**July 2019**

**Examination Paper Solutions**

***School of Science and Technology***

**MIDDLESEX UNIVERSITY**

**EXAMINATION PAPER**

**Academic Year 2018/2019 (July)**

**CSD3939**

###### Developing Artificial Intelligence

**Prof C. Huyck**

Time allowed: 3 Hours

Total number of questions: 4 Questions

Instructions to candidates: Answer all questions. Each question carries 25 marks.

Materials provided: Equipment permitted: None

 Total number of pages:

**EXAM PAPER CAN BE REMOVED FROM THE EXAM ROOM**

|  |
| --- |
| **No books, paper or electronic devices are permitted to be brought into the examination room other than those specified above.****Candidates are warned that credit cannot be given for work that is illegible** |

1. State Spaces

(a) The second coursework represented a character by a 64 dimensional vector of values between 0 and 16 (inclusive). How many possible inputs are there?

(7 marks)

Marking scheme:

3 base of 17, 2 for a base of 16 (which is wrong but a reasonable mistake)

2 for a power between 10 and 100

2 for correct

**Sample answer:**

**Each of the features can take one of 17 values so the answer is 17 to the power of 64 or 17^64. The output feature has 10 values so including that there are 17^64\*10.**

(b) You can define a maze with a graph where the nodes represent choice points, and the arcs represent the connections between the choice points. Describe the depth first search algorithm of mazes. Give an example maze with at least 5 choice points and an example search of that maze.

(10 marks)

Marking scheme:

2 for a maze with 5 or more choice points

2 for answer

1 for marking cycles

2 for example search

**Sample answer: You start at the start state, expand one option, and recall the others in a stack for backtracking. You repeat this process until there are no options, or you reach the goal. If there are no options, you backtrack to the recorded option. If you get to the start state with no options, you can’t reach the goal. The below example has no cycles; if there are cycles you have to mark them, and not add them a second time.**

**Example search of the below graph. S->4->2 <backtrack to>4->3<>4<>S->5->1->G**



(c) Describe Iterative Deepening. When is it useful? Give an example of when it is useful.

(8 marks)

Marking scheme:

2 it’s used in minimax

2 for good example like chess

2 for correct description of deepening.

**Sample answer:**

**Iterative deepening is typically used during a minimax search. The system, playing an agent, looks forward (using minimax). When you get make a choice, you lose all the optional branches except the one you chose. At this point you deepen, by continuing the search from this location. This goes on while the opponent is making their choice, so you may get one or even two extra plys. Minimax is used for adversarial games like chess, so a good chess program uses iterative deepening.**

2. Knowledge Representation

a) Is XML Turing complete? (A simpler version of this question is: does XML do any processing?) Explain why it is or is not.

(6 marks)

Marking scheme:

3 for No

3 for an explanation.

**Sample answer:**

**No, XML is not Turing complete. It is a language for specifying tag sets. HTML is a tag set, but XML is typically used for domain specific tag sets like MathML. In both XML and the languages it defines (like MathML), there is no processing. (Modern HTML enables you to embed code for processing, but that is not part of the XML specification.)**

(b) Write a semantic net for the topics in CSD3939. This should include at least 12 nodes, and 5 types of arcs, including the most important types of arcs.

(10 marks)

Marking scheme:

2 points for IsA

2 points for instance

2 points for other reasonable arcs (mine are not exhaustive)

2 points for 12 nodes

2 points for good topics

**Sample answer:**



(c) There is a project at a social media firm to automatically determine whether a statement is true or false. What would be the benefits and drawbacks of using a logic based reasoning system? Specify the type of logic you would recommend.

(9 marks)

Marking scheme:

3 for selecting a type of logic.

3 for noting that it wouldn’t be perfect (and why)

3 for noting that it could help (and why)

Note the answer below is a sample answer, and wildly different answers will work.

**Sample answer:**

**This is a difficult task because it is difficult to determine whether most statements are true or not. For instance, it is not at all clear that the statement “Donald Trump is a good president.” Is true or not. None the less, there are facts that are generally agreed like “Donald Trump was elected President of the United States in 2016.” Using a relatively simple first order predicate logic system could help determine whether many statements are true or false based on a stored list of agreed upon facts. It would also have to say that it cannot determine that truth value of many statements. One further benefit would be that it could say that statements are vacuous (If Hillary Clinton was elected president of the US in 2016, the world would be a better place now.) Also as the database of facts got bigger, it would have to maintain consistency, and it might take a long time to prove facts that it could in fact prove.**

3. Machine Learning

a) What is the distinction between a parametric and a non-parametric algorithm. Give an example of each.

(10 marks)

Marking scheme:

3 for example non- parametric.

3 for example parametric

4 for description of the difference.

**Sample answer:**

**A parametric algorithm (like a multi layer perceptron learning via back propagation) sets parameters based on the training set. (In the case of the MLP, the weights between perceptrons are the parameters.) It then uses this system on the test set. A non-parametric algorithm (like nearest neighbor using Euclidean distance) does not set any parameters.**

(b) I’ve made a table of mushrooms. The three binary features (colour, spotted, height) determine whether a mushroom is poisonous. Build the decision tree to categorise mushrooms that asks the fewest questions.

|  |  |  |  |
| --- | --- | --- | --- |
| Red | Spots | Tall | Poisonous |
| Red  | Spots | Short | Poisonous |
| Red | No Spots | Tall | Not Poisonous |
| Red | No Spots | Short | Not Poisonous |
| Blue | Spots | Tall | Poisonous |
| Blue | Spots | Short | Poisonous |
| Blue | No Spots | Tall | Poisonous |
| Blue | No Spots | Short | Not Poisonous |

(8 marks)

Marking scheme:

2 for a decision tree

3 for a correct decision tree

3 for a minimal correct decision tree

**Sample answer: I think this is the only minimal tree.**



(c) Would a deep net be a good solution for the second coursework? (The second coursework was a 2-fold categorisation of digits with inputs being 64 dimensions, and there being 5620 data items.) Explain why or why not.

(7 marks)

Marking scheme:

3 for overfitting

3 for small data set

1 for reduction due to 2 fold

(partial credit for yes noting expanding data set including a generative model)

**Sample answer: No, a deep net would not be good. There are many versions of the MNIST task from coursework 2, but the one we used only had 5620 items. That’s not a lot of data. So, if you trained on that with a two fold test, you’d only have 2810 for training. Any reasonable sized deep net would overfit the data. On larger versions of the task (with millions of data items), deep nets are currently have the best results. You can make use of mechanisms to expand the data set, but those do have problems.**

4. Applications

a) Give an example of a mobile robot, and briefly explain how it works.

(9 marks)

Marking scheme:

2 for a good example (there are lots)

3 for sensors

2 for effectors

**Sample answer:**

**An example of a mobile robot is automatic vacuum that sweeps on its own. It moves forward (sweeping) using its effectors (wheel engines), and has sensors to avoid objects. Typically it has a bump sensor (if it runs into something) and might have a visual sensor. It might even use GPS. Similarly, it might build up a map of the environment is sweeping.**

 (b) Describe a McCullouch Pitts neuron. (It is also called an integrate and fire neuron).

(8 marks)

Marking scheme:

3 for integrate

3 for fire

2 for notes about what it does if it doesn’t fire in a cycle. I’ll give full credit for the Lapicque model that retains all the activation making it fire even with low regular input.

**Sample answer:**

**A McCullouch Pitts neuron is a model neuron that integrates activity from incoming synapses. If it collects enough activity, it fires and sends out activity from its own synapses. Typically, it works in time steps, and does not retain activity from earlier steps.**

 (c) How would you use a bag of words technique to represent semantics of a word in a corpus of 1000 documents?

(8 marks)

Marking scheme:

3 for term by document matrix

3 for reduction

2 for showing how similar words have similar vectors

**Sample answer: Some form of Singular Value Decomposition (like latent semantic analysis) can be used. You build up a term by document matrix. The meaning of the words is the documents it is in. (You typically throw away stop words like a and in). You then reduce the matrix to two much smaller matrices. So an initial 1000x3000 matrix can be reduced to two matrices say 1000x100 and 100x3000. (In this example, there are 3000 unique non-stop words in the corpus.) This solves the synonymy problem so that dog, cat and puppy will all be represented by similar 100 dimensional vectors. You typically compare using a cosine measurement. This vector now loosely represents the semantics of the items.**