Natural Language Processing
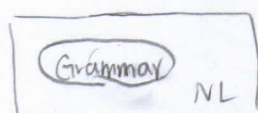## Midterm Examination

Name: 許景皓

Student ID : 60647085S

Date: April 26, 2018

Time: 14:20-16:20

1. Give two issues for evaluating a particular grammar for a language. You should explain for your issues. (10 points)

+10

Grammar NL   undergenerate   此 Grammar 產生屬於 NL
                               部份 NL 不能由 Grammar 產生

NL Grammar   overgenerate   NL 的語句皆可由此 Grammar
                               此 Grammar 所產生的不一定屬於 NL
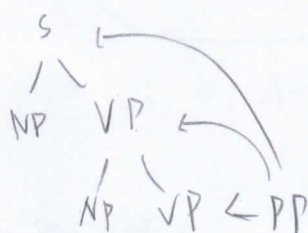
∴ 要依照 NL 給予適當的 Grammar

2. What is part-of speech?　Give 5 examples that correspond to 5 different part of speech symbols.　(10 points)

5 example　5 different port of symbol

1. lexicon ambiguity

2. structual ambiguity

PP attatchment

S

NP VP

NP VP < PP

+2

coordinate
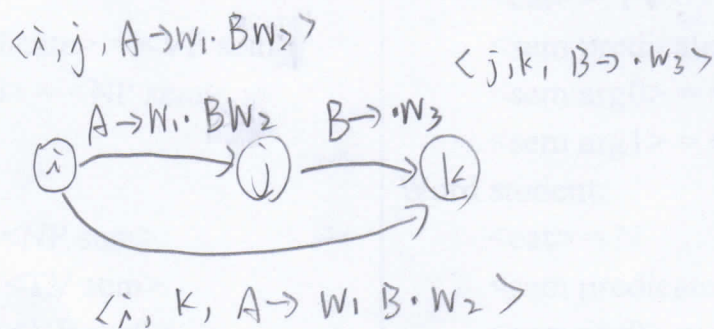
A and B or C

→ (A and B) or C

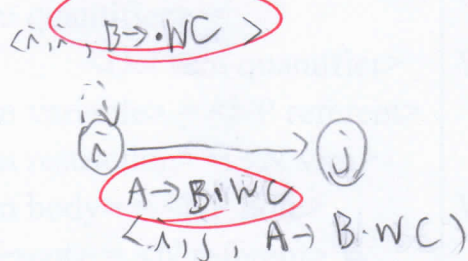→ A and (B or C)

noun — noun compounding

mail box and

post office

3. What is the data structure Chart? Please specify (1) the fundamental rule and (2) the bottom-up rule using Chart. (15 points)

(1) fundamental

$$\langle i, j, A \rightarrow W_1 \cdot B W_2 \rangle \qquad \langle j, k, B \rightarrow \cdot W_3 \rangle$$

$$A \rightarrow W_1 \cdot B W_2 \qquad B \rightarrow \cdot W_3$$

$$\langle i, k, A \rightarrow W_1 B \cdot W_2 \rangle$$

(2) bottom up

$$\langle i, i, B \rightarrow \cdot WC \rangle$$

$$A \rightarrow B \cdot WC$$

$$\langle i, j, A \rightarrow B \cdot WC \rangle$$

4. Consider the semantic representation of the sentence "Every student majored in a subject." Please reference to the corresponding grammar, and fill in blanks of (A), (B), (C), (D), (E), (F), (G), (H), (I) and (J) in the semantic representation. (10 points)

Rule	Word "majored in":
S → NP VP :	<cat> = TV
<S sem predicate> = <VP sem>	<sem predicate> = majored_in
<S sem arg0> = <NP sem>.	<sem arg0> = <arg0>
Rule	<sem arg1> = <arg1>.
VP → TV NP :	Word student:
<VP sem> = <NP sem>	<cat> = N
<NP hole> = <TV sem>	<sem predicate> = student
<TV arg0> = <VP arg0>	<sem arg0> = <referent>.
<TV arg1> = <NP referent>	Word subject:
Rule	<cat> = N
NP → Det N:	<sem predicate> = subject
<NP sem quantifier> =	<sem arg0> = <referent>.
<Det sem quantifier>	Word a:
<NP sem variable> = <NP referent>	<cat> = Det
<NP sem restriction> = <N sem>	<sem quantifier> = exists.
<NP sem body> = <NP hole>	Word every:
<NP referent> = <N referent>	<cat> = Det
	<sem quantifier> = all.

The semantic representation for the sentence: Every student majored in a subject.

sem : quantifier : (A) __All__
    variable : (B) __X__
    restriction : arg0 : (C) __X__
        predicate : student
    body : quantifier : (D) __exists__
   variable : (E) __Y__
        restriction : arg0 : (F) __Y__
            predicate : (G) __a- subject__
        body : arg0 : (H) __X__
            arg1 : (I) __Y__
            predicate : (J) __majored_in__

+10

5. Consider the following figure. It shows three layers of analysis, including syntactic analysis, semantic analysis and pragmatic analysis. A dot (·) indicates a sentence, a syntactic structure, a semantic representation, or a pragmatically acceptable interpretation depending on the layer of analysis. Some cases have been explained in this figure, and some cases are still left open for your answers. Please fill in (A), (B), and (C). Your answers should include the terminology to describe the linguistic phenomenon and an example for each case. (15 points)
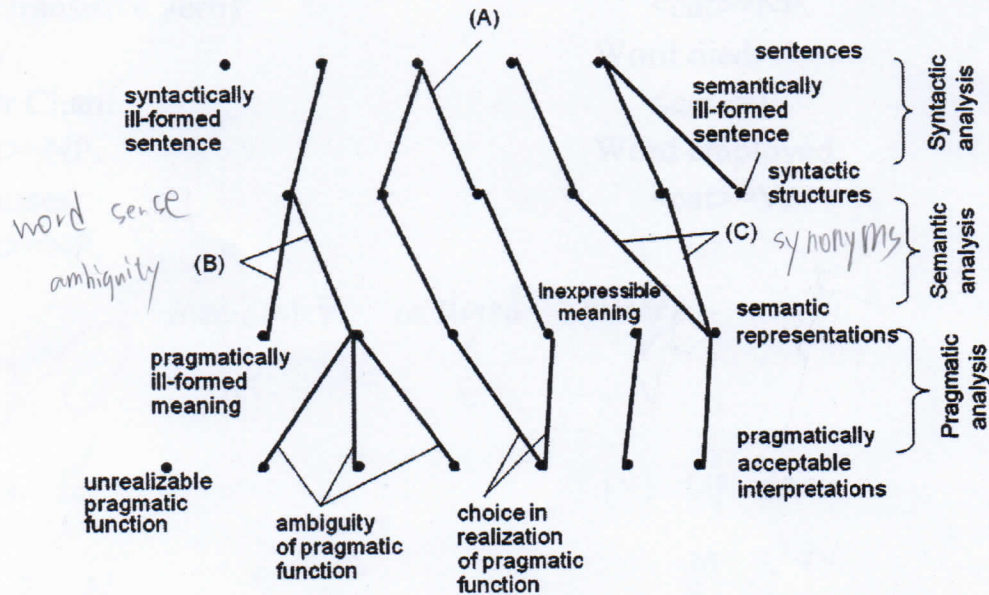
(A) structure ambiguity

同一個語意但句型不同

ex: 我在公園看到一位小女孩

我看到一位小女孩在公園

(B) word sense ambiguity

ex: bank 有河岸和銀行

(C) synonyms 同義

6. If we have the following grammar, show the parsing tree of "MediCenter employed nurses" using (1) bottom-up parsing and (2) top-down parsing.   (20 points)

Rule {simple sentence formation}
  S→NP VP.

Rule {transitive verb}
  VP→V NP.

Rule {intransitive verb}
  VP→V.

Word Dr Chan:
    <cat>=NP.

Word nurses:
    <cat>=NP.

Word MediCenter:
    <cat>=NP.

Word patients:
    <cat>=NP.

Word died:
    <cat>=V.

Word employed:
    <cat>=V.

MediCenter   employed   nurses
     M   E   N

(1) bottom up

M E N
↓
NP E N
↓
NP V N
↓
NP V NP
↓
NP VP
↓
S

```
       S
      / \
    NP   VP
     |  / \
     M  V  NP
        |   |
        E   N
```

(2) top down

M E N
↓
NPa E N
↓
NPa V N
↓
NPa VPa N
↓
S a N (failed
↓
NPa VPa NPb
↓
NPa V NPb
↓
Sb

```
            Sb
            |
   ┌────────┴────────┐
 NPa ---Sa (failed) NPb
   |          |       |
 NPa        VPa     NPb
   |          |       |
   |          V       |
   M          |       N
              E
```
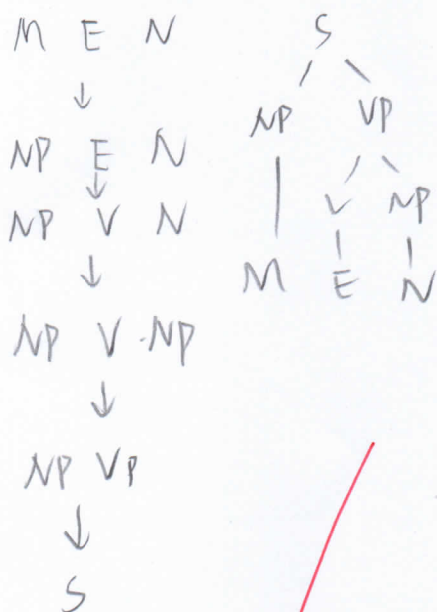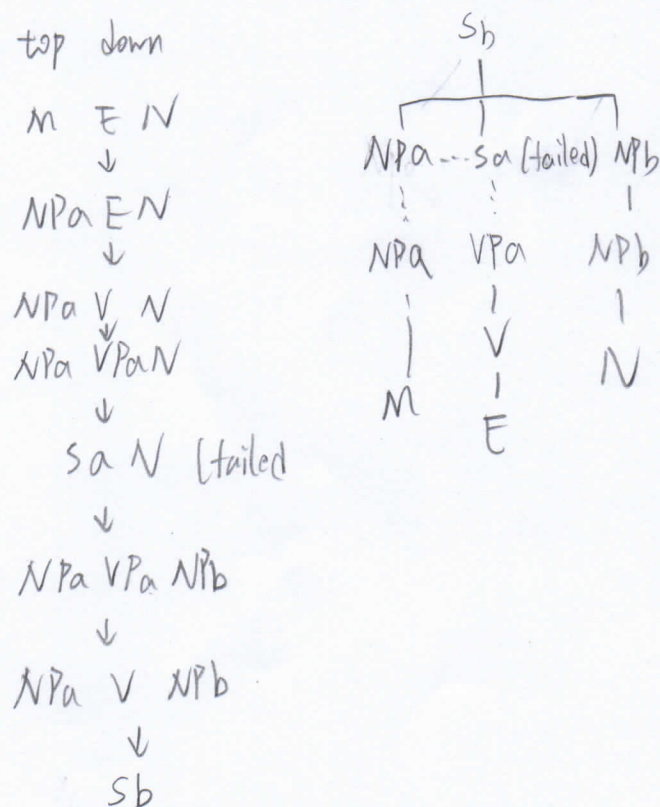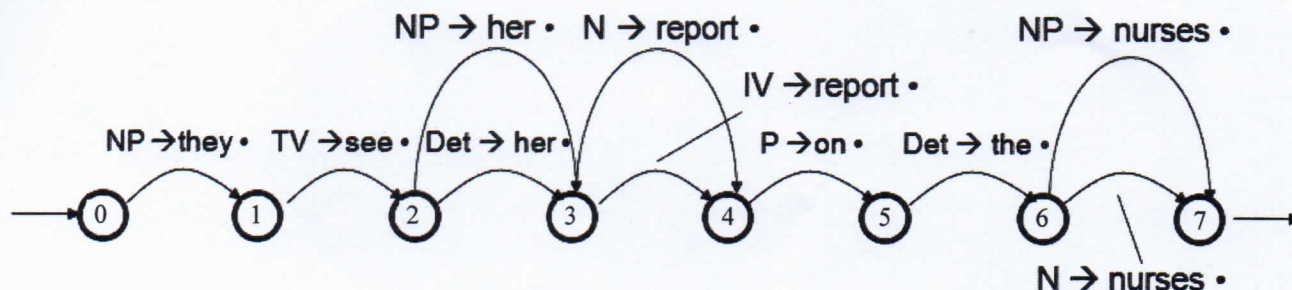
7. What is the top-down strategy in rule invocation for chart data structure? Suppose we have the following chart and the grammar rules. Please draw the new charts when adopting the top-down strategy step by step. For simplification, please only add the new edges to the first node (node 0). You don't need to apply to the other 7 nodes. (20 points)

NP → her • N → report • NP → nurses •

IV → report •

NP → they • TV → see • Det → her • P → on • Det → the •

0 1 2 3 4 5 6 7

N → nurses •

Grammar rules:

Rule {simple sentence formation}
  S → NP VP

Rule {intransitive verb}
  VP → IV

Rule {intransitive verb plus PP complement}
  VP → IV PP

Rule {transitive verb}
  VP → TV NP

Rule {transitive verb plus PP complement}
  VP → TV NP PP

Rule {transitive verb plus VP complement}
  VP → TV NP VP

Rule {simple noun phrase}
  NP → Det N

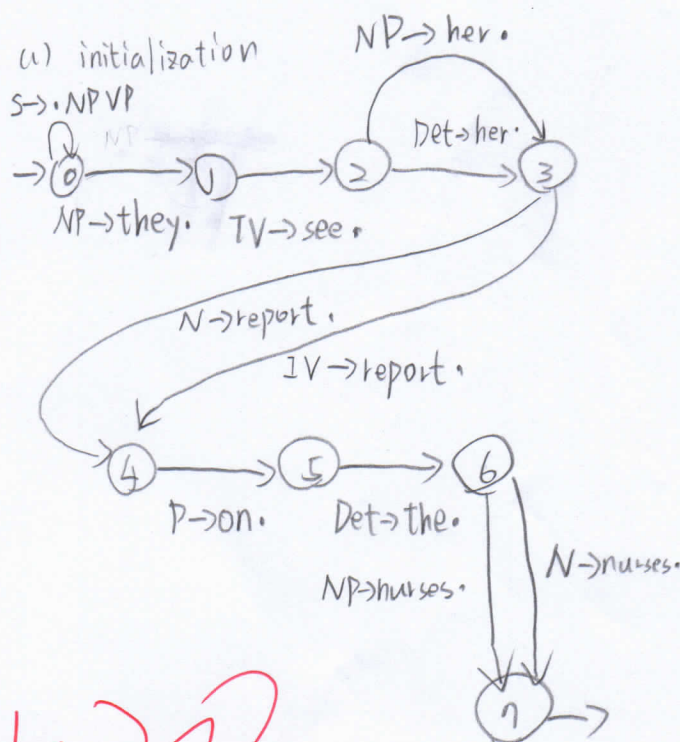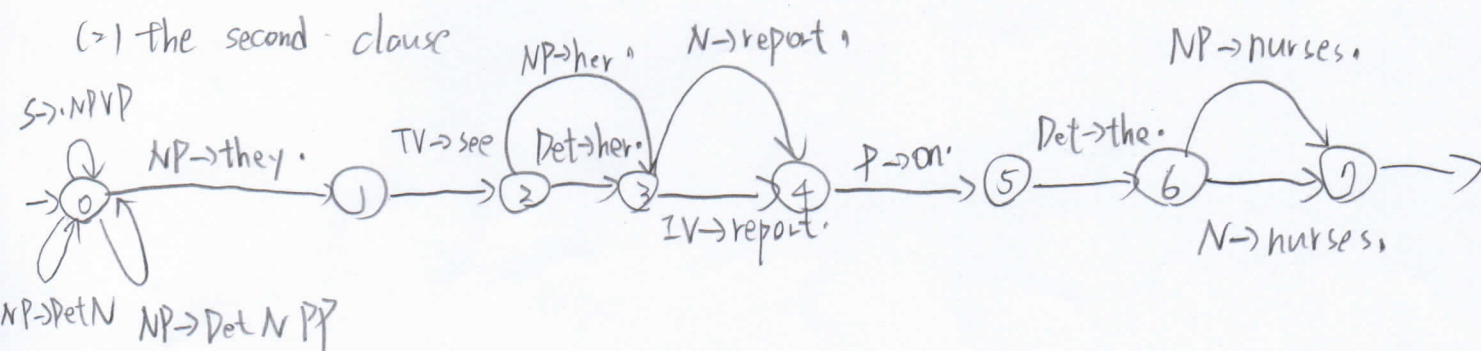Rule {noun phrase with PP complement}
  NP → Det N PP

Rule {simple prepositional phrase}
  PP → P NP

(1) initialization

S→.NPVP

NP→her•

Det→her•

→ 0 1 2 3

NP→they• TV→see•

N→report•

IV→report•

4 5 6

P→on• Det→the•

NP→nurses•

N→nurses•

7

+20

(2) the second clause

S→.NPVP

NP→they• TV→see• Det→her•

NP→her• N→report•

0 1 2 3 4

IV→report•

P→on• Det→the•

NP→nurses•

5 6 7

N→nurses•

NP→DetN NP→Det N PP

8. In question-answering applications, users ask questions with natural language statements, and a system answers the questions based on a database. Suppose you are given the web as the supporting database in your question-answering system. Please propose such a natural language understanding system.   (5 points, bonus)

用 why, how, when , where , what
做詢問句, 再依內容做 Query

+5