CSCI 3130: Formal Languages and Automata Theory — Homework 1
The Chinese University of Hong Kong, Fall 2016 — due 11:59pm Thursday September 22

Collaborating on homework is encouraged, but you must write your own solutions in your own words and list your collaborators. Copying someone else's solution will be considered plagiarism and may result in failing the whole course.

Please answer clearly and concisely. Explain your answers. Unexplained answers will get lower scores or even no credits.

(1) (30 points) Give a DFA for the following languages, specified by a transition diagram. For each one of them, give a short and clear description of how the machine works. The alphabet is $\Sigma = \{\texttt{0}, \texttt{1}, \texttt{2}\}$:

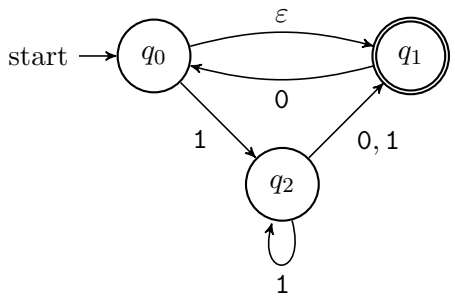
(a) $L_1 = \{w \mid w$ has fewer `0`s than `1`s and at most two `1`s$\}$.

(b) $L_2 = \{w \mid$ the sum of digits of w is divisible by 3$\}$.

(c) L_3 is the language described by $\texttt{0}^*\texttt{1}^*\texttt{2}^*$.

(d) $L_4 = \{w \mid w$ contains the pattern `12` an even number of times$\}$.

(2) (15 points) Convert the following NFA to a DFA using the method described in class. Specify the DFA by its transition diagram. The alphabet is $\Sigma = \{\texttt{0}, \texttt{1}\}$.



(3) (30 points) If w is a string, we say that a string x is an *ending* of w if $w = yx$ for some string y. For example, `bcd` and `d` are both endings of `abcd`. Given a language L, define $L^E = \{x \mid x$ is an ending of some $w \in L\}$. That is, L^E contains the endings of strings in L.

Prove that if L is a regular language, then so is L^E.

Hint: What is the definition of a regular language? If the desired result is true for simpler regular languages, can you show that it is also true for more complex regular languages?

(4) (25 points) In this problem you will design an NFA that checks if an input string correctly represents an integer (say in a spreadsheet application). Examples of correctly formatted integer strings are:

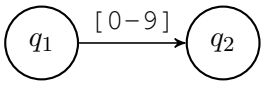
```
10000
12,345,678
-456
```

Examples of incorrectly formatted integer strings are:

```
010000
,123,456,789
4-56
```

For simplicity, assume the input string consist only of digits '`0`' through '`9`', and the special symbols '`,`', and '−'.

When drawing the transition diagram of your NFA, you can use the shorthand notation `[0-9]` to describe transitions labeled by all the digits '`0`', '`1`', ... '`9`'. An example of such a shorthand is



Similarly, you can use the notation `[1-9]` to describe the transitions labeled by all the non-zero digits.

This is a design problem, and part of your job is to figure out a way to distinguish among correct and incorrect integer strings. There may not be a single right answer. You must describe your assumptions and reasoning clearly in your solution. Solutions that only provide a picture of an NFA with no explanation will get no credit.