## T. 51

THE INTERNATIONAL
TELEGRAPH AND TELEPHONE CONSULTATIVE COMMITTEE

TERMINAL EQUIPMENT AND PROTOCOLS FOR TELEMATIC SERVICES

## LATIN BASED CODED CHARACTER SETS FOR TELEMATIC SERVICES

Recommendation T. 51

## 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. 51 was revised by Study Group VIII and was approved under the Resolution No. 2 procedure on the 18 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.

All rights reserved. No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the ITU.

## Recommendation T. 51

# LATIN BASED CODED CHARACTER SETS FOR TELEMATIC SERVICES 

(Malaga-Torremolinos, 1984: Amended at Melbourne 1988; revised 1992)

## 1 Scope

1.1 The CCITT,
considering
(a) the increasing interdependence of the various CCITT character sets and coding schemes in various telematic services;
(b) the introduction of new facilities such as code conversion and interworking between various telematic services;
(c) the advantage of having a single unified repertoire and coding of Latin based character set in a Recommendation to act as a reference for the telematic services;
(d) that Recommendations T. 61 and T.100/T. 101 define the character coding systems for teletex and videotex;
(e) that Recommendation T. 50 specifies the International Reference Version (IRV) of the 7-bit coded character set,

## provides the following Recommendation as a reference document

towards which the Latin based portion of the coded character sets of telematic services should migrate and from which coded character subsets and elements of code extension mechanisms can be derived for individual telematic services.
1.2 This Recommendation specifies a primary set and a 96-character supplementary set of graphic characters. When various telematic services restrict their primary and supplementary sets to be respective subsets of those given in this Recommendation, it will be ensured that no code position in any of the specified code tables is assigned more than one meaning within different telematic services. 94 -character subsets of the supplementary code table can be found in Recommendations of specific telematic services, i.e., in Recommendations T. 61 and T.101.
1.3 This Recommendation gives the escape sequences for designating the primary and supplementary sets of graphic characters, to be used according to the code extension techniques specified.
1.4 Non-Latin based character sets are to be dealt with in Recommendation T.52.
1.5 This Recommendation describes those code extension mechanisms that are relevant to existing telematic services. Additional mechanisms will be included in this Recommendation as the need for such is identified for one or more telematic services. The purpose of this Recommendation is to include an up-to-coding systems in various telematic services.
1.6 In this Recommendation 7-bit code tables are described which can be used either in a 7-bit or in an 8-bit environment, with applicable code extension mechanisms that are given in other Recommendations specific to given telematic services.
1.7 This Recommendation gives a unified superset of the repertoire of Latin based alphanumeric characters (see Annex A).
1.8 This Recommendation gives a table of character and control sets used in CCITT telematic services (see Annex B).
1.9 There is no conformance clause in this Recommendation specifying the mandatory and optional subsets of code extension mechanisms and coded character sets. Conformance requirements will be the subject of other CCITT Recommendations specific to particular telematic services.

## 2 Graphic character sets

2.1 Primary set
2.1.1 The primary set of graphic characters specified in Figure 1/T.51 is identical with the set of graphic characters of the International Reference Version (IRV) of the 7-bit coded character set of Recommendation T.50.
2.1.2 The primary set is designated as G0 by the sequence ESC $2 / 84 / 2$. It can also be alternatively designated as G1, G2 or G3 by the sequences ESC $2 / 94 / 2$, ESC $2 / 104 / 2$ or ESC $2 / 114 / 2$ respectively. See $\S 3$ for details on code extension techniques.

Terminals used for telematic services which make reference to the 1988 version of Recommendation T. 51 use, for the designation of the primary set as G0, the sequence ESC $2 / 84 / 0$ and alternatively as G1, G2 and G3, the sequences ESC $2 / 94 / 0$, ESC $2 / 104 / 0$, ESC $2 / 114 / 0$.

### 2.2 Supplementary set

2.2.1 The supplementary set of graphic characters is specified in Figure 2/T.51.
2.2.2 Unallocated code positions are subject to future standardization and will be allocated when a need for such is identified.
2.2.3 The supplementary set is designated as G2 by the sequence ESC $2 / 145 / 2$. It can be alternatively designated as G1 or G3 by the sequences ESC $2 / 135 / 2$ or ESC $2 / 155 / 2$ respectively.

Terminals used for telematic services which make reference to the 1988 version of Recommendation T. 51 use, for the designation of the supplementary set as G2, the sequence ESC $2 / 106 / 2$ and alternatively as G0, G1 and G3, the sequences ESC $2 / 86 / 2$, ESC $2 / 96 / 2$, ESC 2/11 6/2.

### 2.2.4 Notes on the primary and supplementary sets of graphic characters for Figures 1/T.51 and 2/T.51

In the figures the number of the note being referred to is encircled.
Note 1 - All the characters in column 4 of the supplementary set are non-spacing characters. They are all diacritical marks.

Note 2 - Cross-shaded code positions are reserved for future standardization by the CCITT.
Note 3 - Terminals used for current CCITT defined telematic services may send and receive the codes $2 / 6$ and $2 / 4$ of the supplementary set for the NUMBER SIGN and DOLLAR SIGN, respectively. When receiving codes $2 / 3$ and $2 / 4$ from the primary set of graphic characters, terminals may interpret them as \# and $\propto$ respectively. Future applications in telematic services should code the NUMBER SIGN, CURRENCY SIGN and DOLLAR SIGN in accordance with Figures 1/T. 51 and 2/.T.51.

Note 4 - Terminals used for CCITT defined telematic services should send only the codes $4 / 1$ of the supplementary set followed by SPACE for a stand-alone grave accent, $4 / 3$ of the supplementary set followed by SPACE for a stand-alone circumflex accent, and $4 / 4$ of the supplementary set followed by SPACE for a stand-alone tilde. Whenever a telematic terminal is capable of receiving and interpreting codes $6 / 0,5 / 14$ and $7 / 14$ from the primary set of graphic characters, terminals shall interpret them as GRAVE, CIRCUMFLEX and TILDE respectively.

Note 5 - This code position is reserved and shall not be used.

Note 6 - Current telematic services may interpret this as the non-spacing underline. The non-spacing underline character is never used individually but always in combination with some other graphic character to represent the graphic rendition "underlined" for the associated character. The non-spacing underline character can be used in combination with any graphic character of the repertoire, including an accented letter or an umlaut, or space. It is recommended to implement the "underline" function by means of the control function SGR(4) instead of the "nonspacing underline" graphic character. However, both must be correctly interpreted when received.

|  |  |  |  | $\mathrm{b}_{7}$ | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | $\mathrm{b}_{6}$ | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 |
|  |  |  |  | $\mathrm{b}_{5}$ | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| $\mathrm{b}_{4}$ | $\mathrm{b}_{3}$ | $\mathrm{b}_{2}$ | $\mathrm{b}_{1}$ |  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 0 | 0 | 0 | 0 | 0 |  |  |  | 0 | @ | P | 4 | 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 | 3 | C | S | C | S |
| 0 | 1 | 0 | 0 | 4 |  |  | \$3 | 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 | 1 | 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 | 1 | 1 | \| |
| 1 | 1 | 0 | 1 | 13 |  |  | - | $=$ | M | ] | m | $\}$ |
| 1 | 1 | 1 | 0 | 14 |  |  | . | $>$ | N | (4) | n | $\sim$ |
| 1 | 1 | 1 | 1 | 15 |  |  | / | ? | O | - | 0 |  |

Note - Notes to this figure are contained in § 2.2.4.

FIGURE 1/T. 51
The primary set of graphic characters for telematic services (coded representation when invoked in columns 2-7 of the code table)


Note - Notes to this figure are contained in § 2.2.4.

FIGURE 2/T. 51
The supplementary set of graphic characters for telematic services (coded representation when invoked in columns 2-7 of the code table)

## Code extension technique

## 3.1 <br> General

3.1.1 Code extension techniques are required for the designation of various graphic or control character sets and their invocation in the 7-bit set or 8-bit set in use. Such techniques are derived from ISO Standard 2022.
3.1.2 This Recommendation describes only those code extension techniques currently specified for existing telematic services. Additional techniques will be further incorporated as they are identified for use in one or more telematic services.

### 3.2 Definitions

For the purpose of code extension techniques given in this Recommendation, the following definitions apply.

### 3.2.1 bit combination

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

### 3.2.2 byte

A bit string that is operated upon as a unit and the size of which is independent of redundancy or framing techniques.

### 3.2.3 character

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

### 3.2.4 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.

### 3.2.5 code extension

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

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

### 3.2.7 control character

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

### 3.2.8 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.

### 3.2.9 to designate

To identify a set of characters that are to be represented, in some cases immediately and in others on the occurrence of a further control function, in a prescribed manner.

### 3.2.10 environment

The characteristic that identifies the number of bits used to represent a character in a data processing or data communication system or in part of such a system.

### 3.2.11 escape sequence

A bit string that is used for control purposes in code extension procedures and that consists of two or more bit combinations. The first of these bit combinations represents the character ESCAPE (1/11).

### 3.2.12 final character

The character the bit combination of which terminates an escape sequence.

### 3.2.13 graphic character

A character, other than a control function, that has a visual representation normally handwritten, printed or displayed.

### 3.2.14 intermediate character

A character the bit combination of which occurs between that of the ESCAPE character and that of the Final character in an escape sequence consisting of more than two bit combinations.

### 3.2.15 to invoke

To cause a designated set of characters to be represented by the prescribed bit combinations whenever those bit combinations occur, until an appropriate code extension function occurs.
3.2.16 position

That part of a code table identified by its column and row coordinates.
3.2.17 to represent
a) to use a prescribed bit combination with the meaning of a character in a set of characters that has been designated and invoked; or
b) to use an escape sequence with the meaning of an additional control function.

### 3.3 Code extension facilities

These are depicted in Figure 3/T. 51 for the 7 -bit environment and Figure 4/T. 51 for the 8 -bit environment. They include the following functions:
a) designation and invocation of control sets C 0 and C 1 by means of the relevant escape sequences given in § 3.4;
b) designation of a graphic character set G0 by means of the relevant escape sequence given in § 3.4;
c) designation of up to three additional G-sets called G1, G2 and G3 by means of the relevant escape sequences given in § 3.4;
d) invocation of the designated graphic sets, by means of locking and/or non-locking shift functions, given in § 3.5;
e) designation and invocation of a complete code by means of the relevant escape sequence given in §3.4.

## 3.4 <br> Types of character sets

There are a number of different types of control and graphic character sets that can be designated and invoked for use in the 7 -bit or 8 -bit environment. These are listed in Table 3/T.51 and defined below. A given control or graphic character set can be designated by an escape sequence terminated by a Final character F from bit combination $4 / 0$ to $7 / 14$, specific to the character set to be designated. Final characters are allocated by ISO and registered in the ISO "International register of coded character sets to be used with escape sequences".


FIGURE 3/T. 51
Code extension in a 7-bit environment
(showing all shift functions)
3.4.1 C 0 set of 32 control characters (bit combinations $0 / 0$ to $1 / 15$ ). A C 0 set is designated and invoked by the sequence ESC $2 / 1 \mathrm{~F}$, where F identifies a registered C 0 set.


FIGURE 4/T. 51
Code extension in an 8-bit environment
(showing all shift facilities)
3.4.2 C1 set of 32 control characters (bit combinations $8 / 0$ to $9 / 15$ in an 8 -bit environment, or ESC $4 / 0$ to ESC $5 / 15$ in a 7-bit environment). A C1 set is designated and invoked by the sequence ESC $2 / 2 \mathrm{~F}$, where F identifies a registered C1 set.
3.4.3 G0 set of 94 graphic characters (bit combinations $2 / 1$ to $7 / 14$ ).
3.4.4 G1, G2 and G3 sets of 94 graphic characters (bit combinations $2 / 1$ to $7 / 14$ or $10 / 1$ to $15 / 14$ ). A registered set of 94 graphic characters can be designated by the escape sequence ESC $2 / 8 \mathrm{~F}$, ESC $2 / 9 \mathrm{~F}$, ESC $2 / 10 \mathrm{~F}$ or ESC $2 / 11 \mathrm{~F}$ in order to be used as a G0, G1, G2 or G3 set respectively, where F identifies the designated set.
3.4.5 G1, G2 and G3 sets of 96 graphic characters (bit combinations $2 / 0$ to $7 / 15$ or $10 / 0$ to $15 / 15$ ). A registered set of 96 graphic characters can be designated by the escape sequence ESC $2 / 13 \mathrm{~F}$, ESC $2 / 14 \mathrm{~F}$ or ESC $2 / 15 \mathrm{~F}$ in order to be used as a G1, G2 or G3 set respectively, where F identifies the designated set.
3.4.6 Multiple-byte G0 set of more than 94 graphic characters each represented by more than one bit combinations from $2 / 1$ to $7 / 14$. The designation sequence is ESC $2 / 42 / 8 \mathrm{~F}$.
3.4.7 Multiple-byte G1, G2 and G3 sets of more than 94 graphic characters each represented by more than one bit combination from $2 / 1$ to $7 / 14$ or from $10 / 1$ to $15 / 14$. A registered multiple-byte set of more than 94 graphic characters can be designated by the escape sequence ESC $2 / 4 \mathrm{~F}$, ESC $2 / 42 / 9 \mathrm{~F}$, ESC $2 / 42 / 10 \mathrm{~F}$ or ESC $2 / 42 / 11 \mathrm{~F}$ in order to be used as a G0, G1, G2 or G3 set respectively, where F identifies the designated set.
3.4.8 Multiple-byte G1, G2 and G3 sets of more than 96 graphic characters each represented by more than one bit combination from $2 / 0$ to $7 / 15$ or from $10 / 0$ to $15 / 15$. A registered multiple-byte set of this type identified by a Final character F can be designated by the escape sequence ESC $2 / 42 / 13 \mathrm{~F}$, ESC $2 / 42 / 14 \mathrm{~F}$ or ESC $2 / 42 / 15 \mathrm{~F}$ in order to be used as G1, G2 or G3 set respectively.
3.4.9 Complete code containing all bit combinations $0 / 0$ to $7 / 15$ in 7 -bit environment, or $0 / 0$ to $15 / 15$ in 8 -bit environment. A complete code identified by a Final character F, can be designated and invoked by the escape sequence ESC $2 / 5 \mathrm{~F}$.

Note - Complete code is now referred to in ISO 2022 as "Coding System different from that of ISO 2022".

### 3.5 Invocation functions

Following the designation of a G0, G1, G2 or G3 set as specified in § 3.4 , any one of these sets will require invocation into the 7-bit or 8-bit in-use code table. This is performed by the use of either locking-shift functions or nonlocking shift functions listed in Table 4/T.51. Coding for these functions is given in Table 5/T.51.

### 3.5.1 Use of locking-shift functions

There are seven locking-shift functions as given in Tables 4/T.51 and 5/T.51. A locking-shift function invokes the relevant G0, G1, G2 or G3 set into columns 2 to 7 or into columns 10 to 15 ( 8 -bit code only) in order to replace the previously invoked G-set. The occurrence of a locking-shift function shall not affect those bit combinations that are included in any escape sequence or those that follow single-shift function SS2 or SS3. If a particular set has already been invoked, use of the corresponding locking-shift function has no effect.

### 3.5.2 Use of single-shift functions

The single-shift function SS2 shall invoke one character from the last designated G2 set. Similarly SS3 shall invoke one character from the last designated G3 set. In the case of using single-byte graphic sets, these invocations alter the meaning of the immediately following bit combination only and ascribe to it the meaning of the corresponding bit combination of the G2 or G3 set. The bit combination permitted to follow SS2 or SS3 is limited to one from columns 2 to 7. If a single-shift function is used to invoke a character from a multiple-byte set, the shift function will affect two or more bit-combinations to represent that character. The use of a single-shift function does not affect the current shift status established by a locking-shift function.

Types of character sets

| Description | Character set | Designation sequence | Bit combinations |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  | 7-bit environment | 8-bit environment |
| Sets of 32 control characters (Note 1) | $\begin{aligned} & \mathrm{C} 0 \\ & \mathrm{C} 1 \end{aligned}$ | $\begin{aligned} & \text { ESC } 2 / 1 \mathrm{~F} \\ & \text { ESC } 2 / 2 \mathrm{~F} \end{aligned}$ | 0/0 to $1 / 15$ ESC $4 / 1$ to ESC $5 / 15$ | 00/00 to 01/15 08/00 to 09/15 |
| Sets of 94 graphic characters | G0 | ESC $2 / 8 \mathrm{~F}$ | 2/1 to 7/14 | 02/01 to 07/14 |
|  | $\begin{aligned} & \text { G1 } \\ & \text { G2 } \\ & \text { G3 } \end{aligned}$ | ESC $2 / 9 \mathrm{~F}$ <br> ESC $2 / 10 \mathrm{~F}$ <br> ESC $2 / 11$ | 2/1 to 7/14 | $\begin{gathered} 02 / 01 \text { to } 07 / 14 \\ \text { or } \\ 10 / 01 \text { to } 15 / 14 \end{gathered}$ |
| Sets of 96 graphic characters | $\begin{aligned} & \text { G1 } \\ & \text { G2 } \\ & \text { G3 } \end{aligned}$ | $\begin{aligned} & \text { ESC } 2 / 13 \mathrm{~F} \\ & \text { ESC } 2 / 14 \mathrm{~F} \\ & \text { ESC } 2 / 15 \mathrm{~F} \end{aligned}$ | 2/0 to 7/15 | $\begin{gathered} 02 / 00 \text { to } 07 / 15 \\ \text { or } \\ 10 / 00 \text { to } 15 / 15 \end{gathered}$ |
| Sets of more than 94 graphic characters each represented by more than one byte | G0 | ESC $2 / 4 \mathrm{~F}$ (Note 2) <br> ESC $2 / 42 / 8 \mathrm{~F}$ | 2/1 to 7/14 | 02/01 to 07/14 |
|  | $\begin{aligned} & \text { G1 } \\ & \text { G2 } \\ & \text { G3 } \end{aligned}$ | ESC $2 / 42 / 9 \mathrm{~F}$ <br> ESC 2/4 2/10 F <br> ESC $2 / 42 / 11 \mathrm{~F}$ | 2/1 to 7/14 | $\begin{aligned} & 02 / 01 \text { to } 07 / 14 \\ & \text { or } \\ & 10 / 01 \text { to } 15 / 14 \end{aligned}$ |
| Sets of more than 96 graphic characters each represented by more than one byte | $\begin{aligned} & \text { G1 } \\ & \text { G2 } \\ & \text { G3 } \end{aligned}$ | $\begin{aligned} & \text { ESC } 2 / 42 / 13 \mathrm{~F} \\ & \text { ESC } 2 / 42 / 14 \mathrm{~F} \\ & \text { ESC } 2 / 42 / 15 \mathrm{~F} \end{aligned}$ | 2/0 to 7/15 | $\begin{gathered} 02 / 00 \text { to } 07 / 15 \\ \text { or } \\ 10 / 00 \text { to } 15 / 15 \end{gathered}$ |
| Complete code |  | ESC $2 / 5 \mathrm{~F}$ | 0/0 to 7/15 | 00/00 to 15/15 |

Note 1 - Control sets C 0 and C 1 , and complete codes are simultaneously designated and invoked by the relevant escape sequences indicated.

Note 2 - Multi-byte character sets are to be designated by four-character escape sequences. However, the following threecharacter escape sequences have in the past been registered and their use is valid for the registered multi-byte character sets:

ESC 2/4 4/0;
ESC 2/4 4/1;
ESC 2/4 4/2.

## Allocation of shift functions to the graphic character sets to be invoked

|  | Locking-shift functions |  | Non-locking shift functions |
| :---: | :---: | :---: | :---: |
| A set | Columns 2 to 7 <br> of 7-bit or 8-bit code | Columns 10 to 15 <br> of 8-bit code | Columns 2 to 7 <br> of 7-bit or 8bit code |
| G0 | SI(7-bit), LS0(8-bit) | - | - |
| G1 | SO(7-bit), LS1(8-bit) | LS1R | - |
| G2 | LS2 | LS2R | SS2 |
| G3 | LS3 | LS3R | SS3 |

TABLE 5/T. 51

Coding for shift functions

| Shift functions | Coding |  |
| :--- | ---: | :---: |
| Single-shift two | SS2 | $1 / 9$ |
| Single-shift three | SS3 | $1 / 13$ |
| Shift in SI(7-bit), locking-shift zero | LS0(8-bit) | $0 / 15$ |
| Shift out SO(7-bit), locking-shift one | LS1(8-bit) | $0 / 14$ |
| Locking-shift one right | LS1R | ESC 7/14 |
| Locking-shift two | LS2 | ESC 6/14 |
| Locking-shift two right | LS2R | ESC 7/13 |
| Locking-shift three | LS3 | ESC 6/15 |
| Locking-shift three right | LS3R | ESC 7/12 |

## ANNEX A

## (to Recommendation T.51)

## Superset of the repertoire of the Latin based character set

A. 1 This annex contains a unified superset of the repertoire of Latin based alphanumeric graphic characters.

Each graphic character is identified by the identification system identical in that used in Recommendation T. 61 (see § A.2).

## A. 2 Identification system

A system was developed that allows for the identification and description of each graphic character or control function. The system is shown in Figure A-1/T.51.

Each identifier consists of two letters and two digits.
The first letter indicates the alphabet, the language, etc.
The second letter indicates the letter of an alphabet or, in the case of a non-alphabetic graphic character or a control function, the group of characters or control functions.

The first digit indicates whether the letter in the second position is an accented one, whether the diacritical mark is above or below the letter, etc. It has not special meaning in the case of the first letter being a $\mathrm{C}, \mathrm{N}$ or S .

The second digit indicates whether the letter is a capital or a small (even or odd). If the first letter is a $\mathrm{C}, \mathrm{N}$ or S , this digit being even or odd has no significance.

The numbering is used in a consistent manner so that each diacritical mark is always given the same number.
The numbering principle is shown in Table A-1/T.51.

TABLE A-1/T. 51

## Numbering principle for alphabetic characters

| Item | Small | Capital |
| :--- | :---: | :---: |
| No diacritical mark |  |  |
| Acuse accent | 01 | 02 |
| Grave accent | 11 | 12 |
| Circumflex accent | 13 | 14 |
| Diaeresis or umlaut mark | 15 | 16 |
| Tilde | 17 | 18 |
| Caron | 19 | 20 |
| Breve | 21 | 22 |
| Double acute accent | 23 | 24 |
| Ring | 25 | 26 |
| Dot | 27 | 28 |
| Macron | 29 | 30 |
| Cedilla | 31 | 32 |
| Ogonek | 41 | 42 |
| Diphthong or ligature | 43 | 44 |
| Special form | 51 | 52 |



FIGURE A-1/T. 51

## Identification system

A. 3 Combination of diacritical marks and basic letters

Figure A-2/T. 51 specifies the combinations of diacritical marks and basic letters that are defined in this annex. This figure also shows ligatures and other characters.

| Basic letters | Acute accent | Grave Accent | $\begin{aligned} & \text { Circumflex } \\ & \text { accent } \end{aligned}$ | Diaeresis or umlaut accent | Tilde | Caron or hacek | Breve | Double acute accent | Ring | $\begin{gathered} \text { Dot } \\ \text { abover } \end{gathered}$ | Macron | Cedilla | Ogonek | Ligature | Others |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| a A | á Á | à À | â Â | ä Ä | ã Ã |  | ă Ă |  | å Å |  | $\overline{\mathrm{a}} \overline{\mathrm{A}}$ |  | ą A | æÆ |  |
| b B |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| c C | ćc |  | ç C |  |  | č Č |  |  |  | ċ Ċ |  | ç Ç |  |  |  |
| d D |  |  |  |  |  | d D |  |  |  |  |  |  |  |  | す Đ |
| e E | é É | è̀̇ | ê Ê | ë Ë |  | ěĚ |  |  |  | è Ė | $\overline{\mathrm{e}} \overline{\mathrm{E}}$ |  | e E¢ |  |  |
| f F |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| g G | g |  | ĝ Ĝ |  |  |  | ğ Ğ |  |  | g $\dot{\text { G }}$ |  | Ģ |  |  |  |
| h H |  |  | ĥt |  |  |  |  |  |  |  |  |  |  |  | 万 H |
| il | íl | ì | î | ii | ii |  |  |  |  | i | $\bar{\top}$ |  | i！ | ij リ | 1 |
| j J |  |  | j ${ }^{\text {j }}$ |  |  |  |  |  |  |  |  |  |  |  |  |
| k K |  |  |  |  |  |  |  |  |  |  |  | k K |  |  | к |
| IL | íL |  |  |  |  | ǐĽ |  |  |  |  |  | $!$ ！ |  |  | $\begin{gathered} \hline X \chi \\ \text { I.E } \end{gathered}$ |
| m M |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| n N | ń N |  |  |  | ñ ${ }^{\text {N }}$ | ň Ň |  |  |  |  |  | n N |  |  | n |
| $\bigcirc 0$ | óó | ò Ò | ô Ô | ö | õo |  |  | \％＂ 0 |  |  | $\overline{0} \bar{O}$ |  |  | œ 〇E | $\varnothing \varnothing$ |
| p P |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| q Q |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| r R | r R |  |  |  |  | ř Ř |  |  |  |  |  | ¢ R |  |  |  |
| s S | śs |  | st Ŝ |  |  | šs |  |  |  |  |  | s Ş |  |  | $\beta$ |
| $t \mathrm{~T}$ |  |  |  |  |  | ťT |  |  |  |  |  | ！T |  |  | t $\ddagger \mathrm{p}$ |
| u U | ú Ú | ù Ù | û Û | ü Ü | ũ Ũ |  | ŭ Ŭ | ü Ü | ¢̊ บ̇ |  | $\bar{u} u$ |  | ب |  |  |
| v V |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| w W |  |  | $\hat{w}$ W |  |  |  |  |  |  |  |  |  |  |  |  |
| $\mathbf{x} \mathbf{X}$ |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| y Y | ý Ý |  | $\hat{y} \hat{Y}$ | ÿ Y |  |  |  |  |  |  |  |  |  |  |  |
| z Z | żŻ |  |  |  |  | ž Ž |  |  |  | ż̇̇ |  |  |  |  |  |

FIGURE A－2／T． 51
Use of alphabetical characters with dialectical marks，ligatures or others

## A. 4 Specification of superset of Latin based character repertoire

## Explanations on coded representation

- P-prefix denotes Primary Code Table;
- S-prefix denotes Supplementary Code Table;
- Absence of prefix denotes T. 50 Basic Code Table (For SPACE and DELETE characters).

Notes on implementation of the repertoire in telematic services
(1) Not used in current teletex service (Recommendation T.61);
(2) Not used in any current telematic services;
(3) Currently used in Recommendation T.101, Videotex Data Syntax III only.

## General remarks

a) The Latin based repertoire of teletex (Recommendation T.61), is represented by this superset, with exclusions defined by Notes (1), (2) and (3) above.
b) The Latin based repertoire of Recommendation T.101, Videotex Data Syntax I, is represented by the characters of the Primary Set with the exception that the REVERSE SOLIDUS is replaced by the YEN SIGN.
c) The Latin based repertoire of Recommendation T.101, Videotex Data Syntax II, is represented by this superset with exclusions defined by Notes (2) and (3) above, plus a few characters not within the scope of coding defined by the Primary Set and the Supplementary Set of this Recommendation.
d) The Latin based repertoire of Recommendation T.101, Videotex, Data Syntax III, is represented entirely by this superset.

## A.4.1 Latin alphabet characters

|  |  |  | Coded <br> Identifier |
| :--- | :--- | :--- | :--- |
|  |  | representation |  |


| Identifier | Name | Coded representation |  |
| :---: | :---: | :---: | :---: |
| LC11 | Latin small letter c with acute | S 4/2 | P 6/3 |
| LC12 | Latin capital letter C with acute | S 4/2 | P 4/3 |
| LC15 | Latin small letter c with circumflex | S $4 / 3$ | P 6/3 |
| LC16 | Latin capital letter C with circumflex | S 4/3 | P 4/3 |
| LC21 | Latin small letter c with caron | S 4/15 | P 6/3 |
| LC22 | Latin capital letter C with caron | S 4/15 | P 4/3 |
| LC29 | Latin small letter c with dot above | S 4/7 | P 6/3 |
| LC30 | Latin capital letter C with dot above | S 4/7 | P 4/3 |
| LC41 | Latin small letter c with cedilla | S 4/11 | P 6/3 |
| LC42 | Latin capital letter C with cedilla | S 4/11 | P 4/3 |
| LD01 | Latin small letter d |  | P 6/4 |
| LD02 | Latin capital letter D |  | P 4/4 |
| LD21 | Latin small letter d with caron | S 4/15 | P 6/4 |
| LD22 | Latin capital letter D with caron | S 4/15 | P 4/4 |
| LD61 | Latin small letter d with stroke | S 7/2 |  |
| LD62 | Latin capital letter D with stroke | S 6/2 |  |
| LD63 | Latin small letter eth (icelandic) | S 7/3 |  |
| LE01 | Latin small lettere |  | P 6/5 |
| LE02 | Latin capital letter E |  | P 4/5 |
| LE11 | Latin small letter e with acute | S 4/2 | P $6 / 5$ |
| LE12 | Latin capital letter E with acute | S 4/2 | P 4/5 |
| LE13 | Latin small letter e with grave | S 4/1 | P 6/5 |
| LE14 | Latin capital letter E with grave | S 4/1 | P 4/5 |
| LE15 | Latin small letter e with circumflex | S 4/3 | P $6 / 5$ |
| LE16 | Latin capital letter E with circumflex | S 4/3 | P 4/5 |
| LE17 | Latin small letter e with diaeresis | S 4/8 | P $6 / 5$ |
| LE18 | Latin capital letter E with diaeresis | S 4/8 | P 4/5 |
| LE21 | Latin small letter e with caron | S 4/15 | P 6/5 |
| LE22 | Latin capital letter E with caron | S 4/15 | P 4/5 |
| LE29 | Latin small letter e with dot above | S 4/7 | P $6 / 5$ |
| LE30 | Latin capital letter E with dot above | S 4/7 | P 4/5 |
| LE31 | Latin small letter e with macron | S 4/5 | P 6/5 |
| LE32 | Latin capital letter E with macron | S 4/5 | P 4/5 |
| LE43 | Latin small letter e with ogonek | S 4/14 | P 6/5 |
| LE44 | Latin capital letter E with ogonek | S 4/14 | P 4/5 |
| LF01 | Latin small letter f |  | P 6/6 |
| LF02 | Latin capital letter F |  | P 4/6 |
| LG01 | Latin small letter g |  | P 6/7 |
| LG02 | Latin capital letter G |  | P 4/7 |
| LG11 | Latin small letter g with cedilla | S 4/2 | P 6/7 |
| LG15 | Latin small letter g with circumflex | S 4/3 | P 6/7 |
| LG16 | Latin capital letter G with circumflex | S 4/3 | P 4/7 |
| LG23 | Latin small letter g with breve | S 4/6 | P 6/7 |
| LG24 | Latin capital letter G with breve | S 4/6 | P 4/7 |
| LG29 | Latin small letter g with dot above | S 4/7 | P 6/7 |
| LG30 | Latin capital letter G with dot above | S 4/7 | P 4/7 |
| LG42 | Latin capital letter G with cedilla | S 4/11 | P 4/7 |
| LH01 | Latin small letter h |  | P 6/8 |
| LH02 | Latin capital letter H |  | P 4/8 |
| LH15 | Latin small letter h with circumflex | S 4/3 | P 6/8 |
| LH16 | Latin capital letter H with circumflex | S $4 / 3$ | P 4/8 |
| LH61 | Latin small letter h with stroke | S 7/4 |  |
| LH62 | Latin capital letter H with stroke | S 6/4 |  |
| LI01 | Latin small letter i |  | P 6/9 |
| LI02 | Latin capital letter I |  | P 4/9 |
| LI11 | Latin small letter i with acute | S 4/2 | P 6/9 |
| LI12 | Latin capital letter I with acute | S 4/2 | P 4/9 |
| LI13 | Latin small letter i with grave | S 4/1 | P 6/9 |
| LI14 | Latin capital letter I with grave | S 4/1 | P 4/9 |


| Identifier | Name | Coded representation |  |
| :---: | :---: | :---: | :---: |
| LI15 | Latin small letter i with circumflex | S 4/3 | P 6/9 |
| LI16 | Latin capital letter I with circumflex | S 4/3 | P 4/9 |
| LI17 | Latin small letter i with diaeresis | S 4/8 | P $6 / 9$ |
| LI18 | Latin capital letter I with diaeresis | S 4/8 | P 4/9 |
| LI19 | Latin small letter i with tilde | S 4/4 | P $6 / 9$ |
| LI20 | Latin capital letter I with tilde | S 4/4 | P 4/9 |
| LI30 | Latin capital letter I with dot above | S 4/7 | P 4/9 |
| LI31 | Latin small letter i with macron | S 4/5 | P 6/9 |
| LI32 | Latin capital letter I with macron | S 4/5 | P 4/9 |
| LI43 | Latin small letter i with ogonek | S 4/14 | P 6/9 |
| LI44 | Latin capital letter I with ogonek | S 4/14 | P 4/9 |
| LI51 | Latin small ligature ij | S 7/6 |  |
| LI52 | Latin capital ligature IJ | S 6/6 |  |
| LI61 | Latin small letter dotless i | S $7 / 5$ |  |
| LJ01 | Latin small letter j |  | P 6/10 |
| LJ02 | Latin capital letter J |  | P 4/10 |
| LJ15 | Latin small letter j with circumflex | S 4/3 | P 6/10 |
| LJ16 | Latin capital letter J with circumflex | S 4/3 | P 4/10 |
| LK01 | Latin small letter k |  | P 6/11 |
| LK02 | Latin capital letter K |  | P 4/11 |
| LK41 | Latin small letter k with cedilla | S 4/11 | P 6/11 |
| LK42 | Latin capital letter K with cedilla | S 4/11 | P 4/11 |
| LK61 | Latin small letter kra (greenlandic) | S 7/0 |  |
| LL01 | Latin small letter 1 |  | P 6/12 |
| LL02 | Latin capital letter L |  | P 4/12 |
| LL11 | Latin small letter 1 with acute | S 4/2 | P 6/12 |
| LL12 | Latin capital letter L with acute | S 4/2 | P 4/12 |
| LL21 | Latin small letter 1 with caron | S 4/15 | P 6/12 |
| LL22 | Latin capital letter L with caron | S 4/15 | P 4/12 |
| LL41 | Latin small letter 1 with cedilla | S 4/11 | P 6/12 |
| LL42 | Latin capital letter L with cedilla | S 4/11 | P 4/12 |
| LL61 | Latin small letter 1 with stroke | S 7/8 |  |
| LL62 | Latin capital letter L with stroke | S 6/8 |  |
| LL63 | Latin small letter 1 with middle dot | S $7 / 7$ |  |
| LL64 | Latin capital letter L with middle dot | S 6/7 |  |
| LM01 | Latin small letter m |  | P 6/13 |
| LM02 | Latin capital letter M |  | P 4/13 |
| LN01 | Latin small letter n |  | P 6/14 |
| LN02 | Latin capital letter N |  | P 4/14 |
| LN11 | Latin small letter n with acute | S 4/2 | P 6/14 |
| LN12 | Latin capital letter N with acute | S 4/2 | P 4/14 |
| LN19 | Latin small letter n with tilde | S 4/4 | P 6/14 |
| LN20 | Latin capital letter N with tilde | S 4/4 | P 4/14 |
| LN21 | Latin small letter n with caron | S 4/15 | P 6/14 |
| LN22 | Latin capital letter N with caron | S 4/15 | P 4/14 |
| LN41 | Latin small letter n with cedilla | S 4/11 | P 6/14 |
| LN42 | Latin capital letter N with cedilla | S 4/11 | P 4/14 |
| LN61 | Latin small letter eng (lappish) | S 7/14 |  |
| LN62 | Latin capital letter ENG (lappish) | S 6/14 |  |
| LN63 | Latin small letter n preceded by apostrophe | S 6/15 |  |
| LO01 | Latin small letter o |  | P 6/15 |
| LO02 | Latin capital letter O |  | P 4/15 |
| LO11 | Latin small letter o with acute | S 4/2 | P 6/15 |
| LO12 | Latin capital letter O with acute | S 4/2 | P 4/15 |
| LO13 | Latin small letter o with grave | S 4/1 | P 6/15 |
| LO14 | Latin capital letter O with grave | S 4/1 | P 4/15 |
| LO15 | Latin small letter o with circumflex | S 4/3 | P 6/15 |
| LO16 | Latin capital letter O with circumflex | S 4/3 | P 4/15 |


| Identifier | Name | Coded representation |  |
| :---: | :---: | :---: | :---: |
| LO17 | Latin small letter o with diaeresis | S 4/8 | P 6/15 |
| LO18 | Latin capital letter O with diaeresis | S 4/8 | P 4/15 |
| LO19 | Latin small letter o with tilde | S 4/4 | P 6/15 |
| LO20 | Latin capital letter O with tilde | S 4/4 | P 4/15 |
| LO25 | Latin small letter o with double acute | S 4/13 | P 6/15 |
| LO26 | Latin capital letter O with double acute | S 4/13 | P 4/15 |
| LO31 | Latin small letter o with macron | S 4/5 | P 6/15 |
| LO32 | Latin capital letter O with macron | S 4/5 | P 4/15 |
| LO51 | Latin small ligature ¢ | S 7/10 |  |
| LO52 | Latin capital ligature OE | S 6/10 |  |
| LO61 | Latin small letter o with stroke | S $7 / 9$ |  |
| LO62 | Latin capital letter O with stroke | S 6/9 |  |
| LP01 | Latin small letter p |  | P 7/0 |
| LP02 | Latin capital letter P |  | P 5/0 |
| LQ01 | Latin small letter q |  | P $7 / 1$ |
| LQ02 | Latin capital letter Q |  | P 5/1 |
| LR01 | Latin small letter r |  | P 7/2 |
| LR02 | Latin capital letter R |  | P 5/2 |
| LR11 | Latin small letter r with acute | S 4/2 | P 7/2 |
| LR12 | Latin capital letter R with acute | S 4/2 | P 5/2 |
| LR21 | Latin small letter r with caron | S 4/15 | P 7/2 |
| LR22 | Latin capital letter R with caron | S 4/15 | P 5/2 |
| LR41 | Latin small letter r with cedilla | S 4/11 | P 7/2 |
| LR42 | Latin capital letter R with cedilla | S 4/11 | P 5/2 |
| LS01 | Latin small letters |  | P 7/3 |
| LS02 | Latin capital letter S |  | P 5/3 |
| LS11 | Latin small letter s with acute | S 4/2 | P 7/3 |
| LS12 | Latin capital letter S with acute | S 4/2 | P 5/3 |
| LS15 | Latin small letter s with circumflex | S 4/3 | P 7/3 |
| LS16 | Latin capital letter S with circumflex | S $4 / 3$ | P 5/3 |
| LS21 | Latin small letter s with caron | S 4/15 | P 7/3 |
| LS22 | Latin capital letter S with caron | S 4/15 | P 5/3 |
| LS41 | Latin small letter s with cedilla | S 4/11 | P 7/3 |
| LS42 | Latin capital letter S with cedilla | S 4/11 | P 5/3 |
| LS61 | Latin small letter sharp s (german) | S 7/11 |  |
| LT01 | Latin small letter t |  | P 7/4 |
| LT02 | Latin capital letter T |  | P 5/4 |
| LT21 | Latin small letter t with caron | S 4/15 | P 7/4 |
| LT22 | Latin capital letter T with caron | S 4/15 | P 5/4 |
| LT41 | Latin small letter t with cedilla | S 4/11 | P 7/4 |
| LT42 | Latin capital letter T with cedilla | S 4/11 | P 5/4 |
| LT61 | Latin small letter t with stroke | S 7/13 |  |
| LT62 | Latin capital letter T with stroke | S 6/13 |  |
| LT63 | Latin small letter thorn (icelandic) | S 7/12 |  |
| LT64 | Latin capital letter thorn (icelandic) | S 6/12 |  |
| LU01 | Latin small letter u |  | P $7 / 5$ |
| LU02 | Latin capital letter U |  | P 5/5 |
| LU11 | Latin small letter $u$ with acute | S 4/2 | P $7 / 5$ |
| LU12 | Latin capital letter U with acute | S 4/2 | P 5/5 |
| LU13 | Latin small letter u with grave | S 4/1 | P $7 / 5$ |
| LU14 | Latin capital letter U with grave | S 4/1 | P 5/5 |
| LU15 | Latin small letter u with circumflex | S 4/3 | P $7 / 5$ |
| LU16 | Latin capital letter U with circumflex | S 4/3 | P 5/5 |
| LU17 | Latin small letter $u$ with diaeresis | S 4/8 | P $7 / 5$ |
| LU18 | Latin capital letter U with diaeresis | S 4/8 | P 5/5 |
| LU19 | Latin small letter u with tilde | S 4/4 | P $7 / 5$ |
| LU20 | Latin capital letter U with tilde | S 4/4 | P 5/5 |
| LU23 | Latin small letter u with breve | S 4/6 | P 7/5 |
| LU24 | Latin capital letter U with breve | S 4/6 | P 5/5 |


| Identifier | Name | Coded representation |  |
| :---: | :---: | :---: | :---: |
| LU25 | Latin small letter u with double acute | S 4/13 | P 7/5 |
| LU26 | Latin capital letter U with double acute | S 4/13 | P 5/5 |
| LU27 | Latin small letter $u$ with ring above | S 4/10 | P $7 / 5$ |
| LU28 | Latin capital letter U with ring above | S 4/10 | P 5/5 |
| LU31 | Latin small letter u with macron | S 4/5 | P $7 / 5$ |
| LU32 | Latin capital letter U with macron | S 4/5 | P $5 / 5$ |
| LU43 | Latin small letter u with ogonek | S 4/14 | P $7 / 5$ |
| LU44 | Latin capital letter U with ogonek | S 4/14 | P 5/5 |
| LV01 | Latin small letter v |  | P 7/6 |
| LV02 | Latin capital letter V |  | P 5/6 |
| LW01 | Latin small letter w |  | P $7 / 7$ |
| LW02 | Latin capital letter W |  | P 5/7 |
| LW15 | Latin small letter w with circumflex | S 4/3 | P $7 / 7$ |
| LW16 | Latin capital letter W with circumflex | S 4/3 | P 5/7 |
| LX01 | Latin small letter x |  | P $7 / 8$ |
| LX02 | Latin capital letter X |  | P 5/8 |
| LY01 | Latin small letter y |  | P 7/9 |
| LY02 | Latin capital letter Y |  | P 5/9 |
| LY11 | Latin small letter y with acute | S 4/2 | P 7/9 |
| LY12 | Latin capital letter Y with acute | S 4/2 | P 5/9 |
| LY15 | Latin small letter y with circumflex | S 4/3 | P 7/9 |
| LY16 | Latin capital letter Y with circumflex | S 4/3 | P 5/9 |
| LY17 | Latin small letter y with diaeresis | S 4/8 | P 7/9 |
| LY18 | Latin capital letter Y with diaeresis | S 4/8 | P 5/9 |
| LZ01 | Latin small letter z |  | P 7/10 |
| LZ02 | Latin capital letter Z |  | P 5/10 |
| LZ11 | Latin small letter z with acute | S 4/2 | P 7/10 |
| LZ12 | Latin capital letter Z with acute | S 4/2 | P 5/10 |
| LZ21 | Latin small letter z with caron | S 4/15 | P 7/10 |
| LZ22 | Latin capital letter Z with caron | S 4/15 | P 5/10 |
| LZ29 | Latin small letter z with dot above | S 4/7 | P 7/10 |
| LZ30 | Latin capital letter Z with dot above | S 4/7 | P 5/10 |

## A.4.2 Non-alphabetic characters

## A.4.2.1 Decimal digits

| Identifier | Name | Coded representation |
| :---: | :--- | :---: |
| ND01 | Digit one | $\mathrm{P} 3 / 1$ |
| ND02 | Digit two | $\mathrm{P} 3 / 2$ |
| ND03 | Digit three | $\mathrm{P} 3 / 3$ |
| ND04 | Digit four | $\mathrm{P} 3 / 4$ |
| ND05 | Digit five | $\mathrm{P} 3 / 5$ |
| ND06 | Digit six | $\mathrm{P} 3 / 6$ |
| ND07 | Digit seven | $\mathrm{P} 3 / 7$ |
| ND08 | Digit eight | $\mathrm{P} 3 / 9$ |
| ND09 | Digit nine | $\mathrm{P} 3 / 0$ |
| ND10 | Digit zero |  |

## A.4.2.2 Currency signs

| Identifier | Name | Coded representation |
| :---: | :--- | :--- |
| NC01 | Currency sign | $\mathrm{S} 2 / 8$ |
| NC02 | Pound sign | $\mathrm{S} 2 / 3$ |
| NC03 | Dollar sign | $\mathrm{P} 2 / 4$ |
| NC04 | Cent sign | $\mathrm{S} 2 / 2$ |
| NC05 | Yen sing | $\mathrm{S} 2 / 5$ |

## A.4.2.3 Punctuation marks

| Identifier | Name | Coded representation | Notes |
| :---: | :---: | :---: | :---: |
| SP01 | Espace | 2/0 |  |
| SP02 | Exclamation mark | P $2 / 1$ |  |
| SP03 | Inverted exclamation mark | P $2 / 1$ |  |
| SP04 | Quotation mark | P $2 / 2$ |  |
| SP05 | Apostrophe | P $2 / 7$ |  |
| SP06 | Left parenthesis | P $2 / 8$ |  |
| SP07 | Right parenthesis | P $2 / 9$ |  |
| SP08 | Comma | P 2/12 |  |
| SP09 | Low line | P 5/13 |  |
| SP10 | Hyphen-minus | P 2/13 |  |
| SP11 | Full stop | P 2/14 |  |
| SP12 | Solidus | P 2/15 |  |
| SP13 | Colon | P 3/10 |  |
| SP14 | Semicolon | P 3/11 |  |
| SP15 | Question mark | P 3/15 |  |
| SP16 | Inverted question mark | S 3/15 |  |
| SP17 | Left-pointing double angle quotation mark | S 2/11 |  |
| SP18 | Right-pointing double angle quotation mark | S 3/11 |  |
| SP19 | Left single quotation mark | S $2 / 9$ | (1) |
| SP20 | Right single quotation mark | S 3/9 | (1) |
| SP21 | Left double quotation mark | S 2/10 | (1) |
| SP22 | Right double quotation mark | S 3/10 | (1) |
| SP31 | No-break space | S $2 / 0$ | (2) |
| SP32 | Soft hyphen | S 7/15 | (2) |

Note - In teletex and videotex, quotation mark, apostrophe and comma are independent characters that cannot have the meaning of diacrical marks.

## A.4.2.4 Arithmetic signs

| Identifier | Name | Coded representation |
| :---: | :--- | :--- |
| SA01 | Plus sign | $\mathrm{P} 2 / 11$ |
| SA02 | Plus-minus sign | $\mathrm{S} 3 / 1$ |
| SA03 | Less-than sign | $\mathrm{P} 3 / 12$ |
| SA04 | Equals sign | $\mathrm{P} 3 / 13$ |
| SA05 | Greater-than sign | $\mathrm{P} 3 / 14$ |
| SA06 | Division sign | $\mathrm{S} 3 / 8$ |
| SA07 | Multiplication sign | $\mathrm{S} 3 / 4$ |

## A.4.2.5 Subscripts and superscripts

| Identifier | Name | Coded representation |
| :---: | :--- | :--- |
| NS01 | Superscript one | S $5 / 1$ |
| NS02 | Superscript two | S $3 / 2$ |
| NS03 | Superscript three | S $3 / 3$ |

## A.4.2.6 Fractions

| Identifier | Name | Coded representation | Notes |
| :--- | :--- | :--- | :--- |
| NF01 | Vulgar fraction one half | $\mathrm{S} 3 / 13$ |  |
| NF04 | Vulgar fraction one quarter | $\mathrm{S} 3 / 12$ |  |
| NF05 | Vulgar fraction three quarters | $\mathrm{S} 3 / 14$ |  |
| NF18 | Vulgar fraction one eighth | $\mathrm{S} 5 / 12$ | $(1)$ |
| NF19 | Vulgar fraction three eighths | S $5 / 13$ | $(1)$ |
| NF20 | Vulgar fraction five eighths | $\mathrm{S} 5 / 14$ | $(1)$ |
| NF21 | Vulgar fraction seven eighths | $\mathrm{S} 5 / 15$ | $(1)$ |


| Identifier | Name | Coded representation | Notes |
| :---: | :---: | :---: | :---: |
| SM01 | Number sign | P $2 / 3$ |  |
| SM02 | Percent sign | P $2 / 5$ |  |
| SM03 | Ampersand | P $2 / 6$ |  |
| SM04 | Asterisk | P 2/10 |  |
| SM05 | Commercial AT | P 4/0 |  |
| SM06 | Left square bracket | P 5/11 |  |
| SM07 | Reverse solidus | P 5/12 | (1) |
| SM08 | Right square bracket | P 5/13 |  |
| SM11 | Left curly bracket | P 7/11 | (1) |
| SM12 | Horizontal bar | S 5/0 | (3) |
| SM13 | Verticle line | P 7/12 |  |
| SM14 | Right curly bracket | P 7/13 | (1) |
| SM17 | Micro sign | S 3/5 |  |
| SM18 | Ohm sign | S 6/0 |  |
| SM19 | Degree sign | S 3/0 |  |
| SM20 | Masculine ordinal indicator | S 6/11 |  |
| SM21 | Femenine ordinal indicator | S 6/3 |  |
| SM24 | Section sign | S $2 / 7$ |  |
| SM25 | Pilcrow sign | S 3/6 |  |
| SM26 | Middle dot | S 3/7 |  |
| SM30 | Leftwards arrow | S 2/12 | (1) |
| SM31 | Rightwards arrow | S 2/14 | (1) |
| SM32 | Upwards arrow | S 2/13 | (1) |
| SM33 | Downwards arrow | S 2/15 | (1) |
| SM52 | Copyright sign | S 5/3 | (1) |
| SM53 | Registered sign | S 5/2 | (1) |
| SM54 | Trade mark sign | S 5/4 | (1) |
| SM93 | Music note | S 5/5 | (1) |
| SM95 | Broken bar | S 5/7 | (2) |
| SM96 | Not sign | S 5/6 | (2) |


| Identifier | Name | Coded representation |  |
| :---: | :---: | :---: | :---: |
| SD11 | Acute accent | S 4/2 | 2/0 |
| SD13 | Grave accent | S 4/1 | 2/0 |
| SD15 | Circumflex accent | S 4/3 | 2/0 |
| SD17 | Diaeresis | S 4/8 | 2/0 |
| SD19 | Tilde | S 4/4 | 2/0 |
| SD21 | Caron | S 4/15 | 2/0 |
| SD23 | Breve | S 4/6 | 2/0 |
| SD25 | Double acute accent | S 4/13 | 2/0 |
| SD27 | Ring above | S 4/10 | 2/0 |
| SD29 | Dot above | S 4/7 | 2/0 |
| SD31 | Macron | S 4/5 | 2/0 |
| SD41 | Cedilla | S 4/11 | 2/0 |
| SD43 | Ogonek | S 4/14 | 2/0 |

Note - The grave accent, circumflex accent and tilde are also coded as P 6/0, P $5 / 14$ and P $7 / 14$ respectively.

## A.4.5 Special character

| Identifier | Name | Coded representation |
| :---: | :---: | :---: |
| SM34 | Delete | $7 / 15$ |

ANNEX B
(to Recommendation T.51)
Table of character and control sets

| Register No. | Name of registered set | (Note 1) | Final character | C0 | C1 | G0 | G1 | G2 | G3 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | Primary control set of IRV |  | 4/0 | X |  |  |  |  |  |
| 6 | Rec. T. 50 International reference version (IRV). This is also the primary set of Rec. T. 51 | 94 | 4/2 |  |  | X | X | X | X |
| 132 | Rec. T. 101 Data syntax I Primary control set | C0 | 4/9 | X |  |  |  |  |  |
| 125 | Rec. T. 101 Data Syntax I Supplementary control set | C1 | 4/4 |  | x |  |  |  |  |
| 42 | Rec. T. 101 Data Syntax I <br> Kanji set (2 byte set) JISC 6226 (1978) | $94 \times 94$ | 4/0 |  |  | X | X | X | x |
| 137 | Rec. T. 101 Data Syntax I Mosaic 1 set | 94 | 7/9 |  |  | X | x | x | x |
| 14 | Rec. T. 101 Data Syntax I <br> Primary character set JISC 6220 (1969) | 94 | 4/10 |  |  | X | X | X | X |
| 131 | Rec. T. 101 Data Syntax I Complete code | - | 4/3 |  |  |  |  |  |  |
| 134 | Rec. T. 101 Data Syntax II Primary control set | C0 | 4/10 | X |  |  |  |  |  |
| 56 | Rec. T. 101 Data Syntax II <br> Serial supplementary control set | C1 | 4/0 |  | X |  |  |  |  |
| 73 | Rec. T. 101 Data Syntax II Parallel supplementary control set | C1 | 4/1 |  | X |  |  |  |  |
| 70 | Rec. T. 101 Data Syntax II Supplementary character set | 94 | 6/2 |  |  | X | X | X | X |


| Register No. | Name of registered set | (Note 1) | Final character | C0 | C1 | G0 | G1 | G2 | G3 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 71 | Rec. T. 101 Data Syntax II 2nd supplementary mosaic set | 94 | 6/3 |  |  | X | X | X | X |
| 72 | Rec. T. 101 Data Syntax II 3rd supplementary mosaic set | 94 | 6/4 |  |  | X | X | X | X |
| 145 | Rec. T. 101 Data Syntax II Complete code | - | 4/4 |  |  |  |  |  |  |
| 150 | Rec. T. 52 Greek primary set | 94 | 2/1 4/0 |  |  | X | X | X | X |
| 108 | Rec. T. 101 Data Syntax III Complete code | - | 4/1 |  |  |  |  |  |  |
| 135 | Rec. T. 101 Data Syntax III Primary control set | C0 | 4/11 | x |  |  |  |  |  |
| 136 | Rec. T. 101 Data Syntax III Supplementary control set | C1 | 4/6 |  | X |  |  |  |  |
| 6 | Rec. T. 101 Data Syntax III Primary character set | 94 | 4/2 |  |  | X | X | X | x |
| 128 | Rec. T. 101 Data Syntax III Supplementary character set | 94 | 7/12 |  |  | X | X | X | X |
| Cannot be registered | Rec. T. 101 Data Syntax III PDI set | 96 | (5/7) <br> never <br> to be assigned by ISO |  |  |  | X | X | X |
| 129 | Rec. T. 101 Data Syntax III Mosaic set | 96 | 7/13 |  |  | X | X | X | X |
| 102 | Rec. T. 61 Primary graphic character set (left half of Fig. 2/T.61) | 94 | $7 / 5$ |  |  | X | X | X | X |
| 103 | Rec. T. 61 Supplementary character set (right half of Fig. 2/T.61) | 94 | 7/6 |  |  | X | X | X | X |
| 106 | Rec. T. 61 Primary control set | C0 | 4/5 | X |  |  |  |  |  |
| 107 | Rec. T. 61 Supplementary control set | C1 | 4/8 |  | X |  |  |  |  |


| Register <br> No. | Name of registered set |  | (Note 1) | Final character | C0 | C1 | G0 | G1 | G2 | G3 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 156 | Rec. T. 51 | Supplementary graphic set | 96 | 5/2 |  |  |  | x | x | x |
| $\begin{gathered} \text { Up-dated } \\ 87 \end{gathered}$ | Rec. T. 52 | Kanji set ( 2 byte set) JIS $\times$ 0208-1990 | $94 \times 94$ | $\begin{gathered} 4 / 2 \\ (\text { Note } 2) \end{gathered}$ |  |  | x | x | x | x |
| 164 | Rec. T. 52 | Hebrew supplementary set | 96 | 5/3 |  |  |  | x | x | x |
| 144 | Rec. T. 52 | Cyrillic supplementary set | 96 | 4/12 |  |  |  | x | x | x |
| 13 | Rec. T. 52 | Katakana set | 96 | 4/9 |  |  |  | x | x | x |
| 89 | Rec. T. 52 | Arabic set | 94 | 6/11 |  |  | x | x | x | x |
| 165 | Rec. T. 52 | Chinese set | $94 \times 94$ | 4/5 |  |  | x | x | x | x |
| 161 | Rec. T. 101 | Audio Data Syntax | Complete code | 4/5 |  |  |  |  |  |  |
| 162 | Rec. T. 101 | Photo-Videotex Data Syntax | Complete Code | 4/5 |  |  |  |  |  |  |

Note 1 - The 94 or 96 character set applies only to single byte graphic character sets. The $94 \times 94$ set applies to two byte graphic character sets. Primary control sets are C0. Supplementary control sets are C1. Complete code is also referred to as "Coding System different from that of ISO 2022".

Note 2 - The designation sequences are:

ESC 02/06 04/00 ESC 02/04 04/02 for G0
ESC 02/06 04/00 ESC 02/04 02/09 04/02 for G1
ESC 02/06 04/00 ESC 02/04 02/10 04/02 for G2
ESC 02/06 04/00 ESC 02/04 02/11 04/02 for G3

## ANNEX C

(to Recommendation T.51)

## Alphabetical list of abbreviations used in this Recommendation

## ESC ESCAPE

IRV International reference version

