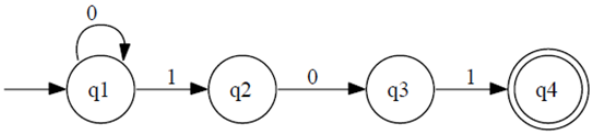
Computational Linguistics 2017-2018 Sheet 1

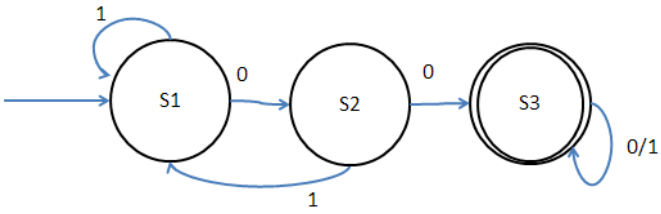
Q1) What is the language recognized by the following FSA



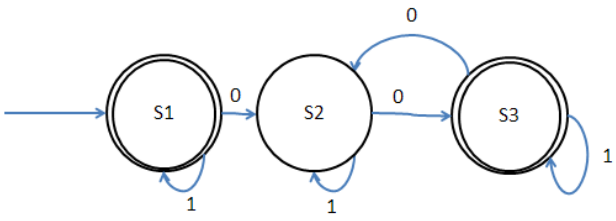
The language accepted by the FSA is the set of strings of 0's and 1's which ends with the sub-string of '101'.

Q2) Draw a transition diagram for a FSA that accepts the language that consists of:

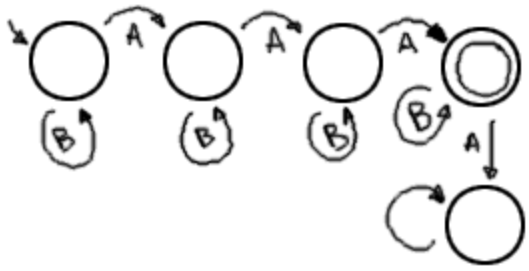
- Strings containing the sequence 00 and Σ = { 0, 1}.



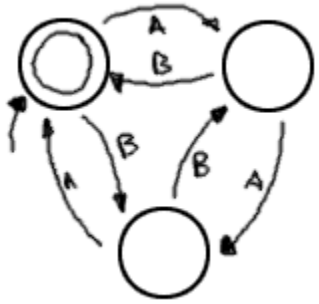
- Strings containing even number of Zero's and Σ = { 0, 1}.



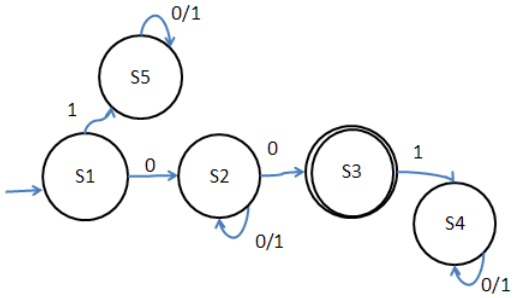
- Strings with exactly 3 a's and Σ = { a, b}.



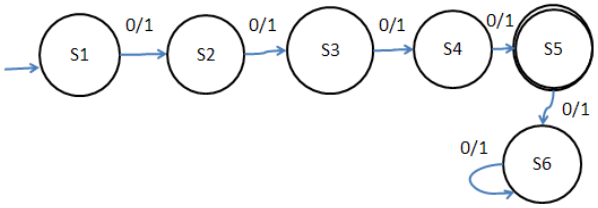
- Strings where (number of a's mod 3) = (number of b's mod 3), and Σ = { a, b}. (e.x. abaabbaaa)



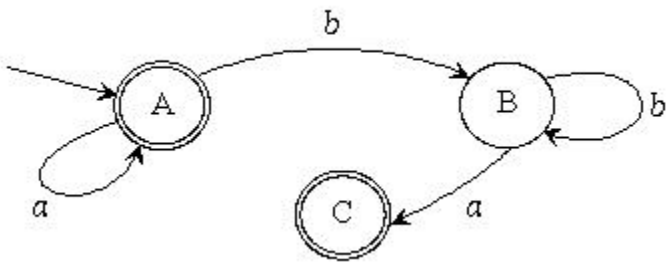
- Strings that start and finish with 0 and Σ = { 0, 1}.



- Binary strings of length 4 and Σ = { 0, 1}. Recognized strings would include 0000, 1111, 0101, but not 01, or 00000.



Q3) Given the following FSA with Σ = { a, b}.



- Write down two words that this machine accepts and two words that it does not accept.

  Accept: aaa, abba

  Not accept: bb, ab

- Describe the language that this machine recognizes.

  The language accepted by the FSA is the set of strings of a's and b's which are empty or must end with a.

- Write down the regular expression representing the language that this machine accepts.

  $a^* \mid a^* b^+ a$

- Give the mathematical model for this machine -i.e. give:
  - Q - the set of states, {A, B, C}
  - I - the start state, {A}
  - F - the set of final states, {C}
  - E - the set of rules that map a state to the next state.

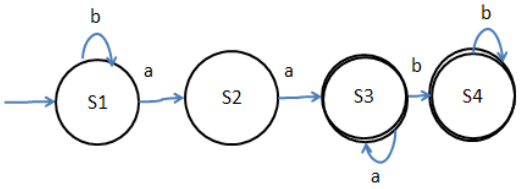
    {A,a,A}, {A,b,B}, {B,a,C}, {B,b,B}

Q4) What strings can be generated from the following regular expressions:

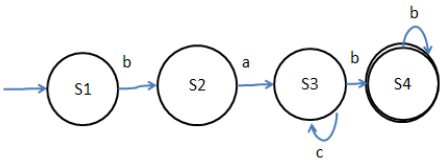
- $(a^*b^*)^*$ → ε, a, b, ab, aab, abb, aa...bb.. , aa…bb…aa…bb…, ….
- $(a^+b)^*$ → ε, ab, aab, aaab, …., abab, abaab, aabaaab,…
- $(abc^*)^+$ → ab, abc, abcc, abccc,…., abab, ababc, ababcc, abcab, abccabccc, ….
- $(a|b|c^*)^+$ → a, b, ε, c, cc, ccc, …, aa, ab, ac, acc, accc, …, ba, bb, bc, bcc, bccc, …, ca, cb, cca, ccb, ccca, cccb, …

Q5) Draw a graph for FSA that represents the following regular expressions:

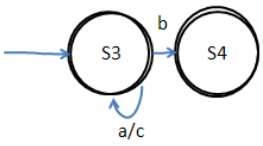
- $b^*a^*(aa)a^*b^*$



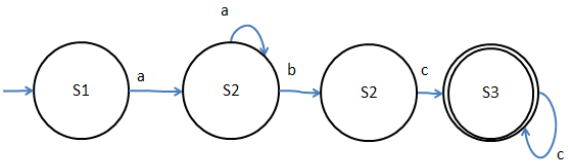
- bac^*b^+



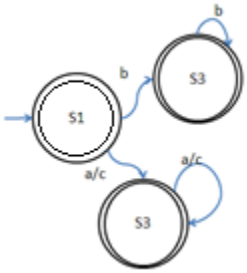
- $(a|c)^*|b$



- a^+bC^+

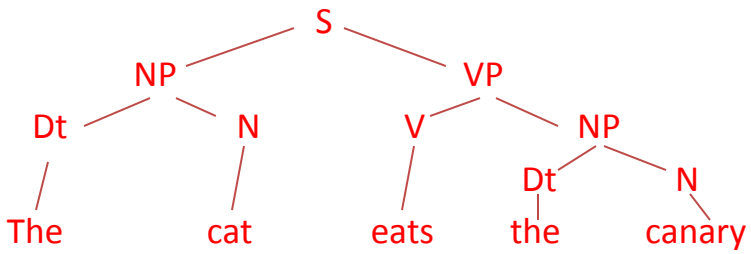


- $(a|c)^{+}|b^{*}$



Q6) Given the following Context-Free Grammar, derive a valid sentence and draw its parse tree:

- N : {S, NP, VP, DT, N, V}
  S : {S}
  Σ: {canary, cat, song, sings, eats, the}
  P:
  S → NP VP
  NP → Dt N
  VP → V NP
  V → sings | eats
  N → cat | song | canary
  DT → the
  The sentence is valid

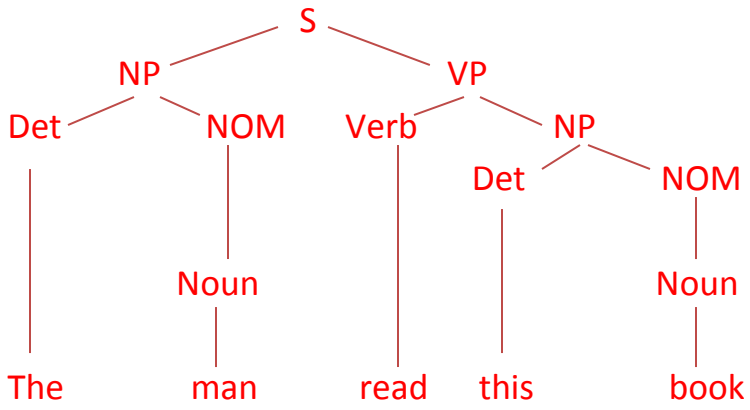


Q7) Given the following CFG:

N {S, NP, NOM, VP, Det, Noun, Verb, Aux}
Σ : {that, this, a, the, man, book, flight, meal, include, read, does}
S : {S}
P:
S → NP VP
S → Aux NP VP
S → VP
NP → Det NOM
NOM → Noun
NOM → Noun NOM
VP → Verb
VP → Verb NP
Det → that | this | a | the
Noun → book | flight | meal | man | fish
Verb → book | include | read
Aux → does

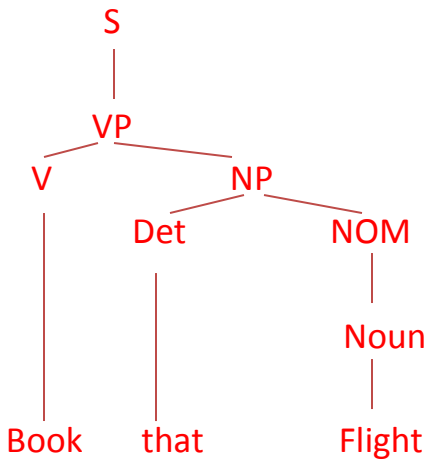
State if the following sentences are valid or not, and draw its parse tree:

- "The man read this book"



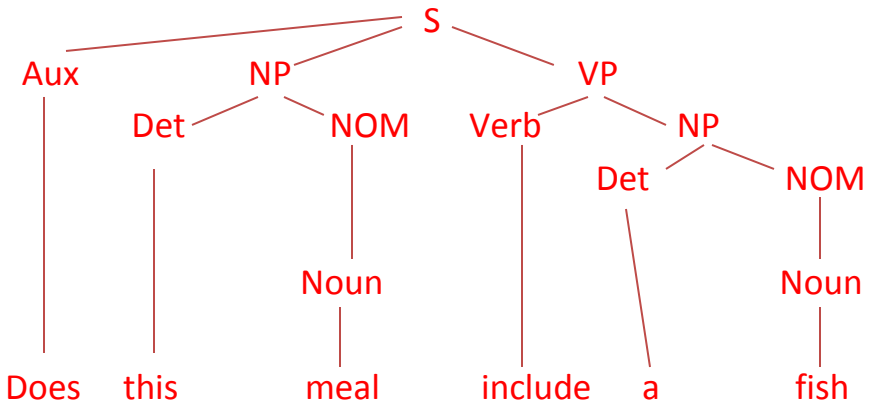
The sentence is valid

- “Book that Flight”



The sentence is valid

- “Does this meal include a fish”



The sentence is valid