CS 331 Mock Final Solution
Spring 2023

You have 110 minutes to complete this exam. You are only allowed to use your textbook, your notes, your assignments and solutions to those assignments during this exam. Please read all the questions carefully and follow directions. For all questions, partial credit can sometimes be granted when you show your work, thought process, assumptions, etc.

Section I: Pre-Midterm questions

1. Circle true or false below each question.

a) Goal-based agents can reason about a sequence of steps required to reach a desired state.

True False

b) Breadth-first search is complete for both the tree-based and graph-based search algorithms.

True False

c) The asymptotic time complexity of depth-first search (DFS) is better than that of breadth-first search (BFS).

True **False**

d) In iterative-deepening depth-first tree search (IDDFS), increasing the depth limit by more than 1 at each iteration would sacrifice the algorithm's optimality guarantee.

True False

e) Suppose you are given two admissible heuristics for A* search: h1 and h2. If h1 dominates h2, then using h1 results in a shorter solution path.

True **False**

f) Hill climbing is the same as beam search with beam width 1 (k=1).

True False

g) The performance of genetic algorithms depends strongly on the state representation.

True False

h) Gradient descent finds the global minimum of a search space by following the gradient to the lowest point of the cost function.

True **False**

i) The minimax algorithm assumes the opponent is playing rationally.

True False

j) The order in which successors are generated affects the average complexity of the alpha-beta pruning algorithm.

True False

k) The expectiminimax algorithm incorporates stochastic elements of game play by adding a branch to the search tree for every possible outcome of the stochastic event.

True False

Section II: Probability

2. One test for COVID-19 antibodies reports a sensitivity (chance of getting a positive result given that antibodies are present) of 88% and a specificity (chance of getting a negative result given that antibodies are absent) of 98.8%. Suppose for a given population, 4% of individuals have antibodies. Define two random variables: A = antibodies present (true/false), T = test result (pos/neg).

a) Compute P(A=false).

1 – P(A=true) = 0.96

b) Compute P(T=pos | A=false).

1 – P(T=neg | A=false) = 0.012

c) Compute P(T=neg).

P(A=true, T=neg) + P(A=false, T=neg)
P(A=true) P(T=neg | A=true) + P(A=false) P(T=neg | A=false)

0.04 * (1 – P(T=pos | A=true)) + 0.96 * 0.988
0.04 * 0.12 + 0.96 * 0.988
0.95328

d) Compute P(A=false | T=pos).

P(T=pos | A=false) P(A=false) / P(T=pos)
0.012 * 0.96 / (1 – 0.95328)
0.24658

III. Logic
3. Are the following statements true or false? Justify your answer.
a) If an inference rule is sound, then it is also complete.

True **False**
Modus ponens is sound but not complete. It cannot infer P from $P \wedge Q$.

b) $\neg(P \Rightarrow Q) \equiv P \wedge \neg Q$.

True False
The l.h.s. and r.h.s. are both true in only one model where P is true and Q is false.

c) Suppose a KB entails a sentence A. Then in the set of all models where A is true, we know that the KB is also true.

True **False**
It is the other way. In all models where KB is true, A is true.

d) $(O \Rightarrow S) \equiv (\neg O \vee S) \wedge (\neg U \vee U)$ is valid.

True False
Because $(\neg U \vee U)$ is always true, the r.h.s. is equivalent to $(\neg O \vee S)$. L.h.s is also equivalent to it.

e) If the resolution algorithm finds the empty clause, then entailment holds. [2 points]

True False
True, because that would prove that the KB is inconsistent with the negated goal.

4. A. Express the following facts in propositional logic using suitable propositions and convert them to clausal form (CNF)

a) You should have an arrow to shoot the Wumpus
$ShootArrow \Rightarrow HaveArrow$ Translates to
$\neg ShootArrow \vee HaveArrow$ ---------------------------[1]

b) If you shoot the arrow and your aim is good the Wumpus dies.
$ShootArrow \wedge GoodAim \Rightarrow WumpusDies$ Translates to
$\neg ShootArrow \vee \neg GoodAim \vee WumpusDies$ -----------------------[2]

c) You shot an arrow
$ShootArrow$ -------------------------[3]

d) The Wumpus did not die
$\neg WumpusDies$ ------------------------[4]

B. Show that your aim is not good from the above clauses using resolution
Negated goal: $GoodAim$ -----------------------[5]

$GoodAim$ [5] $\neg ShootArrow \vee \neg GoodAim \vee WumpusDies$ [2]

$\neg ShootArrow \vee WumpusDies$ $\neg WumpusDies$ [4]

$ShootArrow$ [3]

$\neg ShootArrow$

$\square$

Section IV: Bayesian Networks

5. Consider the following data on how class attendance, turning assignments in on time, and students' class performances are related:

A	**B**	**Number of cases in dataset**
yes	low	1
yes	medium	5
yes	high	2
no	low	4
no	medium	3
no	high	0

a) Compute P(B=medium | A = yes) **without** using uniform Dirichlet priors.

5 / 8 = 0.625

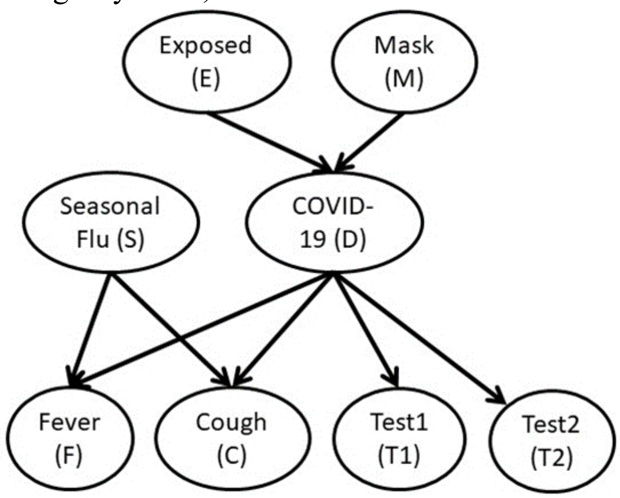
b) Compute P(B=medium | A = yes) using uniform Dirichlet priors.

(5 + 1) / (8 + 3) = 6 / 11 = 0.54545

c) Compute P(A=no | B = high) using uniform Dirichlet priors.

(0 + 1) / (2 + 2) = 1 / 4 = 0.25

6. Consider the following Bayes net, where each variable can be true or false.



a) Give an expression for the joint probability distribution represented by this network in as simplified a form as possible.

P(E) P(M) P(S) P(D | E, M) P(F | S, D) P(C | S, D) P(T1 | D) P(T2 | D)

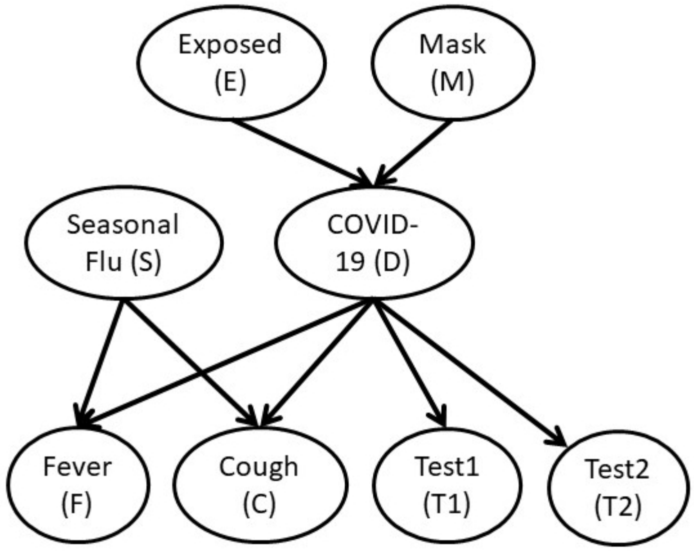
b) Give an expression for P(M=false, D=true, T1=true), simplified as much as possible.

Since S, F, C, and T2 are not ancestors of M, D, or T1, they will not be in the expression. You can show this more rigorously by including them in the expression and then simplifying.

$\sum_e [P(E = e)P(M = false)P(D = true \,|E = e, M = false)\; P(T1 = true \,|P(D = true)]$

$P(M = false)P(T1 = true \,|D = true) \sum_e [P(E = e)P(D = true \,|E = e, M = false)]$

7. Use the Bayesian network below to determine whether or not the following conditional independence relationships hold or not. Show the blocked/unblocked paths for partial credit.



a) I(S, D | { })

True.

S-C-D blocked by C
S-F-D blocked by F

b) I(S, D | {F, M})

False.

S-C-D blocked by C
S-F-D not blocked

c) I(F, C | {S, T1})

False.

F-S-C blocked by S
F-D-C not blocked

Section V: Machine Learning

8. Consider the decision tree learning with a class variable C and two attributes A and B, all of which are binary. Suppose we have the following counts of different data items. For example the first cell in the C=0 column represents the number of items where A = 0, B = 0, C =0. Calculate the Information Gains, I(C; A) and I(C; B). Which variable A or B would the decision tree learning algorithm choose to split the data at the root and why?

	C = 0	C = 1
A = 0, B = 0	10	3
A = 0, B = 1	11	5
A = 1, B = 0	6	10
A = 1, B = 1	8	12
Total	35	30

Recall that $I(C; A) = H(C) - \sum_a P(a)H(C|a)$

$$H(C) = -P(C = 0)\log P(C = 0) - P(C = 1)\log P(C = 1)$$
$$= -\frac{35}{65}\log\frac{35}{65} - \frac{30}{65}\log\frac{30}{65} = 0.9957$$

$$H(C|A = 0) = -P(C = 0|A = 0)\log P(C = 0|A = 0) - P(C = 1|A = 0)\log P(C = 1|A = 0)$$

Therefore, $H(C|A = 0) = -\frac{21}{29}\log\frac{21}{29} - \frac{8}{29}\log\frac{8}{29} = 0.8497$

Similarly, $H(C|A = 1) = -\frac{14}{36}\log\frac{14}{36} - \frac{22}{36}\log\frac{22}{36} = 0.9641$

$$P(A = 0)H(C|A = 0) + P(A = 1)H(C|A = 1) = \frac{29}{65}0.8497 + \frac{36}{65}0.9641 = 0.9131$$

Therefore, $I(C; A) = 0.9957 - 0.9131 = 0.0826$

Similarly, $H(C|B = 0) = -\frac{16}{29}\log\frac{16}{29} - \frac{13}{29}\log\frac{13}{29} = 0.9923$

And, $H(C|B = 1) = -\frac{19}{36}\log\frac{19}{36} - \frac{17}{36}\log\frac{17}{36} = 0.9978$

$$P(B = 0)H(C|B = 0) + P(B = 1)H(C|B = 1) = \frac{29}{65}0.9923 + \frac{36}{65}0.9978 = 0.9953$$

Therefore, $I(C; B) = 0.9957 - 0.9953 = 0.0004$

Since $I(C; A)$ is higher than $I(C; B)$, the learning algorithm chooses to split on A at the root node.

9. Are the following statements true or false? Justify.

   a. Machine learning algorithms can sometimes make mistakes even if they have a lot of training data

True. There may be inherent noise in the data or the hypothesis space, eg, linear functions, may not be able to express the true concept. In both cases, errors are inevitable.

   b. Overfitting happens when the performance on the test set is better than the performance on training set

False. Overfitting happens when the performance on the training set is better.

   c. Naïve Bayes is a machine learning approach based on probability theory.

True. Naïve Bayes is a conditional probability estimator which can be viewed as statistical machine learning algorithm.

   d. Overfitting can happen even when there is abundance of data

True. Noisy data can also cause overfitting.

   e. Machine learning algorithms only work when the target function is deterministic

False. Naïve Bayes is an example where the target function is stochastic. It works well in many cases.