

25.01.2022

PATİKA – INNOVA JAVA SPRİNG ODEV – 3

Bekir Gürkan GÜLDAŞ

İçindekiler

1. ASCII ve Extended ASCII	3
2. UNICODE	5
3. JAR	8
4. WAR	9
5. Absolute Path	10

1. ASCII ve Extended ASCII

ASCII (American Standard Code for Information Interchange) Latin alfabesi üzerine kurulu 7 bitlik bir karakter kümesidir. Harflerin bilgisayar ortamında saklanması ve taşınması ile ilgili geliştirilen ilk sistem ASCII sistemi olmuştur.

Bu sistemde Amerikan İngilizcesi alfabesinde bulunan harflerin her birine bir sayı atanmış ve harfler bilgisayar ortamında sayı olarak saklanmıştır. Harflerin kodlanması için 7 bitlik baytlar kullanılmıştır.

Dec	Hex	Oct	Chr	Dec	Hex	Oct	HTML	Chr	Dec	Hex	Oct	HTML	Chr	Dec	Hex	Oct	HTML	Chr
0	0	000	NULL	32	20	040	 	Space	64	40	100	@	@	96	60	140	`	`
1	1	001	SoH	33	21	041	!	!	65	41	101	A	A	97	61	141	a	a
2	2	002	SoTxt	34	22	042	"	"	66	42	102	B	B	98	62	142	b	b
3	3	003	SoT	35	23	043	#	#	67	43	103	C	C	99	63	143	c	c
4	4	004	SoT	36	24	044	$	\$	68	44	104	D	D	100	64	144	d	d
5	5	005	Enq	37	25	045	%	%	69	45	105	E	E	101	65	145	e	e
6	6	006	Ack	38	26	046	&	&	70	46	106	F	F	102	66	146	f	f
7	7	007	Bell	39	27	047	'	'	71	47	107	G	G	103	67	147	g	g
8	8	010	Bsp	40	28	050	((72	48	110	H	H	104	68	150	h	h
9	9	011	HTab	41	29	051))	73	49	111	I	I	105	69	151	i	i
10	A	012	LFeed	42	2A	052	*	*	74	4A	112	J	J	106	6A	152	j	j
11	B	013	VTab	43	2B	053	+	+	75	4B	113	K	K	107	6B	153	k	k
12	C	014	FFeed	44	2C	054	,	,	76	4C	114	L	L	108	6C	154	l	l
13	D	015	CR	45	2D	055	-	-	77	4D	115	M	M	109	6D	155	m	m
14	E	016	SOOut	46	2E	056	.	.	78	4E	116	N	N	110	6E	156	n	n
15	F	017	SIn	47	2F	057	/	/	79	4F	117	O	O	111	6F	157	o	o
16	10	020	DLE	48	30	060	0	0	80	50	120	P	P	112	70	160	p	p
17	11	021	DC1	49	31	061	1	1	81	51	121	Q	Q	113	71	161	q	q
18	12	022	DC2	50	32	062	2	2	82	52	122	R	R	114	72	162	r	r
19	13	023	DC3	51	33	063	3	3	83	53	123	S	S	115	73	163	s	s
20	14	024	DC4	52	34	064	4	4	84	54	124	T	T	116	74	164	t	t
21	15	025	Nack	53	35	065	5	5	85	55	125	U	U	117	75	165	u	u
22	16	026	Syn	54	36	066	6	6	86	56	126	V	V	118	76	166	v	v
23	17	027	SoTB	55	37	067	7	7	87	57	127	W	W	119	77	167	w	w
24	18	030	Can	56	38	070	8	8	88	58	130	X	X	120	78	170	x	x
25	19	031	SoM	57	39	071	9	9	89	59	131	Y	Y	121	79	171	y	y
26	1A	032	Sub	58	3A	072	:	:	90	5A	132	Z	Z	122	7A	172	z	z
27	1B	033	Esc	59	3B	073	;	;	91	5B	133	[[123	7B	173	{	{
28	1C	034	FSep	60	3C	074	<	<	92	5C	134	\	\	124	7C	174	|	
29	1D	035	GSep	61	3D	075	=	=	93	5D	135]]	125	7D	175	}	}
30	1E	036	RSep	62	3E	076	>	>	94	5E	136	^	^	126	7E	176	~	~
31	1F	037	USep	63	3F	077	?	?	95	5F	137	_	_	127	7F	177		Delete

1. ASCII ve Extended ASCII

ASCII kodlaması kullanılmaya başladıktan sonra ü, ö veya ç gibi harflerin de kodlanması ihtiyacı ortaya çıkmış Genişletilmiş ASCII denilen yeni bir sistem kullanılmıştır. Bu sistemde harfleri kodlamak için kullanılan yedi bite bir bit daha eklenerek 8 bitlik baytlar kullanılmıştır.

İlk çıkarılan genişletilmiş ASCII kümesi olan ISO 8859-1 (Latin-1), Latin alfabesinde kullanılan harflerin büyük çoğunluğu karşılamaktadır. Daha sonra farklı dillerin alfabelerinde bulunan o dile özgü karakterler için farklı genişletilmiş ASCII kümeleri çıkarılmıştır. Bu farklı kümelere Code Page denilmektedir.

DisplayName	Name	CodePage
IBM EBCDIC (US-Canada)	IBM037	37
OEM United States	IBM437	437
IBM EBCDIC (International)	IBM500	500
Arabic (ASMO 708)	ASMO-708	708
Arabic (DOS)	DOS-720	720
Greek (DOS)	ibm737	737
Baltic (DOS)	ibm775	775
Western European (DOS)	ibm850	850
Central European (DOS)	ibm852	852
OEM Cyrillic	IBM855	855
Turkish (DOS)	ibm857	857
OEM Multilingual Latin I	IBM00858	858
Portuguese (DOS)	IBM860	860
Icelandic (DOS)	ibm861	861
Hebrew (DOS)	DOS-862	862
French Canadian (DOS)	IBM863	863
Arabic (864)	IBM864	864
Nordic (DOS)	IBM865	865
Cyrillic (DOS)	cp866	866
Greek, Modern (DOS)	ibm869	869
IBM EBCDIC (Multilingual Latin-2)	IBM870	870
Thai (Windows)	windows-874	874
IBM EBCDIC (Greek Modern)	cp875	875
Japanese (Shift-JIS)	shift_jis	932
Chinese Simplified (GB2312)	gb2312	936
Korean	ks_c_5601-1987	949
Chinese Traditional (Big5)	big5	950
IBM EBCDIC (Turkish Latin-5)	IBM1026	1026

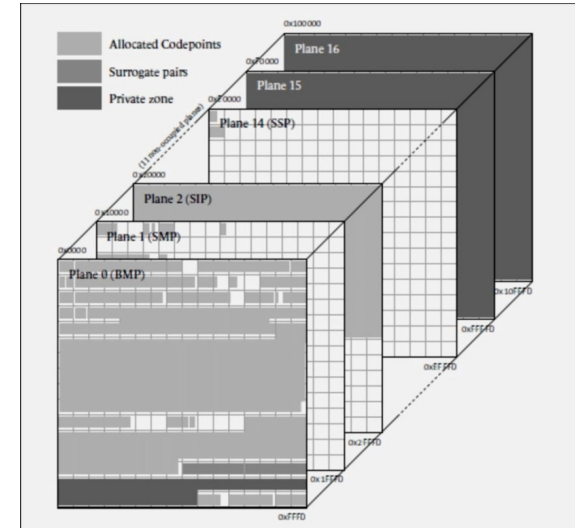
2. UNICODE

CodePage'ler arasında ki karışıklılıkla birlikte binlerce farklı karakter barındıran Çince veya Japonca gibi dillerin harfleri için 8 bitlik karakter kapasitesi yeterli olamamıştır. Bu yüzden ASCII sistemi yerini Unicode'a(Evrensel Kod) bırakmıştır. Sistemin amacı farklı karakter kodlama sistemlerinin birbiriyle tutarlı çalışmasını ve dünyadaki tüm yazım sistemlerinden metinlerin bilgisayar ortamında tek bir standart altında temsil edilebilmesini sağlamaktır.

2. UNICODE

Unicode, dünyadaki bütün yaşayan dillerin tüm karakterlerini kapsayabilmesi amacıyla uzunluğu 16 bite çıkarılmış "Wide-Body ASCII" olarak tanımlanabilir. Aynı zamanda Unicode on yedi adet düzleme(plane) ayrılmıştır. İlk düzlem olan Temel Çokdilli Düzlem (Basic Multilingual Plane, BMP) düzlemindeki karakterlerin kod noktaları U+00000 - U+0FFFF şeklinde yazılır. Sıfıncı düzlem olduğundan U+0000 - U+FFFF şeklinde de belirtilebilir.

V·T·E		Unicode planes and used code point ranges								[hide]
Basic		Supplementary								
Plane 0		Plane 1		Plane 2		Planes 3–13	Plane 14	Planes 15–16		
0000–FFFF		10000–1FFFF		20000–2FFFF		30000–DFFFF	E0000–EFFFF	F0000–10FFFF		
Basic Multilingual Plane		Supplementary Multilingual Plane		Supplementary Ideographic Plane		unassigned	Supplement-ary Special-purpose Plane	Supplement-ary Private Use Area planes		
BMP		SMP		SIP		—	SSP	SPUA-A/B		
0000–0FFF	8000–8FFF	10000–10FFF		20000–20FFF	28000–28FFF		E0000–E0FFF	15: SPUA-A		
1000–1FFF	9000–9FFF	11000–11FFF	1B000–1BFFF	21FFF–21FFF	29000–29FFF			F0000–FFFFF		
2000–2FFF	A000–AFFF	12000–12FFF		22000–22FFF	2A000–2AFFF			16: SPUA-B		
3000–3FFF	B000–BFFF	13000–13FFF	1D000–1DFFF	23FFF–23FFF	2B000–2BFFF			100000–10FFFF		
4000–4FFF	C000–CFFF	14000–14FFF	1E000–1EFFF	24000–24FFF	2C000–2CFFF					
5000–5FFF	D000–DFFF	15000–15FFF	1F000–1FFFF	25000–25FFF	2D000–2DFFF					
6000–6FFF	E000–EFFF	16000–16FFF		26000–26FFF	2E000–2EFFF					
7000–7FFF	F000–FFFF			27000–27FFF	2F000–2FFFF					



2. UNICODE

Her düzlem, blok adı verilen bölümlere ayrılmıştır ve her blokta o blokla ilgili karakterler bulunur. Blokların büyüklüğü değişken olmakla birlikte bu sayı her zaman 16'nın ve genelde de 128'in katıdır. Aynı yazı türünde bulunan karakterler farklı bloklara dağılmış olabilir

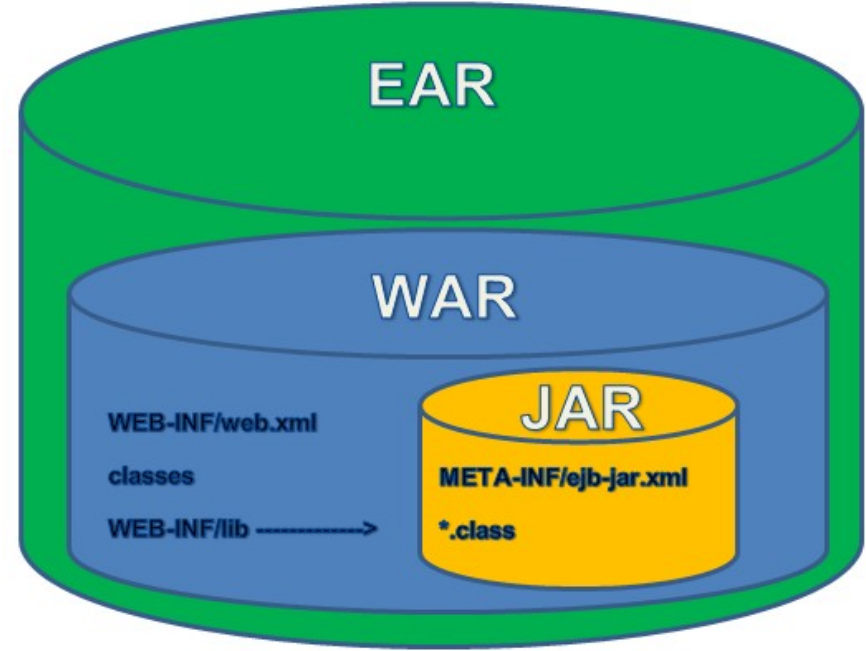
Unicode'un temel farkı genişletilmiş ASCII kümelerinin karakterlere atamak için kullandığı $[0, 255]$ olan sayı aralığını $[0, 1.114.111]$ 'e çıkarmasında yatmaktadır. Unicode sisteminde 255'ten büyük sayılar iki tabanında 8 biti aşmaktadır ve eğer 1.114.111 sayısının hepsi eşit uzunlukta olacak şekilde doğrudan iki tabanına dönüştürülüp kullanılmak istenirse her karakterin 32 bit uzunluğunda olması, iki tabanına dönüştürüldüğünde 32 bitten kısa olan sayıların da başına sıfır eklenerek 32 bite tamamlanması gerekir. Çünkü her karakterin uzunluğu sabit olmazsa kodlamada bir karakterin nerede bitip diğerinin nerede başladığı anlaşılamaz.

3. JAR

JAR dosyası, Java Runtime Environment tarafından Java dosyalarını, metin, resimler, vb. dosyaları ve bunlara bağlı meta verileri saklayan bir paket dosya formatı, arşiv dosyasıdır. İçinde birden çok dosya barındırdığı için ZIP formatında olan JAR dosyası, içinde Java'ya özgü bildirim dosyaları taşır. Java Runtime Environment tarafından kullanılmak üzere tasarlanmış bir formattır.

4. WAR

WAR dosyası, bir web uygulamasının tüm bölümlerini içeren bir arşivdir. Servlet sınıfları, web servisleri, JSP'ler, HTML sayfaları, resimler ve diğer kaynaklar war dosyası kapsamındadır. WAR dosyası, Java kodu için belirtilen dizinlere ve belirlenmiş bir yapılandırma dosyasına sahip bir JAR dosyasıdır. Uygulama sunucusuna neyi çalıştıracağını ve nasıl çalıştıracağını söyleyen web.xml dosyası mevcuttur. WAR dosyaları her zaman .war uzantısına sahiptir, ancak standart jar araçlarıyla oluşturulup okunabilirler.



5. Absolute Path

Path bir işletim sisteminde bir dosya yada klasöre verilen özel bir lokasyondur. Path bir dosya yolunun alfa sayısal karakterlerin birleşiminden oluşur. Absolute path ise bir dosya yada klasörün root(kök) dizinden itibaren verilen path'e denir. Root (/) dizininden itibaren alt klasörlere erişilebilir.

Absolute Path:

```
C:\Windows\system32\MessagingService.dll
```

From C:\Windows\schemas\ the Relative Path is:

```
..\system32\MessagingService.dll
```

From C:\Windows\system32\drivers\ the Relative Path is

```
..\MessagingService.dll
```