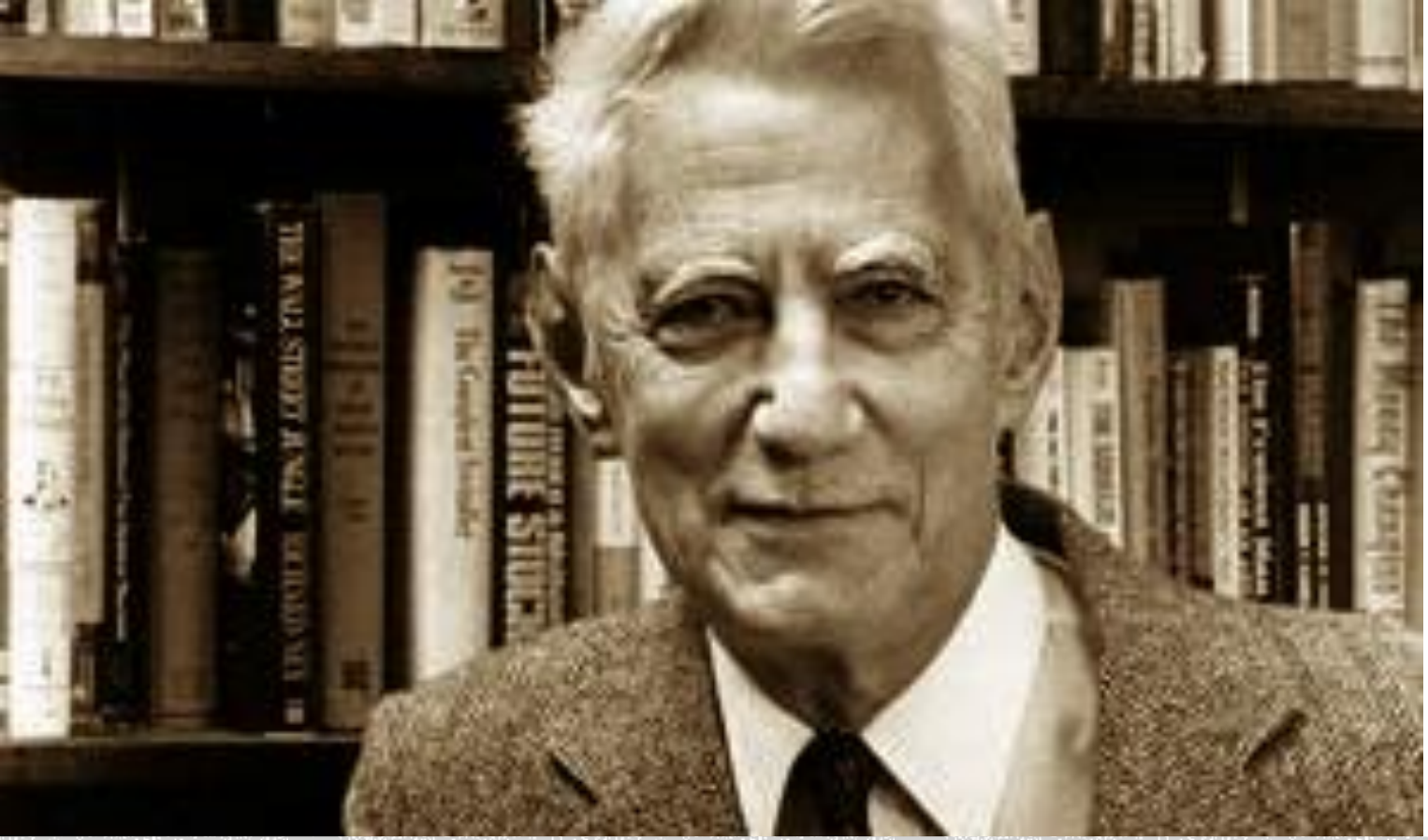
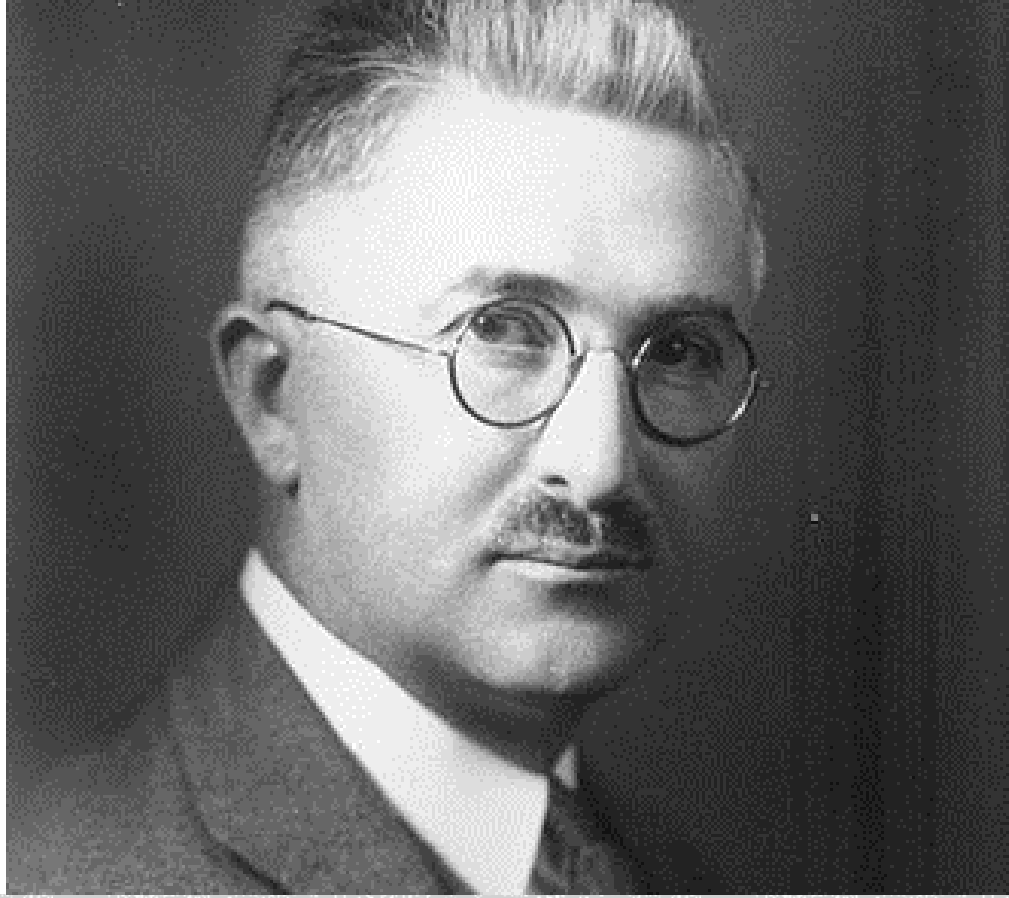




ИНФОРМАТИКА

Лекция 1-2

11, 25 сентября 2018 г.

СТРУКТУРА КУРСА

Лекции

- Основы информатики и информационных технологий

Практики

- Word
- Excel
- PowerPoint

http://prog.tversu.ru

ИНФОРМАТИКА

Опр. 1: Информатика (из БРЭ) — наука о методах и процессах сбора, хранения, обработки, передачи, анализа и оценки информации с применением компьютерных технологий, обеспечивающих возможность её использования для принятия решений.

Возникла во второй половине **XX** века.

Одна из самых молодых естественных и технических наук.

ЭВОЛЮЦИЯ НАУК

Время	Период	Картина мира	Математика	Основной объект
до XVII века	аграрный	материальная картина мира сельскохозяйственное производство	элементарная математика (арифметика), которая включала элементы алгебры, планиметрию, стереометрию, тригонометрию	Число
XVII – XX век	индустриальный	динамические системы: движение небесных тел и движение тел на Земле, мореплавание; технологии базировались на преобразовании энергии из одного вида в другой	период развития высшей математики: анализа бесконечно малых, теории функций действительного переменного, аналитической и дифференциальной геометрии	Функция

ЭВОЛЮЦИЯ НАУК

Время	Период	Картина мира	Математика	Основной объект
с середины XX века по настоящее время	Постиндустриальный (информационный)	информационная картина мира; вместо динамических систем стали доминировать их модели, в центр внимания становятся информационные системы: сначала кибернетические, а затем и интеллектуальные системы, например, экспертные системы	дискретная математика, алгебраические и топологические методы	множества и отношения, алгебра, топология, математическая логика

ОБЛАСТЬ ИНФОРМАТИКИ

4 главных области:

- теория вычислений
- алгоритмы и структуры данных
- методология программирования и языков
- архитектура компьютера

Дополнительные:

- программная инженерия
- искусственный интеллект
- компьютерные сети и коммуникации
- базы данных
- параллельные вычисления
- распределенные вычисления
- человеко-машинное взаимодействие
- компьютерная графика
- операционные системы
- числовые и символьные вычисления

ИНФОРМАЦИЯ

Что такое информация?

В житейском смысле – это сумма сведений (знаний) о чем-либо.

В философии три подхода:

- атрибутивный
- функциональный
- антропоцентрический

Опр. 2: Информация – это сведения об объектах и явлениях окружающего мира, уменьшающие степень неопределенности знаний об этих объектах или явлениях.

СВОЙСТВА ИНФОРМАЦИИ

1. Невоспроизводимость
2. Эмерджентность (от английского “emergency”)
3. Операциональность
4. Объективность
5. Полнота
6. Достоверность
7. Адекватность
8. Доступность
9. Актуальность

ДАННЫЕ

Что такое данные?

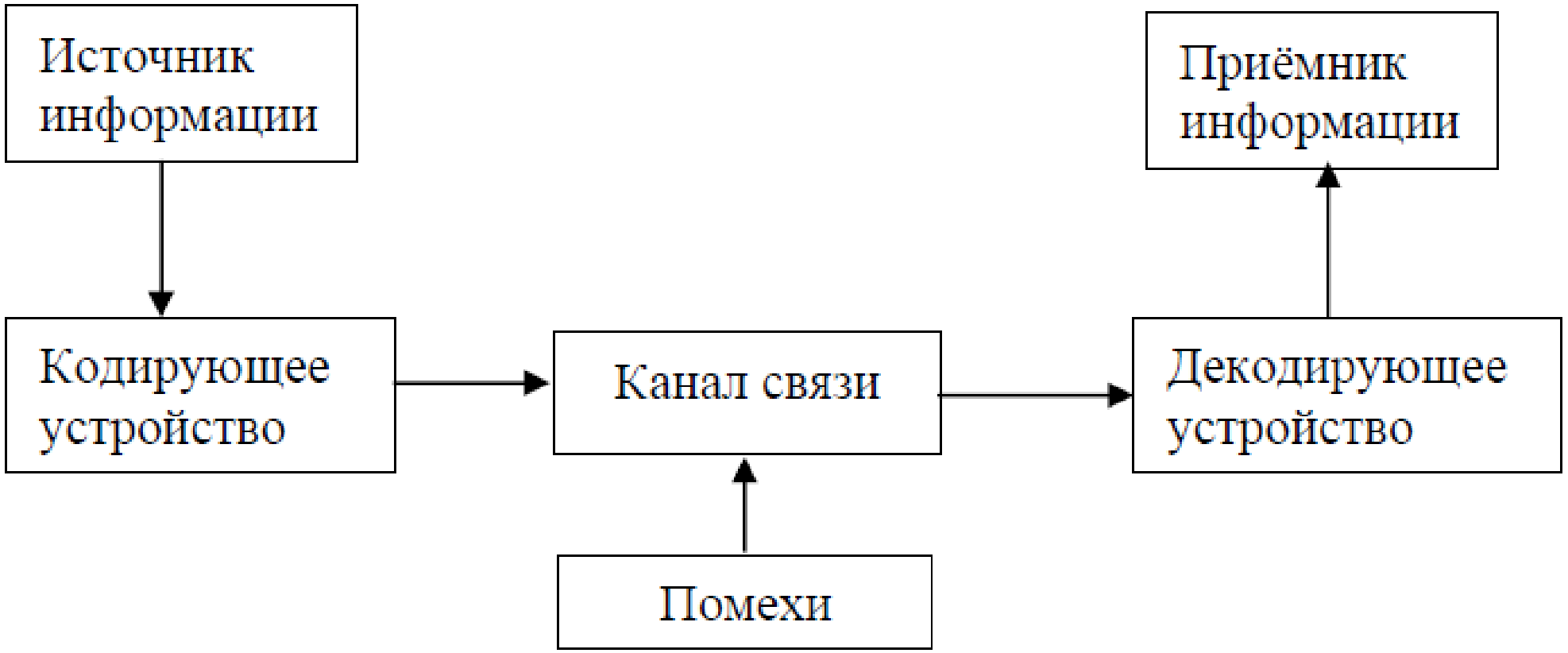
Опр. 3: Данные – это зарегистрированные сигналы.

Сигнал – изменяющийся во времени физический процесс.

Регистрация – запись (фиксация) сигнала на физическом носителе.

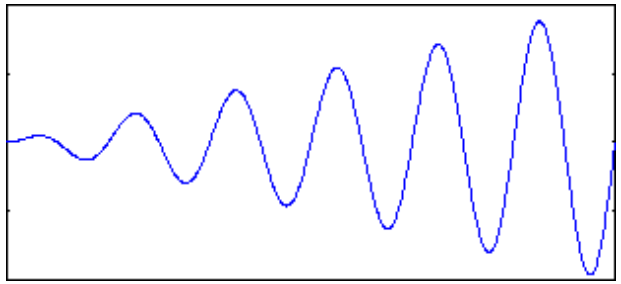
Опр. 3а: Данные – совокупность сведений, зафиксированных на определённом носителе в форме, пригодной для постоянного хранения, передачи и обработки.

ОБЩАЯ СХЕМА ПЕРЕДАЧИ ДАННЫХ

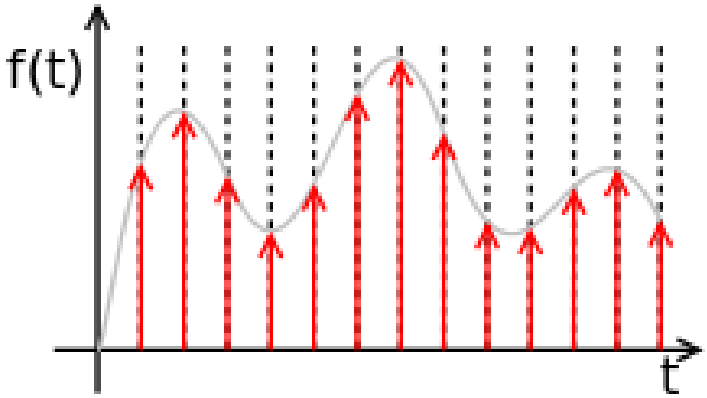


ФОРМЫ ПРЕДСТАВЛЕНИЯ ДАННЫХ

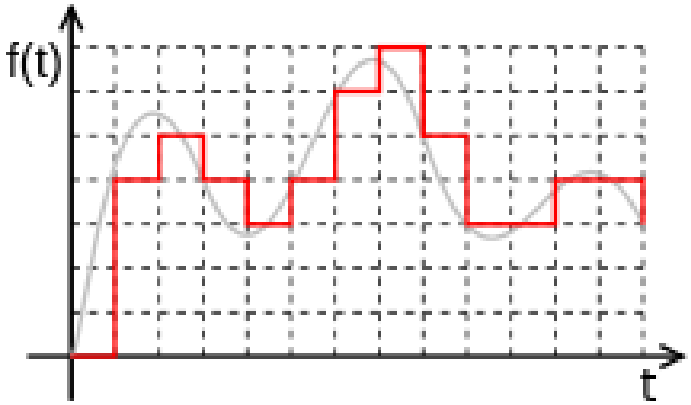
Аналоговая



Дискретная



Цифровая



Это – свойство **данных**. Информация нематериальна и не может быть «аналоговой» или «дискретной».

ФОРМЫ ПРЕДСТАВЛЕНИЯ ДАННЫХ

Принципиальное наиважнейшее свойство цифровых сигналов – их можно **обозначить**, т.е. приписать каждому из конечного набора значений сигнала **знак**, позволяющий их различать.

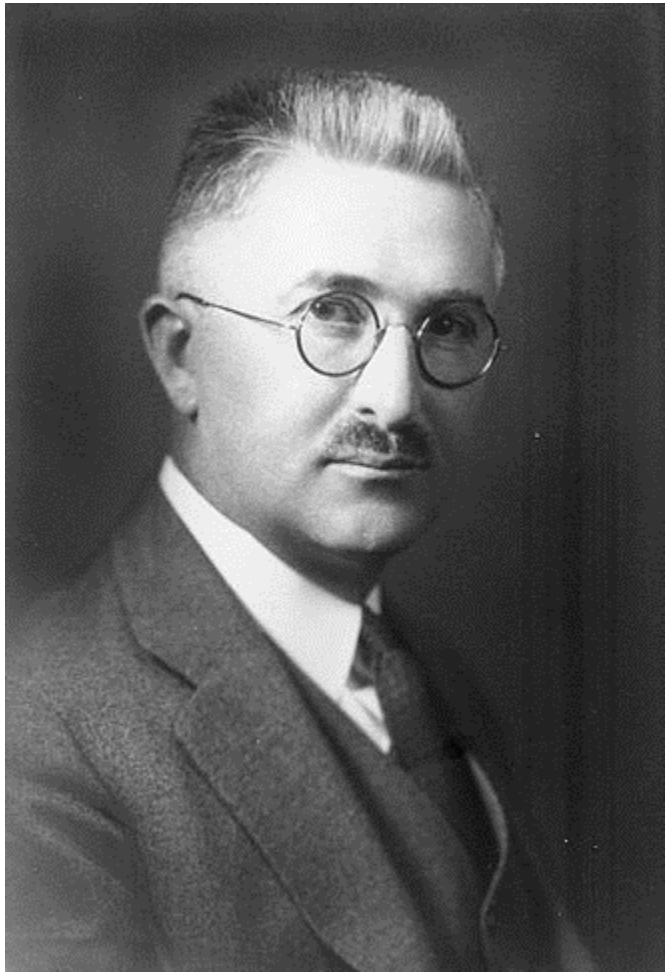
Опр. 4: Знак – это элемент некоторого конечного множества отличных друг от друга сущностей.

Примеры знаков: жест, рисунок, буква, сигнал светофора, определенный звук.

Опр. 5: Вся совокупность знаков, используемых для представления дискретных данных, называется **алфавитом**.

Опр. 6: Мощность алфавита – количество элементов в нем.

ИЗМЕРЕНИЕ ИНФОРМАЦИИ (ФОРМУЛА ХАРТЛИ)



Опр. 7: Количество информации по Хартли (Хартлиевская мера информации), содержащееся в в последовательности из **n** символов из алфавита **A** мощности **m** равно: $I = n * \log_2 m$ **бит**

*Замечание 1. Хартли предполагал, что все символы алфавита **A** могут встретиться с равной вероятностью (частотой) в любом месте сообщения.*

*Замечание 2. Любое сообщение длины **n** в алфавите **A** будет содержать одинаковое количество информации по Хартли.*

ИЗМЕРЕНИЕ ИНФОРМАЦИИ (ФОРМУЛА ХАРТЛИ)

Опр. 8: Информационная энтропия – мера неопределенности источника информации, который может выдать одно из $\mathbf{m}$ значений (символов алфавита). Численно равна значению $\log_2 m$.

Чем больше мощность алфавита ($\mathbf{m}$), тем больше энтропия (неопределенность).

Количество информации (по Хартли):
количество символов, умноженное на энтропию источника: $I = n * \log_2 m$

С практической точки зрения, **энтропия по Хартли** – количество бит, которое необходимо для передачи одного значения от источника приемнику, а **количество информации** – общее количество бит, необходимое для передачи всего сообщения.

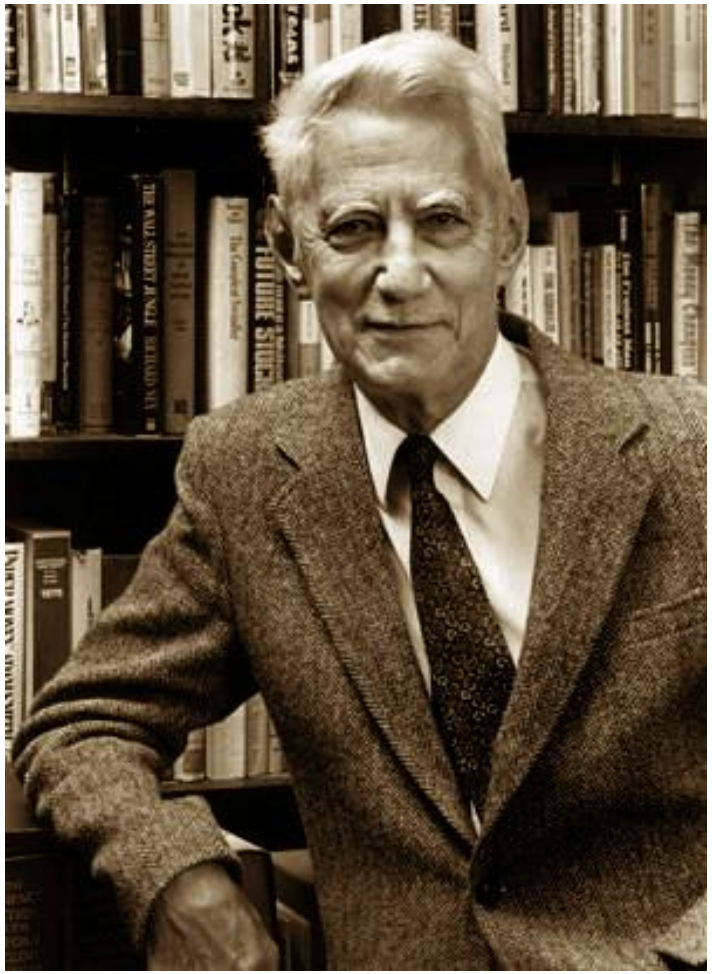
ИЗМЕРЕНИЕ ИНФОРМАЦИИ (ФОРМУЛА ХАРТЛИ)

Опр. 9: Сообщение, уменьшающее неопределенность знания в два раза, несет **1 бит** информации.

Опр. 10: Неопределенность знания о результате некоторого события **–** это количество возможных результатов (исходов) этого события.

Сообщение об одном из двух равновероятных результатов некоторого события несет 1 бит информации.

ИЗМЕРЕНИЕ ИНФОРМАЦИИ (ФОРМУЛА ШЕННОНА)



Опр. 11: Один символ алфавита x_i, вероятность появления которого в сообщении равна p_i, несет количество информации по Шеннону: $\log_2\left(\frac{1}{p_i}\right)$ бит

Опр. 12: Энтропия источника по Шеннону равна среднему количеству информации по всем символам:

$$\mathrm{H(X)} = \sum_{i=1}^{m} p_i \log_2\left(\frac{1}{p_i}\right) \textbf{ бит}$$

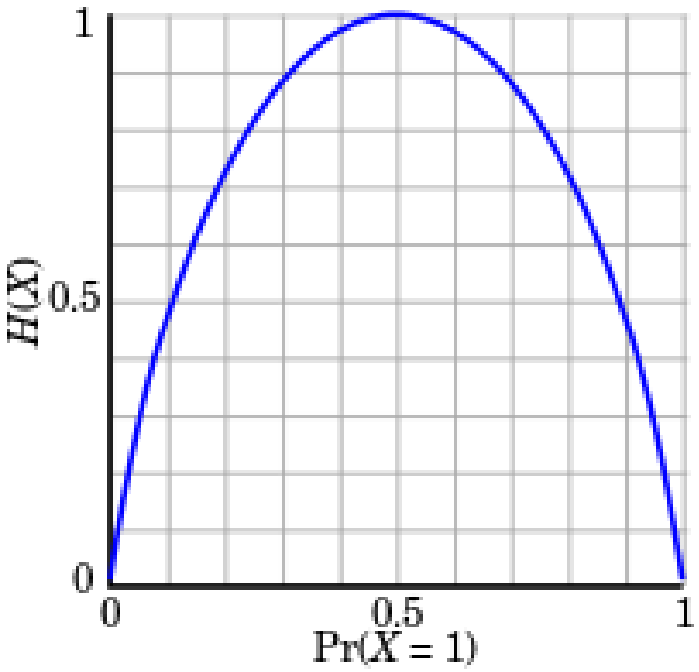
Опр. 13: Количество информации по Шеннону, содержащееся в сообщении из **n** символов, равно:

$$\mathrm{n} * \sum_{i=1}^{m} p_i \log_2\left(\frac{1}{p_i}\right) = n * H(X) \textbf{ бит}$$

ИЗМЕРЕНИЕ ИНФОРМАЦИИ (ФОРМУЛА ХАРТЛИ)

Замечание 3. Если все события равновероятны, то формула Шеннона преобразуется в формулу Хартли.

Теорема 1. Максимум энтропии достигается тогда, когда все исходы равновероятны.



УПРАЖНЕНИЯ

1. Шахматная доска состоит из 64 полей: 8 столбцов на 8 строк. Какое количество бит несет сообщение о выборе одного шахматного поля?
2. При игре в кости используется кубик с шестью гранями. Сколько битов информации получает игрок при каждом бросании кубика?
3. В корзине лежат 8 черных шаров и 24 белых. Сколько бит информации несет сообщение о том, что достали черный шар?
4. Найти энтропию с.в. X, заданной распределением

X	1	2	3	4	5	6	7	8
p	0.1	0.2	0.1	0.05	0.1	0.05	0.3	0.1.

5. Пусть при бросании несимметричной четырехгранной пирамидки вероятности отдельных событий равны: p1=1/2; p2=1/4; p3=1/8; p4=1/8. Чему равна энтропия системы?
6. Чему равна вероятность получить четную сумму очков, большую 5, при выбрасывании двух игральных кубиков? Чему равна энтропия данной с.в.?

УПРАЖНЕНИЯ

7. Найти энтропию с.в. X и среднюю длину каждого из приведенных кодов для этой с.в.

X	1	3	4	5	6
p	0.4	0.2	0.1	0.2	0.1
$code1(X)$	000	001	010	011	111
$code2(X)$	0	100	101	110	111
$code3(X)$	00	01	110	10	111
$code4(X)$	0	10	1110	110	1111.

8. Д.с.в. X равна количеству гербов, выпавших на двух идеальных монетках. Найти энтропию X.
9. Про д.с.в. X известно, что ее значениями являются буквы кириллицы. Произведен ряд последовательных измерений X, результат которых “ТЕОРИЯИНФОРМАЦИИ”. Составить на основании этого результата приблизительный закон распределения вероятностей этой д.с.в. и оценить минимальную среднюю длину кодов для X.