



Подання даних у пам'яті комп'ютера

к.т.н., доцент кафедри прикладної математики
Рижа Ірина Андріївна

Про що ця лекція???

- ▶ Опишемо форми подання чисел у пам'яті комп'ютера.
- ▶ Викладемо особливості подання додатніх та від'ємних чисел.
- ▶ Розглянемо кодування текстових даних.



Форми подання чисел у пам'яті комп'ютера

Форма подання числа

- сукупність правил, що дають змогу встановити взаємну відповідність між записом числа у пам'яті комп'ютера та його кількісним еквівалентом.

Розрядна сітка

- фіксована кількість двійкових розрядів, які використовуються комп'ютером для записування даних (зазвичай дорівнює 32, 64).

Машинне зображення числа

- це подання числа у розрядній сітці комп'ютера:

$$A = [A] K_A, \quad (1)$$

A – число;

$[A]$ – машинне зображення числа;

K_A – коефіцієнт, який залежить від форми подання числа.

Природна форма подання чисел

Природна форма подання (із природною комою)

– форма подання чисел у вигляді послідовності цифр, що розділені комою на цілу та дробову частини.

Наприклад:

- ▶ 12560 – ціле число;
- ▶ 0,003572 – правильний дріб;
- ▶ 4,89760 – неправильний дріб.

Для кожного числа у природній формі:

- ▶ **необхідно** вказати положення його коми в розрядної сітці, виділеній для представлення числа в машині;
- ▶ виникають додаткові апаратні витрати досить великого обсягу.

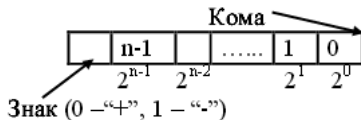
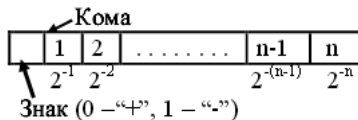
Форма подання чисел з фіксованою комою

Форма подання чисел з фіксованою комою (крапкою)

– форма подання, при якій положення крапки в розрядній сітці машини фіксується апаратними засобами раз і назавжди.

Положення крапки в розрядній сітці строго фіксується:

- ▶ перед старшим (лівим) розрядом числа – **правильний дріб**;
- ▶ після молодшого (правого) розряд – **ціле число**;



У **знакову частину** записується інформація про знак числа:

- ▶ знак позитивного числа "+" – символ 0;
- ▶ знак від'ємного числа "-" – символ 1.

Приклад 1.

Записати двійковий код числа $(+7)_{10}$ у формі з фіксованою крапкою, використовуючи 6-розрядну сітку.

$$(+7)_{10} = (111)_2$$

5	4	3	2	1	0
біт знаку	16	8	4	2	1
0	0	0	1	1	1

Приклад 2.

Записати двійковий код числа $(+23)_{10}$ у формі з фіксованою крапкою, використовуючи 8-розрядну сітку.

$$(+23)_{10} = (10111)_2$$

7	6	5	4	3	2	1	0
біт знаку	64	32	16	8	4	2	1
0	0	0	1	0	1	1	1

Кількість розрядів у розрядній сітці, відведених для зображення числа, визначає:

1. діапазон

- ▶ цілі числа: $1 \leq |A| \leq 2^{n-1} - 1$

2. точність представлення числа.

Машинний нуль

– число, абсолютне значення якого менше одиниці молодшого розряду розрядної сітки.

- ▶ записується в розрядну сітку комп'ютера у вигляді 0, тому що для його подання не вистачає довжини розрядної сітки;
- ▶ насправді не дорівнює 0.

Переповнення розрядної сітки комп'ютера

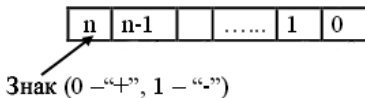
відбувається, коли число, отримане в результаті обчислень, перевищує за абсолютним значенням максимальне машинне число.

*Сьогодні в комп'ютерах у переважній більшості у форматі з фіксованою комою подаються **цілі** числа.*

Подання додатніх і від'ємних чисел

У комп'ютерах для спрощення виконання арифметичних операцій, числа подаються спеціальними двійковими кодами:

1. прямим;
2. оберненим;
3. додатковим.



- ▶ Лівий (старший) розряд зберігає знак числа.
- ▶ Цифрова частина (решта розрядів) використовується для представлення абсолютної величини числа.

Подання додатніх і від'ємних чисел

Додатні числа у прямому, оберненому та додатковому кодах записуються **однаково** – двійковим кодом числа з цифрою 0 у знаковому розряді.

Наприклад, при розмірі розрядної сітки $n = 8$

Число $1_{10} = 1_2$

0	0	0	0	0	0	0	1
---	---	---	---	---	---	---	---

Знак числа "+"

Число $127_{10} = 1111111_2$

0	1	1	1	1	1	1	1
---	---	---	---	---	---	---	---

Знак числа "+"

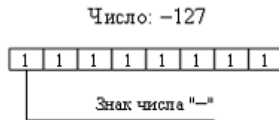
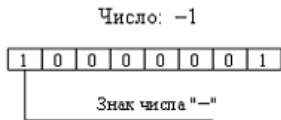
Подання додатніх і від'ємних чисел

Від'ємні числа в прямому, оберненому і додатковому кодах мають різне представлення.

Прямий код від'ємного числа

відрізняється від прямого коду додатного числа тим, що значення його знакового розряду дорівнює не 0, а 1.

Наприклад, прямий код чисел -1 і -127 у 8-розрядній сітці:



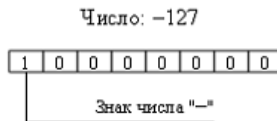
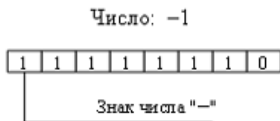
Подання додатніх і від'ємних чисел

Обернений код від'ємного числа

отримується із прямого коду шляхом заміни його цифр на їх доповнення до 1:

- ▶ у всіх розрядах нулі замінюються на 1, а одиниці – на 0;
- ▶ код знака зберігається без змін.

Наприклад, обернений код чисел -1 і -127 у 8-розрядній сітці:



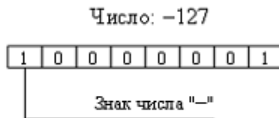
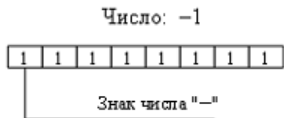
Подання додатніх і від'ємних чисел

Додатковий код від'ємного числа

отримується із оберненого коду збільшенням на 1 його молодшого розряду:

- ▶ перенос із знакового розряду ігнорується.

Наприклад, додатковий код чисел -1 і -127 у 8-розрядній сітці:



Подання додатніх і від'ємних чисел

Додатковий код є математичним доповненням основи p :

$$|X| + X_{\text{д}} = p^n,$$

$|X|$ – абсолютне значення числа;

n – кількість розрядів у представленні числа.

Наприклад, при $n = 8$:

$$(+1)_{\text{д}} = (1)_{10} = (00000001)_2;$$

$$(-1)_{\text{д}} = 2^8 - |-1| = 256 - 1 = (255)_{10} = (11111111)_2;$$

$$(-127)_{\text{д}} = 2^8 - |-127| = 256 - 127 = (129)_{10} = (10000001)_2.$$

- ▶ Для подання від'ємних чисел у форматі з фіксованою комою зазвичай використовується додатковий код.
- ▶ Операція віднімання замінюється на операцію додавання числа, що має знак, протилежний до знака від'ємника.

Наприклад

$$A - B = A + (-B).$$

Форма подання чисел з плаваючою комою

Форма подання чисел з плаваючою комою (крапкою)

– представлення числа у вигляді добутку цифрової частини та деякого степеня основи системи числення:

$$A = M_A p^s, \quad (2)$$

M_A – мантиса числа A ;

p – основа системи числення;

s – порядок числа.

- ▶ у мантисі M_A зберігаються значущі цифри числа;
- ▶ порядок s визначає величину числа;
- ▶ знак мантиси збігається зі знаком числа A .

Ця форма представлення – неоднозначна.

Наприклад

$$73,28 = 73,28 \cdot 10^0 = 7,328 \cdot 10^1 = 0,7328 \cdot 10^2 = 732,8 \cdot 10^{-1} = \dots$$

Форма подання чисел з плаваючою комою

Для однозначності представлення та підвищення точності запису використовують *нормальну форму подання*.

Нормальна форма подання чисел з плаваючою крапкою

– форма подання чисел, для якої виконується умова

$$\frac{1}{p} \leq |M_A| < 1. \quad (3)$$

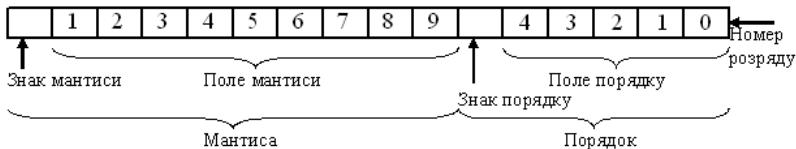
Для запису числа у формі з плаваючою крапкою фіксують 4 поля:



Найчастіше цей формат використовують для представлення дуже великих або дуже малих дійсних чисел.

Приклад 3.

У розрядну сітку, зображену на рисунку, записати двійкове число $A = -10110,1111_2$.

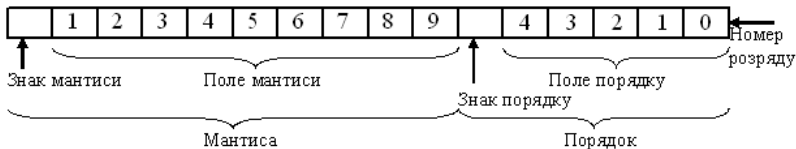


Нормальна форма: $A = -0,101101111 \cdot 2^5$; $\frac{1}{2} \leq |M_A| \leq 1$.

знак мантиси	1	2	3	4	5	6	7	8	9	знак порядку	4	3	2	1	0
1	1	0	1	1	0	1	1	1	1	0	0	0	1	0	1

Приклад 4.

У розрядну сітку, зображену на рисунку, записати двійкове число $A = +0,000110010111_2$.



Нормальна форма: $A = +0,110010111 \cdot 2^{-3}$; $\frac{1}{2} \leq |M_A| \leq 1$.

знак мантиси	1	2	3	4	5	6	7	8	9	знак порядку	4	3	2	1	0
0	1	1	0	0	1	0	1	1	1	1	0	0	0	1	1

Кодування текстових даних

Якщо кожному символу алфавіту поставити у відповідність певне ціле число – його **код**, то за допомогою двійкового коду можна кодувати і текстову інформацію.

Вісьмох двійкових розрядів достатньо для кодування **256** різних символів:

- ▶ усі символи англійської та української мов (великі і малі);
- ▶ розділові знаки;
- ▶ символи основних арифметичних операцій;
- ▶ деякі загальноприйняті спеціальні символи.

Для того, щоб увесь світ однаково кодував текстові дані, потрібні єдині **таблиці кодування**.

8-розрядна система кодування ASCII

(American Standard Code for Information Interchange – стандартний код інформаційного обміну США)

ввів у дію інститут стандартизації США (ANSI — American National Standard Institute) для англійської мови:

- ▶ підтримує кодування **256** (2^8) незалежних значень;
- ▶ **базова таблиця** закріплює значення кодів від 0 до 127;
- ▶ **розширена таблиця** – від 128 до 255.

Базова таблиця ASCII:

- ▶ Коди від 0 до 31 – **управляючі коди**, яким не відповідають жодні символи мов і, відповідно, ці коди не виводяться ні на екран, ні на пристрої друку, але ними можна управляти виведенням інших даних;
- ▶ Коди від 32 до 127 – призначені для кодування символів англійського алфавіту, розділових знаків, цифр, арифметичних операцій і деяких допоміжних символів.

ASCII TABLE

Decimal	Hex	Char	Decimal	Hex	Char	Decimal	Hex	Char	Decimal	Hex	Char
0	0	[NULL]	32	20	[SPACE]	64	40	@	96	60	`
1	1	[START OF HEADING]	33	21	!	65	41	A	97	61	a
2	2	[START OF TEXT]	34	22	"	66	42	B	98	62	b
3	3	[END OF TEXT]	35	23	#	67	43	C	99	63	c
4	4	[END OF TRANSMISSION]	36	24	\$	68	44	D	100	64	d
5	5	[ENQUIRY]	37	25	%	69	45	E	101	65	e
6	6	[ACKNOWLEDGE]	38	26	&	70	46	F	102	66	f
7	7	[BELL]	39	27	'	71	47	G	103	67	g
8	8	[BACKSPACE]	40	28	(72	48	H	104	68	h
9	9	[HORIZONTAL TAB]	41	29)	73	49	I	105	69	i
10	A	[LINE FEED]	42	2A	*	74	4A	J	106	6A	j
11	B	[VERTICAL TAB]	43	2B	+	75	4B	K	107	6B	k
12	C	[FORM FEED]	44	2C	,	76	4C	L	108	6C	l
13	D	[CARRIAGE RETURN]	45	2D	-	77	4D	M	109	6D	m
14	E	[SHIFT OUT]	46	2E	.	78	4E	N	110	6E	n
15	F	[SHIFT IN]	47	2F	/	79	4F	O	111	6F	o
16	10	[DATA LINK ESCAPE]	48	30	0	80	50	P	112	70	p
17	11	[DEVICE CONTROL 1]	49	31	1	81	51	Q	113	71	q
18	12	[DEVICE CONTROL 2]	50	32	2	82	52	R	114	72	r
19	13	[DEVICE CONTROL 3]	51	33	3	83	53	S	115	73	s
20	14	[DEVICE CONTROL 4]	52	34	4	84	54	T	116	74	t
21	15	[NEGATIVE ACKNOWLEDGE]	53	35	5	85	55	U	117	75	u
22	16	[SYNCHRONOUS IDLE]	54	36	6	86	56	V	118	76	v
23	17	[ENG OF TRANS. BLOCK]	55	37	7	87	57	W	119	77	w
24	18	[CANCEL]	56	38	8	88	58	X	120	78	x
25	19	[END OF MEDIUM]	57	39	9	89	59	Y	121	79	y
26	1A	[SUBSTITUTE]	58	3A	:	90	5A	Z	122	7A	z
27	1B	[ESCAPE]	59	3B	;	91	5B	[123	7B	{
28	1C	[FILE SEPARATOR]	60	3C	<	92	5C	\	124	7C	
29	1D	[GROUP SEPARATOR]	61	3D	=	93	5D]	125	7D	}
30	1E	[RECORD SEPARATOR]	62	3E	>	94	5E	^	126	7E	~
31	1F	[UNIT SEPARATOR]	63	3F	?	95	5F	_	127	7F	[DEL]

Розширена таблиця ASCII

– використовується національними системами шляхом завантаження потрібної кодової сторінки.

Windows-1251

– кодування, яке є стандартним 8-бітним кодуванням для всіх локалізованих українських і російських версій Microsoft Windows:

- ▶ містить практично усі символи слов'янської кириличної писемності для звичайного тексту;
- ▶ містить всі символи для російської, української, білоруської, сербської і болгарської мов.

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 І	194 В	210 Т	226 в	242 т
131 ґ	147 “	163 Ј	179 і	195 Г	211 У	227 г	243 у
132 ”	148 ”	164 ђ	180 ґ	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 Љ	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 ў	175 Ā	191 ī	207 П	223 Я	239 п	255 я

Універсальна система кодування UNICODE

- розроблена для уніфікації кодування текстових даних:
 - ▶ підтримує кодування **65 536** (2^{16}) різних символів;
 - ▶ містить символи більшості мов планети;
 - ▶ завдяки уніфікованості впроваджена у багатьох останніх технологіях, включаючи сучасні операційні системи, XML, Java (та інші мови програмування) та .NET Framework.

Дякую за увагу!

Далі буде...