**СОДЕРЖАНИЕ**

Содержание

Аннотация

Введение

Содержание задания

Теоретическая часть

Практическая часть

а) расчеты

б) программа

Заключение

а) результаты работы программы

б) блок-схема

Литература

**АННОТАЦИЯ**

В этой работе по данному числу символов в алфавите рассчитываются их вероятности, количество информации, если символы встречаются с равными вероятностями и с разными вероятностями, недогруженность символов, скорость передачи сообщений и избыточность сообщений. Кроме того, в данной работе строится оптимальный двоичный код по методике Шеннона – Фано. Выполнение этой курсовой работы закрепляет наши знания по дисциплине «Теория информации».

К работе прилагается программа, написанная на языке программирования высокого уровня (Turbo Pascal).

**SUMMARY**

In this work on the given numbeof symbols in the alphabet their probabilities, amount of the information if symbols meet equal probabilities and with different probabilities, speed of message transfer and redundancy of messages pay off. Besides in the given work the optimum binary code by technique of Shennon and Fano is under construction. Performance of this course work fixes our knowledge on discipline «The Theory of the Information».

**ВВЕДЕНИЕ**

Информатика и вычислительная техника – это область науки и техники, которая включает совокупность средств, способов и методов человеческой деятельности, направленных на создание и применение устройств связи, систем сбора, хранения и обработки информации.

Во многих случаях хранимая и передаваемая информация может представлять интерес для лиц, желающих использовать ее в корыстных целях.

Одним из методов защиты является кодирование.

Кодирование – это отображение сообщений кодом по определенному правилу присвоения символов.

Код – это правило, описывающее отображение одного набора знаков в другой набор знаков (или слов). Кодом также называют и множество образов при этом отображении.

Оптимальный код – это наиболее эффективный случай кодирования с нулевой избыточностью. При устранении избыточности существенно снижается количество символов, требуемых для кодируемых сообщений. Вследствие этого уменьшается время передачи, снижается требуемый объем памяти.

Таким образом, знание методов обработки информации является базовым для инженеров, работа которых связана с вычислительными системами и сетями. Избыточность - дополнительные средства, вводимые в систему для повышения ее надежности и защищенности.

Таким образом, информатика занимается изучением обработки и передачи информации.

В работе отражается применение базовых понятий информатики.

**СОДЕРЖАНИЕ ЗАДАНИЯ**

Для проведения расчетов разработать программу на языке ПАСКАЛЬ.

1.1. Число символов алфавита k = m (номер варианта задания) + 10. Определить количество информации на символ сообщения, составленного из этого алфавита:

а) если символы алфавита встречаются с равными вероятностями;

б) если символы алфавита встречаются в сообщении с вероятностями:

р1 = 0,15; p2 = p1/(k-1); p3 = (p1 + p2 )/(k-2) ...

k-1

pk = ∑ pn /(k – k + 1).

n=1

Сумма всех вероятностей должна быть равой единице, поэтому:

pi

рi = -----

k

∑ pj

j=1

Определить, насколько недогружены символы во втором слу­чае.

1.2. Число символов алфавита = m (номер варианта задания). Вероятности появления символов равны соответственно

р1 = 0,15; p2 = p1/(k-1); p3 = (p1 + p2 )/(k-2) ...

k-1

pk = ∑ pn /(k – k + 1).

n=1

Длительности символов τ1 = 1 сек; τ2 = 2 сек;

τk = τk-1 + 1.

Чему равна скорость передачи сообщений, составленных из таких символов?

Определить количество информации на символ сообщения, составленного из этого алфавита:

а) если символы алфавита встречаются с равными вероятностями;

Определить, насколько недогружены символы во втором случае.

1.3. Сообщения составляются из алфавита с числом символов = m. Вероятность появления символов алфавита равна соответственно:

р1 = 0,15; p2 = p1/(k-1); p3 = (p1 + p2 )/(k-2) ...

k-1

pk = ∑ pn /(k – k + 1).

n=1

Найти избыточность сообщений, составленных из данного алфавита.

Построить оптимальный код сообщения.

**ТЕОРЕТИЧЕСКАЯ ЧАСТЬ**

**КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ**

Общее число неповторяющихся сообщений, которое может быть составлено из алфавита m путем комбинирования по n символов в сообщении,

N = mn

Неопределенность, приходящаяся на символ первичного (кодируемого) алфавита, составленного из равновероятных и взаимонезависимых символов,

H = log2 m

Так как информация есть неопределенность, снимаемая при получении сообщения, то количество информации может быть представлено как произведение общего числа сообщений k на среднюю энтропию H, приходящуюся на одно сообщение:

I = k\*H бит

Для случаев равновероятных и взаимонезависимых символов первичного алфавита количество информации в k сообщениях алфавита m равно:

I = k\*log2 m бит

Для неравновероятных алфавитов энтропия на символ алфавита:

m m

H =∑ pi\*log2(1/2pi)=-∑pi\*log2pi бит/символ

i=1 i=1

А количество информации в сообщении, составленном из k неравновероятных символов,

m

I = -k\*∑ pi\*log2pi бит

i=1

**ВЫЧИСЛЕНИЕ СКОРОСТИ ПЕРЕДАЧИ ИНФОРМАЦИИ И ПРОПУСКНОЙ СПОСОБНОСТИ КАНАЛОВ СВЯЗИ**

В условиях отсутствия помех скорость передачи информации определяется количеством информации, переносимым символом сообщения в единицу времени, и равна

C = n\*H,

где n - количество символов, вырабатываемых источником сообщений за единицу времени; H - энтропия (неопределенность), снимаемая при получении одного символа сообщений, вырабатываемых данным источником.

Скорость передачи информации также может быть представлена как

 бит/сек,

где тау - время передачи одного двоичного символа.

Для сообщений, составленных из равновероятных взаимонезависимых символов равной длительности, скорость передачи информации:

C=(1/τ)\*log2 m бит/сек

В случае неравновероятных символов равной длительности:

m

C =(1/τ)\*∑pi\*log2pi бит/сек

i=1

В случае неравновероятных и взаимонезависимых символов разной длительности:



Пропускная способность (или емкость канала связи) – есть максимальная скорость передачи информации по данному каналу связи. Под каналом связи подразумевается совокупность средств, предназначенных для передачи информации от данного источника сообщений к адресату. Выражение для пропускной способности отличается тем, что пропускную способность характеризует максимальная энтропия:

Смакс= бит/сек

Для двоичного кода:

Смакс бит/сек

При наличии помех пропускная способность канала связи вычисляется как произведение количества принятых в секунду знаков n на разность энтропии источника сообщений и условной энтропии источника сообщений относительно принятого сигнала:

 бит/сек (15)

или

 бит/сек

В общем случае



 бит/сек (16)

Если символы источника сообщений неравновероятны и взаи­мозависимы, то энтропия источника считается по формуле общей условной энтропии.

Для симметричных бинарных каналов, в которых сигналы передаются при помощи двух качественных признаков и вероятность ложного приема , а вероятность правильного приема , потери учитываются при помо­щи условной энтропии вида

 бит/сек (17)

пропускная способность таких каналов

 бит/сек (18)

Для симметричного бинарного канала



 бит/сек (19)

Для симметричных дискретных каналов связи с числом качест­венных признаков m > 2 пропускная способность

 бит/сек (20)

**ОПРЕДЕЛЕНИЕ ИЗБЫТОЧНОСТИ СООБЩЕНИЙ. ОПТИМАЛЬНОЕ КОДИРОВАНИЕ**

Если энтропия источника сообщений не равна максимальной энтропии для алфавита с данным количеством качественных признаков (имеются в виду качественные признаки алфавита, при помощи которых составляются сообщения), то это, прежде всего, означает, что сообщения данного источника могли бы нести большее количество информации. Абсолютная недогруженность на символ сообщений такого источника:

∆D=(Нмакс-Н) бит/символ

Для определения количества "лишней" информации, которая заложена в структуре алфавита либо в природе кода, вводится понятие избыточности. Избыточность, с которой мы имеем дело в теории информации, не зависит от содержания сообщения и обычно заранее известна из статистических данных. Информационная избыточность показывает относительную недогруженность на символ алфавита и является безразмерной величиной:

,

где = μ - коэффициент сжатия (относительная энтропия). Н и Нмакс берутся относительно одного и того же алфавита.

Кроме общего понятия избыточности существуют частные виды избыточности (избыточность, обусловленная неравновероятным распределением символов в сообщении, избыточность, вызванная статистической связью между символами сообщения).

Избыточность, которая заложена в природе данного кода, получается в результате неравномерного распределения в сообщениях качественных признаков этого кода и не может быть задана одной цифрой на основании статистических испытаний. Так при передаче десятичных цифр двоичным кодом максимально загруженными бывают только те символы вторичного алфавита, которые передают значения, являющиеся целочисленными степенями двойки. В остальных случаях тем же количеством символов может быть передано большее количество цифр (сообщений). Например, тремя двоичными разрядами мы можем передать и цифру 5, и цифру 8. Фактически для передачи сообщения достаточно иметь длину кодовой комбинации.

Фактически для передачи сообщения достаточно иметь длину кодовой комбинации



где N - общее количество передаваемых сообщений.

L можно представить и как



где  и  - соответственно качественные признаки первичного и вторичного алфавитов. Поэтому для цифры 5 в двоичном коде можно записать

 дв. симв.

Однако эту цифру необходимо округлить до ближайшего целого числа (в большую сторону), так как длина кода не может быть выражена дробным числом.

В общем случае, избыточность от округления:



где , k - округленное до ближайшего целого числа значение . Для нашего примера



Избыточность необходима для повышения помехоустойчивости кодов и ее вводят искусственно в виде добавочных  символов. Если в коде всего n разрядов и  из них несут информационную нагрузку, то  характеризуют абсолютную корректирующую избыточность, а величина  характеризует относительную корректирующую избыточность.

Для уменьшения избыточности используют оптимальные коды. При построении оптимальных кодов наибольшее распространение получили методики Шеннона-Фано и Хаффмена. Согласно методике Шеннона-Фано построение оптимального кода ансамбля из сообщений сводится к следующему:

1) множество из сообщений располагается в порядке убывания вероятностей;

2) первоначальный ансамбль кодируемых сигналов разбивается на две группы таким образом, чтобы суммарные вероятности сообщений обеих групп были по возможности равны.

Если равной вероятности в подгруппах нельзя достичь, то их делят так, чтобы в верхней части (верхней подгруппе) оставались символы, суммарная вероятность которых меньше суммарной вероятности символов в нижней части (нижней подгруппе);

3) первой группе присваивается символ 0, а второй группе - символ 1;

4) каждую из образованных подгрупп делят на две части таким образом, чтобы суммарные вероятности вновь образованных подгрупп были по возможности равны;

5) первым группам каждой из подгрупп вновь присваивается 0, а вторым - 1. Таким образом, мы получаем вторые цифры кода. Затем каждая из четырех групп вновь делится на равные (с точки зрения суммарной вероятности) части до тех пор, пока в каждой из подгрупп не останется по одной букве.

Построенный код называют оптимальным неравномерным кодом (ОНК).

**ПРАКТИЧЕСКАЯ ЧАСТЬ**

**a) Расчеты**

1. рассчитывается первоначальные вероятности для неравновероятных символов алфавита.
2. выполняет нормирование указанных вероятностей.
3. рассчитывается энтропия алфавита из равновероятных символов.
4. производится расчет энтропии алфавита с неравновероятными символами и недогруженность в этом случае.
5. с учетом заданных длительностей символов вычисляется скорость передачи и избыточность.
6. строится оптимальный код по методу Шеннона-Фано.

Расчет вероятностей.

|  |  |
| --- | --- |
| Промежуточные значения:  k-1  ...pk = S pn /(m - k + 1).  n-1 | Окончательный результат:  рi = pi/(pi) |
| p1 = 0,1500  p2 = 0,0065  p3 = 0,0071  p4 = 0,0078  p5 = 0,0086  p6 = 0,0095  p7 = 0,0105  p8 = 0,0118  p9 = 0,0132  p10 = 0,0150  p11 = 0,0171  p12 = 0,0198  p13 = 0,0231  p14 = 0,0273  p15 = 0,0327  p16 = 0,0400  p17 = 0,0500  p18 = 0,0643  p19 = 0,0857  p20 = 0,1200  p21 = 0,1800  p22 = 0,3000  p23 = 0,6000  p24 = 1,8000  рi = 3,6 | p1=0,0417  p2=0,0018  p3=0,0020  p4=0,0022  p5=0,0024  p6=0,0026  p7=0,0029  p8=0,0033  p9=0,0037  p10=0,0042  p11=0,0048  p12=0,0055  p13=0,0064  p14=0,0076  p15=0,0091  p16=0,0111  p17=0,0139  p18=0,0179  p19=0,0238  p20=0,0333  p21=0,0500  p22=0,0833  p23=0,1667  p24=0,5000  рi = 1 |

Определение количества информации на символ сообщения, составленного из данного алфавита.

Количество информации на символ сообщения для символов данного алфавита, встречающихся с равными вероятностями:

Hmax = log2 24 = ln 24/ln 2 = 4,5850 бит/символ

Количество информации на символ сообщения для символов данного алфавита, встречающихся в сообщении с разными вероятностями:

H = – (0,0417\*log20,0417 + 0,0018\*log20,0018 + 0,020\*log2 0,0020 + 0,0022\*log20,0022 + 0,0024\*log20,0024 + 0,0026\*log20,0026 + 0,0029\*log20,0029 + 0,0033\*log20,0033 + 0,0037\*log20,0037 + 0,0042\*log20,0042 + 0,0048\*log20,0048 + 0,0055\*log20,0055 + 0,0064\*log20,0064 + 0,0076\*log20,0076 + 0,0091\*log20,0091 + 0,0111\*log20,0111 + 0,0139\*log20,0139 + 0,0179\*log20,0179 + 0,0238\*log20,0238 + 0,0333\*log20,0333 + 0,0500\*log20,0500 + 0,0833\*log20,0833 + 0,1667\*log20,1667 + 0,5000\*log20,5000) =

= 2,6409 бит/символ

Недогруженность символов в данном случае:

N = Нmax – Н = 4,5850 – 2,6409 = 1,9441 бит/символ

Вычисление скорости передачи информации.

С= – (0,0417\*log20,0417 + 0,0018\*log20,0018 + 0,020\*log2 0,0020 + 0,0022\*log20,0022 + 0,0024\*log20,0024 + 0,0026\*log20,0026 + 0,0029\*log20,0029 + 0,0033\*log20,0033 + 0,0037\*log20,0037 + 0,0042\*log20,0042 + 0,0048\*log20,0048 + 0,0055\*log20,0055 + 0,0064\*log20,0064 + 0,0076\*log20,0076 + 0,0091\*log20,0091 + 0,0111\*log20,0111 + 0,0139\*log20,0139 + 0,0179\*log20,0179 + 0,0238\*log20,0238 + 0,0333\*log20,0333 + 0,0500\*log20,0500 + 0,0833\*log20,0833 + 0,1667\*log20,1667 + 0,5000\*log20,5000) /

(1\*0,0417 + 2\*0,0018 + 3\*0,020 + 4\*0,0022 + 5\*0,0024 + 6\*0,0026 + 7\*0,0029 + 8\*0,0033 + 9\*0,0037 + 10\*0,0042 + 11\*0,0048 + 12\*0,0055 + 13\*0,0064 + 14\*0,0076 + 15\*0,0091 + 16\*0,0111 + 17\*0,0139 + 18\*0,0179 + 19\*0,0238 + 20\*0,0333 + 21\*0,0500 + 22\*0,0833 + 23\*0,1667 + 24\*0,5000) = 0,1244 бит/сек

Избыточность сообщений, составленных из данного алфавита.

D = 1 – (Н/Нmax) = 1 – (2,6409 / 4,5850) = 0,4240

Построение оптимального кода

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | p24=0,5000 | 0,5 | 0 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 0 |
| 2 | p23=0,1667 | 0,5 | 1 | 0,25 | 1 | 0,1666 | 1 |  |  |  |  |  |  |  |  |  |  |  |  | 111 |
| 3 | p22=0,0833 |  | 1 |  | 1 | 0,0833 | 0 |  |  |  |  |  |  |  |  |  |  |  |  | 110 |
| 4 | p21=0,0500 |  | 1 | 0,25 | 0 |  | 0 | 0,05 | 1 0 |  |  |  |  |  |  |  |  |  |  | 1000 |
| 5 | p1=0,0417 |  | 1 |  | 0 |  | 0 | 0,0690 | 1 | 0,0357 | 1 |  |  |  |  |  |  |  |  | 10011 |
| 6 | p20=0,0333 |  | 1 |  | 0 | 0,1190 | 0 |  | 1 | 0,0333 | 0 |  |  |  |  |  |  |  |  | 10010 |
| 7 | p19=0,0238 |  | 1 |  | 0 |  | 1 |  | 1 | 0,0428 | 1 | 0,0178 | 1 |  |  |  |  |  |  | 101111 |
| 8 | p18=0,0179 |  | 1 |  | 0 |  | 1 |  | 1 |  | 1 | 0,025 | 0 | 0,0138 | 0 |  |  |  |  | 1011100 |
| 9 | p17=0,0139 |  | 1 |  | 0 |  | 1 |  | 1 |  | 0 | 0,025 | 1 |  |  |  |  |  |  | 101101 |
| 10 | p16=0,0111 |  | 1 |  | 0 |  | 1 | 0,0666 | 1 |  | 1 |  | 0 |  |  |  |  |  |  | 101110 |
| 11 | p15=0,0091 |  | 1 |  | 0 |  | 1 | 0,0642 | 0 |  | 0 |  | 1 | 0,0090 | 1 |  |  |  |  | 1010011 |
| 12 | p14=0,0076 |  | 1 |  | 0 |  | 1 |  | 0 |  | 0 |  | 1 | 0,0102 | 0 | 0,0054 | 0 |  |  | 10100100 |
| 13 | p13=0,0064 |  | 1 |  | 0 |  | 1 |  | 0 |  | 0 | 0,0166 | 0 | 0,0064 | 1 |  |  |  |  | 1010001 |
| 14 | p12=0,0055 |  | 1 |  | 0 |  | 1 |  | 0 |  | 0 | 0,0166 | 1 | 0,0064 | 1 |  |  |  |  | 1010011 |
| 15 | p11=0,0048 |  | 1 |  | 0 |  | 1 |  | 0 | 0,0333 | 1 |  | 1 |  | 1 | 0,0047 | 1 |  |  | 10101111 |
| 16 | p10=0,0042 |  | 1 |  | 0 |  | 1 |  | 0 |  | 1 |  | 1 | 0,0088 | 1 |  | 0 | 0,0032 | 0 | 101011100 |
| 17 | p9=0,0037 |  | 1 |  | 0 |  | 1 |  | 0 |  | 1 |  | 1 | 0,0078 | 0 | 0,0036 | 1 |  |  | 10101101 |
| 18 | p8=0,0033 |  | 1 |  | 0 |  | 1 |  | 0 |  | 1 |  | 1 | 0,0078 | 1 | 0,0036 | 0 |  |  | 10101110 |
| 19 | p7=0,0029 |  | 1 |  | 0 |  | 1 |  | 0 |  | 1 |  | 0 |  | 1 |  | 0 |  |  | 10101010 |
| 20 | p6=0,0026 |  | 1 |  | 0 |  | 1 |  | 0 |  | 1 | 0,0167 | 0 |  | 1 | 0,0026 | 1 | 0,0026 | 1 | 101010111 |
| 21 | p5=0,0024 |  | 1 |  | 0 |  | 1 |  | 0 |  | 1 | 0,0147 | 0 |  | 1 |  | 1 | 0,0024 | 0 | 101010110 |
| 22 | p4=0,0022 |  | 1 |  | 0 |  | 1 |  | 0 |  | 1 |  | 0 |  | 0 | 0,0022 | 0 |  |  | 10101000 |
| 23 | p3=0,0020 |  | 1 |  | 0 |  | 1 |  | 0 |  | 1 |  | 0 |  | 0 | 0,0038 | 1 | 0,0020 | 1 | 101010011 |
| 24 | p2=0,0018 |  | 1 |  | 0 |  | 1 |  | 0 |  | 1 |  | 0 | 0,0083 | 0 |  | 1 | 0,0018 | 0 | 101010010 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Буква | Вероятность появления буквы | Кодовые слова | Число знаков в кодовом слове | Pi·li |
| A[1] (p24) | 0,5000 | 0 | 1 | 0,5 |
| A[2] (p23) | 0,1667 | 111 | 3 | 0,50001 |
| A[3] (p22) | 0,0833 | 110 | 3 | 0,2500 |
| A[4] (p21) | 0,0500 | 1000 | 4 | 0,2000 |
| A[5] (p 1) | 0,0417 | 10011 | 5 | 0,2083 |
| A[6] (p20) | 0,0333 | 10010 | 5 | 0,1667 |
| A[7] (p19) | 0,0238 | 101111 | 6 | 0,1429 |
| A[8] (p18) | 0,0179 | 1011100 | 7 | 0,1250 |
| A[9] (p17) | 0,0139 | 101101 | 6 | 0,0833 |
| A[10] (p16) | 0,0111 | 101110 | 6 | 0,0667 |
| A[11] (p15) | 0,0091 | 1010011 | 7 | 0,0636 |
| A[12] (p14) | 0,0076 | 10100100 | 8 | 0,0606 |
| A[13] (p13) | 0,0064 | 1010001 | 7 | 0,0449 |
| A[14] (p12) | 0,0055 | 1010011 | 7 | 0,0385 |
| A[15] (p11) | 0,0048 | 10101111 | 8 | 0,0381 |
| A[16] (p10) | 0,0042 | 101011100 | 9 | 0,0375 |
| A[17] (p9) | 0,0037 | 10101101 | 8 | 0,0294 |
| A[18] (p8) | 0,0033 | 10101110 | 8 | 0,0261 |
| A[19] (p7) | 0,0029 | 10101010 | 8 | 0,0234 |
| A[20] (p6) | 0,0026 | 101010111 | 9 | 0,0237 |
| A[21] (p5) | 0,0024 | 101010110 | 9 | 0,0214 |
| A[22] (p4) | 0,0022 | 10101000 | 8 | 0,0173 |
| A[23] (p3) | 0,0020 | 101010011 | 9 | 0,0178 |
| A[24] (p2) | 0,0018 | 101010010 | 9 | 0,0163 |

**Определение количества информации на символ сообщения. Построение оптимального кода.**

С начало множество из сообщений расположим в порядке убывания вероятностей. Затем, разобьем данное множество на две группы таким образом, чтобы суммарные вероятности сообщений обеих групп были по возможности равны. Но поскольку равенство не достигается, то мы их делим так, чтобы в верхней части оставались символы, суммарная вероятность которых меньше суммарной вероятности символов в нижней части. Первой группе присваиваем символ 0, а второй группе = символ 1. каждую из образованных подгрупп делим на две части таким образом, чтобы суммарные вероятности вновь образованных подгрупп были по возможности равны. Первым группам каждой из подгрупп вновь присваиваем 0, а вторым 1. таким образам мы получаем мы получаем вторые цифры кода. Затем каждую из четырех групп вновь делим на равные части до тех пор, пока в каждой из подгрупп не останется по одной букве.

Оптимальный код (получившийся результат):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Буква | Вероятность  появления буквы | Кодовое слово | Число знаков в кодовом слове | pi∙ li |
| P1 | 0,055 | 000 | 3 | 0,165 |
| P2 | 0,053 | 0010 | 4 | 0,212 |
| P3 | 0,051 | 00110 | 5 | 0,255 |
| P4 | 0,050 | 00111 | 5 | 0,250 |
| P5 | 0,048 | 0100 | 4 | 0,192 |
| P6 | 0,046 | 0101 | 4 | 0,176 |
| P7 | 0,044 | 0110 | 4 | 0,114 |
| P8 | 0,043 | 01110 | 5 | 0,215 |
| P9 | 0,041 | 011110 | 6 | 0,246 |
| P10 | 0,040 | 011111 | 6 | 0,240 |
| P11 | 0,039 | 1000 | 4 | 0,156 |
| P12 | 0,038 | 10010 | 5 | 0,190 |
| P13 | 0,036 | 10011 | 5 | 0,180 |
| P14 | 0,035 | 1010 | 4 | 0,140 |
| P15 | 0,033 | 10110 | 5 | 0,165 |
| P16 | 0,032 | 101110 | 6 | 0,192 |
| P17 | 0,030 | 101111 | 6 | 0,180 |
| P18 | 0,029 | 11000 | 5 | 0,145 |
| P19 | 0,027 | 11001 | 5 | 0,135 |
| P20 | 0,026 | 11010 | 5 | 0,130 |
| P21 | 0,025 | 110110 | 6 | 0,150 |
| P22 | 0,023 | 110111 | 6 | 0,138 |
| P23 | 0,022 | 11100 | 5 | 0,110 |
| P24 | 0,020 | 111010 | 6 | 0,120 |
| P25 | 0,019 | 111011 | 6 | 0,114 |
| P26 | 0,018 | 111100 | 6 | 0,108 |
| P27 | 0,017 | 111101 | 6 | 0,102 |
| P28 | 0,016 | 111110 | 6 | 0,096 |
| P29 | 0,013 | 1111110 | 7 | 0,091 |
| P30 | 0,012 | 11111110 | 8 | 0,096 |
| P31 | 0,010 | 11111111 | 8 | 0,080 |

Ручное построение ОНК по методике Шеннона-Фано:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| P1 | 0,010 | 11111111 | 0,520 | 0,277 | 0,147 | 0,086 | 0,051 | 0,035 | 0,022 | 0,010 |
| P2 | 0,012 | 11111110 | 0,012 |
| P3 | 0,013 | 1111110 | 0,013 |  |
| P4 | 0,016 | 111110 | 0,016 |  |  |
| P5 | 0,017 | 111101 | 0,035 | 0,017 |  |  |
| P6 | 0,018 | 111100 | 0,018 |  |  |
| P7 | 0,019 | 111011 | 0,061 | 0,039 | 0,019 |  |  |
| P8 | 0,020 | 111010 | 0,020 |  |  |
| P9 | 0,022 | 11100 | 0,022 |  |  |  |
| P10 | 0,023 | 110111 | 0,130 | 0,074 | 0,048 | 0,023 |  |  |
| P11 | 0,025 | 110110 | 0,025 |  |  |
| P12 | 0,026 | 11010 | 0,026 |  |  |  |
| P13 | 0,027 | 11001 | 0,056 | 0,027 |  |  |  |
| P14 | 0,029 | 11000 | 0,029 |  |  |  |
| P15 | 0,030 | 101111 | 0,243 | 0,130 | 0,095 | 0,062 | 0,030 |  |  |
| P16 | 0,032 | 101110 | 0,032 |  |  |
| P17 | 0,033 | 10110 | 0,033 |  |  |  |
| P18 | 0,035 | 1010 | 0,035 |  |  |  |  |
| P19 | 0,036 | 10011 | 0,113 | 0,074 | 0,036 |  |  |  |
| P20 | 0,038 | 10010 | 0,038 |  |  |  |
| P21 | 0,039 | 1000 | 0,039 |  |  |  |  |
| P22 | 0,040 | 011111 | 0,471 | 0,262 | 0,168 | 0,124 | 0,081 | 0,040 |  |  |
| P23 | 0,041 | 011110 | 0,041 |  |  |
| P24 | 0,043 | 01110 | 0,043 |  |  |  |
| P25 | 0,044 | 0110 | 0,044 |  |  |  |  |
| P26 | 0,046 | 0101 | 0,094 | 0,046 |  |  |  |  |
| P27 | 0,048 | 0100 | 0,048 |  |  |  |  |
| P28 | 0,050 | 00111 | 0,209 | 0,154 | 0,101 | 0,050 |  |  |  |
| P29 | 0,051 | 00110 | 0,051 |  |  |  |
| P30 | 0,053 | 0010 | 0,053 |  |  |  |  |
| P31 | 0,055 | 000 | 0,055 |  |  |  |  |  |

**ТЕКСТ ПРОГРАММЫ:**

uses Crt,Graph;

const k=24;

type

mass=array [1..k] of real;

var

i,x: integer;

s,H,Hmax,deltaD,sum,C,sumTiPi,D: real;

p,a: mass;

code: array [1..k] of string[20];

{Процедура построения оптимального кода по методике Шеннона-Фано}

procedure shannona(b:mass);

procedure divide(var nS:integer; n1,n2:integer);

var

s,s1,s2: real;

begin

s:=0;

for i:=n1 to n2 do s:=s+a[i];

s1:=0; s2:=0;

i:=n1-1;

repeat

inc(i);

s1:=s1+a[i];

s2:=s1+a[i+1];

until abs(s/2-s1)<abs(s/2-s2);

nS:=i;

for x:=n1 to nS do

if (s/2-s1)>=0 then code[x]:=code[x]+'0'

else code[x]:=code[x]+'1';

for x:=nS+1 to n2 do

if (s/2-s1)<0 then code[x]:=code[x]+'0'

else code[x]:=code[x]+'1';

end;

var

tmp: real;

j,n1,n2,nS: integer;

begin

for i:=1 to k do code[i]:='';

for i:=1 to k do a[i]:=b[i];

for i:=1 to k do

for j:=k downto(i+1) do

if a[i]<a[j]

then

begin

tmp:=a[i];

a[i]:=a[j];

a[j]:=tmp;

end;

j:=1;

repeat

divide(nS,j,k);

n1:=nS;

while (nS-j)>0 do divide(nS,j,nS);

j:=nS+1;

n2:=n1;

while (n1-j)>0 do divide(n1,j,n1);

j:=n2+1;

until j>(k-1);

end;

procedure zastavka;

var dr,reg,err:integer;

begin

dr:=detect;reg:=detect;

initgraph(dr,reg,'d:\tp7\tpu\');

err:=graphresult;

if err<>grok then

begin

writeln('Ошибка инициализации графического модуля!');

halt;

end;

setcolor(19);

settextstyle(3,0,4);

outtextxy(150,40,'Расчетно-графическая работа');

outtextxy(240,65,'на тему');

setcolor(14);

settextstyle(4,0,4);

outtextxy(50,125,'''Построение оптимального кода методом Шеннона-Фано''');

settextstyle(6,0,2);

setcolor(19);

outtextxy(325,250,'Выполнил:');

settextstyle(6,0,2);

setcolor(10);

outtextxy(400,250,'Калинин С.А. ПС-11');

outtextxy(200,450,'Нажмите любую клавишу');

readln;

closegraph;

end;

procedure vivod;

begin

textcolor(lightgreen);

writeln('Оптимальный код: '); {вывод конечной таблицы}

writeln('Символ':7,'Вероятность':13,'Оптимальный код':20,'Число зн.':15,'Вероятн.\*Числ.зн.':20);

for i:=1 to k do

begin

write(' p[',i:2,'] ');

write(p[i]:0:4,' ');

write(code[i]:20,' ');

write(length(code[i]):15,' ');

write((p[i]\*length(code[i])):0:4);

if i<>k then writeln;

end;

end;

begin

zastavka;

clrscr;

{1.1 а) Кол-во информации на символ сообщения,

составленного из алфавита равновероятных символов}

Hmax:=ln(k)/ln(2);

{1.1 б) Расчет вероятностей для неравновероятных символов}

p[1]:=0.15;

sum:=p[1];

for i:=2 to k do

begin

p[i]:=sum/(k+1-i);

sum:=sum+p[i];

end;

clrscr;

textcolor(11);

writeln('Промежуточные значения вероятностей: ');

writeln;

textcolor(10);

for i:=1 to 14 do

writeln('Вероятность p[',i:2,'] = ',p[i]:4:4);

readkey;

clrscr;

textcolor(11);

writeln('Промежуточные значения вероятностей: ');

writeln;

textcolor(10);

for i:=15 to k do

writeln('Вероятность p[',i:2,'] = ',p[i]:4:4);

writeln;

textcolor(11);

for i:=1 to k do s:=s+p[i];

writeln('Сумма вероятностей = ',s:4:2);

readkey;

H:=0;

sumTiPi:=0;

for i:=1 to k do

begin

p[i]:=p[i]/sum;

{1.1 б) Расчет энтропии для алфавита неравновероятных символов}

H:=H+p[i]\*(ln(1/p[i])/ln(2));

sumTiPi:=sumTiPi+i\*p[i];

end;

clrscr;

textcolor(11);

writeln('Окончательные значения: ');

writeln;

textcolor(10);

for i:=1 to 14 do

writeln('Вероятность p[',i:2,'] = ',p[i]:4:4);

readkey;

clrscr;

textcolor(11);

writeln('Окончательные значения: ');

writeln;

textcolor(10);

for i:=15 to k do

writeln('Вероятность p[',i:2,'] = ',p[i]:4:4);

writeln;

textcolor(11);

s:=0;

for i:=1 to k do s:=s+p[i];

writeln('Сумма вероятностей = ',s:4:2);

readkey;

{1.1 б) Определение недогруженности символов}

deltaD:=Hmax-H;

{1.2 Расчет скорости передачи сообщения}

C:=H/SumTiPi;

{1.3 Расчет избыточности сообщений}

D:=1-H/Hmax;

{Вызов процедуры построения оптимального кода}

shannona(p);

{Вывод результатов}

clrscr;

textcolor(11);

{ writceln('Количество символов алфавита = ',k,'.');}

writeln('1.1 Количество информации на символ сообщения:');

writeln(' a) для алфавита равновероятных символов: ');

textcolor(10); writeln(' Hmax =',Hmax:7:4,' бит/символ');

textcolor(11); writeln(' b) для алфавита неравновероятных символов: ');

textcolor(10); writeln(' H =',H:7:4,' бит/символ');

textcolor(11); write(' Недогруженность:');

textcolor(10); writeln(' дельтаD =',deltaD:7:4,' бит/символ');

textcolor(11); writeln;

Writeln('1.2 Скорость передачи информации:');

textcolor(10); writeln(' C =',C:7:4,' бит/сек');

textcolor(11); writeln;

Writeln('1.3 Избыточность сообщений:');

textcolor(10); writeln(' D =',D:7:4);

writeln;

TextColor(11);

write(' Нажмите любую клавишу для вывода таблицы резултатов построения.');

readkey;

clrScr;

vivod;

readkey;

end.**Заключение:**

В моей курсовой работе я использовал теоретический материал и разработанную на языке (высокого уровня) Turbo Pascal программу. Мною было рассчитано количество информации на символ сообщения, составленного из алфавита, состоящего из 24 символа, для двух случаев:

1] если символы алфавита встречаются с равными вероятностями;

2] если вероятности не равны.

Также я определил количество недогрузки символов во втором случае, вычислил количество информации на символ сообщения и скорость передачи сообщений, составленных из таких символов, нашел избыточность сообщений, составленных из данного алфавита. Построил оптимальный код сообщения, применяя методику Шеннона-Фано: при помощи последовательного деления множества вероятностей на группы по принципу равенства сумм вероятностей я составил в соответствие каждому символу наиболее оптимальную двоичную комбинацию. Таким образом, был получен оптимальный двоичный код для алфавита из 31 символа.

В результате выполнения работы были получены следующие результаты:

* количество информации на символ для равновероятного алфавита – 4,585 бит/сим;
* количество информации на символ для неравновероятного алфавита - 2,6409 бит/сим;
* недогруженность символов – 1,9441 бит/сим;
* скорость передачи информации – 0,1244 бит/сек;
* избыточность сообщения – 0,4240;
* построен следующий оптимальный код:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Символ | Вероятность  появления | Код | Число знаков |  |
| p[ 1] | 0.0417 | 0 |  |  |
| p[ 2] | 0.0018 | 111 |  |  |
| p[ 3] | 0.0020 | 110 |  |  |
| p[ 4] | 0.0022 | 1000 |  |  |
| p[ 5] | 0.0024 | 10011 |  |  |
| p[ 6] | 0.0026 | 10010 |  |  |
| p[ 7] | 0.0029 | 101111 |  |  |
| p[ 8] | 0.0033 | 1011100 |  |  |
| p[ 9] | 0.0037 | 101101 |  |  |
| p[10] | 0.0042 | 101101 |  |  |
| p[11] | 0.0048 | 1010011 |  |  |
| p[12] | 0.0055 | 10100100 |  |  |
| p[13] | 0.0064 | 1010001 |  |  |
| p[14] | 0.0076 | 1010001 |  |  |
| p[15] | 0.0091 | 10101111 |  |  |
| p[16] | 0.0111 | 101011100 |  |  |
| p[17] | 0,0139 | 10101101 |  |  |
| p[18] | 0,0179 | 10101101 |  |  |
| p[19] | 0,0238 | 10101010 |  |  |
| p[20] | 0,0333 | 101010111 |  |  |
| p[21] | 0,0500 | 101010110 |  |  |
| p[22] | 0,0833 | 10101000 |  |  |
| p[23] | 0,1667 | 101010011 |  |  |
| p[24] | 0,5000 | 101010010 |  |  |

**СПИСОК ИСПОЛЬЗУЕМОЙ ЛИТЕРАТУРЫ:**

1. Бауэр Ф. Информатика, М. 1992.

2. Колесник В.Д. Курс теории информации, М. 1982.

3. Фаронов В. В. Turbo Pascal 7.0. Учебное пособие, М. 2000.

4. Цымбаль В.П. Задачник по теории информации и кодированию, Киев. 1976.

5. Марченко А.И. Программирование в среде Turbo Pascal 7.0.