6.034 Quiz 2

21 October 2015

Name	Neville Longbottom
Email	neville@hogwarts.edu

Circle your TA **(for 1 extra credit point)**, so that we can more easily enter your score in our records and return your quiz to you promptly.

Ryan Alexander	**Nick Flamel**	**Ben Greenberg**
Neil Gurram	**Eeway Hsu**	**Brittney Johnson**
Veronica Lane	**Robert Luo**	**Jessica Noss**
Mycal Tucker	**Sarah Vente**	**Jess Wass**

Problem number	Maximum	Score	Grader
1	50		
2	50		
Total	100		

SRN	6		

Note: We only expected students to write the parts of the solutions that are written in black, or highlighted in green.

There are 14 pages in this quiz, including this one, but not including tear-off sheets. Tear-off sheets with duplicate drawings and data are located after the final page of the quiz. As always, open book, open notes, open just about everything, including a calculator, but no computers.

Problem 1: A Dementor on the Train (50 points)

Three students (**R**on, **H**ermione, **M**alfoy) are looking for train compartments to sit in on the Hogwarts Express train. There are four compartments (1, 2, 3, 4), and each compartment can hold any number of students.

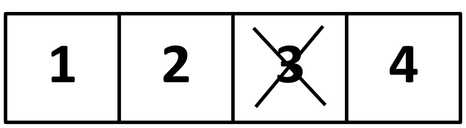
The students have following constraints:

1. **M**alfoy despises **R**on and **H**ermione, so **M**alfoy must not sit in the same compartment as either of them.
2. **H**ermione wants to keep an eye on **M**alfoy, so their compartments must be adjacent (and not the same).
3. **R**on has a crush on **H**ermione, but he also thinks she's an obnoxious know-it-all, so their compartments must be adjacent (and not the same).

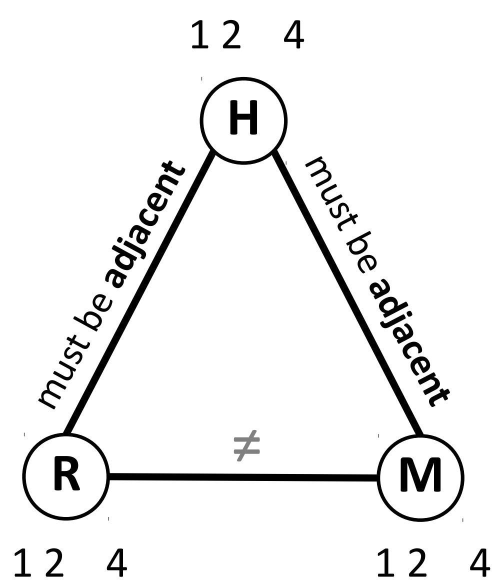
Alas, **a dementor has just arrived in compartment 3**, so no one should sit there. Thus, the initial domains are:

R	1 2 4
M	1 2 4
H	1 2 4

The train compartments and a graph of the constraints are pictured below:



This is the layout of the train compartments. Out of the three remaining compartments, 1 and 2 are adjacent to each other. 4 is not adjacent to any usable compartment.



Note: For your convenience, additional copies of the constraint graph and initial domains table are provided on a tear-off sheet after the last page of the quiz.

Part A: Counting Solutions (6 points)

A1 (3 points) Using the initial domains specified on the previous page, but ignoring the constraints, how many possible assignments of values to variables are there? (Remember that compartment 3 will never be used.)

YOUR ANSWER: $\boxed{3^3=27}$

size of each domain multiplied

A2 (3 points) How many of those possible assignments satisfy all the constraints – that is, how many different solutions does this problem have? (You may use your intuition, or figure out the answer using the later parts of this problem.)

YOUR ANSWER: $\boxed{0}$

there is no solution where both Ron & Malfoy are adjacent to Hermione

Part B: DFS + Forward Checking (23 points)

B1 (20 points) Use **depth-first search with forward checking** and **NO propagation** to assign the students to train compartments in this order: **R**on, **M**alfoy, **H**ermione.

Start with the initial domains shown in the table, and **DO NOT** perform any additional domain reduction before search.

R	1	2		4
M	1	2		4
H	1	2		4

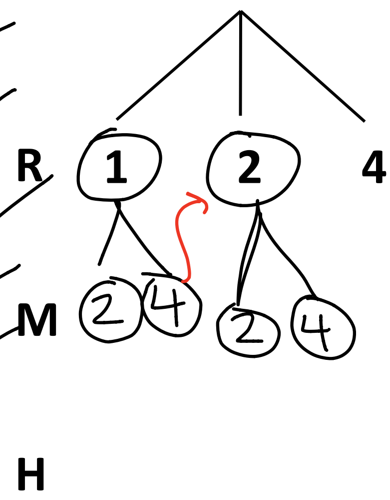
★ On the next page, show your work by simultaneously
(1) filling out the domain worksheet and
(2) drawing the search tree (optional, for partial credit).

Note: Detailed instructions for filling out the domain worksheet can be found on a tear-off sheet after the last page of the quiz.

(Part B1: Depth-first search with forward checking only)

	Var assigned or de-queued	List all values just ELIMINATED from neighboring variables	Back track
1	R = 1	$H \neq 1, 4$, $M \neq 1$	☐
2	M=2	$H \neq 2$	☑
3	M=4	$H \neq 2$	☑
4	R=2	$H \neq 2, 4$, $M \neq 2$	☐
5	M=1	$H \neq 1$	☑
6	M=4	$H \neq 1$	☑
7	R=4	$H \neq 1, 2, 4$	☑
8			☐
9			☐

Search Tree



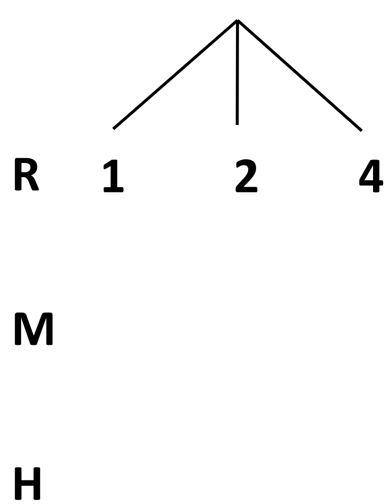
Never cross out values from variables that have already been assigned. Don't cross out the same value again within the same branch.

Note: For your convenience, below is a **duplicate** copy of the domain worksheet and tree. If you write on both copies, please indicate clearly which one you want us to grade.

☐ I want to start over; grade the version below.

	Var assigned or de-queued	List all values just ELIMINATED from neighboring variables	Back track
1	R = 1		☐
2			☐
3			☐
4			☐
5			☐
6			☐
7			☐
8			☐
9			☐

Search Tree



B2 (3 points) Based on the results of your search in Part B1, write the initials of the students (**R**, **M**, **H**) in their assigned train compartments:

1	2	3	4

(If no solution exists, circle NO SOLUTION instead.)

NO SOLUTION

Part C: Propagation through Reduced Domains (21 points)

Repeat the search, following the same instructions as in Part B, but this time using **depth-first search with forward checking and propagation through domains reduced by any number of values**.

Start with the initial domains shown in the table, and **DO NOT** perform any additional domain reduction before search.

R	1	2		4
M	1	2		4
H	1	2		4

★ On the next page, show your work by simultaneously

(1) filling out the domain worksheet and

(2) drawing the search tree (optional, for partial credit).

We accepted both of the solutions below:

(Part C: Depth-first search with forward checking and propagation through reduced domains)

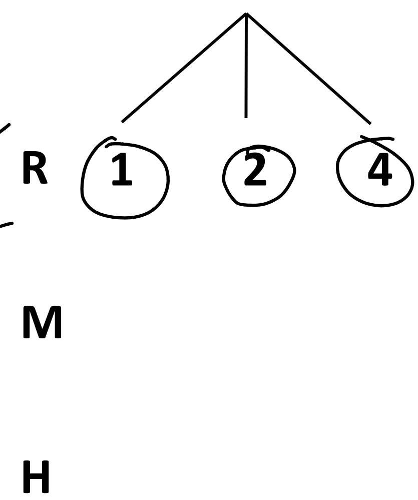
(propagated in alphabetical order, as instructed in the tear off)

propagation never assigns values

never cross out values from the variable that is being propagated

	Var assigned or de-queued	List all values just ELIMINATED from neighboring variables	Back track
1	R = 1	H ≠ 1, 4, M ≠ 1	☐
2	H	M ≠ 2	☑
3	R=2	H ≠ 2, 4, M ≠ 2	☐
4	H	M ≠ 1	☑
5	R=4	H ≠ 1, 2, 4, M ≠ 4	☑
6			☐
7			☐
8			☐
9			☐

Search Tree



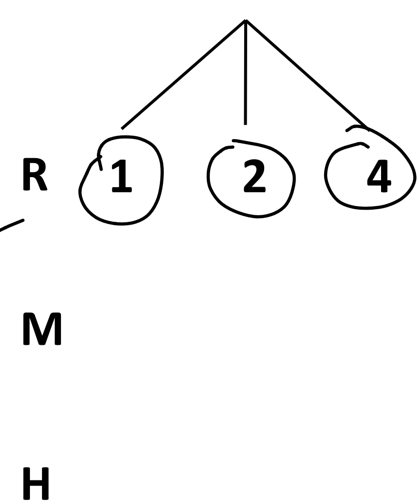
Note: For your convenience, below is a **duplicate** copy of the domain worksheet and tree. If you write on both copies, please indicate clearly which one you want us to grade.

☐ I want to start over; grade the version below.

(Propagated in the order shown on the right)

	Var assigned or de-queued	List all values just ELIMINATED from neighboring variables	Back track
1	R = 1	M ≠ 1, H ≠ 1, 4	☐
2	M	H ≠ 2	☑
3	R=2	M ≠ 2, H ≠ 2, 4	☐
4	M	H ≠ 1	☑
5	R=4	M ≠ 4, H ≠ 1, 2, 4	☑
6			☐
7			☐
8			☐
9			☐

Search Tree



Problem 2: ID Trees & k-Nearest Neighbors (50 points)

The Force has awakened! President Leia Organa Solo has appointed you to investigate some mysterious disturbances in the Force.

Part A: Jedi or Sith? (23 points)

Your first task is to classify newly identified Force users as either Jedi or Sith. You have information about 7 presumed Force users, and whether each one was trained at the Jedi **Academy**, can use **Lightning**, ever **Falls in Love**, and uses a **Red Lightsaber**. You decide to use this information (shown in the table below) to create an identification tree.

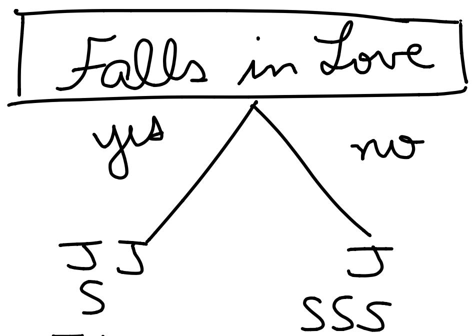
Force User	Classi-fication	Academy	Falls in Love	Lightning	Red Lightsaber
1 Obi-Wan Kenobi	Jedi	Yes	Yes	No	No
2 Qui-Gon Jinn	Jedi	Yes	No	No	No
3 Luke Skywalker	Jedi	No	Yes	No	No
4 Darth Vader	Sith	Yes	Yes	No	Yes
5 Count Dooku	Sith	Yes	No	Yes	Yes
6 Emperor Palpatine	Sith	No	No	Yes	Yes
7 General Grievous	Sith	No	No	No	No

Note: For your convenience, a copy of this data is provided on a tear-off sheet after the last page of the quiz.

A1 (6 points) Compute the disorder of each of the following feature tests. Use the table of logarithms below to express your answer **as sums and products of decimals and fractions only**. Your final answer should have **<u>no logarithms</u>** in it. Space is provided below to show your work.

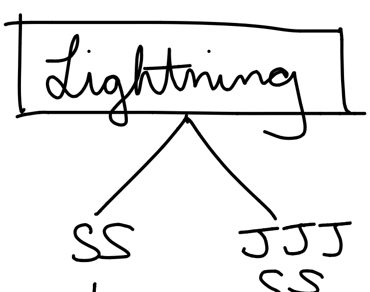
Feature Test	Disorder
Falls in Love	$\frac{3}{7}(.9)+\frac{4}{7}(.8)= 0.84$
Lightning	$\frac{5}{7}(.97)=0.69$

Show your disorder calculations for partial credit.



$$D=\frac{3}{7}\left[\left(\frac{2}{3}\log_2\frac{2}{3}+\frac{1}{3}\log_2\frac{1}{3}\right)\right]+\frac{4}{7}\left[-\left(\frac{3}{4}\log_2\frac{3}{4}+\frac{1}{4}\log_2\frac{1}{4}\right)\right]$$

$$=\frac{3}{7}(.9)+\frac{4}{7}(.8)$$



$$D=\varnothing+\frac{5}{7}\left[-\left(\frac{2}{5}\log_2\frac{2}{5}+\frac{3}{5}\log_2\frac{3}{5}\right)\right]$$

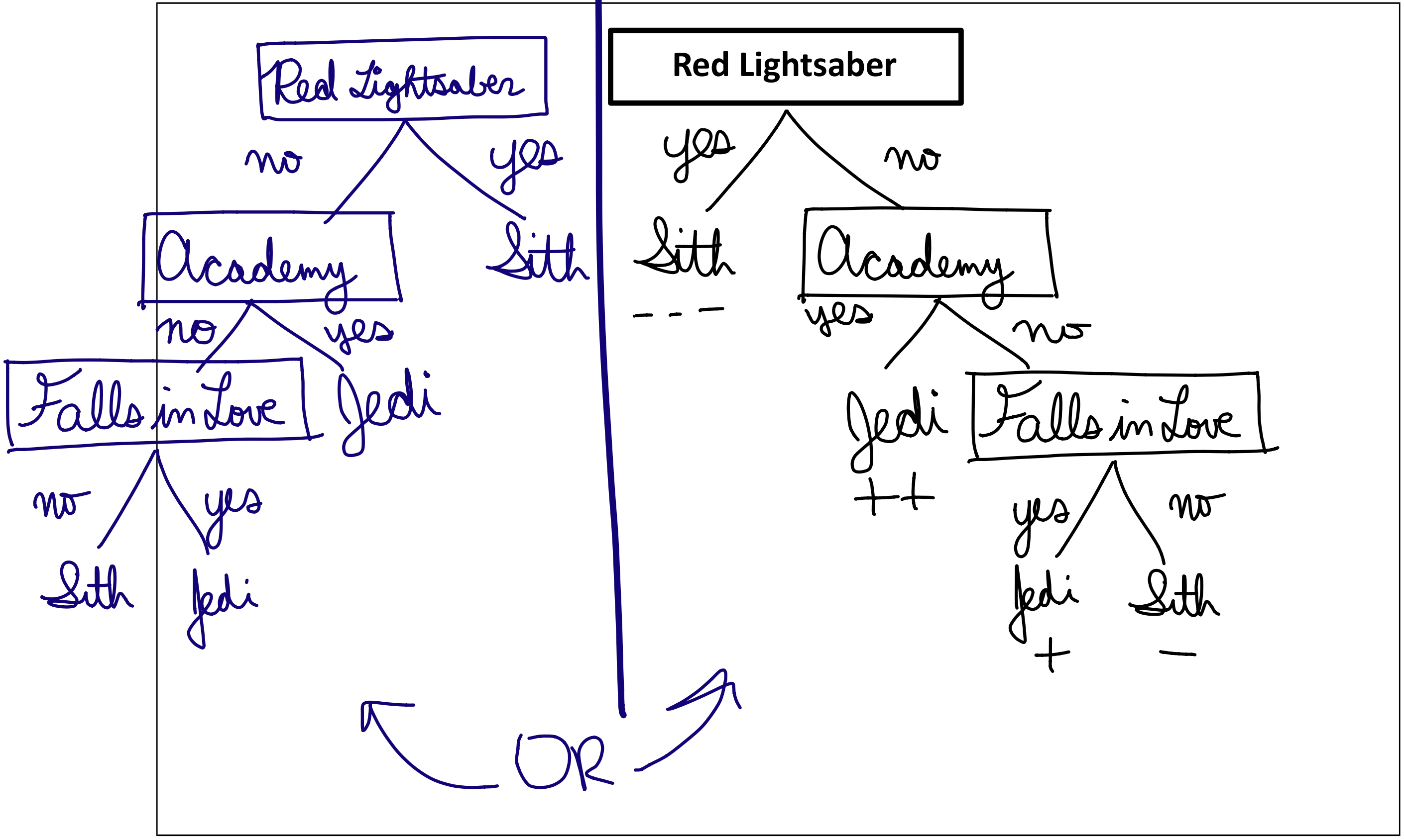
$-\left[\frac{1}{2}\log_2\frac{1}{2}+\frac{1}{2}\log_2\frac{1}{2}\right]=1$ $\quad$ $-\left[\frac{2}{5}\log_2\frac{2}{5}+\frac{3}{5}\log_2\frac{3}{5}\right]\approx 0.97$ $\quad$ $-\left[\frac{1}{3}\log_2\frac{1}{3}+\frac{2}{3}\log_2\frac{2}{3}\right]\approx 0.9$

$-\left[\frac{1}{4}\log_2\frac{1}{4}+\frac{3}{4}\log_2\frac{3}{4}\right]\approx 0.8$ $\quad$ $-\left[\frac{1}{5}\log_2\frac{1}{5}+\frac{4}{5}\log_2\frac{4}{5}\right]\approx 0.72$ $\quad$ $-\left[\frac{1}{6}\log_2\frac{1}{6}+\frac{5}{6}\log_2\frac{5}{6}\right]\approx 0.65$

A2 (14 points) Jedi Master Yoda tells you **"Begin with the Red Lightsaber test, you should."** Regardless of your earlier answers, follow Yoda's advice and **start with the Red Lightsaber feature test**. Choosing from the four feature tests in the table, finish constructing the complete greedy disorder-minimizing identification tree to classify the training data as Jedi or Sith.

Break ties alphabetically: **Academy**, **Falls in Love**, **Lightning**, **Red Lightsaber**.

<u>Draw your identification tree</u> in the space below:



Additional space provided for disorder calculations, for your convenience:

A3 (2 points) To thank Yoda for his help, you want to classify him with your ID tree. Yoda has the following features:

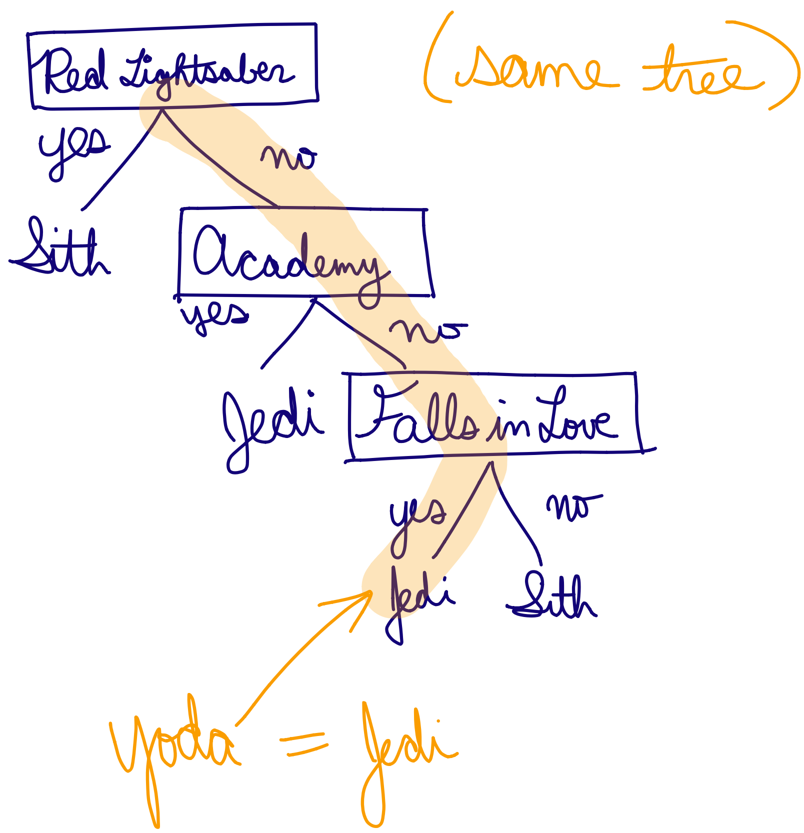
Force User	Classi-fication	Academy	Falls in Love	Lightning	Red Lightsaber
Yoda	Jedi	No	Yes	Unknown	No

Based on these features, how would your identification tree in Part A2 classify Yoda? (Circle one)

Jedi **Sith** **CAN'T TELL**

A4 (1 point) Based on your answer to Part A3 above, does your identification tree classify Yoda correctly? (Circle one)

YES **NO** **CAN'T TELL**

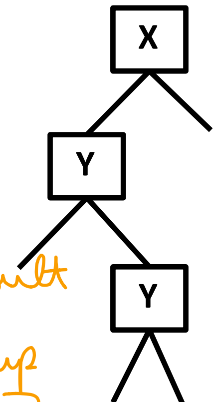


Part B: R2-D2 learns about ID trees (9 points)

Another friend, R2-D2, wants to learn how identification trees work. He displays a sequence of ID trees. Each tree has feature tests labeled with letters, where each different letter represents a distinct test. R2-D2 draws in all the branches, but does not label the branches or the resulting classifications.

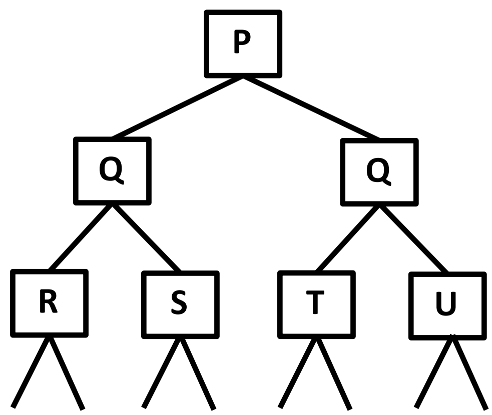
For each tree, R2-D2 asks "¿◈◎◪▤◳◱◓◘▥?", meaning **"Based only on the structure of this tree, could it be a greedy disorder-minimizing ID tree that perfectly classifies its training data?"** For each tree below, circle the **one best response** to help R2-D2 learn about ID trees.

B1 (3 points)



1. **No.** You can't use test Y twice in the same ID tree.
2. **No.** You can use test Y twice, but not in the same branch of the ID tree.
3. **No.** You can use test Y twice in the same branch of the ID tree, but only if there's at least one other test in between. ↳ This would still result in only 1 group in test 2
4. **Yes.** That could be a complete, greedy, disorder-minimizing ID tree.

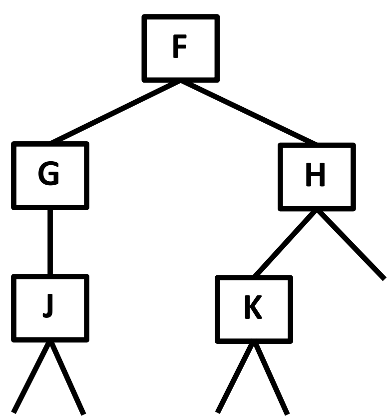
B2 (3 points)



1. **No.** You can't use test Q twice in the same ID tree.
2. **No.** You can use test Q twice, but not in the same level of the ID tree.
3. **No.** At each level, the greedy algorithm must separate out a homogeneous group.
4. **Yes.** That could be a complete, greedy, disorder-minimizing ID tree.

It's possible that no test creates a homogenous group at a given level

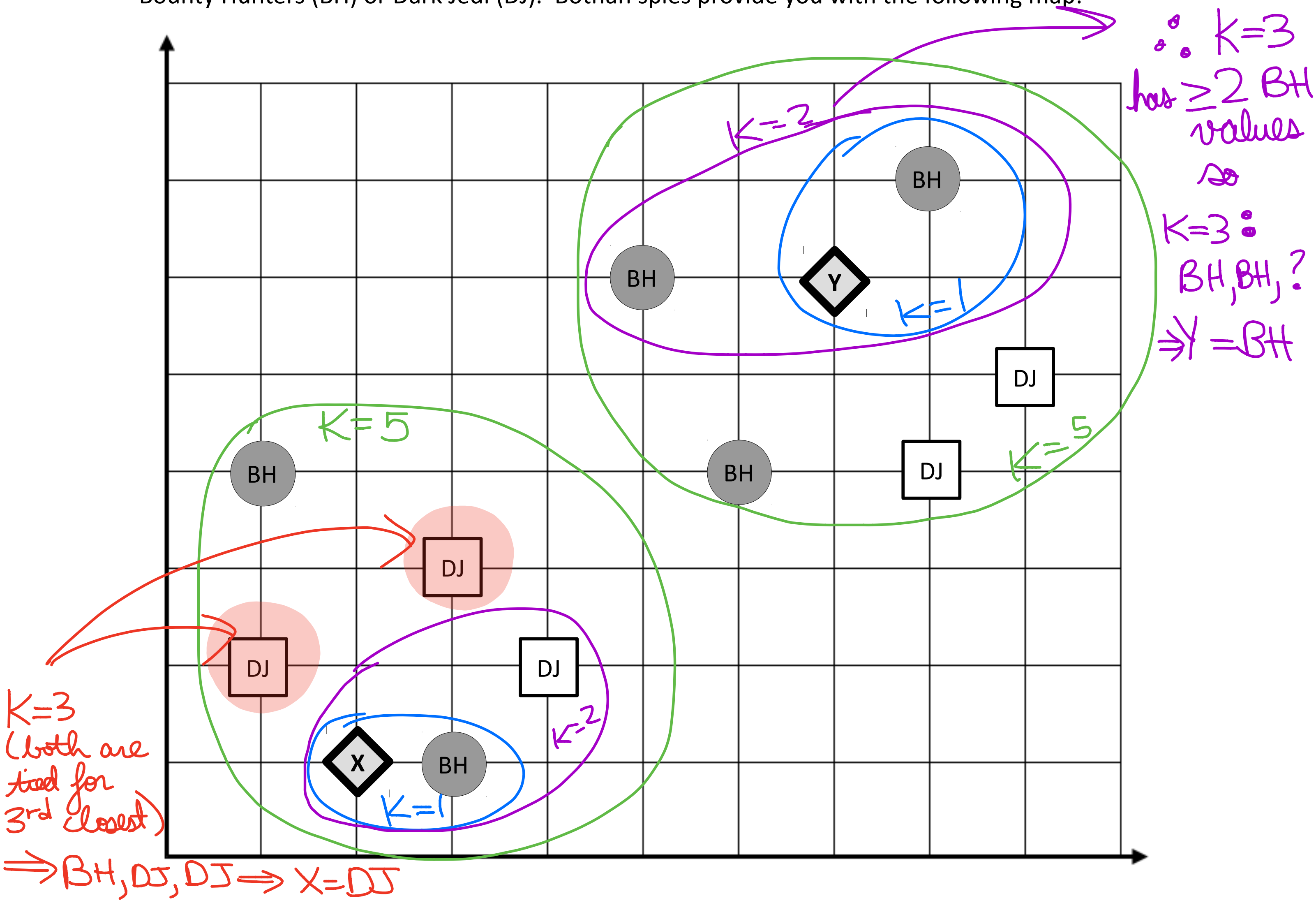
B3 (3 points) This time, R2-D2 adds a constraint to make the exercise more difficult: "Suppose that you have **exactly 5 feature tests and 5 training points**." R2-D2 draws the tree shown here. As before, circle the one best response:



1. **No.** A disorder-minimizing tree must not use all of the feature tests.
2. **No.** The disorder-minimizing algorithm must be given more training points than feature tests.
3. **No.** The tree should never use a feature test with only one branch.
4. **Yes.** That could be a complete, greedy, disorder-minimizing ID tree.

Part C: Nearest neighbors (18 points)

Your next task as an investigator is to determine whether troublemakers in the Outer Rim are Bounty Hunters (BH) or Dark Jedi (DJ). Bothan spies provide you with the following map:



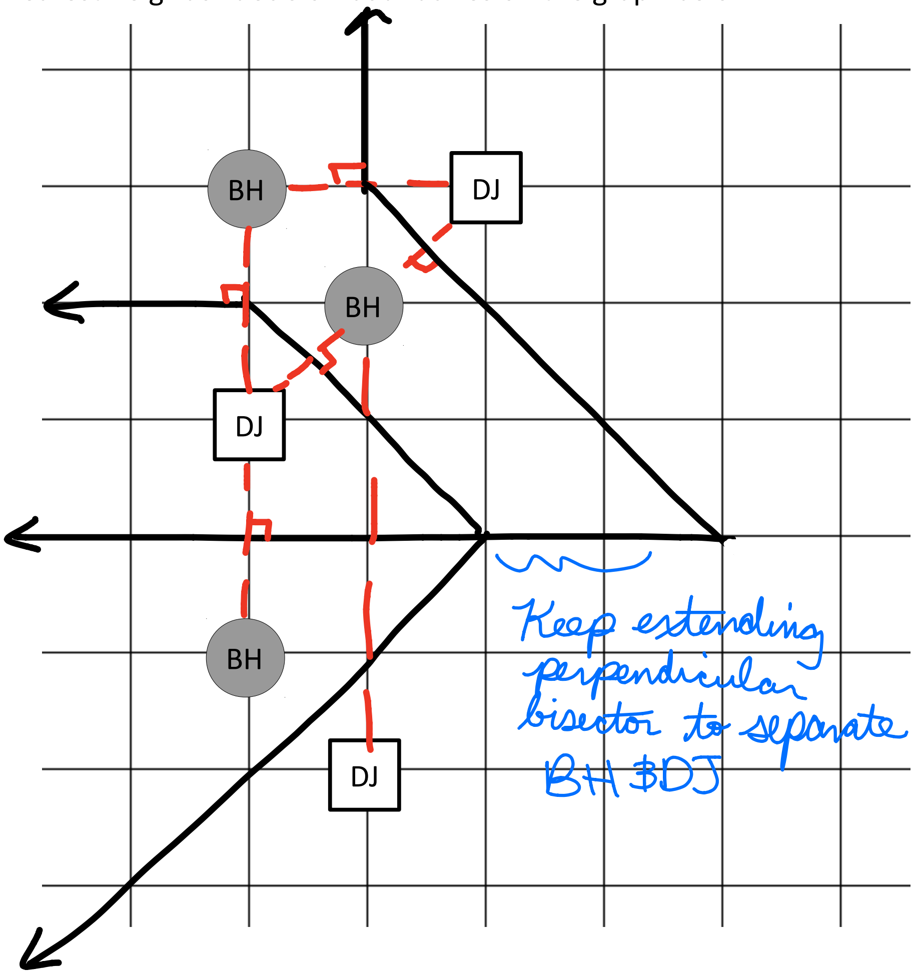
C1 (6 points) On the map above, there are two test points (X and Y) with unknown classifications. Using k-nearest neighbors with a particular value of k, **X is classified as a Dark Jedi (DJ)** and **Y is classified as a Bounty Hunter (BH)**. For which of the following values of k is this true? (All distances are measured using the usual Euclidean distance metric.)

Circle all that apply, or circle NONE OF THESE if none apply:

k = 1 **k = 3** **k = 5** **k = 7**

NONE OF THESE

C2 (12 points) Meanwhile, members of the Rogue Squadron have gathered their own data. Draw the 1-nearest neighbor decision boundaries on the graph below.



extend lines to ∞

Problem 3: Spiritual and Right-Now

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. Sussman's explanation of propagators featured an example involving:

4

   1. Electronic circuit analysis.
   2. Fast implementation of rule chaining.
   3. Calculating the distance to a galaxy.
   4. Programs that explain their own decisions.
   5. Programs that write code.

2. Boyden described a procedure that:

3

   1. Enlarges brains by administering genetically engineered growth hormones.
   2. Enlarges brains by selective breeding.
   3. Enlarges brains by infusing polymers that swell.
   4. Uses transcranial magnetic stimulation to reduce stress in dogs.
   5. Prevents obesity in dogs by genetic manipulations that make fatty foods smell bad.

(MORE QUESTIONS ON NEXT PAGE →)

3. Winston claimed that the success of Waltz's drawing-understanding program is most dependent on:

1. Using hill climbing enhanced by an extended list.
2. Using fast algorithms developed initially for searching maps.
3. Using fast algorithms developed initially for aligning protein sequences.
4. Using constraint propagation to find consistent interpretations.
5. Deploying dropout mechanisms borrowed from machine-learning technology.

4

4. Winston claimed that the success of genetic algorithms is best attributed to:

1. Exploitation of new evolutionary theories.
2. Developing constraint-propagation heuristics.
3. Harnessing evolutionary leaps over short timescales.
4. Using clustering algorithms to infer phenotypes.
5. Picking suitable formulas for survival probability.

5

5. Kanwisher spoke about:

1. Brain regions associated with IQ.
2. Brain regions involved in autism.
3. Brain damage caused by prolonged fatigue.
4. Brain damage caused by electric shock therapy.
5. Brain structure and schizophrenia.

1

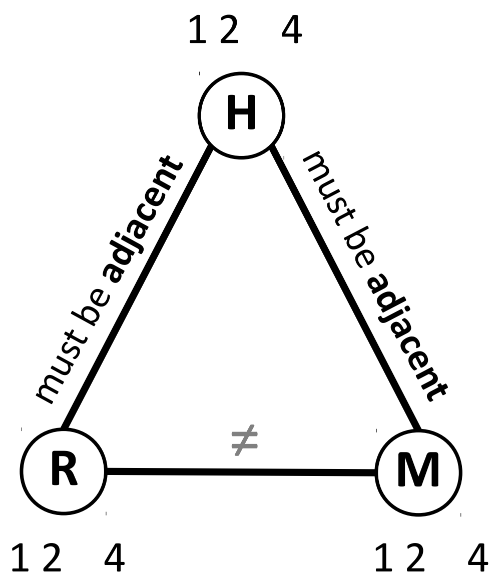
6. Kanwisher claimed:

1. Face recognition ability is correlated with mental arithmetic ability.
2. No brain regions respond selectively to written words because evolution is slow.
3. Response to places, faces, and body parts is concentrated in one brain region.
4. Language loss causes little or no loss of ability to reason logically.
5. Stimulating the brains of epileptic patients can cause them to think they hear music.

4

Tear-off sheet for Problem 1
Initial domains and constraint graph (see reverse for additional copy)

R	1 2 4
M	1 2 4
H	1 2 4



Domain worksheet instructions:

** Note: If you already know how to fill out the domain worksheet, you need not read this. **
Fill out the domain worksheet as you draw your search tree. There may be more rows than you need.

1. Every time you **assign a variable** or **remove a variable from the propagation queue**, fill out a new row in the table. (The same variable might appear in more than one row, especially if you have to backtrack.)
2. In that row, indicate **which variable you assigned or de-queued**; write its **assigned** value if it has one (e.g. X=x), otherwise just write its **name** (X). In the second column, list the **values that were just eliminated from neighboring variables** as a result. If no values were just eliminated, write **NONE** instead.
3. If your search has to backtrack after assigning or de-queuing a variable: first, **finish listing** all values eliminated from neighboring variables in the current row. Next, check the "backtrack" box in that row. Then, continue with the next assignment in the following row as usual.
4. If you add several variables to your propagation queue at once, break ties by adding variables to your propagation queue **in alphabetical order**. Only add variables to your propagation queue **if they are not already in the queue**.

Example row showing an assigned variable

ex	X = 3	Y ≠ 3, 4 Z ≠ 3 (example)	☐

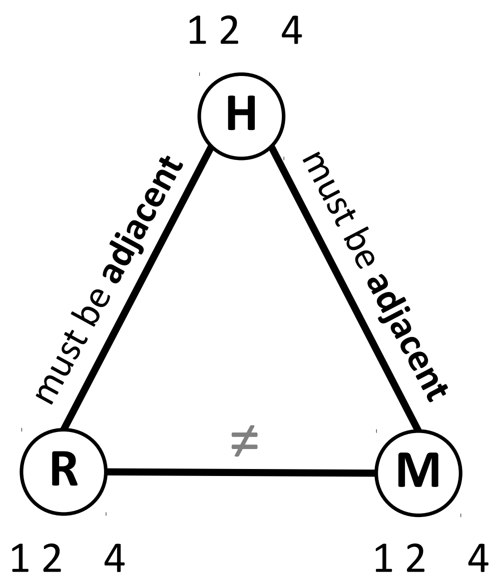
Example row showing a de-queued (propagated) variable

ex	X	W ≠ 1, 4 (example)	☐

Tear-off sheet

Additional copy of Problem 1 initial domains and constraint graph:

R	1 2 4
M	1 2 4
H	1 2 4



Problem 2 training data:

Force User	Classi-fication	Academy	Falls in Love	Lightning	Red Lightsaber
1 Obi-Wan Kenobi	Jedi	Yes	Yes	No	No
2 Qui-Gon Jinn	Jedi	Yes	No	No	No
3 Luke Skywalker	Jedi	No	Yes	No	No
4 Darth Vader	Sith	Yes	Yes	No	Yes
5 Count Dooku	Sith	Yes	No	Yes	Yes
6 Emperor Palpatine	Sith	No	No	Yes	Yes
7 General Grievous	Sith	No	No	No	No

Problem 2 table of logarithms:

$$-\left[\frac{1}{2}\log_2\frac{1}{2} + \frac{1}{2}\log_2\frac{1}{2}\right] = 1 \qquad -\left[\frac{2}{5}\log_2\frac{2}{5} + \frac{3}{5}\log_2\frac{3}{5}\right] \approx 0.97 \qquad -\left[\frac{1}{3}\log_2\frac{1}{3} + \frac{2}{3}\log_2\frac{2}{3}\right] \approx 0.9$$

$$-\left[\frac{1}{4}\log_2\frac{1}{4} + \frac{3}{4}\log_2\frac{3}{4}\right] \approx 0.8 \qquad -\left[\frac{1}{5}\log_2\frac{1}{5} + \frac{4}{5}\log_2\frac{4}{5}\right] \approx 0.72 \qquad -\left[\frac{1}{6}\log_2\frac{1}{6} + \frac{5}{6}\log_2\frac{5}{6}\right] \approx 0.65$$