

Assignment .1 Part B

Name : Ruchita Suresh Rasal

Roll No. : 53

Class : B.F.I.T.

Sem : VII

Subject : Is lab

DOP	Doc	marks	Sign

Assignment 1 Part B

Q.1 Explain PFA's descriptions for WUMPUS World



The PFA's descriptors for WUMPUS world are as follows:-

1) Performance Measure

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

2) Environment

- Empty Rooms
- Room 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 glinty
- Arrow to shoot to shoot the WUMPUS

3) Sensors (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
- Audio mechanism to shoot the arrow

The humpus world agent has following characteristics

1. Fully observable
2. Deterministic
3. Episodic
4. static
5. Discrete
6. single agent

2) Explain various elements of cognitive system

- ① Cognitive computing is new type of computing with goal of more accurate models of how human brain/mind senses, resources & responds to stimulus.
- ② Cognitive computing applications link data analysis and adaptive page display i.e. Adaptive User interface, to adjust content for particular type of audience.

Following are some of features of cognitive system

1. Interactive - They may interact easily with user so that those users can define their needs comfortably. They may also interact with other with other processors, devices and cloud services, as well as with people
2. Adaptive - They may be engineered to feed on dynamic data in real time. They may learn as information changes and as goals and requirements evolve. They may resolve ambiguity and tolerate tips unpredictability.
3. Contextual - They may understand, identify and extract contextual elements such as meaning, syntax, time, location, appropriate domain, regulations, user's profile, process, task and goal. They may draw on

multiple sources of information, including both structured & unstructured digital information, as well as sensory inputs like visual, gestural, auditory or sensor provided.

4. Iterative & stateful - They may aid in defining by a problem by asking questions or finding additional source of input if a problem statement is ambiguous or incomplete. They may "remember" previous interactions in a process and return information that is suitable for the specific application at that point in time.

3) Write note on Language Model.

- ① The goal of a language model is to compute a probability of a token and are useful in many different Natural language applications.
- ② Language model (LM) actually a grammar of a language as it gives probability of word that will follow
- ③ Language Model Definition:

- In case of Probabilistic modeling the probability of a sentence of words is calculated:

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

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

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

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

- There are various language models available.

1. Methods using the Markov assumption:

- Markov Property : A process which is stochastic in nature, is said to have Markov property if conditional probability distribution of future states of process depends only upon the present state, not on sequence of events that happened in past.
- For example, if $k=1$:

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

2. N-gram Models :

- From markov Assumption, define N-gram models where $k=n-1$ as following.

$$P(w_i | w_1, w_2, \dots, w_{i-1}) = P(w_i | w_{i-(n-1)}, \dots, w_{i-1})$$

3. Unigram Model ($k=1$) :

$$P(w_1, w_2, \dots, w_n) \approx \prod_i P(w_i)$$

4. 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 estimating Bigram Probabilities ;

$$(w_i | w_{i-1}) = \frac{\text{Count}(w_{i-1}, \dots, w_i)}{\text{Count}(w_{i-1})}$$

④ Language modeling is one of most important started in our corpus, "i" appeared as first word in two sentences of applications for language modeling.

⑤ Language modeling must also consider correlated ordering of tokens. As every language is based on some grammar, where order has a lot of influence on meaning of a text

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 system also have huge commercial use.
- ② In traditional Machine Translation system, parallel corpus a collection of text is used each of which is translated into one or more language than original.
- ③ It is obvious that, this approach skips hundreds of important details, requires lot of human features engineering consist of many different & independent machine learning problems.

A] Neural Machine Translation (NMT)

- The above process is modelled through one big artificial neural n/w, known as Recurrent Neural Network (RNN). It has connection b/w passes & connections through time.
1. Features of NMT are as follows :- End-to-end training, distributed representations, better exploration of context, more fluent text generation,
 2. long short-term memory (LSTM) - LSTM works as a solution to vanishing gradient problem by introducing gates & an explicitly defined memory cell.
 3. Gate Recurrent Units (GRU) : They are slightly variation on LSTMs and are extensions of Neural Machine Translation. They have one less gate and are wired slightly differently.

5) Explain following terms:

a) Phonology

- It is study of organizing sound systematically.
- It is Natural Language Generation (NLG)

terminology

b) Morphology

- It is a study of construction of words from primitive meaningful units
- It is Natural Language Generation (NLG)

terminology.

c) Lexical analysis

- It involves identifying and analyzing the structure of words. Lexicon of a language means the collection of words and phrases in language.
- lexical analysis is dividing the whole chunk of txt into paragraphs, sentences and words

d) Syntactic analysis

- It involves analysis of words in the sentence for grammar and arranging words in a manner that shows the relationship among the words
- The sentence such as "The school goes to boy" is rejected by English syntactic analyzer

e) Word Sense Disambiguation

- While using words that have more than one meaning ; we have to select the meaning which makes the most sense in context.

Date _____
Page _____

- For this problem, we are typically given a list of words and associated word senses, e.g. from a dictionary or from an online resources such as WordNet