

Name	
email	

6.034 Final Examination

December 16, 2009

Circle your TA and principle recitation instructor so that we can more easily identify with whom you have studied:

Erica Cooper	Matthew Peairs	Mark Seifter
Yuan Shen	Jeremy Smith	Olga Wichrowska

Robert Berwick	Randall Davis	Gregory Martin
----------------	---------------	----------------

Indicate the approximate percent of the lectures, mega recitations, recitations, and tutorials you have attended so that we can better gauge their correlation with quiz and final performance. Your answers have no effect on your grade.

	Lectures	Recitations	Megas	Tutorials
Percent attended				

Quiz	Score	Grader
Q1		
Q2		
Q3		
Q4		
Q5		

There are 38 pages in this final examination, including this one. In addition, tear-off sheets are provided at the end with duplicate drawings and data. As always, open book, open notes, open just about everything.

Quiz 1, Problem 1, Rules (50 points)

The administration, worried about the social habits of its students, agrees to finance cross-school-mixers. The 034 TA's decide to fly to England and mix with the students at Hogwarts School of Witchcraft and Wizardry. A merry old time ensues, but the morning after, due to an accidental confundo charm (and perhaps also a large consumption of butterbeer), no one can remember the events that transpired. The 034 staff, in an attempt to show off the power of Muggle logic, promise they can piece together the important events with a rule based system.

Using their keen sense of logic, Matt, Erica, and Mark are able to piece together the following rules:

RULES:

R0 : IF (?X) goes to MIT,
 THEN (?X) is a muggle,
 (?X) consumed butterbeer

R1: IF (?X) made math jokes AND
 (?X) consumed butterbeer
 THEN (?X) was transfigured into a porcupine

R2: IF (?Y) fancies (?X) AND
 (?X) fancies (?Y) AND
 (?Y) is a muggle
 THEN (?X) snogged (?Y)

R3: IF (?X) fancies (?Y) AND
 (?X) made math jokes,
 THEN (?Y) fancies (?X)

R4: IF (?X) made math jokes
 THEN (?X) goes to MIT

You start with the following list of assertions which **is all you have to go on.**

ASSERTIONS:

A0: Olga made math jokes
A1: Yuan goes to MIT
A2: Jeremy made math jokes
A3: Hermione consumed butterbeer
A4: Jeremy fancies Hermione

Part A: Forward Chaining (24 points)

Run forward chaining on the rules and assertions provided for the first 5 iterations. 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. As usual, break ties using the earliest rule on the list that matches. If the earliest rule matches more than once, break ties by assertion order.

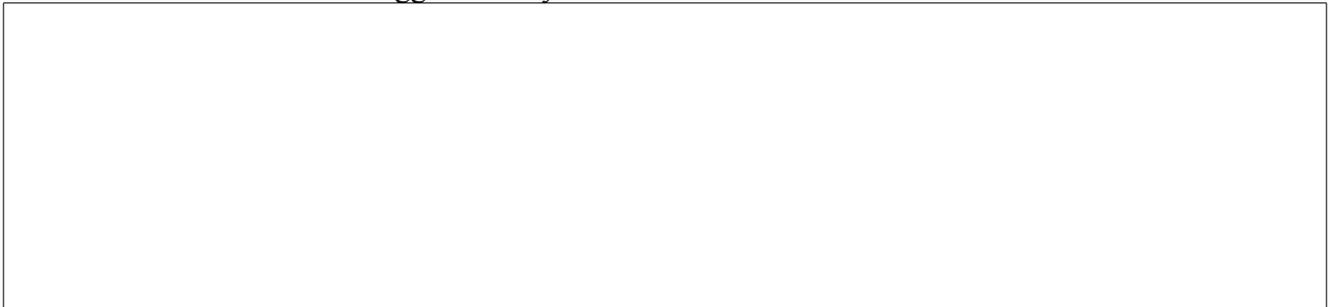
	Matched	Fired	New assertions added to database
1			
2			
3			
4			
5			

Part B: Backward Chaining (26 points)

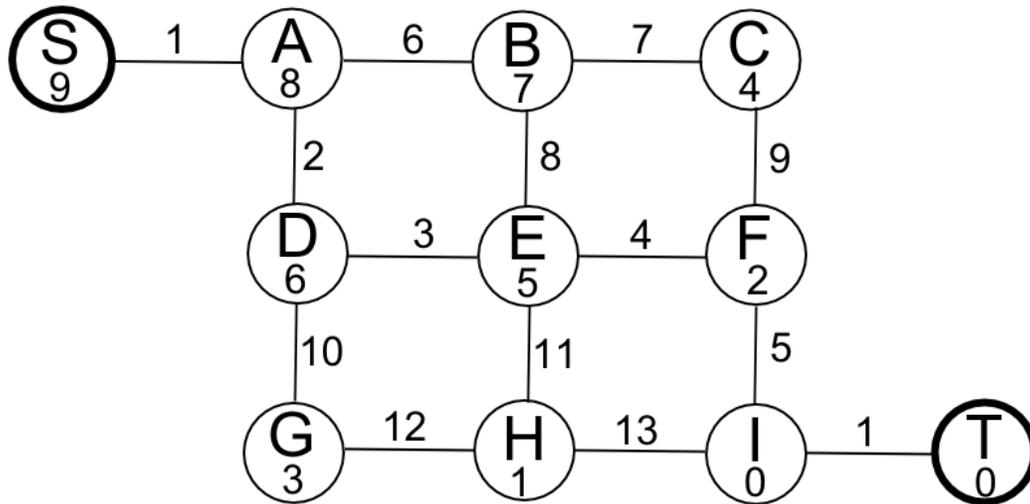
Ron Weasley claims that Hermione snogged Jeremy. Use backward chaining to determine if this event occurred. **Draw the goal tree for this statement.** Partial credit will be given for partial completion of the goal tree.

Hermione snogged Jeremy

Is the claim that Hermione snogged Jeremy true?



Quiz 1, Problem 2: Search (50 points)



Aliens have invaded the MIT campus, thinking that the Stata Center was their mothership. Despite their realization that it was not, they went ahead with their invasion plan anyway. Advanced alien technology has disabled use of phones and internet, so you and your friends decide to go across the river to Boston to get help and to tell everyone what's going on. You decide to use the secret underground network of tunnels underneath the campus to get from your starting location at **S**, to the subway station at **T**.

You have the above graph of the underground tunnels. The edges are labeled with distances, and the nodes are labeled with a heuristic estimate to your destination at **T**. When performing search, ties are broken by choosing the node that is **alphabetically first**.

Note that node G is not the goal node in this problem.

Part A: Depth-First Search. (15 points)

A1: You first attempt to pick your route using **Depth-first search with an extended list**. Draw your search tree:



Also, show your extended list at the time your search terminates:



A2: What is the final path that is found using depth-first search?



Part B: Beam Search (15 Points)

B1: Next, you try to pick your route using **Beam Search with a beam width of 2, using an extended list**. In the event of a tie, use the alphabetical order of the final node on the contending partial paths. Draw your search tree:



Also, show your extended list at the time your search terminates:



B2: What is the route found using beam search?



Part C: Branch and Bound (20 Points)

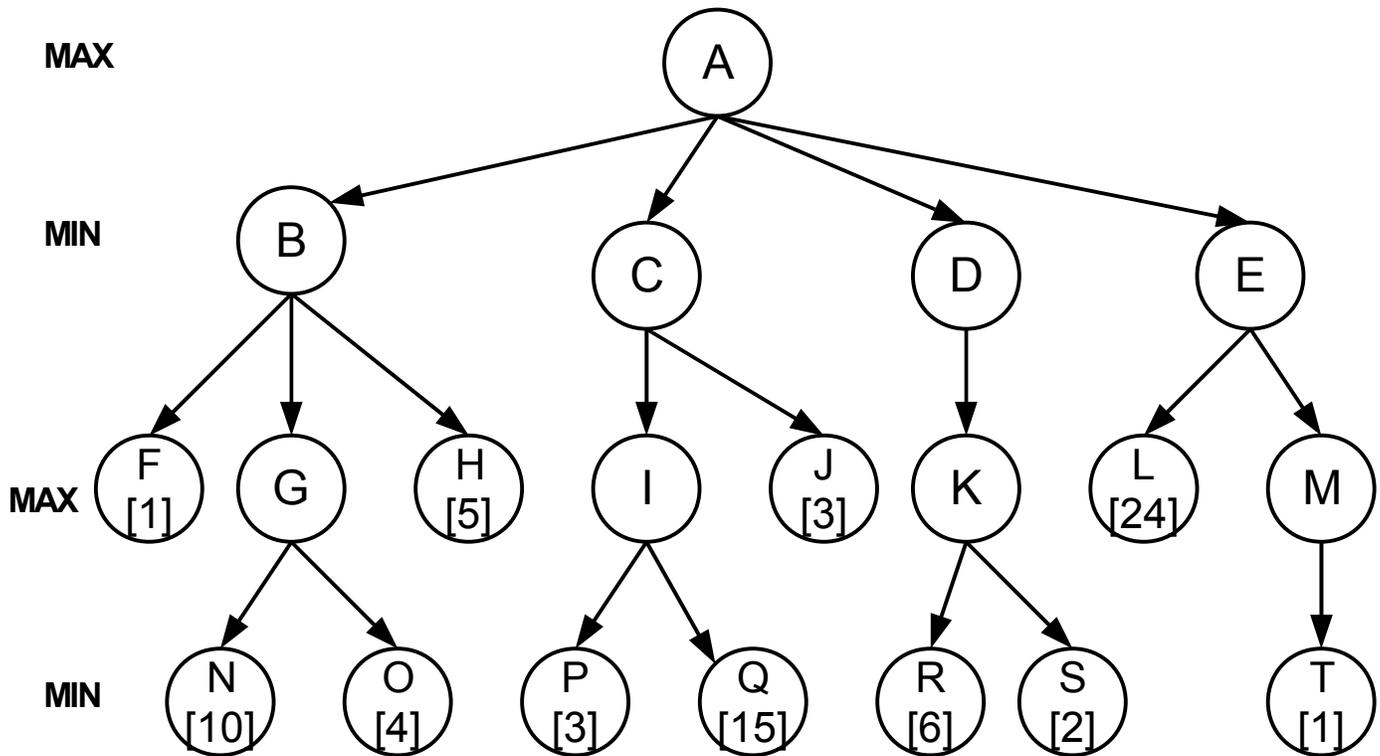
C1: Lastly, you try to plan your route using **Branch and Bound using the path lengths indicated on the graph and also using an extended list.** Draw your search tree and extended list below:

Also, show your extended list at the time your search terminates:

C2: What is the route found using branch and bound?

C3: If you repeat the search using the **A* algorithm with the heuristic values indicated on the graph,** will you find the same route? Include a brief explanation of why or why not.

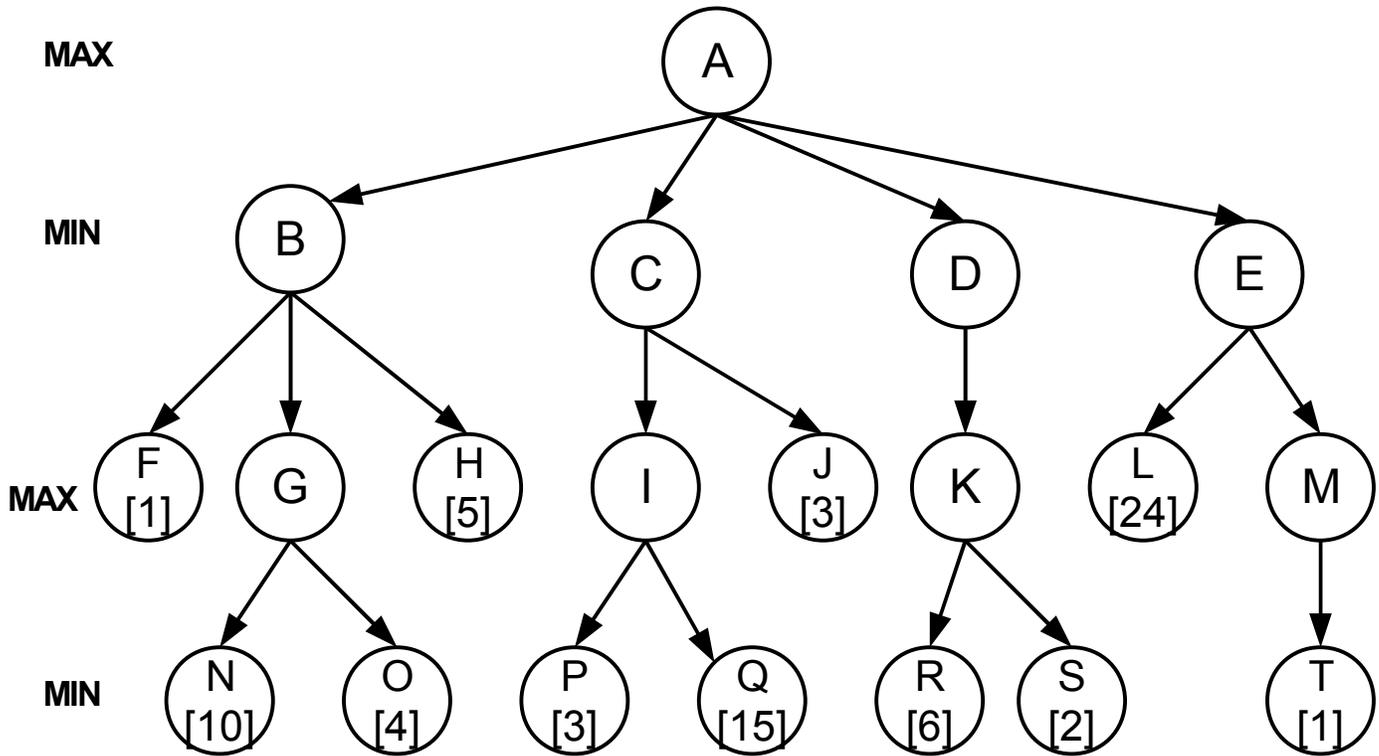
Quiz 2 Problem 1: Games (50 points)



A: Using minimax only, no alpha-beta,, indicate the values of the following nodes. (10 pts)

A =	G =
B =	I =
C =	K =
D =	M =
E =	

B: Using minimax only, what is the best next move from A? (Indicate a letter) (4 pts)



C: Trace the steps of Alpha beta pruning on the same tree above. Note that alpha, betas are updated before pruning occurs, if in doubt consult the reference implementation given on the tear-off sheet.

List the leaf nodes in the order that they are statically evaluated. (10 pts)

--

What are the final Alpha Beta values at node E (4pts)

Alpha =	Beta =
---------	--------

What are the final Alpha Beta values at node A (4pts)

Alpha =	Beta =
---------	--------

D: For a full **binary tree** (branching factor = 2) **of depth 3** (4 layers of nodes including the layer with the root node, 8 nodes in the bottom layer), with the root node being a MAX node, what is the fraction of leaf nodes that are statically evaluated under alpha beta pruning under conditions of maximum pruning? (11 pts)

Use the space below to show your work:

E: For a given depth, d , does the fraction requested in part D increase, decrease, or stay the same with increasing b . Circle your answer (7 pts)

Increase

Decrease

Stay the same

Explain:

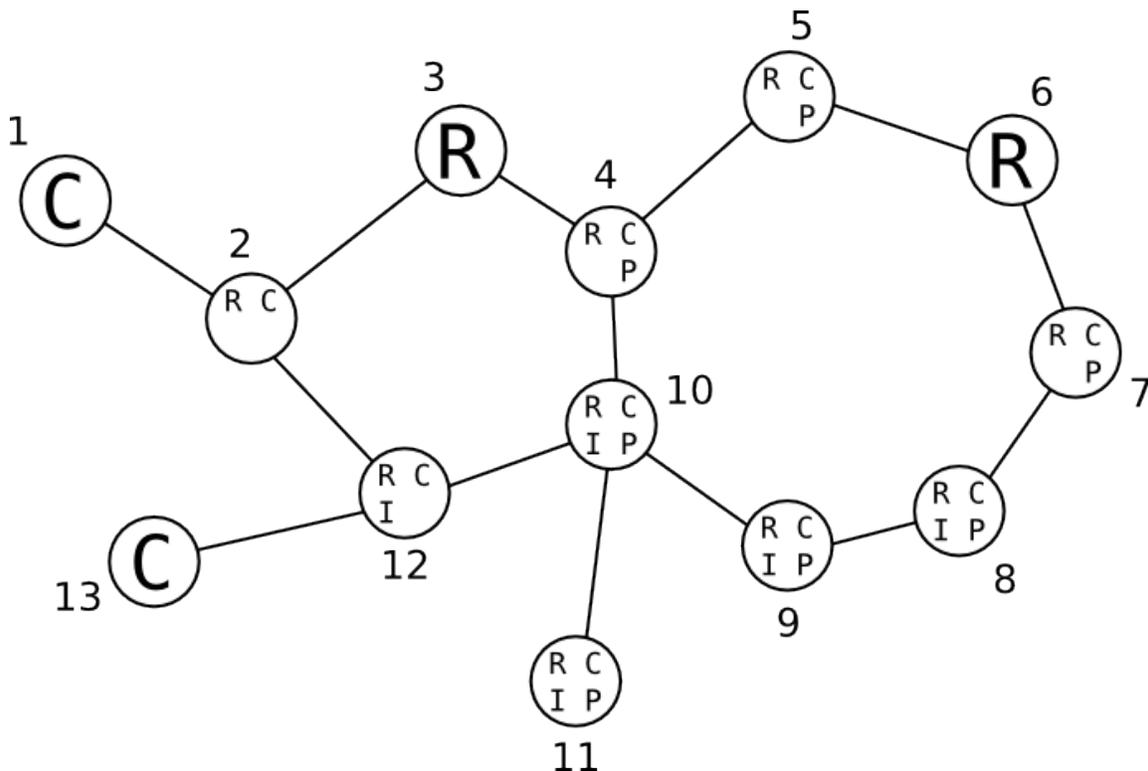
Quiz 2, Problem 2: Constraint Propagation (50 Points)

Important note: there is no search in this problem, only constraint propagation.

On the newly colonized planet Mars, cities consist of pressurized domes connected by a series of tubes. Each dome is designated for a specific type of use, and there are some restrictions on what sorts of domes can be connected to each other:

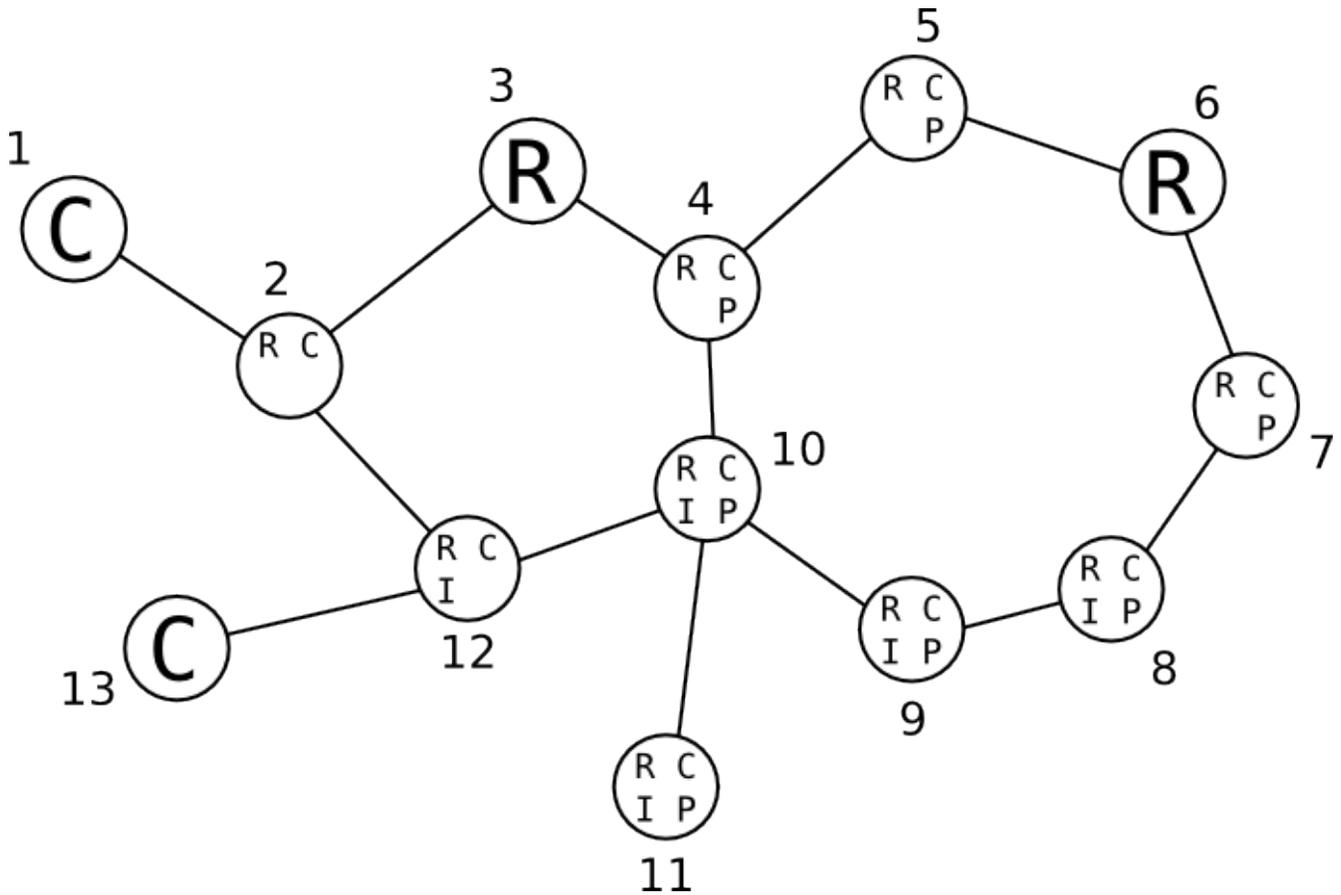
1. A **R**esidential dome can only be connected to another **R**esidential dome, a **C**ommercial dome, or a **P**ark.
2. A **C**ommercial dome can only be connected to a **R**esidential dome, another **C**ommercial dome, or an **I**ndustrial dome.
3. An **I**ndustrial dome can only be connected to a **C**ommercial dome or another **I**ndustrial dome.
4. A **P**ark can only be connected to a **R**esidential dome.

You decide to use constraint propagation to design a new Martian city. You begin with the following partially completed plan, with some domes already designated and their neighboring domains reduced accordingly:



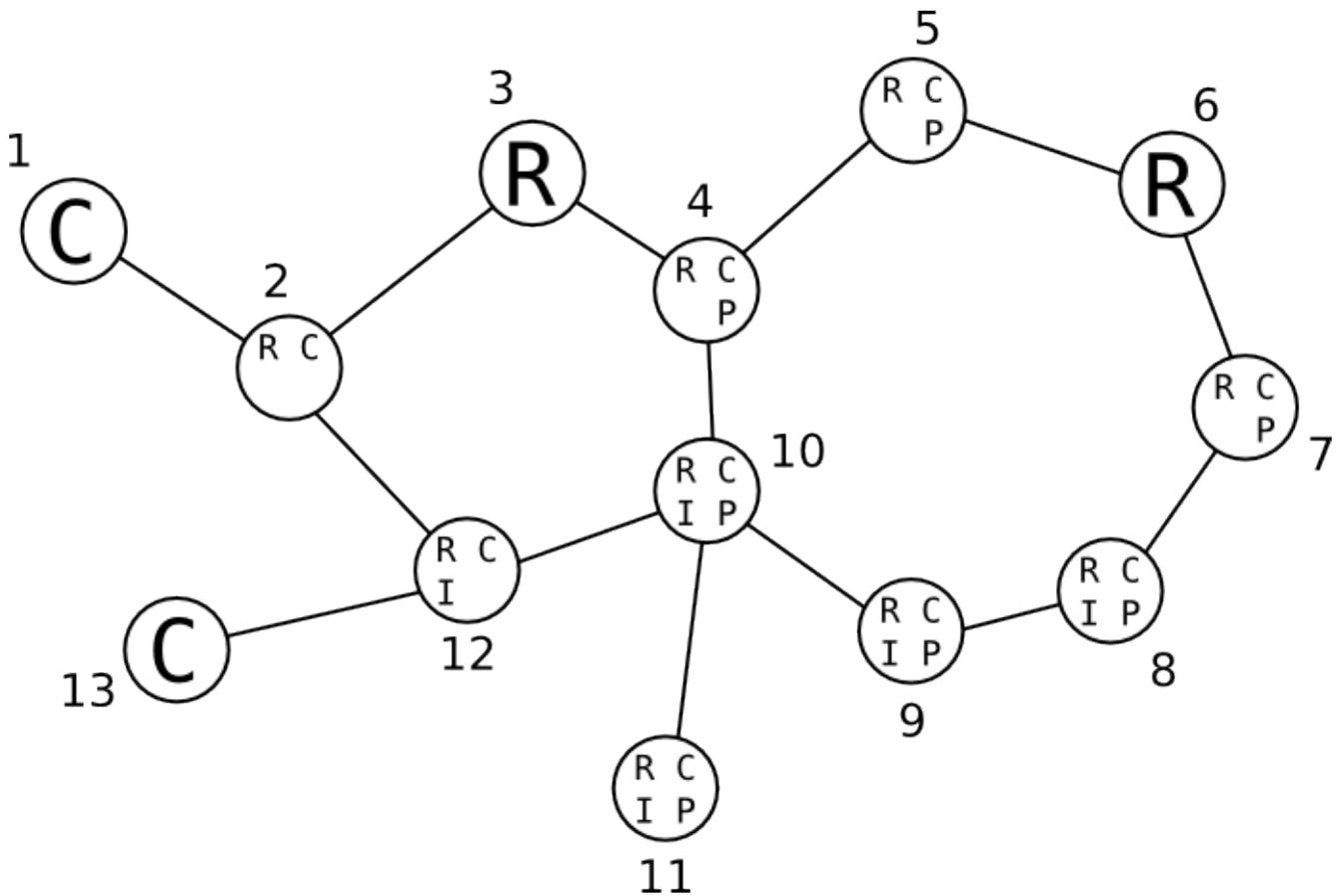
Part A (12 points)

You decide to begin by designating Dome #10 as Industrial. Using forward checking (no propagation beyond the neighbors of the just-assigned variable), show how the domains of the city's domes are reduced by crossing out the appropriate letters in the map below:



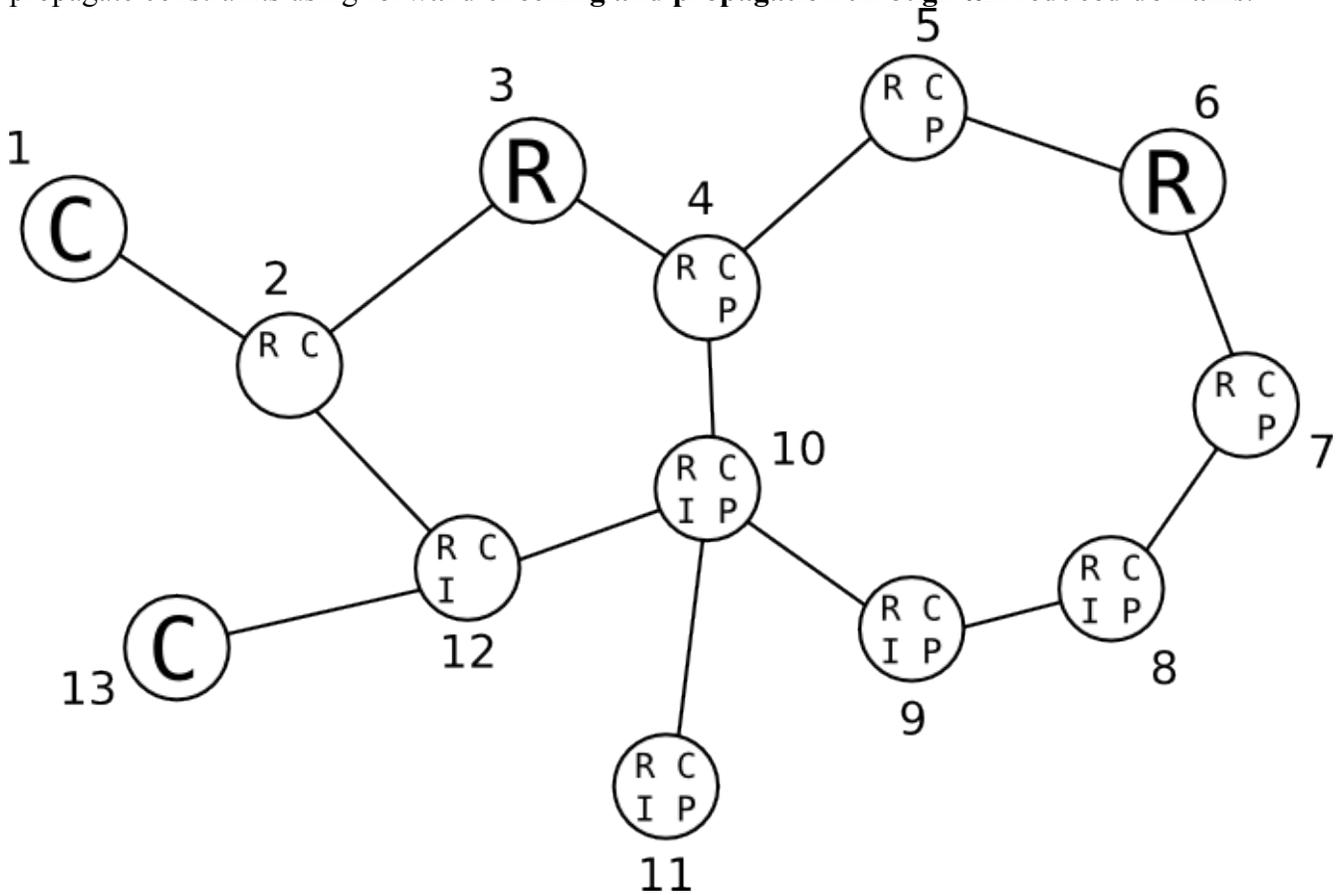
Part B (12 points)

Next, show how the domains are reduced if you designate Dome #10 as Industrial, and then propagate constraints using **forward checking and propagation through singleton domains**.



Part C (12 points)

Finally, show how the domains are reduced if you designate Dome #10 as **Industrial**, and then propagate constraints using **forward checking and propagation through all reduced domains**.



Part D (14 points)

Yuan says that you are right to pick a dome type for Dome #10 first, but Olga advises you to start with Dome #2 instead. What is the reasoning behind each of these suggestions?

Yuan:

Olga:

Quiz 3, Problem 1:

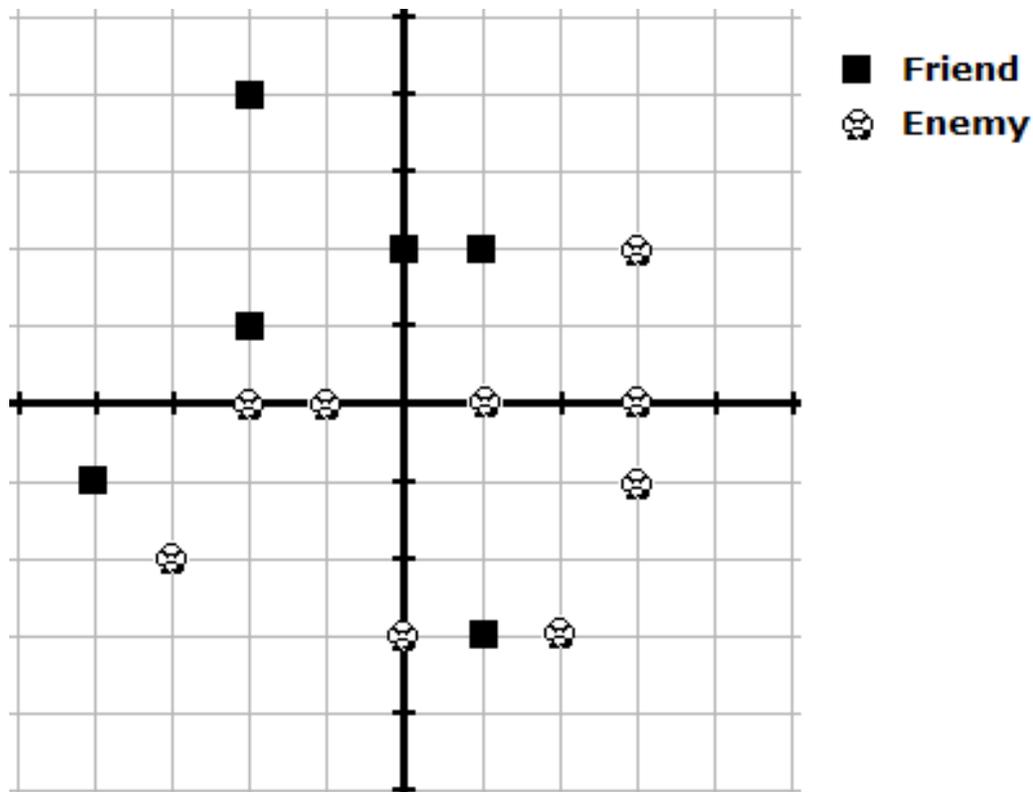
KNN & ID-Trees (50 points)

The Mario Bros. have hired you to help them identify potential dangers during their frequent expeditions to explore strange new worlds, battle evil monsters, and attempt to rescue some princesses.

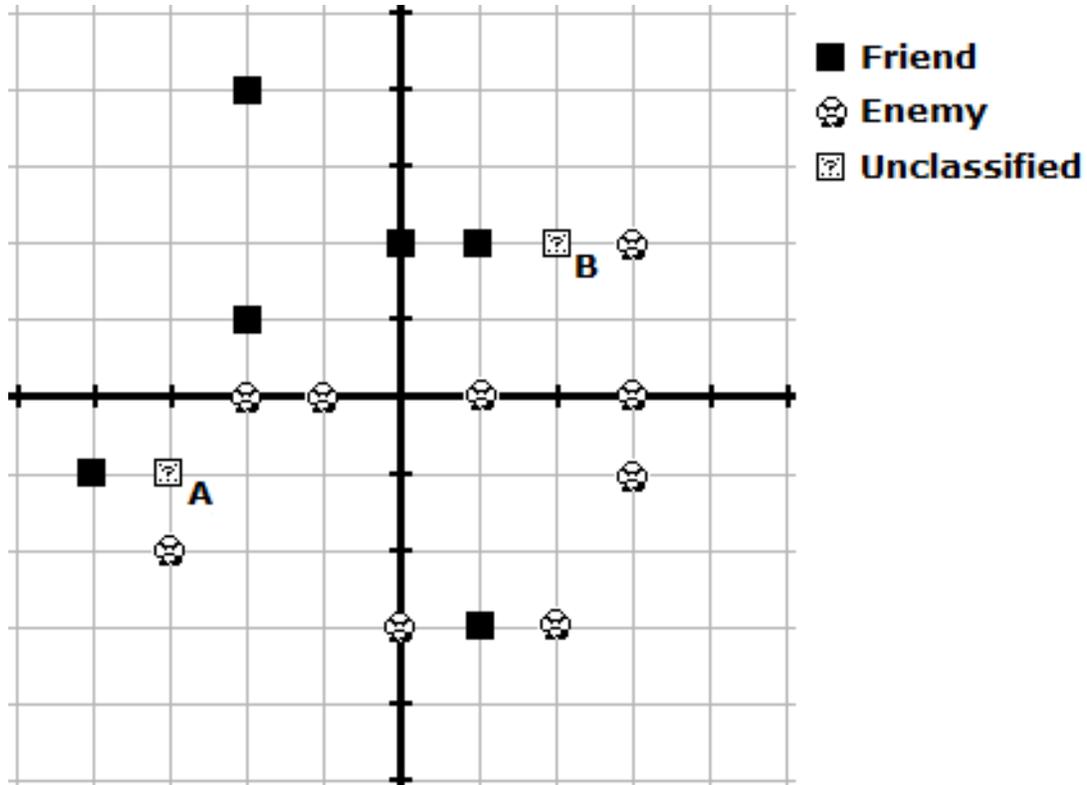
Part A: Nearest Neighbors (20 Points)

You decide to use Nearest Neighbors to classify new creatures that Mario might encounter, based on their positions on a map. In the map, friendly creatures are marked with a square, and enemies are marked with monster icons.

A1 (12 points): On the following graph, draw the decision boundaries produced by 1-nearest-neighbor.



A2 (8 points): The graph below shows two new creatures, marked with a question mark symbols and labeled A and B. Show how these will be classified using 3-nearest-neighbors and 5-nearest-neighbors below.



	Creature A	Creature B
Using 3 NN:		
Using 5 NN:		

Part B: ID-Trees (30 Points)

It turns out that enemies move around the world, so a simple K-Nearest-Neighbors on their location won't do a very good job of protecting Mario. Instead, you decide to use an Identification Tree, based on some characteristics of the creatures in this world, to classify creatures as Enemies (Enemy=Y) and Friends (Enemy=N). The other 5 characteristics you note are in the table below.

	Character	Enemy?	Talks?	Boss?	Annoying?	Killable with Jump?
A	Yoshi	N	N	N	N	N
B	Tree	N	N	N	N	N
C	Luigi	N	N	N	N	N
D	Toad	N	Y	N	N	N
E	Peach	N	Y	N	Y	N
F	Bowser	Y	N	Y	N	N
G	Bob-omb	Y	N	N	N	Y
H	Goomba	Y	N	N	N	Y
I	Piranha	Y	N	N	Y	N
J	Dry Bones	Y	N	N	Y	N
K	Chain Chomp	Y	N	N	Y	N
L	Thwomp	Y	N	N	Y	N
M	Boo	Y	N	N	Y	N
N	Koopa Troopa	Y	N	N	Y	Y

B1 (8 points): What is the disorder of the test "**Boss = Y**"? Leave your answer in terms of fractions, real numbers, and logarithms.

B2 (7 points): What is the disorder of the test "**Annoying = Y**"? Leave your answer in terms of fractions, real numbers, and logarithms.

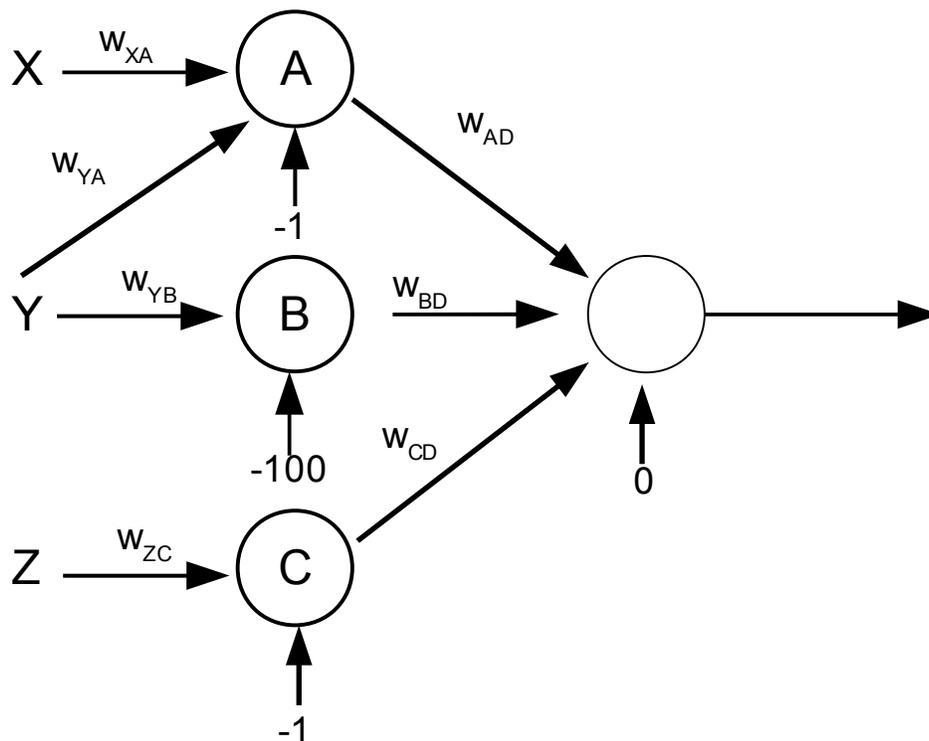
B3 (15 points): Draw the **disorder minimizing** identification tree Mario can use to correctly decide whether any of the above examples is an enemy (Enemy=Y) or friend (Enemy=N). Use the letters provided next to characters' names to show how the characters are separated by each decision.

Hint: neither Boss nor Annoying are the first test.

Quiz 3, Problem 2: Neural Nets (50 points)

Hermione Granger decides to spend a semester abroad at MIT for her Muggle Studies class. Upon arriving, Hermione befriends the coolest people on campus. The 034 staff, over a festive round of Butter Beer, then teach Hermione about neural nets. She decides to practice.

Note that the tear off sheets include the neural net tear-off sheet from Quiz 3.



Hermione creates the above neural net. Each node uses a standard sigmoid function. All threshold units have fixed weights of 1. **They are never updated.**

Part A: (12 Points)

For each node (A, B, C, and D), using only the weight variables provided in the diagram, d (desired output), and the outputs of each node (O_A , O_B , O_C , and O_D), determine δ_A , δ_B , δ_C , and δ_D .

Part B: (12 Points)

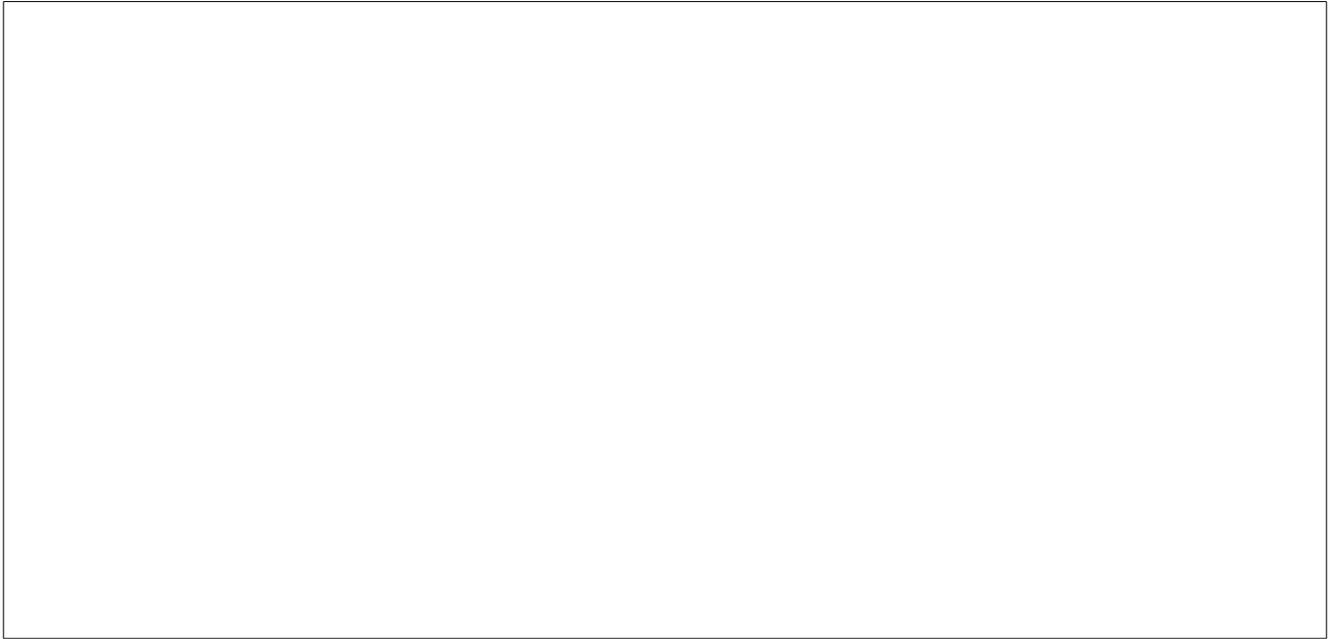
Hermione, as an intellectual exercise, runs the neural net with the following parameters:

$X=0.5$	$Y = 0$	$Z = 5$	$W_{XA} = 2$
$W_{YA} = 2$	$W_{YB} = 1$	$W_{ZC} = 0.2$	$W_{AD} = 1$
$W_{BD} = 1.5$	$W_{CD} = -1$		

Determine the output for each node (O_A , O_B , O_C , and O_D). Approximate sigmoid(x) very simply by letting it evaluate to 0 for $x < -50$ and 1 for $x > 50$.

Part C: (18 Points)

Run back-propagation using these same parameter values for one iteration to determine new weight values for W_{AD} , W_{BD} , and W_{CD} . Use your results from Part A and Part B, and assume a learning rate of $r = 2$ and a desired output, $d=1$.



Part D: (8 Points)

Hermione changes **all the decision functions** in the **original network, before training**, from using a sigmoid function to using a **threshold function** (which outputs 1 if the input is greater than 0, and outputs 0 otherwise).

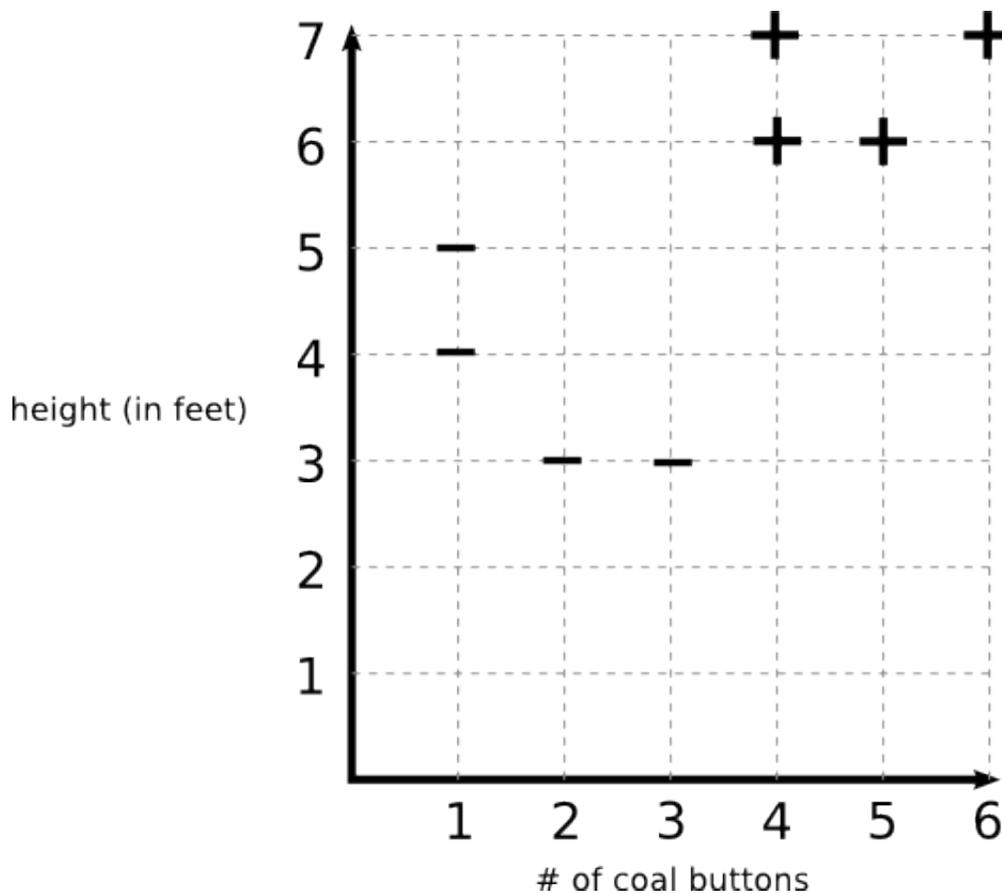
Using X, Y, and Z and the standard logical operators (AND, NOT, and OR), determine a logic expression for the neural net. Assume that the inputs are always 0 (meaning false) or 1 (meaning true).



Quiz 4, Problem 1: SVMs (50 points)

Part A: F.R.O.S.T. and the Snowmen (25 Points)

You have been hired by the Foundation for Research into Occult Snow Transformations (F.R.O.S.T.) to investigate a strange new phenomenon. It seems that certain snowmen can be turned into living, talking creatures when an old silk hat is placed on their heads. So far, F.R.O.S.T. researchers have determined that the factors that best predict whether a particular snowman can be thus transformed are the snowman's height and the number of coal buttons on its chest. Their preliminary data is below: each “+” represents a magical talking snowman, and each “-” represents an ordinary, non-transforming snowman.



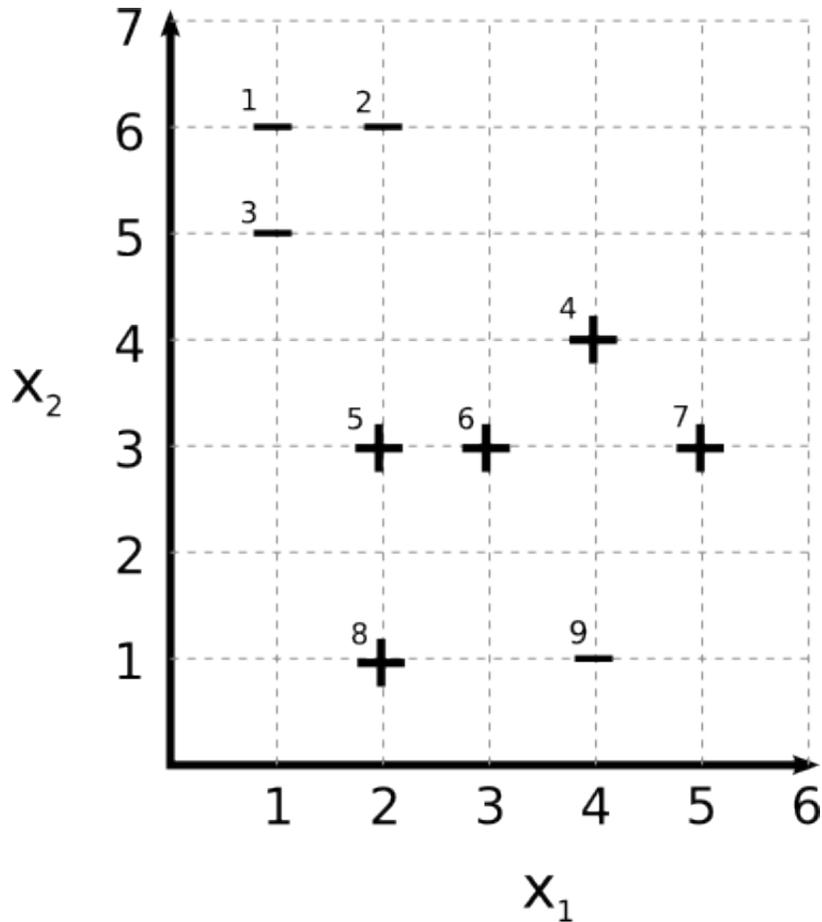
A1 (5 points): You decide to use a linear SVM to classify magical vs. non-magical snowmen. **Draw the resulting decision boundary in the graph above, and circle the support vectors.**

A2 (10 points): In this SVM, what is the vector \mathbf{w} and the constant b ?

A3 (10 points): Suppose the F.R.O.S.T. scientists discover another magical, talking snowman that is five feet tall and has six coal buttons on his chest. What would the α value of this new data point be? Justify your answer.

Part B: Kernels (25 Points)

In this section, you will project the data below into a new space with $\phi(\mathbf{u}) = \langle |x_1 - x_2| \rangle$. That is, you project the two-dimensional vector \mathbf{u} into a one-dimensional vector in a one-dimensional space.



B1 (7 points): What is the kernel function $K(u,v)$ for this transformation?

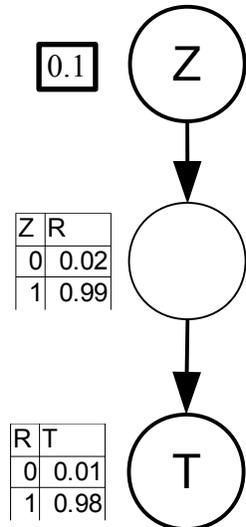
B3 (10 points): In the transformed space, what is the vector w and the constant b ?



B2 (8 points): What is the final classifier produced. Express your answer in the original space, not the transformed space.



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



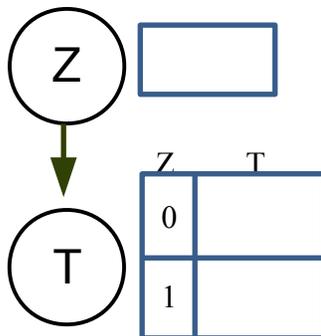
Zombies once again threaten the town, but this time the healthy folks are using science to help keep themselves on a brain-free diet!

When someone first becomes a zombie, their zombieness is latent for a while, as the transformation slowly progresses, but successful isolation at this stage is critical for avoiding new zombie infections.

New zombies quickly begin to produce a protein R that one can easily test for. One percent of infected individuals do not produce protein R (or perhaps just not yet), while protein R has been found in two percent of subjects who turned out not to have turned into zombies later.

The test is 98% sensitive to the presence of the protein, and 99% specific to it, as shown in the Bayes Net on the left.

A (9 points) To make things a little simpler for the general population to understand, you want to give information in terms of how sensitive and specific the test is to zombihood, not to the protein:



Show your work for partial credit:

B (6 points) Of 100 townfolk, how many will test positive:

C (6 points) ...and how many healthy ones will test positive:

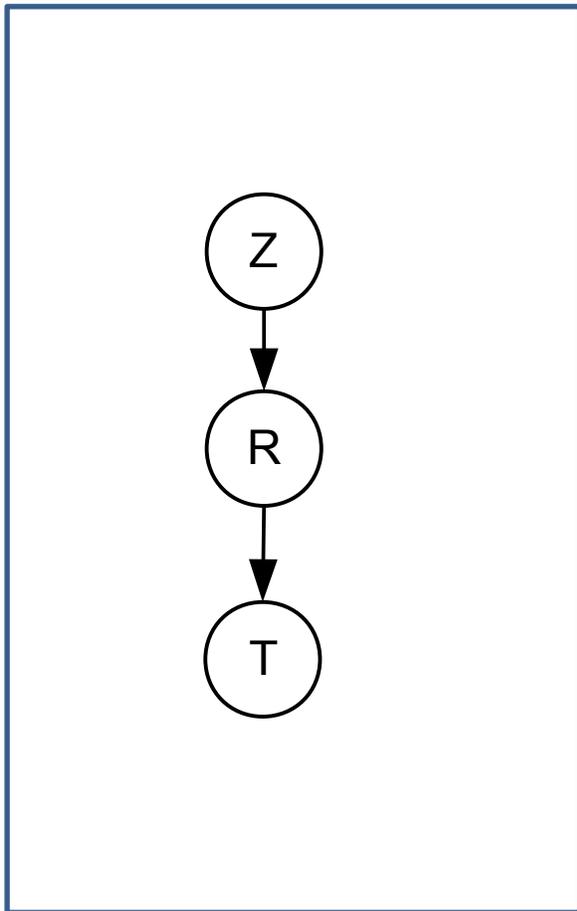
D (6 points) What is the overall accuracy of the test? (the probability that it predicts correctly)

E (5 points) A visitor, Ulric, was tested upon entering the town, and had a positive result. Ulric insists that he is healthy, and hasn't come anywhere near any zombies. To be sure, the guards decide to test him again. What assumption(s) are the guards making about the disease and the testing process?

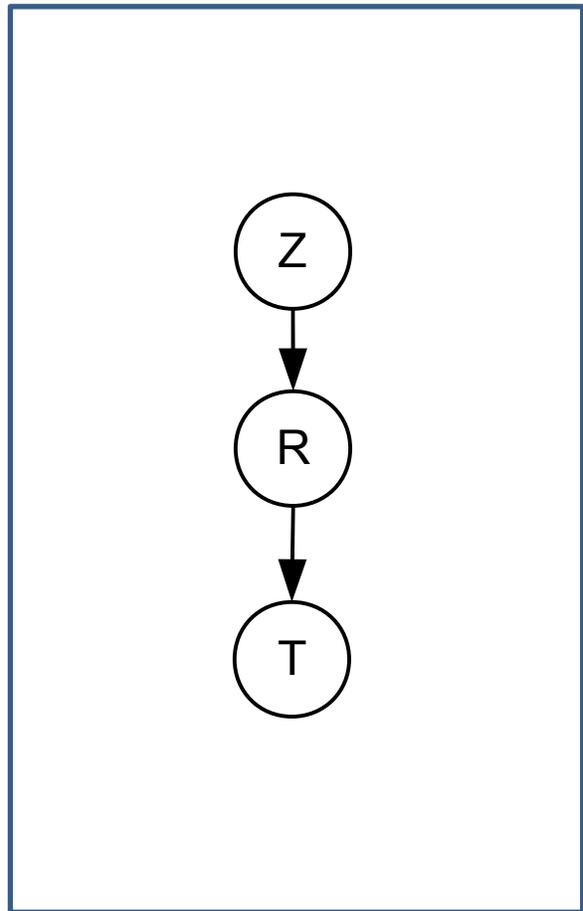
F (5 points) Victor and Wendy both arrive from Brainton, where Ulric came from too, and they both test positive, though they appear quite healthy. Ulric is doing fine in quarantine. The town scientists begin to consider two alternate hypotheses. Both of the alternate hypotheses allow being from Brainton (indicated with B) to explain away a positive test result. Both of the models extend the original three-node model with just the one extra node B. Draw the two possible models they should consider.

Draw only the new node B and its edges — you do not need conditional probability tables (yet):

Alternate Model A



Alternate Model B



Despite **two positive test results each**, Ulric, Victor, and Wendy all turn out to be healthy. **An expedition to (apparently zombie-free) Brainton revealed that 80% of its population test positive using your town's test.** One of the alternate models is very likely!

G (6 points)

Assume that you think a model whose Bayes Net representation has $k+1$ parameters is about **half** as likely *a priori* as a model with k parameters, unless $k = 0$, in which case assume that the model probability is some constant, c .

How likely is the original (three variable) model, *a priori*?

How likely are each of the models you drew, *a priori*?

Alternate Model A

Alternate Model B

H (7 points)

To rule out your original model entirely, you decide that an alternate model has to explain the data not just better, but **at least 100 times better** than the original model does.

Was the expedition necessary to rule out the original model, or did we know enough once we knew the three visitors were actually healthy, despite the tests?

Circle one:

Needed the Expedition

Ulric, Victor, and Wendy were enough

Explain, stating any assumptions you need to make:

Quiz 5, Problem 1, (40 points)

You have made it. You have graduated, started a thriving company, found a spouse, got a home in the suburbs, children. Now it is time to get a dog. Knowing nothing about dogs yourself, you ask a dog-loving, but inarticulate friend to characterize a series of dogs as Good or Bad. You hope to learn from your friend's characterizations.

Your first task is to pick some descriptors. You decide to look at intelligence, breed, exceptional characteristics, and gender. Male and female form a set. Yes and no form a set. Breeds are the leaves of a tree and form groups specified by the American Kennel Club: Bouvier and Collie are part of the Herding group, Beagle and Dachshund are part of the Hound group, and Chihuahua is part of the Toy group. Herding, Hound, Toy, Working, and Sporting all belong to the Recognized Dog category.

Using Arch learning, indicate in the table what is learned from each example and identify the heuristic involved by name, if known. If nothing is learned, put an x in the corresponding 2 columns.

Part A (32 points)

Candidate	Gender	Breed	Exceptional Quality	Intelligent	Heuristic	What is learned
Good	Male	Bouvier		Yes		
Bad	Female	Bouvier		Yes		
Bad	Male	Bouvier	Nasty	Yes		
Good	Male	Collie		Yes		
Good	Male	Collie		No		
Good	Male	Beagle		No		
Good	Male	Dachshund		No		
Bad	Female	Chihuahua		No		
Bad	Male	Chihuahua	Nasty	Yes		

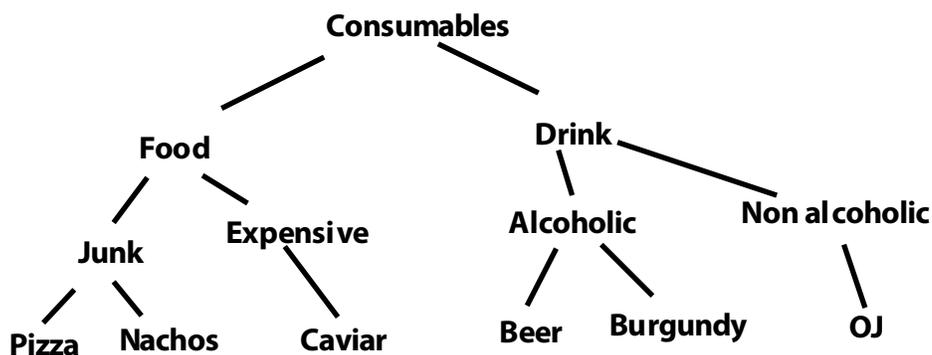
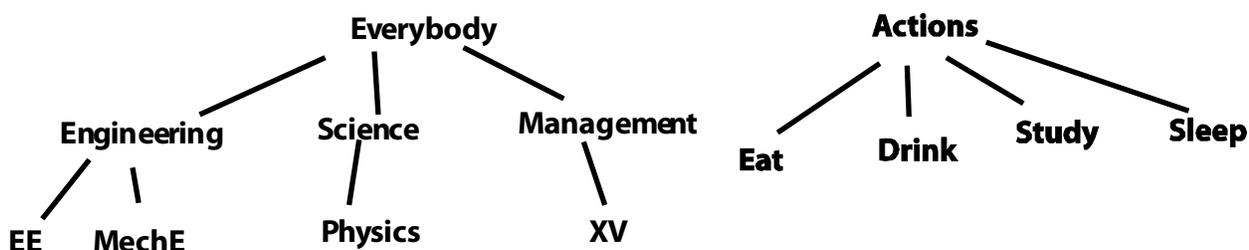
Part B (8 points)

Exhibit an example, which if used as the second example, would teach two characteristics at once.

Candidate	Gender	Breed	Exceptional quality	Intelligent

Quiz 5, Problem 2, (21 points)

President Hockfield, eager to understand student dining preferences, decides to collect information in a self-organizing map. She starts by constructing some classification trees on a paper napkin. As you can see, she views students from the perspective of departments, such as EE, MechE, Physics, and XV:



Next, she develops an initial map with three relations:

Map element	Subject	Relation	Object
1	EE	Eats	Pizza
2	XV	Eats	Caviar
3	MechE	Drinks	Alcoholic beverage

Note that in this map, there is no concept of neighboring cell. She has decided to put that off for a while. Accordingly, each adjustment to the map alters only that cell judged closest to the new relation flowing into the map.

Part A

Now, President Hockfield observes the following:

Subject	Relation	Object
MechE	Eats	Nachos

Not fully remembering Professor Winston's lecture, she asks you to adjust the map to accommodate this information using the classification trees and edit distance (number of up and down steps in the classification tree needed to go from one combination to the other) to determine which element to change. You don't quite remember how to change the selected element, but fortunately, you took a photograph of Professor Winston's practice board:

```
If    sample object classification extends map object classification
Then  extend map classification object by one class
Else if sample object classification and map object classification are the same
Then  do nothing
Else  trim map object classification by one class
```

To save time, do not bother copying in any information that remains unchanged.

Map element	Subject	Relation	Object
1			
2			
3			

Part B

Repeat with the following observation. Note that any change you made in Part A is to be carried into this part:

Subject	Relation	Object
EE	Drinks	Beer

Again, you need only enter the changed row. You need not write in any row that remains the same after the work you did in Part A.

Map element	Subject	Relation	Object
1			
2			
3			

Part C

Repeat with the following observation. Note that any changes you made in previous parts are to be carried into this part:

Subject	Relation	Object
Management	Drinks	Burgundy

Again, you need only enter the changed row. You need not write in any row that remains the same after the work you did in previous parts.

Map element	Subject	Relation	Object
1			
2			
3			

Quiz 5, Problem 3, (15 points)

A: A Japanese Haiku is a poem that consists of 17 syllables arranged in 3 lines of 5, 7, and 5 syllables. These are very popular, so a publishing house has asked you to develop a system that can generate good ones automatically.

The syllables are drawn from the approximately 50 syllables that are found in Japanese.

James, who dropped 6.034 early in the term suggests that you just generate all possible Haikus and hand them to a panel of judges who will pick the best.

Part A (12 points)

Determine how long it would take to calculate all possible Haikus given a computer that produces one Haiku per nanosecond. So you won't think you need a calculator, you may make the following approximations:

$50 = 32$, $17 = 16$, and $1024 = 1000$, and $\text{seconds/year} = 10^7$

Part B (3 points)

Is James suggestion practical? Circle your answer

Yes No

Quiz 5, Problem 4, (24 points)

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

Genetic algorithms, without crossover, is best described as a kind of

1. Instance of General Problem Solver architecture
2. Instance of subsumption architecture
3. Instance of SOAR architecture
4. Hill climbing search
5. None of the above

Crossover is best described as

1. A product of means-ends analysis
2. A product of an abstraction barrier
3. A label for the strange mating behavior of Zebra Finches
4. A means to escape local maxima in a search space
5. None of the above

The SOAR architecture is best described as

1. A programming language
2. A descendant of the General Problem Solver architecture
3. An amalgam of several ideas
4. The design philosophy that led to the Stata center
5. None of the above

The Genesis architecture (Winston's research focus) is best described as

1. A search for a universal representation
2. A commitment to reasoning as the distinguishing feature of human intelligence
3. A search for principles that explain the intelligence of non human primates
4. A demonstration that natural language has little or no role in explaining human intelligence
5. None of the above

Minsky's Emotion Machine/Society of Mind architecture is best described as

1. Focused on explaining visual problem solving
2. Focused on reasoning on multiple levels
3. A commitment to the idea that language is the differentiating feature of human intelligence
4. A commitment to the idea that culture is reflected in the myths associated with the culture
5. None of the above

Intermediate features and the Goldilocks principle is best explained as

1. The observation that eyes are too small and faces are too large
2. The observation that trajectories are too small and transitions are too large
3. The observation that cultures are defined by their fairy tales
4. The observation that the intelligence of a songbird lies half way between insects and humans
5. None of the above

Tear off sheets—you need not hand in the tear off sheets

Q1P1

RULES:

R0 : IF (?X) goes to MIT,
 THEN (?X) is a muggle,
 (?X) consumed butterbeer

R1: IF (?X) made math jokes AND
 (?X) consumed butterbeer
 THEN (?X) was transfigured into a porcupine

R2: IF (?Y) fancies (?X) AND
 (?X) fancies (?Y) AND
 (?Y) is a muggle
 THEN (?X) snogged (?Y)

R3: IF (?X) fancies (?Y) AND
 (?X) made math jokes,
 THEN (?Y) fancies (?X)

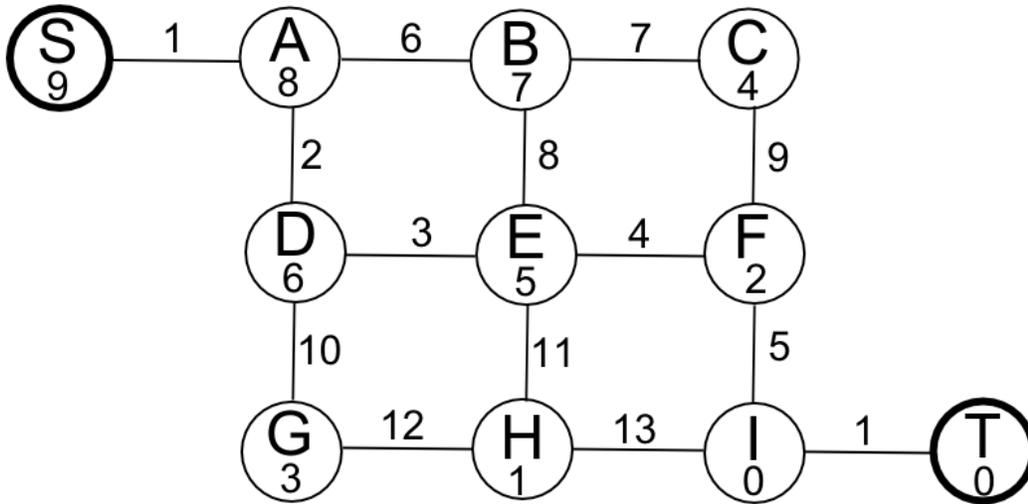
R4: IF (?X) made math jokes
 THEN (?X) goes to MIT

You start with the following list of assertions which **is all you have to go on.**

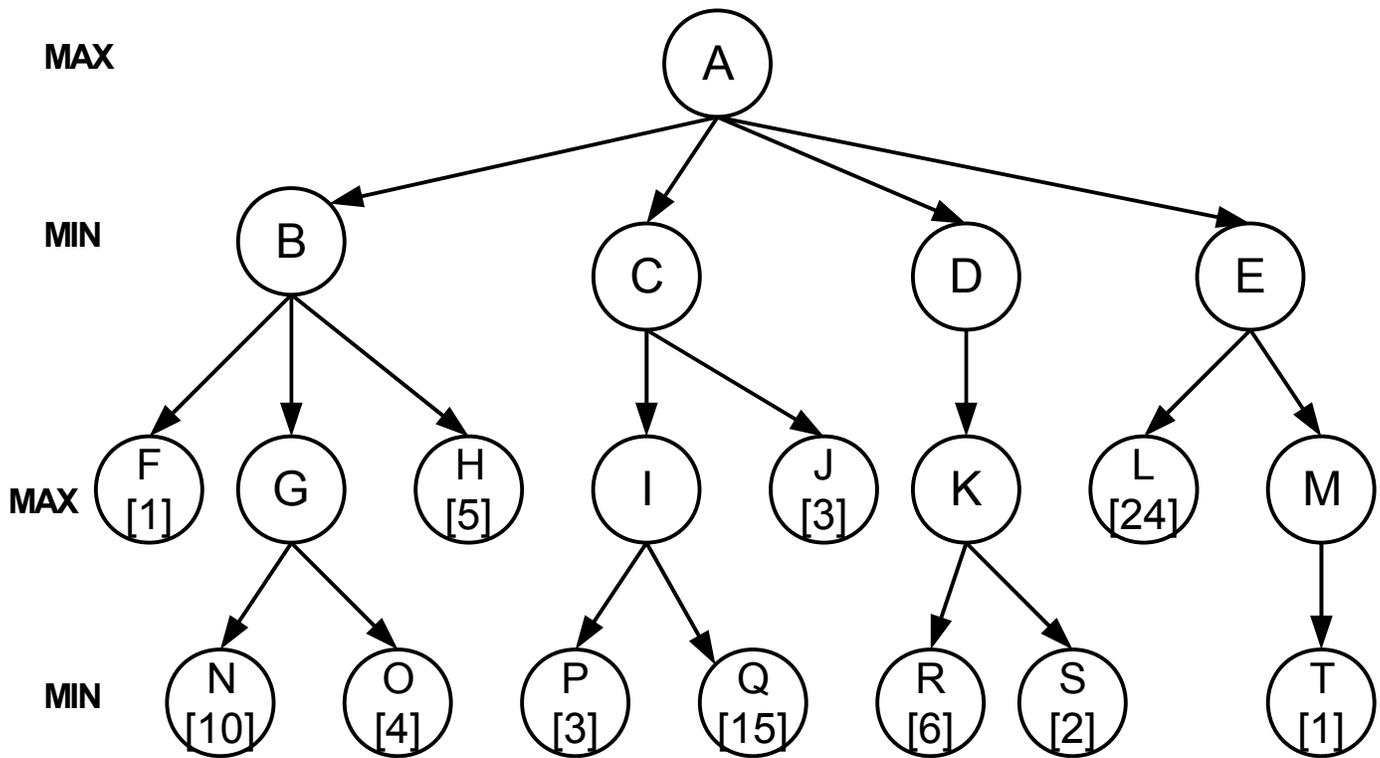
ASSERTIONS:

- A0: Olga made math jokes
- A1: Yuan goes to MIT
- A2: Jeremy made math jokes
- A3: Hermione consumed butterbeer
- A4: Jeremy fancies Hermione

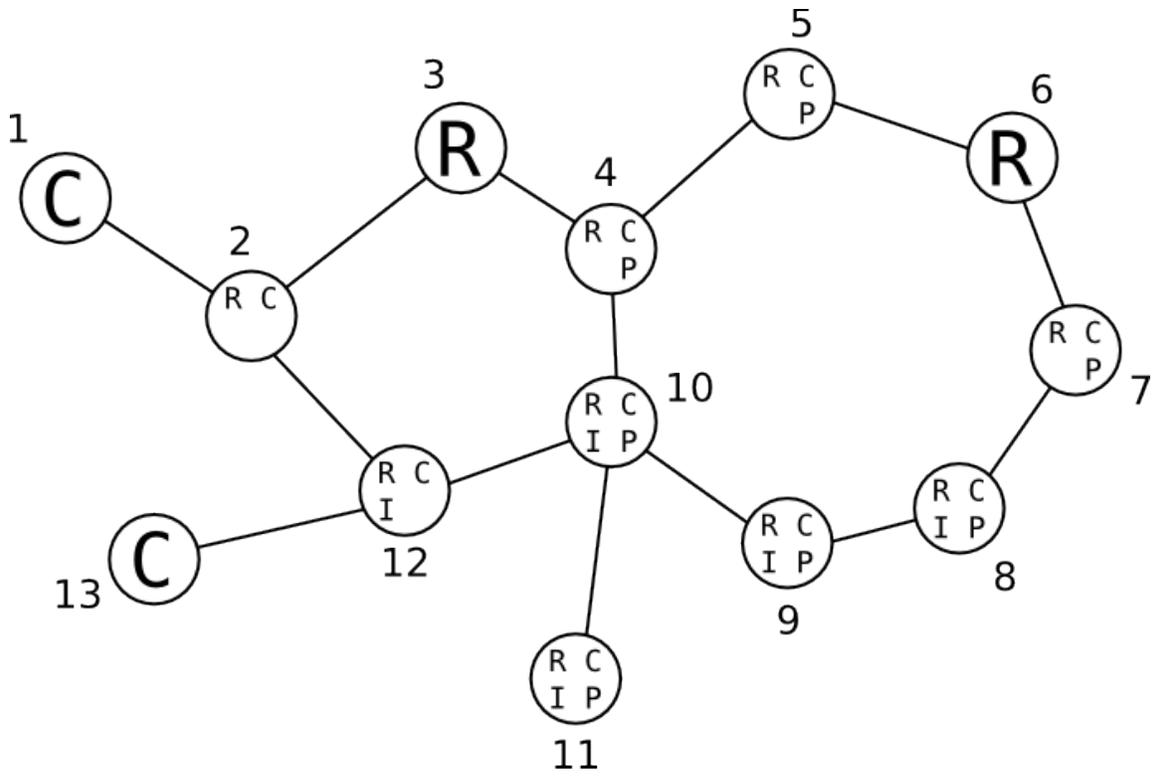
Q1P2



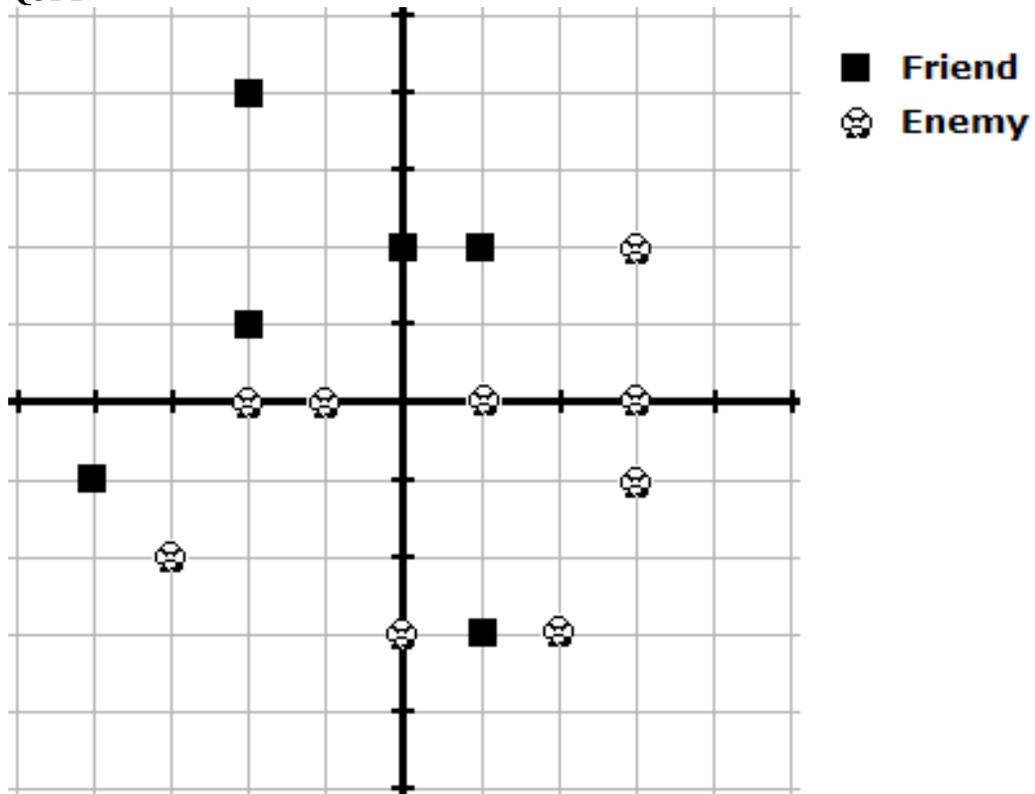
Q2P1



Q2P2



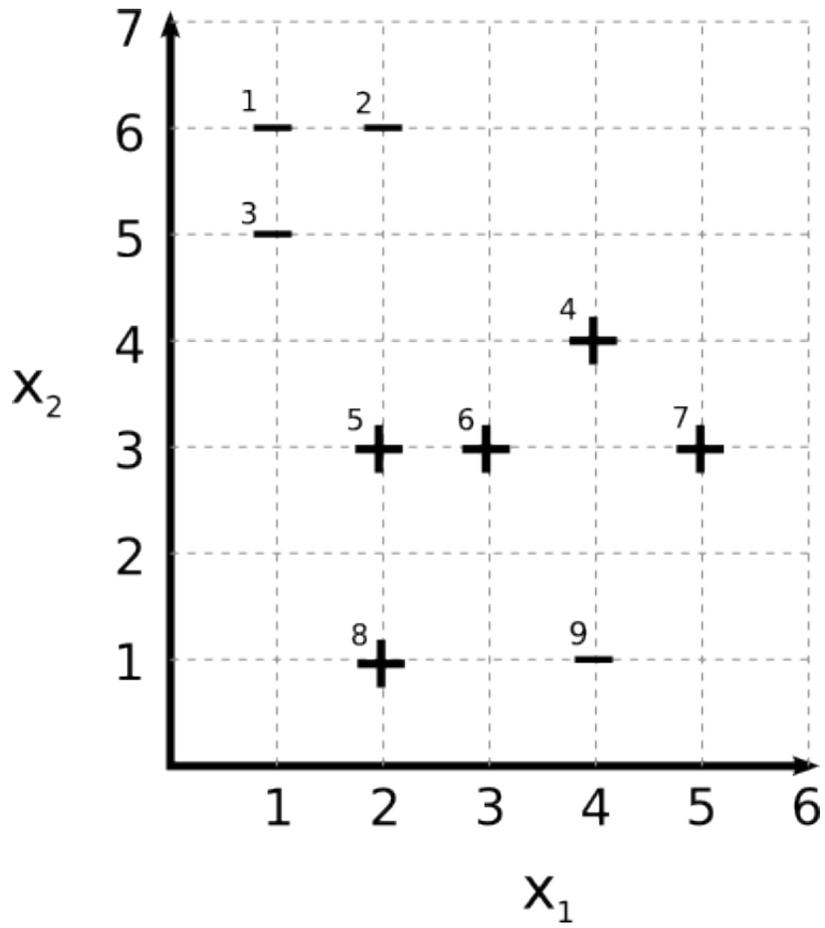
Q3P1



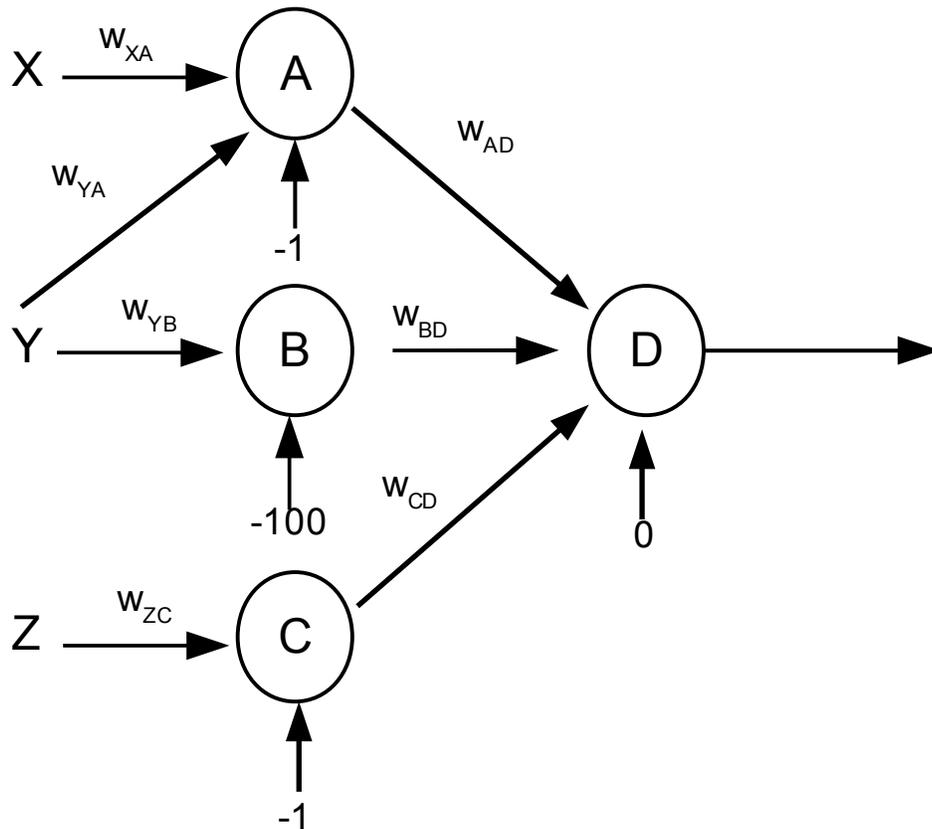
Q3P1

	Character	Enemy?	Talks?	Boss?	Annoying?	Killable with Jump?
A	Yoshi	N	N	N	N	N
B	Tree	N	N	N	N	N
C	Luigi	N	N	N	N	N
D	Toad	N	Y	N	N	N
E	Peach	N	Y	N	Y	N
F	Bowser	Y	N	Y	N	N
G	Bob-omb	Y	N	N	N	Y
H	Goomba	Y	N	N	N	Y
I	Piranha	Y	N	N	Y	N
J	Dry Bones	Y	N	N	Y	N
K	Chain Chomp	Y	N	N	Y	N
L	Thwomp	Y	N	N	Y	N
M	Boo	Y	N	N	Y	N
N	Koopa Troopa	Y	N	N	Y	Y

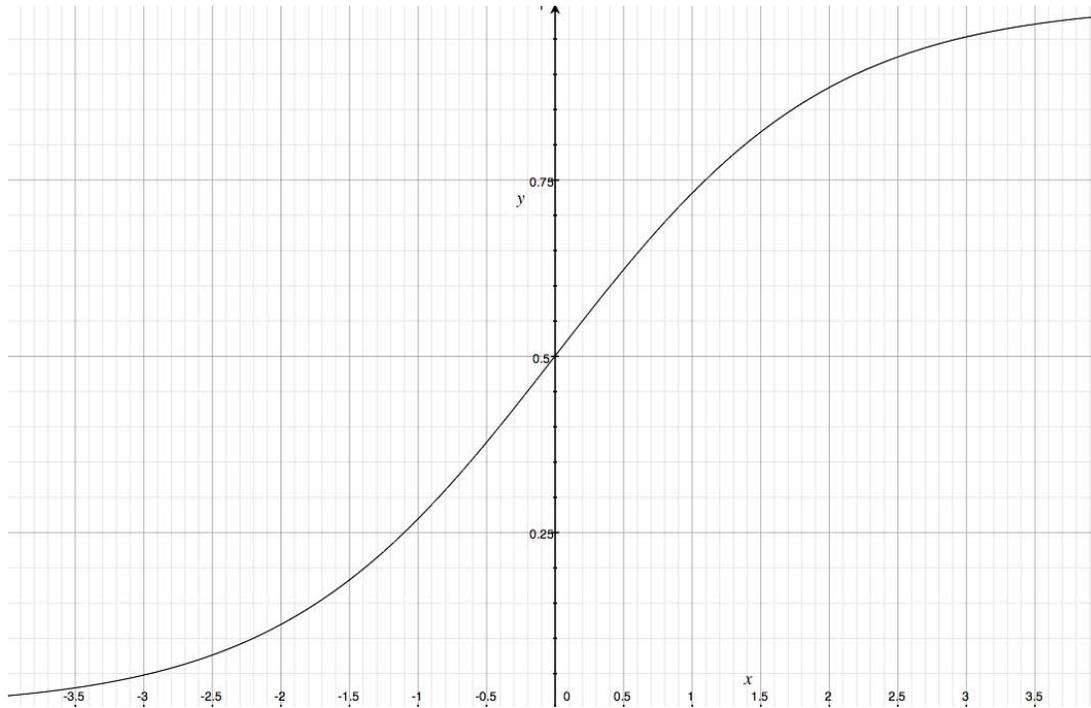
Q4P1



Q3P2



Q3P2



Neural net update:

$$E = \frac{1}{2} \sum_k (o_k - d_k)^2$$
$$w_{i \rightarrow j} = w_{i \rightarrow j} - \Delta w_{i \rightarrow j}$$
$$\Delta w_{i \rightarrow j} = R \times o_i \times \delta_j$$

where R is a rate constant and the δ s are computed with the following formulas:

$$\delta_k = o_k(1 - o_k) \times (o_k - d_k)$$
$$\delta_l = o_l(1 - o_l) \times \sum_j w_{l \rightarrow j} \times \delta_j$$

where

- o_k is output k of the output layer
- d_k is the desired output k of the output layer
- δ_k is a delta associated with the output layer
- o_l is output l of left layer in a left-right pair
- δ_l is a delta associated with the layer l
- δ_r is a delta associated with the adjacent layer to the right, layer r

MIT OpenCourseWare
<http://ocw.mit.edu>

6.034 Artificial Intelligence
Fall 2010

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.