

CSE105 (Winter 2014): Homework 4

Instructor: Daniele Micciancio

Due on Wednesday, February 12, 2014.

Solutions to this homework assignment should consists of a writeup with your answers to problems 1, 4 and 5a-5d (to be submitted in class on the due date) and jflap files for problems 2a-2d, and 4e1-4e3, to be submitted as with turnin using the usual procedure. Depending on your answer, the solution to problem 3 should be submitted either as part of your writeup or as a jflap file. Make sure you name your files as HW4p2a.jff-HW4p2d.jff, (HW4p3.jff), HW4p5e1.jff-HW4p5e3.jff. Zip the files together and submit the zipped file using turnin.)

1 Pumping Lemma

Let L be the set of all strings (over the alphabet $\{0, 1, +, =\}$ of the form $x + y = z$, where x, y and z are the binary representation of three numbers such that the sum of x and y equals z . For example, “110+11 = 1001” is in L . Prove that the language L is **not** regular using the pumping lemma.

2 Context Free Grammars

For each of the following languages design a corresponding context free grammar, and submit your answer as a jflap file HW4p2a.jff-HW4p2d.jff.

- (a) The set of all strings of the form $a^n b^m$ such that $n \neq m$.
- (b) The complement of the set $\{a^n b^n : n \geq 0\}$.
- (c) The set of all strings over the alphabet $\{a, b\}$ that are **not** palindromes. (Remember that a palindrom is a string that reads the same when spelled backward.) For example, *abbaa*, *baba* are in L , but *abba*, *bab*, *a* are not.
- (d) The set of all strings of the form $a^{n_1} b a^{n_2} b \dots b a^{n_k}$ with $k \geq 2$ and $n_1, \dots, n_k \geq 0$ integers such that $n_i = n_j$ for some $i \neq j$. In other words, the language contains strings consisting of (possibly empty) sequences of *as*, separated by single *bs*, such that two sequences of *as* have the same length. For example *abaabaaabaa*, *abbb* are in L (the second string contains 3 sequences of length 0!), while *abaabaaa*, *ab* are not.

3 Encoding grammars

Consider the set of context free grammars (say, over the alphabet $\Sigma = \{a, b\}$) using the following syntactical conventions:

- Variable symbols are indexed by positive integers $n \geq 0$, and they are written as px^nq , i.e., as n copies of 'x' enclosed between a 'p' and a 'q'.

- The start symbol has index 0, so it is written just as 'pq'.
- Rules are written as $rV2Ae$ where V is a variable symbol (written as px^nq), '2' is a separator symbol, A is the concatenation of terminal ('a' or 'b') and nonterminal (variables px^nq) symbols, and 'r' and 'e' are special symbols marking the beginning and end of a rule.
- Rules are concatenated together into a single string.

For example, the grammar G_6 in Example 2.10 from the textbook is given by

$rpq2pxqpqpxqe \quad rpq2pxxe \quad rpxq2pxxe \quad rpxq2pqe \quad rpxxq2be \quad rpxxq2e$

where, for readability, we have broken the string into fragments corresponding to each grammar rule. As you see, each context free grammar is described by a string over the 8-element alphabet $\Sigma = \{a, b, p, q, x, r, e, 2\}$. So, the set of (syntactically valid) context free grammars is a language over Σ .

Is the set of syntactically valid context free grammars a regular language? Prove your answer correct, either by using the pumping lemma (to prove the language nonregular) or by giving a regular expression describing the language (to prove the language regular.) In the first case, include the proof that the language is not regular in your writeup. In the second case, write the regular expression in JFLAP and submit it as file HW4p3.jff.

4 A non-context-free language

Prove that there is a language $L \subseteq \{a, b\}^*$ which is not context free, i.e., it cannot be described by a context free grammar. [Hint: give a proof by diagonalization, along the same lines as the one given in the lecture notes on nonregular languages. You may want to use the encoding of context free grammars from problem 3, but remember, your language L should be over the alphabet $\{a, b\}$.]

5 From Regular Expressions to Context Free Grammars

In this problem you are asked to prove that any regular expression can be transformed into an equivalent context free grammar. The problem is broken into parts.

(a) Prove that the class of context free languages is closed under union by giving a transformation that on input two context free grammars $G_1 = (V_1, \Sigma, R_1, S_1)$ and $G_2 = (V_2, \Sigma, R_2, S_2)$, produces a context free grammar $G = (V, \Sigma, R, S)$ such that $L(G) = L(G_1) \cup L(G_2)$.

(b) Prove that the class of context free languages is closed under concatenation by giving a transformation that on input two context free grammars $G_1 = (V_1, \Sigma, R_1, S_1)$ and $G_2 = (V_2, \Sigma, R_2, S_2)$, produces a context free grammar $G = (V, \Sigma, R, S)$ such that $L(G) = L(G_1) \cdot L(G_2)$.

(c) Prove that the class of context free languages is closed under star by giving a transformation that on input a context free grammar $G_1 = (V_1, \Sigma, R_1, S_1)$, produces a context free grammar $G = (V, \Sigma, R, S)$ such that $L(G) = L(G_1)^*$.

(d) Use parts (a),(b) and (c) to prove that for any regular expression E there is a corresponding context free grammar G such that $L(G) = L(E)$.

(e) Use part (d) to give a context free grammar for each of the following regular expressions:

1. $a(b \cup a)^* \cup b^*$
2. $(ab \cup ba)^*$
3. $((a \cup b^*)c)^*$

Write your context free grammars in jflap and submit them as files HW4p5e1.jff-HW4p5e3.jff.