HW01: Regular Languages and Probability

Hand in at: `http://www.cs.utah.edu/~hal/handin.pl?course=cmsc723`. Remember that only PDF submissions are accepted. We encourage using LaTeX to produce your writeups. See `hw00.tex` for an example of how to do so. You can make a `.pdf` out of the `.tex` by running "pdflatex hw00.tex".

1 Book Exercises

Complete exercises *2.1 (parts 1, 2 and 5), 2.4 and *2.8 from J+M.

Hint: If you want to draw a picture and put it in LaTeX, then use the `\includegraphics` command. For example: `\includegraphics[width=4in]{myimage.jpg}` will do exactly what you think it would. This will work (at least) with jpg and pdf files.

2 Basic Probability

Alice is a student in CL1. In addition to CL1, Alice has found herself taking an "Easy A" class at the same time, taught by Professor Deb. Deb is known for giving out As to almost every student. In fact, only one in every $1,000$ students has ever gotten anything less than an A (Deb has taught a *lot* of classes). When a student *does* get less than an A, they give Deb a bad teaching eval. In particular, of those students that don't get an A, only 5% will *not* give Deb a bad teaching eval. Of course, students often give bad teaching evals regarless of the reason: it happens about a ten percent of the time.

1. Based on the above paragraph, give numeric values for $P(\text{LessThanA})$, $P(\text{BadEval})$ and $P(\text{BadEval} \mid \text{LessThanA})$.

2. What is the probability that a student will give Deb a *good* teaching eval even though they got less than an A? Show your work.

3. After the semester is over, Deb finds a student who gave her a bad teaching eval. What is the probability that Deb gave this student less than an A? Show your work.

3 Additional Exercises

In one or two sentences, answer the following two questions.

1. What is the goal of a language model?

2. Why is it bad to test on the training set?