

---

# Кодирование текстовой информации



---

# Содержание

- Вопросы для повторения
  - Двоичное кодирование текстовой информации в компьютере
  - Кодовая таблица
  - Код ASCII
  - Принцип последовательного кодирования алфавита
  - Различные кодировки знаков
  - Примеры
  - Тестовая проверочная работа
-

---

# Вопросы для повторения

- Какие виды информации по способу представления вы знаете?
  - Кодирование каких видов информации вы изучили на прошлых уроках?
  - Вспомните главную форму информатики?
  - Что принято за единицу количества информации?
  - Назовите производные единицы измерения количества информации.
-

# Двоичное кодирование текстовой информации в компьютере

Для представления текстовой информации достаточно 256 различных символов.

$N = 2^l$ ,  $256 = 2^8$ ,  $2^8 = 2^l$ ,  $l = 8$  битов

Для кодирования каждого знака требуется количество информации, равное 8 битам.

Для представления текста в памяти компьютера необходимо представить его в двоичной знаковой системе.

Каждому знаку необходимо поставить в соответствие уникальный двоичный код в интервале от 00000000 до 11111111 (в десятичном коде от 0 до 255)

---

## Кодовая таблица

Для представления символов и соответствующих им кодов используется *кодовая таблица*.

В качестве стандарта во всем мире принята таблица ASCII (American Standard Code for Information Interchange – Американский стандартный код для обмена информацией). Условно таблица разделена на части:

- от 0 до 32 коды соответствуют операциям;
  - с 33 по 127 соответствуют символам латинского алфавита, цифрам, знакам арифметических операций и знакам препинания;
  - со 128 по 255 являются национальными.
-

|    |             |    |   |          |    |   |          |    |   |          |
|----|-------------|----|---|----------|----|---|----------|----|---|----------|
| 32 | 00100000    | 56 | 8 | 00111000 | 80 | P | 01010000 | 04 | h | 01101000 |
| 33 | ! 00100001  | 57 | 9 | 00111001 | 81 | Q | 01010001 | 05 | i | 01101001 |
| 34 | " 00100010  | 58 | : | 00111010 | 82 | R | 01010010 | 06 | j | 01101010 |
| 35 | # 00100011  | 59 | ; | 00111011 | 83 | S | 01010011 | 07 | k | 01101011 |
| 36 | \$ 00100100 | 60 | < | 00111100 | 84 | T | 01010100 | 08 | l | 01101100 |
| 37 | % 00100101  | 61 | = | 00111101 | 85 | U | 01010101 | 09 | m | 01101101 |
| 38 | & 00100110  | 62 | > | 00111110 | 86 | V | 01010110 | 10 | n | 01101110 |
| 39 | ' 00100111  | 63 | ? | 00111111 | 87 | W | 01010111 | 11 | o | 01101111 |
| 40 | ( 00101000  | 64 | @ | 01000000 | 88 | X | 01011000 | 12 | p | 01110000 |
| 41 | ) 00101001  | 65 | A | 01000001 | 89 | Y | 01011001 | 13 | q | 01110001 |
| 42 | * 00101010  | 66 | B | 01000010 | 90 | Z | 01011010 | 14 | r | 01110010 |
| 43 | + 00101011  | 67 | C | 01000011 | 91 | [ | 01011011 | 15 | s | 01110011 |
| 44 | , 00101100  | 68 | D | 01000100 | 92 | \ | 01011100 | 16 | t | 01110100 |
| 45 | - 00101101  | 69 | E | 01000101 | 93 | ] | 01011101 | 17 | u | 01110101 |
| 46 | . 00101110  | 70 | F | 01000110 | 94 | ^ | 01011110 | 18 | v | 01110110 |
| 47 | / 00101111  | 71 | G | 01000111 | 95 | _ | 01011111 | 19 | w | 01110111 |
| 48 | 0 00110000  | 72 | H | 01001000 | 96 | ' | 01100000 | 20 | x | 01111000 |
| 49 | 1 00110001  | 73 | I | 01001001 | 97 | a | 01100001 | 21 | y | 01111001 |
| 50 | 2 00110010  | 74 | J | 01001010 | 98 | b | 01100010 | 22 | z | 01111010 |
| 51 | 3 00110011  | 75 | K | 01001011 | 99 | c | 01100011 | 23 | { | 01111011 |
| 52 | 4 00110100  | 76 | L | 01001100 | 00 | d | 01100100 | 24 |   | 01111100 |
| 53 | 5 00110101  | 77 | M | 01001101 | 01 | e | 01100101 | 25 | } | 01111101 |
| 54 | 6 00110110  | 78 | N | 01001110 | 02 | f | 01100110 | 26 | ~ | 01111110 |
| 55 | 7 00110111  | 79 | O | 01001111 | 03 | g | 01100111 | 27 | □ | 11111111 |

---

# Принцип последовательного кодирования алфавита:

В кодовой таблице ASCII латинские буквы (прописные и строчные) располагаются в алфавитном порядке. Расположение цифр также упорядочено по возрастанию значений.

Данное правило соблюдается и в других таблицах кодировки. Благодаря этому и в машинном представлении для символьной информации сохраняется понятие «алфавитный порядок».

---

# Различные кодировки знаков

В настоящее время существуют пять различных кодировок для русских букв

(Windows, MS-DOS, KOI-8, Mac, ISO), поэтому тексты созданные в одной кодировке не будут правильно отображаться в другой.

В последние годы широкое распространение получил новый международный стандарт кодирования текстовых символов Unicode, который отводит на каждый символ 2 байта (16 битов), что позволило закодировать многие алфавиты в одной таблице.

$N = 2^1$ ,  $N = 2^{16}$ ,  $N = 65536$



---

## Пример 1

Буква «i» в таблице кодировки символов имеет десятичный код 105. Что зашифровано последовательностью десятичных кодов:

108 105 110 107?

При расшифровке данной последовательности кодов не нужно обращаться к таблице кодировки символов.

Необходимо учесть принцип последовательного кодирования алфавитов и вспомнить порядок букв в латинском алфавите (...i, j, k, l, m, n, o, ...)

Закодировано: «link»

---

## Пример 2

С помощью последовательности десятичных кодов:

99 111 109 112 117 116 101 114 закодировано слово «computer». Какая последовательность десятичных кодов будет соответствовать этому же слову, записанному заглавными буквами?

При шифровке слова не пользоваться таблицей кодировки символов. Необходимо учесть, что разница между десятичным кодом строчной буквы латинского алфавита и десятичным кодом соответствующей заглавной буквы равна 32.

Если «с» имеет код 99, то «С» имеет код  $99-32=67$ .

COMPUTER – 67 79 77 80 85 84 69 82

---

## Пример 3

- Какое количество информации необходимо для кодирования одного символа компьютерного алфавита?
  - Слова «компьютер»?
  - Слова «алфавит»?
-

# Тестовая проверочная работа

## Вариант 1

1. Полный набор символов, используемый для кодирования, называют:  
1) шифром      2) алфавитом      3) синтаксисом      4) семантикой
2. Минимальным объектом, используемым для кодирования текста, является:  
1) бит      2) пиксель      3) символ      4) растр
3. Количество битов, необходимое для кодирования одного символа алфавита в коде Unicode, равно:  
1) 8      2) 16      3) 32      4) 256
4. Сообщение «урок» содержит следующее количество информации:  
1) 4 бита      2) 32 бита      3) 8 байт      4) 32 байта
5. Какие символы могут быть зашифрованы кодами таблицы ASCII 119 и 251?  
1) «д» и «ш»      2) «j» и «s»      3) «d» и «D»      4) «w» и «ы»

## Вариант 2

1. Таблица кодировки символов устанавливает соответствие между:  
1) символами, их десятичными номерами и двоичными кодами  
2) символами разных алфавитов  
3) символами и количеством байтов, которые они занимают  
4) символами и номерами ячеек памяти, которые они занимают
2. Минимальным объектом, используемым для кодирования текста, является:  
1) бит      2) пиксель      3) символ      4) растр
3. Количество битов, необходимое для кодирования одного символа алфавита в коде ASCII, равно:  
1) 1      2) 8      3) 32      4) 256
4. В сообщении «класс!» содержится следующее количество информации:  
1) 5 битов      2) 6 битов      3) 5 байтов      4) 6 байтов
5. Код буквы «О» в некоторой таблице кодировки символов равен 142. Что зашифровано с помощью последовательности кодов: 145 143 142 144 146?  
1) SPORT      2) SPORТ      3) СПОРТ      4) SNORT

# Ответы для проверки

- Вариант 1

|   |   |   |   |   |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| 2 | 3 | 2 | 2 | 4 |

- Вариант 2

|   |   |   |   |   |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 2 | 4 | 3 |

# Домашнее задание

1. Повторить изученный материал, воспользовавшись конспектом лекции и учебником стр. 49- 52, ответить на контрольные вопросы.
2. Выполнить задания 2.1, 2.2 по учебнику
3. В таблице ниже представлена часть кодовой таблицы ASCII:

| Символ         | <b>1</b> | <b>5</b> | <b>A</b> | <b>B</b> | <b>Q</b> | <b>a</b> | <b>b</b> |
|----------------|----------|----------|----------|----------|----------|----------|----------|
| Десятичный код | 49       | 53       | 65       | 66       | 81       | 97       | 98       |
| 16-ричный код  | 31       | 35       | 41       | 42       | 51       | 61       | 62       |

Определить 16-ричный код символа “q”.