

# Измерение и представление информации

## Оглавление

Краткие теоретические сведения .....	3
Измерение информации .....	3
Вероятностный подход .....	3
Алфавитный подход .....	6
Единицы измерения информации .....	6
Представление информации в компьютере .....	7
Кодирование текстовой информации. Кодовые таблицы символов .....	8
Кодирование графической информации .....	10
Кодирование звуковой информации .....	13
Примеры решения заданий .....	15
Измерение информации .....	15
Пример 1 задания с кратким ответом .....	15
Пример 2 задания с кратким ответом .....	15
Пример 3 задания с выбором одного ответа .....	16
Пример 4 задания с кратким ответом .....	16
Пример 5 задания с кратким ответом .....	16
Пример 6 задания с выбором одного ответа .....	17
Пример 7 задания с выбором одного ответа .....	17
Пример 8 задания с кратким ответом .....	17
Пример 9 задания с кратким ответом .....	17
Пример 10 задания с выбором одного ответа .....	18
Пример 11 задания с кратким ответом .....	18
Пример 12 задания с кратким ответом .....	18
Пример 13 задания с кратким ответом .....	19
Пример 14 задания с выбором одного ответа .....	19

Представление информации в компьютере .....	20
Пример 15 задания с выбором одного ответа .....	20
Пример 16 задания с выбором одного ответа .....	20
Пример 17 задания с выбором одного ответа .....	21
Пример 18 задания с выбором одного ответа .....	21
Пример 19 задания с выбором одного ответа .....	22
Пример 20 задания с кратким ответом .....	22
Пример 21 задания с кратким ответом .....	22
Решения заданий демоварианта 2012 .....	23
Задание А8 .....	23
Характеристики задания .....	23
Задание .....	23
Решение .....	23
Задание А11 .....	24
Характеристики задания .....	24
Задание .....	24
Решение .....	24
Задание В1 .....	26
Характеристики задания .....	26
Задание .....	26
Решение: .....	26

## Краткие теоретические сведения

### Измерение информации

С понятиями «вещество» и «энергия» связаны числовые характеристики, выраженные в некоторых единицах измерения. Можно измерить вес, длину, температуру, силу тока, количество тепла. Информацию тоже можно измерить.

Существует несколько подходов к измерению информации. Рассмотрим два из них – содержательный (вероятностный) и алфавитный.

Одними из первых количественное определение информации дали американские инженеры и исследователи Р.Хартли и К.Шеннон, изучавшие теории кодирования и передачи информации. Они предложили рассматривать информацию как снятую при получении некоторого сообщения неопределенность. Количество информации в сообщении зависит от его содержания. В сообщении, не несущем новых сведений, количество информации равно нулю. Такой подход к определению количества информации называют содержательным.

Алфавитный подход не учитывает содержания сообщения. Сообщение рассматривается как последовательность символов (знаков) некоторого алфавита, при этом учитывается информационный вес символа.

### Вероятностный подход

Сообщение несет информацию о наступлении некоторого события из нескольких возможных вариантов, снимая тем самым неопределенность. Например, известно, что проводилась контрольная работа, оценками которой могут быть 2, 3, 4 и 5. Сообщение о том, что ученик получил оценку 4, снимает неопределенность, произошло одно из четырех возможных событий.

Р.Хартли в 1928 г. сформулировал законы, которым должно подчиняться количество информации:

- 1) Если сообщение несет заранее известную информацию, количество информации равно нулю.
- 2) Чем больше количество возможных вариантов событий, тем больше информации содержится в сообщении о наступлении конкретного события.
- 3) Количество информации в сообщении о нескольких независимых событиях должно быть равно сумме количеств информации, содержащейся в сообщениях о каждом из этих событий.

Если обозначить количество событий (вариантов, исходов)  $N$ , а количество информации, зависящее от  $N$ , как  $Info(N)$ , то законы Хартли могут быть формализованы следующим образом:

- 1)  $Info(1) = 0$ ;
- 2)  $N_1 > N_2 \Rightarrow Info(N_1) > Info(N_2)$  (монотонность);
- 3)  $Info(N_1) + Info(N_2) + \dots + Info(N_k) = Info(N_1 + N_2 + \dots + N_k)$

Этим требованиям удовлетворяет функция логарифм  $\log_b N$ , где  $b > 1$ . Таким образом, формула для определения количества информации имеет вид

$$I = \log_b N$$

Минимально возможное количество информации снижает неопределенность в два раза, при этом  $b=2$ , и формула Хартли

$$I = \log_2 N$$

Это единица информации называлась «двоичная цифра». В 1947 г. американский математик Джон Тьюки (John Tukey) использовал термин бит (bit). С 1948 г. Клод Шеннон стал использовать этот термин в своих работах по теории связи.

Заметим, что количество информации по Хартли совпадает с длиной двоичного кодового слова.

Получение информационного сообщения в один бит уменьшает неопределенность нашего знания о чем-либо в два раза.

Хартли учитывал только количество возможных исходов опыта (событий), считая, что они равновероятны, то есть ни одно событие не имеет преимуществ перед другими. Примеры таких опытов – подбрасывание монеты или игральной кости. В первом случае количество исходов равно двум (орел или решка), во втором – шести, по количеству граней кости. Сообщение о том, что выпала решка, несет  $I = \log_2 N = \log_2 2 = 1$  бит информации. Сообщение о том, что выпала одна из шести граней, несет  $I = \log_2 6$  бит.

Если возможное количество равновероятных исходов опыта, или событий, равно  $N$ , то вероятность наступления одного из них определяется как  $p=1/N$ , и формула Хартли имеет вид

$$I = \log_2 1/p = -\log_2 p$$

Значение вероятности может изменяться от 0 до 1. Если известно, что событие произойдет обязательно, то вероятность его наступления равна единице («после зимы наступит весна»), и количество информации равно нулю (отсутствие новизны).

При измерении количества информации Р.Хартли не учитывал вероятность наступления события. К.Шеннон в своих работах предложил учитывать вероятность наступления события при измерении информации. Основная идея заключалась в том, что

сообщение о наступлении маловероятного события несет большее количество информации, чем сообщение о наступлении более вероятного.

Если  $N$  – общее число возможных событий, из них интересующее нас событие может произойти  $K$  раз, то вероятность этого события можно оценить как  $K/N$ . Вероятность выражается в долях единицы. Вероятность достоверного события равна 1. Достоверным называют событие, которое обязательно произойдет. Например, в корзине 30 шаров, все они красного цвета. Событие «из корзины вынут красный шар» достоверное. Вероятность невозможного события равна нулю (из корзины с 30-ю красными шарами вынут белый шар).

Допустим, контрольную работу писал отличник. Пусть вероятность того, что он получит оценку 5, равна  $1/2$ , оценку 4 –  $1/4$ , оценки 3 и 2 – по  $1/8$ <sup>1</sup>. Для каждого возможного сообщения в отдельности формула Хартли справедлива и должна применяться. Если вероятность некоторого  $j$ -го события равна  $p_j$ , то количество информации о наступлении этого события равно  $-\log_2 p_j$ . Таким образом, сообщение о том, что ученик получил оценку 5, несет  $-\log_2(1/2) = 1$  бит, оценку 4  $-\log_2(1/4) = 2$  бита, оценку 2 или 3  $-\log_2(1/8) = 3$  бита информации.

Средневзвешенное количество информации, которая может быть получена, определяется по формуле

$$I = 1/2 \cdot 1 \text{ бит} + 1/4 \cdot 2 \text{ бита} + 1/8 \cdot 3 \text{ бита} + 1/8 \cdot 3 \text{ бита} = 1,75 \text{ бита.}$$

В общем случае среднее количество информации, получаемой при неравновероятных событиях, определяется по формуле Шеннона:

$$I = -\sum_{i=1}^n p_i \log_2 p_i$$

где  $p_i$  – вероятность наступления  $i$ -го события.

Результаты, полученные по этой формуле, совпадают с результатами формулы Хартли при равновероятных событиях.

К. Шеннон показал, что учет вероятности наступления события определяет экономию времени, затрачиваемого на передачу сообщения. В качестве примера он рассмотрел сообщение, алфавит которого состоит из четырех символов – А, В, С, D. Вероятности появления этих символов в тексте сообщения равны соответственно  $1/2$ ,  $1/4$ ,  $1/8$  и  $1/8$ . Это значит, что в длинном тексте символ А будет занимать половину всего текста, символ В – одну четверть и т.д. Все символы кодируются двоичными знаками 0 и 1. Тогда

<sup>1</sup> Эти события образуют полную группу событий, и сумма их вероятностей равна единице:  $1/2 + 1/4 + 1/8 + 1/8 = 1$ .

самый простой код – равномерный: А – 00, В – 01, С – 10, D – 11. Для этого кода требуется два двоичных знака (2 бита) на символ сообщения.

Если же учесть вероятность появления символа, можно построить более хороший (префиксный) код: А – 0, В – 10, С – 110, D – 111. Количество использованных двоичных знаков в среднем уменьшится и будет равно

$$1/2 \cdot 1 + 1/4 \cdot 2 + 1/8 \cdot 3 + 1/8 \cdot 3 = 1,75 \text{ (бит)}.$$

К. Шеннон показал, что не существует метода кодирования, который использует меньшее число двоичных знаков.

## Алфавитный подход

Алфавитный подход называют техническим. В технике информацией считается любая последовательность символов или знаков. Для определения количества информации используют длину последовательности, заданной в исходном алфавите, без учета ее содержания, и количество кодовых символов закодированного сообщения. Набор символов алфавита исходного сообщения можно рассматривать как различные возможные события.

Если допустить, что все символы исходного алфавита встречаются в тексте с одинаковой частотой (равновероятно), то количество информации  $I$ , которое несет каждый символ, вычисляется по формуле:

$$I = \log_2 N,$$

где  $N$  — мощность исходного алфавита. Величину  $I$  называют информационным весом символа (исходного алфавита).

Количество информации, содержащееся в символьном сообщении, вычисляется по формуле:

$$I_c = K \cdot I = K \cdot \log_2 N$$

где  $K$  – количество символов в тексте исходного сообщения. Если  $\log_2 N$  не является целым числом, его нужно округлить до ближайшего большего целого. Часто  $I_c$  обозначают символом  $V$  и называют объемом информации.

Итак, при алфавитном подходе к измерению информации количество информации зависит от объема текста (числа знаков) и от количества символов в алфавите.

## Единицы измерения информации

Повторим, что в вычислительной технике битом называют наименьшую порцию памяти, необходимую для хранения одного из двух знаков «0» и «1», используемых для представления информации (данных и команд) в компьютере.

Бит – очень маленькая единица информации, поэтому существует величина **байт**, представляющая собой последовательность из 8 бит (1 байт –  $2^3$  бит).

Более крупные единицы измерения информации обозначаются с использованием префиксов (приставок), известных нам из физики (кило, мега и т.д.). В международной системе единиц эти приставки означают кратность степени числа 10, поэтому их называют десятичными. При измерении информации каждый следующий префикс соответствует увеличению значения не в 1000, а в  $1024 = 2^{10}$  раз. Приставки называют двоичными и записывают, начиная с прописной буквы<sup>2</sup>. Например:

Килобит      1 Кбит = 1024 бит =  $2^{10}$  бит

Мегабит      1 Мбит =  $2^{10}$  Кбит =  $2^{20}$  бит

Гигабит      1 Гбит =  $2^{10}$  Мбит =  $2^{20}$  Кбит =  $2^{30}$  бит

Терабит      1 Тбит =  $2^{10}$  Гбит =  $2^{20}$  Мбит =  $2^{30}$  Кбит =  $2^{40}$  бит

Петабит      1 Пбит =  $2^{10}$  Тбит =  $2^{20}$  Гбит =  $2^{30}$  Мбит =  $2^{40}$  Кбит =  $2^{50}$  бит

При решении задач следует переходить к одной единице измерения. Вычисления можно выполнить довольно легко, если в значениях величин выделить сомножители – степени двойки. Таблица соответствия единиц измерения информации.

<b>1 байт</b>	$2^3$ бит					
<b>1 Кбайт</b>	$2^{13}$ бит	$2^{10}$ байт				
<b>1 Мбайт</b>	$2^{23}$ бит	$2^{20}$ байт	$2^{10}$ Кбайт			
<b>1 Гбайт</b>	$2^{33}$ бит	$2^{30}$ байт	$2^{20}$ Кбайт	$2^{10}$ Мбайт		
<b>1 Тбайт</b>	$2^{43}$ бит	$2^{40}$ байт	$2^{30}$ Кбайт	$2^{20}$ Мбайт	$2^{10}$ Гбайт	
<b>1 Пбайт</b>	$2^{53}$ бит	$2^{50}$ байт	$2^{40}$ Кбайт	$2^{30}$ Мбайт	$2^{20}$ Гбайт	$2^{10}$ Тбайт

## **Представление информации в компьютере**

Команды программ и данные для обработки поступают в процессор из памяти. Память является электронным устройством и состоит из микросхем, которые в свою очередь состоят из тысяч более мелких электронных компонентов. Подобные электронные компоненты могут находиться только в двух состояниях: «включено» или «выключено», что соответствует двум цифрам двоичной системы счисления 1 или 0, или одному биту.

<sup>2</sup> По международным стандартам для обозначения кратности степеням двойки должны использоваться новые обозначения двоичных приставок. Они начинаются на те же слоги, что десятичные, но второй слог у всех двоичных приставок — би (binary — «двоичный», англ.). То есть килобит становится кибибитом, мегабит — мебибитом, и т. д. Эти обозначения пока «не прижились», и на практике чаще используются привычные десятичные приставки.

Таким образом, любая информация в памяти компьютера представляется в виде последовательности битов, каждый из которых находится в одном из допустимых состояний.

### Кодирование текстовой информации. Кодовые таблицы символов

В компьютерах используются специальные кодовые таблицы, которые ставят в соответствие двоичные последовательности (коды) и символы. Существуют стандартные таблицы кодов. Они могут использовать один, два и более байтов для кодирования одного символа.

Широко используется таблица кодов, известная как **стандарт ASCII (American Standart Code for Information Interchange** – Американский стандартный код для обмена информацией), использующая один байт для кодирования одного символа. ASCII представляет собой кодировку для представления десятичных цифр, символов латинского и национального алфавитов, знаков препинания, символов арифметических операций и управляющих символов. Управляющие символы называют непечатаемыми символами, к ним относятся такие как «перевод строки» (код символа 10), «возврат каретки» (код 13) и др.

Первая половина кодовой таблицы содержит **стандартные** символы ASCII (символы с кодами 0 – 127). Они одинаковые во всех странах и известны как стандарт, принятый Американским национальным институтом стандартов (**American National Standarts Institute – ANSI) ANSI X3.4.**

Первые 32 символа являются управляющими, а остальные печатаемыми. В таблице приведены символы с кодами 32 – 127 первой части кодовой таблицы ASCII. Коды в таблице записаны в шестнадцатеричной системе счисления, как принято в информатике. Код символа "A", например,  $41_{16} = 65_{10}$ .

Первая часть кодовой таблицы ASCII

КОД	СИМВОЛ	КОД	СИМВОЛ	КОД	СИМВОЛ	КОД	СИМВОЛ	КОД	СИМВОЛ	КОД	СИМВОЛ
20	пробел	30	.	40	@	50	P	60	'	70	p
21	!	31	0	41	A	51	Q	61	a	71	q
22	"	32	1	42	B	52	R	62	b	72	r
23	#	33	2	43	C	53	S	63	c	73	s
24	\$	34	3	44	D	54	T	64	d	74	t
25	%	35	4	45	E	55	U	65	e	75	u
26	&	36	5	46	F	56	V	66	f	76	v
27	'	37	6	47	G	57	W	67	g	77	w
28	(	38	7	48	H	58	X	68	h	78	x
29	)	39	8	49	I	59	Y	69	i	79	y
2A	*	3A	9	4A	J	5A	Z	6A	j	7A	z
2B	+	3B	:	4B	K	5B	[	6B	k	7B	{
2C	,	3C	;	4C	L	5C	\	6C	l	7C	



2D	-	3D	<	4D	M	5D	]	6D	m	7D	}
2E	.	3E	>	4E	N	5E	^	6E	n	7E	~
2F	/	3F	?	4F	O	5F		6F	o	7F	DEL

Таблицу кодов не надо запоминать, но следует помнить последовательность символов:

- 1) знаки препинания и арифметических операций;
- 2) цифры от 0 до 9;
- 3) прописные символы латинского алфавита;
- 4) строчные символы латинского алфавита.

Если в Блокноте или текстовом редакторе MS Word нажать и удерживать клавишу Alt, а на цифровой части клавиатуры (калькуляторе) набрать десятичный код символа, то он появится на экране. Например, Alt+1 – это управляющий символ кодовой таблицы, на экране он имеет вид ☺.

Вторая часть кодовой таблицы (символы с кодами 128 – 255) называют **расширенными кодами ASCII**. В расширенные коды ASCII включают символы национальных алфавитов, например, символы кириллицы. Но даже с учетом этих дополнительных знаков алфавиты многих языков не удается охватить при помощи 256 знаков. По этой причине существуют различные варианты кодировки ASCII, включающие символы разных языков.

Отсутствие согласованных стандартов привело к появлению различных кодовых таблиц (вернее, различных вторых частей кодовых таблиц, которые называют кодовыми страницами) для кодирования символов кириллицы, среди которых

- международный стандарт ISO 8859;
- кодовая таблица фирмы Microsoft CP-1251<sup>3</sup> (Windows-1251), используется приложениями, работающими под ОС Windows;
- кодовая таблица, применяемая в интернете, в ОС Unix KOI8-R<sup>4</sup>

и другие.

По этой причине тексты на русском языке, набранные с использованием одной кодовой таблицы, невозможно прочитать при использовании другой кодовой таблицы.

В настоящее время в компьютерах широко применяется стандарт кодирования **Unicode (Юникод)**, в котором для кодирования одного символа отводится 4 байта. Первые 128 символов Юникода совпадают с символами ASCII. Остальная часть кодовой

<sup>3</sup> CP – сокр. от Code Page – кодовая страница

<sup>4</sup> KOI – сокр. от Code for Information Exchange – код обмена информацией.

таблицы включает символы, используемые в основных языках мира, математическую и другую научную символику.

В настоящее время широко распространена кодировка **UTF<sup>5</sup>-8**. Она реализует представление Юникода, но использует один байт для кодирования символов первой части кодовой таблицы ASCII, от 2 до 6 байт для кодирования остальных символов.

## Кодирование графической информации

Изображение на экране монитора формируется набором экранных точек – **пикселей** (pixel – picture element). Каждая экранная точка имеет свой цвет, код цвета каждой точки хранится в памяти компьютера. Качество изображения зависит от разрешающей способности (сокр. разрешение) монитора – количества пикселей, отображаемых по горизонтали и вертикали экрана монитора.

Первые мониторы были монохромными. Для кодирования цвета пикселя использовался один бит памяти, значение 1 соответствует белому цвету, 0 – черному. Подобные экраны используются в недорогих сотовых телефонах, системах видеонаблюдения и других устройствах.

Цвет каждого пикселя современного дисплея определяется компонентами трех основных цветов: красного (Red, R), зеленого (Green, G) и синего (Blue, B). В памяти необходимо хранить информацию о состоянии каждой точки изображения, то есть о состоянии каждой из ее трех составляющих. Управление яркостью (интенсивностью) каждой составляющей позволяет влиять на цвет экранной точки.

**Цветовой моделью** называется представление цвета в виде наборов чисел (обычно трех или четырех). В компьютерной графике используется несколько видов цветовых моделей.

Рассмотрим цветовую модель, связанную с представлением пикселя составляющими красного, зеленого и синего цветов. Она называется **RGB (Red-Green-Blue) моделью**.

В RGB модели происходит сложение цветов и добавление их к черному цвету экрана, поэтому она называется **аддитивной** (additive). Разные цвета образуются смешиванием трех основных цветов в разных пропорциях, то есть с разными яркостями.

**Глубина цвета** (color depth) — это число бит, используемых для хранения кода цвета одного пикселя изображения.

---

<sup>5</sup> UTF – сокр. от Unicode Transformation Format — формат преобразования Юникода

В модели **RGB** каждый цвет может кодироваться тремя байтами (режим **TrueColor**). Каждый байт отвечает за яркость красной, зеленой и синей составляющей пикселя соответственно. Таким образом, глубина цвета в режиме TrueColor составляет 24 бита. Изображения, пиксели которых закодированы таким способом, называются 24-битными изображениями.

Чтобы указать цвет пикселя в модели **RGB**, достаточно перечислить разделенные точками яркости каждой составляющей, например, 255.255.0 – код желтой точки, записанный при помощи десятичных кодов яркостей. Значения яркости варьируются от 0 – выключено до 255 – включено на максимум. Если значения яркостей всех трех составляющих равны, получим оттенки серого цвета.

В языке разметки HTML для обозначения цвета используется шестнадцатеричная запись вида "#XXXXXX". Каждый цвет записывается в виде двух шестнадцатеричных цифр без пробелов. Первая пара цифр – интенсивность красного цвета, вторая – зеленого, последняя – синего. Напомним, что  $FF_{16} = 1111\ 1111_2 = 255_{10}$ . Например, белый цвет обозначается так "#FFFFFF".

Значения некоторых цветов в числовой модели RGB

Цвет	R	G	B	R	G	B
	Десятичная запись			Шестнадцатеричная запись		
Красный (red)	255	0	0	FF	00	00
Зеленый (green)	0	255	0	0	FF	0
Синий (blue)	0	0	255	00	00	FF
Фуксин (magenta)	255	0	255	FF	00	FF
Голубой (cyan)	0	255	255	00	FF	FF
Желтый (yellow)	255	255	0	FF	FF	00
Белый (white)	255	255	255	FF	FF	FF
Черный (black)	0	0	0	00	00	00

На рисунке 1 представлена модель RGB. Отметим, что на рисунке не показан черный фон.

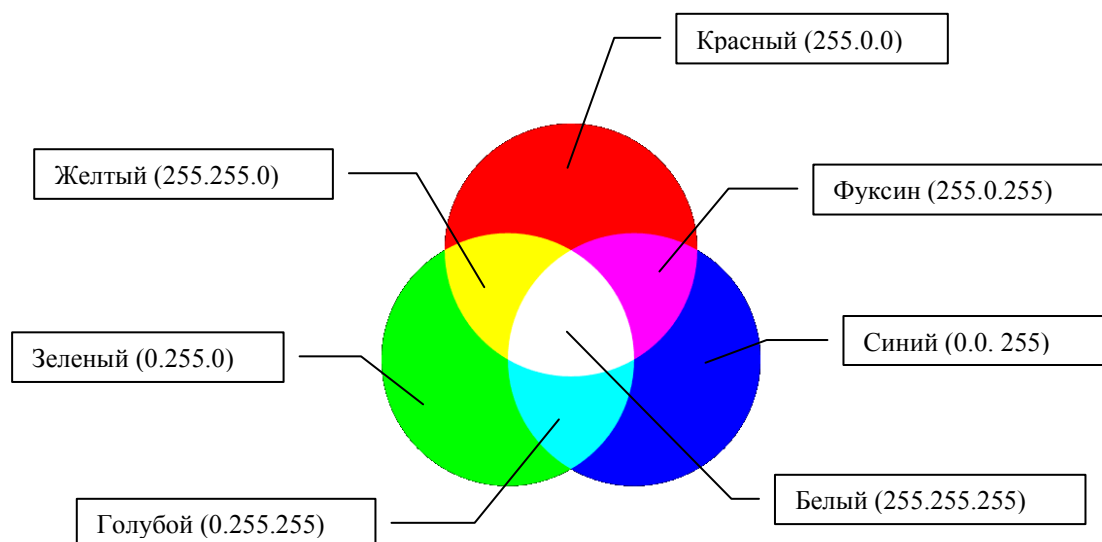
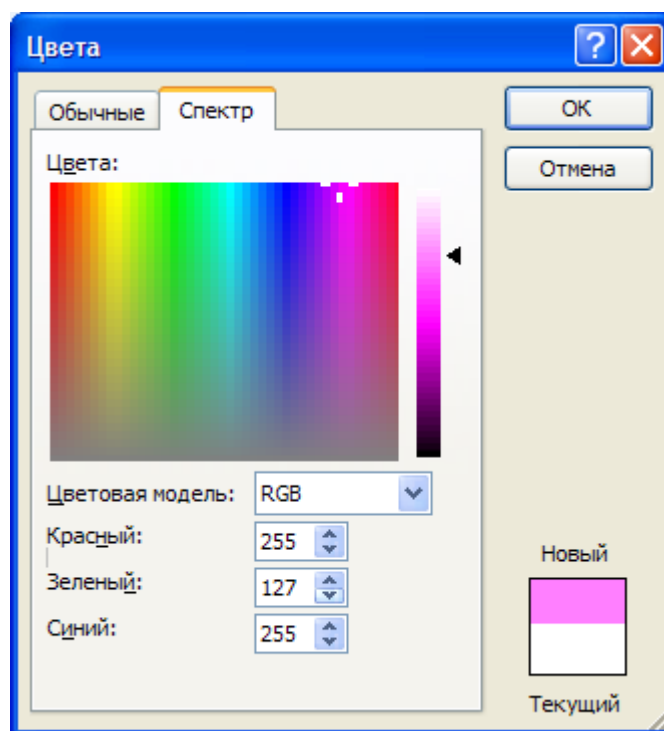


Рис. 1. Цветовая модель RGB

Если изменять интенсивность каждого цвета для смешанных цветов, например, задать цвет 127.127.0, то мы получим на экране болотный цвет, а не более темный оттенок желтого цвета, как можно было ожидать. Это связано с тем, что человеческий глаз более чувствителен к зеленому цвету. Чем ниже интенсивности составляющих, тем темнее цвет на экране. И наоборот – чем выше интенсивности цветов, тем светлее оттенки.

Во многих приложениях есть возможность настройки цвета заданием интенсивности каждой составляющей. На рис. 4.2 показано диалоговое окно настройки



цвета в приложениях MS Office.

Рис. 2. Диалоговое окно настройки цвета

### **Глубина цвета и объем памяти**

Количество различных цветов  $N$  и количество битов для их кодирования (глубина цвета)  $I$  связаны формулой  $N = 2^I$ . При  $I = 24$  бита можно закодировать  $2^{24} = 16\,777\,216$  различных цветов. При известном  $N$  глубина цвета определяется по формуле

$$I = \log_2 N$$

Если известно разрешение экрана (количество точек по горизонтали и вертикали) и глубина цвета, можно определить объем видеопамати для хранения одного кадра (одной страницы) изображения. Например, при разрешении экрана  $640 \times 480$  и использовании 24 бит на точку объем видеопамати равен

$$640 \cdot 480 \cdot 24 \text{ бит} = 2^7 \cdot 5 \cdot 2^5 \cdot 3 \cdot 5 \cdot 2^3 \cdot 3 \text{ бит} = 2^{15} \cdot 9 \cdot 25 \text{ бит} = 2^{15} / 2^{13} \cdot 9 \cdot 25 \text{ Кбайт} = 4 \cdot 9 \cdot 25 \text{ Кбайт} = 900 \text{ Кбайт}.$$

### **Кодирование звуковой информации**

Звук – это распространяющиеся в упругих средах – газах, жидкостях и твёрдых телах – механические колебания с непрерывно меняющейся амплитудой и частотой, воспринимаемые органами слуха. Механические колебания с частотами от 15 Гц до 20 кГц, которые способны воспринимать органы слуха человека, называются звуковыми или акустическими. Чем больше амплитуда сигнала, тем громче звук, чем больше частота сигнала (число колебаний в секунду), тем выше тон.

В настоящее время существует два основных способа записи звука: аналоговый (непрерывный) и цифровой (дискретной). Виноловая пластинка является примером аналогового хранения звуковой информации, так как звуковая дорожка свою форму изменяет непрерывно. Компакт-диски являются примером цифрового хранения звуковой информации, так как звуковая дорожка компакт-диска содержит участки с различной отражающей способностью.

Для того чтобы записать звук на какой-нибудь носитель, его нужно преобразовать в электрический сигнал. Это делается с помощью микрофона. Микрофоны имеют мембрану, которая колеблется под воздействием звуковых волн. К мембране присоединена катушка, перемещающаяся синхронно с мембраной в магнитном поле. В катушке возникает переменный электрический ток. Так звуковые волны преобразуются микрофоном в электрический ток переменного напряжения, который представляет собой аналоговый сигнал. Применительно к электрическому сигналу «аналоговый» обозначает, что этот сигнал непрерывен по времени и амплитуде (рис. 4.5.а)

Для того чтобы компьютер мог обрабатывать звук, непрерывный сигнал должен быть превращен в последовательность электрических импульсов (двоичных нулей и

единиц). В процессе кодирования непрерывного звукового сигнала, производится его дискретизация по времени. **Дискретизация** – это преобразование непрерывных сигналов в набор дискретных значений, каждому из которых присваивается число.

Для дискретизации надо несколько раз в секунду измерять величину аналогового сигнала и кодировать ее, например, с помощью 256 значений.

Фактически плоскость, на которой изображен непрерывный сигнал, разбивается вертикальными и горизонтальными линиями (рис. 4.5.б) и считается, что график проходит строго через узлы полученной сетки, непрерывная плавная линия заменяется ломаной.

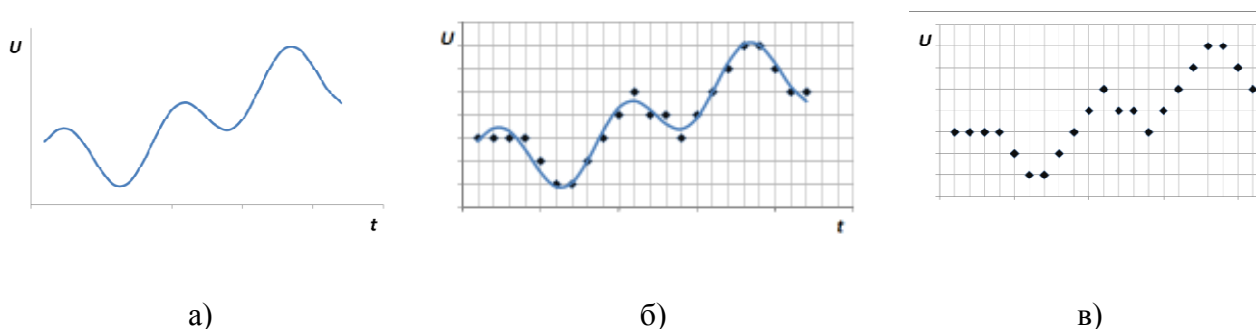


Рис. 4.5. Пример дискретизации аналогового сигнала

Дискретизация по времени соответствует разбиению вертикальными линиями. Она характеризуется **частотой дискретизации**. Частота дискретизации звукового компакт-диска 44,1 кГц, DVD – примерно 96 кГц. Это значит, что величина аналогового сигнала измеряется 44100 и 96000 раз в секунду соответственно. Если кодируется стереозвук, отдельно кодируются два канала.

Горизонтальное разбиение также важно, чем меньше расстояние между горизонтальными линиями сетки, тем качественнее будет цифровой звук. Количество линий сетки определяет количество уровней звука, поэтому горизонтальное разбиение называется **квантованием по уровню**. Для кодирования полученных значений уровней используют двоичные числа. Количество используемых для кодирования бит называется **глубиной звука**. Если глубина звука  $I=8$  бит или  $I=16$  бит, можно закодировать соответственно  $N=2^8=256$  уровней или  $N=2^{16}=65536$  уровней сигналов. Это значит, что интервал от нулевого до максимального напряжения аналогового сигнала разбивается на 256 или 65536 уровней, что соответствует количеству высот звука (тонов).

Из рисунка 4.5 видно, что преобразование аналогового сигнала в цифровой всегда происходит с некоторыми искажениями.

Преобразование непрерывной звуковой волны в последовательность звуковых импульсов различной амплитуды производится с помощью аналого-цифрового преобразователя (АЦП), размещенного на звуковой плате.

С помощью специальных программных средств (редакторов звукозаписей) открываются широкие возможности по созданию, редактированию и прослушиванию звуковых файлов. Но, как видно из примера, звуковые файлы занимают очень много места в памяти. Поэтому используются методы сжатия звуковых файлов. Качество музыки после сжатия несколько ухудшается, но это практически незаметно, так как при разработке алгоритмов сжатия учитываются законы восприятия музыки человеком.

Наиболее популярные форматы сжатых музыкальных файлов – MP3, WMA и др.

## Примеры решения заданий

### Измерение информации

#### Пример 1 задания с кратким ответом.

Сколько бит информации нужно получить, чтобы отгадать одно задуманное целое число из 32 возможных чисел в интервале  $[0;31]$ ?

**Решение:** По формуле Хартли

$$I = \log_2 32 = 5 \text{ или } 2^I = 32.$$

Минимальное количество информации, которое можно получить, содержит ответ на вопрос, допускающий два варианта ответа: «да» (1) или «нет» (0). Количество информации, равное 5 битам, означает, что для отгадывания числа достаточно задать пять вопросов. При получении ответа на каждый из вопросов количество рассматриваемых чисел (неопределенность) должно уменьшаться в два раза. Задав первый вопрос «Задуманное число находится во второй половине чисел?» и услышав ответ, получим один бит информации, а из 32 чисел останется 16 (от 0 до 15 или от 16 до 31). После получения ответа на второй вопрос, из 16 чисел останется 8 и т.д. Последний, пятый, вопрос позволит отгадать число. Заметим, что для кодирования 32 различных чисел двоичными кодами требуется 5 бит.

**Ответ:** 5

#### Пример 2 задания с кратким ответом

Сколько Килобайт информации содержит сообщение объемом  $2^{16}$  бит? В ответе укажите одно число.

**Решение.** Воспользуемся приведенной выше таблицей: 1 Кбайт =  $2^{13}$  бит, значит,  $2^{16}$  бит =  $2^{16-13}$  Кбайт =  $2^3$  Кбайт = 8 Кбайт.

**Ответ:** 8.

### Пример 3 задания с выбором одного ответа

Выберите вариант ответа, в котором объемы памяти расположены в порядке убывания

- 1) 1010 байт, 2 байта, 1 Кбайт, 20 бит, 10 бит
- 2) 1010 байт, 1 Кбайт, 20 бит, 2 байта, 10 бит
- 3) 1010 байт, 1 Кбайт, 2 байта, 20 бит, 10 бит
- 4) 1 Кбайт, 1010 байт, 20 бит, 2 байта, 10 бит

**Решение.** Приведем все величины к единой единице измерения, например, выразим все значения в битах.

$$1010 \text{ байт} = 1010 \cdot 8 \text{ бит};$$

$$2 \text{ байта} = 2 \cdot 8 \text{ бит} = 16 \text{ бит};$$

$$1 \text{ Кбайт} = 1024 \text{ байт} = 1024 \cdot 8 \text{ бит}.$$

Расположим значения по убыванию:

$$1024 \cdot 8 \text{ бит}; \quad 1010 \cdot 8 \text{ бит}; \quad 20 \text{ бит}; \quad 16 \text{ бит}; \quad 10 \text{ бит или}$$

$$1 \text{ Кбайт}; \quad 1010 \text{ байт}; \quad 20 \text{ бит}; \quad 2 \text{ байта}; \quad 10 \text{ бит}$$

**Ответ:** № 4.

### Пример 4 задания с кратким ответом

Сколько бит информации содержит сообщение о том, что на экзамене ученик вытянул билет №14, а всего экзаменационных билетов было 32?

**Решение:**

Выбран один из 32 билетов, произошло одно из 32 равновероятных событий. По формуле Хартли количество информации равно:

$$I = \log_2 32 = 5 \text{ (бит)}$$

**Ответ:** 5

### Пример 5 задания с кратким ответом

При угадывании целого числа из интервала от 10 до  $N$  получено 7 бит информации. Укажите максимально возможное значение  $N$ .

**Решение:**

В указанном интервале находится  $(N-10+1)$  целых чисел. произошло одно из 32 равновероятных событий. По формуле Хартли количество информации равно:

$$I = \log_2(N-9) = 7 \text{ (бит)}$$

$$\text{Тогда } N-9 = 2^7 = 128, N = 128+9 = 137$$

**Ответ:** 137



**Пример 6 задания с выбором одного ответа**

Преподаватель разработал 30 тестовых заданий по одной теме, все задания пронумерованы одинаковым минимально возможным количеством бит. Определите объем памяти, необходимый для хранения всех номеров заданий.

- 1) 30 байт                      2) 30 бит                      3) 19 байт                      4) 5 бит

**Решение:**

Для кодирования одного из 30 различных номеров требуется  $I = \log_2 30 \approx 5$  (бит). Для кодирования всех номеров требуется  $30 \cdot 5 = 150$  бит или  $150/8 \approx 19$  байт.

**Ответ:** № 3

**Пример 7 задания с выбором одного ответа**

В магазине 16 стеллажей с учебниками для 10-11 классов, на каждом стеллаже 8 полок. Автоматизированная система поиска книг выдала покупателю информацию, что сборник заданий для подготовки к ЕГЭ по информатике находится на пятой полке восьмого стеллажа. Какое количество информации получено покупателем?

- 1) 128 бит                      2) 16 бит                      3) 8 бит                      4) 7 бит

**Решение:**

В магазине  $16 \cdot 8$  полок с учебниками. Количество информации, полученное покупателем, равно  $I = \log_2(16 \cdot 8) = 7$  (бит).

**Ответ:** № 4

**Пример 8 задания с кратким ответом**

В корзине лежат 8 подберезовиков и 24 подосиновиков. Какое количество информации (в битах) содержится в сообщении о том, что взятый наугад из корзины гриб оказался подберезовиком?

**Решение:**

Общее количество грибов в корзине  $8+24=32$ . Количество информации, полученное в сообщении о том, что вынут подберезовик, равно  $I = -\log_2(8/32) = 2$  (бит).

**Ответ:** 2

**Пример 9 задания с кратким ответом**

В течении полугодия ученик получал оценки 2, 3, 4 и 5, всего он получил 64 оценки. Сообщение о том, что ученик получил оценку 4, несет 2 бита информации. Сколько четверок получил ученик за полугодие?

**Решение:**

Пусть  $X$  – количество четверок, полученных учеником. Тогда количество информации в сообщении о полученной четверке, равно  $I = -\log_2(X/64) = 2$  (бит). Отсюда  $64/X = 4$ ,  $X = 64/4 = 16$

**Ответ:** 16

### Пример 10 задания с выбором одного ответа

Информационное сообщение объемом 1,5 Кбайт содержит 2048 символов. Сколько символов содержит исходный алфавит сообщения?

- 1) 256                      2) 128                      3) 64                      4) 6

**Решение:**

1) Определим информационный вес одного символа сообщения:

$$I = 1,5 \cdot 2^{13} \text{ бит} / 2048 = 1,5 \cdot 2^{13} \text{ бит} / 2^{11} = 1,5 \cdot 4 \text{ бит} = 6 \text{ бит.}$$

2) Количество символов исходного алфавита равно  $2^I = 2^6 = 64$

**Ответ:** № 3

### Пример 11 задания с кратким ответом

Для равномерного кодирования алфавита из 10 символов используется минимально возможное количество бит на символ. Сколько еще символов можно добавить к исходному алфавиту, не изменяя информационного веса символа?

**Решение:**

Определим информационный вес одного символа алфавита:  $I = \log_2(10)$ , округлим до ближайшего большего целого, получим 4 бита. Используя 4 бита, можно закодировать  $2^4 = 16$  разных символов. Следовательно, можно добавить к исходному алфавиту  $16 - 10 = 6$  символов.

**Ответ:** 6

### Пример 12 задания с кратким ответом

Метеорологическая станция ведет наблюдение за температурой воздуха. Температура представляется целым числом, изменяется в пределах от  $-30$  до  $+30$  градусов. Каждое значение температуры записывается минимально возможным целым количеством бит. Каков информационный объем результатов измерений за сутки в битах?

**Решение:**

1) Определим количество различных значений температуры:  $30 - (-30) + 1 = 61$  (одно значение соответствует температуре 0 градусов).

2) Для хранения результатов одного измерения потребуется  $I = \log_2(61)$ , округлим до ближайшего большего целого, получим 6 бит.

3) Для хранения результатов измерения за сутки потребуется  $24 \cdot 6 = 144$  бита.

**Ответ:** 144

### Пример 13 задания с кратким ответом

Метеорологическая станция ведет наблюдение за температурой воздуха. Температура представляется целым числом, изменяется в пределах от  $-28$  до  $+36$  градусов. Каждое значение температуры записывается минимально возможным целым количеством бит. Результаты измерения за сутки записываются минимально возможным целым количеством байт. Каков информационный объем результатов измерений за сентябрь в байтах?

**Решение:**

1) Определим количество различных значений температуры:  $36 - (-28) + 1 = 65$ .

2) Для хранения результатов одного измерения потребуется  $I = \log_2(65)$ , округлим до ближайшего большего целого, получим 7 бит.

3) Для хранения результатов измерения за сутки потребуется  $24 \cdot 7 = 168$  бит или  $168/8=21$  байт.

4) для хранения результатов измерений за 30 дней потребуется  $21 \cdot 30=630$  байт

**Ответ:** 630

### Пример 14 задания с выбором одного ответа

В некоторой стране автомобильный номер длиной 7 символов составляют из заглавных букв (используются только 20 различных букв) и десятичных цифр в любом порядке. Каждый такой номер в компьютерной программе записывается минимально возможным и одинаковым целым количеством байт (при этом используют посимвольное кодирование и все символы кодируются одинаковым и минимально возможным количеством бит).

Определите объем памяти, отводимый этой программой для записи 32 номеров.

1) 1120 байт      2) 160 байт      3) 140 байта      4) 128 байт

**Решение:**

1) Определим информационный вес одного символа. Количество символом исходного алфавита  $N = 20 + 10 = 30$ . Информационный вес символа, или количество двоичных разрядов для кодирования одного символа, равно

$$I = \log_2 N = \log_2 30$$

Округлим до ближайшего большего целого, получим 5 бит.

2) Для записи автомобильного номера длиной 7 символов потребуется

$$I_c = K \cdot I = 7 \cdot 5 = 35 \text{ бит,}$$

или 5 байт.

3) Для записи одного номера требуется 5 байт, для записи 32 номеров  $32 \cdot 5 = 160$  байт.

**Ответ:** № 2

## Представление информации в компьютере

### Пример 15 задания с выбором одного ответа

В таблице представлены коды некоторых символов в кодовой таблице ASCII:

Символ	1	F	D	2	P	p	f
Десятичный код	49	70	68	50	80	112	102
Шестнадцатеричный код	31	46	44	32	50	70	66

Каков шестнадцатеричный код символа «d»?

- 1) 5E                                      2) 100                                      3) 88                                      4) 64

**Решение:**

В кодовой таблице ASCII коды строчных букв больше кодов заглавных, строчные буквы, как и заглавные, расположены в таблице последовательно. Для того чтобы вычислить «расстояние» между буквами вычтем из кода символа 'f' код символа 'F' (или из кода 'p' код 'P'):  $102 - 70 = 32$ . Для получения кода символа 'd' следует к коду символа 'D' прибавить 32:  $68 + 32 = 100$ .

Представим число 100 в шестнадцатеричной системе:  $100_{10} = 64_{16}$

**Ответ:** № 4

### Пример 16 задания с выбором одного ответа

В кодировке Unicode каждый символ кодируется двумя байтами. Определите информационный объем текста:

**Длина этой строки – 24 бита.**

в этой кодировке

- 1) 48 бит                                      2) 224 бит                                      3) 384 бит                                      4) 448 бит

**Решение:**

Количество символов в заданной строке равно 28. При подсчете необходимо учитывать пробелы, точку и знак «длинное тире». Если на каждый символ отводится 2 байта, а текст состоит из 28 символов, то для его кодирования необходимо  $2 \cdot 28 = 56$  байт. Среди

представленных вариантов такого ответа нет. Значит, верный ответ приведен в битах. Переведем 56 байт в биты:  $56 \cdot 8 = 448$  бит.

**Ответ:** № 4

### Пример 17 задания с выбором одного ответа

Автоматическое устройство осуществило перекодировку текста на русском языке, первоначально записанного в 16-битном коде Unicode, в 8-битную кодировку KOI8-R. При этом информационный объем текста уменьшился на 240 бит. Какова длина текста в символах?

- 1) 15                      2) 30                      3) 60                      4) 240

**Решение:**

Обозначим  $X$  – длину текста в символах. Тогда для 16-битного кода Unicode информационный объем текста составляет:  $16 \cdot X$  бит, а для 8-битной кодировки KOI8-R:  $8 \cdot X$  бит. Зная, что после перекодировки информационный объем текста уменьшился на 240 бит, составим уравнение:  $16 \cdot X - 8 \cdot X = 240$ , откуда  $X = 30$ .

**Ответ:** № 2

### Пример 18 задания с выбором одного ответа

Автоматическое устройство осуществило перекодировку текста, состоящего из символов русского и латинского алфавитов, цифр и знаков препинания, первоначально записанного в кодировке UTF-8, в 16-битный Unicode. В кодировке UTF-8 латинские символы, знаки препинания и другие символы ASCII занимают 1 байт, а остальные символы изображаются последовательностями длиной от 2 до 6 байт. В частности, символы кириллицы занимают в UTF-8 2 байта. В результате преобразования текст стал занимать 15 Кбит вместо первоначальных 10 Кбит. Сколько в тексте символов русского алфавита?

- 1) 320                      2) 640                      3) 625                      4) 5000

**Решение:**

Пусть  $X$  – количество букв русского алфавита,  $Y$  – количество прочих символов ASCII в сообщении. Тогда исходный текст занимает в памяти:  $8 \cdot Y + 16 \cdot X = 10 \cdot 1024$  (бит), а перекодированный:  $16 \cdot Y + 16 \cdot X = 15 \cdot 1024$  (бит). Решим систему двух уравнений, получим  $8 \cdot Y = 5 \cdot 1024$ . Подставим  $8 \cdot Y$  в первое уравнение,  $16 \cdot X = 10 \cdot 1024 - 5 \cdot 1024$ ;  $X = 5 \cdot 1024 / 16$ .

Количество символов русского алфавита  $X = 320$

**Ответ:** № 1

**Пример 19 задания с выбором одного ответа.**

Определить минимальный объем памяти, необходимой для хранения растрового изображения размером  $16 \times 32$  пикселей при 256-цветном кодировании каждого пикселя.

- 1) 128 Кбайт            2) 16 Кбайт            3) 4 Кбайт            5) 512 байт

**Решение:** изображение состоит из  $16 \times 32 = 2^9$  пикселей. На кодирование одного пикселя требуется  $I = \log_2 256 = 8$  бит = 1 байт. На кодирование всего изображения требуется  $2^9 \cdot 1$  байт = 512 байт.

**Ответ:** № 4

**Пример 20 задания с кратким ответом**

Растровое изображение имеет размер  $3 \times 4$  дюйма, разрешение 96 dpi, каждая точка может иметь один из 256 оттенков. Минимальный объем памяти (Кбайт), необходимый для хранения этого изображения, равен \_\_\_\_\_

**Решение:**

- 1) Определим количество точек изображения  $3 \cdot 96 \cdot 4 \cdot 96 = 3 \cdot 3 \cdot 2^5 \cdot 2^2 \cdot 3 \cdot 2^5 = 27 \cdot 2^{12}$ .
- 2) Определим глубину цвета  $I = \log_2 256 = 8$  бит = 1 байт
- 3) Объем памяти  $V = 27 \cdot 2^{12} \cdot 1$  байт =  $27 \cdot 2^2$  Кбайт = 108 Кбайт

**Ответ:** 108

**Пример 21 задания с кратким ответом.**

Какой объем памяти потребуется для хранения одной минуты стереозвуковой записи, если частота дискретизации равна 44,1 кГц, для кодирования звука используется 16 бит.

**Решение:** измерение производится 44100 раз в секунду или  $44100 \cdot 60$  раз в минуту. Для сохранения результата каждого измерения требуется 16 бит = 2 байт. Всего  $44100 \cdot 60 \cdot 2$  байт. Для стереозвука записываются два сигнала, следовательно, потребуется

$$44100 \text{ 1/сек} \cdot 60 \text{ сек} \cdot 2 \text{ байт} \cdot 2 = 10584000 \text{ байт} = 10,1 \text{ Мбайт}$$

**Ответ:** 10,1

## Решения заданий демоварианта 2012

### Задание А8

#### Характеристики задания

Дискретное (цифровое) представление текстовой, графической, звуковой информации и видеoinформации. Единицы измерения количества информации.

Проверяемые элементы содержания	Знание технологии обработки звука
Контролируемый элемент содержания (по кодификатору)	3.3.3. Ввод и обработка звуковых объектов
Требования к уровню подготовки (по кодификатору)	2.8 использовать компьютер для обработки звука
Вид деятельности	Воспроизведение представлений или знаний (при выполнении практических заданий)
Уровень	базовый
Максимальный первичный балл	1
Время выполнения	2 мин.

#### Задание

**А8** Производится одноканальная (моно) звукозапись с частотой дискретизации 16 кГц и 24-битным разрешением. Запись длится 1 минуту, ее результаты записываются в файл, сжатие данных не производится. Какое из приведенных ниже чисел наиболее близко к размеру полученного файла, выраженному в мегабайтах?

- 1) 0,2                      2) 2                      3) 3                      4) 4

#### Решение

Измерение звука производится 16000 раз в секунду или  $16000 \cdot 60$  раз в минуту. Для сохранения результата каждого измерения требуется 24 бит = 3 байт. Всего  $16000 \cdot 60 \cdot 3$  байт. Вычислим размер файла в мегабайтах

$$16000 \cdot 60 \cdot 3 / 2^{20} = 2^7 \cdot 125 \cdot 2^2 \cdot 15 \cdot 3 / 2^{20} = 125 \cdot 15 \cdot 3 / 2^{11} \approx 6000 / 2048 \approx 3.$$

**Ответ:** 3

## Задание А11

### Характеристики задания

Проверяемые элементы содержания	Умение подсчитывать информационный объем сообщения
Контролируемый элемент содержания (по кодификатору)	1.1.3. Дискретное (цифровое) представление текстовой, графической, звуковой информации и видеоинформации. Единицы измерения количества информации.
Требования к уровню подготовки (по кодификатору)	1.4.1. Оценить объем памяти, необходимый для хранения информации
Вид деятельности	Применение знаний и умений в стандартной ситуации
Уровень	повышенный
Максимальный первичный балл	1
Время выполнения	3 мин.

### Задание

**А11** Для регистрации на сайте некоторой страны пользователю требуется придумать пароль. Длина пароля – ровно 11 символов. В качестве символов используются десятичные цифры и 12 различных букв местного алфавита, причём все буквы используются в двух начертаниях: как строчные, так и заглавные (регистр буквы имеет значение!).

Под хранение каждого такого пароля на компьютере отводится минимально возможное и одинаковое целое количество байтов, при этом используется посимвольное кодирование и все символы кодируются одинаковым и минимально возможным количеством битов. Определите объём памяти, который занимает хранение 60 паролей.

- 1) 540 байт                      2) 600 байт                      3) 660 байт                      4) 720 байт

### Решение

**Шаг 1.** Определим минимально необходимое количество бит для кодирования одного символа из заданного алфавита.

Количество возможных символов в пароле  $12 \cdot 2$  букв + 10 цифр = 34. Определим минимально необходимое количество бит для кодирования одного из 34 различных символов. При мощности алфавита равной двум (для кодирования числа используются цифры 0 и 1) количество разрядов (бит) для кодирования одного из 34 различных символов определяется по формуле:

$$I = \lceil \log_2 N \rceil$$



где  $I$  – количество бит для кодирования одного символа, а  $N$  – количество возможных различных символов. Запись  $\lceil X \rceil$  означает, что  $X$  округляется до ближайшего большего целого. Для кодирования одного символа нам потребуется  $\lceil \log_2 34 \rceil = 6$  бит.

**Шаг 2.** Определим минимально необходимое количество байт для кодирования одного пароля.

На кодирование 11 символов пароля необходимо  $6 \cdot 11 = 66$  бит. Для кодирования пароля одного пользователя необходимо  $\lceil 66/8 \rceil = 9$  байт.

**Шаг 3.** Информация о 60 паролях займет  $60 \cdot 9 = 540$  байт.

**Ответ:** № 1

## Задание В1

### Характеристики задания

Проверяемые элементы содержания	Кодирование текстовой информации. Кодировка ASCII. Основные кодировки кириллицы
Контролируемый элемент содержания (по кодификатору)	1.1.3. Дискретное (цифровое) представление текстовой, графической, звуковой информации и видеоинформации. Единицы измерения количества информации.
Требования к уровню подготовки (по кодификатору)	1.4.1. Оценить объем памяти, необходимый для хранения информации
Вид деятельности	Воспроизведение представлений или знаний (при выполнении практических заданий)
Уровень	базовый
Максимальный первичный балл	1
Время выполнения	1 мин.

### Задание

**В1** Автоматическое устройство осуществило перекодировку информационного сообщения на русском языке длиной в 20 символов, первоначально записанного в 2-байтном коде Unicode, в 8-битную кодировку КОИ-8. На сколько бит уменьшилась длина сообщения? *В ответе запишите только число*

### Решение:

Сообщение состоит из 20 символов. Первоначально оно имело объем  $20 \cdot 2$  байт, после перекодирования –  $20 \cdot 1$  байт. Следовательно, объем сообщения уменьшился на

$$20 \cdot 2 - 20 = 20 \text{ байт} = 160 \text{ бит}$$

**Ответ:** 160