

Set A

CSE440: Natural Language Processing II

Date: 22 October 2025

This is an MCQ quiz. Each question is worth 1. Write your answers serially on the other side.

Questions:

1. Which of the following components are part of NLP Annotation?
 - i) Lemmatization/Stemming
 - ii) POS Tagging
 - iii) Named Entity Recognition
 - A) i,ii
 - B) i,iii
 - C) ii,iii
 - D) i,ii,iii
2. "One morning I shot an elephant in my pajamas" demonstrates-
 - A) Coreference error
 - B) Attachment ambiguity
 - C) POS ambiguity
 - D) Semantic role labeling
3. What makes sentence segmentation difficult?
 - A) Long sentences
 - B) Abbreviations
 - C) Lack of punctuation
 - D) Use of adjectives
4. Which statement about tokenization solutions is TRUE?
 - A. One universal algorithm fits all languages
 - B. Each language requires its own tokenization principles
 - C. It is unnecessary for NLP in recent years.
 - D. Tokenization and stemming are the same
5. What problem do stemming and lemmatization try to solve?
 - A) Case sensitivity
 - B) Words that look different but have similar meaning
 - C) Word embedding
 - D) Sentence segmentation
6. Which is an example of an open-class word category?
 - A) Pronoun
 - B) Determiner
 - C) Conjunction
 - D) Noun
7. What challenge exists in POS tagging?
 - A) Words can have different tags depending on context
 - B) When there are too many punctuation marks
 - C) Some words are always proper nouns
 - D) Tag sets change daily

8. Why are rule-based sentence segmentation systems hard to maintain?

- A) They are too simple to implement
- B) New punctuation rules constantly emerge
- C) The number of exceptions grows quickly
- D) They can't recognize capital letters

9. Which scenario best shows a tokenization error?

- A) Treating "New York" as one token
- B) Splitting "isn't" into "is" and "n't"
- C) Considering "e-mail" as two tokens
- D) Recognizing "Mr." as a one token

10. What is the main advantage of lemmatization over stemming?

- A) It is faster
- B) It produces linguistically valid words
- C) It removes punctuation
- D) It does not require POS tagging

Name:

Student ID:

Section:

Answer

Set A

- 1 C**
- 2 B**
- 3 B**
- 4 B**
- 5 B**
- 6 D**
- 7 A**
- 8 C**
- 9 C**
- 10 B**