Artificial Intelligence

1. Answer the following.
   - **A.** Define ‘intelligence’, ‘artificial intelligence’, and ‘intelligence agent.’
   - **B.** Explain why robot vacuum cleaners and ChatGPT are intelligent agents.
   - **C.** Explain the differences between machine learning and deep learning.
   - **D.** Compare and explain how AI software (or AI programs) differs from traditional software

.

2. Compare the following search strategies.
   - **A.** Exhaustive search
   - **B.** Heuristic search
   - **C.** Monte Carlo search
   - **D.** Local Search

3. Consider the following sentences:

   - John likes all kinds of food.
   - Apples are food.
   - Chicken is food.
   - Anything anyone eats and isn’t killed by is food.
   - Bill eats peanuts and is still alive.
   - Sue eats everything Bill eats.

   - **A.** Translate these sentences into formulas in predicate logic.
   - **B.** Prove that John likes peanuts using backward chaining.
   - **C.** Convert the formulas of part **A** into clause form.

4. Describe and compare the following two different approaches to artificial intelligence.
   - **A.** Knowledge-based approach.
   - **B.** Data-oriented approach.

5. Explain why search technique is one of important techniques in AI researches.

6. Provide three specific examples of how the search algorithm plays important roles in real-world applications or academic research.

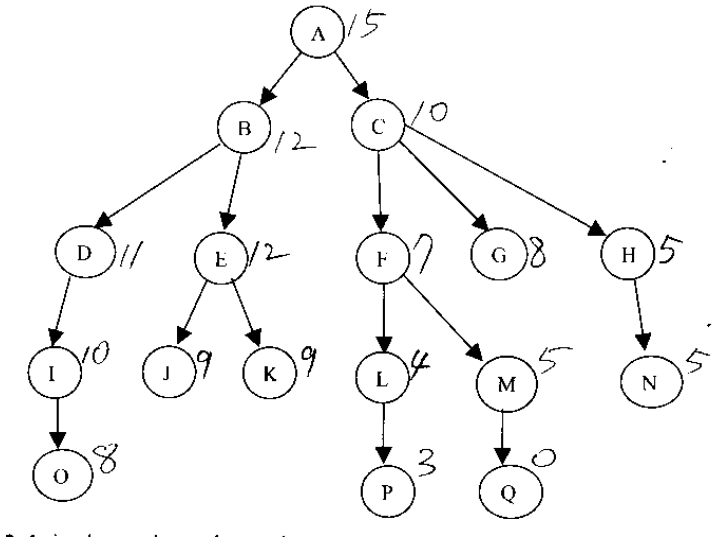
7. Define the following terms in your own words: state, state space, search technique, initial state, goal state, and objective function.

8. Given the following search tree (such that the number next to each node represents the cost of the node),

A. Show the list of nodes visited, using breadth-first search

B. Show the list of nodes visited, using hill-climbing search

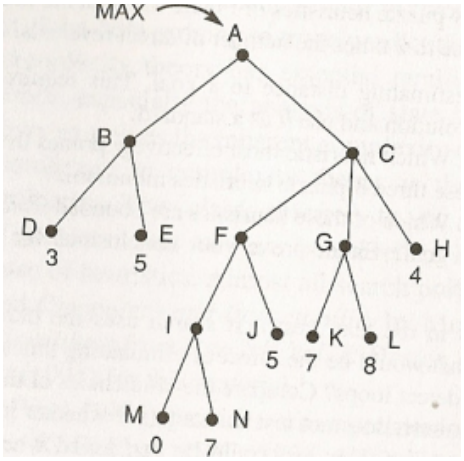
C. Show the list of nodes visited, using best-first-search



9. Given the following sub tree in a game search tree,

A. Apply minmax algorithm to the tree

B. Apply alpha-beta pruning to the tree



10. Answer the following questions about decision tree

A. Build a decision tree (ID3 or C4.5) based on the following dataset.

Class	Size	Color	Shape
A	small	yellow	round
A	big	yellow	round
A	big	red	round
A	small	red	round
B	small	black	round
B	big	black	cube
B	big	yellow	cube
B	big	black	round
B	small	yellow	cube

B. Does the ID3 algorithm always produce a consistent decision tree for a training data set? Justify your answer.

11. Explain the followings.

A. Definition of VC (Vapnik–Chervonenkis) dimension

B. The VC dimension of a linear classifier for n dimensional input space

12. According to the following classification rule, the class of the input x is predicted as the one with the maximum posterior probability.

$$y(x) = \arg\max_i P(\text{class} = i|x)$$

A. For binary classification problems, prove that the above classification rule is optimal. In what sense is it optimal?

B. Is the above classification rule still optimal for multiclass classification? Justify your answer.

13. Consider the following two-dimension samples of training data for naive Bayes classifiers.

x_1	x_2	class
5	4	1
6	5	1
6	5	2
6	6	1
6	6	2
6	6	2
7	6	1
7	6	2
7	6	2
8	7	2

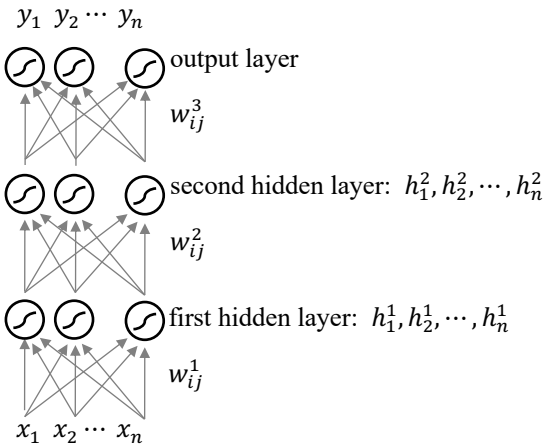
A. Assume that x_1 and x_2 follows categorical distributions, and a naive Bayes classifier is trained using a maximum likelihood estimation method. Compute $p(\text{class} = 1 | x_1 = 7, x_2 = 5)$ and $p(\text{class} = 2 | x_1 = 7, x_2 = 5)$.

B. Assume that x_1 and x_2 follows Gaussian distributions, and a naive Bayes classifier is trained using a maximum likelihood estimation method. Compute the mean vector and covariance matrix for each class.

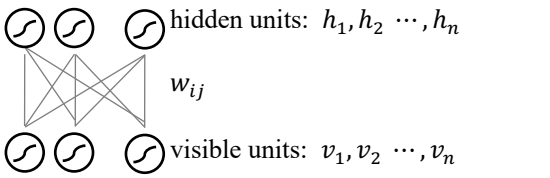
14. There are two coins in an urn. A random experiment of drawing a coin from the urn and tossing it is repeated three times with replacement of the coins. Suppose that the result of the random experiment is $D = \{H, H, T\}$. Run one iteration of the E-step and M-step of the expectation maximization (EM) algorithm using D as a training data set. What are the values of $p(H|C_1)$ and $p(C_1)$? Assume the following initial probabilities.

$p(H|C_1) = 0.6$ (probability of head when tossing coin C_1)
$p(H|C_2) = 0.4$ (probability of head when tossing coin C_2)
$p(C_1) = 0.5$ (probability of drawing coin C_1 from the urn)

15. Derive a weight update rule for the first hidden layer when the following artificial neural network is to be trained using the error back-propagation algorithm.



16. Consider the following restricted Boltzmann machine (RBM).

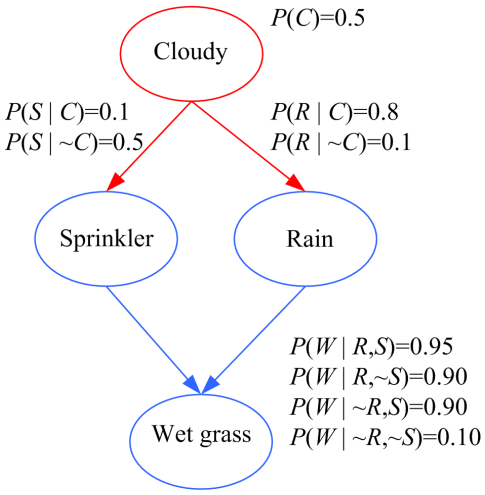


A. Derive a weight update rule when the objective function is maximum likelihood (ML).

B. Derive a weight update rule when the objective function is Kullback-Leibler (KL) divergence.

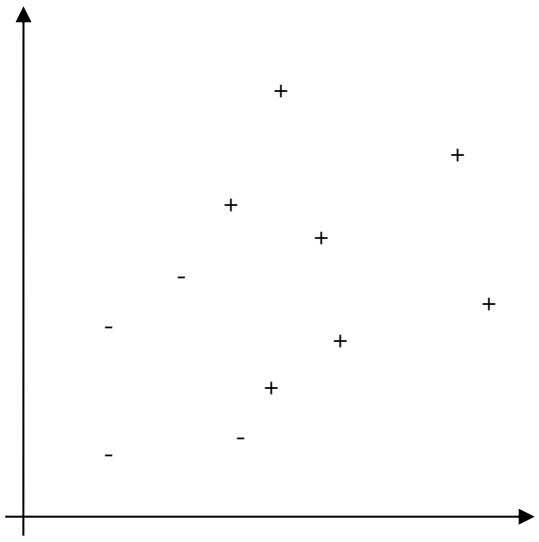
C. Derive a weight update rule when the objective function is contrastive divergence (CD).

17. Compute $p(C|W)$ in the following Bayesian belief network.



18. Find all support vectors in the following figure when C is equal to zero.

$$L_p = \frac{1}{2}\|w\|^2 + C\sum_t \xi^t - \sum_t \alpha^t[\gamma^t(w^T x^t + w_o) - 1 + \xi^t] - \sum_t \mu^t \xi^t$$



19. Describe why NLP is hard with your own example in terms of ambiguities in each phases of NLP.

20. Design your document filtering system by using a statistical model.

21. Explain two major approaches to NLP and their pros and cons.

22. Explain statistical models for part-of-speech tagging, syntactic tagging, and machine translation by using noisy channel.

23. Explain specifically what the number of hidden layers and the number of hidden nodes in neural networks mean and how affect the final results.

24. Explain the "Vanishing gradient problem" in deep neural networks and discuss possible solutions.

25. Discuss when bottlenecks are useful in deep neural networks with examples.

26. Define "skip connection" and discuss when it is beneficial. Show at least two architectures that have skip connections.

27. Explain why ReLU is less likely to suffer from the gradient vanishing problem than the sigmoid function.

28. Draw ReLU, LeakyReLU, and ELU with details such as intercepts and functions values. Point out one problem of ReLU and explain why the problem can be addressed by LeakyReLU or ELU.

29. Describe briefly, from a technical perspective, how data-driven and agent-driven artificial intelligence can pose challenges from a reliability, fairness, and safety perspective. Describe any technical techniques or attempts to avoid them.