Summer 1, 2006 CIS 511

Introduction to the Theory of Computation
Jean Gallier

Homework 4

June 13, 2006; Due June 22, 2006

"A problems" are for practice only, and should not be turned in.

Problem A1. Give constructions proving that for any PDA M, the acceptance modes $T(M)$, $N(M)$, and $L(M)$, are equivalent.

Problem A2. (i) Given any set X, for any two subsets $A, B \subseteq X$, prove that $A \subseteq B$ iff $A \cap \overline{B} = \emptyset$, where $\overline{B}$ denotes the complement of B in X.

(ii) Sketch an algorithm to test whether $L(G) \subseteq L(D)$, where G is a context-free grammar and D is a DFA.

"B problems" must be turned in.

Problem B1 (60 pts). A context-free grammar, $G = (V, \Sigma, P, S)$, is *linear* iff for every production $(A \rightarrow \alpha) \in P$,

$$\alpha \in \Sigma^* N \Sigma^* \cup \Sigma^*,$$

where $N = V - \Sigma$. A language, L, is a *linear context-free language* iff there is some linear context-free grammar, G, such that $L = L(G)$.

(a) Recall that for any string, $w \in \Sigma^*$, the string w^R denotes the reversal of the string w, i.e., the string w written in reverse order ($\epsilon^R = \epsilon$). Prove that the language $L_0 = \{ww^R \mid w \in \{a, b\}^*\}$ is linear context-free.

(b) Given any regular language, $L \subseteq \Sigma^*$, prove that $L^R = \{w^R \mid w \in L\}$ is also regular.

(c) Let $G = (V, \Sigma, P, S)$ be a *linear context-free* grammar. Assume that the set of productions P contains p productions and that it is ordered in some fashion. Each production is named p_i, with $1 \leq i \leq p$. Let $\Delta = \{\delta_1, \ldots, \delta_p\}$ be a set in one-to-one correspondence with the set of productions P and assume that Δ is disjoint from V. The language $R \subseteq \Delta^*$ is defined as follows:

$$R = \{\delta_{i_1} \cdots \delta_{i_n} \mid n \geq 1, \quad \text{there is a derivation } S \Longrightarrow \alpha_1 \Longrightarrow \cdots \Longrightarrow \alpha_n, \\ \text{such that} \quad \alpha_n \in \Sigma^* \quad \text{and the production applied at the } k\text{-th step is } p_{i_k}\}.$$

Let $h_1, h_2 \colon \Delta^* \rightarrow \Sigma^*$ be the homomorphisms defined as follows: For any $p_i \in P$,

(1) If p_i is of the form $A \to uBv$, where $A, B \in N$, and $u, v \in \Sigma^*$, then $h_1(\delta_i) = u$ and $h_2(\delta_i) = v$;

(2) If p_i is of the form $A \to w$, where $A \in N$ and $w \in \Sigma^*$, then choose any $u, v \in \Sigma^*$ such that $w = uv$, and let $h_1(\delta_i) = u$ and $h_2(\delta_i) = v$.

Prove that for any k, with $1 \leq k \leq n - 1$,

$$\alpha_k = h_1(\delta_{i_1} \cdots \delta_{i_k}) B h_2(\delta_{i_k} \cdots \delta_{i_1}),$$

for some $B \in N$, and that

$$\alpha_n = h_1(\delta_{i_1} \cdots \delta_{i_n}) h_2(\delta_{i_n} \cdots \delta_{i_1}).$$

Conclude that $L(G) = \{h_1(w)h_2(w^R) \mid w \in R\}$, for some language R over Δ.

(d) Prove that the language R defined in (c) is regular.

(e) Given a linear context-free language, $L = L(G)$, from questions (c) and (d), we know that there is a regular language R and two homomorphisms h_1, h_2, such that

$$L = \{h_1(w)h_2(w^R) \mid w \in R\}.$$

Let $\Omega = \{\omega_1, \ldots, \omega_p\}$ be a set in one-to-one correspondence with Δ and such that Ω is disjoint from V and Δ. Let $f \colon \Delta^* \to \Omega^*$ be the homomorphism determined by defining $f(\delta_i) = \omega_i$, for all i, with $1 \leq i \leq p$. Let S be the regular language

$$S = \{f(w)^R \mid w \in R\} = f(R)^R,$$

where R is the regular set of question (c). Also, let $g_1 \colon (\Delta \cup \Omega)^* \to \Sigma^*$ be the homomorphism determined by defining

$$g_1(\delta_i) = h_1(\delta_i) \quad \text{and} \quad g_1(\omega_i) = h_2(\delta_i).$$

Finally, let $g_2 \colon (\Delta \cup \Omega)^* \to \{a, b\}^*$ be the homomorphism determined by defining

$$g_2(\delta_i) = a^i b \quad \text{and} \quad g_2(\omega_i) = ba^i,$$

for all i, with $1 \leq i \leq p$.

Prove that

$$g_2^{-1}(L_0) \cap RS = \{w_1 w_2 \mid w_1 \in R, \quad w_2 \in S, \quad \text{and} \quad f(w_1) = w_2^R\},$$

where L_0 is the language of question (a), R is defined in question (c), and S is defined above in (e). From this, prove that

$$L = g_1(g_2^{-1}(L_0) \cap RS).$$

Hence, prove that every linear context-free language can be obtained from the language $L_0 = \{ww^R \mid w \in \{a,b\}^*\}$ by homomorphism, inverse homomorphism, and intersection with regular languages.

Remark: In fact, it can be shown that the family of linear context-free languages is the least class of languages with these properties.

Problem B2 (50 pts). Let $L \subseteq \{a\}^*$ be a context-free language. Prove that L is actually a regular language. Proceed as follows: If L is finite, this is obvious, thus, assume that L is infinite. Let $L = L(G)$, for some CFG G.

(i) Let $K > 1$ be the constant of the pumping lemma for G, and let $r = K!$. Prove the following fact: for every $w \in L$, if $|w| \geq K$, then

$$\{wa^{rn} \mid n \geq 0\} \subseteq L.$$

(ii) For every i such that $0 \leq i < r$, let

$$L_i = \{a^n \mid a^n \in L,\, n \geq K,\, n \equiv i \bmod r\}.$$

Clearly,

$$L = \{a^n \mid a^n \in L,\, n < K\} \cup \bigcup_{i=0}^{r-1} L_i.$$

If $L_i \neq \emptyset$, let z_i be the shortest string in L_i. Prove that

$$L_i = \{z_i a^{rm} \mid m \geq 0\}.$$

Conclude that L is regular.

(iii) Prove that it is decidable whether $L_i = \emptyset$ (i.e., describe (concisely) an algorithm).

(iv) Given a context-free language L over $\{a,b\}$, prove that it is decidable whether $\{a\}^* \subseteq L$ (i.e., describe (concisely) an algorithm).

Problem B3 (60 pts). Give pushdown automata for the following languages:

(a) $L_4 = \{wcw^R \mid w \in \{a,b\}^*\}$ (w^R denotes the reversal of w)

(b) $L_5 = \{ww^R \mid w \in \{a,b\}^*\}$

(c) $L_6 = \{a^m b^n \mid 1 \leq m \leq n \leq 3m\}$

(d) $L_7 = \{a^n c b^n \mid n \geq 1\} \cup \{a^n d b^{2n} \mid n \geq 1\}$

(e) $L_8 = \{a^{2m} b^n a^{2m} b^p \mid m,n,p \geq 1\} \cup \{a^m b^{3n} a^p b^{3n} \mid m,n,p \geq 1\}$

(f) $L_9 = \{xcy \mid |x| = |y|,\, x,y \in \{a,b\}^*\}$

Whenever possible, give a DPDA. In each case, give a brief justification of the fact that your PDA generates the desired language.

Problem B4 (40 pts). Use the pumping lemma (or Ogden's lemma) to show that the following languages are not context-free:

$L_1 = \{a^m b^n c^p \mid 1 \leq m < n < p\}$

$L_2 = \{a^n b^n c^p \mid n, p \geq 1, p \neq n\}$

Problem B5 (30 pts). Given $\Sigma = \{a\}$, give a Turing machine accepting

$$L = \{a^{2^n} \mid n \geq 1\}.$$

Problem B6 (30 pts). *Ackermann's function* A is defined recursively as follows:

$$\begin{aligned} A(0,\, y) &= y + 1, \\ A(x+1,\, 0) &= A(x,\, 1), \\ A(x+1,\, y+1) &= A(x,\, A(x+1, y)). \end{aligned}$$

Prove that

$$\begin{aligned} A(0,\, x) &= x + 1, \\ A(1,\, x) &= x + 2, \\ A(2,\, x) &= 2x + 3, \\ A(3,\, x) &= 2^{x+3} - 3, \end{aligned}$$

and

$$A(4, x) = \left.2^{2^{\cdot^{\cdot^{\cdot^{2^{16}}}}}}\right\}^{x} - 3,$$

with $A(4, 0) = 16 - 3 = 13$. Equivalently (and perhaps less confusing)

$$A(4, x) = \left.2^{2^{\cdot^{\cdot^{\cdot^{2^{2}}}}}}\right\}^{x+3} - 3.$$

TOTAL: 270 points.