Ali Haider
14259
Mid Term Assignment
Course: Natural Language Processing
Date: April 20, 2020

Note: Attempt all Questions.

Question No. 1: (08)

Briefly explain NLP? Write the name of 2 Applications of NLP with example? Write the name of 2 Challenges of NLP with example?

NLP:

Natural Language Processing, usually shortened as NLP, is a branch of artificial intelligence that deals with the interaction between computers and humans using the natural language. The ultimate objective of NLP is to read, decipher, understand, and make sense of the human languages in a manner that is valuable. Most NLP techniques rely on machine learning to derive meaning from human languages.

Application of NLP:

- Machine translation (MT), process of translating one source language or text into another language, is one of the most important applications of NLP.
- Interactive Voice Response (IVR) applications used in call centers to respond to certain users' requests.
- Personal assistant applications such as OK Google, Siri, Cortana, and Alexa.

Challenges:

- Syntax and Ambiguity

  Example: I saw the man with a telescope
  - Who had a telescope.
- Semantics

  Example: The astronomer loves the star.
  - Star in sky
  - Or celebrity (05)

a) Define Phonology and Morphology with the help of example? (03)

Phonology:

Phonology is defined as the study of sound patterns and their meanings, both within and across languages.

Example: An example of phonology is the study of different sounds and the way they come together to form speech and words - such as the comparison of the sounds of the two "p" sounds in "pop-up."

Morphology:

Morphology is the branch of linguistics which seeks to determine the base units of meaning within a given language. A morpheme is the smallest unit of language which is individually meaningful. While sounds may distinguish words through minimal pairs, for example, they are not expressly responsible for the *meaning* of the word. Most correspondences between sounds and meaning are arbitrary.

Example: the word “cat” has just one morpheme but the word “cats” has 2

Question No. 2: (10)

a) What do you mean by regular expressions?

Answer:

A regular expression is a set of characters, or a **pattern**, which is used to find sub strings in a given string. if there’s a pattern in any string, you can easily extract, substitute and do variety of other string manipulation operations using regular expressions. Regular expressions are a language in itself since they have their own compilers and almost all popular programming languages support working with regexes.

b) Specify the text strings using the below regular expressions:

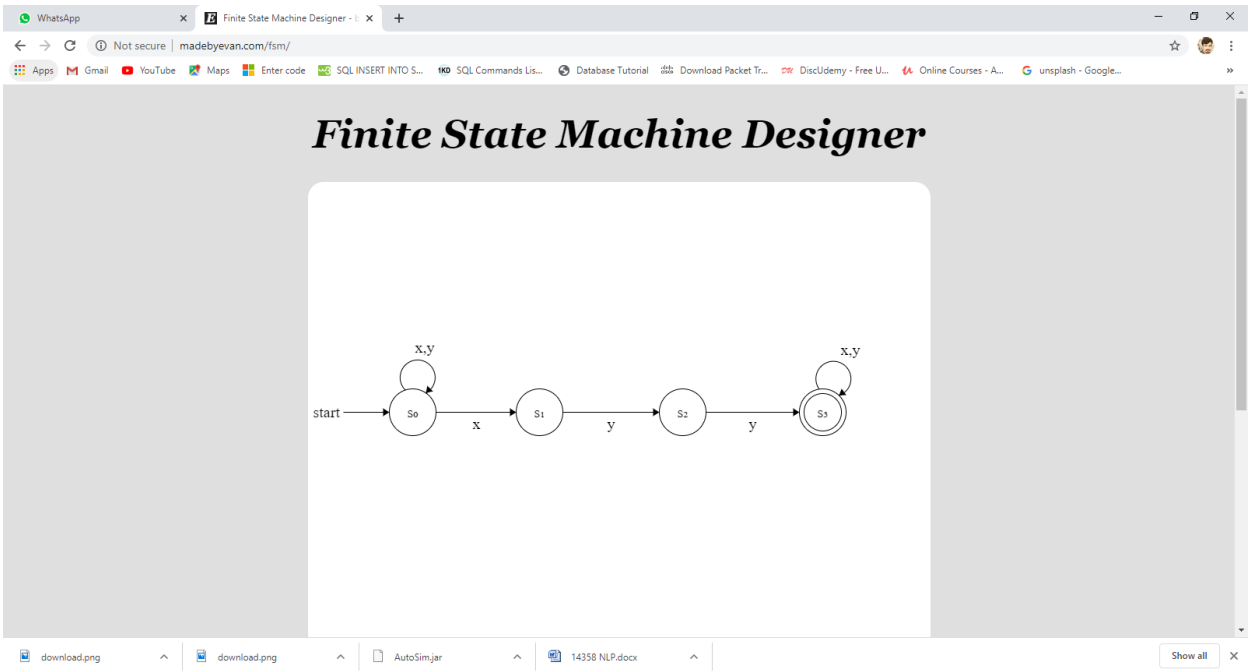
1. /[a-fAF0-9]
   a. Given string: a89 opx cfff
2. /[abc]
   a. Given string abc ac acb a0b a2b a42c A878
3. a(b|c)
   a. Given string abc aa acbaob
4. /abc*
   a. Given string ab abc abcc babc abc abcc babc
5. /abc+
   a. Given string ab abc abcc babc abc abcc babc
6. /[^a-z A-Z]
   a. Given string Price of cat $1
7. /[^a-z A-Z 0-9]
   a. Given string: a89 opx cfff $1!
8. /a(bc)
   a. Given string: ab abc ac acb a0b a2b a42c A87d
9. /a[bc]
   a. Given string abc ac acb a0ba2b
10. a|b|c

a. Given string: ab abc ac acb a0b a2b a42c A87d

Question No. 3: (05)

a) Design an NFA over an alphabet ∑={x,y} such that every string accepted must have a substring --xyy-- ? identify its tuples and also convert it into DFA.

NFA:



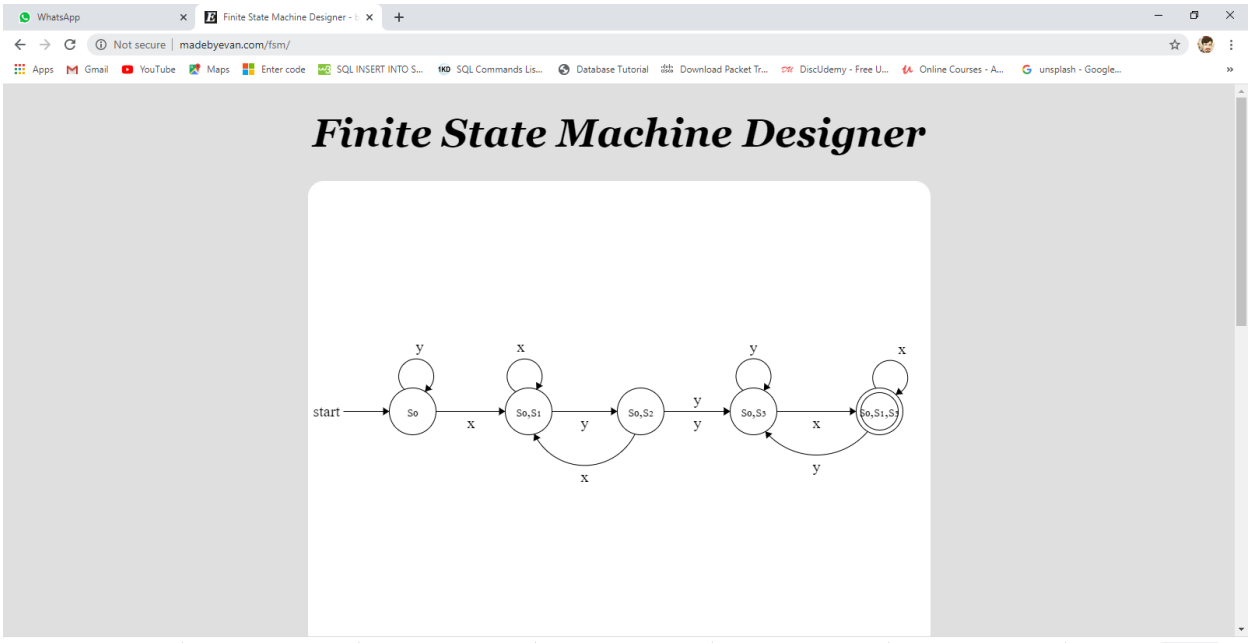
FSA Transition Table:

S	X	Y
s0	s0 , s1	s0
s1	Null	s2
s2	Null	s3
s3	s1	s3

DFSA Transition Table:

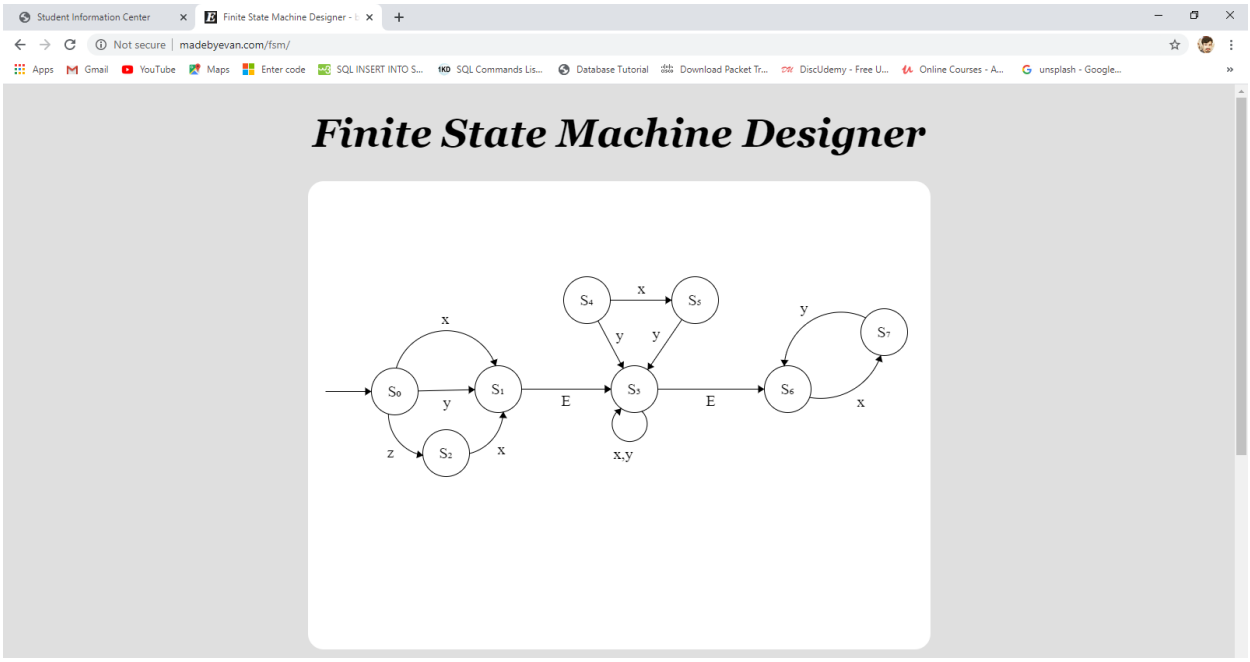
Q	X	Y
→ s0	{s0 , s1}	s0
{s0,s1}	{s0,s1}	{s0,s2}
{s0,s2}	{s0,s1}	{s0,s3}
{s0,s3}	{s0,s1,s3}	{s0,s3}
{s0,s1,s3}	{s0,s1,s3}	{s0,s3}

DFA:



Question No. 4: (05)

a) Design an NFA for the regular expression : $(x+y+zx)\,((yxy)^*+(x+y)^*)^*(xy)^*$



Question No. 5: (02)

Find the Maximum Likelihood Estimation of the below according to the given corpus using conditional probability:

<s> The green eyes </s> <s> The green jungle </s>
<s> The green jungle </s> <s> The green eyes </s>
<s> The green park </s> <s> The green eyes </s>

i. P(jungle|The green)= P(The green jungle) = 2 = 0.333
P(The green) 6

ii. P(eyes|The green) = P(The green eyes) = 3 = 0.5
P(The green) 6

iii. P(park|The green) = P(The green park) = 1 = 0.17
P(The green) 6

iv. P(sea|The green) = P(The green park) = 0 = 0
P(The green) 6

P(S) =2 .3 .1 .0 = 0
6 6 6 6