Computer Science 81, Fall 2006
Assignment 1
Due Thur. Sept. 7
Strings and Regular Languages

I will be in Italy the first class day, but will try to answer questions via email. I'm not sure about the right starting level, but let's try these problems. Don't panic if you can't get them. Just let me know what you think about the level (after doing the reading). Problem 5 will be the least obvious. Also, try some other problems in the book if these seem too hard.

Solve the following problems in conjunction with reading Kozen lectures 1-6 (pages 1-39):

1. Show that, for any set of strings A,
   $$A\,A^* = A^*A$$
   where * is the unary asterate (star) operator.

In 2-4, construct a DFA (deterministic finite-state automaton) that accepts the indicated sets of strings. Don't forget to have your DFA properly classify the null string ε.

2. The set of all strings over $\{a, b\}$ in which each b is followed (but not necessarily immediately followed) by at least two a's. (For example, *bbbaa* is accepted, but *bababa* is not.)

3. The set of all strings of the form $a^n x$ where $n \leq 5$, $x \in \{a, b\}^*$ and each b is followed (in the sense above) by at most n a's. (For example, *aaababababa* is accepted, but *aaaababababaa* is not, because the first b is followed by 4 a's.)

4. The set of all strings over $\{a, b, c\}$ containing either an even number of a's, or both an odd number of b's and an odd number of c's. (For example, *abaca* would be accepted, but *abacaaba* would not.) (Using the product construction might be helpful here.)

5. We say that a string y is an *infix* of a string w, and w is an *outfix* of string y, provided that $w = xyz$ for some strings x and z. If R is a set of strings, the *infixes*(R) is the set of all infixes of strings in R and outfixes(R) is the set of all outfixes of strings in R.

   Establish that if R is any regular set (i.e. is accepted by some DFA), then so are
   a. *outfixes*(R) and
   b. *infixes*(R)

   Here you are allowed to use the result that any set accepted by an NFA (non-deterministic finite-state automaton) is also accepted by some DFA. Using the idea of ε-transitions (Kozen, page 36) might also be helpful.