COMP 202 Formal Methods — Exam 1 Solutions

There are 100 points possible. When a problem involves several steps in applying a procedure, it will be to your advantage to show your work so that proper partial credit can be assigned in the event of a minor error in applying the procedure. (Problems 3 and 4 are of this sort.)

1. Using and understanding regular expressions. (15 points total)

1(a). (5 points) Consider the regular expression

$$r = ((a + \lambda)a^*b)^*a$$

For each of the following words, determine whether the word belongs to the regular language described by r.
(1) the word λ
(2) the word a
(3) the word b^3a
(4) the word $baba$
(5) the word $abab$

Solution. The words a, b^3a and $baba$ belong to $L(r)$. The words λ (the empty word) and $abab$ do not belong to $L(r)$.

1(b). (10 points) Give a regular expression that describes the language L (over the alphabet $\Sigma = \{a, b\}$) consisting of all words that *do not* begin with aab. That is,

$$L = \{x \in \{a, b\}^* : x \neq aaby,\ y \in \Sigma^*\}$$

Examples of words that belong to L include: $\lambda, abab, bb$ and $baab$. Examples of words that do not belong to L include: $aab, aabab, aabbb$ and $aabbbab$.

Solution. One way to solve this is to consider all possible words of length at most 3 organized as a trie of the possible prefixes of words. From this one reads off the solution:

$$\lambda + a + a^2 + (b + ab + a^3)(a + b)^*$$

2. Designing a DFA (20 points). Find a DFA that recognizes the language L over the alphabet $\Sigma = \{a, b\}$ consisting of all words that avoid having three consecutive a's:

$$L = \{x \in \Sigma^* : x \neq uaaav,\ u, v \in \Sigma^*\}$$

Examples of words that belong to L include: $aaba, aabbaa, bbb$ and $abab$. Examples of words that do not belong to L include: $aaa, bbabaaab, bbaaab$ and $babaaabbbb$. Make every state of your DFA explicit (i.e., no implicit *dead state*). Hint: your DFA should have 4 states.

Solution. A solution DFA is given by the following information:
(1) The set of states is $Q = \{0, 1, 2, 3\}$.

(2) The start state is 0.
(3) The accept states are $F = \{0, 1, 2\}$.
(4) The transition function is described by the table:

	a	b
0	1	0
1	2	0
2	3	0
3	3	3

3. Translating a regular expression into an NFA (20 points). Give a diagram of an NFA that accepts the regular language described by the regular expression in 1(a).

Solution. A solution NFA is given by the following information:
(1) The set of states is $\{0, 1, 2, 3\}$.
(2) The start state is 0.
(3) The only accept state is 3.
(4) The transition relation is described by the table:

	λ	a	b
0	$\{2\}$	$\{1\}$	$\emptyset$
1	$\emptyset$	$\{1\}$	$\{2\}$
2	$\{0\}$	$\{3\}$	$\emptyset$
3	$\emptyset$	$\emptyset$	$\emptyset$

4. Converting an NFA to a DFA (20 points). Find a DFA that accepts the same language as the NFA described by the following information:
(1) The set of states of the NFA is $Q = \{1, 2, 3, 4, 5\}$.
(2) The start state of the NFA is 1.
(3) The set F of accept states of the NFA is $F = \{1\}$.
(4) The transition relation is given by the following table:

	λ	a	b
1	$\{2\}$	$\{5\}$	$\{1\}$
2	$\{3\}$	$\{4\}$	$\emptyset$
3	$\emptyset$	$\emptyset$	$\{1\}$
4	$\{5\}$	$\emptyset$	$\{3\}$
5	$\emptyset$	$\{2\}$	$\{4\}$

It is easy to make mistakes, so show your work in calculating the DFA using the subset construction. If time allows, it is a good idea to test your construction. (Hint: a correct solution has 6 states.)

Solution. A solution is given by the following information:
(1) The states of the DFA are: $[123], [45], [23], [345], [12345]$ and $[2345]$.
(2) The start state is $[123]$.

(3) The accept states are [123] and [12345].
(4) The transition function is given by the following table:

	a	b
[123]	[45]	[123]
[45]	[23]	[345]
[23]	[45]	[123]
[345]	[23]	[12345]
[12345]	[2345]	[12345]
[2345]	[2345]	[12345]

5. Understanding and using the Myhill-Nerode Theorem to demonstrate nonregularity. (15 points total)

5(a) (6 points) Let L be the language over the alphabet $\Sigma = \{a, b\}$ consisting of all words x for which the number of a's in x is equal to the number of b's in x:

$$L = \{x \in \Sigma^* : n_a(x) = n_b(x)\}$$

Examples of words that belong to L include: $abba, abbaab, \lambda$ and $bbabaaba$. Examples of words that do not belong to L include: bba, aba, aaa and $ababa$. For each of the following pairs of words x and y, determine if xR_Ly.
(1) $x = aa$ and $y = aaa$
(2) $x = aab$ and $y = ababa$
(3) $x = \lambda$ and $y = ab$

Solution.
(1) aa and aaa are not R_L related: e.g., $aabb \in L$ while $aaabb \notin L$.
(2) aab and $ababa$ are R_L related.
(3) λ and ab are R_L related.
It's not too hard to see from the definition of L and the definition of R_L that xR_Ly if and only if

$$n_a(x) - n_b(x) = n_a(y) - n_b(y)$$

5(b) (9 points) Let L for this problem denote the language over the alphabet $\Sigma = \{a, b, c\}$ consisting of all words x such that the number of a's in x is equal to the sum of the number of b's in x and the number of c's in x:

$$L = \{x \in \Sigma^* : n_a(x) = n_b(x) + n_c(x)\}$$

Give a proof using the Myhill-Nerode Theorem that L is not regular.

Solution. To make an argument using the MN Theorem, there are two steps: (1) choose an infinite set of words S, and (2) argue that any two words in S can be distinguished (i.e., are not in the same R_L equivalence class.

There are several different ways to do this for this language. One simple possibility is to take $S = \{a^n : n \geq 0\}$.

If $i \neq j$ then $a^i b^i \in L$ and $a^j b^i \notin L$. QED

6. Using the Myhill-Nerode Theorem to demonstrate regularity (10 points). In this problem we consider the alphabet to be $\Sigma = \{a, b\}$. The language L that we are interested in consists of all words x such that the first three letters of x are the same as the last three letters of x. (If the length of x is less than three, for example $x = ba$, then we define first3(x) and $last3(x)$ by using a special symbol B to denote a blank position. Thus first3$(ba) = baB$ and last3$(ba) = Bba$ and first3$(\lambda) = BBB =$ last3(λ). Examples of words that belong to L include: $\lambda, aaa, bbb, aabbaab$ and $ababa$. Examples of words that do not belong to L include: $abbaab, bababbb, baab$ and $abaab$.

In order to prove that this language is regular we argue as follows. Define the relation R on Σ^* by xRy if and only if first3$(x) =$ first3(y) and last3$(x) =$ last3(y). It is easy to argue that this is an equivalence relation. In order to apply the Myhill-Nerode Theorem to conclude that L is regular, we have to argue two more things:
(1) That R is a right congruence, that is, for any $x, y, z \in \Sigma^*$, if xRy then $xzRyz$.
(2) That L is a union of equivalence classes of R, that is, if xRy and $x \in L$ then $y \in L$.

Give arguments for these two points. (Hint: both arguments are easy and short if you understand what is required.)

Solution.
(1) Suppose xRy and $z \in \Sigma^*$. Then by the definition of R, the first three letters of x and y are the same and so the first three letters of xz and yz are also the same. Also, since the last three letters of x and y are the same, it is also the case that the last three letters of xz and yz are the same. So $xzRyz$ and therefore R is a right congruence.
(2) Since $x \in L$, first3$(x) =$ last3(x). Since xRy, first3$(x) =$ first3(y) and last3$(x) =$ last3(y). Therefore first3$(y) =$ last3(y) and so $y \in L$.