Cairo University
Faculty of Computers and Information

Final Exam

Department: Computer Science
Course Name: Natural Language Processing   **Date: 5/1/2016**
Course Code: CS462   **Duration: 2 hours**
Instructor(s): Dr. Hanaa Bayomi Aly   **Total Marks: 60**

Question 1 [20 marks]

1. Which of the following strings cannot be derived from the symbol S using the rules
   S → SS | aaa | aaaaa ?

   a. aaaaa
   b. aaaaaa
   c. aaaaaaa
   d. aaaaaaaa
   e. aaaaaaaaa

2. Which of the following pairs of rules involves indirect recursion?

   a. NP → Det Adj N ,   NP   → Pron
   b. VP → V NP   ,   NP   →Det N that VP
   c. N' → N N'   ,   NP   → Det N'
   d. S   → AdvP S   ,   AdvP → Adj Adv
   e. VP → V NP   ,   NP   → Det N PP

3. Which of the following strings is a member of the language over {a, b} defined by the
   regular expression (aa + ba)* (bb)*?

   a. aabbba
   b. aaaabb
   c. babbaa
   d. bbaa
   e. None of the above.

4. Consider the following context-free grammar, with start symbol S and terminals **a, ; , <, >**.

   S → <L | a   L →aR | <LR   R → > | ;L

   How many different parse trees are there for the string << a >; a > ?

   a. 0
   b. 1
   c. 2
   d. 3
   e. More than three.

5. Which of the following English words is not an open-class word?
    a. drink
    b. time
    c. exam
    d. therefore
    e. can

6. Which of the following strings does not match the regular expression written in machine syntax as (0*1?2+)*

    a. ε
    b. 02
    c. 12
    d. 01201
    e. 12222

7. Consider the following finite state transducer with input alphabet {a, b, c} and output alphabet {0, 1, 2}:

    ε:0   a:1   ε:0
    a:0   a:0
    b:2   b:ε   b:ε
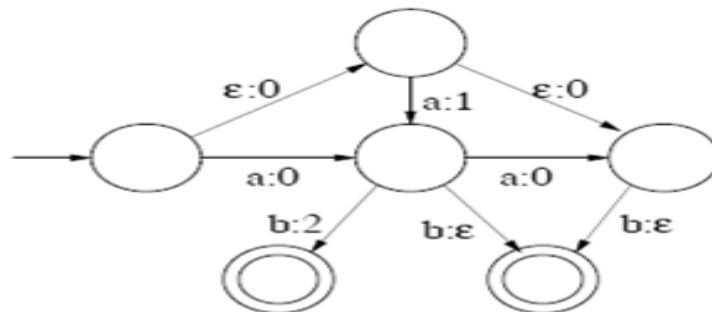
    Which of the following is not a possible output string that could arise from processing the input string "ab"
    a. 0
    b. 00
    c. 01
    d. 02
    e. 012

8. Consider the following probabilistic context-free grammar:

    Which of the following sentences is assigned the highest Probability by this grammar?

    a. clocks tick
    b. tables fly
    c. clocks fly
    d. tables chase clocks
    e. tables hate clocks

$$
\begin{aligned}
S &\rightarrow N\ VP & (1.0) \\
VP &\rightarrow IV & (0.8) \\
VP &\rightarrow TV\ N & (0.2) \\
N &\rightarrow \text{clocks} & (0.7) \\
N &\rightarrow \text{tables} & (0.3) \\
IV &\rightarrow \text{tick} & (0.7) \\
IV &\rightarrow \text{fly} & (0.3) \\
TV &\rightarrow \text{hate} & (0.7) \\
TV &\rightarrow \text{chase} & (0.3)
\end{aligned}
$$

9. Which of the following sets of productions is not in Chomsky normal form?

$G1:$ $S \rightarrow AB$ $G2:$ $S \rightarrow AB$ $G3:$ $S \rightarrow A|B$
       $A \rightarrow AB|a$ $A \rightarrow AB|a$ $A \rightarrow AB|a$
       $B \rightarrow Ba|b$ $B \rightarrow BA|b$ $B \rightarrow BA|b$

   a. G1
   b. G2
   c. G3
   d. G1 and G3
   e. G2 and G3

10. Consider the following grammar: Which of the following part-of-speech tag sequence can be produced by grammar?

$NP \rightarrow art\ NP1$
$NP \rightarrow ppro\ NP1$
$NP1 \rightarrow num\ NP1$
$NP1 \rightarrow NP2$
$NP2 \rightarrow adj\ NP2$
$NP2 \rightarrow adj\ NP3$
$NP3 \rightarrow noun\ NP3$
$NP3 \rightarrow noun$

   a.   art num noun
   b.   art num num noun
   c.   art adj noun
   d.   art noun
   e.   art adj num noun

Question 2 [7 marks]

Suppose that we have a simple grammar G , with a set of non-terminal symbols N={S,NP,VP,DET,N,V},**S** is the start symbol, with a set of terminal symbols Σ={a,the,boy,man,saw}, and with a set of the following productions:

    1. S   → NP VP
    2. NP   → Det N
    3. VP   → V NP
    4. Det   → the | a
    5. N   → man | boy
    6. V   → saw

Suppose that we use the Top down parser. Draw a trace table to explain how the algorithm will parse the sentence "the man saw a boy"

Question 3 [13 marks]

a) Give a brief description about your presentation topic   *[5 mark]*

b) Define NER , and the problems of NER in Arabic language with example *[8 mark]*

Question 4 [14 marks]

Consider the following simple Chomsky Normal Form (CNF) grammar for commands in English.

    S   → G NNP
    NNP → NP NP
    NP   → DT Nom
    Nom → dog | bone | A N
    G   → give
    DT   → the | a
    N   → dog | bone
    A   →large | small

a) Draw a CYK parse chart for the following sentence: *[6 mark]*
        "give the dog a bone"

b) What is the parser type of CKY, define the parser type and its properties?   *[8 mark]*

Question 5 [6 marks]

Assume the following *WordNet* senses with their definitions:

cat^1 : any of several large cats typically able to roar and living in the wild

cat^2 : feline mammal usually having thick soft fur and being unable to roar

cat^3 :   an informal term for a youth or man

mammal : any warm-blooded vertebrate having the skin more or less covered with hair; young are born alive and nourished with milk

Tiger : large feline of forests in most of Asia having a tawny coat with black stripes

Man : an adult male person (as opposed to a woman)

carnivore: terrestrial or aquatic flesh-eating mammal

How is cat^1 related to each of the other senses – is it a **homonym, a synonym, an antonym, a hyponym, a hypernym, a meronym,** or **none of them**? Note that there can be more than one relation that matches. In your answer sheet, Copy the following table and Fill Semantic relations column with *one of previous relation* **OR** *more than one relation name* **OR** *no relation*

Sense 1	Sense 2	Semantic Relations
cat^1	cat^2	
cat^1	cat^3	
cat^1	mammal	
cat^1	Tiger	
cat^1	Man	
cat^1	carnivore	