

Лабораторная работа №5

«Оптимальное кодирование.

Арифметическое кодирование»

Вариант 1

Цель работы: освоить алгоритм арифметического кодирования и декодирования.
Научиться сжимать сообщения с помощью арифметического кодирования.

1

Условие: Вычислить кумулятивные вероятности каждого символа.

a	A	B	C	D	E	F
p	0,25	0,05	0,1	0,25	0,05	0,3

Решение:

Вычислим кумулятивные вероятности каждого символа по правилу:

$$q(a_1) = 0; \quad q(a_2) = p_1; \quad q(a_3) = p_1 + p_2; \quad \dots; \quad q(a_n) = \sum_{i=1}^{n-1} p_i$$

a	A	B	C	D	E	F
p	0,25	0,05	0,1	0,25	0,05	0,3
q	0	0,25	0,3	0,4	0,65	0,7

2

Условие: Закодировать сообщение (ADADBDCBBDCA BFDAFCEB), используя алгоритм арифметического кодирования. Вычислить числа F_{20} и G_{20} . Вычислить Q и k . Перевести число Q в двоичную систему счисления с точностью k знаков после запятой. Закодированное двоичное число перевести в десятичное число Z .

Решение:

Инициализируем $F_i = 0$ и $G_i = 1$, после чего рассчитаем их пользуясь следующим алгоритмом:

Выполнять с первого символа по n -ный:

$$F_i = F_{i-1} + q(x_i) \times G_{i-1}$$

$$G_i = G_{i-1} \times p(x_i)$$

Составим таблицу, демонстрирующую нахождение F и G:

i	x_i	$p(x_i)$	$q(x_i)$	F	G
0				0	1
1	A	0,25	0	0	0,25
2	D	0,25	0,4	0,1	0,0625
3	A	0,25	0	0,1	0,015625
4	D	0,25	0,4	0,10625	0,00390625
5	B	0,05	0,25	0,1072265625	0,0001953125
6	D	0,25	0,4	0,1073046875	0,000048828125
7	C	0,1	0,3	0,1073193359375	0,0000048828125
8	B	0,05	0,25	0,107320556640625	0,000000244140625
9	B	0,05	0,25	0,10732061767578125	0,00000001220703125
10	D	0,25	0,4	0,10732062255859375	0,0000000030517578125
11	C	0,1	0,3	0,10732062347412109375	0,00000000030517578125
12	A	0,25	0	0,10732062347412109375	0,0000000000762939453125
13	B	0,05	0,25	0,107320623493194580078125	0,000000000003814697265625
14	F	0,3	0,7	0,1073206234958648681640625	0,0000000000011444091796875
15	D	0,25	0,4	0,1073206234963226318359375	0,000000000000286102294921875
16	A	0,25	0	0,1073206234963226318359375	0,00000000000007152557373046875
17	F	0,3	0,7	0,107320623496372699737548828125	0,000000000000021457672119140625
18	C	0,1	0,3	0,1073206234963791370391845703125	0,0000000000000021457672119140625
19	E	0,05	0,65	0,107320623496380531787872314453125	0,000000000000000107288360595703125
20	B	0,05	0,25	0,10732062349638055860996246337890625	0,0000000000000000536441802978515625

$$F_{20} = 0,10732062349638055860996246337890625$$

$$G_{20} = 0,00000000000000000536441802978515625$$

Вычислим Q:

$$Q = F + \frac{G}{2} = 0,10732062349638055860996246337890625 + \frac{0,00000000000000000536441802978515625}{2} = 0,107320623496380561292171478271484375$$

Вычислим k:

$$k = \lceil -\log_2 G \rceil + 1 = \lceil -\log_2 (0,00000000000000000536441802978515625) \rceil + 1 = 59$$

Число Q в двоичной системе счисления с точностью до k-го символа (C):

$$C = 0,00011011011110010101110101001000000110100111000011001100101$$

C в десятичной системе счисления (Z):

$$Z = 0,10732062349638056032719024557309239753521978855133056640625$$

3

Условие: Декодировать полученное число Z в сообщение, используя алгоритм арифметического кодирования. Сравнить декодированное сообщение с изначальным и убедиться в их идентичности.

Решение:

Декодируем полученное число Z в сообщение, используя следующий алфавит и алгоритм:

a	A	B	C	D	E	F
p	0,25	0,05	0,1	0,25	0,05	0,3
q	0	0,25	0,3	0,4	0,65	0,7

Полагая $S = 0$; $G = 1$ выполняем, пока не раскодировем сообщение:

$$j = 1$$

пока $S + q_j \times G < F$ выполняем:

$$j = j + 1$$

$$S = S + q_j \times G$$

$$G = p_j \times G$$

раскодированный символ это x_j

Составим таблицу декодирования для демонстрации алгоритма декодирования:

$$Z = 0,10732062349638056032719024557309239753521978855133056640625$$

i	S	G	$x_i?$	$q(x_i)$	$S + q \times G$	x_i	$p(x_i)$
1	0	1	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	$0 \leq Z$ $0,25 > Z$	A	0,25
2	0	0,25	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	$0 < Z$ $0,0625 < Z$ $0,075 < Z$ $0,1 \leq Z$ $0,1625 > Z$	D	0,25
3	0,1	0,0625	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	$0,1 \leq Z$ $0,115625 > Z$	A	0,25
4	0,1	0,015625	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	$0,1 < Z$ $0,10390625 < Z$ $0,1046875 < Z$ $0,10625 \leq Z$ $0,11015625 > Z$	D	0,25

i	S	G	$x_i?$	$q(x_i)$	$S + q \times G$	x_i	$p(x_i)$
5	0,10625	0,00390625	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,10625 < Z 0,1072265625 ≤ Z 0,107421875 > Z	B	0,05
6	0,1072265625	0,0001953125	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,1072265625 < Z 0,107275390625 < Z 0,10728515625 < Z 0,1073046875 ≤ Z 0,107353515625 > Z	D	0,25
7	0,1073046875	0,000048828125	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,1073046875 < Z 0,10731689453125 < Z 0,1073193359375 ≤ Z 0,10732421875 > Z	C	0,1
8	0,1073193359375	0,0000048828125	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,1073193359375 < Z 0,107320556640625 ≤ Z 0,10732080078125 > Z	B	0,05
9	0,107320556640625	0,000000244140625	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,107320556640625 < Z 0,10732061767578125 ≤ Z 0,1073206298828125 > Z	B	0,05

i	S	G	$x_i?$	$q(x_i)$	$S + q \times G$	x_i	$p(x_i)$
10	0,10732061767578125	0,00000001220703125	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,10732061767578125 < Z 0,1073206207275390625 < Z 0,107320621337890625 < Z 0,10732062255859375 ≤ Z 0,1073206256103515625 > Z	D	0,25
11	0,10732062255859375	0,0000000030517578125	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,10732062255859375 < Z 0,107320623321533203125 < Z 0,10732062347412109375 ≤ Z 0,107320623779296875 > Z	C	0,1
12	0,10732062347412109375	0,00000000030517578125	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,10732062347412109375 ≤ Z 0,1073206235504150390625 > Z	A	0,25
13	0,10732062347412109375	0,0000000000762939453125	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,10732062347412109375 < Z 0,107320623493194580078125 ≤ Z 0,10732062349700927734375 > Z	B	0,25

i	S	G	$x_i?$	$q(x_i)$	$S + q \times G$	x_i	$p(x_i)$
14	0,107320623493194580078125	0,0000000000003814697265625	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,107320623493194580078125 < Z 0,10732062349414825439453125 < Z 0,1073206234943389892578125 < Z 0,107320623494720458984375 < Z 0,10732062349567413330078125 < Z 0,1073206234958648681640625 < Z	F	0,3
15	0,1073206234958648681640625	0,00000000000011444091796875	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,1073206234958648681640625 < Z 0,107320623496150970458984375 < Z 0,10732062349620819091796875 < Z 0,1073206234963226318359375 < Z 0,107320623496608734130859375 > Z	D	0,25
16	0,1073206234963226318359375	0,0000000000000286102294921875	A B C D E F	0 0,25 0,3 0,4 0,65 0,7	0,1073206234963226318359375 < Z 0,10732062349639415740966796875 > Z	A	0,25

i	S	G	$x_i?$	$q(x_i)$	$S + q \times G$	x_i	$p(x_i)$
17	0,107320623496322631835937 5	0,0000000000000071525573730 46875	A	0	0,1073206234963226318359375 < Z	F	0,3
			B	0,25	0,1073206234963405132293701171 875 < Z		
			C	0,3	0,1073206234963440895080566406 25 < Z		
			D	0,4	0,1073206234963512420654296875 < Z		
			E	0,65	0,1073206234963691234588623046 875 < Z		
			F	0,7	0,1073206234963726997375488281 25 < Z		
18	0,107320623496372699737548 828125	0,0000000000000021457672119 140625	A	0	0,1073206234963726997375488281 25 < Z	C	0,1
			B	0,25	0,1073206234963780641555786132 8125 < Z		
			C	0,3	0,1073206234963791370391845703 125 < Z		
			D	0,4	0,1073206234963812828063964843 75 > Z		
			E	0,65			
			F	0,7			

i	S	G	$x_i?$	$q(x_i)$	$S + q \times G$	x_i	$p(x_i)$
19	0,1073206234963791370391845703125	0,00000000000000021457672119140625	A	0	0,1073206234963791370391845703125 < Z	E	0,05
			B	0,25	0,107320623496379673480987548828125 < Z		
			C	0,3	0,10732062349637978076934814453125 < Z		
			D	0,4	0,1073206234963799953460693359375 < Z		
			E	0,65	0,107320623496380531787872314453125 < Z		
			F	0,7	0,10732062349638063907623291015625 > Z		
20	0,107320623496380531787872314453125	0,0000000000000000107288360595703125	A	0	0,107320623496380531787872314453125 < Z	B	0,05
			B	0,25	0,10732062349638055860996246337890625 < Z		
			C	0,3	0,1073206234963805639743804931640625 > Z		
			D	0,4			
			E	0,65			
			F	0,7			

Полученное сообщение – ADADBDCBBDCA BFDAFCEB – идентично исходному (ADADBDCBBDCA BFDAFCEB).

4

Условие: Создать подпрограмму для реализации алгоритма арифметического кодирования. На вход подпрограммы передаются символы и их вероятности, а также сообщение, которое нужно закодировать. Для лучшей реализации алгоритма использовать числа повышенной точности. Подпрограмма возвращает число Z .

5

Условие: Создать подпрограмму для реализации алгоритма арифметического декодирования. На вход подпрограмме передаются символы и их вероятности, а также число Z , которое нужно декодировать. Для лучшей реализации алгоритма использовать числа повышенной точности. Подпрограмма возвращает декодированное сообщение.

6

Условие: Создать демонстрационную программу, показывающую работу подпрограмм из п. 4 и п. 5. Данные из п. 1 и п. 2 использовать как тестовые.