

Introduction to the Theory of Computation

Homework 5

March 31, 2009; Due April 14, beginning of class

“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.

Problem A3. Prove that if $L \subseteq \Sigma^*$ is a context-free language, then its reversal, $L^R = \{w^R \in \Sigma^* \mid w \in L\}$, is also context-free.

“B problems” must be turned in.

Problem B1 (40 pts). Use the pumping lemma (or Ogden’s lemma) to show that the following languages are not context-free:

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

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

Hint. For L_1 , consider $a^m b^n c^p$ with m large, and let the a ’s be distinguished. For L_2 , let k be the integer of Ogden’s lemma. Let $p = k!$, and consider $a^{2p} b^{2p} c^p$, with the c ’s distinguished.

Problem B2 (20 pts). Given any two alphabets Σ, Δ , a *substitution* is a function $\tau: \Sigma \rightarrow 2^{\Delta^*}$ assigning some language $\tau(a) \subseteq \Delta^*$ to every symbol $a \in \Sigma$. A substitution $\tau: \Sigma \rightarrow 2^{\Delta^*}$ is extended to a map $\tau: 2^{\Sigma^*} \rightarrow 2^{\Delta^*}$ by first extending τ to strings using the following definition

$$\begin{aligned}\tau(\epsilon) &= \{\epsilon\}, \\ \tau(ua) &= \tau(u)\tau(a),\end{aligned}$$

where $u \in \Sigma^*$ and $a \in \Sigma$, and then to languages by letting

$$\tau(L) = \bigcup_{w \in L} \tau(w),$$

for every language $L \subseteq \Sigma^*$.

For example, let $\tau: \Sigma \rightarrow 2^{\Sigma^*}$ be the substitution defined such that $\tau(a) = \{\epsilon, a\}$ for every $a \in \Sigma$. Explain (in words) what $\tau(L)$ is.

In general, prove that if L is a context-free language and if $\tau(a)$ is a context-free language for every $a \in \Sigma$, then $\tau(L)$ is also a context-free language.

Problem B3 (50 pts). Let $h: \Sigma^* \rightarrow \Delta^*$ be a homomorphism, and let $L \subseteq \Delta^*$. Assume that $\epsilon \notin L$.

(i) Define $\Omega, \Gamma \subseteq \Sigma$ as follows:

$$\begin{aligned}\Omega &= \{a \in \Sigma \mid h(a) \neq \epsilon\}, \\ \Gamma &= \{a \in \Sigma \mid h(a) = \epsilon\}.\end{aligned}$$

Let $\bar{\Omega}$ be the new set of symbols

$$\bar{\Omega} = \{\bar{a} \mid a \in \Omega\},$$

and let E be a new symbol.

Let τ_1 be the substitution on Δ defined such that

$$\tau_1(a) = (\bar{\Omega} \cup \{E\})^* \{a\} (\bar{\Omega} \cup \{E\})^*,$$

for all $a \in \Delta$, and let

$$R = (\{\bar{a}h(a) \mid a \in \Omega\} \cup \{E\})^*$$

and

$$L_2 = \tau_1(L) \cap R.$$

Prove that

$$L_2 \cap (\bar{\Omega} \cup \{h(a) \mid a \in \Omega\})^* = \{\bar{a}_1 h(a_1) \bar{a}_2 h(a_2) \cdots \bar{a}_n h(a_n) \mid a_i \in \Omega, h(a_1 a_2 \cdots a_n) \in L\}.$$

(ii) Let $g: (\Delta \cup \bar{\Omega} \cup \{E\})^* \rightarrow (\Sigma \cup \{E\})^*$ be the homomorphism defined such that

$$\begin{aligned}g(a) &= \epsilon & \text{if } a \in \Delta, \\ g(\bar{a}) &= a & \text{if } a \in \Omega, \\ g(E) &= E,\end{aligned}$$

let

$$L_3 = g(L_2),$$

and let τ_2 be the substitution on $\Sigma \cup \{E\}$ defined such that

$$\begin{aligned}\tau_2(a) &= \{a\} & \text{if } a \in \Sigma, \\ \tau_2(E) &= \Gamma^+.\end{aligned}$$

Observe that for every $w \in L$, if $h^{-1}(w) \neq \emptyset$, then there is some $y \in L_3$ such that $h(y) = w$, and that every $y \in L_3$ is in $h^{-1}(L)$ after the occurrences of E have been erased.

Let \mathcal{L} be a family of languages that is closed under substitution by regular languages, intersection with regular languages, and union with regular languages. Prove that for every $L \in \mathcal{L}$, if $\epsilon \notin L$, then

$$\tau_2(g(\tau_1(L) \cap R)) = h^{-1}(L).$$

Prove that if $\epsilon \in L$, then

$$h^{-1}(L) = \tau_2(g(\tau_1(L - \{\epsilon\}) \cap R)) \cup \Gamma^*.$$

Conclude that \mathcal{L} is closed under inverse homomorphisms.

(iii) Prove that if L is context-free, then so is

$$h^{-1}(L) = \{w \in \Sigma^* \mid h(w) \in L\}.$$

Do not give a proof based on PDA's!

Problem B4 (50 pts). Give pushdown automata for the following languages:

- (a) $L_5 = \{ww^R \mid w \in \{a, b\}^*\}$ (w^R denotes the reversal of w)
- (b) $L_6 = \{a^m b^n \mid 1 \leq m \leq n \leq 3m\}$
- (c) $L_7 = \{a^n b^n \mid n \geq 1\} \cup \{a^{2n} b^n \mid n \geq 1\}$
- (d) $L_8 = \{a^{2m} b^n a^{2m} b^p \mid m, n, p \geq 1\} \cup \{a^m b^{3n} a^p b^{3n} \mid m, n, p \geq 1\}$
- (e) $L_9 = \{xycy \mid |x| = |y|, x, y \in \{a, b\}^*\}$

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

Problem B5 (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

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

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 $L_i = \emptyset$.

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

Problem B6 (30 pts). (1) Given the alphabet $\Sigma_2 = \{a, b, \bar{a}, \bar{b}\}$, define the relation \simeq on Σ_2^* as follows: For all $u, v \in \Sigma_2^*$,

$$u \simeq v \quad \text{iff} \quad \exists x, y \in \Sigma_2^*, \quad u = xa\bar{a}y, \quad v = xy \quad \text{or} \quad u = x\bar{b}b\bar{b}y, \quad v = xy.$$

Let \simeq^* be the reflexive and transitive closure of \simeq , and let $D_2 = \{w \in \Sigma_2^* \mid w \simeq^* \epsilon\}$. Give a context-free grammar for D_2 , and justify your answer.

Note: Strings such as $a\bar{a}b\bar{b}$ and $a\bar{b}b\bar{a}$ are in D_2 .

(2) Given the alphabet $\Sigma_m = \{a_1, \dots, a_m, \bar{a}_1, \dots, \bar{a}_m\}$, define the relation \simeq on Σ_m^* as follows: For all $u, v \in \Sigma_m^*$,

$$u \simeq v \quad \text{iff} \quad \exists x, y \in \Sigma_m^*, \quad u = xa_i\bar{a}_i y, \quad v = xy, \quad \text{for some } i, 1 \leq i \leq m.$$

Let \simeq^* be the reflexive and transitive closure of \simeq , and let $D_m = \{w \in \Sigma_m^* \mid w \simeq^* \epsilon\}$. Give a context-free grammar for D_m , and prove that your grammar is correct *rigorously*.

Note: D_m is known as the *Dyck set* on m letters.

TOTAL: 240 points.