

ASSIGNMENT 1 Part B

Name- Geeta Pramod Sakpal

Roll No-54

Branch- IT (BE)

sub- AI

Sem- VII

DOP	DOC	marks	sign

Q. 2. Explain various element of cognitive system.

→ 1. A way of interpreting input- A cognitive system needs to answer a question for provide a result based on an input. The input might be a search term, text, phrase, a query asked in natural language, or it may be a response to an action of some sort.

The first thing a system needs to do is understand the context of signal examples- location, speed of motion, such context info will enable the system to narrow down the potential responses to those that are more appropriate. Cognitive system needs to start somewhere they need to know or expect something about the user to interpret the input.

2. A body of content/information that supports the decision-

- The purpose of cognitive system is to help humans make choices and solve problems.

But the system does not make up the answer. Even synthesis of new knowledge is based on foundation knowledge. IBM watson for example ingests many structured repositories of information: dictionaries,

news articles, and databases, taxonomies, and ontologies. These sources provide the information needed to respond to questions, forming the corpus of information that Watson draws upon.

3. A way of processing the signal against the content/info corpus- This is where machine learning for eg. comes into play. ML has for long been applied to categorization and classification approaches and advanced text analytics. The processing might be in form of query/matching algorithm or may involve other mechanisms to interpret the query, transform it, reduce ambiguity, derive syntax, define word sense, deduce logical relationships or otherwise parse/process the signal against the corpus.

The key here to iteratively improve the system performance overtime by approximating an output & using that as an input for next round of processing. In some cases, incorrect answer might be input for next time the system encounters the problem or questions.

Q.3. Write note on language model.

- 1. The goal of language model is to compute a probability of a token and are useful in many different Natural language processing applications.
2. LM actually a grammar of a languages as it gives the probability of word that will follow.
3. For example, they have been used in Twitter Bots for 'robot' accounts to form their own sentences.

Language Model Definitions-

1. In case of probabilistic language modeling the probability of a sentence as sequence of words is calculated.

$$P(W) = P(W_1, W_2, W_3, \dots, W_n)$$

2. It can also be used to find the probability of the next word in the sentence.

$$P(W_5 / W_1, W_2, W_3, W_4)$$

3. A model that computes either of these is called a language model.

4. There are various language models in available in practice.

Following are few of them-

- A) Methods using the Markov assumption
- i. The probability of the next word can be estimated given only the previous k number of words.

For example, if $k=1$:
 $P(\text{transparent} / \text{its water is so}) \approx P(\text{transparent} / \text{so})$

or if $k=2$:

$P(\text{transparent} / \text{its water is so}) \approx P(\text{transparent} / \text{is so})$

the Markov Assumption, $k=i$:

$$P(W_i / W_1, W_2, \dots, W_{i-1}) \approx P(W_i / W_{i-k}, \dots, W_{i-1})$$

B> N-gram Models-

From the Markov Assumption, we can formally define N-gram models where $k=n-1$ as follows-

$$P(W_i / W_1, W_2, \dots, W_{i-1}) \approx P(W_i / W_{i-(n-1)}, \dots, W_{i-1})$$

The simplest version of this are defined as unigram model ($k=1$) and the Bigram model ($k=2$).

c> Unigram Model ($k=1$):

$$P(W_1, W_2, \dots, W_n) \approx \prod_i P(W_i)$$

d> Bigram Model ($k=2$):

$$P(W_i / W_1, W_2, \dots, W_{i-1}) \approx P(W_i / W_{i-1})$$

following is maximum Likelihood Estimate model to estimating Bigram Probabilities

$$(W_i / W_{i-1}) = \frac{\text{Count}(W_{i-1}, \dots, W_i)}{\text{Count}(W_{i-1})}$$

5. Language Modeling is one of most important parts of modern Natural language processing. There are many sort of application for language modeling, like: spell correction, speech Recognition, machine Translation, Question Answering, summarization, sentiment analysis, etc. All these tasks require use of language model. Language model is supposed to represent the text to a form understandable from the machine point of view.

6. Moreover, language modelling must also consider the correlated ordering of tokens. As every language is based on some grammar where order has a lot of influence on the meaning of a text.

Q. 4. Write a note on machine translation.

- Machine Translation is classic test of language understand. It consists of both language analysis & generation.
- many machine translation system have huge commercial use following are few of the examples.
- Google translate goes through 100 billion words per day.
- PBay uses machine translation techniques to enable cross border trade & connect buyers/sellers around globe.
- Facebook uses to translate text in posts & comments automatically in order to break language barriers.
- System became the first software provider to launch a neural machine translation engine in more than 30 languages in 2016.
- Microsoft brings AI powered translation to end users & developers on android, ios & Amazon fire whether or not they have access to internet.
- In a traditional machine translation system, parallel corpus a collection of trees is used to each of which is translated into one or more other languages than the original. For example, given the source language. e.g. French & the target language. e.g. English, multiple

statistical models needs to be build, including a probabilistic formulation using the rule, translation model $P(F|e)$ trained on parallel corpus & a language model $P(e)$ trained on the english corpus.

- It is obvious that, this approach skips hundreds of important details, requires a lot of human feature engineering & is overall a complex system.

Q.5. Explain the following terms-

a] Phonology-

- It is the study of organizing sounds systematically in an NLP (Natural Language Processing) system.

b] Morphology-

- It is a study of construction of words from primitive meaningful units.

c] Lexical Analysis-

- Lexicon is the words & phrases in language. Lexical analysis deals with their recognition & identification of structure of sentences. It divides the paragraphs in sentences, phrases & words.

d] Syntactic Analysis-

- In this, the sequences are parsed as noun, verbs, adjective & other parts of sentences. In this phrase the grammar of the sentence is analyzed in order to get relationship among different words in sentences. For example, "mango eats me" will be rejected by analyzer.

e] word sense disambiguation.

- While using words that ~~have~~ have more than one meaning we have to select the meaning which makes the most sense in context. For example, we are typically given a list of words associated words sends (e.g. from a dictionary or from an online resource such as word net).