

Физические основы получения информации

1. Общие вопросы получения информации

1.1. Формы представления информации

Определения:

1. Информация – любые сведения о каком-либо событии или объекте.
2. Информация — сведения, воспринимаемые человеком или специальными устройствами как отражение фактов материального мира в процессе коммуникации.
3. Информация – это знания, которые человек получает из различных источников с помощью органов чувств.
4. Информация – отражение одного объекта другим.
5. Информация – это сведения о чём-либо, когда форма их представления также является информацией, то есть имеет формирующую функцию в соответствии с собственной природой.

Основная форма представления информации – это сообщение.

Сообщением называют информацию, фиксированную в определенной форме. Сообщение отображается в виде сигнала (электрического, звукового, светового и др.).

Передача информации состоит в переносе ее на расстояние посредством сигналов различной физической природы по механическим, оптическим, акустическим, электромагнитным и другим каналам связи. Чаще всего используются электрические и электромагнитные каналы связи.



Сигнал можно характеризовать как средство перенесения информации в пространстве и времени.

Чтобы было соответствие между сообщением и сигналом, сигнал должен формироваться по определенным правилам.

Построение сигнала по определенным правилам называют кодированием.

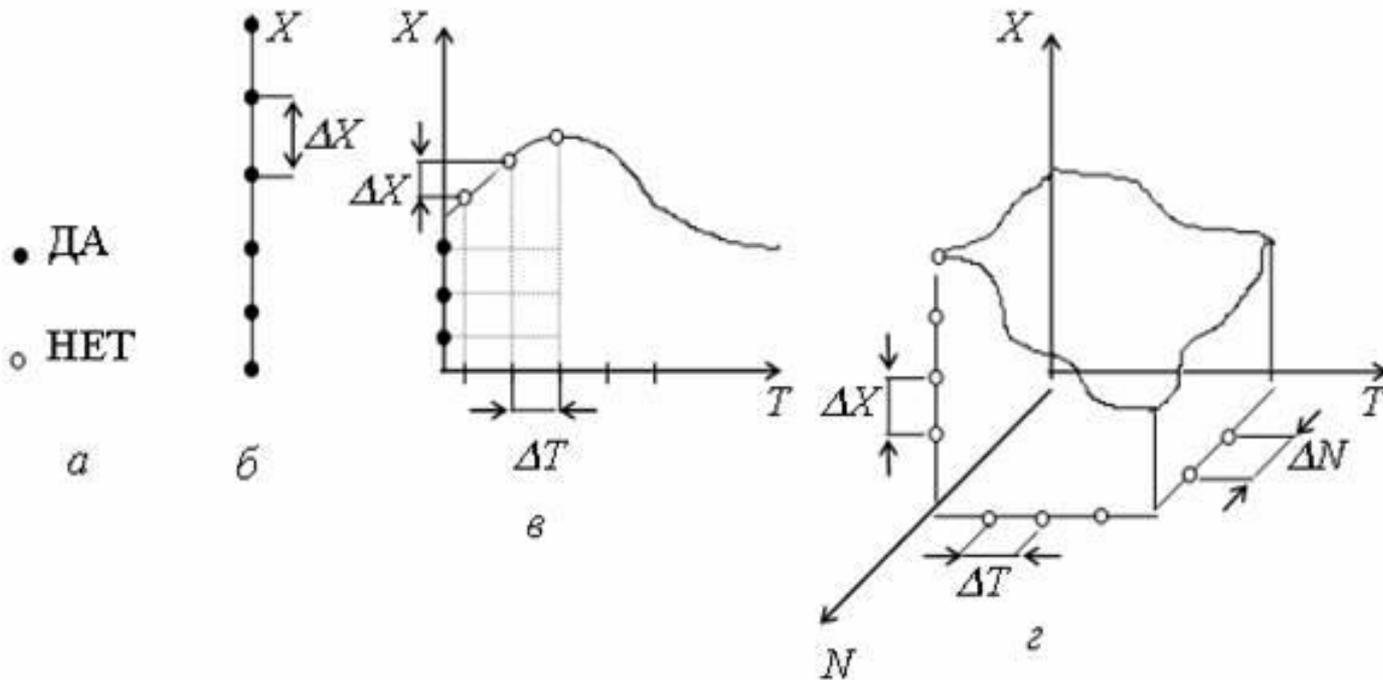
Таким образом, в материально-энергетической форме информация всегда проявляется в виде сигналов.

Первичным и неделимым элементом информации следует считать двоичное событие, то есть утверждение или отрицание.

Двоичное событие условно представляется единицей или нулем, импульсом или паузой.

Информация описывается моделями с различной мерностью:

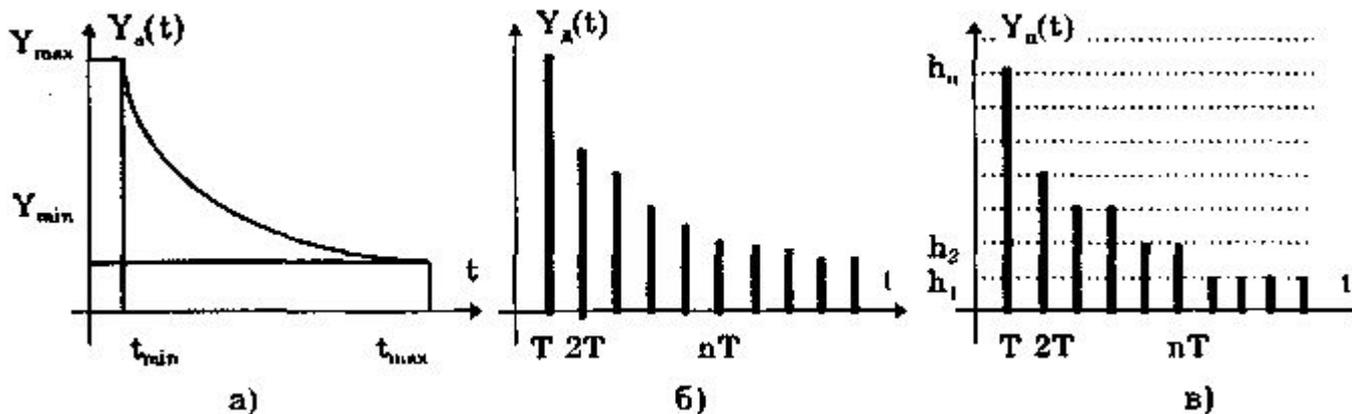
- событие представляет собой нульмерную информацию;
- величина - одномерную;
- функция - двухмерную;
- комплекс - трехмерную.



Источники и формы сообщений и соответствующие им сигналы бывают непрерывные и дискретные.

Непрерывными называются такие сообщения, которые могут принимать любые значения и являются непрерывными функциями времени.

Дискретными называются сообщения, состоящие из отдельных элементов (символов, букв, импульсов), принимающих конечное число различных значений.



Для передачи сообщения по каналу связи ему необходимо поставить в соответствие определенный сигнал.



Сигнал – это физический процесс, отображающий сообщение.

Преобразование сообщения в сигнал, удобный для передачи по данному каналу связи, называют *кодированием*.

Операцию восстановления сообщения по принятому сигналу называют *декодированием*.

Обмен информацией использует систему знаков.

Семиотика - наука о знаках, словах и языках.

Три уровня семиотики:

***синтактический* - рассматривают внутренние свойства текстов, т.е. отношения между знаками, отражающие структуру данной знаковой системы;**

***семантический* уровень - анализируют отношения между знаками и обозначаемыми ими предметами, действиями, качествами, т.е. смысловое содержание текста, его отношение к источнику информации;**

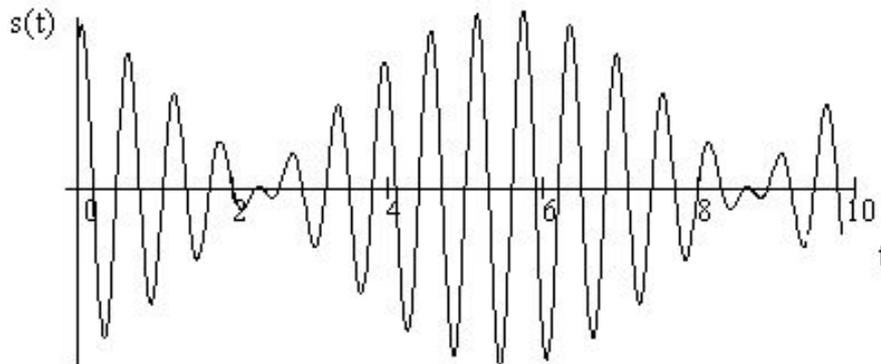
***прагматический* - рассматривают отношения между текстом и теми, кто его использует, т.е. потребительское содержание текста, его отношение к получателю.**

Основу сигнала составляет физический процесс, называемый носителем информации.

Носитель становится сигналом в процессе модуляции.

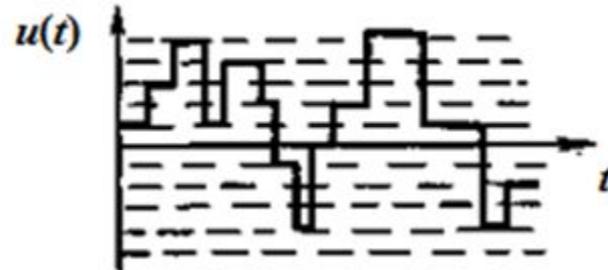
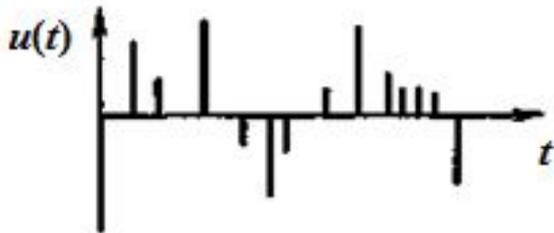
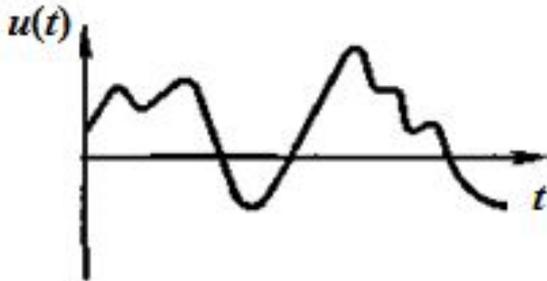
Параметры носителя, изменяемые во времени в соответствии с передаваемым сообщением, называют информативными.

Пример: модулированные колебания:



Сигналы подразделяют на :

- непрерывные,
- дискретные,
- дискретно-непрерывные.



Временное представление сигналов

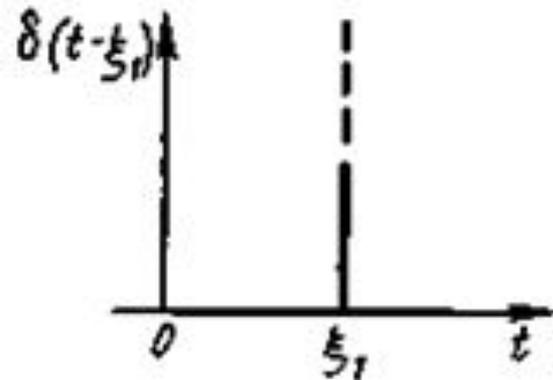
В качестве базисных функций используются единичные импульсные функции - дельта-функции или единичные ступенчатые функции.

Единичная импульсная функция:

$$\delta(t) = \begin{cases} \infty & \text{при } t = 0, \\ 0 & \text{при } t \neq 0, \end{cases} \quad \int_{-\infty}^{\infty} \delta(t) dt = 1,$$

$$\delta(t - \xi_1) = \begin{cases} \infty & \text{при } t = \xi_1, \\ 0 & \text{при } t \neq \xi_1, \end{cases}$$

$$\int_{-\infty}^{\infty} \delta(t - \xi_1) dt = 1.$$



Значение реального сигнала $u(t)$ в конкретный момент времени ξ_1 :

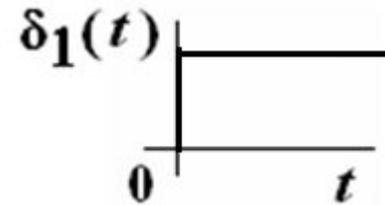
$$u(\xi_1) = \int_{-\infty}^{\infty} u(t) \delta(t - \xi_1) dt$$

Временное представление сигнала $u(t)$ с помощью единичных импульсных функций :

$$u(t) = \int_{-\infty}^{\infty} u(\xi) \delta(\xi - t) d\xi = \int_{-\infty}^{\infty} u(\xi) \delta(t - \xi) d\xi$$

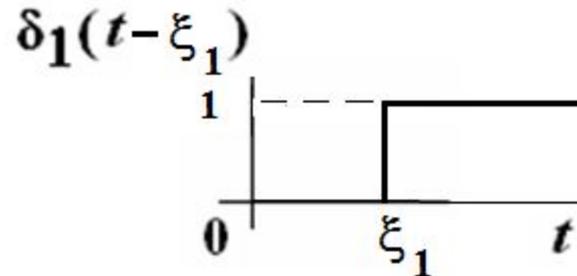
Единичная ступенчатая функция:

$$f(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

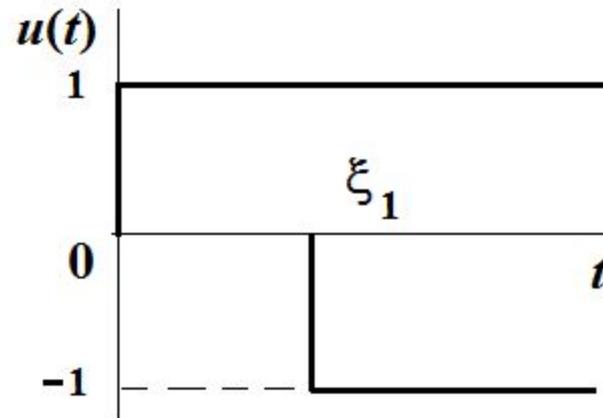
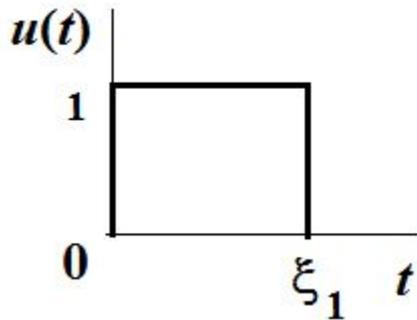


Смещенная единичная ступенчатая функция:

$$f(t) = \begin{cases} 0 & t < \xi_1 \\ 1 & t \geq \xi_1 \end{cases}$$



Временное представление прямоугольного сигнала $u(t)$ с помощью единичных ступенчатых функций:

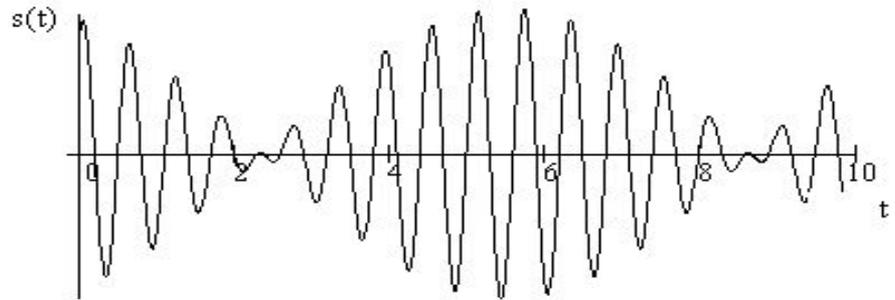


$$u(t) = \delta(t) - \delta(t - \xi_1)$$

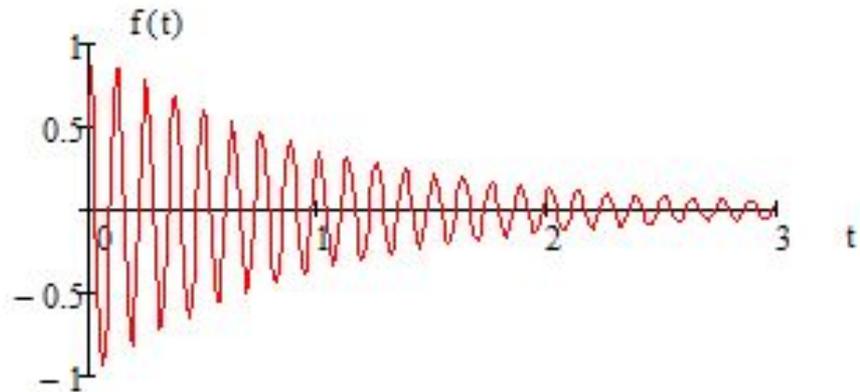
Частотное представление модулированных сигналов

Модулированный по амплитуде сигнал представляет собой произведение двух функций, одна из которых имеет произвольный вид, а другая является синусоидальной.

$$f(t) = A \cos(\omega_1 t) \cdot \cos(\omega_2 t)$$



$$f(t) = A \cdot e^{-at} \cdot \cos(\omega_2 t)$$



Модулированные сигналы представляют в частотной области тоже можно представить в виде частотных спектров.

Для получения спектров модулированных сигналов применяют теорему сдвига в частотной области.

Теорема:

Если в изображении сигнала комплексную частоту сместить на ω_0 , то оригинал сигнала нужно умножить на экспоненту

$$e^{-s_0 t}$$

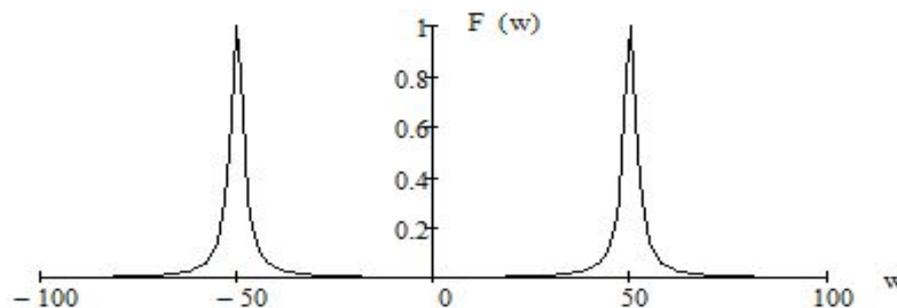
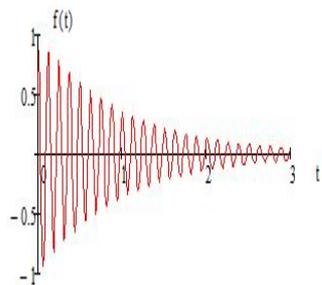
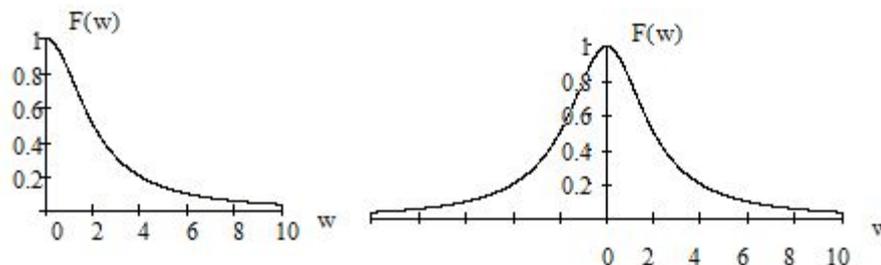
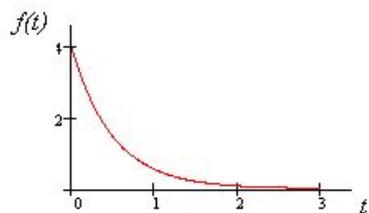
$$F(s + s_0) \div f(t) \cdot e^{-s_0 t}$$

$$F(s) = \frac{A}{s + a}$$

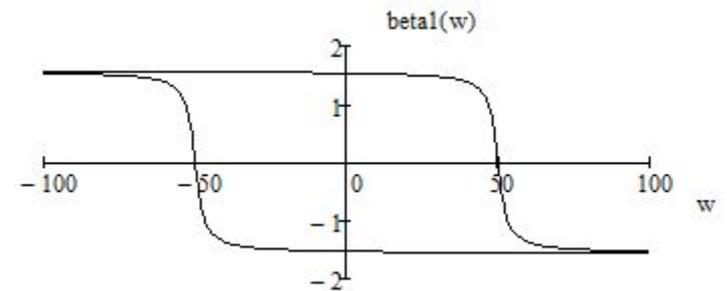
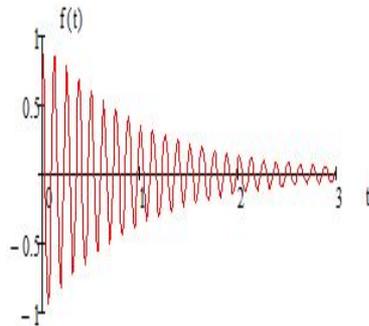
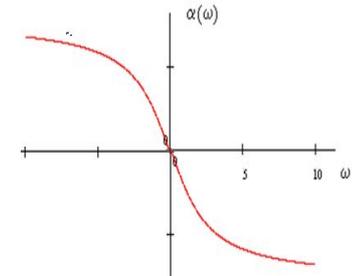
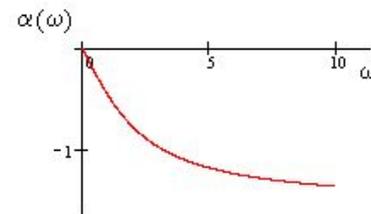
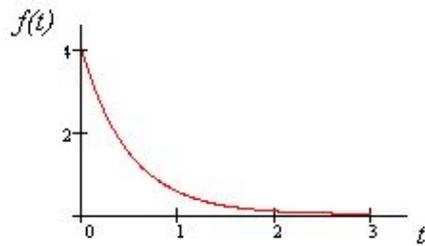
$$f(t) = A \cdot e^{-at} \cdot \delta_1(t)$$

Спектры модулированных по амплитуде синусоидальных колебаний можно получить, если сместить по оси частот вправо и влево на величину частоты колебаний амплитудный и фазовый спектры огибающей функции.

$$F(j\omega \boxplus j\omega_0) = F(\omega \boxplus \omega_0) \cdot e^{j\alpha(\omega \boxplus \omega_0)}$$



Фазовый спектр модулированного сигнала



Мера количества информации

Направления оценки количества информации

В теории информации выделяются три основных направления:

1. Структурное.

Структурное - рассматривает дискретное строение массивов информации и их измерение простым подсчетом информационных элементов.

Используется для оценки возможностей технических средств различных систем переработки информации, независимо от конкретных условий их применения.

2. Статистическое.

Статистические оценки применяются при рассмотрении вопросов передачи данных, определении пропускной способности каналов связи.

Направление оперирует понятием энтропии как меры неопределенности, то есть здесь учитывается вероятность появления тех или иных сообщений.

3. Семантическое.

Семантические используются при решении задач построения систем передачи информации и при оценке эффективности различных устройств.

Учитывает целесообразность, ценность или существенность информации.

Структурные меры информации

Структурные меры учитывают только дискретное строение информации. Элементами информационного комплекса являются кванты - неделимые части информации.

Различают *геометрическую, комбинаторную и аддитивную* структурные меры.

1. *Геометрическая мера* представляет собой измерение длины линии, площади или объема геометрической модели информационного комплекса в количестве квантов.

Максимально возможное число квантов в заданных структурных габаритах определяет информационную емкость системы.

Информационная емкость есть число, указывающее количество квантов в полном массиве информации.

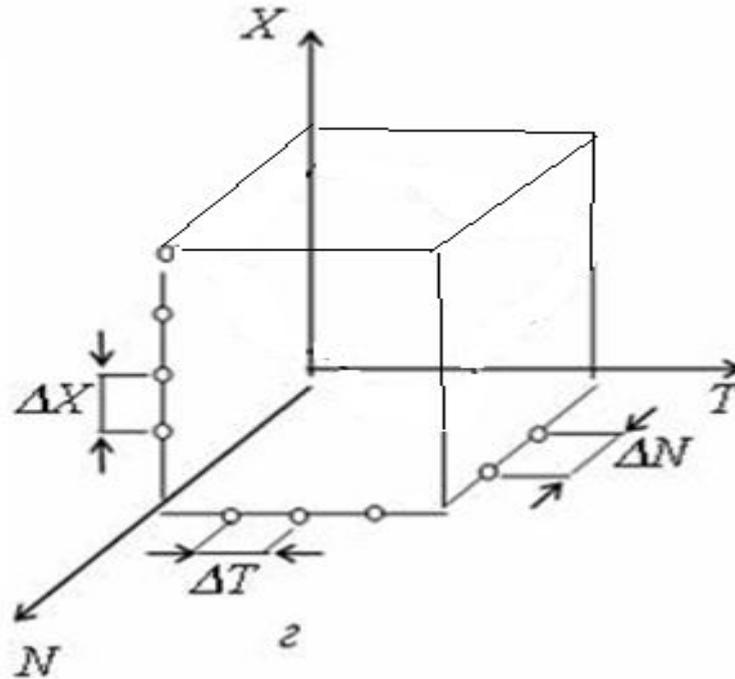
Количество информации M в комплексе $X(T,N)$, определенное геометрическим методом, равняется

$$M = m_X \cdot m_T \cdot m_N$$

$$m_X = \frac{X}{\Delta X}; \quad m_T = \frac{T}{\Delta T}; \quad m_N = \frac{N}{\Delta N}.$$

$\Delta X, T, N$ - интервалы, через которые осуществляются дискретные отсчеты.

Пример геометрической меры информации



$$m_X = 4; \quad m_T = 4; \quad m_N = 3. \quad M = m_X \cdot m_T \cdot m_N = 48$$

2. Комбинаторная мера, при которой количество информации вычисляется как количество комбинаций элементов.

Дискретное сообщение можно рассматривать как слово, состоящее из некоторого количества элементов n , заданных алфавитом, состоящим из m элементов-букв.

Если сообщение состоит из двух элементов ($n=2$), то всего может быть m^2 различных сообщений.

Например, из десяти цифр (0, 1, 2,..., 9) может быть образовано сто различных чисел от 0 до 99.

Если количество элементов равно трем, то количество различных сообщений равно m^3 и т.д.

Таким образом, число возможных сообщений определяется:

$$L = m^n$$

где L - число сообщений; n - число элементов в слове; m - алфавит.

Величина L может быть принята в качестве меры количества информации.

Для k источников общее число возможных различных сообщений равно

$$L = L_1 \cdot L_2 \cdots L_k$$

3. Аддитивная мера

Хартли ввел логарифмическую (аддитивную) меру количества информации, позволяющую оценивать количество информации, содержащейся в сообщении, логарифмом числа возможных сообщений:

$$I = \log_2 L_1 = n \log_2 m$$

Для k источников информации:

$$I = \log_2 L = \log_2 L_1 + \log_2 L_2 + \dots + \log_2 L_k$$

$$I = I_1 + I_2 + \dots + I_k$$

Ральф Винтон Лайон Хартли (1888-1970) — американский учёный. Он сделал вклад в теорию информации, введя в 1928 году логарифмическую меру информации, которая называется *хартлиевским количеством информации* или просто *мерой Хартли*.

Статистические меры информации

Мера Шенона.

Сумма вероятностей всех сообщений в полной группе событий равна единице:

$$P_1 + P_2 + \dots + P_k = \sum_{i=1}^k P_i = 1$$

Рассмотрим сложные сообщения, составляемые из n элементов. Предположим, что в некоторое сообщение вошло n_1 элементов алфавита X_1 , n_2 элементов алфавита X_2 и т.д.

Тип элемента	X_1	X_2	X_3	...	X_i	...	X_m
Число элементов	n_1	n_2	n_3	...	n_i	...	n_m
Вероятности выбора элементов	P_1	P_2	P_3	...	P_i	...	P_m

Вероятность того, что в сообщение войдут n_i элементов X_i равняется $P_i^{n_i}$

Вероятность образования сообщения из n_m элементов будет равна:

$$P = P_1^{n_1} \dots P_i^{n_i} \dots P_m^{n_m}$$

Вероятности появления типичных сообщений P будут одинаковы и могут быть найдены из:

$$P = \prod_{i=1}^m P_i^{n_i} P_i$$

Число типичных сообщений: $L = \frac{1}{P}$

Количество информации I , содержащейся в одном сообщении:

$$I = \log L = -\log \prod_{i=1}^m P_i^{n_i P_i} = -\sum_{i=1}^m n_i P_i \log P_i = -n \sum_{i=1}^m P_i \log P_i$$

- формула Шенона.

Она дает более полное представление об источнике информации, чем аддитивная мера Хартли.

Количество информации, приходящейся на один элемент называется удельной информативностью или энтропией:

$$H = \frac{I}{n} = -\sum_{i=1}^m P_i \log P_i$$

Семантические меры информации

Семантические меры информации оценивают смысл, содержание информации, ее целесообразность и существенность.

Целесообразность, полезность информации для решения какой-то задачи можно оценить по эффекту, который оказывает полученная информация на решение задачи.

Если вероятность достижения цели увеличивается, то информацию следует считать полезной.

Информационная модель и формирование ее элементов

Моделирование — это метод познания, состоящий в создании и исследовании моделей.

Модель — это некий новый объект, который отражает существенные особенности изучаемого объекта, явления или процесса.

Модели бывают:

- статические и динамические;
- предметные и информационные.

Статическая модель — это одномоментный срез информации по объекту.

Динамическая модель позволяет увидеть изменения объекта во времени.

Предметные модели воспроизводят геометрические, физические и другие свойства объектов в материальной форме.

Информационные модели представляют объекты и процессы в образной или знаковой форме.

Образные модели (рисунки, фотографии и др.) представляют собой зрительные образы объектов, зафиксированные на каком-либо носителе информации (бумаге, фото- и киноплёнке и др.).

Знаковые информационные модели строятся с использованием различных языков (знаковых систем).

Знаковая информационная модель может быть представлена в форме текста, формулы или таблицы.

Основные этапы разработки и исследования моделей на компьютере

Компьютерное моделирование является одним из эффективных методов изучения сложных систем.

Процесс разработки моделей и их исследования на компьютере можно разделить на несколько основных этапов:

- 1. Построение описательной информационной модели (выделение существенных параметров).**
- 2. Создание формализованной модели (запись формул).**
- 3. Построение компьютерной модели.**
- 4. Компьютерный эксперимент.**
- 5. Анализ полученных результатов и корректировка исследуемой модели.**

1. Описательная информационная модель

Такая модель выделяет существенные с точки зрения целей проводимого исследования параметры объекта, а несущественными параметрами пренебрегает.

2. Формализованная модель.

Описательная информационная модель записывается с помощью какого-либо формального языка.

В такой модели с помощью формул и уравнений неравенств фиксируются формальные соотношения между значениями свойств объектов.

3. Компьютерная модель.

Формализованную информационную модель необходимо преобразовать в компьютерную на понятном для компьютера языке.

Существуют два принципиально различных пути построения компьютерной модели:

- 1) создание алгоритма решения задачи и его кодирование на одном из языков программирования;**
- 2) формирование компьютерной модели с использованием одного из приложений (электронных таблиц, СУБД и т. д.).**

В процессе создания компьютерной модели полезно разработать удобный графический интерфейс, который позволит визуализировать формальную модель.

4. Компьютерный эксперимент.

Если компьютерная модель существует в виде программы на одном из языков программирования, ее нужно запустить на выполнение и получить результаты.

Если компьютерная модель исследуется в приложении, например в электронных таблицах, можно провести сортировку или поиск данных, построить диаграмму или график и так далее.

5. Получение результатов и корректировка исследуемой модели.