Date
Page

Assignment 1 Part B

Name :- Namra Javid Palte

Roll no :- 41

Branch :- B.E-I.T.

Sub :- A.I.

Sem :- VII

DOP	DOC	marks	Sign

Q.1) Explain PEAS description for WUMPUS World

→ 1) Performance measure
→ +100 for grabing the gold and coming back to the starting position.
→ -200 if the player (agent) is killed
→ -1 per action
→ -10 for using the arrow

2) Environment
→ Empty Rooms
→ Rooms with WUMPUS.
→ Rooms neighbouring to WUMPUS which are Smelly.
→ Rooms with Bottomless pits
→ Rooms neighbouring to bottomless pits which are breezy
→ Room with gold which is glittery.
→ Arrow to shoot the WUMPUS.

3) Sensor (assuming a robotic agent)
→ Camera to get the view.
→ odour sensor to smell the stench
→ Audio sensor to listen to the scream and bump.

4) Effectors (assuming a robotic agent)
→ motor to move left, right
→ Robot arm to grab the gold.
→ Robot mechanism to shoot the arrow

The WUMPUS world agent has following characteristics.

1) Fully observable
2) Deterministic
3) Episodic
4) Static
5) Discrete
6) Single agent.

Q2) Explain various element of cognitive system.

→ 1) A way of interpreting input :- A cognitive system needs to answer a question or 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 systems need 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 artticles and databases, taxonomies and ontologies. These sources provide the information needed to respond to questions, forming the corpus of information that watson draws ø 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 over time by approxi-mating an output f using that as an input for next round of processing. In some cases, incorrect answer might be p input for next time the system

encounters the problem or question.

Q3) 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) Language model (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 probablistic language modeling the probability of a sentence as sequence of words is calculated:

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

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:

1) The Probability of the next word can be estimated given only the previous k number of words.

for example, if k=1:

P(transparent | its water is so) ≈ P(transparent | so)

or if k=2

P(transparent | its water is so) ≈ P(transparet | isso)

2) following is the general equation for the Markov Assumption, k=i:

$P(W_i | W_1 W_2 \ldots W_{i-1}) = P(W_i | W_{i-k} \ldots W_{i-1})$

B) N-gram Models :-

from the markov Assumption, we can formally define N-gram models where k=n-1 as follow:

$P(W_i | W_1 W_2 \cdots W_{i-1}) \approx P(W_i | W_{i-(n-1)} \cdots W_{i-1})$

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

C) Unigram Model (k=1):

$P(W_1 W_2 \ldots W_n) \approx \prod_i P(W_i)$

D) Bigram Model (k=2):

$P(W_i | W_1 W_2 \ldots W_{i-1}) \approx P(W_i | W_{i-1})$

following is maximum Likelihood Estimate model to Estimating Bigram Probabilities:

$$(W_i | W_{i-1}) = \frac{\text{count}(W_{i-1} \ldots W_i)}{\text{count}(W_{i-1})}$$

Example :-

A corpus with following 3 sentence, let's find out probability tha "I" start the sentence. Here "<s" and "/s" denote the start and end of sentence respectively.

<s I am Sam /s>

<s Sam I am /s>

<s I do not like green eggs and ham /s>

- Therefore, we have

$$P(I \mid <s) = \frac{count(<s, I)}{count(<s)} = \frac{2}{3}$$

→ "I" appeared as the first word in two sentence.

5) Language Modelling is one of most important parts of modern Natural Language Processing. There are many sort of application for language modelling, 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 note on Machine Translation.

→ 1) Machine translation is the classic test of language understanding. It consists of both language analysis and language generation. Many machine translation systems have huge commercial use. following are the few example:

a) Google Translate goes through 100 billion words per day.

b) eBay uses Machine Translation techniques to enable cross-border trade and connect buyers and sellers around the world.

c) facebook uses machine translation to translate text in posts and comments automatically, in order to break language barriers and allow people around the world to communicate with each other.

2) In a traditional Machine Translation System, Parallel corpus a collection of texts is used each of which, is translated into one or more language than the original.

3) It is obvious that, this approach skips hundreds of important details, requires a lot of human feature engineering consists of many different and Independent machine learning problems, and overall is a very complex system.

A) Neural Machine Translation (NMT):-

Standard Neural Machine Translation is an end-to-end neural network where

the source sentence is encoded by a Recurrent Neural Network (RNN) called encoder, and the target words are predicted using another RNN known as decoder.

- following are the features of NMT:

1) End-to-end training: All parameters in NMT are simultaneously optimized to minimize a loss function on the network's output.

2) Distributed representation: NMT has a better exploitation of word and phrase similarities. Hence, it forms a robust translator.

3) Better exploration of context: NMT can use a much bigger context for both source and partial target text in order to translate more accurately.

4) More fluent text generation: Deep learning text generation is of much higher quality than the parallel corpus way.

B) Long Short-Term Memory (LSTM):-

LSTM works as a solution to vanishing gradient problem by introducing gates and an explicitly defined memory cell.

Each neuron has a memory cell and three gates: input, output and forget.

The function of these gates is to safeguard the information by stopping or allowing the flow of it.

1) The input gate determines how much of the information from the previous layer get stored in cell.

2) The output layer takes the job on the other

end & determines how much of the next layer get to know about the state of this cell.

3) The forget gate seem like an odd inclusion at first but sometime it's good to forget. If it's learning a book and a new chapter begins, it may be necessary for the network to forget some characters from the previous ~~chap~~ chapter.

c) Gated Recurrent Units (GRU):

They are slight variation on LSTMs and are extensions of Neural Machine Translation. They have one less gate and are wired slightly differently. GRU has an update gate instead of an input, output, and a forget gate. This update gate determines how much information to be kept from the last state and how much information to forget from the previous layer.

Q5) Explain following terms

a) Phonology :- It is the study of the speech sounds of a particular language. The origin of the word can be traced to Greek language, where 'phone' means sound or voice. Phonetics, a subdivision of phonology is the study of speech sounds of human language from the perspective of their production, perception or their physical properties. IPA (International Phonetic Alphabet) is a tool that represents human sounds in a regular way while studying phonology. In IPA, every written symbol represents one & only one speech sound & vice-versa.

b) Morphology :- It is a branch of linguistics that focuses on the way in which words are formed from morphemes. There are two types of morphemes namely lexical morpheme and grammatical morpheme. Steming is the simplest form of morphological processing.

c) Lexical analysis :- Lexicon is the words and phrases in language. Lexicon analysis deals with the recognition and identification of structure of the sentences. It divides the paragraphs in sentences, phrases and words.

d) Syntactic analysis :- In syntactic analysis the Sentence are parsed as noun, verbs, adjectives, and other parts of sentences. In this phase the grammar of sentence is analyzed in order to get the relationships among different words in the sentence. for example, "mango eats me" will be rejected by syntactic analyzer.

e) Word Sence Disambiguation :- Word sense disambi-guation, in natural ~~na~~ language processing may be defined as the ability to determine which meaning of word is activated by the use of word in particular context, a process which appears to be largely unconscious in people.