Automaten und Formale Sprachen

SoSe 2017 in Trier

Henning Fernau
Universität Trier
fernau@uni-trier.de

23. Juni 2017

- Organisatorisches
- Einführung
- Endliche Automaten und reguläre Sprachen
- **Kontextfreie Grammatiken und kontrextfreie Sprachen**
- Chomsky-Hierarchie

Kontextfreie Grammatiken und kontrextfreie Sprachen

1. Automaten mit unendlichem Speicher
2. Kontextfreie Grammatiken
3. **Normalformen**
4. **Nichtkontextfreie Sprachen**
5. Algorithmen für kontextfreie Grammatiken

Eine **kontextfreie Grammatik in Chomsky-Normalform** ist ein Quadrupel $G = (\Sigma, N, R, S)$ mit:

- Σ ist das *Terminalalphabet*,
- N ist das *Nonterminalalphabet* (die *Variablenmenge*),
- $R \subset (N \times N^2) \cup (N \times \Sigma)$ ist das Alphabet der *Regeln* oder *Produktionen*;
- $S \in N$ ist das *Startsymbol* oder *Anfangszeichen*.

Ziel der Vorlesung: o.B.d.A.: Chomsky-Normalform

Chomsky-Normalform

Satz: Zu jeder kfG G gibt es eine kfG G' in Chomsky-Normalform mit $L(G)\setminus\{\lambda\} = L(G')$.

Beweis: Zunächst können wir in der angegebenen Reihenfolge die Lemmata anwenden.

Hinweis: Durch das zweite und dritte Lemma wird die im ersten Lemma sichergestellte Eigenschaft nicht wieder zerstört, und durch das dritte Lemma nicht die im zweiten Lemma hergestellte Eigenschaft.

Dann sind möglicherweise noch rechte Regelseiten länger als zwei.

Dies wird durch den Algorithmus auf der nächsten Folie repariert. $\square$

Algorithmus zur Vermeidung zu langer rechter Regelseiten

Grundidee: Einfügen von “Zwischenzuständen” beim “Buchstabieren”

Eingabe: Regelmenge R (die im Folgenden modifiziert wird)

WHILE $(\exists A \to w \in R : \ell(w) > 2)$ DO:

1. Stelle $w = Bu$ dar.
2. Erzeuge neues Nichtterminalzeichen X.
3. Ersetze $A \to w$ durch $A \to BX$ und $X \to u$ in R.

Korrektheit ist klar. Frage: Warum terminiert das Verfahren überhaupt?

Ein Beispiel

Regelmenge $S \to aSb, S \to \lambda, S \to c$.

1. Lemma $\leadsto S \to X_a S X_b,\ S \to \lambda,\ S \to X_c,\ X_a \to a,\ X_b \to b,\ X_c \to c$.

2. Lemma $\leadsto S \to X_a S X_b,\ S \to X_a X_b,\ S \to X_c,\ X_a \to a,\ X_b \to b,\ X_c \to c$.

3. Lemma $\leadsto S \to X_a S X_b,\ S \to X_a X_b,\ S \to c,\ X_a \to a,\ X_b \to b,\ X_c \to c$.

Konstruktion aus Satz $\leadsto S \to X_a X,\ X \to S X_b,\ S \to X_a X_b,\ S \to c,\ X_a \to a,\ X_b \to b,\ X_c \to c$.

Satz: Es gibt einen Algorithmus, der zu jeder vorgelegten kfG $G = (\Sigma, N, R, S)$ und jedem $w \in \Sigma^*$ entscheidet, ob $w \in L(G)$ gilt.

Beweis: Wird $w = \lambda$ angefragt, so berechnen wir N^λ wie im Beweis zum 2. Lemma. Es gilt: $\lambda \in L(G)$ gdw. $S \in N^\lambda$.

Wird $w \neq \lambda$ angefragt, so überführen wir G in eine Chomsky-Normalform kfG G' mit der Eigenschaft $w \in L(G)$ gdw. $w \in L(G')$.

Da G' keine λ-Regeln enthält, kann man “$w \in L(G)$?” durch Berechnen aller aus G' ableitbaren Satzformen der Länge höchstens $\ell(w)$ beantworten. $\square$

Cocke, Younger und Kasami haben gezeigt, wie man die Komplexität des beschriebenen Verfahrens durch dynamisches Programmieren erheblich vermindern kann.

Das Verfahren von Cocke, Younger und Kasami (*CYK-Algorithmus*)

Satz: Ist eine kfG G in Chomsky-Normalform fixiert, so lässt sich die Frage "$w \in L(G)$?" in einer Zeit beantworten, die sich durch ein kubisches Polynom in $\ell(w)$ abschätzen lässt.

Beweis: Betrachte $w = a_1 \dots a_n$.
Lege 2–dimensionale dreiecksförmige Tabelle T an mit Nonterminalmengen als Einträgen so dass $T[i,k] = \{A \in N \mid A \stackrel{*}{\Rightarrow} a_i \dots a_k\}$. T kann man wie folgt "von oben nach unten" berechnen:

$$T[i,i] = \{A \in N \mid A \rightarrow a_i \in R\}$$

$$T[i,k] = \{A \in N \mid A \rightarrow BC \in R \wedge \exists i \leq j < k : B \in T[i,j], C \in T[j+1,k]\}$$

Die CYK-Tabelle

a_1	a_2	$\dots$	a_n
$T[1,1]$	$T[2,2]$	$\dots$	$T[n,n]$
$T[1,2]$	$T[2,3]$	$\dots$	
$\vdots$	$\vdots$		
$T[1,n-1]$	$T[2,n-1]$		
$T[1,n]$			

$a_1 \dots a_n \in L(G)$ gdw. das Startsymbol erscheint in $T[1,n]$.

Der CYK-Algorithmus (schon detaillierter als normaler Pseudo-Code)

FOR $j = 1, \ldots, n$

$\quad T[i,i] = \{A \in N \mid A \to a_i \in R\}$

FOR $k = 1, \ldots, n-1$

$\quad$ FOR $i = 1, \ldots, n-k$

$\quad\quad T[i, k+i] = \emptyset$

$\quad\quad$ FOR $j = i, \ldots, k+i-1$

$\quad\quad\quad$ IF $\exists A \to BC \in R : B \in T[i,j], C \in T[j+1, k+i]$

$\quad\quad\quad$ THEN $T[i, k+i] = T[i, k+i] \cup \{A\}$

Ein Beispiel

$S \longrightarrow AX$, $X \longrightarrow SB$, $S \longrightarrow c$, $A \longrightarrow a$, $B \longrightarrow b$ erzeugt $L = \{a^n c b^n \mid n \geq 0\}$.

a	a	c	b	b
A	A	S	B	B
		X		
	S			
	X			
S				

Alternative Sichtweise: Hornlogik mit Markierungsalgorithmus/Resolution

Erinnerung: Die Wahrheit einer Aussage kann in Linearzeit bewiesen werden.

Übersetze Eingabewort $w = aacbb$ in Tatsachen:
$A(0,1)$ (weil a das erste Zeichen ist und $A \to a$ Regel),
$A(1,2)$ (weil a das zweite Zeichen ist und $A \to a$ Regel),
$S(2,3)$, $B(3,4)$, $B(4,5)$.

Übersetze Chomsky-NF-Regeln in (kubisch viele) Aussagen:
$S \to AX \rightsquigarrow S(i,k) \leftarrow (A(i,j) \wedge X(j,k))$ für $0 \leq i < j < k \leq 5$,
$X \to SB \rightsquigarrow X(i,k) \leftarrow (S(i,j) \wedge B(j,k))$ für $0 \leq i < j < k \leq 5$.

Jedem Nichtterminal entsprechen quadratisch viele Aussagenvariablen.
Die Indizes beziehen sich auf Wortpositionenbereiche.
Im Beispiel gilt: w ist ableitbar gdw. $S(0,5)$ ist wahr.

Satz: **KF** ist unter Vereinigung abgeschlossen.

Beweis: L_1, L_2 seien durch kfGs G_1 bzw. G_2 spezifiziert.

Genauer sei, für $i = 1, 2$, $G_i = (\Sigma_i, N_i, R_i, S_i)$. O.E. nehmen wir an: $N_1 \cap N_2 = \emptyset$.

Betrachte $G = (\Sigma_1 \cup \Sigma_2, N_1 \cup N_2 \cup \{S\}, R_1 \cup R_2 \cup \{S \to S_1, S \to S_2\}, S)$ mit $S \notin N_1 \cup N_2$.

Beh.: $L(G) = L(G_1) \cup L(G_2)$.

$w \in L(G)$ gdw. es gibt Ableitungsfolge $S \Rightarrow S_i \overset{*}{\Rightarrow}_G w$ ($S \notin N_1 \cup N_2$)

gdw. $S_1 \overset{*}{\Rightarrow}_{G_1} w$ oder $S_2 \overset{*}{\Rightarrow}_{G_2} w$ ($N_1 \cap N_2 = \emptyset$) gdw. $w \in L(G_1)$ oder $w \in L(G_2)$

gdw. $w \in L(G_1) \cup L(G_2)$. $\square$

Satz: **KF** ist unter Konkatenation abgeschlossen.

Beweis: L_1, L_2 seien durch kfGs G_1 bzw. G_2 spezifiziert.

Genauer sei, für $i = 1, 2$, $G_i = (\Sigma_i, N_i, R_i, S_i)$. O.E. nehmen wir an: $N_1 \cap N_2 = \emptyset$.

Betrachte $G = (\Sigma_1 \cup \Sigma_2, N_1 \cup N_2 \cup \{S\}, R_1 \cup R_2 \cup \{S \to S_1 S_2\}, S)$ mit $S \notin N_1 \cup N_2$.

Beh.: $L(G) = L(G_1)L(G_2)$.

$w \in L(G)$ gdw. es gibt Ableitungsfolge $S \Rightarrow S_1 S_2 \overset{*}{\Rightarrow}_G w$ ($S \notin N_1 \cup N_2$)

gdw. $S_1 \overset{*}{\Rightarrow}_{G_1} u$ und $S_2 \overset{*}{\Rightarrow}_{G_2} v$ mit $uv = w$ ($N_1 \cap N_2 = \emptyset$) gdw. $u \in L(G_1)$ und $v \in L(G_2)$

gdw. $w \in L(G_1)L(G_2)$ mit $w = uv$. $\square$

Satz: **KF** ist unter Kleene Stern abgeschlossen.

Beweis: L sei durch kfG $G = (\Sigma, N, R, S)$ spezifiziert.

Betrachte $G' = (\Sigma, N \cup \{\hat{S}\}, R \cup \{\hat{S} \to \hat{S}S, \hat{S} \to \lambda\}, \hat{S})$ mit $\hat{S} \notin N$.

Beh.: $L(G') = L(G)^*$.

$w \in L(G')$ gdw. es gibt Ableitungsfolge $\hat{S} \Rightarrow^k \hat{S}S^k \Rightarrow S^k \stackrel{*}{\Rightarrow}_{G'} w_1 \ldots w_k = w$

gdw. $S^k \stackrel{*}{\Rightarrow}_G w_1 \ldots w_k = w$ (denn $\hat{S} \notin N$)

gdw. $S \stackrel{*}{\Rightarrow}_G w_i$ für $1 \leq i \leq k$

gdw. $w \in L(G)^*$. $\square$

Satz: **KF** ist unter Durchschnitt mit regulären Sprachen abgeschlossen.

Den Beweis führen wir "später" in allgemeinerem Kontext für sogenannte Kellerautomaten als Spezialfall. Auf Grammatikebene ist es auch einfach.

L sei durch kfG $G = (\Sigma, N, R, S)$ in Chomsky-NF spezifiziert; o.E. $\lambda \notin L$, die reguläre Sprache durch DEA $A = (Q, \Sigma, \delta, q_0, F)$. Konstruiere kfG $G' = (\Sigma, N', R', S')$ in Chomsky-NF wie folgt:

$$\begin{aligned}
N' &= Q \times N \times Q \cup \{S'\}, \\
R' &= \{S' \to (q_0, S, q_f) \mid q_f \in F\} \\
&\cup \{(p, A, r) \to (p, B, q)(q, C, r) \mid q \in Q, A \to BC \in R\} \\
&\cup \{(p, A, q) \to a \mid A \to a \in R, \delta(p, a) = q\}
\end{aligned}$$

Es gilt: $L(G') = L \cap L(A)$.

Details zur Übung.

Ein Beispiel zur Schulung der Intuition

Betrachte

$$L = \{a^n b^n c^n \mid n \in \mathbb{N}\}$$

Eine kontextfreie Grammatik für L müsste irgendwie die Länge des a-Blocks mit der des folgenden b-Blocks abgleichen.
Dazu muss sie wohl Regeln der Form $X \to aX'b$ verwenden.
Dann besteht keine Möglichkeit mehr, die Länge des c-Blocks mit der eines vorhergehenden a- oder b-Blocks abzugleichen.

Vermutung: L ist nicht kontextfrei.

Ein **binärer Wurzelbaum** ist gegeben durch ein Tripel $B = (V, \phi, r)$ mit ausgezeichneter *Wurzel* $r \in V$ und einer *Vater-Abbildung* $\phi : V \setminus \{r\} \to V$ mit der Eigenschaft

$$\forall v \in V : \#\underbrace{\{u \in V \mid \phi(u) = v\}}_{\kappa(v):=} \leq 2$$

$\kappa(v)$ liefert also die *Kinder* von v.
Knoten v mit $\kappa(v) = \emptyset$ heißen *Blätter*.

Lemma: Der Ableitungsbaum eines jeden von einer kontextfreien Grammatik in Chomsky-Normalform akzeptierten Wortes kann als binärer Wurzelbaum aufgefasst werden.

Die *Höhe* eines binären Wurzelbaumes $B = (V, \phi, r)$ ist gegeben durch

$$h(B) = \max_{v \in V \setminus \{r\}} \{k \in \mathbb{N} \mid \phi^k(v) = r\}.$$

Sonderfall $h(B) = 0$ bedeutet: $B = (\{r\}, \phi, r)$ mit trivialem ϕ.

Lemma: Hat ein binärer Wurzelbaum B mehr als 2^h Blätter, so gilt $h(B) > h$.

Beweis: Wir zeigen die *Kontraposition* per Induktion:

Gilt $h(B) \leq h$, so hat $B = (V, \phi, r)$ höchstens 2^h viele Blätter.

$h = 0$ ist trivial.

Es gelte $h > 0$. Daher gilt $\kappa(r) \neq \emptyset$.

Betrachte $r' \in \kappa(r)$. $V' := \{v \in V \mid \exists k \in \mathbb{N} : \phi^k(v) = r'\}$.

ϕ' sei die Einschränkung von ϕ auf $V' \setminus \{r'\}$.

$B' = (V', \phi', r')$ ist ein binärer Wurzelbaum der Höhe höchstens $h - 1$.

Auf B' ist die Induktionshypothese anwendbar: B' hat höchstens 2^{h-1} viele Blätter.

$\rightsquigarrow$ B hat höchstens $\#\kappa(r) \cdot 2^{h-1} \leq 2^h$ viele Blätter. $\square$

Lemma: Ist $G = (\Sigma, N, R, S)$ eine kfG in Chomsky-Normalform und ist $w \in L(G)$ mit $\ell(w) > 2^{\#N}$, so gilt für jeden Ableitungsbaum von w bzgl. G, dass es einen Weg von S zu einem Blatt gibt, auf dem mehr als $\#N$ viele Nichtterminalzeichen ersetzt werden.

Beweis: Nach dem vorigen Lemma hat der unterliegende binäre Wurzelbaum eine Höhe größer als $\#N$. Es gibt also einen Weg in besagtem Ableitungsbaum von der mit S beschrifteten Wurzel zu einem Blatt, dem mehr als $\#N$ Regelanwendungen entspricht $\rightsquigarrow$ Behauptung. $\square$

Folgerung: Auf besagtem Weg von der Wurzel zum Blatt im Ableitungsbaum von w finden wir also nach dem Schubfachprinzip zwei Regelanwendungen $A \to v$ und $A \to u$ mit gleicher linker Seite.

Erweiterte Chomsky-Normalform

Satz: Jede $L \in$ **KF** lässt sich beschreiben durch eine kfG $G = (\Sigma, N, R, S)$ mit Regeln der Form $N \times ((NN) \cup (\Sigma))$; zusätzlich darf eine Regel $S \to \lambda$ existieren, wobei dann gefordert ist, dass S in keiner rechten Regelseite vorkommt.

Beweis: Wie vorher erklärt, gibt es kfG $G' = (\Sigma, N', R', S')$ in Chomsky-Normalform für $L(G) \setminus \{\lambda\}$.
Definiere $R'' = \{S \to w \mid S' \to w \in R'\}$ für ein neues $S \notin N'$ und
$R = R' \cup R'' \cup \{S \to \lambda \mid \lambda \in L\}$, $N = N' \cup \{S\}$.
$\rightsquigarrow G = (\Sigma, N, R, S)$ beschreibt L und hat die geforderten Eigenschaften. $\square$

Hinweis: Die vorige Folie hat auch für die erweiterte Chomsky-Normalform Gültigkeit.

Ein Pumping-Lemma für KF

Satz: Zu jeder kfS L gibt es eine Konstante $n > 0$, sodass jedes Wort $w \in L$ mit $\ell(w) \geq n$ als Konkatenation $w = uvxyz$ dargestellt werden kann mit geeigneten u, v, x, y, z mit folgenden Eigenschaften:

1. $\ell(v) > 0$ oder $\ell(y) > 0$;
2. $\ell(vxy) \leq n$;
3. $\forall i \geq 0 : uv^ixy^iz \in L$.

Beweis: L werde durch kfG $G = (\Sigma, N, R, S)$ in erweiterter Chomsky-Normalform beschrieben.
Es sei $n = 2^{\#N}$.
Gibt es kein $w \in L$ mit $\ell(w) \geq n$, so stimmt die Aussage leer.
Sonst wähle ein $w \in L$ mit $\ell(w) \geq n$.
Nach obiger Folgerung gibt es zwei Regelanwendungen $A \to u'$ und $A \to v'$ auf einem Weg in einem Ableitungsbaum von w.
$\rightsquigarrow$ Es gibt Linksableitung

$$S \stackrel{*}{\Rightarrow} uA\zeta \stackrel{*}{\Rightarrow} uvA\eta\zeta \stackrel{*}{\Rightarrow} uvx\eta\zeta \stackrel{*}{\Rightarrow} uvxy\zeta \stackrel{*}{\Rightarrow} uvxyz.$$

Wegen Chomsky-NF (keine Kettenregeln oder λ-Regeln) gilt $\ell(v) > 0$ oder $\ell(y) > 0$.
$\ell(vxy) \leq n$ ergibt sich aus dem Beweis der Folgerung sowie wegen Chomsky-NF.
Daher gibt es auch Ableitungen

$$S \stackrel{*}{\Rightarrow} uA\zeta \stackrel{*}{\Rightarrow} ux\zeta \stackrel{*}{\Rightarrow} uxz$$

und allgemeiner

$$S \to uA\zeta \stackrel{*}{\Rightarrow} uvA\eta\zeta \stackrel{*}{\Rightarrow} uvvA\eta\eta\zeta \stackrel{*}{\Rightarrow} uv^iA\eta^i\zeta \stackrel{*}{\Rightarrow} uv^ixy^iz.$$

□

Ein Beispiel zur Anwendung des Pumping-Lemmas:

$$L = \{a^k b^k c^k \mid k \in \mathbb{N}\} \notin \textbf{KF}$$

Wäre L kontext-frei, so gäbe es Pumping-Konstante n.
Wähle $w = a^n b^n c^n$ als genügend langes Wort.
Diskutiere Zerlegungen $w = uvxyz$.
Wegen $\ell(vxy) \leq n$ enthält vxy höchstens a's gefolgt von b's oder b's gefolgt von c's.
Beide Fälle sind analog; diskutiere also den ersten.
"Nullpumpen" liefert ein Wort, in dem alle a- und b-Vorkommen zusammen weniger als die doppelte Anzahl von c's ausmachen (Widerspruch zur Wortstruktur).

Das Pumping-Lemma kennzeichnet nicht Kontextfreiheit

Betrachte

$$L = \{a^k d^r a^k d^s a^k \mid k, r, s \in \mathbb{N}\}$$

Mit der gleichen Intuition wie vorher ist die Sprache nicht kontextfrei.
L erfüllt aber das Pumping-Lemma:
Ein Wort der Form $w = a^k d^r a^k d^s a^k$ mit $r > 0$ lässt sich zerlegen in: $w = uvxyz$ mit (z.B.) $u = a^k$, $v = d$ und $y = \lambda$.
Dann gilt $uv^ixy^iz \in L$ für alle $i \in \mathbb{N}$.
Der Fall $r = 0$ aber $s > 0$ geht analog.
Gilt $r = s = 0$, so gilt für $w = a^{3k}$, und die Wahl $v = a^3$, $y = \lambda$ führt wiederum auf $uv^ixy^iz \in L$ für alle $i \in \mathbb{N}$.

Wie kann man die **Intuition retten** ?

Betrachte

$$L' = \{a^k d a^k d a^k \mid k \in \mathbb{N}\}$$

Jetzt ermöglicht Nullpumpen einen Widerspruch zur angenommenen Kontextfreiheit. (Übungsaufgabe !)

Idee: Verwende Abschlusseigenschaften,
hier Durchschnitt mit regulären Sprachen:

Wäre $L = \{a^k d^r a^k d^s a^k \mid k, r, s \in \mathbb{N}\} \in$ **KF**, so auch $L' = L \cap \{a\}^*\{d\}\{a\}^*\{d\}\{a\}^*$.

Noch eine Übungsaufgabe. . .

Sei $\#_x w$ = Anzahl der Vorkommen von x in w.

Zeigen Sie, dass

$$L = \{w \in \{a, b, c\}^* \mid \#_a w = \#_b w = \#_c w\}$$

nicht kontextfrei ist.

Zur rechten Wahl des Pumpwortes

Betrachte $L = \{tt \mid t \in \{a, b\}^*\}$.

Wäre $L \in$ **KF**, so gäbe es Pump-Konstante n für L.

Diskutiere $w = a^n b^n a^n b^n \in L$.
Für vxy mit $\ell(vxy) \leq n$ gibt es drei Unterfälle:
(1) $vxy \in \{a\}^*\{b\}^*$, $z \in \{b\}^*\{a^n b^n\}$.
(2) $vxy \in \{b\}^*\{a\}^*$.
(3) $vxy \in \{a\}^*\{b\}^*$, $u \in \{a^n b^n\}\{a\}^*$.

Wir betrachten (3) eingehend (die anderen Fälle sind ähnlich):
$uxz = a^n b^n a^m b^k \in L$ mit $m < n$ oder $k < n$ (Nullpumpen).
$uxz = ss$ mit $\ell(s) < 2n$.
Das “erste s” muss mit a anfangen, aber das zweite muss mit b beginnen (wegen $\ell(s) < 2n$ und da uxz nicht nur aus a’s besteht).

Schlechte Wahl: $w = a^\ell b^\ell a^\ell b^\ell \in L$ für $\ell = \lceil n/4 \rceil$.

Nachtrag zu Abschlusseigenschaften

Satz: **KF** ist nicht unter Durchschnitt abgeschlossen.

Beweis: $L_1 = \{a^n b^n c^k \mid k, n \in \mathbb{N}\}$

$L_2 = \{a^n b^k c^k \mid k, n \in \mathbb{N}\}$

L_1 und L_2 sind kontextfrei.

$L_1 \cap L_2$ ist jedoch nicht kontextfrei (s.o.). $\square$

Satz: **KF** ist nicht unter Komplementbildung abgeschlossen.

Beweis: **KF** ist gegen Vereinigung abgeschlossen. Wäre **KF** gegen Komplement abgeschlossen, so nach de Morgan auch unter Durchschnitt, im Gegensatz zum vorigen Satz. $\square$

Entscheidbarkeitsfragen I

Frage: Gibt es einen Algorithmus, sodass ...

Satz: Das Leerheitsproblem ist für kfG entscheidbar.

1. Beweis: Ersetze alle Terminalzeichenvorkommen in der Grammatik G durch das leere Wort; das liefert neue Grammatik G'. Dann gilt: $L(G) \neq \emptyset$ gdw. $L(G') = \{\lambda\}$. $\lambda \in L(G')$ kann man entscheiden (siehe letzter Foliensatz). $\square$

2. Beweis: o.E. G in erweiterter Chomsky-Normalform.
Die betreffende Pump-Konstante n_G garantiert: Ist $L(G) \neq \emptyset$, so gibt es $w \in L(G)$ mit $\ell(w) < n_G$ (Nullpumpen).
Daher kann man $L(G) = \emptyset$ durch Testen aller Wörter bis zur Länge n_G entscheiden. $\square$

Entscheidbarkeitsfragen II

Endlichkeitsproblem: Gegeben Sprachbeschreibung G für L, ist L endlich ?

Satz: Das Endlichkeitsproblem ist für kfG entscheidbar.

Beweis: Dies folgt wie in obigem 2. Beweis zum Leerheitsproblem aus:

Lemma: Zu jeder Sprache $L \in$ **KF** gibt es eine Konstante $n > 0$, sodass gilt: L ist unendlich gdw. es gibt ein Wort $t \in L$ mit $n \leq \ell(t) < 2n$.

Lemma: Zu jeder Sprache $L \in$ **KF** gibt es eine Konstante $n > 0$, sodass gilt: L ist unendlich gdw. es gibt ein Wort $t \in L$ mit $n \leq \ell(t) < 2n$.

Beweis: Es sei n die Pump-Konstante aus dem Pumping-Lemma.

Ist L unendlich, so gibt es insbesondere Wörter mit Mindestlänge $2n$ in L.

Wähle unter diesen ein Wort t' minimaler Länge.

Nach dem Pumping-Lemma können wir $t' = uvxyz$ schreiben mit $\ell(vxy) \leq n$.

$\rightsquigarrow$ Nullpumpen liefert $t = uxz$ mit $n \leq \ell(t) < 2n$, da t' minimal.

Ein Wort $t \in L$ mit $n \leq \ell(t) < 2n$ können wir "aufpumpen",

d.h., mit der Existenz solch eines t ist L unendlich. $\square$