University of London International Programmes
Computing and Information Systems/Creative Computing

CO3310 Artificial intelligence

Coursework assignment 2 2017–18

Notes

- Any websites cited below were last visited on 6 February 2018.

- In what follows, AIMA refers to *Artificial Intelligence: A Modern Approach* (3rd edition) by Stuart Russell and Peter Norvig.

- Hutton and Barto (2017) refers to *Reinforcement Learning: An Introduction* by Richard S. Sutton and Andrew G. Barto (2nd edition, in progress) (Cambridge, MA: MIT Press, 2017): http://incompleteideas.net/book/the-book-2nd.html

- You should list all references at the end of your work, and they should be properly cited whenever referred to. References to websites should include title and author if available, and date last visited.

- Answers should be expressed **in your own words**: an answer consisting entirely or mostly of quotes is unlikely to gain good marks, whether or not the sources are properly referenced.

There are 100 marks available for this coursewok assignment.

*Your coursework should be submitted as a single PDF file (**not** a zip file; and any additional files will be disregarded), using the following file-naming conventions:*

YourName_SRN_COxxxxcw#.pdf (e.g. MarkZuckerberg_920000000_CO3310cw2.pdf)

- ***YourName** is your full name as it appears in your student record (check your student portal);*
- ***SRN** is your Student Reference Number, for example 920000000;*
- ***COXXXX** is the course number, for example CO3310; and*
- ***cw#** is either cw1 (coursework 1) or cw2 (coursework 2).*

Marks may be deducted if your submission is not in the required format.

Important reminder: It is important that your submitted coursework assignment is your own individual work and, for the most part, written in your own words. You must provide appropriate in-text citation for both paraphrase and quotation, with a detailed reference section at the end of your coursework assignment (this should not be included in any word count). Copying, plagiarism and unaccredited and wholesale reproduction of material from books or from any online source is unacceptable, and will be penalised (see our guide on how to avoid plagiarism on the VLE).

1. Machine learning and AI ethics   **[35 marks]**

(a)

    i.   Explain the following terms in the context of **reinforcement learning**:

- 1. Reward
- 2. Return
- 3. Policy
- 4. Value function
- 5. Temporal-difference learning.

    ii.   What is meant by the Markov property in the context of machine learning, and why is it useful?

    iii. Explain some important differences between reinforcement learning and evolutionary methods, as described by Hutton and Barto (2017, Chapter 1).

(b) Recent research has shown that machine learning applied to human language data can acquire and even amplify discriminatory biases present in the data, including stereotyped association of gender with particular occupations or negative attitudes towards ethnic minorities (Bolukbasi *et al.* 2016, Caliskan *et al.* 2017). Furthermore, the notorious incident of Microsoft's 'Tay' experiment showed that a chatbot can be taught to acquire undesirable attitudes by malicious Internet users (Wolf *et al.* 2017).

With reference to at least two of the papers listed below:

    i.   Summarise and explain the kinds of bias that can be found in the results of ML systems applied to natural language data.

    ii. Discuss what developers of ML systems can do to reduce the acquisition of discriminatory biases and/or minimise their harmful results.

You should write no more than about 1,200 words.

References

1. Tolga Bolukbasi *et al.* (2016) 'Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings', *Advances in Neural Information Processing Systems* (pp.4349–57): https://arxiv.org/pdf/1607.06520.pdf

2. Caliskan, A. *et al.* (2017) 'Semantics derived automatically from language corpora contain human-like biases', *Science*, 356 (6334), pp.183–86. ISSN 0036-8075: http://opus.bath.ac.uk/55288/4/CaliskanEtAl_authors_full.pdf

3. Marty J. Wolf *et al.* (2017) 'Why We Should Have Seen That Coming: Comments on Microsoft's Tay "Experiment," and Wider Implications', ACM *SIGCAS Computers and Society*, 47(3), pp.54–64: https://www.orbit-rri.org/ojs/index.php/orbit/article/view/49/46

2. Planning   [30 marks]

(a) Explain the following terms in the context of **AI planning**:

    i.   Progressive planning
    ii.   Regressive planning
    iii.   Partial-order planning
    iv.   Serializable subgoals
    v.   Subgoal independence.

(b) This question involves PDDL. Suppose you and a friend plan to travel together from London Heathrow airport to Paris Charles de Gaulle. Owing to restrictions on your tickets, you must both travel on the same flight. You are in Manchester and your friend is in Glasgow. Flights are available from both these cities to London.

    i.   Write PDDL sentences for any actions that will be involved in this journey, the initial state and the goal state. Explain the intended meanings of any predicates and constants you use.
    ii.   Construct a **plan** for the journey using the actions you have defined in (i).

3. Natural Language Processing   [35 marks]

(a) Study the phrase structure grammar shown on p.42 of the subject guide. Which of the following sentences are generated by this grammar? Justify your answer, and create a parse tree for each sentence that matches the grammar.

    i.   John smells the wumpus in the pit that stinks.
    ii.   Boston is right ahead.
    iii.   Mary and John see Boston.
    iv.   Mary feels the breeze in Boston.
    v.   The wumpus that Mary smells is nearby.
    vi.   John is in Boston and a smelly dead wumpus is nearby.

(Note that *S Conf S* in the grammar should read *S Conj S*.)

(b) Explain how the grammar can be modified so that it will generate the following sentences, and create a parse tree for each one:

    i.   Miriam sees a wumpus or a small aardvark and walks away.
    ii.   The wumpus and the pit stink.
    iii.   Pedro saw 1345 wumpuses in Boston.
    iv.   Maya very often sees a wumpus in the nearby park.

Are any of the sentences (a) i–vi or (b) i–iv ambiguous according to the original grammar or your modified version? Justify your answer by giving alternate parse trees.

(c) Extend the augmented grammar on p.45 of the subject guide to include verbs with three arguments such as *give(a, b, c)*, meaning 'a gives b to c'. Write grammar rules which will derive *give(john, wumpus, mary)* from both sentences (1) and (2) below, and give augmented parse trees for both.

    1.   John gives Mary a wumpus.
    2.   John gives a wumpus to Mary.

[Total 100 marks]

[END OF COURSEWORK ASSIGNMENT 2]