Karlsruher Institut für Technologie
Institut für Theoretische Informatik

Prof. Dr. Peter Sanders
L. Hübschle-Schneider, T. Maier

1. Übungsblatt zu Theoretische Grundlagen der Informatik im WS 2015/16

http://algo2.iti.kit.edu/TGI2015.php
{sanders,huebschle,t.maier}@kit.edu

Musterlösungen

Aufgabe 1 (*Formale Sprachen,* $1+2+2$ *Punkte*)

a) Beschreiben Sie die kleensche Hülle von $L_1 = ab$ in mathematischer Schreibweise (kein Punkt-Punkt-Punkt).

b) Beschreiben Sie den Test, ob die Zeichen innerhalb einer Eingabe über dem Alphabet $\{a, b, c, d\}$ sortiert sind, als Entscheidungsproblem einer formellen Sprache ($a < b < c < d$).

c) Zeigen Sie, dass auf der Menge der Sprachen das Produkt zweier Sprachen eine assoziative Operation ist. Hat diese Operation ein neutrales Element (Einselement)? Hat sie ein Nullelement? (Ein Nullelement n erfüllt $nx = n = xn$.)

Musterlösung:

a) $\{(ab)^k \mid k \in \mathbb{N}_0\}$

b) $L_2 = \{a^i b^j c^k d^l \mid i, j, k, l \in \mathbb{N}_0\}$ die Sprache L_2 enthält alle sortierten Zeichenfolgen aus $\{a, b, c, d\}$s

c) z.Z.

(a) $(L_1 \cdot L_2) \cdot L_3 = L_1 \cdot (L_2 \cdot L_3)$ zur Übersichtlichkeit schreiben wir: $L_{ij} = L_i \cdot L_j$
Wir zeigen: $L_{(12)3} \subseteq L_{1(23)}$ (Rückrichtung analog)
gegeben $w \in L_{(12)3}$ existiert eine Zerlegung $w = w_{12}w_3 \mid w_{12} \in L_{12},\ w_3 \in L_3$. Da $w_{12} \in L_{12}$ existiert weiterhin eine Zerlegung $w_{12} = w_1 w_2 \mid w_1 \in L_1,\ w_2 \in L_2$. Es gilt, $w_2 w_3 \in L_{23}$ deshalb gilt auch $w_1 \cdot w_2 \cdot w_3 \in L_{1(23)}$.

(b) $\{\varepsilon\}$ ist neutrales Element.
Wir zeigen nur, dass $\{\varepsilon\} \cdot L = L$ (analog: $L \cdot \{\varepsilon\}$). 1. $L \subseteq \{\varepsilon\} \cdot L$ gegeben $w \in L$ das Wort $w = \varepsilon w \in \{\varepsilon\} \cdot L$. 2. $\{\varepsilon\} \cdot L \subseteq L$ jedes $w \in \{\varepsilon\} \cdot L$ lässt sich schreiben als $w = \varepsilon \cdot w_L \mid w_L \in L$ es gilt $w = w_L$ und deshalb $\{\varepsilon\} \cdot L \subseteq L$.

(c) $\{\}$ ist Nullelement.
Die Sprache $\{\} \cdot L$ ist definiert als die Menge aller Worte $w \mid w = w_1 w_2 \mid w_1 \in \{\} \wedge w_2 \in L$ ein solches w_1 kann es nicht geben, deshalb ist $\{\} \cdot L$ leer. Analog für die Multiplikation von rechts.

Aufgabe 2 (*Chomsky-3-Grammatiken,* $4+4$ *Punkte*)

a) Erstellen Sie je eine Chomsky-3-Grammatik für die folgenden formalen Sprachen über dem Terminalalphabet $\Sigma = \{a, b, n\}$ (Ihre Grammatiken sollten nicht mehr als 8 Variablen aufweisen).

(a) $L(G_1) = \{$*alle Wörter, die* ab *enthalten*$\}$

(b) $L(G_2) = \{a, b\}^5$ **hier stand ursprünglich $\{\{a, b\}^5\}$, wir bitten dies zu entschuldigen**

(c) $L(G_3) = L(G_1)^c = \{\textit{alle Wörter, die } \textbf{nicht} \text{ ab } \textit{enthalten}\}$

(d) $L(G_4) = \{\textit{Alle Teilworte des Worts } \text{banana}\}$

b) Geben Sie für die folgenden Grammatiken äquivalente Chomsky-3-Grammatiken an.

(a) $G_5 = (\{S, A, B, C\}, \Sigma, P_5, S)$

$$P_5 = \{S \to B \mid AB \mid BC,\\ A \to a \mid Aa,\\ B \to b \mid BB,\\ C \to nAC \mid \varepsilon\}$$

(b) $G_6 = (\{S, A, B, C\}, \Sigma, P_6, S)$

$$P_6 = \{S \to A \mid B \mid C,\\ A \to AB \mid a,\\ B \to BC \mid b,\\ C \to CC \mid n \mid \varepsilon\}$$

Musterlösung:
Terminal- und Nichtterminalalphabete seien implizit gegeben durch die Produktionsmenge. Vorsicht!: meine Grammatiken sind zum Teil stark minimiert, und natürlich nicht die einzigen Lösungen. Wiederholte kleine Fehler können zu ganzzahligem Punktabzug akkumuliert werden.

$$P_1 = \{S \to aS \mid bS \mid nS \mid aA,\\ A \to bB \mid b,\\ B \to aB \mid bB \mid nB \mid a \mid b \mid n\}$$

$$P_2 = \{S \to aA \mid bA,\\ A \to aB \mid bB,\\ B \to aC \mid bC,\\ C \to aD \mid bD,\\ D \to a \mid b\}$$

$$P_3 = \{S \to bB \mid nB \mid aA \mid a \mid b \mid n \mid \varepsilon,\\ A \to aA \mid nB \mid a,\\ B \to bB \mid nB \mid aA \mid b \mid a\}$$

$$P_4 = \{S \to bA_1 \mid aN_1 \mid nA_2 \mid aN_2 \mid nA_3 \mid\\ a \mid b \mid n \mid \varepsilon,\\ A_1 \to aN_1 \mid a,\\ N_1 \to nA_2 \mid n,\\ A_2 \to aN_2 \mid a,\\ N_2 \to nA_3 \mid n,\\ A_3 \to a\}$$

Korrekturtipp (für b): Falls die erzeugte Grammatik nur einige Worte der Sprache nicht erzeugt (potentielle Randfälle wie ε) oder ungewünscht erzeugt ein Punkt Abzug. Nicht Chomsky 3 oder ∞ falsche Worte -2P. $L(G_5) = b+ \mid a+b+ \mid b+(na+)*$

$$P_5' = \{S \to bB \mid aA_B \mid bB_C \mid b,\\ A \to aA \mid a,\\ B \to bB \mid b,\\ A_B \to aA_B \mid bB \mid b,\\ B_C \to bB_C \mid nC \mid b \mid n,\\ C \to nA_C \mid nA,\\ A_C \to aA_C \mid aC \mid a\}$$

$L(G_6) = a(bn*)* \mid bn* \mid n*$

$$P_6' = \{S \to aB \mid bC \mid nC \mid a \mid b \mid n \mid \varepsilon,\\ B \to bC_B \mid b,\\ C \to nC \mid n,\\ C_B \to bC_B \mid nC_B \mid b \mid n\}$$

Aufgabe 3 (*Induktion über Ableitungsschritte, $2 + 5$ Punkte*)

$G_7 = (\{S, A\}, \{a, b\}, P_7, S)$

$$P_7 = \{S \to A,\\ A \to aAb \mid bAa \mid AA \mid a \mid \varepsilon\}$$

a) Geben Sie 3 unterschiedliche Ableitungswege für das Wort *aaaabbaab* an.

b) Beweisen Sie induktiv, dass alle Worte der Sprache $L(G_7)$ mindestens so viele as wie bs enthalten.

Musterlösung:

a) z.B.
$S \Rightarrow A \Rightarrow AA \Rightarrow aAbA \Rightarrow aaAbbA \Rightarrow aaAbbaAb \Rightarrow aaAAbbAb \Rightarrow aaaAbbAb \Rightarrow aaaabbAAb \Rightarrow aaaabbaab$
$S \Rightarrow A \Rightarrow aAb \Rightarrow aAAb \Rightarrow aaAbAb \Rightarrow aaAbbAab \Rightarrow aaAbbaab \Rightarrow aaAAbbaab \Rightarrow aaaAbbaab \Rightarrow aaaabbaab$
$S \Rightarrow A \Rightarrow AA \Rightarrow AaAb \Rightarrow aAbaAb \Rightarrow aAbaab \Rightarrow aaAbbaab \Rightarrow aaAAbbaab \Rightarrow aaAabbaab \Rightarrow aaaabbaab$
... es gibt noch weitere Ableitungswege

b) z.Z.: $A \Rightarrow^* w \quad \rightarrow \quad n_a(w) \geq n_b(w)$.
IA: $A \Rightarrow^1 w$ demnach: $w = \varepsilon$ oder $w = a$; beide enthalten mindestens gleich viele as wie bs
IV: Für ein beliebiges aber festes n und alle $k \leq n$ gelte $A \Rightarrow^k w \ \rightarrow \ n_a(w) \geq n_b(w)$
IS: Gegeben w mit $A \Rightarrow^{n+1} w$, betrachte den ersten Ableitungsschritt:

- 1. Fall: $A \Rightarrow aAb \Rightarrow^n aw'b = w$: es gilt $n_a(w) = n_a(w') + 1 \geq n_b(w') + 1 = n_b(w)$ nach IV ($n_a(w') \geq n_b(w')$).
- 2. Fall: $A \Rightarrow bAa \Rightarrow^n bw'a = w$: analog
- 3. Fall: $A \Rightarrow AA \Rightarrow^n w$: OBdA können wir davon ausgehen, dass immer das linkeste Nichtterminal abgeleitet wird. Deshalb gilt $A \Rightarrow AA \Rightarrow^i w_1 A \Rightarrow^j w_1 w_2$ mit $i \leq n$ und $j \leq n$. Mit Hilfe der IV folgt: $n_a(w) = n_a(w_1) + n_a(w_2) \geq n_b(w_1) + n_b(w_2) = n_b(w)$
- Alle anderen Ableitungsregeln terminieren die Ableitung, und können deshalb nicht als erstes angewendet werden, es sei denn $w = \varepsilon$ oder $w = a$. Diese Fälle sind von der Induktionsannahme abgedeckt.