Fachbereich Mathematik
Prof. Dr. Ulrich Kohlenbach
Dr. Benno van den Berg

TECHNISCHE
UNIVERSITÄT
DARMSTADT

12. Januar, 2009

3. Hausaufgabe Formale Grundlagen der Informatik I
WS 2008/2009

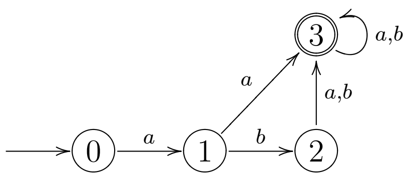
Abgabe: Fr 23. Januar, in der Übung

Geben Sie bitte ihren Namen und die Nummer ihrer Übungsgruppe an.

(H3.1) [Grammatiken]

Sei $\Sigma = \{a, b\}$.

(a) Betrachten Sie den Automat $\mathcal{A}$:



Geben Sie eine reguläre Grammatik an, die die Sprache $L(\mathcal{A})$ erzeugt.

(b) Sei

$$L = \{x \in \Sigma^* \,:\, |x|_a = |x|_b\}.$$

Zeigen Sie, dass die Sprache L kontextfrei ist, indem Sie eine Grammatik angeben, die diese Sprache erzeugt. (Und begründen Sie Ihre Antwort!)

Musterlösung.

(a)

$$\begin{array}{lcl} X_0 & \to & aX_1 \\ X_1 & \to & aX_3 \,|\, bX_2 \\ X_2 & \to & aX_3 \,|\, bX_3 \\ X_3 & \to & aX_3 \,|\, bX_3 \,|\, \varepsilon \end{array}$$

(b)

$$X_0 \rightarrow \varepsilon \mid X_0X_0 \mid aX_0b \mid bX_0a$$

(H3.2) [Chomsky-Normalform]

Betrachten Sie die kontextfreie Grammatik $G = (\{a,b\}, \{X_0\}, P, X_0)$ mit

$$\begin{array}{rlcl} P: & X_0 & \rightarrow & aXY \mid bXb \mid a \\ & X & \rightarrow & aXa \mid bY \mid \varepsilon \\ & Y & \rightarrow & bX_0a \mid aX_0 \end{array}$$

Konstruieren Sie eine zu G äquivalente Grammatik in Chomsky-Normalform.

Musterlösung.

1. Schritt (Variablen vor Buchstaben):

$$\begin{array}{lcl} X_0 & \rightarrow & Z_aXY \mid Z_bXZ_b \mid a \\ X & \rightarrow & Z_aXZ_a \mid Z_bY \mid \varepsilon \\ Y & \rightarrow & Z_bX_0Z_a \mid Z_aX_0 \\ Z_a & \rightarrow & a \\ Z_b & \rightarrow & b \end{array}$$

2. Schritt (elimiere ε-Produktionen):

$$\begin{array}{lcl} X_0 & \rightarrow & Z_aXY \mid Z_aY \mid Z_bXZ_b \mid Z_bZ_b \mid a \\ X & \rightarrow & Z_aXZ_a \mid Z_aZ_a \mid Z_bY \\ Y & \rightarrow & Z_bX_0Z_a \mid Z_aX_0 \\ Z_a & \rightarrow & a \\ Z_b & \rightarrow & b \end{array}$$

3. Schritt (eliminiere $X \rightarrow X_0 \ldots X_k$ mit $k \geqslant 3$):

$$\begin{array}{lcl} X_0 & \rightarrow & VY \mid Z_aY \mid WZ_b \mid Z_bZ_b \mid a \\ X & \rightarrow & VZ_a \mid Z_aZ_a \mid Z_bY \\ Y & \rightarrow & UZ_a \mid Z_aX_0 \\ V & \rightarrow & Z_aX \\ W & \rightarrow & Z_bX \\ U & \rightarrow & Z_bX_0 \\ Z_a & \rightarrow & a \\ Z_b & \rightarrow & b \end{array}$$

(H3.3) [Chomsky-Hierarchie]

Welche der folgenden Sprachen über dem Alphabet $\Sigma = \{a,b,c\}$ sind (i) regulär, (ii) kontexfrei, aber nicht regulär, oder (iii) nicht kontextfrei? Begründen Sie Ihre Antwort!

$$\begin{array}{lcl} L_1 & = & \{x \in \Sigma^* : |x|_a > |x|_b\} \\ L_2 & = & \{x \in \Sigma^* : |x|_a > |x|_b > |x|_c\} \\ L_3 & = & \{x \in \Sigma^* : |x|_a > |x|_b \text{ und } |x|_b \leqslant 2009\} \\ L_4 & = & \{x \in \Sigma^* : |x|_a > |x|_b \text{ und } |x|_b \geqslant 2009\} \end{array}$$

Musterlösung.

L_1: L_1 ist kontextfrei, aber nicht regulär. Eine Grammatik für L_1 wäre

$$P: \begin{array}{lcl} X_0 & \rightarrow & XX_0 \,|\, aX_0 \,|\, aX \\ X & \rightarrow & \varepsilon \,|\, XX \,|\, aXb \,|\, bXa \end{array}$$

L_1 ist nicht regulär. Angenommen L_1 wäre regulär. Sei n die Konstante aus dem Pumping Lemma und setze $x = a^{n+1}b^n$. Nach dem Lemma gibt es eine Zerlegung $x = uvw$ mit $|uv| \leqslant n, v \neq \varepsilon$ und $uv^m w \in L_1$ für alle $m \in \mathbb{N}$. Da $|uv| \leqslant n$ und $v \neq \varepsilon$, ist $v = a^k$ für $k \geqslant 1$. Also ist $uv^0w = a^{n+1-k}b^n \notin L_1$. Widerspruch!

L_2: L_2 ist nicht kontextfrei. Angenommen L_2 wäre kontextfrei. Sei n die Konstante aus dem Pumping Lemma und setze $x = a^{n+2}b^{n+1}c^n$. Nach dem Lemma gibt es eine Zerlegung $x = yuvwz$ mit $|uvw| \leqslant n, uw \neq \varepsilon$ und $yu^m vw^m z \in L_2$ für alle $m \in \mathbb{N}$. Da $|uvw| \leqslant n$ kann uvw nicht sowohl a's als auch c's enthalten. Deshalb gibt es nur die folgenden Möglichkeiten:

- u und w enthalten nur a. Dann enthält yu^0vw^0z nicht mehr a's als b's und ist deshalb nicht in L_2 enthalten. Widerspruch!
- u und w enthalten a und b. Dann enthält yu^0vw^0z nicht mehr b's als c's und ist deshalb nicht in L_2 enthalten. Widerspruch!
- u und w enthalten nur b. Dann enthält yu^0vw^0z nicht mehr b's als c's und ist deshalb nicht in L_2 enthalten. Widerspruch!
- u und w enthalten b und c. Dann enthält yu^2vw^2z nicht mehr a's als b's und ist deshalb nicht in L_2 enthalten. Widerspruch!
- u und w enthalten nur c. Dann enthalt yu^2vw^2z nicht mehr b's als c's und ist deshalb nicht in L_2 enthalten. Widerspruch!

L_3: L_3 ist regulär. Für jedes $n \in \mathbb{N}$ ist

$$L_3^n = \{x \in \Sigma^* \,:\, |x|_a > n \text{ und } |x|_b = n\} = \{x \in \Sigma^* \,:\, |x|_a > n\} \cap \{x \in \Sigma^* \,:\, |x|_b = n\}$$

ein Durchschnitt von zwei regulären Sprachen und deshalb regulär. Es folgt, dass $L_3 = \bigcup_{n \leqslant 2009} L_3^n$ auch regulär ist.

L_4: L_4 ist kontextfrei, aber nicht regulär. $L_4 = L_1 \cap \{x \in \Sigma^* \,:\, |x|_b \geqslant 2009\}$ ist Durchschnitt einer kontextfreien Sprache mit einer regulären Sprache und deshalb kontextfrei. Wäre L_4 auch noch regulär, dann wäre $L_1 = L_3 \cup L_4$ das auch. Wir haben aber oben gezeigt, dass L_1 nicht regulär ist.

(H3.4) [CYK Algorithmus]

Betrachten Sie die kontextfreie Sprache L über dem Alphabet $\Sigma = \{a, b\}$, die von der folgenden Grammatik in Chomsky-Normalform erzeugt wird:

$$\begin{array}{lcl} X_0 & \to & VZ_a \,|\, WZ_b \,|\, Z_aZ_a \,|\, Z_bZ_b \\ V & \to & Z_aX_0 \\ W & \to & Z_bX_0 \\ Z_a & \to & a \\ Z_b & \to & b \end{array}$$

(i) Beschreiben Sie L umgangssprachlich.

(ii) Wenden Sie den CYK Algoritmus an, um zu bestimmen ob $bbaa \in L$ oder $aabbaa \in L$.

Musterlösung.

(i) L ist die Menge der Palindrome gerader Länge.

(ii) Erzeugbare Teilworte der Länge 1:

$$\begin{array}{lcl} a & : & Z_a \\ b & : & Z_b \end{array}$$

Erzeugbare Teilworte der Länge 2:

$$\begin{array}{lcl} aa & : & X_0 \\ ab & : & - \\ bb & : & X_0 \\ ba & : & - \end{array}$$

Erzeugbare Teilworte der Länge 3:

$$\begin{array}{lcl} aab & : & - \\ abb & : & V \\ bba & : & - \\ bab & : & W \end{array}$$

Erzeugbare Teilworte der Länge 4:

$$\begin{array}{lcl} aabb & : & - \\ abba & : & X_0 \\ bbab & : & - \end{array}$$

Erzeugbare Teilworte der Länge 5:

$$\begin{array}{lcl} aabba & : & V \\ abbab & : & - \end{array}$$

Deshalb ist $bbaa$ nicht erzeugbar, und $aabbaa$ ist erzeugbar (mittels $X_0 \to VZ_a$).