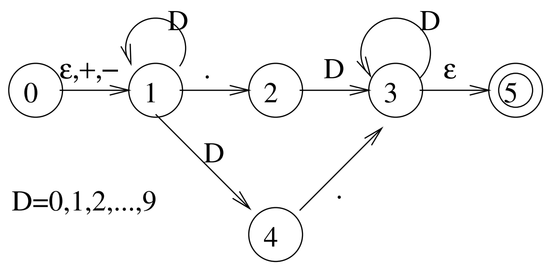
Name ______________________ Student No. ______________________

No aids allowed. Answer all questions on test paper. Use backs of sheets for scratch work.

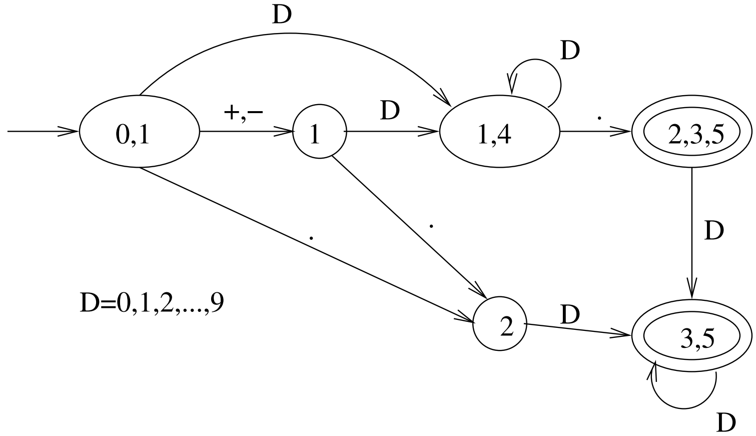
Total Marks: 40

[10] 1. Consider the following ε-NFA for accepting decimal numbers:



Construct the corresponding DFA using the *subset construction* and *ε-closure* technique. Present your result in diagram form.

Hint: Don't compute all 2^6 subsets of states! Start with ε-closure(q_0) to compute the initial state of your DFA, and continue from there to see which subsets of states will be necessary. The result should not have more than about 6 states in total.


D
D
+,−
D
.
0,1
1
1,4
2,3,5
.
.
D
D=0,1,2,...,9
2
D
3,5
D

[10] 2. Let R, S be regular expressions. Prove or disprove each of the following statements.

(a) $(RS + R)^*R = R(SR + R)^*$

TRUE: Use "the Test"; $(ab + a)^*a = a(ba + a)^*$.
We now show that $L((ab+a)^*a) = L(a(ba+a)^*)$. Suppose that $x \in L((ab+a)^*a)$. Then, if x has no (ab)'s in it, $x = a^i$, $i > 0$, so $x \in L(a(ba+a)^*)$. Suppose that x has (ab)'s in it. Then, $x = a^i(ab)z(ab)a^j$, where $i \geq 0$, and $j > 0$. Now, using induction on the number of (ab)'s in the middle segment, or just by an "observation", we say $x = a^{i+1}\underline{(bza)(ba)a}^{j-1}$ where the underlined segment is in $L((ba + a)^*)$. Thus, $x \in L(a(ba+a)^*)$. An analogous argument in the other direction shows the equality of the two languages, and finishes the proof of the statement.

(b) $(R + S)^*S = (R^*S)^*$

FALSE: Use "the Test"; $(a+b)^*b = (a^*b)^*$. Now, ε does not belong to the language of the left-hand side expression, since there every string has to end with a b. However, ε belongs to the language of the right-hand side expression, by definition of "$*$".

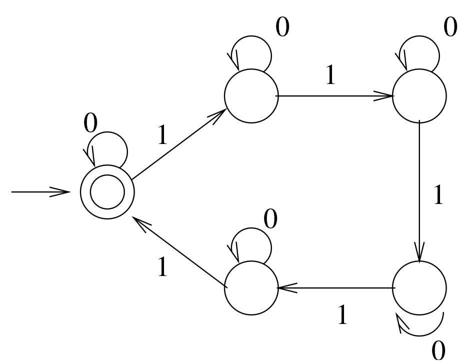
[10] 3. Which of the following languages (over $\Sigma = \{0, 1\}$) is regular? Prove your assertion.

(a) $L_1 = \{w|\ w$ has as many 1s as 0s $\}$.
For example, $010110, 1100 \in L_1$, but $110 \notin L_1$.

L_1 is not regular. Suppose L_1 is regular. Then, by the PL, there exists an n such that $0^n 1^n \in L_1$, and there exist x, y, z such that $xyz = 0^n 1^n$, $|xy| \leq n$, $y \neq \varepsilon$, and $xy^k z \in L_1$ for all k. Since $|xy| \leq n$, $xy = 0^i 0^j$, $i + j \leq n$, and $j \neq 0$. So, $xy^k z = 0^i 0^{kj} 1^n$. Now, let k be big enough so that $i + kj > n$ (i.e., let k be an integer bigger than $(n - i)/j$). Hence, $xy^k z$ has more zeros than 1s, so it is not in L_1. Contradiction, so L_1 is not regular.

(b) $L_2 = \{w|$ the nr of 1s in w is a multiple of 5 $\}$.
For example, $00, 1111111111, 1011011 \in L_2$, but $00110 \notin L_2$.

L_2 is regular. We present below the DFA recognizing L_2.



[10] 4. Give an algorithm to tell whether a regular language L contains at least 100 strings.

Hint: Use the following result shown in class: given a DFA A with n states, $L(A)$ is infinite $\iff$ A accepts a string x of length $n \leq |x| \leq 2n$.

Our procedure works as follows: given a regular language L, we consider the DFA A that accepts it; that is, $L(A) = L$. Suppose that A has n states.

First, test if A accepts a string x of length $n \leq |x| \leq 2n$. Given n, there are finitely many such strings, so after finitely many steps, we will know if A accepts such a string or not.

If A accepts some x_0 such that $n \leq |x_0| \leq 2n$, then we know (by the hint) that $L(A)$ is infinite, so in particular $L(A) = L$ has at least 100 strings, so we answer YES.

If A does not accepts any string x of length $n \leq |x| \leq 2n$, then $L(A) = L$ is finite. Furthermore, by the PL we know that A does not accept any string of length $\geq n$, since if it did, $L(A) = L$ would be infinite. So we test all the strings x of length $|x| \leq n$, and we count how many are accepted. If at least 100 are accepted we answer YES, otherwise we answer NO.