LEBEGŐPONTOS SZÁMÁBRÁZOLÁS

A fixpontos operandusoknak azt a hátrányát, hogy az ábrázolás adott hossza miatt csak korlátozott nagyságú és csak egész számok ábrázolhatók, a lebegőpontos számábrázolás küszöböli ki.

Lebegőpontos ábrázolással a számtartomány lényegesen megnő és tört szám is ábrázolható. Ezért elsősorban műszaki-tudományos számításoknál használják.

Matematikai-, gépi normál alak:

Legyen a szám: 4253,56 tízes számrendszerben. (a számrendszer alapszámát radix-nak nevezzük)

Tízesre normalizálás esetén:

Matematikai normál alakja: $4253{,}56 = 4{,}25356 * 10^{3}$

Gépi normál alak: $4253{,}56 = 0{,}425356 * 10^{4}$

Általánosan: , ahol

- ’A’ az eredeti szám
- ’X’ az együttható, ennek törtrésze a mantissza (piros)
- ’B’ a hatvány alapja (a számrendszer alapszáma, általában 2, 8, 10, 16)
- ’y’ a hatvány kitevője, az un. ’karakterisztika’ (zöld)

Matematikai normál alaknál (10-es számrendszerben):

$0 <= X < 10 \qquad B = 10$

Gépi normál alaknál ((10-es számrendszerben):

$0 <= X < 1 \qquad B = 10$

A gépi alakban előnyös a 0 egészes kezdés, mert így nem kell ábrázolni a szám egész részét!

Kettesre normalizálás:

Legyen a tízes számrendszerbeli szám: $-91{,}125_{(10}$

1. Először a tízes számrendszerbeli számot átváltjuk kettes számrendszerbe.

$$\mathbf{-1011011{,}001_{(2} = 1*2^{6}+1*2^{4}+1*2^{3}+1*2+1+1*2^{-3} = 91{,}125_{(10}}$$

91	:2		0,	125	*2
45	**1**		**0**	250	
22	**1**		**0**	500	
11	**0**		**1**		
5	**1**				
2	**1**				
1	**0**				
0	**1**				

Addig osztunk, míg a hányados **0** nem lesz!

Lebegőpontos számábrázolás

Számrendszerek

2	8	10	16
000	0	**0**	**0**
001	1	**1**	1
010	2	**2**	2
011	3	**3**	3
100	4	**4**	4
101	5	**5**	5
110	6	**6**	6
111	7	**7**	7
1000	10	8	8
1001	11	9	9
1010	12	10	A
1011	13	11	B
1100	14	12	C
1101	15	13	D
1110	16	14	E
1111	17	15	F
10000	20	16	10

Mint látható egy 8-as számrendszerbeli szám ábrázolására 3 bit kell (**triád**)
Mint látható egy 16-os számrendszerbeli szám ábrázolására 4 bit kell (**tetrád**)

Így például: (ha kell elől feltöltjük értéktelen nullákkal)
$001101010111_{(2} = 855_{(10}$ triádokat hozunk létre a bináris vessző mindkét oldalán, és az oktális (8-as) alak leolvasható.

0	0	1	1	0	1	0	1	0	1	1	1
1			1+4=**5**			**2**			1+2+4=**7**		

$1101010111_{(2} = \mathbf{1527_{(8}} = 1*8^3+5*8^2+2*8+7=512+320+16+7=855_{(10}$

0	0	1	1	0	1	0	1	0	1	1	1
1+2=**3**				1+4=**5**				1+2+4=**7**			

$1101010111_{(2} = \mathbf{357}_{(16} = 3*16^2+5*16+7=\mathbf{855_{(10}}$

2. A számot normalizáljuk a megfelelő számrendszerben gépi normál alakra!
A $855_{(10}=0{,}855_{(10}*10^3$ gépi normál alakja:

Kettesre normalizálva:	$1101010111_{(2}$	$= 0{,}1101010111_{(2}*2^{10}$
Nyolcasra normalizálva:	$\mathbf{1527_{(8}}$	$\mathbf{= 0{,}1527_{(8}*8^4}$
Tizenhatosra normalizálva:	$\mathbf{357}_{(16}$	$= 0{,}357_{(16}*16^3$

3. Ábrázoljuk a **$855_{(10}$** számot 2-re normalizálva 32 biten (gépi szó) lebegőpontosan!

A karakterisztikát **(+10)**-t 7 pozíción feszített előjelesen ábrázoljuk.

1	0	0	0	0	1	0	+2 kitevő
1	0	0	0	0	0	1	+1 kitevő
1	0	0	0	0	0	0	**alapérték a 0 kitevőnek felel meg**
0	1	1	1	1	1	1	-1 kitevő
0	1	1	1	1	1	0	-2 kitevő

Tehát az 1000000-hoz vagy hozzáadjuk a kitevő 2-es számrendszerbeli alakját, vagy kivonjuk.

Jelen esetben: a 10 alakja $2+8=1010_{(2}$ és ezt kell hozzáadni az 1 000 000-hoz.

	1	0	0	0	0	0	0
				1	0	1	0
+	1	0	0	1	0	1	0

$$+\ 0,1101010111_{(2} * 2^{+10}$$

A lebegőpontos ábrázolási forma:

<table>
<tr><td>+/- 0/1</td><td>karakterisztika</td><td>mantissza (szám törtrésze)</td></tr>
<tr><td>A mantissza (szám) előjele 1 pozíció</td><td>előjeles kitevő ábrázolása 7 pozíció</td><td>24 pozíció

a bináris/hexadecimális vessző képzeletbeli helye</td></tr>
</table>

összesen 32 pozíció (egy gépi szó)

0	**1001010**	**1**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

A végeredményt 16-os számrendszerben szokták megadni.

<table>
<tr><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td colspan="4">4</td><td colspan="4">A</td><td colspan="4">D</td><td colspan="4">5</td><td colspan="4">C</td><td colspan="4">0</td><td colspan="4">0</td><td colspan="4">0</td></tr>
</table>

4. **Ábrázoljuk a $855_{(10}$** számot 2-re normalizálva 32 biten (gépi szó) lebegőpontosan!
A válasz a feladatra: a megoldás $4AD5C000_{(16}$

Feladat: Ábrázoljuk a következő gépi normál alakra hozott számot!

$$-\ 0,1101010111_{(2} * 2^{+10}$$

A lebegőpontos ábrázolási forma:

+/- 0/1	karakterisztika	mantissza (szám törtrésze)
A mantissza (szám) előjele 1 pozíció	előjeles kitevő ábrázolása 7 pozíció	24 pozíció a bináris/hexadecimális vessző képzeletbeli helye

összesen 32 pozíció (egy gépi szó)

1	**1001010**	**1**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

A végeredményt 16-os számrendszerben szokták megadni.

1	1	0	0	1	0	1	0	1	1	0	1	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C				A				D				5				C				0				0				0			

A válasz a feladatra: a megoldás $CAD5C000_{(16}$

Feladat: Ábrázoljuk a következő gépi normál alakra hozott számot!

$$+\ 0,1101010111_{(2} * 2^{-10}$$

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

A lebegőpontos ábrázolási forma:

+/- 0/1	karakterisztika	mantissza (szám törtrésze)
A mantissza (szám) előjele 1 pozíció	előjeles kitevő ábrázolása 7 pozíció	24 pozíció a bináris/hexadecimális vessző képzeletbeli helye

összesen 32 pozíció (egy gépi szó)

0	**0110110**	**1**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

A végeredményt 16-os számrendszerben szokták megadni.

0	0	1	1	0	1	1	0	1	1	0	1	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3				6				D				5				C				0				0				0			

A válasz a feladatra: a megoldás $36D5C000_{(16}$

Feladat: Ábrázoljuk a következő gépi normál alakra hozott számot!

$$-\,0,1101010111_{(2}*2^{-10}$$

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

A lebegőpontos ábrázolási forma:

+/- 0/1	**karakterisztika**	mantissza (szám törtrésze)
A mantissza (szám) előjele 1 pozíció	**előjeles kitevő ábrázolása 7 pozíció**	24 pozíció **a bináris/hexadecimális vessző képzeletbeli helye**

összesen 32 pozíció (egy gépi szó)

1	**0110110**	**1**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**1**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

A végeredményt 16-os számrendszerben szokták megadni.

<table>
<tr><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td colspan="4">B</td><td colspan="4">6</td><td colspan="4">D</td><td colspan="4">5</td><td colspan="4">C</td><td colspan="4">0</td><td colspan="4">0</td><td colspan="4">0</td></tr>
</table>

A válasz a feladatra: a megoldás $B6D5C000_{(16}$

Ha 64 biten (dupla szó) vagy félszón (16 bit) kell ábrázolni, akkor a megoldás teljesen ez csak a hexadecimális alak duplájára, ill. felére csökken.

Abban az esetben, ha nem 2-re normalizálva, hanem 16-ra normalizálva kell ábrázolni akkor teljesen hasonló a megoldás, csak a 16-ra normalizált alakkal kell dolgozni!

Ha valakinek kérdése van, nagyon szívesen válaszolok e-mailben.

Lebegőpontos számábrázolás