

Veri Yapıları ve Algoritmalar

Verilerin Bellekte Gösterimi

Dr. Öğr. Üyesi M. Ozan AKI

30.10.2020

Bilgi Büyüklüğünün Ölçülmesi

Tüm sayısal sistemlerde bilgi temel bilgi boyutu **BYTE** tır. BYTE'ı aşan veriler için **WORD** ifadesi kullanılır.

Bir Byte, **8** bitten oluşur. Gerektiğinde üst katları kullanılır.
Byte – Kilo Byte – Mega Byte – Giga Byte – Tera Byte

Üst katlar gerçekte 1024 kat artar. Bazı işletim sistemleri kolaylık bakımından 1000 kat olarak alır. Bu durumda gerçek boyut ile görünen boyut arasında fark oluşur.

Örn: 2GB Bir bellekte;

1000 kat hesabı ile; $2 * 1000 * 1000 * 1000 = 2,000,000,000$ Byte,

1024 kat hesabı ile; $2 * 1024 * 1024 * 1024 = 2,147,483,648$ Byte

Bululunur.

Gerçek kapasite 147MB daha büyüktür.

Verilerin Bellekte Gösterimi

Bellekte tutulan tüm bilgiler sadece bit dizileri şeklindedir.

Bu bit dizileri, sistematik bir şekilde yorumlanarak anlamlı bilgilere dönüştürülür.

Örneğin **0 1 0 0 0 0 1** bit dizisi için;

Tamsayı -> 65

Hexadecimal -> 41

BCD -> 41

ASCII -> A

Benzer şekilde, bilgisayar ortamında bulunan tüm veriler (döküman, resim, ses, video, v.s.) tanımlı veri yapılarından türetilmiş ikili kodlar şeklindedir.

Temel Veri Yapıları

Hemen her derleyicide temel veri yapıları küçük farklılıklarla birlikte aynıdır.

Bazı derleyiciler, kendine özgü veri yapıları içerebilir. Eğer algoritmaya uygun veri yapıları bulunmuyor ise bunu kullanıcı kendisi tanımlamalıdır.

- Karakter
- İşaretli / İşaretsiz Tamsayılar
- Gerçel Sayılar

Karakterlerin Kodlanması

Çoğu derleyicilerde bulunan temel veri yapısıdır.

char 1 Byte (8 Bit) uzunluğundadır.

ASCII tablosundaki karakterlerin tanımlanması amacıyla kullanılır.

Ancak bunun yanında 8 bit tamsayı olarak ta işlemlere katılabilir.

8 Bit (256 karakter), Standart ASCII tablosunun tümünü adresler. (*ALT + Keypad ASCII Code*)

ASCII (American Standard Code for Information Interchange)

0	<NUL>	32	<SPC>	64	@	96	`	128	Ä	160	†	192	¿	224	‡
1	<SOH>	33	!	65	A	97	a	129	Å	161	°	193	¡	225	·
2	<STX>	34	"	66	B	98	b	130	Ç	162	¢	194	¬	226	,
3	<ETX>	35	#	67	C	99	c	131	É	163	£	195	√	227	"
4	<EOT>	36	\$	68	D	100	d	132	Ñ	164	§	196	f	228	‰
5	<ENQ>	37	%	69	E	101	e	133	Ö	165	•	197	≈	229	Â
6	<ACK>	38	&	70	F	102	f	134	Ü	166	¶	198	Δ	230	Ê
7	<BEL>	39	'	71	G	103	g	135	á	167	β	199	«	231	Á
8	<BS>	40	(72	H	104	h	136	à	168	®	200	»	232	Ë
9	<TAB>	41)	73	I	105	i	137	â	169	©	201	...	233	È
10	<LF>	42	*	74	J	106	j	138	ä	170	™	202		234	Í
11	<VT>	43	+	75	K	107	k	139	ã	171	'	203	À	235	Î
12	<FF>	44	,	76	L	108	l	140	å	172	¨	204	Ã	236	Ï
13	<CR>	45	-	77	M	109	m	141	ç	173	≠	205	Õ	237	Ì
14	<SO>	46	.	78	N	110	n	142	é	174	Æ	206	Œ	238	Ó
15	<SI>	47	/	79	O	111	o	143	è	175	Ø	207	œ	239	Ô
16	<DLE>	48	0	80	P	112	p	144	ê	176	∞	208	–	240	🍏
17	<DC1>	49	1	81	Q	113	q	145	ë	177	±	209	—	241	Ò
18	<DC2>	50	2	82	R	114	r	146	í	178	≤	210	"	242	Ú
19	<DC3>	51	3	83	S	115	s	147	ì	179	≥	211	"	243	Û
20	<DC4>	52	4	84	T	116	t	148	î	180	¥	212	`	244	Ü
21	<NAK>	53	5	85	U	117	u	149	ï	181	μ	213	'	245	ı
22	<SYN>	54	6	86	V	118	v	150	ñ	182	ð	214	÷	246	ˆ
23	<ETB>	55	7	87	W	119	w	151	ó	183	Σ	215	◇	247	˜
24	<CAN>	56	8	88	X	120	x	152	ò	184	Π	216	ÿ	248	—
25		57	9	89	Y	121	y	153	ô	185	π	217	Ÿ	249	˘
26	<SUB>	58	:	90	Z	122	z	154	ö	186	∫	218	/	250	˙
27	<ESC>	59	;	91	[123	{	155	õ	187	ª	219	€	251	°
28	<FS>	60	<	92	\	124		156	ú	188	º	220	<	252	¸
29	<GS>	61	=	93]	125	}	157	ù	189	Ω	221	>	253	”
30	<RS>	62	>	94	^	126	~	158	û	190	æ	222	fi	254	˚
31	<US>	63	?	95	_	127		159	ü	191	ø	223	fl	255	ˇ



UNICODE Karakter Kodlaması

Ancak günümüzde farklı alfabelerin ve kullanılan özel karakterlerin tanımlarını kapsamak üzere geliştirilen bu kodlama sistemi 16 bitlik bir tabloya tanımlar.

Toplam 65535 karakter içeren bu tablo **UTF-8** gibi kodlama sistemleri kullanılarak kaydedilir.

Böylece, evrensel bir kod tablosu ile, uluslar arası yazılımların kolayca yazılması ve uyumluluk sorunlarının ortadan kaldırılması sağlanmıştır.

Uygulama: Aynı bilgiyi içeren UNICODE ve ASCII metin dosyalarını kaydederek boyutlarını karşılaştırınız. (<https://unicode-table.com/>)

İşaretsiz Tam Sayıların Kodlaması

Tamsayılar, işaretli ya da işaretsiz olarak tanımlanırlar.

İşaretsiz tamsayılar, bellekte **doğal** karşılığı (ikilik taban karşılığı) ile saklanırlar.

Bu sayılar, derleyicinin varsayılan bit sayısı (8, 16, 32, 64) ile sınırlıdırlar.

Örneğin; **39** Sayısı, 8 bit genişliğindeki bellekte **0 0 1 0 0 1 1 1** olarak saklanır.

Eğer bellek hücresinin bit genişliği daha fazla ise, üst bitler 0 (sıfır) ile doldurulur.

İşaretli Tam Sayıların Kodlaması

İşaretli Tamsayılarda durum biraz farklılaşır;

Aynı bit sayısına sahip bir alanda kodlanacak işaretli sayı ile işaretsiz sayının maksimum ve minimum değerleri farklı olur.

Neden ?

Çünkü işaretli sayının değer bitlerinden biri işaret biti olarak kullanılmak zorundadır.

Bu nedenle,

İşaretsiz bir tamsayı, $0 < n < 2^n - 1$ aralığında iken,

İşaretli bir tamsayı, $-2^{n-1} < n < 2^{n-1} - 1$ aralığında olur.



İşaretli Tam Sayıların Kodlaması

İşaretli Tamsayılar iki yöntemle kodlanabilir;

- 1) Sayının en solundaki bit işaret biti kabul edilir ve negatif sayılar için bu bir 1, pozitif sayılar için 0 yapılır.
- 2) İşaret biti için özel bir alan ayrılmaz, negatif sayılar 2'ye tümleyen şeklinde saklanır.



Tam Sayıların Kapladıkları Alanlar

short int	-32,768 -> +32,767 (16 bit)
unsigned short int	0 -> +65,535 (16 bit)
unsigned int	0 -> +4,294,967,295 (32 bit)
int	-2,147,483,648 -> +2,147,483,647 (32 bit)
long int	-2,147,483,648 -> +2,147,483,647 (32 bit)
signed char	-128 -> +127 (8 bit)
unsigned char	0 -> +255 (8 bit)

Uygulama: C ile tanımlanan bir tamsayı değişkeninin maksimum ve minimum değerlerini gösteren c programını yazınız.

Gerçek (Kesirli) Sayılar

İki tip gösterim şekli vardır;

- Sabit Noktalı (Fixed Point)
- Kayan Noktalı (Floating Point)

Bilgisayar uygulamalarında büyük çoğunlukla Kayan Noktalı sayılar kullanılmaktadır.

Ancak kısıtlı bellek ve işlemci gücüne sahip gömülü sistemlerde sabit noktalı sayılar tercih edilir.

Gerçek (Kesirli) Sayılar

Kayan Noktalı (Floating Point) Sayı Formatı (IEEE 754)

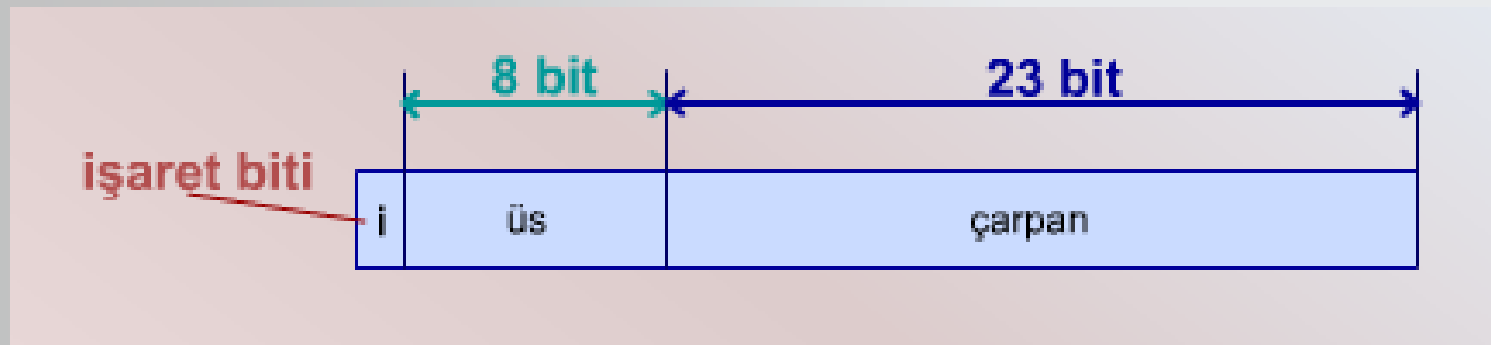
Bu formatta;

1 bit : İşaret biti (sign)

8 bit : Üs (exponent) 2'ye Tümleyen şeklinde bulunur.

23 bit : Çarpan (fraction)

$$\text{Sayı} = \pm \text{Ç} \times T^{\pm \text{Ü}} \quad (T:\text{Taban}=2)$$



The diagram illustrates the IEEE 754 Single Precision Floating Point Format. It consists of a 32-bit structure divided into three main fields:

- Sign (S):** 1 bit (bit 31).
- Exponent (E):** 8 bits (bits 30-23).
- Mantissa (M):** 23 bits (bits 22-0).

Below the diagram, the bit fields are expanded:

- Sign:** 0
- Exponent:** 1 0 0 0 0 1 1 0
- Mantissa:** 0 1 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

$$1 \quad * \quad 10^{111} \quad * \quad 1,010000011$$

10100000,11

160,75



Gerçek Sayıların Çözümlemesi

Örn; 2.5 sayısını IEEE 754 formatında gösteriniz.

$$2.5 = (-1)^i \cdot (1 + k) \cdot 2^{\text{üs}-\text{bias}}$$

Burada;

i : sayının işareti (0:Pozitif ya da 1:Negatif)

$(1+k)$: Çarpan ($1 \leq \text{Ç} < 2$, $k < 0$)

bias : sabit (32 bit formda: 127, 64 bit formda:1023)

$$i=0$$

$$2.5 / 2^n = (1+k) \quad n=1 \text{ için, } 2.5 / 2 = (1 + k), k = 0.25$$

Böylece, $2.5 = \text{IEEE 754 formatında } 1.25 \cdot 2^1$ olarak gösterilir.



Gerçek Sayıların Çözümlemesi

Örn; 3.14 Pi sayısını IEEE 754 formatında gösteriniz.

$$3.14 = (-1)^i \cdot (1 + k) \cdot 2^{\text{üs}-\text{bias}}$$

Burada;

i : sayının işareti (0:Pozitif ya da 1:Negatif)

$(1+k)$: Çarpan ($1 \leq \text{Ç} < 2$, $k < 0$)

bias : sabit (32 bit formda: 127, 64 bit formda:1023)

$i=0$

$$3.14 / 2^n = (1+k) \quad n=1 \text{ için, } 3.14 / 2 = (1 + k), k = 0.57$$

Böylece, $3.14 = \text{IEEE 754 formatında } 1.57 \cdot 2^1 \text{ olarak gösterilir.}$

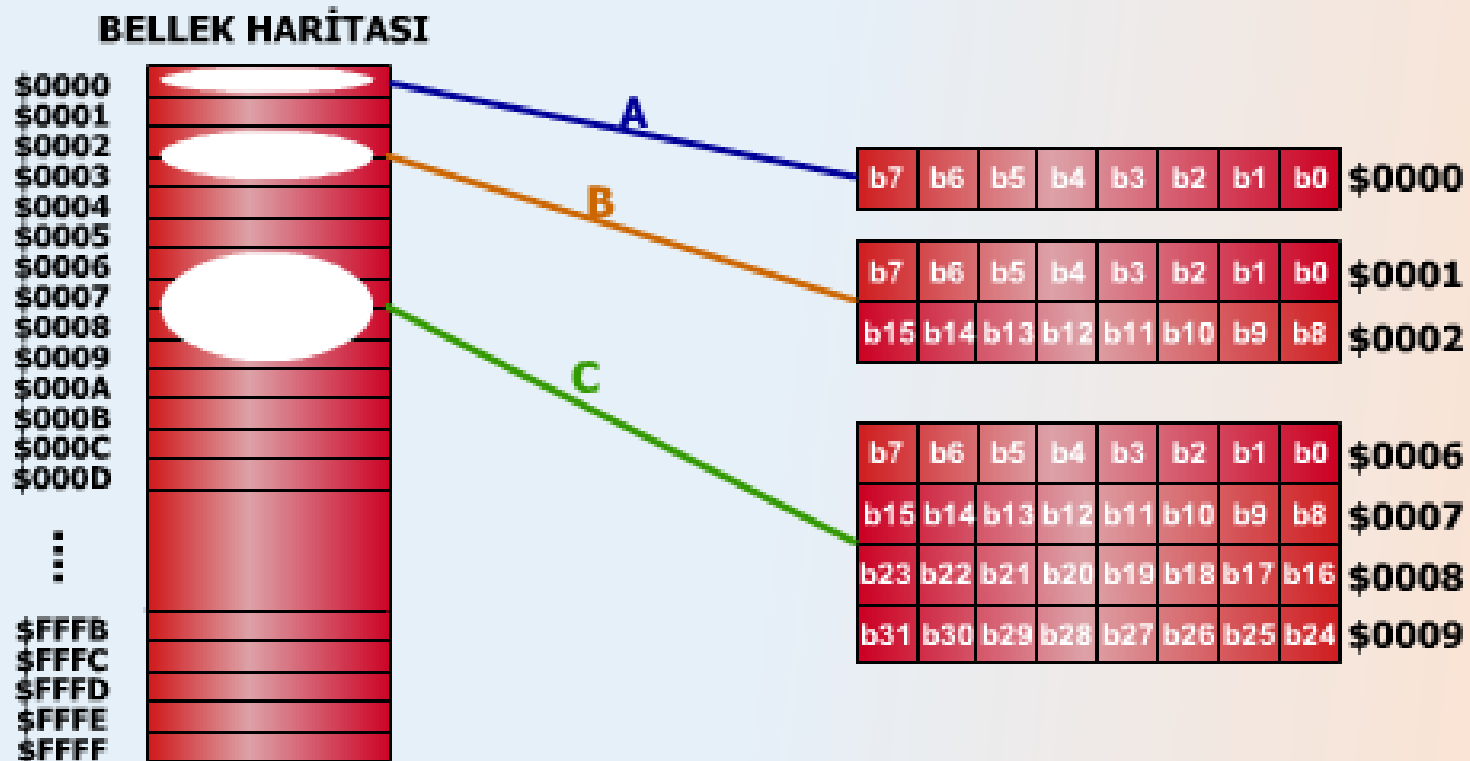
Gerçek (Kesirli) Sayıların Kapladığı Alan

Parametre	32 bitlik kayan nokta	64 bitlik kayan nokta
İşaret bit'i	0, artı; 1 eksi sayı	0, artı; 1 eksi sayı
Çarpan uzunluğu (bit)	23	52
Üs uzunluğu (bit)	8	11
Üssün tabanı	2 tabanı	2 tabanı
Üssün en büyük değeri	127	1023
Üssün en küçük değeri	-126	-1022
10 tabanında en küçük üs değeri	10^{-38}	10^{-308}
10 tabanında en büyük üs değeri	10^{38}	10^{308}
Kesir sayısı/genişliği	$2^{23} = 8$ Mega	$2^{52} = 2048$ Tera
Sıfır sayısı için	$\dot{U}=0, \zeta=0$	$\dot{U}=0, \zeta=0$
Sonsuzun gösterilim için	$\dot{U}=255, \zeta=0$	$\dot{U}=2047, \zeta=0$



Bellek ile Veri Yapısı Bit Sayısı

Eğer veri yapısını bit sayısı, belleğin tek adres hücresinde saklanamayacak kadar büyük ise, bu veri yapısı, ardaşık bellek alanlarını kullanır.





Little-Endian & Big-Endian

Eğer bir veri birden fazla bellek alanı kaplayacaksa, bu bilgi nasıl yazılıp okunmalıdır? *Rafadan yumurtanızı nasıl yersiniz?*

