

COMP 481: Automata, Formal Languages, and Computability
 Spring 2009
 Solutions to Homework Assignment #3

1. (a) L_a is not regular and hence has an infinite number of equivalence classes, where each string $w \in \Sigma^*$ is in an equivalence class by itself.
- (b) L_b has an infinite number of equivalence classes, and they are

$$[a^i] = \{w \in \{a, b\}^* : \#_a(w) - \#_b(w) = i\},$$

for $i \geq 0$, and

$$[b^i] = \{w \in \{a, b\}^* : \#_b(w) - \#_a(w) = i\},$$

for $i \geq 1$.

- (c) The language is regular, and hence has a finite number of equivalence classes, which are:

- all strings that have bbb as a substring.
- all strings that do not have bbb as a substring and do not end with b .
- all strings that do not have bbb as a substring and the last symbol from the right is a b that is not preceded by a b .
- all strings that do not have bbb as a substring and the last two symbols from the right are bb .

It is easy to draw an FSM with four states, each of which corresponds to exactly one of these four equivalence classes.

2. In the first step, the following pairs are marked: $(1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5), (3, 6), (4, 6), (5, 6)$.

In the second step, the following pairs become marked: $(1, 6), (2, 6)$.

In the third step, nothing gets marked, and we stop. Therefore, we have the following groups of equivalent states: $p = \{1, 2\}$, $q = \{3, 4, 5\}$, and $r = \{6\}$. The minimized automaton is $M = (Q, \Sigma, \delta, q_0, F)$, where $Q = \{p, q, r\}$, $\Sigma = \{a, b\}$, $\delta = \{(p, a, p), (p, b, q), (q, a, q), (q, b, r), (r, a, r), (r, b, r)\}$, $q_0 = p$, and $F = \{q\}$.

3. (a) If there exists a string w , where $|w| < n$, such that $w \in L(M)$, then $L(M) \neq \emptyset$. Trivial. We now prove the other direction. Assume $L(M) \neq \emptyset$. Let $w \in L(M)$ be the shortest string in $L(M)$. If $|w| < n$, then we are done. Otherwise, $|w| \geq n$ and the Pumping Lemma applies. In other words, $w = xyz$, where $|xy| \leq n$, $|y| \geq 1$, and $w' = xy^0z = xz \in L(M)$. Clearly, $|w'| < |w|$. However, this contradicts the assumption that w is a shortest string in $L(M)$. Therefore, there must exist a string in $L(M)$ whose length is less than n .

- (b) If M accepts a string w , where $n \leq |w|$, then by the Pumping Lemma, $w = xyz$, where $|xy| \leq n$, $|y| \geq 1$, and $xy^iz \in L$ for every $i \geq 0$. Therefore, $L(M)$ is infinite.

For the other direction, assume $L(M)$ is infinite. Therefore, there must exist at least one string whose length is ≥ 2 (the number of strings whose length is $< n$ is $\sum_{i=0}^n |\Sigma|^i$, which is finite). Let w be a shortest such a string. If $|w| < 2n$, then we are done. Otherwise, assume $|w| = p \geq 2n$. By the Pumping Lemma, $w = xyz$, where $|xy| \leq n$, $|y| \geq 1$, and $xy^iz \in L$. In particular, $w' = xy^0z = xz \in L$. Since $|y| \leq n$, then $|xz| \geq n$. Since $w = xyz$ is a *shortest* string that satisfies $|w| \geq 2n$, then it must be that $|xz| < 2n$. Therefore, $w' \in L(M)$ and $n \leq |w| < 2n$. This completes the proof.

4. (a) L_1 is not regular. Let $L'_1 = L_1 \cap a^+b^*c^*$; then,

$$L'_1 = \{a^ib^jc^j : i \geq 1 \text{ and } j \geq 0\}.$$

We prove that L'_1 is not regular, from which it follows that L_1 is not regular (otherwise, L'_1 would be the intersection of two regular languages, and hence itself regular).

Since regular languages are closed under the *Reverse* operation, it suffice to prove that L'^R_1 is not regular. Assume L'^R_1 were regular, and let k be the constant. Choose string $w = c^kb^ka$, which is clearly in L'^R_1 and satisfies $|w| \geq k$. Any way of writing $w = xyz$ such that $|xy| \leq k$ and $|y| \geq 1$ would have $x = c^p$, $y = c^q$, and $z = c^{k-p-q}b^ka$, where $p+q \leq k$ and $q \geq 1$. If we take $i = 0$, then $xy^iz = c^{k-q}b^ka$, and since $q \neq 0$, it follows that $xy^iz \notin L'^R_1$; a contradiction, and hence L'^R_1 is not regular.

- (b) L_2 is regular. A regular expression E such that $L(E) = L_2$ is

$$1(00)^*.$$

- (c) L_3 is not regular. Assume L_3 were regular, and let k be the constant of the Pumping Theorem. Choose $w^{k!}$, which is clearly in L_3 and satisfied $|w| \geq k$. For every way of writing $w = xyz$, we have $x = a^p$, $y = a^q$, and $z = a^{k!-p-q}$, such that $p+q \leq k$ and $q \geq 1$. Choose $i = 2$. Then, $xy^iz = a^{k!+q}$. Since $q \geq 1$ and $q \leq n$, it follows that

$$k! < k! + 1 \leq k! + q \leq k! + k < (k+1)!.$$

Therefore, $k! + q$ is not of the form $m!$ for any m . Hence, $xy^iz \notin L_3$; a contradiction, and hence L_3 is not regular.

- (d) L_4 is regular. A regular expression E such that $L(E) = L_4$ is

$$(a \cup b)^*.$$

In other words, this is the language of all strings over $\{a, b\}$.

- (e) L_5 is not regular. Assume it were regular, then $\overline{L_5}$ would be regular as well. We prove $\overline{L_5}$ is not regular. Assume it were, and let k be the constant of the Pumping Theorem. Choose string a^pb^p where $p > k$ is a prime number. Any way of writing $w = xyz$ such that $|xy| \leq k$ and $|y| \geq 1$ would have $x = a^r$, $y = a^s$, and $z = a^{p-r-s}b^p$, where $r+s \leq k$ and $s \geq 1$. If we take $i = 0$, then $xy^iz = a^{p-s}b^s$. We know that $p-s > 1$ and $p-s < p$, and since p is a prime number, it follows that $\gcd(p, p-s) = 1$, and hence $xy^iz \notin \overline{L_5}$. A contradiction, and hence $\overline{L_5}$ is not regular.