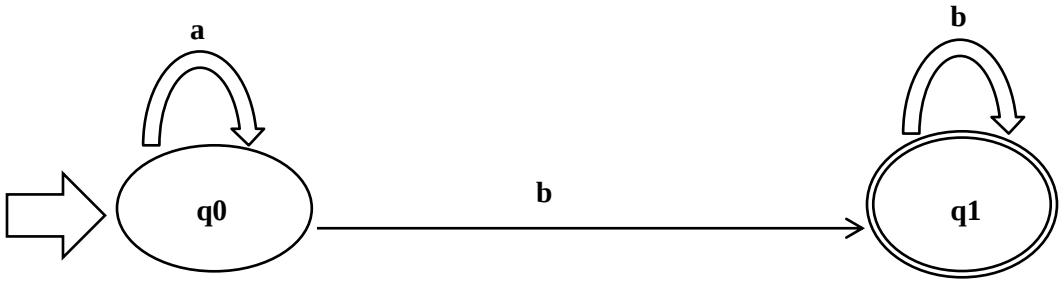
LING 570 – HW2

Q1 - Part (1)

Consider the following DFA of D1 = (∑, Q, q0, F, δ)

∑ = {a, b}
Q = {q0, q1}
F = {q1}

δ = { q0 x a → q0
q0 x b → q1
q1 x b → q1 }



The language accepted by D1 is {ab, aab, abb, abbb, bbb, …} or a*b+

We can create the a regular grammar that generates the same language with the following mappings;

DFA	Regular Grammar
∑	∑
q0	S
δ	P
Q	N

So by substitution into a grammar G = (N, ∑, P, S)

N = {q0, q1}
∑= {a, b}

P = { S → a S
S → b q1
q1 → b q1
q1 → ε }

We can see that the same language L(G) = a*b+ is also accepted by G.

Q1 - Part (2)

Consider the grammar G = ({q1, q2, q3}, {a, b}, P, q1)

P = { S → a S
S → ε S
S → ε q3
S → b q2
q2 → a q2
q2 → a q3
q2 → b q3
q3 → a S }

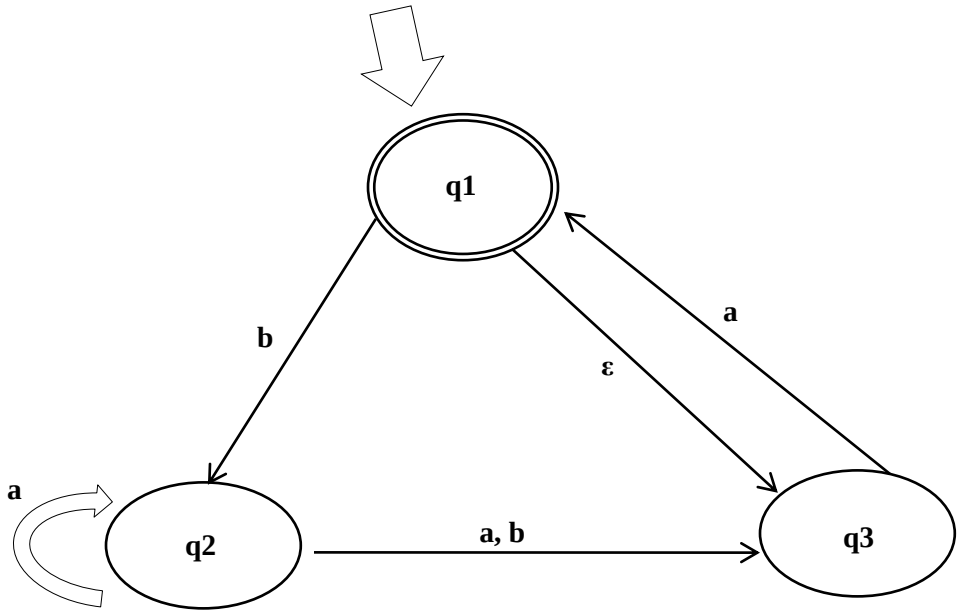
where L(G) = { ε, a, baba, baa, …}.

We can also map the grammar to a NFA of N1 = (Q, ∑, δ, q0, F)

Regular Grammar	NFA
{a, b}	∑
q1	F
P	δ
{q1, q2, q3},	Q

The mapping gives us N1 = ({q1, q2, q3}, {a, b}, δ, q1, q1).

Based on the transition function of P = δ we can construct a NFA for N1 below. In addition, the NFA also takes into account of the grammar showing transition leading to more than one state and transition happening without reading any inputs (ε).



Q2 - Part (2)

Both commands below produced the same outcome of:

(0 -> 0 "they" : "PRO" / 1) (1 -> 1 "can" : "AUX" / 0.99) (2 -> 2 "fish" : "NOUN" / 0.7) 0.693

1. carmel -k 1 fsa7 wfst1
2. cat wfst1_test | carmel -k 1 -sli wfst1

These commands take in a FSA (fsa7) or a word-string sequence (wfst1_test) as input, pass them through a weighted finite state transducer (wfst1) to produce a weighted FSA path that shows the top 1 most likely path consisting of tag sequences based on the input word-string. The ranking of the most likely path is based on the combined conditional probabilistic values for each pair of output tag and input word.

Q2 - Part (3)

The command produced the following output:

(0 -> 0 "they" : "PRO" / 1) (1 -> 1 "can" : "AUX" / 0.99) (2 -> 2 "fish" : "NOUN" / 0.7) 0.693
(0 -> 0 "fish" : "NOUN" / 0.7) (1 -> 1 "they" : "PRO" / 1) (2 -> 2 "can" : "AUX" / 0.99) 0.693

The –b option accepts batch composition of input word-string sequences. The outcome is similar to what were described in Part (2) and that is output the result of the top 1 most likely path for each given input sequence. If there are 3 input sequences of word-string then the output would also consists of 3 lines, with each line output in the same order as the input sequence.

End of HW2 – submitted by Wee Teck Tan

Student ID: 0937003

Course Name: LING 570