CS772: Deep Learning for Natural Language Processing (DL-NLP)

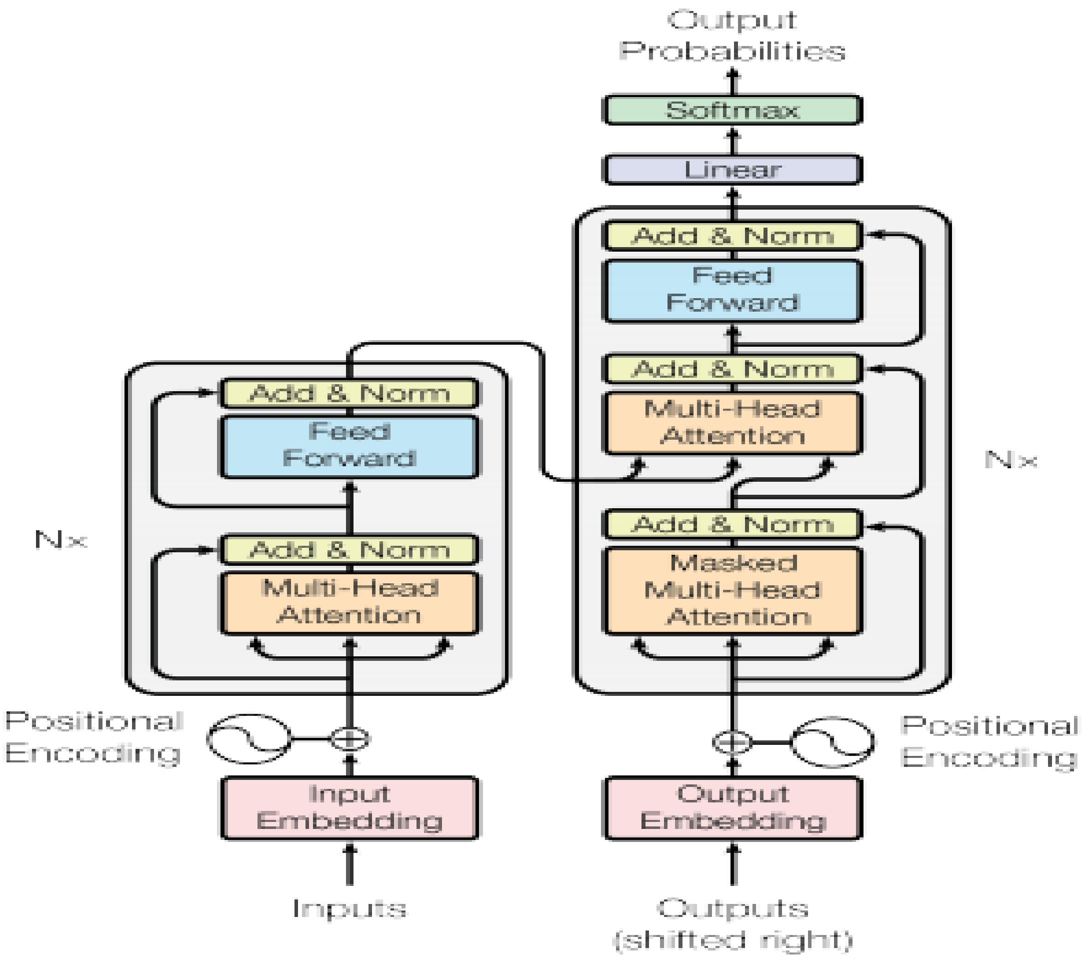
Transformer, LM, MT, CAI

Pushpak Bhattacharyya
Computer Science and Engineering Department
IIT Bombay

Week 11 of 20th March, 2023

Re-cap

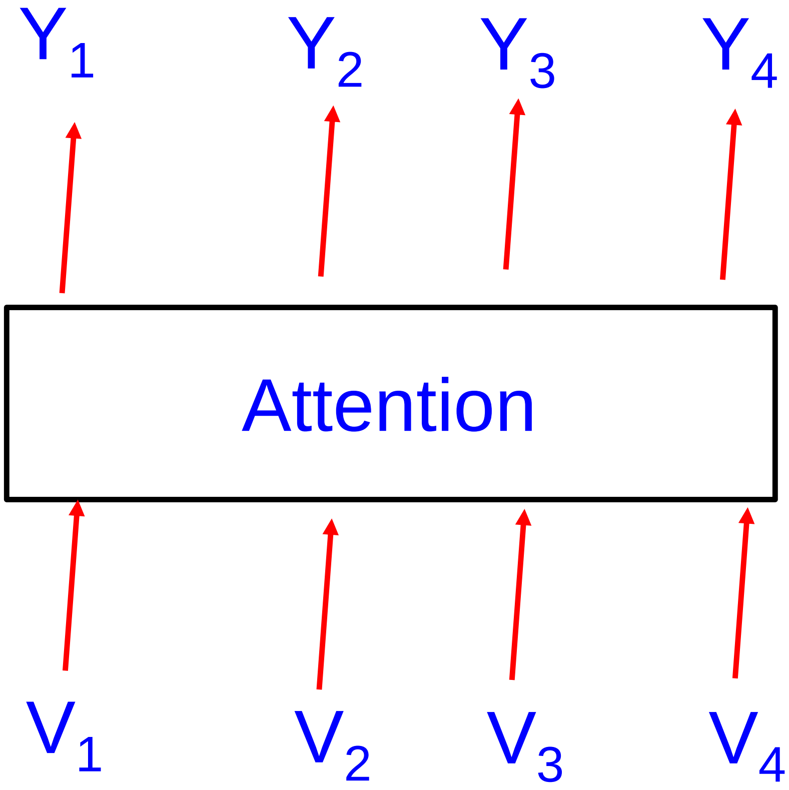
A classic diagram and a classic paper



Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. "Attention is all you need." NeurIPS (2017).

http://nlp.seas.harvard.edu/2018/04/03/attention.html
http://jalammar.github.io/illustrated-transformer/

Self Attention Block



Y_1 Y_2 Y_3 Y_4

Attention

V_1 V_2 V_3 V_4

Bank of the river

Word Embedding and Contextual Word Embedding

- Consider the phrase “*bank of the river*”
- Word embeddings of ‘*bank*’, ‘*of*’, ‘*the*’, ‘*river*’: V_1, V_2, V_3, V_4
- Now create a ‘score’ vector S_i for each word vector
- S_1: $(V_1.V_1, V_1.V_2, V_1.V_3, V_1.V_4)$
- Similarly, S_2, S_3, S_4

S-matrix

$$S = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \\ S_{41} & S_{42} & S_{43} & S_{44} \end{bmatrix}$$

S-scaled matrix

$$S-scaled = \frac{1}{\sqrt{d_k}} \times \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \\ S_{41} & S_{42} & S_{43} & S_{44} \end{bmatrix}$$

W-matrix

$$W = \begin{bmatrix} w_{11} & w_{12} & w_{13} & w_{14} \\ w_{21} & w_{22} & w_{23} & w_{24} \\ w_{31} & w_{32} & w_{33} & w_{34} \\ w_{41} & w_{42} & w_{43} & w_{44} \end{bmatrix}$$

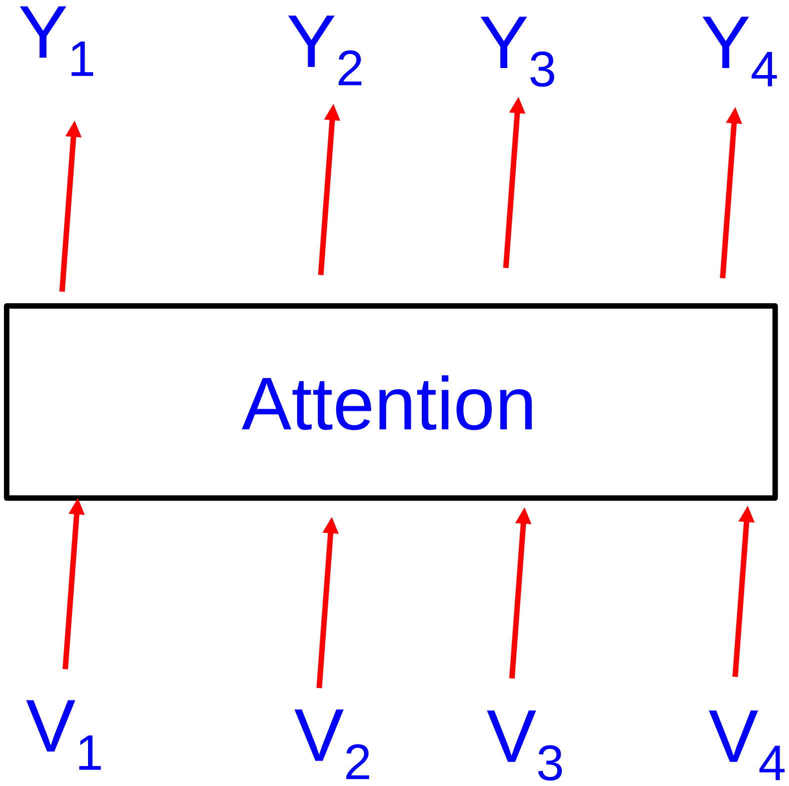
$$W_i - vector = soft\max\left(\frac{S_i - vector}{\sqrt{d_k}}\right)$$

Y-matrix

$$Y = \begin{bmatrix} y_{11} & y_{12} & y_{13} & y_{14} \\ y_{21} & y_{22} & y_{23} & y_{24} \\ y_{31} & y_{32} & y_{33} & y_{34} \\ y_{41} & y_{42} & y_{43} & y_{44} \end{bmatrix}$$

$$Y_i - vector = w_{11}.V_1 + w_{12}.V_2 + w_{13}.V_3 + w_{14}.V_4$$

Attention Block



Y_1 Y_2 Y_3 Y_4

Attention

V_1 V_2 V_3 V_4

Bank of the river

Interpretation of Y1 (1/2)

- Consider the word ‘BANK’ in the example - Bank of a River
- V1 is a member of a cluster in the embedding space which consists of financial words like loan, interest, money, etc.
- We want Y1 to not be a member of this cluster; It should have place property and not financial property
- Now, Y1 is a result of vector addition and is another vector in the same n-dimensional vector space

Interpretation of Y1 (2/2)

- If V4 vector (corresponding to RIVER) was very strong then it will result in greater emphasis in Y1
- How can we make V4 stronger?
- through attention which in turn is learnt learning
- Introduce weights of NN: Query, Key, Value with learnable parameters
- The weight w14 should be learnt in a way that it helps V4 exert more influence to create required Y1

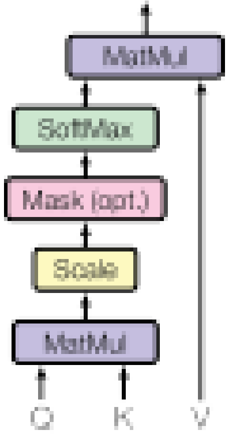
Query, Key and Value

$$attention(Q, K, V) = soft\max\left(\frac{Q\ .K^{T}}{\sqrt{d_{k}}}\right).V$$

Query, Key and Value with LEANABLE Parameter (1/2)

$$attention(Q,K,V) = soft\max\left(\frac{W^Q Q\ .W^K K^T}{\sqrt{d_k}}\right).W^V V$$

Scaled Dot-Product Attention



W^Q, *W*^{K} and *W*^{V} can be the weights of 3 linear layers of neurons which can be learnt by gradient descent

Positional Encoding

Position Vector components

Let the kth component of the tth position vector be denoted as *pos(t,k)*, *k* varying from 0 to *d-1, d* being the dimension of the PE vector. Then for even and odd positions (I varies from 0 to d/2-1):

$$pos(t,2i) = \sin\left(\frac{1}{10000^{\frac{2i}{d}}}t\right)$$

$$pos(t,2i+1) = \cos\left(\frac{1}{10000^{\frac{2i}{d}}}t\right)$$

Example: “Jack saw Jill” (1/2)

Three positions indexed as 0, 1 and 2.

Assume word vector dimension *d* to be *4*

Assume the frequency to be $1/(10^{2i/d})$

i varies from 0 to (4/2-1)=1

Then (cntd. next slide)

Example: “Jack saw Jill” (2/2)

$$pos\_vector('Jack') = < pos(0,0), pos(0,1), \xcancel{pos(0,0)}, pos(0,2), pos(0,3) >$$
$$= < \sin(0), \cos(0), \sin(0), \cos(0) >$$
$$= < 0,1,0,1 >$$

$$pos\_vector('saw') = < pos(1,0), pos(1,1), pos(1,2), pos(1,3) >$$
$$= \left\langle \sin\left(\frac{1}{10^{\frac{2\cdot 0}{4}}}\right), \cos\left(\frac{1}{10^{\frac{2\cdot 0}{4}}}\right), \sin\left(\frac{1}{10^{\frac{2\cdot 1}{4}}}\right), \cos\left(\frac{1}{10^{\frac{2\cdot 1}{4}}}\right) \right\rangle$$
$$= \left\langle \sin(1), \cos(1), \sin\left(\frac{1}{10^{0.5}}\right), \cos\left(\frac{1}{10^{0.5}}\right) \right\rangle$$

$$pos\_vector('Jill') = < pos(2,0), pos(2,1), pos(2,2), pos(2,3) >$$
$$= < \left\langle \sin(2), \cos(2), \sin\left(\frac{2}{10^{0.5}}\right), \cos\left(\frac{2}{10^{0.5}}\right) \right\rangle$$

Machine Translation

Why is MT difficult: Language Divergence

- Languages have different ways of expressing meaning
  - Lexico-Semantic Divergence
  - Structural Divergence

Our work on English-IL Language Divergence with illustrations from Hindi
(Dave, Parikh, Bhattacharyya, Journal of MT, 2002)

Simplified Vauquois Triangle

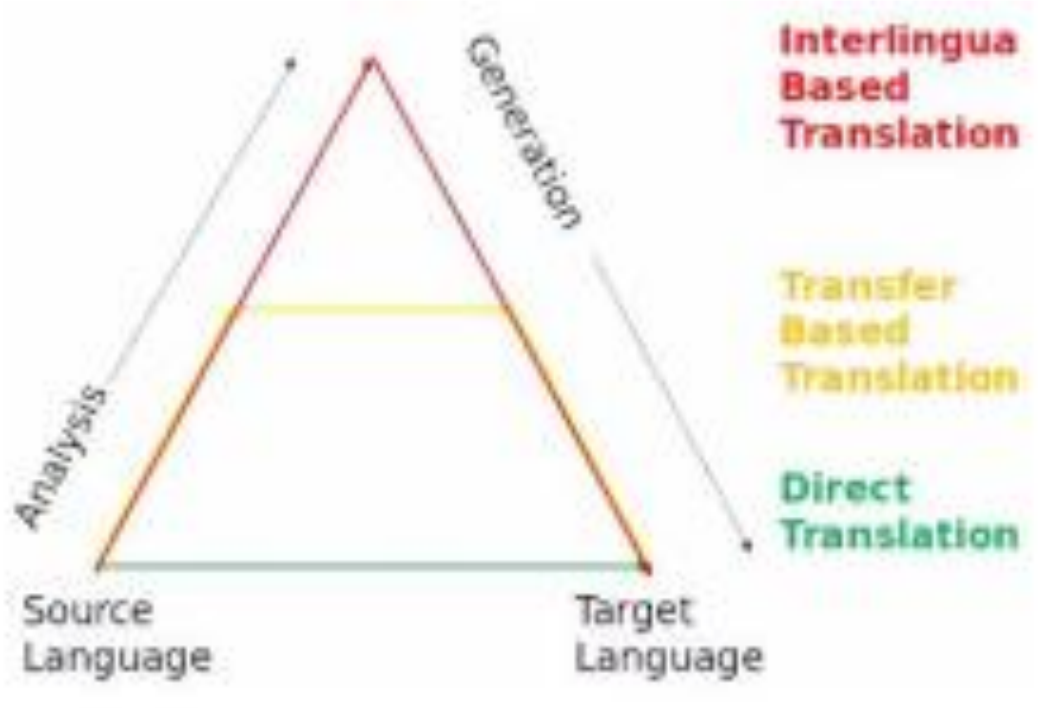


Figure X.6: Abridged Vauquois Triangle

ATG and NMT

- Analysis-Transfer-Generation, the foundation of MT

- NMT addresses this by
  - (a) encoding the input
  - (b) encoded vector is enriched by self attention
  - (c) cross attention and
  - (d) auto regression

MT and LM

Foundation

- Data driven approach
- Goal is to find out the English sentence e given foreign language sentence f whose p(e|f) is maximum.
- Translations are generated on the basis of statistical model
- Parameters are estimated using bilingual parallel corpora

$$\tilde{e} = \operatorname*{argmax}_{e \in e^*} p(e|f) = \operatorname*{argmax}_{e \in e^*} p(f|e)p(e)$$

SMT: Language Model

- To detect *good* English sentences
- Probability of an English sentence $w_1w_2 \ldots\ldots w_n$ can be written as

  $Pr(w_1w_2 \ldots\ldots w_n) = Pr(w_1) * Pr(w_2|w_1) * \ldots * Pr(w_n|w_1\, w_2 \ldots w_{n-1})$
- Here $Pr(w_n|w_1\, w_2 \ldots w_{n-1})$ is the probability that word w_n follows word string $w_1\, w_2 \ldots w_{n-1}$.
  - N-gram model probability
- Trigram model probability calculation

$$p(w_3|w_1w_2) = \frac{count(w_1w_2w_3)}{count(w_1w_2\,)}$$

Argmax(.): a very general framework (1/2)

$$e^* = \arg\max_e P(e \mid f)$$
$$= \arg\max_e [P(e)P(f \mid e)]$$

***P(e)*: language model**

- QA: *f→ question, e→answer*
- Summarization: *f→ text, e→summary*
- ASR: *f→ speech, e→text*
- TTS: *f→ text, e→speech*
- Image captioning: *f→ image, e→caption*
- Question Generation: *f→ <context, answer>, e→question*
- Disfleuncy correction: *f→ disfluent output of ASR, e→fluent output*

Argmax(.): a very general framework (2/2)

$$e^* = \arg\max_e P(e \mid f)$$
$$= \arg\max_e [P(e)P(f \mid e)]$$

***P(e)*: "language" model**

- POS Tagging: *f→ sentence, e→POS tag sequence*
- Chunking: *f→ sentence, e→BIO sequence*
- Parsing: *f→ sentence, e→parse tree*
- Spell Checking: *f→ incorrect spelling, e→ corrected spelling*
- Named Entity Recognition: *f→ sentence, e→ Name Tags*
- Dialogue Intent: *f→ dialogue sequence, e→ dialogue turns with Intent like 'question', 'elaboration', 'affirmation', 'command/request' etc.*

PaLM experiments

Metaphor → Normal

Normal → Metaphor

Metaphor to normal conversion: Prompts

- **Input**: I'm drowning in a sea of grief → **output**: I am having a lot of grief.

- **Input**: He really flared up my temper → **output**: He really made me angry.

- **Input**: I'm feeling blue →

  output: I'm feeling blue.

Metaphor to normal conversion: Testing

- **Input**: It is raining cats and dogs →

  output: It is raining heavily
- **Input**: Mowing the lawn was a walk in the park → **output**: Mowing the lawn was easy
- **Input**: Making chocolate chip cookies is a piece of cake →

  output: Making chocolate chip cookies is easy
- **Input**: Rohan is a night owl →

  output: Rohan is a person who stays up late at night
- **Input**: Bill is an early bird →

  output: Bill is an early riser.

Metaphor generation: Prompts

- **Input**: The area became warm → **Output**: Warmth blanketed the area

- **Input**: This is a very good thing to have additionally→
  output: This is the icing on the cake.

- **Input**: I'm feeling sad →
  Output: I'm feeling blue.

Metaphor generation: : Testing

- **Input**: His words were very harsh → **Output**: His words were like a dagger
- **Input**: I sense the failure coming → **Output**: I sense doom and gloom
- **Input**: She is going through a lot of ups and downs → **Output**: She is on a roller coaster ride
- **Input**: He is a very energetic fellow → **Output**: He is a ball of fire
- **Input**: These days John is very depressed → **Output**: John is down in the dumps these days

More experiments on PaLM: declarative → Interrogative

Thanks to Abhijeet Awasti

Prompts

- 1. Susan must leave. Must Susan leave?
- 2. Harry can swim. Can Harry swim?
- 3. Mary has read the book. Has Mary read the book?
- 4. Bill is sleeping. Is Bill sleeping?
- 5. John will solve the problem. Will John solve the problem?

Test Set

- 10. Bill could be sleeping.
- 11. Mary has been reading.
- 12. Susan should have left.
- 13. John might be late.
- 14. The boy who is here left
- 15 The player who should have played is injured

Results

- 10. Could Bill be sleeping?
- 11. Has Mary been reading?
- 12. Should Susan have left?
- 13. Might John be late?
- 14. Did the boy who is here leave?
- 15. Should the player who have played be injured?

Handling center embedding is always difficult even for humans

Need for Abstract underlying structure.

- Abstract underlying structure required.
- Three other phenomena displaying the same abstract pattern; such as: *Negation, Emphasis* and *Verb phrase Ellipsis:*

NEGATION

(13)	John left	John didn't leave.
	John should leave.	John shouldn't leave.
	John has left.	John hasn't left.
	John is leaving.	John isn't leaving.

Emphasis and Verb Phrase Ellipsis

EMPHASIS

(14) John left. John **did** leave.

John should leave. John **should** leave.

John has left. John **has** left.

John is leaving. John **is** leaving.

VERB PHRASE ELLIPSIS

(15) John left. Mary did too.

John should leave. Mary should too.

John has left. Mary has too.

John is leaving. Mary is too.

Classical Language Modelling: Syntax

Continued from last lecture

Syntax and Grammar

- Syntax is the study of how sentences are structured, the rules underlying

- Grammaticality: study of well formedness and ill formedness; *does the string belong to the language?*

    - *Delhi is the capital of India*
    - **Is of capital the Delhi India*
    - *?India is the capital of Delhi*

NLP= Linguistics + Probability

- “Does the string belong?”

- All strings belong!!! but with different degrees
- Degrees given as probabilities
- Whenever a sentence is spoken, uttered, it has entered the language!

- But in future some sentences have higher LIKELIHOOD of being seen compared to others

More likely vs. less likely

- "*Delhi is the capital of India*": likely to recur in the corpus

- Speech disfluency- "*is ..the…of… Delhi the capital of India?*"

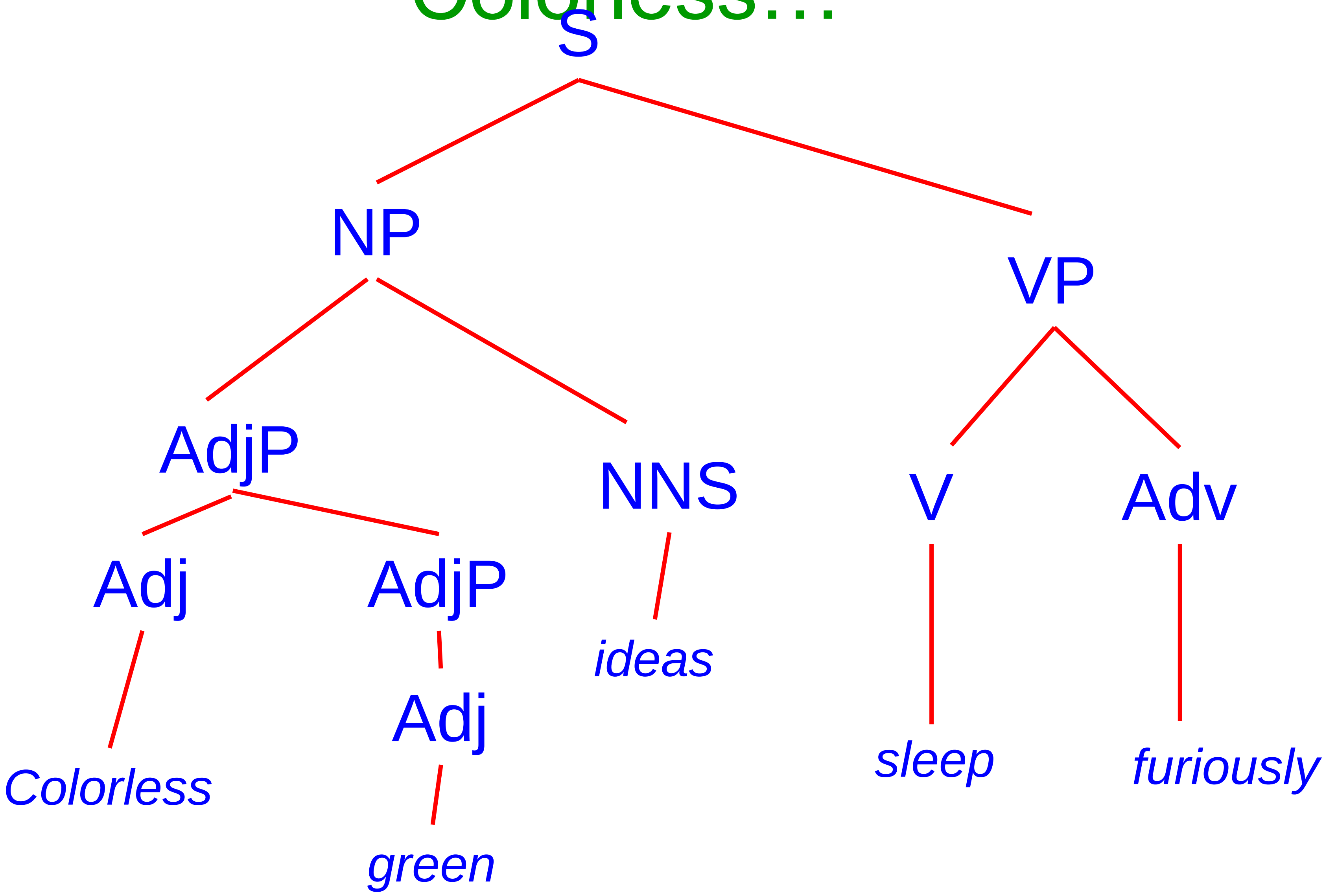
- "*is the of Delhi the capital of India?*": infrequent in the corpus

Syntactic oddity and semantic oddity

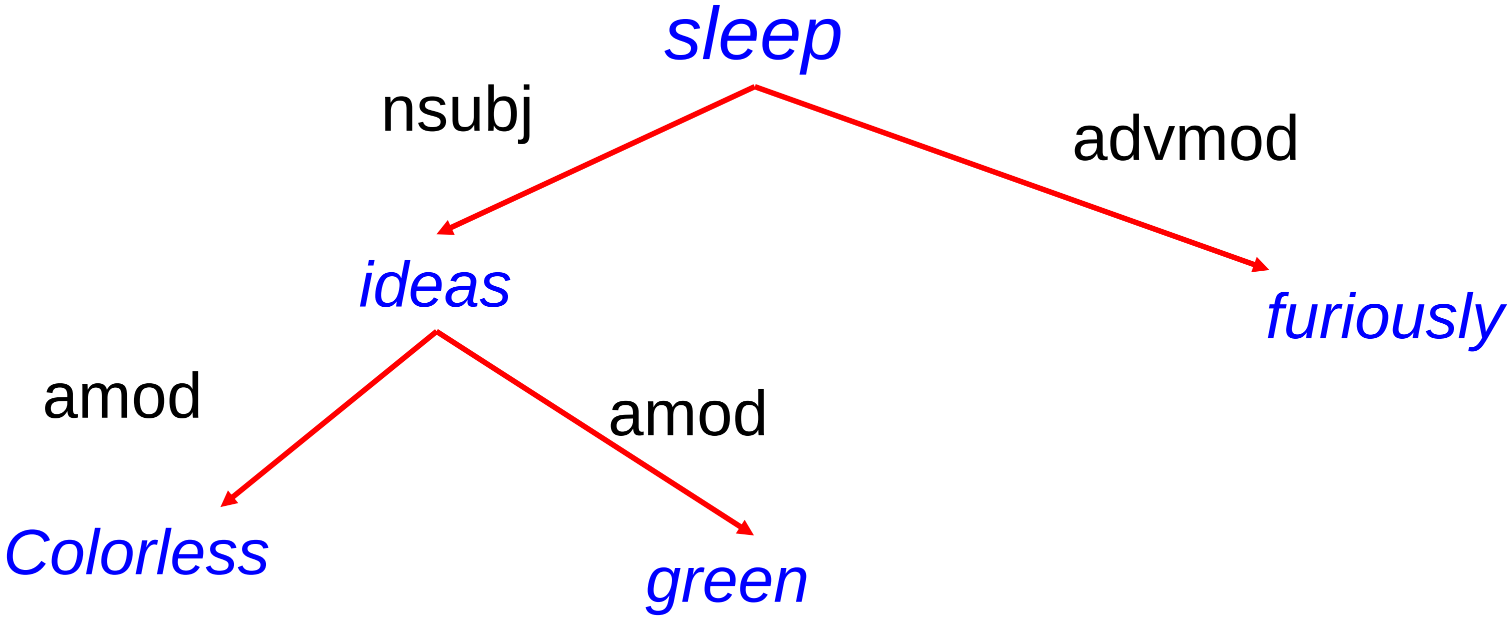
- Called *fluency* and *adequacy* in NLP: NLG situations like MT, QA, Summarization, Paraphrasing and so on

- Famous sentence due to Chomsky:
  - *? Colorless green Ideas sleep furiously*
- No problem with syntax; semantically odd

Constituency parse tree of “Colorless…”



Dependency parse tree of “Colorless…”



Grammar Rules

A segment of English

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

- DT → the 1.0
- NN → gunman 0.5
- NN → building 0.5
- VBD → sprayed 1.0
- NNS → bullets 1.0

Note: The numbers are probability values;
We will not worry about them now

CYK Parsing: Start with (0,1)

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To **From**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	DT						
1	-------						
2	-------	---------					
3	-------	---------	--------				
4	--------	---------	--------	---------			
5	--------	---------	--------	---------	---------		
6	--------	---------	--------	---------	---------	---------	

CYK: Keep filling diagonals

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To **From**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	DT						
1	-------	NN					
2	-------	---------					
3	-------	---------	--------				
4	--------	---------	--------	---------			
5	--------	---------	--------	---------	---------		
6	--------	---------	--------	---------	---------	---------	

CYK: Try getting higher level structures

0 *The* **1** *gunman* **2** *sprayed* **3** *the* **4** *building* **5** *with* **6** *bullets* **7**.

To **From**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	DT	NP					
1	-------	NN					
2	-------	---------					
3	-------	---------	--------				
4	--------	---------	--------	---------			
5	--------	---------	--------	---------	---------		
6	--------	---------	--------	---------	---------	---------	

CYK: Diagonal continues

0 *The* **1** *gunman* **2** *sprayed* **3** *the* **4** *building* **5** *with* **6** *bullets* **7.**

To **From**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	DT	NP					
1	-------	NN					
2	-------	---------	VBD				
3	-------	---------	--------				
4	--------	---------	--------	---------			
5	--------	---------	--------	---------	---------		
6	--------	---------	--------	---------	---------	---------	

CYK (cont…)

0 *The* **1** *gunman* **2** *sprayed* **3** *the* **4** *building* **5** *with* **6** *bullets* **7.**

To **From**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	DT	NP	--------				
1	-------	NN	--------				
2	-------	---------	VBD				
3	-------	---------	--------				
4	--------	---------	--------	---------			
5	--------	---------	--------	---------	---------		
6	--------	---------	--------	---------	---------	---------	

CYK (cont…)

0 *The* **1** *gunman* **2** *sprayed* **3** *the* **4** *building* **5** *with* **6** *bullets* **7.**

To From	1	2	3	4	5	6	7
0	DT	NP	--------				
1	-------	NN	--------				
2	-------	---------	VBD				
3	-------	---------	--------	DT			
4	--------	---------	--------	---------			
5	--------	---------	--------	---------	---------		
6	--------	---------	--------	---------	---------	---------	

CYK (cont…)

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To **From**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	DT	NP	--------	---------			
1	-------	NN	--------	---------			
2	-------	---------	VBD	---------			
3	-------	---------	--------	DT			
4	--------	---------	--------	---------	NN		
5	--------	---------	--------	---------	---------		
6	--------	---------	--------	---------	---------	---------	

CYK: starts filling the 5th column

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To From	1	2	3	4	5	6	7
0	DT	NP	--------	---------			
1	-------	NN	--------	---------			
2	-------	---------	VBD	---------			
3	-------	---------	--------	DT	NP		
4	--------	---------	--------	---------	NN		
5	--------	---------	--------	---------	---------		
6	--------	---------	--------	---------	---------	---------	

CYK (cont…)

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To From	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	**DT**	**NP**	**--------**	**---------**			
1	**-------**	**NN**	**--------**	**---------**			
2	**-------**	**---------**	**VBD**	**---------**	**VP**		
3	**-------**	**---------**	**--------**	**DT**	**NP**		
4	**--------**	**---------**	**--------**	**---------**	**NN**		
5	**--------**	**---------**	**--------**	**---------**	**---------**		
6	**--------**	**---------**	**--------**	**---------**	**---------**	**---------**	

CYK (cont…)

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To From	1	2	3	4	5	6	7
0	**DT**	**NP**	**--------**	**---------**			
1	**-------**	**NN**	**--------**	**---------**	**---------**		
2	**-------**	**---------**	**VBD**	**---------**	**VP**		
3	**-------**	**---------**	**--------**	**DT**	**NP**		
4	**--------**	**---------**	**--------**	**---------**	**NN**		
5	**--------**	**---------**	**--------**	**---------**	**---------**		
6	**--------**	**---------**	**--------**	**---------**	**---------**	**---------**	

CYK: S found, but NO termination!

0 *The* **1** *gunman* **2** *sprayed* **3** *the* **4** *building* **5** *with* **6** *bullets* **7**.

To **From**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	**DT**	**NP**	**--------**	**---------**	**S**		
1	**-------**	**NN**	**--------**	**---------**	**---------**		
2	**-------**	**---------**	**VBD**	**---------**	**VP**		
3	**-------**	**---------**	**--------**	**DT**	**NP**		
4	**--------**	**---------**	**--------**	**---------**	**NN**		
5	**--------**	**---------**	**--------**	**---------**	**---------**		
6	**--------**	**---------**	**--------**	**---------**	**---------**	**---------**	

CYK (cont…)

0 *The* **1** *gunman* **2** *sprayed* **3** *the* **4** *building* **5** *with* **6** *bullets* **7**.

To From	1	2	3	4	5	6	7
0	DT	NP	--------	---------	S		
1	-------	NN	--------	---------	---------		
2	-------	---------	VBD	---------	VP		
3	-------	---------	--------	DT	NP		
4	--------	---------	--------	---------	NN		
5	--------	---------	--------	---------	---------	P	
6	--------	---------	--------	---------	---------	---------	

CYK (cont…)

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To From	1	2	3	4	5	6	7
0	DT	NP	--------	---------	S	---------	
1	-------	NN	--------	---------	---------	---------	
2	-------	---------	VBD	---------	VP	---------	
3	-------	---------	--------	DT	NP	---------	
4	--------	---------	--------	---------	NN	---------	
5	--------	---------	--------	---------	---------	P	
6	--------	---------	--------	---------	---------	---------	

CYK: Control moves to last column

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To From	1	2	3	4	5	6	7
0	DT	NP	--------	---------	S	---------	
1	-------	NN	--------	---------	---------	---------	
2	-------	---------	VBD	---------	VP	---------	
3	-------	---------	--------	DT	NP	---------	
4	--------	---------	--------	---------	NN	---------	
5	--------	---------	--------	---------	---------	P	
6	--------	---------	--------	---------	---------	---------	NP NNS

CYK (cont…)

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To From	1	2	3	4	5	6	7
0	**DT**	**NP**	**--------**	**---------**	**S**	**---------**	
1	**-------**	**NN**	**--------**	**---------**	**---------**	**---------**	
2	**-------**	**---------**	**VBD**	**---------**	**VP**	**---------**	
3	**-------**	**---------**	**--------**	**DT**	**NP**	**---------**	
4	**--------**	**---------**	**--------**	**---------**	**NN**	**---------**	
5	**--------**	**---------**	**--------**	**---------**	**---------**	**P**	**PP**
6	**--------**	**---------**	**--------**	**---------**	**---------**	**---------**	**NP** **NNS**

CYK (cont…)

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To From	1	2	3	4	5	6	7
0	DT	NP	--------	---------	S	---------	
1	-------	NN	--------	---------	---------	---------	
2	-------	---------	VBD	---------	VP	---------	
3	-------	---------	--------	DT	NP	---------	NP
4	--------	---------	--------	---------	NN	---------	---------
5	--------	---------	--------	---------	---------	P	PP
6	--------	---------	--------	---------	---------	---------	NP NNS

CYK (cont…)

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To From	1	2	3	4	5	6	7
0	**DT**	**NP**	**--------**	**---------**	**S**	**---------**	
1	**-------**	**NN**	**--------**	**---------**	**---------**	**---------**	
2	**-------**	**---------**	**VBD**	**---------**	**VP**	**---------**	**VP**
3	**-------**	**---------**	**--------**	**DT**	**NP**	**---------**	**NP**
4	**--------**	**---------**	**--------**	**---------**	**NN**	**---------**	**---------**
5	**--------**	**---------**	**--------**	**---------**	**---------**	**P**	**PP**
6	**--------**	**---------**	**--------**	**---------**	**---------**	**---------**	**NP** **NNS**

CYK: filling the last column

0 *The* **1** *gunman* **2** *sprayed* **3** *the* **4** *building* **5** *with* **6** *bullets* **7.**

To From	1	2	3	4	5	6	7
0	DT	NP	--------	---------	S	---------	
1	-------	NN	--------	---------	---------	---------	---------
2	-------	---------	VBD	---------	VP	---------	VP
3	-------	---------	--------	DT	NP	---------	NP
4	--------	---------	--------	---------	NN	---------	---------
5	--------	---------	--------	---------	---------	P	PP
6	--------	---------	--------	---------	---------	---------	NP NNS

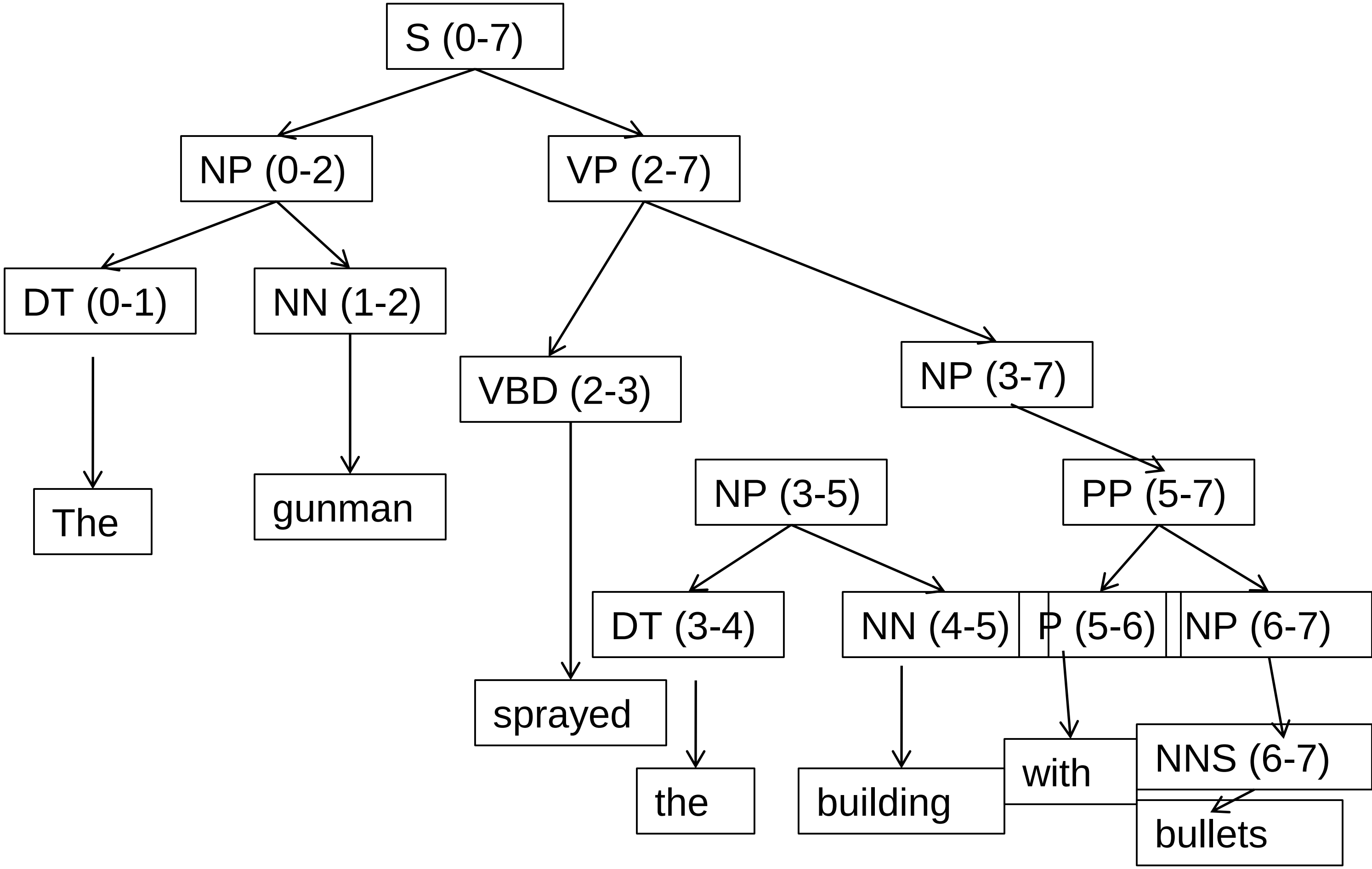
CYK: terminates with S in (0,7)

0 *The* 1 *gunman* 2 *sprayed* 3 *the* 4 *building* 5 *with* 6 *bullets* 7.

To From	1	2	3	4	5	6	7
0	DT	NP	--------	---------	S	---------	S
1	-------	NN	--------	---------	---------	---------	---------
2	-------	---------	VBD	---------	VP	---------	VP
3	-------	---------	--------	DT	NP	---------	NP
4	--------	---------	--------	---------	NN	---------	---------
5	--------	---------	--------	---------	---------	P	PP
6	--------	---------	--------	---------	---------	---------	NP NNS

CYK: Extracting the Parse Tree

- The parse tree is obtained by keeping back pointers.



Conversational AI

Attempts at Automation

- InstructGPT:
  - ***Command/Request/Order → Response***
- ChatGPT:
  - Carry out a ***conversation***
  - Respect context (state), personalization, quality and quantity and respond
    - Input: *I have been promoted*
    - Appropriate response: *I am delighted/congratulations/great* ..
    - Inappropriate: *why did they promote you?*

Gricean Maxims: Cooperative Principle in Converstaion (Wikipedia)

- **Quantity, Quality, Relation, and Manner**
- Paul Grice, philosopher of language
- “*Make your contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged*”.
- Captures the LINK between utterances

Maxim of Quantity (length and depth)

- *Be **informative,*** and submaxims are:
  - Make your contribution as informative as is required (for the current purposes of the exchange).
  - Do not make your contribution more informative than is required.
- Grice's analogy: "If you are assisting me to mend a car, I expect your contribution to be neither more nor less than is required. If, for example, at a particular stage I need four screws, I expect you to hand me four, rather than two or six."

Maxim of Quality (truth)

- Be ***Truthful***
- Submaxims:
  - Do not say what you believe is false.
  - Do not say that for which you lack adequate evidence
- Grice’s analogy: "I expect your contributions to be genuine and not spurious. If I need sugar as an ingredient in the cake you are assisting me to make, I do not expect you to hand me salt; if I need a spoon, I do not expect a trick spoon made of rubber."

Maxim of Relation (relevance)

- Information is ***relevant*** to the current exchange; therefore omitting any irrelevant information
- Grice’s analogy for this maxim: "I expect a partner’s contribution to be appropriate to the immediate needs at each stage of the transaction. If I am mixing ingredients for a cake, I do not expect to be handed a good book, or even an oven cloth (though this might be an appropriate contribution at a later stage)."

Maxim of Manner (clarity)

- Be ***perspicuous***
- Submaxims:
  - Avoid obscurity of expression — i.e., avoid language that is difficult to understand.
  - Avoid ambiguity — i.e., avoid language that can be interpreted in multiple ways.
  - Be brief — i.e., avoid unnecessary prolixity.
  - Be orderly — i.e., provide information in an order that makes sense, and makes it easy for the recipient to process it.

Examples of Gricean Maxims (1/3)

- ***Quantity:***
  - Question: *Do you know where the IITB library is*
  - *Answer: Yes I do, it is near the maths building*
  - *(violation): Yes, I do*
- ***Quality:***
  - Question: *Do you know where the IITB library is*
  - *Answer: Yes I do, it is near the maths building*
  - *(violation, if the answerer purposefully misleads): Yes, it is near the main gate*

Examples of Gricean Maxims (2/3)

- ***Relation:***
  - Question: *Do you know where the IITB library is*
  - *Answer: Yes I do, it is near the maths building*
  - *(violation): Mumbai is a big city*
  - *(violation of a different kind and degree): IITB campus is big! (so I do not know)*
  - *(flouting, but ok): I saw Ram enter the library from the maths buidling*
- ***Manner:***
  - Question: *Do you know where the IITB library is*
  - *Answer: Yes I do, it is near the maths building*

Examples of Gricean Maxims (3/3)

- ***Manner:***
  - Question: *Do you know where the IITB library is*
  - *Answer: Yes I do, it is near the maths building*
  - *(violation): Up yonder in the citadel of learning where polynominals are the bread and operators are the butter and where Hardy and Ramanujam permeate the atmosphere, thither will thee find the storehouse of what bibliophiles love* (maybe used for humorous effect)

What does ChatGP “think”? (1/2)

If for the question, "where is the library?" someone answers "Up yonder in the citadel of learning where polynominals are the bread and operators are the butter and where Hardy and Ramanujam permeate the atmosphere, thither will thee find the storehouse of what bibliophiles love", what would you say?

What does ChatGP “think”? (2/2)

Based on the answer given, it seems that the person is using figurative and poetic language to describe the location of the library. They are painting a vivid picture of the library being located in a place of high learning where mathematics and academic excellence are highly valued.

However, their answer may not be helpful if the person asking the question simply needs a straightforward answer to their question. In such a case, it might be more helpful to ask the person to provide a more direct and concrete answer to the question of where the library is located.