

UNIVERSITI SAINS MALAYSIA

First Semester Examination
Academic Session 2000/2001

September/October 2000

CSC514 - Natural Language Processing

Duration : [3 hours]

INSTRUCTION TO CANDIDATE:

- Please ensure that this examination paper contains **FOUR** questions in **FOUR** printed pages before you start the examination.
 - This is an 'Open Book' Examination.
 - Answer **ALL** questions. You can choose to answer either in Bahasa Malaysia or English.
-

1. Given the following sentence:

"Green herons fly in groups"

where the possible syntactic categories of each word in the sentence are:

fly, groups : **verb**
 green, herons, fly, groups : **noun**
 in : **preposition**
 green, fly : **adjective**

- (a) Draw all possible syntactic trees for the given sentence. [25/100]
- (b) Provide a Context Free Grammar which is capable of generating all the trees as identified in (a) [25/100]
- (c) Construct a detailed chart illustrating the parsing process of the sentence "Green herons fly in groups" based on the bottom-up chart parsing technique. [25/100]
- (d) List out all the excessive partial parse trees (if any) constructed in (c). Discuss briefly on how the parsing process can be improved further based on the top-down chart parsing technique. [25/100]
2. Given the following sentence:

"John painted the bridge"

where the possible senses of each word in the sentence are :

John1 : name of a person
 the1 : determiner used before nouns to refer to things mentioned earlier
 painted1 : cover a surface with paint; e.g. he painted the walls white
 painted2 : describe vividly in words; e.g. she painted a gloomy picture of her life at the orphanage
 painted3 : adorn by using cosmetics etc; e.g. she did her hair and painted her face
 bridge1 : an elevated structure that spans a body of water
 bridge2 : a card game
 bridge3 : a dental device

bridge4 : an abstract notion of providing a connection

- (a) Give the quasi-logical form for the sentence " John painted the bridge"
[25/100]
- (b) Provide a simple Context Free Grammar and the corresponding lexicon both enhanced with SEM (semantic) features which is capable of generating the quasi-logical form given in (a).
[40/100]
- (c) Provide a simple word sense hierarchy and the corresponding selectional restrictions for paint1, paint2 and paint3 such that "John painted the bridge" will be interpreted as "John1 painted1 the1 bridge1".
[35/100]
3. Using the frame-based representation, define the action class DEPOSIT that corresponds to a sense of the verb **deposit** as in the sentence below:

(I) Ali deposited some money into his account.

In particular, your definition should contain enough detail so that each of the following statements could be concluded from sentence (I).

- (II) Ali had some money.
 (III) Ali had an account at a bank.
 (IV) Ali went to a bank.

For each of these sentences, discuss in detail how the necessary knowledge is represented (as a precondition, effect, decomposition, and so on). Identify other conclusions (if any) that can be made from sentence (I), which can be used to generate the expectation needed to decide on the appropriate interpretation of sentence 3b based on the matching algorithm.

- 3a Ali deposited some money into his account.
 3b He left the bank.

Assuming sentence 3b can be interpreted as

Interpretation 1:

LEAVE1(E1) & Agent(E1) = Ali1 & Theme(E1) = Bank1

Interpretation 2:

LEAVE1(E2) & Agent(E2) = Ali1 & Theme (E2) = Bank2

Where

Bank1: A business that keeps money for individuals or companies, exchanges currencies, makes loans, and offers other financial services.

Bank2: The steep side of a river, stream, lake, or canal.

[100/100]

4. (a) Describe in detail how the semantic distance between word senses can be derived from (i) monolingual dictionary, (ii) a training corpus annotated with sense number.

[40/100]

- (b) Discuss in detail how the semantic distance between word senses obtained in (a) can be applied to each of the following NLP applications.

(i) Text summarization

(ii) Text categorization

(iii) Text classification

(iv) Machine translation

[60/100]