CS626: NLP, Speech and the Web

Pushpak Bhattacharyya
CSE Dept.,
IIT Bombay

Lecture 15, 17: Parsing Ambiguity, Probabilistic Parsing, sample seminar

17th and 24th September, 2012

(16th lecture was on UNL by Avishek)

Dealing With Structural Ambiguity

- Multiple parses for a sentence
  - The man saw the boy with a telescope.
  - The man saw the mountain with a telescope.
  - The man saw the boy with the ponytail.

At the level of syntax, all these sentences are ambiguous. But semantics can disambiguate 2^{nd} & 3^{rd} sentence.

Prepositional Phrase (PP) Attachment Problem

$V - NP_1 - P - NP_2$

(Here P means preposition)

NP_2 attaches to NP_1 ?

or NP_2 attaches to V ?

- Xerox Parsers
    - XLE Parser (Freely available)
    - XIP Parser (expensive)

Linguistic Data Consortium
(LDC at UPenn)

Europe

India

European Language Resource Association (ELRA)

Linguistic Data Consortium for Indian Languages (LDC-IL)

Parse Trees for a Structurally Ambiguous Sentence

Let the grammar be –

$S \rightarrow$ NP VP

$NP \rightarrow$ DT N | DT N PP

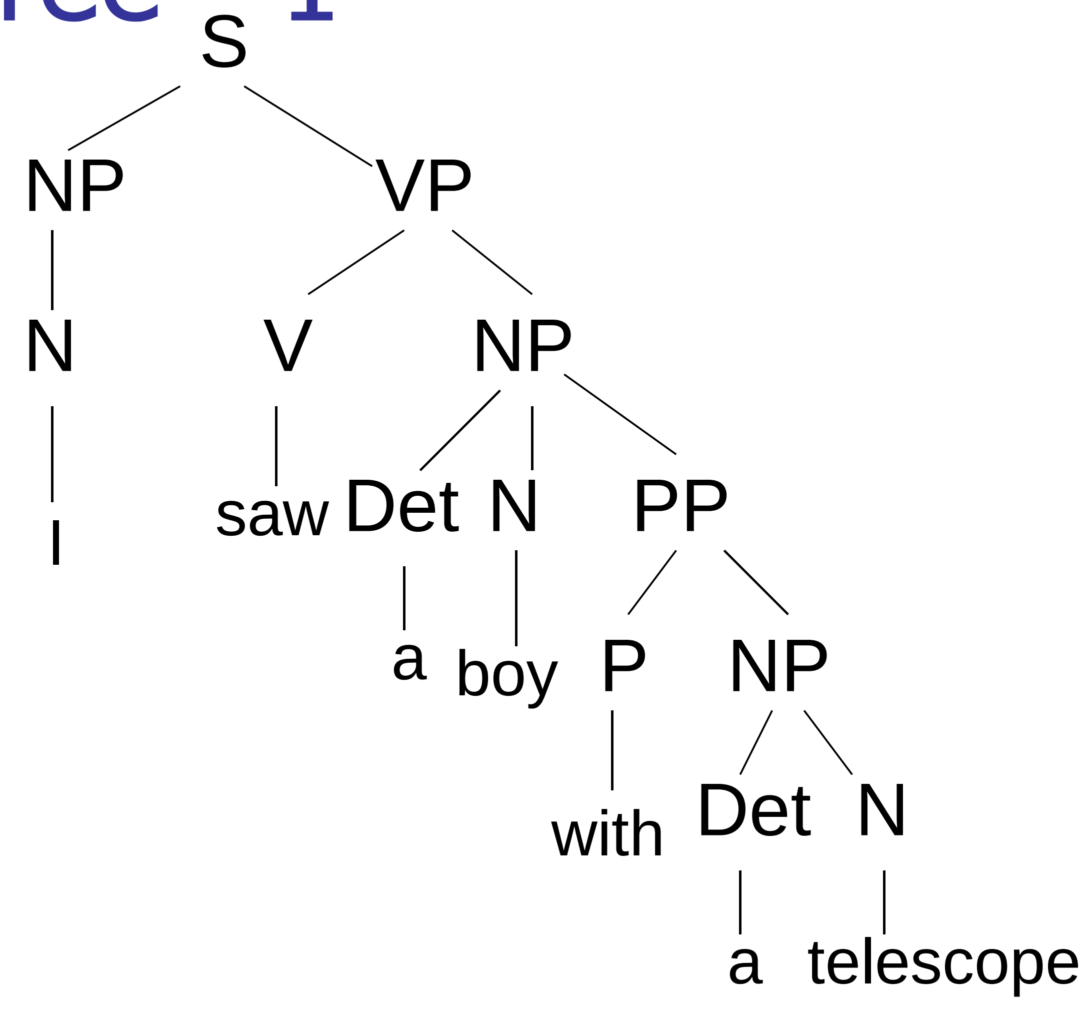
$PP \rightarrow$ P NP

$VP \rightarrow$ V NP PP | V NP

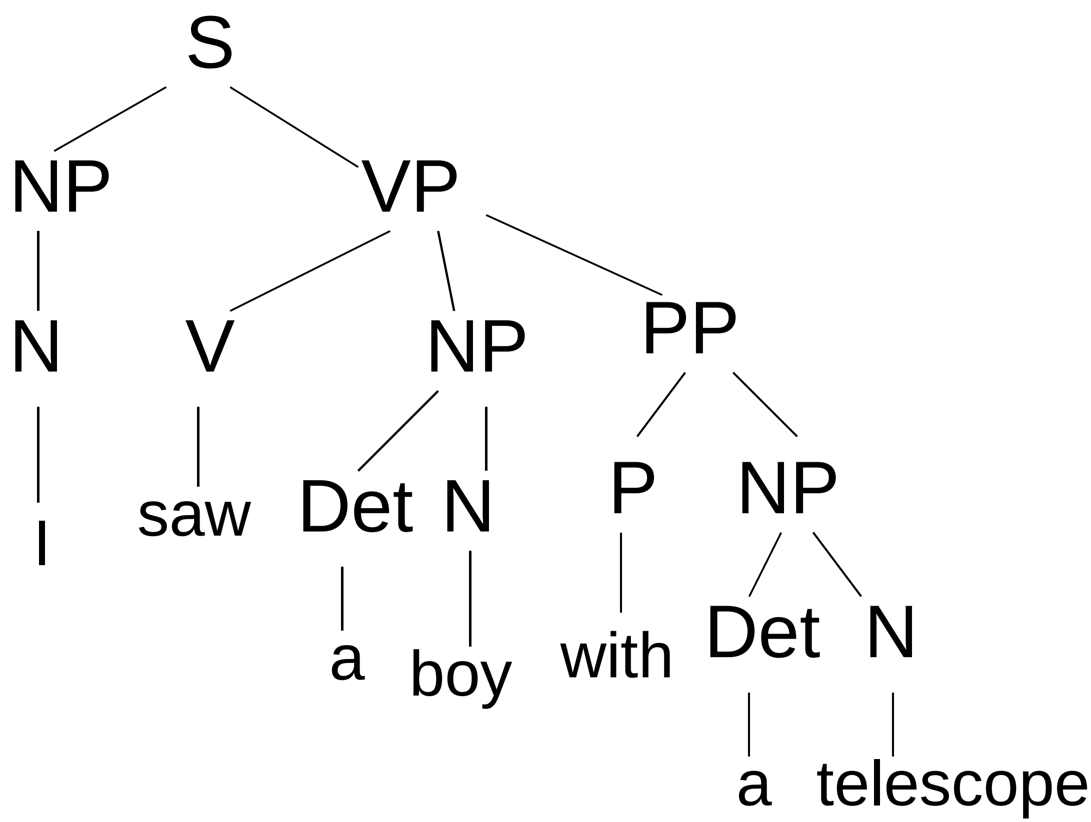
For the sentence,

"I saw a boy with a telescope"

Parse Tree - 1



Parse Tree -2



Parsing Structural Ambiguity

Parsing for Structurally Ambiguous Sentences

- Sentence “I saw a boy with a telescope”
- Grammar:

$$
\begin{array}{lcl}
S & \rightarrow & NP\ VP \\
NP & \rightarrow & ART\ N \mid ART\ N\ PP \mid PRON \\
VP & \rightarrow & V\ NP\ PP \mid V\ NP \\
ART & \rightarrow & a \mid an \mid the \\
N & \rightarrow & boy \mid telescope \\
PRON & \rightarrow & I \\
V & \rightarrow & saw
\end{array}
$$

Ambiguous Parses

- Two possible parses:
  - PP attached with Verb (i.e. *I used a telescope to see*)
  - (S (NP (PRON "I")) (VP (V "saw")
    (NP ((ART "a") (N "boy"))
    (PP (P "with") (NP (ART "a") (N "telescope")))))
  - PP attached with Noun (i.e. *boy had a telescope*)
  - (S (NP (PRON "I")) (VP (V "saw")
    (NP ((ART "a") (N "boy")
    (PP (P "with") (NP (ART "a") (N "telescope"))))))

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
2	((NP VP) 1)	–	–	Use NP → ART N | ART N PP | PRON

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
2	((NP VP) 1)	–	–	Use NP → ART N | ART N PP | PRON
3	((ART N VP) 1)	(a) ((ART N PP VP) 1) (b) ((PRON VP) 1)	–	ART does not match "I", backup state (b) used

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
2	((NP VP) 1)	–	–	Use NP → ART N | ART N PP | PRON
3	((ART N VP) 1)	(a) ((ART N PP VP) 1) (b) ((PRON VP) 1)	–	ART does not match "I", backup state (b) used
3 B	((PRON VP) 1)	–	–	

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
2	((NP VP) 1)	–	–	Use NP → ART N | ART N PP | PRON
3	((ART N VP) 1)	(a) ((ART N PP VP) 1) (b) ((PRON VP) 1)	–	ART does not match "I", backup state (b) used
3 B	((PRON VP) 1)	–	–	
4	((VP) 2)	–	Consumed "I"	

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
2	((NP VP) 1)	–	–	Use NP → ART N | ART N PP | PRON
3	((ART N VP) 1)	(a) ((ART N PP VP) 1) (b) ((PRON VP) 1)	–	ART does not match "I", backup state (b) used
3 B	((PRON VP) 1)	–	–	
4	((VP) 2)	–	Consumed "I"	
5	((V NP PP) 2)	((V NP) 2)	–	Verb Attachment Rule used

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
2	((NP VP) 1)	–	–	Use NP → ART N | ART N PP | PRON
3	((ART N VP) 1)	(a) ((ART N PP VP) 1) (b) ((PRON VP) 1)	–	ART does not match "I", backup state (b) used
3 B	((PRON VP) 1)	–	–	
4	((VP) 2)	–	Consumed "I"	
5	((V NP PP) 2)	((V NP) 2)	–	Verb Attachment Rule used
6	((NP PP) 3)	–	Consumed "saw"	

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
2	((NP VP) 1)	–	–	Use NP → ART N | ART N PP | PRON
3	((ART N VP) 1)	(a) ((ART N PP VP) 1) (b) ((PRON VP) 1)	–	ART does not match "I", backup state (b) used
3 B	((PRON VP) 1)	–	–	
4	((VP) 2)	–	Consumed "I"	
5	((V NP PP) 2)	((V NP) 2)	–	Verb Attachment Rule used
6	((NP PP) 3)	–	Consumed "saw"	
7	((ART N PP) 3)	(a) ((ART N PP PP) 3) (b) ((PRON PP) 3)		

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
2	((NP VP) 1)	–	–	Use NP → ART N | ART N PP | PRON
3	((ART N VP) 1)	(a) ((ART N PP VP) 1) (b) ((PRON VP) 1)	–	ART does not match "I", backup state (b) used
3 B	((PRON VP) 1)	–	–	
4	((VP) 2)	–	Consumed "I"	
5	((V NP PP) 2)	((V NP) 2)	–	Verb Attachment Rule used
6	((NP PP) 3)	–	Consumed "saw"	
7	((ART N PP) 3)	(a) ((ART N PP PP) 3) (b) ((PRON PP) 3)		
8	((N PP) 4)	–	Consumed "a"	

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
2	((NP VP) 1)	–	–	Use NP → ART N | ART N PP | PRON
3	((ART N VP) 1)	(a) ((ART N PP VP) 1) (b) ((PRON VP) 1)	–	ART does not match "I", backup state (b) used
3 B	((PRON VP) 1)	–	–	
4	((VP) 2)	–	Consumed "I"	
5	((V NP PP) 2)	((V NP) 2)	–	Verb Attachment Rule used
6	((NP PP) 3)	–	Consumed "saw"	
7	((ART N PP) 3)	(a) ((ART N PP PP) 3) (b) ((PRON PP) 3)		
8	((N PP) 4)	–	Consumed "a"	
9	((PP) 5)	–	Consumed	

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
2	((NP VP) 1)	–	–	Use NP → ART N | ART N PP | PRON
3	((ART N VP) 1)	(a) ((ART N PP VP) 1) (b) ((PRON VP) 1)	–	ART does not match "I", backup state (b) used
3 B	((PRON VP) 1)	–	–	
4	((VP) 2)	–	Consumed "I"	
5	((V NP PP) 2)	((V NP) 2)	–	Verb Attachment Rule used
6	((NP PP) 3)	–	Consumed "saw"	
7	((ART N PP) 3)	(a) ((ART N PP PP) 3) (b) ((PRON PP) 3)		
8	((N PP) 4)	–	Consumed "a"	
9	((PP) 5)	–	Consumed	

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
…	…	…	…	…
7	((ART N PP) 3)	(a) ((ART N PP PP) 3) (b) ((PRON PP) 3)		
8	((N PP) 4)	–	Consumed "a"	
9	((PP) 5)	–	Consumed "boy"	
10	((P NP) 5)	–	–	
11	((NP) 6)	–	Consumed "with"	

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
…	…	…	…	…
7	((ART N PP) 3)	(a) ((ART N PP PP) 3) (b) ((PRON PP) 3)		
8	((N PP) 4)	–	Consumed "a"	
9	((PP) 5)	–	Consumed "boy"	
10	((P NP) 5)	–	–	
11	((NP) 6)	–	Consumed "with"	
12	((ART N) 6)	(a) ((ART N PP) 6) (b) ((PRON) 6)	–	

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
…	…	…	…	…
7	((ART N PP) 3)	(a) ((ART N PP PP) 3) (b) ((PRON PP) 3)		
8	((N PP) 4)	–	Consumed "a"	
9	((PP) 5)	–	Consumed "boy"	
10	((P NP) 5)	–	–	
11	((NP) 6)	–	Consumed "with"	
12	((ART N) 6)	(a) ((ART N PP) 6) (b) ((PRON) 6)	–	
13	((N) 7)	–	Consumed "a"	

Top Down Parse

	State	Backup State	Action	Comments
1	((S) 1)	–	–	Use S → NP VP
…	…	…	…	…
7	((ART N PP) 3)	(a) ((ART N PP PP) 3) (b) ((PRON PP) 3)		
8	((N PP) 4)	–	Consumed "a"	
9	((PP) 5)	–	Consumed "boy"	
10	((P NP) 5)	–	–	
11	((NP) 6)	–	Consumed "with"	
12	((ART N) 6)	(a) ((ART N PP) 6) (b) ((PRON) 6)	–	
13	((N) 7)	–	Consumed "a"	
14	((–) 8)	–	Consume "telescope" Finish Parsing	

Top Down Parsing - Observations

- Top down parsing gave us the Verb Attachment Parse Tree (i.e., *I used a telescope*)
- To obtain the alternate parse tree, the backup state in step 5 will have to be invoked
- Is there an efficient way to obtain all parses ?

Bottom Up Parse

I	saw	a	boy	with	a	telescope

1 2 3 4 5 6 7 8

Colour Scheme :

- Blue for Normal Parse
- Green for Verb Attachment Parse
- Purple for Noun Attachment Parse
- Red for Invalid Parse

Bottom Up Parse

I saw a boy with a telescope

1 2 3 4 5 6 7 8

$NP_{12} \rightarrow PRON_{12} \bullet$

$S_{1?} \rightarrow NP_{12} \bullet VP_{2?}$

Bottom Up Parse

I saw a boy with a telescope

1 2 3 4 5 6 7 8

$NP_{12} \rightarrow PRON_{12} \bullet$ $VP_{2?} \rightarrow V_{23} \bullet NP_{3?}$

$S_{1?} \rightarrow NP_{12} \bullet VP_{2?}$ $VP_{2?} \rightarrow V_{23} \bullet NP_{3?} PP_{??}$

Bottom Up Parse

I	saw	a	boy	with	a	telescope

1 2 3 4 5 6 7 8

$NP_{12} \rightarrow PRON_{12} \bullet$ $VP_{2?} \rightarrow V_{23} \bullet NP_{3?}$ $NP_{35} \rightarrow ART_{34} \bullet N_{45}$

$S_{1?} \rightarrow NP_{12} \bullet VP_{2?}$ $VP_{2?} \rightarrow V_{23} \bullet NP_{3?} PP_{??}$ $NP_{3?} \rightarrow ART_{34} \bullet N_{45} PP_{5?}$

Bottom Up Parse

	I	saw	a	boy	with	a	telescope
1	2	3	4	5	6	7	8

$NP_{12} \rightarrow PRON_{12} \bullet$ $\quad VP_{2?} \rightarrow V_{23} \bullet NP_{3?}$ $\quad NP_{35} \rightarrow ART_{34} \bullet N_{45}$ $\quad NP_{35} \rightarrow ART_{34} N_{45} \bullet$

$S_{1?} \rightarrow NP_{12} \bullet VP_{2?}$ $\quad VP_{2?} \rightarrow V_{23} \bullet NP_{3?} PP_{??}$ $\quad NP_{3?} \rightarrow ART_{34} \bullet N_{45} PP_{5?}$ $\quad NP_{3?} \rightarrow ART_{34} N_{45} \bullet PP_{5?}$

Bottom Up Parse

I saw a boy with a telescope

1 2 3 4 5 6 7 8

$NP_{12} \rightarrow PRON_{12}\bullet$	$VP_{2?} \rightarrow V_{23}\bullet NP_{3?}$	$NP_{35} \rightarrow ART_{34}\bullet N_{45}$	$NP_{35} \rightarrow ART_{34} N_{45}\bullet$
$S_{1?} \rightarrow NP_{12}\bullet VP_{2?}$	$VP_{2?} \rightarrow V_{23}\bullet NP_{3?} PP_{??}$	$NP_{3?} \rightarrow ART_{34}\bullet N_{45} PP_{5?}$	$NP_{3?} \rightarrow ART_{34} N_{45}\bullet PP_{5?}$

$VP_{25} \rightarrow V_{23} NP_{35}\bullet$

$S_{15} \rightarrow NP_{12} VP_{25}\bullet$

$VP_{2?} \rightarrow V_{23} NP_{35}\bullet PP_{5?}$

Bottom Up Parse

I saw a boy with a telescope

1 2 3 4 5 6 7 8

$NP_{12} \rightarrow PRON_{12} \bullet$ $VP_{2?} \rightarrow V_{23} \bullet NP_{3?}$ $NP_{35} \rightarrow ART_{34} \bullet N_{45}$ $NP_{35} \rightarrow ART_{34} N_{45} \bullet$ $PP_{5?} \rightarrow P_{56} \bullet NP_{6?}$

$S_{1?} \rightarrow NP_{12} \bullet VP_{2?}$ $VP_{2?} \rightarrow V_{23} \bullet NP_{3?} PP_{??}$ $NP_{3?} \rightarrow ART_{34} \bullet N_{45} PP_{5?}$ $NP_{3?} \rightarrow ART_{34} N_{45} \bullet PP_{5?}$

$VP_{25} \rightarrow V_{23} NP_{35} \bullet$ $VP_{2?} \rightarrow V_{23} NP_{35} \bullet PP_{5?}$

$S_{15} \rightarrow NP_{12} VP_{25} \bullet$

Bottom Up Parse

I saw a boy with a telescope

1 2 3 4 5 6 7 8

$NP_{12} \rightarrow PRON_{12}\bullet$ $VP_{2?} \rightarrow V_{23}\bullet NP_{3?}$ $NP_{35} \rightarrow ART_{34}\bullet N_{45}$ $NP_{35} \rightarrow ART_{34}N_{45}\bullet$ $PP_{5?} \rightarrow P_{56}\bullet NP_{6?}$ $NP_{68} \rightarrow ART_{67}\bullet N_{7?}$

$S_{1?} \rightarrow NP_{12}\bullet VP_{2?}$ $VP_{2?} \rightarrow V_{23}\bullet NP_{3?}PP_{??}$ $NP_{3?} \rightarrow ART_{34}\bullet N_{45}PP_{5?}$ $NP_{3?} \rightarrow ART_{34}N_{45}\bullet PP_{5?}$ $NP6_{?} \rightarrow ART_{67}\bullet N_{78}PP_{8?}$

$VP_{25} \rightarrow V_{23}NP_{35}\bullet$ $VP_{2?} \rightarrow V_{23}NP_{35}\bullet PP_{5?}$

$S_{15} \rightarrow NP_{12}VP_{25}\bullet$

Bottom Up Parse

I saw a boy with a telescope

1 2 3 4 5 6 7 8

$NP_{12} \rightarrow PRON_{12}\bullet$

$S_{1?} \rightarrow NP_{12}\bullet VP_{2?}$

$VP_{2?} \rightarrow V_{23}\bullet NP_{3?}$

$VP_{2?} \rightarrow V_{23}\bullet NP_{3?}PP_{??}$

$NP_{35} \rightarrow ART_{34}\bullet N_{45}$

$NP_{3?} \rightarrow ART_{34}\bullet N_{45}PP_{5?}$

$NP_{35} \rightarrow ART_{34}N_{45}\bullet$

$NP_{3?} \rightarrow ART_{34}N_{45}\bullet PP_{5?}$

$PP_{5?} \rightarrow P_{56}\bullet NP_{6?}$

$NP_{68} \rightarrow ART_{67}\bullet N_{7?}$

$NP6_{?} \rightarrow ART_{67}\bullet N_{78}PP_{8?}$

$NP_{68} \rightarrow ART_{67}N_{78}\bullet$

$VP_{25} \rightarrow V_{23}NP_{35}\bullet$

$S_{15} \rightarrow NP_{12}VP_{25}\bullet$

$VP_{2?} \rightarrow V_{23}NP_{35}\bullet PP_{5?}$

Bottom Up Parse

I saw a boy with a telescope

1 2 3 4 5 6 7 8

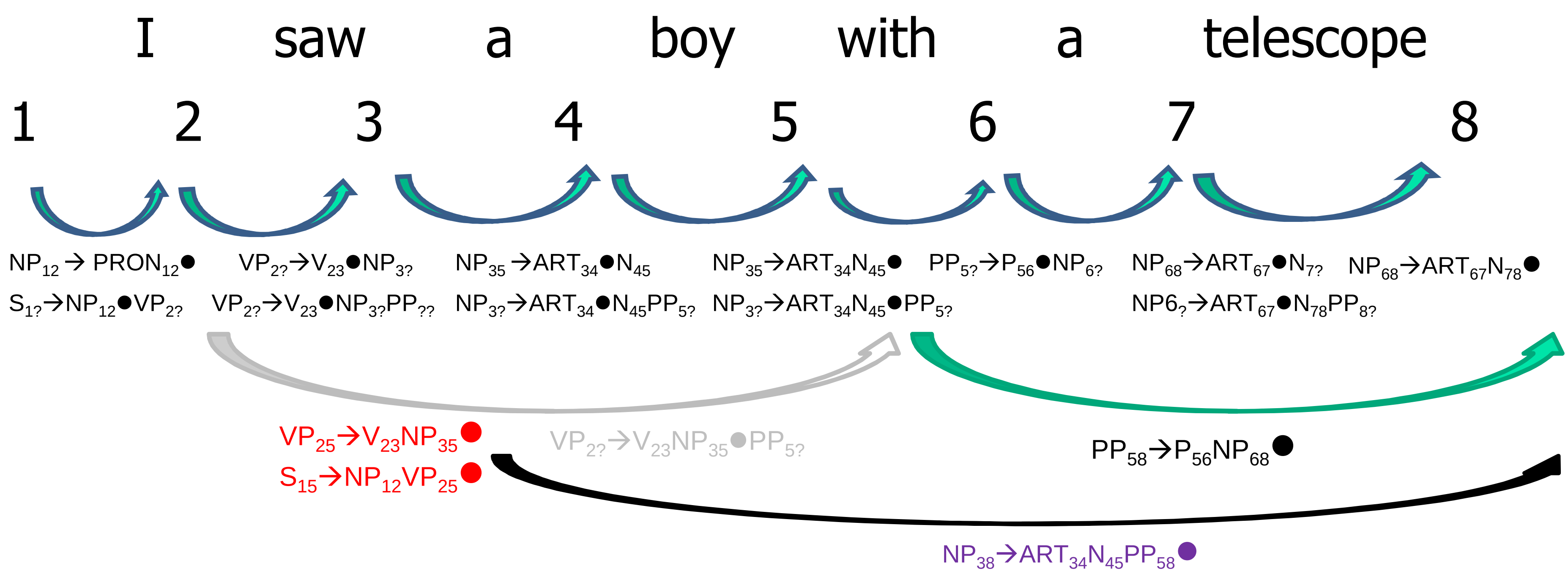
$NP_{12} \rightarrow PRON_{12}\bullet$ $VP_{2?}\rightarrow V_{23}\bullet NP_{3?}$ $NP_{35} \rightarrow ART_{34}\bullet N_{45}$ $NP_{35}\rightarrow ART_{34}N_{45}\bullet$ $PP_{5?}\rightarrow P_{56}\bullet NP_{6?}$ $NP_{68}\rightarrow ART_{67}\bullet N_{7?}$ $NP_{68}\rightarrow ART_{67}N_{78}\bullet$

$S_{1?}\rightarrow NP_{12}\bullet VP_{2?}$ $VP_{2?}\rightarrow V_{23}\bullet NP_{3?}PP_{??}$ $NP_{3?}\rightarrow ART_{34}\bullet N_{45}PP_{5?}$ $NP_{3?}\rightarrow ART_{34}N_{45}\bullet PP_{5?}$ $NP6_{?}\rightarrow ART_{67}\bullet N_{78}PP_{8?}$

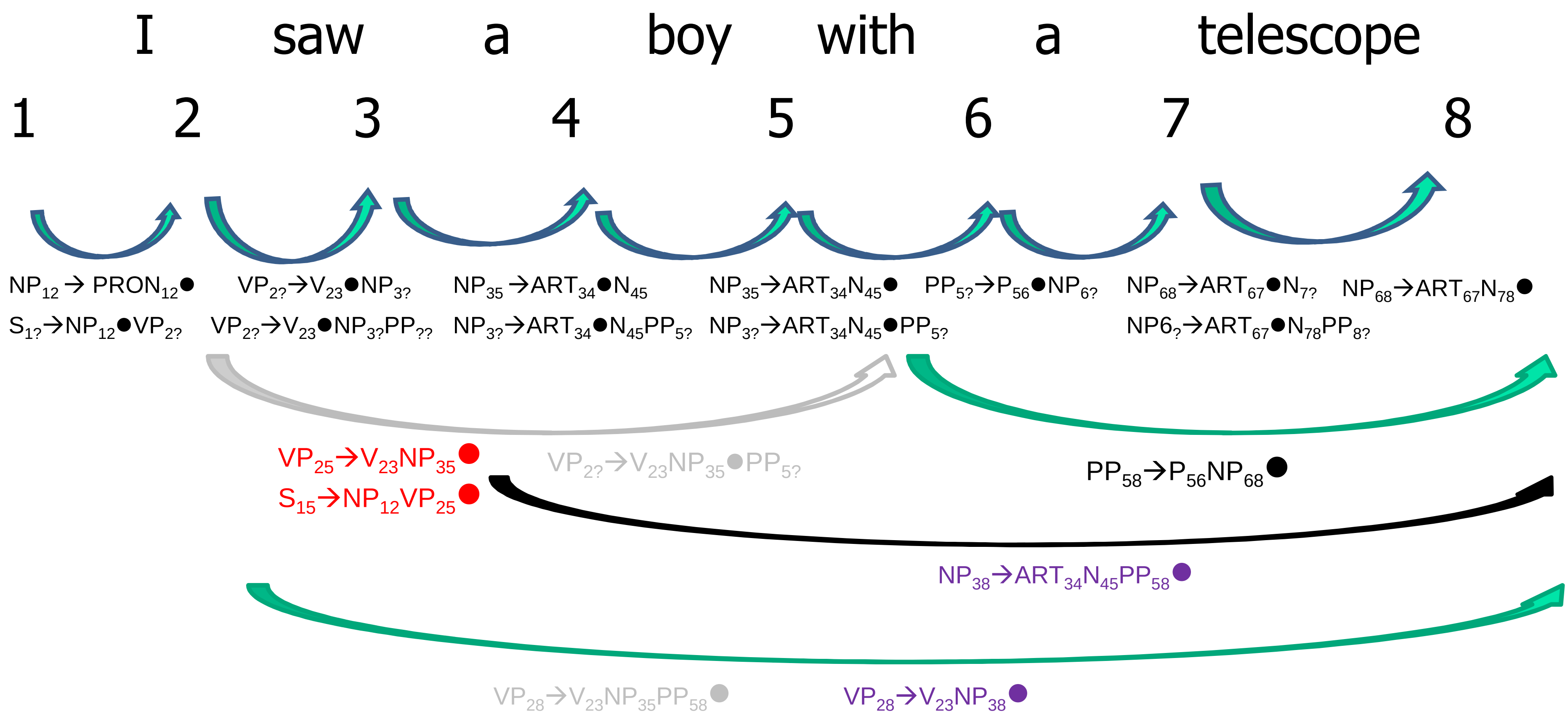
$VP_{25}\rightarrow V_{23}NP_{35}\bullet$ $VP_{2?}\rightarrow V_{23}NP_{35}\bullet PP_{5?}$ $PP_{58}\rightarrow P_{56}NP_{68}\bullet$

$S_{15}\rightarrow NP_{12}VP_{25}\bullet$

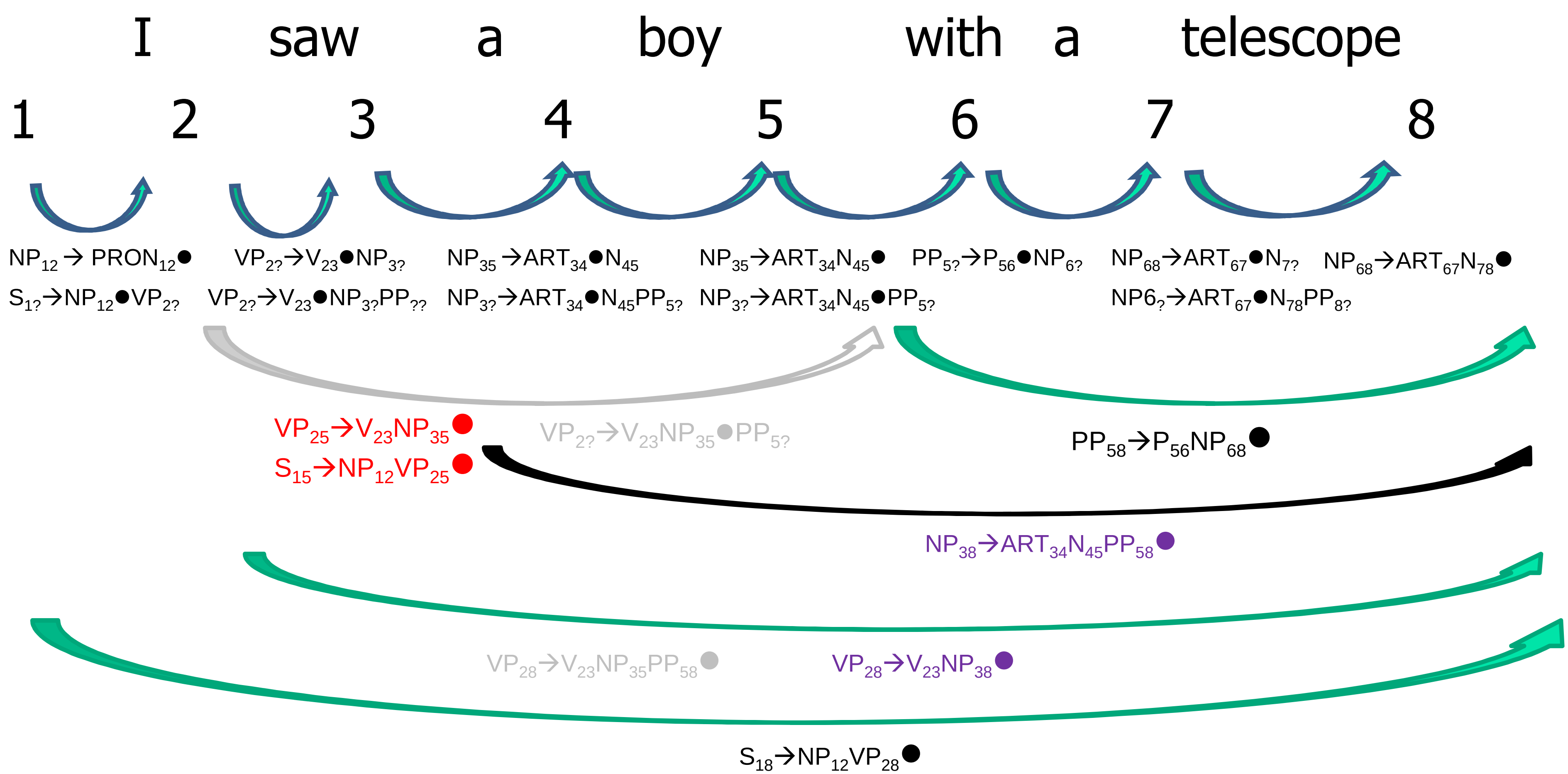
Bottom Up Parse



Bottom Up Parse



Bottom Up Parse



Bottom Up Parsing - Observations

- Both Noun Attachment and Verb Attachment Parses obtained by simply systematically applying the rules
- Numbers in subscript help in verifying the parse and getting chunks from the parse

Exercise

For the sentence,

"The man saw the boy with a telescope"
& the grammar given previously,
compare the performance of top-down,
bottom-up & top-down chart parsing.

Start of Probabilistic Parsing

Example of Sentence labeling: Parsing

$[_{S1}[_{S}[_{S}[_{VP}[_{VB}$Come$][_{NP}[_{NNP}$July$]]]]$

$[_{,},]$

$[_{CC}$ and$]$

$[_{S}\ [_{NP}\ [_{DT}$ the$]\ [_{JJ}$ IIT$]\ [_{NN}$ campus$]]$

$[_{VP}\ [_{AUX}$ is$]$

$[_{ADJP}\ [_{JJ}$ abuzz$]$

$[_{PP}[_{IN}$ with$]$

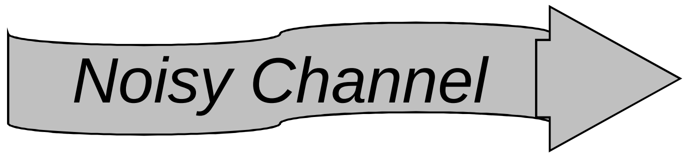
$[_{NP}[_{ADJP}\ [_{JJ}$ new$]\ [_{CC}$ and$]\ [\ _{VBG}$ returning$]]$

$[_{NNS}$ students$]]]]]]$

$[_{.}.]]]$

Noisy Channel Modeling

Source sentence



Target parse

$$T^* = \underset{T}{\operatorname{argmax}} [P(T|S)]$$

$$= \underset{T}{\operatorname{argmax}} [P(T).P(S|T)]$$

$$= \underset{T}{\operatorname{argmax}} [P(T)],$$ since given the parse the sentence is completely determined and $P(S|T)=1$

Corpus

- A collection of text called *corpus*, is used for collecting various language data
- With annotation: more information, but manual labor intensive
- Practice: *label automatically; correct manually*
- The famous *Brown Corpus* contains 1 million tagged words.
- **Switchboard:** very famous corpora 2400 conversations, 543 speakers, many US dialects, annotated with orthography and phonetics

Discriminative vs. Generative Model

$$W^* = \underset{W}{\operatorname{argmax}} \left(P(W|SS)\right)$$

Discriminative Model

Compute directly from $P(W|SS)$

Generative Model

Compute from $P(W).P(SS|W)$

Language Models

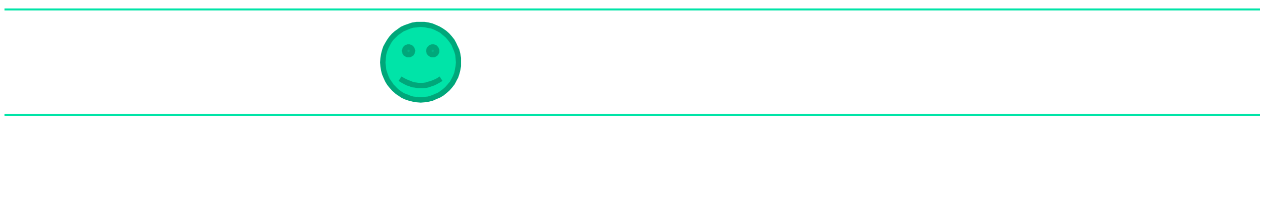
- N-grams: sequence of n consecutive words/characters
- Probabilistic / Stochastic Context Free Grammars:
  - Simple probabilistic models capable of handling recursion
  - A CFG with probabilities attached to rules
  - Rule probabilities → how likely is it that a particular rewrite rule is used?

PCFGs

- Why PCFGs?
  - Intuitive probabilistic models for tree-structured languages
  - Algorithms are extensions of HMM algorithms
  - Better than the n-gram model for language modeling.

Data driven parsing is tied up with what is seen in the training data

- *"Stand right walk left."* Often a message on the airport escalators.

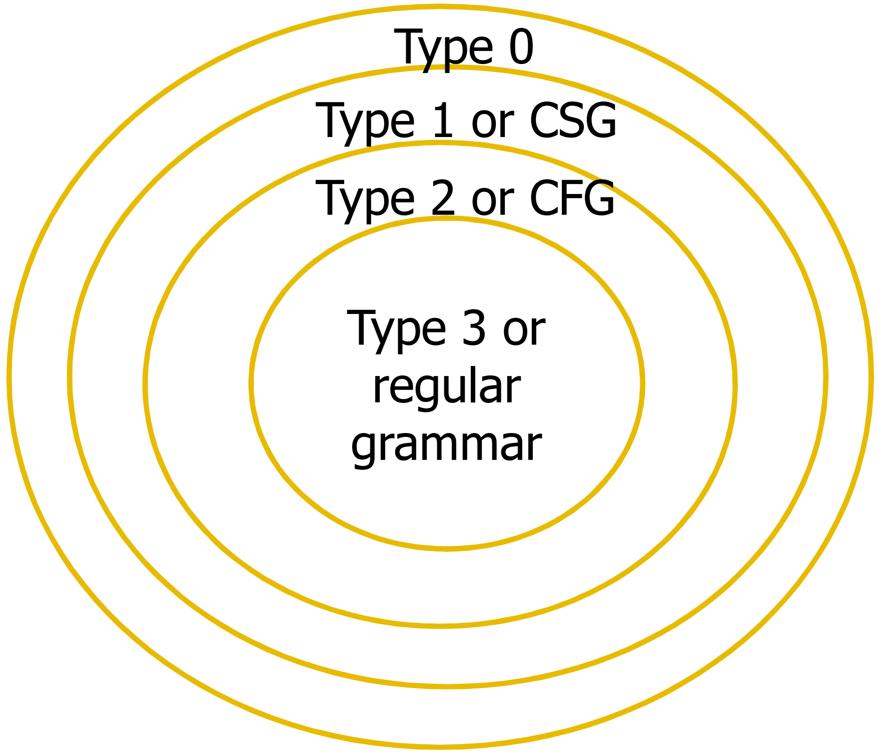


- "Walk left" can come from :- *"After climbing , we realized the walk left was still considerable."*
- "Stand right" can come from: - *"The stand right though it seemed at that time, does not stand the scrutiny now."*
- Then the given sentence "Stand right walk left" will never be parsed correctly.

Chomski Hierarchy

Properties of the hierarchy:-

- *The containment is proper.*
- *Grammar is not learnable even for the lowest type from positive examples alone.*



$$Type\,0 \supset Type\,1 \supset Type\,2 \supset Type\,3$$

Rationalism vs Empiricism

- Scientific paper : - *"The pendulum swings back"* by Ken Church.
- Every 20 years there is a shift from rationalism to empiricism.
- AI approaches are of two types:-
  - Rational (Theory and model driven e.g. Brill Tagger; humans decide the templates)
  - Empirical (Data driven e.g. HMM)
- Is it possible to classify probabilistic CFG into any of these categories?
  - Very difficult to answer.
  - Anything that comes purely from brain can be put into the category of "rationalism".
- There is no pure rationalism or pure empiricism

Laws of machine learning

- "Learning from vaccuum (zero knowledge) is impossible".
- "Inductive bias decides what to learn".
  - *In case of POS tagger , bias is that we assume the tagger to be a HMM model.*

Formal Definition of PCFG

- A PCFG consists of
  - A set of terminals $\{w_k\}$, k = 1,….,V

    $\{w_k\}$ = { child, teddy, bear, played…}
  - A set of non-terminals $\{N^i\}$, i = 1,…,n

    $\{N_i\}$ = { NP, VP, DT…}
  - A designated start symbol N^1
  - A set of rules $\{N^i \rightarrow \zeta^j\}$, where ζ^j is a sequence of terminals & non-terminals

    NP $\rightarrow$ DT NN
  - A corresponding set of rule probabilities

Rule Probabilities

- Rule probabilities are such that

$$\forall i \ \sum_{\mathrm{i}} \mathrm{P}(\mathrm{N}^{i} \rightarrow \zeta^{j}) = 1$$

E.g., P(NP → DT NN) = 0.2

P(NP → NN) = 0.5

P(NP → NP PP) = 0.3

- P(NP → DT NN) = 0.2
  - Means 20 % of the training data parses use the rule NP → DT NN

Probabilistic Context Free Grammars

- $S \rightarrow NP\ VP$ 1.0
- $NP \rightarrow DT\ NN$ 0.5
- $NP \rightarrow NNS$ 0.3
- $NP \rightarrow NP\ PP$ 0.2
- $PP \rightarrow P\ NP$ 1.0
- $VP \rightarrow VP\ PP$ 0.6
- $VP \rightarrow VBD\ NP$ 0.4

- $DT \rightarrow$ the 1.0
- $NN \rightarrow$ gunman 0.5
- $NN \rightarrow$ building 0.5
- $VBD \rightarrow$ sprayed 1.0
- $NNS \rightarrow$ bullets 1.0