Below, for each problem on this Midterm Exam, “Perfect” is the percentage of students who received full credit, “Partial” is the percentage who received partial credit, and “Zero” is the percentage who received zero credit.

(Due to rounding or other exceptional reasons, values below may be only approximate estimates.)

Problem 1 Perfect: ~35.1% (~20 Students), Partial: ~64.9% (~37 Students), Zero: ~0% (~0 Students)

Problem 2 Perfect: ~22.8% (~13 Students), Partial: ~75.4% (~43 Students), Zero: ~1.8% (~1 Students)

Problem 3 Perfect: ~78.9% (~45 Students), Partial: ~21.0% (~12 Students), Zero: ~0% (~0 Students)

Problem 4 Perfect: ~84.2% (~48 Students), Partial: ~15.8% (~9 Students), Zero: ~0% (~0 Students)

Problem 5 Perfect: ~91.2% (~52 Students), Partial: ~7.0% (~4 Students), Zero: ~1.8% (~1 Students)

Problem 6 Perfect: ~29.8% (~17 Students), Partial: ~70.2% (~40 Students), Zero: ~0% (~0 Students)

Problem 7 Perfect: ~66.7% (~38 Students), Partial: ~33.3% (~19 Students), Zero: ~0% (~0 Students)

Problem 8 Perfect: ~31.6% (~18 Students), Partial: ~68.4% (~39 Students), Zero: ~0% (~0 Students)

Problem 9 Perfect: ~35.1% (~20 Students), Partial: ~64.9% (~37 Students), Zero: ~0% (~0 Students)

Problem 10 Perfect: ~50.9% (~29 Students), Partial: ~49.1% (~28 Students), Zero: ~0% (~0 Students)

Problem 11 Perfect: ~57.9% (~33 Students), Partial: ~21.1% (~12 Students), Zero: ~21.1% (~12 Students)

Problem 12 Perfect: ~94.7% (~54 Students), Partial: ~5.3% (~3 Students), Zero: ~0% (~0 Students)

CS-171, Intro to A.I. — Mid-term Exam — Summer Quarter, 2016

YOUR NAME: ____

YOUR ID: __________ ID TO RIGHT: __________ ROW: ________ SEAT: ________

The exam will begin on the next page. <u>Please, do not turn the page until told.</u>

When you are told to begin the exam, please check first to make sure that you have all eleven pages, as numbered 1-11 in the bottom-right corner of each page. We wish to avoid copy problems. We will supply a new exam for any copy problems.

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

<u>Please turn off all cell phones now.</u>

Please clear your desk entirely, except for pen, pencil, eraser, a blank piece of paper (for scratch pad use), and an optional water bottle. Please write your name and ID# on the blank piece of paper and <u>turn it in with your exam</u>. Please show your UCI ID when you turn it in.

<u>This page summarizes the points for each question, so you can plan your time.</u>

1. (9 pts total) Bayesian networks.

2. (12 pts total, 3 pts each) STATE-SPACE SEARCH STRATEGIES

3. (10 pts total, 2 pts each) WUMPUS WORLD MODELS.

4. (12 pts total, 4 pts each) Fun in the kinship domain: English and FOL.

5. (10 pts total, -5 for each error, but not negative) RESOLUTION THEOREM PROVING.

6. (10 pts total, 1 pt each) Probability Rules and Independence.

7. (5 pts total) Hierarchical Agglomerative Clustering.

8. (5 pts total) k-Means Clustering.

9. Logic Concepts (6 pts total, 1 pt each).

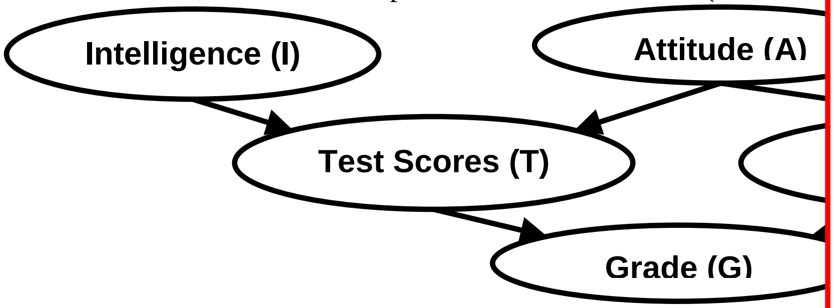
10. Probability concepts and formulae (9 pts total, 1 pt each).

11. (8 pts total, -1 pt each wrong answer, but not negative) SEARCH PROPERTIES.

12. (4 pts total, 1 pt each) TASK ENVIRONMENT.

<u>The Exam is printed on both sides to save trees! Work both sides of each page!</u>

1. (9 pts total) Bayesian networks. Consider the Bayesian network shown below. Recall that each node represents a random variable, and each arc represents a direct influence (either cau



Recall that there is one factor in the expression for each node in the graph. The conditioned variable is given by the node, and the conditioning variables are given by the parents of the node. So:

P(G|T,P) since T and P are parents of G

P(T|I,A) since I and A are parents of T

P(P|A) since A is the only parent of P

P(I)P(A) since I and A have no parents

1.a. (3 pts) Write the factored conditional probability expression that
Use the variable ordering indicated below, (G, T, P, I, A).

P(G, T, P, I, A) = P(G | T, P) P(T | I, A) P(P | A) P(I) P(A)

1.b. (4 pts, 1 pt each) Label ea

(A) ____T____ P(A | I) = P(A

(B) ____F____ Attitude (A) ar

~~**(C)**~~ ____F____ Intelligence (I

(D) ____F____ P(I ∧ A ∧ T) =

(A) See Fig. 14.4(a), p. 518: "A node X is conditionally independent of its non-descendants given its parents." A and I are non-descendants of each other and have no parents (i.e., nothing is give), so they are independent.

(B) G depends indirectly on A through T and also through P.

(C) **Cancelled.** Everyone gets (C) right, regardless of your actual answer.

(D) Both I and A are parents of T in the graph, and so both I and A directly influence T. Thus, I, A, and T are not independent.

1.c. (2 pts, 1 pts each) For each
formula that calculates it correctly. Here, Domain(Attitude) = {good, bad} and Domain(Participation) = {pass, fail}. The first one is done for you as an example.

(example) P(A = good) = Σ g, t, p, i P(G = g ∧ T = t ∧ P = p ∧ I = i ∧ A = good)

(A) P(A = good ∧ P = pass) = ____D________

(A) Σ g, t, p, i P(G = g ∧ T = t ∧ I = i ∧ A = good | P = pass)
(B) Σ g, t, i P(G = g ∧ T = t ∧ I = i | P = pass ∧ A = good)
(C) Σ p,a P(G = g ∧ T = t ∧ P = pass ∧ I = i ∧ A = good)
(D) Σ g, t, i P(G = g ∧ T = t ∧ P = pass ∧ I = i ∧ A = good)

For problem 1.c, the sum is over all possible values of all variables that **do not** appear in the marginal probability. In that sum. variables that **do** appear in the marginal probability are set equal to the value given in in the marginal probability.

(B) P(P = fail) = ____C________

(A) Σ p P(G = g ∧ T = t ∧ P = fail ∧ I = i ∧ A = a)
(B) Σ g, t, i, a P(G = g ∧ T = t ∧ I = i ∧ A = a | P = fail)
(C) Σ g, t, i, a P(G = g ∧ T = t ∧ P = fail ∧ I = i ∧ A = a)
(D) Σ g, t, p, i, a P(P = fail | G = g ∧ T = t ∧ I = i ∧ A = a)

Problem 1.b.(C) should have read: "Intelligence (I) and Grade (G) are conditionally independent given Test Scores (T) **and Attitude (A)**."

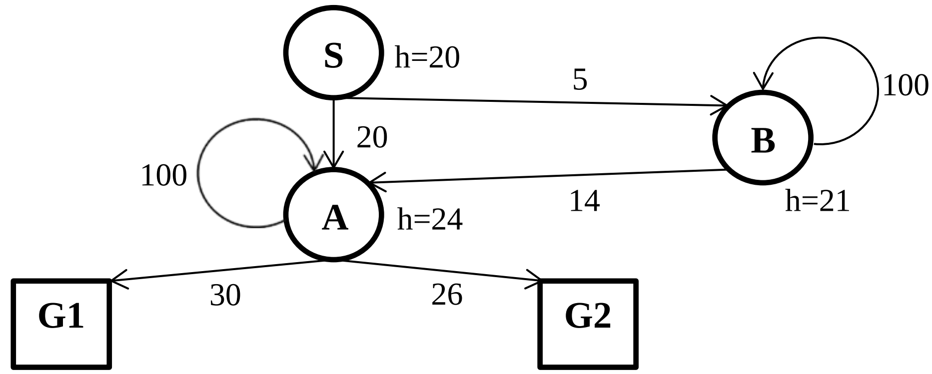
In that case, it would have been True. As stated, it is False. Please see the discussion about Markov blanket in Fig. 14.4(b), p. 518.

We decided that this distinction was too subtle and complicated to be a fair question on a timed closed-book Exam, so 1.b.(C) is cancelled.

2. (12 pts total, 3 pts each) STATE-SPACE SEARCH STRATEGIES. Execute Tree Search through this graph (i.e., do not remember visited nodes). Step costs are given next to each arc. Heuristic values are given next to each node (as h=x). The successors of each node are indicated by the arrows out of that node. Child nodes are returned in left-to-right order, i.e., children of S are (A, B), children of A are (A, G1, G2), and children of B are (A,B), in that order.

For each search strategy below, 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, or indicate the repeating cycle if the search gets stuck in a loop. Show the path from start to goal, or write "None". Give the cost of the path that is found, or write "None"

The first one is done for you as an example. (Note that here, G1 and G2 are different nodes, unlike Quiz 1.)



2.example. DEPTH FIRST SEARCH.

Order of node expansion: S A G1

See Section 3.4.3 and Fig. 3.17.

DFS does the Goal-test before the child is pushed onto the queue.

Path found: None   Cost of path found: None

2.a. (3 pts) UNIFORM COST SEA

UCS does goaltest when node is popped off queue.

See Section 3.4.2 and Fig. 3.14.

(1 pt) Order of node expansion: S B A A G2

(1 pt) Path found: S B A G2   **(1 pt)** Cost of path found: 45

2.b. (3 pts) GREEDY (BEST-FIRST) SEARCH

GBFS has the same behavior whether the goaltest is done before node is pushed or after node is popped, because h=0 for a goal node, so goal nodes always sort to the front of the queue anyway.

(1 pt) Order of node expansion: S B B B B B ...

(1 pt) Path found: None   **(1 pt)** Cost of path found: None

See Section 3.5.1 and Fig. 3.23.

2.c. (3 pts) ITERATED DEEPENING SEARCH.

IDS does the Goal-test before the child is pushed onto the queue.

(1 pt) Order of node expansion: S S A G1

(1 pt) Path found: S A G1   **(1 pt)** Cost of path found: 50

See Sections 3.4.4-5 and Figs. 3.17-19.

2.d. (3 pts) A* SEARCH.

A* does goaltest when node is popped off queue.

(1 pt) Order of node expansion: S B A A G2

(1 pt) Path found: S B A G2   **(1 pt)** Cost of path found: 45

See Section 3.5.2 and Figs. 3.24-25.

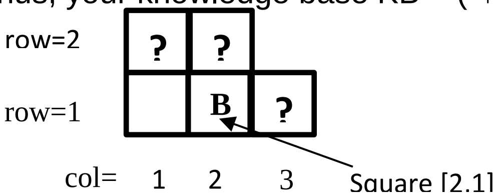
TECHNICAL NOTE: Technically, the goal node is not expanded, because no children of a goal node are generated. The goal node is listed in "Order of node expansion" for your convenience. Your answer is correct if you do not show the goal node in "Order of node expansion" — but it is a nicety to do so. Nevertheless, "Path found" *always* must show the goal node, because a path to a goal always must end in a goal.

3. (10 pts total, 2 pts each) WUMPUS WORLD MODELS.

Recall that a knowledge base KB entails a sentence S (written KB |= S) just in case the set of models that make the knowledge base true is a subset of the models that make S true (a model is a possible world). If this condition holds, it is impossible for KB to be true and S to be false. In such a case, S must be true in all worlds in which KB is true.

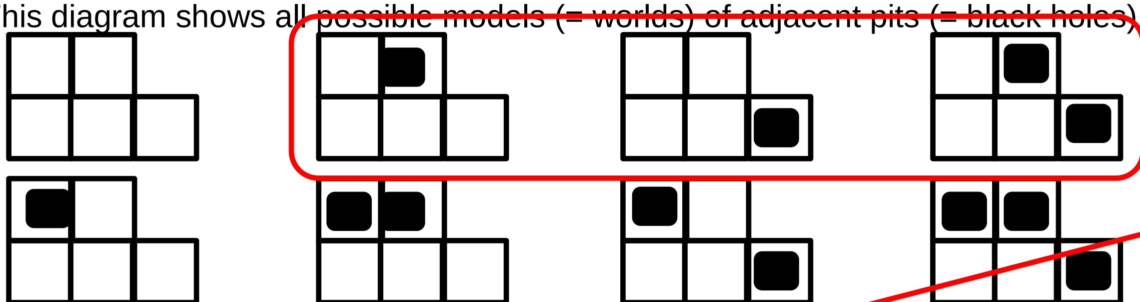
This question will concern only breezes and pits. <u>*Squares next to pits are breezy, and breezy squares are next to squares with pits.*</u> Pits are represented as black boxes in the world drawings below, and breezes are represented by the letter B. We ignore the wumpus, gold, etc.

Your agent did not detect a breeze at square [1,1] **(column, row).** Square [2,1] has a breeze. Thus, your knowledge base KB = ($\neg$ B_1,1) $\wedge$ (B_2,1), where B = Breeze.



This problem is the same as #1 on WQ'2013 Mid-term and #4 on WQ'2015 Final. See Section 7.3 and Fig. 7.5, p. 241, of R&N; and lecture slides for Propositional Logic A and Review for Mid-term Exam.

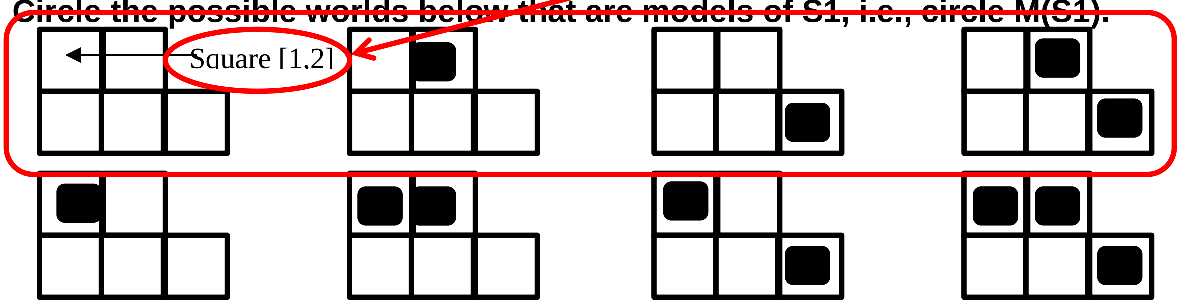
This diagram shows all possible models (= worlds) of adjacent pits (= black holes):



3.a. Circle the possible worlds above that are models of KB, i.e., circle M(KB).

3.b. Consider ONLY the sentence S1 = "Square [1,2] does not have a pit."

Circle the possible worlds below that are models of S1, i.e., circle M(S1).

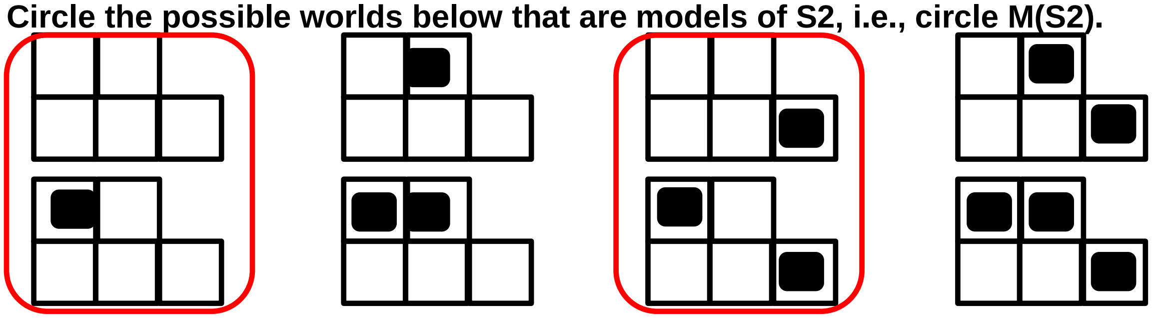


3.c. Does KB |= S1? (Y = yes, N = no) ____Y____

In 3.b, a student said that he was confused and thought the arrow from "Square[1,2]" meant that it was an example, so he ignored it. We decided to temper justice with mercy. Your answer to 3.b will be considered correct if you ignored the world in 3.b that has the arrow from "Square[1,2]."

3.d. Consider ONLY the sentence S2 = "Square [2,2] does not have a pit."

Circle the possible worlds below that are models of S2, i.e., circle M(S2).



3.e. Does KB |= S2? (Y = yes, N = no) ____N____

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

4. (12 pts total, 4 pts each) Fun in the kinship domain: English and FOL. For each English sentence, write the letter of the best or closest FOL sentence (wff, or well-formed formula).

ParentOf(x, y) means x is a parent of y. **MarriedTo(x, y)** means x is married to y. **SiblingOf(x, y)** means x is a sibling of y. **Female(x)** means x is female. Once a predicate has been defined in a problem, it may be used in subsequent problems. *All objects are persons, i.e., no need for Person(x) guard predicates.*

English definitions are "Your ... is/has ... of/with you." FOL definitions are $\forall x \forall y\ (P(x,y) \Leftrightarrow ...)$; y = you. To help you, the intended variable bindings are identified. **The first one is done for you, as an example.**

For the benefit of non-native English speakers, the diagrams below each problem illustrate the intended relationships described by the English statements above. The arc tail is the predicate first argument, the arc head is the second argument, and additional predicates are given as text.

4.example __A__ "Your sister(x) is a female who is a sibling of you (y)."

A. $\forall x \forall y\ SisterOf(x, y) \Leftrightarrow (Female(x) \land SiblingOf(x, y))$

B. $\forall x \forall y\ SisterOf(x, y) \Leftrightarrow (Female(x) \Rightarrow SiblingOf(x, y))$

(and Female(x))

See Section 8.3.2.

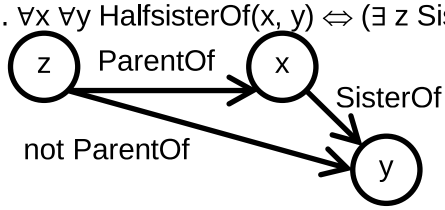
4.a (4 pts) __D____ "Your Half-sister (x) is a sister who has a parent (z) that is not a parent of you (y)."

A. $\forall x \forall y\ HalfsisterOf(x, y) \Leftrightarrow (\exists z\ [SisterOf(x, y) \Rightarrow ParentOf(z, x)] \land \neg ParentOf(z, y))$

B. $\forall x \forall y\ HalfsisterOf(x, y) \Leftrightarrow (\forall z\ SisterOf(x, y) \land [ParentOf(z, x) \Rightarrow \neg ParentOf(z, y)])$

C. $\forall x \forall y\ HalfsisterOf(x, y) \Leftrightarrow (\forall z\ SisterOf(x, y) \land ParentOf(z, x) \land \neg ParentOf(z, y))$

D. $\forall x \forall y\ HalfsisterOf(x, y) \Leftrightarrow (\exists z\ SisterOf(x, y) \land ParentOf(z, x) \land \neg ParentOf(z, y))$



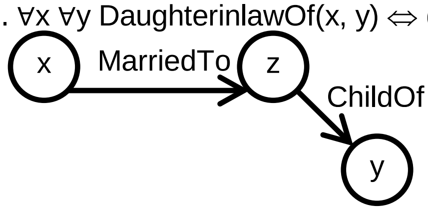
4.b (4 pts) __C____"Your daughter-in-law (x) is a female who is married to a child (z) of you (y)."

A. $\forall x \forall y\ DaughterinlawOf(x, y) \Leftrightarrow (\exists z\ [ChildOf(z, y) \land MarriedTo(x, z)] \Rightarrow Female(x))$

B. $\forall x \forall y\ DaughterinlawOf(x, y) \Leftrightarrow (\forall z\ ChildOf(z, y) \land MarriedTo(x, z) \land Female(x))$

C. $\forall x \forall y\ DaughterinlawOf(x, y) \Leftrightarrow (\exists z\ ChildOf(z, y) \land MarriedTo(x, z) \land Female(x))$

D. $\forall x \forall y\ DaughterinlawOf(x, y) \Leftrightarrow (\forall z\ ChildOf(z, y) \lor [MarriedTo(x, z) \land Female(x)])$



(and Female(x))

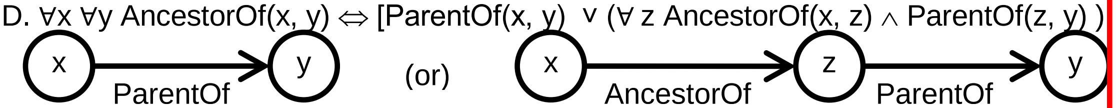
4.c (4 pts) __A____ "Your ancestor (x) is a parent of you (y) or an ancestor of a parent (z) of you (y)."

A. $\forall x \forall y\ AncestorOf(x, y) \Leftrightarrow [ParentOf(x, y) \lor (\exists z\ AncestorOf(x, z) \land ParentOf(z, y))]$

B. $\forall x \forall y\ AncestorOf(x, y) \Leftrightarrow [ParentOf(x, y) \lor (\forall z\ AncestorOf(x, z) \land ParentOf(z, y))]$

C. $\forall x \forall y\ AncestorOf(x, y) \Leftrightarrow [ParentOf(x, y) \lor (\exists z\ AncestorOf(x, z) \lor ParentOf(z, y))]$

D. $\forall x \forall y\ AncestorOf(x, y) \Leftrightarrow [ParentOf(x, y) \lor (\forall z\ AncestorOf(x, z) \land ParentOf(z, y))]$



4.c IS CANCELLED.

Everyone gets 4.c right, regardless of your answer.

The reason was a copy-and-edit bug on the original Exam, which has been fixed in this revised version.

So, for future study, you may use this revised version as a reliable guide.

5. (10 pts total, -5 for each error, but not negative) RESOLUTION THEOREM PROVING. You are engaged in Knowledge Engineering for the Wumpus Cave. You have interviewed an expert on the Wumpus Cave who told you, among other things, "A breeze in square (1,1) is equivalent to a pit in square (1,2) or a pit in square (2,1)." You translated this into propositional logic as, "(B11 $\Leftrightarrow$ P12 $\vee$ P21)," and then into Conjunctive Normal Form as "($\neg$B11 $\vee$ P12 $\vee$ P21) $\wedge$ ($\neg$P12 $\vee$ B11) $\wedge$ ($\neg$P21 $\vee$ B11)."

Now it is time for the first "live" test of your system. An agent has been lowered down into the Wumpus cave, and reports back by radio, "Square (1,1) has a breeze. Also, I went into square (1,2) and I did not die, so it does not have a pit." You translate this knowledge into propositional logic as "(B11) $\wedge$ ($\neg$ P12)" and add it to your knowledge base.

Next your system is asked to perform inference. The agent asks by radio, "Is it true that square (2,1) has a pit?" You translate this query into propositional logic as the goal sentence "(P21)." You form the negated goal as "($\neg$ P21)." Your knowledge base plus negated goal is:

($\neg$B11 $\vee$ P12 $\vee$ P21)
($\neg$P12 $\vee$ B11)
($\neg$P21 $\vee$ B11)
(B11)
($\neg$ P12)
($\neg$ P21)

See Sections 7.2 and 7.5.2.

Run resolution on this knowledge base until you produce the null clause, "( )", thereby proving that the goal sentence is true. The shortest proof I know of is only three lines long. (Bonus points for a shorter proof.) It is OK to use more than lines, if your proof is correct.

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.

Resolve ___($\neg$B11 $\vee$ P12 $\vee$ P21)___ and ___(B11)___ to give ___(P12 $\vee$ P21)___

Resolve ___(P12 $\vee$ P21)___ and ___($\neg$P12)___ to give ___(P21)___

Resolve ___(P21)___ and ___($\neg$P21)___ to give ___( )___

Resolve ______ and ______ to give ______

Resolve ______

Other proofs are OK as long as they are correct. For example, you might perform the resolution steps above in any other order you choose.

Resolve ______ and ______ to give ______

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

6. (10 pts total, 1 pt each) Probability Rules and Independence.
Consider the following full joint distribution for Boolean variables A, B, C:

See Chapter 13.

A	*B*	*C*	*P(a,b,c)*
t	*t*	*t*	*0.03*
t	*t*	*f*	*0.12*
t	*f*	*t*	*0.17*
t	*f*	*f*	*0.18*
f	*t*	*t*	*0.03*
f	*t*	*f*	*0.12*
f	*f*	*t*	*0.24*
f	*f*	*f*	*0.11*

Write an arithmetical expression that will evaluate to the following probabilities. You do not need to add up the numbers to produce a single numerical result. It is sufficient for you to write an arithmetical expression that will evaluate to the correct numerical result. The first one is done for you, as an example.

6.example $P(A = f, B = f, C = f \mid A = f, B = f)$ = 0.11 / (0.24 + 0.11) [≈ 0.3143]

6.a. (1 pt) $P(A = f)$ = 0.03+0.12+0.24+0.11 [= 0.50]

6.b. (1 pt) $P(B = t)$ = 0.03+0.12+0.03+0.12 [= 0.30]

6.c. (1 pt) $P(B = t, C = t)$ = 0.03+0.03 [= 0.06]

6.d. (1 pt) $P(A = f, C = t)$ = 0.03+0.24 [= 0.27]

6.e. (1 pt) $P(A = t \mid B = t)$ = (0.03+0.12) / (0.03+0.12+0.03+0.12) [= 0.5]

6.f. (1 pt) $P(C = f \mid B = t)$ = (0.12+0.12) / (0.03+0.12+0.03+0.12) [= 0.80]

6.g. (1 pt) Are A and B independent of each other? (Y=Yes, N=No): Y

6.h. (1 pt) Are B and C independent of each other? (Y=Yes, N=No): N

~~**6.i**~~ **(1 pt)** Are B and C conditionally independent given A? (Y=Yes, N=No): N

~~**6.j.**~~ **(1 pt)** Are A and C conditionally independent given B? (Y=Yes, N=No): N

Problem 6.i and 6.j are cancelled.
Everyone gets 6.i and 6.j right, regardless of your actual answer.

Problem 6.i and 6.j may require calculating up to eight numeric conditional probabilities before the correct answer is apparent. Yet, we prohibit electronic calculators, so all calculations must be done by hand.

We decided that this laborious manual process was too tedious, time-consuming, and complicated to be a fair question on a timed closed-book no-calculator Exam, so 6.i and 6.j are cancelled.

6.a. P(A=f) = P(A=f,B=t,C=t)+P(f,t,f)+P(f,f,t)+P(f,f,f)
= 0.03+0.12+0.24+0.11 = 0.50

6.b. P(B=t) = P(A=t,B=t,C=t)+P(t,t,f)+P(f,t,t)+P(f,t,f)
= 0.03+0.12+0.03+0.12 = 0.30

6.c. P(B=t, C=t) = P(A=t,B=t,C=t)+P(f,t,t)
= 0.03+0.03 = 0.06

6.d. P(A=f, C=t) = P(A=f,B=t,C=t)+P(f,f,t)
= 0.03+0.24 = 0.27

6.e. P(A=t | B=t) = P(A=t,B=t)/P(B=t)
P(A=t,B=t) = P(A=t,B=t,C=t)+P(t,t,f) = 0.03+0.12=0.15
P(B=t) = 0.30 (above, 6.b)
P(A=t,B=t)/P(B=t) = 0.15/0.30 = 0.50

6.f. P(C=f | B=t) = P(B=t,C=f)/P(B=t)
P(B=t,C=f) = P(A=t,B=t,C=f)+P(f,t,f) = 0.12+0.12=0.24
P(B=t) = 0.30 (above, 6.b)
P(B=t,C=f)/P(B=t) = 0.24/0.30 = 0.80

6.g. **A and B independent requires P(A, B) = P(A) * P(B).**

(1) P(A=t,B=t) = P(A=t,B=t,C=t) + P(t,t,f) = 0.03+0.12 = 0.15
P(A=t) = 1 - P(A=f) = 1 - 0.50 = 0.50 (above, 6.a)
P(B=t) = 0.30 (above, 6.b)
P(A=t,B=t) = 0.15 = 0.50*0.30 = P(A=t)*P(B=t)

(2) P(A=f,B=t) = P(A=f,B=t,C=t) + P(f,t,f) = 0.03+0.12 = 0.15
P(A=f) = 0.50 (above, 6.a)
P(B=t) = 0.30 (above, 6.b)
P(A=f,B=t) = 0.15 = 0.50*0.30 = P(A=f)*P(B=t)

(3) P(A=t,B=f) = P(A=t,B=f,C=t) + P(t,f,f) = 0.17+0.18 = 0.35
P(A=t) = 1 - P(A=f) = 1 - 0.50 = 0.50 (above, 6.a)
P(B=f) = 1 - P(B=t) = 0.70 (above, 6.b)
P(A=f,B=t) = 0.35 = 0.50*0.70 = P(A=f)*P(B=t)

(4) P(A=f,B=f) = P(A=f,B=f,C=t) + P(f,f,f) = 0.24+0.11 = 0.35
P(A=f) = 0.50 (above, 6.a)
P(B=f) = 1 - P(B=t) = 0.70 (above, 6.b)
P(A=f,B=f) = 0.35 = 0.50*0.70 = P(A=f)*P(B=f)

Yes, A and B are independent. P(A, B) = P(A) * P(B) for all values of A=a and B=b.

6.h. **B and C independent requires P(B, C) = P(B) * P(C).**

P(B=t,C=t) = 0.06 (above, 6.c)
P(B=t) = 0.30 (above, 6.b)
P(C=t) = P(A=t,B=t,C=t)+P(t,f,t)+P(f,t,t)+P(f,f,t)
= 0.03+0.17+0.03+0.24 = 0.47
P(B=t,C=t) = 0.06 =/= 0.141 = 0.30*0.47 = P(B=t)*P(C=t)

No, B and C are not independent. P(B=t, C=t) =/= P(B=t) * P(C=t)

Problem 6.i and 6.j may require calculating up to eight numeric conditional probabilities before the correct answer is apparent. Yet, we prohibit electronic calculators, so all calculations must be done by hand.

We decided that this laborious manual process was too tedious, time-consuming, and complicated to be a fair question on a timed closed-book no-calculator Exam, so 6.i and 6.j are cancelled.

However, we have provided the correct answers and the correct derivations on this annotated key, and so problems 6.i and 6.j may be used as reliable study guides for students to go more deeply into the material.

Cancelled. Everyone gets this problem right, regardless of your actual answer.

6.i. **B and C conditionally independent given A requires P(B, C | A) = P(B | A) * P(C | A)**

P(B=t,C=t | A=t) = P(A=t, B=t, C=t) / P(A=t) (definition)
P(A=t, B=t, C=t) = 0.03
P(A=t) = 1 - P(A=f) = 1 - 0.50 = 0.50 (above, 6.a)
P(B=t, C=t | A=t) = P(A=t, B=t, C=t) / P(A=t) = 0.03 / 0.50 = 0.06

P(B=t | A=t) = P(A=t, B=t) / P(A=t) (definition)
P(A=t, B=t) = P(A=t, B=t, C=t)+P(t,t,f) = 0.03+0.12 = 0.15
P(B=t |A=t) = P(A=t, B=t) / P(A=t) = 0.15 / 0.50 = 0.30

P(C=t |A=t) = P(A=t, C=t) / P(A=t) (definition)
P(A=t, C=t) = P(A=t, B=t, C=t)+P(t,f,t) = 0.03+0.17 = 0.20
P(C=t |A=t) = P(A=t,C=t) / P(A=t) = 0.20 / 0.50 = 0.40

P(B=t, C=t | A=t) = 0.06 =/= 0.12 =0.30 * 0.40 = P(B=t | A=t)*P(C=t | A=t)

No, B and C are not conditionally independent given A.
P(B=t, C=t | A=t) =/= P(B=t | A=t) * P(C=t | A=t)

Cancelled. Everyone gets this problem right, regardless of your actual answer.

6.j. **A and C conditionally independent given B requires P(A, C | B) = P(A | B) * P(C | B)**

P(A=t, C=t | B=f) = P(A=t, B=f, C=t) / P(B=f) (definition)
P(A=t, B=f, C=t) = 0.17
P(B=f) = 1 - P(B=t) = 1 - 0.30 = 0.70 (above, 6.b)
P(A=t, C=t | B=f) = P(A=t, B=f, C=t) / P(B=f) = 0.17 / 0.70 ≈ 0.242857

P(A=t | B=f) = P(A=t, B=f) / P(B=f) (definition)
P(A=t, B=f) = P(A=t, B=f, C=t) + P(t,f,f) = 0.17+0.18 = 0.35
P(A=t | B=f) = P(A=t, B=f) / P(B=f) = 0.35 / 0.70= 0.50

P(C=t |B=f) = P(B=f, C=t) / P(B=f) (definition)
P(B=f, C=t) = P(A=t, B=f, C=t)+P(f,f,t) = 0.17+0.24= 0.41
P(C=t |B=f) = P(B=f, C=t) / P(B=f) = 0.41 / 0.70 ≈ 0.585714

P(A=t, C=t | B=f) ≈ 0.242857 ≈/≈ 0.292857 = 0.585714 * 0.50 ≈ P(A=t | B=f)*P(C=t | B=f)

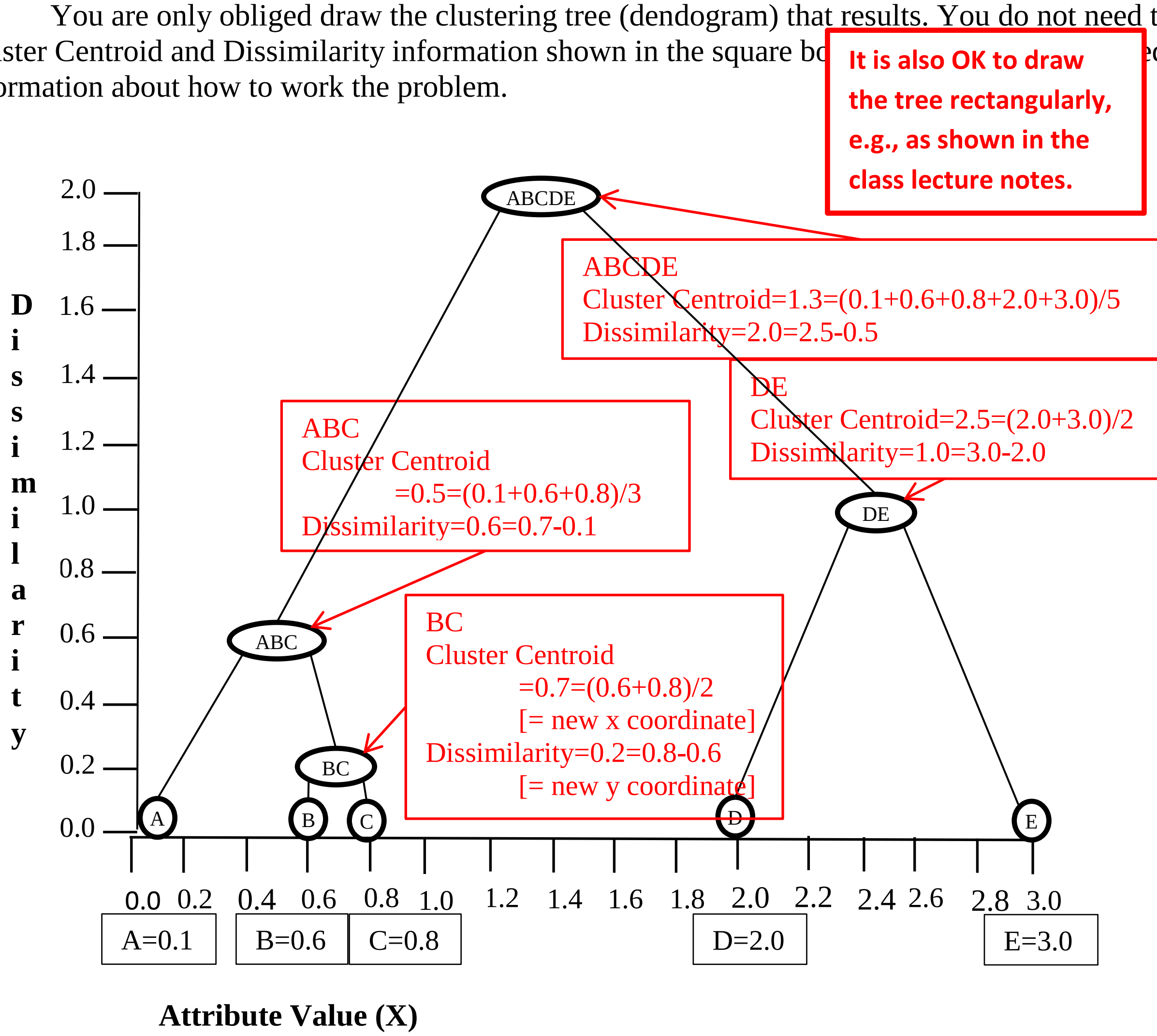
No, A and C are not conditionally independent given B.
P(A=t, C=t | B=f) =/= P(A=t | B=f) * P(C=t | B=f)

7. (5 pts total) Hierarchical Agglomerative Clustering. Consider this training data set (it is the same as in problem 1). Examples are A-E, and the single attribute is X.

Example	A	B	C	D	E
Attribute Value (X)	0.1	0.6	0.8	2.0	3.0

Draw the dendogram (clustering tree) that results from applying hierarchical agglomerative clustering to this data. When two clusters are merged, replace them with their cluster centroid, i.e., the statistical mean of all cluster members. This rule means, (1) each cluster is represented by its cluster centroid which is the numerical mean (average) of all of its cluster members; and (2) dissimilarity between clusters is computed as the distance between their cluster centroids using Euclidean distance. **(Note: A better measure of dissimilarity is the root-mean-squared-deviation [RMSD] of each cluster member from its cluster centroid; but that is infeasible in an exam like this.)** Label the cluster centroids by drawing an oval around the data points that are included in that cluster centroid. The first one is done for you as an example.

You are only obliged draw the clustering tree (dendogram) that results. You do not need to write in the Cluster Centroid and Dissimilarity information shown in the square bo[xes; that is provid]ed only for your information about how to work the problem.



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

8. (5 pts total) k-Means Clustering. Consider this training data set (it is the same data set as in problems 1 and 6). Examples are A-E, and the single attribute is X.

Example	
Attribute Value (X)	

Full credit for each sub-problem if your answers for it would be correct given your answers for previous sub-problems, even if prevuious sub-problems were wrong.

Apply k-Means Clustering to this data set for k=2,

8.a. (1 pt) You have randomly chosen two data points with which to initialize your two clusters. Randomly, you chose example A to initialize cluster #1 and example B to initialize cluster #2.

Write down the cluster assignments that result. Write C, D, and E in the blanks below according to which cluster they are assigned (A and B are already assigned).

cluster #1: ___A_______________ cluster #2: ___B, C, D, E___

C, D, and E are all closer to B than to A.

8.b. (1 pt) After assigning examples to clusters in 8.a, you recompute the cluster centroids (means) to be the mean (average) of the examples currently assigned to each cluster.

For each cluster, ... he nev ... of the examples that were assigned to that clu...

cluster #1 contains only A, so its centroid is A's X.

cluster #2 centroid is the mean X of B, C, D, & E.

cluster #1: ___0.1_______________ cluster #2: ___1.6 = (0.6+0.8+2.0+3.0)/4___

8.c. (1 pt) After recomputing the cluster centroids (means) in 8.b, you reassign the examples to the clusters to which they are closest (i.e., the example is assigned to the closest cluster centroid).

Write down the cluste... ılt. Write A, B, C, ... below according to which cluster th...

A, B, and C are closer to 0.1 than to 1.6.

D and E are closer to 1.6 than to 0.1.

cluster #1: ___A, B, C____________ cluster #2: ___D, E____________________

8.d. (1 pt) After assigning examples to clusters in 8.c, you recompute the cluster centroids (means) to be the mean (average) of the examples currently assigned to each cluster.

For each cluste... the new cluste... xamples that were assigned to that c...

cluster #1 new centroid is the mean X of A, B, & C.

cluster #2 new centroid is the mean X of D & E.

cluster #1: ___0.5 = (0.1+0.6+0.8)/3_ cluster #2: ___2.5 = (2.0+3.0)/2____________

8.e. (1 pt) After recomputing the cluster centroids (means) in ... he clusters to which they are closest (i.e., the example is assigned to the clo...

Write down the cluster assignments that result. W... ks below according to which cluster they are assigned.

cluster membership doesn't change, so the clustering process is quiescent and terminates.

cluster #1: ___A, B, C____________ cluster #2: ___D, E____________________

9. Logic Concepts (6 pts total, 1 pt each). See Chapter 7.

A .	Logic	A	Formal symbol system for representation and inference
C	Valid	B	The idea that a sentence follows logically from other sentences
G	Complete	C	True in every possible world
E	Conjunctive Normal Form	D	True in at least one possible world
F	Sound	E	A sentence expressed as a conjunction of clauses (disjuncts)
D	Satisfiable	F	Inference system derives only entailed sentences
B	Entailment	G	Inference system can derive any sentence that is entailed

See Chapter 13.

10. Probability concepts and formulae (9 pts total, 1 pt each).

A .	Probability Theory	A	Assigns each sentence a degree of belief ranging from 0 to 1
H	Conditional independence	B	Degree of belief accorded without any other information
G	Independence	C	Degree of belief accorded after some evidence is obtained
J	Product rule (chain rule)	D	Gives probability of all combinations of values of all variables
C	Conditional probability	E	Takes values from its domain with specified probabilities
B	Unconditional probability	F	A possible world is represented by variable/value pairs
F	Factored representation	G	$P(a \wedge b) = P(a)\, P(b)$
E	Random variable	H	$P(a \wedge b \mid c) = P(a \mid c)\, P(b \mid c)$
I	Bayes' rule	I	$P(a \mid b) = P(b \mid a)\, P(a) / P(b)$
D	Joint probability distribution	J	$P(a \wedge b \wedge c) = P(a \mid b \wedge c)\, P(b \mid c)\, P(c)$

11. (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 you See Figure 3.21.

Criterion	Complete?	Time complexity	Space complexity	Optimal?
Breadth-First	Yes	O(b^d)	O(b^d)	Yes
Uniform-Cost	Yes	O(b^(1+floor(C*/ε))) O(b^(d+1)) also OK	O(b^(1+floor(C*/ε))) O(b^(d+1)) also OK	Yes
Depth-First	No	O(b^m)	O(bm)	No
Iterative Deepening	Yes	O(b^d)	O(bd)	Yes
Bidirectional (if applicable)	Yes	O(b^(d/2))	O(b^(d/2))	Yes

12. (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<u>erformance (measure)</u>   E<u>nvironment</u>   A<u>ctuators</u>   S<u>ensors</u>

****** THIS IS THE END OF THE MID-TERM EXAM ******