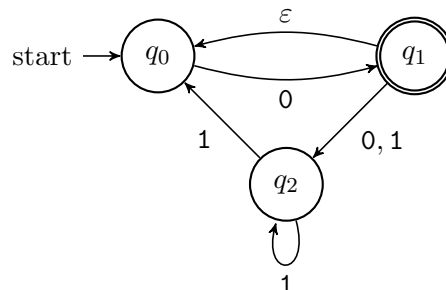


Collaborating on homework is encouraged, but you must write your own solutions in your own words and list your collaborators. Copying someone else's solution will be considered plagiarism and may result in failing the whole course.

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

- (1) (12 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  $\Sigma = \{0, 1, 2\}$ :
  - (a)  $L_1 = \{w \mid w \text{ the sum of the digits in } w \text{ is divisible by } 5\}$ .
  - (b)  $L_2 = \{w \mid w \text{ contains the pattern } 12 \text{ an even number of times}\}$ .
  - (c)  $L_3 = \{w \mid w \text{ begins with } 01 \text{ or ends with } 2 \text{ (or both)}\}$ .
  - (d)  $L_4$  is the language described by  $(0 + 1)^*(1 + 2)^*(0 + 2)^*$ .
- (2) (6 points) Convert the following NFA to a DFA using the method described in class. Specify the DFA by its transition diagram. The alphabet is  $\Sigma = \{0, 1\}$ .



- (3) (12 points) If  $w = w_1w_2 \dots w_n$  is a string, we say that a string  $x$  is an *ending* of  $w$  if  $x = w_iw_{i+1} \dots w_n$  for some  $1 \leq i \leq n$ . For example, `bcd` and `d` are both endings of `abcd`. Given a language  $L$ , define  $L^E = \{x \mid x \text{ is an ending of some } w \in L\}$ . 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) (10 points) In this problem you will design an NFA that checks if a web address is formatted correctly. Examples of correctly formatted web addresses are:

```
http://www.gov.hk/en/about/govdirectory/  
http://www.cse.cuhk.edu.hk/  
http://www.google.com  
http://en.wikipedia.org/wiki/automaton  
http://www.travelchinaguide.com/attraction/hongkong/island/victoria.htm
```

Examples of incorrect addresses are:

```
http://..invalid.domain.name/  
http://no/domain/name/here.html  
http:///too.many.slashes/  
no.protocol.identifier  
the://protocol.must.be.http  
http://www.acme.com/web:page:name:cannot:contain:colons.html
```

In general, web addresses follow the Unified Resource Locator (URL) format, but for this problem you can make some simplifying assumptions about the kind of web addresses you are dealing with. You can assume that the address consists of the protocol identifier `http://`, followed by the domain name (e.g., `www.cse.cuhk.edu.hk`) ending in some suffix like `.com`, `.hk`, possibly followed by a slash `/`, then some optional directory structure (e.g., `en/about/govdirectory/`), and finally an optional web page name (e.g., `victoria.htm`).

For simplicity, assume web addresses consist only of the lowercase letters 'a', 'b', up to 'z' and the special symbols '/', ':', and '.'.

When drawing the transition diagram of your NFA, you can use the shorthand notation `[a-z]` to describe transitions labeled by all the letters 'a', 'b', ... 'z'. You can also label a single transition by multiple symbols: For instance, a transition labeled by `'cat'` stands for three consecutive transitions labeled by `'c'`, `'a'`, and `'t'` respectively.

This is a design problem, and part of your job is to figure out a way to distinguish among correct and incorrect web addresses. 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.