

Calculators may be used in this examination provided they are not capable of being used to store alphabetical information other than hexadecimal numbers

UNIVERSITY OF BIRMINGHAM

School of Computer Science

Artificial Intelligence and Machine Learning

Mock Examination

Time allowed: 2 hours

[Answer all questions]

Note

Answer ALL questions. Each question will be marked out of 20. The paper will be marked out of 60, which will be rescaled to a mark out of 100.

Question 1

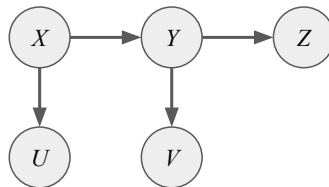


Figure 1: Directed acyclic graph (DAG) for Q1.

This question relates to the combinatorial (optimization) problem of finding all topological orderings of a given directed acyclic graph (DAG) with N vertices.

- (a) What is the worst-case size of the configuration space \mathcal{X} for this combinatorial problem, in terms of N ? Justify your answer. **[2 marks]**
- (b) A DAG with no edges is said to be *disconnected*. Assume the DAG in Figure 1 is disconnected. How many possible topological orderings of the nodes are there? Justify your solution. **[3 marks]**
- (c) You must solve the problem of devising an efficient algorithm for computing all topological orderings of the DAG in Figure 1.
 - (i) Give the steps in an exhaustive SDP algorithm for solving this problem. **[5 marks]**
 - (ii) Use this SDP algorithm to compute the set of all possible topological orderings of this DAG. **[6 marks]**
- (d) Now, assume the DAG in Figure 1 represents a probabilistic graphical model (PGM) over the random variables X, Y, Z, U and V .
 - (i) Using the topological ordering $[X, U, Y, V, Z]$ give the corresponding chain factorization of the joint distribution $P(X, Y, Z, U, V)$. **[2 marks]**
 - (ii) Using your chain factorization in (d)(i) above, remove all redundant dependencies from the conditional distributions, to give the Markov factorization of the joint distribution. **[2 marks]**

Question 2

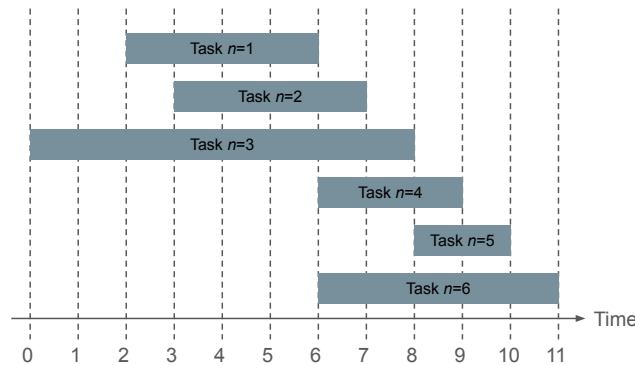


Figure 2: Task schedule chart for Q2.

(a) This question relates to dynamic programming as a method for combinatorial optimization.

(i) Explain the principle of optimality. **[2 marks]**

(ii) Explain how the principle of optimality reduces the computational complexity of an exact algorithm for solving an optimization problem. **[3 marks]**

(iii) Explain the concept of memoization. **[3 marks]**

(b) Consider the planning problem of automatically selecting the best subset of a set of tasks to complete. Each of the $N = 6$ tasks is associated with a *value*, from the list $v = [20, 7, 4, 3, 10, 8]$. The optimal subset has the largest sum of values. Each task begins and ends at the times given in the schedule chart above (Figure 3). Tasks can only be selected together if they do not overlap. The list $p = [0, 0, 0, 1, 3, 1]$ gives the largest task index $n' < n$ such that task n' does not overlap with task n .

(i) Use the following Bellman recursion to compute the total value of the optimal subset of tasks, V_6^* : **[8 marks]**

$$V_0^* = 0$$

$$V_n^* = \max(V_{n-1}^*, V_{p_n}^* + v_n)$$

(ii) What is the time complexity of this recursion? Justify your answer. **[2 marks]**

(iii) What is the worst case time complexity of the exhaustive solution to this problem *in general* (that is, independent of the specific task values and start/end times)? Justify your answer. **[2 marks]**

Question 3

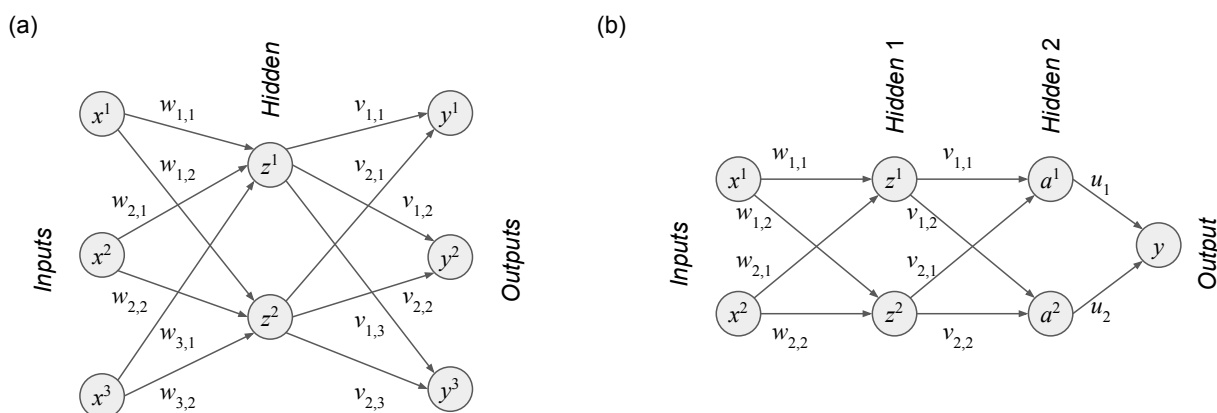


Figure 3: Deep neural networks for Q3.

- (a) For the deep neural network of Figure 2(a) (*autoencoder*) above with softplus activation nonlinearities, write down the formulae for the hidden nodes z^1 , z^2 and the output nodes y^1, y^2, y^3 . **[6 marks]**

- (b) The formulae for the deep network of Figure 2(b) are given in matrix form as,

$$z = \max(0, W^T x), \quad a = \max(0, V^T z), \quad y = \max(0, U^T a)$$

with weights $W = \begin{bmatrix} 0.1 & 0.7 \\ -0.3 & 0.2 \end{bmatrix}$, $V = \begin{bmatrix} 1.3 & -0.5 \\ 0.3 & -1.4 \end{bmatrix}$ and $U = \begin{bmatrix} 0.6 \\ -0.1 \end{bmatrix}$.

- (i) For inputs with value $x = \begin{bmatrix} 3 & 2 \end{bmatrix}^T$, compute the values of the first hidden layer z . **[4 marks]**
- (ii) Given that the values of the second hidden layer are $a = \begin{bmatrix} 0.75 & 0 \end{bmatrix}^T$, compute the value of the output y . **[4 marks]**
- (c) Using dual numbers for automatic differentiation, derive two dual number formulae: one for y paired with y_{u_1} (its derivative with respect to the weight u_1) and the other for y paired with y_{u_2} (derivative with respect to weight u_2). Both expressions should be in terms of the hidden outputs a and parameters U . Show your working. **[6 marks]**

This page intentionally left blank.

Do not complete the attendance slip, fill in the front of the answer book or turn over the question paper until you are told to do so

Important Reminders

- Coats/outwear should be placed in the designated area.
- Unauthorised materials (e.g. notes or Tippex) must be placed in the designated area.
- Check that you do not have any unauthorised materials with you (e.g. in your pockets, pencil case).
- Mobile phones and smart watches must be switched off and placed in the designated area or under your desk. They must not be left on your person or in your pockets.
- You are not permitted to use a mobile phone as a clock. If you have difficulty seeing a clock, please alert an Invigilator.
- You are not permitted to have writing on your hand, arm or other body part.
- Check that you do not have writing on your hand, arm or other body part – if you do, you must inform an Invigilator immediately
- Alert an Invigilator immediately if you find any unauthorised item upon you during the examination.

Any students found with non-permitted items upon their person during the examination, or who fail to comply with Examination rules may be subject to Student Conduct procedures.