

NATIONAL INSTITUTE OF TECHNOLOGY ROURKELA

Department of Computer Science and Engineering

Spring Alternative Mid Semester Examination (March), 2019

Subject: **Natural Language Processing**(CS 6314)

Write neatly and legibly.

Answer **all** the questions.

Figures at the right hand margin indicate marks.

All parts of a question should be answered at one place.

Number your answers correctly according to the numbering system used in this question paper.

Full Marks: 30

Time: 2 Hours

1. Answer the following questions. [2 x 4 = 8]
 - (a) Describe the ambiguity issues in various stages of Natural Language Processing.
 - (b) What is Morphology? Explain various types with suitable examples.
 - (c) How do Finite State Transducers take an advantage over Finite State Machines? Explain with a suitable example.
 - (d) States various methods to compute weights for Vector Space Model.
2. Consider the following documents (D_1 , D_2 , and D_3) and the query Q . [6]

D_1 : Shipment of gold damaged in a fire
 D_2 : Delivery of silver arrived in a silver truck
 D_3 : Shipment of gold arrived in a truck
 Q : gold silver truck

Compute the similarity coefficient between the query and documents and retrieve the documents ranks using Vector Space Model
3. Consider the query and documents list from question #2. Assuming that all three documents are training data and D_2 and D_3 are relevant to the query, retrieve the documents for the query using Probabilistic retrieval strategy given by Roberson and Sparck Jones. Compare this method with the Poisson model given by Robertson and Walker. (For the Poission model assumes the constants $k_1 = 1$, $k_2 = 0$, $k_3 = 8$, $b = 0.6$.) [6]
4. Consider the query and documents list from question #2. Compute the similarity coefficient between query and documents using a language model given by Ponte and Croft. [6]
5. For a given language the probabilities given in the table holds with respect to the occurrences of Noun, Adjective and Verb for machine translation to another language. [4]

	Noun	Adjective	Verb
Noun	0.8	0.05	0.15
Adjective	0.2	0.6	0.2
Verb	0.2	0.3	0.5

Given that the first word is Noun, what's the probability that the next word is Noun and the word after is Adjective?

Given that current word is a Verb, what's the probability that an Adjective will occur two words from now?