# National Institute of Technology Rourkela

## Department of Computer Science and Engineering

Spring End Semester Examination (April), 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: 50 Time: 3 Hours

1. Answer the following questions. [4 x 5 = 20]
   (a) Describe the characteristics of Indian languages. Detail various machine translation approaches.
   (b) Explain various approaches for word sense disambiguation in NLP.
   (c) Give the architecture of Text to Speech (TTS) of speech synthesis and explain in details various sub-steps involved.
   (d) Project various historical research advances in speech synthesis from beginning to till date.
   (e) Describe various knowledge representation formalisms with examples.

2. Consider the following documents ($D_1$, $D_2$, and $D_3$) and the query $Q$. [10]

   $D_1$: 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 Language Model

3. Assume a simplfied model of weather prediction in the following table. (Y-axis as Today's weather and X-axis Tomorrow's weather) [6]

| | Sunny | Rainy | Foggy |
|---|---|---|---|
| Sunny | 0.8 | 0.05 | 0.15 |
| Rainy | 0.2 | 0.6 | 0.2 |
| Foggy | 0.2 | 0.3 | 0.5 |

   Suppose you were locked in a room for several days, and you were asked about the weather outside. The only piece of evidence you have is whether the person who comes into the room carrying your daily meal is carrying an umbrella or not. Assume the following probabilities.

| | P(Umbrella) |
|---|---|
| Sunny | 0.1 |
| Rainy | 0.8 |
| Foggy | 0.3 |

(a) Given that today is sunny, what's the probability that tomorrow is sunny and the day after is rainy?

(b) Given that today is foggy, what's the probability that it will be rainy two days from now?

(c) Suppose the day you were locked in it was sunny. The next day the caretaker carried an umbrella into the room. Assuming that the prior probability of the caretaker carrying an umbrella on any day is 0.5, what's the probability that the second day was rainy?

(d) Suppose the day you were locked in the room it was sunny; the caretaker brought in an umbrella on day 2 but not on day 3. Again assuming that the prior probability of the caretaker bringing an umbrella is 0.5, what's the probability that it's foggy on day 3?

4. Assuming various psycho-physiological signals of an affect like galvanic skin response (R), Electrocardiogram (E), Blood pulse (B) and Electroencephalogram (G) with two affects Happy (H) and Sad (S) the following are the data provided for a decoding problem. The R,B,E,G values for the affect Sad (S) are given as $0.3, 0.2, 0.2, 0.3$ respectively and the R, B, E, G values for the affect Happy (H) are given as $0.2, 0.3, 0.3, 0.2$ respectively. The transition probabilities from S to S is 0.7, S to H is 0.3, H to H is 0.4 and H to S is 0.6. Assuming the sequence $\theta$ = (R G G B), select the best path for decoding using Viterbi algorithm. [6]

5. Considering the following sentences in a discourse [8]

John saw a beautiful 1961 Ford Falcon at the used car dealership. He showed it to Bob. He bought it.

(a) Deploy Resolution of Anophora Procedure algorithm to assign the anaphora to the corresponding nouns.

(b) Apply the centering algorithm to evaluate the central theme.