

Natural Language Processing

ECS763P

Sample Solutions for Coursework (2)
and Review

March 27th 2017

Sample Solutions: Course Work (2)

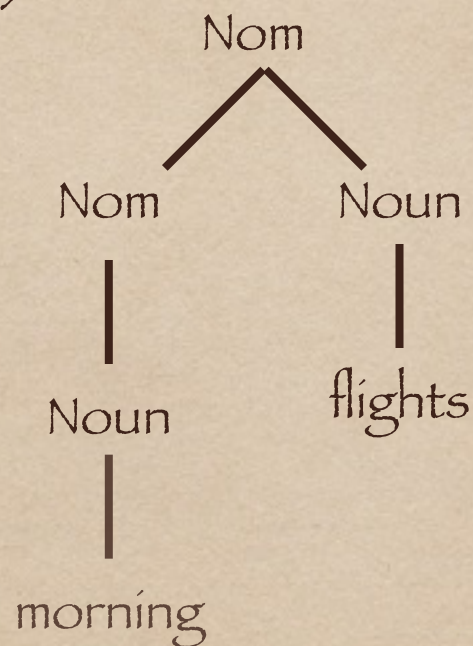
Question (1). Drawing CFG parse trees for 9 expressions, using the rules in the relevant chapters of J&F, listing the corresponding rules.

(e) morning flights

Nom -> Nom Noun | Noun

a nominal consists of one or more nouns

pages 3, ch. 11 (online pdf)



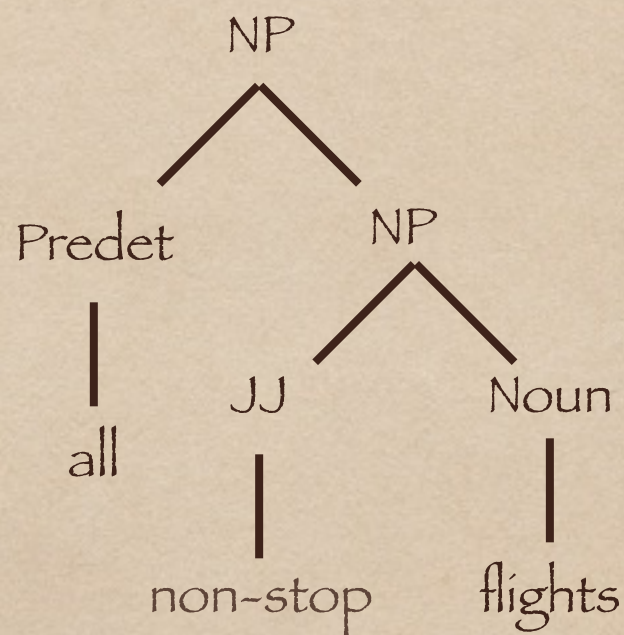
Sample Solutions: Course Work (2)

(f) all non-stop flights

NP -> **Predet NP**, for **Predet** a number, an amount, or a quantifier.

NP -> **JJ Noun**, for **JJ** an adjective

pages 17, Figure 11.10



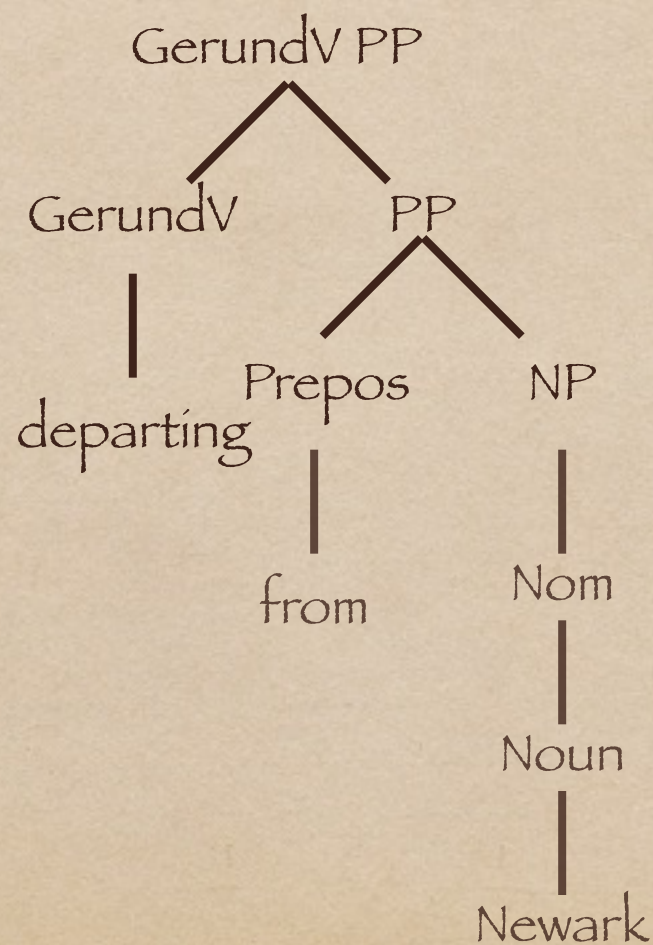
Sample Solutions: Course Work (2)

(d) departing from Newark

Gerund -> GerundV NP | GerundV PP | GerundV | ...

a verb phrase beginning with the gerund (ing) form of the verb

page 11



Sample Solutions: Course Work (2)

(g) arriving by six pm

NP -> CD RB,

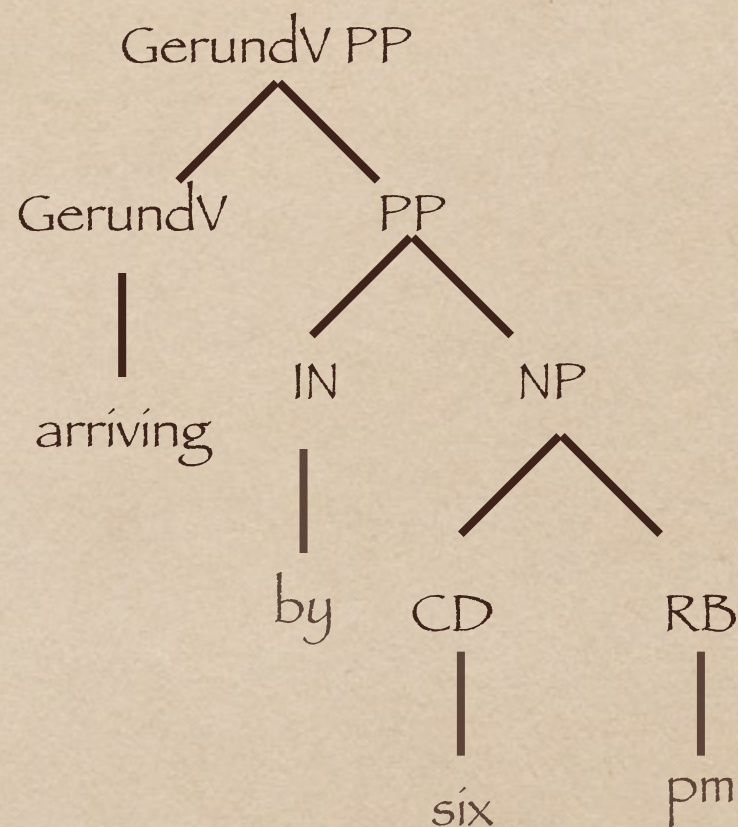
CD -> eleven, six,

RB -> a.m., p.m.

PP -> IN NP,...

IN -> by, before,...

Figure 11.10, page 17



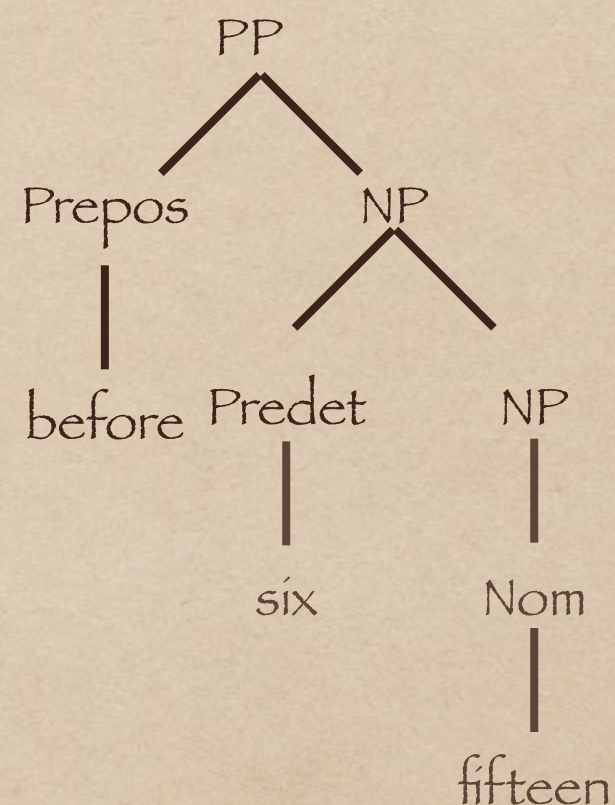
Sample Solutions: Course Work (2)

(c) before six fifteen

NP -> Predet NP,

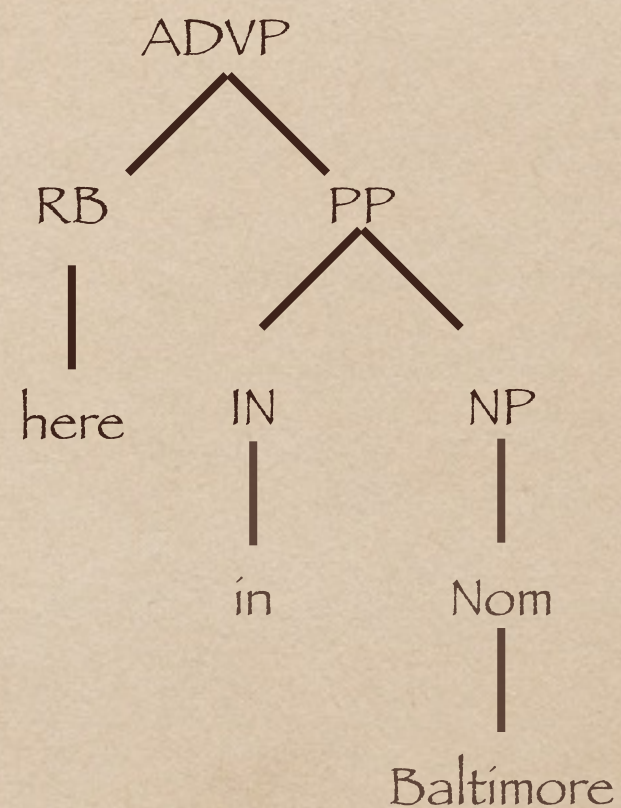
pages 4

... discuss trace: missing am or pm.



Sample Solutions: Course Work (2)

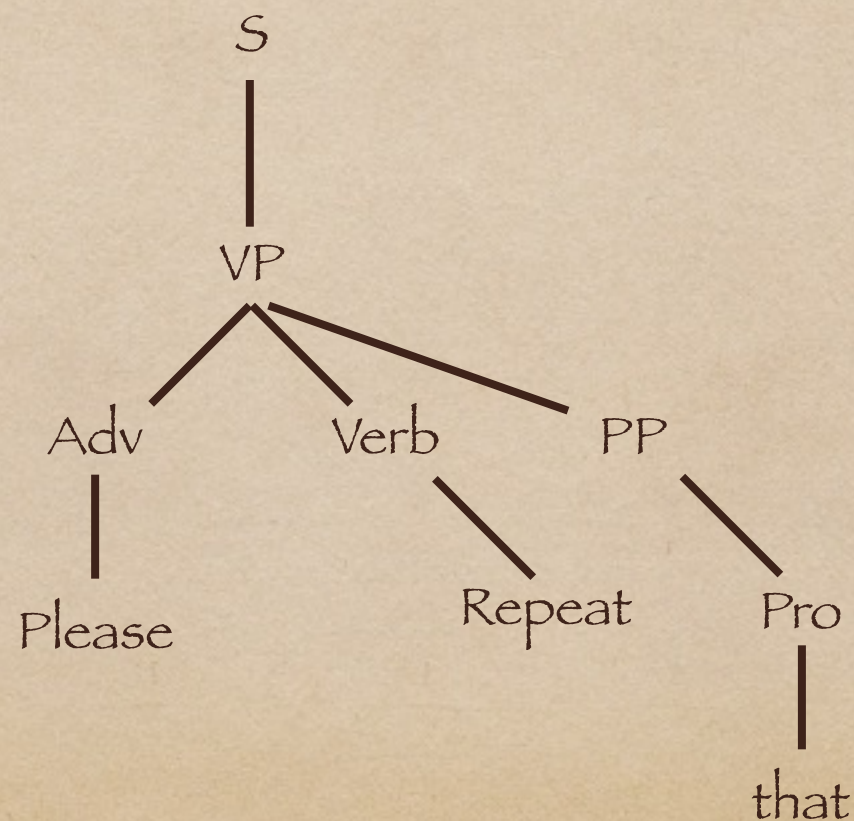
(h) here in Baltimore
from **ADJP** -> **JJ PP** on page 17,
one can derive **ADVP** -> **RB PP**



Sample Solutions: Course Work (2)

Question (2). Drawing CFG parse trees for 5 sentences
(c) Please Repeat that.

S -> VP, VP -> Adv Verb PP, page 18



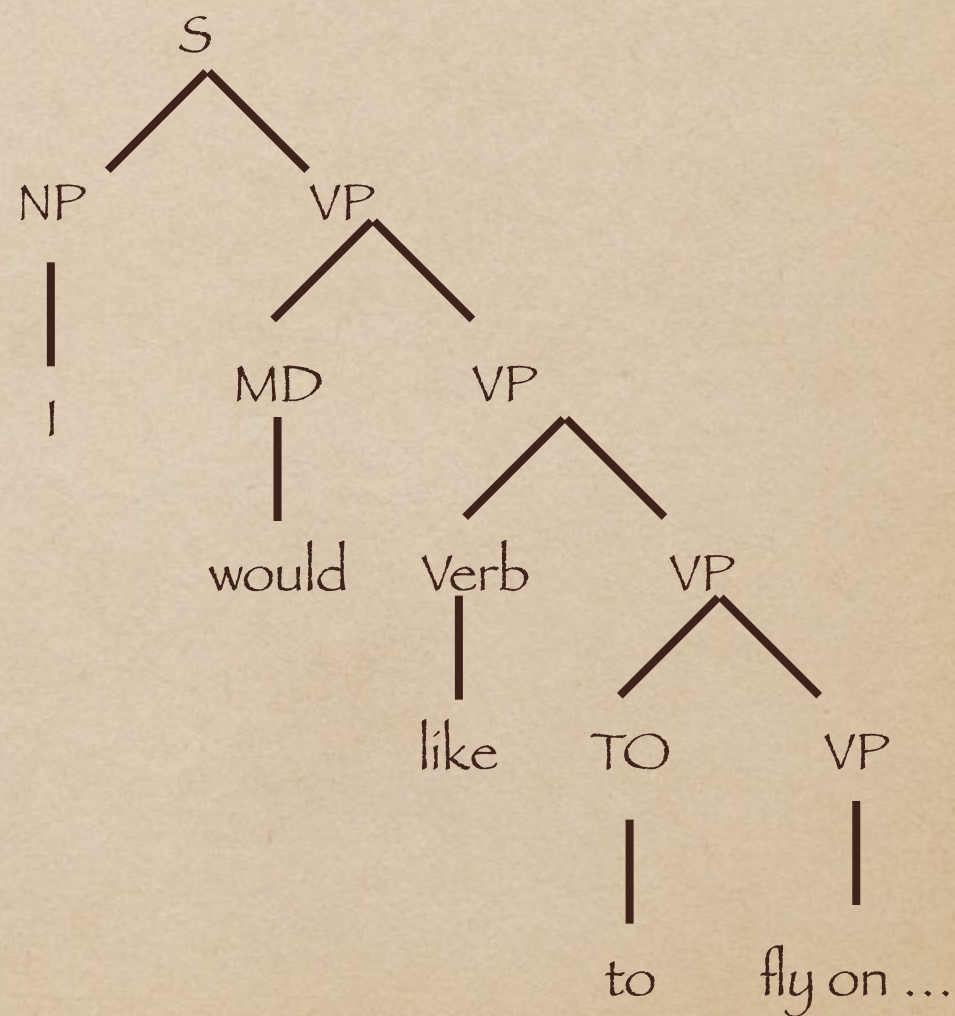
Sample Solutions: Course Work (2)

(b) I would like to fly on American Airlines.

S -> NP VP, VP -> MD VP, VP -> Verb VP

MD -> would, VP -> TO VP, TO -> to

pages 17, Figure 11.10



Sample Solutions: Course Work (2)

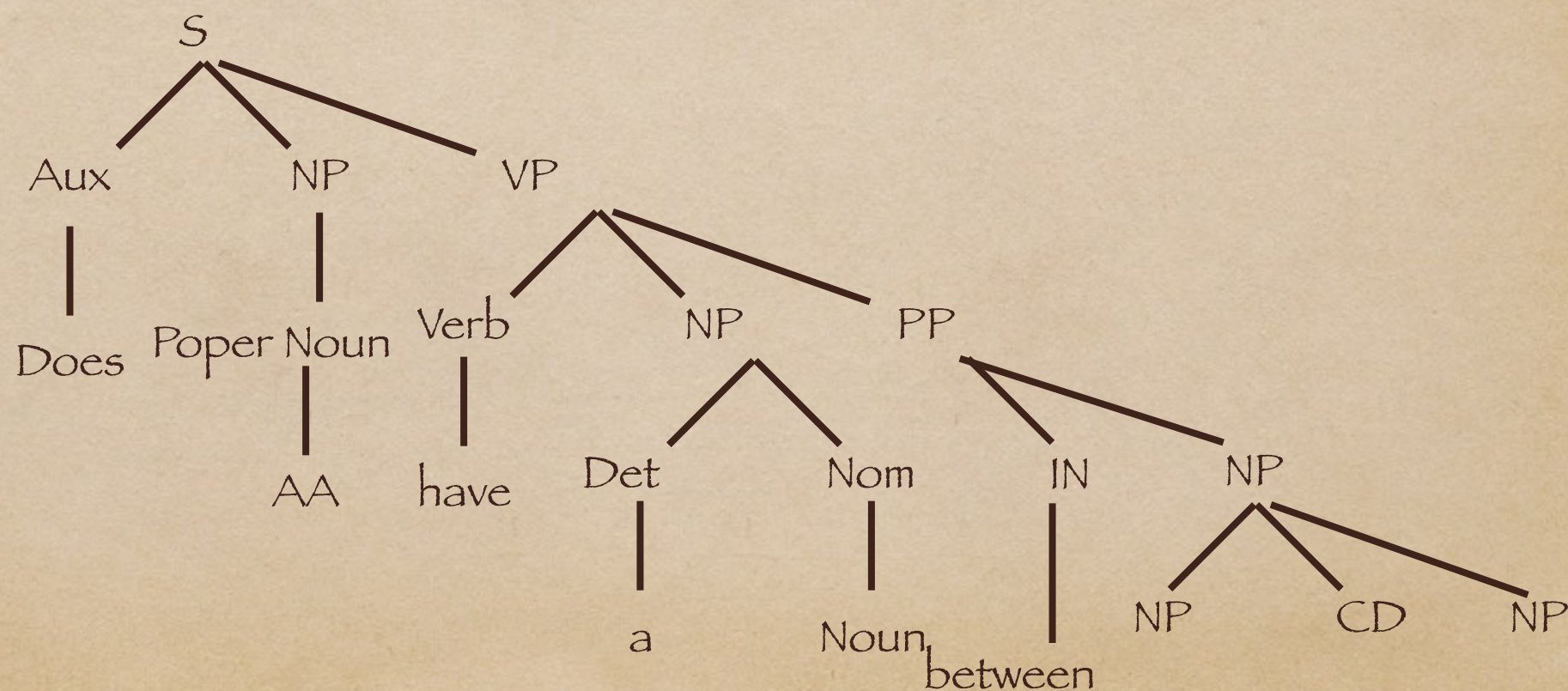
(a) Does American Airlines have a flight between 5 am and 6 am?

S -> Aux NP VP, pages 7,8

Aux is an auxiliary, e.g. does, do, to form yes-no questions

VP -> Verb NP PP, **PP -> IN NP**, **NP -> NP CD NP**

page 17

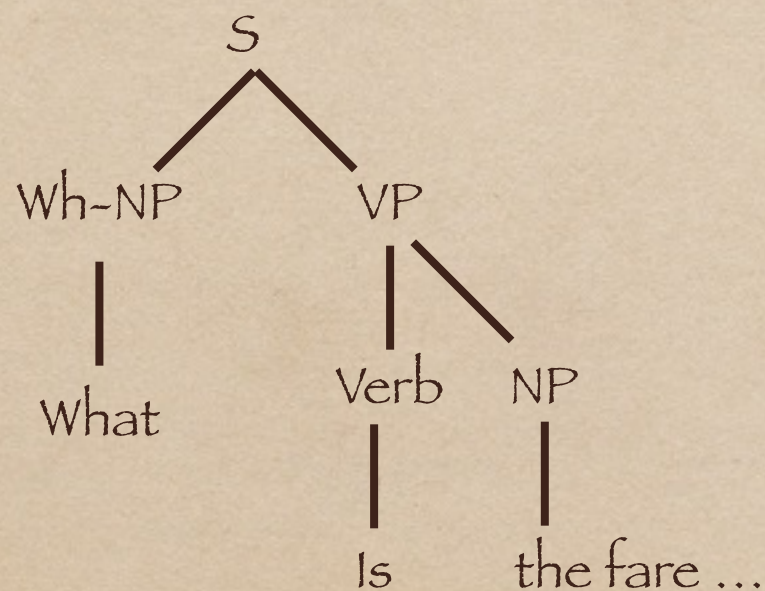


Sample Solutions: Course Work (2)

(e) What is the fare from Atlanta to Denver?

S -> Wh-NP VP, page 7, or **S -> Wh-NP SQ**, **SQ -> Verb NP**

Wh-NP is a wh-word that can appear in subject position, e.g. what, who, which.



Sample Solutions: Course Work (2)

Question (4): 4 different meanings for :

Show me the meals on the flight from SF.

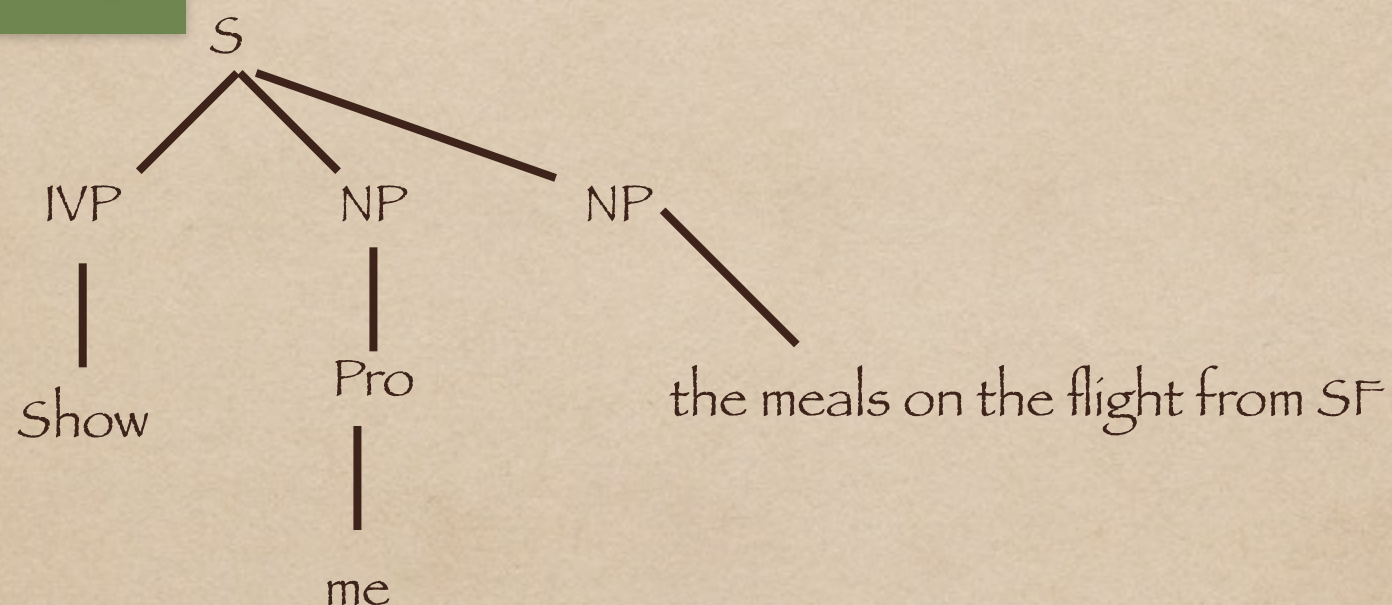
- 1- Show me the meals [**that are served**] on the flight from SF.
- 2- Show me the meals [**when we are**] on the flight from SF.
- 3- Show me the meals on the flight [**when we are in**] SF.
- 4- Show me the meals on the flight [**but only the ones that come from**] SF.

Sample Solutions: Course Work (2)

Question (4): different NLTK-parses of

Show me the meals on the flight from SF.

Parse (1)



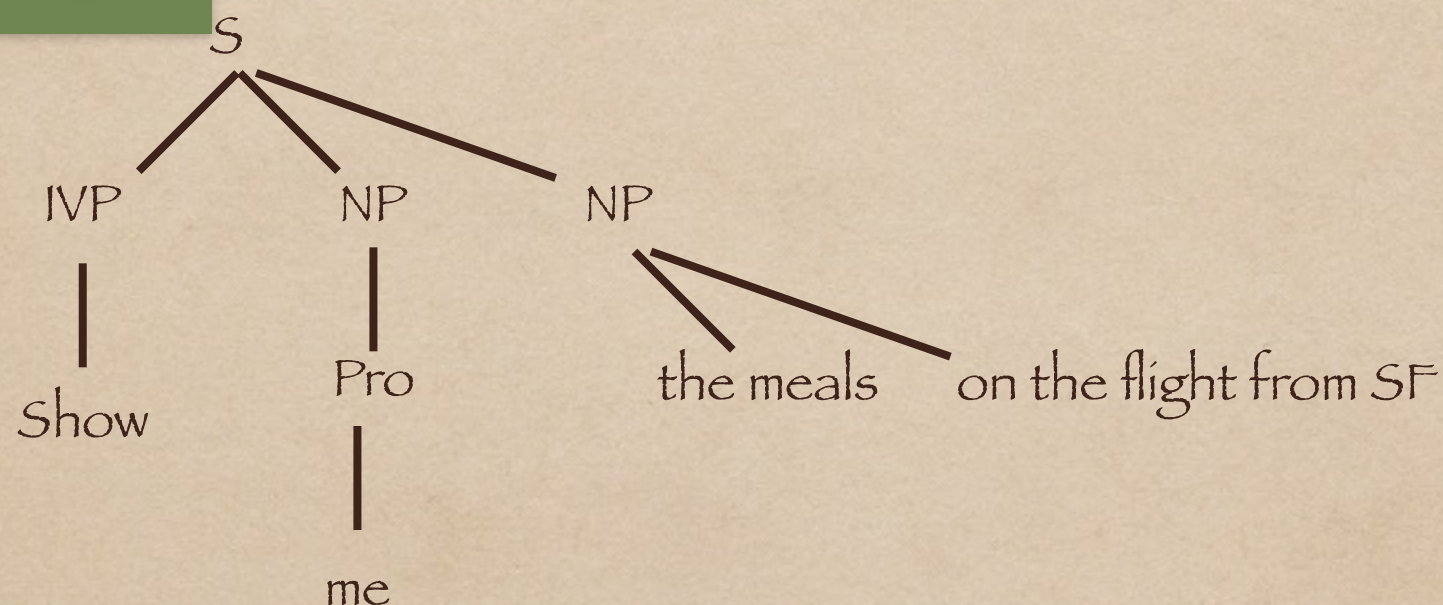
2 possibilities to parse the last NP

Sample Solutions: Course Work (2)

Question (4): different NLTK-parses of

Show me the meals on the flight from SF.

Parse (1-1)



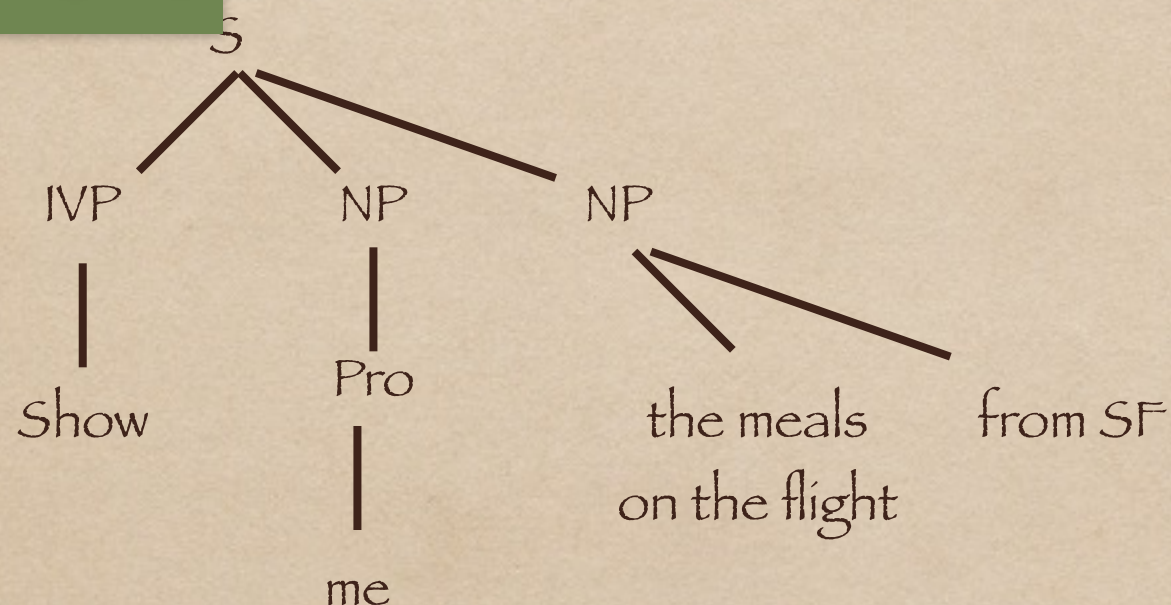
The first one

Sample Solutions: Course Work (2)

Question (4): different NLTK-parses of

Show me the meals on the flight from SF.

Parse (1-2)



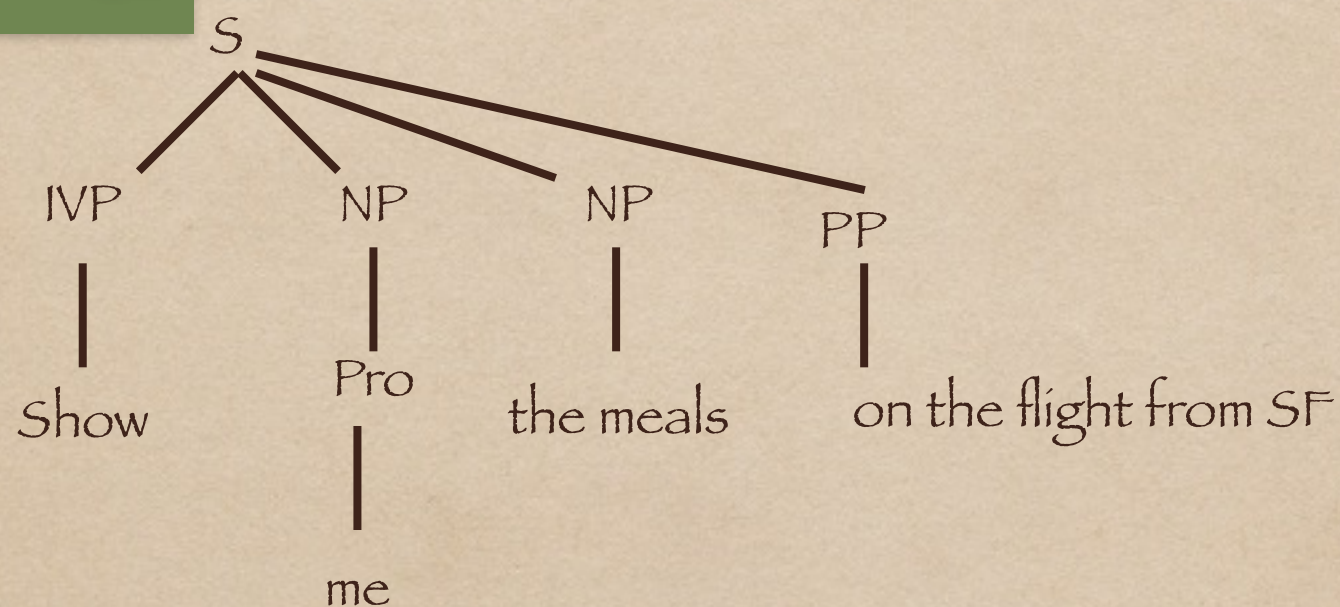
The second one

Sample Solutions: Course Work (2)

Question (4): different NLTK-parses of

Show me the meals on the flight from SF.

Parse (3)

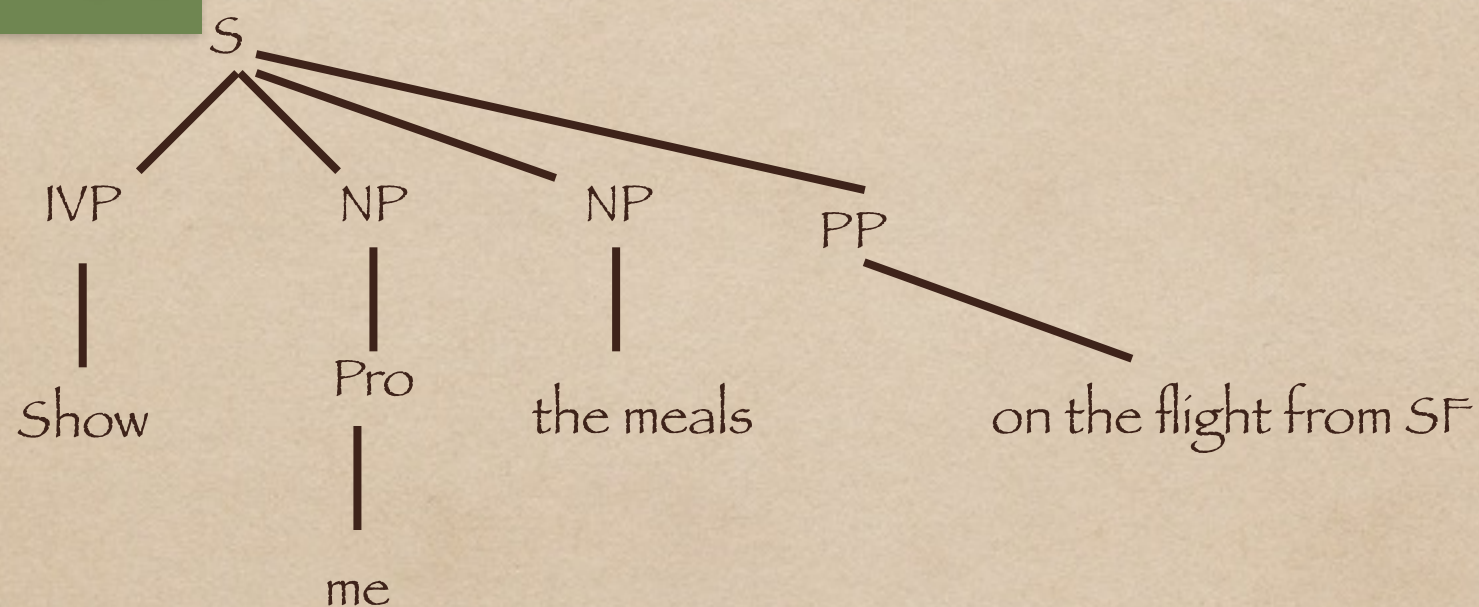


Sample Solutions: Course Work (2)

Question (4): different NLTK-parses of

Show me the meals on the flight from SF.

Parse (4)



Sample Solutions: Course Work (2)

Two different parsing strategies:

These are syntactically the same, but semantically very different. For instance, if you use the Nom PP rule, you are assuming an expression like “on the flights from” is a nominal, which is not correct.

(I)

using the existing rules

PP \rightarrow Prep NP

NP \rightarrow Det Nom

Nom \rightarrow Nom PP

(II)

using the new rule:

NP \rightarrow NP PP

used together with the
existing rule:

NP \rightarrow Det Nom

Sample Solutions: Course Work (2)

Due to the overlap between the existing rule

Nom \rightarrow Nom PP

and the new rule

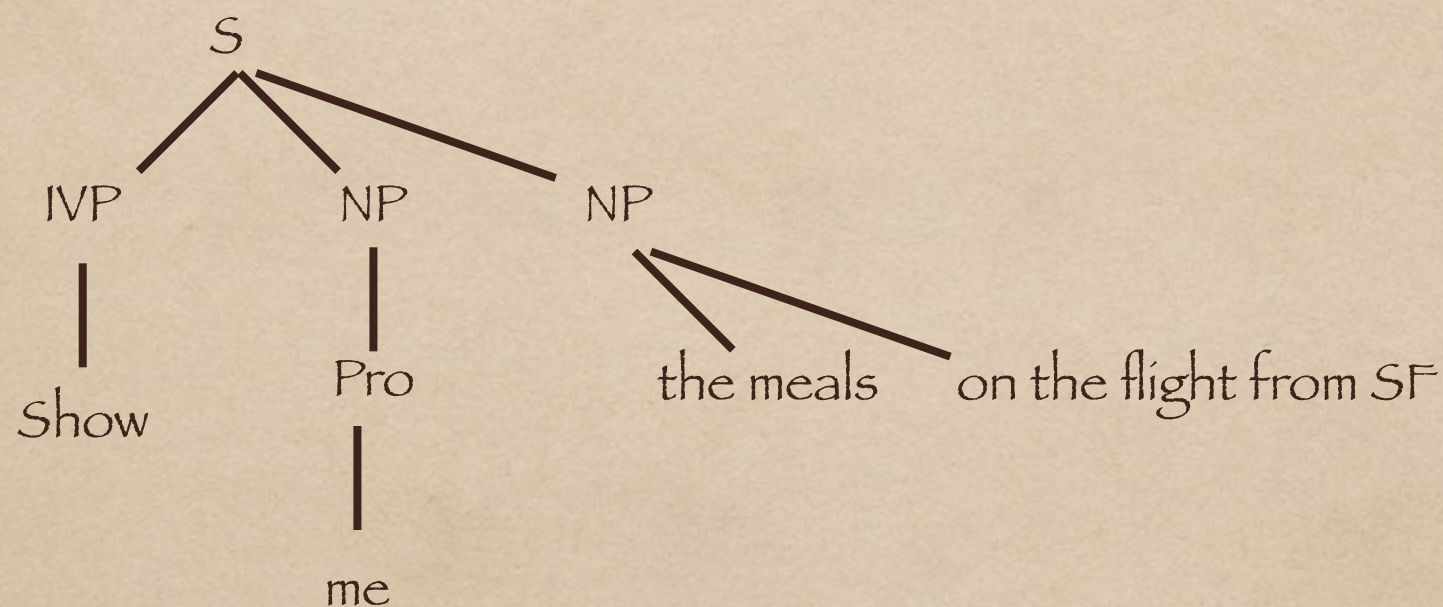
NP \rightarrow NP PP

the NLTK-parser will then generate many more parses than the the four ones we went through here.

Sample Solutions: Course Work (2)

Question (5): learn probabilities for the rules using ATIS corpus.

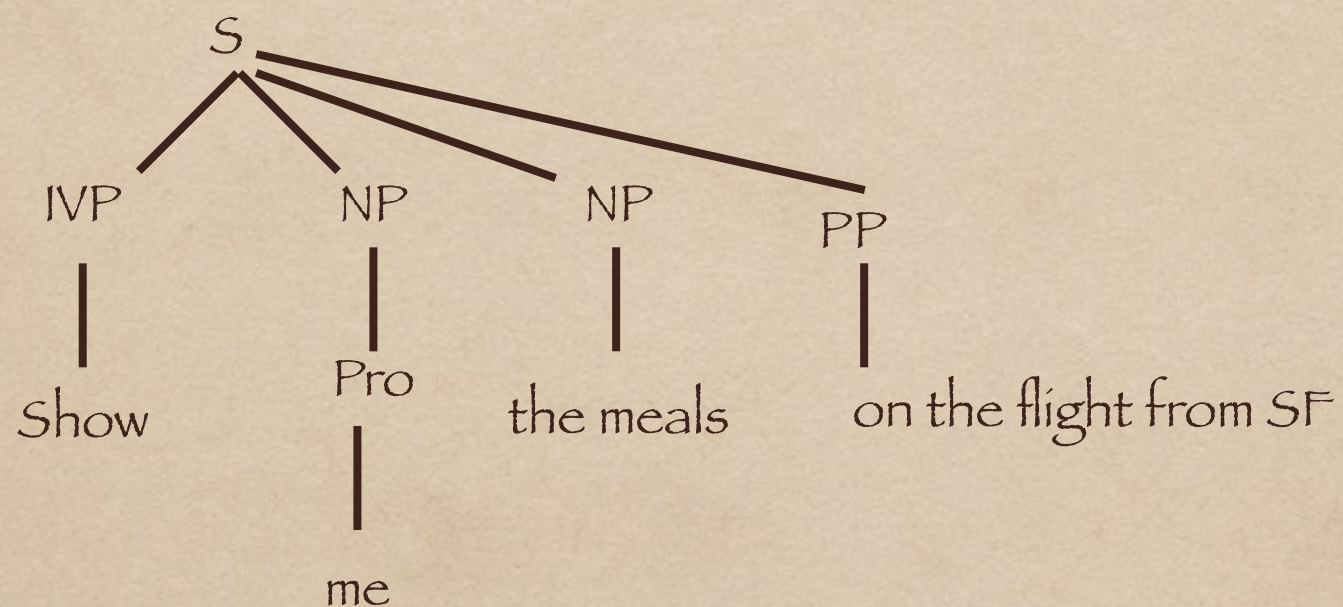
This option will be the most likely.



Sample Solutions: Course Work (2)

Question (5): learn probabilities for the rules

This option will be the 2nd most likely.



General Review of Part (II)

(i) Formal Grammars

- 1- Formal Grammar: historical and etymological architectures
- 2- Different types of Formal Grammars:
 - Context Free Grammars (CFG)
 - Lexical Grammars also known as Categorical Grammars
 - Dependency Grammars
- 3- Formal definition of a CFG
- 4- Definition of a Chomsky-Normal form CFG
- 5- Ambiguity: in meaning and in parsing
- 6- Different types of ambiguity, e.g. pp attachment, coordination
- 7- Parsing algorithms with a CFG

General Review of Part (II)

(ii) Probabilistic Formal Grammars

8- Probabilistic Context Free Grammars (PCFG):

9- Formal Definition of a PCFG

10- How to learn the weights of a PCFG

11- How to apply a PCFG to the challenge of ambiguity

12- Algorithm for parsing with a PCFG

13- Turning a CFG parse into a Dependency parse

General Review of Part (IV)

Distributional Semantics

- 14- Insights and ideas behind the method
- 15- Building a Document-Term and a Word-Word matrix
- 16- Turn any of these matrices to vector representations for documents and words
- 18- Parameters of the model, e.g. window size, normalisation
- 19- Formal definitions of different normalisation schemes such as: conditional probabilities, PMI, PPMI
- 20- Formal definition of cosine similarity
- 21- Applying the above to raw vector representations

The material on neural nets will NOT be in the exam.