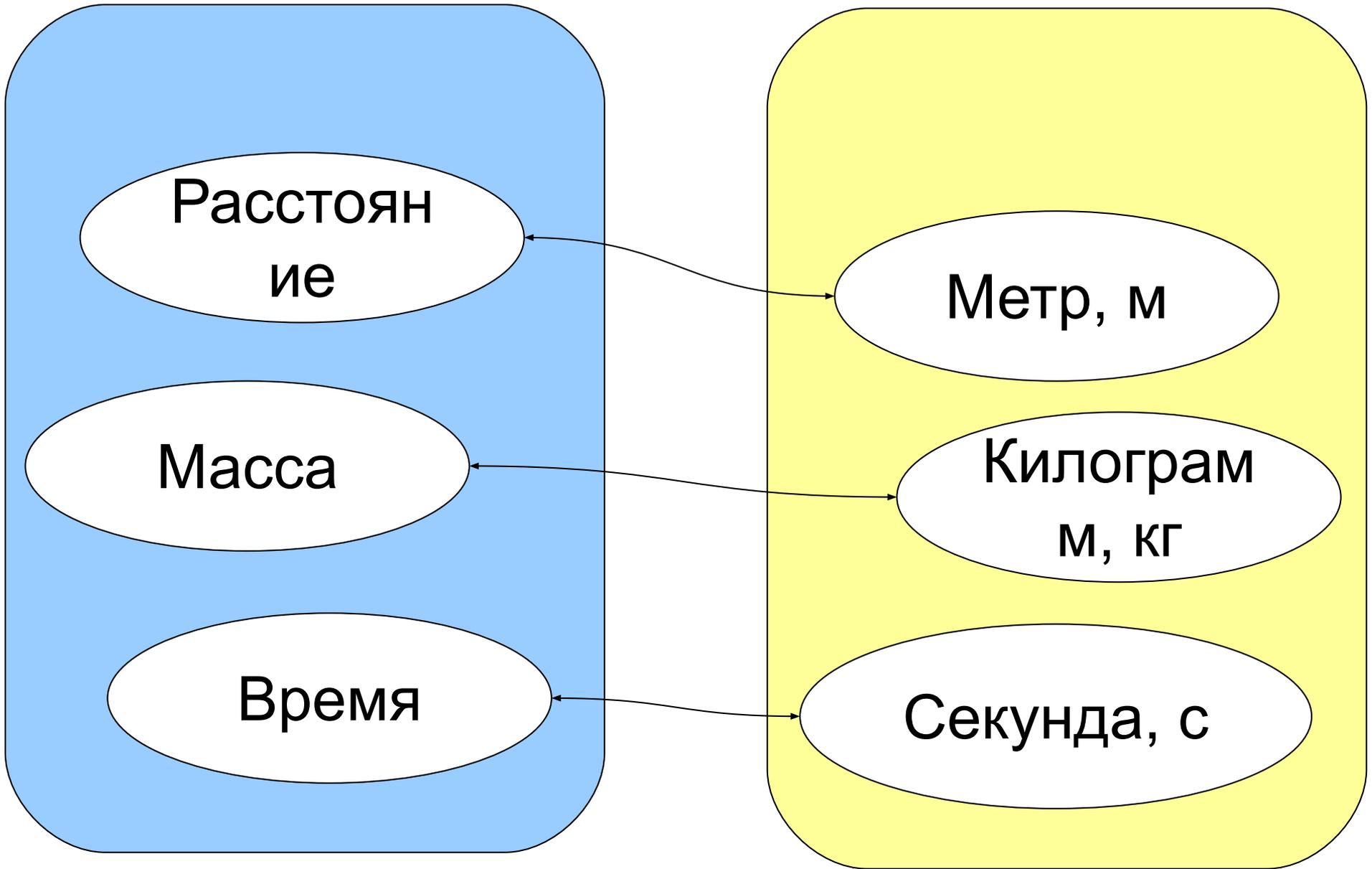


Измерение информации § 4

Алфавитный
подход

Содержательный
подход

Не связан с
содержанием
текста



Алфавитный подход

Алфавит – набор знаков, используемых при кодировании информации с помощью некоторого языка.

Примеры:

АБВГДЕЖЗИКЛМНОПРСТУФХЦЧШЩЪЫЬЭЮЯ 32

ABCDEFGHIJKLMNOPQRSTUVWXYZ 26

× ○

2

0123456789

10

Мощность алфавита – количество символов алфавита.



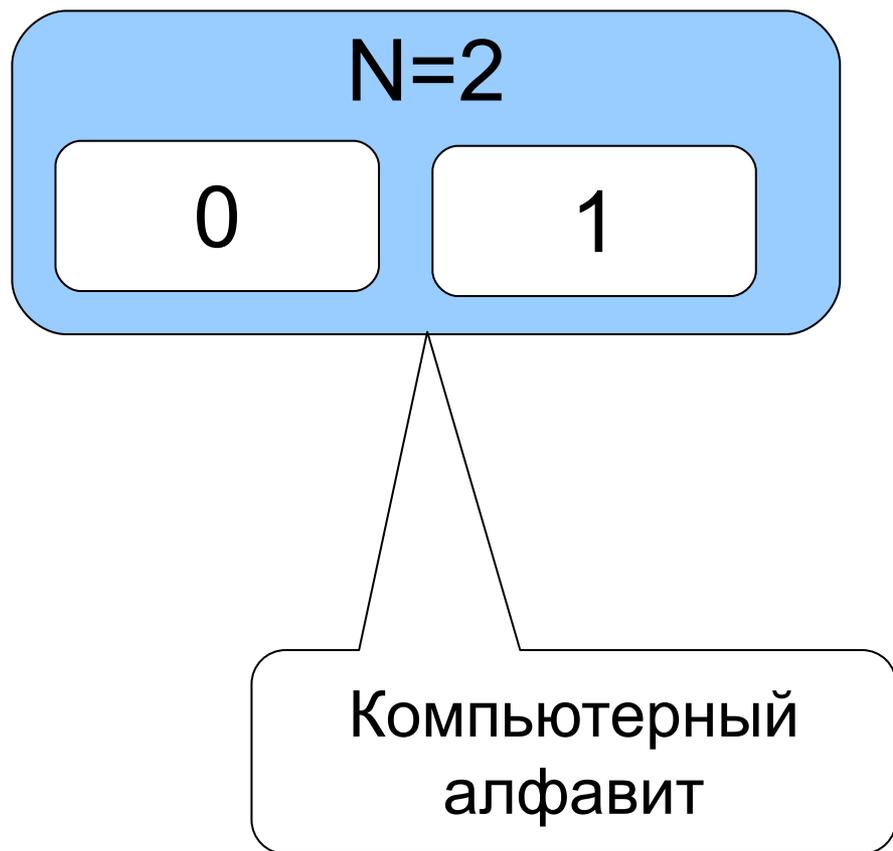
Все символы несут одинаковую информацию:

$$N = 2^I$$

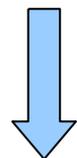
МОЩНОСТЬ
алфавита

информационный
вес символа

Информационный вес СИМВОЛА



$$N=2^I$$



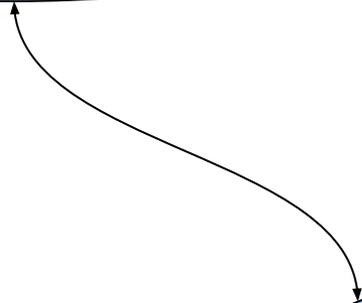
$$2=2^1$$

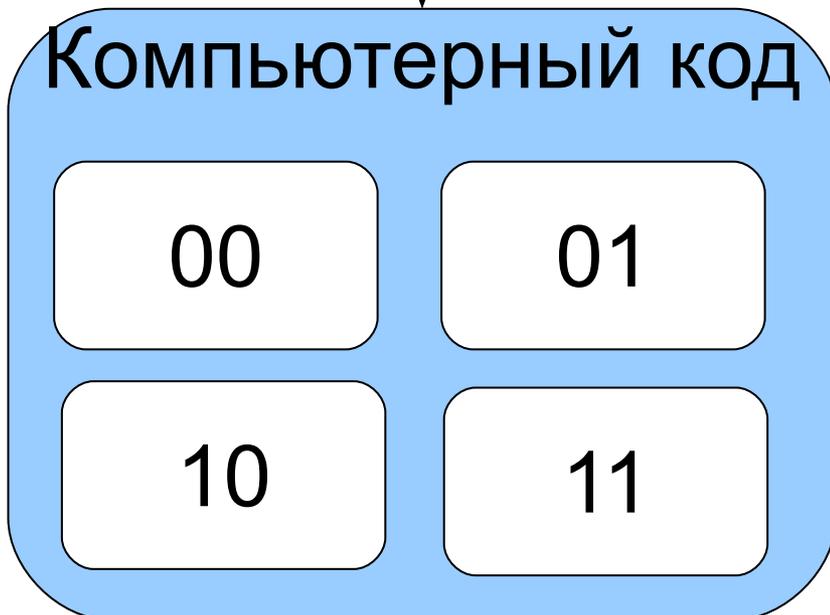
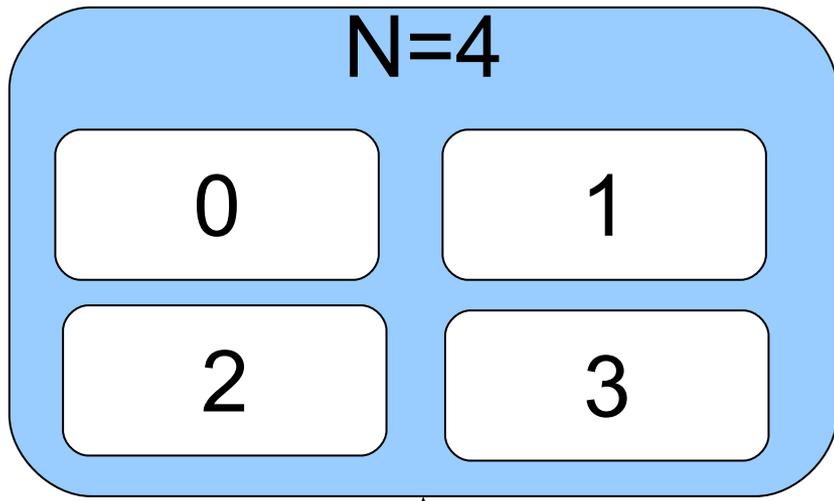
$$I=1 \text{ бит}$$

Информационный вес
символа двоичного
алфавита

Увеличение
мощности
алфавита

Увеличение
информационн
ого
веса символа





$$N=2^I$$



$$4=2^I$$

$$I=2 \text{ бита}$$

N	2	4	8	16
l	1	2	3	4

Зависимость между
мощностью
алфавита и
информационным весом
символа (количеством
знаков в коде)

$$N=2^l$$

Формула
Хартли

Информационный объём текста

Количество символов

X

Информационный вес одного символа

Алфавитный подход

Задача. Определить объем информации в сообщении

ПРИВЕТВАСЯ

для кодирования которого используется русский алфавит (только заглавные буквы).

Решение:

- считаем все символы (здесь **10** символов)
- мощность алфавита – 32 символа ($32=2^5$)
- 1 символ несет **5 бит** информации

Ответ: $10 \cdot 5 \text{ бит} = 50 \text{ бит}$

Подготовка текстового документа

С помощью
ручки и бумаги

Алфавит

Знаки
препин.

С помощью
компьютера

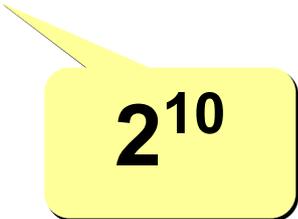
Комп.
алфавит

$N=256$

$I=8$ бит

Единицы измерения

- 1 байт (*byte*) = 8 бит
- 1 Кб (килобайт) = 1024 байта
- 1 Мб (мегабайт) = 1024 Кб
- 1 Гб (гигабайт) = 1024 Мб
- 1 Тб (терабайт) = 1024 Гб
- 1 Пб (петабайт) = 1024 Тб



2^{10}

Домашнее задание: §4