The University of Melbourne
School of Computing and Information Systems
COMP30026 Models of Computation

Selected Tutorial Solutions, Week 9

73. (a) $\{w \mid w$ begins with a 1 and ends with a 0$\}$: $1(0 \cup 1)^*0$
    (b) $\{w \mid w$ contains the substring 0101$\}$: $(0 \cup 1)^*0101(0 \cup 1)^*$
    (c) $\{w \mid w$ has length at least 3 and its third symbol is 0$\}$: $(0 \cup 1)(0 \cup 1)0(0 \cup 1)^*$
    (d) $\{w \mid$ the length of w is at most 5$\}$: $(\epsilon \cup 0 \cup 1)(\epsilon \cup 0 \cup 1)(\epsilon \cup 0 \cup 1)(\epsilon \cup 0 \cup 1)(\epsilon \cup 0 \cup 1)$
    (e) $\{w \mid w$ is any string except 11 and 111$\}$: $\epsilon \cup 1 \cup 11111^* \cup (0 \cup 1)^*0(0 \cup 1)^*$
    (f) $\{w \mid$ every odd position of w is a 1$\}$: $(1(0 \cup 1))^*(\epsilon \cup 1)$
    (g) $\{w \mid w$ contains at least two 0s and at most one 1$\}$: $0^*(00 \cup 001 \cup 010 \cup 100)0^*$
    (h) $\{\epsilon, 0\}$: $\epsilon \cup 0$
    (i) The empty set: $\emptyset$
    (j) All strings except the empty string: $(0 \cup 1)(0 \cup 1)^*$

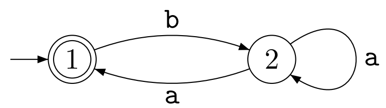
74. If A is regular then *suffix*(A) is regular. Namely, let $D = (Q, \Sigma, \delta, q_0, F)$ be a DFA for A. Assume every state in Q is *reachable* from q_0. Then we can turn D into an NFA N for *suffix*(A) by adding a new state q_{-1} which becomes the NFA's start state. For each state $q \in Q$, we add an epsilon transition from q_{-1} to q.

    That is, we define N to be $(Q \cup \{q_{-1}\}, \Sigma, \delta', q_{-1}, F)$, with transition function

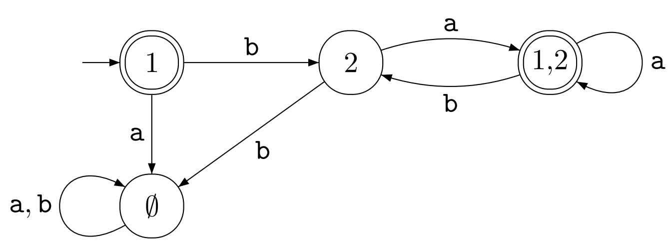
$$\delta'(q, x) = \begin{cases} \{\delta(q, x)\} & \text{for } q \in Q \text{ and } x \in \Sigma \\ Q & \text{for } q = q_{-1} \text{ and } x = \epsilon \\ \emptyset & \text{for } q \neq q_{-1} \text{ and } x = \epsilon \end{cases}$$

    The restriction we assumed, that all of D's states are reachable, is not a severe one. It is easy to identify unreachable states and eliminate them (which of course does not change the language of the DFA). To see why we need to eliminate unreachable states before generating N in the suggested way, consider what happens to this DFA for $\{\epsilon\}$: $(\{q_0, q_1, q_2\}, \{\mathtt{a}\}, \delta, q_0, \{q_0\})$, where $\delta(q_0, \mathtt{a}) = \delta(q_1, \mathtt{a}) = q_1$ and $\delta(q_2, \mathtt{a}) = q_0$.

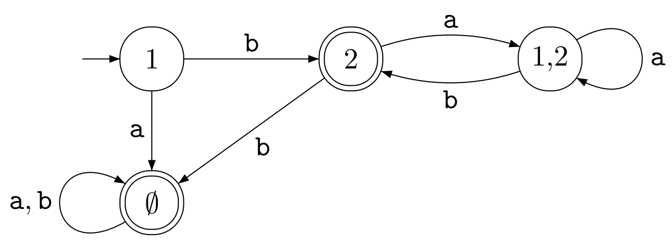
75. (a) Here is an NFA for $(\mathtt{ba}^*\mathtt{a})^*$:



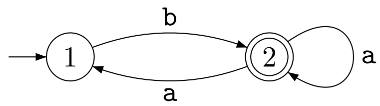
    (b) Here is an equivalent DFA, obtained using the subset construction:



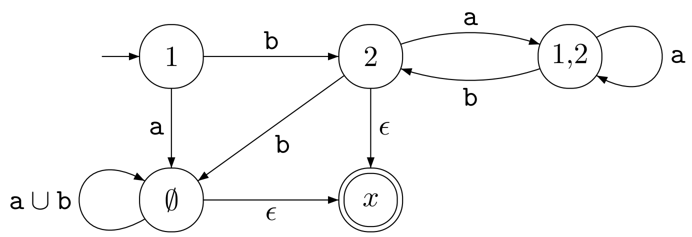
    (c) It is easy to get a DFA for the complement:



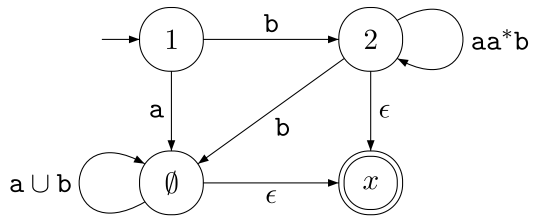
(d) It would be problematic to do the "complement trick" on the NFA, as it is only guaranteed to work on DFAs. We would get the following, which accepts, for example, `baa`:



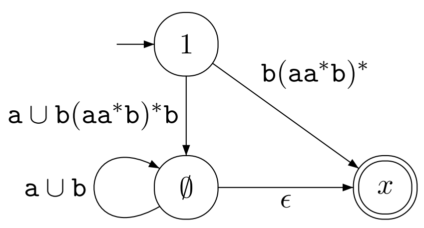
(e) Starting from the DFA that we found in (c), we make sure that we have just one accept state:



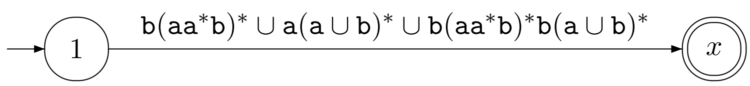
Let us first remove the state labeled 1,2:



Now we can remove state 2:



Finally, eliminating the state labelled $\emptyset$, we are left with:



The resulting regular expression can be read from that diagram.

76. (a) We show that $A = \{\mathtt{0}^n\mathtt{1}^n\mathtt{2}^n \mid n \geq 0\}$ is not regular. Assume it is, and let p be the pumping length. Consider $s = 0^p1^p2^p \in A$. Since $|s| \geq p$, by the pumping lemma, s can be written $s = xyz$, so that $y \neq \epsilon$, $|xy| \leq p$, and $xy^iz \in A$ for all $i \geq 0$. But since $|xy| \leq p$, y must contain 0s only. Hence $xz \notin A$. Thus we have a contradiction, and we conclude that A is not regular.

(b) We use the pumping lemma to show that B is not regular. Assume that it were. Let p be the pumping length, and consider $\mathtt{a}^{p+1}\mathtt{b}\mathtt{a}^p$. This string is in B and of length $2p + 2$. By the pumping lemma, there are strings x, y, and z such that $\mathtt{a}^{p+1}\mathtt{b}\mathtt{a}^p = xyz$, with $y \neq \epsilon$, $|xy| \leq p$, and $xy^iz \in B$ for all $i \in \mathbb{N}$. By the first two conditions, y must be a non-empty string consisting of as only. But then, pumping down, we get xz, in which the number of as on the left no longer is strictly larger than the number of as on the right. Hence we have a contradiction, and we conclude that B is not regular.

(c) We want to show that $C = \{w \in \{\mathtt{a}, \mathtt{b}\}^* \mid w \text{ is not a palindrome}\}$ is not regular. But since regular languages are closed under complement, it will suffice to show that $C^c = \{w \in \{\mathtt{a}, \mathtt{b}\}^* \mid w \text{ is a palindrome}\}$ is not regular. Assume that C^c *is* regular, and let p be the pumping length. Consider $a^pba^p \in C^c$. Since $|s| \geq p$, by the pumping lemma, s can be written $s = xyz$, so that $y \neq \epsilon$, $|xy| \leq p$, and $xy^iz \in C^c$ for all $i \geq 0$. But since $|xy| \leq p$, y must contain as only. Hence $xz \notin C^c$. We have a contradiction, and we conclude that C^c is not regular. Now if C was regular, C^c would be regular too. Hence C cannot be regular.

77. Here are the context-free grammars:

(a) $\{w \mid w \text{ starts and ends with the same symbol}\}$:

$$\begin{array}{lcl} S & \rightarrow & 0\,T\,0 \mid 1\,T\,1 \mid 0 \mid 1 \\ T & \rightarrow & 0\,T \mid 1\,T \mid \epsilon \end{array}$$

(b) $\{w \mid \text{the length of } w \text{ is odd}\}$:

$$S \rightarrow 0 \mid 1 \mid 0\,0\,S \mid 0\,1\,S \mid 1\,0\,S \mid 1\,1\,S$$

(c) $\{w \mid \text{the length of } w \text{ is odd and its middle symbol is } \mathtt{0}\}$:

$$S \rightarrow 0 \mid 0\,S\,0 \mid 0\,S\,1 \mid 1\,S\,0 \mid 1\,S\,1$$

(d) $\{w \mid w \text{ is a palindrome}\}$:

$$S \rightarrow 0\,S\,0 \mid 1\,S\,1 \mid 0 \mid 1 \mid \epsilon$$

78. Here is a context-free grammar for $\{\mathtt{a}^i\mathtt{b}\mathtt{a}^j \mid i > j \geq 0\}$:

$$\begin{array}{lcl} S & \rightarrow & A\,B \\ A & \rightarrow & \mathtt{a} \mid \mathtt{a}\,A \\ B & \rightarrow & \mathtt{b} \mid \mathtt{a}\,B\,\mathtt{a} \end{array}$$

79. The class of context-free languages is closed under the regular operations: union, concatenation, and Kleene star.

Let G_1 and G_2 be context-free grammars generating L_1 and L_2, respectively. First, if necessary, rename variables in G_2 so that the two grammars have no variables in common. Let the start variables of G_1 and G_2 be S_1 and S_2, respectively. Then we get a context-free grammar for $L_1 \cup L_2$ by keeping the rules from G_1 and G_2, adding

$$\begin{array}{lcl} S & \rightarrow & S_1 \\ S & \rightarrow & S_2 \end{array}$$

where S is a fresh variable, and making S the new start variable.

We can do exactly the same sort of thing for $L_1 \circ L_2$. The only difference is that we now just add one rule:

$$S \rightarrow S_1\ S_2$$

again making (the fresh) S the new start variable.

Let G be a context-free grammar for L and let S be fresh. If we add two rules to those from G:

$$\begin{array}{lcl} S & \rightarrow & \epsilon \\ S & \rightarrow & S\ S' \end{array}$$

where S' is G's start variable, then we have a context-free grammar for L^* (it has the fresh S as its start variable).

80. Here are some sentences generated from the grammar:

    (a) A dog runs
    (b) A dog likes a bone
    (c) The quick dog chases the lazy cat
    (d) A lazy bone chases a cat
    (e) The lazy cat hides
    (f) The lazy cat hides a bone

    The grammar is concerned with the structure of well-formed sentences; it says nothing about meaning. A sentence such as "a lazy bone chases a cat" is syntactically correct—its structure makes sense; it could even be semantically correct, for example, "lazy bone" may be a derogatory characterisation of some person. But in general there is no guarantee that a well-formed sentence carries meaning.

81. We can easily extend the grammar so that a sentence may end with an optional adverbial modifier:

$$\begin{array}{lcl} S & \rightarrow & NP\ \ VP\ \ PP \\ & \vdots & \\ PP & \rightarrow & \epsilon \\ PP & \rightarrow & \text{quietly} \\ PP & \rightarrow & \text{all day} \\ & \vdots & \end{array}$$