# Lecture 22

PCFG (Probabilistic CFG)-PART I

# CFG (Context Free Grammar)

1. S→VP NP
2. S→NP VP
3. VP→V
4. NP→Det Adj NP
5. VP→V NP
6. NP→Det N

N→dog|man|cat

Det→the|a

Adj→old|small

V→ate|cried

# Writing the derivation/construct a tree after the parsing procedure is complete

## (test sentence: the giraffe dreams)

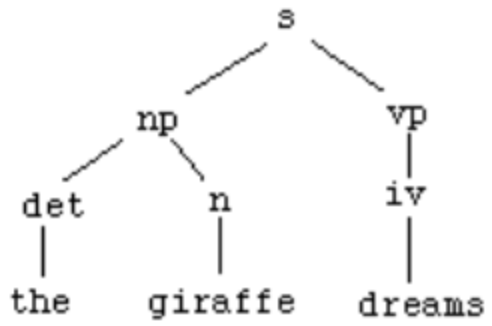
```
s    → np vp

np   → det n
vp   → tv np
     → iv

det  → the
     → a
     → an

n    → giraffe
     → apple

iv   → dreams
tv   → eats
     → dreams
```




- Derivation:

S=>**np** vp=>**det** n vp=>the **n** vp=> the giraffe **vp**=>the giraffe **iv**=>the giraffe dreams

# How to convert CFG to PCFG

- **PCFG: Probabilistic CFG**
- Attach a probability value to each production such that the sum of probabilities of productions with the same LHS $\sum_{same\_LHS} P(productions)$ is equal to 1
- Utility: 1) Find the best tree for a sentence (for ambiguous grammar)

  2) Find the best output sentence (in case of multiple candidates)
- How to find the probability (tree) or probability (sentence)?
- Write the derivation for the test sentence/construct the tree
- Multiply the probabilities of all the productions used for making the tree

$$P(sentence) = \prod_{tree} P(productions)$$

# PCFG (Probabilistic Context Free Grammar):

## Human: Who cried?

## Chatbot: the man cried/a dog cried

## (tree&deriv→P(sent)→winner sentence as per maximum probability)

- S→VP NP (0.1)
- S→NP VP (0.9)
- VP→V (0.6)
- NP→Det Adj NP (0.3)
- VP→V NP (0.4)
- NP→Det N (0.7)

- N→dog (0.2)
- N→man (0.6)
- N→cat (0.2)
- Det→the (0.7)
- Det→a (0.3)
- Adj→old (0.5)
- Adj→small (0.5)
- V→ate (0.6)
- V→cried (0.4)