CS 475/575 Exam 2 **Spring 2012** **Solution**

Your score on this exam is based on your best 7 of these 8 problems.

1. Write a regular expression that generates the set of strings over alphabet {a, b} that do ***not*** contain two consecutive identical characters. That is, substrings aa and bb are forbidden. Remember to include both the cases when the string's length is odd and when it is even. Hint: write out some of the strings that belong in this language.

   $(ab)^* \cup (ab)^*a \cup (ba)^* \cup (ba)^*b$

   Alternative solution:

   $(\varepsilon \cup b)(ab)^*(\varepsilon \cup a)$

2. Determine whether each identity below is true or false, and provide convincing justification.

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

      True. Both sides can generate all strings over alphabet {a, b}. That is, both sides are equivalent to $(a \cup b)^*$.

   b. $(a^+ b^+)^+ = (a^+ \cup b^+)^+$

      False. The right side can generate strings {a, b, aa, bb, ba} but the left side cannot generate any of these strings.

3. Write a regular expression that generates the set of strings that represent number literals such as shown in the examples below. Use alphabet Σ = {d, e, +, –, .} where d is for digit.

1234	1234e56	1234e+56	1234e–56
12.34	12.34e56	12.34e+56	12.34e–56
12.	12.e56	12.e+56	12.e–56
.34	.34e56	.34e+56	.34e–56
+1234	+1234e56	+1234e+56	+1234e–56
+12.34	+12.34e56	+12.34e+56	+12.34e–56
+12.	+12.e56	+12.e+56	+12.e–56
+.34	+.34e56	+.34e+56	+.34e–56
–1234	–1234e56	–1234e+56	–1234e–56
–12.34	–12.34e56	–12.34e+56	–12.34e–56
–12.	–12.e56	–12.e+56	–12.e–56
–.34	–.34e56	–.34e+56	–.34e–56

$\mathbf{(+ \cup - \cup \varepsilon)\ (d^+ \cup d^+ . d^* \cup . d^+)\ (e\ (+ \cup - \cup \varepsilon)\ d^+ \cup \varepsilon)}$

4. Write a regular expression that generates the set of strings that do ***not*** represent number literals over alphabet Σ = {d, e, +, –, .} where d is for digit.
Hint: any string over Σ is a number literal unless one or more of the following occurs:
- The string has one of the characters {d, +, –, .} immediately before one of {+, –}
- The string has at least two e's, or has at least two .'s, or has . somewhere after e
- The string has no d
- The string contains e, but either has no d before e or has no d after e

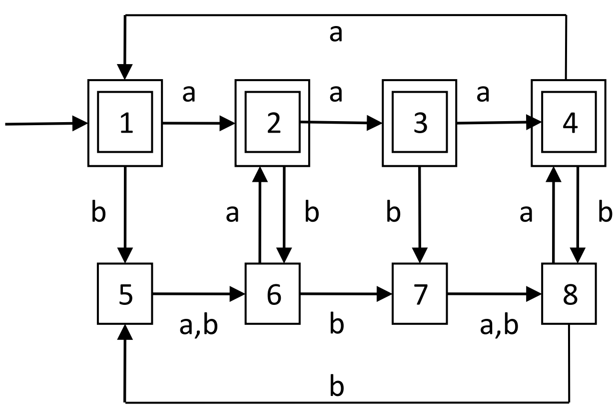
Examples: 1e++2 +3–4 5e6e7 8.9.0 1e2.3 –. +.e–456 –7.89e+

$\mathbf{\Sigma^*\ (d \cup + \cup - \cup .)\ (+ \cup -)\ \Sigma^*}$
$\mathbf{\cup\ \Sigma^*\ e\ \Sigma^*\ e\ \Sigma^*\ \cup\ \Sigma^*\ .\ \Sigma^*\ .\ \Sigma^*\ \cup\ \Sigma^*\ e\ \Sigma^*\ .\ \Sigma^*}$
$\mathbf{\cup\ (e \cup + \cup - \cup .)^*}$
$\mathbf{\cup\ (+ \cup - \cup .)^*\ e\ \Sigma^*\ \cup\ \Sigma^*\ e\ (+ \cup - \cup .)^*}$

5. Let $L = \{ a^m b^n \mid m \geq n \}$. Prove that L is not a regular language.

   Suppose L is regular.
   Let p = the pumping theorem constant.
   Let $s = a^p b^p \in L$.
   Then s = uvw such that $|uv| \leq p$, $|v| \geq 1$, and $u v^k w \in L$ for all $k \geq 0$.
   Hence $u = a^q$, $v = a^r$, and $w = a^{p-q-r} b^p$, where $q+r \leq p$ and $r \geq 1$.
   When k = 0, $u v^k w = u w = a^q a^{p-q-r} b^p = a^{p-r} b^p \notin L$, because $r \geq 1$ implies $p-r < p$.
   Contradiction, so L is not regular.

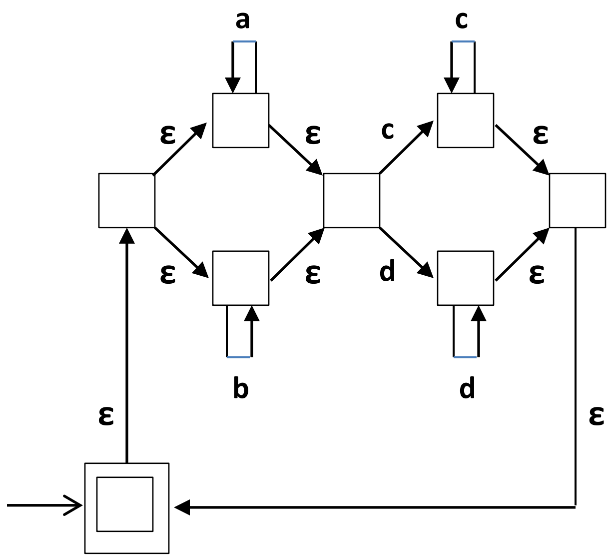
6. Given the deterministic finite-state machine shown below, complete the table of distinguishabilities, and specify which states are equivalent.



2	**X**	–	–	–	–	–	–
3	**O**	**X**	–	–	–	–	–
4	**X**	**O**	**X**	–	–	–	–
5	**X**	**X**	**X**	**X**	–	–	–
6	**X**	**X**	**X**	**X**	**X**	–	–
7	**X**	**X**	**X**	**X**	**O**	**X**	–
8	**X**	**X**	**X**	**X**	**X**	**O**	**X**
	1	2	3	4	5	6	7

Equivalent states are {1, 3}, {2, 4}, {5, 7}, {6, 8}.

7. Draw a non-deterministic finite-state machine that is equivalent to this regular expression:
$((a^* \cup b^*)\ (c^+ \cup d^+))^*$



8. Write a regular expression that is equivalent to this finite-state machine: