Homework #1: CMPT-413

Anoop Sarkar – anoop@cs.sfu.ca

Please write down the answers to this homework into a text file and submit a printed copy. If you cannot create a text file and print it, please write the expressions clearly. It is often hard to read regular expressions that are hand-written. Write down your name and student id on the upper right hand corner of your submission. You can use `Perl` to check whether your regular expressions work.

1. (10pts) Write a regular expression to match the following description :

   (a) A single determiner of English (assume that the only determiners are *a*, *an*, *the*) e.g. it should accept "*an*" but should not accept "*it was a blast*" or "*the the*".

   (b) Arithmetic expressions containing integers, addition and multiplication, e.g. $2 + 8 * 3$ or $12 * 8 + 100$

2. (25pts) Do Exercise **2.1** except for **2.1(c)** and **2.1(g)** from *Speech and Language Processing* by Jurafsky and Martin (starts on page 53).

3. (25pts) For the textual examples given below, write down a single regular expression that will accept the examples in the left column and will *not* accept the examples in the right column.

Accept this column	Do not accept this column
1-8-2003	1-8/7
1/8/2003	1/8 = .125
1.8.03	2003401
2003	1.8

4. (40pts) Draw a finite-state transducer that maps the language specified by the regular expression `(ab|aba)+` to the language specified by the regular expression `(01|010)+`.