

ТЕОРЕТИЧЕСКИЙ МАТЕРИАЛ

1 Основные понятия информатики

Информация – сведения о чем-либо.

История человечества – это история того, как человек стремился получить информацию об окружающем его мире, сохранить, использовать, передать ее следующим поколениям. Наскальная живопись, клинопись, устная речь, музыкальные звуки, нотные знаки для их записи, алфавит, телеграф, радио, телефон, телевидение, компьютеры – вот лишь некоторые звенья цепи попыток совершенствовать способы получения, сохранения, обработки и передачи информации.

Применительно к компьютерной обработке данных под информацией понимают некоторую последовательность символьических обозначений (букв, цифр, закодированных графических образов и звуков и т.п.), несущую смысловую нагрузку и представленному в понятном компьютеру виду.

Информационные процессы – действия, выполняемые с информацией: получение, хранение, обработка, передача.

Компьютер открывает перед человеком новые возможности и перспективы развития этих процессов и коммуникации между людьми. Информация – основное понятие информатики.

Информационные технологии – технологии накопления, обработки и передачи информации с использованием определенных технических средств.

Информационные ресурсы – это идеи человечества и указания по их реализации, накопленные в форме, позволяющей воспроизводить их (книги, статьи, патенты, диссертации, научно-исследовательская и опытно-конструкторская документация, технические переводы, данные о передовом производственном опыте и др.).

Информатика – наука, изучающая методы сбора, хранения, обработки и передачи информации с помощью средств вычислительной техники.

Термин «информатика» происходит от французских слов *information* (информация) и *automatique* (автоматика).

2 Основные свойства информации

Среди основных свойств информации можно выделить следующие:

Объективность, полнота, достоверность, адекватность, доступность и актуальность информации.

Объективность и субъективность информации. Понятие объективности информации является относительным. Это понятно, если учесть, что методы являются субъективными. Более объективной принято считать ту информацию, в которую методы вносят меньший субъективный элемент. Так, например, принято считать, что в результате наблюдения фотоснимка природного объекта или явления образуется более объективная информация, чем в результате наблюдения рисунка того же объекта, выполненного человеком. В ходе информационного процесса степень объективности информации всегда понижается. Это свойство учитывают, например, в правовых дисциплинах, где по-разному обрабатываются показания лиц, непосредственно наблюдавших события или получивших информацию косвенным путем (посредством умозаключений или со слов третьих лиц).

Полнота информации. Полнота информации во многом характеризует качество информации и определяет достаточность данных для принятия решений или для создания новых данных на основе имеющихся.

Чем полнее данные, тем шире диапазон методов, которые можно использовать, тем проще подобрать метод, вносящий минимум погрешностей в ход информационного процесса.

Достоверность информации. Данные возникают в момент регистрации сигналов, но не все сигналы являются «полезными» - всегда присутствует какой-то уровень посторонних сигналов, в результате чего

полезные данные сопровождаются определенным уровнем «информационного шума».

Если полезный сигнал зарегистрирован более четко, чем посторонние сигналы, достоверность информации может быть более высокой. При увеличении уровня шумов достоверность информации снижается. В этом случае для передачи того же количества информации требуется использовать либо больше данных, либо более сложные методы.

Адекватность информации – это степень соответствия реальному объективному состоянию дела.

Неадекватная информация может образовываться при создании информации на основе неполных или недостоверных данных. Однако и полные, и достоверные данные могут приводить к созданию неадекватной информации в случае применения к ним неадекватных методов.

Доступность информации – мера возможности получить ту или иную информацию.

На степень доступности информации влияют одновременно как доступность данных, так и доступность адекватных методов для их интерпретации. Отсутствие доступа к данным или отсутствие адекватности методов обработки данных приводят к одинаковому результату: информация оказывается недоступной. Отсутствие адекватных методов для работы с данными во многих случаях приводит к применению неадекватных методов, в результате чего образуется неполная, неадекватная или недостоверная информация.

Актуальность информации – это степень соответствия информации текущему моменту времени.

Нередко с актуальностью, как и с полнотой, связывают коммерческую ценность информации. Поскольку информационные процессы растянуты во времени, то достоверная и адекватная, но устаревшая информация может приводить к ошибочным решениям. Необходимость поиска (или

разработки) адекватного метода для работы с данными может приводить к такой задержке в получении информации, что она становится неактуальной и ненужной. На этом, в частности, основаны многие современные системы шифрования данных с открытым ключом. Лица, не владеющие ключом (методом) для чтения данных, могут заняться поиском ключа, поскольку алгоритм его работы доступен, но продолжительность этого поиска столь велика, что за это время работы информация теряет актуальность и, соответственно, связанную с ней практическую ценность.

Будучи объектом преобразования и использования, информация характеризуется следующими свойствами:

Синтаксис – свойство, определяющее способ представления информации на носителе (в сигнале).

Так, данная информация представлена на электронном носителе с помощью определенного шрифта. Здесь же можно рассматривать такие параметры представления информации, как стиль и цвет шрифта, его размеры, междустрочный интервал и т.д. Выделение нужных параметров как синтаксических свойств, очевидно, определяется предполагаемым способом преобразования. Например, для плохо видящего человека существенным является размер и цвет шрифта. Если предполагается вводить данный текст в компьютер через сканер, важен формат бумаги;

Семантика – свойство, определяющее смысл информации как соответствие сигнала реальному миру.

Так, семантика сигнала “информатика” заключается в данном ранее определении. Семантика может рассматриваться как некоторое соглашение, известное потребителю информации, о том, что означает каждый сигнал (так называемое правило интерпретации). Например, именно семантику сигналов изучает начинающий автомобилист, штудирующий правила дорожного движения, познавая дорожные знаки (в этом случае сигналами выступают сами знаки). Семантику слов (сигналов) познаёт обучаемый

какому-либо иностранному языку. Можно сказать, что смысл обучения информатике заключается в изучении семантики различных сигналов – суть ключевых понятий этой дисциплины;

Прагматика – свойство, определяющее влияние информации на поведение потребителя.

Так прагматика информации, получаемой читателем настоящего учебного пособия, заключается, по меньшей мере, в успешной сдаче экзамена по информатике. Хочется верить, что этим прагматика данного труда не ограничится, и он послужит для дальнейшего обучения и профессиональной деятельности читателя.

3 Кодирование информации

Кодирование информации – это процесс представления информации в виде, удобном для её хранения, передачи и обработки. Чаще всего мы имеем дело с преобразованием информации в последовательность символов. Для того чтобы это преобразование было однозначным, то есть чтобы любой мог его понять и восстановить исходную информацию, используются специальные системы правил – коды. Одним из самых фундаментальных и исторически важных кодов является код ASCII, что расшифровывается как American Standard Code for Information Interchange (Американский стандартный код для обмена информацией). Этот код был создан для того, чтобы разные компьютеры и устройства могли понимать друг друга. ASCII представляет собой таблицу, в которой каждому символу (букве, цифре, знаку препинания) сопоставлен уникальный числовой код. Изначально таблица ASCII использовала 7 бит, что позволяло закодировать 128 символов. Эти символы делятся на две группы: управляющие символы (непечатаемые, например, символ перевода строки или звукового сигнала) и печатаемые символы (те, что мы видим на экране или на бумаге). Например, заглавной латинской букве 'A' соответствует десятичный код 65, строчной 'a' – код 97, цифре '0' – код 48, а знаку пробела – код 32. Когда вы

нажимаете на клавиатуре клавишу 'A', компьютер не хранит её изображение, а сохраняет в память число 65. Именно этот числовой код и является закодированной информацией.

Рассмотрим пример кодирования слова "САТ" (кот на английском) с помощью таблицы ASCII. Чтобы закодировать это слово, мы должны заменить каждую букву на её числовой код. В таблице ASCII символ 'С' имеет код 67, символ 'A' – код 65, символ 'T' – код 84. Таким образом, процесс кодирования будет выглядеть так: исходное слово "САТ" превращается в последовательность чисел: 67, 65, 84. Именно в таком виде компьютер и будет хранить и обрабатывать это слово. Теперь рассмотрим процесс раскодирования. Допустим, мы получили от компьютера последовательность чисел: 72, 101, 108, 108, 111. Наша задача – превратить её обратно в читаемый текст. Мы берем таблицу ASCII и для каждого числа ищем соответствующий символ. Число 72 соответствует заглавной латинской букве 'Н', число 101 – строчной 'е', число 108 – строчной 'т', и снова 108 – 'т', и, наконец, число 111 – строчной 'о'. Соединив все эти символы, мы получаем слово "Hello". Важно понимать, что кодирование и раскодирование возможны только тогда, когда и отправитель, и получатель используют одну и ту же кодовую таблицу.

Хотя ASCII была очень популярна, она имела серьёзное ограничение – в ней не было русских букв и символов других языков. Это привело к созданию расширенных кодировок, таких как KOI-8 и Windows-1251, которые использовали уже 8 бит (256 символов). А сегодня самым распространенным в мире стандартом является Unicode (например, кодировка UTF-8), который использует несколько байт для кодирования одного символа и может представить практически все письменные языки мира, а также множество специальных символов. Таким образом, кодирование — это краеугольный камень информатики, позволяющий

преобразовать любую информацию (текст, звук, изображение) в универсальный цифровой вид, понятный современным компьютерам.

На рисунке 1 представлена таблица ASCII.

Dec	Hex	Char	Cmd	Dec	Hex	Char	Cmd	Dec	Hex	Char	Cmd	Dec	Hex	Char	Cmd
0	00		NUL	32	20		(sp)	64	40	@		96	60	`	
1	01	⌚	SOH	33	21	!		65	41	A		97	61	a	
2	02	❶	STX	34	22	"		66	42	B		98	62	b	
3	03	♥	ETX	35	23	#		67	43	C		99	63	c	
4	04	◆	EOT	36	24	\$		68	44	D		100	64	d	
5	05	♣	ENQ	37	25	%		69	45	E		101	65	e	
6	06	♦	ACK	38	26	&		70	46	F		102	66	f	
7	07	•	BEL	39	27	'		71	47	G		103	67	g	
8	08	▣	BS	40	28	(72	48	H		104	68	h	
9	09	○	TAB	41	29)		73	49	I		105	69	i	
10	0A	▣	LF	42	2A	*		74	4A	J		106	6A	j	
11	0B	♂	VT	43	2B	+		75	4B	K		107	6B	k	
12	0C	♀	FF	44	2C	,		76	4C	L		108	6C	l	
13	0D	♪	CR	45	2D	-		77	4D	M		109	6D	m	
14	0E	♪	SO	46	2E	.		78	4E	N		110	6E	n	
15	0F	☼	SI	47	2F	/		79	4F	O		111	6F	o	
16	10	▶	DLE	48	30	0		80	50	P		112	70	p	
17	11	◀	DC1	49	31	1		81	51	Q		113	71	q	
18	12	↑	DC2	50	32	2		82	52	R		114	72	r	
19	13	!!	DC3	51	33	3		83	53	S		115	73	s	
20	14	¶	DC4	52	34	4		84	54	T		116	74	t	
21	15	§	NAK	53	35	5		85	55	U		117	75	u	
22	16	—	SYN	54	36	6		86	56	V		118	76	v	
23	17	↑	ETB	55	37	7		87	57	W		119	77	w	
24	18	↑	CAN	56	38	8		88	58	X		120	78	x	
25	19	↓	EM	57	39	9		89	59	Y		121	79	y	
26	1A	→	SUB	58	3A	:		90	5A	Z		122	7A	z	
27	1B	←	ESC	59	3B	;		91	5B	[123	7B	{	
28	1C	└	FS	60	3C	<		92	5C	\		124	7C		
29	1D	↔	GS	61	3D	=		93	5D]		125	7D	}	
30	1E	▲	RS	62	3E	>		94	5E	^		126	7E	~	
31	1F	▼	US	63	3F	?		95	5F	_		127	7F	△	DEL

Рисунок 1 – Таблица ASCII

4 Мощность алфавита и информационный объём

Понятие мощность алфавита это просто общее количество символов в нем. Символами считаются все буквы заглавные и строчные цифры знаки препинания скобки пробел и другие специальные знаки, которые можно использовать в тексте. Например, мощность алфавита для русского языка обычно принимают равной 54 символа это 33 буквы 10 цифр и 11 знаков препинания и других символов. А мощность алфавита для английского языка обычно 26 букв. В информатике, когда мы говорим о кодировании компьютеру не важно какие это буквы русские или английские ему важно общее количество разных символов, которые нужно различать. Чем больше

мощность алфавита, тем больше информации несет один символ этого алфавита.

Теперь перейдем к информационному весу символа и объему текста. Компьютер различает символы по их уникальным двоичным кодам. Чтобы закодировать один символ из алфавита мощностью N нужно определенное количество бит. Бит – это самая маленькая единица информации. Это один разряд в двоичном коде, который может быть равен нулю или единице. Существует простая формула, которая связывает мощность алфавита N и информационный вес одного символа i который измеряется в битах. Эта формула выглядит так N равно два в степени i . То есть мощность алфавита это количество различных символов, которые можно закодировать с помощью i бит. Давай рассмотрим пример. Возьмем алфавит из четырех символов, например, А Б Г. Сколько бит нужно чтобы закодировать каждый символ. Подставляем в формулу 4 равно два в степени i . Два в какой степени равно четырем. Это два во второй степени. Значит i равно 2 бита. И правда двумя битами можно составить четыре разные комбинации 00 01 10 11 и каждой комбинации мы можем сопоставить свой символ.

Теперь более сложный и важный пример алфавит, который используется в компьютере для представления текста. Современные компьютеры обычно используют алфавит мощностью 256 символов. Это все английские и русские буквы цифры знаки препинания и много специальных символов. Найдем сколько бит весит один символ такого алфавита. Воспользуемся формулой 256 равно два в степени i . Два в восьмой степени как раз равно 256. Значит i равно 8 бит. Вот откуда берется известная единица измерения информации байт. Один байт – это как раз 8 бит и этого достаточно чтобы закодировать один символ из стандартного компьютерного алфавита.

Информационный объем всего текста это просто сумма информационных весов всех символов в этом тексте. То есть чтобы найти

информационный объем текста нужно умножить количество символов в тексте K на информационный вес одного символа i . I равно K умножить на i . Давай посчитаем информационный объем слова КОТ в кодировке, где один символ весит 2 байта или 16 бит. В слове три буквы. Значит объем информации I равен 3 символа умножить на 16 бит равно 48 бит. Или если считать в байтах, то 3 символа умножить на 2 байта равно 6 байт.

Рассмотрим еще один пример с другой мощностью алфавита. Допустим у нас есть алфавит из 16 символов, например, для системы счисления от 0 до 9 и букв A до F как в шестнадцатеричной системе. Найдем вес одного символа. 16 равно два в степени i . Два в четвертой степени равно 16 значит i равно 4 бита. Теперь закодируем слово CAFE. Оно состоит из 4 символов. Информационный объем этого слова будет равен 4 символа умножить на 4 бита равно 16 бит.

Последний пример для закрепления. Сообщение, записанное 32 символьным алфавитом, содержит 80 символов. Сколько бит информации оно несет. Сначала найдем вес одного символа. 32 это два в пятой степени значит один символ весит 5 бит. Теперь умножаем 80 символов на 5 бит получаем 400 бит. Такой информационный объем у этого сообщения.

ПЕРЕЧЕНЬ ЗАДАНИЙ

1. Дайте определение понятию «информация» и приведите примеры её передачи в разные исторические эпохи.
2. Перечислите основные свойства информации и объясните каждое из них на примере.
3. Объясните различие между синтаксисом, семантикой и прагматикой информации.
4. Что такое информационные ресурсы? Приведите примеры из разных областей науки и техники.
5. Раскройте понятие «информационные технологии» и объясните их роль в современном обществе.
6. Закодируйте слово DOG с помощью таблицы ASCII.
7. Раскодируйте последовательность чисел: 72, 101, 108, 108, 111.
8. Преобразуйте слово КОТ в двоичный код, если используется кодировка UTF-8 (каждый символ занимает 2 байта).
9. Закодируйте слово INFO в ASCII и запишите числовую последовательность.
10. Раскодируйте последовательность чисел: 67, 65, 84 и укажите полученное слово.
11. Найдите информационный объём слова КОТ, если используется алфавит мощностью 32 символа (каждый символ весит 5 бит).
12. Рассчитайте информационный объём текста из 50 символов, если мощность алфавита равна 64.
13. Определите, сколько бит необходимо для кодирования алфавита мощностью 128 символов.
14. Сообщение содержит 100 символов, закодированных алфавитом мощностью 256. Найдите информационный объём сообщения.
15. Рассчитайте информационный объём слова CAFE, если используется шестнадцатеричный алфавит (16 символов).

ПРИМЕР ОФОРМЛЕНИЯ РАБОТЫ

Данная работа должна быть выполнена на листках А4.

Далее представлен пример оформления работы.

Комплексная проверочная работа

ФИО, дата выполнения

Задание 1.

Решение.

Задание 2.

Решение.

...