24. Kontextfreie Sprachen

Obwohl das Wortproblem für kontextsensitive Sprachen entscheidbar ist, ist nicht bekannt, ob dieses auch tatsächlich – d.h. in Polynomialzeit – entscheidbar ist. Da man allgemein vermutet, dass $\mathrm{NSPACE}(O(n))$ nicht in P enthalten ist, ist dies wegen der gezeigten Gleichheit $\mathrm{KS} = \mathrm{NSPACE}(O(n))$ jedoch unwahrscheinlich. Für kontextfreie Grammatiken, die wir uns in diesem Abschnitt näher ansehen möchten, kann man dagegen die Lösbarkeit des Wortproblems in Polynomialzeit zeigen. Da auf der anderen Seite viele wichtige Sprachen (wie z.B. die üblichen Programmiersprachen, wenn man von der Typenkonsistenz absieht) kontextfrei sind, wurden die Klasse KF und Teilklassen hiervon sehr eingehend untersucht.

Aus Zeitgründen werden wir uns hier auf einige Aspekte beschränken müssen. Mit Hilfe von Normierungen kontextfreier Grammatiken, insbesondere der Chomsky-Normalform, werden wir zeigen, dass jede kontexfreie Sprache auch kontextsensitiv ist, und eine Struktureigenschaft (Pumping Lemma) der kontextfreien Sprachen herleiten, die die Echtheit der Inklusion von KF in KS zeigt. Weiter werden wir kurz auf Abschlusseigenschaften und Entscheidbarkeitsfragen für KF eingehen. Eine Maschinencharakterisierung von KF werden wir dagegen nur erwähnen, aber nicht beweisen.

Ein wichtiges Hilfsmittel für die Untersuchung kontextfreier Grammatiken ist die Darstellung von Herleitungen mit Hilfe von Herleitungsbäumen, wie wir es schon in Abschnitt 21 an einem Beispiel (s. Beispiel 21.4) getan haben. Möglich ist dies, da die Prämisse einer kontextfreien Regel nur aus einer Variablen besteht. Sei $G = (N,T,P,S)$ eine kontextfreie Grammatik. In einem *G-Herleitungsbaum* eines (Terminal-)Wortes w ist die Wurzel mit dem Axiom S markiert, die Markierungen der Söhne eines mit einer Variablen X markierten Knotens ergeben die Konklusion v einer Regel von G mit Prämisse X[1], und die Blätter von links nach rechts gelesen ergeben das hergeleitete Wort w. Jeder G-Herleitung kann man eindeutig einen G-Herleitungsbaum zuordnen. Umgekehrt entsprechen einem G-Herleitungsbaum jedoch i. Allg. mehrere Herleitungen. Diese unterscheiden sich jedoch nur in der Reihenfolge der Regelanwendungen. Führen zwei Herleitungen zu demselben Herleitungsbaum, nennt man diese daher *äquivalent*.

Verlangt man, dass in einer Herleitung stets Regeln auf die am weitesten links stehende Variable angewendet werden, so spricht man von einer *Linksherleitung*. Man kann zeigen, dass man jedem Herleitungsbaum genau eine Linksherleitung zuordnen kann, da durch die Forderung, stets die erste Variable zu ersetzen, die Reihenfolge der Regelanwendungen durch den Baum festgelegt ist. Insbesondere besitzt also jedes herleitbare Wort eine Linksherleitung.

Wir wollen diese Aussagen hier nicht beweisen, sondern begnügen uns mit einem erläuternden Beispiel.

[1] Genauer: Im Falle einer λ-Regel $X \to \lambda$ hat der mit X markierte Knoten einen mit λ markierten Sohn. Im Falle einer Regel $X \to Y_1 \dots Y_n$ ($n \geq 1$) sind es n Söhne, die von links nach rechts mit $Y_1, \dots, Y_n$ markiert sind. Wir gehen also davon aus, dass die Söhne linear (d.h. von links nach rechts) geordnet sind.

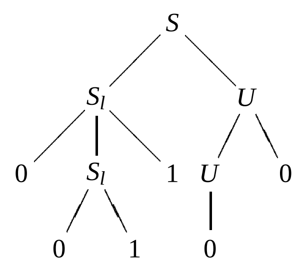
24.1 BEISPIEL. Sei $L = \{0^m1^m0^n : m,n \geq 1\} \cup \{0^m1^n0^n : m,n \geq 1\}$ (vgl. Beispiel 21.10 in Abschnitt 21). Eine kontextfreie Grammatik $G = (N, \{0,1\}, P, S)$, die L erzeugt, kommt mit den Variablen $N = \{S, S_l, S_r, U\}$ und den Regeln

$$
\begin{array}{llcll}
\text{(R1)} & S & \to & S_lU & \\
\text{(R2)} & S & \to & US_r & \\
\text{(R3,4)} & S_l & \to & 0S_l1 & \mid\ 01 \\
\text{(R5,6)} & S_r & \to & 1S_r0 & \mid\ 10 \\
\text{(R7,8)} & U & \to & U0 & \mid\ 0
\end{array}
$$

aus. Hierbei sind die Regeln so gewählt, dass man mit den Regeln R1, R3, R4, R7, R8 Wörter der Form $0^m1^m0^n$, und mit den Regeln R2, R5, R6, R7, R8 Wörter der Form $0^m1^n0^n$ herleiten kann. Gilt $n = m$, so stehen beide Möglichkeiten offen. So besitzt das Wort $w = 0^21^20^2$ z.B. folgende Herleitungen:

$$
\begin{array}{lccccccccccc}
\text{(H1)} & S & \Rightarrow & S_lU & \Rightarrow & 0S_l1U & \Rightarrow & 0S_l1U0 & \Rightarrow & 0011U0 & \Rightarrow & 001100 \\
\text{(H2)} & S & \Rightarrow & S_lU & \Rightarrow & S_lU0 & \Rightarrow & S_l00 & \Rightarrow & 0S_l100 & \Rightarrow & 001100 \\
\text{(H3)} & S & \Rightarrow & US_r & \Rightarrow & U0S_r & \Rightarrow & 00S_r & \Rightarrow & 001S_r0 & \Rightarrow & 001100
\end{array}
$$

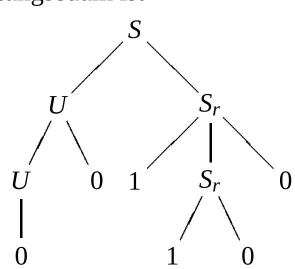
Die ersten beiden Herleitungen sind äquivalent, und führen zu dem Herleitungsbaum



Die zugehörige (zu H1 und H2 äquivalente) Linksherleitung ist

$$S \Rightarrow S_lU \Rightarrow 0S_l1U \Rightarrow 0011U \Rightarrow 0011U0 \Rightarrow 001100$$

Der zu H3 gehörende Herleitungsbaum ist



Die Herleitung H3 (die bereits eine Linksherleitung ist) ist also zu H1 und H2 nicht äquivalent.

Besitzt ein Wort w nichtäquivalente Herleitungen (d.h. verschiedene Linksherleitungen) in einer Grammatik G, so heißt es *mehrdeutig* (bzgl. G), und die Grammatik G heißt *mehrdeutig*, wenn es bzgl. G mindestens ein mehrdeutiges Terminalwort gibt. Eine kontextfreie Sprache L heißt *inhärent mehrdeutig*, wenn jede kontextfreie Grammatik, die L erzeugt, mehrdeutig ist. Die Grammatik G ist, wie das Beispielwort $w = 0^21^20^2$ zeigt, mehrdeutig. Die mehrdeutigen Wörter sind (wie man leicht zeigen kann) gerade die Wörter der Form $w = 0^m1^m0^m$ $(m \geq 1)$. Man kann weiter zeigen (nicht so einfach), dass die Sprache L inhärent mehrdeutig ist. $\diamond$

Um zu zeigen, dass jede kontextfreie Sprache auch kontextsensitiv ist, geben wir zu jeder kontextfreien Grammatik G eine äquivalente λ-treue kontextfreie Grammatik G' an. Die Grammatik G' ist dann insbesondere kontextsensitiv.

24.2 SATZ. *Zu jeder kontextfreien Grammatik $G = (N,T,P,S)$ kann man effektiv eine äquivalente kontextfreie Grammatik $G' = (N',T,P',S')$ angeben, die λ-treu und separiert ist.*

24.3 KOROLLAR. KF $\subseteq$ KS. □

Zum Beweis von Satz 24.2 zeigen wir zunächst, dass man für eine kontextfreie Grammatik G die *eliminierbaren*, d.h. in das leere Wort überführbaren Variablen effektiv bestimmen kann.

24.4 LEMMA. *Die Menge*

$$E \;=\; \{X \in N : X \overset{*}{\underset{G}{\Rightarrow}} \lambda\}$$

der eliminierbaren Variablen einer kontextfreien Grammatik $G = (N,T,P,S)$ lässt sich effektiv bestimmen.

BEWEIS. Die Mengen E_n ($n \geq 1$) seien induktiv definiert durch

$$\begin{aligned} E_1 &= \{X \in N : X \to \lambda \in P\} \\ E_{n+1} &= E_n \cup \{X \in N : \exists w \in E_n^*(X \to w \in P)\}. \end{aligned}$$

Offensichtlich lassen sich die Mengen E_n effektiv angeben, und es gilt

$$E_1 \subseteq E_2 \subseteq E_3 \subseteq \ldots$$

Da die Mengen E_i in der endlichen Menge N enthalten sind, wird diese Kette schließlich stationär, d.h. es gibt n_0 minimal mit $E_n = E_{n_0}$ für alle $n \geq n_0$, also $E_{n_0} = \bigcup_{n\geq 1} E_n$. Nach Definition der Mengen E_n können wir n_0 (und damit E_{n_0}) effektiv bestimmen. Es gilt nämlich

$$n_0 \;=\; \mu n(E_n = E_{n+1}),$$

da $E_n = E_{n+1}$ impliziert, dass $E_n = E_m$ für alle $m \geq n$ gilt (Induktion!).

Es genügt also, $E = E_{n_0}$ zu zeigen. Hierzu zeigt man zunächst $E_n \subseteq E$ für alle n durch eine triviale Induktion nach n. Zum Nachweis der Umkehrung $E \subseteq E_{n_0}$ hat man für gegebenes $X \in E$ zu zeigen, dass $X \in E_n$ für ein geeignetes n und damit auch $X \in E_{n_0}$ gilt. Wegen $X \in E$ gibt es eine Herleitung

$$X = w_n \;\Rightarrow\; w_{n-1} \;\Rightarrow\; \ldots \;\Rightarrow\; w_1 \;\Rightarrow\; w_0 = \lambda,$$

die X in n Schritten ($n \geq 1$ geeignet) in λ überführt. Mit einer (einfachen) Induktion zeigt man hierfür, dass $w_i \in E_i^*$ gilt, also insbesondere $X \in E_n$. □

Wir können nun Satz 24.2 beweisen.

BEWEIS VON SATZ 24.2. Wegen Lemma 23.2 können wir davon ausgehen, dass die Grammatik $G = (N,T,P,S)$ bereits separiert ist. Es bleibt die λ-Treue sicherzustellen.

Wegen Lemma 24.4 können wir effektiv feststellen, ob λ in der von G erzeugten Sprache liegt, da $\lambda \in L(G)$ genau dann gilt, wenn das Axiom S eliminierbar ist, d.h. $S \in E$ gilt. Es genügt also, eine λ-freie kontextfreie Grammatik $\tilde{G} = (\tilde{N}, T, \tilde{P}, \tilde{S})$ anzugeben mit $L(\tilde{G}) = L(G) - \{\lambda\}$. Ist nämlich $\lambda \notin L(G)$, so ist $\tilde{G}$ die gewünschte, zu G äquivalente λ-treue und separierte Grammatik G'. Gilt $\lambda \in L(G)$, so erhält man G' aus $\tilde{G}$, indem man ein neues Axiom hinzufügt, das in das alte Axiom überführt oder eliminiert werden kann. D.h. $G' = (\tilde{N} \cup \{S'\}, T, \tilde{P} \cup \{S' \to \tilde{S}, S' \to \lambda\}, S')$.

Die Grammatik $\tilde{G}$ erhält man aus der Grammatik G, indem man alle λ-Regeln von G weglässt. Für jede Umformungsregel $X \to Y_1 \ldots Y_n$ ($n \geq 1$) nimmt man dafür alle Varianten $X \to w$ hinzu, wobei $w \neq \lambda$ aus $Y_1 \ldots Y_n$ durch Weglassen von frei gewählten Vorkommen eliminierbarer Variablen entsteht.

Es ist klar, dass in $\tilde{G}$ nur nichtleere Wörter aus $L(G)$ herleitbar sind. Benutzt man nämlich in einer $\tilde{G}$-Herleitung von z eine Variante $X \to w$ einer Regel $X \to Y_1 \ldots Y_n$ aus P, so kann man diese Regelanwendung in G simulieren, indem man zunächst die Regel $X \to Y_1 \ldots Y_n$ anwendet und dann die in w weggelassenen Variablen Y_i eliminiert. (Formal beweist man die entsprechende Aussage für Satzformen durch Herleitungsinduktion.)

Zum Nachweis der Umkehrung $L(G) - \{\lambda\} \subseteq L(\tilde{G})$ führen wir die Grammatik $\hat{G} = (N, T, \hat{P}, S)$ mit $\hat{P} = P \cup \tilde{P} \cup \{X \to \lambda : X \in E\}$ ein und zeigen

$$w \in L(\hat{G}) - \{\lambda\} \;\Rightarrow w \text{ besitzt eine } \hat{G}\text{-Herleitung, in der keine } \lambda\text{-Regel angewendet wird.} \tag{24.1}$$

Dies genügt, da nach Definition von $\hat{G}$ die Grammatiken G und $\tilde{G}$ „Teilgrammatiken" von $\hat{G}$ sind, wobei $\tilde{G}$ alle $\hat{G}$-Regeln mit Ausnahme der λ-Regeln enthält. Ist also $w \in L(G) - \{\lambda\}$, so gilt auch $w \in L(\hat{G}) - \{\lambda\}$ und die nach (24.1) existierende λ-Regel-freie $\hat{G}$-Herleitung von w ist zugleich eine $\tilde{G}$-Herleitung, weshalb $w \in L(\tilde{G})$.

Zum Nachweis von (24.1) zeigen wir durch Induktion nach n, dass sich jede $\hat{G}$-Herleitung H der Länge n eines Terminalwortes $w \neq \lambda$ in eine λ-Regel-freie $\hat{G}$-Herleitung H' desselben Wortes überführen lässt.

Sei also solch eine Herleitung H der Länge n gegeben, wobei wir o.B.d.A. davon ausgehen können, dass in H eine λ-Regel $Y \to \lambda$ angewendet wird. (Andernfalls können wir $H' = H$ wählen.) Da H ein Wort $w \neq \lambda$ herleitet, kann die Regel $Y \to \lambda$ nicht die erste Regelanwendung sein (also insbesondere muss $n \geq 1$ sein). Da $\hat{G}$ separiert ist, muss also Y zuvor durch eine Umformungsregel $X \to Y_1 \ldots Y_m$ erzeugt worden sein, d.h. $Y = Y_i$ für ein $i \in \{1, \ldots, m\}$. Die beiden Regelanwendungen $X \to Y_1 \ldots Y_m$ und $Y \to \lambda$ können wir jedoch durch die Anwendung der Regel $r = X \to Y_1 \ldots Y_{i-1} Y_{i+1} \ldots Y_m$ ersetzen. (Man beachte, dass, für $m \geq 2$, $r \in \tilde{P} \subseteq \hat{P}$, da $Y = Y_i \in E$ gilt. Ist $m = 1$, d.h. $X \to Y_i$ mit $Y = Y_i$, so ist mit $Y \in E$ auch $X \in E$, weshalb $r = X \to \lambda \in \hat{P}$ nach Definition von $\hat{P}$.) Dies liefert eine Herleitung H_0 von w der Länge $n - 1$. Nach Induktionsvoraussetzung gibt es dann aber eine λ-Regel-freie Herleitung $(H' :=) H_0'$ von w. □

Satz 24.2 lässt sich durch Einschränkung der zulässigen Umformungsregeln noch verschärfen.

24.5 DEFINITION. Eine kontextfreie Grammatik $G = (N, T, P, S)$ ist in *Chomsky-Normalform*, wenn G λ-treu ist und neben der eventuellen λ-Regel $S \to \lambda$ nur Regeln der Form

$$X \to YZ \quad \text{und} \quad X \to a \quad (X, Y, Z \in N, a \in T)$$

enthält.

24.6 SATZ. *Jede kontextfreie Grammatik $G = (N,T,P,S)$ lässt sich effektiv in eine äquivalente kontextfreie Grammatik $G' = (N',T,P',S')$ in Chomsky-Normalform überführen.*

BEWEIS. Nach Satz 24.2 dürfen wir annehmen, dass die gegebene Grammatik G λ-treu und separiert ist. Es genügt dann, Umformungsregeln

$$X \to Y_1 \dots Y_n \quad \text{mit } n = 1 \text{ oder } n \geq 3$$

durch Regeln der zulässigen Gestalt zu ersetzen.

Wir erreichen dies in zwei Schritten, indem wir zunächst eine zu G äquivalente Grammatik $\tilde{G} = (\tilde{N},T,\tilde{P},S)$ angeben, in der Umformungsregeln mit zu langen Konklusionen eliminiert sind, und dann die zu $\tilde{G}$ äquivalente Grammatik $G' = (N',T,P',S)$ angeben, die auch keine Variablenumbenennungen mehr enthält.

Die Grammatik $\tilde{G}$ entsteht aus G, indem man jede Regel $X \to Y_1 \dots Y_n$ mit $n \geq 3$ durch die Regeln

$$\begin{array}{rcl} X & \to & Y_1 Z_2 \\ Z_2 & \to & Y_2 Z_3 \\ & \vdots & \\ Z_{n-1} & \to & Y_{n-1} Y_n \end{array}$$

ersetzt, wobei die Variablen $Z_1, \dots, Z_{n-1}$ neu (und für jede Regel verschieden) sind.

Zur Elimination der Variablenumbenennungen in $\tilde{G}$ definieren wir für jede Variable $X \in \tilde{N}$ die Menge

$$U(X) \;=\; \{Y \in \tilde{N} : X \overset{*}{\underset{G}{\Rightarrow}} Y\}$$

der Variablen, in die X überführt werden kann, und ersetzen alle Umbenennungsregeln

$$X \to Y_1 \mid Y_2 \mid \dots \mid Y_k$$

mit Prämisse X durch die Regeln

$$\begin{array}{ll} X \to a, & \text{wobei } \exists Y \in U(X)(Y \to a \in \tilde{P}) \\ X \to Z_1 Z_2, & \text{wobei } \exists Y \in U(X)(Y \to Z_1 Z_2 \in \tilde{P}). \end{array}$$

Die so definierte Grammatik G' kann man effektiv aus $\tilde{G}$ gewinnen, da man die Mengen $U(X)$ effektiv berechnen kann. Hierzu beobachtet man (ähnlich wie im Beweis von Lemma 24.4), dass $U(X)$ der kleinste Fixpunkt der durch

$$\begin{array}{rcl} U_0(X) & = & \{X\} \\ U_{n+1}(X) & = & U_n(X) \cup \{Y \in \tilde{N} : \exists Z \in U_n(X)(Z \to Y \in \tilde{P})\} \end{array}$$

definierten aufsteigenden Folge $U_0(X) \subseteq U_1(X) \subseteq \dots$ ist.

Auf den formalen Beweis, dass $L(G) = L(\tilde{G}) = L(G')$ gilt, verzichten wir. □

Der folgende Satz ist ein wichtiges Werkzeug zum Nachweis der Nichtkontextfreiheit von Sprachen.

24.7 SATZ. (PUMPING LEMMA FÜR KF) *Zu jeder kontextfreien Sprache* $L \subseteq \Sigma^*$ *gibt es eine Zahl* $p \in \mathbb{N}$*, sodass jedes Wort* $z \in L$ *der Länge* $|z| \geq p$ *eine Zerlegung* $z = uvwxy$ *in Teilwörter* $u, v, w, x, y \in \Sigma^*$ *mit folgenden Eigenschaften besitzt:*

(i) $vx \neq \lambda$

(ii) $|vwx| < p$

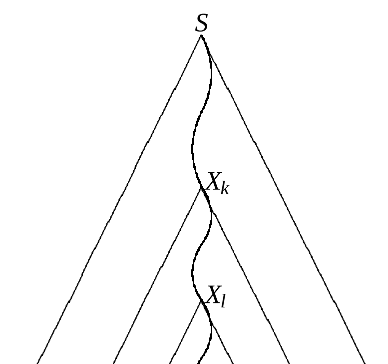
(iii) $\forall n \geq 0\, (z_n = uv^n wx^n y \in L)$

Hierbei lässt sich solch eine Zahl p *aus jeder kontextfreien Grammatik* G*, die* L *erzeugt, effektiv berechnen.*

BEWEIS. Sei $G = (N, T, P, S)$ eine kontextfreie Grammatik in Chomsky-Normalform, die L erzeugt. Aufgrund der Normalform ist jeder G-Herleitungsbaum ein Binärbaum. Da in einem Binärbaum B die Anzahl b der Blätter in der Tiefe t des Baumes (d.h. der Länge t des längsten Pfades, t die Anzahl der Kanten, d.h $t+1$ die Anzahl der Knoten des Pfades) durch

$$b \leq 2^{t+1}$$

beschränkt ist, folgt hieraus, dass jedes Wort $z \in L$ mit $|z| \geq 2^{t+1}$ einen Herleitungsbaum der Tiefe $\geq t$ hat. Wählen wir also $p = 2^{||N||+2} + 1$, ein Wort $z \in L$ mit $|z| \geq p$ und einen G-Herleitungsbaum B von z, so hat B die Tiefe $t \geq ||N|| + 1$. Ist also α ein Pfad maximaler Länge in B, so sind dessen Knoten mit $S = X_1, \ldots, X_t, a$ $(X_i \in N, a \in T)$ beschriftet. Da $t \geq ||N|| + 1$, muss zumindest eine Variable mehrfach auf dem Pfad vorkommen. Wir können also $1 \leq k < t$ maximal und hierzu $k < l \leq t$ wiederum maximal wählen, so dass $X_k = X_l$ gilt. Die gewünschte Zerlegung $z = uvwxy$ des Wortes z erhalten wir nun dadurch, dass wir u und y als die (Inschriften der) Blätter wählen, die links bzw. rechts von X_k (d.h. von den Blättern unterhalb von X_k) liegen, v und x als die Blätter, die unterhalb von X_k aber links bzw. rechts von X_l liegen, und w als die Blätter, die unterhalb von X_l liegen:



Um zu zeigen, dass die Zerlegung die gewünschten Eigenschaften hat, beobachten wir, dass (durch geeignetes Anordnen der Regelanwendungen) dem Baum B eine Herleitung H von z entspricht

$$S \overset{*}{\Rightarrow} uX_k y \overset{*}{\Rightarrow} uvX_l xy \overset{*}{\Rightarrow} uvwxy$$

mit Teilherleitungen

$$S \overset{*}{\Rightarrow} uX_ky \tag{24.2}$$
$$X_k \overset{*}{\Rightarrow} vX_lx \tag{24.3}$$
$$X_l \overset{*}{\Rightarrow} w \tag{24.4}$$

entspricht. Da $k < l$, hat hierbei die Herleitung (24.3) die Länge ≥ 1, weshalb wegen der Normalform von G, $vx \neq \lambda$, also die Anforderung (i) an die Zerlegung erfüllt ist. Da α ein Pfad maximaler Länge ist, und k maximal gewählt wurde, kommt es echt unterhalb von X_k auf keinem Pfad zu Variablenwiederholungen. D.h. der Teilbaum von B mit Wurzel X_k hat Tiefe $\leq ||N|| + 1$, und damit hat das aus X_k (gemäß (24.3) und (24.4)) hergeleitete Teilwort vwx von z die Länge

$$|vwx| \leq 2^{||N||+2} < p$$

Hiermit ist also auch (ii) erfüllt. Zum Nachweis von (iii) müssen wir $z_n = uv^nwx^ny \in L$ für jedes $n \geq 0$ zeigen. Da $X_k = X_l$, erhalten wir jedoch aus (24.2)-(24.4) folgende Herleitung für z_n:

$$\begin{array}{lll} S & \overset{*}{\Rightarrow} uXy & ((24.2)) \\ & \overset{*}{\Rightarrow} uv^nXx^ny & (n \text{ mal } (24.3)) \\ & \overset{*}{\Rightarrow} uv^nwx^ny & ((24.4)) \end{array}$$

wobei wir X für $X_k = X_l$ geschrieben haben. □

24.8 KOROLLAR. KF $\subsetneq$ KS.

BEWEIS. Da wir KF $\subseteq$ KS bereits gezeigt haben, genügt es, eine Sprache $L \in$ KS $-$ KF anzugeben. Wir wählen hierfür $L = \{0^n1^n0^n : n \geq 1\}$, für die wir in Beispiel 21.11 in Abschnitt 21 bereits eine Grammatik vom Erweiterungstyp angegeben haben. Nach Satz 23.3 gilt also $L \in$ KS. Den Beweis, dass L nicht kontextfrei ist, führen wir indirekt. Für einen Widerspruch nehmen wir an, dass $L \in$ KF gelte. Nach dem Pumping Lemma gibt es dann eine Zahl p, sodass jedes $z \in L$ mit $|z| \geq p$ eine Zerlegung $z = uvwxy$ mit den oben angegebenen Eigenschaften (i), (ii) und (iii) besitzt. Wir wählen nun für $z = 0^p1^p0^p$ solch eine Zerlegung und betrachten das Wort $z_0 = uv^0wx^0y = uwy$, das nach (iii) wiederum zu L gehört, also die Gestalt $z_0 = uwy = 0^q1^q0^q$ für ein $q \geq 1$ hat. Da $|vwx| \leq p$ nach (ii), muss jedoch der erste Block 0^p von z ein Anfangsstück von u oder der zweite Block 0^p von z ein Endstück von y sein (oder beides), weshalb das Weglassen des Teils vx aus z in z_0 zumindest einen dieser Blöcke unverändert lässt. Dies bedeutet aber, dass $q = p$, also $z_0 = z$ gelten muss. Der weggelassene Teil vx von z ist also leer, im Widerspruch zu (i). □

Ein Beispiel für eine Sprache über dem unären Alphabet, die kontextsensitiv, aber nicht kontextfrei ist, ist die Sprache $L = \{0^{n^2} : n \geq 0\}$. Um $L \in$ KS zu zeigen, gibt man entweder eine kontextsensitive Grammatik an, oder man überzeugt sich, dass man L mit linearem Platzbedarf erkennen kann. Dass $L \notin$ KF gilt, zeigt man mit Hilfe des Pumping Lemmas. Hierzu beobachtet man, dass für jede unendliche kontextfreie Sprache L die Lücken zwischen den Längen der Wörter in L linear beschränkt sind, d.h. es eine Zahl p gibt, sodass es für alle n ein Wort $w \in L$ mit $np \leq |w| \leq (n+1)p$ gibt. Für eine kontextfreie Sprache $L = \{0^{f(n)} : n \geq 0\}$ über $\{0\}^*$ bedeutet dies aber, dass $f(n) \in O(n)$ gilt (Übung!).

Aus dem Pumping-Lemma folgt weiter, dass – im Gegensatz zum Fall der kontextsensitiven Grammatiken – das Leerheits- und Unendlichkeitsproblem für kontextfreie Grammatiken entscheidbar sind.

24.9 SATZ. *Für eine kontextfreie Grammatik kann man effektiv feststellen, ob $L(G)$ leer oder unendlich ist.*

BEWEIS. Gegeben eine kontextfreie Grammatik G, können wir eine Zahl p wie im Pumping Lemma für $L = L(G) \subseteq \Sigma^*$ effektiv bestimmen. Wir behaupten, dass

$$L \neq \emptyset \quad \Leftrightarrow \quad \exists w \in \Sigma^* \, (|w| < p \,\&\, w \in L) \tag{24.5}$$

$$L \text{ unendlich} \quad \Leftrightarrow \quad \exists w \in \Sigma^* \, (p \leq |w| < 2p \,\&\, w \in L) \tag{24.6}$$

Da das Wortproblem für kontextfreie Grammatiken entscheidbar ist[2], können wir die rechten Seiten in (24.5) und (24.6) effektiv überprüfen, da die Mitgliedschaft in L für eine vorgegebene endliche Menge von Wörtern zu testen ist. Daher liefern (24.5) und (24.6) Entscheidungsverfahren für das Leerheits- bzw. Unendlichkeitsproblem für KF.

Zum Beweis der nichttrivialen Richtung „$\Rightarrow$" in (24.5) nehmen wir $L \neq \emptyset$ an und wählen ein Wort $z \in L$ minimaler Länge, und nehmen für einen Widerspruch $|z| \geq p$ an. Nach dem Pumping Lemma gibt es dann eine Zerlegung $z = uvwxy$ von z mit $vx \neq \lambda$ und $z_0 = uwy \in L$. Da $|z_0| = |z| - |vx| < |z|$ gilt, widerspricht dies der Minimalität von z.

Zum Beweis von „$\Rightarrow$" in (24.6) nehmen wir L unendlich an. L enthält dann beliebig lange Wörter, weshalb wir ein Wort z mit minimaler Länge $\geq p$ wählen können. Dass dann $|z| < 2p$ gelten muss, zeigt man wieder mit dem Pumping Lemma. Gilt nämlich $|z| \geq 2p$, so wählen wir wieder eine Zerlegung $z = uvwxy$ mit $|vx| > 0$, $|vwx| \leq p$ und $z_0 = uwy \in L$. Da dann aber $p \leq |z_0| < |z|$ gilt, widerspricht dies der Minimalität von z.

Schließlich zum Beweis der Richtung „$\Rightarrow$" von (24.6) genügt es zu beobachten, dass das Pumping Lemma zu einem Wort $z \in L$ mit $|z| \geq p$ unendlich viele Wörter $z_n \in L$ $(n \geq 1)$ durch „Aufpumpen" von Teilen von z liefert. □

Die Verschiedenheit von KF und KS lässt sich auch durch unterschiedliche Abschlusseigenschaften dieser Klassen zeigen (vgl. Lemmata 23.14 und 23.15).

24.10 SATZ. KF *ist gegen Vereinigung und homomorphe Bilder, nicht aber gegen Durchschnitt und Komplement abgeschlossen.*

BEWEIS. Sind $L_1, L_2 \in$ KF und $G_i = (N_i, T_i, P_i, S_i)$ Grammatiken, die L_i erzeugen, wobei o.B.d.A. $N_i \cap (N_j \cup T_j) = \emptyset$ $(\{i, j\} = \{1, 2\})$, so erzeugt die kontextfreie Grammatik $G = (N_1 \cup N_2 \cup \{S\}, T_1 \cup T_2, P_1 \cup P_2 \cup \{S \to S_1|S_2\}, S)$ die Vereinigung $L_1 \cup L_2$ von L_1 und L_2.

Ist $L \subseteq \Sigma^*$ kontextfrei und $h : \Sigma^* \to T^*$ ein Homomorphismus, so zeigt man $h(L) = \{h(w) : w \in L\} \in$ KF wie folgt. Man wählt eine kontextfreie Grammatik $G = (N, \Sigma, P, S)$ in Chomsky-Normalform, die L erzeugt, und deren Variablenmenge N zu T disjunkt ist. Ersetzt man in P alle Substitutionsregeln $X \to a$ durch $X \to h(a)$, so erhält man eine kontextfreie Grammatik $G' = (N, T, P', S)$, die $h(L)$ erzeugt.

[2] Um $w \in L(G)$ zu entscheiden, überführen wir G effektiv in eine äquivalente kontextfreie Grammatik G' in Chomsky-Normalform. Diese ist dann auch kontextsensitiv, weshalb $w \in L(G')$, wie im vorhergehenden Abschnitt gezeigt, entscheidbar ist.

Dass KF nicht gegen den Durchschnitt abgeschlossen ist, zeigt das Beispiel der Sprachen $L_1 = \{0^m 1^m 0^n : m, n \geq 1\}$ und $L_2 = \{0^m 1^n 0^n : m, n \geq 1\}$. Diese sind kontextfrei (vgl. Beispiel 21.10 in Abschnitt 21), aber $L_1 \cap L_2 = \{0^n 1^n 0^n : n \geq 1\}$ ist, wie im Beweis von Korollar 24.8 gezeigt, nicht kontextfrei.

Da sich der Durchschnitt zweier Sprachen mit Hilfe der Vereinigung und des Komplements ($L_1 \cap L_2 = \overline{\overline{L_1} \cup \overline{L_2}}$) darstellen lässt, ist KF auch nicht gegen Komplementbildung abgeschlossen. □

Im Gegensatz zu der Klasse KS der kontextsensitiven Sprachen lässt sich die Klasse KF der kontextfreien Sprachen durch keine Komplexitätsklasse beschreiben. Man kann zeigen, dass

$$\mathrm{KF} \subseteq \mathrm{DTIME}(n^3)$$

gilt, da das Wortproblem für jede kontextfreie Grammatik in Chomsky-Normalform in Kubikzeit lösbar ist (Algorithmus von Cocke-Kasami-Younger, s. Literatur). Für die nichtdeterministische Zeitkomplexität kann man die Abschätzung verbessern zu

$$\mathrm{KF} \subseteq \mathrm{NTIME}(O(n)).$$

Hierzu beobachtet man, dass für eine kontextfreie Grammatik $G = (N, T, P, S)$ in Chomsky-Normalform die Länge l einer Herleitung linear in der Länge n des hergeleiteten Wortes w beschränkt ist: Da G λ-treu ist und keine Variablenumbenennungen erlaubt, gilt $l \leq 2n$. Eine nichtdeterministische Turingmaschine M kann daher mit linearem Zeitbedarf die Sprache L dadurch erkennen, dass sie eine Linksherleitung des Eingabewortes w „errät“. Es gibt jedoch sehr einfache Sprachen, die nicht kontextfrei sind, nämlich

$$\mathrm{DTIME}(O(n)) \not\subseteq \mathrm{KF}.$$

Ein Beispiel hierfür ist die Sprache $\{0^n 1^n 0^n : n \geq 1\}$, die in $\mathrm{DTIME}(O(n)) - \mathrm{KF}$ liegt.

Wenn man KF auch nicht mit Ressourcenbeschränkten Turingmaschinen beschreiben kann, so kann man doch eine Maschinen-Charakterisierung von KF mit Hilfe von nichtdeterministischen Kellermaschinen, d.h. 1-Stapel-Maschinen geben. Wir verzichten hier auf den nicht ganz einfachen Beweis und eine Formalisierung des Kellermaschinenkonzepts.