

UNIVERSITI SAINS MALAYSIA

First Semester Examination
Academic Session 1999/2000

September 1999

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:

"Time flies like an arrow"

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

time : verb, noun
 flie : verb, noun
 like : verb, preposition
 an : article
 arrow : noun

- (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) Based on the grammar provided in (b), give a chart structure representing all the trees in (a). [25/100]
- (d) What are the parsing techniques that can be used to produce the chart structure given in (c)? Give a brief description of how each technique works and the main difference between these techniques. [25/100]

2. Given the following sentence:

"The dishwasher saw the article"

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

the1 : determiner used before nouns to refer to things mentioned earlier
 dishwasher1 : a machine for washing dishes
 dishwasher2 : someone who washes dishes
 saw1 : cut with a power tool for cutting wood
 saw2 : past tense of SEE
 article1 : a piece of writing on a particular subject in a newspaper
 article2 : a determiner that may indicate the specificity of reference of a noun phrase

- (a) Give all possible logical forms (with thematic roles) for the given sentence considering only word sense ambiguity (i.e. ignore ambiguity due to the scoping of 'the'). Give also the corresponding quasi-logical form. [25/100]

(b) Provide a simple Context Free Grammar and the corresponding lexicon both enhanced with SEM (semantic) features that are capable of generating the quasi-logical form given in (a). [40/100]

(c) Based on the grammar provided in (b), give the parse tree for "The dishwasher saw the article" showing the SEM features. [35/100]

3. Using the frame-based representation, define the action class DRIVE that corresponds to a sense of the verb **drive** in

(I) Ali drove a car to school.

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

(II) Ali was inside the car at some time.

(III) Ali had the car keys.

(IV) The car was at school for some time.

(V) Ali opened the car door.

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 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 expectation vs. interpretation matching algorithm.

3a Ali drove a car to school.

3b He parked under a tree.

Assuming sentence 3b can be interpreted as

Interpretation 1:

PARK1(E1) & Agent(E1) = Ali1 & UNDER-LOC(E1) = Tree1

Interpretation 2:

PARK2(E2) & Agent(E2) = Ali1 & UNDER-LOC(E2) = Tree1

Where

PARK1 : to put or place (a car or other vehicle) for a time

PARK2 : to settle (oneself) in a particular place for a long time, e.g. he parked himself beside me.

[100/100]

4. Discuss in detail how the lexical semantic research results obtained at UTMK can be applied to each of the following NLP applications.
- (a) Text summarization [20/100]
 - (b) Text categorization and classification [20/100]
 - (c) Machine translation [20/100]
 - (d) Message filtering and routing system [20/100]
 - (e) Information retrieval (search) system [20/100]

- oooOooo -