Binary Coding Schemes

- The code is made up of fixed size and in a standard format.
- A combination of bits represent a unique symbol in the data.
- Using the standard codes, the programmer uses the combination of bits to represent a symbol in the data.
- Some commonly used coding schemes are
  - Binary Coded Decimal (BCD)
  - American Standard Code for Information Interchange (ASCII)
  - Extended Binary Coded Decimal Interchange Code (EBCDIC)
  - Unicode

BCD

- A method that uses binary digits 0 and 1 to represent decimal digits.
- The idea of this coding scheme is to convert each digit of a decimal number instead of converting the entire decimal value.
- It is represented using 4 bits.
- For example, the BCD rendition of the base-10 number 1895 is

0001 1000 1001 0101

The binary equivalents of 1, 8, 9, and 5, always in a four-digit format, go from left to right.

DECIMAL	BCD
	8421
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001

BCD

- Advantages:
  - There is no limit to the size of a number.
  - Fast way to convert numbers from decimal to binary
  - If 4 bit is not sufficient, we can go to 6 bit BCD. This will represent 2^6 different characters.
- Disadvantages:
  - Modern computers donot use this coding scheme.

EBCDIC

- Extended Binary Coded Decimal Interchange Code
- It is an 8 bit alphanumeric code that was developed by IBM to represent alphabets, numbers, special symbols including control characters.
- This was mainly used by IBM mainframe computers.
- Here, 1100 is the zone combination used for A to I.
- 1101, is used for J to R.
- 1110 is used for S to Z.
- 1111 is used for 0 to 9.

EBCDIC Code Table

B8 →	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
B7 →	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
B6 →	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
B5 →	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
B4 ↓ B3 ↓ B2 ↓ B1 ↓ HEX-1 \ HEX-0	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0 0 0 0 0	NUL	DLE	DS		SP	&	—									0
0 0 0 1 1	SOH	SBA	SOS				/		a	j			A	J		1
0 0 1 0 2	STX	EUA	FS	SYN					b	k	s		B	K	S	2
0 0 1 1 3	ETX	IC							c	l	t		C	L	T	3
0 1 0 0 4	PF	RES	BYP	PN					d	m	u		D	M	U	4
0 1 0 1 5	PT	NL	LF	RS					e	n	v		E	N	V	5
0 1 1 0 6	LC		ETB	UC					f	o	w		F	O	W	6
0 1 1 1 7	DEL	IL	ESC	EOT					g	p	x		G	P	X	7
1 0 0 0 8		CAN							h	q	y		H	Q	Y	8
1 0 0 1 9		EM							i	r	z		I	R	Z	9
1 0 1 0 A	SMM	CC	SM		¢	!	¦	:								
1 0 1 1 B	VT				.	$	,	#								
1 1 0 0 C	FF	DUP		RA	<	*	%	@								
1 1 0 1 D	CR	SF	ENQ	NAK	(	)	_	'								
1 1 1 0 E	SO	FM	ACK		+	;	>	=								
1 1 1 1 F	SI	ITB	BEL	SUB		¬	?	"								

ASCII

- American Standard Code for Information Interchange
- It is a 7 bit code.
- It is used for transmitting and processing data amongst different computers.
- Later, IBM developed new version of ASCII, known as ASCII-8.

It consist of 256 symbols.

- The concept is same as EBCDIC only the difference is in the combinations assigned to represent the various alphabets, numeric and special characters.

Dec	Hx	Oct	Char	Dec	Hx	Oct	Html	Chr	Dec	Hx	Oct	Html	Chr	Dec	Hx	Oct	Html	Chr
0	0	000	NUL (null)	32	20	040		Space	64	40	100	@	@	96	60	140	`	`
1	1	001	SOH (start of heading)	33	21	041	!	!	65	41	101	A	A	97	61	141	a	a
2	2	002	STX (start of text)	34	22	042	"	"	66	42	102	B	B	98	62	142	b	b
3	3	003	ETX (end of text)	35	23	043	#	#	67	43	103	C	C	99	63	143	c	c
4	4	004	EOT (end of transmission)	36	24	044	$	$	68	44	104	D	D	100	64	144	d	d
5	5	005	ENQ (enquiry)	37	25	045	%	%	69	45	105	E	E	101	65	145	e	e
6	6	006	ACK (acknowledge)	38	26	046	&	&	70	46	106	F	F	102	66	146	f	f
7	7	007	BEL (bell)	39	27	047	'	'	71	47	107	G	G	103	67	147	g	g
8	8	010	BS (backspace)	40	28	050	(	(	72	48	110	H	H	104	68	150	h	h
9	9	011	TAB (horizontal tab)	41	29	051	)	)	73	49	111	I	I	105	69	151	i	i
10	A	012	LF (NL line feed, new line)	42	2A	052	*	*	74	4A	112	J	J	106	6A	152	j	j
11	B	013	VT (vertical tab)	43	2B	053	+	+	75	4B	113	K	K	107	6B	153	k	k
12	C	014	FF (NP form feed, new page)	44	2C	054	,	,	76	4C	114	L	L	108	6C	154	l	l
13	D	015	CR (carriage return)	45	2D	055	-	-	77	4D	115	M	M	109	6D	155	m	m
14	E	016	SO (shift out)	46	2E	056	.	.	78	4E	116	N	N	110	6E	156	n	n
15	F	017	SI (shift in)	47	2F	057	/	/	79	4F	117	O	O	111	6F	157	o	o
16	10	020	DLE (data link escape)	48	30	060	0	0	80	50	120	P	P	112	70	160	p	p
17	11	021	DC1 (device control 1)	49	31	061	1	1	81	51	121	Q	Q	113	71	161	q	q
18	12	022	DC2 (device control 2)	50	32	062	2	2	82	52	122	R	R	114	72	162	r	r
19	13	023	DC3 (device control 3)	51	33	063	3	3	83	53	123	S	S	115	73	163	s	s
20	14	024	DC4 (device control 4)	52	34	064	4	4	84	54	124	T	T	116	74	164	t	t
21	15	025	NAK (negative acknowledge)	53	35	065	5	5	85	55	125	U	U	117	75	165	u	u
22	16	026	SYN (synchronous idle)	54	36	066	6	6	86	56	126	V	V	118	76	166	v	v
23	17	027	ETB (end of trans. block)	55	37	067	7	7	87	57	127	W	W	119	77	167	w	w
24	18	030	CAN (cancel)	56	38	070	8	8	88	58	130	X	X	120	78	170	x	x
25	19	031	EM (end of medium)	57	39	071	9	9	89	59	131	Y	Y	121	79	171	y	y
26	1A	032	SUB (substitute)	58	3A	072	:	:	90	5A	132	Z	Z	122	7A	172	z	z
27	1B	033	ESC (escape)	59	3B	073	;	;	91	5B	133	[	[	123	7B	173	{	{
28	1C	034	FS (file separator)	60	3C	074	<	<	92	5C	134	\	\	124	7C	174			|
29	1D	035	GS (group separator)	61	3D	075	=	=	93	5D	135	]	]	125	7D	175	}	}
30	1E	036	RS (record separator)	62	3E	076	>	>	94	5E	136	^	^	126	7E	176	~	~
31	1F	037	US (unit separator)	63	3F	077	?	?	95	5F	137	_	_	127	7F	177	&#127;	DEL

Source: www.LookupTables.com

128	Ç	144	É	160	á	176	▒	192	└	208	╨	224	α	240	≡								
129	ü	145	æ	161	í	177	▒	193	┴	209	╤	225	ß	241	±								
130	é	146	Æ	162	ó	178	▓	194	┬	210	╥	226	Γ	242	≥								
131	â	147	ô	163	ú	179	│	195	├	211	╙	227	π	243	≤								
132	ä	148	ö	164	ñ	180	┤	196	─	212	╘	228	Σ	244	⌠								
133	à	149	ò	165	Ñ	181	╡	197	┼	213	╒	229	σ	245	⌡								
134	å	150	û	166	ª	182	╢	198	╞	214	╓	230	µ	246	÷								
135	ç	151	ù	167	º	183	╖	199	╟	215	╫	231	τ	247	≈								
136	ê	152	ÿ	168	¿	184	╕	200	╚	216	╪	232	Φ	248	°								
137	ë	153	Ö	169	⌐	185	╣	201	╔	217	┘	233	Θ	249	∙								
138	è	154	Ü	170	¬	186	║	202	╩	218	┌	234	Ω	250	·								
139	ï	155	¢	171	½	187	╗	203	╦	219	█	235	δ	251	√								
140	î	156	£	172	¼	188	╝	204	╠	220	▄	236	∞	252	n								
141	ì	157	¥	173	¡	189	╜	205	═	221	▌	237	φ	253	²								
142	Ä	158	₧	174	«	190	╛	206	╬	222	▐	238	ε	254	■								
143	Å	159	ƒ	175	»	191	┐	207	╧	223	▀	239	∩	255									

Source: www.LookupTables.com

Unicode

- It is a 16 bit universal character coding standard for the representation of text which includes numbers, text and symbols.
- The Unicode consortium based in California developed the new unicode standard.
- It uses 32 bit to represent a symbol in the data.
- It is capable of representing approximately 1 million characters.
- It can uniquely represent any character or symbol present in any language like Chinese, Japanese, etc.
- In addition to the letters; mathematical and scientific symbols are also separated in Unicode codes.

Unicode

- The main advantage of Unicode is that it is compatible with ASCII.
- The first 256 characters are same as ASCII