Natural Language Processing

Sunil

27 January 2017

Problem 1

I use Python notation for regular expression.

a

The set of all lower case alphabetic strings ending in a b.

$$\Sigma = \{a, b, ...z\}$$

$$/[a - z] * ab/$$

b

The set of all strings from the alphabet a, b such that each a is immediately preceded by and immediately followed by a b;

$$\Sigma = \{a, b\}$$

$$/(abaab) * /$$

c

All strings that have both the word grotto and the word raven in them (but not, e.g., words like grottos that merely contain the word grotto). By "word", we mean an alphabetic string separated from other words by whitespace, any relevant punctuation, line breaks, and so forth.

$$\Sigma = \{a, b, ...z, ...\}$$

$$/grotto. * raven/$$

Problem 2

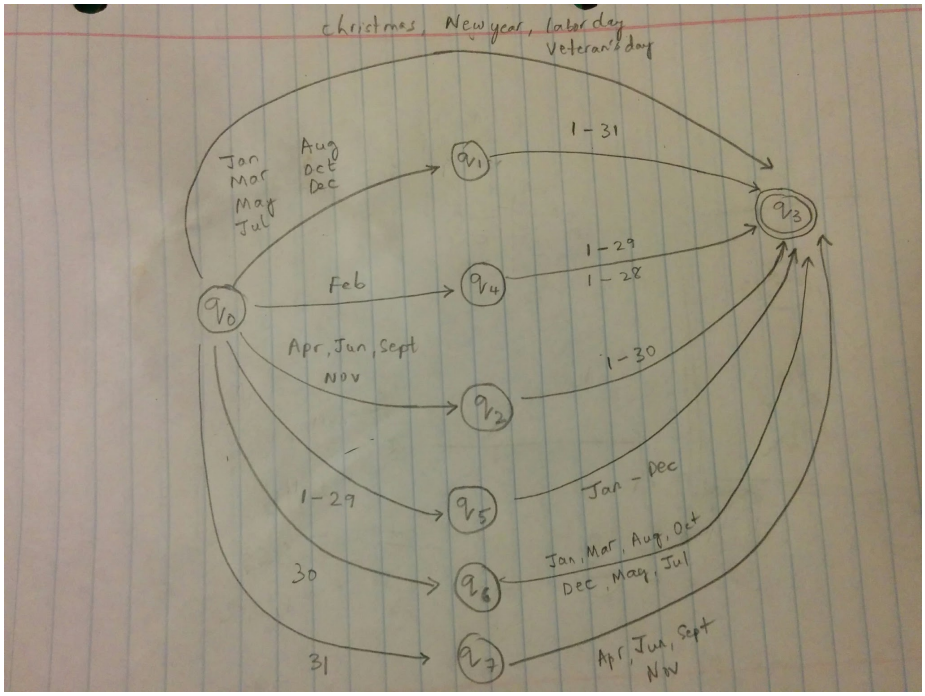

Figure 1: FSA Recognize date