CSE 2001—Winter 2008

Problem Set No. 2
Posted: February 15, 2008
Due: TBA on the web site

All reports must be typed (except for diagrams). All assignments are due by 2:00pm on the due date **in the course box**.

1. The following *must* be done by faithfully following the **proof** of Kleene's theorem, namely, that "for any Regular Expression α there is an NFA M such that $L(\alpha) = L(M)$".

Start with the regular expression

$$a(a + b)^* aa$$

over $\{a, b\}$.

(a) (**2 MARKS**) Construct an NFA that accepts it, *using the technique used in the proof of Kleene's theorem.* **State-diagram notation please!**

(b) (**2 MARKS**) Transform the NFA to a DFA. **State-diagram (graph) notation please!**

(c) (**3 MARKS**) Minimise the DFA.

2. (**5 MARKS**) Provide an algorithm that checks whether or not

$$(\exists x)(x \notin L_1 \cup L_2)$$

for any given regular languages L_1, L_2 and string x, all over some fixed Σ.

3. (**10 MARKS**) Define for any language L its "initial segment", $init(L)$, by

$$init(L) = \{w : (\exists y)wy \in L\} \tag{1}$$

Prove: If L is regular, then so is $init(L)$. Do so in two ways, (A) and (B):

(A) Assume for some FA M that $L = L(M)$. Using M, build, in detail, and justify your design, an FA N such that $init(L) = L(N)$. "Build" means in terms of M. You will explain how M is to be modified to get N, no matter what M you started with.

(B) Assume that $L = L(\alpha)$ for some regular expression α. Now use a proof by induction on α's length.

4. (**5 MARKS**) By induction on regular expression length prove: "For every α, there is a CFG, G, such that $L(G) = L(\alpha)$".

"$S \to \emptyset$" cannot be a rule. $\emptyset$ is a *set*, not a *string*, so, it has *no business* on the rhs of a rule!

5. (**5 MARKS**) Show that a grammar that *mixes* regular productions of the forms $A \to Ba$ *and* $A \to aB$ *may* produce a *non*-regular context free language.

(*Hint.* Find a *simple* CFG with mixed productions that generates a CFL that is known *not* to be regular. As always, unless your grammar **obviously** produces the language in question, you have to give a good *general* argument that it does.)

6. (**5 MARKS**) Find a Chomsky Normal Form grammar for $\{0^n 1^{4n} : n \geq 1\}$, and then proceed to build, **using the method shown in the text or in class**, a 2-state PDA that accepts the language by **Empty-Stack**.

All steps must be shown.

7. (**5 MARKS**) Prove that the language over $\{a, b\}$ given by $L = \{a^n b a^{n^2} : n \geq 0\}$ is *not* a CFL.