
2. ПРЕДСТАВЛЕНИЕ ИНФОРМАЦИИ

2.1 Уровни передачи информации

Сообщение – форма представления информации в виде совокупности знаков (символов), используемая для передачи.

Уровни изучения сообщений:

- синтаксический;
 - семантический;
 - прагматический.
-

2.1 Уровни передачи информации

- На *синтаксическом уровне* рассматриваются внутренние свойства сообщений, т.е. отношения между знаками, отражающие структуру данной знаковой системы.
- На *семантическом уровне* анализируются отношения между знаками, т.е. смысловое содержание сообщения, его отношение к источнику информации.
- На *прагматическом уровне* рассматриваются отношения между сообщением и получателем, т.е. потребительское содержание сообщения, его отношение к получателю.

2.1. Меры информации

Различают меры информации:

- *синтаксического уровня,*
 - *семантического уровня,*
 - *прагматического уровня.*
-

2.1. Меры информации

Параметры оценки информации синтаксического уровня:

- объем информации (данных) V_D
- количество информации I

2.1. Меры информации

Объем информации (данных) V_D в сообщении измеряется количеством символов (разрядов) в этом сообщении.

- В десятичной системе счисления один разряд имеет вес, равный 10. Единица измерения информации будет **дит** (десятичный разряд).
- В двоичной системе счисления один разряд имеет вес, равный 2, и соответственно единицей измерения информации будет – **бит** (*bit – binary digit* – двоичный разряд).

2.1. Меры информации

Единицы измерения информации:

- 1 Килобайт=1024 байт= 2^{10} байт;
- 1 Мегабайт=1024 Килобайт= 2^{20} байт;
- 1 Гигабайт=1024 Мегабайт = 2^{30} байт;
- 1 Терабайт=1024 Гигабайт= 2^{40} байт;
- 1 Петабайт=1024 Терабайт= 2^{50} байт;
- 1 Экзабайт=1024 Петабайт= 2^{60} байт;
- 1 Зеттабайт=1024 Экзабайт= 2^{70} байт;
- 1 Йоттабайт =1024 Зеттабайт= 2^{80} байт .

2.1. Меры информации

Количество информации I

Энтропия системы это мера неопределенности системы.

- α - априорные сведения о системе;
- $H(\alpha)$ - мера неосведомленности потребителя;
- $I_{\beta}(\alpha)$ - информация после получения сообщения;
- $H_{\beta}(\alpha)$ - апостериорная неопределенность состояния системы;

Количество информации о системе

$$I_{\beta}(\alpha) = H(\alpha) - H_{\beta}(\alpha)$$

2.1. Меры информации

Энтропия дискретной системы

Дискретный источник информации - физическая система, имеющую конечное множество возможных состояний $\{a_i\}, i = \overline{1, N}$
Множество возможных сообщений $A = \{a_1, a_2, \dots, a_i\}$
называется **абстрактным алфавитом**.

Мера неопределенности источника с равновероятными состояниями:

$$H(A) = \log N$$

2.1. Меры информации

Единицы измерения неопределенности:

- **Биты** – основание логарифма равно 2:

$$H(A) = \log_2 N$$

- **Натты** – основание логарифма e :

$$H(A) = \ln N$$

- **Диты** - основание логарифма 10:

$$H(A) = \lg N$$

2.1. Меры информации

Если сообщение состоит из n равновероятных состояний, то количество информации равно $\log_2 n$

Примеры $\log_2 32 = 5, \log_2 16 = 4, \log_2 8 = 3$

Количество информации 2^I , где $I = 4$ (бита)

Энтропия:

$$H = -\sum_{i=1}^N p_i \log p_i \quad (2.3); \quad H = -\sum_{i=1}^N p_i \log_2 p_i \quad (2.4);$$

$$H = -\sum_{i=1}^N p_i \log_2 p_i = -N * \frac{1}{N} \log_2 \frac{1}{N} = -N * \frac{1}{N} (\log_2 1 - \log_2 N) = \log_2 N$$

2.1. Меры информации

Семантическая мера.

Для измерения количества информации на семантическом уровне, наибольшее распространение получила тезаурусная мера, которая связывает семантические свойства информации со способностью пользователя принимать поступившие сообщения.

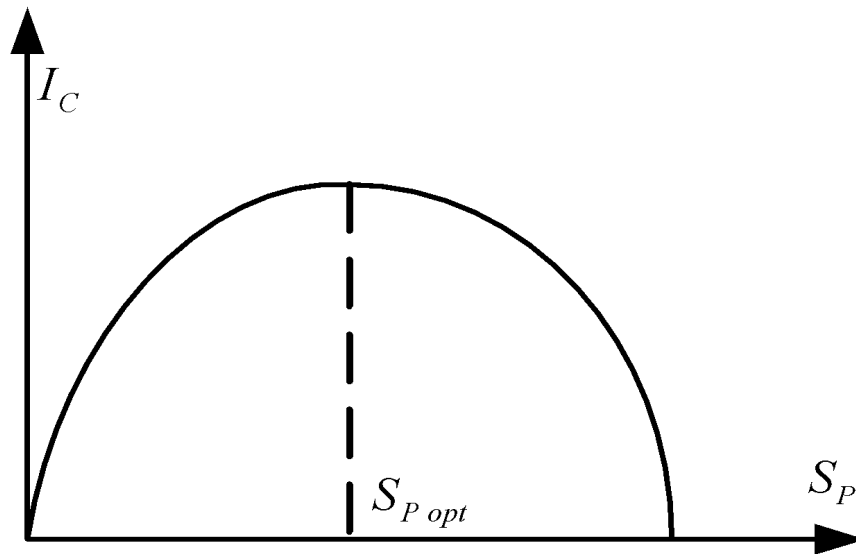
- **Тезаурус** - совокупностью сведений, которыми располагает пользователь или система.
-

2.1. Меры информации

S - смысловое содержание информации;

S_P - тезаурус пользователя;

I_C - количество семантической информации.



$$C = \frac{I_C}{V_D}$$

2.1. Меры информации

Прагматическая мера определяет полезность (ценность) информации для достижения пользователем поставленной цели.

$$I_{n\beta}(\gamma) = \Pi(\gamma/\beta) - \Pi(\gamma)$$

2.3. Показатели качества информации

- Репрезентативность
 - Содержательность
 - Достаточность
 - Доступность
 - Актуальность
 - Своевременность
 - Точность
 - Достоверность
 - Устойчивость
-

2.3. Показатели качества информации

- **Репрезентативность** информации связана с правильностью ее отбора и формирования в целях адекватного отражения свойств объекта.
 - **Содержательность** информации отражает семантическую емкость, равную отношению количества семантической информации в сообщении к объему обрабатываемых данных.
-

2.3. Показатели качества информации

- **Достаточность** (полнота) информации означает, что она содержит минимальный, но достаточный для принятия правильного решения состав (набор показателей).
 - **Доступность** информации восприятию пользователя обеспечивается выполнением соответствующих процедур ее получения и преобразования.
-

2.3. Показатели качества информации

- **Актуальность** информации определяется степенью сохранения ценности информации для управления в момент ее использования и зависит от динамики изменения ее характеристик и от интервала времени, прошедшего с момента возникновения данной информации.
 - **Своевременность** информации означает ее поступление не позже заранее назначенного момента времени, согласованного с временем решения поставленной задачи.
-

2.3. Показатели качества информации

- **Точность** информации определяется степенью близости полученной информации к реальному состоянию объекта, процесса, явления и т.п.
 - а) формальная точность;
 - б) реальная точность;
 - в) максимальная точность;
 - г) необходимая точность.
-

2.3. Показатели качества информации

- **Достоверность** информации определяется ее свойством отражать реально существующие объекты с необходимой точностью.
 - **Устойчивость** информации отражает ее способность реагировать на изменения исходных данных без нарушения необходимой точности.
-

2.4 Системы классификации информации

Классификация – система распределения объектов (предметов, явлений, процессов, понятий) по классам в соответствии с определенными признаками.

Объект это любой предмет, процесс, явление материального или нематериального свойства.

Реквизит – логически неделимый информационный элемент, описывающий определенные свойства объекта, процесса, явления и т.п.

2.4 Системы классификации информации

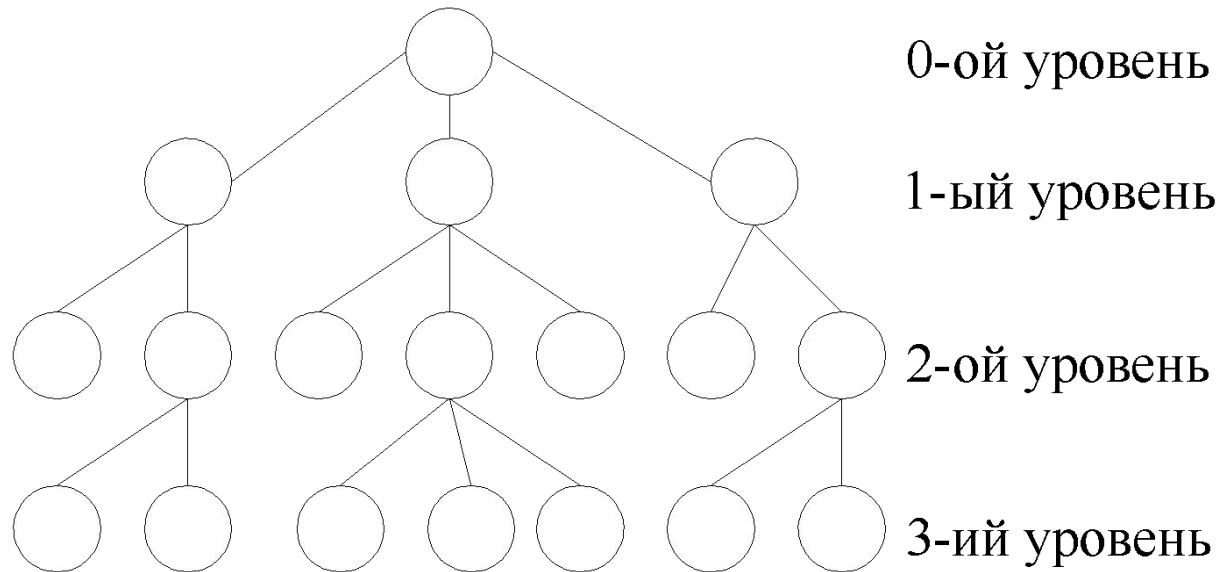
Классификатор – систематизированный свод наименований и кодов классификационных группировок.

Методы классификации объектов:

- *иерархический;*
 - *фасетный;*
 - *дескрипторный.*
-

2.4 Системы классификации информации

Иерархическая система классификации



2.4 Системы классификации информации

Фасетная система классификации

		Фасеты				
		Φ_1	Φ_2	Φ_3	...	Φ_n
Значения фасет	1	●	●		●	●
	2		●	●		●
	...					
	k	●			●	

2.4 Системы классификации информации

Дескрипторная система классификации

1. Отбирается совокупность ключевых слов или словосочетаний.
2. Выбранные ключевые слова и словосочетания подвергаются нормализации.
3. Создается словарь дескрипторов.
4. Между дескрипторами устанавливаются связи.

Связи могут быть трех видов:

- синонимические;
 - родо-видовые;
 - ассоциативные.
-