

CS171, Intro to A.I. — Final Exam — Summer Quarter, 2016

YOUR NAME: _____

YOUR ID: _____ ID TO RIGHT: _____ ROW: _____ SEAT NO.: _____

The exam will begin on the next page. **Please, do not turn the page until told.**

When you are told to begin the exam, please check first to make sure that you have all 13 pages, as numbered 1-13 in the bottom-left corner of each page.

The exam is closed-notes, closed-book. No calculators, cell phones, electronics.

Please clear your desk entirely, except for pen, pencil, eraser, an optional blank piece of paper (for optional scratch pad use), and an optional water bottle.

Please turn off all cell phones now.

This page summarizes the points available for each question so you can plan your time.

1. (5 pts total, -1 for each wrong answer, but not negative) Mini-Max, Alpha-Beta Pruning.
2. (6 pts total, 2 pts each) Machine Learning Classifier Decision Boundaries.
3. (12 pts total) Decision tree, underfitting and overfitting.
4. (10 pts total, 2 pts each) Constraint Satisfaction Problems. NorCal Counties
5. (10 pts total, 2 pts each) Execute Tree Search through this graph.
6. (10 pts total, 1 pt each) Bayesian Networks.
7. (5 pts total, -1 pt for each error, but not negative) Resolution Theorem Proving: Cake Theft.
8. (5 pts total, 1 pt each) English to FOL Conversion.
9. (3 pts total, 1 pt each) Model Complexity, Underfitting, Overfitting.
10. (8 pts total, -1 pt each wrong answer, but not negative) Search Properties.
11. (4 pts total, 1 pt each) Task Environment.
12. (12 pts total, 1 pt each) Mini-Max Game Search.
13. (6 pts total, 1 pt each) Valid, Unsatisfiable, Satisfiable.
14. Probability formulae (4 pts total, 1 pt each).

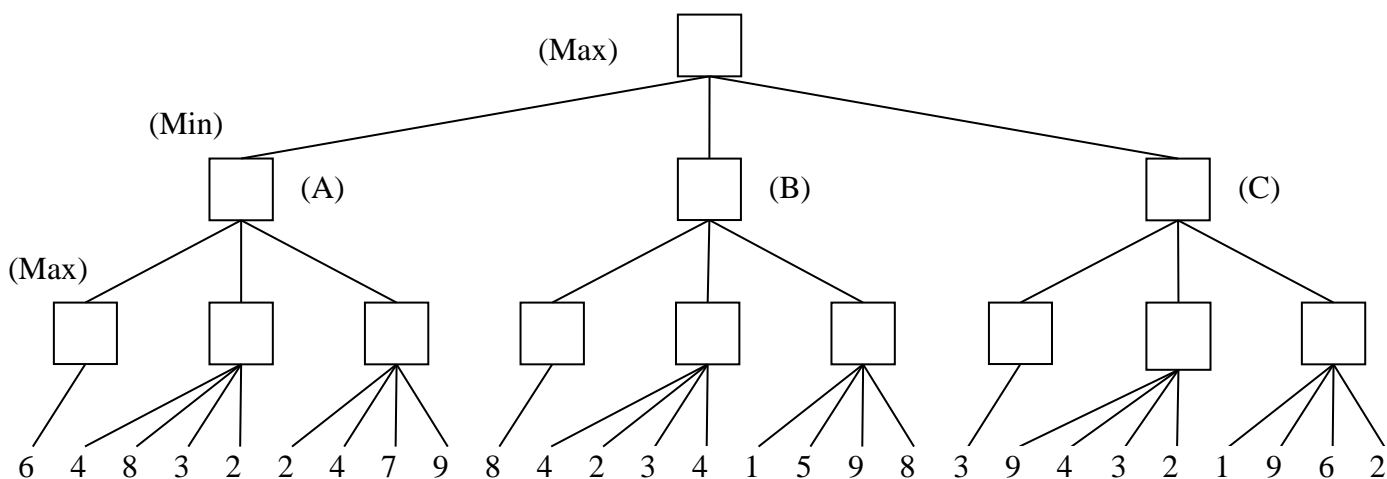
The Exam is printed on both sides to save trees! Work both sides of each page!

1. (5 pts total, -1 for each error, but not negative) Mini-Max, Alpha-Beta Pruning.

In the game tree below it is **Max's** turn to move. At each leaf node is the estimated score of that resulting position as returned by the heuristic static evaluator.

(1) Perform Mini-Max search and label each branch node with its value.

(2) Cross out each leaf node that would be pruned by alpha-beta pruning.



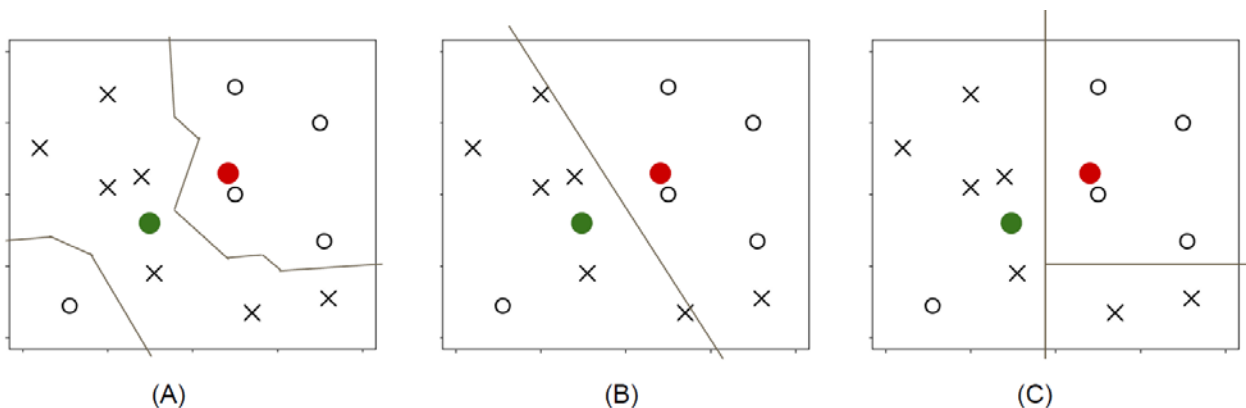
2. (6 pts total, 2 pts each) Machine Learning Classifier Decision Boundaries.

Match each decision boundary below with the classifier type most likely to produce such a decision boundary. The class means (centroids) are shown as filled circles.

2.a. (2 pts) Decision Tree Classifier. (Write A, B, or C) _____

2.b. (2 pts) Minimum Distance Classifier. (Write A, B, or C) _____

2.c. (2 pts) Nearest Neighbor Classifier. (Write A, B, or C) _____



****** TURN PAGE OVER AND CONTINUE ON THE OTHER SIDE ******

3. (12 pts total) Decision tree, underfitting and overfitting. Consider the following set of training examples. There are two features: A and B. The binary target variable (also known as the class label) is $Y \in \{1, 0\}$. You came up with a decision tree classifier that is shown in Figure 1. (It is labeled DT2 because it has two splits, as will be used below).

A	B	Y
3	3	0
3	6	0
5	5	0
7	2	0
9	7	0
9	3	0
2	7	1
7	6	1
8	5	1
10	5	1
10	7	1
11	2	1

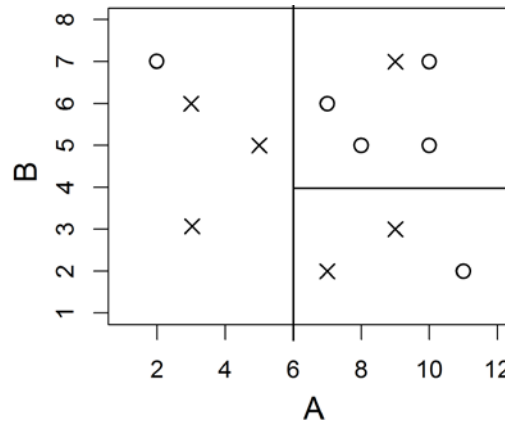
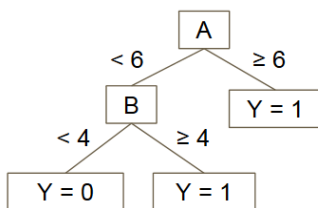


Figure 1: Decision Tree DT2 (2 splits)

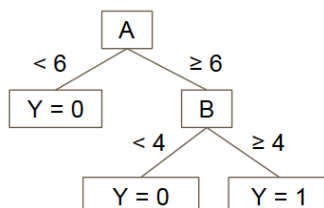
3.a. (2 pts) In Figure 1, do circles represent $Y=0$ or $Y=1$? (Write 0 or 1) _____

3.b. (2 pts) Consider the three decision trees shown below (labeled 1, 2, and 3). Which tree below represents the same decision tree as that which is shown in Figure 1? When a leaf node is not pure, take the majority class as its label.

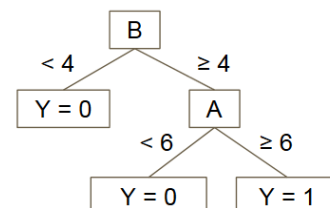
(Write 1, 2, or 3) _____



(1)



(2)



(3)

3.c. (2 pts total, 1 pt each) Classify the new examples below using the decision tree shown in Figure 1. When a decision region is not pure, take the majority class as its label. Indicate your answer as a circle \bigcirc or an X.

3.c.i. (1 pt) What class is $[A = 2, B = 7]$? (write circle \bigcirc or X) _____

3.c.ii. (1 pt) What class is $[A = 10, B = 2]$? (write circle \bigcirc or X) _____

**** THIS PROBLEM CONTINUES ON THE NEXT PAGE ****

Your friends came up with 2 different decision trees, shown below in Figures 2 and 3. Together with Figure 1, you now have 3 decision trees. These decision trees are named by their number of splits: The decision tree in Figure 1 has 2 splits and is named DT2, the decision tree in Figure 2 has 1 split and is named DT1, and the decision tree in Figure 3 has 3 splits and is named DT3.

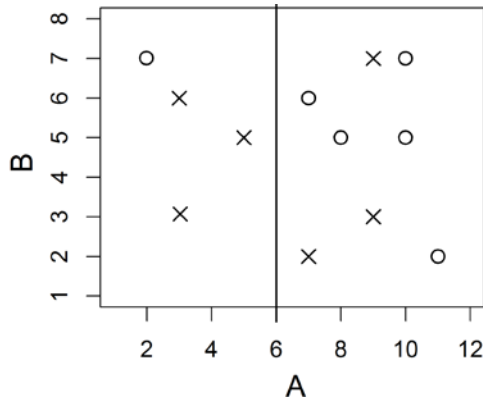


Figure 2: Decision Tree DT1 (1 split)

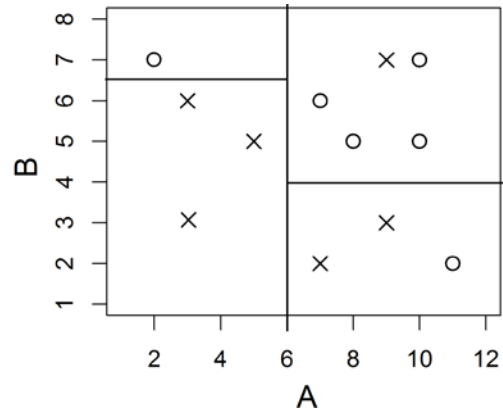


Figure 3: Decision Tree DT3 (3 splits)

In order to evaluate whose classifier performs the best, you decided to examine the training error and the test error of the 3 decision trees. Here you may simply use the fraction of misclassified data cases as “error” (eg, there are 12 training examples, so if 1 training data case is classified incorrectly, then training error = 1/12) . Two test data cases are given to you in the table below. (Assume you don’t know their class label before classifying them into classes. Only use the label to examine your test error).

A	B	Y
4	7	0
7	3	0

3.d. (4 pts total, 1 pt each) Fill in the table below. Use fractions of the form N/12 (training set error) or N/2 (test set error). When a decision region is not pure, take the majority class as its label. DT1 (Figure 2) is done for you as an example.

	DT1 (1 split)	DT2 (2 splits)	DT3 (3 splits)
Training error	4/12		
Test error	1/2		

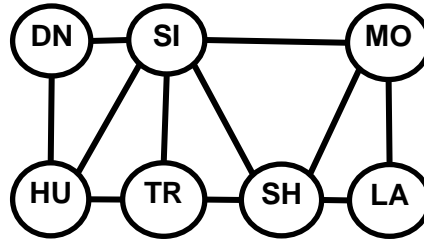
3.e. (2 pts total, 1 pt each)

Which decision tree classifier is likely overfitting? (Write DT1, DT2, or DT3) _____

Which decision tree classifier is likely underfitting? (Write DT1, DT2, or DT3) _____

**** TURN PAGE OVER AND CONTINUE ON THE OTHER SIDE ****

4. (10 pts total, 2 pts each) Constraint Satisfaction Problems. NorCal Counties.



DN = Del Norte
 HU = Humboldt
 LA = Lassen
 MO = Modoc
 SH = Shasta
 SI = Siskiyou
 TR = Trinity

You are a map-coloring robot assigned to color this Northern California map. Adjacent counties must be colored a different color (R=Red, B=Blue, G=Green). The constraint graph is shown.

4.a. (2pts total, -1 each wrong answer, but not negative) FORWARD CHECKING.

Cross out all values that would be eliminated by Forward Checking, after variable SH has just been assigned value G as shown:

DN	HU	LA	MO	SH	SI	TR
R G B	R G B	R G B	R G B	G	R G B	R G B

4.b. (2pts total, -1 each wrong answer, but not negative) ARC CONSISTENCY.

SH and SI have been assigned values, but no constraint propagation has been done. Cross out all values that would be eliminated by Arc Consistency (AC-3 in your book).

DN	HU	LA	MO	SH	SI	TR
R G B	R G B	R G B	R G B	G	B	R G B

4.c. (2pts total, -1 each wrong answer, but not negative) MINIMUM-REMAINING-VALUES HEURISTIC.

Consider the assignment below. HU is assigned and constraint propagation has been done. List all unassigned variables that might be selected by the Minimum-Remaining-Values (MRV) Heuristic:

DN	HU	LA	MO	SH	SI	TR
R B	G	R G B	R G B	R G B	R B	R B

4.d. (2pts total, -1 each wrong answer, but not negative) DEGREE HEURISTIC.

Consider the assignment below. (It is the same assignment as in problem 4.c above.) HU is assigned and constraint propagation has been done. List all unassigned variables that might be selected by the Degree Heuristic:

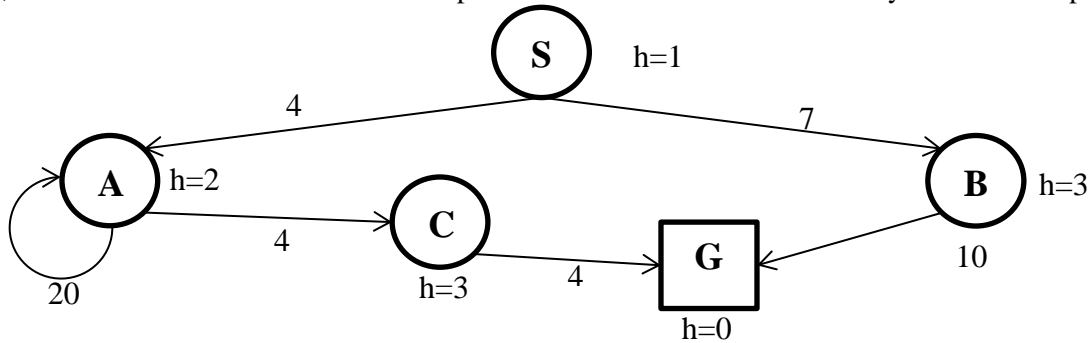
DN	HU	LA	MO	SH	SI	TR
R B	G	R G B	R G B	R G B	R B	R B

4.e. (2pts total) MIN-CONFLICTS HEURISTIC. Consider the complete but inconsistent assignment below. SI has just been selected to be assigned a new value during local search for a complete and consistent assignment. What new value would be chosen below for SI by the Min-Conflicts Heuristic?

DN	HU	LA	MO	SH	SI	TR
B	G	R	G	G	?	B

5. (10 pts total, 2 pts each) Execute Tree Search through this graph. (Do not remember visited nodes.) Step costs are given next to each arc. Heuristic values are next to each node (as $h=x$). The successors of each node are indicated by the arrows out of that node. Successors are returned in left-to-right order.

For each search strategy, show the order in which nodes are expanded (i.e., to expand a node means that its children are generated), ending with the goal node that is found. Show the path from start to goal, or write "None". Give the cost of the path found. The first one is done for you as an example.



5.a. BREADTH FIRST SEARCH:

Order of node expansion: S A B G

Path found: S B G Cost of path found: 17

5.b. (2 pts) DEPTH FIRST SEARCH:

Order of node expansion: _____

Path found: _____ Cost of path found: _____

5.c. (2 pts) UNIFORM COST SEARCH:

Order of node expansion: _____

Path found: _____ Cost of path found: _____

5.d. (2 pts) GREEDY (BEST-FIRST) SEARCH:

Order of node expansion: _____

Path found: _____ Cost of path found: _____

5.e. (2 pts) ITERATED DEEPENING SEARCH:

Order of node expansion: _____

Path found: _____ Cost of path found: _____

5.f. (2 pts) A* SEARCH:

Order of node expansion: _____

Path found: _____ Cost of path found: _____

****** TURN PAGE OVER AND CONTINUE ON THE OTHER SIDE ******

6. (10 pts total, 1 pt each) Bayesian Networks.

Draw the Bayesian Network that corresponds to the conditional probability equation.

6.a. $P(B|A,C) P(A) P(C|D) P(D)$ A B C D

6.b. $P(A) P(B) P(C) P(D)$ A B C D

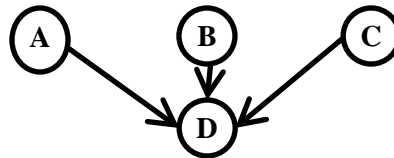
6.c. $P(A|B) P(C|B) P(B) P(D)$ A B C D

6.d. $P(D|C) P(C|B) P(B|A) P(A)$ A B C D

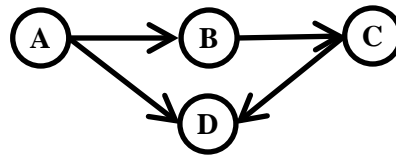
6.e. $P(B|A) P(A) P(C|D) P(D)$ A B C D

Write down the factored conditional probability equation that corresponds to the graphical Bayesian Network shown.

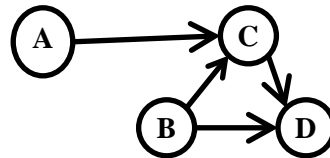
6.f. _____



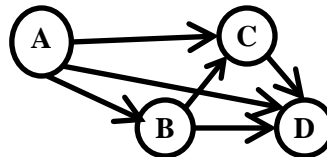
6.g. _____



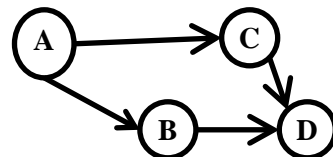
6.h. _____



6.i. _____



6.j. _____



7. (5 pts total, -1 pt for each error, but not negative) Resolution Theorem Proving: Cake Theft.
 (http://www.brainbashers.com) Chief Inspector Parker interviewed five local burglars to identify who stole Mrs. Archer's cake.

It was well known that each suspect told exactly one lie:

Arnold: It was not Edward. It was Brian. *Brian: It was not Charlie. It was not Edward.*
Charlie: It was Edward. It was not Arnold. *Derek: It was Charlie. It was Brian.*
Edward: It was Derek. It was not Arnold.

Use these propositional variables:

A=It was Arnold. **B**=It was Brian. **C**=It was Charlie. **D**=It was Derek. **E**=It was Edward.

You translate the evidence into propositional logic (recall that each suspect told exactly one lie):

Arnold: $(E \wedge B) \vee (\neg E \wedge \neg B)$ *Brian:* $(C \wedge \neg E) \vee (\neg C \wedge E)$
Charlie: $(\neg E \wedge \neg A) \vee (E \wedge A)$ *Derek:* $(\neg C \wedge B) \vee (C \wedge \neg B)$
Edward: $(\neg D \wedge \neg A) \vee (D \wedge A)$

At most one burglar stole the cake:

$(A \Rightarrow \neg B \wedge \neg C \wedge \neg D \wedge \neg E)$ $(B \Rightarrow \neg A \wedge \neg C \wedge \neg D \wedge \neg E)$ $(C \Rightarrow \neg A \wedge \neg B \wedge \neg D \wedge \neg E)$
 $(D \Rightarrow \neg A \wedge \neg B \wedge \neg C \wedge \neg E)$ $(E \Rightarrow \neg A \wedge \neg B \wedge \neg C \wedge \neg D)$

After converting to Conjunctive Normal Form, your Knowledge Base (KB) consists of:

$(E \vee \neg B)$ $(\neg E \vee B)$ $(C \vee E)$ $(\neg C \vee \neg E)$ $(\neg E \vee A)$ $(E \vee \neg A)$
 $(\neg C \vee \neg B)$ $(C \vee B)$ $(\neg D \vee A)$ $(D \vee \neg A)$
 $(\neg A \vee \neg B)$ $(\neg A \vee \neg C)$ $(\neg A \vee \neg D)$ $(\neg A \vee \neg E)$ $(\neg B \vee \neg C)$
 $(\neg B \vee \neg D)$ $(\neg B \vee \neg E)$ $(\neg C \vee \neg D)$ $(\neg C \vee \neg E)$ $(\neg D \vee \neg E)$

From Brian, it was Charlie or Edward. From Derek, it was Charlie or Brian. **Thus, it was Charlie.**

You will be asked to prove, **“It was Charlie.”** The goal is (C) . You adjoin the negated goal to your KB:
 $(\neg C)$

Produce a resolution proof, using KB and the negated goal, that “It was Charlie.”

Repeatedly choose two clauses, write one clause in the first blank space on a line, and the other clause in the second. Apply resolution to them. Write the resulting clause in the third blank space, and insert it into the knowledge base. Continue until you produce (C) . If you cannot produce (C) , then you have made a mistake. The shortest proof I know is only three lines. It is OK to use more lines, if your proof is correct.

(Note that in a real reasoning context, you would start five proofs simultaneously, one each trying to prove (A) , (B) , (C) , (D) , or (E) . Only the proof of (C) would succeed, answering the question.)

Resolve _____ and _____ to give _____.

Resolve _____ and _____ to give _____.

Resolve _____ and _____ to give _____.

Resolve _____ and _____ to give _____.

Resolve _____ and _____ to give _____.

Resolve _____ and _____ to give _____.

****** TURN PAGE OVER AND CONTINUE ON THE OTHER SIDE ******

8. (5 pts total, 1 pt each) English to FOL Conversion. For each English sentence below, write the FOL sentence that best expresses its intended meaning. Use Person(x) for “x is a person,” Food(x) for “x is food,” and Likes(x, y) for “x likes y.”

The first one is done for you as an example.

8.example. “Every person likes every food.”

$$\forall x \forall y [\text{Person}(x) \wedge \text{Food}(y)] \Rightarrow \text{Likes}(x, y)$$

8.a. (1 pt) “For every food, there is a person who likes that food.”

8.b. (1 pt) “There is a person who likes every food.”

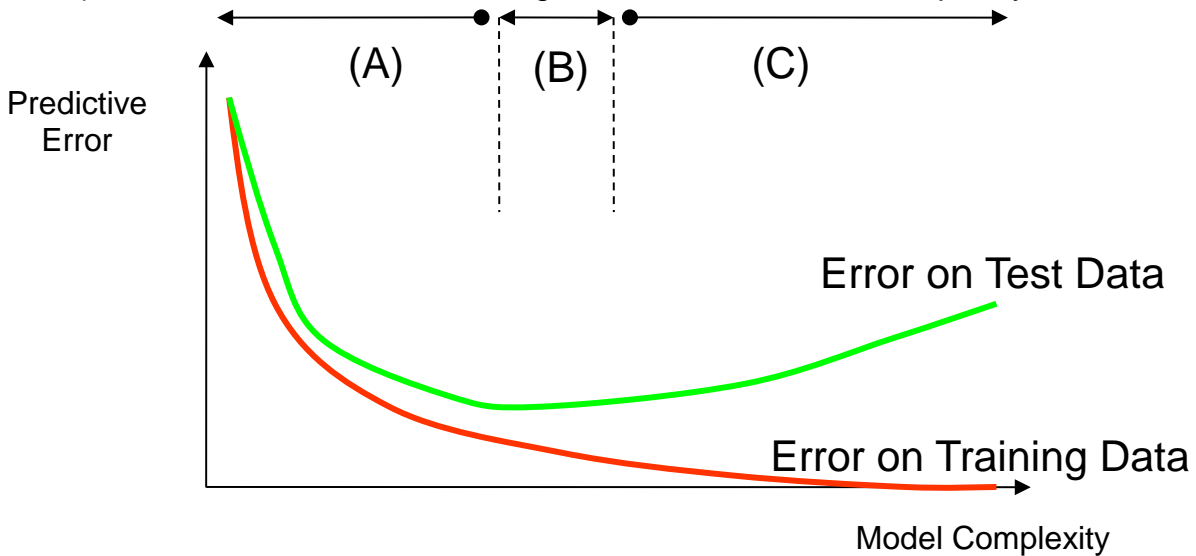
8.c. (1 pt) “Some person likes some food.”

8.d. (1 pt) “There is a food that every person likes.”

8.e. (1 pt) “For every person, there is a food that the person likes.”

9. (3 pts total, 1 pt each) Model Complexity, Underfitting, Overfitting.

Consider the hypothetical graph below of Predictive Error (y-axis) vs. Model Complexity (x-axis), and how Error on Test/Training Data varies as Model Complexity increases.



Label the following regions as A, B, or C, depending on whether they correspond to Overfitting, Underfitting, or Ideal Model Complexity:

9.a. Overfitting (1 pt) (Write A, B, or C) _____

9.b. Underfitting (1 pt) (Write A, B, or C) _____

9.c. Ideal Model Complexity (Write A, B, or C) _____

10. (8 pts total, -1 pt each wrong answer, but not negative) Search Properties.

Fill in the values of the four evaluation criteria for each search strategy shown. Assume a tree search where b is the finite branching factor; d is the depth to the shallowest goal node; m is the maximum depth of the search tree; C^* is the cost of the optimal solution; step costs are identical and equal to some positive ϵ ; and in Bidirectional search both directions use breadth-first search.

Note that these conditions satisfy all of the footnotes of Fig. 3.21 in your book.

Criterion	Complete?	Time complexity	Space complexity	Optimal?
Breadth-First				
Uniform-Cost				
Depth-First				
Iterative Deepening				
Bidirectional (if applicable)				

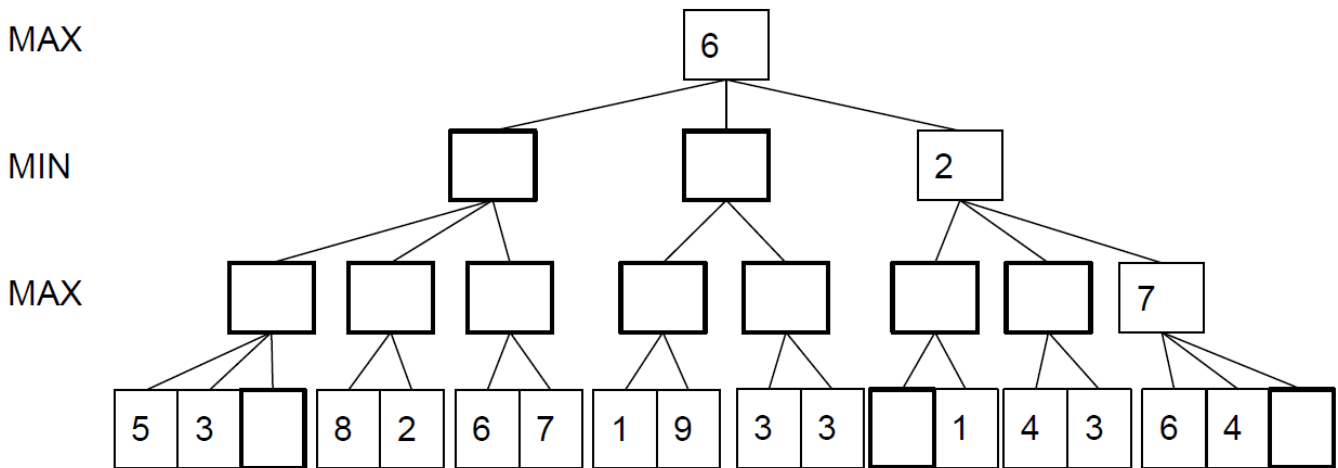
**** TURN PAGE OVER AND CONTINUE ON THE OTHER SIDE ****

11. (4 pts total, 1 pt each) Task Environment. Your book defines a task environment as a set of four things, with the acronym PEAS. Fill in the blanks with the names of the PEAS components.

P _____ E _____ A _____ S _____

12. (12 pts total, 1 pt each) Mini-Max Game Search. In the game tree below, it is MAX's turn to move. At each leaf node is the estimated score of that resulting position as returned by the heuristic static evaluator. Each empty box in the game tree below has a single unique value that is consistent with the rest of the tree.

Fill in the values of the empty boxes below so that the tree is consistent.



13. (6 pts total, 1 pt each) Valid, Unsatisfiable, Satisfiable.

For each sentence below, write V=valid if the sentence is valid, U=unsatisfiable if the sentence is unsatisfiable, and S=satisfiable if the sentence is satisfiable but is neither valid nor unsatisfiable.

- 13.a. (1 pt) (write V, U, or S) _____ (A OR (\neg A))
- 13.b. (1 pt) (write V, U, or S) _____ (A AND (\neg A))
- 13.c. (1 pt) (write V, U, or S) _____ (A AND (\neg B))
- 13.d. (1 pt) (write V, U, or S) _____ ((A AND (\neg A)) => B)
- 13.e. (1 pt) (write V, U, or S) _____ ((A OR (\neg A)) => B)
- 13.f. (1 pt) (write V, U, or S) _____ ((A OR (\neg A)) => (A AND (\neg A)))

14. Probability formulae (4 pts total, 1 pt each).

For each term on the left, write the letter of the best description on the right.

The first one is done for you as an example.

A	Probability Theory	A	Assigns each sentence a degree of belief ranging from 0 to 1
	Conditional independence	B	$P(a \wedge b) = P(a) P(b)$
	Independence	C	$P(a \wedge b c) = P(a c) P(b c)$
	Product rule (chain rule)	D	$P(a b) = P(b a) P(a) / P(b)$
	Bayes' rule	E	$P(a \wedge b \wedge c) = P(a b \wedge c) P(b c) P(c)$

**** THIS IS THE END OF THE FINAL EXAM. HAPPY SUMMER!!! ****