

# 6.034 Quiz 4

2 December 2015

Name	
Email	

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**

**Ben Greenberg**

**Neil Gurram**

**Eeway Hsu**

**Brittney Johnson**

**Veronica Lane**

**Alice Liddell**

**Robert Luo**

**Jessica Noss**

**Mycal Tucker**

**Sarah Vente**

**Jess Wass**

Problem number	Maximum	Score	Grader
1	50		
2	50		
Total	100		

SRN	6		
-----	---	--	--

There are 8 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: Adaboost (50 points)

## Part A: Furthest from $\frac{1}{2}$ (35 points)

This is a boosting question with four training points (A, B, C, D). There are four weak classifiers ( $h_1, h_2, h_3, h_4$ ) that make the following misclassifications:

Classifier	Misclassified training points (A, B, C, D)
$h_1$	A
$h_2$	B
$h_3$	C D
$h_4$	A B C

**A1 (23 points)** Perform three rounds of boosting with these classifiers and training data. In each round, pick the classifier with the error rate furthest from  $\frac{1}{2}$ . Break ties by picking the classifier that comes first in this list:  $h_1, h_2, h_3, h_4$ . Space for scratch work is provided on the following page.

In any round, if Adaboost would terminate instead of choosing a classifier, write "NONE" for the weak classifier ( $h$ ) and for the voting power ( $\alpha$ ). Then, leave all remaining spaces blank.

	Round 1	Round 2	Round 3
weight A	1/4		
weight B			
weight C			
weight D			
Error rate of $h_1$			
Error rate of $h_2$			
Error rate of $h_3$			
Error rate of $h_4$			
weak classifier chosen ( $h$ )			
weak classifier error ( $\epsilon$ )			
voting power ( $\alpha$ )			

Space provided for scratch work:

**A2 (8 points)** After three rounds of boosting, how does the ensemble classifier  $H(x)$  classify each training point? Circle the one best answer in each case. If the answer can't be determined from the available information, circle "Can't tell".

- |    |                      |               |            |
|----|----------------------|---------------|------------|
| A: | Correctly classified | Misclassified | Can't tell |
| B: | Correctly classified | Misclassified | Can't tell |
| C: | Correctly classified | Misclassified | Can't tell |
| D: | Correctly classified | Misclassified | Can't tell |

The following rules for manipulating logarithms may be helpful:

$$\log(x) + \log(y) = \log(x * y)$$

$$\log\left(\frac{1}{x}\right) = -\log(x)$$

$$\log(1) = 0$$

**A3 (4 points)** For each statement below, circle the one best answer (true or false).  
(Hint: consider this problem.)

- |      |       |  |
|------|-------|--|
| TRUE | FALSE | 1. If there are exactly 3 weak classifiers that make non-overlapping errors, it is always possible to use them to construct a perfect ensemble classifier (either by using Adaboost or by picking values of $\alpha$ by hand). |
| TRUE | FALSE | 2. Whenever possible, Adaboost will pick a classifier that makes non-overlapping errors with the previously picked classifiers.  |

### Part B: More True/False (15 points)

This section consists of questions about Adaboost *in general*—they do not rely on the preceding section. Decide whether each of the statements below is true or false. Circle the one best answer in each case.

- |      |       |   |
|------|-------|---|
| TRUE | FALSE | 1. In each round, the error rate $\epsilon$ of each weak classifier $h$ is in the range $0 \leq \epsilon \leq 0.5$ .        |
| TRUE | FALSE | 2. If a weak classifier has an error rate $\epsilon \leq 1/3$ , it can only misclassify up to $1/3$ of the training points. |
| TRUE | FALSE | 3. Given 3 classifiers, a weak classifier can never be picked two rounds in a row.  |
| TRUE | FALSE | 4. The error rate of the ensemble classifier $H(x)$ never increases from one round to the next.                             |
| TRUE | FALSE | 5. Adaboost terminates only if a specific number of rounds is specified.  |

## Problem 2: Bayes Nets

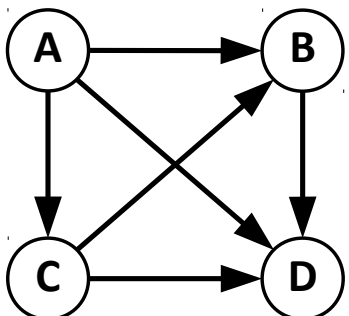
### Part A: Bayes Nets (16 points)

For each Bayes Net described below, answer this question:

**Assuming all of the variables are boolean, how many parameters does the Bayes net have?**

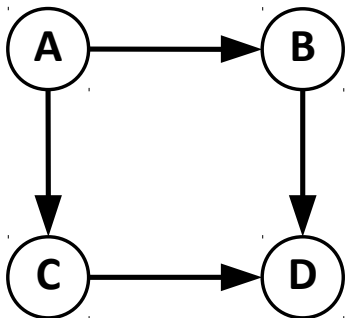
(The number of parameters is the total number of entries in all probability tables.)

**A1 (4 points)**



Number of parameters =

**A2 (4 points)**



Number of parameters =

**A3 (4 points)** A Bayes net with four variables, in which NO variables are assumed to be independent.

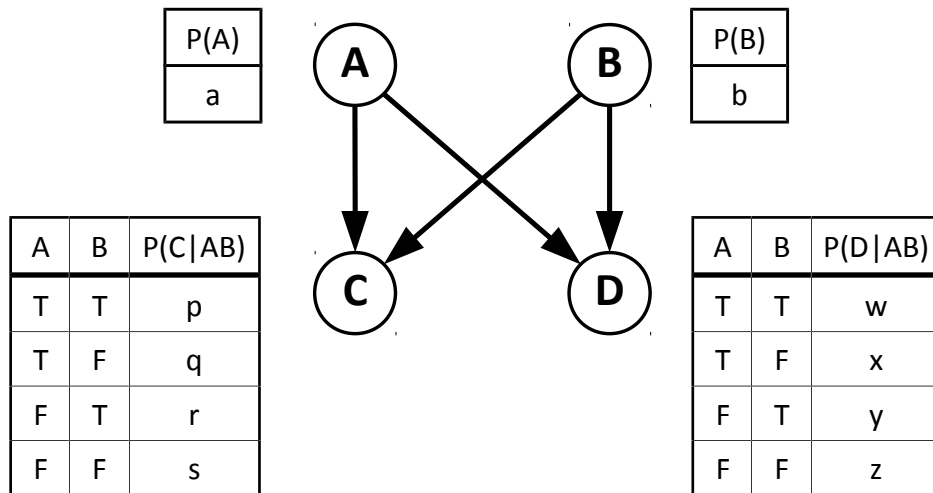
Number of parameters =

**A4 (4 points)** A Bayes net with four variables, in which ALL variables are assumed to be independent.

Number of parameters =

### Part B: Some Probability (18 points)

Here is a Bayes net with 4 boolean variables and their associated probability tables. Each probability is represented by a lowercase variable. Below, show your work for partial credit.

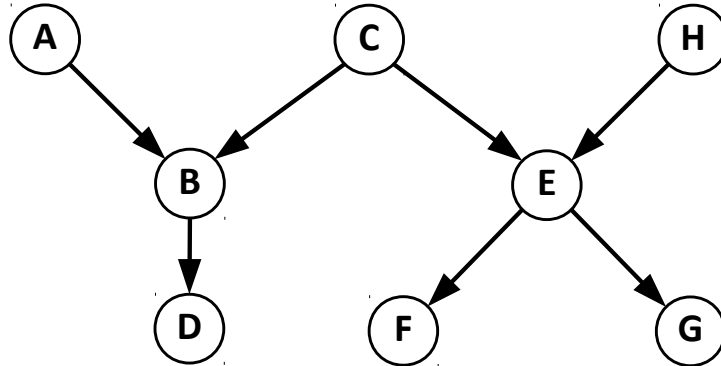


**B1 (8 points)** Write an expression for  $P(\bar{A}B\bar{C}\bar{D})$  in terms of the probabilities specified in the Bayes net.

**B2 (10 points)** Write an expression for  $P(B\bar{C}\bar{D})$  in terms of the probabilities specified in the Bayes net.

### Part C: Independence (16 points)

Here is a Bayes net with 8 variables, which are not necessarily boolean. Assume that the only independence statements that are true are the ones enforced by the shape of the network. Show your work for partial credit.



**C1 (6 points)** List every variable that is marginally independent of A.

**C2 (5 points)** Is  $P(D|C) = P(D|CG)$ ? (Circle one)

YES

NO

**C3 (5 points)** Are A and H conditionally independent, given B and E? (Circle one)

YES

NO

( SRN QUESTIONS ON NEXT PAGE → )

## 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 According to Winston, Minsky believes that we will not really understand intelligence until:

1. We develop provably a universal representation.
2. We develop a provably universal theorem prover.
3. We build robots that operate in the real world.
4. We understand how humans deploy many kinds of reasoning.
5. We develop intelligent machines that can pass the Turing test.

2 Jegelka noted that submodularity has been used with success to:

1. Analyze effects of vaccination programs.
2. Place water-quality sensors.
3. Improve race-horse breeding programs.
4. Program vacuum-cleaner robots.
5. Develop proofs of program correctness.

3 Winston described how the Genesis story-understanding system:

1. Tells stories persuasively by filtering content.
2. Uses Bayesian inference to infer conceptual content.
3. Deploys machine-learning techniques to cluster stories into tragedies and comedies.
4. Authors new stories by combining precedents using a support vector machine.
5. Extends its story database by crawling the web.

4 Sinha explained that a newly sighted adult:

1. Initially sees in red-green, and only later gains the ability to see blue.
2. Can be modeled by a convolutional neural net with few recurrent connections.
3. Tends to see more objects in drawings than a normally sighted person would.
4. Has trouble identifying moving elements in drawings.
5. Exhibits the same kind of errors seen in classifiers based on deep neural net technology.

5 Mansinghka explained how probabilistic programming has been used to:

1. Score SAT essays.
2. Place bets on fantasy sports sites such as Draft Kings.
3. Select answers on 6.034 SRN questions.
4. Guide self-driving cars.
5. Recognize faces in various postures.

6 Katz noted that his START system translates English questions into:

1. Statistical inference models.
2. Sets of ternary expressions.
3. Logical statements in predicate calculus.
4. An inner language based on Esperanto.
5. Transition space diagrams.