

තොරතුරු හා සන්නිවේදන
තාක්ෂණය

අ.පො.ස (සා පෙළ)

රැවින් විජේරත්න

පරිගණකයෙහි අක්ෂර , සංවිධාන සහ
සංකේත නිරූපණය කරන ආකාරය

Education NCP - www.educationcp.lk

ඩිජිටල් පගණක හා සම්බන්ධ ද්වීමය , දශමය , අෂ්ටක හා ඡඩි දශමක සංඛ්‍යා පද්ධති ගැන හා එම සංඛ්‍යා පද්ධති අතර පරිවර්ථන ගැන ඉහත දී අප අධ්‍යනය කලෙමු. ද්වීමය සංඛ්‍යා පද්ධති මගින් නිරූපනය වන්නා වූ 1 හෝ 0 යනු **Bit 01** ලෙස හැඳින්විය හැක. එවැනි **Bit 08** ක් එකතු වී **Byte 01** ක් හෙවත් පරිගණකයෙහි අක්ෂරයක්, සංඛ්‍යාවක් නිර්මාණය කරන්නා වූ ඒකකය තැනේ. මෙවැනි **Byte 01** ක් කොටස් දෙකකට බෙදූ විට **Bit 04** බැගින් වූ කොටස් දෙකකට බෙදේ. එසේ බෙදෙන **Bit 04** ක කොටසක් **Nibble 01** ක් ලෙස හඳුන්වයි.



ද්වීමය කේතක දශම (Binary Coded Decimal)

මෙහිදී ඉලක්කම් නියෝජනය කිරීම සඳහා **Nibble 01** ක් හෙවත් **Bit 04** ක් භාවිතා වේ. 0 සිට 9 දක්වා වූ සංඛ්‍යා මෙසේ **BCD** ක්‍රමයට අනුව

1 – 0001	2 – 0010	3 – 0011	4 – 0100
5 – 0101	6 – 0110	7 – 0111	8 – 1000
9 – 1001			

ලෙස ලිවිය හැක. ඉන් අනතුරුව ලියන වෙනත් අගයන් පහත ලෙස දක්වයි.

Education
NCP
www.education.ncp.lk

1	8	6	5
0001	1000	0110	0101

නමුත් සැබැවින්ම පරිගණකය තුළ දත්ත ගබඩා කිරීමේ දී ඉලක්කමක් හෝ අක්ෂරයක් නිරූපනය සඳහා **Byte 01** ක් යොදා ගනී. එවිට **BCD** ක්‍රමයේ දී භාවිතා වන්නේ නිබලයක් බැවින් සෑම **Byte** එකකින්ම ඉදිරි නිබලය අපතේ යයි.

1	8	6	5				
0000	0001	0000	1000	0000	0110	0000	0101

මෙමගින් අනු ලක්ෂණ 16 ක් නිරූපණය කළ හැකි වුවත් 0-9 දක්වා මෙය භාවිතා කළ ද 10-15 දක්වා සංඛ්‍යා නියෝජනයට මෙය භාවිතා නොවේ. එබැවින් මෙය අතීතයේ භාවිතා වුවත් වර්ථමානයේ භාවිතා නොවන තරමි ය.

1268₁₀ BCD බවට පරිවර්ථනය කරන්න.

1268₁₀ - **0001 0010 0110 1000**_{BCD}

විස්තෘත ද්වීමය කේතක දැශම හුවමාරු කේත (Extended Standard Code For Information Interchange-EBCDIC)

මෙම කේත ක්‍රමයට අනුව අනුලක්ෂණයක් නියෝජනය කිරීම සඳහා බිටු 8 ක් වෙන් කර ගනී. එනම් මින් අනු ලක්ෂණ 256 ($2^8 = 256$) ක් නියෝජනය වේ. (වගුව 3.7) මෙම කේත ක්‍රමය IBM Main Frame පරිගණකවල භාවිත කරයි.

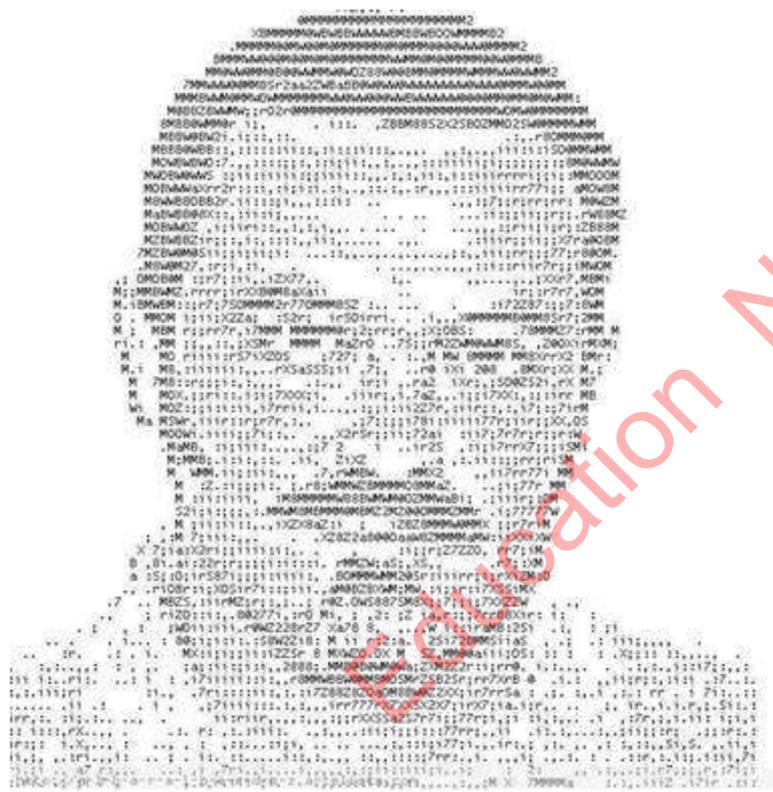
Character	Hexadecimal	EBCDIC Code	
0	F0	1111	0000
1	F1	1111	0001
2	F2	1111	0010
3	F3	1111	0011
4	F4	1111	0100
5	F5	1111	0101
6	F6	1111	0110
7	F7	1111	0111
8	F8	1111	1000
9	F9	1111	1001
A	C1	1100	0001

Character	Hexadecimal	EBCDIC Code	
A	C1	1100	0001
B	C2	1100	0010
C	C3	1100	0011
D	C4	1100	0100
E	C5	1100	0101
F	C6	1100	0110
G	C7	1100	0111
H	C8	1100	1000
I	C9	1100	1001
J	D1	1101	0001
K	D2	1101	0010
L	D3	1101	0011
M	D4	1101	0100
N	D5	1101	0101
O	D6	1101	0110
P	D7	1101	0111
Q	D8	1101	1000
R	D9	1101	1001
S	E2	1110	0010
T	E3	1110	0011
U	E4	1110	0100
V	E5	1110	0101
W	E6	1110	0110
X	E7	1110	0111
Y	E8	1110	1000
Z	E9	1110	1001

අෆස්කි (ASCII) කේත ක්‍රමය

දත්ත ලෙස පරිගණකයට ඇතුළු කරනු ලබන්නේ සංඛ්‍යා පමණක් නොවේ. අකුරු, සංකේත ආදිය ද දත්ත ලෙස පරිගණකයට ලබා දිය යුතු ය. මෙම දත්ත ද්විමය සංඛ්‍යා ඛවට පරිවර්තනය කිරීමේ දී ජාත්‍යන්තර වශයෙන් පිළිගනු ලැබූ සම්මතයකට අනුව එය සිදුකෙරේ. එම සම්මතය අෆස්කි (ASCII - American Standard Code for Information Interchange) කේත ක්‍රමය ලෙස හැඳින්වේ. මෙම කේත ක්‍රමයට අනුව ඔහුම අකුරකට හෝ සංකේතයකට අදාළ සම්මත දශමය අගයක් ඇත. එය අෆස්කි (ASCII) අගය ලෙස හැඳින්වේ. මෙම අෆස්කි ක්‍රමය මගින් අනුලක්ෂණයක් (Characters) නියෝජනය කිරීම සඳහා බිටු 7ක් වෙන් කර ඇත.

මේ මගින් අනු ලබන්නා 128 ක් නියෝජනය කළ හැකි ය.
මෙම ඇස්ති කේත ක්‍රමය පිළිබඳ මුල් ම යෝජනාව ගෙන එන
ලද්දේ ඇන්සි (ANSI - American National Standard
Institute) ආයතනයයි.



Education NCP - www.ednbp.lk

ASCII Table

Dec	Hex	Oct	Char	Dec	Hex	Oct	Char	Dec	Hex	Oct	Char	Dec	Hex	Oct	Char
0	0	0		32	20	40	[space]	64	40	100	@	96	60	140	`
1	1	1		33	21	41	!	65	41	101	A	97	61	141	a
2	2	2		34	22	42	"	66	42	102	B	98	62	142	b
3	3	3		35	23	43	#	67	43	103	C	99	63	143	c
4	4	4		36	24	44	\$	68	44	104	D	100	64	144	d
5	5	5		37	25	45	%	69	45	105	E	101	65	145	e
6	6	6		38	26	46	&	70	46	106	F	102	66	146	f
7	7	7		39	27	47	'	71	47	107	G	103	67	147	g
8	8	10		40	28	50	(72	48	110	H	104	68	150	h
9	9	11		41	29	51)	73	49	111	I	105	69	151	i
10	A	12		42	2A	52	*	74	4A	112	J	106	6A	152	j
11	B	13		43	2B	53	+	75	4B	113	K	107	6B	153	k
12	C	14		44	2C	54	,	76	4C	114	L	108	6C	154	l
13	D	15		45	2D	55	-	77	4D	115	M	109	6D	155	m
14	E	16		46	2E	56	.	78	4E	116	N	110	6E	156	n
15	F	17		47	2F	57	/	79	4F	117	O	111	6F	157	o
16	10	20		48	30	60	0	80	50	120	P	112	70	160	p
17	11	21		49	31	61	1	81	51	121	Q	113	71	161	q
18	12	22		50	32	62	2	82	52	122	R	114	72	162	r
19	13	23		51	33	63	3	83	53	123	S	115	73	163	s
20	14	24		52	34	64	4	84	54	124	T	116	74	164	t
21	15	25		53	35	65	5	85	55	125	U	117	75	165	u
22	16	26		54	36	66	6	86	56	126	V	118	76	166	v
23	17	27		55	37	67	7	87	57	127	W	119	77	167	w
24	18	30		56	38	70	8	88	58	130	X	120	78	170	x
25	19	31		57	39	71	9	89	59	131	Y	121	79	171	y
26	1A	32		58	3A	72	:	90	5A	132	Z	122	7A	172	z
27	1B	33		59	3B	73	;	91	5B	133	[123	7B	173	{
28	1C	34		60	3C	74	<	92	5C	134	\	124	7C	174	
29	1D	35		61	3D	75	=	93	5D	135]	125	7D	175	}
30	1E	36		62	3E	76	>	94	5E	136	^	126	7E	176	~
31	1F	37		63	3F	77	?	95	5F	137	_	127	7F	177	

උදාහරණ :- දැනටමත් සංඛ්‍යා උපයෝගී කර ගෙන 'Hello.'
යන්නට අදාළ ඇස්කි කේතය ලියන්න.

ඉහත වගුව අනුව

$H = 72, e = 101, l = 108, o = 111$ සහ $. = 46$ වේ. ඒ අනුව
'Hello.' යන්න,

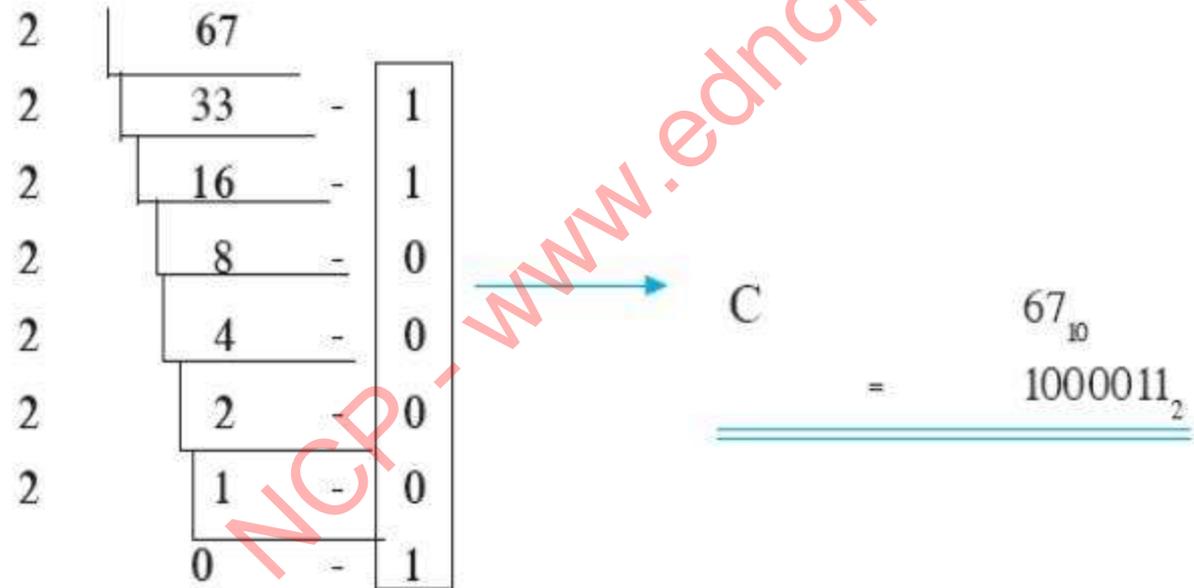
72 101 108 108 111 සහ 46 ලෙස ලිවිය හැකි ය.

උදාහරණ 'COMPUTER' යන වචනය සලකා බලමු. ඉහත
වගුවට අනුව මෙහි ඇතුළත්ව ඇති අදාළ දැනටමත් ඇගයීම්
පහත දැක්වේ.

$C = 67, O = 79, M = 77, P = 80, U = 85, T = 84, E = 69,$

$R = 125$ උදාහරණ මෙම එක් එක් ඇතුළත්ව ඇති අදාළ දැනටමත් ඇගයීම්
ද්විමය සංඛ්‍යා බවට පරිවර්තනය කරමු.

උදාහරණ :



මෙම අනිත් අකුරුවලට අදාළ දූෂමය සංඛ්‍යාවන් ද, ද්විමය සංඛ්‍යා බවට හරවන්න. අප 'COMPUTER' යන වචනය යතුරු පුවරුවෙන් ඇතුළු කළ විට පරිගණකය එය තේරුම් ගන්නේ මෙම අකුරුවලට අදාළ ද්විමය සංඛ්‍යා මගිනි.

UNICODE

මුල් යුගවල දී පරිගණකවල භාවිත වූ ASCII හා EBCDIC යනාදී කේතයන්හි ප්‍රධාන අවාසිය වූයේ ඒවා බිටු 8 කට සීමාවීම ය. බිටු 8 කින් දැක්විය හැකි උපරිම අනුලක්ෂණ ගණන 256 (2^8) ක් නිසා වර්තමාන තොරතුරු යුගයේ අවශ්‍යතාවනට ඒවා ප්‍රමාණවත් නොවනු ඇත. අද අප ජීවත් වනුයේ ගෝලීයකරණයට භාෂ්‍යය වෙමින් පවතින සමාජ ක්‍රමයක් තුළ බැවින් විශ්ව දැනුම බෙදා ගැනීම, ජාත්‍යන්තර වෙළෙඳාම යනාදියට අත්‍යවශ්‍ය වූ සන්නිවේදනය සඳහා විවිධ අනුලක්ෂණ සහ සංකේත විශාල වශයෙන් යොදා ගත යුතුය. විශේෂයෙන් ම ඉංග්‍රීසි භාෂාවට අමතර ව විවිධ ජාතීන් විසින් භාවිතා කරනු ලබන ඒ ඒ භාෂාවන්ට ආවේණික වූ අනුලක්ෂණ

පරිගණකයන්හි භාවිත වීමත් ඒ ඒ භාෂාවන්ට ආවේණික වූ අනුලක්ෂණයන් සඳහා කේත වෙන් කළ යුතු විය. මේ අනුව විවිධ භාෂාවන්හි අක්ෂර, විවිධ ජාතීන් භාවිත කරන්නා වූ සංකේත ආදිය නිරූපණය කිරීම සඳහා අවම වශයෙන් බිටු 16 ක කේතයක් භාවිත කළ යුතු බව පැහැදිලි ය. ඒ අනුව 1991 ඔක්තෝබර් මාසයේ දී බිටු 16 කින් එකිනෙකට වෙනස් සංකේත 65536 ක් නිරූපණය කළ හැකි ජාත්‍යන්තර ප්‍රමිතියකට අනුකූල ව සකස් කරන ලද යුනිකේතය නම් කේතයක් බිහිවිය. ,The Unicode Standard,



මෙම කේත ක්‍රමයෙහි ප්‍රවර්ධනය සම්බන්ධීකරණය කරනුයේ ලාභ අපේක්ෂාවෙන් තොරව කටයුතු කරන **Unicode Consortium**, නැමැති සංවිධානයකි. ශ්‍රී ලංකාවේ භාවිත වන සිංහල සහ දුමිළ භාෂාවන්හි අක්ෂර මාලාවන් ද මේ වන විට යුනිකේත සම්මතයට ඇතුළත් කර ඇත. යුනිකේත ක්‍රමය අද ඉතා ජනප්‍රිය තාක්ෂණයක් වන **XML, Microsoft - NET** යනාදියටත් පයිතන් (**Python**), ජාවා (**Java**) වැනි ප්‍රචලිත පරිගණක භාෂා යනාදියටත්, නවීන පරිගණක මෙහෙයුම් පද්ධතිවලටත් අනුගත කර තිබීම එහි තවත් වාසියක් ලෙස සැලකිය හැක.

The End

Education NCP - www.edncp.lk