Name: ____

DEPARTMENT OF EECS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.02: Digital Communication Systems, Fall 2012

Quiz I

October 11, 2012

"×" your section	Section	Time	Recitation Instructor	TA
☐	1	10-11	Victor Zue	Ruben Madrigal
☐	2	11-12	Victor Zue	Cassandra Xia
☐	3	12-1	Jacob White	Kyu Seob Kim
☐	4	1-2	Yury Polyanskiy	Shao-Lun Huang
☐	5	2-3	Yury Polyanskiy	Rui Hu
☐	6	3-4	Jacob White	Eduardo Sverdlin Lisker

Please read and follow these instructions:

0. Please write your name in the space above and × your section.
1. One two-sided "crib sheet" and a calculator are allowed. No other aids.
2. There are **** questions** in ** sections, and **12 pages** in this quiz booklet.
3. Your total allotted time is **120 minutes**, but we have designed the test to be considerably shorter, to allow you time to work carefully and check your answers.
4. **Please write legibly. Explain your answers, especially when we ask you to.** If you find a question ambiguous, write down your assumptions. Show your work for partial credit.
5. Use the empty sides of this booklet if you need scratch space. *If you use the blank sides for answers, make sure to say so!*

PLEASE NOTE: SOME STUDENTS WILL TAKE THE MAKE-UP QUIZ TOMORROW AND MONDAY AT 9 AM. PLEASE DON'T DISCUSS THIS QUIZ WITH ANYONE IN THE CLASS, UNLESS YOU'RE SURE THEY HAVE TAKEN IT WITH YOU TODAY.

Do not write in the boxes below

1-4 (*/15)	5-7 (*/18)	8 (*/20)	9-15 (*/26)	16-18 (*/21)	Total (*/100)

I Who Said What?

Ali and Bob are communicating on a two-way channel. At every transmission slot, Ali and Bob each independently sends to the other a randomly chosen symbol from a specified set. Ali transmits one of two distinct symbols, a_1 or a_2, with respective probabilities α and $1-\alpha$, while Bob transmits a fixed symbol b_1 (i.e., transmits b_1 with probability 1).

1. [2+2=4 points]: What are the respective entropies, H_{Ali} and H_{Bob}, of Ali's and Bob's transmissions at any slot, in bits? (For H_{Ali}, your answer will be an expression in terms of α, while for H_{Bob} your answer will be a number.)

$H_{\text{Ali}} =$

$H_{\text{Bob}} =$

Suppose Cat is listening in on the channel through a flaky switch that, at each transmission slot, connects her to Ali's transmission with probability p, and to Bob's transmission otherwise, i.e., with probability $(1-p)$. The switch has lights to indicate whether Cat is connected to Ali or to Bob. The fact that Cat is listening has no effect on Ali's and Bob's transmissions.

2. [4 points]: With Ali and Bob transmitting as in Problem 1, Cat can announce one of three possible messages: "Ali transmitted a_1", "Ali transmitted a_2", or "Bob transmitted b_1". What is the entropy, H_{Cat}, of this set of messages?

$H_{\text{Cat}} =$

3. [3 points]: Your friend believes that the general expression relating the entropy H_{Cat} of Cat's possible messages to the entropies H_{Ali} and H_{Bob} of Ali's and Bob's transmissions is

$$H_{\mathrm{Cat}} = pH_{\mathrm{Ali}} + (1-p)H_{\mathrm{Bob}} \ ,$$

as long as Ali's and Bob's transmissions are independent. It turns out that your friend in not quite right. If you've answered Problem 2 correctly, you know that he's missing one or more additional terms on the right that **depend only on** p. Specify what's missing in your friend's expression. **Be sure to simplify your expression** to make clear that it **only depends on** p.

Missing term(s) on right side of preceding equation:

4. [2+2=4 points]: Suppose Cat's switch lights up to tell her she's listening to Ali.

Before seeing what symbol Ali sent, how much information, in bits, has Cat obtained by recognizing that the transmission comes from Ali?

Knowing now that the transmission comes from Ali, how much uncertainty, in bits, does she still have about the actual intercepted symbol?

II Shaking The Tree

In this problem ϵ denotes a small positive number, $0 < \epsilon \ll 1$.

5. [8 points]: Find a Huffman code and its expected code length L for a source whose symbols A, B, C, D have respective probabilities

$$(\frac{4}{11}, \frac{3-\epsilon}{11}, \frac{2+\epsilon}{11}, \frac{2}{11}) .$$

Be sure to draw the code tree!

The binary codewords for A, B, C, D are respectively:

The expected code length $L =$

6. [8+1=9 points]: Repeat the preceding Huffman coding problem for the case where the symbol probabilities are respectively

$$(\frac{4}{11}, \frac{3+\epsilon}{11}, \frac{2-\epsilon}{11}, \frac{2}{11}) \ .$$

Be sure to draw the code tree!

The binary codewords for A, B, C, D are now respectively:

The expected code length now is $L =$

How much shorter is this expected length than what would have been obtained if the code from the previous problem had been used instead?

7. [1 points]: What conclusions are suggested (though of course note proved!) by the preceding calculations, regarding the possible sensitivity of the Huffman code tree and of the expected code length to small perturbations in the symbol probabilities?

III I'm a Webster, You're a Webster

A particular source uses the Lempel-Ziv-Welch algorithm to communicate with a receiver. The message that the source wishes to communicate is made up of just three symbols: a, b, c. These symbols are respectively stored in positions 1, 2 and 3 in the dictionary at both the source and the receiver. The subsequent dictionary entries, as they are built by the LZW algorithm, are assigned to positions numbered 4, 5, 6, ...

8. [5+15=20 points]: Suppose the transmitted sequence is

2, 3, 3, 1, 3, 4, 5, 10, 11, 6, 10, 1

Decode the sequence, and write down the *receiver's entire dictionary* at the end of the transmission.

IV Secret Letter

Members of the Hesperian Order communicate with each other in a binary block code of length 7, with the individual codeword bits sent on consecutive days of the week by messenger, and the color of the messenger's horse — white or black — signaling the bit. The Order makes allowance for the fact that on occasion one of the messengers in a block (and never more than one) is intercepted and induced to swap a white horse for black, or vice versa. The Order's "H-code" is derived by arranging the codeword bits in the form of the letter **H**:

$$\begin{array}{ccc} x_1 & & x_2 \\ x_3 & x_7 & x_4 \\ x_5 & & x_6 \end{array}$$

The set of codewords comprises exactly those words $(x_1, ..., x_7)$ whose GF(2) sum in each of the two vertical columns and in the single horizontal row is 0, i.e.,

$$x_1 + x_3 + x_5 = 0 \ , \qquad x_2 + x_4 + x_6 = 0 \ , \qquad x_3 + x_7 + x_4 = 0 \ .$$

To get full credit, you need to provide careful explanations in addition to correct answers!

9. [5 points]: A vector $\mathbf{c} = (x_1, ..., x_7)$ is a valid codeword if and only if it satisfies the equation $\mathbf{Hc}^T = \mathbf{0}$ for some matrix $\mathbf{H}$. Write down an appropriate matrix $\mathbf{H}$:

H=

10. [2 points]: Is the code linear?

Explanation:

11. [4 points]: What is the rate of the code?

Explanation:

12. [2+1+2=5 points]: What is the minimum weight of a nonzero codeword?

Explanation (**include an example of a codeword with this weight**):

What is the minimum Hamming distance d_{min} of the code?

Explanation:

13. [2+2=4 points]: How many bit errors per block can this code be guaranteed to **detect**?

Explanation:

How many bit errors per block can this code be guaranteed to **correct**?

Explanation:

14. [3 points]: Name a specific block code of the same length and rate, but with better error correcting properties, and specify in what respect the error correcting properties are better.

15. [2+1=3 points]: The Order consults you about the possibility of adding one more bit, x_8, to their codewords, and one more constraint to the existing three constraints that define their H-code. They wish to thereby improve the error correction properties of their code, even if it means getting a somewhat lower rate. The fact that they have two column parity relations and one row parity relation in their existing code reminds you of a code you studied in 6.02. Suggest what *additional parity relation* they should add, and state what *minimum Hamming distance* you expect the resulting code will have (no proof needed).

V Amazing Ariadne

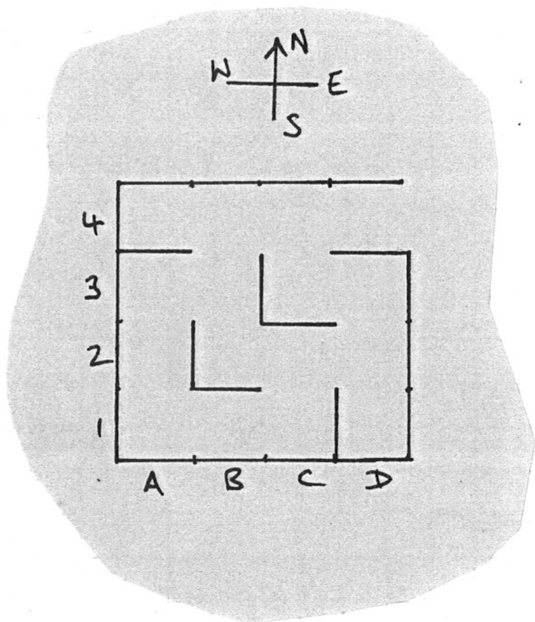


Figure 1: Ariadne's sketch of the maze that Theseus escaped from.

Your archaeologist friend has stumbled on some Cretan figures and inscriptions. She copies the two most intriguing figures, translating their inscriptions, and tells you what she thinks they represent. Figure 1 shows Ariadne's sketch of a 4×4 maze that Theseus managed to escape from, going from location A1 to location D4 in 8 steps (possibly taking steps back and returning to cells he had previously been to), before exiting the maze. Ariadne seems to have asked Theseus to recall the sequence of directions — **N**orth or **E**ast or **S**outh or **W**est — that he took to make his escape. His response is the following, though perhaps not entirely reliable in view of the fact that he was being stalked by the Minotaur:

N E N W N E E E

16. [1 points]: Use a dashed line to draw Theseus's stated path on the 4x4 maze above, and verify that his memory cannot be completely correct.

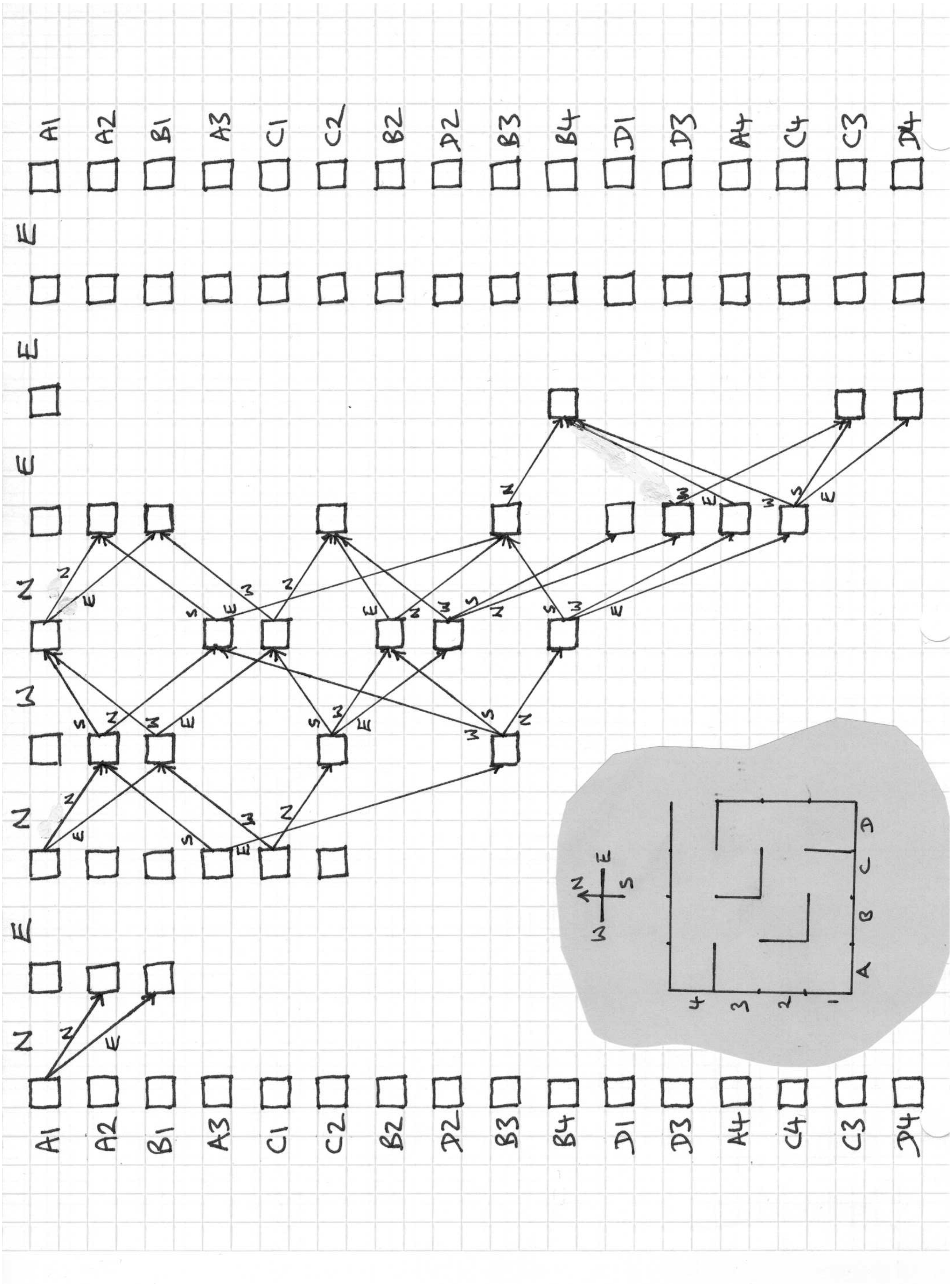
In Figure 2 on the next page you'll see Ariadne's incomplete sketch of the relevant part of a trellis that describes possible 8-step paths through the maze. (A copy of the maze is overlaid there for your convenience.) Your first task — described more specifically in the problem below — is to complete her sketch with whatever states and edges you think are necessary; **avoid drawing states and edges you don't need**, to keep your trellis from looking like a hopeless tangle! You will then run a Viterbi algorithm to search through all possible paths of 8 steps from A1 to D4, finding the one that **matches Theseus's recollected set of directions in as many positions as possible**.

17. [8 points]: Mark in the **relevant** missing states and edges in the trellis in Figure 2 for Theseus's possible 2nd, 7th and 8th steps. Also put in any necessary labels. You **need NOT mark in** those states or edges that you are sure will not be traversed.

18. [10+1+1=12 points]: Run the Viterbi algorithm, filling in the path metric in all the relevant states (i.e., boxes) in the trellis diagram in Figure 2. You need not compute the metric for any boxes that are clearly not going to be needed. Keep track of which edges were used on the optimal path to each relevant state. Then mark in the optimum reconstructed path on the trellis **AND in the maze**. Also summarize your answer by completing the statements below.

The *sequence of directions on the optimum reconstructed path* is as follows:

The **number of positions** in which the directions on the optimum reconstructed path *differ* from the directions on Theseus's recollected path:


A1 A2 B1 A3 C1 C2 B2 D2 B3 B4 D1 D3 A4 C4 C3 D4
N E N W N E E
N E S W
A B C D
1 2 3 4