

CST 2013 Part IA
Regular Languages and Finite Automata
Exercise Sheet

1 Regular Expressions

Exercise 1.1. Write down an ML data type declaration for a type constructor 'a regexp whose values correspond to the regular expressions over an alphabet 'a.

Exercise 1.2. Find regular expressions over $\{0, 1\}$ that determine the following languages:

(a) $\{u \mid u \text{ contains an even number of } 1\text{'s}\}$

(b) $\{u \mid u \text{ contains an odd number of } 0\text{'s}\}$

Exercise 1.3. For which alphabets Σ is the set Σ^* of all finite strings over Σ itself an alphabet?

Tripos questions 2012.2.8 2005.2.1(d) 1999.2.1(s) 1997.2.1(q) 1996.2.1(i)

2 Finite State Machines

Exercise 2.1. For each of the two languages mentioned in Exercise 1.2 find a DFA that accepts exactly that set of strings.

Exercise 2.2. The example of the subset construction given on Slide 17 in the lecture notes constructs a DFA with eight states whose language of accepted strings happens to be $L(a^*b^*)$. Give a DFA with the same language of accepted strings, but fewer states. Give an NFA with even fewer states that does the same job.

Tripos questions 2010.2.9 2009.2.9 2004.2.1(d) 2001.2.1(d) 2000.2.1(b) 1998.2.1(s) 1995.2.19

3 Regular Languages

Exercise 3.1. Why can't the automaton $\text{Star}(M)$ required in step (iv) of Section 3.1 be constructed simply by taking M , making its start state the only accepting state and adding new ε -transitions back from each old accepting state to its start state?

Exercise 3.2. Construct an NFA ^{ε} M satisfying $L(M) = L((\varepsilon|b)^*aab^*)$.

Exercise 3.3. Show that any finite set of strings is a regular language.

Exercise 3.4. Use the construction in Section 4.1 to find a regular expression for the DFA M whose state set is $\{0, 1, 2\}$, whose start state is 0, whose only accepting state is 2, whose alphabet of input symbols is $\{a, b\}$, and whose next-state function is given by the following table.

$\delta_M :$	a	b
0	1	2
1	2	1
2	2	1

Exercise 3.5. The construction $M \mapsto \text{Not}(M)$ given on Slide 26 applies to both DFA and NFA; but for $L(\text{Not}(M))$ to be the complement of $L(M)$ we need M to be deterministic. Give an example of an alphabet Σ and a NFA M with set of input symbols Σ , such that $\{u \in \Sigma^* \mid u \notin L(M)\}$ is not the same set as $L(\text{Not}(M))$.

Exercise 3.6. Let $r = (a|b)^*ab(a|b)^*$. Find a complement for r over the alphabet $\Sigma = \{a, b\}$, i.e. a regular expressions $\sim(r)$ over the alphabet Σ satisfying $L(\sim(r)) = \{u \in \Sigma^* \mid u \notin L(r)\}$.

Tripos questions 2003.2.9 2000.2.7 1995.2.20

4 The Pumping Lemma

Exercise 4.1. Show that there is no DFA M for which $L(M)$ is the language on Slide 33. [Hint: argue by contradiction. If there were such an M , consider the DFA M' with the same states as M , with alphabet of input symbols just consisting of a and b , with transitions all those of M which are labelled by a or b , with start state $\delta_M(s_M, c)$ (where s_M is the start state of M), and with the same accepting states as M . Show that the language accepted by M' has to be $\{a^n b^n \mid n \geq 0\}$ and deduce that no such M can exist.]

Tripos questions 2011.2.8 2006.2.8 2004.2.9 2002.2.9 2001.2.7 1999.2.7 1998.2.7
1996.2.1(j) 1996.2.8 1995.2.27

5 Grammars

Exercise 5.1. Why is the string `the dog` a not in the language generated by the context-free grammar in Section 6.1?

Exercise 5.2. Give a derivation showing that $x + (x'')$ is in the language generated by the context-free grammar on Slide 37. Prove that $x + (x)''$ is not in that language. [Hint: show that if u is a string of terminals and non-terminals occurring in a derivation of this grammar and that $'$ occurs in u , then it does so in a substring of the form v' , or v'' , or v''' , etc., where v is either x or `id`.]

Exercise 5.3. Give a context-free grammar generating all the palindromes over the alphabet $\{a, b\}$.

Exercise 5.4. Give a context-free grammar generating all the regular expressions over the alphabet $\{a, b\}$.

Exercise 5.5. Using the construction given in the proof of part (a) of the Theorem on Slide 40, convert the regular grammar with start symbol q_0 and productions

$$\begin{aligned} q_0 &\rightarrow \varepsilon \\ q_0 &\rightarrow abq_0 \\ q_0 &\rightarrow cq_1 \\ q_1 &\rightarrow ab \end{aligned}$$

into an NFA ^{ε} whose language is that generated by the grammar.

Exercise 5.6. Is the language generated by the context-free grammar on Slide 35 a regular language? What about the one on Slide 37?

Tripos questions 2008.2.8 2005.2.9 2002.2.1(d) 1997.2.7 1996.2.1(k)

6 Pushdown Automata

Exercise 6.1. *Show that if M is the NPDA from Slide 46, then $L(M)$ is the context-free language $\{a^n b^n \mid n \geq 1\}$.*

Exercise 6.2. *Give a NPDA accepting the language of palindromes over the alphabet $\{a, b\}$.*