Formal languages, grammars, and automata

Assignment 4, Wednesday, Dec. 3 2014

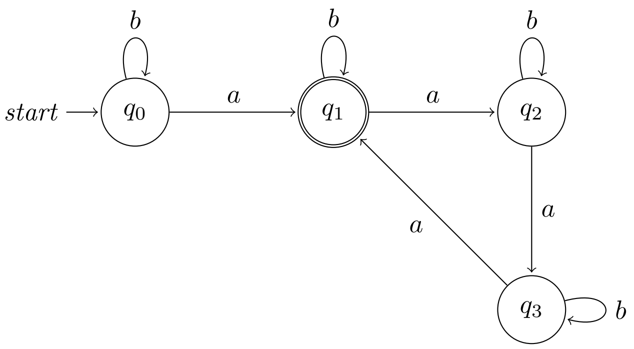
Exercises with answers

1. Let $\Sigma = \{a, b\}$.
   (a) (**5 points**) Construct a DFA that accepts L where $L = \mathcal{L}(b^*ab^*(ab^*ab^*ab^*)^*)$. (You may write down the DFA directly, without first constructing an NFA.)
   (b) (**5 points**) Derive (just like in the pumping lemma) from your automaton a number k, for which you can prove:
   for every $w \in L$ with $|w| \geq k$, there are words u_1, v, u_2 such that
   - $w = u_1 v u_2$ and
   - $|v| \geq 1$,
   - $|u_1 v| \leq k$
   - $\forall n \in \mathbb{N}(u_1 v^n u_2 \in L)$.

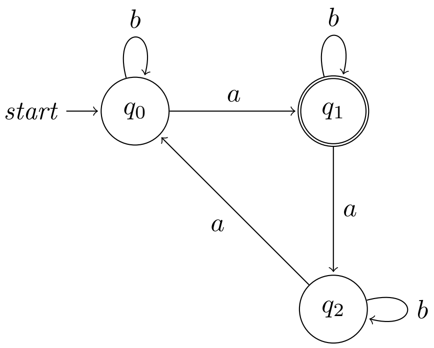
   Prove this.

Answer Info: ....

(a) A possible DFA accepting L is



Or:



(b) Proof in the style of the pumping lemma:
Let $k = 4$ be equal to the number of states in M. Given $w \in L$ with $|w| \geq k$, since w is accepted by M the automaton must end up in an accepting state when given w. In doing so it passes through $|w| + 1 > k$ states, so at least one state is visited twice. Let q be the first such state. There are 5 cases

$q = q_0$ *with a b-loop at q_0. Then $w = bu_2$, $u_1 = \lambda$, $v = b$, $|u_1 v| \leq 4$ and indeed $b^n u_2 \in L$ for every $n \geq 0$.*

$q = q_1$ *with a b-loop at* q_1*. Then* $w = abu_2$*,* $u_1 = a$*,* $v = b$*,* $|u_1 v| \leq 4$ *and indeed* $ab^n u_2 \in L$ *for every* $n \geq 0$*.*

$q = q_2$ *with a b-loop at* q_2*. Then* $w = aabu_2$*,* $u_1 = aa$*,* $v = b$*,* $|u_1 v| \leq 4$ *and indeed* $aab^n u_2 \in L$ *for every* $n \geq 0$*.*

$q = q_3$ *with a b-loop at* q_3*. Then* $w = aaabu_2$*,* $u_1 = aaa$*,* $v = b$*,* $|u_1 v| \leq 4$ *and indeed* $aab^n u_2 \in L$ *for every* $n \geq 0$*.*

$q = q_1$ *with a aaa-loop at* q_1 *(via* q_2, q_3*). Then* $w = aaaau_2$*,* $u_1 = a$*,* $v = aaa$*,* $|u_1 v| \leq 4$ *and indeed* $a(aaa)^n u_2 \in L$ *for every* $n \geq 0$*.*

End Answer Info ....

2. (a) (**5 points**) Prove that the language L_1 is not regular, where

$$L_1 := \{a^n b^p \mid n < p\}$$

(b) (**5 points**) Prove that the language L_2 is not regular, where

$$L_2 := \{a^n b^p \mid n > p\}$$

Give a proof using the pumping lemma and (for **2 bonus points**) also try to give a proof without using the pumping lemma (using closure properties of regular languages and languages that we know to be non-regular).

(c) Prove, without using the pumping lemma, that the language L_3 is not regular, where

$$L_3 := \{a^n b^p \mid n \leq p\}.$$

(So only use closure properties of regular languages and languages that we know to be non-regular.)

Answer Info: ....

(a) Assume that L_1 *is regular (towards a contradiction). By the pumping lemma, there exists a* k *such that for all* $w \in L_1$ *there are* u_1, v, u_2 *such that* $|u_1 v| \leq k$*,* $|v| \geq 1$ *and* $u_1 v^n u_2 \in L_1$ *for all* $n \in \mathbb{N}$*.*
Take $w = a^k b^{k+1}$*, which is in* L_1*. Then we get* $w = u_1 v u_2$*, and because* $|u_1 v| \leq k$ *we know that* $u_1 = a^{|u_1|}$ *and* $v = a^{|v|}$*. So* $u_1 v^2 u_2 = a^{k+|v|} b^{k+1} \in L$*. Because* $|v| \geq 1$ *we have* $k + |v| \geq k + 1$*, so* $u_1 v^2 u_2 \notin L_1$*. Which is a contradiction, so* L_1 *is not regular.*

(b) Proof with the pumping lemma.
Assume that L_1 *is regular (towards a contradiction). Let* k *be as in the pumping lemma. Take* $w = a^{k+1} b^k$*, which is in* L_2*. By the pumping lemma* $w = u_1 v u_2$ *with* $|u_1 v| \leq k$*,* $|v| \geq 1$ *and* $u_1 v^n u_2 \in L_2$ *for all* $n \geq 0$*. Because* $|u_1 v| \leq k$ *we know that* u_1 *and* v *consist entirely of as. Taking* $n = 0$ *we get that* $w' = a^{k+1-|v|} b^k \in L$*, but because* $|v| \geq 1$ *it is not the case that* $k + 1 - |v| > k$*, so* $w' \notin L$*. This is a contradiction, so* L_2 *is not regular.*
Proof without the pumping lemma.
Assume that L_2 *is regular (towards a contradiction). Then* $L_2^R = \{b^p a^n \mid n > p\}$ *is also regular. And if a language is regular, then applying a bijection to the alphabet still gives a regular language (this can be seen immediately by considering the DFA for the language), so*

$$\{a \mapsto b, b \mapsto a\}(L_2^R) = \{a^p b^n \mid n > p\}$$

is also regular. But this language is the same as L_1*, which we know from (a) is not regular. This is a contradiction. Therefore* L_2 *is not regular.*

(c) *The concatenation of two regular languages is regular. We have* $L_1 = L_3\mathcal{L}(b) = \{a^n b^{p+1} \mid n \leq p+1\}$. *So, if* L_3 *is regular then so is* L_1. *But* L_1 *is not regular, so* L_3 *must also not be regular.*
Alternative proof: If L_3 *is regular, then so is* $L_3^R = \{b^p a^n \mid n \leq p\}$ *and so is* L_3^R *with a and b exchanged, which is* $\{a^p b^n \mid n \leq p\}$. *But then also* $L_3 \cap \{a^p b^n \mid n \leq p\}$ *is regular, but that is* $\{a^p b^n \mid n = p\}$, *which we know to be non-regular.*

End Answer Info ....

3. Given L over Σ that is regular, prove that the following language L' is regular:

$$L' := \{w \in \Sigma^* \mid \exists v \in L(w \text{ is a sufffix of } v)\}.$$

NB. w is a *suffix of* v if $v = uw$ for some $u \in \Sigma^*$.

Answer Info: ....
If L *is regular, then it is accepted by some DFA* $M = \langle S, \delta, q_1, F\rangle$. *Define the NFA*$_\lambda$ $M' = \langle S \cup q_0, \delta', q_0, F\rangle$, *where*

$$\begin{aligned}
\delta'(q_0, \lambda) &= \{q \mid \exists u \in \Sigma^*.\ M \text{ is in state } q \text{ after consuming } u\} && \\
\delta'(q_0, a) &= \emptyset && \text{for all } a \in \Sigma \\
\delta'(q, \lambda) &= \emptyset && \text{for all } q \in S \\
\delta'(q, a) &= \{\delta(q, a)\} && \text{for all } q \in S, a \in \Sigma
\end{aligned}$$

That is, M' *is the same as* M *with one extra state* q_0 *that has* λ *transitions to all states of* M *that are reachable from* q_0.

Suppose that $w \in L'$. *Then* w *is the suffix of some word in* $v \in L$, *that is* $v = uw \in L$. *When starting in* q_1 *and consuming* u *the automaton* M *is in some state* q_u. *By construction there is then a* λ *transition from* q_0 *to* q_u *in* M'. *And we know that* M *accepts* w *when starting in* q_u, *and then so does* M', *so* M' *accepts* w.

Conversely, suppose that $w \in \mathcal{L}(M')$. *Since* q_0 *is not accepting, there must be a* λ *transition taken from* q_0 *to some* q_u, *such that* M' *accepts* w *starting from* q_u. *Since the rest of the automaton* M' *is the same as* M, M *also accepts* w *starting from* q_u. *By definition there is then an* $u \in \Sigma^*$ *such that* M *ends in* q_u *after consuming* u. *So* M *accepts* uw *and hence* $w \in L$.

Therefore M' *accepts precisely the language* L', *which proves that* L' *is regular.*
End Answer Info ....

4. Now, $\Sigma = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, +, \times, (,)\}$ Prove that the language L_4 is not regular, where

$$L_4 := \{e \in \Sigma^* \mid e \text{ is a well-formed arithmetical expression}\}$$

NB. In a *well-formed arithmetical expression* the brackets should "match", so $3\times(5+(3+0))$ is well-formed and so is $((((4+5)\times 7)))$, but $5+9)+3)$ and $(4\times(3\times 7)$ are not.

Answer Info: ....
By contradiction. Assume that L' *is regular. Then by the pumping lemma there exists a* k. *The word* $w = \underbrace{((\cdots((}_{k}0\underbrace{))\cdots))}_{k}$ *is a well-formed arithmetical expression. By the pumping lemma* $w = u_1 v u_2$ *such that* $|u_1 v| \leq k$. *So* u_1 *and* v *consist entirely of opening parentheses. So* $w' = \underbrace{((\cdots((}_{k+|v|}0\underbrace{))\cdots))}_{k} \in L'$. *But because* $|v|$ *we know that* w' *is not a well-formed arithmetical expression, so* $w' \notin L$. *This is a contradiction, so* L' *must not be regular.*
End Answer Info ....

5. (More challenging; This exercise is taken from lecture notes on Languages and Automata by Andy Pitts.)
This exercise shows an example of a language that *can* be pumped, but is *not regular*. Let $\Sigma = \{a, b, c\}$.

(a) Show that the following language can be pumped:

$$L = \{c^m a^n b^n \mid m \geq 1 \text{ and } n \geq 0\} \cup \{a^m b^n \mid m, n \geq 0\}$$

(Show that is satisfies the pumping lemma property with $k = 1$.)

(b) Show that L is non-regular.
[Hint: argue by contradiction. If there is a DFA M accepting L, consider the DFA M' with the same states as M, with Σ just $\{a, b\}$, with transitions all those of M which are labelled by a or b, with start state $\delta(q_o, c)$, (where q_0 is the start state of M), and with the same accepting states. Show that the language accepted by M' is not regular.
]

Answer Info: ....

(a) Let $k = 1$. Now suppose we are given a $w \in L$ with $|w| \geq k$. Then either $w = c^m a^n b^n$ for $m \geq 1$ and $n \geq 0$, or $w = a^m b^n$ with $m, n \geq 0$.

In the first case, take $u_1 = \lambda$, $v = c$, $u_2 = c^{m-1} a^n b^n$, so $w = u_1 v u_2$. This is possible because $m \geq 1$. It is the case that $|u_1 v| = 1 \leq k$ and $|v| \geq 1$. We need to prove that $u_1 v^p u_2 \in L$ for all $p \geq 0$. Let $p \geq 0$ be given. If $p = 0$ and $m = 1$ then we need to show that $a^n b^n \in L$. This is the case by the second part of the union. Otherwise we need to show that $c^{p+m} a^n b^n \in L$, which is the case by the first part of the union.

In the second case, if $m > 0$ take $u_1 = \lambda$, $v = a$ and $u_2 = a^{m-1} b^n$. If $m = 0$ take $u_1 = \lambda$, $v = b$ and $u_2 = b^{n-1}$, which is possible because $|w| = n \geq k = 1$. In both cases we have $|u_1 v| = 1 \leq k$ and $|v| \geq 1$. We need to prove that $u_1 v^p u_2 \in L$ for all $p \geq 0$. Let $p \geq 0$ be given. Then in both cases $u_1 v^p u_2$ is of the form $a^{m'} b^{n'}$, so $u_1 v^p u_2 \in L$

Hence, the language L can be pumped.

(b) Assume that L is regular (towards a contradiction). Then there is a DFA $M = \langle S, \delta, q_0, F \rangle$ accepting L. Consider the DFA $M' = \langle S, \delta', \delta(q_0, c), F \rangle$. Where δ' is δ restricted to the alphabet $\{a, b\}$. Then M' accepts a word $w \in \{a, b\}^*$ if and only if M accepts cw. This means that cw is of the form $c^m a^n b^n$ for some $m \geq 1$ and $n \geq 0\}$, hence w is of the form $a^n b^n$. So M' accepts the language $L' = \{a^n b^n \mid n \geq 0\}$. But we know that L' is not regular. This is a contradiction, so L must not be regular.

An alternative proof, without considering an automaton that accepts L, is as follows. Assume that L is regular (towards a contradiction). Then also $L \cap \mathcal{L}(ca^*b^*)$ is regular, which is $\{ca^n b^n \mid n \geq 0\}$. This can be shown to be non-regular, by an immediate adaptation of the proof that $\{a^n b^n \mid n \geq 0\}$ is non-regular.

End Answer Info ....