

Кодирование текста

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| sp | ! | " | # | \$ | % | & | ' | (|) | * | + | , | - | . | / |
| 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | : | ; | < | = | > | ? |
| 48 | 49 | 50 | 51 | 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 |
| @ | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
| 64 | 65 | 66 | 67 | 68 | 69 | 70 | 71 | 72 | 73 | 74 | 75 | 76 | 77 | 78 | 79 |
| P | Q | R | S | T | U | V | W | X | Y | Z | [| \ |] | ^ | _ |
| 80 | 81 | 82 | 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 | 94 | 95 |
| ` | a | b | c | d | e | f | g | h | i | j | k | l | m | n | o |
| 96 | 97 | 98 | 99 | 100 | 101 | 102 | 103 | 104 | 105 | 106 | 107 | 108 | 109 | 110 | 111 |
| p | q | r | s | t | u | v | w | x | y | z | { | | } | ~ | |
| 112 | 113 | 114 | 115 | 116 | 117 | 118 | 119 | 120 | 121 | 122 | 123 | 124 | 125 | 126 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 115 | 116 | 117 | 118 | 119 | 120 | 121 | 122 | 123 | 124 | 125 | 126 | 127 | 128 | 129 | 130 |
| b | d | l | z | [| n | ^ | m | x | λ | Σ | [| |] | | |

Цели урока:

- 1. Познакомиться с методом измерения информации в символьном сообщении.*
- 2. Рассмотреть понятие - Таблица кодировки.*
- 3. Научиться вычислять количество информации, заключенное в тексте, составленном из символов определенного алфавита*



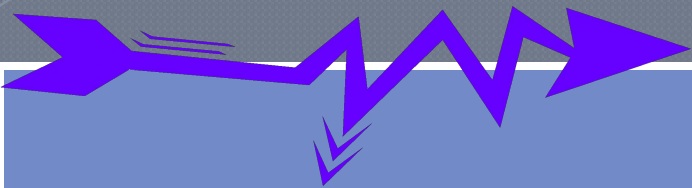


В традиционных кодировках для кодирования одного символа используется **8 бит**. Легко подсчитать по формуле :

$$N=2^i \quad \text{где}$$

i – количество информации;

N – количество возможных событий; что такой 8-разрядный код позволяет закодировать **256** различных символов.



Присвоение символу определенного числового кода - это вопрос соглашения. В качестве международного стандарта принята кодовая таблица **ASCII**, кодирующая первую половину символов с числовыми кодами от 0 до 127 (коды от 0 до 32 отведены не символам, а функциональным клавишам).



Кодовая таблица ASCII (American Standart Code for Information Interchang)

| | | | | | | | | | | | | | | | |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| sp 32 | ! 33 | " 34 | # 35 | \$ 36 | % 37 | & 38 | ' 39 | (40 |) 41 | * 42 | + 43 | , 44 | - 45 | . 46 | / 47 |
| 0 48 | 1 49 | 2 50 | 3 51 | 4 52 | 5 53 | 6 54 | 7 55 | 8 56 | 9 57 | : 58 | ; 59 | < 60 | = 61 | > 62 | ? 63 |
| @ 64 | A 65 | B 66 | C 67 | D 68 | E 69 | F 70 | G 71 | H 72 | I 73 | J 74 | K 75 | L 76 | M 77 | N 78 | O 79 |
| P 80 | Q 81 | R 82 | S 83 | T 84 | U 85 | V 86 | W 87 | X 88 | Y 89 | Z 90 | [91 | \ 92 |] 93 | ^ 94 | _ 95 |
| ` 96 | a 97 | b 98 | c 99 | d 100 | e 101 | f 102 | g 103 | h 104 | i 105 | j 106 | k 107 | l 108 | m 109 | n 110 | o 111 |
| p 112 | q 113 | r 114 | s 115 | t 116 | u 117 | v 118 | w 119 | x 120 | y 121 | z 122 | { 123 | 124 | } 125 | ~ 126 | |




Национальные стандарты кодировочных таблиц включают международную часть кодовой таблицы без изменений, а во второй половине содержат коды национальных алфавитов, символы псевдографики и некоторые математические знаки. К сожалению, в настоящее время существуют пять различных кодировок кириллицы (KOI8-P, Windows, MSDOS, Macintocs и ISO), что вызывает дополнительные трудности при работе с русскоязычными документами.



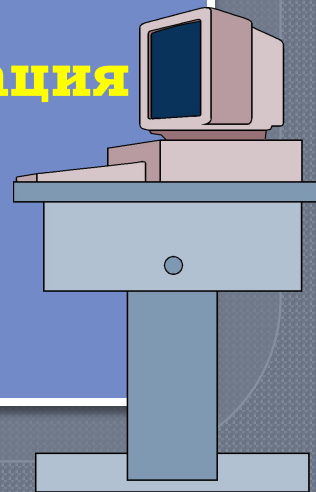



Хронологически одним из первых стандартов кодирования русских букв на компьютерах был КОИ8 («Код обмена информацией, 8-битный»). Эта кодировка применялась еще в 70-ые годы прошлого века на компьютерах серии ЕС ЭВМ, а с середине 80-х стала использоваться в первых русифицированных версиях операционной системы UNIX.





Наиболее распространенной в настоящее время является кодировка Windows, обозначаемая сокращением CP1251 («CP» означает «кодовая страница»). От начала 90-х годов, времени господства операционной системы MSDOS, остается кодировка CP866. Компьютеры фирмы Apple, работающие под управлением операционной системы Mac OS, используют свою собственную кодировку Mac. Кроме того, Международная организация по стандартизации утвердила в качестве стандарта для русского языка еще одну кодировку под названием ISO.





В конце 90-х годов появился новый международный стандарт Unicode, который отводит под один символ не один байт, а два, и поэтому с его помощью можно закодировать не 256, а 65536 различных символов. Полная спецификация стандарта Unicode включает в себя все существующие, вымершие и искусственно созданные алфавиты мира, а также множество математических, музыкальных, химических и прочих символов.

▪



Пример . Представьте в форме шестнадцатеричного кода слово «ЭВМ» в 4-х кодировках.

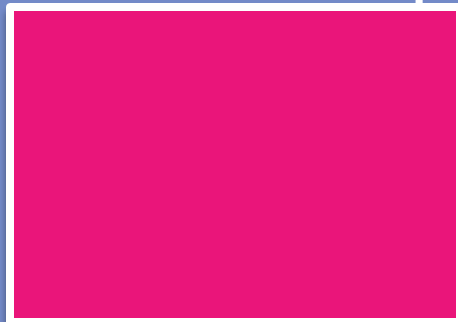
Последовательности десятичных кодов слова «ЭВМ» в различных кодировках составляем на основе кодировочных таблиц:

КОИ8-Р:

CP1251:

Mac:

ISO:



Пример . Представьте в форме шестнадцатеричного кода слово «ЭВМ» в 4-х кодировках.

Последовательности десятичных кодов слова «ЭВМ» в различных кодировках составляем на основе кодировочных таблиц:

КОИ8-Р: 252 247 237

CP1251: 221 194 204

Mac: 157 130 140

ISO: 205 178 188

Пример . Представьте в форме шестнадцатеричного кода слово «ЭВМ» в 4-х кодировках.

Переводим с помощью калькулятора последовательности кодов из десятичной системы в шестнадцатеричную:

КОИ8-Р:

CP1251:

Mac:

ISO:



Пример . Представьте в форме шестнадцатеричного кода слово «ЭВМ» в 4-х кодировках.

Переводим с помощью калькулятора последовательности кодов из десятичной системы в шестнадцатеричную:

| | |
|---------|----------|
| КОИ8-Р: | FC F7 ED |
| CP1251: | DD C2 CC |
| Mac: | 9B 82 8C |
| ISO: | CD B2 BC |

**Спасибо
за внимание!**

Автор

презентации:

Ширяева Ольга Мухадинновна,
Муниципальное Общеобразовательное
Учреждение
«Средняя общеобразовательная школа №2»,
учитель информатики,
1 квалификационной категории.

<http://www.shiryaeva.86sch2-nyagan.edusite.ru>

В презентации использованы материалы учебного пособия: «ПРАКТИКУМ по информатике и информационным технологиям», Н. Угринович, Л. Босова, Н. Михайлова - М.: БИНОМ. Лаборатория знаний, 2004.

