UNIVERSITY OF BIRMINGHAM

School of Computer Science

Artificial Intelligence 2

Main Summer Examinations 2022

Artificial Intelligence 2

Exam paper

Question 1 (Probability)

Mary is taking the exam of AI2, which has three questions: question A, B and C. For each question, Mary either knows how to solve it and gets the full marks, or does not know and gets 0 marks. Suppose question A has 20 marks, question B has 30 marks, and question C has 50 marks. Suppose Mary knows how to solve question A with probability 0.6, question B with probability 0.5 and question C with probability 0.4. Assume Mary solves these three questions independently.

(a) Mary can get the first-class degree if she gets at least 70 marks. What is the probability of Mary getting a first-class degree? **[6 marks]**

(b) What is the expectation of the marks Mary can get from the exam? **[6 marks]**

(c) Let X_1 = "the marks Mary gets from question A", X_2 = "the marks Mary gets from question B" and X_3 = "the marks Mary gets from question C". Let $X = \max\{X_1, X_2, X_3\}$ (the maximum among X_1, X_2, X_3). Write down the probability mass function of X. **[8 marks]**

Question 2 (Probabilistic Machine Learning)

As a machine learning expert for an AI cyber security company, your task is to design an automated network intrusion detection system. You have collected a large number of records of network activities. Each record includes the log information about network activity, such as protocol types, duration, number of failed logins, which are random variables, denoted as $\boldsymbol{X} = [X_1, X_2, \ldots, X_n]^\intercal$. Each record also includes a binary random variable Y called label that was labelled by cyber security experts as intrusions ($Y = 1$) or normal connections ($Y = 0$).

(a) Consider feature selection based on mutual information to reduce the number of independent variables.

(i) Explain to your colleague, who knows nothing about information theory, the concept of mutual information. **[2 marks]**

(ii) Explain the loop, i.e., lines 4-7 of the pseudocode. Note $I(Y; X_i)$ is the mutual information between Y and X_i.

[3 marks]

1: **Initialisation:** Set $F \leftarrow \boldsymbol{X}$ and $S \leftarrow \emptyset$
2: $f_{\max} = \text{argmax}_{X_i \in \boldsymbol{X}} I(Y; X_i)$
3: Set $F \leftarrow F \setminus \{f_{\max}\}$ and $S \leftarrow f_{\max}$
4: **Repeat** until $|S| = K$:
5: $f_{\max} = \text{argmax}_{X_i \in F} I(Y; X_i) - \beta \sum_{X_s \in S} I(X_s; X_i)$
6: Set $F \leftarrow F \setminus \{f_{\max}\}$ and $S \leftarrow f_{\max} \cup S$
7: **End**
8: **End**

(b) After applying your feature selection algorithm, assume you selected four random variables as features, denoted as F_1, F_2, F_3, F_4. Based on these features, you now work with a cyber security expert to construct a Bayesian network to harness the domain knowledge of cyber security. The expert first divides intrusions into three cyber attacks, A_1, A_2, A_3, which are marginally independent from each other. The expert suggests the presence of the four features are used to find the most probable type of cyber attacks. The four features are conditionally dependent on the three types cyber attacks as follows: F_1 depends only on A_1, F_2 depends on A_1 and A_2. F_3 depends on A_1 and A_3, whereas F_4 depends only on A_3. We assume all these random variables are binary, i.e., they are either 1 (true) or 0 (false).

(i) Draw the Bayesian network according to the expert's description. **[2 marks]**

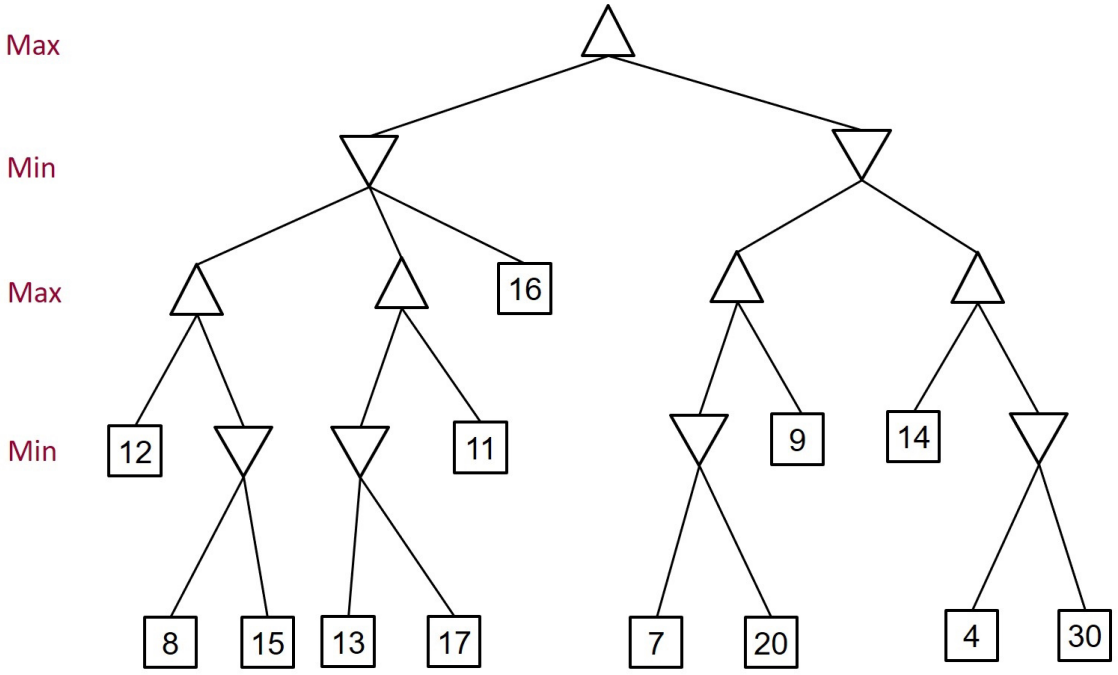
(ii) Write down the joint probability distribution represented by this Bayesian network. **[3 marks]**

(iii) How many parameters are required to describe this joint probability distribution? Show your working. **[5 marks]**

(iv) Suppose in a record we observe F_2 is true, what does observing F_4 is true tell us now? If we observe F_3 is true, what does observing F_4 is true tell us? **[5 marks]**

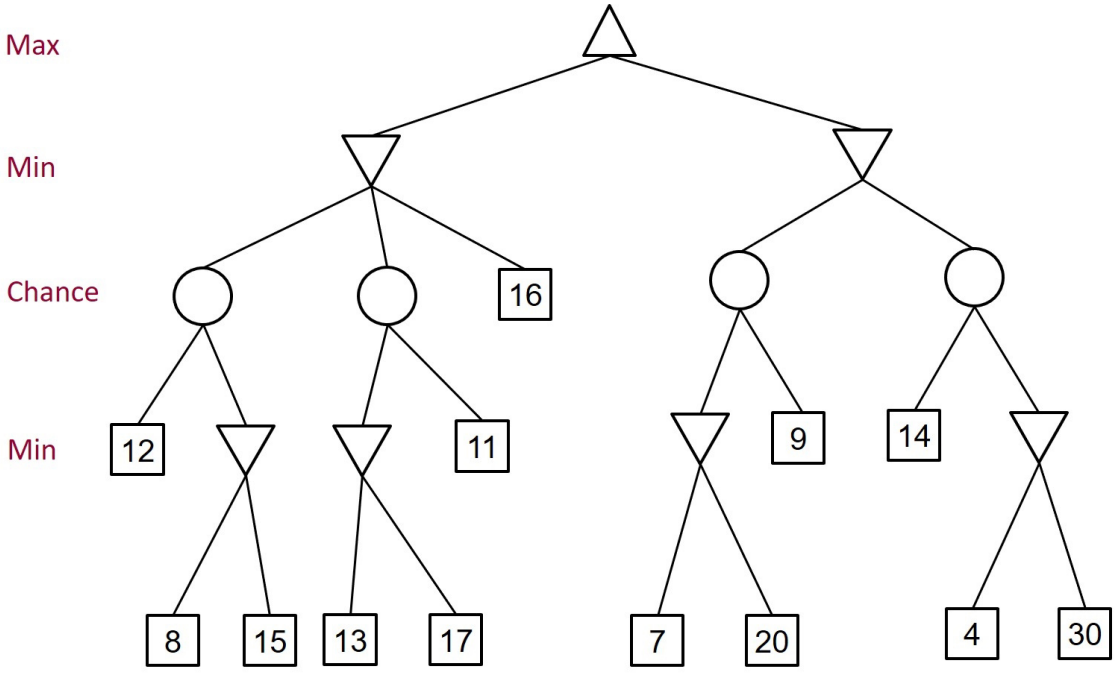
Question 3 Games

(a) Consider the following minimax game tree. There are two players Max and Min; the player Max wants to maximise the utility and the player Min wants to minimise the utility. The tree has five layers and we can use Lm-n to denote the nth node from left to right in the layer m; for example, the root node can be denoted by L1-1, the first node at the bottom layer (with value 8) can be denoted by L5-1, and the fourth node at the fourth layer (with value 11) can be denoted by L4-4. Give the value of the root node (L1-1) and the values of the two nodes at the second layer (L2-1 and L2-2). **[2 marks]**



(b) We use the alpha-beta pruning algorithm to prune the tree. List all the pruned nodes. Assume that child nodes are visited from left to right. **[10 marks]**

(c) Now suppose that the Max player in the third layer is replaced by a Chance player. We then have the following game tree. Let us say that the chance nodes go their different children with the same probability. Give the value of the root node (L1-1) and the values of the two nodes at the second layer (L2-1 and L2-2). **[2 marks]**



(d) Can this tree get pruned (assume that child nodes are visited from left to right)? If yes, list all the pruned nodes; If no, give your explanation. **[6 marks]**