



**ИНФОРМАТИКА**

**СПО**

# ОБРАБОТКА ИНФОРМАЦИИ

ИНФОРМАЦИЯ И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ

# КЛЮЧЕВЫЕ СЛОВА

- ◆ информационный процесс
- ◆ обработка информации
- ◆ кодирование
- ◆ код
- ◆ пре́фиксный код
- ◆ метод половинного деления

# ИНФОРМАЦИОННЫЙ ПРОЦЕСС

**Информационный процесс** — совокупность последовательных действий (операций), производимых над информацией (в виде данных, идей, гипотез, теорий) для получения какого-либо результата (достижения цели).

обработка информации

хранение информации

передача информации

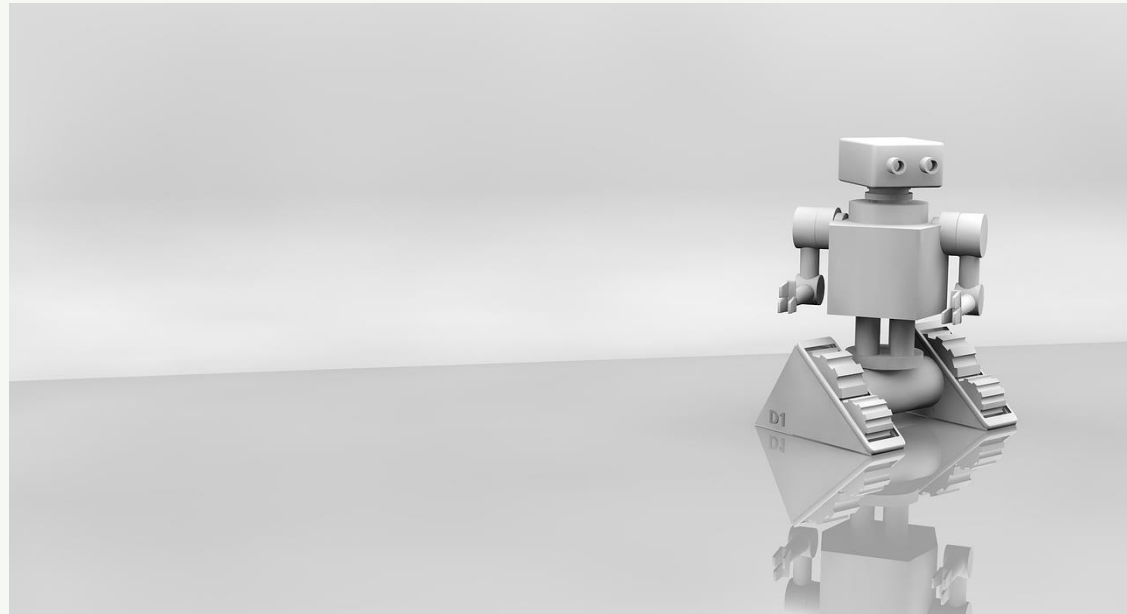
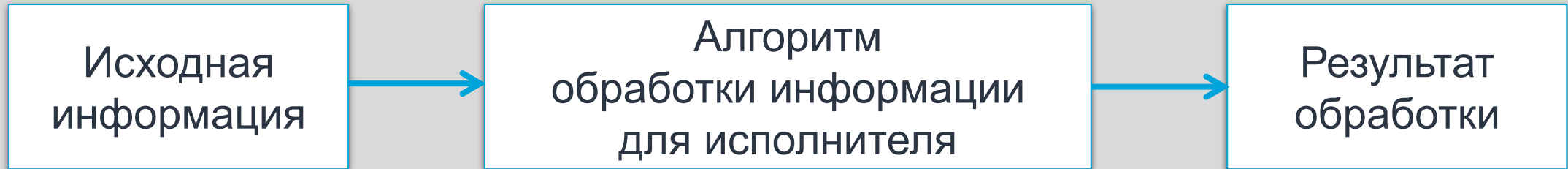
# ОБРАБОТКА ИНФОРМАЦИИ

Обработка информации — целенаправленный процесс изменения содержания или формы представления информации.



# СХЕМА ПРОЦЕССА ОБРАБОТКИ ИНФОРМАЦИИ

В процессе обработки информации всегда решается некоторая информационная задача.



**Исполнитель** – человек или компьютер, который осуществляет обработку информации

**Алгоритм** – последовательность действий, которую нужно выполнить, чтобы достичь нужного результата

# КОДИРОВАНИЕ ИНФОРМАЦИИ

**Кодирование** — обработка информации, заключающаяся в её преобразовании в некоторую форму, удобную для хранения, передачи, обработки информации в дальнейшем.

**Код** — система условных обозначений (кодовых слов), используемых для представления информации.

**Кодовая таблица** — совокупность используемых кодовых слов и их значений.



# АЗБУКА МОРЗЕ

**Азбука Морзе**, названная так в честь американского изобретателя и художника Сэмюэля Морзе, – самый известный пример неравномерного кода, в котором цифры и буквы алфавита представляются последовательностями длинных («тире») и коротких («точек») сигналов.

Сигналы отделяются друг от друга паузами — отсутствием сигналов.

Фактически, пауза является третьим знаком в азбуке Морзе, а сам код — троичным.

A	● —	U	● ● —
B	— ● ● ●	V	● ● ● —
C	— — ● — ●	W	● — — —
D	— ● ●	X	— ● ● —
E	●	Y	— ● — —
F	● ● — ●	Z	— — ● ●
G	— — ●		
H	● ● ● ●		
I	● ●		
J	● — — — —		
K	— ● — —	1	● — — — — —
L	● — ● ●	2	● ● — — — —
M	— — —	3	● ● ● — — —
N	— ●	4	● ● ● ● —
O	— — — —	5	● ● ● ● ●
P	● — — ●	6	— ● ● ● ●
Q	— — — ● —	7	— — — ● ● ●
R	● — — ●	8	— — — — ● ●
S	● ● ●	9	— — — — — ●
T	—	0	— — — — — —



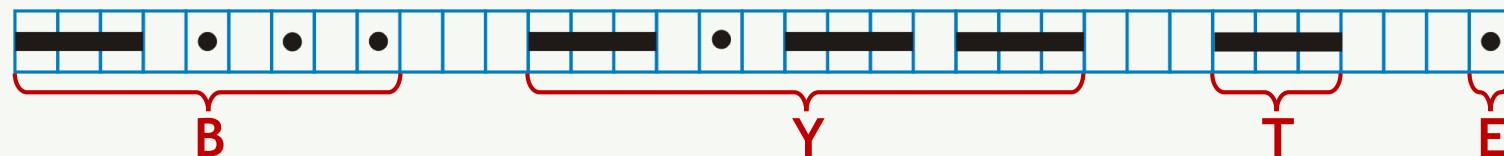
# МЕЖДУНАРОДНАЯ АЗБУКА МОРЗЕ

## Правила кода Морзе

1. Длина точки – одна единица.
2. Тире – три единицы.
3. Пауза между частями одного знака – одна единица.
4. Пауза между знаками – три единицы.
5. Пауза между словами – семь единиц.

A	• —	U	• • —
B	— • • •	V	• • • —
C	— • — •	W	• — — •
D	— • •	X	— • • —
E	•	Y	— • — —
F	• • — •	Z	— — • •
G	— — •		
H	• • • •		
I	• •		
J	• — — —		
K	— • — —		
L	• — • •	1	• — — — —
M	— —	2	• • — — —
N	— •	3	• • • — —
O	— — —	4	• • • • —
P	• — — •	5	• • • • •
Q	— — • —	6	• • • • •
R	• — • •	7	• — — • •
S	• • •	8	• — — — •
T	—	9	• — — — —
		0	— — — — —

*Расшифруйте слово, закодированное с помощью азбуки Морзе, представленное на «временной» шкале следующим образом:*





# ЗАДАЧА 1

Кодовый замок имеет три кольца с цифрами от 0 до 9. Сколько различных комбинаций можно на нем закодировать?



**Решение:**

0 1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8 9

0 1 2 3 4 5 6 7 8 9

Всего:  $10 \cdot 10 \cdot 10 = 1000$   
вариантов

# ПРАВИЛО УМНОЖЕНИЯ

Если элемент  $A$  можно выбрать  $n$  способами, и при любом выборе  $A$  элемент  $B$  можно выбрать  $m$  способами, то пару  $(A, B)$  можно выбрать  $n \cdot m$  способами.



## ПРИМЕР 1

Светодиодная панель содержит восемь излучающих элементов, каждый из которых может светиться или красным, или жёлтым, или синим, или зелёным цветом. Сколько различных сигналов можно передать с помощью панели (все излучающие элементы должны гореть, порядок цветов имеет значение)?

Цвета для пары элементов (1, 2) можно выбрать  $4 \cdot 4 = 4^2 = 16$  способами;

Цвета для тройки элементов (1, 2, 3) можно выбрать  $16 \cdot 4 = 4^3 = 64$  способами и т. д.

Цвета для восьми элементов (1, 2, 3, 4, 5, 6, 7, 8) можно выбрать  $4^8 = 65536$  способами.



## ПРИМЕР 2

Имеющаяся информация должна быть закодирована в четырёхбуквенном алфавите {A, B, C, D}. Выясним, сколько существует различных последовательностей из 7 символов четырёхбуквенного алфавита {A, B, C, D}, которые содержат ровно пять букв A.

1	2	3	4	5	6	7
A	A	A	A	A		

Так как на 6-м и 7-м местах могут стоять любые из трёх оставшихся букв B, C, D, то всего существует 9 ( $3 \cdot 3 = 9$ ) разных семибуквенных последовательностей, в которых первые пять позиций заняты буквой A.

1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	
A	A	A	A			A			A	A	A	A	A				A	A	A	A	A

А сколько таких вариантов всего?



# КОМБИНАТОРИКА

В комбинаторике набор  $k$  элементов, выбранных из данного множества, содержащего  $n$  различных элементов, называется сочетанием из  $n$  по  $k$ .

Для вычисления значения этой величины применяется формула

$$C_n^k = \frac{n!}{k!(n-k)!}$$

Здесь  $n! = 1 \cdot 2 \cdot \dots \cdot n$ .



## ПРИМЕР 2

Множество состоит из мест записи символов в последовательности, его элементы можно обозначить 1, 2, 3, 4, 5, 6 и 7.

Требуется выбрать из этого множество пять мест для размещения буквы А (число сочетаний из 7 по 5).

$$C_7^5 = \frac{7!}{5!(7-5)!} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 1 \cdot 2} = \frac{6 \cdot 7}{2} = 21$$

Для каждого из этих 21 вариантов имеется 9 разных вариантов заполнения двух оставшихся мест.

Всего существует 189 ( $21 \cdot 9 = 189$ ) различных последовательностей из 7 символов четырёхбуквенного алфавита {А, В, С, D}, которые содержат ровно пять букв А.

# ПРЕФИКСНЫЙ КОД

Главное условие использования неравномерных кодов — возможность однозначного декодирования записанного с их помощью сообщения.

**Пре́фиксный код** — код со словом переменной длины, обладающий тем свойством, что никакое его кодовое слово не может быть началом другого (более длинного) кодового слова.

Префиксальный код

*0, 10, 11*

Не префиксальный код

*0, 10, 11, 100*

# ПРАВИЛА ФАНО

Для того чтобы сообщение, записанное с помощью неравномерного кода, однозначно декодировалось, достаточно, чтобы никакое кодовое слово не было началом другого (более длинного) кодового слова.

Для возможности однозначного декодирования достаточно выполнения одного из условий Фано — прямого или обратного.

Обратное условие Фано также является достаточным условием однозначного декодирования неравномерного кода. В нём требуется, чтобы никакой код не был окончанием другого (более длинного) кода.



**Роберт Марио Фано**

итальяно-американский учёный, известный по работам в области теории информации.



## ЗАДАЧА 2

Двоичные коды для 5 букв латинского алфавита представлены в таблице:

A	B	C	D	E
000	01	100	10	011

Какое сообщение (какой набор букв) закодировано с помощью этих кодов двоичной строкой: **0110100011000**.

**Решение:**

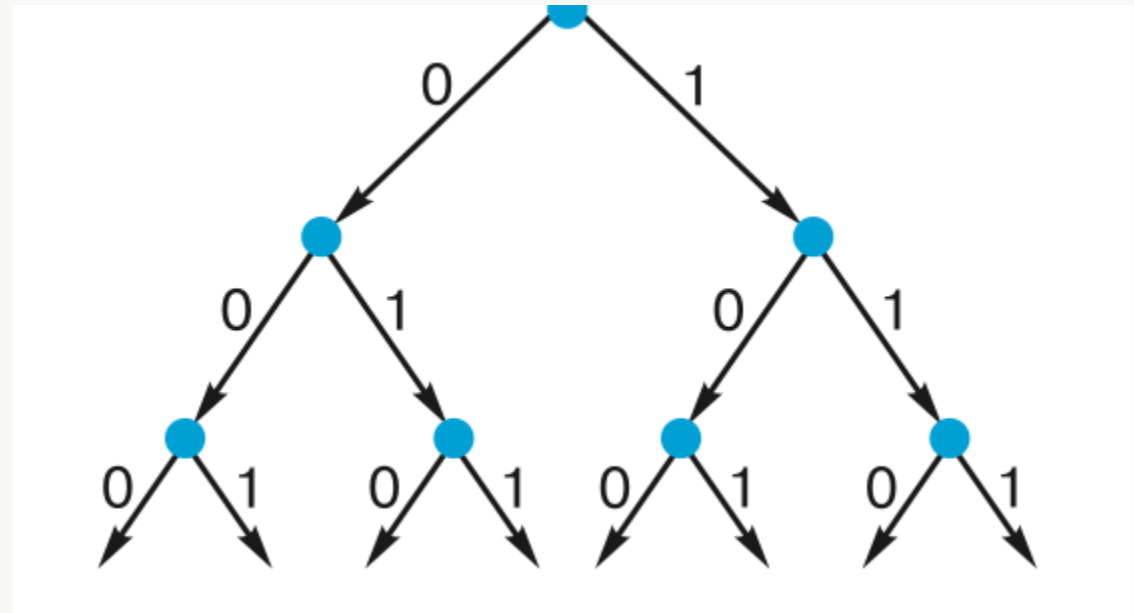
**B D C E A**

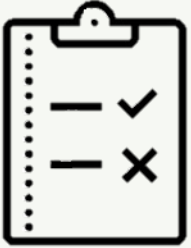
**Ответ: BDCEA**

# КОДОВЫЕ ДЕРЕВЬЯ

**Дерево** — иерархическая структура, состоящая из набора вершин и рёбер. Вершина, в которую не входит ни одного ребра, называется корнем; вершины, из которых не выходит ни одного ребра, называются листьями. Дерево, из вершин которого выходят только два ребра, называется двоичным (бинарным) деревом. Комбинации, соответствующие листьям бинарного дерева, являются кодовыми комбинациями префиксного кода.

Префиксные коды можно наглядно представить с помощью кодовых деревьев.





## ЗАДАЧА 3

Для кодирования некоторой последовательности, состоящей из букв А, Б, В и Г, решили использовать неравномерный двоичный код, позволяющий однозначно декодировать полученную двоичную последовательность. Для букв А, Б и В использовали такие кодовые слова: А — 0, Б — 10, В — 110.

Каким кодовым словом может быть закодирована буква Г? Код должен удовлетворять свойству однозначного декодирования. Если можно использовать более одного кодового слова, надо указать кратчайшее из них.

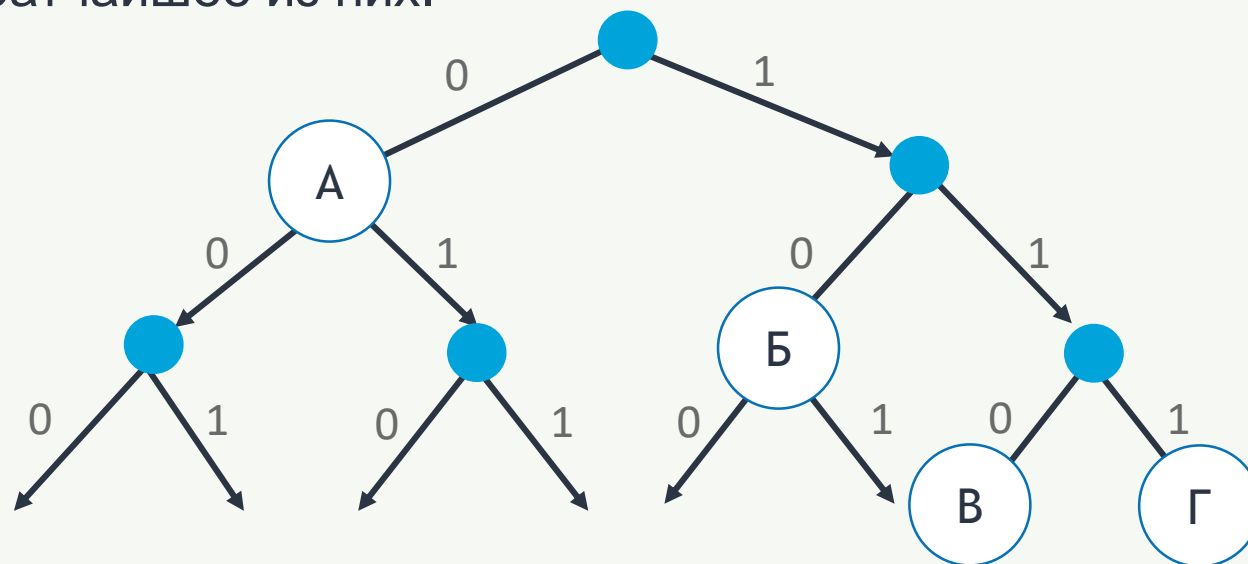




## ЗАДАЧА 3

Для кодирования некоторой последовательности, состоящей из букв А, Б, В и Г, решили использовать неравномерный двоичный код, позволяющий однозначно декодировать полученную двоичную последовательность. Для букв А, Б и В использовали такие кодовые слова: А — 0, Б — 10, В — 110.

Каким кодовым словом может быть закодирована буква Г? Код должен удовлетворять свойству однозначного декодирования. Если можно использовать более одного кодового слова, надо указать кратчайшее из них.



# ПОИСК ИНФОРМАЦИИ

Важнейшая задача обработки информации — поиск информации. Алгоритм поиска зависит от способа организации информации.

## МЕТОД ПОСЛЕДОВАТЕЛЬНОГО ПЕРЕБОРА

- неструктурированный набор данных
- поиск завершается, когда найден искомый элемент или когда просмотрены все элементы набора данных, но искомого элемента в нем нет
- длительность поиска ( $L$ ):  $L = N/2$ , где  $N$  — размер набора данных; если искомый элемент окажется последним или его не окажется вообще, то длительность поиска равна  $N$

## МЕТОД ПОЛОВИННОГО ДЕЛЕНИЯ

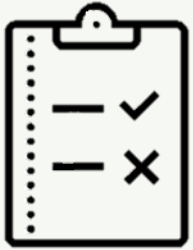
- структурированный набор данных (упорядоченный список)
- искомый элемент сравнивается с центральным элементом последовательности, номер которого находится как  $[N/2] + 1$ ; если значения искомого элемента и центрального совпадают, то поиск завершается, в противном случае поиск продолжается в одной из двух частей последовательности
- длительность поиска ( $L$ ):  $N = 2^L$ , где  $N$  — размер набора данных



## ПРИМЕР 3. МЕТОД ПЕРЕБОРА

Закрывая спортивный магазин, продавец обнаружил отдельно стоящую кроссовку. В магазине осталось только девять коробок с обувью той же модели и того же размера. Помогите продавцу найти пару для этой кроссовки.





# ПРИМЕР 4. МЕТОД ПОЛОВИННОГО ДЕЛЕНИЯ

У плотника в Бобровой деревне 9 складов, пронумерованных от 1 до 9. Плотник не может вспомнить, сколько складов уже заполнил, но помнит, что заполнял их в порядке возрастания номеров. Помогите плотнику найти первый из незаполненных складов за меньшее число ходов.



**Обработка информации** — это целенаправленный процесс изменения содержания или формы представления информации. Существует два различных типа обработки информации:

- 1) обработка, связанная с получением нового содержания, новой информации;
- 2) обработка, связанная с изменением формы представления информации, не изменяющая её содержания.

**Кодирование** — это обработка информации, заключающаяся в её преобразовании в некоторую форму, удобную для хранения, передачи, обработки информации в дальнейшем.

**Код** — это система (список) условных обозначений (кодовых слов), используемых для представления информации.

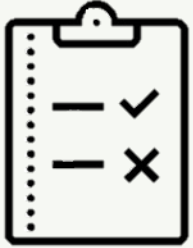
**Префиксный код** — код со словом переменной длины, обладающий тем свойством, что никакое его кодовое слово не может быть началом другого (более длинного) кодового слова. Сообщение, закодированное с помощью префиксного кода, может быть однозначно декодировано.

Задача поиска информации состоит в том, чтобы в некотором хранилище информации (информационном массиве) найти информацию, удовлетворяющую определённым условиям поиска.

Время поиска зависит от способа организации набора данных и используемого алгоритма поиска.

Для осуществления поиска в неструктурированном наборе данных применяется метод последовательного перебора.

Поиск информации в упорядоченном по неубыванию (невозрастанию) наборе данных может быть осуществлён методом половинного деления.



# ВОПРОСЫ И ЗАДАНИЯ

Приведите примеры процессов обработки информации, которые чаще всего вам приходится выполнять в жизни. Для каждого примера определите исходные данные, алгоритм (правила) обработки и получаемые результаты. К каким типам обработки информации относятся эти процессы?

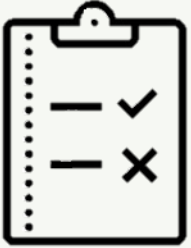




# ВОПРОСЫ И ЗАДАНИЯ

Поясните суть понятий «кодирование», «код», «кодированная таблица».

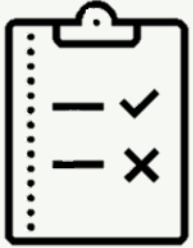




## ВОПРОСЫ И ЗАДАНИЯ

Светодиодная панель содержит шесть излучающих элементов, каждый из которых может светиться или красным, или жёлтым, или зелёным цветом. Сколько различных сигналов можно передать с помощью панели (все излучающие элементы должны гореть, порядок цветов имеет значение)?

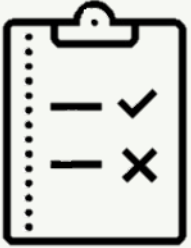




# ВОПРОСЫ И ЗАДАНИЯ

Автомобильный номер состоит из нескольких букв (количество букв одинаковое во всех номерах), за которыми следуют три цифры. При этом используются 10 цифр и только 5 букв: А, В, С, D и F. Требуется не менее 100 тысяч различных номеров. Какое наименьшее количество букв должно быть в автомобильном номере?





# ВОПРОСЫ И ЗАДАНИЯ

Сколько существует различных последовательностей из 6 символов четырёхбуквенного алфавита  $\{A, B, C, D\}$ , которые содержат не менее двух букв  $A$  (т.е. две и более буквы  $A$ )?

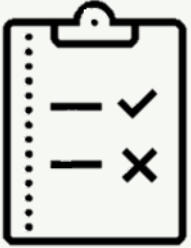




## ВОПРОСЫ И ЗАДАНИЯ

Александр составляет слова, переставляя буквы в слове ЦИРКУЛЯЦИЯ. Сколько слов, в которых есть хотя бы две подряд идущие гласные может составить Александр?

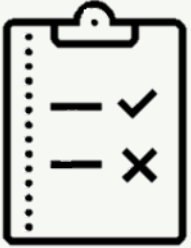




# ВОПРОСЫ И ЗАДАНИЯ

Для передачи закодированных сообщений используется таблица кодовых слов из четырёх букв. Причём используются только буквы А, Р и У. Сколько различных кодовых слов может быть в такой таблице, если ни в одном слове нет трёх одинаковых букв, идущих подряд?

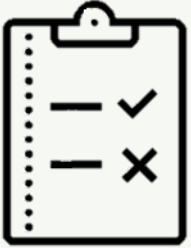




# ВОПРОСЫ И ЗАДАНИЯ

Сравните равномерные и неравномерные коды. Каковы их основные достоинства и недостатки?





# ВОПРОСЫ И ЗАДАНИЯ

Какие коды называют префиксными? Почему они так важны? В чём суть прямого и обратного условий Фано?





# ВОПРОСЫ И ЗАДАНИЯ

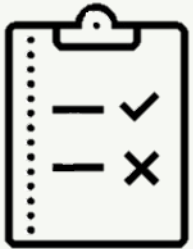
Двоичные коды для 5 букв латинского алфавита представлены в таблице:

A	B	C	D	E
000	01	10	11	001

Из четырёх сообщений, закодированных этими кодами, только одно пришло без ошибки. Найдите его:

1. 110100000100110011;
2. 111010000010010011;
3. 110100001001100111;
4. 110110000100110010.





## ВОПРОСЫ И ЗАДАНИЯ

Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г и Д, используется неравномерный двоичный код, позволяющий однозначно декодировать полученную двоичную последовательность. При этом используются следующие коды: А — 1110, Б — 0, В — 10, Г — 110. Каким кодовым словом может быть закодирована буква Д? Код должен удовлетворять свойству однозначного декодирования. Если можно использовать более одного кодового слова, укажите кратчайшее из них.

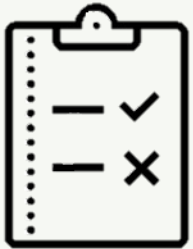




## ВОПРОСЫ И ЗАДАНИЯ

Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г и Д, используется неравномерный троичный код, позволяющий однозначно декодировать полученную троичную последовательность. Вот этот код: А — 0, Б — 11, В — 20, Г — 21, Д — 22. Можно ли сократить для одной из букв длину кодового слова так, чтобы закодированную последовательность по-прежнему можно было декодировать однозначно? Коды остальных букв меняться не должны.

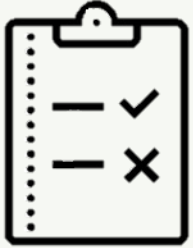




## ВОПРОСЫ И ЗАДАНИЯ

Для передачи сообщений, составленных из заглавных букв русского алфавита, используется неравномерный двоичный код, в котором никакое кодовое слово не является началом другого кодового слова. Это условие обеспечивает возможность однозначной расшифровки закодированных сообщений. Какое наименьшее количество двоичных знаков может содержать сообщение, кодирующее слово КОЛОКОЛ?





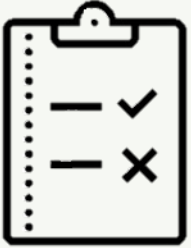
# ВОПРОСЫ И ЗАДАНИЯ

Методом половинного деления в последовательности чисел

061 087 154 180 208 230 290 345 367 389 456 478 523 567 590 612

требуется найти число 590. Опишите процесс поиска.

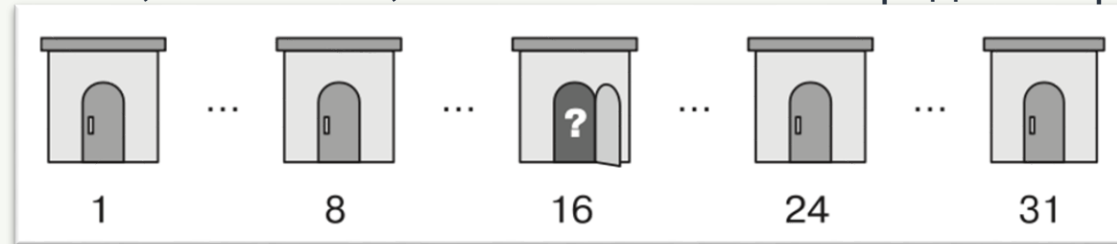




# ВОПРОСЫ И ЗАДАНИЯ

В Международном конкурсе по информатике «Бобёр» школьникам была предложена задача «Склад», подготовленная специалистами из Японии. Вот её условие.

Плотник в Бобровой Деревне использует 31 склад, пронумерованный от 1 до 31. Однажды он забыл, сколько складов уже заполнил, но помнит, что заполнял их в порядке возрастания номеров.



Чтобы уменьшить количество открывания дверей, он действует следующим образом.

Сначала открывает склад со средним номером — склад № 16. Затем:

- если склад № 16 пуст, он решает искать первый незаполненный склад в промежутке от № 1 до № 15, открывает опять средний склад — склад № 8 и повторяет процедуру;
- если склад № 16 заполнен, то нужный склад он ищет между № 17 и № 31, открывает средний склад — склад № 24 и повторяет процедуру.

После всех действий плотник обнаружил, что заполнены были склады от № 1 до № 15 включительно. Сколько дверей ему пришлось открыть?

Решите эту задачу. Какой из рассмотренных нами методов поиска был использован героем этой задачи?

