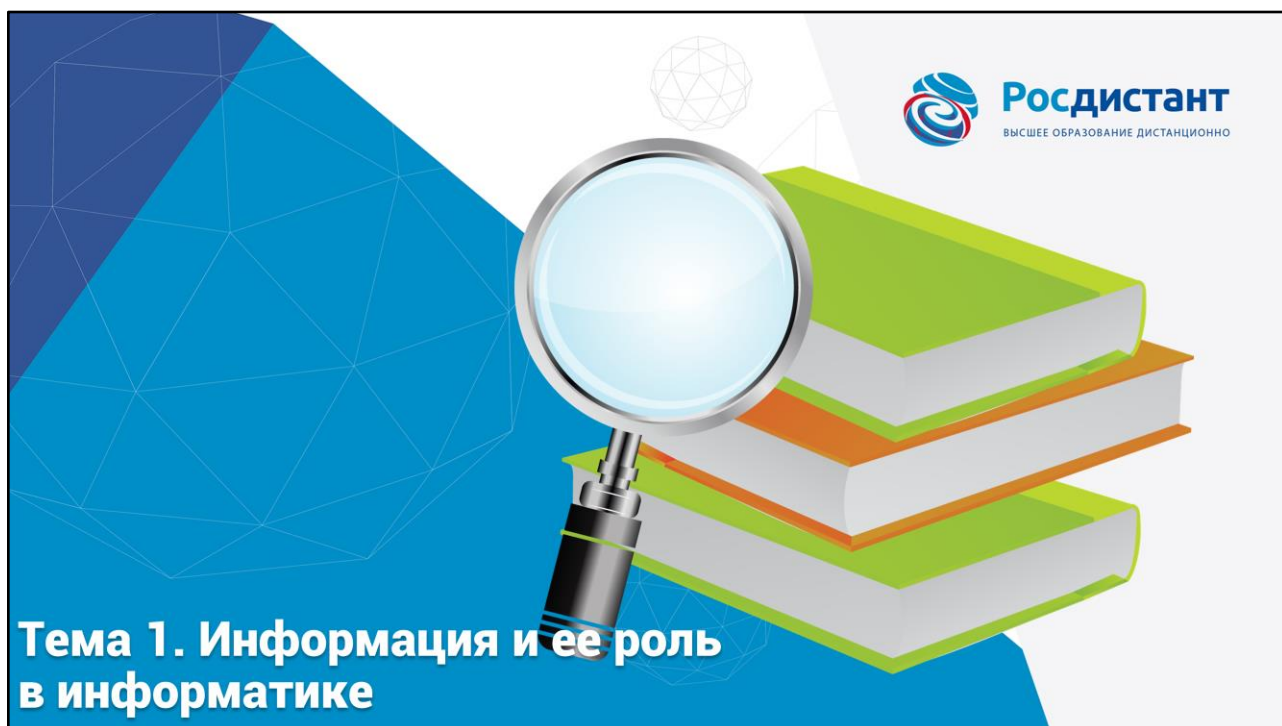




Росдистант
ВЫСШЕЕ ОБРАЗОВАНИЕ ДИСТАНЦИОННО

МЕТОДЫ РЕШЕНИЯ ПРОБЛЕМ В

ИНФОРМАТИКЕ 1 ЧАСТЬ



Тема 1. Информация и ее роль в информатике

В рамках данной темы формируется представление о роли информации в развитии информатики. Представленный материал знакомит с основными признаками информации, подходами к ее измерению, измерению ее объема. Рассматриваются понятие кодирования информации, правила расчета объема информации, цикл обработки информации и формы, которые принимает информация на каждом этапе обработки.

ИНФОРМАТИКА И ИНФОРМАЦИЯ

Информатика	<ul style="list-style-type: none">▪ наука, которая изучает информационные процессы▪ сосредоточена на всех аспектах процессов и методов информации
Объект изучения информатики	<ul style="list-style-type: none">▪ информация
Информация	<ul style="list-style-type: none">▪ выражается в виде сообщения или наблюдения▪ может быть закодирована
Формы информации	<ul style="list-style-type: none">▪ аналоговая▪ непрерывная▪ дискретная



Ознакомление с любой дисциплиной следует начинать с определения предмета изучения этой науки.

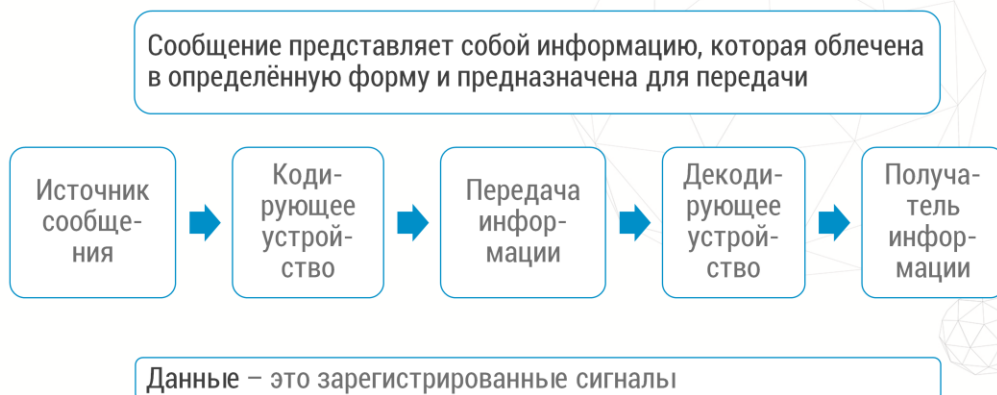
Информатика – это наука, которая изучает информационные процессы и способы управления ими. Информатика сосредоточена на всех аспектах процессов и методов информации. Информатика рассматривает взаимодействие между людьми и информацией наряду с построением интерфейсов, организаций, технологий и систем.

В связи с тем, что основным объектом изучения информатики является информация, следует более подробно остановиться на природе этого понятия. Слово «информация» произошло от латинского *informatio* – «осведомление, представление, изложение».

Информация представляется в виде знаков, иллюстраций, фотоизображений, звуковых и световых волн, радиоволн, нервных и электрических импульсов, ощущений запаха, вкуса и так далее. Информация связана с данными, поскольку данные представляют собой значения, приписываемые параметрам, а информация – это данные в контексте и со смыслом. С точки зрения коммуникации информация выражается либо в виде содержания сообщения, либо посредством прямого или косвенного наблюдения. То, что воспринимается, может быть истолковано как сообщение само по себе, и в этом смысле информация всегда передается как содержание сообщения. Информация может быть закодирована в различных формах для передачи и интерпретации. Например, информация может быть закодирована в

последовательности из признаков или передается через сигнал. Так как информационными носителями выступают сигналы, то выделяют соответствующие им виды информации – аналоговую, или непрерывную, и дискретную. В качестве сигналов могут выступать физические процессы: движение электрического тока, физическое движение тела, движение световых волн и другие. Информация отображается значением одного или сочетанием нескольких показателей сигнала.

ИНФОРМАЦИЯ И ДАННЫЕ



Информация поступает от источника информации к получателю в виде сообщений. Сообщение представляет собой информацию, которая облечена в определённую форму и предназначена для передачи. Источник отправляет сообщение. При этом сообщение источника кодируется в сигнал, передаваемый по каналу связи. В приёмнике сигнал декодируется в исходное сообщение. Сообщение, содержащее информацию о прогнозе погоды, посылается получателю, например зрителю у экрана телевизора, от источника – метеоролога с помощью канала передачи – телеаппаратуры и телевидения. Трансляция информации по каналам передачи обычно сопровождается влиянием помех, которые искажают сигнал и могут привести к потере информации.

Данные – это зарегистрированные сигналы. В зависимости от метода регистрации данные хранятся на разных носителях: бумаге, жёстком диске, флешке и так далее.

Данные и информация настолько тесно взаимосвязаны, что существование одного без другого либо невозможно, либо не имеет смысла.

Информация основывается на данных. Данные – это совокупность сигналов. Однако после восприятия неким получателем данные становятся информацией. Информация появляется лишь тогда, когда есть некоторый источник, который содержит данные, и приёмник. Данные преобразовываются в информацию разными способами: с помощью подсчёта, изменения, сжатия, разбиения на категории и так далее.

Мозг человека обрабатывает полученные данные и выдаёт некоторую информацию. Если необходимо её применить для иного мыслительного процесса, то для него уже эта информация является данными, из которых получается новая информация. Итоговым этапом преобразования информации, которая прошла ряд обработок в течение какого-то интервала времени, становятся знания.

КАЧЕСТВО ИНФОРМАЦИИ



Качество информации – это концепция с множеством атрибутов. Если атрибуты, определяющие качество информации, имеют хорошее качество или высокую ценность, то считается, что информация имеет хорошее качество. Признаками качества информации являются:

- репрезентативность – степень верности отбора информации для наиболее полного отображения объекта или процесса, описываемого этой информацией;
- достаточность – минимальный, но достаточный объём информации для реализации задач, преследуемых получателем информации;
- доступность – степень вероятности получения информации;
- актуальность – мера сохранения важности информации в процессе её применения, которая может меняться со сменой свойств информации во времени;
- надежность – достоверность информации – ключевой атрибут качества. Только если информация достоверна, она может быть полезной. Понимание надежности исходит из прошлого опыта, статуса и надежности источника, методологии, принятой для сбора и обработки информации, а также канала доставки.
- своевременность – получение информации не позже определённого момента, связанного со временем решения задачи;
- точность – мера приближения поступающей информации к настоящему состоянию объекта;

- адекватность – мера соответствия рассматриваемому объекту модели этого объекта, генерируемой с помощью поступающей информации;
- устойчивость – готовность информации отвечать на получение новых данных с требуемой точностью.

Эти атрибуты определяют качество информации. Высокий балл по каждому из атрибутов указывает на хорошее качество информации.

СИНТАКСИЧЕСКАЯ МЕРА ИНФОРМАЦИИ

$$I_{\beta}(\alpha) = H(\alpha) - H_{\beta}(\alpha) \quad (1)$$

- $I_{\beta}(\alpha)$ — количество информации в сообщении β
- $H(\alpha)$ — энтропия (мера неосведомлённости) до получения сообщения β
- $H_{\beta}(\alpha)$ — энтропия после получения сообщения β

Информация понимается и интерпретируется по-разному в разных предметных областях. Поэтому существуют разные подходы к измерению информации.

Количество информации — величина, характеризующая информацию по разнообразию, сложности, упорядоченности, определённости, выбору состояний отображаемой системы. Для измерения количества информации применяется понятие меры информации. Имеются разные подходы к определению меры информации.

При синтаксическом подходе мера информации, которая характеризует количество информации в сообщении, есть числовая характеристика, отражающая неопределённость, исчезающую после получения сообщения. Эта мера неопределённости в теории информации и кодирования называется энтропией.

Бит — объём информации, при получении которой происходит выбор из двух равновероятных вариантов. К примеру, результат бросания монеты даёт нам информацию размером в один бит.

Предположим, что до получения сведений у адресата уже имеются некие сведения о системе α . Мерой его неосведомлённости о состоянии системы является некоторая функция $H(\alpha)$, называемая энтропией. После получения некоторого сообщения β получатель приобрёл дополнительную информацию, уменьшившую его неосведомлённость. Это уменьшение энтропии определяет количество информации в сообщении. Иначе говоря, количество информации в сообщении β есть разность между величинами энтропии до и после получения

сообщения. Формула, отражающая этот факт, на слайде обозначена номером 1. Основным препятствием для применения этой формулы на практике часто является сложность или невозможность вычисления значения энтропии без введения дополнительных ограничений на рассматриваемую систему.

ФОРМУЛА ШЕННОНА

$$H(\alpha) = - \sum_{i=1}^n p_i \cdot \log_2 p_i = -(p_1 \cdot \log_2 p_1 + p_2 \cdot \log_2 p_2 + \dots + p_n \cdot \log_2 p_n) \quad (2)$$

- n – число возможных состояний системы
- p_i – вероятность состояния с номером i

В частном случае для системы, имеющей конечное число n возможных состояний, энтропия этой системы может быть вычислена по формуле Шэннона. Формула обозначена на слайде номером 2.

Согласно этой формуле, сообщение о наступлении события, имеющего меньшую вероятность, содержит больше информации, нежели сообщение о наступлении события, вероятность которого больше. Сообщение о наступлении достоверно наступающего события не содержит никакой информации.

Главным недостатком формулы Шеннона является то, что множество реальных систем не укладываются в предлагаемую модель. Далеко не всегда заранее возможно определить все состояния системы и найти их вероятности.

Теория энтропии была первоначально создана Шенноном как часть его теории коммуникации, в которой система передачи данных состоит из трех элементов: источника данных, канала связи и приемника. В теории Шеннона «фундаментальная проблема связи» заключается в том, чтобы приемник мог определить, какие данные были сгенерированы источником, на основе сигнала, который он получает через канал. Шеннон рассмотрел различные способы кодирования, сжатия и передачи сообщений из источника данных. Он доказал в своей знаменитой теореме кодирования источника, что энтропия представляет собой абсолютный математический предел того, насколько хорошо данные из источника могут быть обработаны без потерь. Шеннон значительно усилил этот результат для зашумленных каналов в своей теореме кодирования зашумленных каналов.

ФОРМУЛА ХАРТЛИ

$$I = \log_2 n \quad (3)$$

Пример:

Из восьми карточек с номерами от 1 до 8 случайным образом извлекли одну. Определить, какое количество информации будет заключено в сообщении о том, какое число записано на карточке

Решение: $I = \log_2 8 = \log_2 2^3 = 3$ бит

Использовано соотношение:

$$\log_2 2^k = k \quad (4)$$

$$I_\beta = \log_2 \frac{n1}{n2} \quad (5)$$



Рассмотрим случай, когда количество состояний системы конечно и все состояния системы равновероятны. В этом случае для вычисления количества информации, содержащейся в сообщении об одном состоянии системы, используется формула Хартли.

В соответствии с этой формулой, получение информации можно рассматривать как выбор одного сообщения из n равновероятных сообщений. Количество информации I , содержащееся в одном сообщении, вычисляется как логарифм от n по основанию два. Формула обозначена на слайде номером 3.

Рассмотрим применение формулы Хартли на примере. Из восьми карточек с номерами от 1 до 8 случайным образом извлекли одну. Определить, сколько информации будет заключено в сообщении о том, какое число записано на карточке.

Рассматриваемое событие, которое заключается в извлечении карточки, имеет восемь возможных исходов, вероятности которых одинаковы. То есть для рассматриваемой системы справедлива формула Хартли. Расчет по этой формуле дает результат 3 бита. Ход вычислений показан на слайде.

В вычислениях мы воспользовались одним из свойств логарифмической функции, согласно которому логарифм от числа 2 в степени k по основанию 2 равен k . Соответствующая формула приведена на слайде и обозначена номером 4.

Предположим, что число возможных равновероятных состояний системы равно $n1$. Получено некоторое сообщение β о системе, в результате чего число

возможных состояний системы уменьшилось и стало равным n_2 . В этом случае количество информации в сообщении I_β можно определить по формуле Хартли, содержащей вместо n [эн] дробь, в числителе которой n_1 , а в знаменателе – n_2 . Запись формулы для этого случая обозначена на слайде номером 5. Примеры её применения будут показаны на следующих слайдах.

ПРИМЕР РАСЧЕТОВ ПО ФОРМУЛЕ ХАРТЛИ

Пример 1

- Карточная колода включает 16 чёрных и 16 красных карт; из колоды случайным образом берётся одна карта; получено сообщение о том, что извлечена карта черной масти; какое количество информации содержит это сообщение?

Решение: $I_{\beta} = \log_2 \frac{n_1}{n_2}$

- Расчётная формула:
 - $n_1=32$; $n_2=16$
 - $I = \log_2 \frac{n_1}{n_2} = \log_2 \frac{32}{16} = \log_2 2^1 = 1$ бит



Функция Хартли зависит только от количества элементов в наборе и, следовательно, может рассматриваться как функция от натуральных чисел. Рассмотрим пример применения формулы Хартли при решении задачи. Карточная колода включает 16 чёрных и 16 красных карт. Из этой колоды случайным образом берётся одна карта. Получено сообщение об извлечении чёрной карты. Определить количество информации в этом сообщении. Вероятность извлечения любой карты из колоды равна для всех карт. Значит, для этой системы может быть применена формула Хартли. Количество возможных состояний системы до получения сообщения соответствует числу карт, то есть $n_1 = 32$.

После получения сообщения количество возможных состояний стало соответствовать количеству чёрных карт в колоде, то есть $n_2 = 16$.

В расчетах применим формулу Хартли. Количество информации в полученном сообщении составит 1 бит. Порядок вычислений показан на демонстрируемом слайде.

ПРИМЕР РАСЧЕТОВ ПО ФОРМУЛЕ ХАРТЛИ

Пример 2

- Карточная колода включает 32 карты, при этом в колоде находятся все 4 туза и 1 пиковая дама; из колоды случайным образом берётся одна карта; определить, какое из двух сообщений I_1 = «извлечён туз» или I_2 = «извлечена дама пик» содержит большее количество информации

Решение:

- $n_1=32$; $n_{21} = 4$; $n_{22} = 1$
- $I_1 = \log_2 \frac{n_1}{n_{21}} = \log_2 \frac{32}{4} = \log_2 8 = \log_2 2^3 = 3 \text{ бит}$
- $I_2 = \log_2 \frac{n_1}{n_{22}} = \log_2 \frac{32}{1} = \log_2 32 = \log_2 2^5 = 5 \text{ бит}$



Рассмотрим еще один пример. Карточная колода включает 32 карты, при этом в колоде находятся 4 туза и 1 пиковая дама. Из этой колоды случайным образом берётся одна карта. Определить, какое из двух сообщений: I_1 = «извлечён туз» или I_2 = «извлечена дама пик» – содержит большее количество информации. В этом примере, как и в предыдущем, $n_1 = 32$. Для сообщения «извлечён туз» количество возможных состояний системы после получения сообщения, обозначим его через n_{21} , равно 4, так как среди карт имеется четыре туза. Аналогично для сообщения «извлечена пиковая дама» $n_{22} = 1$, так как среди карт имеется только одна пиковая дама.

Порядок вычислений демонстрируется на слайде. Количество информации в сообщении «извлечён туз» составит 3 бита. Количество информации в сообщении «извлечена пиковая дама» – 5 бит. Полученные значения позволяют утверждать, что во втором сообщении содержится больше информации, чем в первом.

Рассмотренный пример иллюстрирует одну важную характеристику синтаксического подхода к измерению информации. В рамках этого подхода сообщение о наступлении менее вероятного события содержит больше информации, нежели сообщение о наступлении более вероятного события.

ДРУГИЕ ПОДХОДЫ К ИЗМЕРЕНИЮ ИНФОРМАЦИИ



Итак, мы рассмотрели синтаксический подход к измерению информации. Существуют и другие подходы и соответствующие им меры информации. Семантическая мера применяется для оценки смысла информации. Используется тезаурусная мера, связывающая семантические признаки информации со способностью получателя принимать переданное сообщение, а также с новизной данных, имеющих в сообщении. Тезаурус представляет собой набор информации, имеющийся у получателя информации. Наибольшее количество информации, согласно семантическому подходу, потребитель получает при согласовании её смысла с тезаурусом, когда новые сведения доступны получателю и содержат ранее неизвестную ему информацию. При прагматическом подходе к измерению информации оценивается значимость сведений для достижения получателем поставленной цели, то есть оценивается полезность или ценность сведений. Прагматическая мера, как и семантическая, является относительной величиной, зависящей от особенностей применения получаемых сведений. Полезность сведений обычно оценивают в тех же единицах измерения, в которых измеряется целевая функция. Например, в экономике ценность сведений зависит от эффективности управления, которое осуществляется на базе получаемой информации. В этом случае для измерения информации используются денежные единицы. Существует еще одно видение измерения информации, согласно которому выделяют следующие подходы: содержательный, алфавитный и вероятностный. Содержательный подход означает, что к измерению информации подходят с

точки зрения её содержания, т. е. определяют, в какой мере пришедшая информация-знание уменьшает незнание.

Алфавитный подход используется тогда, когда для преобразования, хранения и передачи информации применяют технические средства.

Вероятностный подход полностью разработан в середине XX века, применяется для измерения количества информации при наступлении событий, имеющих разную вероятность.

ИЗМЕРЕНИЕ ОБЪЕМА ДАННЫХ

bit = binary digit
dit = decimal digit

Единицы измерения объёма данных		
Обозначение	Название	Значение в байтах
Кбайт	килобайт	2^{10}
Мбайт	мегабайт	2^{20}
Гбайт	гигабайт	2^{30}
Тбайт	терабайт	2^{40}
Пбайт	петабайт	2^{50}



Как уже было сказано, информация появляется только при преобразовании данных. Закодированная информация, будучи записанной на носителях информации, таких как магнитный диск или флешка, представляется в виде данных. В этой связи представляет интерес вопрос об измерении объёма данных.

Объём данных в сообщении оценивается числом содержащихся в нём знаков. В соответствии с применяемым кодированием каждый знак имеет разный вес. При двоичном кодировании в качестве единицы измерения используется бит. Название этой единицы происходит от английского *binary digit*, что переводится как «двоичная цифра».

Помимо бита применяется более крупная единица измерения – байт, равная восьми битам. Для обозначения производных от байта единиц измерения объёма данных применяются килобайт, мегабайт, гигабайт, терабайт, петабайт. Один килобайт равен двум в десятой степени байтам, а один мегабайт равен двум в десятой степени килобайтам, что составляет два в двадцатой степени байта, и так далее.

Например, объём современного жёсткого диска измеряется в терабайтах, объём флешки – в гигабайтах, объём файла с текстом книги – в мегабайтах. Заметим, что если использовать не двоичное, а десятичное кодирование, то потребуются другие единицы измерения объёма данных. Здесь в качестве единицы измерения используется дит – от английского *decimal digit*, что переводится как «десятичная цифра».

КОДИРОВАНИЕ ТЕКСТОВОЙ ИНФОРМАЦИИ

Таблица кодировки ASCII ($2^7 = 128$ символов)

	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9	.A	.B	.C	.D	.E	.F
0.	NU L	SOH	STX	ETX	EOT	EN Q	AC K	BEL	BS	HT	LF	VT	FF	CR	SO	SI
1.	DLE	DC1	DC2	DC3	DC 4	NA K	SY N	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2.		!	"	#	\$	%	&	'	()	*	+	,	-	.	/
3.	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4.	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5.	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6.	`	a	b	c	d	e	F	G	h	i	j	k	l	m	n	o
7.	p	q	r	s	t	u	V	W	x	y	z	{		}	~	DEL



В окружающем нас мире информация представляется в самых разных формах. Однако для того, чтобы компьютер мог воспринять информацию и применить к ней какие-либо операции, эта информация должна быть представлена в виде двоичного кода, состоящего из единиц и нулей. Именно такой способ представления данных используется внутри компьютера.

Процесс замены информации исходного сообщения условными обозначениями называется кодированием. Код представляет собой сочетание условных знаков для отображения информации.

Рассмотрим способы кодирования информации текстового типа. Так как исходное текстовое сообщение состоит из отдельных символов, то для двоичного кодирования текста применяется метод, при котором все знаки заменяются числами. Таким образом, код текстового сообщения представляет собой совокупность чисел – кодов знаков, которые его составляют.

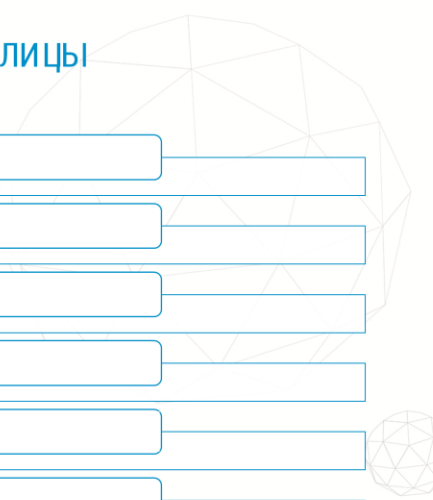
В 1963 году для кодирования текстовой информации была разработана таблица ASCII, в которой для записи кода одного знака отводилось семь бит. В этом случае число всех комбинаций из нулей и единиц равно 128. Данное число является результатом возведения числа 2 в степень 7. Это позволяло закодировать 128 разных знаков – десятичные цифры, заглавные и строчные буквы латинского алфавита, управляющие конструкции. Первые тридцать два кода обозначают управляющие конструкции, необходимые для передачи команд управления.

Таблица кодов ASCII показана на слайде. В первом столбце таблицы указаны

возможные значения первой цифры кода символа. В первой строке – значения второй цифры кода. Заметим, что эти значения представлены в таблице не в десятичной, а в шестнадцатеричной системе счисления.

Возвращаясь к таблице кодирования, можно отметить, что, например, на пересечении четвёртого столбца и четвёртой строки находится заглавная латинская буква *D*. Это означает, что буква *D* имеет шестнадцатеричный код 44. Этот код будет храниться в памяти компьютера в виде соответствующего этому числу двоичного кода.

КОДИРОВАНИЕ СИМВОЛОВ КИРИЛЛИЦЫ



КОИ-8	
Windows-1251	
CP866	
MacCyrillic	
ISO 8859-5 ($2^8 = 256$ символов)	
Unicode ($2^{16} = 65\,536$ символов)	

В вычислениях, хранении данных и передачах данных кодировка символов используется для представления репертуара символов с помощью каких-то кодирующих систем. В зависимости от уровня абстракции и контекста соответствующие кодовые точки и результирующее кодовое пространство могут рассматриваться как битовые комбинации, октеты, натуральные числа, электрические импульсы и т. д. Кодирование символов используется при вычислениях, хранении данных и передаче текстовых данных.

Ранние коды символов, связанные с оптическим или электрическим телеграфом, могли представлять только подмножество символов. Эти символы использовались в письменных языках, иногда ограничиваясь только прописными буквами, цифрами и некоторыми знаками пунктуации. Низкая стоимость цифрового представления данных в современных компьютерных системах позволяет использовать более сложные коды символов, например Unicode, которые представляют большинство символов, используемых во многих письменных языках. Кодирование символов с использованием международных признанных стандартов позволяет обмениваться текстами в электронной форме во всем мире.

Для кодирования символов кириллицы применяют пять таблиц кодировки:

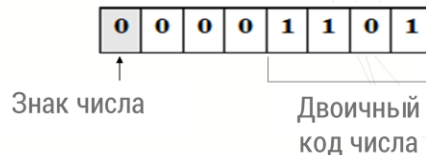
- КОИ-8;
- windows-1251;
- CP866;
- MacCyrillic ;

- *ISO 8859-5*.

Важно отметить, что текстовая информация, закодированная с помощью одной из кодировок, не будет корректно отображаться в другой кодировке, так как в разных кодировках одному и тому же коду соответствуют разные символы. С проявлением этого факта наверняка многие сталкивались при открытии в браузере страниц из Интернета, когда вместо букв отображались непонятные символы. В современных браузерах эта проблема легко решается выбором другой кодировки.

КОДИРОВАНИЕ ЦЕЛЫХ ПОЛОЖИТЕЛЬНЫХ ЧИСЕЛ

Представление положительного десятичного числа 13 в виде восьмибитового кода:



Число	13
Прямой код	0,0001101
Обратный код	0,0001101
Дополнительный код	0,0001101



При компьютерной обработке целые числа и числа с дробной частью кодируются по-разному. Чтобы не усложнять изложение материала, ограничимся рассмотрением кодирования только целых чисел.

Для записи кода целого числа со знаком можно использовать 8, 16 или 32 бита. При этом чем большее число битов используется, тем больший диапазон чисел можно закодировать.

Для простоты будем рассматривать кодирование в памяти компьютера целых чисел со знаком с использованием восьмибитового формата. В этом случае крайний слева бит отвечает за знак числа, а остальные семь битов – за значение числа по модулю.

Для представления чисел используются следующие формы: прямой, обратный и дополнительный коды. Особенно широко используются обратный и дополнительный коды. Они дают возможность свести все арифметические действия, выполняемые процессором, к сочетанию только операций сдвига и сложения. Это позволяет существенно упростить устройство процессора. Положительные целые числа во всех формах представления кодируются одинаково, то есть прямой код числа равен обратному и равен дополнительному коду. В старший разряд кода заносится ноль, а в остальные – двоичный код числа. Если при этом остаются незаполненные разряды, которых всего должно быть восемь, то они заполняются нолями.

Процедура определения двоичных кодов чисел, получаемых переводом числа из десятичной системы счисления в двоичную, будет рассмотрена в следующей

теме нашего курса. Пока примем на веру тот факт, что, например, двоичный код десятичного числа 13 включает четыре разряда: 1101 [один один ноль один]. На слайде показана запись прямого, обратного и дополнительного кодов для десятичного числа 13. Крайний левый старший разряд содержит ноль, указывающий на то, что число положительное. Для наглядности знаковый разряд отделен от остальных запятой. Последние 4 разряда справа занимает двоичный код числа. В оставшиеся незанятыми три разряда записаны ноли.

КОДИРОВАНИЕ ЦЕЛЫХ ОТРИЦАТЕЛЬНЫХ ЧИСЕЛ

Представление отрицательного десятичного числа – 13
в виде восьмибитового кода:

дополнение до единиц имеет
два представления 0:

- 00000000 (+0)
- 11111111 (-0)

1 0 0 0 1 1 0 1

↑
Знак
числа

Двоичный
код числа

Число	-13
Прямой код	1,0001101
Обратный код	1,1110010
Дополнительный код	1,1110011



В отличие от положительных, отрицательные целые числа во всех формах представления кодируются по-разному.

При записи прямого кода, в отличие от положительных чисел, в старший разряд заносится единица, а в остальные – модуль двоичного кода числа. Оставшиеся незанятыми разряды заполняются нолями.

Обратный код отрицательного числа получается из прямого кода заменой всех двоичных разрядов на противоположные: единиц на ноли, ноли на единицы. В знаковый разряд заносится единица.

Дополнительный код отрицательного числа образуется из обратного кода путем прибавления к младшему разряду единицы. Правила выполнения арифметических действий с числами в двоичной системе счисления будут рассмотрены в нашем курсе несколько позже.

На слайде показан процесс получения прямого, обратного и дополнительного кодов для десятичного числа –13.

В качестве альтернативы для представления отрицательных чисел можно использовать систему, известную как дополнение до единиц. Форма дополнения до единиц отрицательного двоичного числа – это побитовое НЕ, примененное к нему, то есть «дополнение» его положительного аналога. Как и представление знака и величины, дополнение до единиц имеет два представления ноля: +0 и –0, хотя в записи они представляются двойной тетрадой нолей и единиц.

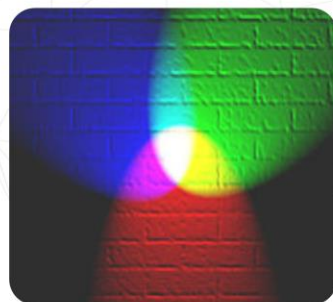
КОДИРОВАНИЕ ГРАФИЧЕСКОЙ ИНФОРМАЦИИ

Примеры кодирования цвета:

- чёрный цвет – (0, 0, 0)
- белый цвет – (255, 255, 255)

Для кодирования можно использовать следующие коды:

- 00 – чёрный цвет
- 10 – зеленый
- 01 – красный
- 11 – коричневый



Модель RGB
(Red, Green, Blue)



На экране монитора графическая информация представляется в виде отдельных точек. Код точки изображения содержит информацию о её цвете.

Графическая информация может быть представлена как в аналоговой, так и в дискретной формах. Примером аналогового изображения является картина, а дискретного – изображение, напечатанное на струйном принтере, состоящее из отдельных точек.

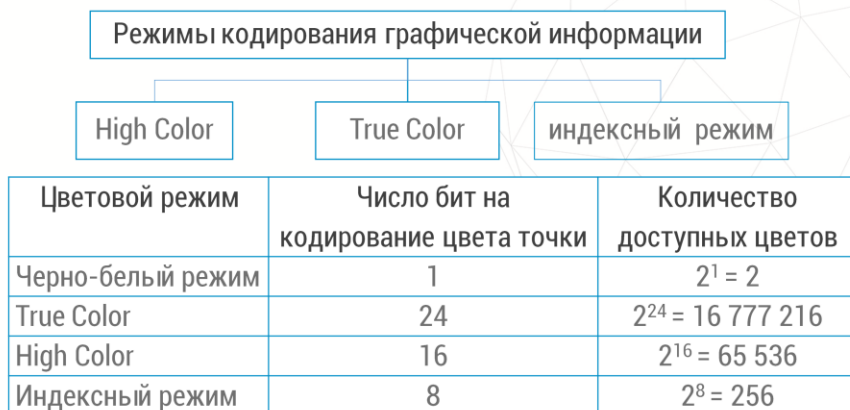
Для черно-белого изображения информационный объем одной точки равен одному биту. В чёрно-белом изображении точка может быть только двух цветов: чёрного и белого. Для кодирования одной точки достаточно одного бита. Белый цвет обозначают единицей, чёрный – через ноль.

В цветном изображении точка может принимать различные цвета, поэтому одного бита уже недостаточно. Для кодирования четырёхцветного изображения потребуются два бита на каждую точку, так как два бита могут принимать четыре разных состояния. К примеру, для кодирования можно использовать следующие коды: 00 – чёрный цвет, 10 – зеленый, 01 – красный, 11 – коричневый.

При кодировании многоцветных изображений используется разложение цвета на компоненты. Например, при выводе графической информации на экран монитора применяется модель цвета *RGB*. Согласно этой модели, любой цвет и оттенок цвета представляются с помощью сочетания трёх основных цветов: красного, зелёного и синего. Модель *RGB* получила своё название по первым буквам английских названий основных цветов. Эта модель цвета является аддитивной, так как любой цвет получается с помощью добавления к чёрному

цвету основных цветов в разных пропорциях. При совмещении одного основного цвета с другим яркость суммируется. Если добавить все три компонента в равных пропорциях, то получится ахроматический серый цвет, при увеличении яркости которого происходит приближение к белому цвету. Нулевые уровни всех основных цветов дают чёрный цвет. Максимальные уровни основных цветов дают белый цвет.

КОДИРОВАНИЕ ГРАФИЧЕСКОЙ ИНФОРМАЦИИ



Обычно на каждый основной цвет отводится восемь битов, в результате на каждую точку приходится 24 бита. Такой способ кодирования даёт возможность использовать 256 уровней каждого основного цвета, а всего получить более 16 миллионов комбинаций цветов, что вполне достаточно для человеческого глаза. Такой режим представления графической информации носит название True Color. Режим представления цветной графики называется полноцветным и позволяет зафиксировать около 16,5 млн различных цветовых оттенков при помощи 3 байт. Каждому из основных цветов можно поставить в соответствие дополнительный цвет, т. е. цвет, дополняющий основной до белого. Иногда используются другие режимы – High Color и индексный. Кодирование цветной графики 16-разрядными двоичными числами называется режимом High Color. High Color позволяет уменьшить в полтора раза размер памяти, необходимой для хранения изображения, но при этом в 256 раз уменьшается количество используемых цветов. В случае индексного кодирования используется всего 256 цветов. На каждую точку отводится восемь бит. В этом случае к данным о коде цвета каждой точки изображения добавляется справочная таблица цветов. Код точки означает номер цвета в этой таблице. Здесь код каждой точки раstra хранит не цвет, а только его номер – индекс в некоей справочной таблице, называемой палитрой, которая должна обязательно прикладываться к графическим данным.

РАСЧЕТ ОБЪЕМА ГРАФИЧЕСКОГО ФАЙЛА

Определение битовой глубины	Определение объема файла в битах
$K = 2^b$ (6) K – количество возможных цветов b – битовая глубина	$V = N \times M \times b$ (7) V – объём файла в битах N – количество точек по горизонтали M – количество точек по вертикали b – битовая глубина



Для получения на экране монитора картинки её надо где-то хранить. Для этого существует видеопамять – специальная оперативная память, в которой формируется графическое изображение.

Экран монитора может работать в двух основных режимах – текстовом и графическом. В графическом режиме экран разделяется на отдельные светящиеся точки, количество которых зависит от типа монитора. Например, 640 точек по горизонтали и 480 по вертикали. Эти значения характеризуют разрешающую способность выбранного графического режима.

Точки на экране обычно называют пикселями, их цвет и яркость могут меняться. Именно в графическом режиме на экране монитора появляются сложные графические изображения, создаваемые специальными программами, которые управляют параметрами каждого пикселя экрана.

Глубина цвета – количество битов, используемых для кодирования цвета точки. Каждый цвет можно рассматривать как возможное состояние точки. Тогда количество цветов, отображаемых на экране монитора, может быть вычислено по формуле $K = 2^b$, где K – количество цветов, b – глубина цвета, или битовая глубина. Формула обозначена на слайде номером 6.

Цветовая палитра – количество цветов, которые используются для воспроизведения изображения. Зная количество возможных цветов в палитре, можно найти значение битовой глубины.

Размер видеопамати, требуемой для вывода графического изображения на экран, вычисляется как произведение числа точек по вертикали на число точек

по горизонтали и на количество битов, используемых для кодирования цвета одной точки.

Объём файла, содержащего графическую информацию, вычисляется аналогично. Расчетная формула обозначена на слайде номером 7.

ПРИМЕР РАСЧЕТА ОБЪЕМА ГРАФИЧЕСКОГО ФАЙЛА

Исходные данные:

- Количество точек по горизонтали $N = 640$; количество точек по вертикали $M = 480$; количество возможных цветов $K = 16$

Решение:

- Находим битовую глубину b :
$$K = 16 = 2^b = 2^4 \Rightarrow b = 4$$
- Находим объем файла V :
$$V = N \times M \times b = 640 \times 480 \times 4 = 1228800 \text{ бит}$$
- Выражаем значение V в килобайтах:
$$1228800 \text{ (бит)} / 8 = 153600 \text{ (байт)} / 1024 = 150 \text{ (Кбайт)}$$



Расчёт информационного объёма растрового графического изображения, количества информации, содержащейся в графическом изображении, основан:

- на подсчёте количества пикселей в этом изображении;
- на определении глубины цвета, информационного веса одного пикселя.

В процессе дискретизации могут использоваться различные палитры цветов, т. е. наборы цветов, в которые могут быть окрашены точки изображения. Каждый цвет можно рассматривать как возможное состояние точки.

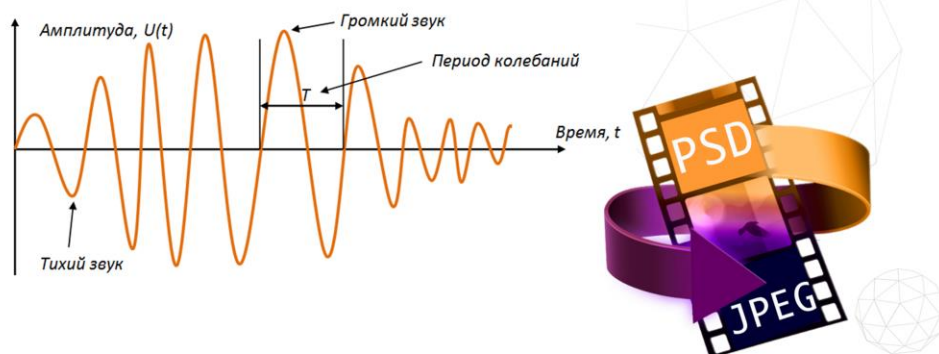
Рассмотрим пример расчета объема графического файла.

Задача. Требуется определить в килобайтах объем файла, требуемого для хранения изображения. Известны размеры изображения по горизонтали – 640 пикселей и по вертикали – 480 пикселей. Для представления изображения используется цветовая палитра, содержащая 16 цветов.

Расчетные формулы и ход решения представлены на слайде. Сначала определим битовую глубину, то есть количество битов информации, используемых для кодирования цвета одной точки. В нашем случае это значение равно 4.

Затем найдём объём файла в битах как произведение числа точек по горизонтали на число точек по вертикали и на значение битовой глубины. Для перевода результата вычислений в килобайты сначала переведём полученное значение в байты, разделив на восемь. Затем полученное число разделим на 1024. В итоге получаем результат решения задачи, равный 150 килобайтам. Аналогично рассчитывается объём файлов для других графических режимов.

КОДИРОВАНИЕ АУДИО- И ВИДЕОИНФОРМАЦИИ



Любой звук, слышимый человеком, является колебанием воздуха, которое характеризуется двумя основными показателями: частотой и амплитудой. Для обработки аудиоинформации этот непрерывный сигнал необходимо преобразовать в последовательность нулей и единиц. Для этого можно измерить напряжение через одинаковые временные интервалы и сохранить полученные значения. Такой процесс получил название дискретизации звука. Для дискретизации аудиоинформации применяются аналого-цифровые преобразователи. Для выполнения преобразования цифровых значений в аудиоинформацию применяются цифро-аналоговые преобразователи. Имеет значение и число битов, отводимых на кодирование одного значения. Этот параметр называют глубиной кодирования. Для записи звука на компакт-диск обычно используется глубина кодирования, равная 16 битам. Наиболее популярный формат звукозаписи MP3 позволяет с приемлемым качеством кодировать аудиоинформацию, при этом для кодирования одной музыкальной композиции требуется около четырёх мегабайт. Это даёт возможность интернет-пользователям быстро скачивать MP3-файлы. Видеоданные содержат видеоряд и наложенный звук, поэтому необходимо проводить не только дискретизацию, но и синхронизацию видеоизображения со звуком. На этом мы завершаем рассмотрение вопросов, связанных с кодированием при обработке на компьютере информации разных типов. Ключевыми здесь являются два положения:

- информация любого типа кодируется в числовом виде;
- для представления кода в устройствах компьютера используется двоичный код.

Позже в курсе будут рассмотрены вопросы представления чисел в разных системах счисления, которые были положены в основу описания функционирования компьютеров.

ХАРАКТЕРИСТИКА ОБРАБОТАННОЙ ИНФОРМАЦИИ

Своевременность

Точность

Полнота

Понятность

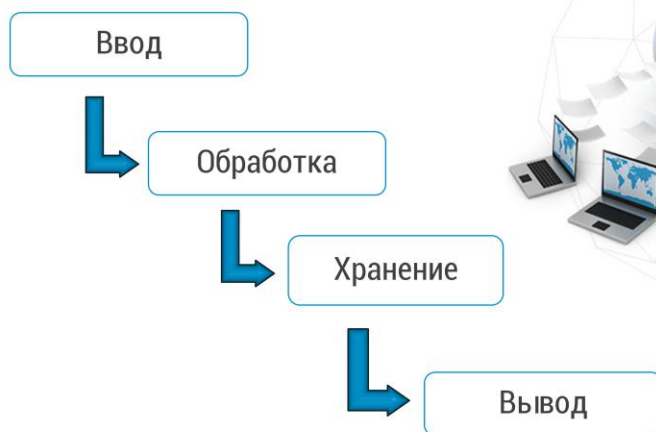


Рассмотрев кодирование информации и ее представление, делаем вывод, что информация – это организованные или классифицированные данные. Эти данные имеют некоторые значимые значения для получателя; это обработанные данные, на которых основываются решения и действия. Чтобы решение было значимым, обработанные данные должны соответствовать следующим характеристикам:

- своевременность – информация должна быть доступна при необходимости, задержка с получением информации делает ее бесполезной;
- точность – оказывает значительное влияние на принятие решений, возможность даже малейших ошибок должна быть сведена к минимуму;
- полнота – информация должна быть полной, неполная информация приводит к неверным и непредвиденным результатам;
- понятность – непонятная информация бесполезна для получателя, становится случаем информационного сбоя, поскольку бесполезная для получателя информация не считается «информацией».

Цикл обработки информации в контексте компьютеров и компьютерной обработки состоит из четырех этапов: ввод, обработка, вывод и хранение. Однако на некоторых уровнях компьютера некоторые устройства обработки фактически используют только три из этих стадий – ввод, обработку и вывод – без необходимости хранить данные. Каждый из этих этапов играет важную роль в действиях по сбору, анализу и распространению, выполняемых компьютерной системой.

ЦИКЛ ОБРАБОТКИ ИНФОРМАЦИИ



Дадим более подробное описание цикла обработки информации.

Цикл обработки информации – это последовательность событий, состоящая из ввода, обработки, хранения и вывода. Чтобы лучше понять, что такое цикл обработки информации, неплохо было бы также изучить цикл обработки данных. События обоих циклов аналогичны.

Чтобы компьютер мог выполнять полезную работу, он должен получать инструкции и данные из внешнего мира.

Компьютер получает данные и инструкции во время ВХОДА стадии цикла обработки информации. Полезные информационные результаты получаются, когда к данным применяются соответствующие входные данные. Применение инструкций к данным происходит во время ОБРАБОТКИ как этапа цикла обработки информации. Чтобы избежать повторного ввода данных и инструкций или повторной обработки информации, компьютеры могут сохранять информацию. Сохранение информации на компьютере происходит во время фазы ХРАНЕНИЯ цикла обработки информации. Затем следует результат на этапе ВЫВОД. Цикл компьютерной обработки – это аналогичный процесс с аналогичными шагами, с помощью которых данные передаются в компьютер.

Ввод. На этом этапе входные данные подготавливаются в некоторой удобной форме для обработки. Форма будет зависеть от обрабатывающей машины.

Например, когда используются электронные компьютеры, входные данные могут быть записаны на любом из нескольких типов носителей ввода, таких как магнитные диски, ленты и так далее.

Обработка. На этом этапе входные данные изменяются для создания данных в более удобной форме. Например, платежные чеки могут быть рассчитаны на основе табелей учета рабочего времени, или сводка продаж за месяц может быть рассчитана на основе заказов на продажу.

Вывод. На этом этапе собирается результат текущего этапа обработки.

Конкретная форма выходных данных зависит от использования данных.

Например, выходными данными могут быть зарплатные чеки для сотрудников.

ОБРАБОТКА ВВОДА

Ввод данных в компьютер характеризуется следующим:



Обработка ввода:

- подача необработанных данных
- ввод данных в цикл для обработки

Использование различных устройств:

- клавиатура, мышь,
- планшетные сканеры,
- планшеты с цифровыми данными,
- считыватели штрих-кодов,
- джойстики и другие



Рассмотрим первый этап – обработка ввода.

Данные должны поступить в систему, прежде чем они могут быть преобразованы в сохраненные данные или вывод информации. Стадия ввода предоставляет средства и механизмы, через которые данные попадают в модель ввода. Некоторые эксперты считают, что сам процесс ввода можно разделить на три этапа: сбор, подготовка и ввод. Однако общий взгляд на этап ввода состоит в том, что данные вводятся в систему с использованием некоторой формы устройства ввода.

Устройство ввода способно собирать данные в их источнике или точке измерения. Источник данных, вводимых человеком в систему, – это клавиатура, микрофон или, возможно, даже движение глаз или другой части тела. Другие формы устройств ввода, такие как термометры, датчики и часы, также соответствуют общему определению устройств ввода. Стадия ввода также может называться стадией кодирования.

Ввод данных в компьютер характеризуется следующим:

- обработка ввода включает подачу собранных необработанных данных или ввод данных в цикл для обработки. Это необработанные данные, которые используются для обработки и получения информации;
- для ввода можно использовать различные устройства, такие как клавиатуры, мыши, планшетные сканеры, считыватели штрих-кодов, джойстики, планшеты с цифровыми данными, электронные кассовые аппараты и т. д.

ОБРАБОТКА ДАННЫХ

Обработка данных



Ввод исходных данных →

Обработка необработанных данных
подходящим или выбранным
методом

Обработка →
центральный процессор компьютера



Следующий этап – обработка данных.

После того, как данные попадают в компьютер, они преобразуются в сохраненные данные или информацию. Агент обработки обычно представляет собой некоторую форму программного обеспечения или микропрограмм с определенным действием, предпринимаемым с определенным типом данных. В портативных или настольных компьютерах агент обработки обычно активен даже до того, как данные поступят. Фактически программное обеспечение обработки также часто запрашивает данные и направляет процесс ввода. Обработка может варьироваться от относительно небольшой и простой до очень большой и сложной. Тем не менее единственной целью этапа обработки является преобразование необработанных входных данных в некоторую форму. Данная форма может быть сохранена для последующего использования или обеспечения вывода информации для дальнейшей обработки или интерпретации.

Обработка данных – выполнение операций с данными – включает в себя следующее:

- после ввода исходных данных необработанные данные обрабатываются подходящим или выбранным методом обработки. Это наиболее важный шаг, поскольку он позволяет обрабатывать данные в виде вывода, который будет использоваться в дальнейшем;
- обработка обычно выполняется центральным процессором компьютера, рассматриваемым как ключевой компонент для выполнения операций.

Скорость обработки зависит от вычислительной мощности процессора.

ОБРАБОТКА ВЫВОДА

Вывод



Результат, при котором
необработанные данные →

- 1) «обрабатываются»
- 2) получаются полезные и предоставляющие информацию данные

Результат можно использовать:

- 1) как есть
- 2) для дальнейшей обработки вместе с другими данными



Обработка вывода отправляет информацию на экран дисплея, принтер, плоттер, динамик или какой-либо другой носитель, который может интерпретировать человеческие чувства. Однако этап вывода может хранить данные в новом формате или преобразовывать обработанные данные во входные данные для другого модуля компьютерной системы. Для большинства пользователей вывод означает либо отображение на экране монитора, либо распечатанный документ или графическое изображение. Вывод также может означать данные, информацию или кодирование.

Вывод как этап получения результатов, то есть новой информации, характеризуется следующими параметрами:

- это результат, на котором необработанные данные, предоставленные на первом этапе, теперь «обрабатываются», и данные становятся полезными и предоставляющими информацию. Они больше не называются данными. Это может быть использовано для визуализации данных;
- вывод можно использовать как есть или использовать для дальнейшей обработки вместе с другими данными. Вывод может быть дан на бумаге или на экрана дисплея монитора или телефона.

ХРАНЕНИЕ

Хранение



Хранение данных – это использование носителя записи для хранения данных с помощью компьютеров или других устройств

Хранение может осуществляться на:

- 1) внешний жесткий диск,
- 2) встроенный жесткий диск,
- 3) флеш-накопители,
- 4) карты micro SD,
- 5) компакт-диски



Последний этап – это хранение информации. Рассмотрим его.

Стадия хранения может происходить непосредственно из стадий обработки или вывода. Этап хранения может служить в качестве этапа псевдоввода или псевдовывода для этапа обработки. На этапе обработки может потребоваться сохранить данные для последующего использования, или нужно будет вызвать ранее сохраненные данные для обработки новых данных из этапа ввода. Выходной каскад может хранить обработанные данные как информацию для отображения другим модулем IPOС, когда это необходимо. Этап хранения не только сохраняет данные или информацию на фиксированном носителе данных, таком как жесткий диск, но также может хранить данные и информацию на съемных носителях.

Хранение, как этап сохранения данных в программном, физическом виде, характеризуется следующим.

Под хранением данных понимается использование носителя записи для хранения данных с помощью компьютеров или других устройств. Наиболее распространенными формами хранения данных являются хранение файлов, хранение блоков и хранение объектов. Каждая из этих форм используется для своей цели.

Хранение может осуществляться на различных устройствах хранения, таких как внешний жесткий диск, встроенный жесткий диск, флеш-накопители, карты micro SD, компакт-диски, или даже в регистрах.

Таким образом, обработка данных – важная часть обработки информации.

Использование обработанной информации в различных процессах принятия решений остается основной целью обработки информации. Обработка сложных данных, полученных из различных источников данных, требует сортировки и фильтрации данных. Ее также можно объединить с существующими источниками данных, чтобы перейти к получению нового набора данных. Дальнейший анализ этих наборов данных помогает в принятии решений и последующих решениях, относящихся к другим аспектам процесса принятия решений. Используемая информация может быть частью кратковременной памяти или долговременной памяти всего процесса.

ВЫВОДЫ

С самого начала информатика была предметом споров

Информатика основала вычисления на совокупности математических знаний

Информатика привела к разработке мощных инструментов

Информатика вызвала повышенное внимание к компьютеру

Все направления развития информатики могут сходиться



С самого начала информатика была предметом споров. Практики расходились во мнениях относительно того, могут ли и должны ли компьютеры и вычислительная техника быть предметом науки, и если да, то какой должна быть эта наука, какую роль в развитии информатики будут играть компьютеры. По мере развития предмета информатика основывала вычисления на совокупности глубоких и элегантных математических знаний, связывающих абстрактные машины с вычислительными процессами. Эта теория привела к разработке мощных инструментов для программирования и мониторинга работы компьютеров, таким образом передав власть компьютера в руки профессиональных и непрофессиональных пользователей.

Наряду с этим основным направлением развития в настоящее время является повышенное внимание к компьютеру. Он рассматривается не просто как инструмент, позволяющий автоматизировать труд. Его рассматривают как полноценного партнера, способного решать большой круг задач теории информатики, охватывая социальную и политическую теорию.

До сих пор не удалось создать жизнеспособную альтернативу информатике в ее нынешнем понимании. При этом последние тенденции в разработке программного обеспечения предполагают, что все направления развития информатики могут сходиться.

Посмотрим на историю развития информатики и на то, какую роль в этом процессе сыграли компьютеры.