



**PES UNIVERSITY**

**UE15CS333**

SRN

--	--	--	--	--	--	--	--	--	--

**B. Tech. 6th Semester: Jan - May 2018**  
**End Semester Examination**  
**Natural Language Processing**

**Prof. NSK**

Time: 3 Hr

Answer All Questions

Max Marks: 100

You can bring one or more hand written notes. No printed matter is allowed. Answer precisely and briefly.  
Please make suitable assumptions if needed and state them.

1	a	Rule: "i before e except after c" Write a small program segment or a pseudocode to find misspelt words amongst a given list of words based on the above rule.	4
	b	Stemming Rule: "ing" → "" if stem contains a vowel Write a small program segment or a pseudocode to stem a given list of words	6
	c	A finite state automaton for English nominal inflection is as follows. 1. $\delta(q_0, \text{reg\_noun}) = q_1$ 2. $\delta(q_1, \text{plural}(-s)) = q_2$ 3. $\delta(q_0, \text{irreg\_sg\_noun}) = q_2$ 4. $\delta(q_0, \text{irreg\_pl\_noun}) = q_2$ $q_1$ and $q_2$ are final states.  We can reach final state $q_1$ in one way and $q_2$ in 3 ways. Give an example of an English word in each case. In rule 3, why not the destination state be $q_1$ ?	4 + 2
	d	Algo min_distance( $w_1, w_2$ ) gives the minimum distance between a pair of words. This is given to you. As we key in a word, match this word against a list of known words and prompt the likely word. If more than one word has the same minimum distance, select one of them at random.	4
2	a	"Never trouble trouble till trouble troubles you" We would like to analyze the word trouble using N-grams. i) using wordform ii) using lemma iii) using POS find the count using unigram and bigram technique.	6
	b	Based on a corpus, we have computed bigrams. Given a bag of words, write an algorithm or a program segment to guess the right sentence.	4
	c	Give an example of a word with multiple POS tags. How is the tag of the word resolved in a sentence using a) Ruled based tagging b) Stochastic tagging	6
	d	How can unknown words be tagged?	4

3	a	There are two forms of you in Indian languages – Hindi(tum, aap), Kannada(neevu, taavu). There should be agreement between the word for you in Indian language and its corresponding verb. Express this as grammar rules. Assume the simplest form of the rule is $S \rightarrow N V$ where N is noun and V verb.	4
	b	These are examples for different sentence level constructions. i. declarative : I prefer a morning flight ii. imperative : show the lowest fare iii. Yes-no-question: is there a morning flight? iv. wh-phrase: why there is no morning flight? For each of these sentences, give a simple grammatical structure.	6
	c	In Earley's algorithm, what does this state mean? $A \rightarrow x.By, [i, j]$ How do we indentify the successful parse based on a state in a chart entry?	4
	d	The professor said on Monday he would give an exam. Show that the above sentence has structural ambiguity by drawing multiple parse trees. Use a suitable grammar.	3 + 3
4	a	"bird eats worms" Express eats as a predicate. What is the restriction on the arguments of the predicate? How do we take care of phrases like: "He eats his own words" "chicken ate the monkey" - translation of the kannada phrase "kodugana koli nungitta"	2 + 2 + 2
	b	What are forward and backward chaining with respect to inferencing? What advantages and disadvantages of these approaches?	4
	c	1. exam(student) 2. exam(student, elective3) 3. exam(student, NLP) 4. exam(student, ESA) 5. exam(student, NLP, ESA) 6. exam(stduent, NLP, elective3) 7. exam(student, ESA, elective3) 8. exam(student, ESA, NLP, elective3) Given the above facts, express facts 7, 3 and 1 using 8 as meaning postulates using quantifiers and implication.	6
	d	Semantic attachment for a sentence can be done as follows. For a declarative statement, $S \rightarrow NP VP \{ DCL(VP.sem(NP.sem)) \}$ For an imperative statement given below, what could be the noun phrase? $S \rightarrow VP \{ IMP(VP.sem(???)) \}$ For Yes-NO-Query type of sentence, the following is the representation. $YNQ(\exists e \text{ raining}(e) \wedge \text{place}(\text{bangalore}) \wedge \text{time}(\text{now}))$ What sentence does this represent?	2 + 2

5	a	<p>State true or false. Give reasons. You may quote one or two examples if needed.</p> <p>i) wood and would are not good candidates for homonymy.</p> <p>ii) in spelling correction, words like founded cause difficulties.</p> <p>lii) In speech recognition, homophones do not cause problems</p> <p>iv) A word may have multiple pronunciation. This causes problems in speech to text.</p> <p>v) In information retrieval, homographs may result in wrong or unwanted information.</p> <p>Hint:</p> <p>homonymy: are words which sound alike or are spelled alike, but have different meanings</p> <p>homophones: are words pronounced the same; have different meanings</p> <p>homographs: are words with same written form; have different meanings</p>	10
	b	<p>"washing dishes" and "stir-fry dishes". In these phrases, the verb is used to disambiguate the word dishes.</p> <p>How about such an approach in the following cases?</p> <p>i) What dishes do you recommend?</p> <p>ii) you cannot eat gold if you are hungry.</p> <p>lii) He literally eats glass</p>	2+ 2+2
	c	<p>Hyponymy relationship for car could be:</p> <p>car =&gt; 4 wheeler =&gt; automobile =&gt; vehicle =&gt; object</p> <p>Would this mean, every occurrence of a car can be replaced by any of these superclasses in any sentence? Give your reasons.</p>	4