Assignment-2

what is parsing? Explain various type of parsing.

The parsing algorithm in Natural Language processing (NLP) are used to analyze the syntactic structure of sentences and extract meaning from them. The main goal of parsing is to determine the grammatical relationship between the words in a sentences and to create tree like representation of syntactic structure

Following are types of parsing.

Recursive decent parsing.

Recursive decent parsing is top-down parsing technique that start from the root of parse and recursively apply grammer rule to generate parse tree for given sentence.

Earley parsing.

Earley parsing is a bottom-up parsing algorithm that builds the parse tree by using chart data structure to store partial parse result.

It is efficient and can handle grammar with arbitrary recursion and ambiguity.

CYK parsing.

CYK (Cocke-Younger-Kasami) parsing is the dynamic programming algorithm that can parse any context-free grammer in chomsky Normal form (CNF).

It work by building table of possible constituents for a given sentence and uses table

to generate the parse tree.

④ Shift Reduce Parsing-

1. Shift reduce parsing is bottom up parsing technique that operates on stack and input buffer of tokens.
2. It starts with an empty stack and gradually reduce the input sentence to parse tree.

⑤ Chart parsing-

1. Chart parsing is a parsing algorithm that uses a chart data structure to keep track of partial parse result.
2. It is generalization of Earley parsing and can handle arbitary context-free grammer.

2. Explain the terms:
a) Non-Structure b) Non-noun structure
c) Arguments and Adjuncts.

→ ① Noun Structure-

1. In NLP, nouns play critical role in structure and meaning.
2. Nouns are typically defined as word that refer to people, places, things, or ideas, and they can be used as subjects, objects, or in other grammatical role in sentence.

One important aspect of Noun structure in NLP is the concept of Noun phrases.

For example, in sentence
"The big brown dog chased the cat!"
the noun phrase "the big brown dog"

include the noun "dog" as its headword, along with the determiner "the" and the adjective "big" the "brown".

Non-Noun structure-

while noun are certainly important in natural language processing, non noun structure are also critical for understanding and analyzing language. Here some examples of non-noun struct. use in NLP.

Verbs-

Verbs are word that express the action or state of being.

They are often main element of in sentence and determine tense, aspect, and mood of sentence.

Adjectives-

Adjectives are word that modify nouns or pronouns by providing more information about their qualities and characteristics.

Adverbs-

Adverb are word that modify verb, adjective, or other adverbs by providing more information about the manner, time, place, or degree of the action or state.

Prepositions-

Prepositions are words that typically indicate a relationship between a noun or pronoun and other element in sentence. They are often use to express location, direction, time, or possession.

5. conjunctions.

1. conjunctions are words that connect words, phrases, or clauses in a sentences.
2. They often used to join elements of equal or unequal and can indicate logical relationship, such as addition, contrast, or consequence.

6. pronouns -

1. pronoun are words that are used in place of noun to avoid repetition or to provide coherence to sentence.
2. They can refer to a specific person or thing or they can be more general.

③ Adjuncts -

1. In linguistics and NLP, adjuncts are optional element in sentence that modify the meaning of sentence but are not essential to its basic structure.
2. They are some time referred as "adverbials" because they often function as adverb, indicating time, location, manner or other circumstances.
3. Inside-outside probabilities are used in probabilistic parsing to estimate the probability of specific parse tree.
4. These probabilities are calculated using a dynamic programming algorithm that computes the probability of the parse tree from the "inside" and "outside" probabilities of tree sub trees.

Arguments

In NLP, argument refers to mandatory part of sentence that are required to make the sentence grammatically correct.

Arguments are typically associated with specific syntactic role, such as object, and indirect object and necessary for the sentences to convey complete coherence meaning.

For example,

"John ate an apple"

"Jhon" is subject and "an apple" is the direct object In this case, both "Jhon" and "an apple" are arguments.

Explain Evidence for Deeper structure?

one piece of evidence for the existence of deeper structure in natural language processing is the phenomenon of long-distance dependencies.

Long-distance dependencies refers to syntactic relationshiphip between the word or phrase that are separated by intervening material.

These dependencies cannot be accounted for by simple linear structure, and requires more complex structure to be properly represented.

For example, consider the following sentence:

"Mary heard that the book she wanted to read had been checked out of the library".

The subject of main clause "Mary" is

separated from its corresponding verb ("heard") by an embedded clause ("that the book she wanted to read had been checked out of the library").

e. In order to properly understand the sentence, the parser must recognize nested structure of the sentence and identify the subject-verb relationship across the intervening material.

4. What do you mean by scope ambiguity and Attachment ambiguity resolution.

→ ① Scope Ambiguity.

1. Scope ambiguity is type of Semantic ambiguity when scope of the word or phrase in sentence is unclear & an unambiguous.
2. The scope of word or phrase refers to the range of elements in the sentence that it modifies or affects.
3. In order words, it refers set of expansions that word or phrase applies to.
4. For example, consider the sentences:
   "Every man loves a women".
5. This is sentence is ambiguous because it is unclear whether the phrase "a woman" is included in the scope of "every man" or not.
6. Depending on their interpretation, the sentence could mean:

1. For Every man, there exist a woman that he loves.
2. There exist a woman that every man

loves.

In the first interpretation, the scope of "every man" includes "a woman".

Attachment ambiguity Resolution-
Attachement ambiguity is a common issues in natural language processing where the intended attachment of word or phrase to a particular syntactic structure is unclear.
This often occur with prepositional phrases, relatives clauses, and other type of phrases that can be attached to different part of sentence.
One way to resolve the attachment ambiguity is to use probabilistic models that design assign probabilities to different attachment option based on the context of the sentence.
For example, in the sentences.

"The man saw the boy with the telescope".

the phrase "with the telescope" can be attachement to either "the man" or "the boy". resulting in two different interpretations:

The man saw the boy who had a telescope.
The man used a telescope to see the boy.