Examiners' commentaries 2016–17

CO3311 Neural networks – Zone A

General remarks

There were six questions on this examination paper, of which candidates were required to answer four. The time allowed was two hours and 15 minutes. Question 1 was the most popular and the highest marks achieved were for Question 5. Lower marks were gained for Questions 3 and 6. You should ensure you cover all parts of a question. Presentation such as sketches/diagrams and tables can aid in gaining marks.

Comments on specific questions

Question 1

1a) required candidates to show how all problems can be thought of as search problems, with specific reference to artificial neural networks. Here we are looking for a set of weights and an architecture that reproduces the training data as faithfully as possible. The two examples requested, that of AND and XOR were chosen as simple examples where one or more than one unit was required. Good answers gave appropriate networks for each and noted that in the case of XOR a single unit would not suffice because the problem is not linearly separable.

1b) required candidates to show their understanding of terms such as 'extended truth table' and to explain the Backpropagation algorithm. This proved easy for most.

1c) required formulae for the operation of the given threshold unit – candidates had to recognise this unit type from the representation introduced during the course.

1d) tested knowledge of this representation further, as candidates needed to draw a specified Backpropagation network. Good answers included a key explaining the symbols used.

1e) allowed candidates to demonstrate their understanding of the training of backpropagation networks. As part of a good answer, it is essential for you to include **a definition of terms** used.

Of the six questions on this paper, this question was the most popular. Overall, it was well answered, justifying the choice of most candidates to answer it.

Question 2

2a) explored the issue of feedback from the output of a unit to its inputs. Good answers showed that (except in very special circumstances) inconsistent requirements could be made on such a unit.

2b) was bookwork, where you needed to define the concept of architecture in the context of artificial neural networks and to give its five key features. Strong answers included a paragraph on each of these, and examples of what they might be.

2c) required a comparison of five main network types (perceptron, backpropagation, Kohonen–Grossberg, Boltzmann and Hopfield networks). Good answers included a point by point comparison of each of the five aspects. This could be easily achieved especially if presented in the form of a table. Less good answers omitted some comparisons, perhaps because candidates who chose not to use a table found it difficult to remember to include each of the five aspects for each of the five network types.

This question was chosen by around 50 per cent of candidates, and was answered well overall.

Question 3

3a) required you to explain how any truth table can be implemented as a network of perceptrons (this is important to recognise, notwithstanding the normal use of artificial neural networks in the context of machine learning). Good answers showed how an arbitrary truth table can be implemented. Weak answers gave a simple example without explaining how the algorithm could be extended to any truth table.

3b) invited candidates to show their algorithm working on a simple truth table; that of the XOR function. Good answers related the answer here to their answer to Part 3a). Weaker answers offered solutions with little explanation or relation to the algorithm.

3c) Having shown that all truth tables are implementable, 3c asked why training is such an important part of neural networks. Good answers included issues of lack of truth tables, either because these are unknown or because they do not exist for many problems. Weaker candidates seemed unaware of the breadth of neural network applications.

3d) considered enhancements to the backpropagation algorithm. Candidates were asked to explain the use of **momentum terms, adaptive learning rates and different learning rates for each weight**. Good answers explained terms, and the difficulties that the unenhanced algorithm might face, and showed how these might be overcome using one or more of these three enhancements.

Overall this question was less well answered than the others, perhaps because it required explanations to some parts, which were sometimes overlooked.

Question 4

4a) required a diagram showing the layers of a Kohonen-Grossberg network, and its essential parameters. Good answers explained the two layers and the fact that the Grossberg layer is often missing.

4b) required the typical net and activation functions of each layer of such a network. Standard answers gave their names; for an excellent answer you should also include a sketch of the activation functions.

4c) required explanation that the Grossberg layer often plays a minor role.

4d) required candidates to state the training algorithm for the Kohonen layer. Poor answers often omitted one or both of the normalization steps. Strong answers included a pseudocode account of the algorithm, with an explanation of the terms included. They also explained what was being sought at each step; for example, with normalization.

4e) asked for a straightforward application of the algorithm. Minor slips were not heavily penalised.

4f) required an explanation of the strategies that might be used to choose the initial classes. One of the hardest aspects of applying Kohonen-Grossberg networks is in the initial choice of classes, as no foolproof strategy is known.

Good answers provided a list of strategies, explaining how they worked, together with a note that none were likely to work in every circumstance. Poor answers just gave a list or partial list.

This question was answered by about 50 per cent of candidates. The mark distribution showed the cohort divided into two – those who were able to express their ideas clearly, particularly in respect of explanations, obtained high marks; and those who had difficulty doing this and obtained low marks.

Question 5

5a) required a description of the process of implementing a 4 unit Hopfield network. Good answers carefully walked the reader through the steps relevant for the chosen tool. Weaker answers often omitted steps – sometimes these were obvious, with the weakest answers giving little more than a sketch of the process.

For part 5b you were required to provide a statement of the fact that Hopfield networks have symmetric weight tables with zeros on the diagonal; and to make use of this fact to complete a partially filled table. Poor answers often omitted a statement of the rule.

5c) carried the majority of the marks for this question (16 marks out of 25). It required the careful working out of the state transition table of the network given by the weight in 5c) and from this the production of a state transition diagram. Good answers worked through using the format given in the course. The ordering of nodes as 3 2 1 in the table rather than 1 2 3 is crucial, and a common mistake, besides the odd arithmetic slip, was incorrect ordering in the table. A standard feature of correct tables is some sort of symmetry, so it is helpful to check those entries, if any, that seem to 'buck the trend'. An all too common error in the diagram was the omission of circles showing self-transitions, and the opposite error also occurred a number of times – that of circling all states – including ones that do not transition to themselves.

This question was answered by around 65 per cent of candidates. It was the best answered question in terms of marks achieved. Once the technique of calculating the table and the diagram were mastered, candidates were proficient in producing these in the limited time of an examination.

Question 6

6a) simply required the names of two tools for implementing artificial neural networks.

6b) asked for a description of the process of implementing a simple network using each of the tools. It was clear from the answers that some candidates did not know the implementation processes. Of course, there were very good answers too, which gave step by step instructions.

6c) required a comparison of the tools. Good answers gave two main advantages and two main disadvantages of each. Poor answers were confused and rambling, making it difficult for examiners to identify the differences.

6d) required a reasoned explanation of where each tool might be used and why. It is important that your answer to this is not just a repetition of that for part c). Please read through the whole question to ensure you are clear about what is being asked in each part. Good answers gave careful reasoning while poorer ones often omitted the reasoning.

Few candidates attempted this question and the marks obtained were generally disappointing. This question was intended to allow candidates an opportunity to display the knowledge gained from coursework, which included a comparison of tools for implementing artificial neural networks. The nature of the answers required – explanations and reasoning rather than calculations – may have been a deterrent.

Examiners' commentaries 2016–17

CO3311 Neural networks – Zone B

General remarks

There were six questions on this examination paper, of which candidates were required to answer four. The time allowed was two hours and 15 minutes. Questions 1 and 4 were popular with almost all candidates choosing them. Question 4 showed the highest average score. Weaker scores were achieved on Questions 3 and 6. You should ensure you cover all parts of a question. Presentation such as sketches/diagrams and tables can aid in gaining marks.

Comments on specific questions

Question 1

1a) concerned threshold units, with a simple requirement to define the commonly used terms: net, step unit, clamped, energy and bipolar activation. Many good answers included sketches to illustrate their points.

1b) required candidates to explain how perceptrons can be used to implement simple logic circuits, specifically by designing AND, NAND, NOR and NOT gates and showing how each of these works. Good answers gave concise and comprehensive explanations, while poorer ones omitted much detail. Some very poor answers addressed only the particular examples and did not explain how the design is produced in general.

1c) covered the uses and limitations of networks of threshold units; these have their limitations both in what they can model and in how the appropriate networks can be found. Good answers systematically covered areas such as limitations to binary outputs, lack of a training algorithm in all but the simplest cases, and often cumbersome solutions. Poorer answers omitted one or more of these aspects.

Almost all candidates chose this question, and on the whole candidates scored well, though there were a few poor answers.

Question 2

The original inspiration of artificial neural networks was biology, but other sciences such as physics have also had their impact.

2a) asked for a comparison (including contrasts) of the inspirations for perceptrons versus Boltzmann machines. Good answers explained how these can be motivated by referring to their respective antecedents while poor answers often simply stated the antecedents without motivation.

2b) required a comparison of the Widrow-Hoff rule and Hebb's rule. Good answers gave the overall strategies and illustrated them with reference to particular network types and formulae. Poorer answers gave either the strategy or the particulars but not both.

2c) asked candidates to state which of the three motivations for studying AI/ANNs discussed during the course they judge to be the most important, and to give their reasons. Good answers considered all three and justified their choice of the most important by comparing them. Poorer answers simply stated their choice with little or no justification or explanation.

2d) asked for a description of the extent to which ANNs contribute to each of the motivations. Good answers included points for each motivation and included limitations as well as successes.

Question 2 was less well answered than the most popular questions (1, 4 and 5) but candidates achieved higher scores than those answering Questions 3 and 6.

Question 3

This question concerned some of the problems that need to be addressed when applying artificial neural networks to real applications. These include choice of the many parameters associated with the architecture: initial values, speed, overfitting and network paralysis.

3a) and 3b) asked candidates to define overfitting and network paralysis, and then to explain some of the strategies that attempt to overcome these problems. Good candidates gave succinct definitions and followed these with an explanation of the strategies. Weak answers often omitted to note that there are no foolproof solutions.

3c) required candidates to list five other problems and write about the strategies that might be used to overcome them. Whilst many good answers were produced, some poor ones did little more than give a list with the briefest of detail.

Around 33 per cent of candidates attempted this question, and answers were generally weak overall.

Question 4

4a) asked about the roles of Kohonen and Grossberg layers in a typical Kohonen-Grossberg network. Good answers talked about both layers and mentioned the fact that the Grossberg layer is often omitted. This aspect was missing in weaker answers.

4b) asked for an explanation of the purpose of normalization, using a suitable example (candidates typically forget to normalize when asked to train a Kohonen-Grossberg network, so this was intended as a useful reminder).

4c) required candidates to explain the process of training such a network. Good answers gave all of the formulae involved (as asked) and a key to the terms appearing in the formulae. Poorer answers often omitted a key and gave sketchy accounts of the algorithm. It was disappointing to find that despite the prompting in 4b) some candidates forgot about normalization.

4d) asked about the difficulty that normalization faces with vectors with zero length either in the initial units or as targets. Applications will avoid this occurrence by using suitable coding. Candidates were in general unable to think through this possibility, though good answers explained that normalization of such an example or unit is not possible.

4e) gave a simple 3-unit network and asked candidates to work through the training algorithm for just one given example. Except for the occasional arithmetic slip most candidates were able to complete this successfully. Some candidates, however, omitted one or more of the normalization steps.

Almost all candidates chose this question; it showed the highest average score, with very few candidates scoring under 10.

Question 5

5a) asked for the architecture of a typical Hopfield network. Good answers mentioned that the network is fully connected with symmetric weights but no self-connections. They also detailed the summation for net and threshold activation function. Poorer answers omitted one or more of these features. A sketch of a small network was included in some good answers.

5b) asked candidates to explain what a 'stable state' is in the context of Hopfield networks. Good answers started with the concept of state and explained that some have successors while some do not.

5c) required an explanation of the concept of energy of such a network. Good explanations included a formula and a statement that this expression reduces when going from one state to another. Weaker answers often omitted the latter statement.

5d) and 5e) required you to calculate the state transition table and diagram of a network given its weight table. This is a straightforward (if tedious) calculation which many candidates were easily able to complete.

Minor arithmetic slips occurred in some answers but marks were given for correct method. Good answers marked the stable states in both the table and the diagram. By looking for a pattern in the state table, strong candidates were able to check their working for minor slips. Poor candidates often used circles inconsistently in the state diagram, either omitting them for some states that transition to themselves or including them for all states.

This was another question that the majority of candidates were able to answer well; many excellent scores were awarded. Sadly, a handful only achieved single digit scores.

Question 6

It is important that you are aware of some of the tools available that help apply artificial neural networks. Question 6 tests knowledge of some essential features in two tools of your choice.

6a) asks for a list of such essential features, and good answers explained **why** they are essential. Poorer answers simply stated the features without giving any rationale. Clearly any tool must include a means of building, training, testing, visualising and evaluating a variety of artificial neural network types.

6b) required evaluations of the performance of two tools against the chosen features. Few candidates answered this question and scores were low. It is difficult to judge whether the low scores were due to lack of knowledge, lack of writing ability under examination conditions or for some other reason.