

ECMA

Standardizing Information and Communication Systems

7-Bit coded Character Set

ECMA

Standardizing Information and Communication Systems

7-Bit coded Character Set

Brief History

Technical Committee TC1 of ECMA met for the first time in December 1960 to prepare standard codes for Input/Output purposes. On 30th April 1965, Standard ECMA-6 was adopted by the General Assembly of ECMA.

Subsequent international activities necessitated three revisions of ECMA-6, which were issued in June 1967, July 1970 and August 1973. In the 3rd edition, TC1 introduced a new concept, that of versions of the code and proposed at the same time an International Reference Version. This concept and this specific version were eventually adopted by ISO for the new International Standard ISO 646-1972 and by CCITT for the International Telegraph Alphabet No. 5 (CCITT Rec. V.3).

In parallel to these developments, further work on extension of the 7-bit code was achieved. The techniques for extending the repertoire of the 7-bit code, remaining in a 7-bit environment, and for expanding it to 8-bit codes are described in Standard ECMA-35 (ISO 2022). In application of these techniques the structure of, and rules for, 8-bit codes are defined in Standard ECMA-43 (ISO 4873).

This 6th Edition corresponds to the 3rd edition of ISO 646 issued in 1991, a revision of the 1983 issue prepared by WG-7 of ISO/IEC/JTC1/SC2 in which several members of ECMA/TC1 were major contributors. The differences between the 5th and the 6th editions are specified in annex C.

Adopted by the General Assembly of ECMA in December 1991.

Table of contents

1 Scope	1
2 Conformance and Implementation	1
2.1 Conformance	1
2.1.1 Conformance of information interchange	1
2.1.2 Conformance of devices	1
2.2 Implementation	1
3 References	2
4 Definitions	2
4.1 Active position	2
4.2 Bit combination	2
4.3 Character	2
4.4 Character position	2
4.5 Coded character set; code	2
4.6 Coded-character-data-element (CC-data-element)	2
4.7 Code extension	2
4.8 Code table	2
4.9 Control character	3
4.10 Control function	3
4.11 Device	3
4.12 Escape sequence	3
4.13 Final byte	3
4.14 Graphic character	3
4.15 Graphic symbol	3
4.16 Repertoire	3
4.17 User	3
5 Notation, Code Table and Names	3
5.1 Notation	3
5.2 Code table	4
5.3 Names	4
6 Specification of the Coded Character Set	4
6.1 Structure	4
6.2 Control characters	5
6.3 Character SPACE	5
6.4 Graphic characters	5
6.4.1 Unique graphic character allocations	5
6.4.2 Alternative graphic character allocations	7
6.4.3 National or application-oriented graphic character allocations	8
6.5 Character DELETE	8
7 Composite Graphic Characters	8
8 Versions of the Coded Character Set	8
8.1 General	8
8.2 International Reference Version (IRV)	9
8.3 National versions	9

8.4 Application-oriented versions	10
9 Identification of Versions	10
9.1 Purpose and context of identification	10
9.2 Identification of a version	10
10 Explanation of Code Tables No. 4 and No. 5	10
Annex A - Specification of the C0 set	13
Annex B - Guidelines for Standards derived from Standard ECMA-6	15
Annex C - Differences between the 5th Edition (1983) and the present (6th) Edition of ECMA-6	17

1 Scope

- 1.1 This ECMA Standard specifies a set of 128 characters, (control characters and graphic characters such as letters, digits and symbols) with their coded representation. Most of these characters are mandatory and unchangeable, but provision is made for some flexibility to accommodate national and other requirements.
- 1.2 This ECMA Standard specifies a 7-bit coded character set with a number of options. It also provides guidance on how to exercise the options to define specific national versions and application-oriented versions. Furthermore it specifies the International Reference Version (IRV) in which such options have been exercised.
- 1.3 This character set is primarily intended for the interchange of information among data processing systems and associated equipment, and within data communication systems. The need for graphic characters and control functions in data processing has also been taken into account in determining this character set.
- 1.4 This character set is applicable to alphabets of the Latin script.
- 1.5 This character set allows the use of control characters for code extension where its character set is insufficient for particular applications. Procedures for the use of these control characters are specified in Standard ECMA-35.
- 1.6 The definitions of the control characters mentioned in this ECMA Standard are specified in Standard ECMA-48. It is assumed that data associated with them are to be processed serially in a forward direction. When they are included in strings of data which are processed other than serially in a forward direction or when they are included in data formatted for fixed-record processing they may have undesirable effects or may require additional special treatment to ensure that they result in their desired function.

2 Conformance and Implementation

2.1 Conformance

2.1.1 Conformance of information interchange

A coded-character-data-element (CC-data-element) within coded information for interchange is in conformance with this ECMA Standard if all the coded representations of characters within that CC-data-element conform to the requirements of 8.1.

A claim of conformance shall identify the version adopted in accordance with 8.2 to 8.4.

2.1.2 Conformance of devices

A device is in conformance with this ECMA Standard if it conforms to the requirements of 2.1.2.1, and either or both of clauses 2.1.2.2 and 2.1.2.3 below. A claim of conformance shall identify the version adopted.

2.1.2.1 Device description

A device that conforms to this ECMA Standard shall be the subject of a description that identifies the means by which the user may supply characters to the device, or may recognize them when they are made available to him, as specified respectively in clauses 2.1.2.2 and 2.1.2.3 below.

2.1.2.2 Originating devices

An originating device shall allow its user to supply any sequence of characters from the version adopted, and shall be capable of transmitting their coded representations within a CC-data-element.

2.1.2.3 Receiving devices

A receiving device shall be capable of receiving and interpreting any coded representations of characters that are within a CC-data-element, and that conform to clause 2.1.1 of this ECMA Standard, and shall make the corresponding characters available to its user in such a way that the user can identify them from among those of the version adopted, and can distinguish them from each other.

2.2 Implementation

The use of this character set requires definitions of its implementation in various media. For example, these could include punched tapes, punched cards, magnetic and optical interchangeable media and transmission channels, thus permitting interchange of data to take place either indirectly by means of an intermediate

recording on a physical medium, or by local connection of various units (such as input and output devices and computers) or by means of data transmission equipment.

The implementation of this coded character set in physical media and for transmission, taking into account the need for error checking, is the subject of other standards.

3 References

- ECMA-35: Code Extension Techniques, 4th Edition
- ECMA-43: 8-Bit Coded Character Set - Structure and Rules
- ECMA-48: Control Functions for Coded Character Sets, 5th Edition
- ISO/IEC 10538: Control functions for text communication

4 Definitions

For the purpose of this Standard the following definitions apply.

4.1 Active position

The character position which is to image the graphic symbol representing the next graphic character or relative to which the next control function is to be executed.

NOTE 1

In general, the active position is indicated in a display by a cursor.

4.2 Bit combination

An ordered set of bits used for the representation of characters.

4.3 Character

A member of a set of elements used for the organization, control or representation of data.

4.4 Character position

The portion of a display that is imaging or is capable of imaging a graphic symbol.

4.5 Coded character set; code

A set of unambiguous rules that establishes a character set and the one-to-one relationship between the characters of the set and their bit combinations.

4.6 Coded-character-data-element (CC-data-element)

An element of interchanged information that is specified to consist of a sequence of coded representations of characters, in accordance with one or more identified standards for coded character sets.

NOTE 2

In a communication environment according to the Reference Model for Open Systems Interconnection of ISO 7498, a CC-data-element will form all or part of the information that corresponds to the Presentation-Protocol-Data-Unit (PPDU) defined in that International Standard.

NOTE 3

When information interchange is accomplished by means of interchangeable media, a CC-data-element will form all or part of the information that corresponds to the user data, and not that recorded during formatting and initialization.

4.7 Code extension

The techniques for the encoding of characters that are not included in the character set of a given code.

4.8 Code table

A table showing the character allocated to each bit combination in a code.

4.9 Control character

A control function the coded representation of which consists of a single bit combination.

4.10 Control function

An action that affects the recording, processing, transmission, or interpretation of data, and that has a coded representation consisting of one or more bit combinations.

4.11 Device

A component of information processing equipment which can transmit, and/or receive, coded information within CC-data-elements.

NOTE 4

It may be an input/output device in the conventional sense, or a process such as an application program or a gateway function.

4.12 Escape sequence

A string of bit combinations that is used for control purposes in code extension procedures. The first of these bit combinations represents the control function ESCAPE.

4.13 Final byte

The bit combination that terminates an escape sequence or a control sequence.

4.14 Graphic character

A character, other than a control function, that has a visual representation normally hand-written, printed or displayed, and that has a coded representation consisting of one or more bit combinations.

4.15 Graphic symbol

A visual representation of a graphic character or of a control function.

4.16 Repertoire

A specified set of characters that are represented by means of one or more bit combinations of a coded character set.

4.17 User

A person or other entity that invokes the services provided by a device.

NOTE 5

This entity may be a process such as an application program if the "device" is a code converter or a gateway function, for example.

NOTE 6

The characters, as supplied by the user or made available to him, may be in the form of codes local to the device, or of non-conventional visible representations, provided that clause 2.1.2 above is satisfied.

5 Notation, Code Table and Names

5.1 Notation

The bits of the bit combinations of the 7-bit code are identified by b₇, b₆, b₅, b₄, b₃, b₂ and b₁, where b₇ is the highest-order, or most-significant, bit and b₁ is the lowest-order, or least-significant, bit.

The bit combinations may be interpreted to represent integers in the range 0 to 127 in binary notation by attributing the following weights to the individual bits:

Bit :	b ₇	b ₆	b ₅	b ₄	b ₃	b ₂	b ₁
Weight :	64	32	16	8	4	2	1

In this ECMA Standard, the bit combinations are identified by notations of the form x/y , where x is a number in the range 0 to 7 and y is a number in the range 0 to 15. The correspondence between the notations of the form x/y and the bit combinations consisting of the bits b_7 to b_1 is as follows:

- x is the number represented by b_7 , b_6 and b_5 where these bits are given the weights 4, 2 and 1 respectively;
- y is the number represented by b_4 , b_3 , b_2 and b_1 where these bits are given the weights 8, 4, 2 and 1 respectively.

The notations of the form x/y are the same as those used to identify code table positions, where x is the column number and y the row number (see clause 5.2).

5.2 Code table

A 7-bit code table consists of 128 positions arranged in 8 columns and 16 rows. The columns are numbered 0 to 7 and the rows 0 to 15.

The code table positions are identified by notations of the form x/y , where x is the column number and y is the row number.

The positions of the code table are in one-to-one correspondence with the bit combinations of the code. The notation of a code table position, of the form x/y , is the same as that of the corresponding bit combination.

5.3 Names

This ECMA Standard assigns a unique name to each character. In addition, it specifies an acronym for each control character and for the characters SPACE and DELETE, and a graphic symbol for each graphic character. By convention, only capital letters, space and hyphen are used for writing the names of the characters. For acronyms only capital letters and digits are used. It is intended that the acronyms and this convention be retained in all translations of the text.

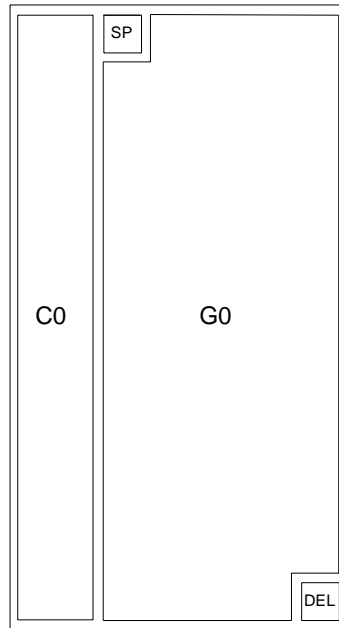
The names chosen to denote graphic characters are intended to reflect their customary meaning. However, this ECMA Standard does not define and does not restrict the meanings of graphic characters. Neither does it specify a particular style or font design for the graphic characters when imaged.

6 Specification of the Coded Character Set

6.1 Structure

The coded character set of this ECMA Standard shall have the following structure.

- a C0 control character set of up to 32 control characters;
- the graphic character SPACE (SP);
- a G0 graphic character set of up to 94 graphic characters;
- the character DELETE (DEL).



97-0171

Figure 1 - Structure of the 7-bit coded character set

6.2 Control characters

The control characters of the C0 set shall be represented by bit combinations as specified in annex A.

6.3 Character SPACE

The acronym of the character SPACE is SP and it is represented by bit combination 2/0.

This character is a graphic character, it has a visual representation consisting of the absence of a graphic symbol.

6.4 Graphic characters

The 94 bit combinations 2/1 to 7/14 are used for the representation of graphic characters as specified in 6.4.1, 6.4.2 and 6.4.3.

All graphic characters shall be spacing characters, that is, they cause the active position to advance by one character position.

6.4.1 Unique graphic character allocations

A unique graphic character is allocated to each of the 82 bit combinations 2/1, 2/2, 2/5 to 3/15, 4/1 to 5/10, 5/15 and 6/1 to 7/10. These graphic characters are specified in table 1.

Table 1 - Unique graphic character allocations

Graphic Symbol	Name	Coded representation
!	EXCLAMATION MARK	2/1
"	QUOTATION MARK	2/2
%	PERCENT SIGN	2/5
&	AMPERSAND	2/6
'	APOSTROPHE	2/7
(LEFT PARENTHESIS	2/8
)	RIGHT PARENTHESIS	2/9
*	ASTERISK	2/10
+	PLUS SIGN	2/11
,	COMMA	2/12
-	HYPHEN-MINUS	2/13
.	FULL STOP	2/14
/	SOLIDUS	2/15
0	DIGIT ZERO	3/0
1	DIGIT ONE	3/1
2	DIGIT TWO	3/2
3	DIGIT THREE	3/3
4	DIGIT FOUR	3/4
5	DIGIT FIVE	3/5
6	DIGIT SIX	3/6
7	DIGIT SEVEN	3/7
8	DIGIT EIGHT	3/8
9	DIGIT NINE	3/9
:	COLON	3/10
;	SEMICOLON	3/11
<	LESS-THAN SIGN	3/12
=	EQUALS SIGN	3/13
>	GREATER-THAN SIGN	3/14
?	QUESTION MARK	3/15
A	LATIN CAPITAL LETTER A	4/1
B	LATIN CAPITAL LETTER B	4/2
C	LATIN CAPITAL LETTER C	4/3
D	LATIN CAPITAL LETTER D	4/4
E	LATIN CAPITAL LETTER E	4/5
F	LATIN CAPITAL LETTER F	4/6
G	LATIN CAPITAL LETTER G	4/7
H	LATIN CAPITAL LETTER H	4/8
I	LATIN CAPITAL LETTER I	4/9
J	LATIN CAPITAL LETTER J	4/10
K	LATIN CAPITAL LETTER K	4/11
L	LATIN CAPITAL LETTER L	4/12
M	LATIN CAPITAL LETTER M	4/13
N	LATIN CAPITAL LETTER N	4/14
O	LATIN CAPITAL LETTER O	4/15
P	LATIN CAPITAL LETTER P	5/0
Q	LATIN CAPITAL LETTER Q	5/1

Table 1 - Unique graphic character allocations (cont'd)

Graphic Symbol	Name	Coded representation
R	LATIN CAPITAL LETTER R	5/2
S	LATIN CAPITAL LETTER S	5/3
T	LATIN CAPITAL LETTER T	5/4
U	LATIN CAPITAL LETTER U	5/5
V	LATIN CAPITAL LETTER V	5/6
W	LATIN CAPITAL LETTER W	5/7
X	LATIN CAPITAL LETTER X	5/8
Y	LATIN CAPITAL LETTER Y	5/9
Z	LATIN CAPITAL LETTER Z	5/10
_	LOW LINE	5/15
a	LATIN SMALL LETTER A	6/1
b	LATIN SMALL LETTER B	6/2
c	LATIN SMALL LETTER C	6/3
d	LATIN SMALL LETTER D	6/4
e	LATIN SMALL LETTER E	6/5
f	LATIN SMALL LETTER F	6/6
g	LATIN SMALL LETTER G	6/7
h	LATIN SMALL LETTER H	6/8
i	LATIN SMALL LETTER I	6/9
j	LATIN SMALL LETTER J	6/10
k	LATIN SMALL LETTER K	6/11
l	LATIN SMALL LETTER L	6/12
m	LATIN SMALL LETTER M	6/13
n	LATIN SMALL LETTER N	6/14
o	LATIN SMALL LETTER O	6/15
p	LATIN SMALL LETTER P	7/0
q	LATIN SMALL LETTER Q	7/1
r	LATIN SMALL LETTER R	7/2
s	LATIN SMALL LETTER S	7/3
t	LATIN SMALL LETTER T	7/4
u	LATIN SMALL LETTER U	7/5
v	LATIN SMALL LETTER V	7/6
w	LATIN SMALL LETTER W	7/7
x	LATIN SMALL LETTER X	7/8
y	LATIN SMALL LETTER Y	7/9
z	LATIN SMALL LETTER	7/10

6.4.2 Alternative graphic character allocations

Two alternative graphic characters are allocated to each of the bit combinations 2/3 and 2/4. These characters are specified in table 2.

Table 2 - Alternative graphic character allocations

Graphic Symbol	Name	Coded representation
#	NUMBER SIGN	2/3
£	POUND SIGN	2/3
\$	DOLLAR SIGN	2/4
¤	CURRENCY SIGN	2/4

Either the character NUMBER SIGN or the character POUND SIGN shall be allocated to bit combination 2/3 and either the character DOLLAR SIGN or the character CURRENCY SIGN shall be allocated to bit combination 2/4 (see clause 8).

Unless otherwise agreed between sender and recipient, the graphic symbols £, \$ and ¤ do not designate the currency of a specific country.

6.4.3 National or application-oriented graphic character allocations

No specific graphic character is allocated to the ten bit combinations 4/0, 5/11 to 5/14, 6/0 and 7/11 to 7/14. These bit combinations are available for national or application-oriented use. Either a unique graphic character shall be allocated to each of these bit combinations, or the bit combination shall be declared unused (see clause 8.1).

6.5 Character DELETE

The acronym of the character DELETE is DEL and it is represented by bit combination 7/15. DEL was originally used to erase or obliterate an erroneous or unwanted character in punched tape. DEL may be used for media-fill or time-fill. DEL characters may be inserted into, or removed from, a data stream without affecting the information content of that stream, but such action may affect the information layout and/or the control of equipment.

7 Composite Graphic Characters

Whilst all graphic characters specified in this International Standard are spacing characters, it is possible, by using BACKSPACE or CARRIAGE RETURN to image two or more graphic characters at the same character position (see also annex B.2).

For example, SOLIDUS and EQUALS SIGN may be combined to image "not equals". The character LOW LINE, that may be used as a free-standing character, may also be associated with other character(s) to represent the graphic rendition "underlined".

Diacritical marks may be allocated to the bit combinations specified in 6.4.3 and be available for composing accented letters. For such composition a sequence of three characters, the first or last of which is the letter to be accented and the second of which is BACKSPACE may be used. Furthermore, QUOTATION MARK, APOSTROPHE or COMMA can be associated with a letter by means of BACKSPACE for the composition of an accented letter with a diaeresis, an acute accent or a cedilla, respectively.

8 Versions of the Coded Character Set

8.1 General

In order to use the 7-bit coded character set for information interchange, it is necessary to exercise the options left open in clause 6:

- to specify the C0 set of control characters according to annex A.
- to specify the G0 set:
 - to allocate to each of the bit combinations 2/3 and 2/4 one of the alternative graphic characters specified in 6.4.2,

- to allocate to each of the bit combinations 4/0, 5/11 to 5/14, 6/0, and 7/11 to 7/14 a unique graphic character, or to declare the bit combination unused.

A graphic character allocated to a bit combination specified in 6.4.1 and 6.4.2 shall not be allocated to any other bit combination. For example, the POUND SIGN, if not allocated to bit combination 2/3, shall not be allocated to any other bit combination.

A character set completed in this way is called a "version of ECMA-6".

Such a version will generally be the subject of a specification document which states how the above options have been exercised. Such a specification is said to be in accordance with this Standard.

8.2 International Reference Version (IRV)

This version is available for use when there is no requirement to use a national or an application-oriented version. In information interchange, the IRV is assumed unless an agreement exists between sender and recipient of the data (see also clause 10). The graphic characters allocated to the G0 set of the IRV shall be as specified in tables 1 and 3, see table 5. Agreement between sender and recipient of the data shall be required if composite characters are used.

If the C0 set of the IRV is used, it shall be as specified in A.2 of annex A.

Table 3 - IRV graphic character allocations

Graphic Symbol	Name	Coded representation
#	NUMBER SIGN	2/3
\$	DOLLAR SIGN	2/4
@	COMMERCIAL AT	4/0
[LEFT SQUARE BRACKET	5/11
\	REVERSE SOLIDUS	5/12
]	RIGHT SQUARE BRACKET	5/13
^	CIRCUMFLEX ACCENT	5/14
`	GRAVE ACCENT	6/0
{	LEFT CURLY BRACKET	7/11
	VERTICAL LINE	7/12
}	RIGHT CURLY BRACKET	7/13
~	TILDE	7/14

8.3 National versions

- 8.3.1** The responsibility for defining national versions lies with the National Standardization Bodies. These bodies shall exercise the options available and make the required selection (see annex B).
- 8.3.2** If so required, more than one national version can be defined within a country. The different versions shall be separately identified. In particular, when alternative graphic characters are required, for a given bit combination, for example 5/12, two different versions shall be identified, even if they differ only by this single character.
- 8.3.3** If, in a country, there is no special demand for specific graphic characters, it is strongly recommended that the characters of the G0 set of the International Reference Version (IRV) be selected and allocated to the same bit combinations as in the IRV.

However, when graphic characters that are different from the characters of the IRV are required, they shall have distinct forms and be given distinctive names which are not in conflict with any of the forms or the names of any of the graphic characters in the IRV.

8.4 Application-oriented versions

Within national or international industries, organizations or professional groups, application-oriented versions may be used. They require precise agreement among the interested parties, who will have to exercise the options available and to make the required selection.

9 Identification of Versions

9.1 Purpose and context of identification

CC-data-elements conforming to a version of this ECMA Standard are intended to form all or part of a composite unit of coded information that is interchanged between an originator and a recipient. The identification of the version of this ECMA Standard that has been adopted by the originator shall also be available to the recipient. The route by which such identification is communicated to the recipient is outside the scope of this ECMA Standard.

However, some standards for interchange of coded information may permit, or require, that the coded representation of the identification applicable to the CC-data-elements forms a part of the interchanged information. This clause specifies a coded representation for the identification of a version of this International Standard. Such coded representations form all or part of an identifying data element, which may be included in information interchange in accordance with the relevant standard.

9.2 Identification of a version

The identification of a version of this ECMA Standard shall comprise a set of identifications, one for the C0 set and G0 set that constitute the version. Each identification in the set shall consist of a designating escape sequence of the type shown below.

ESC 2/1 F shall identify the C0 set.

ESC 2/8 F shall identify the G0 set.

The final byte F of these escape sequence shall be obtained from the International Register ISO 2375. If the C0 set is empty, the identification shall be the same escape sequence in which the final byte F is 7/14.

The IRV shall be identified by the following escape sequences:

For the C0 set, if provided, by Registration ISO IR No. 1 : ESC 2/1 4/0

For the C0 set, if not provided, by ESC 2/1 7/14

For the G0 set by Registration ISO IR No. 6 : ESC 2/8 4/2

Such a version will generally be the subject of a specification document which states how the options in clauses 6, 7 and 8 have been exercised. Such a specification is said to be in accordance with this ECMA Standard.

10 Explanation of Code Tables No. 4 and No. 5

Table 4 is the basic 7-bit code table. It shows the 7-bit coded character set specified in clause 5 and indicates the options related to alternative graphic characters (6.4.2) and national or application-oriented use (6.4.3).

Bit combinations 0/0 to 1/15 are reserved for control characters and are shown empty. Bit combinations 2/0 to 7/15 represent a graphic symbol, an acronym, or a reference to a clause of this International Standard. A reference to 6.4.2 and 6.4.3 is denoted by ① or ②, respectively.

Table 5 is a code table showing a version which includes the G0 set of the IRV and no C0 set.

Table 4 - Basic Code Table

					b7	0	0	0	0	1	1	1	1
					b6	0	0	1	1	0	0	1	1
					b5	0	1	0	1	0	1	0	1
						0	1	2	3	4	5	6	7
b4	b3	b2	b1										
0	0	0	0	0			SP	0	⓪	P	⓪	p	
0	0	0	1	1			!	1	A	Q	a	q	
0	0	1	0	2			"	2	B	R	b	r	
0	0	1	1	3			# ⓪ £	3	C	S	c	s	
0	1	0	0	4			⓪ ⓪ \$	4	D	T	d	t	
0	1	0	1	5			%	5	E	U	e	u	
0	1	1	0	6			&	6	F	V	f	v	
0	1	1	1	7			'	7	G	W	g	w	
1	0	0	0	8			(8	H	X	h	x	
1	0	0	1	9)	9	I	Y	i	y	
1	0	1	0	10			*	:	J	Z	j	z	
1	0	1	1	11			+	;	K	⓪	k	⓪	
1	1	0	0	12			,	<	L	⓪	l	⓪	
1	1	0	1	13			-	=	M	⓪	m	⓪	
1	1	1	0	14			.	>	N	⓪	n	⓪	
1	1	1	1	15			/	?	O	—	O	DEL	

C0 Set

Table 5 - Version with the G0 Set of the IRV

					b7	0	0	0	0	1	1	1	1
					b6	0	0	1	1	0	0	1	1
					b5	0	1	0	1	0	1	0	1
						0	1	2	3	4	5	6	7
b4	b3	b2	b1										
0	0	0	0	0				SP	0	@	P	·	p
0	0	0	1	1				!	1	A	Q	a	q
0	0	1	0	2				"	2	B	R	b	r
0	0	1	1	3				#	3	C	S	c	s
0	1	0	0	4				\$	4	D	T	d	t
0	1	0	1	5				%	5	E	U	e	u
0	1	1	0	6				&	6	F	V	f	v
0	1	1	1	7				'	7	G	W	g	w
1	0	0	0	8				(8	H	X	h	x
1	0	0	1	9)	9	I	Y	i	y
1	0	1	0	10				*	:	J	Z	j	z
1	0	1	1	11				+	;	K	[k	{
1	1	0	0	12				,	<	L	\	l	
1	1	0	1	13				-	=	M]	m	}
1	1	1	0	14				.	>	N	^	n	~
1	1	1	1	15				/	?	O	_	o	DEL

Annex A
(normative)

Specification of the C0 set

A.1 C0 set of a version of Standard ECMA-6

The C0 set of a version of Standard ECMA-6 shall be a sub-set of the C0 set of Standard ECMA-48. The selected control functions shall be used according to their definitions in Standard ECMA-48. Bit combinations corresponding to control functions not selected shall be declared unused.

A.2 C0 set of the IRV

If the C0 set of the International Reference Version (IRV) is used, it shall be the C0 set of Standard ECMA-48.

A.3 Table A.1

Table A.1 shows, for convenience, the C0 set of ISO 6429. For the definition of these control functions see Standard ECMA-48.

Table A.1 - C0 Set of ECMA-48

				b7	0	0
				b6	0	0
				b5	0	1
				0		1
b4	b3	b2	b1			
0	0	0	0	0	NUL	DLE
0	0	0	1	1	SOH	DC1
0	0	1	0	2	STX	DC2
0	0	1	1	3	ETX	DC3
0	1	0	0	4	EOT	DC4
0	1	0	1	5	ENQ	NAK
0	1	1	0	6	ACK	SYN
0	1	1	1	7	BEL	ETB
1	0	0	0	8	BS	CAN
1	0	0	1	9	HT	EM
1	0	1	0	10	LF	SUB
1	0	1	1	11	VT	ESC
1	1	0	0	12	FF	IS ₁
1	1	0	1	13	CR	IS ₂
1	1	1	0	14	SO	IS ₃
1	1	1	1	15	SI	IS ₄

Annex B

(informative)

Guidelines for Standards derived from Standard ECMA-6

B.1 General

When drafting national or application-orientated standards based on ECMA-6, it is recommended that the following considerations be taken into account.

B.2 Structure of a Standard

It is recommended that the same structure and editorial style as implemented for ECMA-6 be adopted. All facilities, restrictions and specifications of the standard should be stated clearly in sentences using plain language, rather than being summarized in tables with notes.

B.2.1 Control functions

The standard should contain the selection of control characters from the C0 set of ECMA-48 and an explicit description of the corresponding control functions. Even where the descriptions are identical with those of ISO 6429, they should be stated explicitly, and not merely by reference to ECMA-48. For application-oriented standards the specific meaning of the Information Separators and of the Device Controls should be defined.

B.2.2 Graphic characters (see 8.3.3)

Where there is no need for particular characters, the graphic characters of the International Reference Version (IRV) should be allocated to the same bit combinations and with the same names as in ECMA-6.

B.2.3 Composite graphic characters and repertoire

ECMA-6 permits the construction of composite graphic characters by using the control characters BACKSPACE and CARRIAGE RETURN to image two or more graphic characters at the same character position.

The total number of graphic characters which can be obtained from any version of the character set, with or without using this facility, is called the repertoire. ECMA-6 does not define a particular repertoire. However, as the interpretation and/or the imaging of composite characters may cause difficulties, agreement between sender and recipient of the data may be required. To minimize the need for such agreements and to facilitate interchange, national or application-oriented standards may specify a standard repertoire of graphic characters which permit only a limited number of composite graphic characters. Such limitations are considered fully compatible with ECMA-6.

B.2.4 Versions

In a standard, one or more versions can be specified. It should be noted that a version is not a standard but only part of a standard. The standard itself consists of the C0 set and the well-defined version or versions and a set of clauses as mentioned above. The definition of a version requires that the options mentioned in 8.1 be properly exercised.

Annex C
(informative)

Differences between the fifth Edition (1983) and the present (sixth) Edition of this ECMA Standard

- 1) The new text for the conformance clause adopted for all code standards has been introduced.
- 2) The wording of all definitions has been aligned with that of ECMA-43, "8-Bit Coded Character Set, Structure and Rules" and ISO/IEC 10538, "Information technology - Control functions for text communication".
- 3) In the G0 set of the IRV the character DOLLAR SIGN replaces the character CURRENCY SIGN in position 2/4. Otherwise this G0 set is unchanged.
- 4) The requirements for the C0 set have been relaxed. The sole requirement is that the C0 set be a sub-set of the C0 set of the fifth edition of this Standard. This C0 set is shown in normative annex A. For the definitions of the 32 control characters reference is made to ECMA-48.

As a consequence the Basic Code Table (table 4) and table 5 showing a version of ECMA-6 with the G0 set of the IRV do not show any specific control characters in columns 0 and 1.

Printed copies can be ordered from:

ECMA

114 Rue du Rhône
CH-1204 Geneva
Switzerland

Fax: +41 22 849.60.01

Internet: documents@ecma.ch

A description of this Standard can be found on the ECMA web site. www.ecma.ch. From there, files E006-DOC.EXE (MSWord, self-expanding) and E006-PDF.PDF (Acrobat PDF) can be freely downloaded.

The ECMA web site, <http://www.ecma.ch>, gives full information on ECMA, ECMA activities, ECMA Standards and Technical Reports.

ECMA

**114 Rue du Rhône
CH-1204 Geneva
Switzerland**

This Standard ECMA-6 is available free of charge in printed form and as a file.

See inside cover page for instructions.