

**NAME :MUJAHID AYAZ**

**ID: 14304**

**SUB: NLP**

**DEP:BS(SE)5<sup>TH</sup>**

**Question No. 1:**

**(08)**

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

**ANSWER:**

**Natural language processing** involves the reading and **understanding** of spoken or written **language** through the medium of a computer. Through **natural language processing**, computers learn to accurately manage and apply overall linguistic meaning to text excerpts like phrases or sentences.

Applications:

## Summarization

- Generating a short summary from one or more documents, sometimes based on a given query



This is a  sentence summary of <http://hpi.de/en/news/jahrgaenge/2015/des...>

Summary processing at low priority. Upgrade to **BOOST**

### Design Thinking Week: Students Improve the Daily Life Experience for People with Illiteracies

On the occasion of the World Literacy Day on September 8 more than 40 young innovators applied their Design Thinking skills in order to make life easier for these people.

Here, the focus was especially on the possibilities of using digital technologies and computers to better the daily obstacles in life of the people concerned.

Under the guidance of the D-School's coaches the teams researched, developed and prototyped - and could present many versatile solutions in the end: e.g. one of the groups came up with an idea for a software program that lets internet browsers read texts, functions and links out loud so that people with reading problems can still use news sites or social networks like Facebook.

<http://smmry.com/>

13

## Text Categorization

- Assigning one (or more) pre-defined category to a text



The screenshot shows a PubMed article page. The article title is "Coupling of angiogenesis and osteogenesis by a specific vessel subtype in bone." The abstract discusses the role of mesenchymal stem cells and osteoblasts in bone formation and regeneration. On the right side, a list of MeSH terms is displayed, including "Aging/metabolism", "Aging/pathology", "Animals", "Blood Vessels/anatomy & histology", "Blood Vessels/cytology", "Blood Vessels/growth & development", "Blood Vessels/physiology\*", "Bone and Bones/blood supply\*", "Bone and Bones/cytology", "Endothelial Cells/metabolism", "Hypoxia-inducible Factor-1, alpha Subunit/metabolism", "Male", "Mice", "Mice, Inbred C57BL", "Neovascularization, Physiologic/physiology\*", "Osteoblasts/cytology", "Osteoblasts/metabolism", "Osteogenesis/physiology\*", "Oxygen/metabolism", "Stem Cells/cytology", and "Stem Cells/metabolism".

11

CHALLENGES:

## Syntax and ambiguity

- I saw the man with a telescope.
  - Who had the telescope?



(<http://www.realtytrac.com/landing/2009-year-end-foreclosure-report.html>)

## Semantics

- The astronomer loves the **star**.
  - Star in the sky
  - Celebrity



(<http://en.wikipedia.org/wiki/Star#/media/File:Starsinthesky.jpg>)



(<http://www.businessnewsdaily.com/2023-celebrity-hiring.html>)

54

b) Define Phonology and Morphology with the help of example?

## Phonetics and phonology

- The study of linguistic sounds and their relations to words

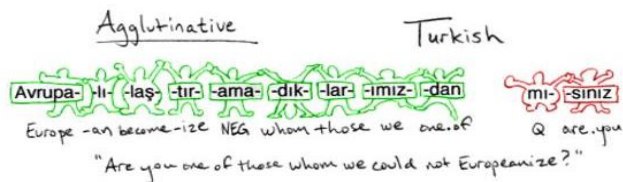
<http://german.about.com/library/blfunkabc.htm>

43

Das Funkalphabet - German Phonetic Spelling Code compared to the international ICAO/NATO code Listen to AUDIO for this chart (below)		
Germany*	Phonetic Guide	ICAO/NATO**
A wie Anton	AM-TO-ne	Alfa/Alpha
Ä wie Ärger	ÄHR-geh-er	(1)
B wie Berta	BARF-ber-tuh	Bravo
C wie Cäsar	SÄH-zar	Charlie
Ch wie Charlotte	shar-LOT-tuh	(1)
D wie Dera	DODE-uh	Delta
E wie Emil	ay-MFAM	Echo
F wie Friedrich	FREED-rih-ohd	Foxtrot
G wie Gustav	GOOS-tah-t	Golf
H wie Heinrich	HINE-reach	Hotel
I wie Ida	IFID-uh	India/Indigo
J wie Julius	YUL-see-ooz	Juliett
K wie Kaufmann	KOOF-mann	Kilo
L wie Ludwig	LOOD-veg	Lima
AUDIO 1 > <a href="#">Listen to mp3 for A-I</a>		
M wie Martha	MAR-tuh	Mike
N wie Nordpol	NORT-pole	November
O wie Otto	AHT-oh	Oscar
Ö wie Ökonom (2)	UEH-ko-nomic	(1)
P wie Paula	POW-huh	Papa
Q wie Quelle	KVEL-uh	Quebec
R wie Richard	RFF-eh-ert	Romeo
S wie Siegfried (3)	SEEG-freed	Sierra
Sch wie Schule	SHOO-uh	(1)
ß (Ezsett)	ES-TSET	(1)
T wie Theodor	TAYN-thee	Tango
U wie Ulrich	OOL-reach	Uniform
Ü wie Übermut	ULI-ber-moot	(1)
V wie Viktor	VICK-tor	Victor
W wie Wilhelm	VIL-helm	Whiskey
X wie Xanthippa	KSAN-tipp-uh	X-Ray
Y wie Ypsilon	IPP-see-lohn	Yankee
Z wie Zeppelin	TSEF-puh-lemn	Zulu

## Morphology

- The study of internal structures of words and how they can be modified
- Parsing complex words into their components



44

(<http://allthingslinguistic.com/post/50939757945/morphological-typology-illustrations-from>)

## Question No. 2:

(10)

a) What do you mean by regular expressions?

ANSW: A formal language for specifying or searching text strings .

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

1. `/[a-zA-F0-9]`

a. Given string: a89opxcfff

Ans: a89 cfff

2. `/[abc]`

a. Given string abc ac acb a0b a2b a42c A878

Ans: abc ac acb ab ab ac

3. `a(b|c)`

a. Given string abc aa acbaob

Ans: ab ac

4. `/abc*`

a. Given string ab abc abcc babc abc abcc babc

Ans: abc abc abcc abc abc abcc abc.

5. `/abc+`

a. Given string ab abc abcc babc abc abcc babc

Ans: abc abcc abc abc abcc abc

6. `/[^a-zA-Z]`

a. Given string Price of cat \$1

Ans: \$1

7. `/[^a-zA-Z 0-9]`

a. Given string: a89 opx cfff \$1!

Ans: \$!

8. `/a(bc)`

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

Ans: abc

9. `/a[bc]`

a. Given string abc ac acb a0ba2b

Ans: ab ac ac

10. `a|b|c`

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

Ans: ab abc ac acb ab ab ac.

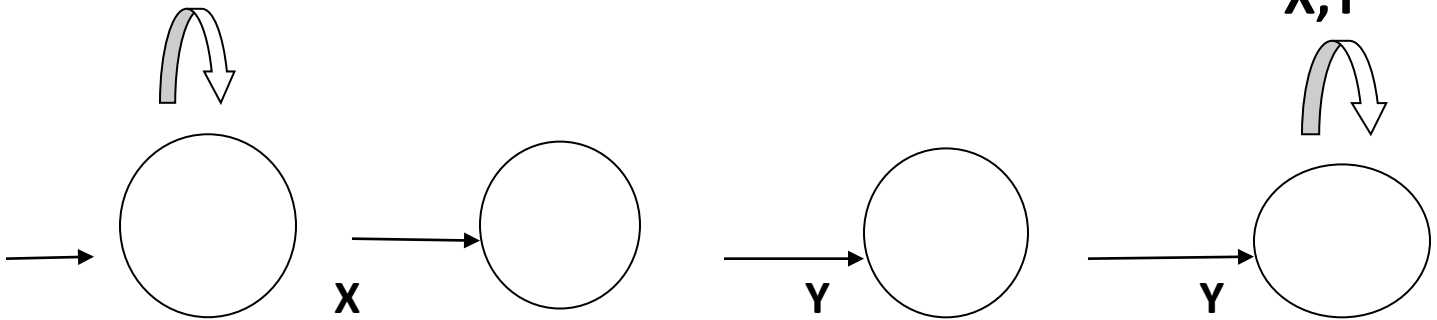
**Question No. 3:**

(05)

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

**Ans: NFA:**

**X,Y**

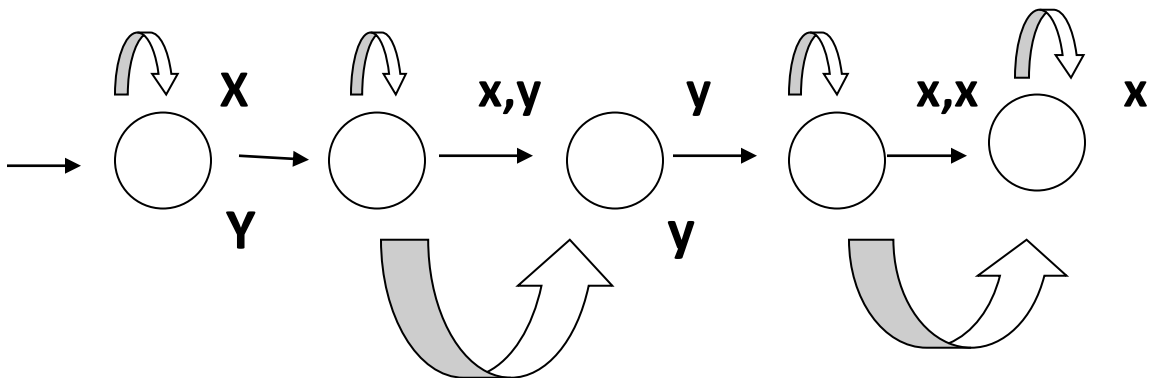


Q	x	y
Q0	Q0,q1	Q0
Q1	NULL	Q2

Q2	NULL	Q3
Q3	Q1	Q4

Q	X	Y
Q0	{Q0,Q1}	Q0
{Q0,Q1}	{Q0,Q1}	{ Q0,Q2 }
{Q0,Q2}	{Q0,Q1}	{ Q0,Q3 }
{Q0,Q3}	{Q0,Q1,Q3}	{ Q0, Q3 }
{Q0,Q1,Q3}	{Q0,Q1,Q3}	{ Q0, Q3 }

# DFA:





**x**

**y**

**Question**

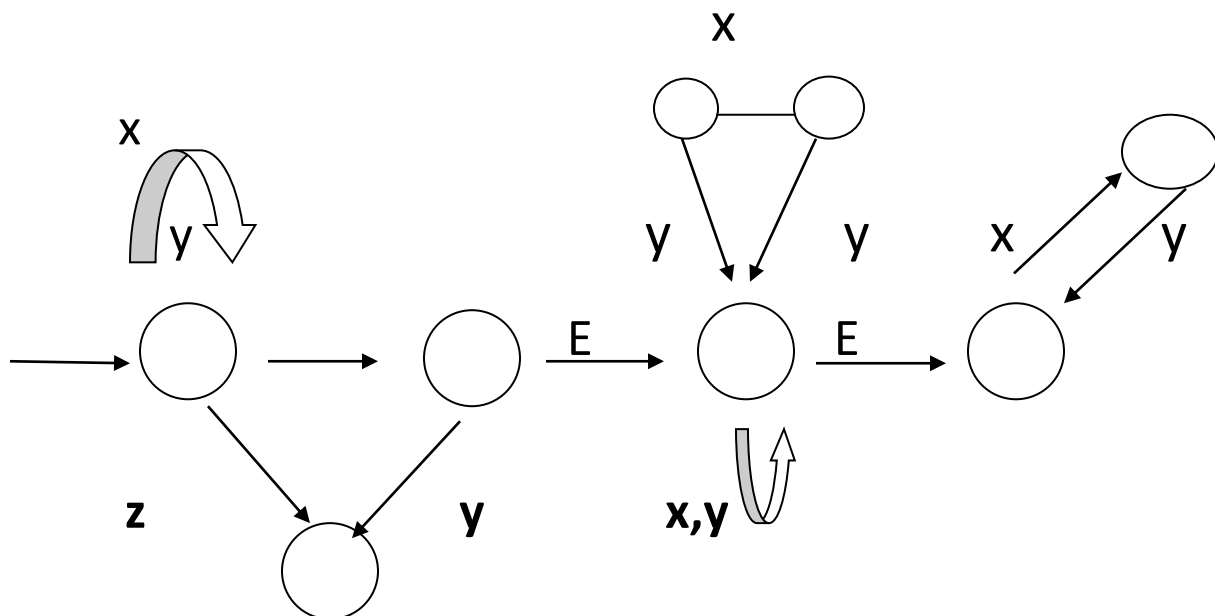
**No.**

**4:**

(05)

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

ANS: Design NFA for the regular expression.



**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.  $P(\text{jungle} | \text{The green})$

$$\text{Ans: } \frac{P(\text{The green jungle})}{P(\text{The green})} = \frac{2}{6} = 0.333$$

ii.  $P(\text{eyes} | \text{The green})$

$$\text{Ans: } \frac{P(\text{The green eyes})}{P(\text{The green})} = \frac{3}{6} = 0.5$$

iii.  $P(\text{park} | \text{The green})$

$$\frac{P(\text{The green park})}{P(\text{The green})} = \frac{1}{6} = 0.17$$

iv.  $P(\text{sea} | \text{The green})$

$$\frac{P(\text{The green park}) \cdot 0}{P(\text{The green})} = 0$$

$$P(\text{The green}) = \frac{1}{6}$$

$$\frac{P(S) = 2 \cdot \frac{1}{6} \cdot 3 \cdot \frac{1}{6} \cdot 1 \cdot \frac{1}{6} \cdot 0 \cdot \frac{1}{6}}{P(\text{The green})} = 0$$