Учреждение образования

«Белорусский государственный технологический университет»

**Защита информации и надежность информационных систем**

**Лабораторная работа №2**

ЭЛЕМЕНТЫ ТЕОРИИ ИНФОРМАЦИИ.

ПАРАМЕТРЫ И ХАРАКТЕРИСТИКИ

ДИСКРЕТНЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

Выполнил:

Студент 3 курса 2 группы ФИТ

Максимова Вера Владимировна

Проверил:

Ржеутская Надежда Викентьевна

**2022 г.**

**1. Закрепить теоретические знания по основам теории инфор-**

**мации.**

Передача информации (данных) осуществляется между двумя

абонентами, называемыми источником сообщения (ИcС) и получа-

телем сообщения (ПС). Третьим элементом информационной сис-

темы является канал (среда) передачи, связывающий ИсС и ПС.

Отметим также, что и в системах с хранением информации

всегда можно выделить ИcС и ПС. В данном случае каналом пере-

дачи здесь выступает устройство хранения информации (память).

Например, при записи данных в ОЗУ (оперативное запоминающее

устройство) компьютера в качестве ИcС и ПС может выступать про-

цессор (соответственно при записи и чтении данных).

Таким образом, простейшая информационная система состоит

из трех элементов: источника сообщения, канала передачи сооб-

щения и получателя сообщения.

Отображение сообщения обеспечивается изменением какой-

либо физической величины, характеризующей процесс (например,

амплитуда, частота, фаза). Эта величина является информационным

параметром сигнала (в общем случае – информационной системы).

Сигналы, как и сообщения, могут быть непрерывными и дис-

кретными. Информационный параметр непрерывного сигнала с

течением времени может принимать любые мгновенные значения

в определенных пределах. Непрерывный сигнал часто называют

аналоговым, а каналы и устройства, функционирующие на основе

такого типа сигналов, – аналоговыми.

Дискретный сигнал (устройство или канал передачи) характе-

ризуется конечным числом значений информационного параметра.

Дискретные сообщения состоят из последовательности дис-

кретных знаков. Часто этот параметр принимает всего два значе-

ния (0 или 1). Сообщение или канал его передачи на основе этих

двух значений сигнала называют двоичным или бинарным.

Построение сигнала по определенным правилам, обеспечи-

вающим соответствие между сообщением и сигналом, называют

кодированием.

Кодирование в широком смысле – преобразование сообщения

в сигнал.

Кодирование в узком смысле – представление исходных зна-

ков, называемых символами, в другом алфавите с меньшим чис-

лом знаков. Оно осуществляется с целью повышения надежности

и преобразования сигналов к виду, удобному для передачи по ка-

налам связи. Последний тип кодирования относится к так назы-

ваемой прикладной теории кодирования информации, занимаю-

щейся поиском и реализацией методов и средств обнаружения не-

соответствий (ошибок) между переданным Xk и принятым Yk со-

общениями.

Рассмотрим основные характеристики и параметры двоичных

систем.

Важнейшая характеристика источника, получателя или канала –

алфавит.

Алфавит, А – это общее число знаков или символов (N), ис-

пользуемых для генерации или передачи сообщений. Символы

алфавита будем обозначать через {аi}, где 1 ≤ i ≤ N; N – мощность

алфавита.

Минимальное число элементов алфавита Nmin = 2, А = {0, 1} –

двоичный код. Один дискретный знак представляет собой элемен-

тарное сообщение, последовательность знаков – сообщение.

Набор элементов алфавита, создаваемых дискретным источ-

ником сообщений, заранее, априори (до опыта) известен получа-

Элементы теории информации. Параметры и характеристики дискретных систем

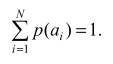
телю. ИсС в каждый дискретный момент времени выдает один

элемент алфавита. Этот элемент сообщения является одним из

символов алфавита. Понятно, что ПС заранее не известно, какой

это элемент. Если обозначить вероятность выбора каждого эле-

мента алфавита p(аi), то



Вероятности p(аi) могут быть получены в результате анализа

частотных свойств символов алфавита, если на входе такого ана-

лизатора принять документ на основе соответствующего алфави-

та. Причем объем документа должен быть таким, чтобы от часто-

сти (частоты) появления каждого символа в анализируемом доку-

менте можно было перейти к вероятности соответствующего со-

бытия. Можно предположить, что указанному требованию будет

соответствовать объем электронного документа не менее несколь-

ких десятков килобайт.

Двоичный канал передачи информации строится на основе

двоичного алфавита: А = {0, 1}. При этом канал, в котором веро-

ятности искажения переданного 0 (принята соответственно 1; это-

му событию соответствует условная вероятность р(1|0)) и пере-

данной 1 (принят соответственно 0; этому событию соответствует

условная вероятность р(0|1)) равны, как и равны вероятности пе-

редачи 0 (р(0)) и 1 (р(1)), называют двоичным симметричным ка-

налом (ДСК).

В общем случае, если передается сообщение Хk = х1, х2, …, хk, а

принимается сообщение Yk = y1, y2, …, yk, то данные условные веро-

ятности можно рассматривать с двух точек зрения: p(хi|yj) и p(yj|хi).

На рисунке далее представлен ДСК.

На рисунке обозначены: q – вероятность правильной (безоши-

бочной) передачи бита сообщения, р – вероятность передачи бита

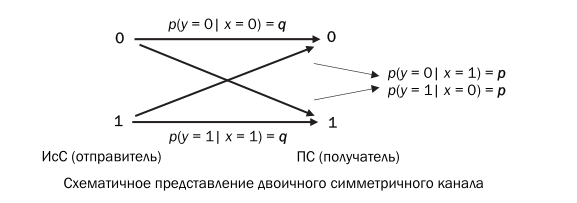
с ошибкой. Понятно, что p + q = 1.

Информационной характеристикой алфавита (источника со-

общений на основе этого алфавита) является энтропия.

Этот термин применительно к техническим системам был вве-

ден К. Шенноном и Р. Хартли.



Энтропию алфавита А = {ai} по К. Шеннону рассчитывают по

следующей формуле:



С физической точки зрения энтропия алфавита показывает,

какое количество информации приходится в среднем на

один символ алфавита.

Частным случаем энтропии Шеннона считается энтропия Харт-

ли. Дополнительным условием при этом является то, что все веро-

ятности одинаковы и постоянны для всех символов алфавита.

С учетом этого формулу (2.1) можно преобразовать к виду:



**2. Приложение**

char[] EngAlph = { 'a', 'b', 'c', 'd', 'e', 'f', 'g', 'h', 'i', 'j', 'k', 'l', 'm', 'n', 'o', 'p', 'q', 'r', 's', 't', 'u', 'v', 'w', 'x', 'y', 'z' };

char[] BelRusAlph = { 'а', 'б', 'в', 'г', 'д', 'е', 'ё', 'ж', 'з', 'i', 'й', 'к', 'л', 'м', 'н', 'о', 'п', 'р', 'с', 'т', 'у', 'Y', 'ф', 'ч', 'ц', 'ч', 'ш', 'ы', 'ь', 'э', 'ю', 'я' };

char[] Eng = { 'h', 'e', 'l', 'l', 'o' };

char[] Belrus = { 'п', 'р', 'ы', 'в', 'i', 'т',};

double lengEng = EngAlph.Length;

double lengBelRus = BelRusAlph.Length;

double Hbinary = 0;//Бинарный язык

double Heng = 0;//Энтропия для каждого языка

double HBelrus = 0;

Console.WriteLine("Кол-во букв в английском алфавите :" + lengEng);

Console.WriteLine("Кол-во букв в белорусском алфавите :" + lengBelRus);

foreach (char i in EngAlph)

{

double count = 0;

foreach (char z in Eng)

{

if (z == i)

count++;

}

double Pai = count / lengEng;

// Console.WriteLine(Pai);

if (Pai != 0)

{

Heng += Pai \* Math.Log2(Pai);

}

}

Heng = Heng \* (-1);

Console.WriteLine("Энтропия английского алфавита:");

Console.WriteLine(Heng);

foreach (char i in BelRusAlph)

{

double count = 0;

foreach (char z in Belrus)

{

if (z == i)

count++;

}

double Pai = count / lengBelRus;

// Console.WriteLine(Pai);

if (Pai != 0)

{

HBelrus += Pai \* Math.Log2(Pai);

}

}

HBelrus = HBelrus \* (-1);

Console.WriteLine("Энтропия белорусского алфавита:");

Console.WriteLine(HBelrus);

int[] countLetter = new int[2];

double[] probabilityLetters = new double[2];

string letters = "01";

using (StreamReader streamReader = new StreamReader(@"D:\3 курс 1 сем\ЗАЩИТА ИНФОРМАЦИИ И НАДЕЖНОСТЬ ИНФОРМАЦИОННЫХ СИСТЕМ\2 Лабораторная работа\binary.txt"))

{

string file = streamReader.ReadToEnd();

double binaryleng = file.Count();

for (int j = 0; j < 2; j++)

{

countLetter[j] = file.Count(x => x == letters[j]);

probabilityLetters[j] = (double)countLetter[j] / binaryleng;

Console.WriteLine();

Hbinary += (probabilityLetters[j] \* (Math.Log(probabilityLetters[j])) / Math.Log(2));

}

}

Hbinary = Hbinary \* (-1);

Console.WriteLine("Энтропия бинарного алфавита:");

Console.WriteLine(Hbinary);

Console.WriteLine("Введите Ваше ФИО: ");

string FIO = Console.ReadLine();

double countBelInformation = HBelrus \* FIO.Replace(" ", "").ToLower().Count();

double countBinaryInformation = Hbinary \* FIO.Replace(" ", "").ToLower().Count();

Console.WriteLine("Количество информации с использованием энтропии беларусского алфавита:");

Console.WriteLine(countBelInformation);

Console.WriteLine("Количество информации с использованием энтропии бинарного алфавита:");

Console.WriteLine(countBinaryInformation);

Console.WriteLine();

double ascii = FIO.ToLower().Count() \* 8;

Console.WriteLine();

Console.WriteLine("Количество информации с использованием ASСII для белорусского алфавита:");

double ASIIBelrInformation = ascii \* HBelrus;

Console.WriteLine(ASIIBelrInformation);

Console.WriteLine();

Console.WriteLine("Количество информации с использованием ASСII для бинарного алфавита:");

double ASIIbinaryInformation = ascii \* Hbinary;

Console.WriteLine(ASIIbinaryInformation);

Console.WriteLine("Вероятность ошибочной передачи единичного бита сообщения состовляет 0,1");

float p = 0.1F;

float q = 1 - p;

Console.WriteLine("Условная энтропия ");

double H01 = -1 \* p \* (Math.Log2(p) / Math.Log(2)) - q \* (Math.Log2(q) / Math.Log(2));

Console.WriteLine(H01);

Console.WriteLine("Эффективная энтропия");

double effectH01 = 1 - H01;

Console.WriteLine(effectH01);

Console.WriteLine("Вероятность ошибочной передачи единичного бита сообщения состовляет 0,5");

p = 0.5F;

q = 1 - p;

Console.WriteLine("Условная энтропия ");

double H05 = -1 \* p \* (Math.Log2(p) / Math.Log(2)) - q \* (Math.Log2(q) / Math.Log(2));

Console.WriteLine(H05);

Console.WriteLine("Эффективная энтропия");

double effectH05 = 1 - H01;

Console.WriteLine(effectH05);

Console.WriteLine("Вероятность ошибочной передачи единичного бита сообщения состовляет 1,0");

p = 1;

q = 1 - p;

Console.WriteLine("Условная энтропия ");

double H1 = -1 \* p \* Math.Log2(p) - q \* Math.Log2(q);

Console.WriteLine(H1);

Console.WriteLine("Эффективная энтропия");

double effectH1 = 1 - H1;

Console.WriteLine(effectH1);

}

}

3**. Ответы на контрольные вопросы**

1. Что такое алфавит источника сообщения?

Набор символов

1. Что такое мощность алфавита источника сообщения?

Количество символов

1. Какова мощность алфавита белорусского языка?

32

1. Какова мощность алфавита русского языка?

33

1. Какова мощность алфавита «компьютерного» языка?

2

1. Что такое энтропия алфавита?

С физической точки зрения энтропия алфавита показывает,

какое количество информации приходится в среднем на

один символ алфавита.

1. Что такое энтропия сообщения?

С физической точки зрения энтропия сообщения показывает,

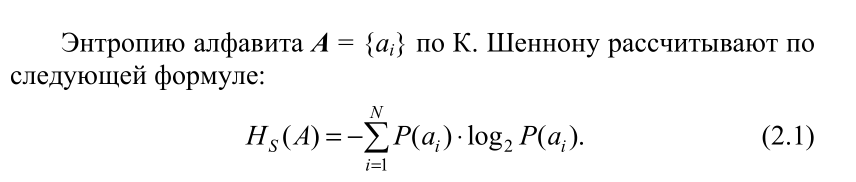
какое количество информации приходится в среднем на

один символ алфавита.

1. От чего зависит энтропия алфавита?

От вероятности появления символа алфавита.

1. Записать формулу для вычисления энтропии.





1. Что нужно знать для вычисления энтропии алфавита?

Мощность алфавита

Символы алфавита

11. Как рассчитываются энтропия Шеннона и энтропия Харт-

ли? В чем принципиальное различие между этими характеристи-

ками? Дайте толкование физического смысла энтропии.

Вероятность по Хартли постоянная

1. Пояснить назначение знака «минус» в формулах (2.1) и (2.4).

Энтропия не может быть отрицательной, а логарифм может быть отрицательным. Поэтому чтобы избежать отриц. Энтропию ставят - .

13. Что такое избыточность алфавита и избыточность сооб-

щений, сформированных в компьютерных системах? Принцип

действия каких систем основан на существовании данной избы-

точности?

Избыточностью алфавита называют уменьшение информаци-

онной нагрузки на один символ вследствие разной вероятности и

взаимозависимости появления его символов в сообщениях.

14. Расположить в порядке возрастания энтропии известные

Вам алфавиты.

Английский

Белоруский

Русский

15. Вычислить энтропию алфавита белорусского (русского)

языка.

5 - Белорусского

16. Вычислить энтропию Шеннона бинарного алфавита, если

вероятность появления в произвольном документе на основе этого

алфавита одного из символов составляет 0,25, другого – 0,75; либо

0 и 1,0; либо 0,5 и 0,5.

----

17. Чему равна энтропия алфавита по Хартли, если мощность

этого алфавита равна: а) 1 символ; б) 2 символа; в) 8 символов?

А) 0

Б) 1

В) 3