Übungsserie 5 zu Automaten und formalen Sprachen — Uni Leipzig
abzugeben am 13. 1. vor der Vorlesung (autotool: bis 15 Uhr) — Inst. f. Informatik

(schriftlich) Beweisen Sie: Die Grammatik G mit Regelmenge $\{S \to \varepsilon, S \to aSbS, S \to bSaS\}$ erzeugt die Sprache $E = \{w \mid w \in \{a,b\}^*, |w|_a = |w|_b\}$.

- Zeigen Sie durch Induktion nach Länge der Ableitung, daß für jede aus S erzeugbare Satzform $u \in \{a,b,S\}^*$ gilt: $|u|_a = |u|_b$. $\boxed{\text{1 P.}}$

  Induktionsanfang, d. h. Ableitungen mit 0 Schritten: von dieser Länge gibt es nur eine Ableitung, nämlich $S \to^0 S$, die erzeugt also die Satzform S, für die gilt die Behauptung, weil $h(S) = |S|_a = |S|_b = 0$.

  Induktionsschritt: bekannt ist, daß $\forall n' \le n : \forall u : S \to^n u \Rightarrow h(u) = 0$.
  Zu zeigen ist $\forall u : S \to^{n+1} u \Rightarrow h(u) = 0$,

  Für jedes u mit $S \to^{n+1} u$ gilt: $\exists v : S \to^n v \to u$. Nach Annahme ist $h(v) = 0$. Wie entsteht u aus v? Fallunterscheidung:

  - durch Regel 1: dann $\exists l, r : v = lSr, u = lr$, also $h(u) = h(lr) = h(lSr) = h(v) = 0$.
  - durch Regel 2: dann $\exists l, r : v = lSr, u = laSbSr$, also $h(u) = h(laSbSr) = h(lSr) + h(a) + h(b) + h(S) = h(lSr) + 1 - 1 + 0 = h(lSr) = h(v) = 0$.
  - durch Regel 3: dann $\exists l, r : v = lSr, u = lbSaSr$, also $h(u) = h(lbSaSr) = h(lSr) + h(b) + h(a) + h(S) = h(lSr) - 1 + 1 + 0 = h(lSr) = h(v) = 0$.

- Wieso folgt daraus $L(G) \subseteq E$? $\boxed{\text{1 P.}}$

  Wenn $w \in L(G)$, dann $S \to^* w$. Nach der vorigen Rechnung gilt $h(w) = 0$, also $w \in E$. (So simpel ist es. Keine Geheimnisse, keine Tricks. Ein Punkt für das Hinschreiben einer einfachen Wahrheit.)

- Zeigen Sie: jedes nicht leere Wort $w \in E$ besitzt eine Zerlegung $w = aubv$ oder $w = buav$ mit $u \in E$. $\boxed{\text{1 P.}}$

  Wegen $w \neq \varepsilon$ ist die folgende Fallunterscheidung vollständig:

  - w beginnt mit a.

    Für $w = w_1 w_2 \dots w_k$ mit $w_i \in \{a,b\}$ betrachte die Folge der Höhen
    $[f_1, f_2, \dots, f_k] = [h(w_1), h(w_1 w_2), \dots, h(w_1 w_2 \dots w_k)]$.

    Wir betrachten die Menge $I = \{i \mid f_i = 0\}$ der Nullstellen von f. Wegen $w \in E \Rightarrow h(w) = 0 \Rightarrow f_k = 0$ ist $I \neq \emptyset$. Also besitzt I ein Minimum $i = \min I$.

    Behauptung: $i > 1$. Beweis: aus $w_1 = a$ folgt $f_1 = 1$, also ist $1 \notin I$.

    Behauptung: $w_i = b$. Weil i die kleinste Nullstelle ist, und $f_1 > 0$, gilt $\forall 1 \le j < i : f_j > 0$, also insbesondere $f_{i-1} > 0$. Wegen $f_i = 0$ muß aber der Buchstabe auf Position i ein b sein.

    Behauptung: $u = w_2 \dots w_{i-1}, v = w_{i+1} \dots w_k$ erfüllen $w = aubv$ und $h(u) = 0$. Gilt nach Konstruktion.

  - w beginnt mit b: Analog. Deswegene gibt es auch nur einen Punkt.

- Was folgt daraus für v? $\boxed{\text{1 P.}}$

  $w \in E \Rightarrow 0 = h(w) = h(aubv) = h(a) + h(u) + h(b) + h(v) = 1 + 0 - 1 + h(v)$ deswegen $h(v) = 0$, also $v \in E$.

- Beweisen Sie damit $E \subseteq L(G)$. $\boxed{\text{1 P.}}$

  Zu zeigen ist: $\forall w \in E : w \in L(G)$. Es ist also zu begründen, daß jedes Wort $w \in E$ tatsächlich in G aus S ableitbar ist.

  Durch Induktion nach der Länge von w:

  - $|w| = 0$. Dann $S \to \varepsilon = w \in L(G)$.
  - $|w| > 0$. Dann $w \neq \varepsilon$. Dann $\exists u, v \in E : w = aubv \vee w = buav$. Die Aussage $u, v \in E$ ist wesentlich (und wurde in den vorigen Schritten gezeigt).

    Es gilt $|u| < |w|$ und $|v| < |w|$, nach Induktion gilt also $u \in L(G)$ und $v \in L(G)$, d. h. es gibt Ableitungen $S \to^* u, S \to^* v$. Dann gibt es auch eine Ableitung $S \to aSbS \to^* aubv = w$ bzw. eine Ableitung $S \to bSaS \to^* buav = w$. In beiden Fällen folgt $w \in L(G)$.

- Erzeugt G' mit Regelmenge $\{S \to \varepsilon, S \to aSbS, S \to SbSa\}$ auch die Sprache E? $\boxed{\text{1 P.}}$

  Nein. Für jedes $w \neq \varepsilon$ mit $w \in L(G)$ gilt: w beginnt oder endet mit a. Beweis: falls es eine Ableitung $S \to^* w \neq \varepsilon$ gibt, dann ist deren erster Schritt entweder $S \to aSbS$ oder $S \to SbSa$, und die am Anfang bzw. Ende erzeugten a bleiben bei jedem weiteren Schritt stehen.

  Deswegen kann G' keine Wörter der Form $b\Sigma^*b$ erzeugen, solche sind aber in E, zum Beispiel $baab$.

Hinweis: benutzen Sie die Funktion $h : \{a, b\}^* \to \mathbb{N}$ (Höhe) definiert durch $h(w) = |w|_a - |w|_b$.