Universität Karlsruhe (TH)
Lehrstuhl Informatik für Ingenieure und Naturwissenschaftler
Dr. Thomas Worsch
Matthias Schulz

16. Juni 2008

Einführung in die Informatik

Lösungen zu Übungsblatt 8

– Kontextfreie Grammatiken –

Aufgabe 1:

a) $N = \{S, A, B\}, T = \{a, b, c\}$,
$P = \{S \to BcAAA$,
$B \to Ba \mid Bb \mid Bc \mid \epsilon$,
$A \to a \mid b \mid c\}$.

Ableitung: $S \Rightarrow BcAAA \Rightarrow BacAAA \Rightarrow acAAA \Rightarrow acaAA \Rightarrow acabA \Rightarrow acabc$

b) $N = \{S, A, B, C\}, T = \{a, b, c\}$,
$P = \{S \to aAa \mid bBb \mid cCc \mid a \mid b \mid c$
$A \to a \mid aAa \mid aAb \mid aAc \mid bAa \mid bAb \mid bAc \mid cAa \mid cAb \mid cAc$
$B \to b \mid aBa \mid aBb \mid aBc \mid bBa \mid bBb \mid bBc \mid cBa \mid cBb \mid cBc$
$C \to c \mid aCa \mid aCb \mid aCc \mid bCa \mid bCb \mid bCc \mid cCa \mid cCb \mid cCc\}$

Ableitung: $S \Rightarrow aAa \Rightarrow abAba \Rightarrow abcAcba \Rightarrow abcabca.$

c) $N = \{S\}, T = \{a, b, c\}$,
$P = S \to \epsilon \mid b \mid SS \mid aSc \mid cSa\}$

Ableitung: $S \Rightarrow SS \Rightarrow SSS \Rightarrow aScSS \Rightarrow abcSS \Rightarrow abcbS \Rightarrow abcbcSa \Rightarrow abcbca$

d) $N = \{S\}, T = \{a, b, c\}$,
$P = S \to \epsilon \mid a \mid b \mid c \mid aSa \mid bSb \mid cSc\}$

Ableitung: $S \Rightarrow aSa \Rightarrow abSba \Rightarrow abcba.$

e) $N = \{S, X_{i>j}, X_{i<j}, X_{j>k}, X_{j<k}, A, C\}, T = \{a, b, c\}$,
$P = S \to X_{i>j}C \mid X_{i<j}C \mid AX_{j>k} \mid AX_{j<k}$
$X_{i>j} \to aX_{i>j} \mid aX_{i>j}b \mid a$
$X_{i<j} \to X_{i<j}b \mid aX_{i<j}b \mid b$
$X_{j>k} \to bX_{j>k} \mid bX_{j>k}c \mid b$
$X_{j<k} \to X_{j<k}c \mid bX_{j<k}c \mid c$
$A \to aA \mid \epsilon$
$C \to cC \mid \epsilon\}$

Erklärung: Von S aus wird in einen der vier Fälle $i < j, i > j, j < k, j > k$ "gesprungen": Aus $X_{i>j}$ lassen sich genau die Wörter ableiten, die die Form $a^i b^j$ mit $i > j$ haben; die Anzahl der c ist beliebig, und c^* beschreibt gerade die Wörter, die aus C ableitbar sind. Somit kann man aus $X_{i>j}C$ genau die Wörter $a^i b^j c^k$ mit $i > j$ ableiten.

Die anderen Fälle gehen entsprechend, und man erhält die Menge aller Wörter $a^i b^j c^k$ mit $i > j \vee i < j \vee j > k \vee j < k$, was gerade die gesuchte Sprache ergibt.

Ableitung: $S \Rightarrow AX_{j>k} \Rightarrow aAX_{j>k} \Rightarrow aaAX_{j>k} \Rightarrow aaX_{j>k} \Rightarrow aabX_{j>k}c \Rightarrow aabbc.$

Aufgabe 2:

a) Aus S kann man nur zwei X ableiten; sobald aus einem X ein Wort, bestehend aus Terminalsymbolen, abgeleitet wird, entsteht genau ein b. Somit enthält jedes $w \in L(G)$ genau 2 b.

b) **Fehler in der Aufgabenstellung:** Eigentlich sollte die Anzahl von a in w abgefragt werden.

   Nach dem ersten Ableitungsschritt liegen zwei X vor. In jedem weiteren Ableitungsschritt, in dem kein b erzeugt wird, werden zwei a erzeugt. Es gibt zwei Ableitungsschritte, in denen ein b erzeugt wird. Somit werden in $n-3$ Ableitungsschritten jeweils zwei a erzeugt, also enthält w $2(n-3)$ a.

c) $L(G) = \{a^n b a^{n+m} b a^m | n, m \in \mathbb{N}_0\}$

Aufgabe 3: Ohne Beschränkung der Allgemeinheit gelte $N_1 \cap N_2 = \emptyset$ (lässt sich notfalls durch Umbenennungen von Nichtterminalen erreichen).

a) Idee: Aus einem neuen Startsymbol lassen sich S_1 und S_2 ableiten. Dann lässt sich jedes Wort aus $L(G_1)$ und jedes Wort aus $L(G_2)$ ableiten, und dies sind die einzigen Wörter, die sich ableiten lassen. Somit gilt $L(G) = L(G_1) \cup L(G_2)$.

   Formal: Sei $S \notin N_1 \cup N_2$. Dann gilt für $G = (N_1 \cup N_2 \cup \{S\}, T_1 \cup T_2, S,$
   $P_1 \cup P_2 \cup \{S \to S_1, S \to S_2\})$ $L(G) = L(G_1) \cup L(G_2)$.

b) Idee: Das Startsymbol S_1 kann sich "vervielfachen" oder zu ϵ abgeleitet werden. Es können dann genau das leere Wort und Wörter, die sich aus Teilwörtern aus $L(G_1)$ zusammensetzen, abgeleitet werden, und es gilt $L(G) = L(G_1)^*$.

   Formal: $G = (N_1, T_1, S_1, P_1 \cup$
   $\{S_1 \to S_1 S_1, S_1 \to \epsilon\})$.

c) Sei $S \notin N_1 \cup N_2$. $G = (N_1 \cup N_2 \cup \{S\}, T_1 \cup T_2, S, P_1 \cup P_2 \cup \{S \to S_1 S_2\})$.

Aufgabe 4: Wir konstruieren Grammatiken G_1 und G_2 mit $L(G_1) = \{a^n b^n c^m : n \geq 0, m \geq 0\}$ und $L(G_2) = \{a^m b^n c^n : n \geq 0, m \geq 0\}$.
Sei $w \in \{a^n b^n c^n : n \geq 0\}$; dann gilt $w \in L(G_1)$ und $w \in L(G_2)$, also gilt $w \in L(G_1) \cap L(G_2)$.
Sei $w \in L(G_1) \cap L(G_2)$. Dann ist w von der Gestalt $a^i b^j c^k$ mit $i = j$ (da $w \in L(G_1)$) und $j = k$ (da $w \in L(G_2)$). Somit gilt $w \in \{a^n b^n c^n : n \geq 0\}$, und es folt $L(G_1) \cap L(G_2) = \{a^n b^n c^n : n \geq 0\}$.
$G_1 = (\{S, A, C\}, \{a, b, c\}, S,$
$\{S \to AC, A \to \epsilon \mid aAb, C \to \epsilon \mid cC\})$.
$G_2 = (\{S, A, C\}, \{a, b, c\}, S,$
$\{S \to AC, A \to \epsilon \mid aA, C \to \epsilon \mid bCc\})$.
Hinweis: Die Sprache $\{a^n b^n c^n : n \geq 0\}$ lässt sich nicht von einer kontextfreien Grammatik erzeugen.
(Beweisskizze: Angenommen, dem ist nicht so. Sei dann G eine Grammatik mit minimaler Anzahl von Nichtterminalen, die $\{a^n b^n c^n : n \geq 0\}$ erzeugt.
Sei k die größte Anzahl an bs, die es auf der rechten Seite einer Regel gibt. Man betrachte das letzte Nichtterminal X, das bei der Ableitung des Wortes $a^{k+1} b^{k+1} c^{k+1}$ ersetzt wird. Dann ändert sich bei der letzten Ersetzung die Anzahl der as, bs oder cs nicht. (Würde sich die Anzahl jedes Symboles ändern, müssten $k+1$ bs produziert werden, was im Widerspruch zur Maximalität von k steht).
Somit kann es keine verschiedenen Möglichkeiten geben für die Ersetzung von X geben, da die Gestalt des Wortes durch die Anzahl des Symbols, dessen Anzahl sich nicht ändert, festgelegt ist.
Ein Nichtterminal X, für das es nur eine Ersetzung w_X gibt, kann jedoch eliminiert werden, indem man auf jeder rechten Seite, in der X vorkommt, X durch w_X ersetzt. Dies steht aber in Widerspruch dazu, dass G eine minimale Menge an Nichtterminalen besitzt.
Somit kann $\{a^n b^n c^n : n \geq 0\}$ nicht von einer kontextfreien Grammatik erzeugt werden.)
Dies bedeutet, dass es im Allgemeinen nicht möglich ist, zu zwei kontextfreien Grammatiken G_1 und G_2 eine Grammatik G zu konstruieren, für die $L(G) = L(G_1) \cap L(G_2)$. Man sagt auch: Die Menge der Sprachen, die von kontextfreien Grammatiken erzeugt werden können, ist nicht abgeschlossen unter der Schnittmengenbildung.

Aufgabe 5:

a) $N = \{S\}, T = \{(,)\}$,
$P = \{S \to SS \mid (S) \mid \epsilon\}$.

b) Wir zeigen: Das Wort $w \in (N \cup T)^*$, das sich nach k Ableitungsschritten ergibt, erfüllt $N_((w) = N_)(w)$ und für jedes Präfix v von w gilt $N_((v) \geq N_)(v)$.

Wenn sich das Wort w in einem Ableitungsschritt ableiten lässt, muss $w = \epsilon$ oder $w = (S)$ oder $w = SS$ gelten, und die Bedingung ist erfüllt.

Wir betrachten das Wort w', das sich aus dem mit k Schritten abgeleiteten Wort w in einem Schritt ableiten lässt.

Falls die angewandte Produktion $S \to \epsilon$ oder $S \to SS$ ist, ändern sich die Werte von $N_((w), N_)(w), N_((v), N_)(v)$ für beliebige Präfixe v von w nicht.

Falls die angewandte Produktion $S \to (S)$ ist, so erhöht sich sowohl $N_((w')$ als auch $N_)(w')$ um 1 gegenüber den Werten für w, so dass die Werte wieder gleich sind.

Falls v ein Präfix von w' ist, das das durch die Ersetzung erzeugte Teilwort (S) nicht enthält, ist $N_((v) \geq N_)(v)$, da dieses Präfix dann auch Präfix von w ist.

Falls v ein Präfix von w' ist, das ein Präfix des durch die Ersetzung erzeugten Teilworts (S) enthält (zum Beispiel auch das ganze Teilwort), hat sich $N_((v)$ um mindestens 1 erhöht, $N_)(v)$ um höchstens 1 gegenüber dem entsprechenden Präfix v' in w, das S anstatt (S) enthält und für das die Bedingung $N_((v) \geq N_)(v)$ gilt.

Somit gilt auch in diesem Fall die geforderte Bedingung, und die Behauptung ist durch Induktion gezeigt.