F. Rubin, "Arithmetic stream coding using fixed-precision registers," *IEEE Transactions on Information Theory,* vol. IT-25, no. 6, pp. 672–675, Nov., 1979.

W.-W. Lu and M. Gough, "A fast-adaptive Huffman coding algorithm," *IEEE Transactions on Communications,* vol. 41, no. 4, pp. 535–538, Apr., 1993.

R. Gallager, "Variations on a theme by Huffman," *IEEE Transactions on Information Theory,* vol. IT-24, no. 6, pp. 668–674, Nov., 1978.

EXERCISES

1.2.1: What is the cardinality for each of the following sets?

a) $A = \{0, 1, 2, 3\}$

b) $A = \{$ cat, dog, mouse $\}$

c) $A = \{$ states in the United States of America $\}$

d) $A = \{x$ such that $2 < x < 40$ and x is a prime number $\}$.

1.2.2: The sets A and B are defined

$$A = \{0, 1, 2, 3, 4\}\ B = \{0, 2, 4, 6, 8\}.$$

Find the cardinality of

a) $A \cup B$ **b)** $A \cap B$.

1.2.3: Let A and B be two sets. The set difference $A\ \ B$ is defined

$$A - B \equiv \{x \in A | x \notin B\}.$$

For find $A = \{a, b, c, d, 0, 1, 2, 3\}$, $B = \{+, /, a, 2, \#\}$ find

a) $A - B$ **b)** $B - A$.

1.2.4: A binary source has symbol probabilities $p_0 = 0.9, p_1 = 0.1$. Find the source entropy.

1.2.5: A binary source has symbol probabilities $p_0 = p, p_1 = 1 - p$. Plot the entropy as a function of p.

1.2.6: A source has an alphabet with $|A| = 8$ and $P_A = \{.25, .20, .15, .12, .10, .08, .05, .05\}$. Find the information efficiency of this source.

1.2.7: Given two information sources with $|A| = 4$, $|B| = 3$. The joint probabilities of symbols from these sources are given in the following table:

	b_0	b_1	b_2
a_0	0.10	0.08	0.13
a_1	0.05	0.03	0.09
a_2	0.05	0.12	0.14
a_3	0.11	0.04	0.06

$\Pr(a_i, b_j)$

Find

a) $H(A)$ **b)** $H(B)$ **c)** $H(A, B)$.

1.2.8: Find the conditional entropies $H(A|B)$, $H(B|A)$ for the sources in Exercise 1.2.7.

1.2.9: Prove: $H(A, B) = H(B, A)$.

1.2.10: Prove: $H(A, B) = H(B) + H(A|B)$.

1.2.11: A discrete memoryless source has a symbol alphabet with $|A| = 10$. Find the upper and lower bounds of its entropy.

1.2.12: The entropy of printed English can be upper bounded by assuming each letter emitted by the source is statistically independent. Find the upper bound on entropy for printed English from the measured probabilities given in the following table:

Letter	Probability	Letter	Probability	Letter	Probability
A	0.0642	J	0.0008	S	0.0514
B	0.0127	K	0.0049	T	0.0796
C	0.0218	L	0.0321	U	0.0228
D	0.0317	M	0.0198	V	0.0083
E	0.1031	N	0.0574	W	0.0175
F	0.0208	O	0.0632	X	0.0013
G	0.0152	P	0.0152	Y	0.0164
H	0.0467	Q	0.0008	Z	0.0005
I	0.0575	R	0.0484	space	0.1859

1.2.13: Explain why your result in Exercise 1.2.12 is an upper bound.

1.3.1: For the 4-ary source and encoder of Example 1.3.1, let $A\{a_0\ a_1\ a_2\ a_3\}$. If the source emits the symbol sequence $a_0\ a_0\ a_1\ a_0\ a_1\ a_0\ a_0\ a_2\ a_3$, what encoded sequence is output by the encoder C ?

1.3.2: Design an algorithm for a decoder for the system of Example 1.3.1. Using your decoder algorithm, decode the sequence 0 1 0 1 1 0 0 0 1 1 0 1 1 1 1 1 0.

1.4.1: Construct the Huffman coding tree for Example 1.3.2.

1.4.2: A discrete memoryless source has an alphabet $\{ a, b, c, d \}$ with symbol probabilities 0.2, 0.4, 0.2, 0.2, respectively.

- **a)** Find the entropy of this source
- **b)** Construct a Huffman code for this source
- **c)** Calculate the efficiency of the code.

1.4.3: Construct a Huffman code for the source in Exercise 1.4.2, which encodes two source symbols at a time. What is the efficiency of this "expanded" code.

1.4.4: Write a computer program to perform Huffman encoding of the English text alphabet given in Exercise 1.2.12, and have your program calculate $\overline{L}$.

1.4.5: Specify a complete ROM lookup table for a Huffman encoder implementation of Example 1.4.2.

1.4.6: (Lab exercise) Design a Huffman encoder circuit for the code in Example 1.3.2. Build and test your design to verify its completeness and correctness.

1.4.7: Specify a ROM lookup table and draw a block diagram for an encoder for Example 1.3.2.

1.5.1: A discrete memoryless source with $A = \{ a, b, c \}$ emits the following string.

$$b\,c\,c\,a\,c\,b\,c\,c\,c\,c\,c\,c\,c\,c\,c\,c\,c\,a\,c\,c\,c\,a$$

Using the Lempel-Ziv algorithm, encode this sequence and find the code dictionary and the transmitted sequence.

1.5.2: A source with $A = \{ a, b, c \}$ is encoded using the Lempel–Ziv algorithm. The transmitted code word sequence is

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

Construct the dictionary and decode this sequence.