

ECMA

EUROPEAN COMPUTER MANUFACTURERS ASSOCIATION

---

STANDARD ECMA - 43

for an

8-BIT CODED CHARACTER SET

December 1974

Free copies of this ECMA standard are available from  
ECMA European Computer Manufacturers Association  
114 Rue du Rhône - 1204 GENEVA (Switzerland)

ECMA

EUROPEAN COMPUTER MANUFACTURERS ASSOCIATION

---

STANDARD ECMA - 43

for an

8-BIT CODED CHARACTER SET

December 1974

## FOREWORD

This Standard has been drafted by Technical Committee TC1 of ECMA. Following considerations lead to the development of this Standard.

Most interchange media being designed for 8-bit working, it would have been inconsequent to leave code standardization at the 7-bit level. For interchange purposes, TC1 wished to foster a standard way of encoding the characters of the 7-bit code in an 8-bit environment. This will assist in making different manufacturers' equipment compatible and will allow for the orderly progression from existing to future equipment in the simplest way. Without such standardization different implementors would solve the problem in individual ways and confusion would ensue.

TC1 had originally intended to standardize the full complement of 256 characters. It has then become clear that this may not be desirable. The 128 characters of the 7-bit code provide a general purpose character set which is adequate for most applications. To go beyond this leads to difficulties on two counts. Firstly some equipment has to be more specialized or sophisticated. Secondly applications requiring additional characters tend themselves to be specialized and therefore each application may require its own special set.

Although ECMA-35, "Extension of the ECMA 7-bit coded Character Set" anticipated standardization of an 8-bit code there was still a need for a separate 8-bit document. ECMA-35 is a very large document covering more ground than is strictly necessary to define an 8-bit code standard. It was considered essential therefore to extract the relevant material and draw it together within the confines of one simple document specifically addressed to the 8-bit code. Furthermore, ECMA-35 was designed to contain all current requirements and future possibilities in the field of code extension, as envisaged by the DP standardizing community, worldwide. It provides a general framework for future standardization. Just the reverse approach is needed for the 8-bit code, which must be as specific as possible. Real benefits arise from the very act of selecting from the multiplicity of possibilities in ECMA-35.

This Standard ECMA-43 has been adopted by the General Assembly of December 17, 1974.

TABLE OF CONTENT

	<u>Page</u>
1. SCOPE AND FIELD OF APPLICATION	1
2. IMPLEMENTATION	1
3. BASIC CODE TABLE	1
4. VERSIONS OF TABLE 1	2
4.1 General	2
4.2 Definition of the Code Table	3
4.3 Default Version	3
5. NOTES ON TABLE 1	4
 TABLE 1	 6
 TABLE 2	 7
 APPENDIX	
Relevant ECMA Standards	8

## 1. SCOPE AND FIELD OF APPLICATION

- 1.1 The 8-bit coded character set specified in this Standard ECMA-43 is derived from, and compatible with, the 7-bit coded character set defined in Standard ECMA-6. Furthermore, its characteristics are in conformance with the code extension techniques described in Standard ECMA-35.
- 1.2 Standard ECMA-35 defines, among other things, a family of 8-bit codes, each of which accommodates a basic set of 128 characters together with 32 additional control characters and 94 additional graphic characters. The present Standard is limited to those members of the 8-bit code family which have the characters of a version of 7-bit code of ECMA-6 as their basic set.
- 1.3 While the 7-bit coded character set of ECMA-6 is the standard code for general information interchange, the 8-bit code described in this Standard is intended for interchange of information among data processing systems and associated equipment operating within an 8-bit environment. This 8-bit code is also recommended for storage of information.
- 1.4 This Standard essentially consists of a table defining the character set and its coded representation, complemented by notes, legend and guidance on the exercise of the options to define specific versions.

## 2. IMPLEMENTATION

The implementation on physical media and in transmission, taking into account the need for error checking, is the subject of other ECMA Standards (see Appendix).

## 3. BASIC CODE TABLE

Table 1 shows a 16 by 16 array of columns numbered 00 to 15 and rows numbered 00 to 15 containing 256 code positions. Each position contains a character and/or a reference to Notes. This Table together with the Notes constitutes the basic 8-bit coded character set. In order to use it, versions according to 4. below must be defined. The requirements, definitions and explanatory notes specified in ECMA-6 are equally applicable to this Standard, although they are not repeated here.

- 3.1 Columns 00 to 07 of the array contain 128 characters which are in one-to-one correspondence with the 128 characters of the basic set defined in Standard ECMA-6. Their coded representation is the same as in the 7-bit code with the addition of an 8th most significant bit which is ZERO.

In particular, a national or an application-oriented version derived in accordance with ECMA-6, or the International Reference Version of the 7-bit code, is an appropriate candidate for utilization in this part of the table.

- 3.2 Columns 08 to 15 of the array contain a further 128 code positions; the 8th bit of their coded representation is ONE.

- 3.2.1 Columns 08 and 09 are reserved for a set of 32 control characters. If this set has to be designated, the rules relevant to C<sub>1</sub> sets defined in ECMA-35 apply. This set shall not include transmission control characters.

- 3.2.2 Columns 10 to 15, with the exception of positions 10/00 and 15/15, are reserved for a set of 94 graphic characters. If this set has to be designated, the rules relevant to G<sub>1</sub> sets defined in ECMA-35 apply.

- 3.2.3 Positions 10/00 and 15/15 have no meaning assigned to them by this Standard. The use of these bit-combinations requires agreement on their meaning and, when required, on their representation in a 7-bit environment.

#### 4. VERSIONS OF TABLE 1

##### 4.1 General

In order to use Table 1 for information interchange, it is necessary to exercise the options left open, i.e. those specified by Notes ② to ⑤. A single character must be allocated to each of the positions for which these options apply or the position must be declared to be unused.

If there is no demand for specific characters, it is strongly recommended that the characters of the International Reference Version of the 7-bit code are allocated to the corresponding code positions.

Furthermore, it is necessary to define the set of control characters allocated to columns 08 and 09, in accordance with Note ⑨, and the set of graphics allocated to columns 10 to 15, in accordance with Note ⑩. When there is no requirement for either or both of the sets, they must be declared to be unused.

A code table completed in this way is called a "version".

#### 4.2 Definition of the Code Table

The definition of a complete code table (version), as described in 4.1 above, may be the subject of agreement between the interchanging parties, or may be specified by means of the ESCAPE sequences described below. In the absence of such agreement or ESCAPE sequences, the default version described in 4.3 is assumed.

Within the limits allowed by 4.1 and to avoid misinterpretation of the interchanged data, the code table in actual use can be defined by means of the three-character ESCAPE sequences described in ECMA-35, i.e.:

- a) ESC 02/01 (F), to designate the set of controls in columns 00 and 01;
- b) ESC 02/08 (F) or ESC 02/12 (F), to designate the set of graphics in columns 02 to 07;
- c) ESC 02/02 (F), to designate the set of controls in columns 08 and 09;
- d) ESC 02/09 (F) or ESC 02/13 (F), to designate the set of graphics in columns 10 to 15.

In the foregoing (F) stands for one of the bit-combinations from 03/00 to 07/14.

#### 4.3 Default Version

This version is available for use when there is no requirement to use a national or an application-oriented version in columns 00 to 07, or no requirement to use columns 08 to 15.

This version (Table 2) is assumed, unless a particular agreement exists between sender and recipient of the data, or a specific designation in accordance with 4.2 above is used. Columns 00 to 07 of this version are identical to columns 0 to 7 of the International Reference Version (IRV) of the 7-bit code (ECMA-6) and columns 08 to 15 are not used.



5. NOTES ON TABLE 1

- ① The Format Effectors are intended for equipment in which horizontal and vertical movements are effected separately. If equipment requires the action of CARRIAGE RETURN to be combined with a vertical movement, the Format Effector for that vertical movement may be used to effect the combined movement. For example, if NEW LINE (symbol NL, equivalent to CR + LF) is required, FE<sub>2</sub> shall be used to represent it. This substitution requires agreement between the sender and the recipient of the data.

The use of these combined functions may be restricted for international transmission on general switched telecommunication networks (telegraph and telephone networks).

- ② The symbol £ is assigned to position 02/03 and the symbol \$ is assigned to position 02/04. In a situation where there is no requirement for the symbol £ the symbol # (NUMBER SIGN) may be used in position 02/03. Where there is no requirement for the symbol \$ the symbol ¤ (CURRENCY SIGN) may be used in position 02/04. The chosen allocations of symbols to these positions for international information interchange shall be agreed between the interested parties. It should be noted that, unless otherwise agreed between sender and recipient, the symbols £, \$ or ¤ do not designate the currency of a specific country.
- ③ National use positions. The allocation of characters to these positions lies within the responsibility of national standardization bodies. These positions are primarily intended for alphabet extensions. If they are not required for that purpose, they may be used for other graphics.
- ④ Positions 05/14, 06/00 and 07/14 provide for the characters UPWARD ARROW HEAD, GRAVE ACCENT and OVERLINE. However, these positions may be used for other graphic characters when it is necessary to have 8, 9 or 10 positions for national use.
- ⑤ Position 07/14 is used for the character      (OVERLINE), the graphic representation of which may vary according to national use to represent ~ (TILDE) or another diacritical sign provided that there is no risk of confusion with another graphic symbol included in the Table.

- ⑥ The graphic characters in positions 02/02, 02/07, 02/12 and 05/14, have respectively the significance of QUOTATION MARK, APOSTROPHE, COMMA and UPWARD ARROW HEAD, however, these characters take on the significance of the diacritical signs DIAERESIS, ACUTE ACCENT, CEDILLA and CIRCUMFLEX ACCENT when they are preceded or followed by BACKSPACE (pos. 00/08), (see also 3.2 of ECMA-6).
- ⑦ The control characters SO (SHIFT OUT) and SI (SHIFT IN) have been allocated to positions 00/14 and 00/15 for consistency with the 7-bit code. However, attention is called to the fact that there is no requirement for a shift function of the type defined by these characters in the 8-bit code, and that ECMA-35 positively proscribes the use of the characters SO and SI, when interchanging data in 8-bit form.
- ⑧ To ensure full compatibility with the code extension techniques described in ECMA-35, position 10/00 and 15/15 are left undefined. The use of these two positions for data interchange, requires agreement between the sender and the recipient of the data.
- ⑨ Positions from 08/00 to 09/15 are reserved for the allocation of a set of up to 32 control characters.
- ⑩ Positions from 10/01 to 15/14 are reserved for the allocation of a set of up to 94 graphic characters.

# Table 1

b <sub>6</sub>	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
b <sub>7</sub>	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
b <sub>8</sub>	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
b <sub>9</sub>	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>	b <sub>4</sub>													
0	0	0	0	00	NUL TC <sub>7</sub> (DLE)	SP	0	③	P	`	p					
0	0	0	1	01	TC <sub>1</sub> (SOH)	DC <sub>1</sub>	!	1	A	Q	a	q				
0	0	1	0	02	TC <sub>2</sub> (STX)	DC <sub>2</sub>	"	2	B	R	b	r				
0	0	1	1	03	TC <sub>3</sub> (ETX)	DC <sub>3</sub>	£(#)	3	C	S	c	s				
0	1	0	0	04	TC <sub>4</sub> (EOT)	DC <sub>4</sub>	\$ (α)	4	D	T	d	t				
0	1	0	1	05	TC <sub>5</sub> (ENQ)	TC <sub>8</sub> (NAK)	%	5	E	U	e	u				
0	1	1	0	06	TC <sub>6</sub> (ACK)	TC <sub>9</sub> (SYN)	&	6	F	V	f	v				
0	1	1	1	07	BEL	TC <sub>10</sub> (ETB)	'	7	G	W	g	w				
1	0	0	0	08	FE <sub>0</sub> (BS)	CAN	(	8	H	X	h	x				
1	0	0	1	09	FE <sub>1</sub> (HT)	EM	)	9	I	Y	i	y				
1	0	1	0	10	① FE <sub>2</sub> (LF)	SUB	*	:	J	Z	j	z				
1	0	1	1	11	① FE <sub>3</sub> (VT)	ESC	+	;	K	③	k	③				
1	1	0	0	12	① FE <sub>4</sub> (FF)	IS <sub>4</sub> (FS)	⑥ ,	<	L	③	l	③				
1	1	0	1	13	① FE <sub>5</sub> (CR)	IS <sub>3</sub> (GS)	-	=	M	③	m	③				
1	1	1	0	14	⑦ SO	IS <sub>2</sub> (RS)	.	>	N	④⑥	n	④⑤				
1	1	1	1	15	⑦ SI	IS <sub>1</sub> (US)	/	?	O	-	o	DEL				



# Table 2

				b <sub>0</sub>	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
				b <sub>1</sub>	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
				b <sub>2</sub>	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
				b <sub>3</sub>	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
					00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
b <sub>4</sub>	b <sub>3</sub>	b <sub>2</sub>	b <sub>1</sub>																	
0	0	0	0	00	NUL TC <sub>7</sub> (DLE)	SP	0	@	P	`	p									
0	0	0	1	01	TC <sub>1</sub> (SOH)	DC <sub>1</sub>	!	1	A	Q	a	q								
0	0	1	0	02	TC <sub>2</sub> (STX)	DC <sub>2</sub>	"	2	B	R	b	r								
0	0	1	1	03	TC <sub>3</sub> (ETX)	DC <sub>3</sub>	#	3	C	S	c	s								
0	1	0	0	04	TC <sub>4</sub> (EOT)	DC <sub>4</sub>	α	4	D	T	d	t								
0	1	0	1	05	TC <sub>5</sub> (ENQ)	TC <sub>8</sub> (NAK)	%	5	E	U	e	u								
0	1	1	0	06	TC <sub>6</sub> (ACK)	TC <sub>9</sub> (SYN)	&	6	F	V	f	v								
0	1	1	1	07	BEL TC <sub>10</sub> (ETB)		'	7	G	W	g	w								
1	0	0	0	08	FE <sub>0</sub> (BS)	CAN	(	8	H	X	h	x								
1	0	0	1	09	FE <sub>1</sub> (HT)	EM	)	9	I	Y	i	y								
1	0	1	0	10	FE <sub>2</sub> (LF)	SUB	*	:	J	Z	j	z								
1	0	1	1	11	FE <sub>3</sub> (VT)	ESC	+	;	K	[	k	{								
1	1	0	0	12	FE <sub>4</sub> (FF)	IS <sub>1</sub> (FS)	/	<	L	\	l									
1	1	0	1	13	FE <sub>5</sub> (CR)	IS <sub>3</sub> (GS)	-	=	M	]	m	}								
1	1	1	0	14	SO	IS <sub>2</sub> (RS)	.	>	N	^	n	~								
1	1	1	1	15	SI	IS <sub>1</sub> (US)	/	?	O	_	o	DEL								

APPENDIX

RELEVANT ECMA STANDARDS

- ECMA-5 : Data Interchange on 7-Track Magnetic Tape, 3rd Edition, June 1970
- ECMA-6 : 7-Bit Input/Output Coded Character Set, 4th Edition, Aug. 1973
- ECMA-10 : Data Interchange on Punched Tape, 2nd Edition, July 1970
- ECMA-12 : Data Interchange on 9-Track Magnetic Tape at 32 bits per mm, 2nd Edition, June 1970
- ECMA-13 : Magnetic Tape Labelling and File Structure for Information Interchange, 2nd Edition, Aug. 1973
- ECMA-14 : Rules for the Definition of 4-Bit Sets Derived from the ECMA 7-Bit Coded Character Set, Nov. 1967
- ECMA-16 : Basic Mode Control Procedures for Data Communication Systems using the ECMA 7-Bit Code, 2nd Edition, June 1973
- ECMA-17 : Graphic Representation of the Control Characters of the ECMA 7-Bit Coded Character Set for Information Interchange, Nov. 1968
- ECMA-20 : Implementation of the ECMA 7-Bit Coded Character Set on Punched Cards, June 1969
- ECMA-23 : Keyboards Generating the Code Combinations of the Characters of the ECMA 7-Bit Coded Character Set, 2nd Edition, January 1975
- ECMA-24 : Code Independent Information Transfer (An Extension to the Basic Mode Transmission Control Procedures), Dec. 1969
- ECMA-25 : Representation of 8-Bit Combinations on 12-Row Punched Cards, June 1970
- ECMA-26 : Recovery Procedures (An Extension to the Basic Mode Control Procedures for Data Communication Systems), April 1971
- ECMA-27 : Abort and Interrupt Procedures (An Extension of the Basic Mode Control Procedures for Data Communication Systems), April 1971
- ECMA-28 : Multiple Station Selection Procedures (An Extension of the Basic Mode Control Procedures for Data Communication Systems), April 1971
- ECMA-29 : Conversational Information Transfer (An Extension of the Basic Mode Control Procedures for Data Communication Systems), Sept. 1971
- ECMA-33 : Track Format Characteristics of Interchangeable 6-Disk Packs, Sept. 1971
- ECMA-34 : Data Interchange on 3,81 mm Magnetic Tape Cassette (32 b/mm, Phase-Encoded), 2nd Edition, July 1973
- ECMA-35 : Extension of the 7-Bit Coded Character Set, Dec. 1971
- ECMA-36 : Data Interchange on 9-Track Magnetic Tape at 63 b/mm (1600 bpi) Phase-Encoded, Dec. 1971
- ECMA-37 : Supplementary Transmission Control Functions (An Extension of the Basic Mode Control Procedures for Data Communication Systems), June 1972
- ECMA-39 : Track Format Characteristics of Interchangeable Single Disk Cartridges (Top Loaded), Sept. 1973
- ECMA-41 : Magnetic Tape Cassette Labelling and File Structure for Information Interchange, Dec. 1973

