KGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCEKGCE

Assignment 1 Part B

NAME: Ajay Kumar Prasad

CLASS: B.E. IT

ROLL NO. 49

SUBJECT: IS Lab

DOA	DOC	MARKS	SIGN

Q.1) Explain PEAS description for WUMPUS world

Ans:

1) Performance Measure
→ +100 for grabbing the gold and returning to the starting point
→ -200 if the player is killed
→ -1 per action
→ -10 for using the arrow

2) Environment
→ Empty rooms (cells)
→ Rooms with WUMPUS
→ Rooms neighbouring to WUMPUS which are smelly
→ Rooms with pits
→ Rooms neighbouring to pits which are breezy
→ Room with gold which is glittery
→ Arrow to shoot WUMPUS

3) Sensor (assuming robotic agent)
→ Camera to get the view
→ Odour sensor to sense the stench
→ Audio sensor to sense the scream

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

The WUMPUS world agent has the following characteristics:

(i) Fully observable
(ii) Deterministic
(iii) Episodic
(iv) Static
(v) Discrete
(vi) Single agent

Q.2) Explain the various elements of cognitive system.

Ans:

(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 may be a response to an action of some sort.

→ From The system needs to understand the context of signal like speed of motion, location etc.

→ Such context inp 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 interprete the input.

B) 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 instance, ingests many 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.

C) Way of processing the signal against the content corpus.

→ This is where machine learning comes into play
→ ML has for long been applied to categorization and classification approach and advanced text analytics.
→ The processing might be in the 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 process the signal against the corpus.
→ The key here is to iteratively improve the system performance over time by approximating an output of 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.

Q-3) Write a note on Language Model.

Ans:

→ The goal of Language Model is to compute a probability of a token and are useful in many different NLP applications.

→ Language Model is actually a grammar of ~~as~~ languages as it gives the probability of word that will follow.

→ For instance, they have been used in Twitter bots for robot accounts to form their own ~~sys~~ sentences.

Language Model Definitions

→ In case of probabilistic language modelling the probability of a sentence as sequence of words is calculated.

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

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

$$P(W_5 \mid W_1, W_2, W_3, W_4)$$

→ A model that computes either of these is called a Language Model.

→ There are various Language Models available in practise.

A) Methods using the Markov assumption

→ The probability of the next word can be estimated given only the previous k number of words

for instance if $k=1$

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

or if $k=2$

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

→ Following is the general equation for the Markov assumption $k = i$:

$$P(w_i \mid w_1, w_2, \ldots w_{i-1}) = P(w_i \mid w_{i-k} \cdots w_{i-1})$$

(B) N-gram models

→ From the Markov assumption we can formally define N-gram models where $k = n-1$ as follows:

$$P(w_i \mid w_1, w_2 \ldots w_{i-1}) \approx P(w_i \mid w_{i-(n-1)} \ldots w_{i-1})$$

→ The simplest version of this is 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 $k = 2$

$$P(w_i \mid w_1 w_2 \ldots w_{i-1}) \approx P(w_i \mid w_{i-1})$$

Following is the maximum likelihood estimate model to estimating bigram probabilities

$$(w_i \mid w_{i-1}) = \frac{\text{count}(w_{i-1} \ldots w_i)}{\text{count}(w_{i-1})}$$

Example

→ A corpus with following 3 sentences. Let's find out probability that 'I' starts 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>

∴ We have

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

→ 'I' appeared as the first word in two sentences

→ Language modeling is one of the most important part of modern NLP.

→ There are many sort of application for language modeling, like spelling 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.

→ Moreover, the 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.

Q4) Write a note on Machine Translation.

Ans:

→ Machine translation is the classic test of language understanding.

→ It consists of both language & analysis and language ~~gen~~ generation.

→ Many machine translation systems have huge commercial use.

→ Following are some examples:

(a) Google Translate goes through 100 billion words a day

(b) eBay uses machine translation techniques to enable cross-border trade and connect ~~by~~ buyers and sellers around the world.

(c) Facebook uses machine translation to translate text in posts and comments automatically

→ In a traditional machine translation system, parallel corpus is ~~a~~ is a collection of text, each of which is translated into one or more language than the original.

→ It is obvious that this approach skips hundreds of important details, requires a lot of human feature engineering consists of many different and independent ML problems and overall is a complex system.

(A) Neural Machine Translation (NMT)

→ Standard Neural Machine Translation is an end-to-end neural network when the source sentence is encoded by a Recurrent Neural Network (RNN) called

called encoder and the target words are predicted using another RNN known as decoder.

→ Following are the features of NMT:

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

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

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

(iv) 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 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 the flow of it.

→ The input gate determines how much of the information from the previous layer gets stored in the cell.

→ 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.

→ The forget gate seem like an odd inclusion at first but sometimes it's good to forget.

→ If its learning a book and a new chapter begins, it may be necessary for the network to forget some characters from the previous chapters.

(c) Gated Recurrent Units (GRU)

→ They are slight variations on LSTMs and are extensions of NMT.

→ 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 info to forget from the previous layer.

Q-5) Explain the 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 like studying phonology.
→ In IPA every written symbol represents one speech sound and 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.
→ Stemming is the simplest form of morphological processing.

(C) Lexical analysis:

→ Lexicon is the words and phrase in language.
→ Lexical analysis deals with the recognition and identification of structure of the sentences.
→ It divides the paragraphs in sentences, phrases and words.

(D) Syntactical analysis
→ In Syntactical analysis the sentences are passed as noun, verb, and other parts of sentences.

→ In this phase the grammar of phrase is analyzed in order to get the relationships among different words in the sentence.

→ For instance, 'Mango eats me' will be rejected by syntactic analyzer.

(ii) Word Sense Disambiguation

→ Word Sense Disambiguation in NLP may be defined as the ability to determine which meaning of the word is activated by the use of word in particular context. a process which appears to be largely unconscious in people.