МИНЕСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ

РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ

ВЫСШЕГО ОБРАЗОВАНИЯ «САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ»

КАФЕДРА № 43

ОТЧЕТ   
ЗАЩИЩЕН С ОЦЕНКОЙ

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Ст. преподаватель  Ассистент |  |  |  | А.А.Фоменкова  М.В.Величко |
| должность, уч. степень, звание |  | подпись, дата |  | инициалы, фамилия |

|  |
| --- |
| ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ |
| «Оценка количества информации в сообщении и эффективное кодирование» |
| по курсу: ИНФОРМАТИКА |
|  |
|  |

РАБОТУ ВЫПОЛНИЛ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| СТУДЕНТ ГР. № | 4135К |  |  |  | Столяров Н. С. |
|  |  |  | подпись, дата |  | инициалы, фамилия |

Санкт-Петербург 2021

**Цель работы:** Получение практических навыков численного определения количества информации, содержащегося в сообщении. Освоение методов построения кодов дискретного источника информации используя конструктивный метод, предложенный К.Шенноном и Н.Фано, и метод Хаффмана. На примере показать однозначность раскодирования имеющегося сообщения.

**1 часть**

1 Таблица

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № | Символ | Код символа | Число вхождений  символа в текст | Вероятнисть вхождения  символа | I |
| 1 | 0 | 48 | 2447 | 0,013275176 | 6,235125 |
| 2 | 1 | 49 | 1726 | 0,009363692 | 6,738707 |
| 3 | 2 | 50 | 1060 | 0,005750587 | 7,442075 |
| 4 | 3 | 51 | 833 | 0,004519094 | 7,789751 |
| 5 | 4 | 52 | 932 | 0,005056177 | 7,627737 |
| 6 | 5 | 53 | 598 | 0,003244199 | 8,267922 |
| 7 | 6 | 54 | 600 | 0,003255049 | 8,263105 |
| 8 | 7 | 55 | 637 | 0,003455777 | 8,176774 |
| 9 | 8 | 56 | 461 | 0,002500963 | 8,643301 |
| 10 | 9 | 57 | 2439 | 0,013231776 | 6,239849 |
| 11 | й | 233 | 1940 | 0,01052466 | 6,570083 |
| 12 | ц | 246 | 1716 | 0,009309441 | 6,74709 |
| 13 | у | 243 | 3518 | 0,01908544 | 5,711384 |
| 14 | к | 234 | 3761 | 0,020403735 | 5,615023 |
| 15 | е | 229 | 11806 | 0,064048522 | 3,964691 |
| 16 | н | 237 | 9628 | 0,052232693 | 4,258903 |
| 17 | г | 227 | 1822 | 0,0098845 | 6,660616 |
| 18 | ш | 248 | 811 | 0,004399742 | 7,828365 |
| 19 | щ | 249 | 455 | 0,002468412 | 8,662201 |
| 20 | з | 231 | 2221 | 0,012049108 | 6,37493 |
| 21 | х | 245 | 1227 | 0,006656576 | 7,231004 |
| 22 | ъ | 250 | 22 | 0,000119352 | 13,03249 |
| 23 | ф | 244 | 1025 | 0,005560709 | 7,490515 |
| 24 | ы | 251 | 2333 | 0,012656717 | 6,303953 |
| 25 | в | 226 | 4671 | 0,02534056 | 5,302408 |
| 26 | а | 224 | 10322 | 0,0559977 | 4,158489 |
| 27 | п | 239 | 3452 | 0,018727384 | 5,738707 |
| 28 | р | 240 | 6626 | 0,035946595 | 4,798001 |
| 29 | о | 238 | 11602 | 0,062941805 | 3,989838 |
| 30 | л | 235 | 4868 | 0,026409301 | 5,24281 |
| 31 | д | 228 | 3238 | 0,017566417 | 5,831036 |
| 32 | ж | 230 | 738 | 0,004003711 | 7,964447 |
| 33 | э | 253 | 235 | 0,001274894 | 9,615407 |
| 34 | я | 255 | 2907 | 0,015770714 | 5,986608 |
| 35 | ч | 247 | 2666 | 0,014463269 | 6,111462 |
| 36 | с | 241 | 5457 | 0,029604674 | 5,078031 |
| 37 | м | 236 | 5217 | 0,028302654 | 5,142919 |
| 38 | и | 232 | 12306 | 0,066761063 | 3,904849 |
| 39 | т | 242 | 7593 | 0,04119265 | 4,601469 |
| 40 | ь | 252 | 1673 | 0,009076163 | 6,783702 |
| 41 | б | 225 | 1146 | 0,006217144 | 7,329532 |
| 42 | ю | 254 | 908 | 0,004925975 | 7,665375 |
| 43 | . | 46 | 6205 | 0,033662636 | 4,892708 |
| 44 | , | 44 | 3678 | 0,019953453 | 5,647218 |
| 45 | : | 58 | 781 | 0,004236989 | 7,882745 |
| 46 | ; | 59 | 667 | 0,00361853 | 8,110381 |
| 47 | - | 45 | 1058 | 0,005739737 | 7,4448 |
| 48 | ( | 40 | 1554 | 0,008430578 | 6,890153 |
| 49 |  | 32 | 30743 | 0,166783306 | 2,583953 |
|  |  | Всего символов в тексте | 184329 |  |  |
|  |  |  | Полная вероятность | 1 |  |
|  |  |  |  | Энтропия источника | 6,542299 |
|  |  |  |  | Максимальная энтропия | 5,61471 |

2 таблица

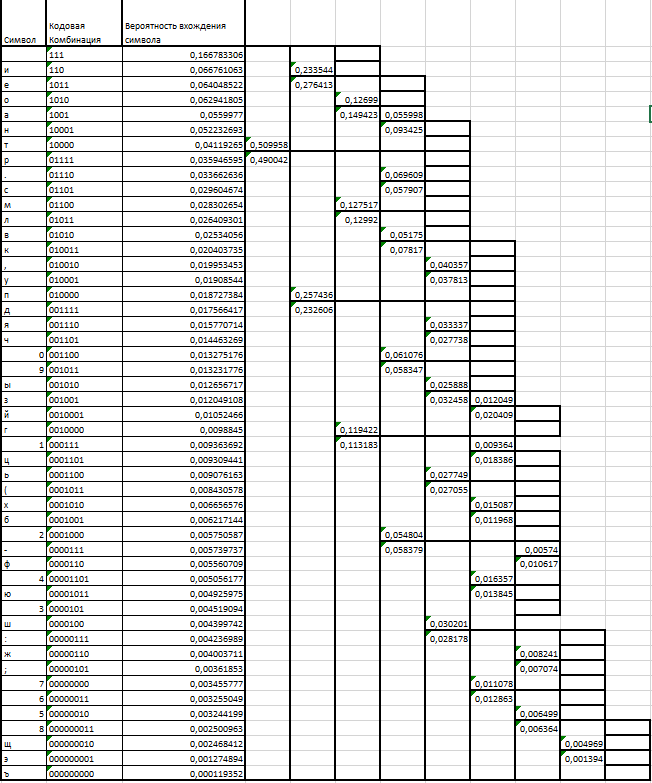
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Неопределённость | Разрядность кода | Абсолютная избыточность | Относительная избыточность |
| Стандартная кодовая  таблица ASCII | 8 | 8 | 3,348870456 | 0,418608807 |
| Мера Хартли | 5,614709844 | 6 | 0,927588973 | 0,141783339 |

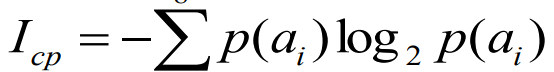
**Формулы используемые в работе:**

|  |  |
| --- | --- |
| **Ii=-log2pi** | **Ii** — количество информации  **p**-вероятность вхождения символа |
| **D=(Hmax-H)/Hmax** | **Hmax**— максимальное значение энтропии  **D** —относительная избыточность |
| **Dabc=Hmax-H** | **Dabc**- абсолютная избыточность |
| **Pi = ki/K** | **Pi** – вероятность вхождения символа в текст  **ki** — число вхождений символов в текст  **K** — всего символов в текст |
| **Iср= Iср =∑ pi (-log2 pi) = H** | **H**- энтропия источника |
| **H1 = loga N** | **N**— число возможных значений принятого слова после получения сообщения |

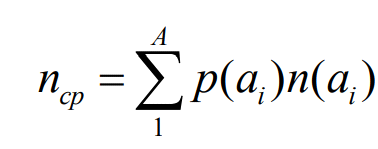
**Вывод:** Мы научились определять количество информации содержащейся в тексте.

**2 часть**

****



=4,651129809



=4,68847697958486

**ПРИМЕР ДЕКОДИРОВАНИЯ СООБЩЕНИЯ:**

Информатика - 110100010000110101001111011001001100001100100111001

И - 110

Н - 10001

Ф - 0000110

О - 1010

Р - 01111

М - 01100

А - 1001

Т - 110

И - 110

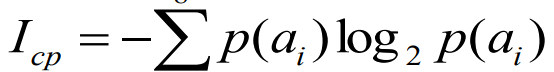
К - 010011

А – 1001

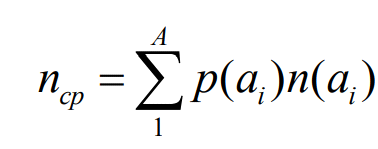
**Вывод:** Мы научились кодировать дискретные источники информации методом Шеннона-Фано.

**3 часть**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Вероятнисть вхождения  символа | Шаг 1 | Шаг 2 | Шаг 3 | Шаг 4 | Шаг 45 | Шаг 46 | Шаг 47 |  |
| 0,166783306 | 0,166783 | 0,166783 | 0,166783 | 0,166783 | 0,166783 | 0,166783 | 0,166783 | 1 |
| 0,066761063 | 0,066761 | 0,066761 | 0,066761 | 0,066761 | 0,33973 | 0,493487 | 0,833217 |  |
| 0,064048522 | 0,064049 | 0,064049 | 0,064049 | 0,064049 | 0,263724 | 0,33973 |  | 1 |
| 0,062941805 | 0,062942 | 0,062942 | 0,062942 | 0,062942 | 0,229763 |  | 0,833217 |  |
| 0,0559977 | 0,055998 | 0,055998 | 0,055998 | 0,055998 |  | 0,493487 |  |  |
| 0,052232693 | 0,052233 | 0,052233 | 0,052233 | 0,052233 | 0,33973 |  |  |  |
| 0,04119265 | 0,041193 | 0,041193 | 0,041193 | 0,041193 |  |  |  |  |
| 0,035946595 | 0,035947 | 0,035947 | 0,035947 | 0,035947 |  |  |  |  |
| 0,033662636 | 0,033663 | 0,033663 | 0,033663 | 0,033663 |  |  |  |  | |  |  |
| 0,029604674 | 0,029605 | 0,029605 | 0,029605 | 0,029605 |  |  |  |  | |  |  |
| 0,028302654 | 0,028303 | 0,028303 | 0,028303 | 0,028303 |  |  |  |  | |  |  |
| 0,026409301 | 0,026409 | 0,026409 | 0,026409 | 0,026409 |  |  |  |  | |  |  |
| 0,02534056 | 0,025341 | 0,025341 | 0,025341 | 0,025341 |  |  |  |  | |  |  |
| 0,020403735 | 0,020404 | 0,020404 | 0,020404 | 0,020404 |  |  |  |  | |  |  |
| 0,019953453 | 0,019953 | 0,019953 | 0,019953 | 0,019953 |  |  |  |  | |  |  |
| 0,01908544 | 0,019085 | 0,019085 | 0,019085 | 0,019085 |  |  |  |  | |  |  |
| 0,018727384 | 0,018727 | 0,018727 | 0,018727 | 0,018727 |  |  |  |  | |  |  |
| 0,017566417 | 0,017566 | 0,017566 | 0,017566 | 0,017566 |  |  |  |  | |  |  |
| 0,015770714 | 0,015771 | 0,015771 | 0,015771 | 0,015771 |  |  |  |  | |  |  |
| 0,014463269 | 0,014463 | 0,014463 | 0,014463 | 0,014463 |  |  |  |  | |  |  |
| 0,013275176 | 0,013275 | 0,013275 | 0,013275 | 0,013275 |  |  |  |  | |  |  |
| 0,013231776 | 0,013232 | 0,013232 | 0,013232 | 0,013232 |  |  |  |  | |  |  |
| 0,012656717 | 0,012657 | 0,012657 | 0,012657 | 0,012657 |  |  |  |  | |  |  |
| 0,012049108 | 0,012049 | 0,012049 | 0,012049 | 0,012049 |  |  |  |  | |  |  |
| 0,01052466 | 0,010525 | 0,010525 | 0,010525 | 0,010525 |  |  |  |  | |  |  |
| 0,0098845 | 0,009884 | 0,009884 | 0,009884 | 0,009884 |  |  |  |  | |  |  |
| 0,009363692 | 0,009364 | 0,009364 | 0,009364 | 0,009364 |  |  |  |  | |  |  |
| 0,009309441 | 0,009309 | 0,009309 | 0,009309 | 0,009309 |  |  |  |  | |  |  |
| 0,009076163 | 0,009076 | 0,009076 | 0,009076 | 0,009076 |  |  |  |  | |  |  |
| 0,008430578 | 0,008431 | 0,008431 | 0,008431 | 0,008431 |  |  |  |  | |  |  |
| 0,006656576 | 0,006657 | 0,006657 | 0,006657 | 0,006711 |  |  |  |  | |  |  |
| 0,006217144 | 0,006217 | 0,006217 | 0,006217 | 0,006657 |  |  |  |  | |  |  |
| 0,005750587 | 0,005751 | 0,005751 | 0,005751 | 0,006217 |  |  |  |  | |  |  |
| 0,005739737 | 0,00574 | 0,00574 | 0,005745 | 0,005751 |  |  |  |  | |  |  |
| 0,005560709 | 0,005561 | 0,005561 | 0,00574 | 0,005745 |  |  |  |  | |  |  |
| 0,005056177 | 0,005056 | 0,005056 | 0,005561 | 0,00574 |  |  |  |  | |  |  |
| 0,004925975 | 0,004926 | 0,004926 | 0,005056 | 0,005561 |  |  |  |  | |  |  |
| 0,004519094 | 0,004519 | 0,004519 | 0,004926 | 0,005056 |  |  |  |  | |  |  |
| 0,004399742 | 0,0044 | 0,0044 | 0,004519 | 0,004926 |  |  |  |  | |  |  |
| 0,004236989 | 0,004237 | 0,004237 | 0,0044 | 0,004519 |  |  |  |  | |  |  |
| 0,004003711 | 0,004004 | 0,004004 | 0,004237 | 0,0044 |  |  |  |  | |  |  |
| 0,00361853 | 0,003619 | 0,003863 | 0,004004 | 0,004237 |  |  |  |  | |  |  |
| 0,003455777 | 0,003456 | 0,003619 | 0,003863 | 0,004004 |  |  |  |  | |  |  |
| 0,003255049 | 0,003255 | 0,003456 | 0,003619 | 0,003863 |  |  |  |  | |  |  |
| 0,003244199 | 0,003244 | 0,003255 | 0,003456 | 0,003619 |  |  |  |  | |  |  |
| 0,002500963 | 0,002501 | 0,003244 | 0,003255 |  |  |  |  |  | |  |  |
| 0,002468412 | 0,002468 | 0,002501 |  | 0,006711 |  |  |  |  | |  |  |
| 0,001274894 | 0,001394 |  | 0,005745 |  |  |  |  |  | |  |  |
| 0,000119352 |  | 0,003863 |  |  |  |  |  |  | |  |  |
|  | 0,001394 |  |  |  |  |  |  |  | |  |  |



=4,651129809



=4,686256086

|  |  |
| --- | --- |
| Символ | Кодовая Комбинация |
|  | 111 |
| и | 110 |
| е | 1011 |
| о | 1010 |
| а | 1001 |
| н | 10001 |
| т | 10000 |
| р | 01111 |
| . | 01110 |
| с | 01101 |
| м | 01100 |
| л | 01011 |
| в | 01010 |
| к | 010011 |
| , | 010010 |
| у | 010001 |
| п | 010000 |
| д | 001111 |
| я | 001110 |
| ч | 001101 |
| 0 | 001100 |
| 9 | 001011 |
| ы | 001010 |
| з | 001001 |
| й | 0010001 |
| г | 0010000 |
| 1 | 000111 |
| ц | 0001101 |
| ь | 0001100 |
| ( | 0001011 |
| х | 0001010 |
| б | 0001001 |
| 2 | 0001000 |
| - | 0000111 |
| ф | 0000110 |
| 4 | 00001101 |
| ю | 00001011 |
| 3 | 0000101 |
| ш | 0000100 |
| : | 00000111 |
| ж | 00000110 |
| ; | 00000101 |
| 7 | 00000000 |
| 6 | 00000011 |
| 5 | 00000010 |
| 8 | 000000011 |
| щ | 000000010 |
| э | 000000001 |
| ъ | 000000000 |

Информатика – 110100010000110101001111011001001100001100100111001

И – 110

Н - 10001

Ф - 0000110

О - 1010

Р - 01111

М - 01100

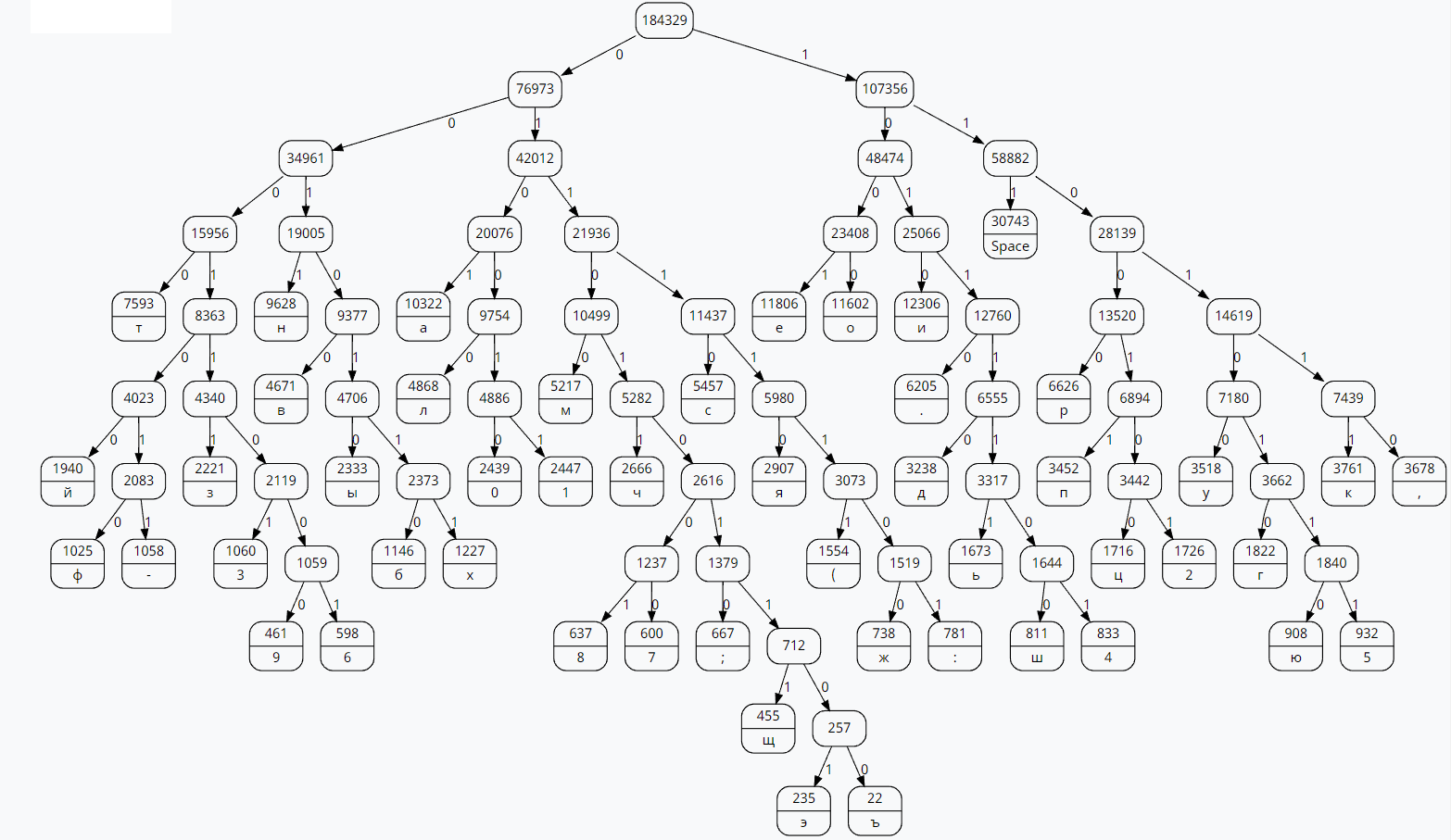
А - 1001

Т - 10000

И - 110

К -010011

А - 1001

****

**Вывод:** По итогам лабораторной работы я получил практические навыки численного определения количества информации, содержащегося в сообщении. Также можно сделать вывод, что в разных алфавитах, на примере рассмотренных стандартной кодовый ASCII и меры Хартли их неопределенность, разрядность кода, абсолютная избыточность и относительная избыточность отличаются. Если нам не известен алфавит, то нужно использовать стандартную кодовую таблицу ASCII, а если не известен, то нужно использовать меру Хартли.

Мы освоили метод построения дискретного источника информации, используя конструктивный метод, предложенный К.Шенноном и Н.Фано и методику Д.Хаффмана и на примере показали однозначность раскодирования имеющихся сообщений. Методика Хаффмана лучше, поскольку кодирование однозначно. Символ будет кодироваться меньшим количеством нулей и единиц.