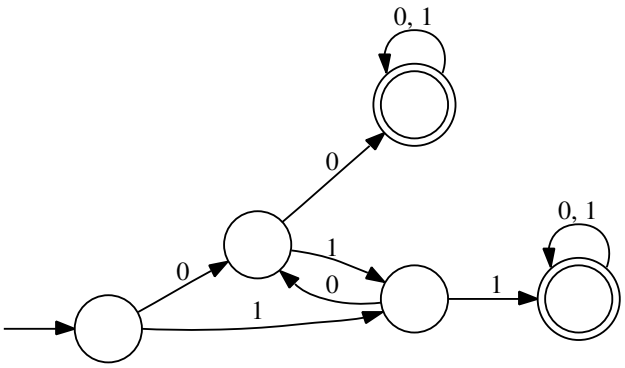
Written Assignment 1
Due Thursday, January 25, 2007

This assignment asks you to prepare written answers to questions on regular languages, finite automata, and lexical analysis. Each of the questions has a short answer. You may discuss this assignment with other students and work on the problems together. However, your write-up should be your own individual work. Written assignments can be turned in at the start of lecture. Alternatively, assignments can be turned in at Professor Aiken's office in Gates 411, or submitted electronically in PDF format by following the electronic submission instructions at `http://www.stanford.edu/class/cs143/policies/submit.html`, by 5:00 PM on the due date.

1. Consider the following deterministic finite automaton (DFA) over the alphabet $\Sigma = \{0, 1\}$.



   Give a one-sentence description of the language recognized by the DFA. Write a regular expression for this language.

2. Consider the following languages over the alphabet $\Sigma = \{0, 1\}$.

   - L_1: All strings that contain at least two 0s
   - L_2: All strings that contain at least one 1
   - L_3: All strings that contain at least two 0s and at least one 1
   - L_4: All strings that contain at most one 0 or no 1s

   Give DFAs for each of the languages L_1, L_2, L_3, and L_4.

   Aside: This example illustrates that the regular languages are closed under intersection and complementation. Note that $L_3 = L_1 \cap L_2$ and $L_4 = \Sigma^* - L_3$, where Σ^* represents the language containing all strings over the alphabet Σ.

3. Let E_3 be the language over the alphabet $\Sigma = \{a_1, a_2, a_3\}$ defined as follows.

E_3: All strings in which a_i occurs an even number of times for some $i \in \{1, 2, 3\}$

Give a non-deterministic finite automaton (NFA) for the language E_3.

4. Write regular expressions for the following languages over the alphabet $\Sigma = \{0, 1\}$:

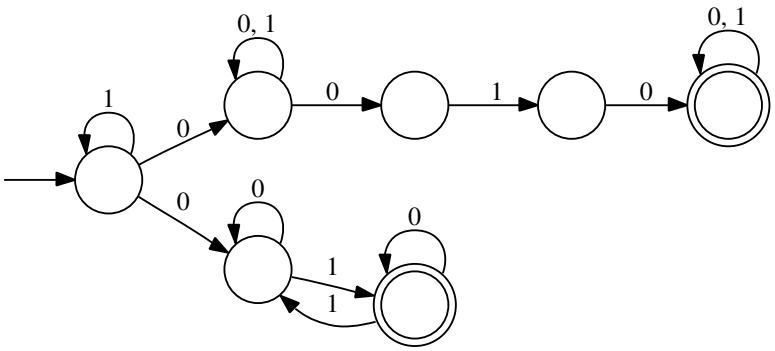
   (a) All strings that contain at least one 0 and at least one 1 and that also end with at least two 1s.

   (b) All strings that do not begin with 01.

   (c) All strings that contain an odd number of 1s.

5. Give a DFA for each of the following languages over the alphabet $\Sigma = \{0, 1\}$.

   (a) The language of the following NFA.



   (b) The language of the regular expression $(0 + 01)^*1^*$.

6. Consider the string

```
aaabaabbbababbb
```

   and its tokenization

```
aa a b aabb a b a bbb
```

   Give a flex specification with the minimum number of rules that produces this tokenization. Each flex rule should be as simple as possible as well. You may not use regular expression union (i.e., $R_1 + R_2$) in your solution. Do not give any actions; just assume that the rule returns the string that it matches.