

Тема 1.4

Кодирование данных в ЭВМ



Кодирование данных



Прием *кодирования*, то есть выражение данных одного типа через данные другого типа.

С О М Р U Т Е R

43 4F 4D 50 55 54 45 52

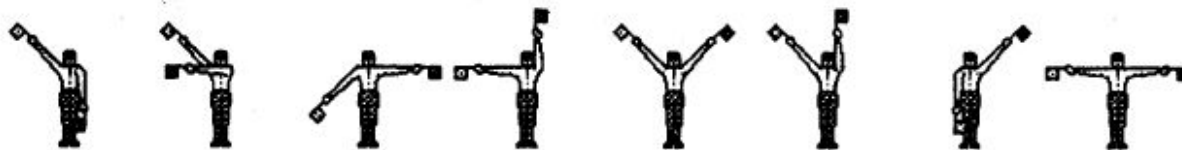
Код ASCII

· · · - - - - - · · · - - - - - - · · ·

Код Морзе

●● ●●● ●●● ●●●● ●●● ●●●● ●●● ●●●●

Код Брайля



Код морской сигнальный

Кодирование данных в двоичной форме



Двоичные цифры (англ. – *binary digit* или, сокращенно, *bit* (бит)).

Одним битом могут быть выражены два понятия:

0 или 1 (*да* или *нет*)

Двумя битами уже можно выразить четыре различных понятия:

- 00 01 10 11

Тремя битами можно закодировать восемь различных значений:

- 000 001 010 011 100 101 110 111

Вывод: Увеличивая на единицу количество разрядов в системе двоичного кодирования, мы увеличиваем в два раза количество значений, которое может быть выражено в данной системе:

$$N=2^m,$$

- N – количество независимых кодируемых значений;
- m – разрядность двоичного кодирования, принятая в данной системе.

Формы представления чисел

С целью упрощения схем вычитание в ЭВМ заменяется сложением специально построенных кодов чисел.

Применяются *прямой, обратный* и *дополнительный* коды чисел

Формы представления чисел

1) *Целые числа*

1.1 *Двоично-десятичное представление целых чисел:*

$$(123)_{10} = (0001.0010.0011)_{2-10}$$

1.2 *Прямой код:*

- число переводится в двоичную систему счисления;
- двоичную запись слева дополняют необходимым числом незначащих нулей.

$$(37)_{10} = (100101)_2 = (0000000000100101)_{\text{прямой код}}$$

Формы представления чисел

Замечание:

*В случае представления величины со знаком самый левый (старший) разряд указывает:
на **положительное число**, если содержит
ноль,
и на **отрицательное число**, если – единицу.*

Формы представления чисел

1.3 *Дополнительный код:*

- для положительного числа совпадает с его прямым кодом;
- для отрицательных чисел:
 - записывают прямой код модуля числа;
 - инвертируют его;
 - прибавляют к инверсному коду 1.

$(-37)_{10}$

(0000000000100101) *прямой код*

(1111111111011010) *инверсия*

(1111111111011011) *дополнительный код*

Формы представления чисел

Замечание:

При получении числа по его дополнительному коду, сначала определяют знак числа:

- *если число положительное просто переводят его в десятичную систему счисления;*
- *если число отрицательное:*
 - *вычитают из кода 1;*
 - *инвертируют код;*
 - *переводят в десятичную систему счисления.*

(1111111111000000) дополнительный код

(1111111110111111)

(0000000001000000)

$(1000000)_2 = (64)_{10}$ \square **ответ: -64**

Формы представления чисел

2) *Натуральная с фиксированным положением запятой*

Недостатками представления чисел с фиксированной запятой являются:

- необходимость предшествующего расчета и введения масштабных коэффициентов для исключения возможности переполнения разрядной сетки (когда число по модулю превышает единицу), а также потери младших разрядов (число по модулю меньше единицы младшего разряда);
- зависимость относительной точности от значения поступающих чисел. Максимальная относительная точность достигается при выполнении действий над максимально возможными числами.

Преимуществом является простота и высокое быстродействие выполнения операций.

Формы представления чисел

3) Полулогарифмическая *с плавающей запятой*

$$A = m \cdot q^n,$$

где q – основание СС;

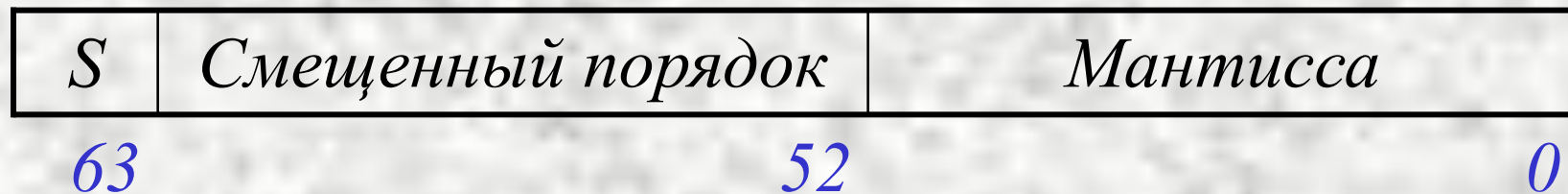
n – целое число, называемое **порядком** числа A ;

m – **мантисса** числа A ($1 < |m| < 10$).

$$\begin{aligned} 31,415926 &= \\ &= 3,1415926 \cdot 10^1 \end{aligned}$$

$$\begin{aligned} 0,00125 &= \\ &= 1,25 \cdot 10^{-3} \end{aligned}$$

Способ хранения мантиисы с плавающей точкой подразумевает, что двоичная запятая находится на фиксированном месте, а именно следует после первой двоичной цифры.



Смещение выбирается так, чтобы минимальному значению порядка соответствовал нуль.

Бит *S* указывает на знак числа.

Алгоритм для получения представления действительного числа в памяти ЭВМ:

- перевести модуль числа в двоичную систему счисления;
- нормализовать двоичное число;
- прибавить к порядку смещение и перевести смещенный порядок в двоичную систему счисления;
- учитывая знак заданного числа (0 – положительное; 1 – отрицательное) записать его представление в памяти ЭВМ.

Пример: Найти код числа: -312,3125.

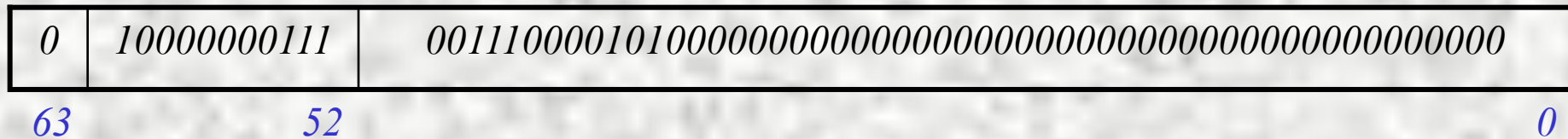
1. $(312,3125)_{10} = (100111000,0101)_2$

2. $(100111000,0101)_2 = (1,001110000101 \cdot 2^8)_2$

3. смещенный порядок:

$8+1023 = (1031)_{10} = (10000000111)_2$

Ответ:



Кодирование текстовых данных

Кодировка Windows 1251

128 Ъ	144 ђ	160	176 •	192 А	208 Р	224 а	240 р
129 Ѓ	145 ‘	161 Ў	177 ±	193 Б	209 С	225 б	241 с
130 ‚	146 ’	162 ў	178	194 В	210 Т	226 в	242 т
131 ґ	147 “	163 Ј	179 ¡	195 Г	211 У	227 г	243 у
132 ”	148 ”	164 ђ	180 г	196 Д	212 Ф	228 д	244 ф
133 ...	149 •	165 Г	181 μ	197 Е	213 Х	229 е	245 х
134 †	150 –	166	182 ¶	198 Ж	214 Ц	230 ж	246 ц
135 ‡	151 —	167 §	183 ·	199 З	215 Ч	231 з	247 ч
136 ‘	152 ’	168 Ё	184 ё	200 И	216 Ш	232 и	248 ш
137 ‰	153 ™	169 ©	185 №	201 Й	217 Щ	233 й	249 щ
138 Љ	154 щ	170 €	186 €	202 К	218 Ъ	234 к	250 ъ
139 ‹	155 ›	171 «	187 »	203 Л	219 Ы	235 л	251 ы
140 Њ	156 њ	172 ¬	188 ј	204 М	220 Ь	236 м	252 ь
141 Ќ	157 ќ	173 -	189 S	205 Н	221 Э	237 н	253 э
142 Ћ	158 ћ	174 ®	190 s	206 О	222 Ю	238 о	254 ю
143 Ў	159 џ	175 ĩ	191 ĩ	207 П	223 Я	239 п	255 я

Кодирование текстовых данных

Кодировка КОИ-8

128		144	▣	160	—	176	┌	192	ю	208	п	224	Ю	240	П
129		145	▤	161	Ё	177	┐	193	а	209	я	225	А	241	Я
130	┌	146	▥	162	Г	178	└	194	б	210	р	226	Б	242	Р
131	└	147	┌	163	ё	179	Ё	195	ц	211	с	227	Ц	243	С
132	┐	148	▧	164	г	180	┌	196	д	212	т	228	Д	244	Т
133	┘	149	•	165	Г	181	┐	197	е	213	у	229	Е	245	У
134	└	150	√	166	г	182	┐	198	ф	214	ж	230	Ф	246	Ж
135	┌	151	≈	167	Г	183	┐	199	Г	215	в	231	Г	247	В
136	┐	152	≪	168	Г	184	┐	200	х	216	ь	232	Х	248	Ь
137	┘	153	≫	169	Г	185	┘	201	и	217	ы	233	И	249	Ы
138	+	154	┘	170	Г	186	┘	202	й	218	з	234	Й	250	З
139	■	155	┘	171	Г	187	┘	203	к	219	ш	235	К	251	Ш
140	■	156	•	172	Г	188	+	204	л	220	э	236	Л	252	Э
141	■	157	²	173	Г	189	┌	205	м	221	щ	237	М	253	Щ
142	■	158	•	174	Г	190	┌	206	н	222	ч	238	Н	254	Ч
143	■	159	÷	175	Г	191	ё	207	о	223	ъ	239	О	255	Ъ

Кодирование текстовых данных

Кодировка ISO

В ISO не определены	160		176	А	192	Р	208	а	224	р	240	№
	161	Ё	177	Б	193	С	209	б	225	с	241	ё
	162	Ъ	178	В	194	Т	210	в	226	т	242	ђ
	163	Ѓ	179	Г	195	У	211	г	227	у	243	ѓ
	164	Є	180	Д	196	Ф	212	д	228	ф	244	є
	165	Ѕ	181	Е	197	Х	213	е	229	х	245	ѕ
	166	І	182	Ж	198	Ц	214	ж	230	ц	246	і
	167	Ї	183	З	199	Ч	215	з	231	ч	247	ї
	168	Ј	184	И	200	Ш	216	и	232	ш	248	ј
	169	Љ	185	Й	201	Щ	217	й	233	щ	249	љ
	170	Њ	186	К	202	Ъ	218	к	234	ъ	250	њ
	171	Ћ	187	Л	203	Ы	219	л	235	ы	251	ћ
	172	Ќ	188	М	204	Ь	220	м	236	ь	252	ќ
	173	-	189	Н	205	Э	221	н	237	э	253	ѕ
	174	Ў	190	О	206	Ю	222	о	238	ю	254	ў
	175	Џ	191	П	207	Я	223	п	239	я	255	џ

Кодирование текстовых данных

ГОСТ-альтернативная кодировка

128	А	144	Р	160	а	176	⋮	192	⌒	208	⌒	224	р	240	Ё
129	Б	145	С	161	б	177	■	193	Г	209	⌒	225	с	241	ё
130	В	146	Т	162	в	178	■	194	⌒	210	⌒	226	т	242	Є
131	Г	147	У	163	г	179	⌒	195	⌒	211	⌒	227	у	243	є
132	Д	148	Ф	164	д	180	⌒	196	—	212	⌒	228	ф	244	ї
133	Е	149	Х	165	е	181	⌒	197	⌒	213	⌒	229	х	245	і
134	Ж	150	Ц	166	ж	182	⌒	198	⌒	214	⌒	230	ц	246	ÿ
135	З	151	Ч	167	з	183	⌒	199	⌒	215	⌒	231	ч	247	ÿ
136	И	152	Ш	168	и	184	⌒	200	⌒	216	⌒	232	ш	248	·
137	Й	153	Щ	169	й	185	⌒	201	⌒	217	⌒	233	щ	249	·
138	К	154	Ъ	170	к	186	⌒	202	⌒	218	⌒	234	ъ	250	·
139	Л	155	Ы	171	л	187	⌒	203	⌒	219	■	235	ы	251	√
140	М	156	Ь	172	м	188	⌒	204	⌒	220	■	236	ь	252	№
141	Н	157	Э	173	н	189	⌒	205	⌒	221	■	237	э	253	α
142	О	158	Ю	174	о	190	⌒	206	⌒	222	■	238	ю	254	■
143	П	159	Я	175	п	191	⌒	207	⌒	223	■	239	я	255	

Кодирование текстовых данных



Универсальная система кодирования текстовых данных

UNICODE (Юникод)

**основана на 16-разрядном кодировании символов,
что позволяют обеспечить уникальные коды для
65 536
различных символов**

Кодирование графических данных

- 1) Ч/Б изображения - 256 градаций серого (8 бит)
- 2) Цветные изображения
 - а) RGB (Red, Green, Blue) – 24 бит (16.5 млн. цв.)
режим полноцветный *True Color*
 - б) CMYK (Cyan, Yellow, Magenta, Black) – 32 бит
режим полноцветный *True Color*
 - в) 16 бит - режим высококачественный *High Color*

Кодирование звуковой информации

1) *FM (Frequency Modulation – Частотная модуляция)*

**Кодирование на основе АЦП
(Аналогово-цифрового преобразования)**

2) *Wave-Table (Таблично-волновой синтез)*

**Кодирование на основе таблиц, содержащих образцы
звука (сэмплы)**

3) Для хранения 5-минутной аудиозаписи CD-качества
необходимый объем памяти составит:

$$\begin{aligned} & 48.000 \text{ Гц} \times 16 \text{ бит} \times 2 \text{ канала} \times 300 \text{ с} = \\ & = 57.600.000 \text{ байт} = 56.250 \text{ Кбайт} = 55 \text{ Мбайт} \end{aligned}$$