CSC 512 Compiler Construction
Fall 2022 Xu Liu HW 1

1

Describe informally the languages accepted by the following `FAS`:

b.

It only has one end state. It represents any pattern of `1`s or `0`s, but doesn't have a pattern for `10` or `01` after s_1 and s_2 which will be followed by a sequence of `00` and `11` respectively.

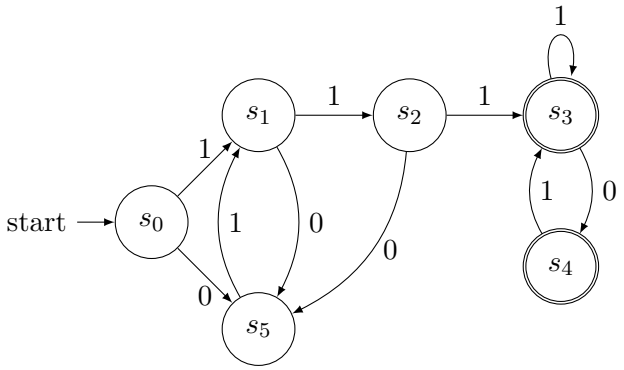
c.

Will contain a lot of `a`'s and `b`'s. Subset of `aab` and `baa` will appear.

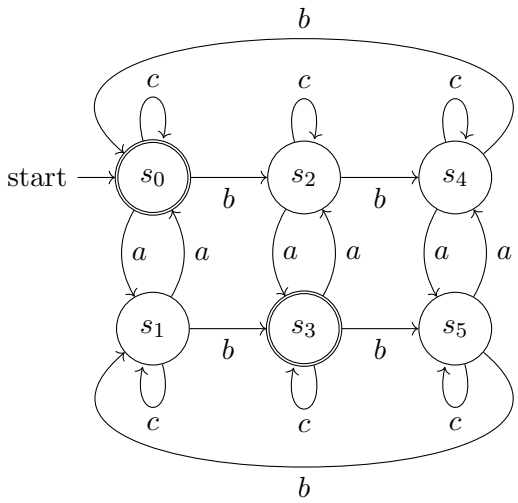
2

Construct an `FA` accepting each of the following languages:

b. $w \in \{0,1\}^* | w$ contains '111' as a substring and does not contain '00' as a substring



c. $w \in \{a,b,c\}^* | \in w$ the number of a's modulo 2 is equal to the number of b's modulo 3



4

Different programming languages use different notations to represent integers. Construct a regular expression for each one of the following:

c. Currency, in dollars, represented as a positive decimal number rounded to the nearest one-hundredth. Such numbers begin with the character $, have commas separating each group of three digits to the left of the decimal point, and end with two digits to the right of the decimal point, for example, $8,937.43 and $7,777,777.77.

$$\$\Big([1\cdots9]([0\cdots9]|\epsilon)([0\cdots9]|\epsilon)(,[0\cdots9][0\cdots9][0\cdots9])^*\Big|0\Big).[0\cdots9][0\cdots9]$$

5

Write a regular expression for each of the following languages:

e. Given an alphabet $\Sigma = \{+, -, \times, \div, (,), \texttt{id}\}$, L is the set of algebraic expressions using addition, subtraction, multiplication, division, and parentheses over `id`s.

$$L = \texttt{id}(+|-|\times|\div)\texttt{id}\Big((+|-|\times|\div)\texttt{id}\Big)^*$$

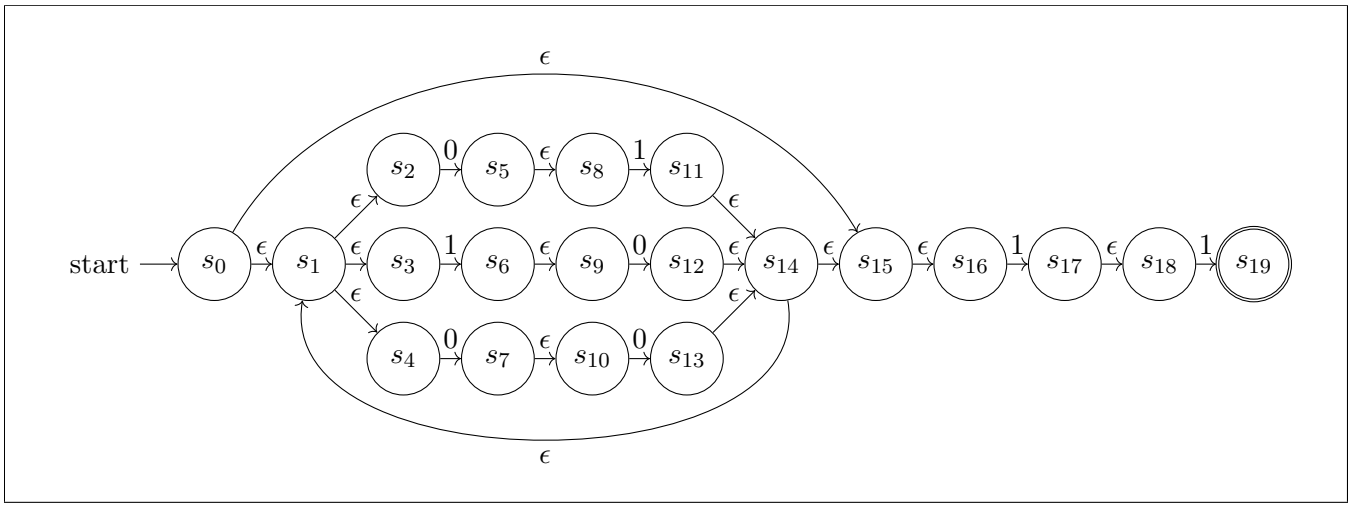
An exception is that parenthesis cannot be matched with regular expression, thus the language cannot be fully represented by regular expressions.

7

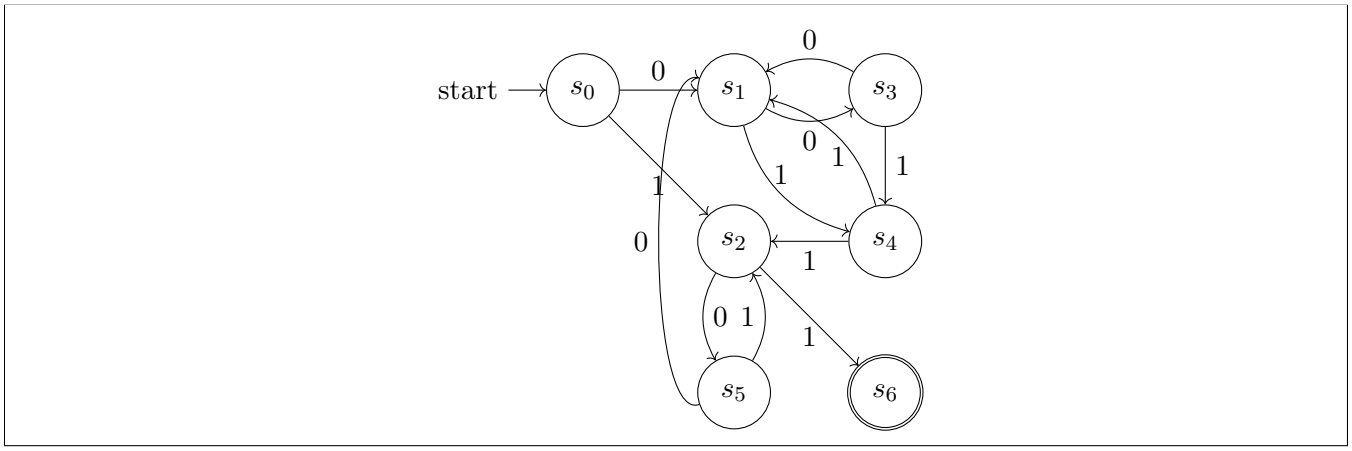
Consider the regular expression:

$$(01|10|00)^*11$$

a. Use Thompson's construction to construct an NFA for RE.



b. Convert the NFAs to DFAs.



c. Minimize the DFAs.