PAGE NO.
DATE : / /

Stages of NLP

- Phonetics and Phonology
- Morphology
- Lexical Analysis
- Syntactic Analysis
- Semantic Analysis
- Pragmatics
- Discourse

1. Phonetics and Phonology

- It is concerned with processing of speech and sound
- concerns processing accents, pauses, amplitude, tone etc.

Challenges

Homophones: words which sound similar.
e.g. bank(finances) V/s bank(river)
bat V/s bat

Near Homophones: words which have very close sound
e.g. Peas, Peace
to, two, too
macrac, macac
write, right
knows, nose

Teacher's Signature.....

PAGE NO.
DATE: / /

Word-boundry: break word at proper place

aajaayenge → aaj aayenge
(will come today)

aa jaayenge
(will come)

I got [ua] plate → I got up late

I got a plate

Dissluency: ah, um, ahem etc...

- No meaning at all, speaker only use to organize her/his thoughts

2. Morphology

- It deals with word formation rules from root words

- Nouns: plural (boy - boys)
  gender marking (czar - czarina)

- Verbs: Tense (stretch - stretched)
  Aspect (sit - had sat)
  modality (khana - khaile)

- Detect all morphenes contained in large word string. Break the words and obtained properties

morphenes : A smallest meaningful unit in the grammar of a language

townhall : "town", "hall"
dogs : "dog", "s"
happiness : "happy", "ness"

- Language with rich morphology have the advantage of easier processing at higher stages of processing

3. Lexical Analysis
- Essentially refers to dictionary access and obtaining the properties of the word

e.g. dog
noun (lexical property)
take-s in plural (morph property)
animate (semantic property)
4-legged (semantic property)
carnivore (semantic property)

Why would we require such properties?

- During morphological stage dogs is divided into 'dog' & 's'. Root word with its properties is stored in dictionary.

Teacher's Signature

- So, during lexical we can identify that 'dog' is a noun and it takes 's' in plural and other details.

- So, during question-answering or conversation such properties are useful, is they are stored for each root word

Lexical Ambiguity

- Lexical ambiguity can occur when a word carries different sense i.e. having more than one meaning and the sentence in which it is contained can be interpreted differently depending on its correct sense

Examples

The word silver can be used as noun, an adjective or a verb

- She bagged two silver medals [Noun]
- She made a silver speech [Adjective]
- His worries had silvered his hair [verb]

How to resolve Lexical Ambiguity?

- Lexical ambiguity can be resolved by lexical category disambiguation i.e. part-of-speech tagging (POS)

- As many words may belong to more than one lexical category

- Part-of-speech tagging is the process of assigning a part-of-speech or lexical category such as noun, verb, pronoun, adverb, adjective etc. to each word in a sentence

PAGE NO.
DATE : / /

Lexical semantic Ambiguity

→ This type of lexical ambiguity, which occurs when a single word is associated with multiple senses

eg. bank, pen, fast, bat, cricket etc

The tank was full of water.
I saw a military tank

- The occurrence of tank in both sentences corresponds to the syntactic category noun, but their meanings are different.

→ Lexical semantic ambiguity resolved using word sense disambiguation (WSD)

Challenge

Lexical or word sense disambiguation

Lexical disambiguation

First step : Part of speech disambiguation
- Dog as a noun (animal)
- Dog as a verb (to pursue)

Sense disambiguation
- Dog (as animal)
- Dog (as a very detestable person)

- Select meaning per context
- Everytime word has only one sense, whenever it appears in document, so, identify that exact sense

Need word relationship in a context
e.g. The chair emphasised the need for adult education

Teacher's Signature.....

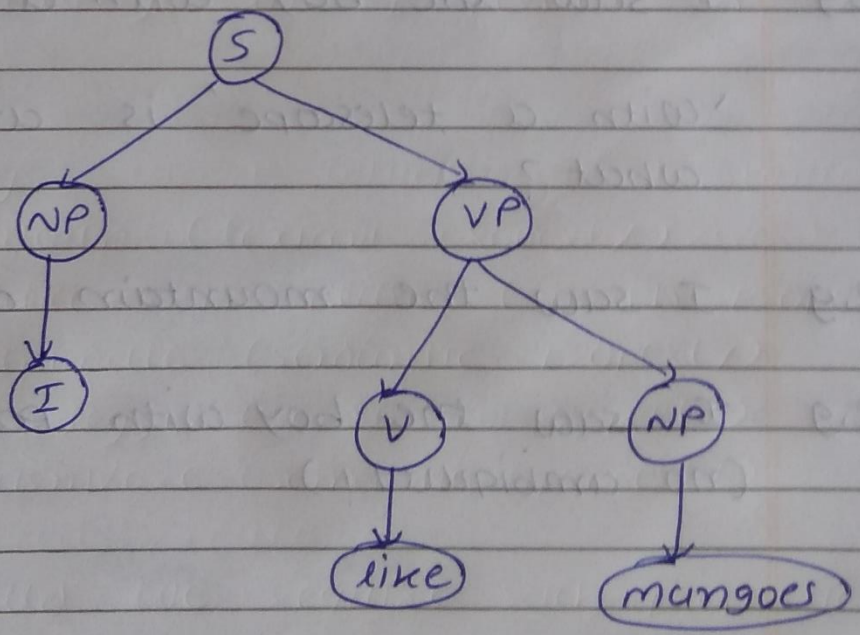
chair : - a piece of furniture
- person holding a chair

e.g watch what you want, when you want
(2 senses of watch)

watch : To take care (be careful)
watch : To see

- So, words are having multiple meaning, we need to identify that, at given context which meaning is suitable to use.

4. Syntactic Analysis
- After words are processed
- structure detection



Teacher's Signature.....

Challenges in syntatic processing : structural ambiguity

SCOPE

1. The old men and women were taken to safe locations

(old men and women) vs ((old men) and women)

2. No smoking areas will allow Hookas inside

Preposition phrase attachment
(possibility of multiple phrase attachments)

e.g. I saw the boy with a telescope

'with a telescope' is attached to what?

e.g. I saw the mountain with a telescope

e.g I saw the boy with pony-tail
(no ambiguity!)

PAGE NO.
DATE : / /

Semantic analysis

- It derives an absolute (dictionary definition) meaning from context. it determines the possible meaning of a sentence in context.

- The structures created by the syntactic analyzer are assigned meaning

- The structure for no such mapping is possible are rejected

I saw the boy with a pony tail

[only one meaning is found as others are not possible]

- Even after the syntax and the meaning of the individual words have been resolved, there are two ways of reading a sentence

a particular 'dog'

The dog is chasing the cat

The dog has been domesticated for 10k years

→ The specifies 'dog'

Teacher's Signature

PAGE NO.
DATE: / /

Pragmatics

- It derives knowledge from external commonsense information

- (cold) men) and women) as oppose to (old men and women) in "old men and women were taken to safe locations", since women both young and old were likely to taken to safe locations

Discourse

- The meaning of an individual sentence may depend on the sentence they precede it and may influence the meaning of the sentence they follow it

No smoking areas allow Hookas inside, but not sigals

John is sleeping. He is lazy

Teacher's Signature.....