INTERNATIONAL TELECOMMUNICATION UNION

CCITT

THE INTERNATIONAL
TELEGRAPH AND TELEPHONE
CONSULTATIVE COMMITTEE

T.50

(09/92)

TERMINAL EQUIPMENT AND PROTOCOLS FOR TELEMATIC SERVICES

INTERNATIONAL REFERENCE ALPHABET (IRA)

(FORMERLY INTERNATIONAL ALPHABET No. 5 or IA5)

INFORMATION TECHNOLOGY – 7-BIT CODED CHARACTER SET FOR INFORMATION INTERCHANGE

Recommendation T.50

FOREWORD

The CCITT (the International Telegraph and Telephone Consultative Committee) is a permanent organ of the International Telecommunication Union (ITU). CCITT is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The Plenary Assembly of CCITT which meets every four years, establishes the topics for study and approves Recommendations prepared by its Study Groups. The approval of Recommendations by the members of CCITT between Plenary Assemblies is covered by the procedure laid down in CCITT Resolution No. 2 (Melbourne, 1988).

Recommendation (revised) T.50 was prepared by Study Group VIII and was approved under the Resolution No. 2 procedure on the 18th of September 1992.

CCITT NOTES

1) In this Recommendation, the expression “Administration” is used for conciseness to indicate both a telecommunication administration and a recognized private operating agency.

2) A list of abbreviations used in this Recommendation can be found in Annex C.

Recommendation T.50

INTERNATIONAL REFERENCE ALPHABET (IRA)

(Formerly International Alphabet No. 5 or IA5)

Information technology – 7-bit coded character set for information interchange

(revised 1992)

1 Scope

1.1 This Recommendation 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 Recommendation 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 ISO 2022.

1.6 The definitions of the control characters mentioned in this Recommendation are specified in Annex A. 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.

1.7 This is a paired Recommendation/International Standard, equivalent in technical content to ISO/IEC 646 (1991).

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 Recommendation 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 Recommendation if it conforms to the requirements of § 2.1.2.1, and either or both of §§ 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 Recommendation 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 the user, as specified respectively in §§ 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 § 2.1.1, 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 Recommendations.

3 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this Recommendation are encouraged to investigate the possibility of applying the most recent editions of the standards listed below.

ISO 2022: 1986 *Information processing – ISO 7-bit and 8-bit coded character sets – Coded extension techniques*.

4 Definitions

For the purpose of this Recommendation 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 – In general, the active position is indicated 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 representation by one or more 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 1 – In a communication environment according to the Reference Model for Open Systems Interconnection of Recommendation X.200 (ISO 7498) a CC-data-element will form all or part of the information that corresponds to the application user data and the presentation protocol data unit (PPDU) defined in that Recommendation.

Note 2 – 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 combinations.

4.11 **device**

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

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

4.12 **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.

4.13 **graphic symbol**

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

4.14 **position**

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

4.15 **repertoire**

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

4.16 **user**

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

Note 1 – This entity may be a process such as an application program if the "device" is a code convertor or a gateway function, for example.

Note 2 – The characters, as supplied by the user or made available to the user, may be in the form of codes local to the device, or of non-conventional visible representations, provided that § 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 b_7, b_6, b_5, b_4, b_3, b_2, and b_1, where b_7 is the highest-order, or most-significant, bit and b_1 is the lowest order, or least-significant, bit.

The bit combination 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 Recommendation, 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 § 9).

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 position of the code table is 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.

The allocation of individual characters to the bit combinations of the 7-bit code is specified in §§ 6.2 to 6.5.

5.3 *Names*

This Recommendation assigns 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, 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 translation of the text.

The names chosen to denote graphic characters are intended to reflect their customary meaning. However, this Recommendation 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 Recommendation shall have the following structure:

– a C0 control character set of up to 32 control characters allocated to positions 0/0 to 1/15;

– the graphic character SPACE (SP) allocated to position 2/0;

– a G0 graphic character set of up to 94 graphic characters allocated to positions 2/1 to 7/14;

– the character DELETE (DEL) allocated to position 7/15.

6.2 *Control characters*

The control characters allocated to positions 0/0 to 1/15 shall be as specified in Annex A.

6.3 *Character SPACE*

The acronym of the character SPACE is SP and its coded representation is 2/0.

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

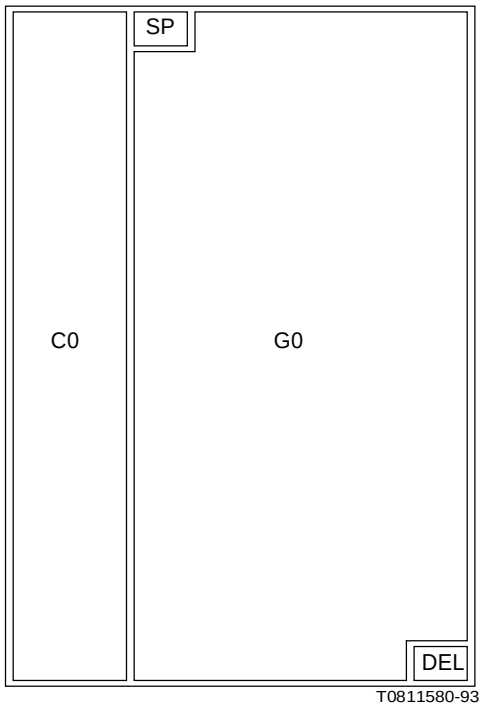

FIGURE 1/T.50

Structure of the 7-bit coded character set

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 characters are specified in Table 1/T.50.

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/T.50.

TABLE 1/T.50

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 sign	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
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, underline	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 z	7/10

TABLE 2/T.50

Alternative graphic character allocations

Graphic symbol	Name	Coded representation
£	Pound sign	2/3
#	Number sign	2/3
$	Dollar sign	2/4
¤	Currency sign	2/4

Either the character POUND SIGN or the character NUMBER 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 § 8).

Unless otherwise agreed between sender and recipient, the graphic symbols £, $, ¤ 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 § 8.1).

6.5 *Character DELETE*

The acronym of the character DELETE is DEL and its coded representation is 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 Recommendation 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 § 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*

The basic code table is shown in Table 4/T.50.

In order to use the 7-bit coded character set for information interchange, it is necessary to exercise the options left open in § 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 T.50”.

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 § 10). The graphic characters allocated to the G0 set of the IRV shall be as specified in Tables 1/T.50 and 3/T.50. See Table 5/T.50. Agreement between sender and recipient of the data shall be required if composite characters are used.

The C0 set of the IRV is as specified in Annex A.

TABLE 3/T.50

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 Recommendation 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 Recommendation 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 Recommendation.

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 Recommendation. Such coded representation 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 Recommendation 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:

– Registration ISO IR No. 1: ESC 2/1 4/0;

– Registration ISO IR No. 6: ESC 2/8 4/2.

Note – This IRV defers from the previous IRV of IA5.

10 Explanation of code Tables 4/T.50 and 5/T.50

Table 4/T.50 is the basic 7-bit code table. It shows the 7-bit coded character set specified in § 5 and indicates the options related to alternative graphic characters (see § 6.4.2) and national or application-oriented use (see § 6.4.3).

Positions 0/0 to 1/15 reserved for control characters are shown empty. Positions 2/0 to 7/15 contain a graphic symbol, an acronym, or a reference to a section of this Recommendation. A reference to §§ 6.4.2 and 6.4.3 is denoted by 1 or 2 respectively.

Table 5/T.50 is a code table showing a version which includes the G0 set of the IRV and no C0 set.

TABLE 4/T.50

Basic code table

b_7	0	0	0	0	1	1	1	1
b_6	0	0	1	1	0	0	1	1
b_5	0	1	0	1	0	1	0	1
b_4 b_3 b_2 b_1	0	1	2	3	4	5	6	7
0 0 0 0	0		SP	0	②	P	②	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		‼ ② £	3	C	S	c	s
0 1 0 0	4		¤ ① $	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	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
1 1 1 1	15		/	?	O	_	o	DEL

C0 set

T0811590-93

TABLE 5/T.50

Version with the G0 set of the IRV

b_7	0	0	0	0	1	1	1	1
b_6	0	0	1	1	0	0	1	1
b_5	0	1	0	1	0	1	0	1
b_4 b_3 b_2 b_1	0	1	2	3	4	5	6	7
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

T0811600-93

ANNEX A

(to Recommendation T.50)

(Normative)

Specification of the C0 set

A.1 *Definition of control set*

Table A-1/T.50 shows the C0 set.

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

A.1.1 **ACK acknowledge**

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

A.1.2 **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.1.3 **BS backspace**

A format effector which causes the active position to move one character position in the direction opposite to the transmission of one or more texts.

A.1.4 **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.1.5 **CR carriage return**

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

A.1.6 **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. When used for data flow control, DC1 is sometimes called “X-ON”.

A.1.7 **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.1.8 **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. When used for data flow control, DC3 is sometimes called "X-OFF".

A.1.9 **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.1.10 **DEL delete**

A character used primarily to erase or obliterate an erroneous or unwanted character in punched tape. DEL characters may also 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.

A.1.11 **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.1.12 **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.

A.1.13 **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.1.14 **EOT end of transmission**

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

A.1.15 **ESC escape**

A control character which is used for code extension purposes. It causes the meanings of a limited number of bit combinations following it in the data stream to be changed. The use of ESC is defined in ISO 2022.

A.1.16 **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.1.17 **ETX end of text**

A transmission control character which indicates the end of a text.

A.1.18 **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.1.19 **HT horizontal tabulation**

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

A.1.20 **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.1.21 **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.1.22 **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.1.23 **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.1.24 **LF line feed**

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

A.1.25 **NAK negative acknowledge**

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

A.1.26 **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.1.27 **SI shift-in**

A control character which is used for code extension purposes. It causes the meanings of the bit combinations following it in the data stream to be changed. The use of SI is defined in ISO 2022.

A.1.28 **SO shift-out**

A control character which is used for code extension purposes. It causes the meanings of the bit combinations following it in the data stream to be changed. The use of SO is defined in ISO 2022.

A.1.29 **SOH start of heading**

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

A.1.30 **STX start of text**

A transmission control character which is used to indicate the beginning of a text and the end of a heading.

A.1.31 **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.1.32 **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.1.33 **VT vertical tabulation**

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

A.2 *C0 set of a version of Recommendation T.50*

The C0 set of a version of this Recommendation shall be a subset of the C0 set defined in § A.1. Code positions corresponding to control functions not selected shall be declared unused.

A.3 *C0 set of the International Reference Version (IRV)*

The C0 set of the International Reference Version (IRV) is as shown in Table A-1/T.50

TABLE A-1/T.50
C0 set of the IRV

b_7	0	0
b_6	0	0
b_5	0	1
	0	1

b_4	b_3	b_2	b_1			
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	IS4
1	1	0	1	13	CR	IS3
1	1	1	0	14	SO	IS2
1	1	1	1	15	SI	IS1

T0811610-93

ANNEX B

(to Recommendation T.50)

(Informative)

Guidelines for standards derived from Recommendation T.50

B.1 *General*

When drafting national or application-oriented standards based on this Recommendation, 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 this Recommendation 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 and an explicit description of the corresponding control functions. Even where the descriptions are identical with those of this Recommendation, they should be stated explicitly, and not merely by reference to this Recommendation. 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 positions and with the same names as in this Recommendation.

B.2.3 *Composite graphic characters and repertoire*

This Recommendation 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. This Recommendation 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 this Recommendation.

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

(to Recommendation T.50)

Alphabetical list of abbreviations used in this Recommendation

ACK	Acknowledge
BEL	Bell
BS	Backspace
CAN	Cancel
CR	Carriage return
DC1	Device control one
DC2	Device control two
DC3	Device control three
DC4	Device control four
DEL	DELETE
DEL	Delete
DLE	Data link escape
EM	End of medium
ENQ	Enquiry
EOT	End of transmission
ESC	Escape
ETB	End of transmission block
ETX	End of text
FF	Form feed
FS	File separator
GS	Group separator
HT	Horizontal tabulation
IA5	International Alphabet No. 5
IRA	International reference alphabet
IRV	International reference version
IS1	Information separator one
IS2	Information separator two
IS3	Information separator three
IS4	Information separator four
LF	Line feed
NAK	Negative acknowledge

NUL	Null
RS	Record separator
SI	Shift-in
SO	Shift-out
SOH	Start of heading
SP	SPACE
STX	Start of text
SUB	Substitute character
SYN	Synchronous idle
US	Unit separator
VT	Vertical tabulation