Languages and Automata Exercises

Thomas Forster

October 5, 2003

1. Prove that $L((r|s)^*) = L((r^*s^*)^*)$ (Use induction on word length)

2. Prove that $L((rs^*)^*) \subseteq L((r^*s^*)^*)$ but that the reverse inclusion does not hold.

3. Describe[1] deterministic automata to recognise the following subsets of $\{0, 1\}^*$:

   (a) The set of all strings with three consecutive 0's; provide a regular expression corresponding to this set as well;

   (b) The set of all strings w such that every set of five consecutive characters in w contains at least two 0's;

   (c) The set of all strings such that the 10th character from the right end is a '0'; provide a regular expression corresponding to this set as well. *For pedants:* This could mean one of two things. Answer both of them.

4. Let L be a regular language over an alphabet Σ. Which of the following are regular languages?

   (a) $\{w \in \Sigma^* : (\exists u \in \Sigma^*)(wu \notin L)\}$

   (b) $\{w \in L : (\forall u \in \Sigma^*)((\text{length}(u) > 0) \to wu \notin L)\}$

   (c) $\{w \in L : (\forall u, v \in \Sigma^*)((w = uv \wedge \text{length}(u) > 0) \to u \notin L)\}$

   (d) The preceding question has a typo in it. Find it.

   (e) S, an arbitrary subset of L.

   (f) $\{w \in \Sigma^* : (\exists u, v \in \Sigma^*)((w = uv) \wedge (vu \in L))\}$ (*hint*: needs a different approach ...)

5. A combination lock has three 1-bit inputs and opens just when it receives the input sequence `101`, `111`, `011`, `010`. Design a finite deterministic automaton with this behaviour (with accepting state(s) corresponding to the lock being open).

[1]This word is very carefully chosen!

6. Let Σ be an alphabet and let B and C be subsets of Σ^* such that the empty string is not in B. Let $X \subseteq \Sigma^*$ and show that if X satisfies the equation $X = BX \cup C$, then $B^*C \subseteq X$ and $X \subseteq B^*C$, i.e. the unique solution is $X = B^*C$. [Hint: use induction on number of "blocks".]

7. Show that if in the previous question we allow $\Lambda \in B$, then $X = B^*D$ is a solution for any $D \supset C$.

8. Let $A = \{b, c\}$, $B = \{b\}$, $C = \{c\}$. Find the solutions $X_1, X_2 \subset A^*$ of the following pairs of simultaneous equations: (i) $X_1 = BX_1 \cup CX_2$; $X_2 = (B \cup C)X_1 \cup CX_2 \cup \{\Lambda\}$ (ii) $X_1 = (BX_1 \cup \{\Lambda\})$; $X_2 = BC(X_1 \cup \{\Lambda\})$.

9. There is an alphabet Σ with six letters a, b, c, d, e and f that represent the six rotations through $\pi/2$ radians of each face of the Rubik cube. Everything you can do to the Rubik cube can be represented as a word in this language. Let L be the set of words in Σ^* that take the cube from its initial state back to its initial state. Is L regular? If you have had a sleepless night over this you may consult the footnote for a hint.[2]

10. Can you construct an FDA to recognise binary representations of multiples of 3? You may assume the machine starts reading the most significant bit first.

11. For which primes p can you build a FDA to recognise decimal representations of multiples of p? How many states do your machines have?

12. Let q be a number between 0 and 1. Let L be the set of sequences $s \in \{0,1\}^*$ such that the binary number between 0 and 1 represented by s is less than or equal to q. Show that L is a regular language iff q is rational. What difference would it have made if we had defined L to be be the set of sequences $s \in \{0,1\}^*$ such that the binary number between 0 and 1 represented by s is less than q.

13. Give regular grammars for the two following regular expressions over the alphabet $\Sigma = \{a, b\}$ and construct finite non-deterministic automata accepting the regular language denoted by them:

    (a) $ba|(a|bb)a^*b$

    (b) $((a|b)(a|b))^*|((a|b)(a|b)(a|b))^*$

14. For each of the following languages either show that the language is regular (for example by showing how it would be possible to construct a finite state machine to recognise it) or use the pumping lemma to show that it is not.

    (a) The set of all words not in a given regular language L.

[2] If this is to be a regular language, there must be a FDA that recognises it. What might this FDA be?

(b) The set of all palindromes over the alphabet a, b, c (a palindrome is a word that is unchanged when reversed, for example, *abcbabcba*).

(c) If L is a regular language, the language which consists of reversals of the words in L; thus if L contains the word *abcd*, then the reversed language L^R contains *dcba*.

(d) Given regular languages L and M, the set of strings that contain within them first a substring that is part of language L, then a substring from M; arbitrary characters from the alphabet a, b, c are allowed before, between and after these strings.

(e) Given regular languages L and M, the set of strings that contain within them some substring which is part of both L and M.

15. What is the language of boolean (propositional) logic? Is it regular? What about the version without infixes ("Polish notation") What about reverse polish notation?

16. A "Muller C-element" is a device which receives two streams of binary digits ($x_0x_1x_2\ldots$ and $y_0y_1y_2\ldots$) and outputs a stream of binary digits $z_0z_1z_2\ldots$ satisfying the relation

$$z_{n+1} = \begin{cases} x_{n+1} & \text{if } x_{n+1} = y_{n+1}, \\ z_n & \text{if } x_{n+1} \neq y_{n+1} \end{cases}$$

Design a Moore machine with this behaviour.

17. Give context-free grammars generating the following languages:

(a) $\{a^p b^q c^r | p \neq q \text{ or } q \neq r\}$

(b) $\{w \in \{a, b\}^* |\ w$ contains exactly twice as many as as bs$\}$

18. Let M be a finite deterministic automaton with n states. Prove that $L(M)$ is an infinite set if and only if it contains a string of length l with $n \leq l < 2n$.

19. The **interleaving** of two languages L and M is the set of all words that can be obtained from a word in L and a word in M by interleaving the two words in the way that people shuffle together two halves of a pack of cards. Prove that

(a) The interleaving of two regular languages is regular

(b) The interleaving of a regular and a context free language is context free

(c) The interleaving of two context free languages is not always context-free.[3]

20. Does the set of strings in $\{a, b, c\}^*$ which have as many as as bs and cs put together make a regular language?

21. Let K, L and M be regular languages. Is $\{u \in L : (\exists v \in K)(uv \in M)\}$ a regular language?

22. Is the language of Roman numerals regular?

[3]There is a pumping lemma for context-free languages which is not in the course. With the hint that $a^n b^m c^n d^m$ is not context-free it all becomes terribly easy!