Reg. No.										

MANIPAL INSTITUTE OF TECHNOLOGY

MANIPAL

A Constituent Institution of Manipal University

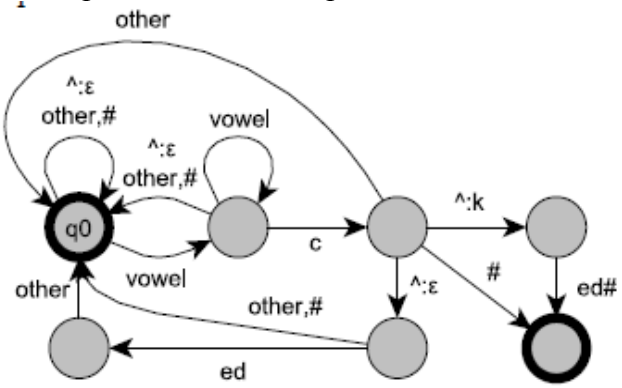
VII SEMESTER B.TECH. (COMPUTER SCIENCE AND ENGINEERING)
MAKE-UP EXAMINATIONS, NOV/DEC 2017
SUBJECT: **PE-V NATURAL LANGUAGE PROCESSING (CSE 4011)**
REVISED CREDIT SYSTEM
(30/12/2017)

Time: 3 Hours MAX. MARKS: 50

Instructions to Candidates:

- Answer **ALL** the questions.
- Missing data may be suitably assumed.

1A. Given the following transducer for the orthographic rule find the specific insertion rule applied by the transducer. Explain with an example.



4

1B. Give one example each for an English sentence with syntactic ambiguity, semantic ambiguity. In each case specify the ambiguity clearly 3

1C. Explain the difference between inflectional and derivational morphology. 3

2A. Consider the following Training data:
<s> I am Sam </s> <s> Sam I am </s> <s> Sam I like </s> <s> Sam I do like </s> <s> do I like Sam </s>
Assume that we use a bigram language model based on the above training data.
i) What is the most probable next word predicted by the model for the following word sequences
(a) <s> Sam . . . (b) <s> Sam I do . . . (c) <s> Sam I am Sam . . . (d) <s> do I like . . .
ii) Which of the following sentences gets a higher probability with this model?
(a) <s> Sam I do I like </s> (b) <s> Sam I am </s> (c) <s> I do like Sam I am </s> 4

2B. Find and tabulate the bi-gram count for the training data given in Q 2A. Apply add-one smoothing to get add one-smoothed probabilities assuming vocabulary size V=50. Compute the probability of the sentence given below using smoothed probabilities. Mention the main disadvantage of this technique.
<s> Sam I do I like </s> 4

2C. Differentiate between Interpolation and back-off in N-gram model. 2

3A. Explain HMM model with all necessary formal definition. Write the Viterbi algorithm for assigning part of speech to a given sentence. 4

3B. What are the three main methods used for part-of-speech (POS) tagging? What are their main differences? 3

3C. Find one tagging error in each of the following sentences that are tagged with the Penn Treebank tag-set given in Table Q3C:
i) I/PRP need/VBP a/DT flight/NN from/IN Atlanta/NN
ii) Does/VBZ this/DT flight/NN serve/VB dinner/NNS
iii) I/PRP have/VB a/DT friend/NN living/VBG in/IN Denver/NNP

Table Q3C

Tag	Description	Tag	Description
CC	Coordinating conjunction	PRP$	Possessive pronoun
CD	Cardinal number	RB	Adverb
DT	Determiner	RBR	Adverb, comparative
EX	Existential there	RBS	Adverb, superlative
FW	Foreign word	RP	Particle
IN	Preposition or subordinating conjunction	SYM	Symbol
JJ	Adjective	TO	to
JJR	Adjective, comparative	UH	Interjection
JJS	Adjective, superlative	VB	Verb, base form
LS	List item marker	VBD	Verb, past tense
MD	Modal	VBG	Verb, gerund or present participle
NN	Noun, singular or mass	VBN	Verb, past participle
NNS	Noun, plural	VBP	Verb, non-3rd person singular present
NNP	Proper noun, singular	VBZ	Verb, 3rd person singular present
NNPS	Proper noun, plural	WDT	Wh-determiner
PDT	Predeterminer	WP	Wh-pronoun
POS	Possessive ending	WP$	Possessive wh-pronoun
PRP	Personal pronoun	WRB	Wh-adverb

3

4A. i) Complete the context-free grammar below so that it generates (at least) the sentences listed.

Table Q4A

S -> NP VP	N -> mouse
NP -> Det N	N -> cat
NP -> PN	PN -> Tom
VP -> TV NP	PN -> Jerry
VP -> DV NP PP	N -> cheese
Det -> the | The	IV -> sleeps
Det -> a	TV -> eats
	DV -> steals

Sentences:

The cat sleeps
The mouse eats the cheese
Tom steals the cheese from the mouse

ii) Convert the grammar into Chomsky Normal Form (CNF).
iii) Complete the CKY parse table for the sentence ”Tom steals the cheese from Jerry”

5

4B. Draw the parse trees to all three sentences given Q4A, using the modified grammar (with added rules) in Table Q4A. Apply both top-down and bottom-up approach. 3

4C. Assume we have the following context-free grammar G in Chomsky normal form: 2

S -> NP VP	NP -> workers | sacks | garbage | junk
VP -> V NP	N -> bin | sack
NP -> D N	V -> dumped
PP -> P NP	P -> of | into
VP -> VP PP	D -> a | the
NP -> NP PP	C -> and
NP -> NP CNP	

Reg. No.										

CNP -> C NP

Construct the parse tree for the following sentence:

workers dumped sacks of garbage and junk into a bin

5A. Explain the four common and important Sentence-Level Constructions of English along with their CFG rules. 5

5B. Given the Probabilistic Context Free Grammar (PCFG) Table, construct the possible parse tree for the sentence *people fish tanks with rods.* Also find the most likely parse tree with the given PCFG. 3

Table Q5B.

S → NP VP	1.0	N → *people*	0.5
VP → V NP	0.6	N → *fish*	0.2
VP → V NP PP	0.4	N → *tanks*	0.2
NP → NP NP	0.1	N → *rods*	0.1
NP → NP PP	0.2	V → *people*	0.1
NP → N	0.7	V → *fish*	0.6
PP → P NP	1.0	V → *tanks*	0.3
		P → *with*	1.0

5C. With an example for each, list any four prenominal-modifiers of head noun in a Noun Phrase. Write the production rule to handle these pre-modifiers. 2