CST 2020–2021 Part IA Discrete Mathematics
Formal Languages and Automata
Exercise Sheet

Revision 12 of 2021-01-04 04:12:41 +0000 (Mon, 04 Jan 2021)

1 Inductive definitions

Exercise 1.1. Let L be the subset of $\{a, b\}^*$ inductively defined by the axiom $\dfrac{}{\varepsilon}$ and the rule $\dfrac{u}{aub}$ (for any $u \in \{a, b\}^*$).

(a) Use rule induction to prove that every string in L is of the form $a^n b^n$ for some $n \in \mathbb{N}$.

(b) Use mathematical induction to prove $\forall n \in \mathbb{N}.\ a^n b^n \in L$.
Conclude that $L = \{a^n b^n \mid n \in \mathbb{N}\}$.

(c) Suppose we add the string a to L, that is, consider $L' = L \cup \{a\}$. Is L' closed under the axiom and rule? If not, characterise the strings that would be in the smallest set containing L' that is closed under the axiom and rule.

Exercise 1.2. Suppose $R \subseteq X \times X$ is a binary relation on a set X. Let $R^\dagger \subseteq X \times X$ be inductively defined by the following axioms and rules:

$$\frac{}{(x, x) \in R^\dagger}\ (x \in X), \qquad \frac{(x, y) \in R^\dagger}{(x, z) \in R^\dagger}\ (x \in X \text{ and } (y, z) \in R).$$

Rule induction can be used in parts (b) and (c). In both parts the base case will involve showing $(x, x), \forall x \in X$ are members of the relevant set.

(a) Show that $R^\dagger$ is reflexive and that $R \subseteq R^\dagger$.

(b) Use rule induction to show that $R^\dagger$ is a subset of

$$S \triangleq \{(y, z) \in X \times X \mid \forall x \in X.\ (x, y) \in R^\dagger \Rightarrow (x, z) \in R^\dagger\}.$$

Deduce that $R^\dagger$ is transitive.

(c) Suppose $S \subseteq X \times X$ is a reflexive and transitive binary relation and that $R \subseteq S$. Use rule induction to show that $R^\dagger \subseteq S$.

(d) Deduce from (a)–(c) that $R^\dagger$ is equal to R^*, the reflexive-transitive closure of R.

Exercise 1.3. Let L be the subset of $\{a, b\}^*$ inductively defined by the axiom $\dfrac{}{ab}$ and the rules $\dfrac{au}{au^2}$ and $\dfrac{ab^3u}{au}$ (for all $u \in \{a, b\}^*$).

(a) Is ab^5 in L? Give a derivation, or show there isn't one.

(b) Use rule induction to show that every $u \in L$ is of the form ab^n with $n = 2^k - 3m \geq 0$ for some $k, m \in \mathbb{N}$.

(c) Is ab^3 in L? Give a derivation, or show there isn't one.

(d) Can you characterize exactly which strings are in L?

Tripos questions y2016p2q10(a) y2019p2q10(b) y2020p2q10(c)

2 Regular expressions

Exercise 2.1. Find regular expressions over $\{0,1\}$ that determine the following languages:

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

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

Exercise 2.2. Show that the inductive definition of the matching relation between strings and regular expressions given on the slide of the lecture notes entitled "Inductive defintion of matching" has the properties listed on the preceding slide entitled "Matching".

Exercise 2.3. Show that $b^*a(b^*a)^*$ and $(a|b)^*a$ are equivalent regular expressions, that is, a string matches one iff it matches the other.

Tripos questions y2014p2q10(a.ii) y2012p2q8(a–d) y2005p2q1(d) y1999p2q1(s) y1997p2q1(q) y1996p2q1(i) y2019p2q10(a)(i)

3 Finite automata

Exercise 3.1. For each of the two languages mentioned in Exercise 2.1 find a DFA that accepts exactly that set of strings.

Exercise 3.2. Given an NFA$^\varepsilon$ $M = (Q, \Sigma, \Delta, s, F, T)$, we write $q \stackrel{u}{\Rightarrow} q'$ to mean that there is a path in M from state q to state q' whose non-ε labels form the string $u \in \Sigma^*$. Show that $\{(q, u, q') \mid q \stackrel{u}{\Rightarrow} q'\}$ is equal to the subset of $Q \times \Sigma^* \times Q$ inductively defined by the axioms and rules

$$\frac{}{(q, \varepsilon, q)}$$

$$\frac{(q, u, q')}{(q, u, q'')} \text{ if } q' \xrightarrow{\varepsilon} q'' \text{ in } M$$

$$\frac{(q, u, q')}{(q, ua, q'')} \text{ if } q' \xrightarrow{a} q'' \text{ in } M.$$

Exercise 3.3. The example of the subset construction given 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 y2015p2q10(a) y2014p2q10 y2010p2q9 y2009p2q9 y2004p2q1(d) y2001p2q1(d) y2000p2q1(b) y1998p2q1(s) y1995p2q19 y2019p2q10(a)(ii) y2020p2q10(a)

4 Regular Languages

Exercise 4.1. Why can't the automaton $Star(M)$ used in step (iv) of the proof of part (a) of Kleene's Theorem 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 4.2. Construct an NFA$^{\varepsilon}$ M satisfying $L(M) = L((\epsilon|b)^*aab^*)$.

Exercise 4.3. Show that any finite set of strings is a regular language.

Exercise 4.4. Use the construction given in the proof of part (b) of Kleene's Theorem 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.

δ	a	b
0	1	2
1	2	1
2	2	1

Exercise 4.5. If $M = (Q, \Sigma, \Delta, s, F)$ is an NFA, let $Not(M)$ be the NFA $(Q, \Sigma, \Delta, s, Q \setminus F)$ obtained from M by interchaning the role of accepting and non-accepting states. Give an example of an alphabet Σ and an NFA M with set of input symbols Σ, such that $\{u \in \Sigma^* \mid u \notin L(M)\}$ is *not* the same set as $L(Not(M))$.

Exercise 4.6. Let $r = (a|b)^*ab(a|b)^*$. Find a complement for r over the alphabet $\{a, b\}$, i.e. a regular expressions $\sim r$ over the alphabet $\{a, b\}$ satisfying $L(\sim r) = \{u \in \{a, b\}^* \mid u \notin L(r)\}$.

Exercise 4.7. Given DFAs $M_i = (Q_i, \Sigma, \delta_i, s_i, F_i)$ for $i = 1, 2$, let $And(M_1, M_2)$ be the DFA $(Q_1 \times Q_2, \Sigma, \delta, (s_1, s_2), F_1 \times F_2)$ where $\delta : (Q_1 \times Q_2) \times \Sigma \to (Q_1 \times Q_2)$ is given by:

$$\delta((q_1, q_2), a) = (\delta_1(q_1, a), \delta_2(q_2, a))$$

for all $q_1 \in Q_1$, $q_2 \in Q_2$ and $a \in \Sigma$. Show that $L(And(M_1, M_2))$ is the intersection of $L(M_1)$ and $L(M_2)$.

Tripos questions y2018p2q10 y2017p2q10 y2016p2q10(b–c) y2003p2q9 y2000p2q7 y1995p2q20 y2019p2q10(c)

5 The Pumping Lemma

Exercise 5.1. Consider

$$L \triangleq \{c^m a^n b^n \mid m \geq 1 \,\&\, n \geq 0\} \cup \{a^m b^n \mid m, n \geq 0\}$$

The notes show that this language has the pumping lemma property. Show that there is no DFA M which accepts L. [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 y2015p2q10(c) y2014p2q10(b) y2011p2q8 y2006p2q8 y2004p2q9 y2002p2q9 y2001p2q7 y1999p2q7 y1998p2q7 y1996p2q1(j) y1996p2q8 y1995p2q27 y2020p2q10(b)