CSE556: NLP - Quiz 1 (Non-mandatory)

Marks: 15
Duration: 30 mins **Date:** 19-Feb-2024

1. Distinguish the following based on their density, semantic, and dimensionality. a. One-hot, Co-occurrence, and Word embeddings.	[3]
One-hot: Very sparse, no semantic, |V|x|V|. Co-occurrence: Partial sparse/dense, Low-degree semantic, |V|x|V|. Word embeddings: Dense, High-degree semantic, |V| x d, where d << |V| 1 mark (0.33 * 3) each.	
2. Is the following grammar in Chomsky's normal form (CNF)? If not, convert it into CNF. G: E→E+E E→E–E	[3]
E → EA E → EB A → PE B → ME P → + M → - Note: Other solutions are also possible as long as we get similar outcomes from both grammar. No partial marking.	
3. In the formulation of HMM, justify the application of markov assumption. Give equations.	[2]
Probability of a state (s_t) depends on the previous few states instead of the entire sequence of states. For Markov assumption of size 2, it depends on the previous state only. $P(s_t \mid s_1 s_2 \ldots s_{t-1}, O) \Rightarrow P(s_t \mid s_{t-1}, O)$ 1 mark for justification and 1 for equation.	
4. How do we evaluate a smoothing technique? Mathematically describe it for a bi-gram language model.	[2]
We evaluate a smoothing technique by computing its reconstruction count (C*). A smoothing technique is good if it approximates the reconstructed count (C*) as close to the original count (C). For bigram model, $P(x\|y) = C(y,x) / C(y)$ $C^*(y,x) = P(x\|y) * C(y)$ 1 mark for justification and 1 for equation.	

5. What is the fundamental difference between lemmatization and stemming?	[1]
Both lemmatization and stemming normalize the given word; however, a lemma of a word is always a valid word but a stem of a word may or may not be a valid word. E.g., Files → Fly (lemmatization) Flies → Fli (stemming) Example is optional. 1 mark for correct answer.	
6. Identify semantic roles in the below sentence w.r.t. the predicate. "I saw a girl with a telescope."	[4]
saw → Predicate I → Agent girl → Patient Telescope → Instrument/Equipment 1 mark for each role.	