

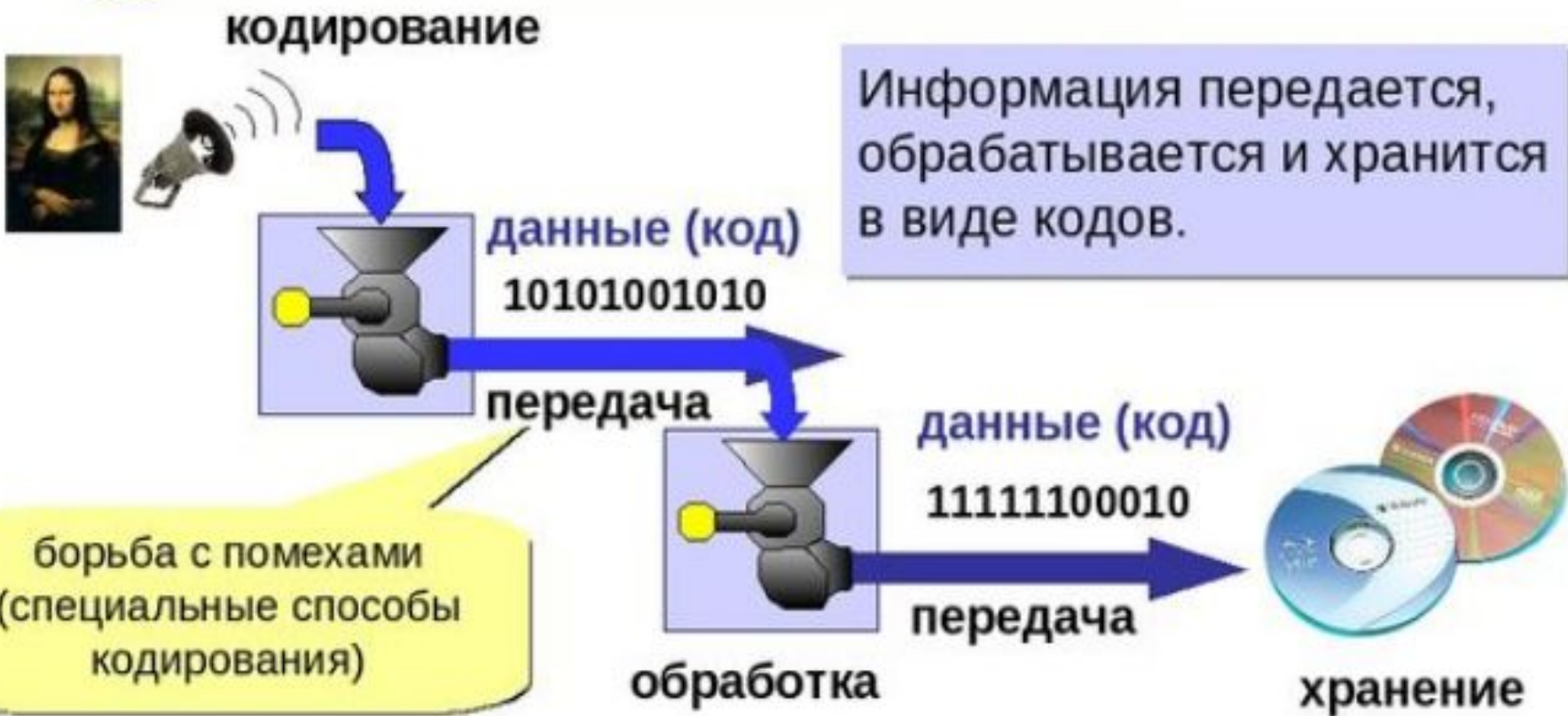
Кодирование информации

— [Тема 1. Кодирование информации](#)

Что такое кодирование?

Кодирование – это запись информации с помощью некоторой знаковой системы (языка).

? Зачем кодируют информацию?

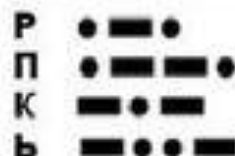
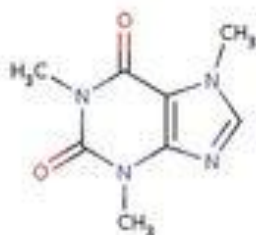


ЯЗЫКИ

Язык – знаковая система, используемая для хранения и передачи информации.

- **естественные** (русский, английский, ...)
есть правила и исключения
- **формальные** (строгие правила)

$$E = mc^2$$



$$16 = 10_{16} = 20_8 = 10000_2$$

```
program qq;  
begin  
  writeln("Привет!");  
end.
```

Грамматика – правила по которым из символов алфавита строятся слова.

Синтаксис – правила, по которым из слов строятся предложения.

Азбука Морзе

Задача 1. Закодируйте свое имя с помощью азбуки Морзе.

А	● —	П	● — — ●	Ь	— ● ● —
Б	— ● ● ●	Р	● — ●	Ы	— ● — —
В	● — —	С	● ● ●	Й	● — — —
Г	— — ●	Т	—		
Д	— ● ●	У	● ● —	1	● — — — —
Е	●	Ф	● ● — ●	2	● ● — — —
Ж	● ● ● —	Х	● ● ● ●	3	● ● ● — —
З	— — — ● ●	Ц	— ● — ●	4	● ● ● ● —
И	● ●	Ч	— — — — ●	5	● ● ● ● ●
К	— ● — —	Ш	— — — — —	6	— ● ● ● ●
Л	● — — ● ●	Щ	— — — ● —	7	— — — ● ● ●
М	— — —	Э	● ● — — ● ●	8	— — — — ● ●
Н	— ●	Ю	● ● — — —	9	— — — — — ●
О	— — — —	Я	● — — ● —	0	— — — — — —

ВАСЯ



Код неравномерный, нужен разделитель!

Кодовые таблицы

Задача 2. Закодируйте свое имя с помощью кодовой таблицы (*Windows-1251*):

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
C	A	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
D	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я

В А С Я

ВАСЯ

C2 C0 D1 DF



Код равномерный, разделитель **HE** нужен!

Цели и способы кодирования

Текст:

- в России: *Привет, Вася!*
- Windows-1251: *CFF0E8E2E52C20C2E0F1FF21*
- передача за рубеж (транслит): *Privet, Vasya!*
- стенография:
- шифрование: *Рсйгжу-!Гбта”*

Числа:

- для вычислений: *25*
- прописью: *двадцать пять*
- римская система: *XXV*



Как зашифровано?



Информация (смысл сообщения) может быть закодирована разными способами!

Декодирование

Декодирование – это восстановление сообщения из последовательности кодов.

М	А	Ы	Л	У	пробел
00	1	01	0	10	11

МАМА МЫЛА ЛАМУ f 00 1 00 1 11 00 01 0 1 11 0 1 00 10

Приняли сообщение:

0010011100010111010010 → ???

ЛЛАЛЛАААЛЛЛАЛАААЛЛАЛЛАЛ

Равномерные коды

Равномерные коды – все кодовые слова (коды отдельных букв) имеют одинаковую длину.

М	А	Ы	Л	У	пробел
000	001	010	011	100	101

МАМА МЫЛА ЛАМУ:

000 001 000 001 101 000 010 011 001 101 011 001 000
100



Равномерные коды позволяют однозначно декодировать сообщения!

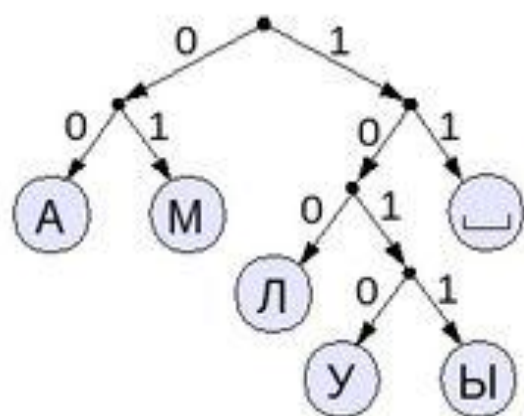


сообщения получаются длинными

Неравномерные коды

кодированные слова имеют разную длину

М	А	Ы	Л	У	пробел
01	00	1011	100	1010	11



0100010011011011100001110000011010
М А М А Л М Ы Л А Л А М У

Префиксный код – ни одно кодированное слово не совпадает с началом другого кодированного слова (условие Фано).



Любой префиксный код позволяет однозначно декодировать сообщения!

Постфиксные коды

Постфикс = окончание слова.

Постфиксный код – ни одно кодовое слово не совпадает с концом другого кодового слова («обратное» условие Фано).

М	А	Ы	Л	У	пробел
10	00	1101	001	0101	11



Любой постфиксный код позволяет однозначно декодировать сообщения (с конца)!



для декодирования нужно получить всё сообщение целиком

Задачи на построение кода

Для передачи по каналу связи сообщения, состоящего только из букв А, Б, В, Г, решили использовать неравномерный по длине код:

А	Б	В	Г
1	000	001	?

Как нужно закодировать букву Г, чтобы длина кода была минимальной и допускалось однозначное разбиение кодированного сообщения на буквы?

1) 00

2) 01

3) 11

4) 010

Решение:

1) для букв А-Б-В выполняются условие Фано

2) при Г=00 условие Фано нарушится (пары Г-Б, Г-В)

3) при Г=01 условие Фано выполняется

4) при Г=11 условие Фано нарушится (пара А-Г)

5) при Г=010 условие Фано выполняется (но длиннее 01)

