Morphological classes, lexical tags, word classes

POS Tagging

Algorithm to assign POS to a word

Challenge : Many words have single tags so we need to find correct tag and make it unambiguous.

We can refer to external words or surrounding words to remove ambiguity.

URBAN EDGE ___/___/___

19/02/2020

(i) Rule based

(ii) Markov model model

Transformation based : combo of (i) and (ii)

Penn Tag Set

Tokenizer = ?, etc

Tokenization must be done compulsorily when question is given.

Tokenization rules in English

(i) Divide at spaces and hyphens.

(ii) Divide before punctuation that is followed by a space or the end of the line.

(iii) Break off the following as separate tokens when followed by a space or end of line.

(iv) Abbreviations

infinitive form -(followed by to)

Important Penn Treebank

(i) Existential 'there' = hint = present at the beginning of the sentence.

(ii) CC, CD, DT

HMM is superior

(iii) JJ, JJR, JJS

(iv) NN, NNS, $

(v) PDT

(vi) POS, PP, PPS

POS Tagging

Challenge : words often have more than one POS

- On my back [NN]
- The back [JJ] door
- Win the voters back [RB]

Open classes : Noun, Adverb, Adjectives, Verbs | 7 tags that are added to [illegible]

URBAN EDGE

- Promised to back[VB] the bill.

Hint: Own logic, Work as per the statement given.

Most tagging can be made ambiguous
→ because

The input to a tagging algorithm is a string of words and a specified

Ambiguity for can

Plays well with others
Plays (NNS/VBZ)
well (UH/JJ/NN/RB)
with (IN)

This is a nice day = PP$
This day is nice = DT
You can go this far = RB

Tagsets distinguish between

Performance
Baseline model
- Tagging unambiguous words with the correct label
- Tagging ambiguous words with their most frequent label
- Tagging unknown words as a noun.

Does/VBZ that/DT flight/NN serve/VB dinner/NN ?

Two sources of information
- lexical information
- Contextual information

Defining noun
- determinant (a, an, the)
- plural
- possessive

Approaches
- Rule based,
- ~~Probabilistic~~ Stochastic = using probability — HMM / Transformation

Verb are not getting added
↓ as exception
Auxillary verb

This is a book → verb actually but because of 'a' ~~before book, it becomes~~ noun

Mass noun : Stars, salt
Count noun : two goats, two people

Mass nouns can also appear without articles where singular count nouns cannot.
- Snow is white but not Goat is white

22/02/2020

Noun (~~Plurals, possessive~~) → PP$
- ~~Com~~ Common noun
  - Mass noun (without articles) ↓ Ex: Clouds, Stars
  - Count noun
- Proper noun (with articles) (capitalized)

Any locative places is a adverb

Tag for article - RP
Tag for adverb - RB

each word has a tag

Phrasel verb :

Particle Vs Preposition (IN)

Prepositions - what kind of verb does it go?

Particle - end of sentence, $\boxed{\text{V + Preposition}}$ URBAN EDGE

(RP) class

Conjunctions

Coordinating

Subordinating

Ex: that - I thought that you might like some milk

is a subordinating conjunction.

Pronouns

PP

Personal (you, she, I)

PP$

Possessive (my, your, his, her)

wh-pronoun (who, whom, what, whoever)

Phrasal verb: combination of verb and particle.

Ex: turn down, rule out, step out

NN - singular or mass noun (time, world)

NN$ - possessive singular common noun (father's)

NNS - plural common noun (years, people)

NNS$ - possessive plural noun (parent's)

NNP - singular proper noun (Robert) (Scotland)

NP$ - possessive singular proper noun (Plato's)

NPS - plural proper noun (Eurpidies')

NPS$ - possessive plural proper noun Chinese

NR - adverbial noun (home, yesterday)

NR$ - possessive adverbial noun (today's)

NRS - plural adverbial noun (Sundays, Fridays)

to + verb.

VB - verb (make, understand, try, determine)

VBD - verb, past tense (looked)

VBG - verb, present participle, gerund (increasing) writing

VBN - verb, past participle (made, given, found)

VBZ - verb 3rd singular present (says, follows, requires).

VBP - present tense verb (refuse, permit)

3 preposition = IN ⇒ subordinating conjunction

WDT = wh-determiner (what, which)
WP$ = Possessive wh-pronoun (whose)
WPO = objective wh-pronoun (whom, which)
WRB = wh-adverb (how, when)

RB = adverb
RBR = comparative adverb (later, more, better)
RBT = superlative adverb (best, most)
RN = nomial verb (here, then)

JJ = adjective
JJR = comparative adjective (better, higher)
JJT = superlative adjective (chief, key, principal)

PPS = 3rd. sg. nominative pronoun (he, she, it)
PP$ = possessive personal pronoun (his, their)
PP$$ = second possessive personal pronoun (mine, his, ours, yours)
RP = particle (up, off)
and CC = coordinating conjunction
~~CS = subordinating conjunction.~~
DT = singular determiner (this, that)
IN = preposition
CD = cardinal digit (one's)

2+2+2

Adverb

directional degree manner temporal

Tag for interjection - UH

Auxiliary verb - be, have, do

is, are, was, were, being, been, I

Example of around PENN word vocabulary.

[Particles → up, down, on, off, in, out]
RP

Three Methods of POS Tagging.

1. Rule - based
Hand-coded rules

2. Probabilistic / Stochastic
Sequence (n-gram) models ; machine learning
→ HMM (Hidden Markov Model)

3. Transformation

Rule - Based POS Tagging

"Two level morphology"

Ex : I am in race [NN]
Race is noun as it is preceded by preposition "in".

Ex : She promised to back the bill
PPS (PP) VBD VBN to [VB] [NN] [JJ] [RB] [DT] [NN] [VB]

rules
(1) Here correct is VBD as it is preceded by pronoun.
Assign VBD tag if it is preceded by pronoun

(2) Assign VB tag to back when it is preceded by 'to'
(3) Assign NN tag to bill when it is preceded by "the".

1. Lexical information
- look up all possible POS for a word in dictionary
- "table" : (noun, verb) but not a (adj, prep ...)
- "rose" : (noun, adj, verb) but not (prep, ...)

Adv and Disadvantages of Rule-Based

2. Syntagmatic information
- some tag sequences are more probable than others.
- DET + N occur frequently

Ex : Secretariat [NNP] is [VBZ] expected [VBN] to [TO] race [VB] tomorrow [NN]

Probabilistic POS Tagging

→ People [NNS] continue [VBP] to [TO] inquire [VB] the [DT] reason [NN] for [IN] the [DT] race [NN] for [IN] outer [JJ] space. [NN]

$P(A|B) = P(A) \cdot P(B|A) / P(B)$

- $P(\text{race}|VB)$ $\cdot$ $P(\text{race}|NN)$

N-gram Tagger

- Uses the preceding (N-1) predicted tags
- Also uses the unigram estimate for current word

~~words~~

Tokens :	w_{n-2}	w_{n-1}	w_n	w_{n+1}
	|	|	|	|
Tags :	t_{n-2}	t_{n-1}	t_n	t_{n+1}

$\frac{C(t_{n-1}, t_n)}{C(t_{n-1})} = P(t_n | t_{n-1})$ ⇒ Tag transformation probability (prior probability)

$P(w_i | t_i)$ ⇒ Observation likelihood probability

$\hat{t}_1$ ⇒ best estimate

$$\boxed{\hat{t}_1^n = \arg\max_{t_1^n} P(t_1^n | w_1^n)}$$

Probability Questions

T^N

$p(w_i|x) =$ $P(A|B) = P($

$P(A|B) = \frac{P(A)\,P(B|A)}{P(B)}$.

$P(w_1^n \cap t_1^n)$

$$P(A|B) = \frac{P(A)\cdot P(B|A)}{P(B)}$$

$$P(t_1^n|w_1^n) = \frac{P(w_1^n|t_1^n)\,P(t_1^n)}{P(w_1^n)}$$

$$P(t_1 \ldots t_n) = P(t_1)\,P(t_2|t_1) \ldots P(t_n|t_{n-1})$$
$$= \prod_{i=1}^{tn} P(t_i|t_{i-1})$$

$$P(w_1^n|t_1^n) = \prod_{j=1}^{n} P(w_1|t_1)$$

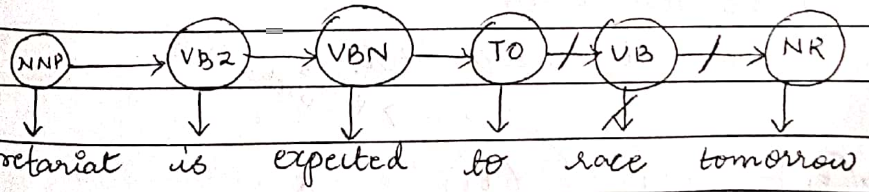
Simplification : Drop the denominator
- Denominator is same for all tag sequences
- Also $P(w_{1,n})$ does not change for each tag sequence

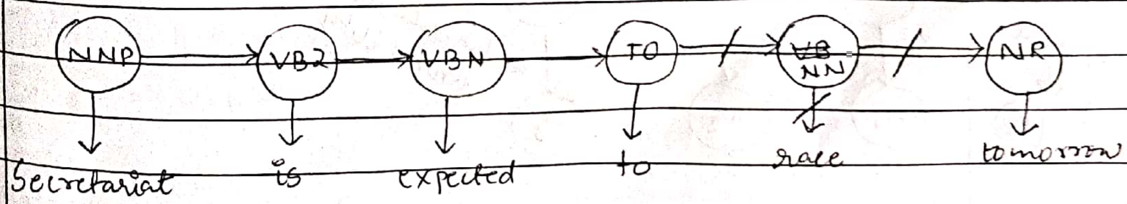
$$P(t_i|t_{i-1}) = \frac{C(t_{i-1}, t_i)}{C(t_{i-1})}$$

$$P(w_i|t_i) = \frac{C(t_i, w_i)}{C(t_i)}$$

$$P(ball|NN) = \frac{C(NN, ball)}{C(NN)}$$

HMM Part of Speech Tagging

(a)


(b)


Procedure:

P(race | VB) => How many times race was tagged as VB ?

(a) $P(VB|TO) * P(NR|VB) * P(race|VB)$ ⇐ most probable

$0.83 * 0.00021 * 0.00012$ $= 2.6 \times 10^{-7}$

(b) $P(\cancel{VB}\ NN|TO) * P(race|NN) * P(NR|\cancel{VB}\ NN)$

$0.00047 * 0.00057 * 0.0012$ $= 3.2 * 10^{-10}$

Hidden Markov Model (HMM)

~~FSA → WFSA → Markov chain → HMM~~

~~↳ outside words~~

Inside tags

HMM

Set of all states

set of words (squares)

start probability (default value = 1)

Two types of probability

- Prior probability (between (states) tags)
- observation likelihood probability (between squares and stages)

Three basic ~~properties~~ problems for HMMs

(i) Likelihood of the i/p

Forward algo - How likely the sentence "I love cat" occurs

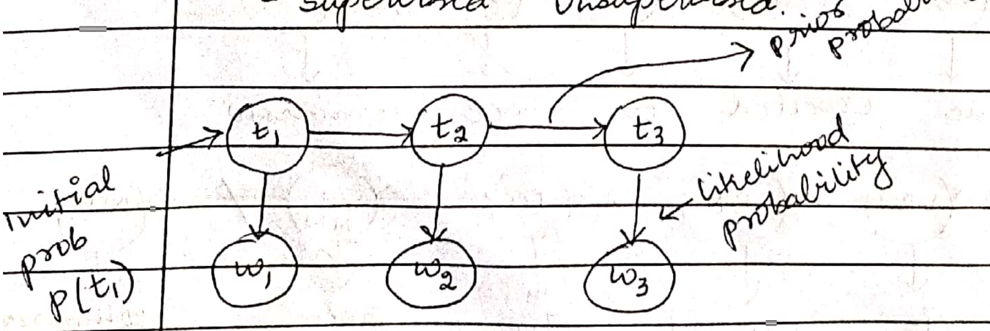
(ii) Decoding the i/p

Viterbi algo - POS tags of "I love cat" occurs

(iii) Estimation (learning):

Find the best model

- Supervised - Unsupervised.



Prediction in generative model

URBAN EDGE __/__/__

Inference - what is most likely tags for given sequence of words w.

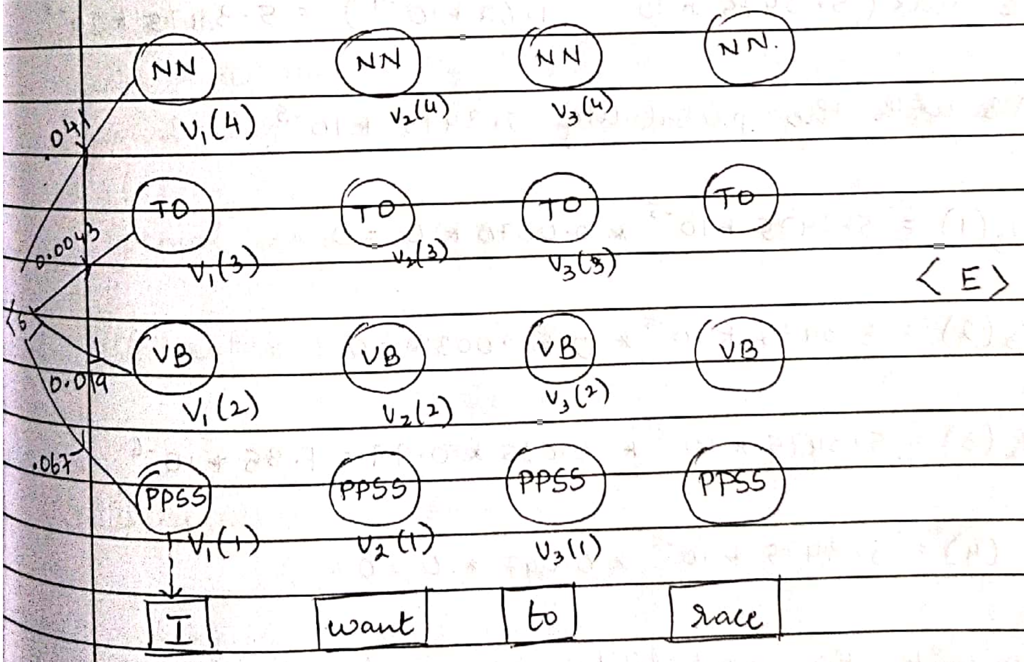
$P(w_i|t_i)$ ⇐ emission probability.

Decoding ⇒ ~~Viterbi Algorithm~~

For any model like HMM that contains hidden variables, the task of determining which sequence of variables is the underlying source of some sequence of observations.

Input : A single HMM and set of observed words

Returns : most probable state | tag sequence $Q = (q_1, q_2 \ldots q_T)$ together with its probability.



$V_1(1) = P(\text{start}) \cdot P(PPSS|\text{start})\ P(I|PPSS)$

$= 1 * 0.067 * 0.37$

$= 0.025$

PPSS, VB, TO,

Probability tables for ~~Pe~~ Tags and Emission will be given

URBAN EDGE

$V_1(2) = 0.019 * 0 = 0$

$V_1(3) = 0.0043 * 0 = 0$

$V_1(4) = 0$

I will have to get its tag PPSS with prob 0.025

→ Take max of previous stage

$V_2(1) = 0.025 * 0.00014 * 0 = 0$

$V_2(2) = 0.025 * 0.23 * 0.0093 = 5.3475 * 10^{-5}$

$V_2(3) = 0.025 * 0.00079 * 0 = 0$

$V_2(4) = 0.025 * 0.000054 * 0.0012 = 1.62 * 10^{-9}$

$\max(5.3475 * 10^{-5}, 1.62 * 10^{-9}) = 5.3475 * 10^{-5}$

VB with the probability $5.3475 * 10^{-5}$

$V_3(1) = 5.3475 * 10^{-5} * 0.0070 * 0 = 0$

$V_3(2) = 5.3475 * 10^{-5} * 0 * .0038 = 0$

$V_3(3) = 5.3475 * 10^{-5} * 0.035 * 0.99 = 1.85 * 10^{-6}$

$V_3(4) = 5.3475 * 10^{-5} * 0.047 * 0 = 0$

TO with the probability $1.85 * 10^{-6}$

$V_4(1) = 1.85 * 10^{-6} * 0 * 0 = 0$

$V_4(2) = 1.85 * 10^{-6} * 0.83 * 0.00012 = 1.8426 * 10^{-10}$

$V_4(3) = 1.85 * 10^{-6} * 0 * 0.99 = 0$

$V_4(4) = 1.85 * 10^{-6} * 0.00047 * 0.00057 = 4.95615 * 10^{-6}$

$\max = 1.8426 * 10^{-10}$

URBAN EDGE ___/___/___

Result :

PPSS VB TO VB
I want to race

Backtrack diagram

$$\begin{bmatrix} 0 & 1.62 \times 10^{-9} & 0 & 4.95 \times 10^{-13} \\ 0 & 0 & 1.85 \times 10^{-6} & 0 \\ 0 & 5.34 \times 10^{-5} & 0 & 1.84 \times 10^{-10} \\ 0.025 & 0 & 0 & 0 \end{bmatrix}.$$

Transformation based Tagger

- Brill tagging
- an instance of transformation based learning (TBL) approach to ML
- Draws inspiration from both rule based and stochastic taggers.
- Like rule based tagging, based on rules that specify what tags would be assigned to what words.

Formal grammars of English

(i) Declarative → some fact
→ There is class today
S → NP VP

(ii) Imperative → command
S → VP → verb phrase
start symbol

(iii) Interrogative → Auxilliary verb
→ Questions

~~S → Aux NP PP~~
~~→ Usually starts~~
with do, does

(iv) Wh question ~~cons~~ structures

CFG example

S → NP VP
NP → DET Nominal
Nominal → Noun
VP → verb

Det → a
Noun → flight
verb → left

Grammar rules

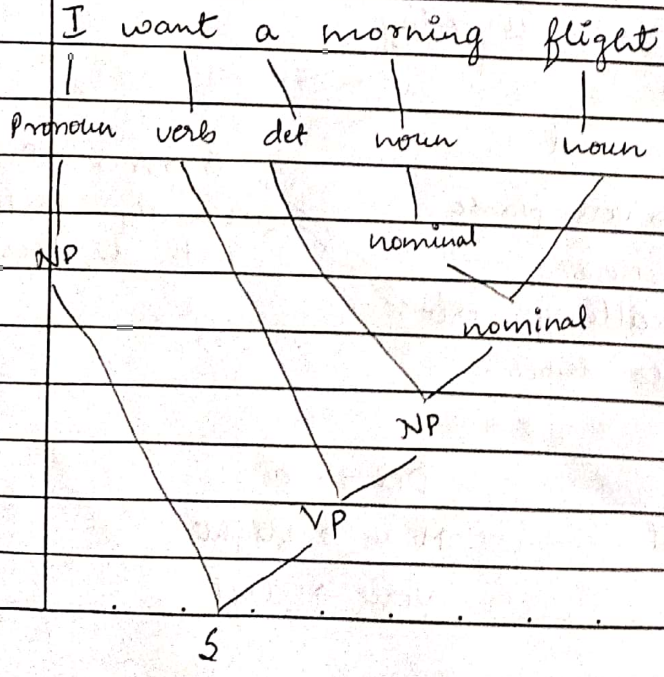
	Example
S → NP VP	I + want a morning flight
NP → Pronoun	I
| proper - noun	Los Angeles
| det nominal	a + flight
Nominal → Nominal Noun	morning + flights
| Noun	flights
VP → verb	do
| verb NP	want + a flight
| verb NP PP	leave + Boston + in morn
| verb PP	leaving + on thursday
PP → Preposition NP	from + Los Angeles

Top down
bottom up

Wh-question structure
- wh subject → S → Wh-NP VP
- wh - non subject → S → Wh-NP Aux NP

Ex : Which class I am attending?

Bottom-up



PP ⇒ Any preposition followed by nomial

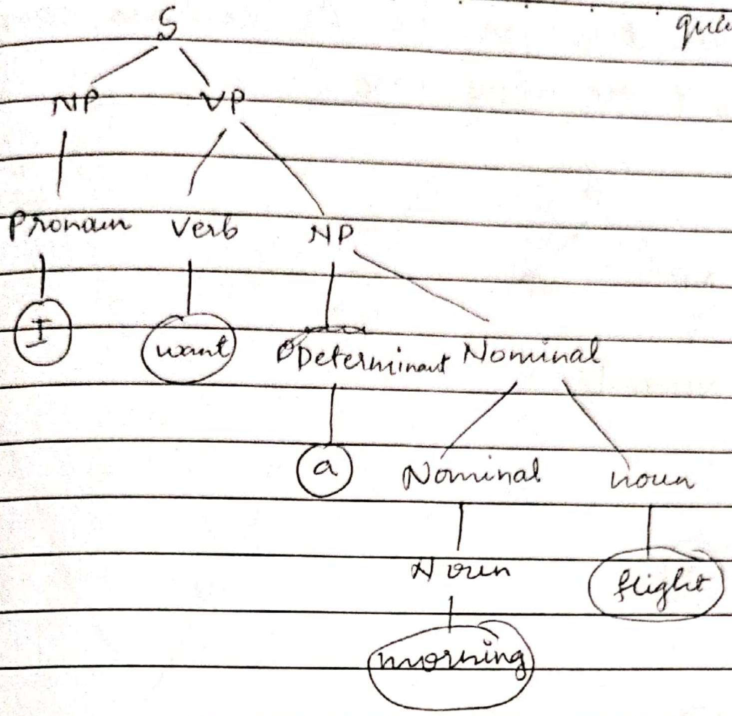
Auxiliary verbs are closed classes

Cardinal - one, two

ordinal - last, first

quantifier - more, less

Top down



Ex: What flight do you have from Burbank to Italy?

Wh - NP AUX NP VP

~~Classes~~ Clauses and Sentences

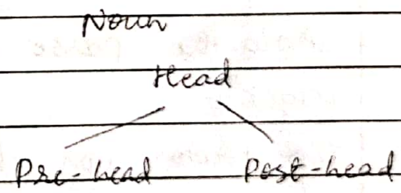
NP's

NP → Pronoun

| Proper - Noun

| Det Noun

| Nominal



Nominal → Noun Noun

* Parse tree for x's y's sentence construction

S → NP VP

NP → Pronoun

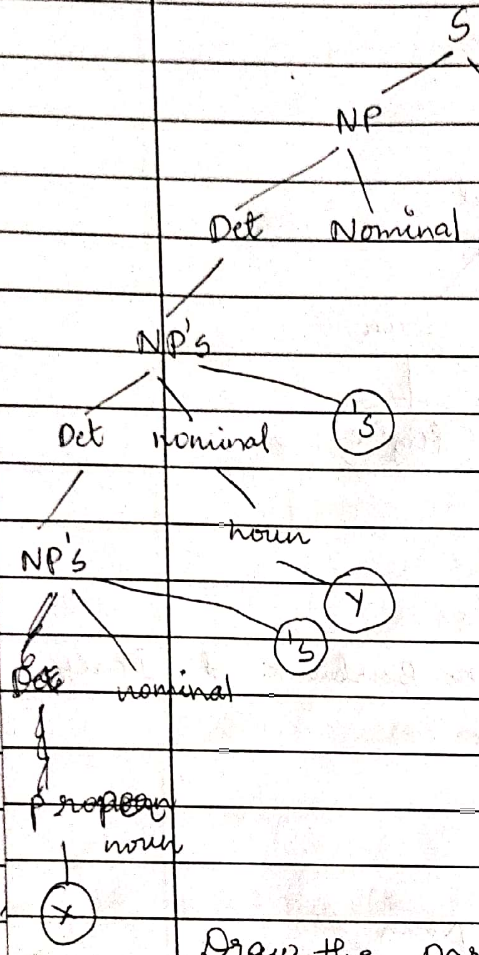
Propnoun - Noun

Det Nominal

Det → NP's

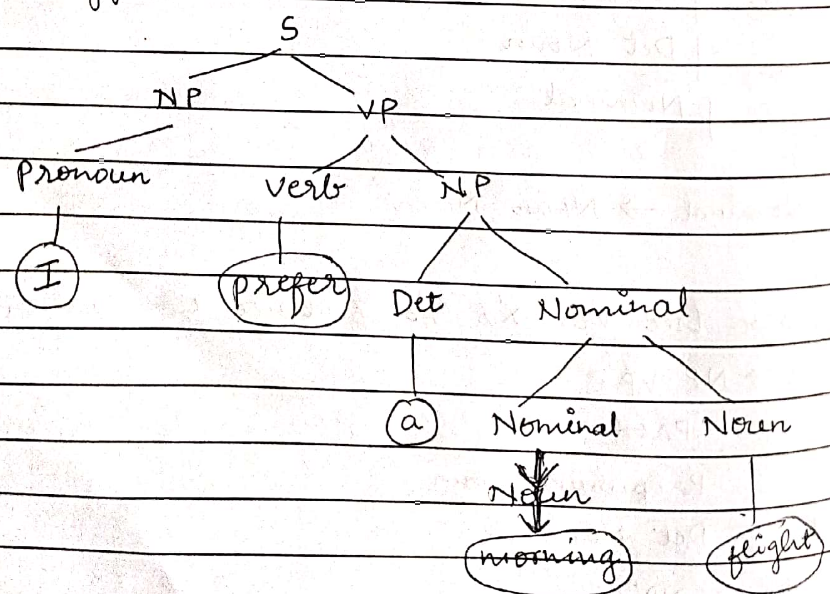
Det → the | a | an | this | these | that

- Draw parse tree for X's Y's sentence construction
- Note X, Y are nouns here

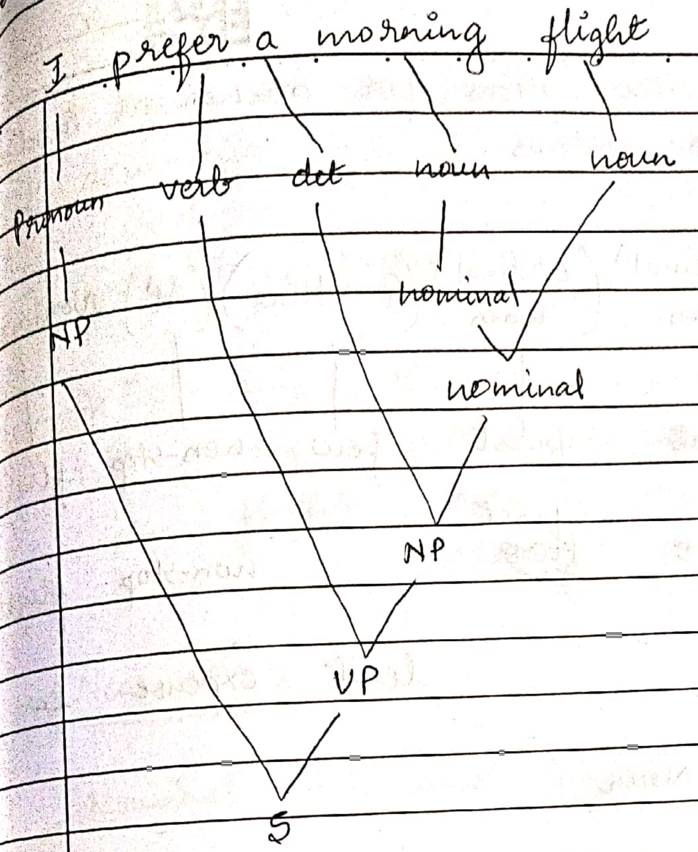


Draw the parse tree for "I prefer a morning flight".

Top-down approach



bottom-up approach



Postdeterminers : Different kinds of word classes can appear before the head noun (the "post determiners") in a nominal. These include cardinal numbers, ordinal numbers and quantifiers.

Ex: Cardinal numbers : two friends, one stop.

Ordinal numbers : the first one, the next day, the second leg, the last flight

Quantifiers : much, little

Adjectives appear after quantifiers but before nouns.

Ex : a first-class fare
a nonstop flight
the longest layover

Adjectives can be grouped into a phrase called an adjective phrase or AP.

We combine all the options for prenomial modifiers with one rule as follows :

NP → (Det) $\begin{pmatrix}\text{Cardinal}\\ \text{noun}\end{pmatrix}$ $\begin{pmatrix}\text{Ordinal}\\ \text{noun}\end{pmatrix}$ (Quantifier) (AP) Nominal

Ex:

The	~~first~~	first	few	non-stop	flights
The	two	~~two~~ first		non-stop	fligers
The			least	expense	fare

<u>After the Head Noun</u>

A head noun can be followed by post modifiers. Three kinds of post modifiers

(i) Prepositional phrases
Ex : All flights [from Cleveland] [boundaries indicate post modifiers]

(ii) non-finite clauses
Ex : any flights [arriving after elevan a.m]

(iii) relative clauses
Ex : a flight [that serves breakfast]

<u>Ex of prepositional phrases</u>

- any stopovers [for Delta seven fifty one]
- all flights [from Cleveland] [to Newark]
- arrival [in San Jose] [before seven p.m]
- a reservation [on flight six oh six] [from Tampa] [to Montreal]

3 types of non-finite postmodifiers

(i) gerundive (-ing) → they consist of verb phrase that begins with (-ing) form of verb

(ii) -ed

(iii) infinitive forms

→ Ex: any of those [leaving on Thursday]

any flights [arriving after 11 a.m.]

flights [arriving within thirty minutes of each other]

New CFG's

Nominal → Nominal Gerund VP

Gerund VP → GerundV NP

| GerundV PP | GerundV | GerundV NP PP

Ex's for infinitives and -ed forms.

- the last flight [to arrive in Boston]
- I need to have dinner [served]

A postnominal relative clause is a clause that often begins with a relative pronoun such as that, who

Ex: a flight [that serves breakfast]

flights [that leave in the morning]

New CFG's

Nominal → Nominal Rel Clause

RelClause → (who | that) VP

Various postnominal modifiers can be combined

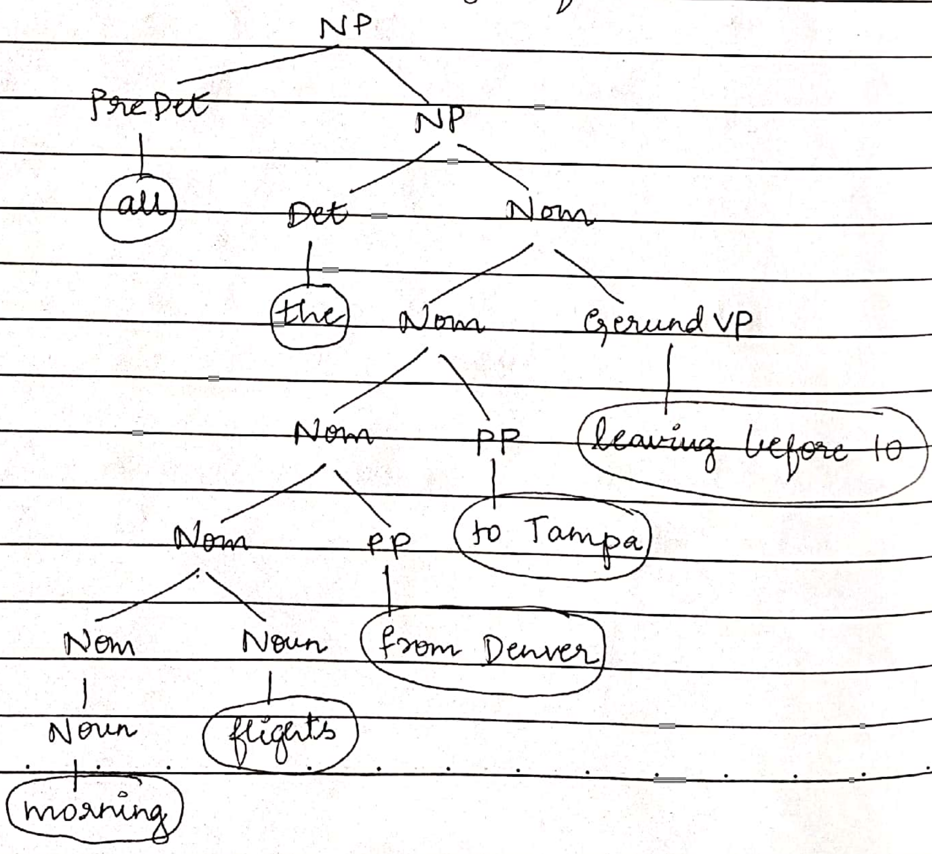
Ex:

- a flight [from Phoenix to Detroit] [leaving Monday evening]
- I need ~~a~~ a flight [to Seattle] [leaving from Baltimore] [making a stop in New York]
- evening flights [from Nashville to Houston] [that servers dinner]

Before the Noun Phrase

Word classes that modify and appear before NPs are called predetermines.
A common predeterminer is all:
- all the flights - all flights

Parse tree for "all the morning flights from Denver to Tampa leaving before 10".



Det → post det → pre head
nominal → post head

det nom PP

Show me the meal [on flight Boeing-780] [from P1] [to P2.]
PP PP

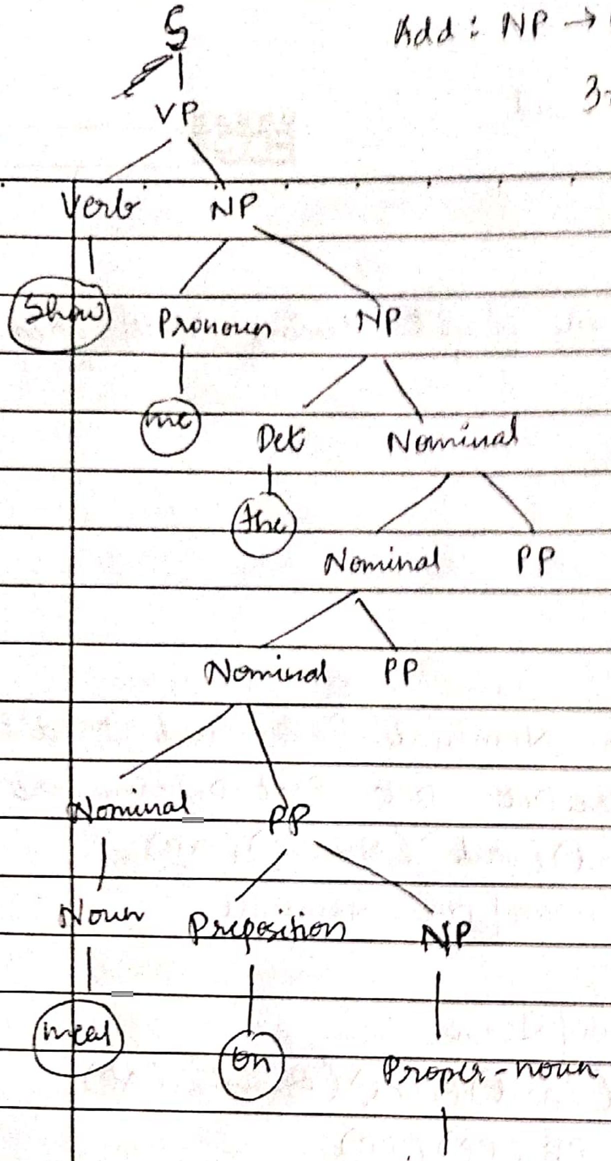
CFG's

S → NP VP
NP → Pre-Head Modifier Nominal Post-Head-Modifiers
Pre-Head-Modifiers → Pre Det Det Post-Determiners
Post-Determiners → (Card)(Ord)(Quant)(AP)
NP → Pronoun | Proper-noun | Det Nominal
Det → NP's
Nominal → Nominal Noun | Noun
Pre-head-Modifiers → (Card)(Ord)(Quant)(AP)
Nominal → Nominal PP (PP)(PP)
Nominal → Nominal GerundVP
Nominal → Nominal RelClause
RelClause → (who | that) VP
Gerund VP → Gerund V | GerundV NP | GerundV PP | GerundV NP PP
VP → Verb | Verb NP | Verb PP | Verb NP PP
PP → Preposition NP

~~Noun~~ m → (~~Noun~~ m PP(PP(PP))) | (~~Noun~~ m GerundVP) | (~~Noun~~ m RelClause)

Add : NP → Pronoun NP

3rd singular form - She, he



First Person Vs Second person

First Person $\quad S \rightarrow Aux\ NP\ VP$

Agreement → Subject - verb agreement or NV agreem

Determiner - Noun agreement

One way is to extend our grammar with multiple sets of rules

- one rule set for 3sg subjects, and
- one for non - 3 sg subjects

Ex : the rule handling yes-no question

He was found
to be guilty

URBAN EDGE

3SgNP → Det SgNominal
Non3SgNP → Det PlNominal
SgNominal → SgNoun
PlNominal → PlNoun
SgNoun → flight | fare | dollar
PlNoun → flights | fares | dollars

The Verb Phrase

VP → Verb
VP → Verb NP
VP → V NP PP
VP → V PP

correction →

VP → VP S (NP)
VP → VP VP

want
- ~~NP complement (I want a flight)~~
- VP complement (I want to ~~flight~~ fly)

find → NP complement
transitive verb – find

Ex : I found (a NLP class)(from coursera)

I — Pronoun, found — verb, (a NLP class) — NP complement, (from coursera) — PP