Theoretische Informatik und Logik

3. Übung am 31.10.2007

Aufgabe 3.1

Sei $L = \left\{(ww|\, w \in \{\underline{0}, \underline{1}\}^*\right\}$. Zeigen Sie mit Hilfe des Pumping Lemmas für reguläre Sprachen, dass L nicht regulär ist.

Grundlagen

Pumping Lemma: Sei L eine unendliche reguläre Sprache. Es gibt eine Schranke $m \geq 0$, sodass jedes Wort in L mit $|w| \geq m$ geschrieben werden kann als $w = xyz$ mit $|xy| \leq m$ und $y \neq \varepsilon$, sodass $w_i = xy^iz$ ebefalls in L liegt für alle $i \geq 0$.

Lösung

Wählt man m für die Schranke aus dem Pumping Lemma (PL), und sei $v = \underline{0}^m\underline{10}^m\underline{1} \in L$, dann ist $|v| = 2m+2 \geq m$. Sei nun weiter xyz eine Zerlegung von v und $|xy| < m$. Dann ist $xy \in \{\underline{0}\}^+$, $y \neq \varepsilon$. Dem PL folgend, muß eine Zerlegung xyz für v existieren, sodass $|xy| < m$ und $y \neq \varepsilon$. Dann ist $xy \in \{\underline{0}\}^+$, und gemäss PL, $w_i = xy^iz \in L$ für alle $i \geq 0$. Also muß $\underline{0}^{m-|i|}\underline{10}^m$ in L sein.

Ist es aber nicht $\rightarrow L$ ist nicht regulär! $\square$

Aufgabe 3.2

Zeigen Sie mit Hilfe des Pumping Lemmas für reguläre Sprachen, dass die Sprache $L = \left\{\,\underline{a}^{2^n} \,\middle|\, n > 0\right\}$ nicht regulär ist.

Grundlagen

Pumping Lemma: Sei L eine unendliche reguläre Sprache. Es gibt eine Schranke $m \geq 0$, sodass jedes Wort in L mit $|w| \geq m$ geschrieben werden kann als $w = xyz$ mit $|xy| \leq m$ und $y \neq \varepsilon$, sodass $w_i = xy^iz$ ebefalls in L liegt für alle $i \geq 0$.

Lösung

Für $0 < m < 2$ lässt sich kein Wort $|w| \geq m$ finden, das sich in die Form xyz zerlegen lässt. Für $m > 2$, wählt man $w = a^{2^{m+1}}$, wobei die Bedingung $|w| \geq m$ erfüllt ist. Es wird nun das Symbol $\prod$ eingeführt, welches eine wiederholte Verkettung kennzeichnet:

$$\underbrace{\underline{b} \cdot \underline{b} \cdot \ldots \cdot \underline{b}}_{a} = \prod_{i=1}^{a} \underline{b}$$

Nun kann $a^{2^{i+1}}$ auch geschrieben werden als $a^{2^{i+1}} = a^2 \cdot \prod_1^n a^{2^n}$. Nun sei xyz eine Zerlegung von w mit $xy = a^2$, bzw. $x = a$, $y = a$ und $z = \prod_1^m a^{2^m}$. Dies erfüllt die Bedingung des PL, dass

$|xy| \leq m$, und $y \neq \varepsilon$. Gemäss PL ist nun $w_i = xy^i z \in L$. Für $i = 0$ ist jedoch $w_0 = a \cdot \prod_1^m a^{2^m} \notin L$. Folglich ist L nicht regulär!

Aufgabe 3.3

Sei $G = \langle \{A, B, C\}, \{\underline{a}, \underline{b}\}, \{A \Rightarrow B\underline{a}C, B \Rightarrow \underline{a}B \mid \varepsilon, C \Rightarrow \underline{a}C \mid \underline{b}C \mid \varepsilon\}, A\rangle$. Geben Sie eine reguläre Grammatik an, welche dieselbe Sprache wie G erzeugt.

Grundlagen

Grammatik: Eine Grammatik (unbeschränkte oder Typ-0-Grammatik) wird durch ein Quadrupel $\langle V, T, P, S\rangle$ beschrieben, wobei V das Alphabet der Variablen (Nonterminale), T das Alphabet der Terminalsymbole, P die Menge der Produktionen und $S \in V$ das Startsymbol ist. V und T müssen disjunkt sein: $V \cap T = \{\}$. Die Menge der Produktionen, P, ist eine Teilmenge von $(V \cup T)^+ \times (V \cup T)^*$, d.h., eine Produktion ist ein Paar (α, β), mit der Einschränkung $\alpha \neq \varepsilon$. Statt (α, β) wird überlicherweise $\alpha \Rightarrow \beta$ geschrieben. Die Abkürzung $\alpha \Rightarrow \beta_1 | \ldots | \beta_n$ steht für $\alpha \Rightarrow \beta_1, \ldots, \alpha \Rightarrow \beta_n$.

reguläre Grammatik: Eine Grammatik heißt regulär (oder Typ-3-Grammatik), wenn alle ihre Produktionen die Form $A \Rightarrow aB$ oder $A \Rightarrow \varepsilon$ besitzen, wobei A, B Nonterminale und α ein Terminalsymbol ist.

Lösung

$G:\quad V = \{A, B, C\} \quad T = \{\underline{a}, \underline{b}\} \quad S = \{A\} \quad P = \{A \Rightarrow B\underline{a}C, B \Rightarrow \underline{a}B \mid \varepsilon, C \Rightarrow \underline{a}C \mid \underline{b}C \mid \varepsilon\}$

Aus $A \Rightarrow B\underline{a}C$, $B \Rightarrow \underline{a}B \mid \varepsilon$ und $C \Rightarrow \underline{a}C | \varepsilon$ folgt: $A \Rightarrow B\underline{a}C \Leftrightarrow A \Rightarrow \underline{a}C$.

$G' = \langle \{A, B\}, \{\underline{a}, \underline{b}\}, \{A \Rightarrow \underline{a}C, C \Rightarrow \underline{a}C \mid \underline{b}C \mid \varepsilon\}, A\rangle$

Die akzeptierte Sprache ist dabei für G' und G: $L = \underline{a}\,\{\underline{a}, \underline{b}\}^*$. $\square$

Aufgabe 3.4

Sei $G = \langle \{S\}, \{\underline{a}, \underline{b}\}, \{S \Rightarrow \underline{a}S \mid \underline{a}S\underline{b} \mid \underline{ab}\}, S\rangle$. Zeigen Sie, dass G mehrdeutig ist, indem Sie für ein Wort w aus $\mathcal{L}$(G) (mit $|w| \geq 5$) zwei verschiedene Linksableitungen und die zugehörigen Ableitungsbäume angeben. Konstruieren Sie eine eindeutige kontexfreie Grammatik G' mit $\mathcal{L}(G') = \mathcal{L}(G)$ und geben Sie für das zuvor gewählte Wort aus $\mathcal{L}(G)$ die Linksableitung, die Rechtsableitung sowie die Parallelableitung in Ihrer Grammatik G' an.

Grundlagen

kontexfreie Grammatik: Eine Grammatik heißt kontexfrei (oder Type-2-Grammatik), wenn die linke Seite jeder Produktion ein einzelnes Nonterminalsymbol ist, d.h., wenn alle Produktionen die Gestalt $A \Rightarrow \beta$ besitzen, wobei $A \in V$ und $\beta \in (V \cup T)^*$. Eine Sprache L heißt kontexfrei, wenn es eine kontexfreie Grammatik G gibt, sodass $\mathcal{L}(G) = L$.

Mehrdeutigkeit: Eine Grammatik heißt mehrdeutig, wenn es ein Wort mit mehreren Linksableitungen gibt.

Linksableigung: $xAy \vdash_L x\beta y$ gilt, falls $A \Rightarrow \beta$ eine Produktion ist, und x nur aus Terminalsymbolen besteht ($y \in T^*$). y ist ein beliebiges Wort aus $(V \cup T)^*$.

Lösung

Für $w \in \mathcal{L}(G) = \underline{aaab}$ finden sich folgende (unterschiedliche) Linksableitungen:

1. $S \vdash_L \underline{a}S \vdash_L \underline{aa}S\underline{b} \vdash_L \underline{aaab}$
2. $S \vdash_L \underline{a}S\underline{b} \vdash_L \underline{aa}S\underline{b} \vdash_L \underline{aaab}$

Um G' zu finden, ersetzt $\underline{a}S$ durch $\underline{a}A \mid \varepsilon$ und erweitere S um $\underline{a}A\underline{b}$:

$$G' \left\langle \{S, A\}, \{\underline{a}, \underline{b}\}, \{S \Rightarrow \underline{a}S\underline{b} \mid \underline{a}A\underline{b}, A \Rightarrow \underline{a}A \mid \varepsilon\}, S \right\rangle$$

Ableitungen für $w \in \mathcal{L}(G')$:

1. $S \vdash_L \underline{a}A \vdash_L \underline{aa}S\underline{b} \vdash_L \underline{aaa}A\underline{b} \vdash_L \underline{aaab}$
2. $S \vdash_R \underline{a}A \vdash_R \underline{aa}S\underline{b} \vdash_R \underline{aaa}A\underline{b} \vdash_R \underline{aaab}$
3. $S \vdash_P \underline{a}A \vdash_P \underline{aa}S\underline{b} \vdash_P \underline{aaa}A\underline{b} \vdash_P \underline{aaab}$

Aufgabe 3.5

Sei Σ ein binäres Alphabet und $L = \{v \in \Sigma^* \mid |v| = 2n, n \geq 1, und\ v \neq ww, w \in \Sigma^*\}$. Geben Sie eine kontexfreie Grammatik an, die L erzeugt.

Ungelöst!