



ПОДХОДЫ К ИЗМЕРЕНИЮ ИНФОРМАЦИИ

ИНФОРМАЦИЯ И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ

10 класс

Ключевые слова

- измерение информации
- содержательный подход
- алфавитный подход
- кодовая таблица



Подходы к измерению информации



Информация и ее свойства

Информатика

изучение процессов сбора, преобразования, хранения, защиты, поиска и передачи всех видов информации и средств их автоматизированной обработки

Тех
инфор

ория массов
коммуникации

Кибернетика

исследование СМИ и их влияния на общество

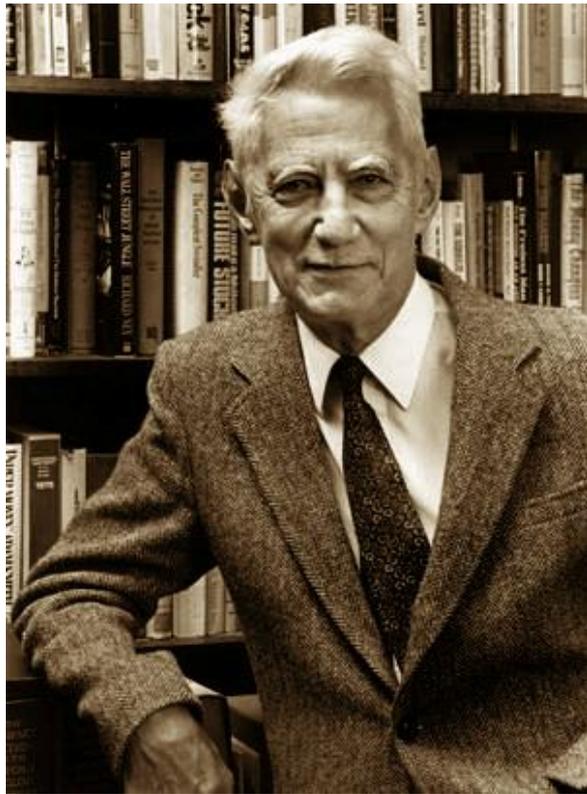
Семиотика

наука о закономерностях процессов управления и передачи информации в машинах, живых организмах и обществе

Содержательный подход



Информация – это снятая неопределенность. Величина неопределённости некоторого события – это количество возможных результатов данного события.

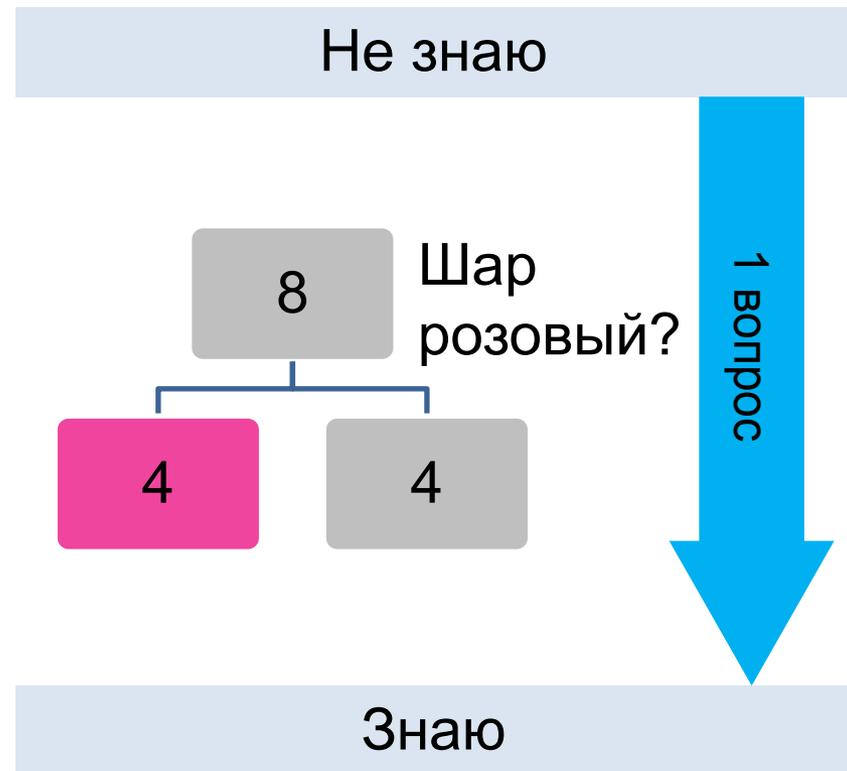


Клод Элвуд Шеннон (1916-2001) – американский инженер и математик. Является основателем теории информации, нашедшей применение в современных высокотехнологических системах связи.



В 1948 году предложил использовать слово «*бит*» для обозначения наименьшей единицы информации.

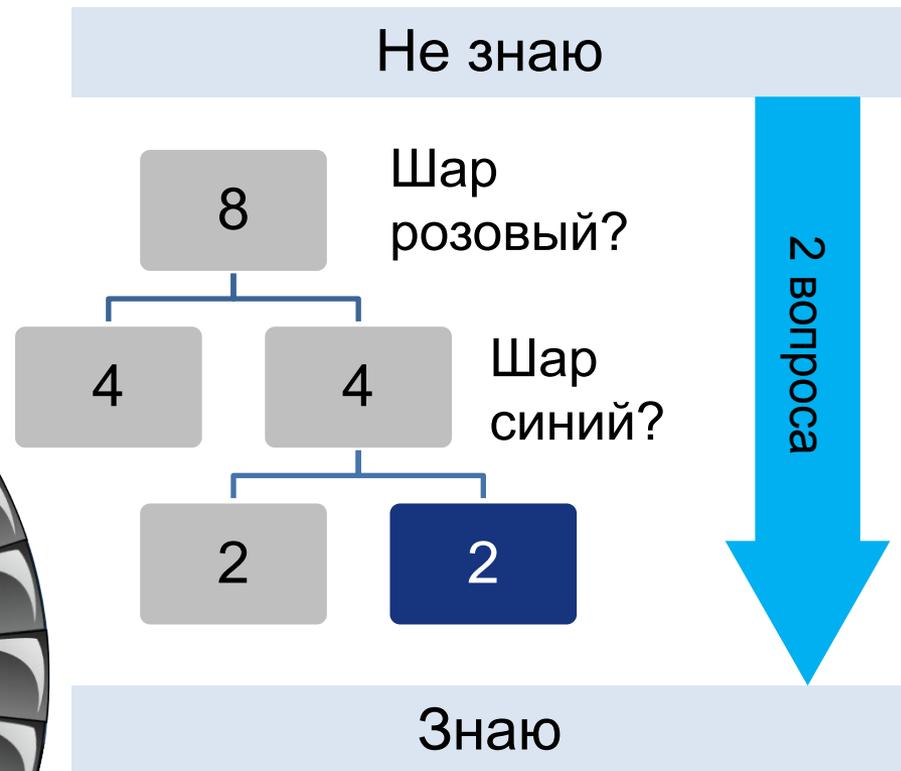
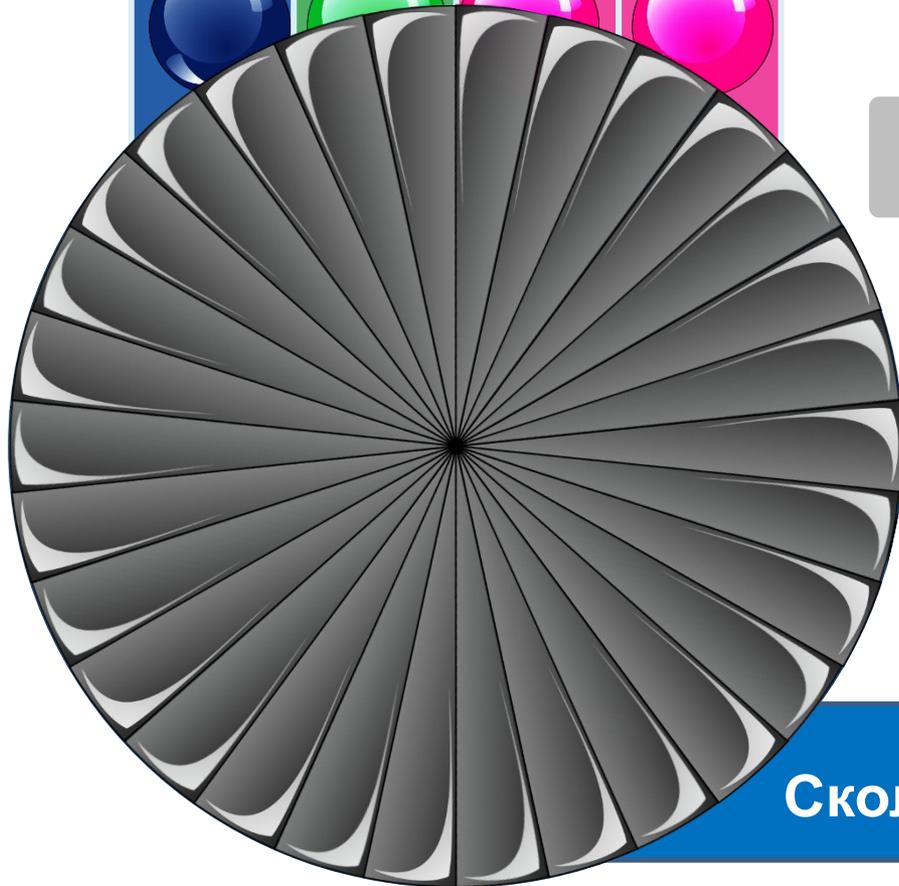
Содержательный подход



$i = 1$ бит

Сколько информации?

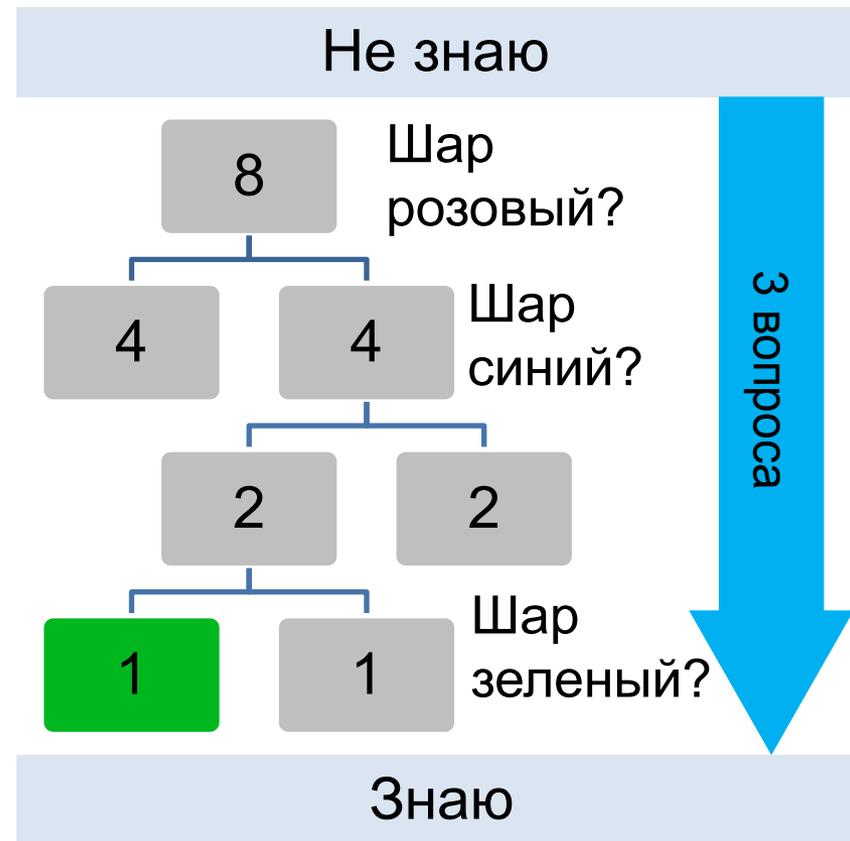
Содержательный подход



$i = 2$ бита

Сколько информации?

Содержательный подход



$i = 3$ бита

Сколько информации?

Метод половинного деления



Исследуйте, сколько вопросов с ответами *Да/Нет* надо задать, чтобы определить цифру на шаре, если начать с вопроса: «Шар синий?»

- 1, 3, 4, 6, 7, 8
- 1, 4, 6, 7



Количество информации, содержащееся в сообщении об одном из N равновероятных результатов некоторого события, определяется из решения уравнения

$$2^i = N.$$



Содержательный подход



Стол телевизионной игры «Что? Где? Когда?» разбит на 13 равных секторов. Какое количество информации содержит сообщение ведущего: «Волчок указывает на супер-блиц».



•13

Не знаю

$$2^i \approx N$$

•2

$$13 \leq 16 = 2^4$$

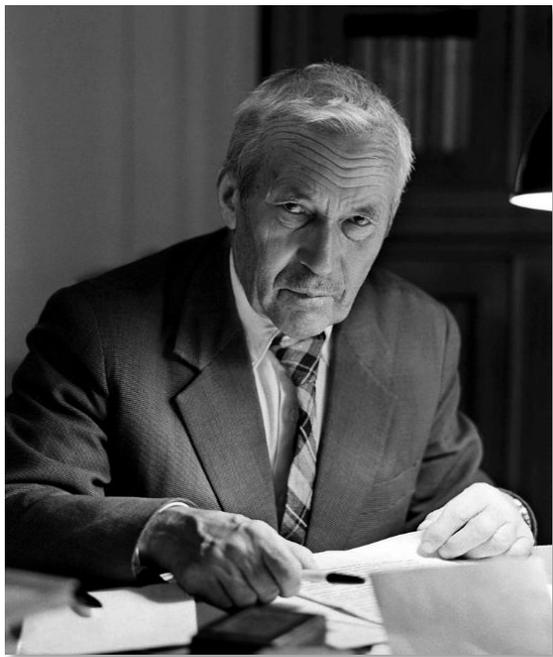
•1

$$i = 4 \text{ бита}$$

4 вопроса

Знаю

Алфавитный подход



Андрей Николаевич Колмогоров
(1903-1987)

Один из крупнейших математиков XX века. Им получены основополагающие результаты в математической логике, теории сложности алгоритмов, теории информации, теории множеств и ряде других областей математики и её приложений.



Информация – последовательность символов (букв, цифр, кодов точек изображения) некоторого алфавита.

Количество информации представленное в сообщении не зависит от содержания сообщения.

Кодирование

Равномерный код		
А	192	11000000
Б	193	11000001
В	194	11000010
Г	195	11000011
Е	197	11000101

Неравномерный код	
А	· -
Б	- · · ·
В	· - -
Г	- - ·
Е	·



Алфавит используемый в азбуке Морзе состоит из 33 букв 10 цифр и 15 специальных символов. Сколько знаков неравномерного кода (точек и тире) в самом длинном символе?



Буквы, которые чаще употребляются, кодируются более простыми сочетаниями точек и тире. Что позволяет сократить длину сообщения.



Почему ФЫВАПРОЛЖЭ



В отличие от Азбуки Морзе, раскладку клавиатуры специально сделали такой, как мы привыкли, чтобы снизить скорость печати.

Буквы первой коммерчески успешной печатной машинки, изобретенной в 1867 году Кристофером Шоулзом, расположили в алфавитном порядке.

Частой поломкой стало «перепутывание» рычажков печатной машинки. Буквы, образующие устойчивые комбинации «разбросали» по разным рядам клавиатуры.

На клавиатуре компьютера технологической проблемы нет. Но зачем нарушать традицию?



Алфавитный подход

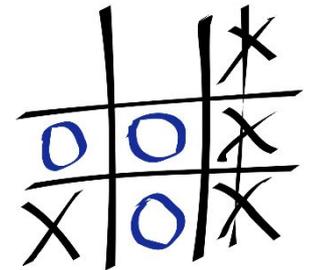


Алгоритм вычисления информационного объёма сообщения:

- определить N – мощность используемого алфавита
- определить i – информационный вес символа алфавита в битах ($2^i = N$)
- вычислить информационный объём сообщения I ($I = i \cdot K$)



Подсчитайте информационный объём необходимый для записи текущего состояния игры в крестики-нолики на поле 3x3 клетки.



Дано:

$$N = 3 \text{ («х», «о», пустая)}$$

$$K = 3 * 3 = 9 \text{ (клеток)}$$

$$I - ?$$

$$2^i = N$$

$$I = K \cdot i$$

$$3 \leq 4 = 2^2, i = 2 \text{ бита}$$

$$I = 2 * 9 = 18 \text{ бит}$$

Ответ: 18 бит

Единицы измерения

информации

$$1 \text{ Кбайт} = 2^{10} \text{ байт}$$

$$1 \text{ Мбайт} = 2^{10} \text{ Кбайт} = 2^{20} \text{ байт}$$

$$1 \text{ Гбайт} = 2^{10} \text{ Мбайт} = 2^{20} \text{ Кбайт} = 2^{30} \text{ байт}$$

$$1 \text{ Тбайт} = 2^{10} \text{ Гбайт} = 2^{20} \text{ Мбайт} = 2^{30} \text{ Кбайт} = 2^{40} \text{ байт}$$

$$1 \text{ Пбайт} = 2^{10} \text{ Тбайт} = 2^{20} \text{ Гбайт} = 2^{30} \text{ Мбайт} = 2^{40} \text{ Кбайт} \\ = 2^{50} \text{ байт}$$



Каждый математик уверен, что $1 \text{ Кбайт} = 1000 \text{ байт}$, а
каждый программист уверен, что $1 \text{ Кбайт} = 1024 \text{ байта}$
каждый программист уверен, что $1 \text{ кг} = 1024 \text{ г}$



Во сколько раз изменится ответ при использовании
приближенного равенства $1 \text{ Кбайт} \approx 1000 \text{ байт}$

Самое главное

Содержательный подход

Информация – это снятая неопределенность. Сообщение, уменьшающее неопределённость знания в 2 раза, несёт **1 бит** информации.

Количество информации i , содержащееся в сообщении об одном из N **равновероятных** результатов некоторого события, определяется из решения уравнения $2^i = N$.

Алфавитный подход

Информация – последовательность символов некоторого алфавита. В двоичном коде один двоичный разряд несёт 1 бит информации.

Информационный объём сообщения вычисляется по формуле: $I = K \cdot i$



Давайте обсудим



Дано множество карточек с географическими названиями:



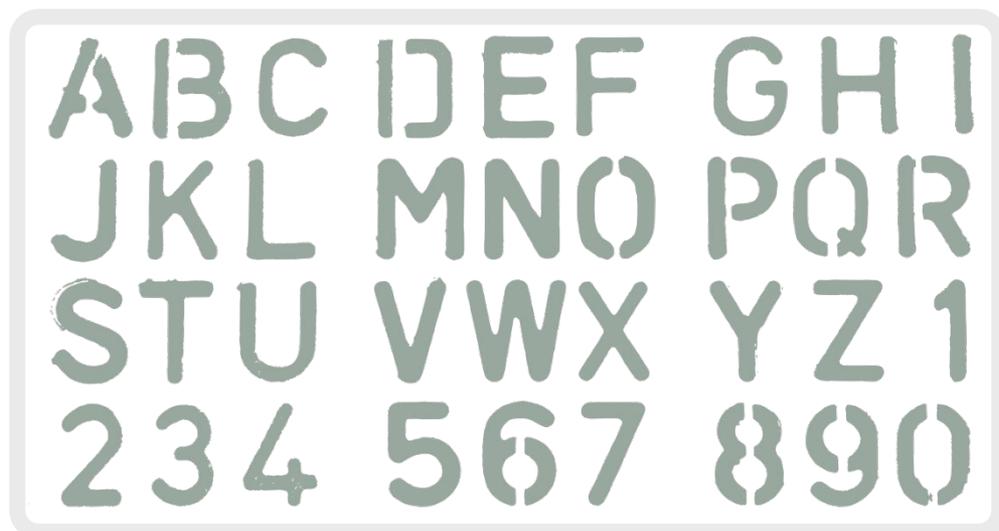
Придумайте сообщение, информационный объем которого равен 1 биту.

Подсказка

Давайте обсудим



При кодировании ID-номера используется посимвольное кодирование. Алфавит содержит 26 букв латинского алфавита и цифры 0...9. Сколько символов можно добавить к алфавиту без увеличения информационного объема необходимого для хранения 100 номеров?



Ответ

Вопросы и задания



1. Что такое неопределённость знания о результате какого-либо события? Приведите пример.
2. В чём состоит суть содержательного подхода к определению количества информации?
3. Что такое бит с точки зрения содержательного подхода?
4. Партию игры на шахматном поле записали в виде последовательности выполненных ходов. Сравнить два способа кодирования:
 - а) традиционное название клеток (пример одного хода: $a1-b2$);
 - б) нетрадиционное именование клеток, при котором клетки имеют «сквозную» нумерацию каждой клетке соответствует одно число (пример хода: $62-18$).

Решите эту задачу для двух видов полей: 8×8 и 10×10 клеток

Вопросы и задания



5. В школе 750 учащихся. Коды учащихся записаны в школьной информационной системе с помощью минимального количества бит. Каков информационный объем в байтах сообщения о кодах 180 учащихся начальных классов?
6. При регистрации в компьютерной системе каждому пользователю выдаётся пароль, состоящий из 6 символов и содержащий только символы из шестибуквенного набора А, В, С, D, Е, F. Для хранения сведений о каждом пользователе отведено одинаковое и минимально возможное целое число байт. При этом используют посимвольное кодирование паролей и все символы кодируются одинаковым и минимально возможным количеством бит. Определите объём памяти в байтах, необходимый для хранения сведений о 120 пользователях.