

Exercise 5 – Regular Expressions

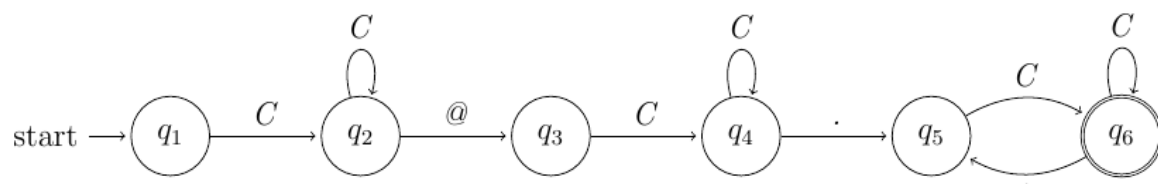
A software team is implementing a new website. On the user registration form, they need to validate several input fields. One of these fields is the date of birth, which should be entered in the format YYYY-MM-DD, with the leading zeroes possibly omitted.

You are asked to provide increasingly more refined versions of a regular expression that can be used to validate the date of birth.

- Provide a regular expression to validate the date format.
- Provide a version that validates limits (i.e., the month must be between 1 and 12 and the day between 1 and 31).
- Provide a version that validates the days according to the month (i.e., take into account months with 30, 31 or 28 days).

Note: You can simplify your expressions by using symbols to represent sets of digits. For instance, if you use D to represent any digit (0-9) then you can write $(0+1+2+3+4+5+6+7+8+9)^*$ simply as D^* .

Another member of the team in charge of validating the new user's e-mail address came up with the following automata (where C stands for any character):



- Convert it into a regular expression using the state elimination method. Show all the intermediate steps.