

Кодирование информации

§ 4. Язык – средство кодирования

Определения

Кодирование — это представление информации в форме, пригодной для её хранения, передачи и автоматической обработки.

Код — это правило, по которому сообщение преобразуется в цепочку знаков.

Язык — это система знаков и правил, используемая для записи и передачи информации.

Естественные языки – сформировались в результате развития общества.

Иероглифы


Египетское письмо	
	рука
	дом
	кобра
	лев
	вода

Иероглифы (Китай)	
日	солнце
月	луна
雨	дождь
山	гора
马	лошадь

Алфавитное письмо

Алфавит — это набор знаков, который используется в языке.


Мощность алфавита — это количество знаков в алфавите.

 Какова мощность русского алфавита? латинского?

АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЪЫЬЭЮЯ
0123456789 . , ; ? ! - : ... « » ()

МОЩНОСТЬ 56

Какие бывают языки?

<ul style="list-style-type: none"> • русский • английский • китайский • шведский • суахили • ... 	$y = 3 \sin x + 1$ $2H_2 + O_2 = 2H_2O$  <p>1. e2-e4 e7-e5...</p>

Формальный язык – это язык, в котором однозначно определяется значение каждого слова, а также правила построения предложений и придания им смысла.

Сообщения

Сообщение — это любая последовательность символов некоторого алфавита.



Сколько различных сообщений длины L можно построить, используя алфавит мощностью M ?

Комбинаторика — это наука, изучающая комбинации объектов.

Пример: алфавит $\{0, 1\}$.

Сообщения длины 2:

00 01 10 11

всего 4

Сообщения

Пример: алфавит $\{ @, \#, \$, \% \}$.

Сообщения длины 1: @ # \$ %.

всего 4

Сообщения длины 2:

@@	@#	@\$	@%
#@	##	#\$	#%
\$@	\$#	\$\$	\$%
%@	%#	%%	%%

всего 16



Сколько сообщений длины L ?

Количество возможных сообщений

Если алфавит языка состоит из M символов (имеет мощность M), количество различных сообщений длиной L знаков равно

$$N = M^L$$

Сколько

- возможных 5-буквенных слов в русском языке?
- возможных 3-буквенных слов в английском языке?
- возможных сообщений длиной L символов в алфавите $\{+, -\}$?

$$33^5$$

$$26^3$$

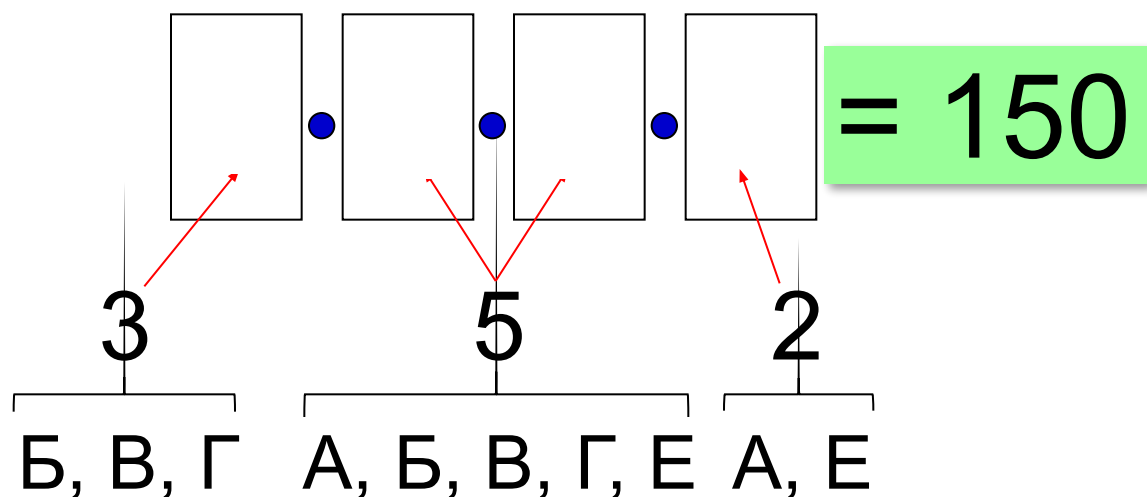
$$2^L$$

Правило умножения

Задача. Сколько различных сообщений длиной 4 знака можно записать с помощью алфавита

$\{A, B, B, \Gamma, E\}$

если слова должны начинаться с согласной буквы и заканчиваться на гласную?



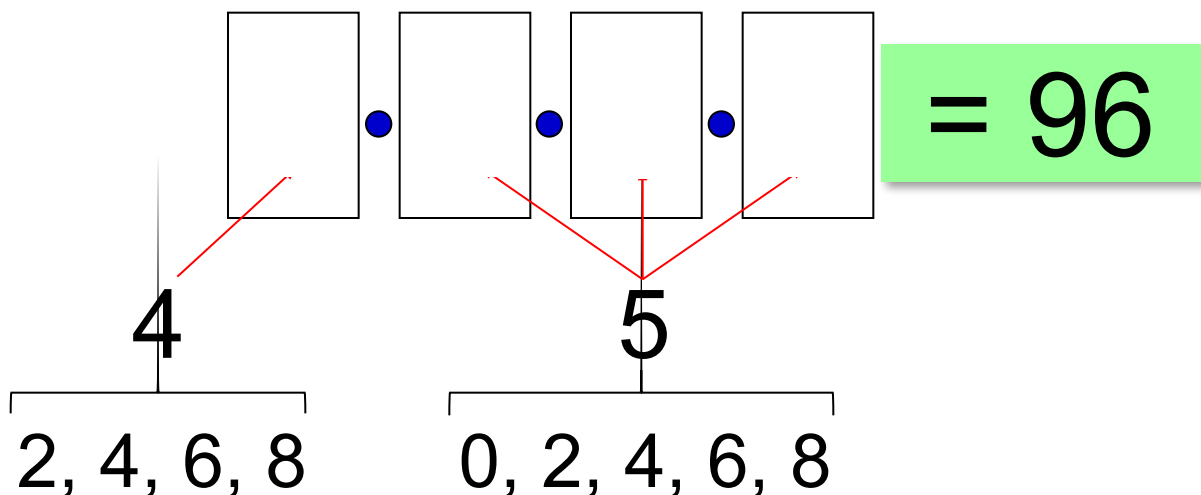
$$N = M_1 \cdot M_2 \cdot M_3 \cdot M_4$$



Правило умножения!

Правило умножения

Задача. Сколько существует четырёхзначных чисел, составленных из чётных цифр, в которых **цифры не повторяются**?



одна цифра уже
использована!

Правило сложения

Задача. Сколько сообщений длиной от 2 до 5 символов можно записать с помощью алфавита $\{0, 1\}$?

$$\begin{array}{ll} L = 2: & N_2 = 2^2 = 4 \\ L = 3: & N_3 = 2^3 = 8 \\ L = 4: & N_4 = 2^4 = 16 \\ L = 5: & N_5 = 2^5 = 32 \end{array}$$

$$N = 4 + 8 + 16 + 32 = 60$$

$$N = N_2 + N_3 + N_4 + N_5$$



Правило сложения!