

Problem Set 1: Regular expressions and finite automata (Solutions)

1. Give a regular expression that can match street addresses. In our definition, a street address consists of a number, followed by an optional letter, for the house/apartment number, then one or more capitalized words for the street name, then either “Dr.”, “Rd.” or “St.” for the street type.

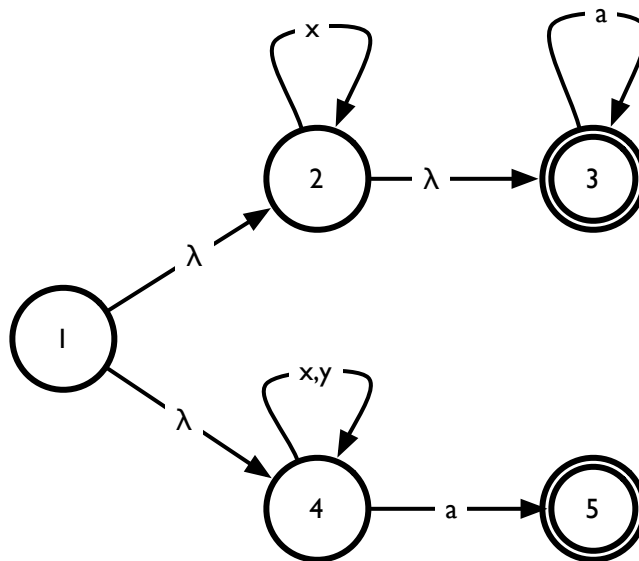
Answer: There are many ways to write a regular expression to capture street addresses. Here is one possible answer, with the symbol σ representing a space:

$$[0 - 9]^+[A - Z a - z]^? \sigma ([A - Z][a - z]^* \sigma)^* (Dr|Rd|St).$$

2. Give a *non-deterministic* finite automaton that matches the following regular expression:

$$((x|y) * a)|(x * a*)$$

Answer: Here is one possible NFA. Note that there are many other possibilities, too.



3. Give a *deterministic* version of the finite automaton, using the construction we described in class. You only need to show the state transition diagram.

Answer: Here is the state transition diagram, built up in the order it would be created during the subset construction algorithm.

State	a	x	y	final?
1, 2, 3, 4	3, 5	2, 3, 4	4	yes
3, 5	3	err	err	yes
2, 3, 4	3, 5	2, 3, 4	4	yes
4	5	4	4	no
3	3	err	err	yes
5	err	err	err	yes

4. Derive the reduced DFA. Show both the graphical representation of the automaton and the state transition diagram.

Answer: We can combine the states $\{3\}$ and $\{3, 5\}$ because they are both final and they behave the same in all cases. We can also combine the states $\{1, 2, 3, 4\}$ and $\{2, 3, 4\}$ because they are both non-final and behave the same in all cases. This gives us the following reduced DFA:

State	a	x	y	final?
2, 3, 4	3	2, 3, 4	4	yes
3	3	err	err	yes
4	5	4	4	no
5	err	err	err	yes

The graphical representation of this DFA is below:

