6.034 Quiz 2

15 October 2014

Name	MARVIN THE MIND SCULPTOR
Email	minsky@tolaria.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.

Josh Blum	**Elisa Castañer**	**Pedro Cattori**
Jack Florey	**Malcom Gilbert**	**Dylan Holmes**
Jessica Noss	**Duncan Townsend**	**Siyao Xu**

Problem number	Maximum	Score	Grader
1	50		
2	50		
Total	100		

SRN	6		

There are 15 pages in this quiz, including this one. As always, open book, open notes, open just about everything, including a calculator, but no computers.

Problem 1: Identification of Trees (50 points)

Part A: Identification Trees (28 points)

While walking along the Charles River in the fall, your friend points out six trees and tells you whether they are Maple or Oak trees. Eager to identify other trees yourself, you observe what appear to be their key features. In particular, you record whether each tree **Has leaves (Yes or No)**, whether it has **Orange foliage (Yes or No)**, its **Leaf shape (Pointy or Rounded)**, and its **Bark texture (Glossy, Furrowed, or Smooth)**. Your resulting training data is summarized in the table below:

Tree type	Has leaves*	Orange foliage	Leaf shape	Bark texture
Oak	Yes	Yes	Pointy	Glossy
Oak	Yes	No	Pointy	Furrowed
Oak	Yes	No	Rounded	Furrowed
Maple	Yes	Yes	Pointy	Furrowed
Maple	Yes	Yes	Pointy	Smooth
Maple	Yes	No	Pointy	Smooth

*Although some trees lose their leaves in the fall, all six of these trees still have leaves.

A1 (4 points) Based on these six training points, order the four feature tests (**Has leaves**, **Orange foliage**, **Leaf shape**, and **Bark texture**) from least to greatest disorder. (You don't need to show your calculations; you may use your intuition. If you prefer not to use your intuition, see part A2 below.) Break any ties alphabetically.

BARK TEXTURE	LEAF SHAPE	ORANGE FOLIAGE	HAS LEAVES
LEAST DISORDER			GREATEST DISORDER

A2 (8 points) Next, compute the disorder of each feature test. Use the table of logarithms below to express your answer as sums and products of decimals and fractions only. Your final answer should have **no logarithms** in it. Space is provided on the next page to show your work.

Test	Has leaves	Orange foliage	Leaf shape	Bark texture
Disorder	1	0.9	$\frac{5}{6}(0.97) \approx 0.8$	$\frac{1}{2}(0.9) = 0.45$

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

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

Show your disorder calculations for partial credit. For your convenience, a copy of the training data is provided on the next page.

Has leaves

HAS LEAVES

yes: OOO MMM — disorder = 1

no: X

Orange foliage

ORANGE LEAVES

yes: OMM — disorder ≈ 0.9

no: OOM — disorder ≈ 0.9

$$\frac{3}{6}(0.9) + \frac{3}{6}(0.9) = \boxed{0.9}$$

Leaf shape

LEAF SHAPE

pointy: OOMMM — disorder ≈ 0.97

rounded: O — disorder = 0

$$\frac{5}{6}(0.97) + \frac{1}{6}(0) \approx \boxed{0.8}$$

Bark texture

BARK TEXTURE

glossy: O — disorder = 0

furrowed: OOM — disorder ≈ 0.9

smooth: MM — disorder = 0

$$\frac{1}{6}(0) + \frac{3}{6}(0.9) + \frac{2}{6}(0) = \boxed{0.45}$$

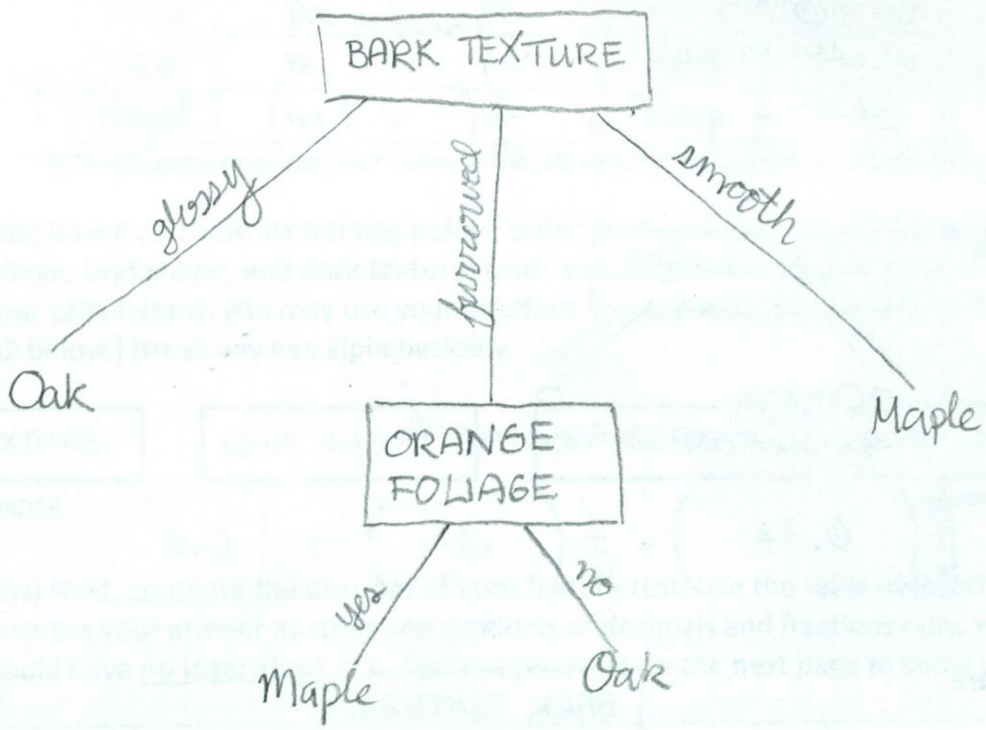
A3 (14 points)
Choosing from the four feature tests (**Has leaves**, **Orange foliage**, **Leaf shape**, and **Bark texture**), construct the complete disorder-minimizing identification tree for classifying the training data as Maple or Oak.

(For your convenience, a copy of the training data is included here.)

	Tree type	Has leaves*	Orange foliage	Leaf shape	Bark texture
1	Oak	Yes	Yes	Pointy	Glossy
2	Oak	Yes	No	Pointy	Furrowed
3	Oak	Yes	No	Rounded	Furrowed
4	Maple	Yes	Yes	Pointy	Furrowed
5	Maple	Yes	Yes	Pointy	Smooth
6	Maple	Yes	No	Pointy	Smooth

*Although some trees lose their leaves in the fall, all six of these trees still have leaves.

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



Additional space provided for disorder calculations, if you need it.

For the first round, the tests have the following disorders:

lowest = best

HAS LEAVES	ORANGE FOLIAGE	LEAF SHAPE	BARK TEXTURE
1	0.9	0.8	0.45

We choose the BARK TEXTURE test, which completely classifies training points with TEXTURE = glossy and TEXTURE = smooth.

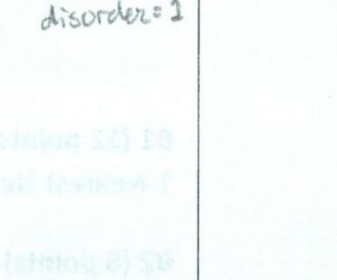
Next, we need a second test to separate the remaining training points with TEXTURE = furrowed. We <u>cannot</u> just use the second-best test from the first round because in general the disorder of a test will change when you look at only a particular subset of the training data. So, we recalculate:

For the second round, when we consider just the training points that have TEXTURE = furrowed, the tests have the following disorders:

So we pick ORANGE FOLIAGE as our second test, which completely classifies the remaining training data.

A4 (2 points)

How would your identification tree classify a specimen with the following features?

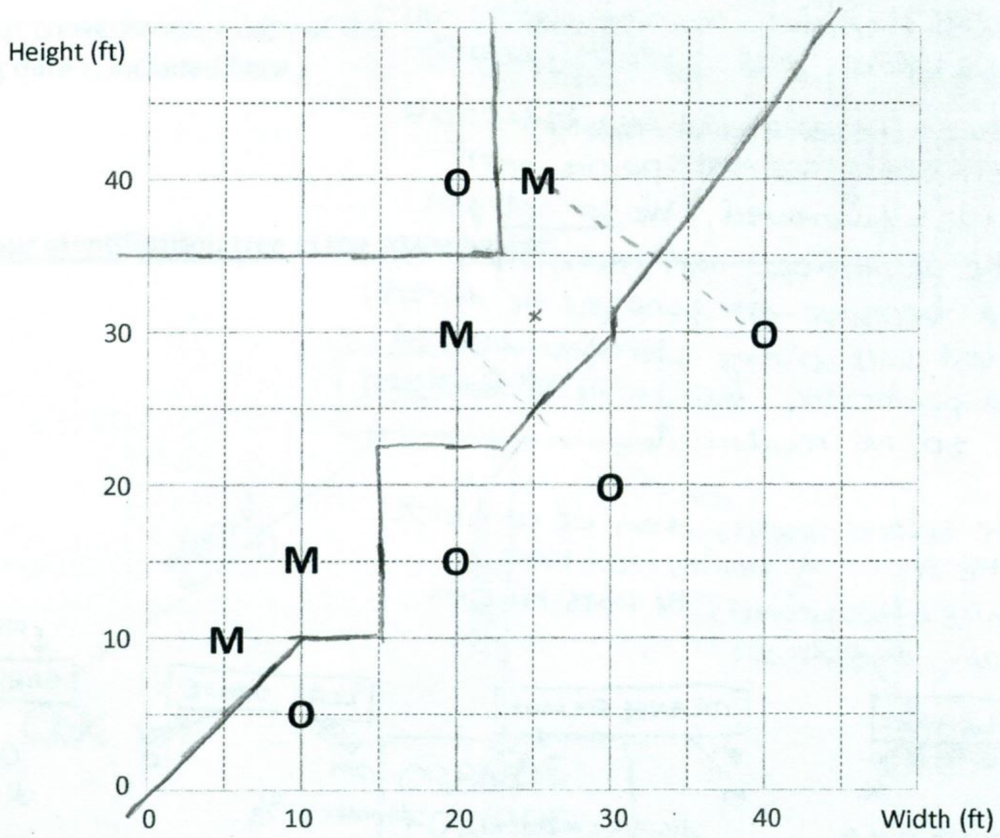


Has leaves	Orange foliage	Leaf shape	Bark texture
Yes	Yes	Rounded	Furrowed

Circle one: MAPLE OAK CAN'T TELL

Part B: Nearest neighbors (22 points)

You plan to go out and identify more trees the next day, but that night there is an early snowstorm! To your dismay, you find that a blanket of snow has obscured the leaves and bark of all the trees, so you can no longer use those as distinguishing features. Undaunted, you decide to venture out into the fresh snow and use trigonometry to measure the width and height of trees that are known to be Maple (**M**) or Oak (**O**). Your measurements are plotted on the following graph:



B1 (12 points) On the graph above, **draw the decision boundaries** produced by 1-Nearest Neighbors.

B2 (8 points) One of your classmates hears that you've been identifying trees. She wants to know what type of tree is knocking snow onto her second-story windowsill, so she calls you with the information that her tree is **25 ft wide** and **32 ft tall.** How would this tree be classified with each of the following classifiers? (Circle the best answer in each case.)

1-Nearest Neighbors?	**MAPLE**	**OAK**	**CAN'T TELL**
3-Nearest Neighbors?	**MAPLE**	**OAK**	**CAN'T TELL**
5-Nearest Neighbors?	**MAPLE**	**OAK**	**CAN'T TELL**
9-Nearest Neighbors?	**MAPLE**	**OAK**	**CAN'T TELL**

B3 (2 points) Finally, you want to use Nearest Neighbors to classify the pre-snowstorm data in Part A. You realize it's possible, but you must first: (Circle the **one** best answer.)

(A) Pick just two of the features to plot.

B (B) Define a distance metric for non-numeric features.

(C) Collect twice as many training points.

(D) Use cross-validation to guarantee overfitting.

(E) Devise an admissible heuristic.

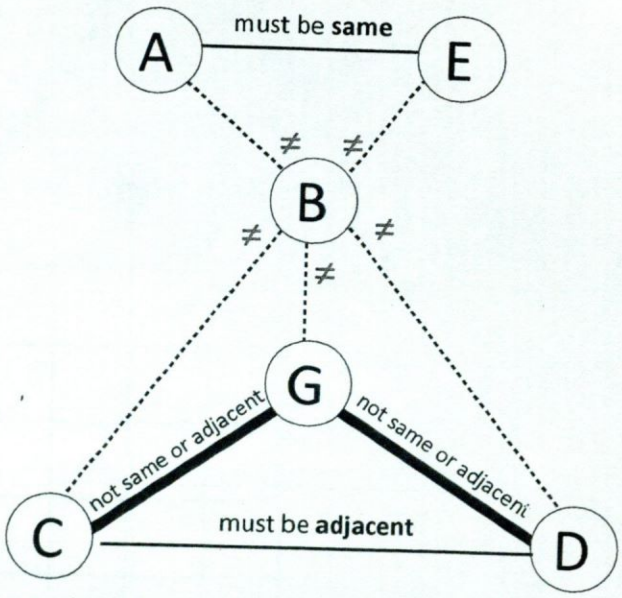
Problem 2: Dominian Domiciles (50 points)

Six powerful mages (**A**jani, **B**olas, **C**handra, **D**ack, **E**lspeth, and **G**ideon) have enlisted you to help them settle down and find homes in the multiverse. They have five neighboring worlds to choose from, arranged in a row as shown. Some worlds could be empty, and some worlds could have more than one occupant.

1	2	3	4	5

Naturally, mages are very particular about where they live, so your assignments must satisfy the following requirements:

1. **A**jani is **E**lspeth's mentor. They must live on the **<u>same world</u>**.
2. Everyone hates **B**olas. No one can live on the same world as him.
3. **C**handra wants to spy on **D**ack secretly—she must live on a world that's **<u>adjacent to</u>** **D**ack's world.
4. **G**ideon is a law mage. He despises the criminals **C**handra and **D**ack so much that **G**ideon **<u>cannot live on the same world—or even on a world adjacent to</u>** **C**handra's or **D**ack's.



(For your convenience, a graph of these constraints is shown here.)

Finally, the mages inform you that you can simplify the domains of these variables, as follows:

- **D**ack **must live on either World 1 or 4**, because he's a wanted criminal everywhere else.
- **E**lspeth **must live on World 3**, because she's stuck there and cannot leave.
- **B**olas **can't live on World 5**, because he fears its strange magic.

Thus, the initial domains of your variables are as follows:

A	1 2 3 4 5
B	1 2 3 4
C	1 2 3 4 5
D	1 4
E	3
G	1 2 3 4 5

Part A. Prestidigitation (9 points)

A1 (3 points) In general (not just on this problem), when you are solving constraint satisfaction problems, which of the following strategies can require you to backtrack? (Circle **all answers** that apply.)

(A) Depth-first search (check assignments only)

(B) Depth-first search with forward checking

(C) Depth-first search with forward checking and propagation through singleton domains.

(D) Depth-first search with forward checking and propagation through reduced domains.

(E) None of the above methods ever require backtracking.

A2 (6 points) Consider just three of the mages: **A**jani, **B**olas, and **C**handra. In what order should you assign values to them so that constraint propagation will be maximally efficient? Write their initials (**A, B, C**) in the order they should be assigned below:

B	C	A
Assign 1st	Assign 2nd	Assign 3rd

most edges ... fewest edges

Part B. Homelands (35 points)

Regardless of your earlier answers, the stubborn mages insist that you assign variables in the order **A, E, D, B, G, C**. Use **depth-first search with forward checking and propagation through singleton domains (domains reduced to size 1)** to find worlds for the six mages.

★ On the next two pages, show your work by simultaneously

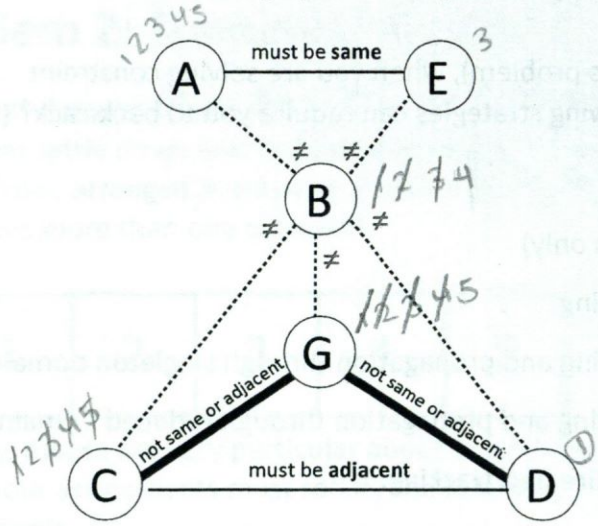
(1) filling out the domain worksheet and

(2) drawing the search tree.

☆ For your convenience, a **duplicate** copy of the domain worksheet and tree is provided on pages 12-13.

If you write on both copies, please indicate clearly which one you want us to grade.

Constraint graph for this problem



Domains for this problem

(*Do not reduce these domains before starting search.*)

A	1 2 3 4 5
B	1 2 ~~3~~ 4
C	1 2 3 4 5
D	1 4
E	3
G	1 2 3 4 5

Fill out this 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**.

TA NOTE: This row is optional depending on whether the convention is to propagate any neighbor of size 1, or just neighbors that have been reduced to size 1 by fwd checking. (better)

	Var assigned or de-queued	List all values just eliminated from neighboring variables	Back track		Var assigned or de-queued	List all values just eliminated from neighboring variables	Back track
1	A=1	B≠1 E≠3	☒	8	B	G≠4 !	☐
2	A=2	B≠2 E≠3	☒	9	G	NONE	☐
3	A=3	B≠3	☐	10	B=4	NONE	☐
4	(E	NONE)	☐	11	G=5	NONE	☐
5	E=3	NONE	☐	12	C=2	NONE	☐
6	D=1	B≠1 G≠1,2 C≠1,3,4,5!	☐	13			☐
7	C	B≠2! G≠3	☐	14			☐

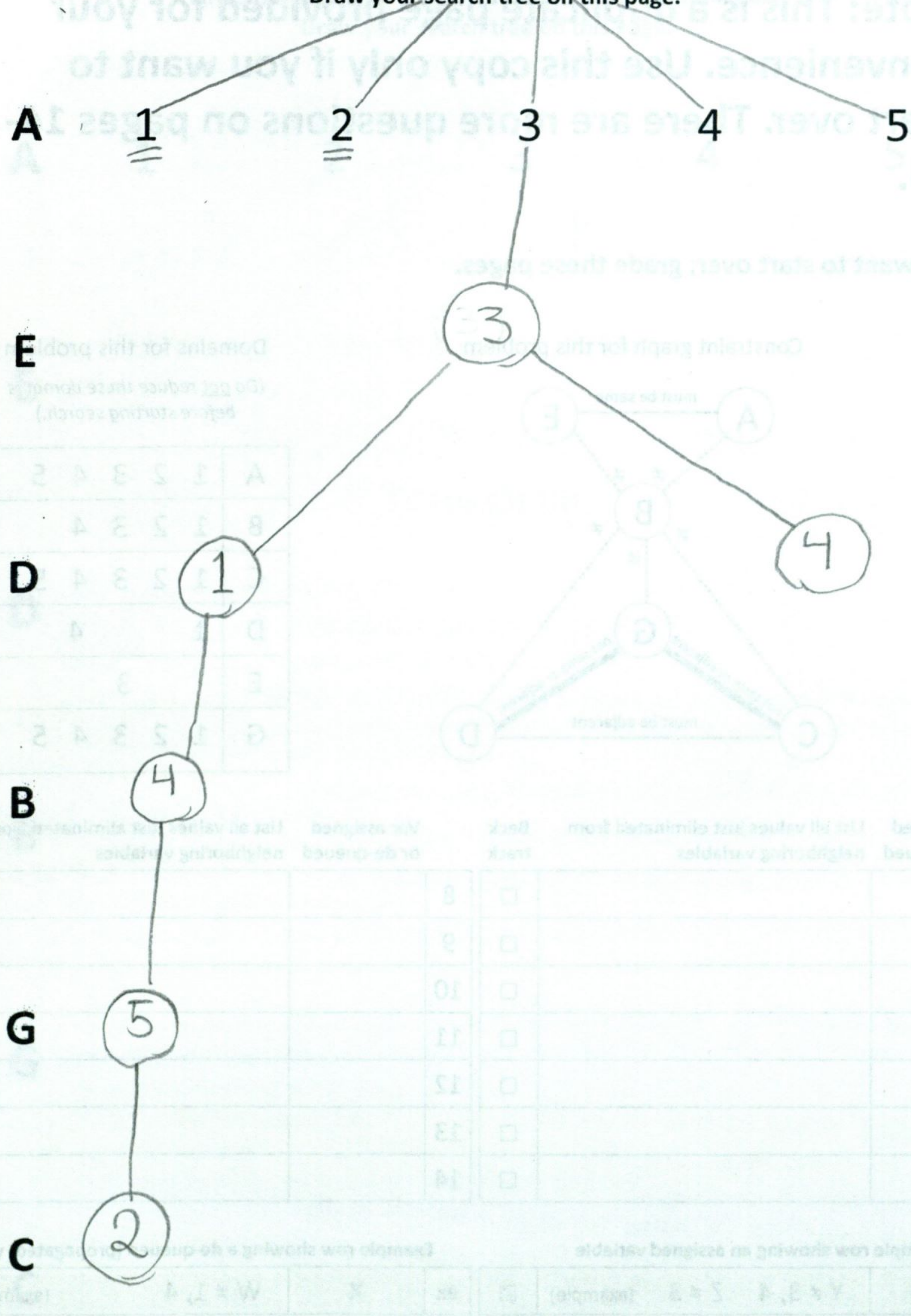
Example row showing an assigned variable

ex	X = 3	Y ≠ 3, 4 Z ≠ 3 (example)	☑

Example row showing a de-queued (propagated) variable

ex	X	W ≠ 1, 4 (example)	☐

Draw your search tree on this page.



Part C (6 points)

C1 (3 points) Based on the results of your search in Part B, write the initials of the mages (**A, B, C, D, E, G**) in their assigned homes among the five worlds.

1	2	3	4	5
D	C	A E	B	G

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

NO SOLUTION

C2 (3 points) Based <u>only on the tree</u> you drew in Part B, what can you conclude about the number of solutions to this problem (i.e. consistent ways of assigning mages to worlds)?

(A) There are **no solutions** to this problem—it's over-constrained.

(B) There is **exactly one** solution to this problem.

(C) There is **at least one** solution to this problem, maybe more.

(D) There is **definitely more than one** solution to this problem.

(A) would be true if we had to backtrack until no solutions were left.

(B) would be true if we drew the whole tree and saw that only one solution worked.

(C) is true because we drew enough of the tree to find at least one solution.

(D). More search is required to tell if there is more than one solution or if all other paths are dead ends.

Problem 3: Spiritual and Right-Now

Circle the **one 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.

1. Sussman is motivated by a desire to understand:

4

1. How English emerged from Germanic and French roots.
2. How to write more secure software.
3. How to write provably correct software.
4. How thinking works.
5. How to treat autism.

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

2

1. How stories change in time as they are retold.
2. Fast implementation of rule chaining.
3. Calculating the distance to a galaxy.
4. Proving a program to be correct.
5. Tweets.

3. In Sussman's propagator system, cells:

4

1. Pass on information distorted by a random number generator.
2. Perform complex logical operations using table lookup.
3. Achieve great speed by using multiple cores.
4. Hold "information about values" rather than "values."
5. Have values that reflect cultural beliefs.

4. Boyden describe how rats can be:

2

1. Trained to stand on two legs.
2. Respond to light without normal retinal photoreceptors.
3. Rendered vicious by cauterization of areas in the forebrain.
4. Shown to dream about movement through a familiar maze.
5. Made obese by electrical stimulation of brain areas associated with hunger.

5. Boyden described a procedure involving direct electrical stimulation of brain tissue so as to:

4

1. Cure chronic indigestion.
2. Suppress disturbing memories.
3. Recall childhood experiences.
4. Stop hand tremors.
5. Reduce symptoms associated with Tourette's syndrome.

6. Optogenetics has to do with:

3

1. Evolution of the mammalian visual system.
2. Optimization of brain circuitry by survival of the fittest.
3. Optical stimulation of genetically altered neurons.
4. Manipulation of the genetic code using lasers.
5. Poor visual acuity associated with old age.