Page No. :
Date. :

Topic : Assignment 1 Part B.

Name :- Pranay Arvind Malte

Roll No :- 56

Subject :- AI

STD / Branch :- B.E / I.T

For Educational Use Only

Page No. :
Date. :
Topic :

Q1 The Wumpus World's agent is an example of a knowledge-based agent that represents knowledge representation, reasoning and planning.

Knowledge-Based agent links general knowledge with current ~~pre~~percepts to infer hidden characters of current state before selecting actions.

PEAS represents performance measures, environment, Actuators & sensors.
The PEAS description helps in grouping the agents.

PEAS Description for the Wumpus World problem:-

1. Performance measures:-

- Agents gets the gold and return back safe = + 1000 points
- Agent dies = -1000 points.
- Each move of the agent = -1 point
- Agent uses the arrow = -10 points.

2. Enviroment:-
- A cave with 16 (4x4) rooms
- Rooms adjacent to the Wumpus are stinking.
- Rooms adjacent to the pit are breezy.
- The room with the gold gliters.

Topic :

3. Actuators :-

Devices that allow the agent to perform the following actions in the environment.

- Move forward
- Turn right
- Turn left
- Shoot
- Grab
- Release

4. Sensors :-

Devices which helps the agent in sensing the following from the environment.

- Breeze
- Stench
- Glitter
- Scream
- Bump

Wumpus World Characterization :-

Partially Observable :- knows only the local perceptions.
Deterministic :- Outcome is precisely specified.
Sequential :- subsequent level of actions performed.
Static :- Wumpus, pits are immobile

Q 2 Cognitive Computing :-

Cognite computing is a new type of computing with the goal of more accurate models of how the human brain / mind senses, reasons & responds to stimules.

The term cognitive computing is used to refer to new hardware and/or software that mimic the functioning of the human brain thereby improving human decision-making.

Cognitive Computing applications link data analysis and adaptive page displays.

<u>features</u> of <u>Cognitive</u> systems:-

1) <u>Interactive</u>:- They may interact easily with users so that those users can define their needs comfortably. They may also interact with other processors, devices & cloud devices & cloud services as well as with people.

2) <u>Adaptive</u>:- 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 unpredictability

Topic :

Page No.
Date. :

3. Contextual :-

They may understand, identity & extract contextual elements such as meaning syntax, time location, appropriate domain, regulations, user's profile, process, task & goal.

They may draw on multiple sources of information, including both structured digital information, as well as sensory inputs like visual, gestral, auditory, or sensor - provided.

4. Iteractive & stateful :-

They may aid in defining a problem by asking questions of finding additional source of input of a problem statement is ambiguous or incomplete.

They may "remember" precious interactions in a process and return information that is suitable for the specific application at that point in time.

Topic : Page No. : Date. :

Design principles for Cognitive Systems:-

1. Standardize:-
Many errors are caused by inconsistencies in how things work, whether how information is displayed or how controls are activated.
To prevent mistakes, a general rule is to insure that similar devices work the same way.
Aggreeing upon a standard helps prevent errors.

2. Use stereotypes:-
A sterotype is commonly held ~~[illegible]~~ expectation of what people think is supposed to happen when they recognize a signal or active a control.

3. Match controls to equipment layout.

4. Simplify presentation of information:-
Too much information is sometimes provided, or it is provided in too complex a fashion.
In general, good designs provide simplified displays although it can depend on the situation.

5. Present information in appropriate detail:-
The design of signs, instruction manuals & controls panels all can benefit from evaluation.

Page No.:
Date.:

Topic:

6. Present clear images:-
Another common problem is exhibing an Image poorly so that the user cannot distinguish or interpret the message.
There are three issues in presenting clear images, namely being visible, distinguishable & interpretable.

7. Use redundancies:-
Sometimes, one message is insufficient.
Because mistakes are easy to make & humans have many limitations it is important to provide the same information is more than one way.

8. Use pattern:-
The human eye grasps patterns well, information presented as a pattern can often be understood much more quickly & accurately than otherwise

9. Provide variable stimuli:-
Humans detect a novel stimulus more readily than a constant one because our senses fatigue easily with continuous exposure.

10. Provide Instantaneous feedback:-
An additional principle that helps prevent errors is to provide feedback to the user on the course of action taken. the sooner the feedback is given the easier it is to determine if an error has been made or not.

Topic : Page No. : Date. : / /

Q3 Language Models :-

The goal of a language model is to compute a probability of a token and are useful in many different Natural Languge Processing applications.

- Language Model (LM) actually a grammer of a language as it gives the probability of word that will follow.

Language Model Definition :-

In case of Probabilistic language 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 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 is available in practice.

Page No.:
Date.: / /

Topic:

1. Methods using the Markov assumption:-

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

A process with this property is called a Markov process.

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

Ex. if $k = 1$:

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

or if $k = 2$:

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

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

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

Topic : ____ Page No. : ____ Date. : / /

2. N-gram Models:-

From the Markov Assumption, we can formully define N-gram models where $k = n-1$ as the following:

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

The simplest versions of this are defined as the Unigram Model ($k=1$) & the Bigram Model ($k=2$).

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

$$P(w_1 w_2 \ldots w_n) \approx \prod_i P(w_i)$$

4. Bigram Model ($k=2$):-

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

These equations can be extended to compute trigrams, 4-grams, 5-grams, etc.
This is an insufficient model of languge because sentences often have long distance dependencies.

Following is the Maximum Likelihood Estimate model to Estimating Bigram Probabilities:-

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

Topic : Page No. : Date. :

Q4 Machine Translation:-

Machine translation is the classic test of language understanding.
It consist of both language analysis and language generation.

Many machine translation systems have huge commercial use.

- Google Translate goes through 100 billion words per day.
- eBay uses Machine Translation techniques to enable cross-border trade and connect buyers and sellers around the world.

- facebook uses machine translation to translate text in posts and comments automatically buyers and sellers around the world.

In a traditional Machine Translation system, parallel corpus a collection of text is used each of which, is translated into one or more other languages than the original.

It is obivious that, this approach skips hundreads of important details, requires a lot of human feature engineering consists of many different and independent machine learning problems, and overall

Page No. :
Date. : / /

Topic :

is a very complen system.

A. Neural Machine Translation (NMT):-

- The above process is modelled through one big artical neural network, known as a Recurrent Neutral Network (RNN) which, is a stateful neural network.

It has connection between passes and connections through time.

Neurons are fed information not just from the previous layer and train the network matters.

- Standard Neural Machine is an end-to-end neural network where the source sentence is encoded by a RNN called encoder, & the target words are predicted using another RNN known as decoder.

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 content:-
NMT can use a much bigger contect for both source

Topic : Page No. : Date. :

and partial target text in order to translate more accurately.

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

~~The main problem with RNNs is the generation is of much higher quality than the parallel co~~

The main problem with RNNs is the vanishing or exploding gradient problem where, depending on the activation functions used, information rapidly gets lost over time.

Intuitively, this wouldn't be much of a problem because these are just weights and not neuron states but the weights through time is actually where the information from the past is stored.

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

- LSTM works as a solution to the vanishing gradient problem by introducing gates and an explicity defined memory cell.
Each neuron has a memory cell and three gates:
The function of these gates is to safeguard the information by stopping or allowing the flow of it.

For Educational Use Only

Topic:

Page No.:
Date.:

1. The input gate determine how much of other information from the previous layer gets stored in the cell.

2. The output layers takes the job on the other end & determines how much of the next layer gets to know about the state of this cell.

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

LSTMs are able to learn complex sequences, such as writing like shakespeare or composing primitive music.

It is the default model for most sequence labeling tasks, which have lots of data.

c. Gated Recurrent Units (GRU):-

- They are a 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 a 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.

The reset gate functions much like the forget gate of an LSTM, but it's located slightly differently. They don't have an output gate.

There have been further improvements in neural machine translation systems over the past few years:-

- Sequence to Sequence, learning with Neutral networks proved the effectiveness of LSTM for Neural Machine Translation.

It presents a general end-to-end approach to sequence learning that mistake minimal assumptions on the sequential. to vector of a fixed dimensionality, and then another deep LSTM to decode the target sequence from the vector.

Neural Machine Translation (NMT) by jointly Learning to Align and Translate introduced the attention mechanism in NLP.

Convolutional over Reurrent Encoder for Neural Machine Translation arguments the standard RNN encoder in NMT with additional convolutional layers in order to capture wider content in the encoder output.

Topic: ______ Page No.: ____ Date.: / /

Q5 Phondogy:-

Phonology mainly deals with the sound system of language. It ~~consists~~ considers sounds in languages are organized systematically in languages.

All the words we pronounce in languages are systematic combination of sounds.

There are more than 5000 languges around the world and these languages have different sound combinations.

Phonology studies of these various combinations.

Morphology:-

Morphology is the study of words or morphemes, the smallest units in a language.
Every language has its own system of sound combination and theses sounds together form a word.

Morpheme is known as the smallest unit in a particular language.
While sounds join to make words connect to form phrases or sentenses.

Words play an important role in any language and linguists have defined words in many ways.

Topic : ____________ Page No. : ____ Date. : / /

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.

Syntactic Analysis :-

In syntactic analysis the sentences are parsed as noun, verbs, adjectives and other parts of sentences.

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

Eg. "mango eats me" will be rejected by syntactic analyzer.

Word Sense Disambiguation :-
Word sense disambiguation, in natural language processing (NLP), may be defined as the ability to determine which meaning of word is activated by the use of word in particular content.

Lexical, ambiguity, syntactic or semantic, is one of the very first problem that any NLP system faces.

Page No. :
Date. : / /

Topic :

Part-of-speech (POS) taggers with high level of accuracy can solve world's syntactic ambiguity.

Resolving semantic ambiguity is harder than resolving syntactic ambiguity.

Eg.
Consider the two examples of distinct sense that exist for the word "bass" -

- I can hear bass sound.
- He likes to eat grilled bass.

The occurrence of the word bass clearly denotes the distinct meaning.

In first sentence, it means frequency and in second, it means fish.

Eg:-
- I can hear bass/frequency sound.
- He likes to eat grilled bass/fish.