nutzloser Symbole (nach 12.2.)
3. Beseitigung der Kettenregeln (nach 12.3.)
4. Für alle Regeln $A \to X_1 X_2 \dots X_n$ mit $n > 1$ führe aus:
   Wenn X_i ein terminales Symbol a ist, ersetze es in der Regel durch ein neues Nichtterminal C_a und füge die Regel $C_a \to a$ hinzu.

Chomsky-Normalform (CNF)

5. Anschließend führe für alle Regeln $A \to B_1 B_2 \ldots B_n$ mit $n > 2$ neue Nichtterminale $D_1, D_2, \ldots, D_{m-2}$ ein.
   Ersetze dann die Regel $A \to B_1 B_2 \ldots B_n$ durch folgende Regeln:

$$A \to B_1 D_1,\, D_1 \to B_2 D_2, \ldots ,\; D_{m-3} \to B_{m-2} D_{m-2},\; D_{m-2} \to B_{n-1} B_n$$

- ➢ In den Schritten 4. und 5. werden keine nutzlosen Symbole und weder ε- noch Kettenregeln eingeführt.
- ➢ Die entstandene Grammatik ist in CNF und erzeugt $L(G) \setminus \{\varepsilon\}$. ■

Fortsetzung des Beispiels

$S \to ABC \mid AB \mid AC \mid b$
$A \to b$
$B \to \cancel{Bb} \mid b$ $\rightsquigarrow$ $B \to BC_b$
$C \to BB \mid \cancel{Bb} \mid b$ $\rightsquigarrow$ $C \to BC_b$ $\qquad C_b \to b$

$S \to \cancel{ABC} \mid AB \mid AC \mid b$ $\rightsquigarrow$ $S \to AD_1,\ D_1 \to BC$
$A \to b$
$B \to BC_b \mid b$
$C \to BB \mid BC_b \mid b$
$C_b \to b$

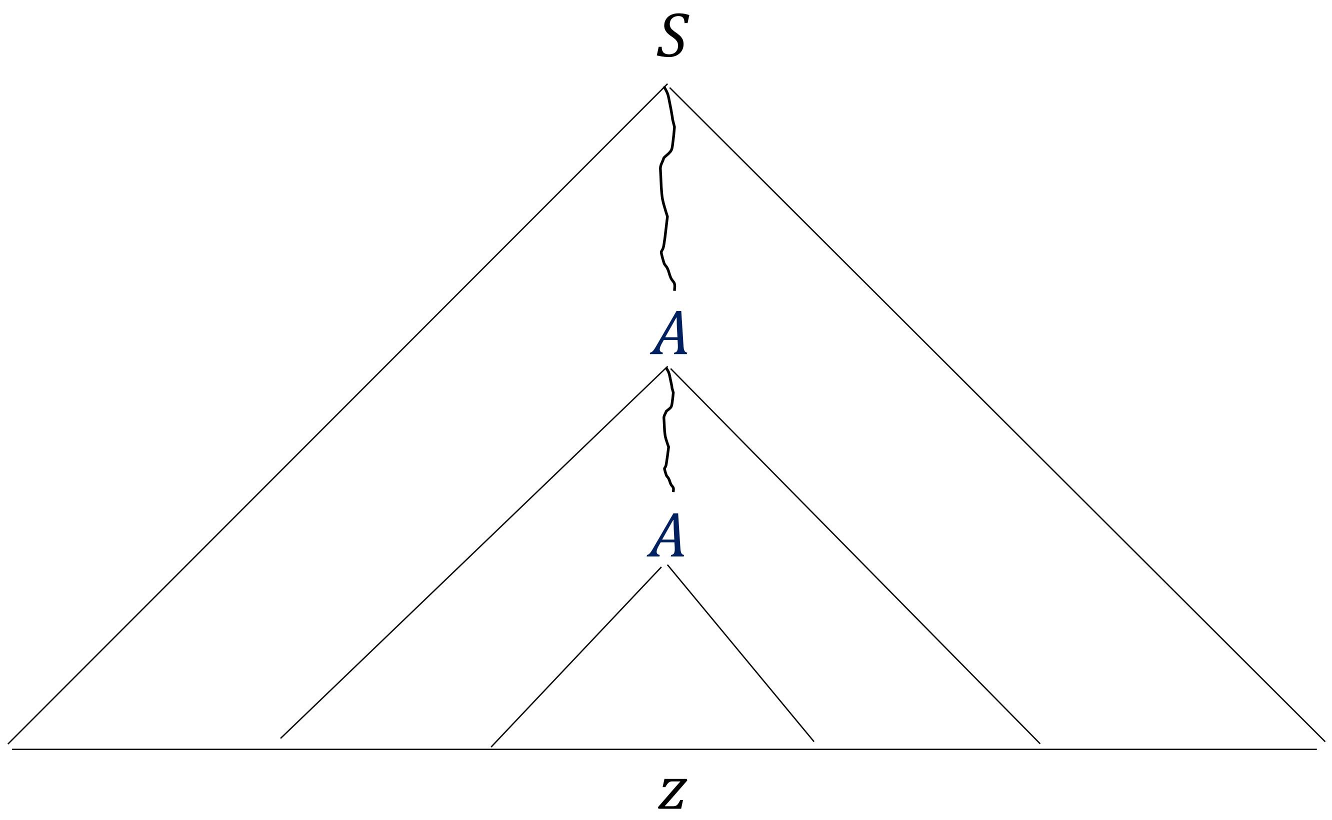
Die Pumping-Eigenschaft von kfG

- Betrachten Wörter aus der Sprache, die so lang sind, dass in der Ableitung ein Nichtterminal mehrfach vorkommen muss.
- Sei $G = (N, T, P, S)$ eine kfG **in CNF** und $|N| = n$.
- Jeder Ableitungsbaum ist ein binärer Baum, wobei
  - alle inneren Knoten mit einem Blatt als Kind nur dieses eine Kind
  - und alle anderen inneren Knoten genau zwei Kinder haben.

➢ Hat der längste Pfad im Ableitungsbaum die Länge i, dann gibt es maximal 2^{i-1} viele Blätter (wie ein Binärbaum der Tiefe $i-1$).

- Betrachten Wörter $z \in L(G)$ mit $|z| \geq 2^n$.
  - ➢ Es gibt mindestens einen Pfad der Länge $n+1$ in jedem Ableitungsbaum für z.
  - ➢ Mindestens zwei der n inneren Knoten haben die gleiche Markierung.

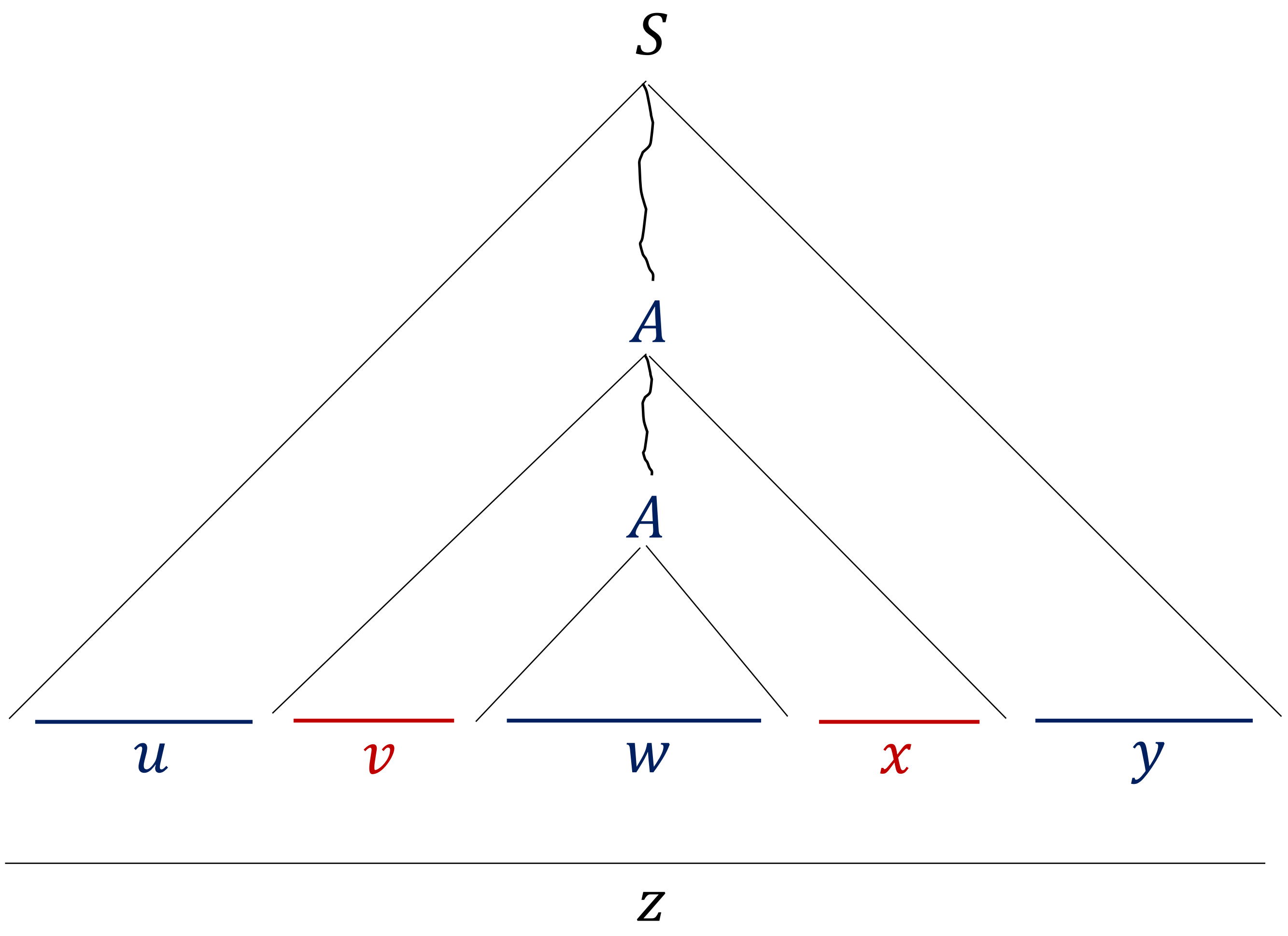
Die Pumping-Eigenschaft von kfG



Seien die A die beiden <u>letzten</u> Auftreten einer wiederholten Markierung auf dem längsten Pfad im Ableitungsbaum.

Pfadlänge vom ersten dieser A bis zu einem Blatt ist höchstens n (nicht mehr als $n + 1$ Knoten).

Die Pumping-Eigenschaft von kfG



Pfadlänge vom ersten dieser A bis zu einem Blatt ist höchstens n (nicht mehr als $n+1$ Knoten).

⬇

$$|vwx| \leq 2^n$$
$$|vx| > 0$$

Setzen $k = 2^n$.

Pumping-Lemma für kfG

Lemma 12.5. Sei L eine kontextfreie Sprache.
Dann gibt es eine (von L abhängige) Konstante $k > 0$, so dass jedes $z \in L$ mit $|z| \geq k$ eine Zerlegung $z = uvwxy$ hat mit

1. $|vwx| \leq k$,
2. $|vx| > 0$,
3. $uv^iwx^iy \in L$ für alle $i \geq 0$.

Beweis: Nach Betrachtungen zum Ableitungsbaum für z wissen wir
$S \overset{*}{\Longrightarrow} uAy \overset{*}{\Longrightarrow} uvAxy \overset{*}{\Longrightarrow} uvwxy$. Wegen $A \overset{*}{\Longrightarrow} w$ und $A \overset{*}{\Longrightarrow} vAx$ sind auch
$S \overset{*}{\Longrightarrow} uAy \overset{*}{\Longrightarrow} uwy$ und $S \overset{*}{\Longrightarrow} uAy \overset{*}{\Longrightarrow} uv^iAx^iy \overset{*}{\Longrightarrow} uv^iwx^iy$ Ableitungen für alle i. ■

Anwendung des Pumping-Lemmas

- $L = \{ a^n b^n c^n \mid n \geq 0 \} \notin \mathcal{L}(\mathrm{CF})$
- **Annahme**: $L \in \mathcal{L}(\mathrm{CF})$.
  - Dann gibt es eine Konstante k aus dem PL für kfG.
  - Für das Wort $z = a^k b^k c^k$ gilt $z \in L$ und $|z| \geq k$.
  - Daher gilt $z = uvwxy$ mit den Eigenschaften 1. bis 3. aus dem PL für kfG.
  - **1. Fall:** $vwx \in \{a\}^*$. Da $|vx| > 0$, gilt $|uv^2wx^2y|_a > |uv^2wx^2y|_b$, also $uv^2wx^2y \notin L$.
  - **2. Fall:** $vwx \in \{b\}^*$ oder $vwx \in \{c\}^*$. Analog.
  - **3. Fall:** $|vwx|_a > 0$ und $|vwx|_b > 0$. Dann gilt $uv^2wx^2y \in L(a^+b^+a^+b^+c^+)$, also $uv^2wx^2y \notin L$.
  - **4. Fall:** $|vwx|_b > 0$ und $|vwx|_c > 0$. Analog.
  - Da es wegen 1. (PL) keine weiteren Fälle gibt, liegt ein **Widerspruch** zu 3. (PL) vor.

Hierarchie

Folgerung 12.6. $\mathcal{L}(\mathrm{REG}) \subset \mathcal{L}(\mathrm{CF}) \subset \mathcal{L}(\mathrm{REC}) \subset \mathcal{L}(\mathrm{RE})$

Beweisskizze: $L = \{\ a^n b^n c^n \mid n \geq 0\ \} \notin \mathcal{L}(\mathrm{CF})$.

Es genügt also zu zeigen, dass L entscheidbar ist.

Die DTM, die $\{\ a^n b^n \mid n \geq 0\ \}$ entscheidet, kann leicht zu einer DTM erweitert werden, die $\{\ a^n b^n c^n \mid n \geq 0\ \}$ entscheidet.

Es gilt $\{\ a^n b^n c^n \mid n \geq 0\ \} \in \mathcal{L}(\mathrm{REC})$. $\square$