

# Model Answer of Mid term Exam

## 2015/2016

### Question 1

**B** is a corpus which only contains one single bitstring:

1 1 0 1 1 1 0 0 1 0 1 1 1 0 1 1 1 1 0 0 0

- 1.1) Calculate the following bigram probabilities from the corpus **B** using MLE (Maximum Likelihood Estimation). Answer with a ratio  $p/q$ , not a floating point number.

(a)  $P(0 | 1)$  ( 1/2 Mark)

$$C(10) / C(1) = 5/13$$

(b)  $P(0 | 0)$  ( 1/2 Mark)

$$C(00) / C(0) = 3/8$$

- 1.2) Assume a bigram language model created from corpus **B**. For each of the following bit strings, decide if it is more probable that  $x_1$  resp  $x_2$  is 0 or 1.

(c) 1 0 1 0 1 0 1  $x_1$  ( 1 Mark)

$x_1 = 1$  is more probable  
since  $P(0 | 1) = 5/13 < P(1 | 1) = 8/13$

(d) 0 1 0 1 0 1 0  $x_2$  ( 1 Mark)

$x_2 = 1$  is more probable  
since  $P(0 | 0) = 3/8 < P(1 | 0) = 4/8$

## Question 2

Write regular expressions that recognize the following languages.

- (a) any string that contains at least three digits ( 1/2 Mark)

`.*\d.*\d.*\d.*`

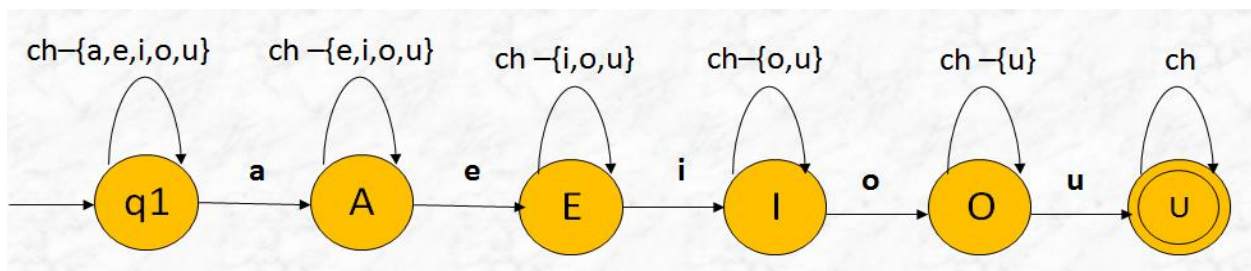
- (b) find a word ending in *ility* , example *accessibility* ( 1/2 Mark)

`(\w*)ility` or `[0-9 A-Z a-z]*ility`

- (c) find a word **w** that can find five vowels, a, e, i, o, and u, that occur in **w** in alphabetical order}. *abstemious*, *facetious*, *sacrilegious* (**Draw the FSA that accept this language**)

`[^aeiou]* a [^eiou]* e [^iou]* i [^ou]* o .*u` ( 1 Mark)

( 1 Mark)



## Question 3

- 1- Define derivation and inflection (with examples) and show the main criteria that differentiate them?

( 1 Mark)

**Inflection** is the form variation of a word under certain grammatical conditions, these conditions consist notably of the number, gender, conjugation, or tense

Ex: help + ed → helped

**Derivation** combines affixes to an existing root or stem to form a new word. Derivation is more irregular and complex than inflection. It often results in a change in the part of speech for the derived word.

EX: acceptable

(1 Mark)

There are three main criteria:

- 1- **Category change:** Inflection does not change grammatical category; derivation sometimes does (thereby creating new words).
- 2- **Order:** Derivational affixes must combine with the base before an inflectional affix does (root - affder- affinf " teachroot-erder-sinf).
- 3- **Productivity:** Inflectional affixes tend to be highly productive (i.e., easily applied to new appropriate stems); derivational affixes apply to restrictive classes of bases.

- 2- Define the Natural Language Processing? What are stages of language processing?

( 1 Mark)

**NLP is a field of computer science, artificial intelligence, and computational linguistics . Concerned with the interactions between computers and human (natural) languages.**

( 1 Mark)

**Language Processing stages are**

- 1- **Phonetics and phonology**
- 2- **Lexical Analysis**
- 3- **Morphology & Lexicon**
- 4- **Syntactic Analysis**
- 5- **Semantic Analysis**
- 6- **Pragmatics**
- 7- **Discourse**