School of Mathematical and Computing Sciences
COMP 202: Formal Methods of Computer Science

Test Solutions

1. Understanding regular expressions (20 points total).

1(a). (6 points) In each case, find a simpler regular expression that defines the same language as the given regular expression.

(i) $(a \mid b \mid ab \mid ba)^*$
simplifies to $(a \mid b)^*$

(ii) $(aab \mid \epsilon)^*$
simplifies to $(aab)^*$

1(b). (6 points) Consider the two regular expressions:

$$R_1 = 0^* \mid 1^* \quad \text{and} \quad R_2 = 01^* \mid 10^* \mid 1^*0 \mid (0^*1)^*$$

(i) Find a string that is in $L(R_1)$ but not in $L(R_2)$.
00, or 000, or 0000 ...

(ii) Find a string that is not in $L(R_1)$ and not in $L(R_2)$.
1100, or 010

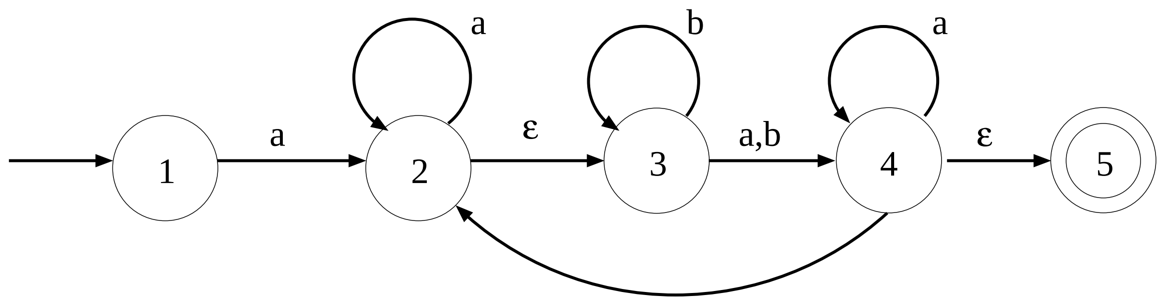
1(c). (8 points) Find a regular expression corresponding to each of the following subsets of $\{0, 1\}^*$.

(i) The language of all strings containing both 11 and 010 as substrings.
$(0|1)^*(11(0|1)^*010 \mid 010(0|1)^*11)(0|1)^*$

(ii) The language of all strings that do not begin with 110.
$\epsilon \mid 1 \mid 11 \mid (0|10|111)(0|1)^*$

2. Understanding finite automata (16 points total).

Consider the transition diagram of an ϵ-NFA shown here.



2(a). (6 points) For each string below, say whether or not the ϵ-NFA accepts it.

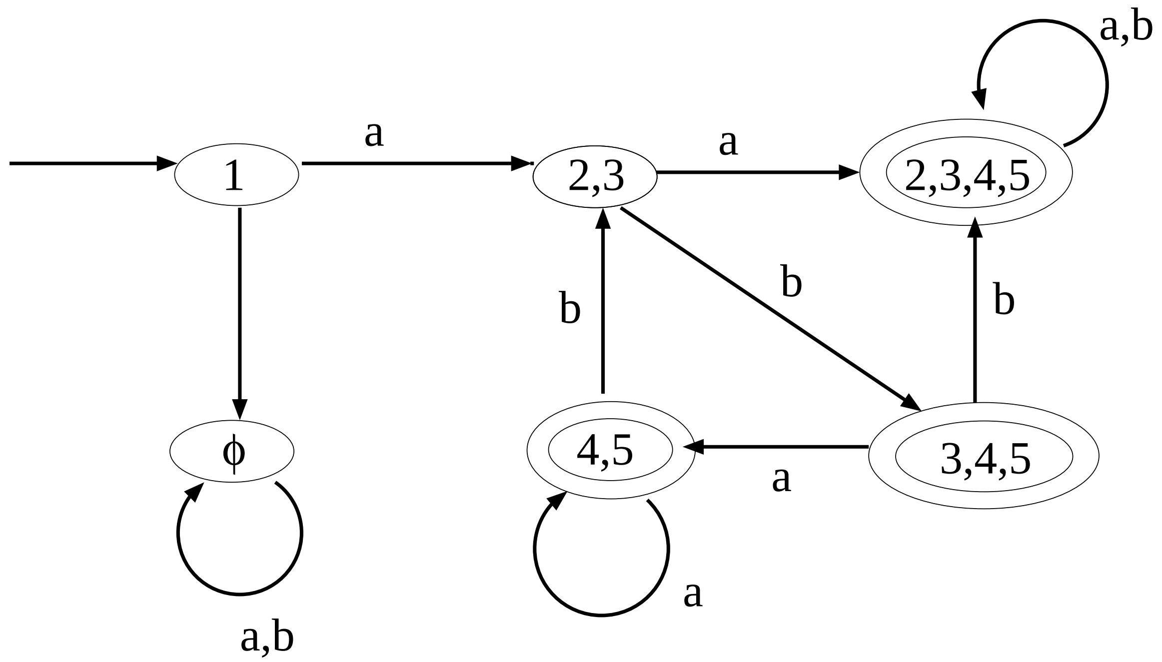
(i) *aba* accepted

(ii) *abab* not accepted

(iii) *aaabbb* accepted

2(b). (10 points) Using the extended version of the subset construction, where we calculate the ϵ-closure of the current state at each step, find a DFA that accepts the same language as this ϵ-NFA.

Transition diagram:



3. Understanding properties of regular languages (14 points total).

3(a). (6 points) Let L be the regular language over the alphabet $\{a, b\}$ represented by the regular expression $(aa \mid bb)^*(ab \mid ba)$.

For each of the following pairs of strings x and y, determine whether x and y are **distinguishable** or **indistinguishable** with respect to L. Give justification for your answers.

x and y are **distinguishable** with respect to L iff $\exists z \in \Sigma^*$ such that $(xz \in L, yz \notin L)$ or $(xz \notin L, yz \in L)$.

x and y are **indistinguishable** with respect to L iff $\forall z \in \Sigma^*(xz \in L \Longleftrightarrow yz \in L)$.

(i) $x = aa$ and $y = aaa$
distinguishable
choose $z = b$, for instance, then $xz = aab \notin L$ and $yz = aaab \in L$

(ii) $x = ab$ and $y = ba$
indistinguishable
$\forall z \in \Sigma^*$ either $xz \in L, yz \in L$ or $xz \notin L, yz \notin L$
if $z = \epsilon$ then $xz \in L, yz \in L$, if $z \neq \epsilon$ then $xz \notin L, yz \notin L$

3(b) (8 points)

Let L be the language over $\{0, 1\}^*$ defined as follows:

$L = \{0^i 1^j \mid j = i \text{ or } j = 2i\}$

To prove that L is not regular by showing that any two distinct elements of the infinite set $\{0^n \mid n \geq 0\}$ are distinguishable with respect to L:

Choose any pair of elements from the set, $0^i, 0^j, i \neq j$. Without loss of generality, we can assume that $i < j$.

Now choose $z = 1^i$, $xz = 0^i 1^i$, so $xz \in L$, $\quad yz = 0^j 1^i, i < j$, so $yz \notin L$.

To prove that L is not regular using the pumping lemma:

Given n, choose $x = 0^n 1^n$, $x \in L$.

For any u, v, and w with $x = uvw$, $|uv| \leq n$ and $|v| > 0$, we have $uv = 0^s$, $v = 0^k$, $0 < k \leq s \leq n$, and $w = 0^r 1^n$, where $s + r = n$.

Now choose $i > 1$, say $i = 2$. $\quad uv^i w = 0^{n+k} 1^n$, so $uv^i w \notin L$.

Suppose we chose $x = 0^n 1^{2n}$, again $x \in L$, and for any u, v, and w with $x = uvw$, $|uv| \leq n$ and $|v| > 0$, we have $uv = 0^s$, $v = 0^k$, $0 < k \leq s \leq n$, and $w = 0^r 1^n$, where $s + r = n$.

Note, however, that k can be any number between 1 and n, so we have to be careful pumping in more zeros.

If we choose $i > n$, then $uv^i w = 0^{n+ik} 1^{2n}$, where ik is at least $n + 1$, so $uv^i w \notin L$.

Alternatively, we can choose $i = 0$, $uv^i w = 0^{n-k} 1^{2n}$, so $uv^i w \notin L$.