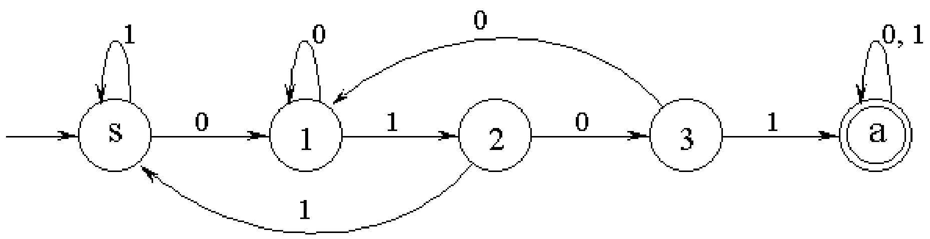
Solution to HW1 of Honors Theory

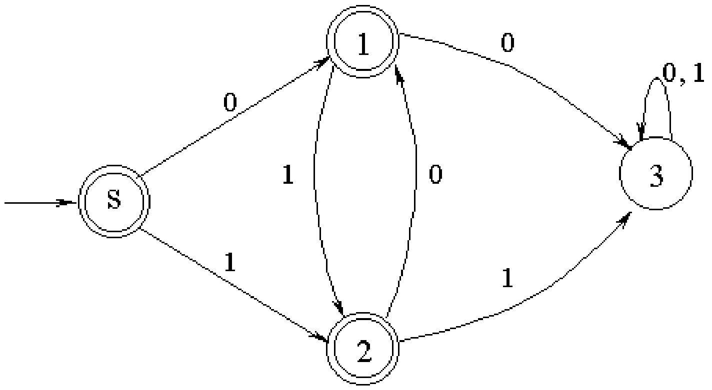
February 22, 2002

1. (10+10 Points)

   (i)



   (ii)



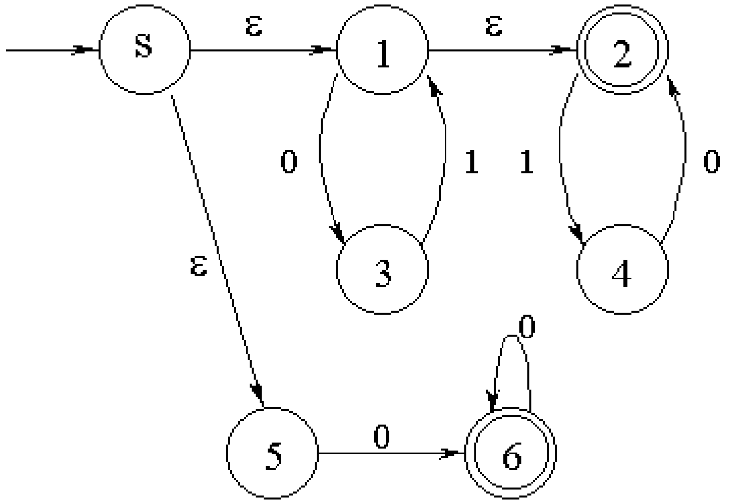
2. (10+10 Points)

   (i) $(0+1)*0101(0+1)*$

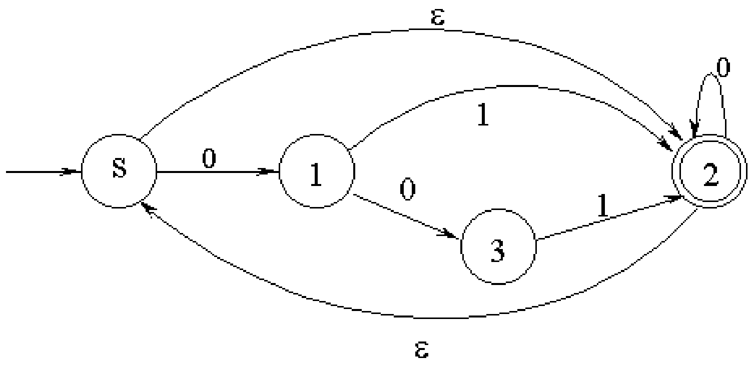
   (ii) $(01)*+(01)*0+(10)*+(10)*1$

3. (10+10 Points)

(i)

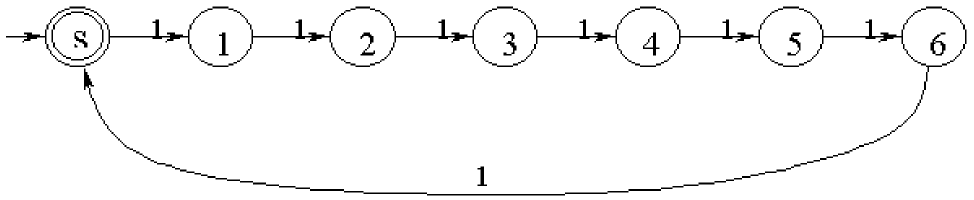


(ii)



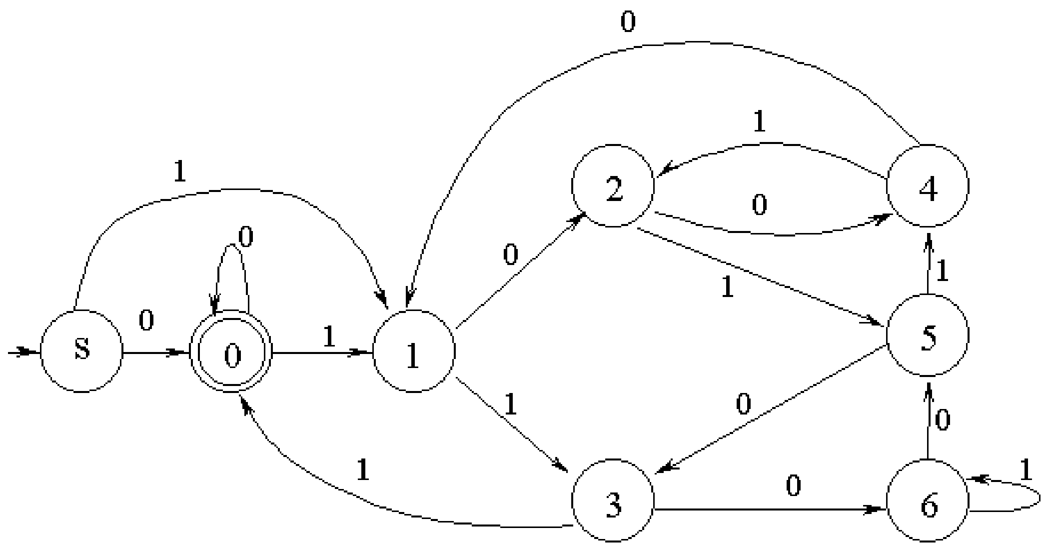
4. (10+10+10+10 Points)

(i) It's regular, as there is a NFA M1 that recognizes it.
The regular expression (1111111)$*$ describes it.



NFA M_1

(ii) It's regular, as there is a DFA M2 that recognizes it.



DFA M_2

To understand this, note that the state labeled i (for $i = 0, 1, \ldots, 6$) indicates the binary string read so far is equal to $i \bmod 7$. It is easy to verify this for each of the arrows. For instance, from state 1, if we see 0, then we go to state 2. This is because if the current string w has a binary value of $7m + 1$, then the new string $w0$ has value $14m + 2$ which is congruent to $2 \bmod 7$.

Most student's solution is slightly defective in that their dfa accepts the empty string, which does not represent any binary number.

(iii) Use the pumping lemma to prove the language A is not regular. Assume to the contrary that the language A is regular. Let p be the pumping length given by the pumping lemma. Choose s to be the string $0^p10^m10^{p+m}$. Because s is a member of A and s has length more than p, the pumping lemma guarantees that s can be split into three pieces, $s = xyz$, where for any $i \geq 1$, the string xy^iz is in A. By conditin 3, y consists only of 0's, let $y = 0^k, k > 0$. The string $xy^2z = 0^{p+k}10^m10^{p+m}$ can not be a member of A, as $p + k + m \neq p + m$ for any $k > 0$. Thus we obtain a contradiction.

(iv) Use the pumping lemma to prove the language B is not regular. Assume to the contrary that B is regular. Let p be the pumping length given by the pumping lemma. Choose s to be the string 0^p1^p. Because s is a member of B and s has length more than p, the pumping lemma guarantees that s can be split into three pieces, $s = xyz$, where for any $i \geq 0$ the string xy^iz is in B. By condition 3, y consists only 0's, let it be 0^k where $k > 0$. The string $xy^2z = 0^{p+k}1^p$ can not be a member of B. Thus we obtain a contradiction.

5. (20+20 Points)

(i) If A is regular, then some regular expression R describes it. We construct a regular expression R' from R by replacing each symbol $a \in \Sigma$ in R with the string $h(a)$. We show R' describes $h[A]$ in both directions, i.e, for any $w \in A, h(w) \in L(R')$; and for any $w' \in L(R'), w' \in h[A]$.
1)If $R = \emptyset$ which means A and $h[A]$ are empty sets, then $R' = \emptyset$, so $h[A] = L(R')$.
If $R = \epsilon$ which means $A = \{\epsilon\}$, and $h[A] = \{h(\epsilon)\} = \{\epsilon\}$, then $R' = \epsilon$, so $h[A] = L(R')$.
If $R = a$ where $a \in \Sigma$, which means $A = \{a\}, h[A] = \{h(a)\}$, then $R' = h(a)$, so $h[A] = L(R')$.
2) Assume for any regular expression R with $|R| \leq k$, let R describe a regular language A, then the regular expression R' constructed from R describes $h[A]$. We need to show that for any regular expression R_n with $|R_n| = k + 1$ or $|R_n| = k + 2$ (let it describe A_n), the regular expression R'_n constructed from R_n describes $h[A_n]$.
a) If $R_n = R^*$, where R is a regular expression of language A with length at most k, then $R'_n = R'^*$:
We already know $L(R') = h[A]$, which implies
$w' \in h[A]$ for any $w' \in L(R')$, and $h(w) \in L(R')$ for any $w \in A$.
For each $w_n = w^* \in L(R_n)$ where $w \in L(R)$, we have
$h(w_n) = h(w^*) = (h(w))^*$, which means $h(w_n) \in L(R'_n)$;
For each $w'_n \in L(R'_n)$, we have $w'_n = w'^* = h(w)^* = h(w^*)$.
That is $w'_n \in h[A_n]$.
So R'_n describes $h[A_n]$.
b) If $R_n = R_1 + R_2$, where R_1, R_2 are regular expressions of languages A_1, A_2, and $|R_1| \leq k, |R_2| \leq k$, then $R'_n = R'_1 + R'_2$:
For each $w_n \in A_1 \cup A_2$,

$h(w_n) \in h[A_1] \cup h[A_2]\} = L(R'_n)$.
For each $w'_n \in L(R'_n) = L(R'_1) \cup L(R'_2)$, $w'_n \in h[A_1] \cup h[A_2] = h[A_n]$.
So R'_n describes $h[A_n]$.
c) If $R_n = R_1 R_2$, where R_1, R_2 are regular expressions of languages A_1, A_2, and $|R_1| \leq k, |R_2| \leq k$, then $R'_n = R'_1 R'_2$:
For each $w_n = w_1 w_2$, where $w_1 \in A_1, w_2 \in A_2$, we have
$h(w_n) = h(w_1 w_2) = h(w_1)h(w_2) \in L(R'_n)$;
For each $w'_n \in L(R'_n)$, we have
$w'_n = w'_1 w'_2 = h(w_1)h(w_2) = h(w_1 w_2)$ where $w'_1 \in h[A_1], w'_2 \in h[A_2]$,
that is $w'_n \in h[A_n]$.
So R'_n describes $h[A_n]$.
Therefore for any regular expression R of language A, the constructed regular expression R' describes $h[A]$.

(ii) If B is regular, then there is a DFA $M = (Q, \Gamma, \delta, q_0, F)$ that accepts B. We construct DFA $M' = (Q, \Sigma, \delta', q_0, F)$ with
$\delta'(q, a) = \delta^*(q, h(a))$ for any $q \in Q, a \in \Sigma$.

We show that $\delta'^*(q,w) = \delta^*(q,h(w))$ by induction.
1. $|w| = 0$, $\delta'(q,\epsilon) = q = \delta(q,\epsilon) = \delta(q,h(\epsilon))$
$|w| = 1$, $\delta'(q,a) = \delta^*(q,h(a))$
2. Assume $\delta'^*(q,w) = \delta^*(q,h(w))$ is true for every w with $|w| \leq k$
For any $w' = wa$ with $|w'| = k+1$
$\delta'^*(q,w')$=$\delta'^*(q,wa)$=$\delta'(\delta'^*(q,w),a)$
$= \delta'(\delta^*(q,h(w)),a) = \delta^*(\delta^*(q,h(w)),h(a))$
$= \delta^*(q,h(w)h(a)) = \delta^*(q,h(wa))$
$= \delta^*(q,h(w'))$.

$\delta'^*(q_0,w) = \delta^*(q_0,h(w))$ for every string w
$\Leftrightarrow w \in L(M')$ iff $h(w) \in L(M) = B$
$\Leftrightarrow M'$ recognizes $h^{-1}[B]$.