**Name:** ______________________________________________

DEPARTMENT OF EECS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

## 6.02 Fall 2011

# Quiz I

October 18, 2011

| “×” your section | Section | Time | Room | Recitation Instructor | “×” your TA |
|---|---|---|---|---|---|
| | | | | | ☐ Mukul Agarwal |
| ☐ | 1 | 10-11 | 36-112 | Paul Ampadu | ☐ Jason Cloud |
| ☐ | 2 | 11-12 | 36-112 | Sidhant Misra | ☐ Shuo Deng |
| ☐ | 3 | 12-1 | 36-112 | Sidhant Misra | ☐ Lyla Fischer |
| ☐ | 4 | 1-2 | 36-112 | Paul Ampadu | ☐ Rui Li |
| ☐ | 5 | 2-3 | 26-168 | Karl Berggren | ☐ Ruben Madrigal |
| ☐ | 6 | 3-4 | 26-168 | Karl Berggren | ☐ Surapap R. |
| | | | | | ☐ Xiawa Wang |

There are **14 questions** (some with multiple parts) and **10 pages** in this quiz booklet. Answer each question according to the instructions given. You have **120 minutes** to answer the questions.

If you find a question ambiguous, please be sure to write down any assumptions you make. **Please be neat and legible.** *Please show your work for partial credit.* And please write your name above.

Use the empty sides of this booklet if you need scratch space. You may also use them for answers, although you shouldn't need to. *If you use the blank sides for answers, make sure to say so!*

**One two-sided “crib sheet” and a calculator allowed. No other notes, books, computers, cell phones, PDAs, information appliances, carrier pigeons, etc.!**

**PLEASE NOTE: SEVERAL STUDENTS WILL TAKE THE MAKE-UP QUIZ TOMORROW AT 7.30 PM. UNTIL THEN, PLEASE DO NOT DISCUSS THIS QUIZ WITH ANYONE IN THE CLASS, UNLESS YOU ARE CERTAIN THEY HAVE TAKEN IT WITH YOU TODAY. THANKS!**

*Do not write in the boxes below*

| 1-3 (x/11) | 4-5 (x/14) | 6-7 (x/11) | 8-11 (x/16) | 12-13 (x/28) | 14 (x/20) | Total (x/100) |
|---|---|---|---|---|---|---|
| | | | | | | |

# I Information, Entropy, and Source Coding

Suppose the temperature tomorrow can take on one of $N \geq 3$ possible distinct values, $V_1, V_2, \ldots, V_N$, whose respective probabilities are $p_1 \geq p_2 \ldots \geq p_N > 0$, with $\sum_{i=1}^{N} p_i = 1$. Call this distribution $D$.

**1. [1 points]:** Write the expression for the entropy of $D$ in terms of the quantities defined above.

**2. [4 points]:** The weatherperson, who is never wrong, announces that the temperature tomorrow is guaranteed to be one of three distinct values.

**A.** What set of three values for tomorrow's temperature will convey the **most** information?

**B.** How much information about tomorrow's temperature has the weatherperson given you if the three values are the most informative ones that you specified in Part A? Give **both** the **expression** and the **units** in which this information is expressed.

**3. [6 points]:** Suppose we take the distribution $D$ and combine two values, $V_1$ and $V_2$, into a single "value", $V_{12}$, standing for the event "*the temperature can be either* $V_1$ *or* $V_2$", while keeping all the other $N-2$ values and probabilities the same (note that $N \geq 3$). Call the resulting distribution $D'$.

Which of the following statements is true? Select the best answer from the choices below (circle your choice). You **must** correctly explain your choice in the space below to receive full credit.

**A.** There exists some distribution $D$ for which $D$ has the same entropy as $D'$.

**B.** $D$ always has strictly higher entropy than $D'$.

**C.** There exists some distribution $D$ for which $D$ has lower entropy than $D'$.

**4. [6 points]:** Ben Bitdiddle and Alyssa P. Hacker are taking the 6.02 quiz and encounter the following Huffman coding problem. They are given four symbols, $A$, $B$, $C$, $D$, and the associated symbol probabilities, $p_A \geq p_B \geq p_C \geq p_D$, and are asked to construct the Huffman code tree. They construct two different trees, both correct constructions. The code trees have the following properties:

**Ben's tree:** The length of the longest path from the root to a symbol is 3.

**Alyssa's tree:** The length of the longest path from the root to a symbol is 2.

Derive one constraint relating the symbol probabilities, which *ensures* that both Ben and Alyssa are correct. Write down the constraint, and show Ben's and Alyssa's trees below.

**5. [8 points]:** Circle **True** or **False** for each of these statements about LZW. Briefly explain each answer to receive credit. Recall that a codeword in LZW is an index into the string table.

**A. True / False** Suppose the sender adds two strings with corresponding codewords $c_1$ and $c_2$ in that order to its string table. Then, it **may** transmit $c_2$ for the first time **before** it transmits $c_1$.

**B. True / False** Suppose the string table never gets full. If there is an entry for a string $s$ in the string table, then the sender **must** have previously sent a distinct codeword for every non-null prefix of string $s$. (If $s \equiv p + s'$ where $+$ is the string concatenation operation and $s'$ is some non-null string, then $p$ is said to be a prefix of $s$.)

# II Digital Signaling

**6. [6 points]:** The lab tasks in PSet 2 involved implementing the 8b/10b line coding scheme, which ensured frequent 0/1 and 1/0 transitions of the transmitted bits. The task also used SYNC bits.

**A.** Briefly describe one problem that the frequent 0/1 and 1/0 transitions help to solve.

**B.** Briefly describe one use of the SYNC bits.

**7. [5 points]:** The cable television signal in your home is poor. The receiver in your home is connected to the distribution point outside your home using two coaxial cables in series, as shown in the picture below. The power of the cable signal at the distribution point is $P$. The power of the signal at the receiver is $R$.


The first cable attenuates (i.e., reduces) the signal power by 7 dB. The second cable attenuates the signal power by an **additional** 13 dB. Calculate $\frac{P}{R}$ as a numeric ratio.

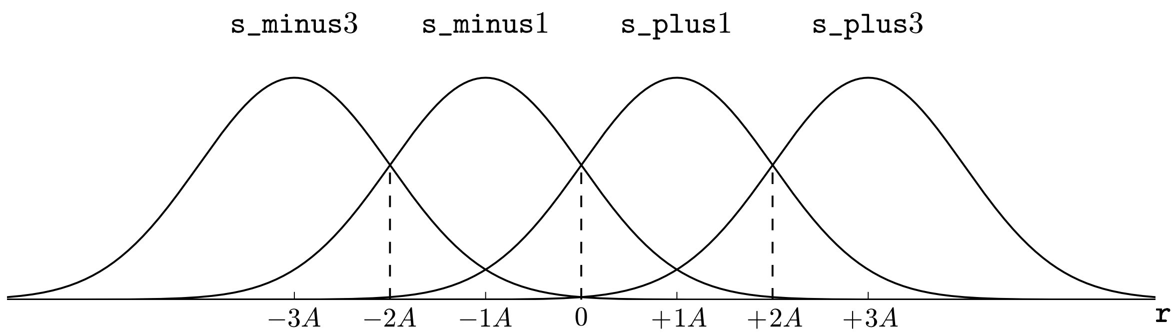
**(Explain your answer in the space below.)**

# III Noise

Ben Bitdiddle studies the bipolar signaling scheme from 6.02 and decides to extend it to a **4-level signaling scheme**, which he calls Ben's Aggressive Signaling Scheme, or **BASS**. In BASS, the transmitter can send four possible signal levels, or voltages: $(-3A, -A, +A, +3A)$, where $A$ is some positive value. To transmit bits, the sender's mapper maps consecutive pairs of bits to a fixed voltage level that is held for some fixed interval of time, creating a **symbol**. For example, we might map bits "00" to $-3A$, "01" to $-A$, "10" to $+A$, and "11" to $+3A$. Each distinct pair of bits corresponds to a unique symbol. Call these symbols s_minus3, s_minus1, s_plus1, and s_plus3. Each symbol has the same prior probability of being transmitted.

The symbols are transmitted over a channel that has no distortion but does have additive noise, and are sampled at the receiver in the usual way. Assume the samples at the receiver are perturbed from their ideal noise-free values by a zero-mean additive white Gaussian noise (AWGN) process with noise intensity $N_0 = 2\sigma^2$, where $\sigma^2$ is the variance of the Gaussian noise on each sample. In the time slot associated with each symbol, the BASS receiver digitizes a selected voltage sample, `r`, and returns an estimate, `s`, of the transmitted symbol in that slot, using the following intuitive digitizing rule (written in Python syntax):

```
def digitize(r):
    if r < -2A: s = s_minus3
    elif r < 0: s = s_minus1
    elif r < 2A: s = s_plus1
    else: s = s_plus3
    return s
```




Ben wants to calculate the **symbol error rate** for BASS, i.e., the probability that the symbol chosen by the receiver was different from the symbol transmitted. Note: we are **not** interested in the **bit** error rate here. Help Ben calculate the symbol error rate by answering the following questions.

**8. [4 points]:** Suppose the sender transmits symbol **s_plus3**. What is the **conditional** symbol error rate given this information; i.e., what is $\mathbb{P}(\text{symbol error} \mid \text{s\_plus3 sent})$? Express your answer in terms of $A$, $N_0$, and the erfc function, defined as $\text{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-x^2}\, dx$.

**(Derive the expression in the space below.)**

**9. [6 points]:** Now suppose the sender transmits symbol **s_plus1**. What is the **conditional** symbol error rate given this information, in terms of $A$, $N_0$, and the erfc function?

**(Derive the expression in the space below.)**

**10. [2 points]:** The conditional symbol error rates for the other two symbols don't need to be calculated separately.

**A.** The symbol error rate when the sender transmits symbol **s_minus3** is the same as the symbol error rate of which of these symbols?

**(Circle ALL that apply.)**

(a) s_minus1.
(b) s_plus1.
(c) s_plus3.

**B.** The symbol error rate when the sender transmits symbol **s_minus1** is the same as the symbol error rate of which of these symbols?

**(Circle ALL that apply.)**

(a) s_minus3.
(b) s_plus1.
(c) s_plus3.

**11. [4 points]:** Combining your answers to questions 8–10, what is the symbol error rate in terms of $A$, $N_0$, and the erfc function? Recall that all symbols are equally likely to be transmitted.

**(Explain your answer in the space below.)**

# IV Linear Block Codes

**12. [1 + 1 + 1 + 2 + 2 + 1 + 2 + 2 + 2 = 14 points]:**
For the following questions, consider an $(n, k, d)$ linear block code in systematic form.

Fill in the blanks.

**A.** Each codeword has a length of ____________________ bits.

**B.** The rate of this code is ____________________.

**C.** The number of parity bits in each codeword is ____________________.

**D.** If the code can correct all single-bit errors, then the number of parity bits **must** be at least ________________________________________.

**E.** Suppose we transmit codewords over a binary symmetric channel with non-zero error probability. Then, the total number of possible **invalid** received words (i.e., those not equal to a valid codeword) is ________________________________________.

**F.** The weight of a codeword is the number of ones in it. The weight of each non-zero codeword is at least ____________________.

**G.** If the code can correct all single-bit errors, but not all 2-bit errors, then the possible value(s) of $d$ is/are ________________________________________.

**H.** If $d$ is odd, then this code can be transformed into a $(n + 1, k, d + 1)$ code by ______________________________________________________________________________.

**I.** Suppose two codewords $c_1$ and $c_2$ from this code have a Hamming distance of 4. If an even parity bit (i.e., a bit equal to the sum of the other bits) is appended to each codeword to produce two new codewords, $c_1'$ and $c_2'$, the Hamming distance between $c_1'$ and $c_2'$ is ____________________.

**13. [6 + 4 + 4 = 14 points]:** Parity Inc. has developed a simple linear block code with three message bits, $D_1, D_2$, and $D_3$, and three parity bits, $P_1, P_2$, and $P_3$. The parity equations are:

$$\begin{aligned} P_1 &= D_1 + D_2 + D_3 \\ P_2 &= D_2 + D_3 \\ P_3 &= D_3 + D_1 \end{aligned}$$

All additions are modulo 2, as usual. Parity Inc. has implemented a maximum-likelihood decoder that is able to correct a certain number of bit errors with this code. Assume that the decoder returns the most-likely message sequence if the number of bit errors is no larger than this correctable number, and returns "uncorrectable errors" otherwise. The sender transmits coded bits in the order $D_1 D_2 D_3 P_1 P_2 P_3$.

**A.** What is the minimum Hamming distance of Parity Inc.'s code?

**(Explain your answer in the space below.)**

**B.** If the receiver gets the word 101101, what will the decoder return?

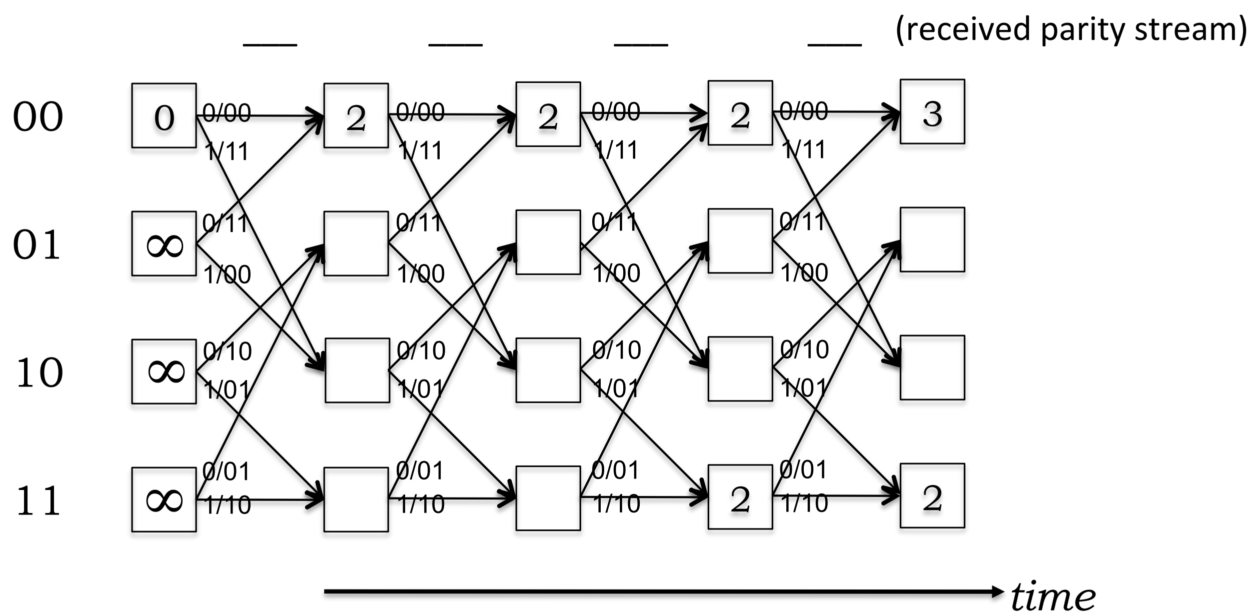
**(Explain your answer in the space below.)**

**C.** If the receiver gets the word 111111, what will the decoder return?

**(Explain your answer in the space below.)**

# V Convolutional Codes

**14. [1 + 1 + 8 + 6 + 4 = 20 points]:** Ben enters the 6.02 lecture hall just as the professor is erasing the convolutional coding with Viterbi decoding example he has worked out, and finds the trellis shown below with several of the numbers and bit strings erased (yeah, the professor is strangely obsessive about how he erases items!). Your job is to help Ben piece together the missing information.




**A.** The constraint length of this convolutional code is ______________.

**B.** The rate of this convolutional code is ______________.

**C.** Assuming hard-decision Viterbi decoding, write the missing path metric numbers inside the boxes in the picture above. (*Work carefully and check your work to avoid a cascade of errors!*)

**D.** What is the **received parity stream**, which the professor has erased? Write your answer in the four blanks shown at the top of the picture above.

**E.** What is the maximum-likelihood message returned by the Viterbi decoder, if we stopped the decoding process at the last stage in the picture shown? Explain your answer below.