

[Show All Answers](#)[Hide All Answers](#)

## 6.02 Practice Problems: Information, Entropy, & Source Coding

---

**Problem 1.** Huffman coding is used to compactly encode the species of fish tagged by a game warden. If 50% of the fish are bass and the rest are evenly divided among 15 other species, how many bits would be used to encode the species when a bass is tagged?

[Show Answer](#)

---

**Problem 2.** Several people at a party are trying to guess a 3-bit binary number. Alice is told that the number is odd; Bob is told that it is not a multiple of 3 (i.e., not 0, 3, or 6); Charlie is told that the number contains exactly two 1's; and Deb is given all three of these clues. How much information (in bits) did each player get about the number?

[Show Answer](#)

---

**Problem 3.** X is an unknown 8-bit binary number. You are given another 8-bit binary number, Y, and told that Y differs from X in exactly one bit position. How many bits of information about X have you been given?

[Show Answer](#)

---

**Problem 4.** In Blackjack the dealer starts by dealing 2 cards each to himself and his opponent: one face down, one face up. After you look at your face-down card, you know a total of three cards. Assuming this was the first hand played from a new deck, how many bits of information do you now have about the dealer's face down card?

[Show Answer](#)

---

**Problem 5.** The following table shows the current undergraduate and MEng enrollments for the School of Engineering (SOE).

Course (Department)	# of students	Prob.
I (Civil & Env.)	121	.07
II (Mech. Eng.)	389	.23
III (Mat. Sci.)	127	.07
VI (EECS)	645	.38
X (Chem. Eng.)	237	.13
XVI (Aero & Astro)	198	.12
Total	1717	1.0

- A. When you learn a randomly chosen SOE student's department you get some number of bits of information. For which student department do you get the *least* amount of information?

[Show Answer](#)

- B. Design a variable length Huffman code that minimizes the average number of bits in messages encoding the departments of randomly chosen groups of students. Show your Huffman tree and give the code for each course.

Show Answer

- C. If your code is used to send messages containing only the encodings of the departments for each student in groups of 100 randomly chosen students, what's the average length of such messages? Write an expression if you wish.

Show Answer

---

**Problem 6.** You're playing an on-line card game that uses a deck of 100 cards containing 3 Aces, 7 Kings, 25 Queens, 31 Jacks and 34 Tens. In each round of the game the cards are shuffled, you make a bet about what type of card will be drawn, then a single card is drawn and the winners are paid off. The drawn card is reinserted into the deck before the next round begins.

- A. How much information do you receive when told that a Queen has been drawn during the current round?

Show Answer

- B. Give a numeric expression for the average information content received when learning about the outcome of a round (aka the entropy).

Show Answer

- C. Construct a variable-length Huffman encoding that minimizes the length of messages that report the outcome of a sequence of rounds. The outcome of a single round is encoded as A (ace), K (king), Q (queen), J (jack) or X (ten). Specify your encoding for each of A, K, Q, J and X.

Show Answer

- D. Using your code from part (C) what is the expected length of a message reporting the outcome of 1000 rounds (i.e., a message that contains 1000 symbols)?

Show Answer

- E. The Nevada Gaming Commission regularly receives messages in which the outcome for each round is encoded using the symbols A, K, Q, J and X. They discover that a large number of messages describing the outcome of 1000 rounds (i.e. messages with 1000 symbols) can be compressed by the LZW algorithm into files each containing 43 bytes in total. They immediately issue an indictment for running a crooked game. Briefly explain their reasoning.

Show Answer

---

### Problem 7.

Consider messages made up entirely of vowels (A, E, I, O, U). Here's a table of probabilities for each of the vowels:

I	p(I)	log <sub>2</sub> (1/p(I))	p(I)*log <sub>2</sub> (1/p(I))
A	0.22	2.18	0.48
E	0.34	1.55	0.53
I	0.17	2.57	0.43
O	0.19	2.40	0.46
U	0.08	3.64	0.29
Totals	1.00	12.34	2.19

A. Give an expression for the number of bits of information you receive when learning that a particular vowel is either I or U.

[Show Answer](#)

B. Using Huffman's algorithm, construct a variable-length code assuming that each vowel is encoded individually. Please draw a diagram of the Huffman tree and give the encoding for each of the vowels.

[Show Answer](#)

C. Using your code from above, give an expression for the expected length in bits of an encoded message transmitting 100 vowels.

[Show Answer](#)

D. Ben Bitdiddle spends all night working on a more complicated encoding algorithm and sends you email claiming that using his code the expected length in bits of an encoded message transmitting 100 vowels is 197 bits. Would you pay good money for his implementation?

[Show Answer](#)

**Problem 8.** Describe the contents of the string table created when encoding a very long string of all a's using the simple version of the LZW encoder shown below. In this example, if the decoder has received E encoded symbols (i.e., string table indices) from the encoder, how many a's has it been able to decode?

```

initialize TABLE[0 to 255] = code for individual bytes
STRING = get input symbol
while there are still input symbols:
    SYMBOL = get input symbol
    if STRING + SYMBOL is in TABLE:
        STRING = STRING + SYMBOL
    else:
        output the code for STRING
        add STRING + SYMBOL to TABLE
        STRING = SYMBOL
output the code for STRING

```

[Show Answer](#)

**Problem 9.** Consider the pseudo-code for the LZW decoder given below:

```

initialize TABLE[0 to 255] = code for individual bytes
CODE = read next code from encoder
STRING = TABLE[CODE]

```

output STRING

```
while there are still codes to receive:
  CODE = read next code from encoder
  if TABLE[CODE] is not defined:
    ENTRY = STRING + STRING[0]
  else:
    ENTRY = TABLE[CODE]
  output ENTRY
  add STRING+ENTRY[0] to TABLE
  STRING = ENTRY
```

Suppose that this decoder has received the following five codes from the LZW encoder (these are the first five codes from a longer compression run):

```
97 -- index of 'a' in the translation table
98 -- index of 'b' in the translation table
257 -- index of second addition to the translation table
256 -- index of first addition to the translation table
258 -- index of third addition to the translation table
```

After it has finished processing the fifth code, what are the entries in TABLE and what is the cumulative output of the decoder?

[Show Answer](#)

---

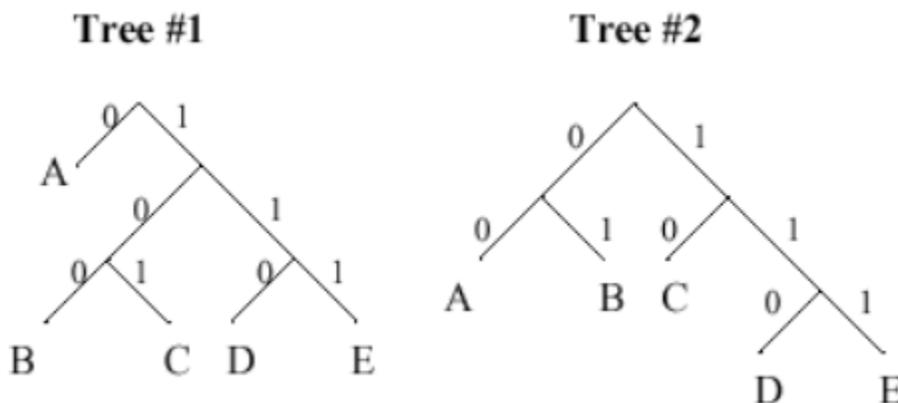
**Problem 10.** Huffman and other coding schemes tend to devote more bits to the coding of

- (A) symbols carrying the most information
- (B) symbols carrying the least information
- (C) symbols that are likely to be repeated consecutively
- (D) symbols containing redundant information

[Show Answer](#)

---

**Problem 11.** Consider the following two Huffman decoding trees for a variable-length code involving 5 symbols: A, B, C, D and E.



A. Using Tree #1, decode the following encoded message: "01000111101".

[Show Answer](#)

B. Suppose we were encoding messages with the following probabilities for each of the 5 symbols:  $p(A) = 0.5$ ,  $p(B) = p(C) = p(D) = p(E) = 0.125$ . Which of the two encodings above (Tree #1 or Tree #2) would yield the shortest encoded messages averaged over many messages?

Show Answer

C. Using the probabilities of part (B), if you learn that the first symbol in a message is "B", how many bits of information have you received?

Show Answer

D. Using the probabilities of part (B), If Tree #2 is used to encode messages what is the average length of 100-symbol messages, averaged over many messages?

Show Answer

**Problem 12.** Ben Bitdiddle has been hired by the Registrar to help redesign the grades database for WebSIS. Ben is told that the database only needs to store one of five possible grades (A, B, C, D, F). A survey of the current WebSIS repository reveals the following probabilities for each grade:

Grade	Probability of occurrence
A	$p(A) = 18\%$
B	$p(B) = 27\%$
C	$p(C) = 25\%$
D	$p(D) = 15\%$
F	$p(E) = 15\%$

A. Given the probabilities above, if you are told that a particular grade is "C", give an expression for the number of bits of information you have received.

Show Answer

B. Ben is interested in storing a sequence of grades using as few bits as possible. Help him out by creating a variable-length encoding that minimizes the average number of bits stored for a sequence of grades. Use the table above to determine how often a particular grade appears in the sequence.

Show Answer

**Problem 13.** Consider a sigma-delta modulator used to convert a particular analog waveform into a sequence of 2-bit values. Building a histogram from the 2-bit values we get the following information:

Modulator value	# of occurrences
00	25875
01	167836
10	167540
11	25974

- A. Using probabilities computed from the histogram, construct a variable-length Huffman code for encoding these four values.

Show Answer

- B. If we transmit the 2-bit values as is, it takes 2 bits to send each value (doh!). If we use the Huffman code from part (A) what is the average number of bits used to send a value? What compression ratio do we achieve by using the Huffman code?

Show Answer

- C. Using Shannon's entropy formula, what is the average information content associated with each of the 2-bit values output by the modulator? How does this compare to the answers for part (B)?

Show Answer

---

**Problem 14.** In honor of Daisuke Matsuzaka's first game pitching for the Redsox, the Boston-based members of the Search for Extraterrestrial Intelligence (SETI) have decided to broadcast a 1,000,000 character message made up of the letters "R", "E", "D", "S", "O", "X". The characters are chosen at random according to the probabilities given in the table below:

Letter	p(Letter)
R	.21
E	.31
D	.11
S	.16
O	.19
X	.02

- A. If you learn that one of the randomly chosen letters is a vowel (i.e., "E" or "O") how many bits of information have you received?

Show Answer

- B. Nervous about the electric bill for the transmitter at Arecibo, the organizers have asked you to design a variable length code that will minimize the number of bits needed to send the message of 1,000,000 randomly chosen characters. Please draw a Huffman decoding tree for your code, clearly labeling each leaf with the appropriate letter and each arc with either a "0" or a "1".

Show Answer

- C. Using your code, what bit string would they send for "REDSOX"?

Show Answer

---

**Problem 15. "Information, please"**

- A. You're given a standard deck of 52 playing cards that you start to turn face up, card by card. So far as

you know, they're in completely random order. How many new bits of information do you get when the first car is flipped over? The fifth card? The last card?

**Show Answer**

B. Suppose there three alternatives ("A", "B" and "C") with the following probabilities of being chosen:

$$p("A") = 0.8$$

$$p("B") = 0.1$$

$$p("C") = 0.1$$

We might encode the of "A" with the bit string "0", the choice of "B" with the bit string "10" and the choice of "C" with the bit string "11".

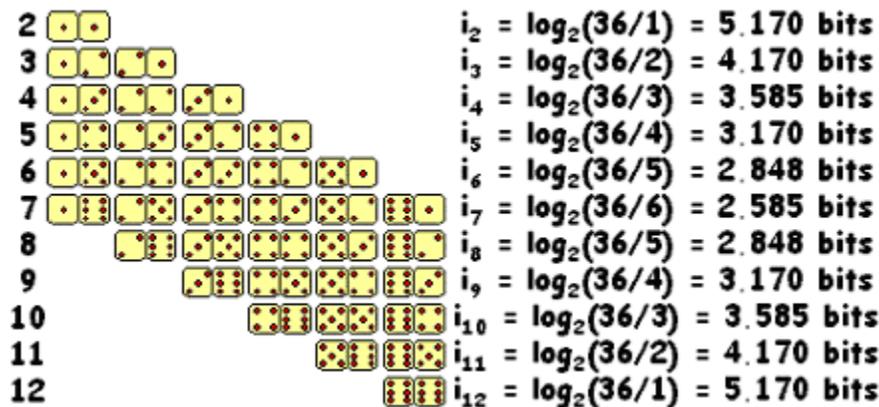
If we record the results of making a sequence of choices by concatenating in left-to-right order the bit strings that encode each choice, what sequence of choices is represented by the bit string "00101001100000"?

**Show Answer**

C. Using the encoding of the previous part, what is the expected length of the bit string that encodes the results of making 1000 choices? What is the length in the worst case? How do these numbers compare with  $1000 \cdot \log_2(3/1)$ , which is the information content of 1000 *equally-probable* choices?

**Show Answer**

D. Consider the sum of two six-sided dice. Even when the dice are "fair" the amount information conveyed by a single sum depends on what the sum is since some sums are more likely than others, as shown in the following figure:



What is the average number of bits of information provided by the sum of 2 dice? Suppose we want to transmit the sums resulting from rolling the dice 1000 times. How many bits should we expect that transmission to take?

**Show Answer**

E. Suppose we want to transmit the sums resulting from rolling the dice 1000 times. If we use 4 bits to encode each sum, we'll need 4000 bits to transmit the result of 1000 rolls. If we use a variable-length binary code which uses shorter sequences to encode more likely sums then the expected number of bits need to encode 1000 sums should be less than 4000. Construct a variable-length encoding for the sum of two dice whose expected number of bits per sum is less than 3.5. (Hint: It's possible to find an

encoding for the sum of two dice with an expected number of bits = 3.306.)

Show Answer

F. Can we make an encoding for transmitting 1000 sums that has an expected length smaller than that achieved by the previous part?

Show Answer

---

### Problem 16.

Consider messages comprised of four symbols -- A, B, C, D -- each with an associated probability of occurrence:  $p(A)$ ,  $p(B)$ ,  $p(C)$ ,  $p(D)$ . Suppose  $p(A) \geq p(B) \geq p(C) \geq p(D)$ . Write down a single condition (equation or inequality) that is both necessary and sufficient to guarantee that the Huffman algorithm will generate a two-bit encoding for each symbol, i.e., the corresponding Huffman code will actually be a fixed-length encoding using 2 bits for each symbol.

Show Answer

---

**Problem 17.** Consider a Huffman code over four symbols: A, B, C, and D. For each of the following encodings indicate if it is a valid Huffman code, i.e., a code that would have been produced by the Huffman algorithm given some set of probabilities  $p(A)$ , ...,  $p(D)$ .

A. A:0, B:11, C:101, D:100

Show Answer

B. A: 1, B:01, C:00, D:010

Show Answer

C. A:00, B:01, C:110, D:111

Show Answer

---

### Problem 18.

After careful data collection, Alyssa P. Hacker observes that the probability of HIGH or LOW traffic on Storow Drive is given by the following table:

Condition	p(HIGH traffic)	p(LOW traffic)
If the Redsox are playing	0.999	.001
If the Redsox are not playing	0.25	0.75

A. If it is known that the Red Sox are playing, then how many bits of information are conveyed by the statement that the traffic level is LOW.

Show Answer

B. Suppose it is known that the Red Sox are *not* playing. What is the entropy of the corresponding

probability distribution of traffic?

Show Answer

---

**Problem 19.**

Consider Huffman coding over four symbols (A, B, C and D) with probabilities  $p(A)=1/3$ ,  $p(B)=1/2$ ,  $p(C)=1/12$  and  $p(D)=1/12$ .

The entropy of the discrete random variable with this probability distribution was calculated to be 1.62581 bits.

We then used the Huffman algorithm to build the following variable length code:

A: 10  
B: 0  
C: 110  
D: 111

which gave an expected encoding length of 1.666 bits/symbol, slightly higher than the entropy bound.

- A. Suppose we made up a new symbol alphabet consisting of all possible pairs of the original four symbols. Enumerate the new symbols and give the probabilities associated with each new symbol.

Show Answer

- B. What is the entropy associated with the discrete random variable that has the probability distribution you gave in part (A). Is it the same, bigger, or smaller than the entropy of 1.626 calculated above? Explain.

Show Answer

- C. Derive the Huffman encoding for the new symbols and compute the expected encoding length expressed in bits/symbol. How does it compare with the 1.666 bits/symbol for the original alphabet? Note: this is tedious to do by hand -- try using the Huffman program you wrote!

Show Answer

MIT OpenCourseWare  
<http://ocw.mit.edu>

6.02 Introduction to EECS II: Digital Communication Systems  
Fall 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.