Natural Language Processing
Midterm Examination

Name: _____ Student ID：_____

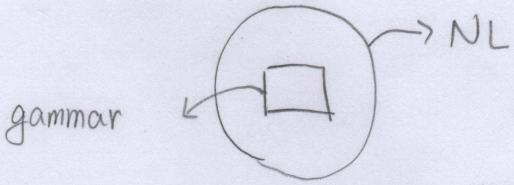
Date: April 20, 2017 Time: 14:20-16:20

98

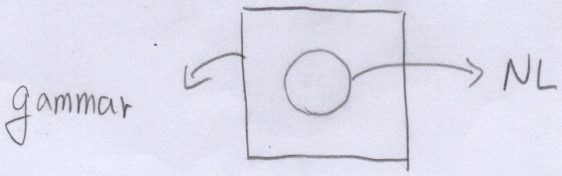
1. Give two issues for evaluating a particular grammar for a language. You should explain for your issues. (10 points)

+10

1. undergenerate : parser 所產生的 sentence, gammar 符合 NL, 可被理解



2. over generate : parser 所產生的 sentence, gammer 無法完全被 NL 解構, 可能會部份不理解



2. Consider the semantic representation of the sentence "Every finance holdings took over a bank". 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 S → NP VP : <S sem predicate> = <VP sem> <S sem arg0> = <NP sem>. Rule VP → TV NP : <VP sem> = <NP sem> <NP hole> = <TV sem> <TV arg0> = <VP arg0> <TV arg1> = <NP referent> Rule NP → Det N: <NP sem quantifier> = <Det sem quantifier> <NP sem variable> = <NP referent> <NP sem restriction> = <N sem> <NP sem body> = <NP hole> <NP referent> = <N referent>	Word "took over": <cat> = TV <sem predicate> = took_over <sem arg0> = <arg0> <sem arg1> = <arg1>. Word bank: <cat> = N <sem predicate> = bank <sem arg0> = <referent>. Word "finance holdings": <cat> = N <sem predicate> = finance_holdings <sem arg0> = <referent>. Word an: <cat> = Det <sem quantifier> = exists. Word every: <cat> = Det <sem quantifier> = all.

The semantic representation for the sentence: Every finance holdings took over a bank.

sem : quantifier : (A) all

variable : (B) X

restriction : arg0 : (C) X

predicate : ~~e~~ finance_holdings

body : quantifier : (D) exists

variable : (E) Y

restriction : arg0 : (F) Y

predicate : (G) bank

body : arg0 : (H) X

arg1 : (I) Y

predicate : (J) took over

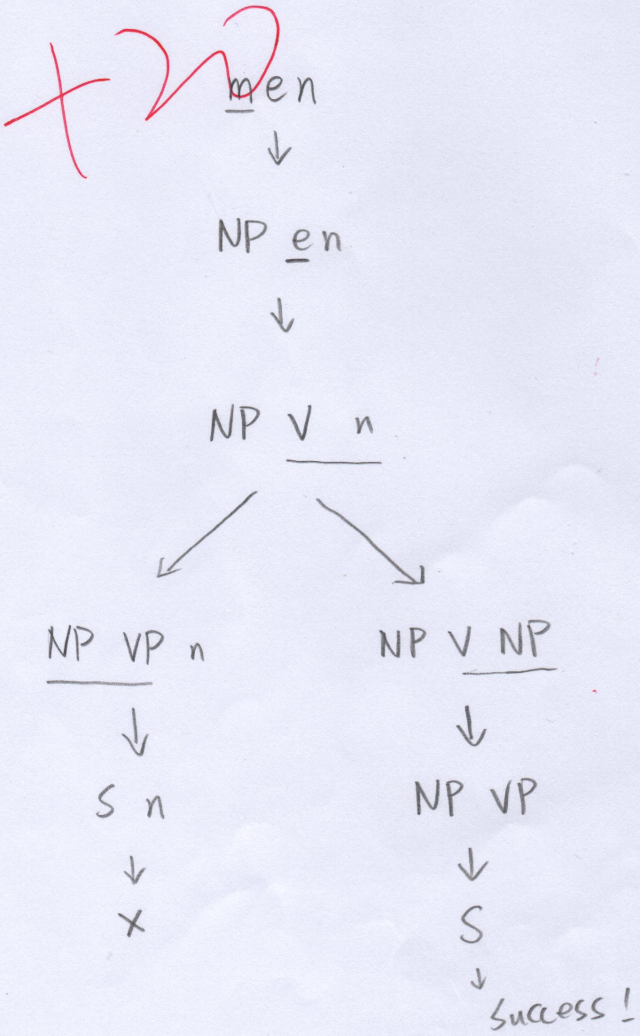
+11

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→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.

(1) bottom-up parsing.

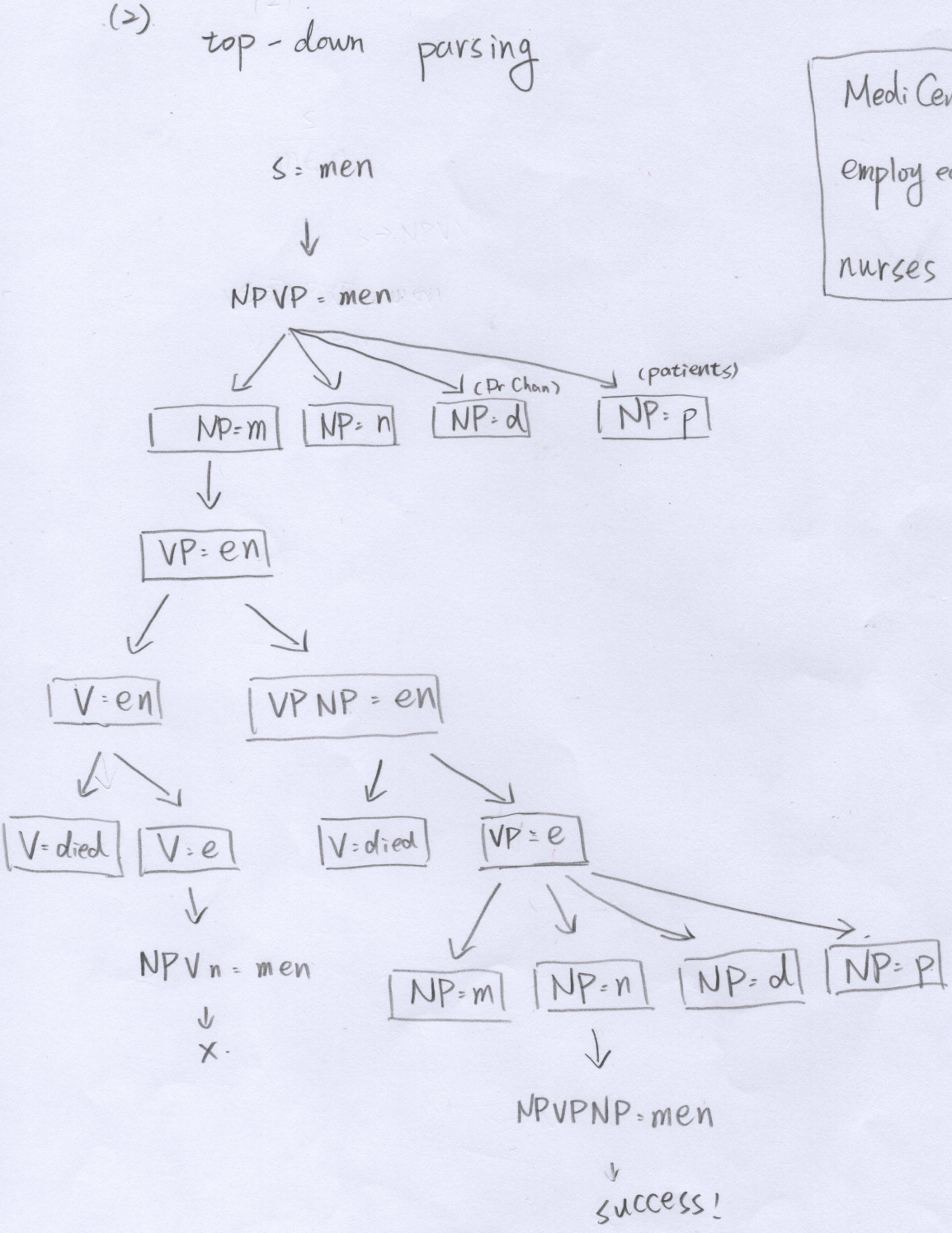
+22



MediCenter = m
employed = e
nurses = n

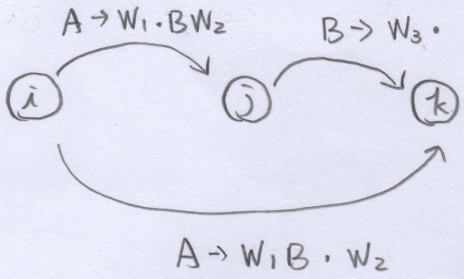
(2) top-down parsing

Medi Center	= m
employed	= e
nurses	= n



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

(1) the fundamental rule



+10 If you are adding $\langle i, j, A \to W_1 \cdot B W_2 \rangle$ and $\langle j, k, B \to W_3 \cdot \rangle$, then you can add $\langle i, k, A \to W_1 B \cdot W_2 \rangle$

A, B are category, W_1, W_2, W_3 are words

* 被增加的 edge 可以是 active or inactive

fundamental rule 只會增加 edge, 不會刪除任一 edge

(2) bottom-up rule

If you are adding $\langle i, j, C \to W_1 \rangle$ and $\langle j, j, B \to W_3 \cdot \rangle$, then you can add $\langle i, i, C \to \cdot W_1 B \rangle$

* 如果 Chart 有很多 inactive edge, 則可增加自己 node 的 active edge

5. (1) What are three major sources of structure ambiguity? Give the example for each source to show the ambiguity.
(2) Given the examples for lexical ambiguity. (20 points)

(1) PP - attachment - PP可修飾的句子太多，造成parser很難判斷語法結構的ambiguity

coordination - 因為 and, or ... 等的連接詞順序造成的ambiguity

Ex. A and B or C → (A and B) or C
↘ A and (B or C)

noun-noun - 因為多個名詞連接造成的ambiguity

Ex. the dog food

→ the dog food 狗的食物

→ the dog food 狗食.

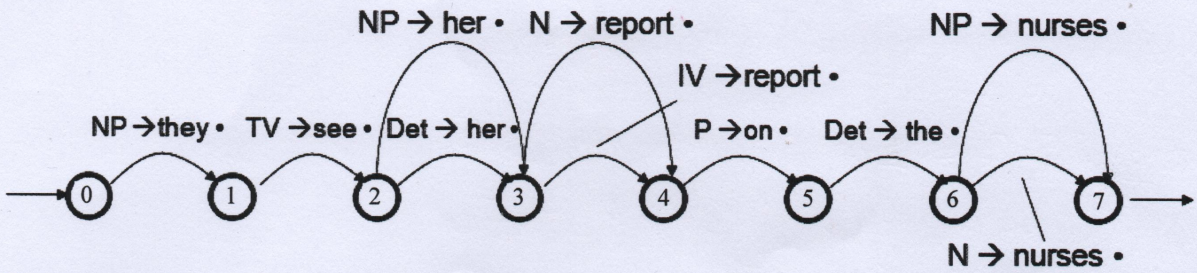
+20

(2) lexical ambiguity

因為同音不同義 / 一字多義造成的ambiguity

Ex. bank → n. 銀行.
↘ V. 儲存

6. 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 → 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

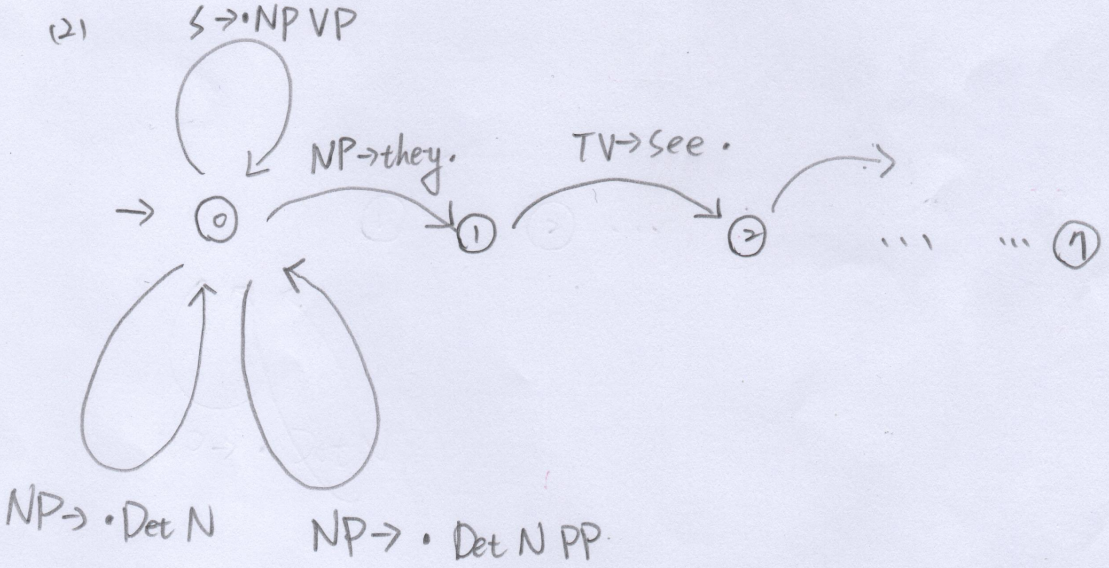
+20

(1) top-down strategy rule.

(1) Initially. S → • NP VP

(2) 根據 category and rule 增加 active edge.
於自己的 node

(2)



7. Give the syntactic and semantic information of a transitive verb "took over" in the lexicon as follows. Please write the corresponding stack structure of the word. (10 points)

Word "took over":
<cat> = TV → transitive
<sem predicate> = took_over
<sem arg0> = <arg0>
<sem arg1> = <arg1>.

+2

Sentence : <arg 0> took_over <arg 1>.
<cat = TV>.
Ex. the airplain took over every hotle chain.
<cat> = category
<sem> = semantics.

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)

提供一個 Corpus-based 的 NLP 在系統.
讓使用者在網頁上 input box 寫入問題.
傳到後端 NLP 系統解析後回傳.
合適的對話句子給 user.

+5