International Standard

ISO 4873

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION•МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ•ORGANISATION INTERNATIONALE DE NORMALISATION

Information processing — ISO 8-bit code for information interchange — Structure and rules for implementation

Traitement de l'information — Code ISO à 8 éléments pour l'échange d'information — Structure et règles de matérialisation

Second edition — 1986-07-15

UDC 681.3.04 : 003.62

Ref. No. ISO 4873-1986 (E)

Descriptors : data processing, information interchange, data transmission, data codes, coded character sets, structures, implementation.

ISO 4873-1986 (E)

Price based on 17 pages

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work.

Draft International Standards adopted by the technical committees are circulated to the member bodies for approval before their acceptance as International Standards by the ISO Council. They are approved in accordance with ISO procedures requiring at least 75 % approval by the member bodies voting.

International Standard ISO 4873 was prepared by Technical Committee ISO/TC 97, *Information processing systems.*

This second edition cancels and replaces the first edition (ISO 4873-1979), of which it constitutes a technical revision.

Users should note that all International Standards undergo revision from time to time and that any reference made herein to any other International Standard implies its latest edition, unless otherwise stated.

Contents

Information processing — ISO 8-bit code for information interchange — Structure and rules for implementation

1 Scope and field of application

This International Standard specifies an 8-bit code derived from, and compatible with, the 7-bit coded character set specified in ISO 646.

The characteristics of this code are also in conformance with the code extension techniques specified in ISO 2022.

This International Standard specifies an 8-bit code with a number of options. It also provides guidance on how to exercise the options to define specific versions.

This character set is primarily intended for general information interchange within an 8-bit environment 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.

This character set includes the 52 small and capital letters of the basic Latin alphabet and may include accented letters, special Latin letters and/or the letters of one or several non-Latin alphabet(s).

2 References

ISO 646, *Information processing — ISO 7-bit coded character set for information interchange.*

ISO 1177, *Information processing — Character structure for start/stop and synchronous character-oriented transmission.*

ISO 2022, *Information processing — ISO 7-bit and 8-bit coded character sets — Code extension techniques.*

ISO 6429, *Information processing — Additional control functions for character-imaging devices.*

ISO 6937/2, *Information processing — Coded character sets for text communication — Part 2: Latin alphabetic and non-alphabetic graphic characters.*

3 Conformance and implementation

3.1 Conformance

An 8-bit code is in conformance with this International Standard if it is a version in accordance with clause 8. Equipment

claimed to implement this International Standard shall be able to interchange information by means of a version of the 8-bit code at a specified level according to clauses 8 and 9; this version and level shall be identified in any such claim.

3.2 Implementation

The use of this code requires definitions of its implementation in various media. For example, these could include magnetic and optical media and transmission channels, thus permitting interchange of data to take place either indirectly by means of an intermediate recording in a physical medium, or by means of data transmission equipment.

The implementation of this code in physical media and for transmission, taking into account the need for error checking, is the subject of other International Standards.

4 Definitions

For the purpose of this International Standard the following definitions apply.

4.1 bit combination: An ordered set of bits that represents a character or is used as part of the representation of a character.

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

4.3 coded character set; code: A set of unambiguous rules that establishes a character set and the one-to-one relationship between each character of the set and its coded representation by one or more bit combinations.

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

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

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

4.7 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.

4.8 graphic character: A character, other than a control function, that has a visual representation normally handwritten, printed or displayed, and that has a coded representation consisting of one or more bit combinations.

4.9 graphic symbol: A visual representation of a graphic character or of a control function.

4.10 position: That part of a code table identified by its column and row co-ordinates.

5 Notation, code table and names

5.1 Notation

The bits of the bit combinations of the 8-bit code are identified by b_8, b_7, b_6, b_5, b_4, b_3, b_2 and b_1, where b_8 is the highest-order, or most-significant bit and b_1 is the lowest-order, or least-significant bit.

The bit combinations may be interpreted to represent numbers in the range 0 to 255 in binary notation by attributing the following weights to the individual bits:

Bit	b_8	b_7	b_6	b_5	b_4	b_3	b_2	b_1
Weight	128	64	32	16	8	4	2	1

In this International Standard, the bit combinations are identified by notations of the form xx/yy, where xx and yy are numbers in the range 00 to 15. The correspondence between the notations of the form xx/yy and the bit combinations consisting of the bits b_8 to b_1, is as follows:

— xx is the number represented by b_8, b_7, b_6 and b_5 where these bits are given the weights 8, 4, 2 and 1 respectively;

— yy 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 xx/yy are the same as the ones used to identify code table positions, where xx is the column number and yy is the row number (see 5.2).

5.2 Code table

An 8-bit code table consists of 256 positions arranged in 16 columns and 16 rows. The columns and rows are numbered 00 to 15.

The code table positions are identified by notations of the form xx/yy, where xx is the column number and yy 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 xx/yy, is the same as that of the corresponding bit combination.

5.3 Names

This International Standard assigns at least one name to each character. In addition, it specifies an acronym for each control character and for the character SPACE, and a graphic symbol for each graphic character. By convention, only capital letters, the graphic symbols for small letters and hyphens are used for writing the names of the characters. It is intended that the acronyms and this convention be retained in all translations of the text of this International Standard.

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

6 Structure of the 8-bit code

6.1 Elements of the 8-bit code

The 8-bit code consists of the following parts:

a) A C0 set: a set of 30 control characters represented by bit combinations 00/00 to 01/15, except 00/14 and 00/15 which shall be unused.

b) The character SPACE: a character represented by the bit combination 02/00, which may be interpreted as a control character, as a graphic character, or as both.

c) A G0 set: a set of up to 94 graphic characters represented by bit combinations 02/01 to 07/14.

d) The character DELETE: a control character represented by the bit combination 07/15.

e) A C1 set: a set of up 32 control characters represented by bit combinations 08/00 to 09/15.

f) A G1 set: a set of up to 96 graphic characters represented by bit combinations 10/00 to 15/15.

g) A G2 set: a set of up to 96 graphic characters.

h) A G3 set: a set of up to 96 graphic characters.

6.2 Initial condition

At the beginning of any information interchange, a default C0 set containing only the character ESCAPE shall be assumed.

6.3 Designation and invocation

For the sake of simplicity, this International Standard shows escape sequences of three characters. Escape sequences with more characters may also be used or required. The format and use of escape sequences are described in detail in ISO 2022.

6.3.1 C0 set

The C0 set shall be allocated to columns 00 and 01 of the code table. It shall be designated and invoked by an escape sequence of the type ESC 02/01 F.

6.3.2 Character SPACE

The character SPACE shall be allocated to column 02 in position 02/00. It is not explicitly designated or invoked.

6.3.3 G0 set

The G0 set shall be allocated to columns 02 to 07 in positions 02/01 to 07/14. It shall be designated by an escape sequence of the type ESC 02/08 F and invoked as specified in clause 9.

6.3.4 Character DELETE

The character DELETE shall be allocated to column 07 in position 07/15. It is not explicitly designated or invoked.

6.3.5 C1 set

The C1 set shall be allocated to columns 08 and 09. It shall be designated and invoked by an escape sequence of the type ESC 02/02 F.

6.3.6 G1 set

The G1 set shall be allocated to columns 10 to 15 when invoked. It shall be designated by an escape sequence of the type ESC 02/09 F or ESC 02/13 F and invoked as specified in clause 9.

6.3.7 G2 set

The G2 set shall be designated by an escape sequence of the type ESC 02/10 F or ESC 02/14 F. Either the set as a whole shall be invoked by the locking-shift function LS2R into columns 10 to 15, or individual characters of it shall be invoked by means of the single-shift function SS2.

6.3.8 G3 set

The G3 set shall be designated by an escape sequence of the type ESC 02/11 F or ESC 02/15 F. Either the set as a whole shall be invoked by the locking-shift function LS3R into columns 10 to 15, or individual characters of the set shall be invoked by means of the single-shift function SS3.

6.3.9 Summary of structure, designation and invocation

Figure 1 summarizes, in a schematic form, the structure of the 8-bit code and designation and allocation of its different parts.

7 Specification of the characters of the 8-bit code

7.1 C0 set

The requirements for the C0 set are

— bit combinations 00/14 and 00/15 shall not be used (see annex D);

— the control character ESCAPE shall be allocated to bit combination 01/11;

— any control characters can be allocated to the other bit combinations subject to the restrictions specified in annex A.

NOTE — A C0 set comprising only ESCAPE allocated to position 01/11 has been registered, and is identified by ESC 02/01 04/07.

7.2 Character ESCAPE

ESCAPE is a control character used to form escape sequences. In this International Standard escape sequences are used to announce the level at which a version is defined and to designate and invoke character sets.

Acronym	Name	Coded representation
E S C	ESCAPE	01/11

7.3 Character SPACE

This character may be interpreted as a graphic character, as a control character, or as both.

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

As a control character, it acts as a format effector that causes the active position to be advanced one character position.

Acronym	Name	Coded representation
S P	SPACE	02/00

7.4 G0 set

The 94 bit combinations 02/01 to 07/14 are used to represent graphic characters. All graphic characters allocated to bit combinations in the range 02/01 to 07/14 are spacing characters, i.e. they cause the active position to advance. Sub-clause 7.4.1 specifies the graphic characters allocated by this International Standard to 82 of these 94 bit combinations. Sub-clauses 7.4.2, 7.4.3 and 7.4.4 specify how further graphic characters may be allocated to the remaining twelve bit combinations.

Each of the sub-clauses 7.4.1, 7.4.2 and 7.4.4 contains a table consisting of three columns. The first column specifies the graphic symbol of each graphic character, the second column specifies the standard name of the graphic character and the third column specifies the bit combination representing the graphic character concerned.

7.4.1 Unique graphic character allocations

A unique graphic character is allocated to each of the 82 bit combinations 02/01, 02/02, 02/05 to 03/15, 04/01 to 05/10, 05/15 and 06/01 to 07/10. These characters are specified in table 1.
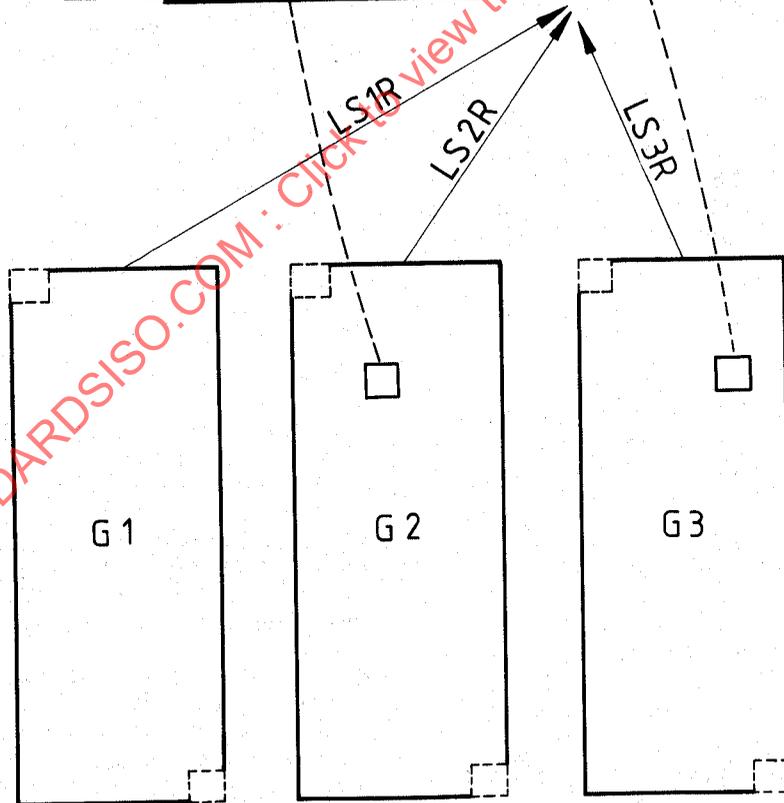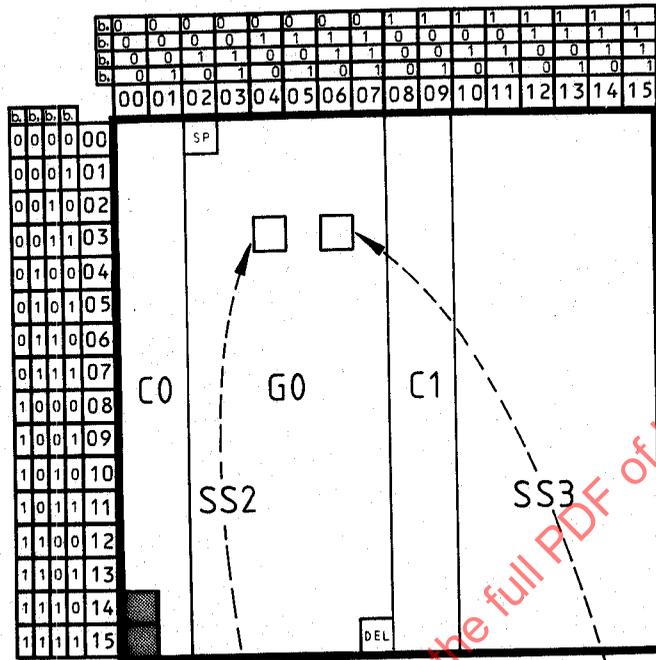
Figure 1 — Structure of the 8-bit code

Table 1 — Unique graphic character allocations

Graphic symbol	Name	Coded representation	Graphic symbol	Name	Coded representation
!	EXCLAMATION MARK	02/01	M	CAPITAL LETTER M	04/13
"	QUOTATION MARK	02/02	N	CAPITAL LETTER N	04/14
%	PERCENT SIGN	02/05	O	CAPITAL LETTER O	04/15
&	AMPERSAND	02/06	P	CAPITAL LETTER P	05/00
'	APOSTROPHE	02/07	Q	CAPITAL LETTER Q	05/01
(	LEFT PARENTHESIS	02/08	R	CAPITAL LETTER R	05/02
)	RIGHT PARENTHESIS	02/09	S	CAPITAL LETTER S	05/03
*	ASTERISK	02/10	T	CAPITAL LETTER T	05/04
+	PLUS SIGN	02/11	U	CAPITAL LETTER U	05/05
,	COMMA	02/12	V	CAPITAL LETTER V	05/06
−	HYPHEN, MINUS SIGN	02/13	W	CAPITAL LETTER W	05/07
.	FULL STOP	02/14	X	CAPITAL LETTER X	05/08
/	SOLIDUS	02/15	Y	CAPITAL LETTER Y	05/09
0	DIGIT ZERO	03/00	Z	CAPITAL LETTER Z	05/10
1	DIGIT ONE	03/01	_	LOW LINE, UNDERLINE	05/15
2	DIGIT TWO	03/02	a	SMALL LETTER a	06/01
3	DIGIT THREE	03/03	b	SMALL LETTER b	06/02
4	DIGIT FOUR	03/04	c	SMALL LETTER c	06/03
5	DIGIT FIVE	03/05	d	SMALL LETTER d	06/04
6	DIGIT SIX	03/06	e	SMALL LETTER e	06/05
7	DIGIT SEVEN	03/07	f	SMALL LETTER f	06/06
8	DIGIT EIGHT	03/08	g	SMALL LETTER g	06/07
9	DIGIT NINE	03/09	h	SMALL LETTER h	06/08
:	COLON	03/10	i	SMALL LETTER i	06/09
;	SEMICOLON	03/11	j	SMALL LETTER j	06/10
<	LESS-THAN SIGN	03/12	k	SMALL LETTER k	06/11
=	EQUALS SIGN	03/13	l	SMALL LETTER l	06/12
>	GREATER-THAN SIGN	03/14	m	SMALL LETTER m	06/13
?	QUESTION MARK	03/15	n	SMALL LETTER n	06/14
A	CAPITAL LETTER A	04/01	o	SMALL LETTER o	06/15
B	CAPITAL LETTER B	04/02	p	SMALL LETTER p	07/00
C	CAPITAL LETTER C	04/03	q	SMALL LETTER q	07/01
D	CAPITAL LETTER D	04/04	r	SMALL LETTER r	07/02
E	CAPITAL LETTER E	04/05	s	SMALL LETTER s	07/03
F	CAPITAL LETTER F	04/06	t	SMALL LETTER t	07/04
G	CAPITAL LETTER G	04/07	u	SMALL LETTER u	07/05
H	CAPITAL LETTER H	04/08	v	SMALL LETTER v	07/06
I	CAPITAL LETTER I	04/09	w	SMALL LETTER w	07/07
J	CAPITAL LETTER J	04/10	x	SMALL LETTER x	07/08
K	CAPITAL LETTER K	04/11	y	SMALL LETTER y	07/09
L	CAPITAL LETTER L	04/12	z	SMALL LETTER z	07/10

7.4.2 Alternative graphic character allocations

Two alternative graphic characters are allocated to each of the bit combinations 02/03 and 02/04. These characters are specified in table 2.

Table 2 — Alternative graphic character allocations

Graphic symbol	Name	Coded representation
£	POUND SIGN	02/03
#	NUMBER SIGN	02/03
$	DOLLAR SIGN	02/04
¤	CURRENCY SIGN	02/04

Either the character POUND SIGN or the character NUMBER SIGN shall be allocated to bit combination 02/03 and either the character DOLLAR SIGN or the character CURRENCY SIGN shall be allocated to bit combination 02/04.

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

7.4.3 National or application-orientated graphic character allocations

No specific graphic characters are allocated to the ten bit combinations 04/00, 05/11 to 05/14, 06/00 and 07/11 to 07/14. A unique graphic character shall be allocated to each of these bit combinations, or the bit combination shall be declared unused.

7.4.4 Recommendation for the allocation of graphic characters

It is recommended that the bit combinations listed in 7.4.2 and 7.4.3 have allocated to them the characters in existing national or International Standards. Those of the International Reference Version (IRV) of ISO 646, are shown in table 3.

Table 3 — IRV graphic character allocations

Graphic symbol	Name	Coded representation
#	NUMBER SIGN	02/03
¤	CURRENCY SIGN	02/04
@	COMMERCIAL AT	04/00
[	LEFT SQUARE BRACKET	05/11
\	REVERSE SOLIDUS	05/12
]	RIGHT SQUARE BRACKET	05/13
^	CIRCUMFLEX ACCENT	05/14
`	GRAVE ACCENT	06/00
{	LEFT CURLY BRACKET	07/11
|	VERTICAL LINE	07/12
}	RIGHT CURLY BRACKET	07/13
~	TILDE, OVERLINE	07/14

7.5 Character DELETE

DELETE is a character that was originally used to erase or obliterate an erroneous or unwanted character in punched tape. DEL characters may serve to accomplish media-fill or time-fill. They may be inserted into, or removed from, a stream of data without affecting the information content of that stream, but such action may affect the information layout and/or the control of equipment.

Acronym	Name	Coded representation
DEL	DELETE	07/15

7.6 C1 set

The C1 set is available for up to 32 control characters in addition to those provided by the C0 set. It shall not include any of the control characters listed in annex A.

No specific control characters are allocated to positions 08/00 to 08/13 and 09/00 to 09/15.

When the single-shift functions SS2 and SS3 are used, they shall be allocated to positions 08/14 and 08/15, respectively, otherwise these positions shall not be used.

7.7 G1 set

The G1 set shall be either a 94-character or a 96-character set of graphic characters.

This set is available for graphic characters in addition to those provided by the G0 set.

A unique graphic character shall be allocated to each position or the position shall be declared unused.

The characters of the G1 set are represented by bit combinations 10/01 to 15/14 if the G1 set is a 94-character set, or by bit combinations 10/00 to 15/15 if the G1 set is a 96-character set.

7.8 G2 set

The G2 set shall be either a 94-character or a 96-character set of graphic characters.

This set is available for graphic characters in addition to those provided by the G0 and the G1 sets.

A unique graphic character shall be allocated to each position or the position shall be declared unused.

If the G2 set is a 94-character set, then no characters shall be allocated to positions 10/00 and 15/15.

The characters of the G2 set shall be invoked either by the single-shift function SS2 or by the locking-shift function LS2R.

— When invoked by SS2, each character is represented by the bit combination of SS2 followed by one of the bit combinations in the range 02/01 to 07/14 if the G2 set is a 94-character set, or 02/00 to 07/15 if the G2 set is a 96-character set.

— When invoked by LS2R, the characters of the G2 set are represented by bit combinations 10/01 to 15/14 if the G2 set is a 94-character set, or by bit combinations 10/00 to 15/15 if the G2 set is a 96-character set.

7.9 G3 set

The G3 set shall be either a 94-character or a 96-character set of graphic characters. This set is available for graphic characters in addition to those provided by the G0, the G1 and the G2 sets.

A unique graphic character shall be allocated to each position or the position shall be declared unused.

If the G3 set is a 94-character set, then no character shall be allocated to positions 10/00 and 15/15.

The characters of the G3 set shall be invoked either by the single-shift function SS3 or by the locking-shift function LS3R.

— When invoked by SS3, each character is represented by the bit combination of SS3 followed by one of the bit combinations in the range 02/01 to 07/14 if the G3 set is a 94-character set, or 02/00 to 07/15 if the G3 set is a 96-character set.

— When invoked by LS3R, the characters of the G3 set are represented by bit combinations 10/01 to 15/14 if the G3 set is a 94-character set, or by bit combinations 10/00 to 15/15 if the G3 set is a 96-character set.

7.10 Summary of the specification of the 8-bit code

Figure 2 summarizes, in a schematic form, the specification of the elements of the 8-bit code.

8 Versions of the 8-bit code

A version of the 8-bit code is a coded character set comprising a C0, a G0, a C1 and a G1 set and, optionally a G2 and a G3 set, with a specification of the necessary escape sequences to designate them. In a version the same character shall not be allocated to more than one of the G0, G1, G2 or G3 sets. In specifying the G0 set the following options shall be exercised:

a) one of the alternative graphic characters specified in 7.4.2 shall be allocated to each of the bit combinations 02/03 and 02/04, and

b) up to ten unique graphic characters shall be allocated to the bit combinations 04/00, 05/11 to 05/14, 06/00 and 07/11 to 07/14 specified in 7.4.3, and any of these bit combinations to which no character is allocated shall be declared unused.

9 Levels

This International Standard specifies three nested levels of implementation each of which may be identified by an announcer sequence.

ESC 02/00 04/12 shall identify Level 1

ESC 02/00 04/13 shall identify Level 2

ESC 02/00 04/14 shall identify Level 3

When announcer sequences are used (see clauses 10 and 11) the announcer sequence of a given level shall be correctly interpreted at any other level.

9.1 Level 1

Level 1, see figure 3, comprises the following facilities:

— the C0 set, designated and invoked by ESC 02/01 F;

— the character SPACE in position 02/00;

— the G0 set, designated by ESC 02/08 F;

— the character DELETE in position 07/15;

— a C1 set, designated and invoked by ESC 02/02 F;

— a G1 set, designated by ESC 02/09 F or ESC 02/13 F.

At Level 1 no shift functions shall be used and the G0 and G1 sets are assumed to be invoked permanently in columns 02 to 07 and 10 to 15, respectively. Therefore an escape sequence that designates the G0 or the G1 set also implicitly invokes that set.

At Level 1 the C1 set and/or the G1 set may be empty if there is no requirement for control characters in addition to those provided by the C0 set and/or for graphic characters in addition to these provided by the G0 set. An empty C1 set shall be designated and invoked by ESC 02/02 07/14 and an empty G1 set by ESC 02/09 07/14 or ESC 02/13 07/14.

At Level 1, G2 and G3 sets shall not be designated.

9.2 Level 2

Level 2, see figure 4, comprises the facilities of Level 1 and in addition to them

— a G2 set, designated by ESC 02/10 F or ESC 02/14 F, the characters of which shall be invoked individually by SS2;

— a G3 set, designated by ESC 02/11 F or ESC 02/15 F, the characters of which shall be invoked individually by SS3.
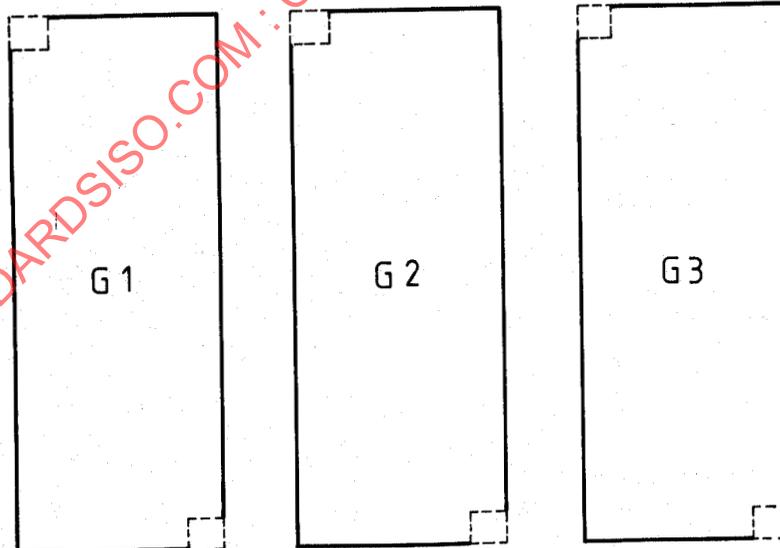
At Level 2 no other shift functions shall be used.

The G1 set shall not be empty, either the G2 or the G3 set can be designated as empty but not both. The G2 set may be empty if there is no requirement for graphic characters in addition to those provided by the G0, G1 and G3 sets. In this case the G2 set shall be designated by ESC 02/10 07/14 or ESC 02/14 07/14. The G3 set may be empty if there is no requirement for graphic characters in addition to those provided by the G0, G1 and G2 sets. In this case the G3 set shall be designated by ESC 02/11 07/14 or ESC 02/15 07/14.

The C1 set shall not be empty, it shall contain at least the single-shift functions SS2 and SS3 in positions 08/14 and 08/15, respectively.

NOTE — A C1 set comprising only SS2 and SS3 allocated to these positions has been registered, and is identified by ESC 02/02 04/07.

7

				00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
0	0	0	0	00		SP	0	②	P	②	p								
0	0	0	1	01		!	1	A	Q	a	q								
0	0	1	0	02		"	2	B	R	b	r								
0	0	1	1	03		£(#)①	3	C	S	c	s								
0	1	0	0	04		$(¤)①	4	D	T	d	t								
0	1	0	1	05		%	5	E	U	e	u								
0	1	1	0	06		&	6	F	V	f	v								
0	1	1	1	07		'	7	G	W	g	w								
1	0	0	0	08		(	8	H	X	h	x								
1	0	0	1	09		)	9	I	Y	i	y								
1	0	1	0	10		*	:	J	Z	j	z								
1	0	1	1	11	ESC	+	;	K	②	k	②									
1	1	0	0	12		,	<	L	②	l	②								
1	1	0	1	13		−	=	M	②	m	②								
1	1	1	0	14		.	>	N	②	n	②	③							
1	1	1	1	15		/	?	O	_	o	DEL	③							

G 1   G 2   G 3

① See 7.4.2 and 7.4.4

② See 7.4.3 and 7.4.4

③ See 7.6

Figure 2 — Elements of the 8-bit code

9.3 Level 3

Level 3, see figure 5, comprises all the facilities of Level 2, with the addition of the following three shift functions:

— LS1R

— LS2R

— LS3R

The G1 set shall not be empty, either the G2 or the G3 set can be designated as empty as specified in 9.2, but not both.

The G1, G2 and G3 sets can be invoked explicitly by LS1R, LS2R and LS3R, respectively. Individual characters of the G2 and G3 sets can be invoked by SS2 and SS3, respectively.

The C1 set shall not be empty, it shall contain at least the single-shift functions SS2 and SS3 in positions 08/14 and 08/15, respectively.

10 Switching from one version to another

In information interchange, re-designation of any of the C or G sets requires an announcer sequence as specified in clause 9, followed by designating escape sequences as specified in 9.1, 9.2 or 9.3.

11 Switching from one level to another

Switching from one level to another, whether or not the contents of the C and G sets are changed, requires the new level to be announced and all sets to be designated.
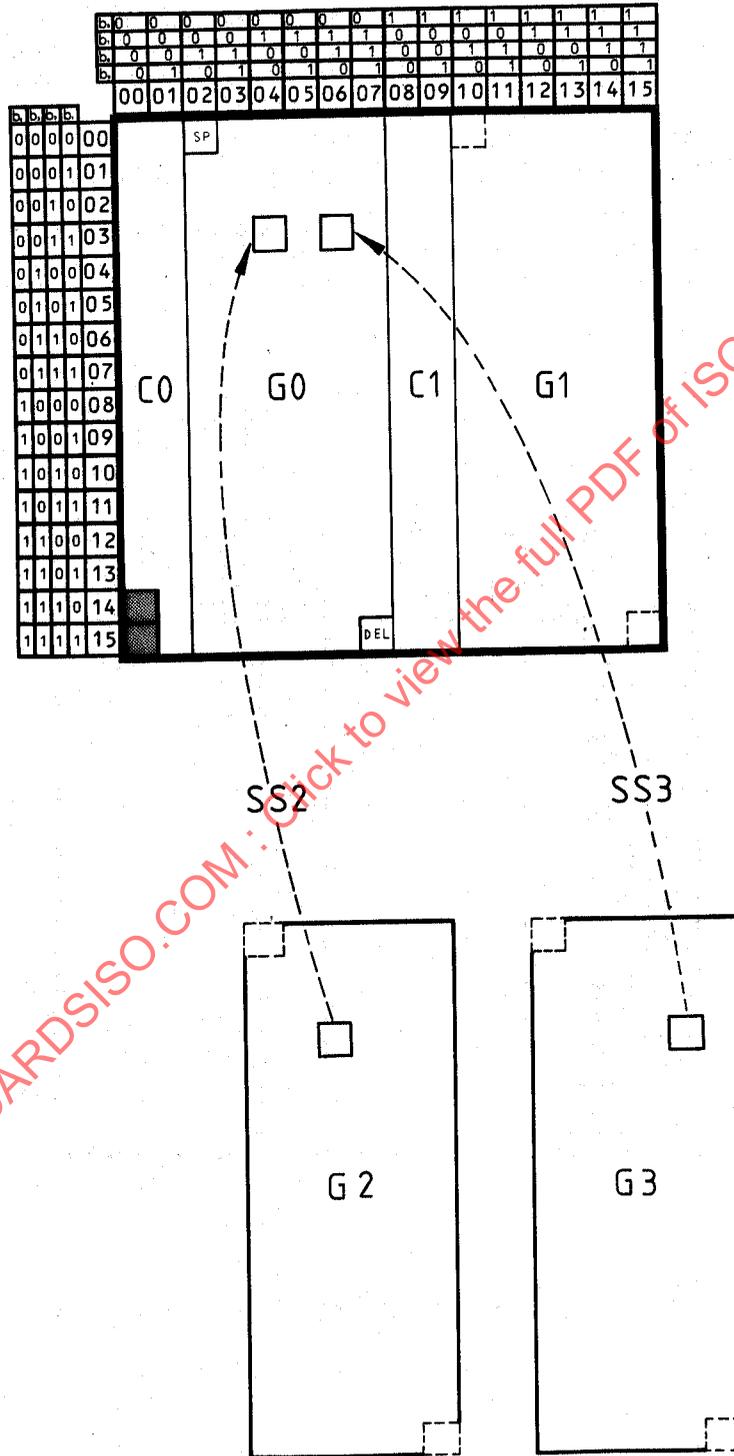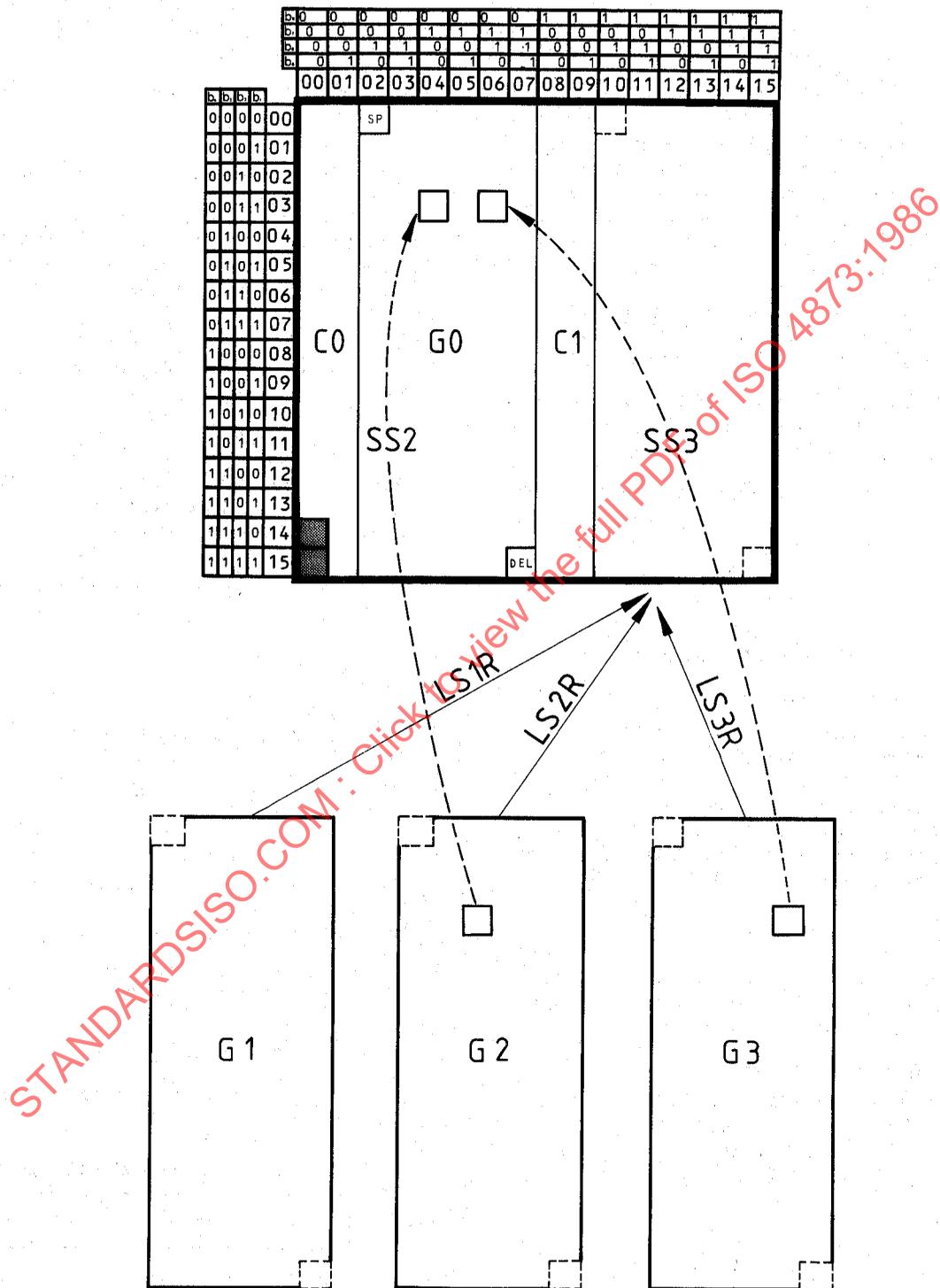
Figure 3 — Level 1

Figure 4 — Level 2

Figure 5 — Level 3

Annex A

Restrictions applicable to the C0 and C1 sets

(This annex forms part of the standard.)

The definitions of some control functions in this International Standard assume that data associated with them *is 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 these control functions 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.

Whilst this International Standard specifies requirements for bit combinations 00/14, 00/15 and 01/11 of the C0 set, it places the following restrictions on the use of the remaining 29 positions:

   a)   If *control characters* described in ISO 646 are used, they shall have the coded representation and the definitions specified below.

   b)   None of these control characters may be allocated to the C1 set.

A.1   General

The control characters are classified in the following categories:

   a)   *Transmission control characters*

Transmission control characters are intended to control or facilitate *transmission* of information over telecommunication networks. Procedures for the use of the transmission control characters on telecommunication networks are the subject of other International Standards. In particular, ISO 1177 specifies that transmission control characters shall be transmitted in 7-bit form plus *a parity bit in the eighth bit* position.

   b)   Format effectors

Format effectors are mainly intended for the control of the layout and positioning of information on character-imaging devices such as printing and display devices.

   c)   Code extension control characters

Code extension control characters shall be used only for designation and invocation of the character sets of the code. Procedures for the use of the code extension control characters are specified in ISO 2022.

   d)   Device control characters

Device control characters are intended for the control of local or remote devices or ancillary devices connected to a data processing or data communication system. These control characters are not intended to control data communication systems; this should be achieved by the use of transmission control characters.

   e)   Information separators

Information separators are used to separate and qualify data logically. There are four such characters. They may be used either in hierarchical order or non-hierarchically;

in the latter case, their specific meanings depend on the application.

   f)   Other control characters

Other *control characters are* those that fall outside the preceding categories.

The composition of each category, and the allocation of the individual control characters in each category to bit combinations in the range 00/00 to 01/15 are specified in clauses A.2 to A.6.

A.2   Transmission control characters

The *transmission control* characters and their coded representations are specified in table 4.

Table 4 — Transmission control characters — Coded representation

Acronym	Name	Coded representation
SOH	START OF HEADING	00/01
STX	START OF TEXT	00/02
ETX	END OF TEXT	00/03
EOT	END OF TRANSMISSION	00/04
ENQ	ENQUIRY	00/05
ACK	ACKNOWLEDGE	00/06
DLE	DATA LINK ESCAPE	01/00
NAK	NEGATIVE ACKNOWLEDGE	01/05
SYN	SYNCHRONOUS IDLE	01/06
ETB	END OF TRANSMISSION BLOCK	01/07

A.3   Format effectors

The format effectors and their coded representations are specified in table 5.

Table 5 — Format effectors — Coded representation

Acronym	Name	Coded representation
BS	BACKSPACE	00/08
HT	HORIZONTAL TABULATION	00/09
LF	LINE FEED	00/10
VT	VERTICAL TABULATION	00/11
FF	FORM FEED	00/12
CR	CARRIAGE RETURN	00/13

A.3.1 Concepts

The definitions of the format effectors use the following concepts:

a) A page is composed of a number of lines, each being composed of a number of character positions.

b) Each character position is capable of imaging SPACE or a graphic symbol.

c) The graphic symbol imaged at a character position represents a graphic character, a control function, or a combination of two or more graphic characters.

d) The active position is the character position at which the action required by the next character in the data stream is to be effected. If the next character is a graphic character, it is imaged at that position; if it is a control character, the corresponding function is performed relative to that position.

e) Movements of the active position are effected as follows:

1) The active position is advanced one character position immediately after imaging a SPACE or a graphic symbol [see c) above], and upon the execution of the function corresponding to a control character for which a graphic symbol is required to be imaged.

2) The active position is moved to a specified character position upon the execution of the function corresponding to a control character that is defined to cause a movement of the active position (i.e. a format effector).

f) The active position is not moved upon the execution of the function corresponding to a control character that is neither required to be imaged by a graphic symbol nor defined to cause a movement of the active position.

g) The effect of an attempt to move the active position beyond the boundaries of a line or a page is not defined by this International Standard.

A.3.2 Combined horizontal and vertical movements of the active position

The format effectors of the C0 set are defined for applications in which horizontal and vertical movements of the active position are effected separately.

This International Standard does not allow the use of one of these format effectors for effecting a combined vertical and horizontal movement. If a single control character is required for such a combined movement, it shall be allocated to the C1 set (see for example ISO 6429).

A.4 Device control characters

The device control characters and their coded representations are specified in table 6.

Table 6 — Device control characters — Coded representation

Acronym	Name	Coded representation
DC1	DEVICE CONTROL ONE	01/01
DC2	DEVICE CONTROL TWO	01/02
DC3	DEVICE CONTROL THREE	01/03
DC4	DEVICE CONTROL FOUR	01/04

A.5 Information separators

The information separators and their coded representations are specified in table 7.

Table 7 — Information separators — Coded representation

Acronym	Name	Coded representation
IS4 (FS)	INFORMATION SEPARATOR FOUR (FILE SEPARATOR)	01/12
IS3 (GS)	INFORMATION SEPARATOR THREE (GROUP SEPARATOR)	01/13
IS2 (RS)	INFORMATION SEPARATOR TWO (RECORD SEPARATOR)	01/14
IS1 (US)	INFORMATION SEPARATOR ONE (UNIT SEPARATOR)	01/15

Each information separator is given two names. The names INFORMATION SEPARATOR FOUR, INFORMATION SEPARATOR THREE, INFORMATION SEPARATOR TWO and INFORMATION SEPARATOR ONE are the general names. The names FILE SEPARATOR, GROUP SEPARATOR, RECORD SEPARATOR and UNIT SEPARATOR are the specific names and are intended mainly for applications where the information separators are used hierarchically. The ascending order is then US, RS, GS, FS. In this case data normally delimited by a particular separator cannot be split by a higher-order separator but will be considered as delimited by any higher-order separator.

A.6 Other control characters

The control characters outside the categories in A.2 to A.5 and their coded representations, are specified in table 8.

Table 8 — Other control characters — Coded representation

Acronym	Name	Coded representation
NUL	NULL	00/00
BEL	BELL	00/07
CAN	CANCEL	01/08
EM	END OF MEDIUM	01/09
SUB	SUBSTITUTE CHARACTER	01/10

A.7 Description of the control characters

The control characters are listed below in the alphabetical order of their acronyms.

A.7.1 ACK ACKNOWLEDGE

A transmission control character transmitted by a receiver as an affirmative response to the sender.

A.7.2 BEL BELL

A control character that is used when there is a need to call for attention; it may control alarm or attention devices.

A.7.3 BS BACKSPACE

A format effector which causes the active position to move one character position backwards.

A.7.4 CAN CANCEL

A character, or the first character of a sequence, indicating that the data preceding it is in error. As a result, this data shall be ignored. The specific meaning of this character shall be defined for each application and/or between sender and recipient.

A.7.5 CR CARRIAGE RETURN

A format effector which causes the active position to move to the first character position on the same line.

A.7.6 DC1 DEVICE CONTROL ONE

A device control character which is primarily intended for turning on or starting an ancillary device. If it is not required for this purpose, it may be used to restore a device to the basic mode of operation (see also DC2 and DC3), or for any other device control function not provided by other DCs.

A.7.7 DC2 DEVICE CONTROL TWO

A device control character which is primarily intended for turning on or starting an ancillary device. If it is not required for this purpose, it may be used to set a device to a special mode of operation (in which case DC1 is used to restore the device to the basic mode), or for any other device control function not provided by other DCs.

A.7.8 DC3 DEVICE CONTROL THREE

A device control character which is primarily intended for turning off or stopping an ancillary device. This function may be a secondary level stop, for example wait, pause, stand-by or halt (in which case DC1 is used to restore normal operation). If it is not required for this purpose, it may be used for any other ancillary device control function not provided by other DCs.

A.7.9 DC4 DEVICE CONTROL FOUR

A device control character which is primarily intended for turning off, stopping or interrupting an ancillary device. If it is not required for this purpose, it may be used for any other device control function not provided by other DCs.

A.7.10 DLE DATA LINK ESCAPE

A transmission control character which will change the meaning of a limited number of contiguously following bit combinations. It is used exclusively to provide supplementary transmission control functions. Only graphic characters and transmission control characters can be used in DLE sequences.

A.7.11 EM END OF MEDIUM

A control character that may be used to identify the physical end of a medium, or the end of the used portion of a medium, or the end of the wanted portion of data recorded on a medium. The position of this character does not necessarily correspond to the physical end of the medium.

A.7.12 ENQ ENQUIRY

A transmission control character used as a request for a response from a remote station — the response may include station identification and/or station status. When a "Who are you" function is required on the general switched transmission network, the first use of ENQ after the connection is established shall have the meaning "Who are you" (station identification). Subsequent use of ENQ may, or may not, include the function "Who are you", as determined by agreement.

A.7.13 EOT END OF TRANSMISSION

A transmission control character used to indicate the conclusion of the transmission of one or more texts.

A.7.14 ETB END OF TRANSMISSION BLOCK

A transmission control character used to indicate the end of a transmission block of data where data is divided into such blocks for transmission purposes.

A.7.15 ETX END OF TEXT

A transmission control character which terminates a text.

A.7.16 FF FORM FEED

A format effector which causes the active position to advance to the corresponding character position on a pre-determined line of the next form or page.

A.7.17 HT HORIZONTAL TABULATION

A format effector which causes the active position to advance to the next pre-determined character position.

A.7.18 IS1 (US) INFORMATION SEPARATOR ONE (UNIT SEPARATOR)

A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a unit.

A.7.19 IS2 (RS) INFORMATION SEPARATOR TWO (RECORD SEPARATOR)

A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a record.

A.7.20 IS3 (GS) INFORMATION SEPARATOR THREE (GROUP SEPARATOR)

A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a group.

A.7.21 IS4 (FS) INFORMATION SEPARATOR FOUR (FILE SEPARATOR)

A control character used to separate and qualify data logically; its specific meaning has to be defined for each application. If this character is used in hierarchical order as specified in the general definition of IS, it delimits a data item called a file.

A.7.22 LF LINE FEED

A format effector which causes the active position to advance to the corresponding character position of the next line.

A.7.23 NAK NEGATIVE ACKNOWLEDGE

A transmission control character transmitted by a receiver as a negative response to the sender.

A.7.24 NUL NULL

A control character used to accomplish media-fill or time-fill. NUL characters may be inserted into, or removed from, a stream of data without affecting the information content of that stream, but such action may affect the information layout and/or the control of equipment.

A.7.25 SOH START OF HEADING

A transmission control character used as the first character of a heading of an information message.

A.7.26 STX START OF TEXT

A transmission control character which precedes a text and which is used to terminate a heading.

A.7.27 SUB SUBSTITUTE CHARACTER

A control character used in the place of a character that has been found to be invalid or in error. SUB is intended to be introduced by automatic means.

A.7.28 SYN SYNCHRONOUS IDLE

A transmission control character used by a synchronous transmission system in the absence of any other character (idle condition) to provide a signal from which synchronism may be achieved or retained between data terminal equipment.

A.7.29 VT VERTICAL TABULATION

A format effector which causes the active position to advance to the corresponding character position on the next pre-determined line.