Spring, 2013 CIS 511

Introduction to the Theory of Computation
Jean Gallier

Homework 5

March 21, 2013; Due April April 9, 2013

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

Problem A1. Prove that the context-free languages are closed under concatenation.

Problem A2. Given a homomorphism $h\colon \Sigma^* \to \Delta^*$, for any any context-free language $L \subseteq \Delta^*$, prove that $h(L)$ is 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 (50 pts). Given a context-free language L and a regular language R, prove that $L \cap R$ is context-free.

Do not use PDA's to solve this problem!

Hint. Without loss of generality, assume that $L = L(G)$, where $G = (V, \Sigma, P, S)$ is in Chomsky normal form, and let $R = L(D)$, for some DFA $D = (Q, \Sigma, \delta, q_0, F)$. Use a kind of cross-product construction as sketched below. Construct a CFG G_2 whose set of nonterminals is $Q \times N \times Q\ \cup \{S_0\}$, where S_0 is a new nonterminal, and whose productions are of the form:

$$S_0 \to (q_0, S, f),$$

for every $f \in F$;

$$(p, A, \delta(p, a)) \to a \quad \text{iff} \quad (A \to a) \in P,$$

for all $a \in \Sigma$, all $A \in N$, and all $p \in Q$;

$$(p, A, s) \to (p, B, q)(q, C, s) \quad \text{iff} \quad (A \to BC) \in P,$$

for all $p, q, s \in Q$ and all $A, B, C \in N$;

$$S_0 \to \epsilon \quad \text{iff} \quad (S \to \epsilon) \in P \text{ and } q_0 \in F.$$

Prove that for all $p, q \in Q$, all $A \in N$, all $w \in \Sigma^+$, and all $n \geq 1$,

$$(p, A, q) \underset{lm}{\overset{n}{\Longrightarrow}}_{G_2} w \quad \text{iff} \quad A \underset{lm}{\overset{n}{\Longrightarrow}}_{G} w \quad \text{and} \quad \delta^*(p, w) = q.$$

Conclude that $L(G_2) = L \cap R$.

Problem B3 (60 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 $L_i = \emptyset$.

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

Problem B4 (40 pts). (1) Given the alphabet $\Sigma_2 = \{a, b, \overline{a}, \overline{b}\}$, define the relation $\simeq$ on Σ_2^* as follows: For all $u, v \in \Sigma_2^*$,

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

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 $a\overline{a}b\overline{b}$ and $ab\overline{b}\overline{a}$ are in D_2.

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

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

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 justify your answer.

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

Problem B5 (40 pts). (1) Prove that the function, $f \colon \Sigma^* \to \Sigma^*$, given by

$$f(w) = w^R$$

is RAM computable. ($\Sigma = \{a_1, \ldots, a_N\}$).

(2) Prove that the function, $f \colon \Sigma^* \to \Sigma^*$, given by

$$f(w) = www$$

is RAM computable. ($\Sigma = \{a_1, \ldots, a_N\}$).

Problem B6 (100 pts). Write a computer program implementing a RAM program interpreter. You may want to assume that the instructions have five fields

N	X	`opcode`	j	Y
N	X	`opcode`	j	$N1$

with $j \in \{1, \ldots, k\}$, where k is the number of symbols in Σ, and that the opcodes are

`add` `tail` `clr` `assign` `gotoa` `gotob` `jmpa` `jmpb` `continue`

where `gota` corresponds to jump above, `gotob` to jump below, `jpma` corresponds to jump above if condition is satisfied, and `jmpb` to to jump below if condition is satisfied. Depending on the opcode, some of the fields may be irrelevant (set them to 0).

The number of input registers is n (so your memory must have at least n registers), and the total number of registers is p. The number k, n p are input to your interpreter, as well as the program to be executed (a sequence of instructions). Assume that line numbers are integers.

Your program should output.

1. The input RAM program P
2. The input strings $w_1, \ldots, w_n$ to the RAM program P.
3. The value of the function being computed.

4. The sequence consisting of the memory contents and the current program counter as your interpreter executes the RAM program.

Test your interpreter on several RAM programs (and input strings), including the programs of B5.

TOTAL: 330 points.