

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ БЮДЖЕТНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«СИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ И
ИНФОРМАТИКИ»

Теория информации

Практическая работа №3
«Блочное кодирование»

Выполнил:
студент гр. ИП-013
Копытина Т.А.

Проверила:
доцент кафедры ПМиК
Мачикина Е.П.

Новосибирск 2024 г.

Результат работы:

```
Хаффман для блока размером 1:  
Р: 0.6025558978 - 1  
И: 0.2990480381 - 01  
М: 0.0983960642 - 00  
Средняя длина кодовых слов (L ср.): 1.3974441022359105  
Энтропия для 1 символа подряд: 1.2903350868165782  
Избыточность для 1 символа подряд: 0.10710901541933238
```

Рис. 1 Результат работы кода для кодирования блока 1 символом

```
Хаффман для блока размером 2:  
РР: 0.3632654694 - 0  
ИР: 0.1847126115 - 111  
РИ: 0.1751129955 - 110  
ИИ: 0.0893164273 - 1011  
РМ: 0.0600375985 - 1001  
МР: 0.0587176513 - 1000  
МИ: 0.0302387904 - 10100  
ИМ: 0.0293988240 - 101011  
ММ: 0.0091996320 - 101010  
Средняя длина кодовых слов (L ср.): 2.657813687452502  
Энтропия для 2 символов подряд: 1.2902870289779254  
Избыточность для 2 символов подряд: 0.038619814748325654
```

Рис. 2 Результат работы кода для кодирования блока 2 символами

```
Хаффман для блока размером 3:  
RRR: 0.2183825294 - 00  
RRI: 0.1110511159 - 100  
RIR: 0.1095512359 - 011  
IRR: 0.1093712503 - 010  
IIR: 0.0530957523 - 1010  
IRI: 0.0524958003 - 11111  
RII: 0.0517758579 - 11110  
RRM: 0.0362371010 - 11011  
RMR: 0.0357571394 - 11010  
MRR: 0.0343772498 - 10111  
III: 0.0275977922 - 111011  
MRI: 0.0188984881 - 111001  
IMR: 0.0178785697 - 111000  
MIR: 0.0176985841 - 110011  
РИМ: 0.0176985841 - 110010  
РМИ: 0.0173986081 - 110000  
ИРМ: 0.0164986801 - 101101  
ИИМ: 0.0089392849 - 1110100  
ИМИ: 0.0088192945 - 1100011  
МИИ: 0.0088192945 - 1100010  
РММ: 0.0058795296 - 11101011  
ММР: 0.0057595392 - 11101010  
МРМ: 0.0046796256 - 10110011  
ИММ: 0.0038396928 - 10110001  
ММИ: 0.0034797216 - 10110000  
МИМ: 0.0029997600 - 101100101  
МММ: 0.0009599232 - 1011001001  
И: 0.0000599952 - 1011001000  
Средняя длина кодовых слов (L ср.): 3.906647468202543  
Энтропия для 3 символов подряд: 1.290177825237757  
Избыточность для 3 символов подряд: 0.012037997496424127
```

Рис. 3 Результат работы кода для кодирования блока 3 символами

```

Хаффман для блока размером 4:
RRRR: 0.1335093193 - 101
RRIР: 0.0676745860 - 1100
IRRR: 0.0664746820 - 1001
RIРR: 0.0634349252 - 0110
RRRI: 0.0615950724 - 0101
ИРИР: 0.0357571394 - 11011
РРИИ: 0.0334373250 - 10001
ИРРИ: 0.0330373570 - 10000
РИРИ: 0.0325573954 - 01111
ИИРР: 0.0323974082 - 01110
РИИР: 0.0299176066 - 01000
РРРИ: 0.0225581953 - 00100
РМРР: 0.0219182465 - 00010
РРМР: 0.0217582593 - 00001
МРРР: 0.0191184705 - 111001
ИИИР: 0.0170386369 - 110101
РИИИ: 0.0165586753 - 110100
ИРИИ: 0.0161587073 - 010011
ИИРИ: 0.0148788097 - 001111
МРРИ: 0.0122390209 - 001101
РРМИ: 0.0116790657 - 001010
РМРИ: 0.0115190785 - 000111
ИМРР: 0.0109591233 - 000110
РМИР: 0.0107991361 - 000000
ИРРМ: 0.0107991361 - 1111111
МРИР: 0.0107191425 - 1111110
РРИМ: 0.0105591553 - 1111100
РИМР: 0.0101591873 - 1111010
ИРМР: 0.0101591873 - 1111001
МИРР: 0.0099192065 - 1110111
РИРМ: 0.0090392769 - 1110001
ИИИИ: 0.0075193984 - 0100100
РМИИ: 0.0060795136 - 0011000
ИМИР: 0.0053595712 - 0000010
МИИР: 0.0053595712 - 11111011
ИИРМ: 0.0051195904 - 11110111
ИРИМ: 0.0051195904 - 11110110
РИМИ: 0.0051195904 - 11110001
ИИМР: 0.0049596032 - 11110000
МИРИ: 0.0049596032 - 11101101
ИМРИ: 0.0048796096 - 11101011
МРИИ: 0.0048796096 - 11101010
ИРМИ: 0.0047996160 - 11101001
РИИМ: 0.0043196544 - 11100001
МРРМ: 0.0039196864 - 11100000
РММР: 0.0034397248 - 00111011
МРМР: 0.0031997440 - 00111001
РРММ: 0.0031197504 - 00110011
ИИИМ: 0.0030397568 - 00110010
ИМИИ: 0.0029597632 - 00101111
РМРМ: 0.0029597632 - 00101110
ММРР: 0.0028797696 - 00101100
РММИ: 0.0027997760 - 00000111
МИИИ: 0.0027997760 - 00000110
ИММР: 0.0027197824 - 111110101
ММИР: 0.0025597952 - 111110100
ИИМИ: 0.0023998080 - 111011000
МРИМ: 0.0023198144 - 111010000
МИРМ: 0.0019998400 - 010010111
ИРММ: 0.0019198464 - 010010110
РИММ: 0.0017598592 - 010010100
МИМР: 0.0017598592 - 001110101
ММРИ: 0.0016798656 - 001110100
МРМИ: 0.0015998720 - 001110001
РМИМ: 0.0015198784 - 001110000
ИМРМ: 0.0014398848 - 001011010
МИМИ: 0.0012798976 - 1110110011
ИММИ: 0.0009599232 - 0100101011
МИИМ: 0.0009599232 - 0100101010
ММРМ: 0.0007199424 - 0010110110
ИИИМ: 0.0006399488 - 11101000111
МММР: 0.0006399488 - 11101000110
МРММ: 0.0006399488 - 11101100101
ИИИМ: 0.0006399488 - 11101100100
РМММ: 0.0004799616 - 11101000100
МММИ: 0.0003999680 - 00101101111
МИММ: 0.0003199744 - 111010001011
ИМММ: 0.0003199744 - 111010001010
ММИИ: 0.0002399808 - 001011011101
ММИМ: 0.0000799936 - 0010110111001
МИ: 0.0000799936 - 0010110111000
Средняя длина кодовых слов (L ср.): 5.17966562674986
Энтропия для 4 символов подряд: 1.2899477187874582
Избыточность для 4 символов подряд: 0.0049686879000068185
PS C:\TI>

```

Рис. 4 Результат работы кода для кодирования блока 4 символами

Результаты:

	Длина блока $n=1$	Длина блока $n=2$	Длина блока $n=3$	Длина блока $n=4$
Оценка избыточности кодирования на один символ входной последовательности	0.1059443	0.03646749	0.01229021	0.00496868

$H(n=1)$: 1.2903350

$L_{\text{ср}}(n=1)$: 1.397444

$H(n=2)$: 1.290287028

$L_{\text{ср}}(n=2)$: 2.65781368

$H(n=3)$: 1.290177825

$L_{\text{ср}}(n=3)$: 3.9066474

$H(n=4)$: 1.289947718

$L_{\text{ср}}(n=4)$: 5.1796656

Вывод:

Из анализа доступных данных можно сделать следующие заключения:

Избыточность кодирования отражает, насколько излишне большой объем информации требуется для передачи данных по сравнению с минимально необходимым. В данном контексте, эту избыточность оценивают по тому, насколько больше битов кодирования требуется для каждого символа входной последовательности, чем минимально возможное количество, достижимое при использовании идеального кода.

При увеличении размера блока n , мы рассматриваем больше символов за один раз при кодировании. Это помогает лучше использовать статистические свойства языка и снижать избыточность кодирования. Например, при маленьком значении n (например, $n=1$), мы кодируем каждый символ независимо, что может привести к большой избыточности из-за недостаточного учета контекста. Однако, при увеличении n до 2, мы начинаем

учитывать пары символов, что позволяет более эффективно адаптироваться к статистике текста. Избыточность кодирования уменьшается, так как мы лучше используем информацию о последовательности символов.

Следовательно, увеличение длины блока n приводит к уменьшению избыточности кодирования на один символ входной последовательности. Это означает, что мы более эффективно используем биты для кодирования информации, что в целом является желаемым результатом при разработке эффективных методов сжатия данных или кодирования.