6.034 Final Examination

19 December 2013

Name:		**e-mail:**	

Indicate which sections of the final you will be taking. We will grade only those sections.

☐ **Quiz 1**		☐ **Quiz 2**		☐ **Quiz 3**		☐ **Quiz 4**	
Problem 1	Problem 2	Problem 1	Problem 2	Problem 1	Problem 2	Problem 1	Problem 2
Quiz 1 Total		Quiz 2 Total		Quiz 3 Total		Quiz 4 Total	
☐ **Bonus SRN**		☐ **SRN 2**		☐ **SRN 3**		☐ **SRN 4**	

Survey

Question 1. Indicate the approximate percent of the lectures, right-now talks, and tutorials that you have attended so that we can better gauge the correlation of quiz and final performance with attendance. Your answers have no effect on your grade.

	Lectures	Right-now talks	Mega Recitations	Tutorials
Percent attended				

Question 2. Characterize your note taking habits. Your answers have no effect on your grade.

Regular lectures:	None taken	Rare	Occasional	Moderate	Extensive
Right-now talks:	None taken	Rare	Occasional	Moderate	Extensive

There are 33 pages on this exam, not including blank pages and tear-off sheets. As always, this exam is open book, open notes, open almost everything—including a calculator—but no computers.

Quiz 1, Problem 1 : Rule-based systems (50 points)

For your birthday, you received the video game you wanted—but alas, the terminology and rules of the game are immensely complex. You decide to write down what you've learned as a database of if-then rules so you can figure out some of this stuff automatically.

Rules:

R0	IF ‘(?x) has wings' THEN ‘(?x) can fly'
R1	IF ‘(?x) is a hero of (?y)' THEN ‘(?x) masters (?y)'
R2	IF ‘(?x) uses (?y)' THEN ‘(?x) is a hero of (?y)'
R3	IF (?x) casts spells THEN ‘(?x) grows silver wings' ‘(?x) uses Magic'
R4	IF ‘(?x) masters wind' THEN ‘(?x) can fly'
R5	IF (AND `(?x) masters Magic' `(?x) is a hero of (?y)' `(?x) can fly') THEN ‘(?x) ascends to godhood'

According to the game's handbook, the following two facts are true:

Assertions:

A0: John is a hero of wind A1: Magic casts spells

★For your convenience, a copy of these rules and assertions is provided on a tear-off sheet after the end of the exam.

Part A: Backward chaining (25 points)

Make the following assumptions about backward chaining:

- The backward chainer tries to find a matching assertion in the list of assertions. If no matching assertion is found, the backward chainer tries to find a rule with a matching consequent. In case none are found, then the backward chainer assumes the hypothesis is false.
- The backward chainer never alters the list of assertions, so it can derive the same result multiple times.
- Rules are tried in the order they appear.
- Antecedents are tried in the order they appear.
- Lazy evaluation/short circuiting is in effect

It's time to test your rule-based system. First, you want to find out whether your favorite character, Magic, will ascend to godhood. To find out, simulate backward chaining with the hypothesis

'Magic ascends to godhood'.

Write all the hypotheses the backward chainer looks for in the database in the order that the hypotheses are looked for. The table has more lines than you need.

Note: We also recommend you use the next page to draw the goal tree that would be created by backward chaining from this hypothesis. The goal tree will help us to assign partial credit in the event you have mistakes in the list.

1 *Magic ascends to godhood*	11
2	12
3	13
4	14
5	15
6	16
7	17
8	18
9	19
10	20

Space provided to draw the goal tree.

Magic ascends to godhood

Part B: Forward chaining (25 points)

Make the following assumptions about forward chaining:

- Assume rule-ordering conflict resolution
- New assertions are added to the bottom of the list of assertions.
- If a particular rule matches assertions in the list of assertions in more than one way, the matches are considered in the order corresponding to the top-to-bottom order of the matched assertions. Thus, if a particular rule has an antecedent that matches both A1 and A2, the match with A1 is considered first.

Next, you want to learn everything you can about your new game, based on the facts from the handbook and the rules you've developed. To do so, you'll need to **run forward chaining using the rules and assertions on page 2.** (An additional copy of the rules and assertions is also provided on a tear-off sheet after the last page of the exam.)

For the first two iterations, fill out the first two rows in the table below, noting the rules whose antecedents match the data, the rule that fires, and the new assertions that are added by the rule. For the remainder, supply only the fired rules and new assertions. There are more rows in the table than you'll need.

	Matched	Fired	New Assertions Added to the List of Assertions
1			
2			
3			
4			
5			
6			
7			
8			
9			

Quiz 1, Problem 2 : Search & Games (50 points)

Part A: Rearrangement Ramifications (5 points):

Game-playing prodigy Maxine has just discovered an enormous game tree, and has decided to write a program to figure out the best move to make. First, she runs the MINIMAX algorithm (without any pruning); it produces a path that ends up at a leaf node *X*. To check her answer, she randomly reorders the children of each node and runs the MINIMAX WITH ALPHA-BETA PRUNING algorithm. It produces a path that ends up at a different leaf node *Y.*

"Drat!" she says. "My code must be buggy."

In fact, her code is working perfectly—explain what special case must have occurred with her tree. (For some credit, you may instead explain why she *thought* her code must be buggy.)

Part B: Overkill (10 points):

Suppose you have a *complete* game tree—this means that the root node of the tree is MAX's first turn, and all the leaves of the tree represent states where the game has finished (win/lose or draw). You have a reliable static evaluator, so a leaf has a positive value if MAX wins (larger values mean better victories for MAX), a negative value if MIN wins (smaller values mean better victories for MIN), or a value of zero if the game ends in a draw.

When you use minimax with alpha-beta pruning on this tree, you get a path whose last node *Z* has a static value of 1, meaning that MAX wins the game. Now suppose you increase the static value of *Z* to 100 (so the victory is even better for MAX) while leaving all other static values the same. What effect will this change have?

The minimax game will certainly still end at state Z. YES NO

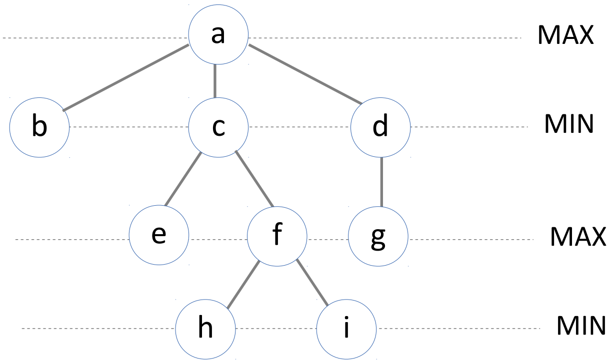
Explain.

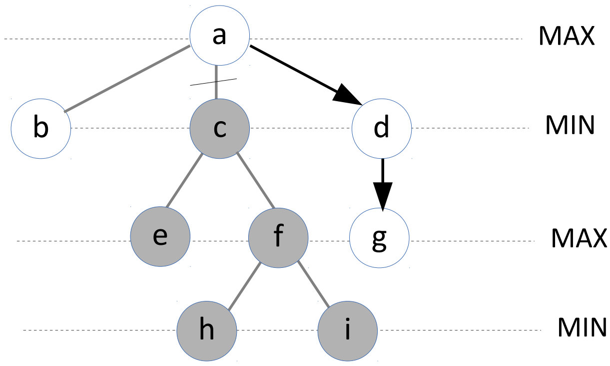
MAX will certainly still win the game. YES NO

Explain.

Part C: Non-secateurs (15 points):

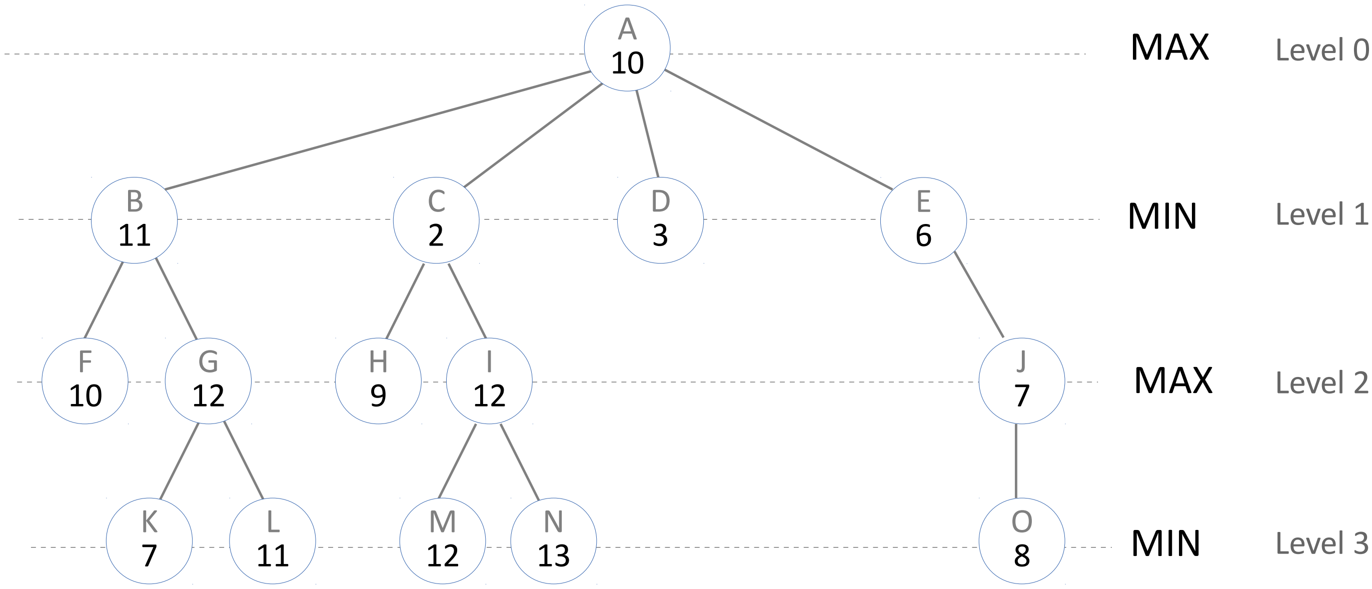
You are playing with the game tree shown on the left. Each node is labelled with its static value (for example, the static value of the root node is *a*). The diagram on the right shows what happens when you run MINIMAX WITH ALPHA-BETA PRUNING on that game tree: the path leading to *c* is pruned,neither *h* nor *i* are statically evaluated, and the minimax path is *a→d→g.*





Based on the results of alpha-beta pruning shown on the right, list all inequalities that must hold among the static values *a, b, c, d, e, f, g, h, i*. (For example, you might write expressions of the form $x < y$ *or* $y \leq z$)

Part D: Zeroing in (20 points):



Your final task is to find the best move for MAX to make for the game tree shown above. Use MINIMAX WITH ALPHA-BETA PRUNING , and perform static evaluation at level 3 (or at higher levels in parts of the tree that do not go all the way to level 3).

D1. List the nodes you statically evaluated, **<u>in the order you evaluated them.</u>**

D2. Which move does MAX choose?

D3. Suppose that after you finish this minimax problem, you overhear two TAs talking about it:

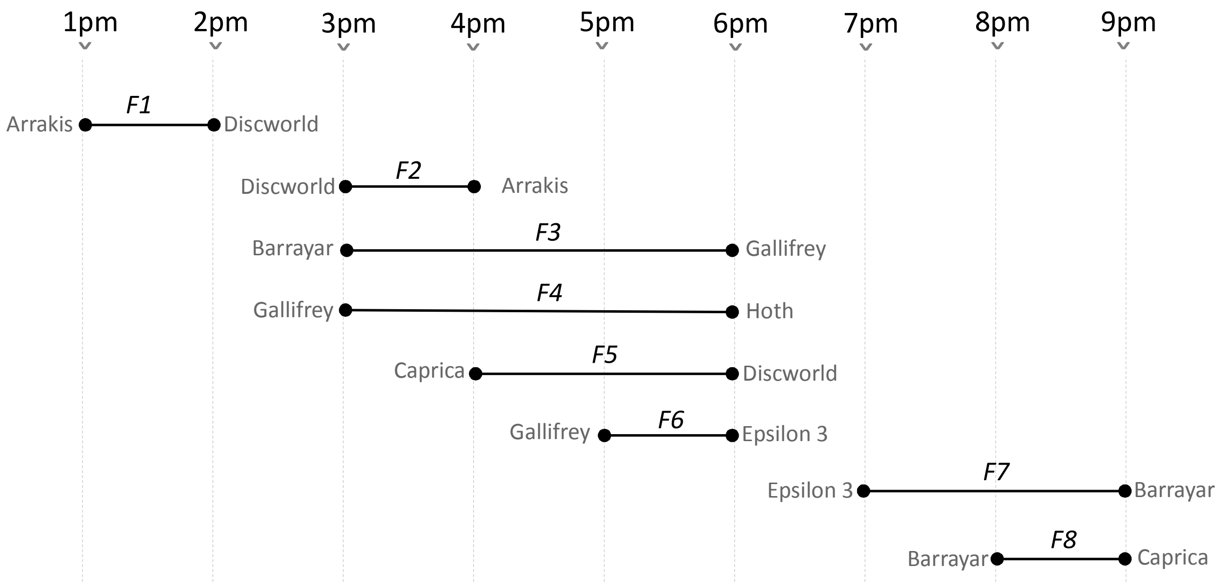
> TA 1: "Oops. I made a typo when drawing the game tree for part D—I meant for *that node* to have a static value of zero. At least all of the other static values are correct."
>
> TA 2: "No worries—look, you get all the same answers whether that node has its current static value or a static value of zero."

List **<u>all</u>** the nodes in the tree they might have been referring to.

Quiz 2, Problem 1: Constraints in interdimensional transport (50 points)

You've just taken over management of the floundering Interdimensional Transport Agency, whose spaceships ferry customers between dimensions via wormholes.

Currently, you offer eight flights (F1-F8) between various exotic locales, as shown below.



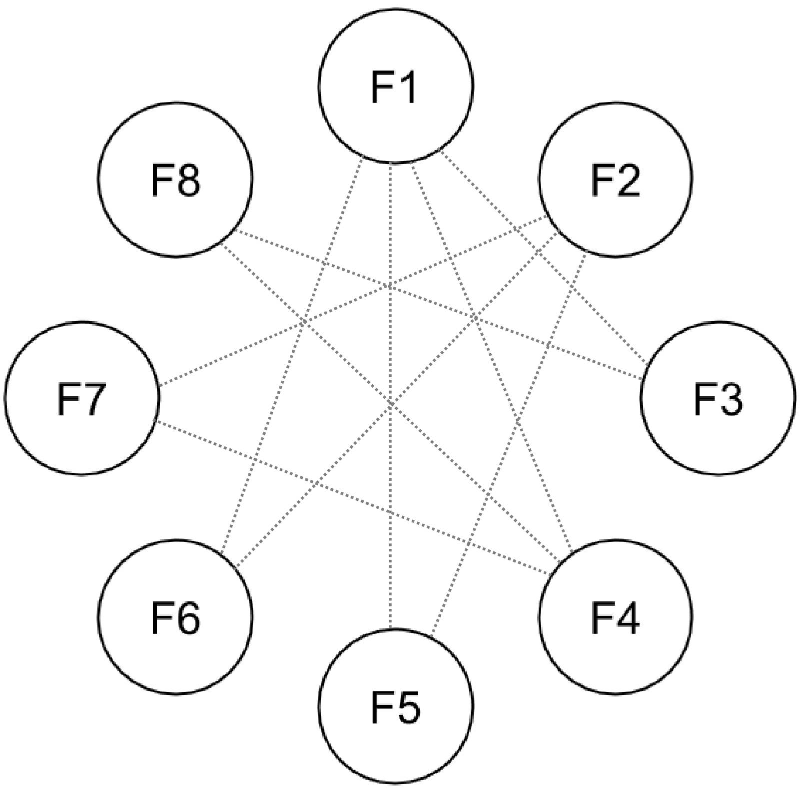
You refuse to change this schedule lest your customers start a riot. Instead, you want to hire the minimum number of spaceships necessary in order to satisfy the demands of this schedule.

Part A: Time and relative positions in space (5 points)

You decide to use the graph on the right to keep track of the constraints between flights.

Alyssa P. Hacker, one of your interns, was already able to add some edges to the graph by employing her advanced knowledge of wormhole physics.

But Ben Bitdiddle, another one of your interns, points out that Alyssa has overlooked a second, obvious constraint: modern day spaceships cannot be in two places at once. Therefore, if two flights have overlapping times, a single spaceship cannot handle them both.



<u>Take Ben's advice:</u> Referring to the schedule above, complete the constraint graph by adding an edge between any two flights that have overlapping times.

Part B: The Final Five (35 pts)

You decide to see if perhaps *five* spaceships (the **A**ntelope, the **B**ear, the **C**at, the **D**og, and the **E**lephant) can handle all eight flights. Use **constraint propagation with forward checking and propagation through singleton domains** to search for an assignment of your five spaceships to the eight flights.

<!> Important: In order to avoid overusing your spaceships, you decide to *rotate the ships after every assignment*. This means that when expanding a node in the tree, you will:

Write the children of **A** in this order:	B, C, D, E, A
Write the children of **B** in this order:	C, D, E, A, B
Write the children of **C** in this order:	D, E, A, B, C
Write the children of **D** in this order:	E, A, B, C, D
Write the children of **E** in this order:	A, B, C, D, E

This establishes a tie-breaking order.

B1. (5pts) Alyssa tells you to start from flight 4. Explain why this is a good idea.

B2. (25pts) Draw your tree below. **<u>Remember, you're starting with flight F4</u>**.

F4	A	B	C	D	E
F5					
F6					
F7					
F8					
F1					
F2					
F3					

B3. (5 pts) Give the assignment of spaceships you found in Part B2, or write NO SOLUTION if search terminated without finding a solution.

Part C: Argument is the Mind-Killer (10 pts)

Now you know whether five spaceships is enough, but you still haven't learned how many spaceships you need *at minimum* in order to handle all eight flights. Your eager interns decide to offer their advice.

Ben suggests the following strategy:

> Pick a number of spaceships that is guaranteed to be large enough to solve the problem. Run constraint propagation assuming you have that many ships. After you find a complete solution, report the number of ships your solution actually required. This will be the minimum number of ships necessary to solve the problem.

C1. (5pts) Alyssa quickly claims that there is a bug in this strategy: If you break ties by rotating the ships after every assignment (like you did in part B2), this strategy won't help you find the minimum number of spaceships. Instead, you should break ties another way, e.g. alphabetically.

Ben, surprised, points out that tiebreaking order is irrelevant since it can't affect whether there is a solution; it can only affect which solution you find, and perhaps how long it takes to find.

Who is correct, Alice or Ben? Explain.

C2.(5pts)

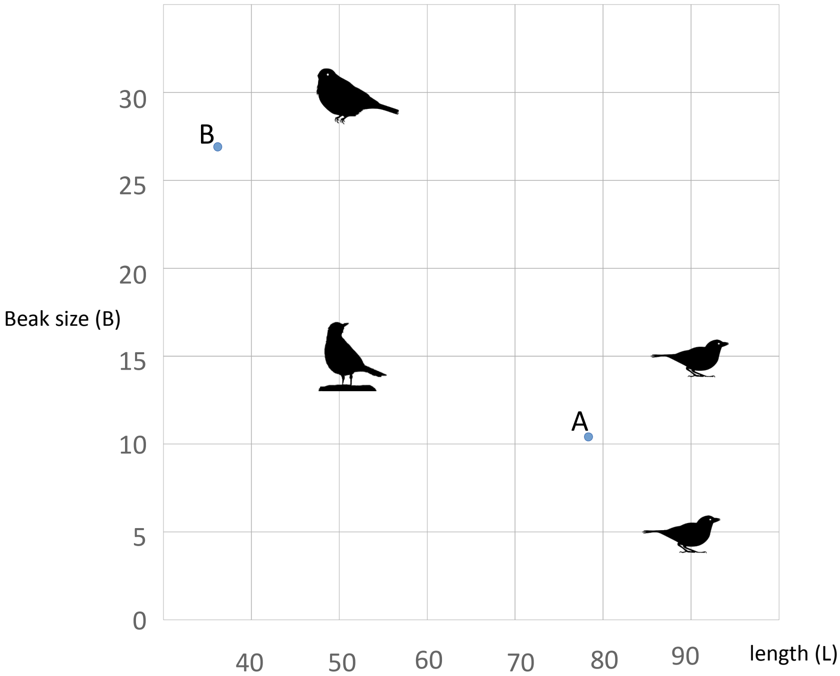
After you finish part C, Alyssa and Ben excitedly report a breakthrough: just by looking at the graph of constraints in Part A, you can tell that you need at least five spaceships to solve the problem. (So in fact, you do not need to do constraint propagation to learn that.)

Explain their reasoning:

Quiz 2, Problem 2 : Nearest neighbors & Identification Trees (50 points)

Part A: For the birds (24 points)

Your friends, who are avid bird watchers, have invited you on a bird-watching expedition—but before you go, you must prove your skill at identifying birds. Remembering an old adage about proximity in feature space, you decide you'll use nearest neighbors classification to classify them.



The chart above shows several well-known examples of birds, plotted with respect to **Length** (L) and **Beak size** (B). All points are at the intersection of gridlines. The following three classes of bird are represented:

Sparrow **Robin** **Finch**

A1. (16 points) On the chart above, draw the decision boundaries produced by the one nearest neighbor algorithm for this training data set. (**<u>Ignore</u>** the points marked A and B for now).

A2. (8 points) Now refer to the points marked A and B in the chart. Determine the classification of each point (**S**parrow, **R**obin, or **F**inch) using one nearest neighbor and three nearest neighbors. If the classification can't be decided, put "CAN'T DECIDE" instead.

	Using one nearest neighbor classification	Using three nearest neighbor classification
A		
B		

Part B: Avian similitude (13 points)

After proving your mettle, you've now joined your friends on their birdwatching expedition. You still don't know much about birds, but your friends have given you a field notebook with five examples they've seen recently:

#	Classification	Length (L)	Beak size (B)	Vibrantly colored (V) (+1 = yes, -1 = no)
1	Finch	50	1	-1
2	Robin	80	2	+1
3	Sparrow	90	3	+1
4	Robin	55	15	+1
5	Finch	65	30	-1

Unlike the previous problem, these examples have a third feature—whether they're Vibrantly colored (V) or not. You still want to use nearest neighbors classification, but computing the Euclidean distance in three dimensions is tricky. Therefore, you decide to *transform this three dimensional feature space into a one-dimensional space* using the following transformation:

$$transform = L + (B \cdot V)$$

★ Using this transformation, you can now easily compute the distance between any two points by transforming each point then computing the <u>one</u>-dimensional euclidean distance between the results (that is $d(x, y) = |x - y|$).

B1. (3pts) Apply the transformation to the five example birds in the table above. Some of the answers are already filled in for you.

Example #	1	2	3	4	5
Transformed value	49			70	

B2. (5pts). Suddenly, your intense data manipulation is interrupted by the appearance of a bird! You quickly determine its characteristics using your keen intuitive sense of length and vibrancy. This specimen has a length of L = 70, a beak size of B = 10, and is vibrantly colored (V = +1).

What is its transformed value? [ ]

Using the five examples and the one-dimensional euclidean metric, how would you classify this new bird using

1 nearest neighbor :	Finch	Robin	Sparrow	CAN'T DECIDE
3 nearest neighbors :	Finch	Robin	Sparrow	CAN'T DECIDE

(Circle the best answer in each case)

B3. (5pts) In general, when using *k* Nearest Neighbors classification, what problem occurs when you choose a value for *k* that is too high (that is, when you take into account too many nearest neighbors)? A one-word answer is all you need.

Part C: Birds in trees (13 points)

For comparison, you decide to train an identification tree classifier using the same collection of birds. (The table is reproduced below)

#	Classification	Length (L)	Beak size (B)	Vibrantly colored (V) (+1 = yes, -1 = no)
1	Finch	50	1	-1
2	Robin	80	2	+1
3	Sparrow	90	3	+1
4	Robin	55	15	+1
5	Finch	65	30	-1

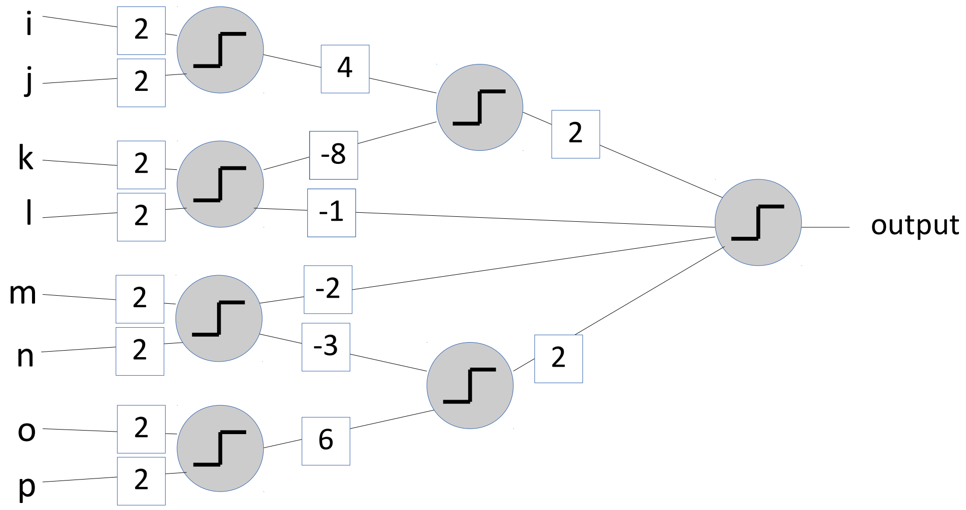
C1. (11pts) Construct the greedy disorder-minimizing ID tree for this dataset. You may use any of the following tests: Length > threshold, Beak size > threshold, Vibrantly colored = Yes/No. Draw the ID tree in the space below.

C2. (2 pts) How would your ID tree classify the bird you spotted in part B? (remember, it had a length of L = 70, a beak size of B = 10, and was vibrantly colored [V = +1]).

Finch Robin Sparrow CAN'T DECIDE

Quiz 3, Problem 1 : Neural Nets

Part A: Foward propagation (15 points)



The neural net depicted above consists of seven neurons which use the *unit step function with a threshold of zero* $u(x)=\begin{cases}0 & x<0\\1 & x\geq 0\end{cases}$. For each of the inputs described below, determine whether the output of this neural net will be zero, one, or whether the output can't be determined from the information in the input description.

For your convenience, duplicate copies of this neural net are provided on a tear-off sheet after the last page of the exam.

Input description	**Output (Circle the best answer)**
All inputs are equal to 1	ZERO ONE CAN'T BE DETERMINED
k, l, o, p are all positive and $m+n<0$	ZERO ONE CAN'T BE DETERMINED
$i+j>0$ and $o+p>0$	ZERO ONE CAN'T BE DETERMINED
$m+n+k+l<0$ and $o+p>0$	ZERO ONE CAN'T BE DETERMINED
i, j, k, l are all equal to 2 and m, n, o, p are all equal to 1	ZERO ONE CAN'T BE DETERMINED

Part B: Minimum network size (20 points)

For each of the following behaviors, indicate the **minimum** number of artificial neurons required to build a neural net that can exhibit the behavior for all possible input combinations. You may set the weights and thresholds of each neuron however you like. If no neural net can exhibit the behavior, write NONE instead.

Assume that each neuron uses the unit step threshold function, and that each network has a single output.

1. Return the sum of 1024 integers.

2. Compute XOR of 2 binary inputs.

3. Draw a plane in three dimensions; classify all the points above the plane as 1 and all the points below the plane as 0.

4. Given ten binary inputs, count the number of inputs that are equal to 1; If exactly five inputs are equal to 1, return 1. Otherwise, return 0.

5. Given 42 integer inputs, always return 1.

Part C: Gradient Ascent (15 points)

Recently, you've pulled your favorite neural net off the shelf and have been experimenting with the following unusual performance function: you pick a particular weight w_p from your neural net, and define the performance as $P=\frac{1}{2}(w_p)^2$.

C1. Now, if the current value of w_p= +1 and the training rate is $r=\frac{1}{2}$ determine the updated value of w_p after one round of training with back propagation. (Remember, for any weight in the network $\Delta w=r\cdot\frac{dP}{dw}$)

C2. A friend suggests that if you perform several additional rounds of back propagation using this process, the value of the performance function will continue to increase. Do you agree or disagree? Explain your answer.

Agree Disagree

C3. You perform 1000 rounds of training on various representative inputs, using the same performance function and rate parameter. $P = \frac{1}{2}(w_p)^2$, $r = \frac{1}{2}$. Which of following can you deduce about the neural net after these 1000 rounds have been completed? **Note that there is only one correct deduction.**

A The updated value of w_p does not depend on the particular training data used.

B The neural net may overfit the data

C Every other weight in the network will converge to the final value of w_p.

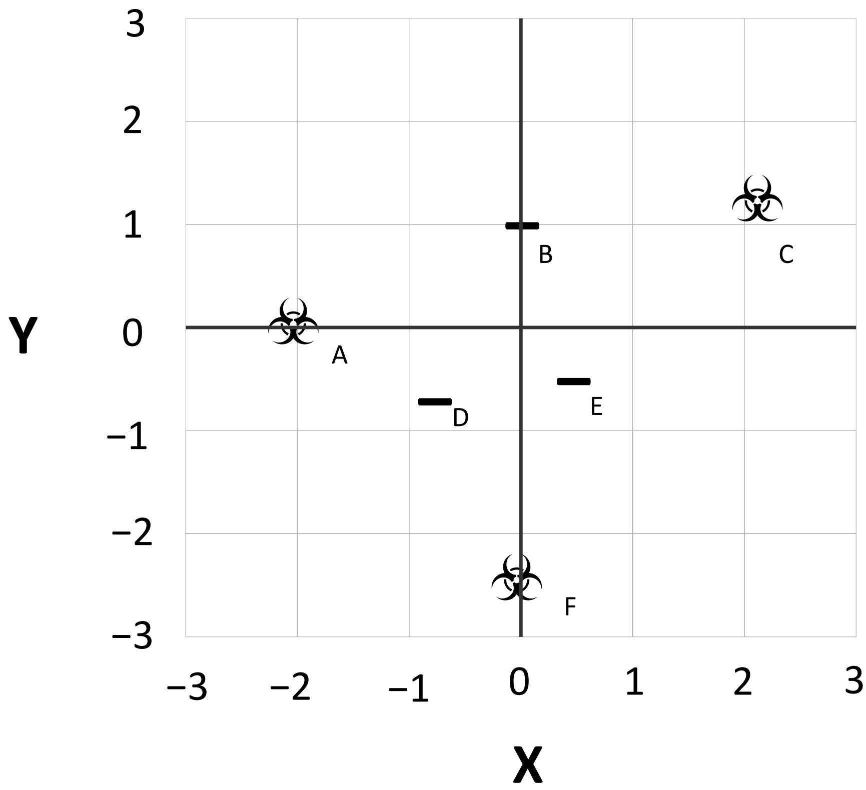
D The outputs of the neural net will not change.

E The entire network behaves like a constant function, outputting the same value regardless of its inputs.

For credit, briefly explain your answer:

Quiz 3, Problem 2 : Support Vector Machines (50 points)

Intrepid navigator Norma Cenva is attempting to travel through foldspace while avoiding all hazardous spots. Below is a map of foldspace, marked with previously detected *hazardous* spots (☣) and *non-hazardous* spots (−).



Part A (10 points)

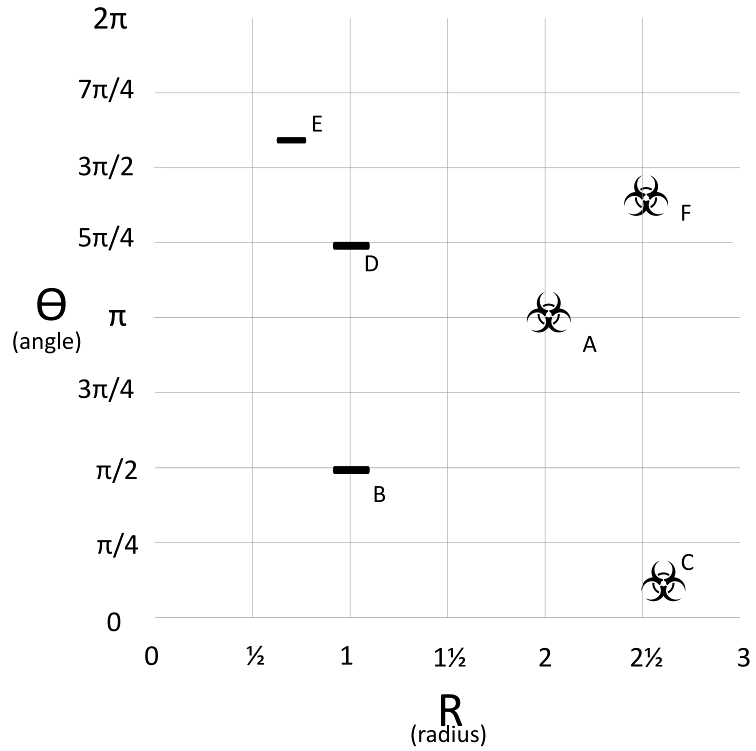
Invigorated by the spice melange, Norma must quickly decide how to separate the hazardous and non-hazardous spots. Which of the following kernel functions can separate this training data?

(Circle **all** answers that apply. If none of these kernels apply, circle "None of these" instead.)

A	Linear kernel $K(\vec{u},\vec{v})=\vec{u}\cdot\vec{v}$
B	Quadratic kernel $K(\vec{u},\vec{v})=(\vec{u}\cdot\vec{v}-d)^2$
C	Radial basis kernel $K(\vec{u},\vec{v})=\exp(-\|\vec{u}-\vec{v}\|/2\sigma^2)$
D	Decomposable kernel #1 $K(\vec{u},\vec{v})=\phi(\vec{u})\cdot\phi(\vec{v})$ where $\phi(\langle x,y\rangle)=x^2+y^2$
E	Decomposable kernel #2 $K(\vec{u},\vec{v})=\phi(\vec{u})\cdot\phi(\vec{v})$ where $\phi(\langle x,y\rangle)=x\cdot y$
F	None of these.

Part B (10 points)

Norma decides instead to use a polar transformation on her data: $\phi(\langle x, y\rangle) = \langle r, \theta\rangle$. The results of this transformation are shown below.

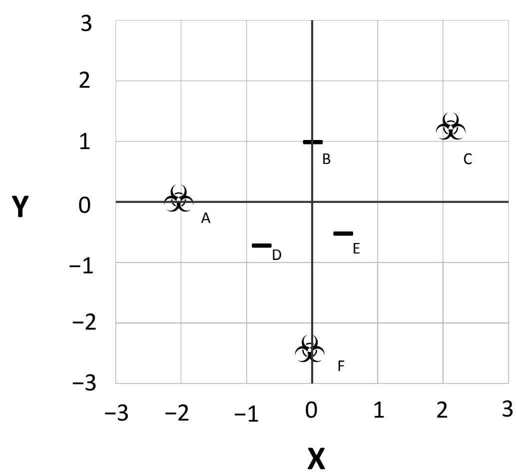


B1. Construct the SVM classifier in this transformed space. On the diagram above,

1. Draw the SVM boundary **with a solid line.**
2. Draw the gutters **as dashed lines**.
3. **Circle** the support vectors.

B2. What does this boundary look like *in the original, pre-transformed space* (shown on the right)?

Sketch the boundary on the diagram, and briefly **describe** its shape below:



Part C (10 points)

The law forbids using complex computers, so Norma must rely on you to calculate the SVM parameters by hand. Compute the following parameters for the boundary $\vec{w} \cdot \vec{x} + b \geq 0$ (**<u>in the transformed space)</u>**. Assume that hazardous points (☣) are positive and non-hazardous points (−) are negative.

The width of the margin =

(Reminder: check the units on the coordinate axes)

The normal vector $\vec{w}$ =

The offset parameter b =

$\alpha_A =$ $\alpha_B =$ $\alpha_C =$

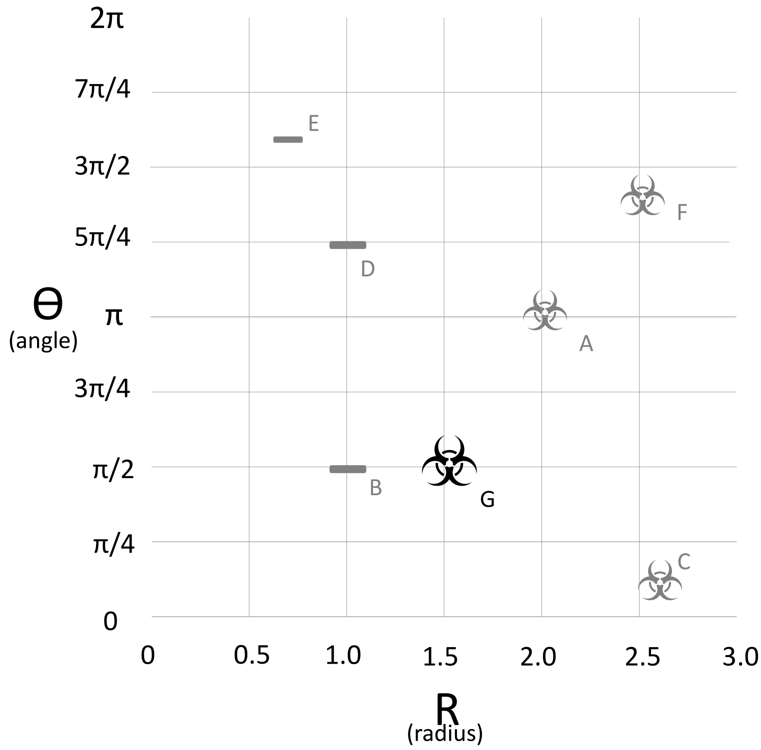
$\alpha_D =$ $\alpha_E =$ $\alpha_F =$

Show your work for partial credit.

Part D (10 points) Norma has just learned that her heighliner ship is located at the point X=1, Y=1. (Of course, this is in the original *pre-transformed* space). According to the SVM classifier, is this a hazardous spot, a non-hazardous spot, or indeterminate? (Circle one.)

HAZARDOUS (☣) NON-HAZARDOUS (−) CAN'T BE DETERMINED

Part E (10 points): Finally, suppose Norma learns that there is a hazardous point **G** located at X =0, Y = 1.5 in the original space. She wants to determine a new SVM boundary using point **G** as a seventh training point. (The *polar transformation* of G is plotted below.)



With the addition of this new training point G, how will the supportiveness values α_i change relative to their original values in part C?

α_A	INCREASE	DECREASE	NO CHANGE
α_B	INCREASE	DECREASE	NO CHANGE
α_C	INCREASE	DECREASE	NO CHANGE
α_D	INCREASE	DECREASE	NO CHANGE
α_E	INCREASE	DECREASE	NO CHANGE
α_F	INCREASE	DECREASE	NO CHANGE

Quiz 4, Problem 1 : Adaboost (50 points)

Instructions: The following boosting questions are all independent of each other. (They might have different numbers of training points, different classifiers, etc.)

★There are 5 problems. Each one is worth 10pts.★

❶ **(10 pts)** I have fifteen training points (A, B, C, …, M, N, O) and six classifiers (h_1, h_2, h_3, h_4, h_5, h_6) . The following table shows which points each classifier misclassifies: there's an ***✗*** in a cell if the classifier misclassifies the training point, otherwise the cell is blank. (For example, h_1 misclassifies points A, B, C, D, E.) Finally, notice that all the classifiers make overlapping errors.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
h_1	✗	✗	✗	✗	✗										
h_2	✗					✗	✗	✗	✗						
h_3		✗				✗				✗	✗	✗			
h_4			✗				✗			✗			✗	✗	
h_5				✗				✗			✗		✗		✗
h_6					✗				✗			✗		✗	✗

A) Is it possible to construct a perfect ensemble classifier out of these classifiers (h_1, h_2, h_3, h_4, h_5, h_6) , either by using Adaboost or by picking values of α by hand? (Circle one)

YES NO

Explain.

B) Regardless of your above answer, if you performed five rounds of Adaboost, which weak classifiers would Adaboost choose? (Break ties by picking the weak classifier that comes first in this list: **h_1, h_2, h_3, h_4, h_5, h_6**).

Note: You may solve this problem by inspection.

	Round 1	Round 2	Round 3	Round 4	Round 5
Weak Classifier					

❷ **(10 pts)** Lily says "I always write my training weights so that they all have the same denominator (and integer numerators). Imagine my surprise when I started working on my latest pset and found that all of my training weights had the same *odd* denominator!"Megan replies "Surely that's impossible because a subset of the weights must always sum to ½."

It isn't impossible in one particular round; what were the weights of Lily's points?

weight_A	**weight_B**	**weight_C**	**weight_D**	**weight_E**

Next, if just points A and B were misclassified that round, what were the updated weights in the following round?

weight_A	**weight_B**	**weight_C**	**weight_D**	**weight_E**

❸ **(10 pts)** The following weights cannot possibly be right. Why not?

weight_A	**weight_B**	**weight_C**	**weight_D**	**weight_E**
1/14	1/14	?	3/14	8/14

❹ **(10 pts)**. I have an unspecified number of training points. In the first round of boosting, I chose a weak classifier with $0 < \varepsilon < ½$. After updating the weights for the second round, all of my training points had weights (strictly) less than ½.

A) What is the smallest number of training points I could have had?

B) Suppose I had *exactly* that many training points—what voting power (α) was assigned to the classifier I picked in the first round?

Show your work for partial credit. In particular, be sure to justify your answer to (A) by explaining why your answer works, and why no smaller answer will work

❺ **(10 pts)** Find the values of the missing weights below, **assuming weight$_A$ is less than or equal to weight$_E$ in each round.**

Round #	**weight$_A$**	**weight$_B$**	**weight$_C$**	**weight$_D$**	**weight$_E$**
6033		2/9	3/9	5/18	
6034		2/12	3/12	5/12	

Hint #1: Solve for the weights in round 6034 first.

Hint #2: Remember the formula for updating the weights.

$$w_{new} = \begin{cases} \frac{1}{2}\frac{1}{\epsilon} w_{old}, & \textit{if misclassified} \\ \frac{1}{2}\frac{1}{1-\epsilon} w_{old} & \textit{otherwise} \end{cases}$$

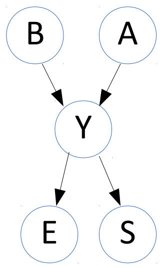
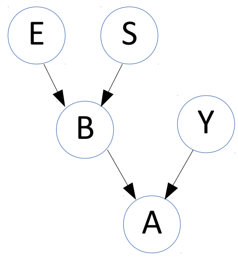
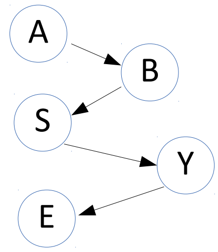
Show your work for partial credit:

Quiz 4, Problem 2 : Probabilistic Inference (50 points)

Part A: BAYES nets (20 points)

In the figure below, there are three Bayes nets and some independence statements. For each of the statements below and each Bayes net, circle True if the statement is true for the net, or False if the statement is false for the net.

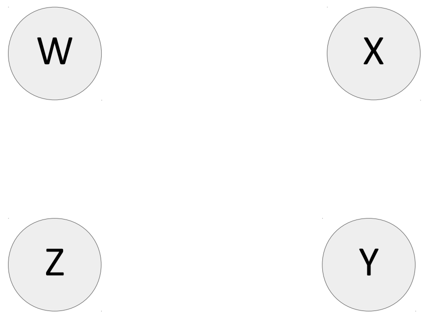
<u>**Note:**</u> Assume that all entries in all conditional probability tables are distinct; this means that the only independence statements that are true are the ones enforced by the shape of the network.

	 **Net 1**		 **Net 2**		 **Net 3**	
E is independent of S	True	False	True	False	True	False
$P(B\|Y,A) \neq P(B\|Y)$	True	False	True	False	True	False
Given B and Y, A is independent of E	True	False	True	False	True	False
S is independent of B given Y	True	False	True	False	True	False
$P(B\|E) = P(B)$	True	False	True	False	True	False

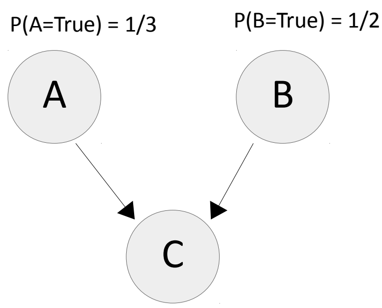
Part B (5 points)
Your friend tells you that she can compute every joint probability between four variables (W, X, Y, Z) as the following product of conditional probabilities:

$$P(W, X, Y, Z) = P(X)P(W|X)P(Y|X)\,P(Z|W,X)$$

Draw arrows in the Bayes net below to reflect the conditional independence properties described by this equation. (You don't need to draw the conditional probability tables; just the arrows are enough.)



Part C (25 points)
Compute the following probabilites for the Bayes net shown on the right. Express your answers in terms of fractions, integers, and *x*.



A	B	P(C=True|A,B)
True	True	1-x
True	False	x
False	True	x
False	False	x/2

C1 (5 pts). The probability that C is true.

C2 (5 pts). The probability that B is true, given that C is true.

C3 (10 pts). The probability that B is true, given that C is true and A is false.

C4 (5 pts). Consider the inequality $P(B|C) < P(B|C\overline{A})$. Find a value of *x* which makes this inequality true and which satisfies the requirements of the Bayes net on the previous page. If it's impossible, write IMPOSSIBLE instead.

Show your work.

SRN, bonus points

Circle the one **best** answer for each of the following questions. There is **no penalty for wrong answers**, so it pays to guess in the absence of knowledge.

1. You have just constructed an image of the red white and blue US flag with a drawing program. You decide to use it as the source in a self-organizing map experiment. The target starts off as a matrix of random colors. You expect the target, eventually, to

1. Be entirely white, the color with the most area in the US flag.
2. Be entirely covered with red, white, and blue patches of varying size.
3. Look like a blurry version of the US flag.
4. Be unchanged, except for 3 points and their neighbors.
5. Be entirely covered by three regions, one red, another blue, a third white.

2. Self-organizing maps have properties that invite comparison with:

1. Social networks.
2. Boosting.
3. Sensory and motor cortex.
4. Start up companies.
5. Bee hives.

3. Harrell's research on games includes work on:

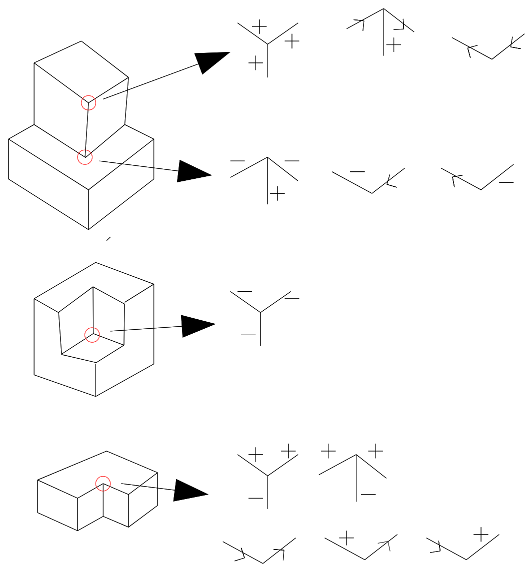
1. Better modeling of physics.
2. Better modeling of light reflectance.
3. Automatic generation of grotesque monsters.
4. Correlation between gaming and mental illness.
5. Enriched models of character traits and biases.

4. When Harrell talks about concept mapping, he is most likely referring to:

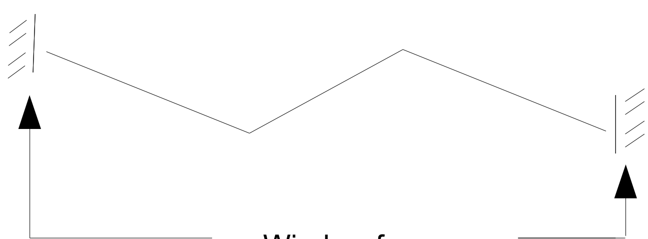
1. Research relating game characteristics to popularity.
2. Analysis of games high-school students like to play in various cities.
3. Using business games in middle school to motivate interest in math.
4. The role of metaphors grounded in sensation, such as “high society.”
5. Cataloging brain centers actuated when we use abstract words, such as “retribution.”

SRN, Quiz 2, Constraint propagation

You have attended Winston's lecture on line drawing analysis and dutifully recorded the table of legal junction arrangements for three-faced vertexes and the vertexes from which they come:



You see a fragment of an object through a window. The T junctions,  , provide no constraint. Note that there are two L junctions; for each L, there are six possible junction labels. Remember: junctions may be rotated however you like, but not mirrored.



How many of the 12 junction labels are eliminated using pure constraint propagation (no search)?

How many ways can the fragment be labeled such that the labels are consistent with the three-faced vertex assumption?

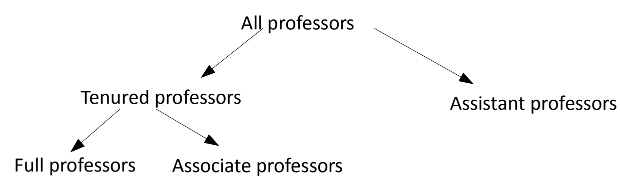
Can the fragment come from an object manufactured by Acme, a company that manufactures only objects that have only convex edges (a straight line connecting any two points in or on the object lies entirely inside the object)? If not, write no; if so, indicate how many ways the fragment can be labeled.

SRN, Quiz 3, Near-miss learning

President Reif is growing tired of interviewing candidates for MIT's next Provost. You tell him that if he gives you his decisions about who should be a finalist after a few interviews, you can use near miss learning to build a model, so he won't have to waste more time; he can use your model to decide who should be a finalist.

Every candidate has a particular **Professorship Rank**, identifies as a particular **Gender**, excels at one particular **Specialty**, works in a particular **Department**, and was born in a particular **Region** of the world, as follows:

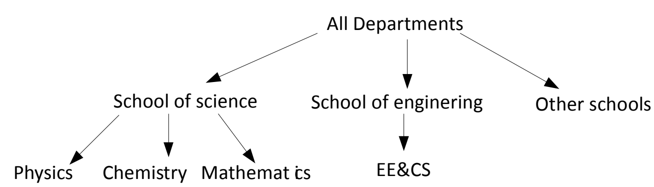
Professor ranks form a hierarchy



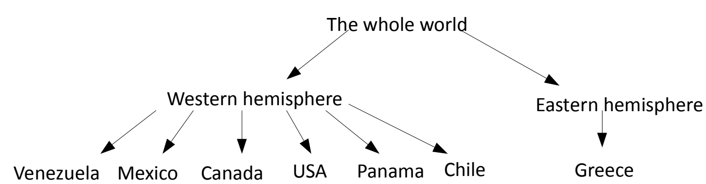
Gender forms a set { Male, Female }

Specialty form a set {money, education, research}

Departments form a hierarchy:



Regions form a hierarchy



President Reif provides you with his notes about each of the first eight candidates:

Candidate	**Finalist?**	**Prof. Rank**	**Gender**	**Specialty**	**Department**	**Region**
1	Yes	Full	Male	Money	Physics	Venezuela
2	Yes	Full	Female	Money	Chemistry	Venezuela
3	No	Full	Male	Education	Math	Venezuela
4	Yes	Full	Male	Money	Physics	Canada
5	Yes	Associate	Female	Money	Physics	Panama
6	No	Assistant	Male	Research	Math	Greece
7	Yes	Associate	Female	Money	Chemistry	USA
8	No	Assistant	Female	Money	EE&CS	Chile

Your task is to apply near-miss learning to President Reif's notes in order to describe which candidates should be finalists (and which ones shouldn't). Start by using Candidate 1 as a seed model, then train using the rest of the candidates in order.

- For each example, fill out the table describing what definite rule you learned, such as "Finalists must be from the School of *X*" or "Region doesn't matter".
- Also include the heuristic(s) you used; choose from
  - require link (feature must be present)
  - forbid link (feature must not be present)
  - extend set (include more varieties)
  - drop link (include all possible varieties)
  - climb tree (generalize to a parent category)
- **Note:** Some examples may not teach you anything new. In that case, put "nothing" as the rule learned and "none" as the heuristic. Others may teach you more than one rule.

Candidate	Definite rule learned	Heuristic(s) used
1		
2		
3		
4		
5		
6		
7		
8		

Finally, summarize these results: describe the candidates that meet the critera for being finalists.

SRN, Quiz 4, Miscellaneous

Circle the one **best** answer for each of the following questions. There is **no penalty for wrong answers**, so it pays to guess in the absence of knowledge.

1. The “Strong story hypothesis” states that

1. Stories are more memorable if they contain surprises.
2. Stories are more memorable if the characters have strong emotional reactions.
3. Story understanding is a defining characteristic of human intelligence.
4. Stories are more understandable if the teller uses a listener model.
5. Stories are more understandable if the teller uses strong, active, Anglo-Saxon verbs.

2. The “Exotic engineering hypothesis” states that:

1. We will be able to determine how brains do what they do with newly emerging instruments.
2. The brain makes mysterious use of massive top-down information flow.
3. The brain makes mysterious use of mechanisms used in probabilistic programming.
4. The brain compensates for slow speed by using massive memory.
5. The brain is amazingly energy efficient.

3. Adult humans do poorly in the rectangular room, one blue wall, disorientation experiment when:

1. Subjected to Rolling Stones music, played loud.
2. Subjected to bright lights flashing at the frequency of the alpha rhythm.
3. They have a damaged frontal lobe.
4. They have had no sleep for 24 hours or more.
5. They say back to the experimenter what the experimenter says.

4. The Genesis story understanding system discovers concepts, such as “revenge” by:

1. Noting emotional impact of events.
2. Hard-wired programs, one per concept.
3. Search, driven by concept patterns expressed in English.
4. Looking for concept indicating words, such as *retribution* and *retaliation*.
5. Using dot products on word vectors as in information retrieval.

5. The genesis story understanding system summarizes stories by:

1. Including events involving verbs on a list of high-impact verbs.
2. Using a rule-based summary system.
3. Including events in which the story's hero is the actor.
4. Eliminating events that would be inferred by the reader.
5. Eliminating events that are common to many stories.

6. Ian Tattersall, the often-mentioned paleoanthropologist, believes:

1. Human intelligence developed slowly and incrementally over about 10 million years.
2. Human intelligence emerged suddenly not much more than 50,000 years ago.
3. Human intelligence emerged when homo sapiens mated with Neanderthals.
4. Humans outcompeted Neanderthals because our mitochondria produce more energy
5. Humans are smarter than all other primates, living and dead, because our brains are bigger.

[This verso page intentionally left blank so the tear-off sheets will be tear-offable]

Tear off sheet for Quiz 1, Problem 1: Rule-based systems

Rules:

R0	IF ‘(?x) has wings' THEN ‘(?x) can fly'
R1	IF ‘(?x) is a hero of (?y)' THEN ‘(?x) masters (?y)'
R2	IF ‘(?x) uses (?y)' THEN ‘(?x) is a hero of (?y)'
R3	IF (?x) casts spells THEN ‘(?x) grows silver wings' ‘(?x) uses Magic'
R4	IF ‘(?x) masters wind' THEN ‘(?x) can fly'
R5	IF (AND `(?x) masters Magic' `(?x) is a hero of (?y)' `(?x) can fly') THEN ‘(?x) ascends to godhood'

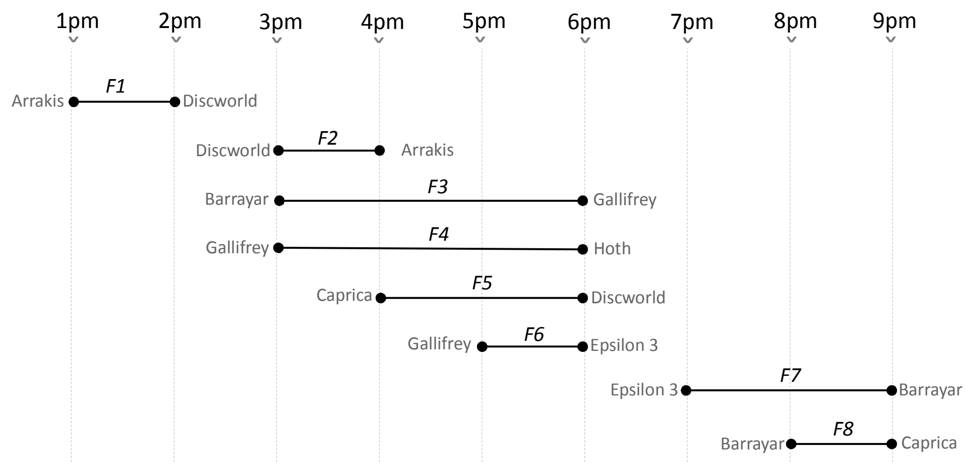
Assertions:

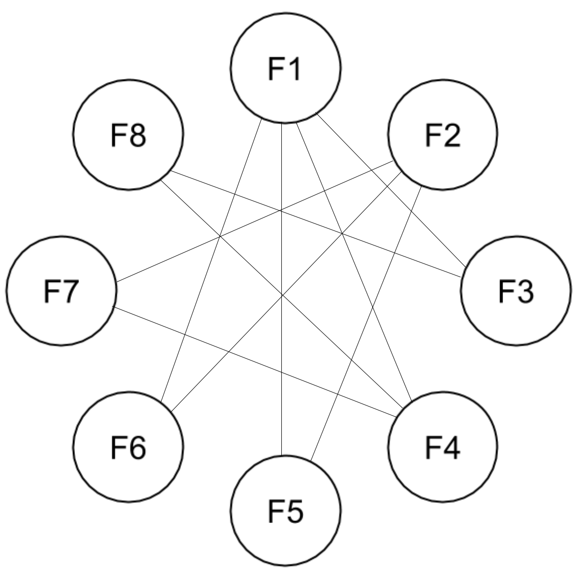
A0: John is a hero of wind

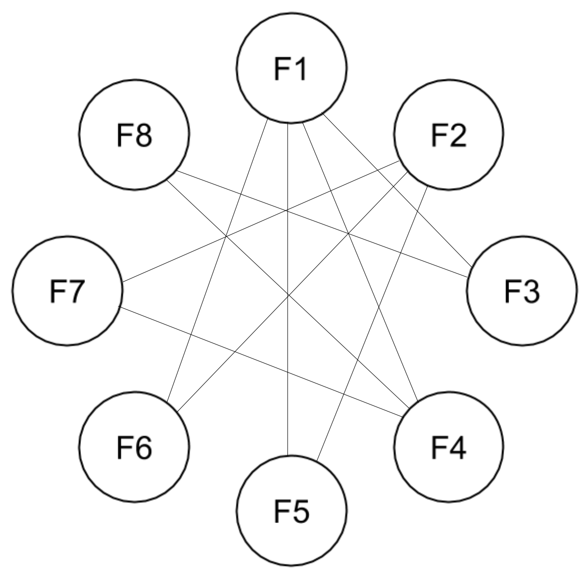
A1: Magic casts spells

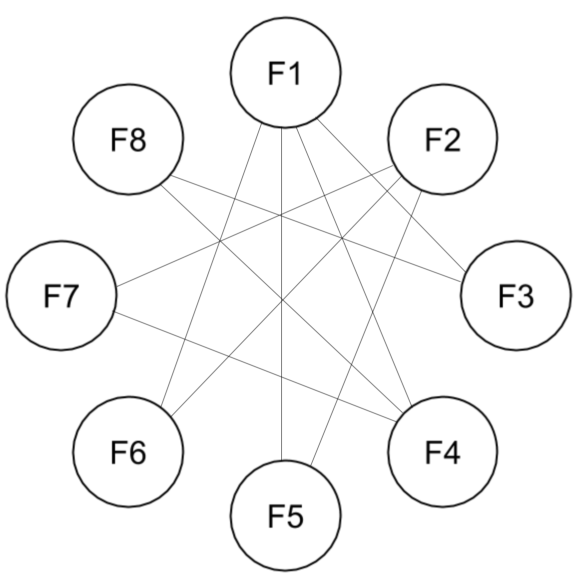
[The back of this Quiz 1 tear-off sheet intentionally left blank]

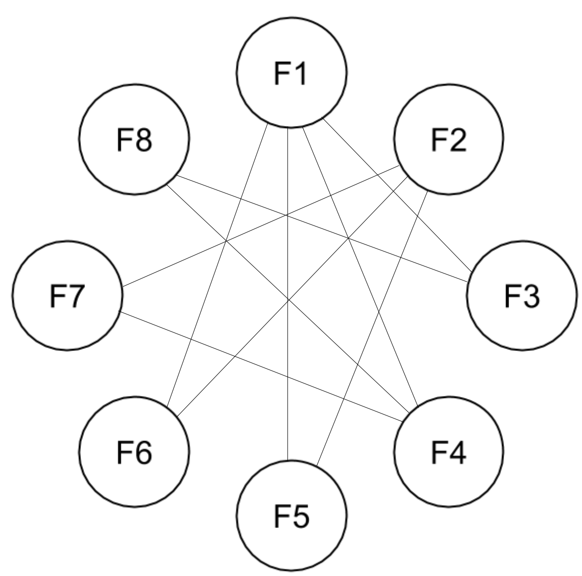
Tear off sheet for Quiz 2, Problem 1: Constraint propagation

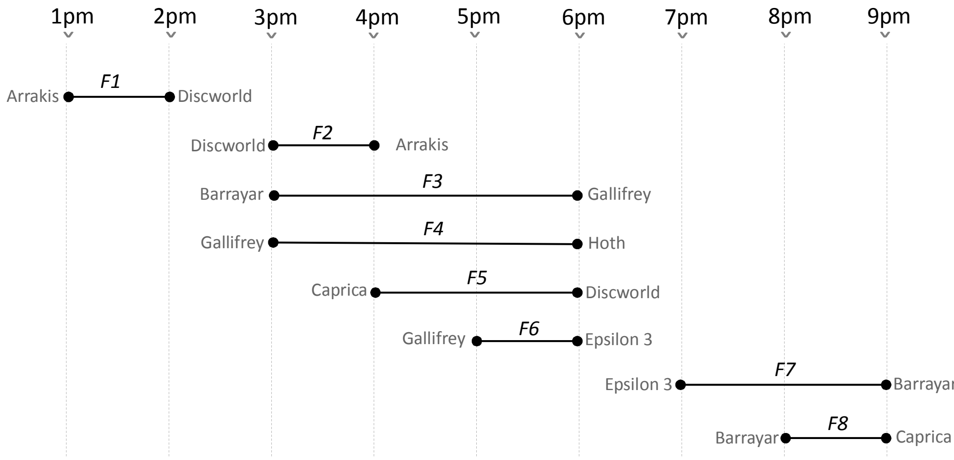


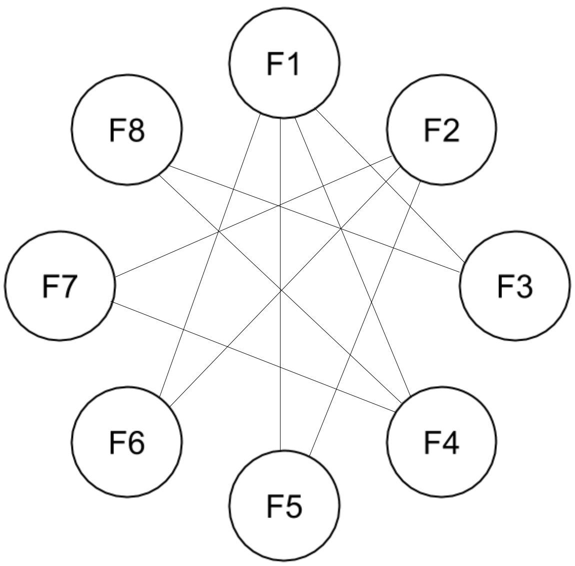


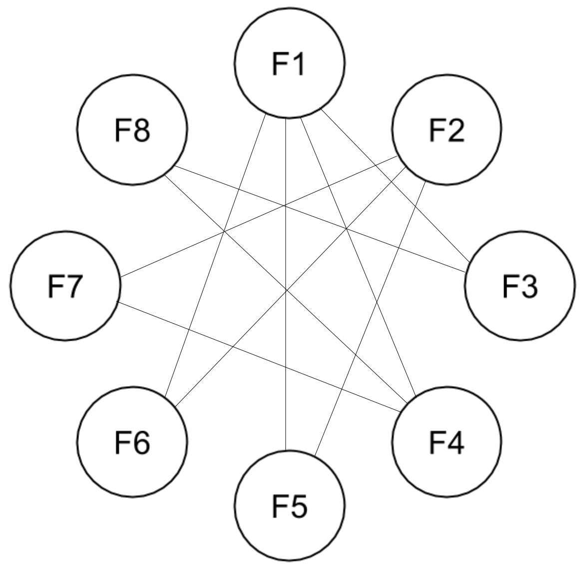


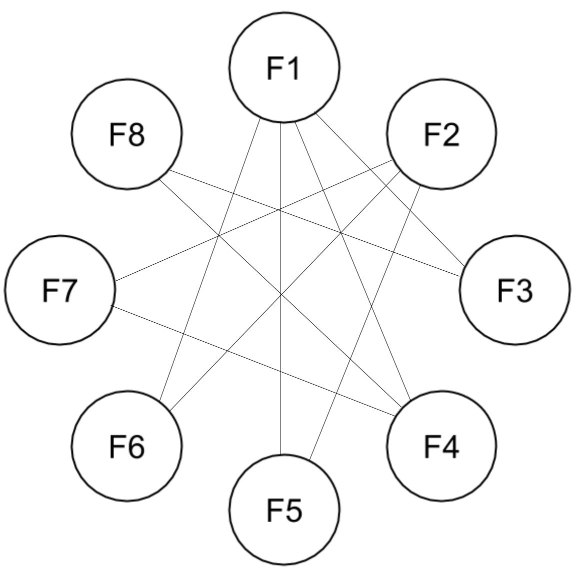


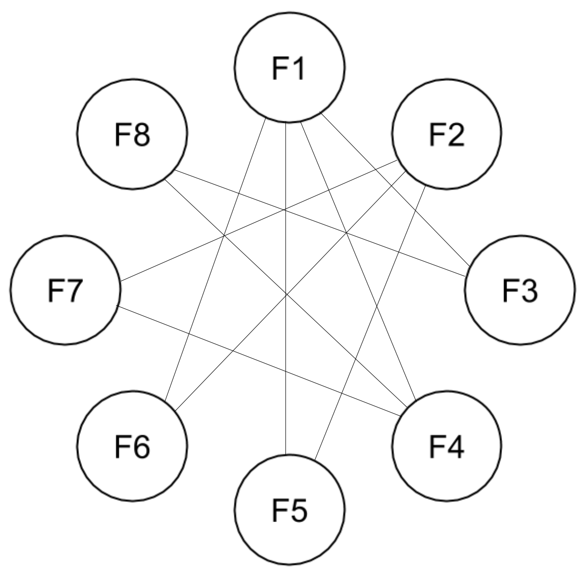



1pm
2pm
3pm
4pm
5pm
6pm
7pm
8pm
9pm
F1
Arrakis
Discworld
F2
Discworld
Arrakis
F3
Barrayar
Gallifrey
F4
Gallifrey
Hoth
F5
Caprica
Discworld
F6
Gallifrey
Epsilon 3
F7
Epsilon 3
Barrayar
F8
Barrayar
Caprica


F1
F2
F3
F4
F5
F6
F7
F8


F1
F2
F3
F4
F5
F6
F7
F8


F1
F2
F3
F4
F5
F6
F7
F8


F1
F2
F3
F4
F5
F6
F7
F8

Tear off sheet for Quiz 3, Problem 1: Neural Nets