

Представление текстовой информации в компьютере.

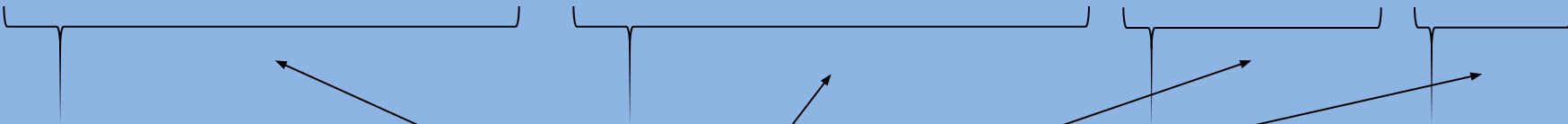
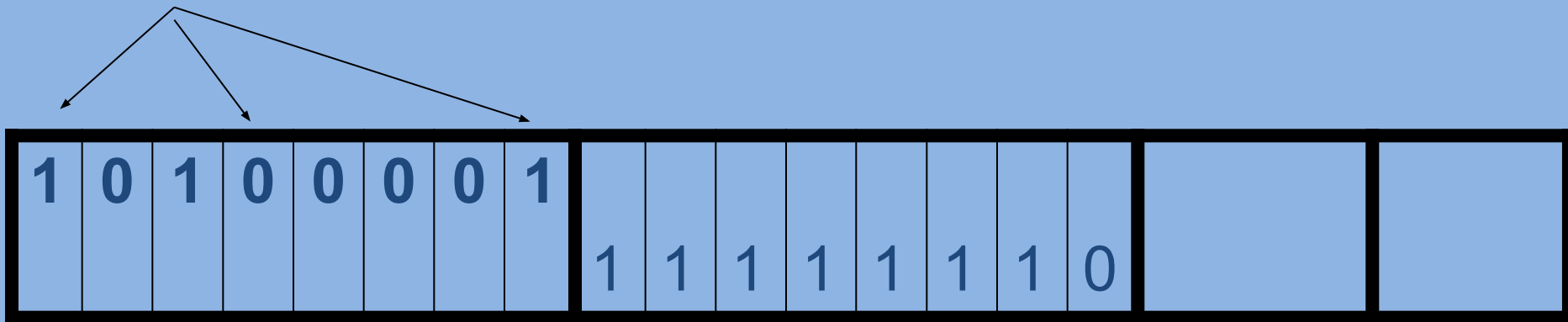
Бит-маленькая единица информации.

Работать с каждым битом отдельно, конечно, можно, но это малопродуктивно.

Обработкой информации в компьютере занимается специальная микросхема, которая называется процессор. Эта микросхема устроена так, что может обрабатывать группу битов одновременно (параллельно).

Байт - это группа из восьми битов.

Биты памяти



Байты памяти

В байте важно не только, включен бит или выключен, но и то, в каком месте стоят включенные биты. Байты 0000 0001, 0000 1000 и 1000 0000 - не одинаковые, а разные.

Если учесть, что важны не только нули и единицы, но и позиции, в которых они стоят, то с помощью одного байта можно выразить 256 различных единиц информации (от 0 до 255).

| Биты | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|--------------------------------|---|---|---|----|----|----|-----|-----|
| Количество кодируемых значений | 2 | 4 | 8 | 16 | 32 | 64 | 128 | 256 |

$$N=2^m ,$$

где N - количество независимых кодируемых значений;

m - разрядность двоичного кодирования, принятая в данной системе.

Единицы измерения информации.

$$1 \text{ Kb (кило)} = 2^{10} \text{ b} = 1.024 \text{ b}$$

$$1 \text{ Mb (мега)} = 2^{10} \text{ Kb} = 2^{20} \text{ b} = 1.048.576 \text{ b}$$

$$1 \text{ Gb (гига)} = 2^{10} \text{ Mb} = 2^{30} \text{ b} = 1.073.741.824 \text{ b}$$

$$1 \text{ Tb (тера)} = 2^{10} \text{ Gb} = 2^{40} \text{ b} = 1.099.511.627.776 \text{ b}$$

Кодирование текстовой информации.

- В русском языке 33 буквы (символа) - для их кодирования достаточно 33 различных байтов.
- Если мы хотим различить прописные (заглавные) и строчные буквы, то потребуется 66 байтов.
- Для строчных и прописных букв английского языка хватит еще 52 символа - получается 118.

- Добавим сюда цифры (от 0 до 9), все возможные знаки препинания: точку, запятую, тире, восклицательный и вопросительный знаки.
- Добавим скобки: круглые, квадратные и фигурные, а также знаки математических операций: +, -, =, /, *.
- Добавим специальные символы, например такие, как: %, #, &, @, - мы видим, что все их можно выразить восемью битами, и при этом еще останутся свободные коды, которые можно использовать для других целей.

Таблица кодов ASCII

Для кодирования текстовой информации

используется таблица символов

ASCII (American Standard Code of Information Interchange).

| КОД | СИМВОЛ | КОД | СИМВОЛ | КОД | СИМВОЛ | КОД | СИМВОЛ | КОД | СИМВОЛ | КОД | СИМВОЛ |
|-----|--------|-----|--------|-----|--------|-----|--------|-----|--------|-----|--------|
| 32 | Пробел | 48 | . | 64 | @ | 80 | P | 96 | ' | 112 | p |
| 33 | ! | 49 | 0 | 65 | A | 81 | Q | 97 | a | 113 | q |
| 34 | " | 50 | 1 | 66 | B | 82 | R | 98 | b | 114 | r |
| 35 | # | 51 | 2 | 67 | C | 83 | S | 99 | c | 115 | s |
| 36 | \$ | 52 | 3 | 68 | D | 84 | T | 100 | d | 116 | t |
| 37 | % | 53 | 4 | 69 | E | 85 | U | 101 | e | 117 | u |
| 38 | & | 54 | 5 | 70 | F | 86 | V | 102 | f | 118 | v |
| 39 | ' | 55 | 6 | 71 | G | 87 | W | 103 | g | 119 | w |
| 40 | (| 56 | 7 | 72 | H | 88 | X | 104 | h | 120 | x |
| 41 |) | 57 | 8 | 73 | I | 89 | Y | 105 | i | 121 | y |
| 42 | * | 58 | 9 | 74 | J | 90 | Z | 106 | j | 122 | z |
| 43 | + | 59 | : | 75 | K | 91 | [| 107 | k | 123 | { |
| 44 | , | 60 | ; | 76 | L | 92 | \ | 108 | l | 124 | |
| 45 | - | 61 | < | 77 | M | 93 |] | 109 | m | 125 | } |
| 46 | . | 62 | > | 78 | N | 94 | ^ | 110 | n | 126 | ~ |
| 47 | / | 63 | ? | 79 | O | 95 | _ | 111 | o | 127 | DEL |

KOI-8

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|----|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | | ⊙ | ⊕ | ♥ | ♦ | ♣ | ♠ | ● | ○ | | | | | | | |
| 1 | ▶ | ◀ | | | | | | ↑ | ↓ | → | ← | ↔ | ▲ | ▼ | | |
| 2 | ! | " | # | \$ | % | & | ' | { | } | * | + | , | - | . | / | |
| 3 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | : | ; | < | = | > | ? |
| 4 | @ | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
| 5 | P | Q | R | S | T | U | V | W | X | Y | Z | [| \ |] | ^ | _ |
| 6 | · | a | b | c | d | e | f | g | h | i | j | k | l | m | n | o |
| 7 | p | q | r | s | t | u | v | w | x | y | z | { | | } | ~ | □ |
| 8 | А | Б | В | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П |
| 9 | Р | С | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я |
| A | а | б | в | г | д | е | ж | з | и | й | к | л | м | н | о | п |
| B | ⊗ | ⊙ | ⊕ | ⊖ | ⊗ | ⊘ | ⊙ | ⊚ | ⊛ | ⊜ | ⊝ | ⊞ | ⊟ | ⊠ | ⊡ | ⊣ |
| C | Ⓕ | Ⓖ | Ⓗ | Ⓘ | Ⓚ | Ⓛ | Ⓜ | Ⓝ | Ⓟ | Ⓡ | Ⓢ | Ⓣ | Ⓥ | Ⓦ | Ⓩ | ⓑ |
| D | Ⓗ | Ⓖ | Ⓗ | Ⓘ | Ⓚ | Ⓛ | Ⓜ | Ⓝ | Ⓟ | Ⓡ | Ⓢ | Ⓣ | Ⓥ | Ⓦ | Ⓩ | ⓑ |
| E | р | с | т | у | ф | х | ц | ч | ш | щ | ъ | ы | ь | э | ю | я |
| F | ё | ё | ≥ | ≤ | |] | ÷ | ≈ | ° | · | √ | » | » | ■ | □ | |

Win-1251

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | A | B | C | D | E | F |
|---|---|---|---|---|----|---|---|---|---|---|----|----|----|----|----|----|
| 0 | | | | ⊙ | ⊕ | ⊗ | ⊖ | | ° | | | | | | | |
| 1 | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 2 | | ! | " | # | \$ | % | & | | (|) | ' | + | | | | / |
| 3 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | : | ; | < | = | > | ? |
| 4 | @ | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
| 5 | P | Q | R | S | T | U | V | W | X | Y | Z | [| \ |] | ^ | _ |
| 6 | · | a | b | c | d | e | f | g | h | i | j | k | l | m | n | o |
| 7 | p | q | r | s | t | u | v | w | x | y | z | { | | } | ~ | □ |
| 8 | Б | Г | Д | Е | Ж | З | И | Й | К | Л | М | Н | О | П | Р | С |
| 9 | Т | У | Ф | Х | Ц | Ч | Ш | Щ | Ъ | Ы | Ь | Э | Ю | Я | а | б |
| A | в | г | д | е | ж | з | и | й | к | л | м | н | о | п | р | с |
| B | т | у | ф | х | ц | ч | ш | щ | ъ | ы | ь | э | ю | я | а | б |
| C | в | г | д | е | ж | з | и | й | к | л | м | н | о | п | р | с |
| D | т | у | ф | х | ц | ч | ш | щ | ъ | ы | ь | э | ю | я | а | б |
| E | в | г | д | е | ж | з | и | й | к | л | м | н | о | п | р | с |
| F | т | у | ф | х | ц | ч | ш | щ | ъ | ы | ь | э | ю | я | а | б |

D=68=1000100

Какие слова закодированы
следующими байтами?

67 79 77 80 85 84 69 82

99 111 109 112 117 116 101 114

COMPUTER

computer

Слово WATER, закодированное в
кодах ASCII, выглядит так:

| | | | | |
|-----|-----|-----|-----|-----|
| W | A | T | E | R |
| 087 | 065 | 084 | 069 | 082 |

А если учесть, что в двоичном представлении для кодирования каждого символа используется 8 – разрядный код, то получится двоичный код длиной в 40 символов.

W

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
|---|---|---|---|---|---|---|---|

A

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
|---|---|---|---|---|---|---|---|

T

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
|---|---|---|---|---|---|---|---|

E

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
|---|---|---|---|---|---|---|---|

R

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
|---|---|---|---|---|---|---|---|