

Natural Language Processing

Midterm Examination

Name: 張昇誠Student ID : 6044 70145

Date: April 21, 2016

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)

lexical particular : 同一個單字有多種意思

ex. He has many fans
(扇子、粉絲)

combination particular : 根據句子的前後組合而有不同意思

ex. $A(B C) \neq (A B) C$

2. Consider the semantic representation of the sentence "An airline took over every hotel chain." 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)

<p>Rule</p> <p>$S \rightarrow NP VP :$ $S \rightarrow Det N TV Det N$</p> <p>$\langle S \text{ sem predicate} \rangle = \langle VP \text{ sem} \rangle$</p> <p>$\langle S \text{ sem arg0} \rangle = \langle NP \text{ sem} \rangle.$</p> <p>Rule</p> <p>$VP \rightarrow TV NP :$</p> <p>$\langle VP \text{ sem} \rangle = \langle NP \text{ sem} \rangle$</p> <p>$\langle NP \text{ hole} \rangle = \langle TV \text{ sem} \rangle$</p> <p>$\langle TV \text{ arg0} \rangle = \langle VP \text{ arg0} \rangle$</p> <p>$\langle TV \text{ arg1} \rangle = \langle NP \text{ referent} \rangle$</p> <p>Rule</p> <p>$NP \rightarrow Det N:$</p> <p>$\langle NP \text{ sem quantifier} \rangle =$</p> <p style="margin-left: 150px;">$\langle Det \text{ sem quantifier} \rangle$</p> <p>$\langle NP \text{ sem variable} \rangle = \langle NP \text{ referent} \rangle$</p> <p>$\langle NP \text{ sem restriction} \rangle = \langle N \text{ sem} \rangle$</p> <p>$\langle NP \text{ sem body} \rangle = \langle NP \text{ hole} \rangle$</p> <p>$\langle NP \text{ referent} \rangle = \langle N \text{ referent} \rangle$</p>	<p>Word "took over":</p> <p>$\langle cat \rangle = TV$</p> <p>$\langle sem \text{ predicate} \rangle = \text{took_over}$</p> <p>$\langle sem \text{ arg0} \rangle = \langle arg0 \rangle$</p> <p>$\langle sem \text{ arg1} \rangle = \langle arg1 \rangle.$</p> <p>Word airline:</p> <p>$\langle cat \rangle = N$</p> <p>$\langle sem \text{ predicate} \rangle = \text{airline}$</p> <p>$\langle sem \text{ arg0} \rangle = \langle referent \rangle.$</p> <p>Word "hotel chain":</p> <p>$\langle cat \rangle = N$</p> <p>$\langle sem \text{ predicate} \rangle = \text{hotel_chain}$</p> <p>$\langle sem \text{ arg0} \rangle = \langle referent \rangle.$</p> <p>Word an:</p> <p>$\langle cat \rangle = Det$</p> <p>$\langle sem \text{ quantifier} \rangle = \text{exists}.$</p> <p>Word every:</p> <p>$\langle cat \rangle = Det$</p> <p>$\langle sem \text{ quantifier} \rangle = \text{all}.$</p>
---	---

The semantic representation for the sentence: (An)(airline)(took over)(every)(hotel chain.)

sem : quantifier : (A) exist

variable : (B) ~~Det~~

restriction : arg0 : (C) X

predicate : airline

body : quantifier : (D) all

variable : (E) ~~Det~~

restriction : arg0 : (F) Y

predicate : (G) hotel chain

body : arg0 : (H) X

arg1 : (I) Y

predicate : (J) took over

3. 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}

$S \rightarrow NP VP$.

Rule {transitive verb}

$VP \rightarrow V NP$.

Rule {intransitive verb}

$VP \rightarrow V$.

Word Dr Chan:

$\langle cat \rangle = NP$.

✓ Word nurses:

$\langle cat \rangle = NP$.

✓ Word MediCenter: M

$\langle cat \rangle = NP$.

Word patients: P

$\langle cat \rangle = NP$.

Word died: D

$\langle cat \rangle = V$.

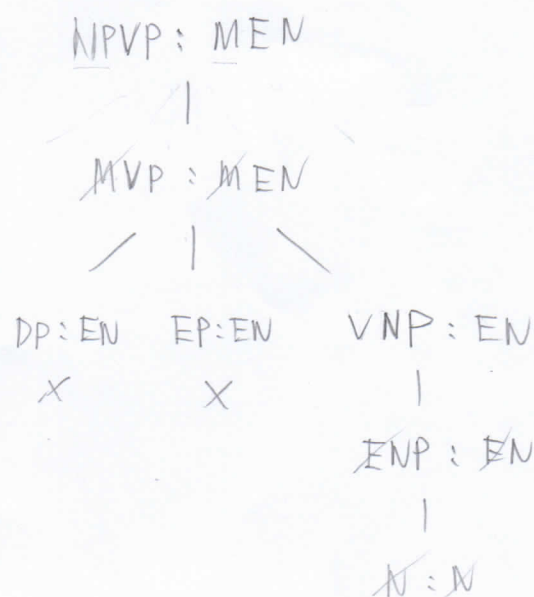
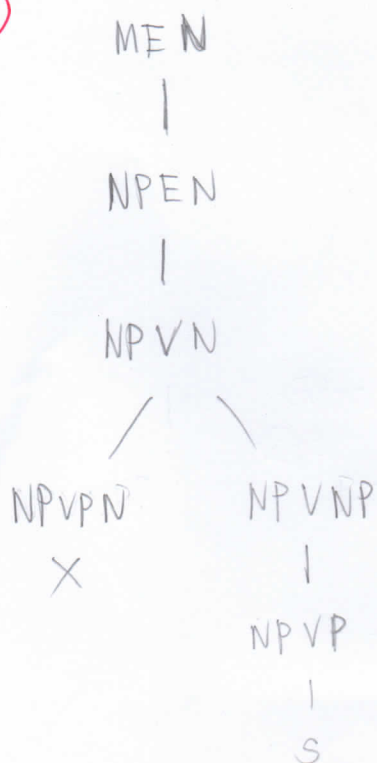
✓ Word employed: E

$\langle cat \rangle = V$.

MediCenter employed nurses \Rightarrow MEN

(1) bottom-up parsing

(2) Top-down parsing



4. Specify (1) the fundamental rule and (2) the bottom-up rule using Chart. (10 points)

(1) fundamental rule :

$(i, j, A \rightarrow w_i \cdot B)$

$A \rightarrow w_i \cdot B$



(2)

5. (1) Please write down Freg's principle.

(2) Explain Freg's principle when given a rule $S \rightarrow NP VP$. (10 points)

(1) 一種語法組合的原則

(2) S 這個 sentence 是由 NP 和 VP 組合而成

+10

6. Give four reasons why we should adopt the corpus-based approach to NLP. (10 points)

(1) 簡單

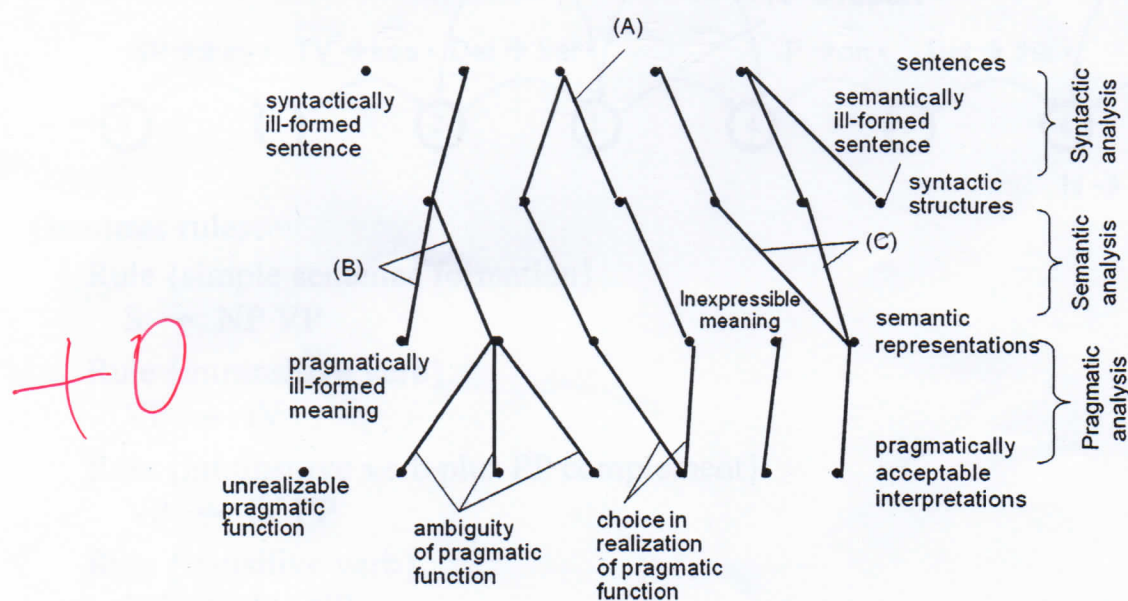
(2) 易讀

(3) 有效率

(4)

+8

7. 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 a brief explanation for each case. (10 points)

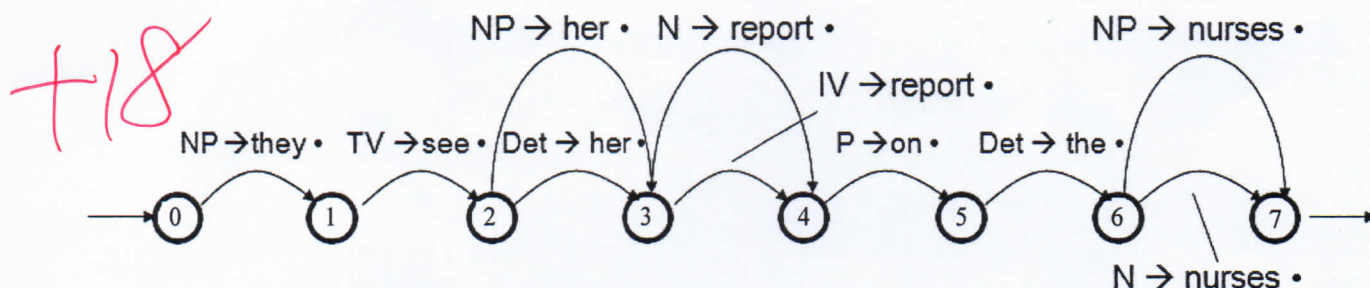


(A) structural ambiguity

(B) word sense ambiguity

(C) synonyms

8. 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)



Grammar rules:

Rule {simple sentence formation}

$S \rightarrow NP VP$

Rule {intransitive verb}

$VP \rightarrow IV$

Rule {intransitive verb plus PP complement}

$VP \rightarrow IV PP$

Rule {transitive verb}

$VP \rightarrow TV NP$

Rule {transitive verb plus PP complement}

$VP \rightarrow TV NP PP$

Rule {transitive verb plus VP complement}

$VP \rightarrow TV NP VP$

Rule {simple noun phrase}

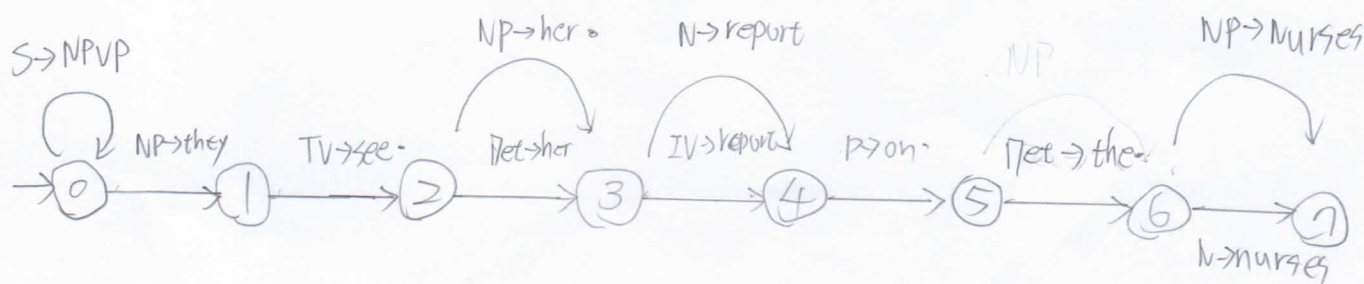
$NP \rightarrow Det N$

Rule {noun phrase with PP complement}

$NP \rightarrow Det N PP$

Rule {simple prepositional phrase}

$PP \rightarrow P NP$



they see her report on the nurses

9. 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)

先找出句子中的 who, when, what, how, why 等疑問詞，
再去分析剩下的文句來解析問題的內容，之後

+5 從 Database 中找出等應這些問題的 Answer 作回應