Coursework commentary 2018–2019

CO3310 Artificial intelligence

Coursework assignment 2

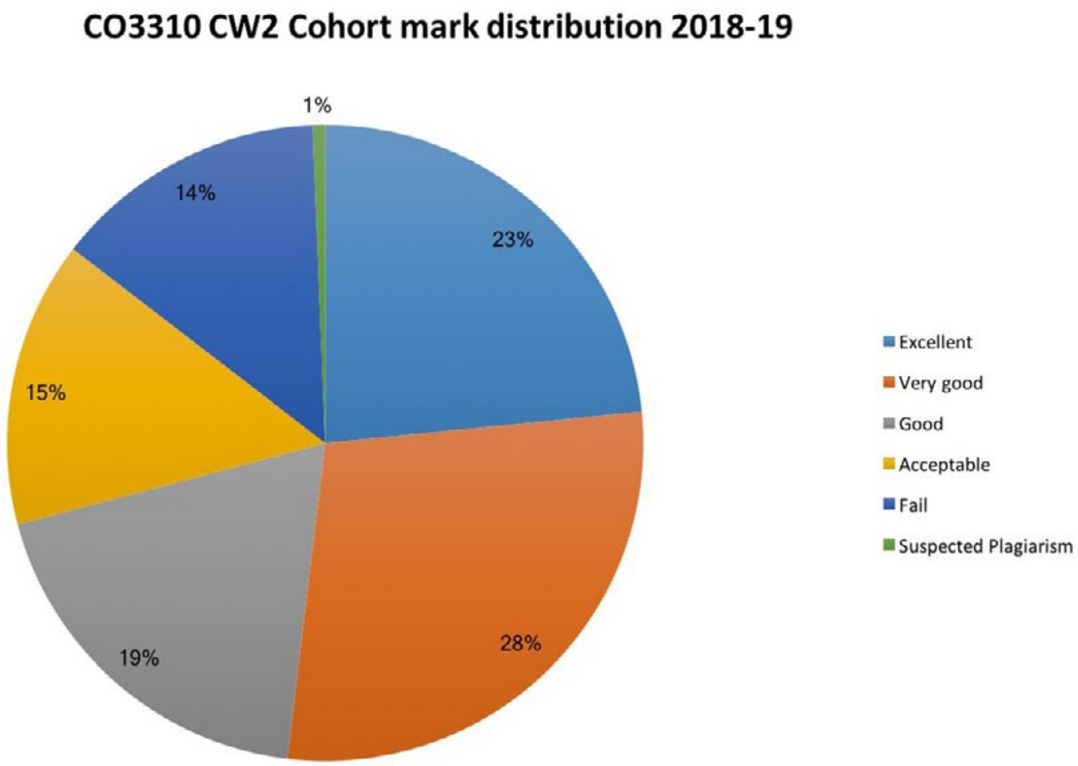
General remarks

- In what follows, 'AIMA' refers to *Artificial Intelligence: A Modern Approach*. (3rd edition, 2010) by Stuart Russell and Peter Norvig.
- All websites cited below were last visited on 8 February 2019.

Students were reminded:

> *It is important that your submitted 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 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).*

See 2018–2019 CW2 cohort mark distribution below:



Comments on specific questions

Question 1: Search and planning [40 marks]

This question posed a route-finding problem applied to real-world data, to be obtained from the Transport for London website.

a. Many students were puzzled by the phrase 'problem formulation'. Section 3.1 of the subject guide shows how to formulate a task as a well-defined problem, so it should be evident that what follows can be seen as a template for a problem formulation. Some students confused a *state* of the agent with a *location* in physical (geographic) space. A location is not a state, but an expression such as *At(Location)* denotes a state. The majority of answers were marked as good, very good or excellent.

b. Students generally showed confident understanding of the different search algorithms, and most answers were assessed as at least good, while more than half were marked as very good or excellent. Some students lost marks by answering the question about depth-limited search only in terms of DFS without proposing or motivating a maximum depth.

   Some confused the traversal of the search space with the actual route selected, e.g. the BFS solution was given as:

   > ***Shortest Path from Victoria to Leicester Square:***
   >
   > *Victoria, Green Park, St. James's Park, Pimlico, Sloane Square, Oxford Circus, Piccadilly Circus, Westminster, Hyde Park Corner, Bond Street, Westminster, Vauxhall, South Kensington, Regent's Park, Warren Street, Tottenham Court Road, Piccadilly Circus, Bond Street, Oxford Circus, Leicester Square*

   which is clearly absurd. For full marks students should give the actual result and indicate how it is calculated, rather than leaving the examiner to read the solution off a tree diagram.

c. Most students answered this question confidently, and over half were marked as 'excellent'. Some students lost marks through misunderstanding or misreading the question: the heuristic is the estimated time from each station to Leicester Square, not to intermediate stations. The question states that the fastest walking time should be used, but some students chose instead to use the **medium** walking time and so lost marks.

d. This question involved the **situation calculus**. The idea was that the literals given in the subject guide would be decomposed in terms of actions, relations, objects and axioms, e.g.:

   have(Agt, Bread, init)

   have(Agt, Jam, init)

   boiled(Kettle, do(boil(Agt, Kettle), S)) ←

   poss(boil(Agt, Kettle), S)

   The main point of this question was to get students to think hard about the challenges of designing ontologies for even simple everyday tasks. Generally, students did quite well on this question, with most answers marked as good or better and a small number gaining full marks.

Question 2: Machine Learning [30 marks]

This was an essay question concerning the strengths and limitations of 'deep learning' (DL), with students required to refer to two specific articles as well as (optionally) their own independent reading. Many students did not in fact refer to the specified articles, preferring to rely on non-academic sources. This is not advised, as there is limited quality control over such sources. Marks were awarded for:

- Evidence of relevant technical knowledge/understanding; there was no need to give lengthy background about AI, ANNs, etc., but answers were expected to show adequate understanding.
- Quality and coherence of argumentation, addressing issues posed in the question and arguing for a specific thesis.
- Scholarly practice and clarity of expression.

Most answers obtained good, very good or excellent marks for technical content, but scored somewhat less well for their quality of argumentation. Good answers would consider and evaluate arguments in support of DL as well as more critical positions, and would aim to reach a conclusion.

Some common shortcomings included over-reliance on non-scholarly sources, a lack of critical assessment of claims that have been made for DL, weak or speculative conclusions, and saying *e.g.* '*I believe*' or even '*I strongly believe*' in place of reasoned argumentation.

Other issues:

- Many students didn't actually say what DL is, perhaps because they didn't fully understand the concept.
- Several students gave as an advantage of DL systems that they do not get tired, which surely applies to any computer system.
- Most answers showed little concern for privacy issues.

Question 3: Natural Language Processing [30 marks]

a. This question concerned the **Chomsky hierarchy** applied to the grammatical analysis of natural language. Chomsky's distinctions between different classes of languages have been influential within theoretical computer science as well as linguistics, but students were specifically required to focus on the implications for natural language. Hardly any actually did this, meaning that relatively few answers were marked as 'very good' or better. A very few students submitted well-researched answers addressing issues to do with NLP such as the expressive power of different levels of context-free versus regular grammars, evidence that some natural languages may require context-sensitive grammars, and the kinds of NLP tasks that are within the scope of finite state methods. However, many only provided technical details of some of the formal characteristics of different classes of languages, which were not pertinent to the question and did not obtain high marks.

b. This question required students to amend or extend a provided set of formal grammar rules. Answers generally showed reasonable understanding, with most assessed as good, very good or excellent. Many students gave away marks by failing to explain their answers as required. Common flaws included over-generating (e.g. using unitary categories for verbs or conjunctions), and assuming informally stated constraints on the order of conjunctions rather than covering this explicitly in the formal rules. Marks were given for compactness and generality of the proposed grammar rules; some students lost marks through proposing *ad hoc* rules that were too closely tailored to the provided example sentences.

c. This question involved drawing syntax trees. This should have been fairly straightforward and in fact the majority of students obtained full marks, though some students presented trees that didn't actually match their proposed grammar rules.