

Информатика. Информация и информационные процессы





План

1. Разделы информатики
2. Основные понятия
 - ❖ Информатика
 - ❖ Информация
 - ❖ Алфавит, знак, слово
 - ❖ Данные, знания
 - ❖ Кодирование
3. Структуры данных



Информатика

Информатика не более наука о компьютерах, чем астрономия — наука о телескопах.

Эдсгер Дейкстра



Зарождение информатики

Информатика развивалась в составе:

- ❖ Математики;
- ❖ Электроники;
- ❖ Других технических наук.

Некоторые начала информатики можно обнаружить даже в лингвистике.

Как Вы думаете, какие именно?



Зарождение информатики

Отдельной наукой информатика была признана лишь в **70-х** годах XX века.

С этого момента информатика начала разрабатывать собственные методы и терминологию.



Разделы информатики

1

Математические основы

2

Теория вычислений

3

Алгоритмы и структуры данных

4

Языки программирования и трансляторы

5

Базы данных

6

Параллельные и распределенные системы



Математические основы

❖ Системы счисления

Целочисленные: двоичная, троичная, четверичная и т.д. Нецелочисленные: с основанием «e», с другими основаниями.

❖ Теория графов

Основы структур данных и алгоритмов поиска.

❖ Математическая логика

Булева логика и другие способы моделирования логических запросов.



Математические основы

◆ Криптография

Алгоритмы и методы защиты информации, как программные, так и аппаратные.

◆ Теория типов

Формальный анализ типов данных и использование этих типов для понимания свойств программ, в частности, их безопасности.



Теория вычислений

- ◆ **Теория автоматов**
Разные логические структуры для решения задач.
- ◆ **Теория вычислений**
Что можно вычислить, используя современные модели компьютеров.
- ◆ **Теория сложности вычислений**
Основные ограничения (в особенности время и размер для хранения данных) классов вычислений.



Алгоритмы и структуры данных

◆ Алгоритмы

Формальные логические процессы, используемые для вычислений и эффективность этих процессов.

◆ Структуры данных

Организация и правила управления данными.

◆ Генетические алгоритмы

Генетические алгоритмы – это способ поиска приблизительных решений задач поиска и оптимизаций.



Языки программирования и трансляторы

◆ Трансляторы

Способы трансляции компьютерных программ, как правило, из языков программирования высокого уровня в языки низкого уровня.

◆ Языки программирования

Формальные парадигмы языков для выражения алгоритмов и свойств этих языков (к примеру, на каком языке лучше решить данную задачу).



Базы данных

◆ Поиск данных

Изучение алгоритмов для поиска и обработки информации в документах и базах данных.



Параллельные и распределенные системы

❖ **Распределенные вычисления**

Вычисления, которые используют несколько компьютерных приборов через сеть для решения общей задачи.

❖ **Компьютерные сети**

Алгоритмы и протоколы для надежной передачи данных через разные разделенные или выделенные ресурсы.

❖ **Параллельные вычисления**

Вычисление, использующее несколько конкурентных потоков исполнения.



Информатика

Информатика – это техническая наука, систематизирующая приемы создания, хранения, воспроизведения, обработки и передачи данных средствами вычислительной техники, а также принципы функционирования этих средств и методы управления ими.



Какое понятие, по-вашему мнению, является центральным в информатике?



Понятие информации

«Информация» - лат. «informatio» - сведение, разъяснение, ознакомление.

Какие примеры информации Вы можете привести?



Информация в других науках

Математика:

Абстракция, абстрактная модель рассматриваемой системы

Биология:

Мера разнообразия в рассматриваемой системе

Теория

Вероятностей:

Вероятность выбора в рассматриваемой системе

Кибернетика:

Часть знаний, использующаяся в целях сохранения, совершенствования, развития системы

Информация

Техника:

Сообщения, передаваемые в форме знаков и сигналов

Теория информации:

Сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые уменьшают имеющуюся о них степень неопределённости (*энтропии*) и неполноты знаний



Понятие информации

Информация — это некоторая упорядоченная последовательность сообщений, отражающих, передающих и увеличивающих наши знания.



Методы получения информации

- ❖ Эмпирические методы или методы получения эмпирических данных.
- ❖ Теоретические методы или методы построения различных теорий.
- ❖ Эмпирико-теоретические методы (смешанные) или методы построения теорий на основе полученных эмпирических данных об объекте, процессе, явлении.



Эмпирические методы

Наблюдение

сбор первичной информации об объекте, процессе, явлении

Сравнение

Обнаружение и соотнесение общего и различного

Измерение

поиск с помощью измерительных приборов эмпирических фактов

Эксперимент

преобразование, рассмотрение объекта, процесса, явления с целью выявления каких-то новых свойств



Эмпирические методы

Кроме классических форм их реализации, в последнее время используются:

- ❖ Опрос
- ❖ Интервью
- ❖ Тестирование
- ❖ и другие



Теоретические методы

- ◆ **Восхождение от абстрактного к конкретному** – получение знаний о целом или о его частях на основе знаний об абстрактных проявлениях в сознании, в мышлении.
- ◆ **Идеализация** – получение знаний о целом или его частях путем представления в мышлении целого или частей, не существующих в действительности.



Теоретические методы

- ❖ **Формализация** – получение знаний о целом или его частях с помощью языков искусственного происхождения (формальное описание, представление).
- ❖ **Виртуализация** – получение знаний о целом или его частях с помощью искусственной среды, ситуации.



Теоретические методы

- ◆ **Аксиоматизация** – получение знаний о целом или его частях с помощью некоторых аксиом (не доказываемых в данной теории утверждений) и правил получения из них (и из ранее полученных утверждений) новых верных утверждений.



Эмпирико-теоретические методы

- ❖ **Абстрагирование** – выделение наиболее важных для исследования свойств, сторон исследуемого объекта, процесса, явления и игнорирование несущественных и второстепенных.
- ❖ **Анализ** – разъединение целого на части с целью выявления их связей.
- ❖ **Декомпозиция** – разъединение целого на части с сохранением их связей с окружением.
- ❖ **Синтез** – соединение частей в целое с целью выявления их взаимосвязей.



Эмпирико-теоретические методы

- ❖ **Композиция** — соединение частей целого с сохранением их взаимосвязей с окружением.
- ❖ **Индукция** – получение знания о целом по знаниям о частях.
- ❖ **Дедукция** – получение знания о частях по знаниям о целом.
- ❖ **Эвристики, использование эвристических процедур** – получение знания о целом по знаниям о частях и по наблюдениям, опыту, интуиции, предвидению.



Эмпирико-теоретические методы

- ❖ **Моделирование (простое моделирование), использование приборов** – получение знания о целом или о его частях с помощью модели или приборов.
- ❖ **Исторический метод** – поиск знаний с использованием предыстории, реально существовавшей или же мыслимой.
- ❖ **Логический метод** – поиск знаний путем воспроизведения частей, связей или элементов в мышлении.



Эмпирико-теоретические методы

- ❖ **Макетирование** – получение информации по макету, представлению частей в упрощенном, но целостном виде.
- ❖ **Актуализация** – получение информации с помощью перевода целого или его частей (а следовательно, и целого) из статического состояния в динамическое состояние.
- ❖ **Визуализация** – получение информации с помощью наглядного или визуального представления состояний объекта, процесса, явления.



Эмпирико-теоретические методы

А также:

- ❖ Мониторинг
- ❖ Деловые игры и ситуации
- ❖ Экспертные оценки
- ❖ Имитация
- ❖ и другие формы



Классификация информации

По отношению к источнику или приемнику:

Входная

Выходная

Внутренняя



Классификация информации

По отношению к конечному результату:

Исходная

Промежуточная

Результирующая



Классификация информации

По стадии использования:

Постоянная

Переменная

Смешанная



Классификация информации

По полноте:

Избыточная

Достаточная

Недостаточная



Классификация информации

По изменчивости:

Первичная

Вторичная



Классификация информации

По доступу:

Открытая

Закрытая



Основные свойства информации:

Полнота

Достаточность
для понимания
и принятия
решения

Понятность

Представление
на языке,
доступном
читателю

Достоверность

Отражение
истинного
положения дел

Актуальность

(своевременность)
существенность
для настоящего
времени

Объективность

отражение
внешнего
мира независимо от
нашего сознания



Формальное определение алфавита

Алфавит – конечное множество различных знаков, символов, для которых определена операция **конкатенации**



Формальное определение алфавита

Операция **конкатенации** — это приписывание, присоединение символа к символу или цепочке символов.

С ее помощью по определенным правилам соединения символов и слов можно получать слова (цепочки знаков) и словосочетания (цепочки слов) в этом алфавите (над этим алфавитом).



Примеры алфавитов

- ❖ Множество из десяти цифр
- ❖ Множество из знаков русского языка
- ❖ Множество из знаков латинского языка
- ❖ Точка и тире в азбуке Морзе
- ❖ и др.

В алфавите цифр знак 5 связан с понятием «быть в количестве пяти элементов».



Определение буквы - знака

Буквой или **знаком** называется любой элемент x алфавита X , где $x \in X$.

Понятие знака неразрывно связано с тем, что им обозначается («со смыслом»), они вместе могут рассматриваться как пара элементов (x, y) , где x – сам знак, а y – обозначаемое этим знаком.



Понятие слова

Конечная последовательность букв алфавита называется **словом** в алфавите (или над алфавитом).



Понятие слова

- ❖ Длиной $|p|$ некоторого слова p над алфавитом X называется число составляющих его букв.
- ❖ Слово (обозначаемое символом \emptyset) имеющее нулевую длину, называется **ПУСТЫМ СЛОВОМ**: $|\emptyset| = 0$.



Понятие слова

- ❖ Множество различных слов над алфавитом X обозначим через $S(X)$ и назовем **словарным запасом (словарем) алфавита (над алфавитом) X** .
- ❖ В отличие от конечного алфавита, словарный запас может быть и бесконечным. Объясните, почему?
- ❖ Слова над некоторым заданным алфавитом определяют **сообщения**



Примеры:

❖ Слова над алфавитом кириллицы:

- “Информатика”
- “инто”
- “ииии”
- “и”



Примеры:

❖ Слова над алфавитом десятичных цифр и знаков арифметических операций:

- "1256"
- "23+78"
- "35-6+89"
- "4"

Примеры:

❖ Слова над алфавитом азбуки Морзе:

■ ". "

■ ". . _"

■ "_ _ _"



Алфавит

В алфавите должен быть определен порядок следования букв (порядок типа "предыдущий элемент – последующий элемент")

Любой алфавит имеет упорядоченный вид $X = \{x_1, x_2, \dots, x_n\}$.



Понятие данных

Данные – факт, понятие или инструкции, представленные в условной форме, удобной для пересылки, интерпретации и обработки человеком или автоматизированными средствами



Понятие знаний

Знания – совокупность сведений (данных или программ), отражающих знания человека – специалиста (эксперта) в определенной предметной области и предназначенных для хранения в базах знаний.

Условно можно записать, что

"знания = факты + убеждения + правила"



Кодирование данных

Для автоматизации работы с данными, относящимися к различным типам, очень важно унифицировать их форму представления – для этого используется прием **кодирования**, то есть выражение данных одного типа через данные другого типа.



Понятие кодирования

Кодирование – это преобразование сообщения в код, то есть в совокупность символов, отображающих сообщение, передаваемое по каналу связи.



Кодирование данных ДВОИЧНЫМ КОДОМ

Двоичное кодирование, используемое в вычислительной технике, основано на представлении данных последовательностью двух знаков: 0 и 1.

Эти знаки называются двоичными цифрами, по английски – **binary digit** или, сокращенно, **bit** (бит).



Кодирование данных ДВОИЧНЫМ КОДОМ

Одним битом могут быть выражены два понятия: **0** или **1** (да или нет, черное или белое, истина или ложь и т.п.).

Если количество битов увеличивается до двух, то можно выразить четыре различных понятия:

00 01 10 11

Тремя битами можно закодировать восемь различных значений:

000 001 010 011 100 101 110 111



Кодирование данных ДВОИЧНЫМ КОДОМ

Количество разрядов в системе
двоичного кодирования **+1**



Количество значений, которое может
быть выражено в данной системе ***2**



Кодирование данных ДВОИЧНЫМ КОДОМ

Общая формула имеет вид:

$$N=2^i$$

N — количество независимых кодируемых значений;

i — разрядность двоичного кодирования, принятая в данной системе.



Кодирование целых чисел

Принцип кодирования целых чисел двоичным кодом: деление целого числа пополам, пока в остатке не образуется ноль или единица. Совокупность остатков от каждого деления, записанная справа налево вместе с последним остатком, и образует двоичный аналог десятичного числа.



Пример:

$$19 : 2 = 9 + 1$$

$$9 : 2 = 4 + 1$$

$$4 : 2 = 2 + 0$$

$$2 : 2 = 0 + 1$$

Таким образом, $19_{10} = 10011_2$



Кодирование целых чисел

Для кодирования целых чисел от 0 до 255 достаточно иметь **8 разрядов** двоичного кода (8 бит).

16 бит позволяют закодировать целые числа от 0 до 65535

24 бита – уже более 16,5 миллионов различных значений.



Кодирование текстовых данных

Если каждому символу алфавита сопоставить определенное целое число (например, порядковый номер), то с помощью двоичного кода можно кодировать и текстовую информацию.

Восьми двоичных разрядов достаточно для кодирования **256** различных СИМВОЛОВ.



Кодирование текстовых данных

Институт стандартизации США (**ANSI** – **American National Standard Institute**) ввел в действие систему кодирования *ASCII* (**American Standard Code for Informational Interchange** – американский стандартный код для обмена информацией).



Кодирование текстовых данных

В системе ASCII закреплены две таблицы кодирования – базовая (символы с номерами 0 - 127) и расширенная (128 - 255).

	!	"	#	\$	%	&	'	()	*	+	,	-	.	/
32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95
'	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111
p	q	r	s	t	u	v	w	x	y	z	{		}	~	□
112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127



Расширенная таблица кодировки ASCII

	☺	☹	♥	♦	♣	♠	●		○						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
▶	◀	!				_		↑	↓	→	←	↔	▲	▼	
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
	!	"	#	\$	%	&	'	()	*	+	,	-	.	/
32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95
'	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111
p	q	r	s	t	u	v	w	x	y	z	{		}	~	□
112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127
Ç	ü	é	â	ã	à	ä	ç	ê	ë	è	ï	î	í	Ë	Ä
128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143
Ë	æ	Æ	ó	ö	ò	ú	ù	ÿ	Ö	Ü	€	£	¥	ℳ	ƒ
144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
	í	ó	ú	ñ	Ñ	ª	º	¿	¬	½	¼	¡	«	»	
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
☐	☐	☐													
176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191
┌	└	┐	┌	┐	┌	┐	┌	┐	┌	┐	┌	┐	┌	┐	┌
192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207
┌	┐	┌	┐	┌	┐	┌	┐	┌	┐	┌	┐	┌	┐	┌	┐
208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223
α	β	Γ	κ	Σ	σ	μ	τ	φ	θ	Ω	δ	∞	φ	ε	η
224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
≡	±	≥	≤			÷	≈	°	·	·	√	π	²	■	□
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255



Кодировка Windows 1251

			©	Ё	§	Є	.		°						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
	!	"	#	\$	%	&		()	^	+				/
32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47
0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63
@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95
'	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111
p	q	r	s	t	u	v	w	x	y	z	{	 	}	~	□
112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127
Ъ	Г	,	Г	„	…	†	‡		%	Ь	«	Ъ	К	Ъ	Ц
128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143
ђ	'	'	"	"		—	—		™	Ь	»	Ъ	К	Ъ	Ц
144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
	Ў	ў	Ј	□	Г	 	§	Ё	©	Є	«		-	@	І
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
°	±	І	і	Г	μ		·	ё	№	є	»	ј	Ѕ	ѕ	ї
176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191
А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207
Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	• 223
а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п
224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255



Кодирование текстовых данных

Использование множества одновременно действующих кодировок объясняется ограниченным набором кодов (256).

Система, основанная на 16-разрядном кодировании символов, получила название **универсальной** – **UNICODE**.

Позволяет обеспечить уникальные коды для 65536 различных символов.

Долгое время переход на данную систему сдерживался недостаточными ресурсами вычислительной техники.



Кодирование графических данных

Графическое изображение состоит из мельчайших точек (пикселей), образующих характерный узор, называемый **растром**.

Пример:



Часть изображения
при увеличении в 7 раз



Кодирование графических данных

Точки растрового изображения, имеющие только два возможных цвета черный и белый, можно закодировать двумя цифрами – *0* или *1*.

Для черно-белых иллюстраций общепринятым является кодирование точки с 256 градациями серого цвета.



Кодирование графических данных

Цветное изображение на экране
получается путем смешивания трех
базовых цветов:

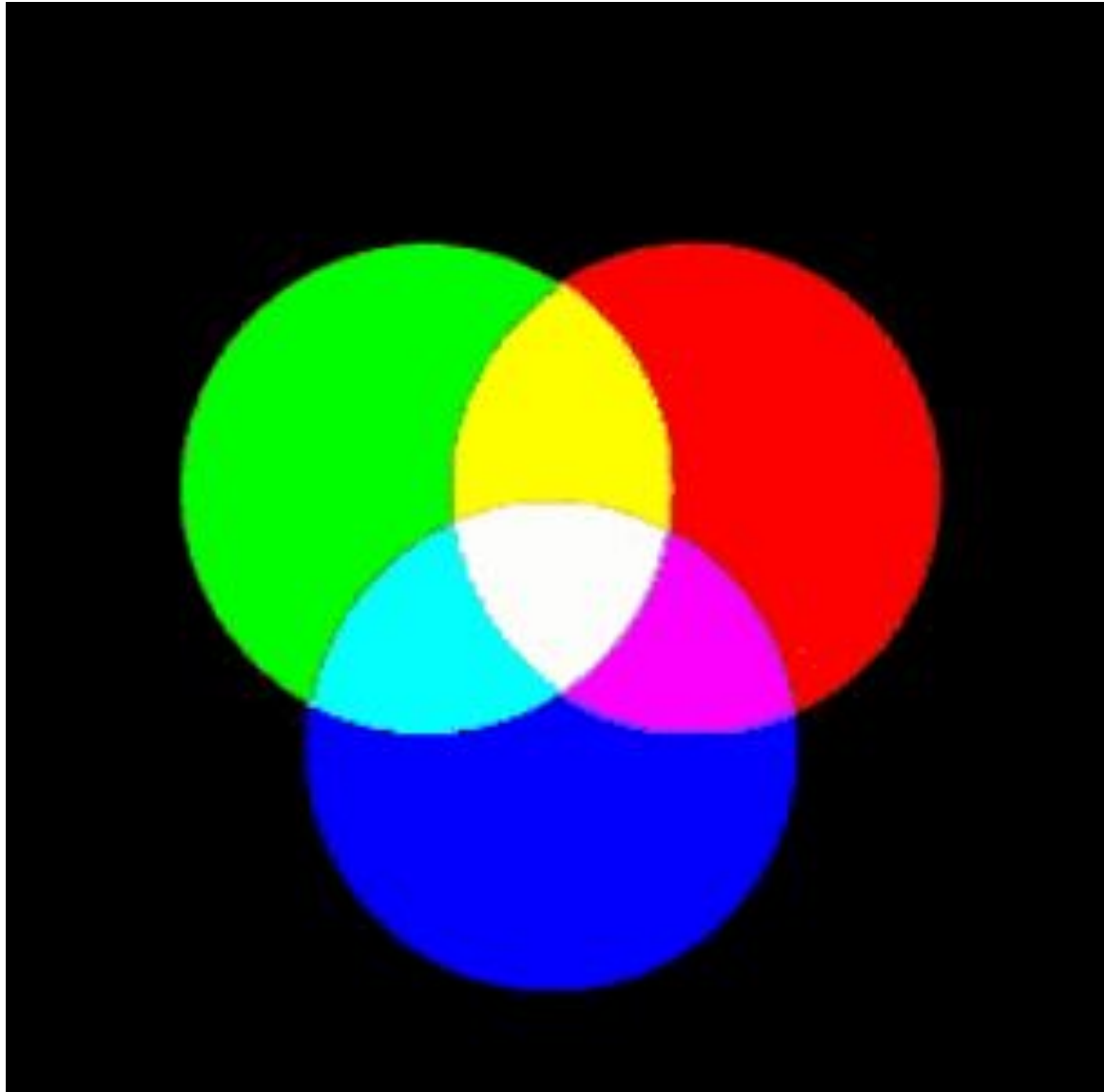
красного (Red)

зеленого (Green)

синего (Blue)



Кодирование графических данных





Кодирование графических данных

Для кодирования цветных графических изображений применяется принцип декомпозиции произвольного цвета на базовые составляющие. Такая система кодирования называется **RGB**.

Если для кодирования яркости каждой из основных составляющих использовать по 256 значений, то система RGB обеспечивает однозначное определение **16777216** различных цветов.



Основные структуры данных

Проблема, возникающая при работе с большими наборами данных – это поиск нужных элементов.

Работа с большими наборами данных автоматизируется проще, когда данные упорядочены, то есть образуют заданную структуру.



Основные структуры данных

- ◆ **Линейная**
- ◆ **Табличная**
- ◆ **Иерархическая**



Линейные структуры данных

Форма представления для пользователя:

№	ФИО
1	Андреев Борис Борисович
2	Борисов Андрей Андреевич
3	Володин Алексей Алексеевич
...	...
27	Санин Николай Николаевич



Линейные структуры данных

Форма хранения:

Андреев Борис Борисович * Борисов Андрей
Андреевич * Володин Алексей Алексеевич *
... * Санин Николай Николаевич



Линейные структуры данных

Линейные структуры данных (списки) – это упорядоченные структуры данных, в которых адрес элемента однозначно определяется его номером.



Табличные структуры данных

Планета	Расстояние до Солнца, у.е.	Относительная масса	Количество спутников
Меркурий	0,39	0,056	0
Венера	0,67	0,88	0
Земля	1,0	1,0	1
Марс	1,51	0,1	2
Юпитер	5,2	318	16

Меркурий*0,39*0,056*0 # Венера*0,67*0,88*0
Земля*1,0*1,0*1 #



Многомерные таблицы

Номер факультета: 3

Номер курса (на факультете): 2

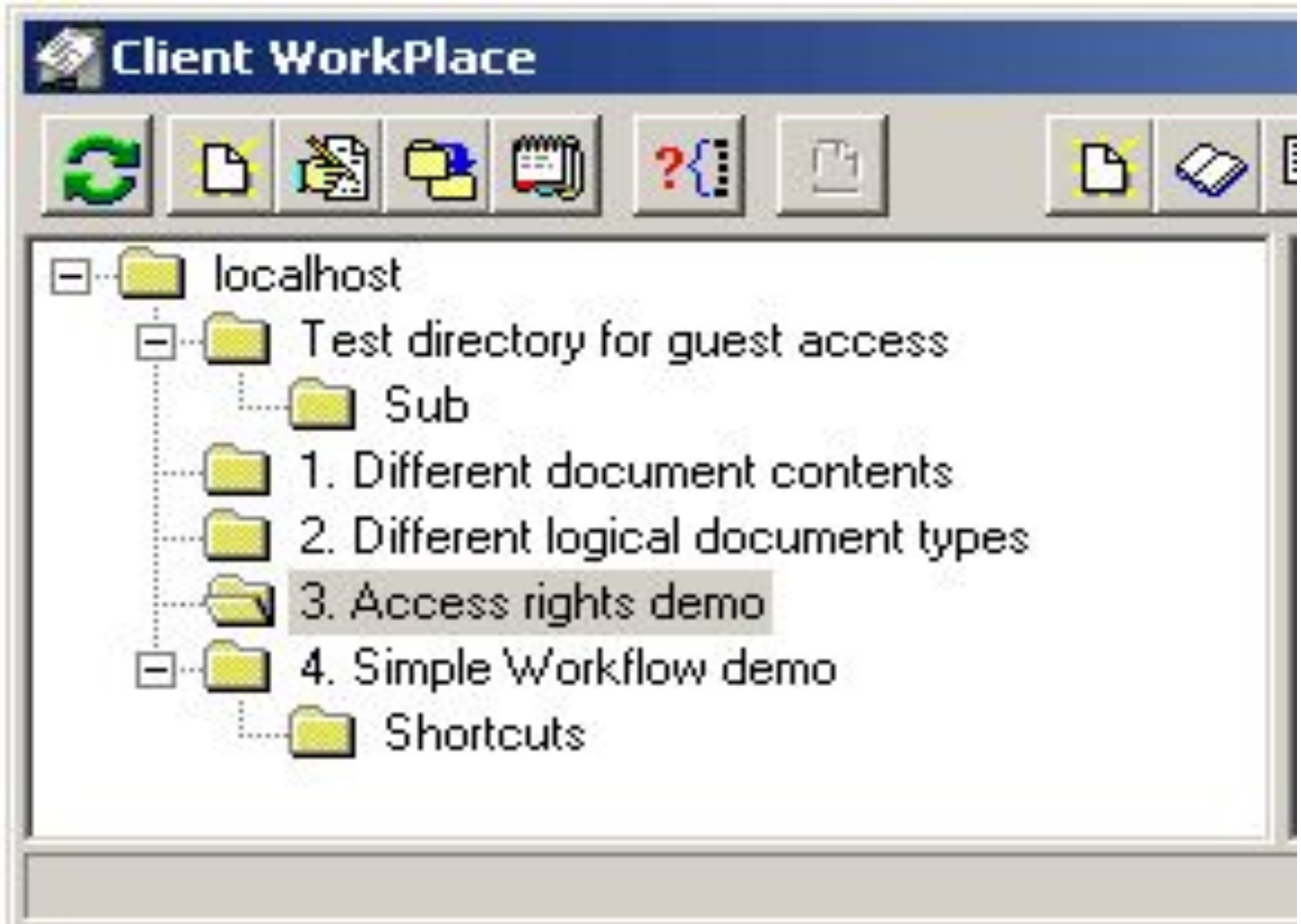
Номер специальности (на курсе): 2

Номер группы в потоке одной специальности: 1

Номер учащегося в группе: 19

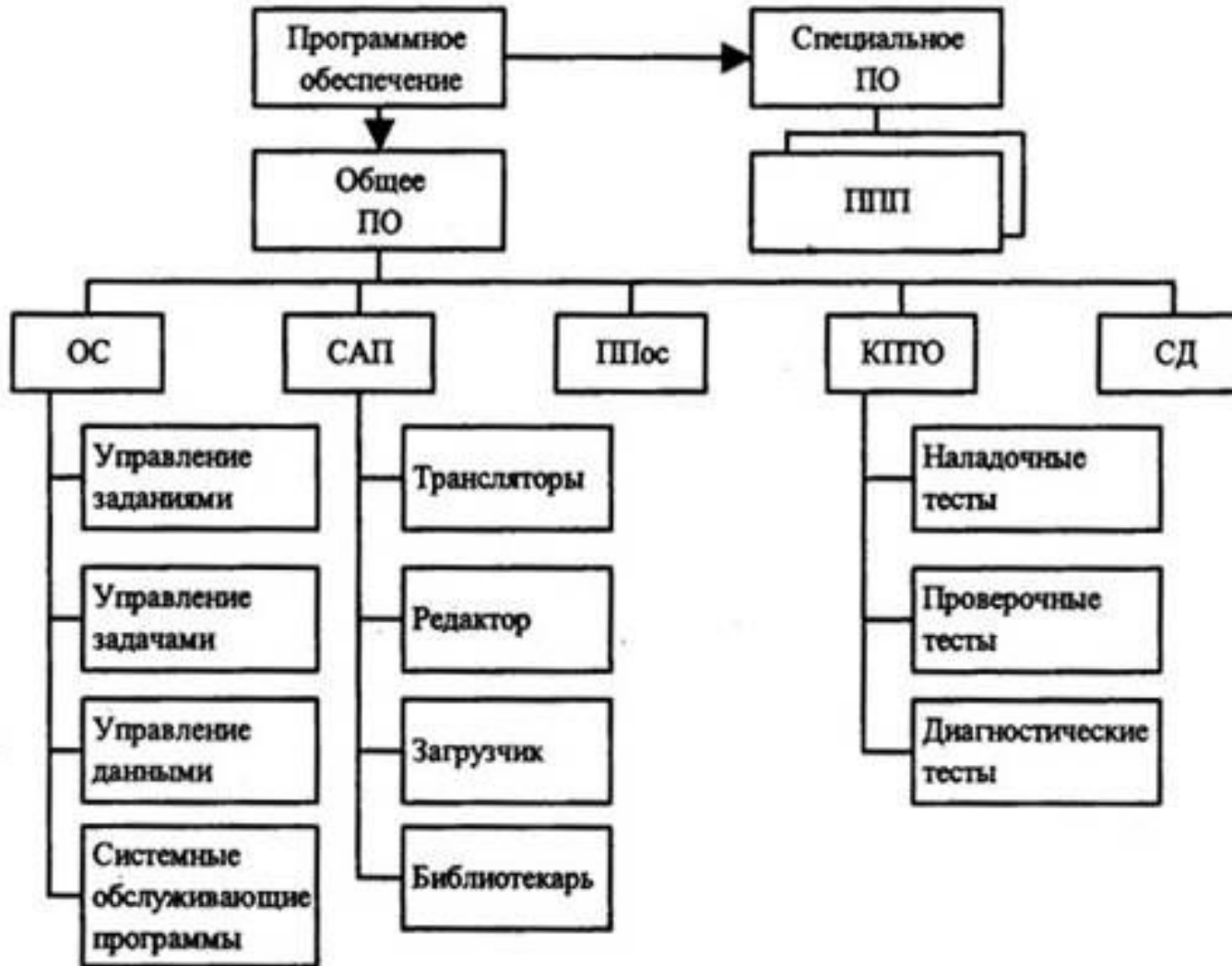


Иерархические структуры данных





Иерархические структуры данных





Иерархические структуры данных

В иерархической структуре данных адрес каждого элемента определяется путем доступа (маршрутом), ведущим от вершины структуры к данному элементу.

Недостаток иерархических структур данных – трудоемкость записи адреса элемента данных и сложность упорядочения.



Вопросы для самостоятельного изучения

- ❖ Структура предметной области информатики (национальный доклад Российской Федерации на II международном конгрессе ЮНЕСКО)
- ❖ Отличия чувашского и татарского варианта кодировки Windows 1251 от стандартного
- ❖ Отличие многомерных структур данных от иерархических
- ❖ Другие цветовые модели синтеза цвета кроме RGB