UNIVERSITI SAINS MALAYSIA

Second Semester Examination
Academic Session 2004/2005

March 2005

CPT315 - Natural Language Processing

Duration : 2 hours

INSTRUCTIONS TO CANDIDATES:

- Please ensure that this examination paper contains **SIX** questions in **THREE** printed pages before you start the examination.

- Answer **ALL** questions.

ENGLISH VERSION OF THE QUESTION PAPER

1. Why is it relevant to distinguish between the notion "word", "word-form", "lemma", "token", and "type"? Illustrate your answer with the help of this example: "I have a can opener but I can't open these cans."

(10/100)

2. Explain the difference between the terms "dictionary" and "lexicon" in NLP.

(10/100)

3. Morphological analysis is the identification of a word-stem from a full word form. Full word forms can be simple (= stem), derived, inflected, or compound.

   (a)   Classify the words in the following list as 'stem', 'derived word', 'inflected word', 'compound word', or 'other'. (N = Noun, V = Verb, Adj = Adjective, Adv = Adverb).

   duck_N/V, ducks_N/V, duckling_N('a baby duck'), ducked_V, ducking_V
   kind_N/Adj, kinds_N, kindly_Adv/Adj, kindness_N, kind_hearted_Adj
   drop_N/V, dropped_V, dropping_V, droppings_N
   hide_V/N, hiding_V, hidden_N, hide-and-seek_N

   (b)   In the list above, some words end with "-ing" (e.g., ducking). Using this set of words, outline the rules required to generate the stems (= stem words).

(20/100)

4. This context-free grammar is ambiguous. Explain why this grammar is ambiguous? Give the different kinds of linguistic ambiguity that you can find?

   S   →   NP VP
   NP →   NP PP
   VP →   saw NP
   VP →   saw NP PP
   NP →   the spy
   NP →   the cop
   PP →   with binoculars

(S = Sentence, NP = Noun Phrase, PP = Prepositional Phrase, VP = Verb Phrase)

Two derivations of the sentence: "The cop saw the spy with binoculars."

S	S
NP VP	NP VP
the cop VP	the cop VP
the cop saw NP	the cop saw NP PP
the cop saw NP PP	the cop saw the spy PP
the cop saw the spy PP	the cop saw the spy with binoculars
the cop saw the spy with binoculars	

(20/100)

5. 'Example-based' is a technique for machine translation. Describe the general architecture of an example-based machine translation.

(20/100)

6. Some TTS systems include a 'pre-processing module'. Explain clearly with some examples what this module does.

(20/100)

-oooOooo-