Part I: For each of the following 9 questions, give a short reply in the space provided.

1. [Linguistic Foundations] Provide a definition of prescriptive grammar and a definition of descriptive grammar.

2. [Cognitive Foundations] Name, and provide definition for, the two parsing principles proposed in Lyn Frazier's Garden-Path Theory.

3. [Technological Foundations] Describe inputs and outputs for Text Summarization and Report Generation.

4. [Finite State Methods] In many regular expression frameworks, such as Perl regex, you have an additional operator: (α {2, 4} which means that you can repeat whatever the regex α extends to from two to four times. Write down a regular expression with only the basic operators that accept the same language.

5. [Parsing] Name the basic operations in the Earley parsing algorithm and briefly describe their functionalities.

6. [Statistical NLP] Write down the formula for a unigram backing-off language model using absolute discounting. Give a one-line description for each element of the formula.

7. [Speech] The task of the duration model in a text-to-speech system is to predict the duration of each speech sound in the synthetic utterance, given an input text. Why is this task difficult?

8. [Semantics] Why is word-sense disambiguation an extremely difficult problem (both for knowledge-based and supervised machine learning approaches)? Explain with an example (maybe taken from the course's slides), why the treatment of WSD via inference is a particular challenge.

9. [Discourse] Name three salience factors used by the RAP system and provide examples for each.

Part II: Answer any 3 of the following 5 questions. If you answer more, you will be graded only on the first three you answer. If you wish an answer to not be graded, clearly draw a line through your answer.

Q1 [Linguistic Foundations]: Answer each of the following two questions:

1. Explain the cause of ambiguity in each of the following sentences:
    a. "*The little boy hit the child with the toy.*"
    b. *"We'll give you a ring tonight."*
2. Explain how lexical ambiguity affects structural ambiguity in the sentence : *"I made her duck"* (Hint: you might draw syntax trees)

Q2 [Parsing]: Using the following CFG (with S as the start symbol):

(1) S->NP VP
(2) NP->D N | N N | time
(3) VP->V NP|VP PP | flies
(4) PP->P NP
(5) N -> time | flies | arrow
(6) V -> like
(7) P -> like
(8) D -> an

Draw the complete parsing chart with Earley's algorithm for the following sentence:

"time flies like an arrow"

Q3 [Statistical NLP]: As a homework task, you implemented a spam mail classifier including a simple training section and a testing section. Please give the most essential elements in pseudo code.

Q4 [Speech] Explain the key properties of diphone synthesis and unit selection synthesis, respectively. What do these approaches have in common and how do they differ from each other? Please list at least 3 commonalities and 3 differences. Unit selection is considered as state-of-the art in speech synthesis – are there any advantages of diphone synthesis over unit selection?

Q5 [Semantics] Give appropriate predicate-logic representations (1) and (2) for sentences (1') and (2'), respectively:

(1') Digs aren't vegetarians (2') Mammals aren't vegetarians

Specifiy the predicate-logic axiom (3) that expresses the fact that "mammal" stands in a hypernymy relation to "dog".

Compute the truth conditions of formulas (1), (2), and (3), respectively, using the interpretation function introduced in the course.

Now consider the following entailment problem: {(1), (3)} |= (2)

On the basis of the truth conditions, decide whether entailment really holds in this case, and give an explicit justification for your decision.

Hint: Since (1) and (2) are structurally identical, you can save writing time by deriving the explicit interpretation for one of the two only; you can obtain the other by just replacing one predicate.