Semantic Interpretation:-

* Semantic Interpretation is the process of assigning meaning to a piece of language, such as word, phrase, sentence, or text.
* It is a fundamental task in Natural Language processing (NLP) and involves analyzing language in its context to infer its itended meaning.
* The goal of semantic Interpretation is to enable computers to understand the meaning of the natural language and to perform tasks based on that understanding.

Scanned with OKEN Scanner

① Structural Ambiguity:-

* Structural ambiguity can have a significant impact on semantic Interpretation because it can leads to multiple possible interpretations of a sentence, each with a different meaning.

* Resolving structural Ambiguity is therefore an important step in semantic Interpretation, as it helps to ensure that the correct meaning of a sentence is understood.

Ex: "The old man the boat".

The sentence is structurally ambiguous because it is unclear whether the man is old or boat is old.

① The man is old :- In this interpretation, "the old man" is a noun phrase.

② The boat is old :- In this interpretation, "the old" is an adjective modifying "boat".

* In this example, resolving the structural Ambiguity is crucial in semantic interpretation.

② Word sense :-

* Word sense is a crucial aspect of semantic Interpretation as it determines meaning of a word in a given context.

* Words often have different meaning, sense and understanding which sense of a word being used in particular context is essential for accurate semantic Interpretation

Scanned with OKEN Scanner

Ex: of word sense in semantic Interpretation

Sentence: I saw a bat in the park"

* The word "bat" has multiple senses/meanings including a flying mammal or a piece of sports equipment.
* In this context, the word "bat" likely refers to the animal sense, but without further context it is not entirely clear
* If the sentence were "I hit the bat with a stick" the sense of bat would be more apparent

③ Entity and event resolution :-

* Entity and event resolution are important aspects of semantic Interpretation which involves identifying and extracting information about entities (such as place, people and organization) and events (actions, states and processes) from text

Here is an example of entity resolution in S.I

Sentence: "John smith works at Google in New York City"

In this sentence, "John smith" is an entity (a person), "Google" is an entity (an organization) and "New York City" is an entity (a location/place).

* Entity resolution involves identifying & extracting these entities from the text & linking them to their corresponding types (Ex: person, organization, location)

Scanned with OKEN Scanner

Event resolution involves identifying and extracting information about events from text

Ex

sentence: "The dog chased the cat up the tree"

In this sentence, event is the action of the dog chasing the cat.

Event resolution involves identifying & extracting ~~these entities from the text~~ information about the action, including the actors (the dog & the cat) & the location (up the tree)

(4) Predicate-argument:-

Predicate argument structure is an important aspect of semantic Interpretation that involves identifying the relationships between the main verb (predicate) and its arguments (subjects, objects and other complements).

* Understanding the predicate-argument structure of a sentence is critical for accurate semantic Interpretation.

Here is an example for predicate-argument structure in S.I

sentence: "The cat chased the mouse"

In this sentence predicate is "chased" and its arguments are "cat (the subj) and "mouse" (the obj).

P-A-S can be represented as follows

Predicate: chased subject: cat object: Mouse

Scanned with OKEN Scanner

Ex2 "John gave Mary the book"

predicate : gave
subject : John
object : book (Direct obj)
obj : Mary (Indirect obj)

(5) Meaning representation :-

* Meaning representation is a process of representing meaning ~~of~~ a sentence or a text in a structured way.
* The goal of meaning representation is to capture the underlying meaning of a text, independent of the surface form of the language used to express it

Ex) -
"John gave mary the book"
In this, capturing the predicate-argument structure of the sentence and relationships b/w the various elements.

the logical form for this could be :

give(John, book, mary)

capturing the predicate (give) + arguments (John, book, mary) in structured way / formal way.

Scanned with OKEN Scanner

18/12/23
Monday

UNIT-III

System Paradigms in NLP

In computer science, a system paradigm is a fundamental approach or model for designing and implementing computer systems.

1. Imperative Paradigm
2. Functional Paradigm
3. Object-oriented Paradigm
4. Even-driven paradigm
5. concurrent paradigm

1. Imperative Paradigm:-

The Imperative paradigm is based on the imperative program where a program is sequence of statements that change the program state.

* The key constructs in imperative programming including variables, assignment, loops and conditions
* Most programming languages such as C, C++, Java, python are based on imperative paradigm

Ex: (a) Request

1. please send the report
2. Ask him to call

(b) command

1. Open the door
2. Turnoff the light

(c) Instruction

1. Follow the steps to install the software
2. Complete the form and submit it

* In NLP, understanding these statements involves identifying the verbs open, send, Turnoff, Ask, follow, complete and their associative objects or targets.

Scanned with OKEN Scanner

targets:- the door, the lights, the report, him, the steps, form.

* This is essential for application like virtual assistance and chart boards
* Enabling them and respond appropriately to user directives
* NLP systems use syntactic and semantic analysis to extract information from imperative statements

② Functional Paradigms:-

* In NLP, the functional paradigms involves treating language processing as series of functions and transformation applied to linguistic data.

Ex: ⓐ Tokenization

functional approach: tokens = tokenize(sentence)

Here, tokenize is the function that transforms a sentence into list of tokens (words)

ⓑ Pos tag:-

functional approach: pos_tags = pos_tag(tokens)

Here, pos tag function assign grammatic categories ex: verb, noun, To each token in a list

ⓒ NER: (Named Entity Recognition)

functional approach: entity = ner_extract(tokens)

Here, ner_extract is the function identifies and extracts named entities like person name, location name, from a list of tokens.

ⓓ Sentiment Analysis:-

functional approach: sentiment_analyze(text)

Here, sentiment_analyze is the function which evaluates the sentiment is positive, negative or neutral of a given text.

Scanned with OKEN Scanner

* Each NLP task is treated as distinct function that takes i/p and produces o/p
* This modular approach enhances code maintainability and allows for easier integration of different language processing tags

③ Object-oriented paradigm:-

In NLP, the object oriented paradigm can be applied to represent linguistic elements as objects with properties and method.

Examples

ⓐ Token object:-

class definition:- (self in instance of class)

python code:- __init__ constructor

```
class Token:-
    def __init__(self, word, pos_tag): self.word=word
        self.pos_tag = pos_tag

    def get_word(self): return self.word

    def get_pos_tag(self): return self_pos_tag
```

object

to define variable

Here Token object with attributes like word, pos_tag(pos)
methods allow retrieval of word and part-of-speech information

ⓑ Document object

class definition

```
class document:
    def __init__(self, text):
        self.text = text
        self.tokens = tokenize(text)
```

Scanned with OKEN Scanner

```
def get_tokens(self): return self.tokens
```

document object contains text
list of Token object obtained through Tokenization

The get_tokens method retrieves the tokenized representation.

③ NER

class definition

```
class NER:
    def __init__(self, entity_type, text):
        self.entity_type = entity_type
        self.text = text
    def get_entity_type(self):
        return self.entity_type
    def get_text(self):
        return self.text
```

Entity with attributes such as entity_type and text.
methods provide access to the type and text information.

19/12/23
Tuesday

Even-driven Paradigm :

In NLP, the even-driven paradigm involves designing system that respond to specific events or triggers rather than executing a predefined sequence of steps

For instance, in NLP applications events could be user inputs, changes in data or for specific linguistic patterns.

Scanned with OKEN Scanner

Ex: chatbots employee this paradigm must respond differently based on the user inputs, adapting its behavior to the context of conversation

User Input event:-

event: User i/p enters message

response: the chatbot analyze the message using NLP techniques to understand user intent and provides a relevant response

Context change event:

event:- User provides additional information changing ~~in the~~ context

response: chatbot adapts its response based on the updated context.

data update event:-

event: external data source updates information for that response

response: the chatbot receives the event updates its data and informs the user.

* chatbot can dynamically adjust its responses based on user interactions, context providing a more adaptive and user friendly experience (Ex: asking weather conditions (past, present, future) changing dynamically)

Scanned with OKEN Scanner

5. Concurrent Paradigm:-

It involves the simultaneous execution of multiple tasks or processes to improve efficiency and responsiveness.

(i) Parallel processing for document analysis:-

Analyzing a large collection of documen and dividing the document collection and proce each segments concurrently on multiple processors or threads

(ii) Multithreading for real time language translation

Real-time translation of incoming text messages.

Word sense

Resources:-

(i) Wordnet:-

Wordnet is one of the most widely used for wordsense information. It is a lexical database of english language that provides a structured inventory of words and their meanings

(ii) Extended wordnet:-

Some projects extended wordnet to cover additional languages or incorporate specialized domains.

Ex: euro wordnet (for European language)

Scanned with OKEN Scanner

(iii) babelnet

It integrates information from various sources including wordnet. It provides multilingual semantic network and includes links to other resources such as wikipedia

Systems:

Building a wordsense system involves implementing methods for wordsense disambiguition to determine the correct sense of a word in a given context

(i) supervise learning :-

Train a machine learning model using the corpus and extended features.

Common supervised learning algorithms are:-

1. Decision trees
2. support vector machines (SVM)
3. Neutral networks

(ii) Unsupervised learning:-

clustering algorithms are used to prove words with similar context into clusters

Corpus and annotations:-

starts with corpus of text that is annotated with wordsense.

Annotations serves as training data for systems.

Feature extraction :-

Extract relevant feature from the text to help identify wordsense which includes POS, surrounding words, syntactic patterns and more

Scanned with OKEN Scanner

Software:-

(i) NLTK

(ii) SPACY

(iii) SENSE CLUSTERS

(iv) Wordsense Disambiguition model

Scanned with OKEN Scanner