MANIPAL INSTITUTE OF TECHNOLOGY

MANIPAL (A constituent unit of MAHE, Manipal)

VII SEMESTER B.TECH

MAKEUP EXAMINATION, NOV/DEC 2023

SUBJECT: PE-III NATURAL LANGUAGE PROCESSING [CSE 4061]

(-/-/2023)

Time: 3 Hours

MAX. MARKS: 50

Instructions to Candidates:

✤ Missing data may be suitably assumed.

Q No.	Questions		CO	AHEP	Blooms
1 Δ	Illustrate the various types of linguistic knowledge essential for	4M	CO1	125	level 2
17.	analyzing a written text for English language.	-1141		1,2,3	2
1B.	Design a Finite State Automaton (FSA) for English derivational morphology that transforms verbs/adjectives into nouns. Outline the FSA's operation using examples with single/multiple suffixes, and the corresponding nominals for each derivation in a tabular format.	4M	CO1	1,2,5	5
1C.	Compare Finite State Automata (FSA) with Finite State Transducers (FST).	2M	CO1	1,2,5	4
2A.	Illustrate the three main methods used for Part-of-Speech (PoS) tagging?	4M	CO2	1, 2, 3, 4,5	2
2B.	If the bi-gram count of the word sequence <i>time flies</i> in a training corpus is 1020, the unigram count of the word <i>time</i> is 728 and the unigram count of the word <i>flies</i> is 1233, then find the add-one smoothed probability $P_{Lapalace}(flies/time)$ assuming vocabulary size $V = 2433$. For the given bigram <i>time flies</i> compute the relative discount. Present the formulae for all the calculations.	3M	CO5	1, 2, 3, 4	3
2C.	Formulate an equation to calculate the likelihood of a full word sequence employing a bi-gram approximation.	3M	CO5	1, 2, 3, 4	5
3A.	Given the transition probabilities and the observation likelihood, find the highest probable part of-speech (POS) tagging for the sentence <i>Janet will back the bill</i> using Viterbi algorithm. Clearly show the Viterbi values in the form of a matrix and mark the back- trace for up to first three words. Use the row header in the same order as given in the table. Transition Probabilities	5M	CO2	1, 2, 3, 4, 5	3

r			MD	X7D	**	A TA T	DP	D/D				
		NNP	MD	VB	JJ	NN	RB	DT				
	<s></s>	0.2767	0.0006	0.0031	0.0453	0.0449	0.0510	0.2026				
	NNP	0.3777	0.0110	0.0009	0.0084	0.0584	0.0090	0.0025				
	MD	0.0008	0.0002	0.7968	0.0005	0.0008	0.1698	0.0041				
	VB	0.0322	0.0005	0.0050	0.0837	0.0015	0.0514	0.2231				
	JJ NN	0.0300	0.0004	0.0001	0.0733	0.4509	0.0036	0.0056				
	PR	0.0090	0.0170	0.0014	0.0080	0.1210	0.0177	0.0008				
	DT	0.1147	0.0021	0.0002	0.2157	0.0120	0.0728	0.0017				
	21	0.1117	0.0021	0.0002	0.2107	0.1711	0.0102	0.0017				
	Emission Probabilities											
			Janet	will	back	the	e b	ill				
	NNI	P	0.000032	0	0	0.0	00048 0					
	MD		0	0.30843	1 0	0	0					
	VB		0	0.00002	8 0.000	672 0	0	.000028				
	JJ		0	0	0.000	340 0	0					
	NN		0	0.00020	0 0.000	223 0	0	.002337				
	RB		0	0	0.010	446 0	0					
	DT		0	0	0	0.5	06099 0					
3B.	Distinguis	h betwe	een oper	n and c	losed	word	classes	of English.	3M	CO4	1, 2, 4	4
	Taking ver	rb as an	example	e word	class e	xplain	its ope	n and closed				
	form.		1			1	1					
3C	"Distingui	shing he	etween n	articles	(RP) a	nd pre	position	ns (IN) poses	2M	CO2	1234	2
	a challengi	ing task'	" Illustra	te with	an exa	mnle	position		2111	002	I,2,0, 1 ,	L 2
	d chancing		· musue								5	
4A.	Apply CK	Y to pa	arse the	sentence	e He s	saw the	e man 1	with the lens	4M	CO3	1,2,3,4,	3
	using the g	grammai	r given b	elow. S	how th	e pars	e table a	and the parse			5	
	tree.											
	S -> NP V	Р										
	NP -> DE1	ΓΝ										
	NP -> NP	рр										
	YY->YNY											
	VP -> V NP											
	VP -> VP PP											
	DET -> the											
	NP -> He											
	N -> man											
	N -> lens											
	P -> with											
	$V \rightarrow saw$											
	N -> cat											
	N -> roof											
	P -> from											
	P -> on											
	P -> in											
	How do yo	ou conv	ert a giv	en gran	nmar to	o Chor	nsky N	ormal Form?				
4B.	What kind	of ambi	iguity ex	ists whi	le pars	ing the	senten	ces "I ate the	3M	CO3	1,2,3,4	2
·	spaghetti	with ch	opsticks	" and	"I ate	the s	paghet	ti with meat			_,_,_,_,	_
	balls". She	ow the c	orrespon	iding pa	rse tree	es.	1 . 0.101				ر ا	
40	Compare t	on-dow	n and bot	ttom_up	narcin	o with	relaven	t examples	21/	CO3	1724	Δ
TO .	List the ad	vantage	s and lim	itations	Puisin	5 "		e chumpies.	5111		1, 2 ,0, 1 ,	
	List the au	vantage	s and III	manons	•						5	

5A.	Explain agreement phenomenon in English grammar with suitable example. How do we handle them in Context Free Grammar (CFG)? Give an example.	4M	CO4	1,2,3,4	2
5B.	What are the pre-modifiers of a head noun? Give its Context Free Grammar (CFG) structure along with suitable examples.	4M	CO2	1,2,3,4, 5	2
5C.	Given the parse trees for a given Probabilistic Context Free Grammar (PCFG), compute the probabilities P(t1) and P(t2). Which is parse is more likely? $t_1:$ $S_{1.0}$ $V_{P_{0.7}}$ $t_2:$ $S_{1.0}$ $V_{P_{0.3}}$ $s_{astronomers}$ $V_{P_{0.7}}$ $V_{P_{0.3}}$ $V_{P_{0.7}}$ $V_{P_{0.3}}$ $V_{P_{0.7}}$ $V_{P_{0.3}}$ $V_{P_{0.7}}$ V_{P_{0	2M	CO5	1,2,3,4	3