6.034 Quiz 2

19 October 2016

Name	
Email	

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

Jake Barnwell	**Michaela Ennis**	**Rebecca Kekelishvili**
Vinny Chakradhar	**Phil Ferguson**	**Nathan Landman**
Alex Charidis	**Stevie Fine**	**Samarth Mohan**
Brian Copeland	**Brittney Johnson**	**Jessica Noss**

Problem number	Maximum	Score	Grader
1 - Constraints	50		
2 - ID Trees	35		
3 - kNN	15		
Total	100		

SRN	6		

Survey

We are investigating the false-positive rate for SRN questions. We will **not** consider your answer to this question when determining grades.

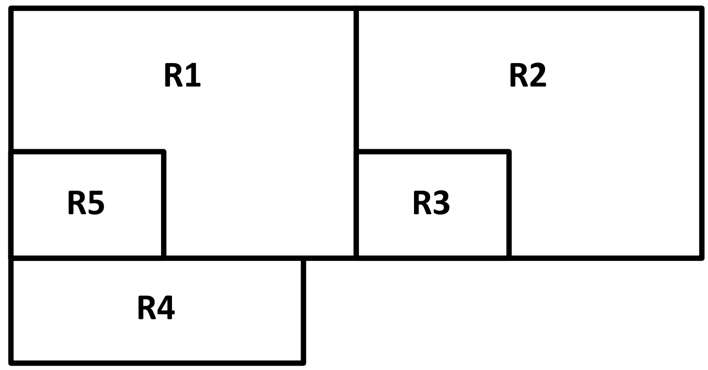
Out of 5 SRN lectures (Sussman, Boyden, Kanwisher, line drawings, deep nets), how many did you attend?

There are 14 pages in this quiz, including this one, but not including tear-off sheets. A tear-off sheet with duplicate drawings and data is located after the final page of the quiz.

As always, open book, open notes, open just about everything, including a calculator, but no computers or phones.

Problem 1: The Architect and the Client (50 points)

You are an architect, and a rich client asks you to design a summer house. Your task will be to assign a function to each room: **B**edroom, **L**iving Room, **K**itchen, or **W**ater Closet (restroom). Here is the floorplan, with five rooms labeled R1 through R5:

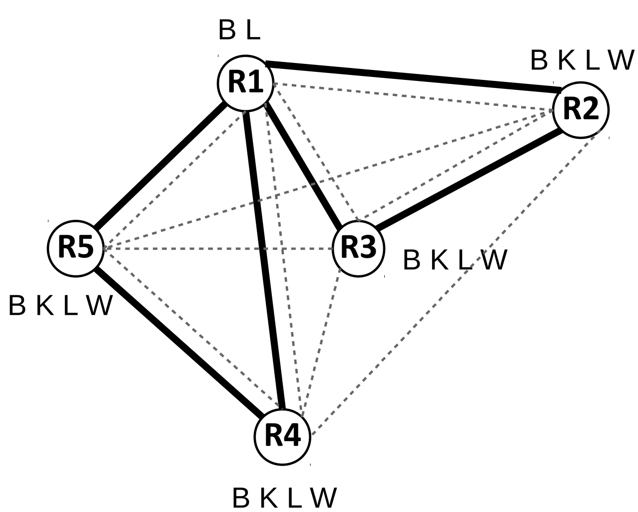


The requirements are:

1. The house must have exactly 1 **B**edroom, 1 **K**itchen, 1 **L**iving Room, and 2 **W**ater Closets.
2. Room R1 must be either the **B**edroom or the **L**iving Room (because that corner of the house has the best view).
3. Adjacent rooms must NOT have the same function, e.g. two adjacent rooms cannot both be Water Closets.

To help you assign functions to rooms, we've drawn the constraint graph for you.

- Solid black lines represent physical adjacencies (can't-be-equal constraints).
- Dashed gray lines represent the global constraints (e.g. only one room can be the Kitchen).
- The domains have been initialized for you, taking into account the fact that R1 must be either the **B**edroom or the **L**iving Room.



(Extra copies of this graph are available on the tear-off sheet after the last page of the quiz.)

Part A: Design (26 points)

A1 (20 points) Perform **<u>Depth-First Search with Forward Checking</u>** (but no propagation) to find a solution. Make assignments to rooms in lexicographic order (R1, R2, …). Do NOT reduce domains before search.

Hint: Make sure to keep track of global constraints! For example, you can write a list **B**, **K**, **L**, **W**, **W** and cross out a letter each time you make an assignment.

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

Domain worksheet instructions:

1. Every time you **<u>assign a variable</u>** or **<u>remove a variable from the propagation queue</u>** (if applicable), fill out a new row in the table. (There may be more rows than you need.)
2. In that row, indicate **which variable you assigned or de-queued**; write its **<u>assigned value</u>** if it has one (e.g. X=x), otherwise just write its **<u>name</u>** (e.g. X). In the second column, list the **values that were just eliminated from neighboring variables** as a result (or "NONE" if no values were eliminated). Do not eliminate values from variables that have already been assigned.
3. If your search has to backtrack after assigning or de-queuing a variable: First, **<u>finish listing</u>** 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 lexicographic order (e.g. R1 before R2).

Example row showing an assigned variable (with backtracking)

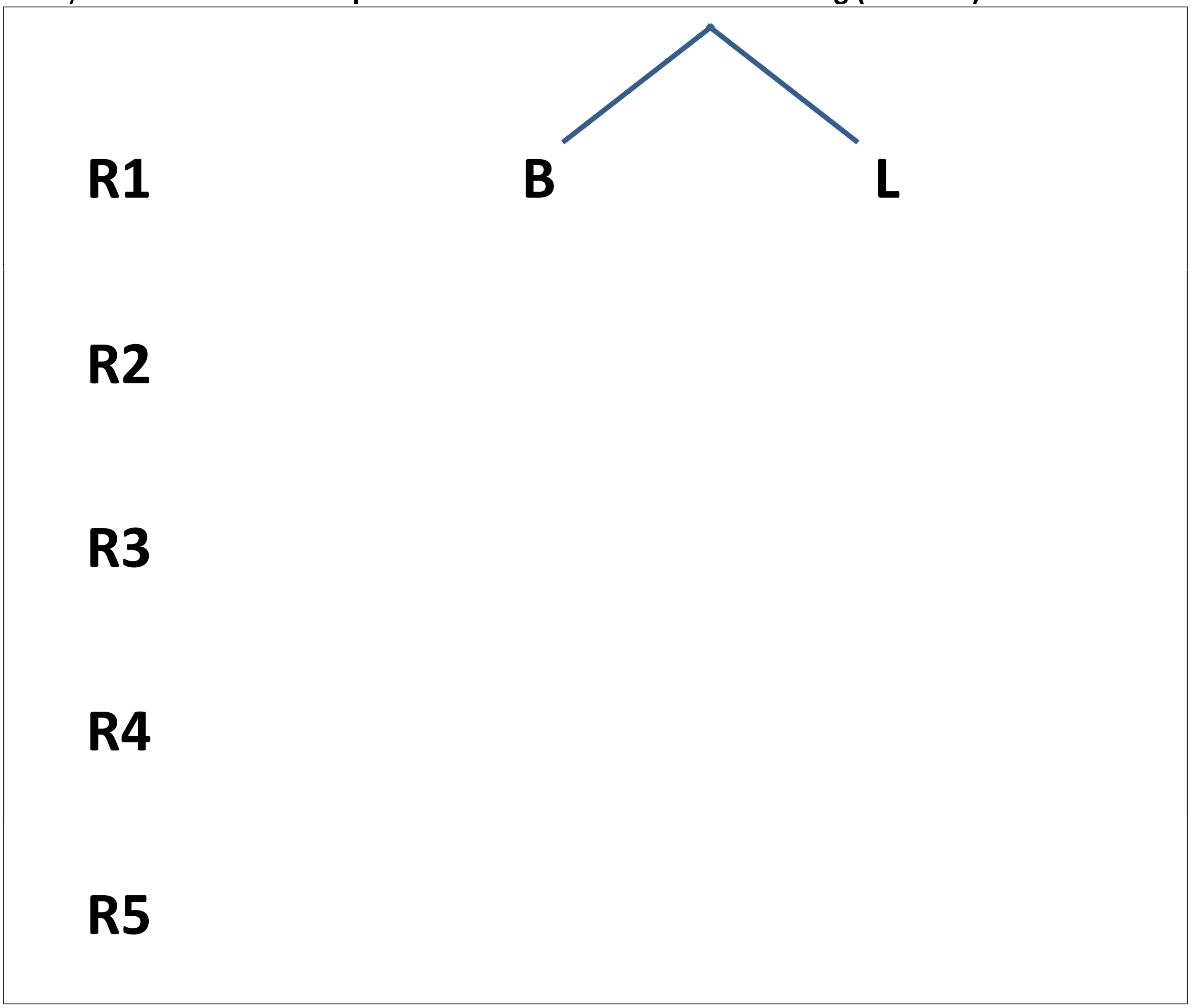
ex	X = 3	Y ≠ 3, 4 Z ≠ 3 (example)	☑

Example row showing a de-queued (propagated) variable

ex	X	W ≠ 1, 4 (example)	☐

(If you want to start over, an extra copy of this tree and domain worksheet are available on the next page. If you write on both, please indicate clearly which copy you want us to grade.)

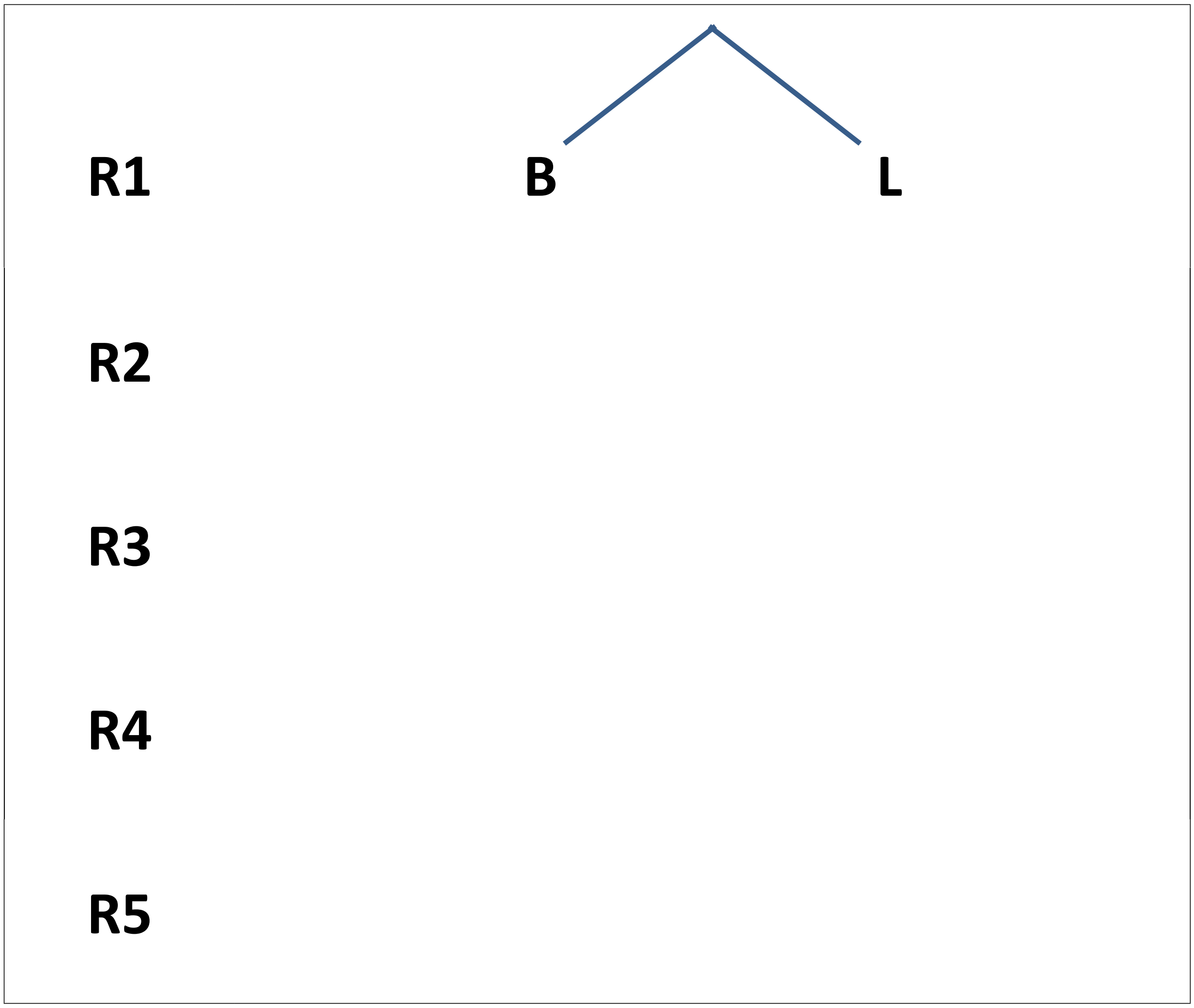
Draw your search tree for **Depth-First Search with Forward Checking (DFS + FC)** here:



	Var assigned or de-queued	List all values just eliminated from neighboring variables	Backtrack?
1			☐
2			☐
3			☐
4			☐
5			☐
6			☐
7			☐
8			☐
9			☐
10			☐

This is a duplicate copy of Part A1 (DFS+FC). If you want to have this copy graded instead, check the box:

☐ **I want to start over; grade this copy**

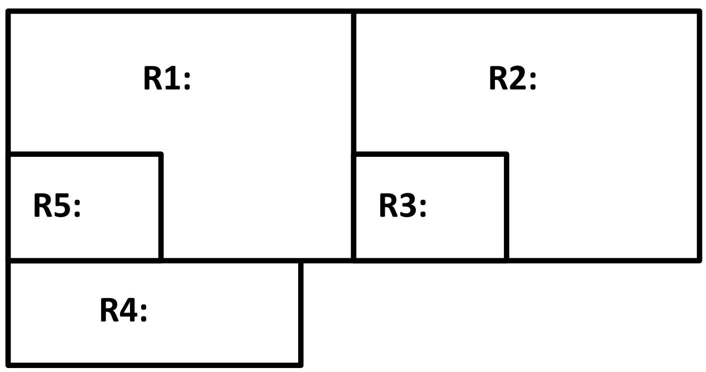


	Var assigned or de-queued	List all values just eliminated from neighboring variables	Backtrack?
1			☐
2			☐
3			☐
4			☐
5			☐
6			☐
7			☐
8			☐
9			☐
10			☐

A2 (2 points) How many times did you backtrack?

A3 (2 points) In total, how many variable assignments did you undo when backtracking?

A4 (2 points) Based on the results of your search, write the assigned room function (**B**, **K**, **L**, or **W**) in each room below, or circle **NO SOLUTION** if no solution was found:



NO SOLUTION

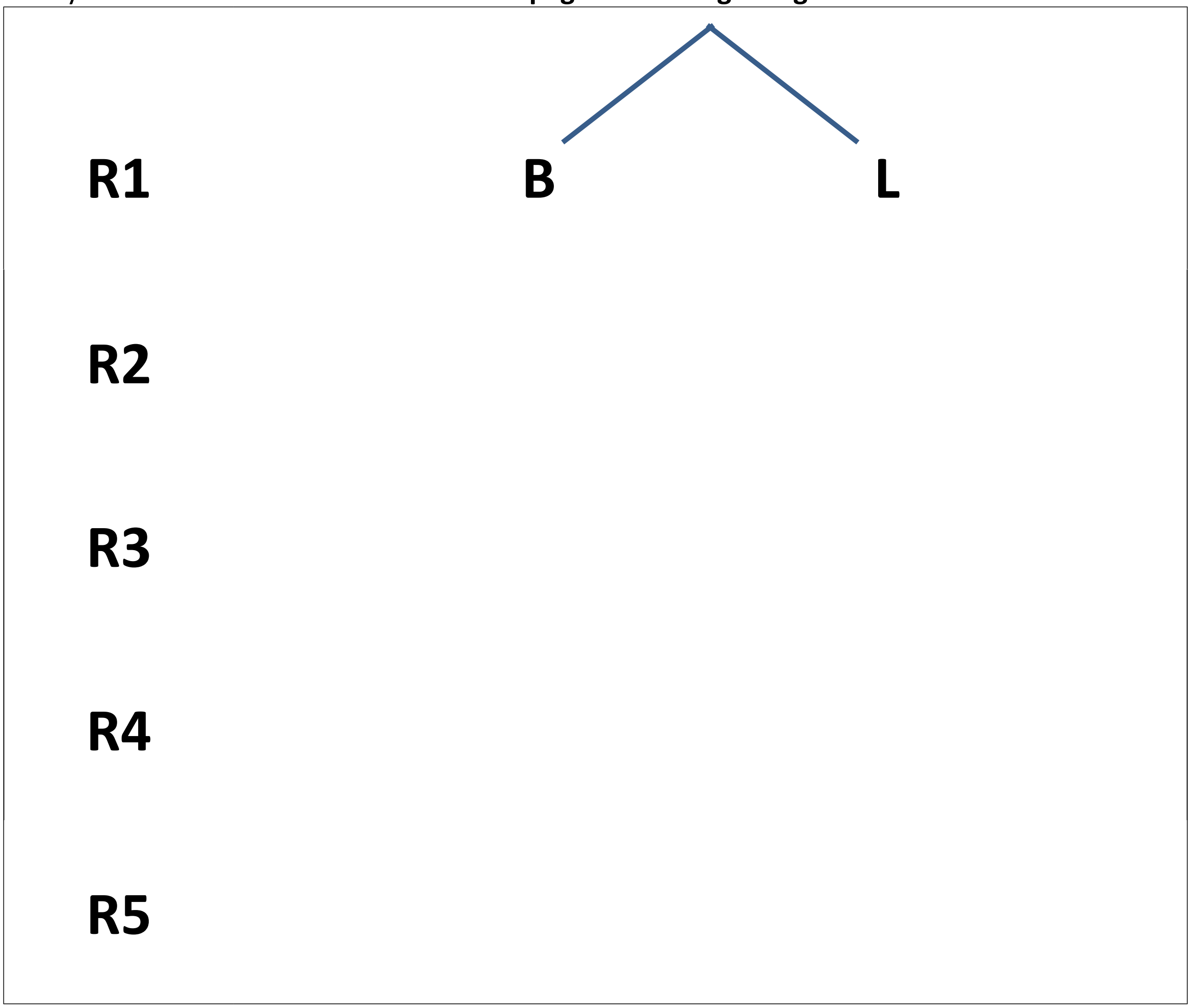
Part B: Design faster! (24 points)

B1 (20 points) Your client is eager to vacation in the new house, so you want a faster way to complete the design process! Perform **Depth-First Search with Forward Checking and Propagation through Singleton Domains**. As before, assign variables in lexicographic order (R1, R2, …). Do NOT reduce domains before search.

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

(If you want to start over, an extra copy of this tree and domain worksheet are available on the next page. If you write on both, please indicate clearly which copy you want us to grade.)

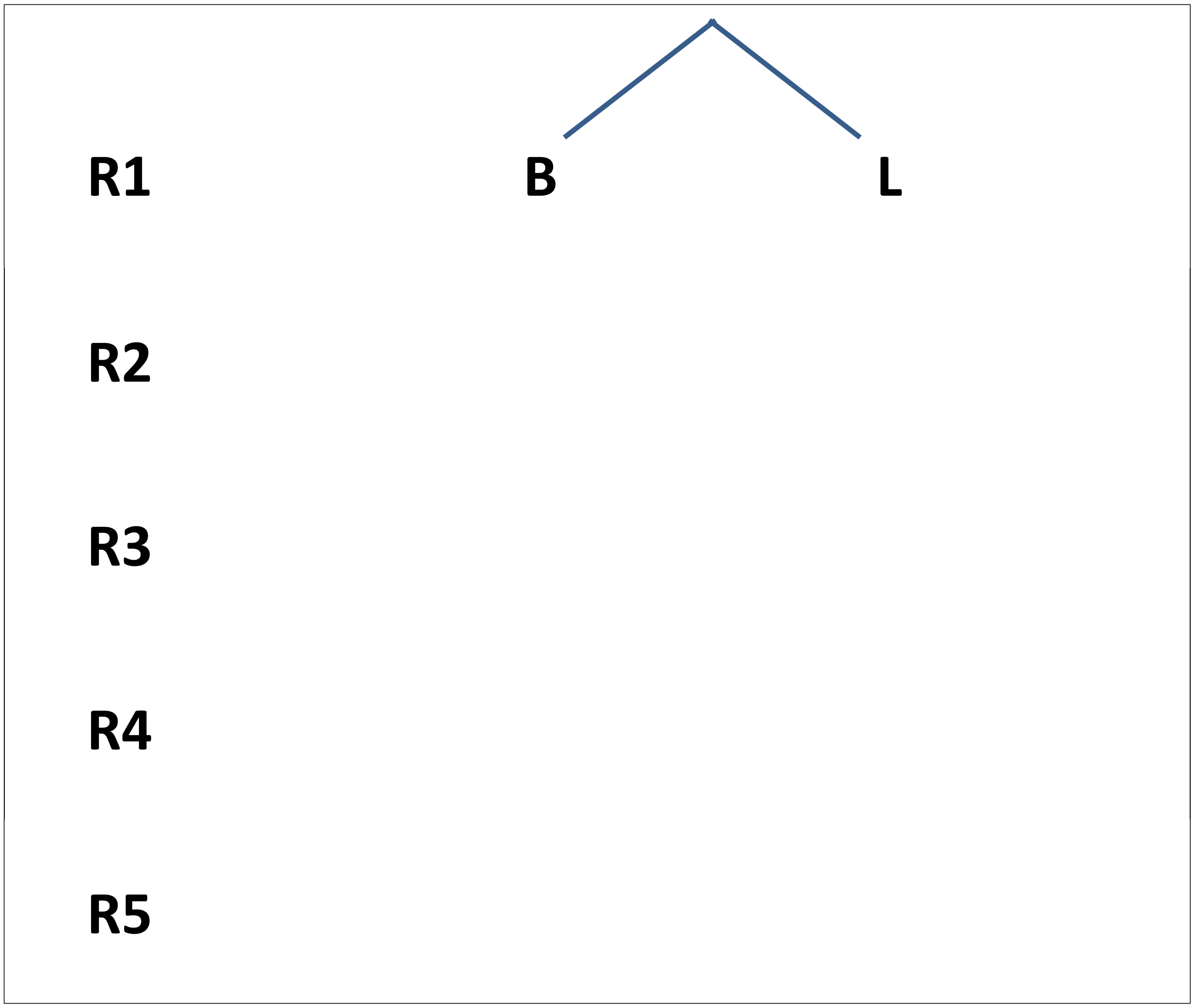
Draw your search tree for **DFS + FC and Propagation through Singleton Domains** here:



	Var assigned or de-queued	List all values just eliminated from neighboring variables	Backtrack?
1			☐
2			☐
3			☐
4			☐
5			☐
6			☐
7			☐
8			☐
9			☐
10			☐

This is a duplicate copy of Part B1 (DFS+FC+PROP-1). If you want to have this copy graded instead, check the box:

☐ **I want to start over; grade this copy**



	Var assigned or de-queued	List all values just eliminated from neighboring variables	Backtrack?
1			☐
2			☐
3			☐
4			☐
5			☐
6			☐
7			☐
8			☐
9			☐
10			☐

B2 (2 points) How many times did you backtrack?

B3 (2 points) In total, how many variable assignments did you undo when backtracking?

Problem 2: Identification Trees (35 points)

Part A: Brexit or Bremain? (35 points)

It is June 2016 and the people of the UK are voting on whether to ***Brexit*** (leave) or ***remain*** in the European Union. Bored with your summer internship, you decide to construct an identification tree to predict how people will vote.

With your outstanding hacking skills, you scrape a database to get census data and recent poll data. The data is shown below, sorted first by **Vote**, then by Annual **Income** (in thousands of £).

Last Name	Vote	Nationality	Income
Longstocking	remain	Scot	19
Oldershaw	remain	Scot	21
McCutherton	remain	English	25
Wellington	remain	Scot	85
Ainsworth	Brexit	Scot	15
Cartwright	Brexit	English	17
Popplewell	Brexit	English	23

(A copy of this data and a table of logarithms are available on the tear-off sheet.)

A1 (7 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 **no logarithms** in it. Space is provided below to show your work.

Feature Test	Disorder
Nationality	
Income > 24	

Show your disorder calculations for partial credit:

$-\left[\frac{1}{2}\log_2\frac{1}{2} + \frac{1}{2}\log_2\frac{1}{2}\right] = 1$

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

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

$-\left[\frac{1}{4}\log_2\frac{1}{4} + \frac{3}{4}\log_2\frac{3}{4}\right] \approx 0.8$

$-\left[\frac{1}{5}\log_2\frac{1}{5} + \frac{4}{5}\log_2\frac{4}{5}\right] \approx 0.72$

$-\left[\frac{1}{6}\log_2\frac{1}{6} + \frac{5}{6}\log_2\frac{5}{6}\right] \approx 0.65$

A2 (20 points) Construct the greedy, disorder-minimizing identification tree to correctly classify all the voters according to their vote. In constructing your identification tree, choose from the following feature tests:

Nationality, Income > 24, Income > 22, Income > 18

Break ties in left-to-right order (first Nationality, then Income > 24, etc.).
For partial credit, clearly show your work in the space provided below.

For credit, draw your ID tree in this box:

For partial credit, show your work here. Additional space is provided on the next page.

Additional space to show work for part A1:

A3 (3 points) You trained your model on individuals, but you want to test whether it will also work for geographic regions. Census data for the Scottish region of Renfrewshire shows that its people have a median annual income of £ 19k:

Region	**Vote**	**Nationality**	**Income**
Renfrewshire	?	Scot	19

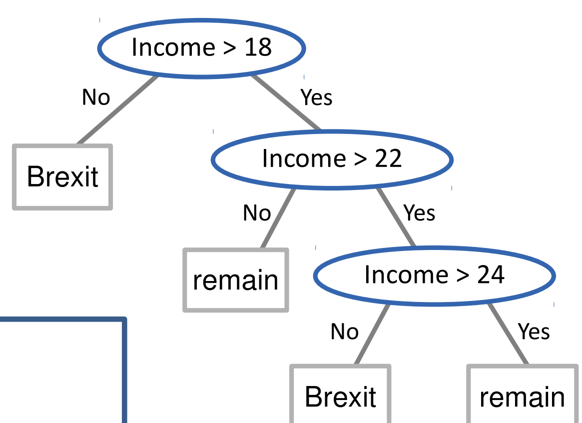
What vote does your ID tree (from part A2) predict for Renfrewshire? (Circle one)

Brexit **remain** **Can't tell**

A4 (2 points) After the election, you learn that the majority of Renfrewshire voted "**remain**". Is this result consistent with your ID tree's prediction? (Circle one)

YES **NO**

A5 (3 points) You read that Annual **Income** is the best indicator for a person's vote. Consider the ID tree on the right, consisting of three **Income** feature tests.



This tree correctly classifies all the training data, but it is still a flawed model for the data. Why?

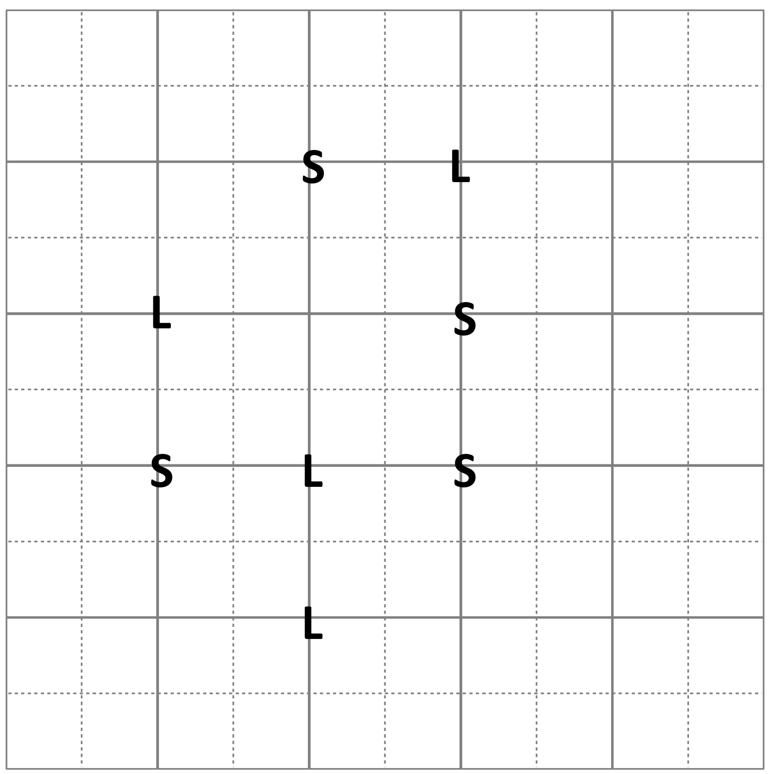
(Write your answer, 10 words or less)

Problem 3: k-Nearest Neighbors (15 points)

Stiwen Jialch finally jumped on the bandwagon and started watching Game of Thrones. He quickly developed a soft corner for the Starks (**S**) and a deep dislike for the Lannisters (**L**). His friend, Kenny Bone, invites him to a viewing party for the Season 7 premiere. Given Stiwen's feelings towards the Lannisters, he does not want to sit next to (closest to) anyone who sympathizes with them.

Kenny gives Stiwen a map (below) of where everyone will be sitting and their family affiliation.

On the graph below, **draw the decision boundaries** produced by **1-Nearest Neighbors**.



See reverse for SRN questions →

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 used electronic circuits to illustrate:
    1. Why it is important to study nth-order linear differential equations.
    2. The power of rule-based expert systems.
    3. How a program can explain its own decisions.
    4. Using analogous systems to understand stellar dynamics.
    5. A way of thinking focused on modeling and abstraction.

2. According to Boyden, light can be used as a reward mechanism for mice because:
    1. Like Pavlov's dogs, mice can be conditioned by presenting food and light at the same time.
    2. Fiber optic implants can be directed toward genetically engineered photoreceptor neurons.
    3. Mice can be “rewired” at birth, redirecting visual stimuli toward reward centers.
    4. Ultraviolet light has a tendency to make all mammals feel happier.
    5. fMRI studies of mice indicate that their reward centers are close to the visual cortex.

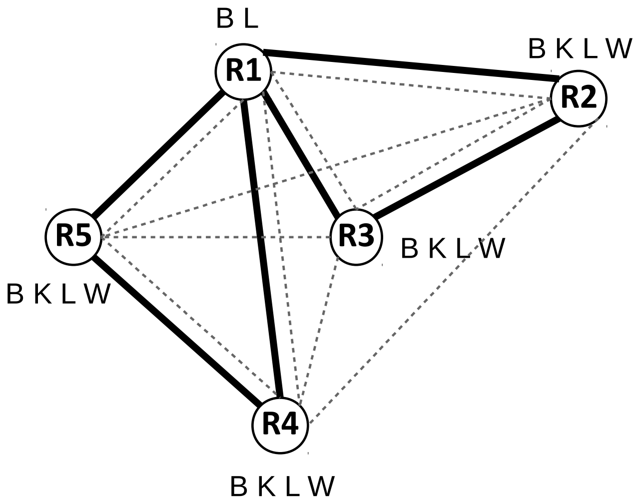
3. In deep neural net terminology, pooling refers to:
    1. A repository used by neural net researchers to share code.
    2. A mechanism that reduces input size by passing along only one value from each group.
    3. A training technique by which one example becomes many via reflection and translation.
    4. The training error due to highly similar test examples, as in breeds of dogs.
    5. The practice of replacing the stairstep function with a sigmoid in back propagation.

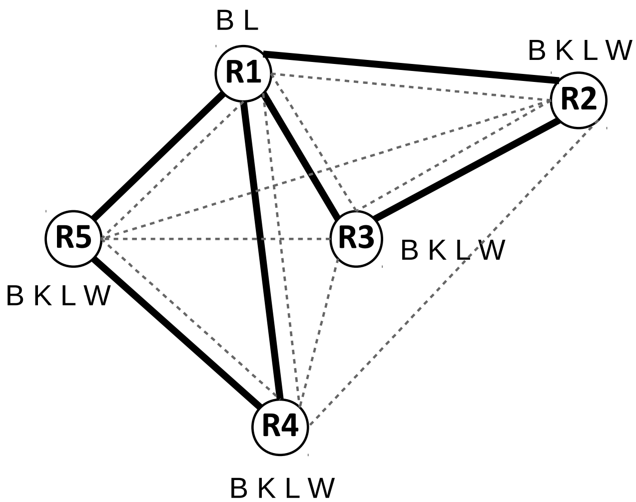
4. According to Winston, the sudden success of deep neural nets is best attributed to:
    1. Increases in computing speed.
    2. Google’s recent invention of the back propagation algorithm.
    3. Improved understanding of animal neurobiology.
    4. Training on small subsets of hand-curated images.
    5. Increasing the use of misclassified images as training progresses.

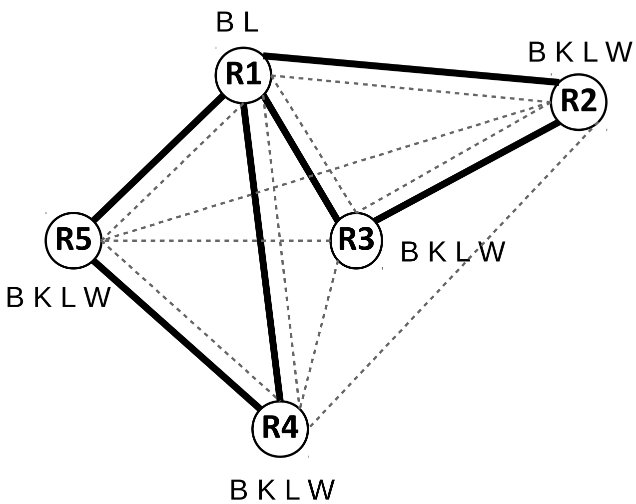
5. Winston claimed that the success of Waltz's drawing understanding program is best attributed to:
    1. Waltz’s superior command of mathematics.
    2. Waltz’s distaste for mathematics and mathematicians.
    3. Joint efforts of people with diverse characteristics.
    4. Increases in computing speed.
    5. Use of a large number of training samples.

6. According to Kanwisher, electrical stimulation of the face area (found using fMRI) in epileptic patients:
    1. Sometimes initiates a seizure.
    2. Causes perceived faces to lose color and appear in sepia tone or grayscale.
    3. Causes perceived faces to seem to speak.
    4. Interferes with the ability to process spoken language.
    5. Causes the perception of faces on inanimate objects.

(Tear-off sheet for Problem 1)







(Tear-off sheet for Problem 2)

Last Name	Vote	Nationality	Income
Longstocking	remain	Scot	19
Oldershaw	remain	Scot	21
McCutherton	remain	English	25
Wellington	remain	Scot	85
Ainsworth	Brexit	Scot	15
Cartwright	Brexit	English	17
Popplewell	Brexit	English	23

$$-\left[\frac{1}{2}\log_2\frac{1}{2} + \frac{1}{2}\log_2\frac{1}{2}\right] = 1$$

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

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

$$-\left[\frac{1}{4}\log_2\frac{1}{4} + \frac{3}{4}\log_2\frac{3}{4}\right] \approx 0.8$$

$$-\left[\frac{1}{5}\log_2\frac{1}{5} + \frac{4}{5}\log_2\frac{4}{5}\right] \approx 0.72$$

$$-\left[\frac{1}{6}\log_2\frac{1}{6} + \frac{5}{6}\log_2\frac{5}{6}\right] \approx 0.65$$