

Extracts from ISO/IEC 10646-1 : 1993

Information technology — Universal Multiple-Octet Coded Character Set (UCS)

Part 1: Architecture and Basic Multilingual Plane

Committees responsible for this British Standard

The preparation of this British Standard was entrusted by the Information Systems Technology Standards Policy Committee (IST/-) to Technical Committee IST/2, upon which the following bodies were represented:

- British Computer Society
- British Federation of Printing Machinery and Supplies Ltd.
- British Telecommunications plc
- Digital Equipment Co. Ltd.
- IBM United Kingdom Ltd.
- Information Systems Committee of the Universities Funding Council
- Institution of Electrical Engineers
- International Computers Limited
- Philips Research Laboratories

This British Standard, having been prepared under the direction of the Information Technology Systems Standards Policy Committee, was published under the authority of the Standards Board and comes into effect on 15 August 1993

© BSI 1993

The following BSI references relate to the work on this standard:
Committee reference IST/2
Draft for comment 90/68644 DC

ISBN 0 580 22228 4

Amendments issued since publication

Amd. No.	Date	Text affected

Contents

	Page
Committees responsible	Inside front cover
National foreword	ii
1 Scope	1
2 Conformance	1
3 Normative references	2
4 Definitions	2
5 General structure of the UCS	3
6 Basic structure and nomenclature	4
7 Special features of the UCS	7
8 The Basic Multilingual Plane	7
9 Other planes	7
10 The Restricted Use zone	7
11 Private Use groups and planes	8
12 Revision and updating of the UCS	8
13 Subsets	8
14 Coded representation forms of the UCS	8
15 Implementation levels	9
16 Use of control functions with the UCS	9
17 Declaration of identification of features	9
18 Structure of the code tables and lists	11
19 Block names	11
20 Characters in bi-directional context	12
21 Special characters	12
22 Order of characters	12
23 Combining characters	13
24 Hangul syllable composition method	14
25 Code tables and lists of character names	14
Annexes	
A Collections of graphic characters for subsets	152
B List of combining characters	154
C Mirrored characters in Arabic bi-directional context	157
D Alternate format characters	159
E Alphabetically sorted list of character names	162
F The use of 'signatures' to identify UCS	196
G UCS transformation format (UTF-1)	197
H Recommendation for combined receiving/originating devices with internal storage	199
J Notations of octet value representations	200
K Character naming guidelines	201
L Sources of characters	203
M External references to character repertoires	205
N Scripts under consideration for future editions of ISO/IEC 10646	207

National foreword

This Published Document has been prepared under the direction of the Information Systems Technology Policy Committee, by technical committee IST/2.

It provides extracts from ISO/IEC 10646-1 : 1993¹⁾ *Information technology — Universal Multiple — Octet Coded Character Set (UCS) — Part 1: Architecture and Basic Multilingual Plane* for the convenience of those users who do not need the code tables of unified Han and Korean Hangul characters. We expect that this information will be most frequently required in the UK.

Extracts from the following pages are provided:

pages 1-129;

pages 240-261;

pages 699-754.

Compliance with a British Standard does not of itself confer immunity from legal obligations.

¹⁾ ISO/IEC 10646-1 : 1993 is implemented as BS ISO/IEC 10646-1 : 1993.

Information technology — Universal Multiple-Octet Coded Character Set (UCS) —

Part 1:

Architecture and Basic Multilingual Plane

1 Scope

ISO/IEC 10646 specifies the Universal Multiple-Octet Coded Character Set (UCS). It is applicable to the representation, transmission, interchange, processing, storage, input and presentation of the written form of the languages of the world as well as additional symbols.

This part of ISO/IEC 10646 specifies the overall architecture, and

- defines terms used in ISO/IEC 10646;
- describes the general structure of the coded character set;
- specifies the Basic Multilingual Plane (BMP) of the UCS, and defines a set of graphic characters used in scripts and the written form of languages on a world-wide scale;
- specifies the names for the graphic characters of the BMP, and the coded representations;
- specifies the four-octet (32-bit) canonical form of the UCS: UCS-4;
- specifies a two-octet (16-bit) BMP form of the UCS: UCS-2;
- specifies the coded representations for control functions;
- specifies the management of future additions to this coded character set.

The UCS is a coding system different from that specified in ISO 2022. The method to designate UCS from ISO 2022 is specified in 17.2.

2 Conformance

2.1 General

Whenever Private Use characters are used as specified in ISO/IEC 10646, the characters themselves shall not be covered by these conformance requirements.

2.2 Conformance of Information Interchange

A coded-character-data-element (CC-data-element) within coded information for interchange is in conformance with ISO/IEC 10646 if

- a) all the coded representations of graphic characters within that CC-data-element conform to clauses 6 and 7, to an identified form chosen from clause 14, and to an identified implementation level chosen from clause 15;
- b) all the graphic characters represented within that CC-data-element are taken from those within an identified subset (clause 13);
- c) all the coded representations of control functions within that CC-data-element conform to clause 16.

A claim of conformance shall identify the adopted form, the adopted implementation level and the adopted subset by means of a list of collections and/or characters.

2.3 Conformance of devices

A device is in conformance with ISO/IEC 10646 if it conforms to the requirements of item a) below, and either or both of items b) , and c).

NOTE - The term *device* is defined (in 4.17) as a component of information processing equipment which can transmit and/or receive coded information within CC-data-elements. A device may be a conventional

input/output device, or a process such as an application program or gateway function.

A claim of conformance shall identify the document that contains the description specified in a) below, and shall identify the adopted form(s), the adopted implementation level, the adopted subset (by means of a list of collections and/or characters), and the selection of control functions adopted in accordance with clause 16.

a) Device description: A device that conforms to ISO/IEC 10646 shall be the subject of a description that identifies the means by which the user may supply characters to the device and/or may recognise them when they are made available to the user, as specified respectively, in subclauses b), and c) below.

b) Originating device: An originating device shall allow its user to supply any characters from an adopted subset, and be capable of transmitting their coded representations within a CC-data-element in accordance with the adopted form and implementation level.

c) Receiving device: A receiving device shall be capable of receiving and interpreting any coded representation of characters that are within a CC-data-element in accordance with the adopted form and implementation level, and shall make any corresponding characters from the adopted subset available to the user in such a way that the user can identify them.

Any corresponding characters that are not within the adopted subset shall be indicated to the user in a way which need not allow them to be distinguished from each other.

NOTES

1 An indication to the user may consist of making available the same character to represent all characters not in the adopted subset, or providing a distinctive audible or visible signal when appropriate to the type of user.

2 See also annex H for receiving devices with re-transmission capability.

3 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO/IEC 10646. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO/IEC 10646 are encouraged to investigate the possibility of applying the most recent

editions of the standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards.

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

ISO/IEC 6429:1992 *Information technology — Control functions for coded character sets*.

4 Definitions

For the purposes of ISO/IEC 10646, the following definitions apply :

4.1 Basic Multilingual Plane (BMP) : Plane 00 of Group 00.

4.2 block : A contiguous collection of characters that share common characteristics, such as script.

4.3 canonical form : The form with which characters of this coded character set are specified using four octets to represent each character.

4.4 CC-data-element (Coded-Character-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.

4.5 cell : The place within a row at which an individual character may be allocated.

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

4.7 character boundary : Within a stream of octets the demarcation between the last octet of the coded representation of a character and the first octet of that of the next coded character.

4.8 coded character : A character together with its coded representation.

4.9 coded character set : A set of unambiguous rules that establishes a character set and the relationship between the characters of the set and their coded representation.

4.10 code table : A table showing the characters allocated to the octets in a code.

4.11 combining character : A member of an identified subset of the coded character set of ISO/IEC 10646 intended for combination with the preceding non-combining graphic character, or with a sequence of combining characters preceded by a non-combining character (see also 4.13).

NOTE - This part of ISO/IEC 10646 specifies several subset collections which include combining characters.

4.12 compatibility character : A graphic character included as a coded character of ISO/IEC 10646 primarily for compatibility with existing coded character sets.

4.13 composite sequence : A sequence of graphic characters consisting of a non-combining character followed by one or more combining characters (see also 4.11).

NOTES

1 A graphic symbol for a composite sequence generally consists of the combination of the graphic symbols of each character in the sequence.

2 A composite sequence is not a character and therefore is not a member of the repertoire of ISO/IEC 10646.

4.14 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 octets.

4.15 default state : The state that is assumed when no state has been explicitly specified.

4.16 detailed code table : A code table showing the individual characters, and normally showing a partial row.

4.17 device : A component of information processing equipment which can transmit and/or receive coded information within CC-data-elements. (It may be an input/output device in the conventional sense, or a process such as an application program or gateway function.)

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

4.19 graphic symbol : The visual representation of a graphic character or of a composite sequence.

4.20 group : A subdivision of the coding space of this coded character set; of 256 x 256 x 256 cells.

4.21 interchange : The transfer of character coded data from one user to another, using telecommunication means or interchangeable media.

4.22 interworking : The process of permitting two or more systems, each employing different coded character sets, meaningfully to interchange character coded data; conversion between the two codes may be involved.

4.23 octet : An ordered sequence of eight bits considered as a unit.

4.24 plane : A subdivision of a group; of 256 x 256 cells.

4.25 presentation; to present : The process of writing, printing, or displaying a graphic symbol.

4.26 presentation form : In the presentation of some scripts, a form of a graphic symbol representing a character that depends on the position of the character relative to other characters.

4.27 private use planes : Planes within this coded character set the contents of which are not specified in ISO/IEC 10646 (see 10.1).

4.28 repertoire : A specified set of characters that are represented in a coded character set.

4.29 row : A subdivision of a plane; of 256 cells.

4.30 script : A set of graphic characters used for the written form of one or more languages.

4.31 supplementary planes : Planes that accommodate characters which have not been allocated to the Basic Multilingual Plane.

4.32 user : A person or other entity that invokes the service provided by a device. (This entity may be a process such as an application program if the "device" is a code converter or a gateway function, for example.)

4.33 zone : A sequence of cells of a code table, comprising one or more rows, either in whole or in part, containing characters of a particular class (see clause 8).

5 General structure of the UCS

The general structure of the Universal Multiple-Octet Coded Character Set (referred to hereafter as "this coded character set") is described in this explanatory clause, and is illustrated in figures 1 and 2. The normative specification of the structure is given in later clauses.

The value of any octet is expressed in hexadecimal notation from 00 to FF in ISO/IEC 10646 (see annex J).

The canonical form of this coded character set — the way in which it is to be conceived — uses a four-dimensional coding space, regarded as a single entity, consisting of 128 three-dimensional groups.

NOTE - Thus, bit 8 of the most significant octet in the canonical form of a coded character can be used for internal processing purposes within a device as long as it is set to zero within a conforming CC-data-element.

Each group consists of 256 two-dimensional planes.

Each plane consists of 256 one-dimensional rows, each row containing 256 cells. A character is located and coded at a cell within this coding space or the cell is declared unused.

In the canonical form, four octets are used to represent each character, and they specify the group, plane, row and cell, respectively. The canonical form consists of four octets since two octets are not sufficient to cover all the characters in the world, and a 32-bit representation follows modern processor architectures.

The four-octet canonical form can be used as a four-octet coded character set in which case it is called UCS-4.

The first plane (Plane 00 of Group 00) is called the Basic Multilingual Plane. The Basic Multilingual Plane includes characters in general use in alphabetic, syllabic and ideographic scripts together with various symbols and digits. The BMP also has a Restricted Use (RU) zone in which the characters have special characteristics.

The subsequent planes are regarded as supplementary or private use planes, which will accommodate additional graphic characters.

The 32 planes with Plane-octet values E0 to FF of Group 00 are for Private Use. The 32 groups with Group-octet values 60 to 7F of this coded character set are also for Private Use. The contents of the cells in Private Use zones are not specified in ISO/IEC 10646.

Each character is located within the coded character set in terms of its Group-octet, Plane-octet, Row-octet, and Cell-octet.

In addition to the canonical form, a two-octet BMP form is specified. Thus, the Basic Multilingual Plane can be used as a two-octet coded character set identified as UCS-2.

Subsets of the coding space may be used in order to give a sub-repertoire of graphic characters.

A UCS Transformation Format (UTF-1) is specified in annex G which can be used to transmit text data through communication systems which are sensitive to octet values for control characters coded according to the structure of ISO 2022.

6 Basic structure and nomenclature

6.1 Structure

The Universal Multiple-Octet Coded Character Set as specified in ISO/IEC 10646 shall be regarded as

a single entity.

This entire coded character set shall be conceived of as comprising 128 groups of 256 planes. Each plane shall be regarded as containing 256 rows of characters, each row containing 256 cells. In a code table representing the contents of a plane (such as in figure 2), the horizontal axis shall represent the least significant octet, with its smaller value to the left; and the vertical axis shall represent the more significant octet, with its smaller value at the top.

Each axis of the coding space shall be coded by one octet. Within each octet the most significant bit shall be bit 8 and the least significant bit shall be bit 1.

Accordingly, the weight allocated to each bit shall be

bit 8	bit 7	bit 6	bit 5	bit 4	bit 3	bit 2	bit 1
128	64	32	16	8	4	2	1

6.2 Coding of characters

In the canonical form of the coded character set, each character within the entire coded character set shall be represented by a sequence of four octets. The most significant octet of this sequence shall be the group-octet. The least significant octet of this sequence shall be the cell-octet. Thus this sequence may be represented as

m.s.			l.s.
Group-octet	Plane-octet	Row-octet	Cell-octet

where m.s. means the most significant octet, and l.s. means the least significant octet.

For brevity, the octets may be termed

m.s.			l.s.	
G-octet	P-octet	R-octet	C-octet	

Where appropriate, these may be further abbreviated to G, P, R, and C.

The value of any octet shall be represented by two hexadecimal digits, for examples: 31 or FE. When a single character is to be identified in terms of the values of its group, plane, row and cell, this shall be represented such as:

0000 0030 for DIGIT ZERO

0000 0041 for LATIN CAPITAL LETTER A

When referring to characters within a plane, the leading four zeros (for G-octet and P-octet) may be omitted. For example, 0030 may be used to refer to DIGIT ZERO.

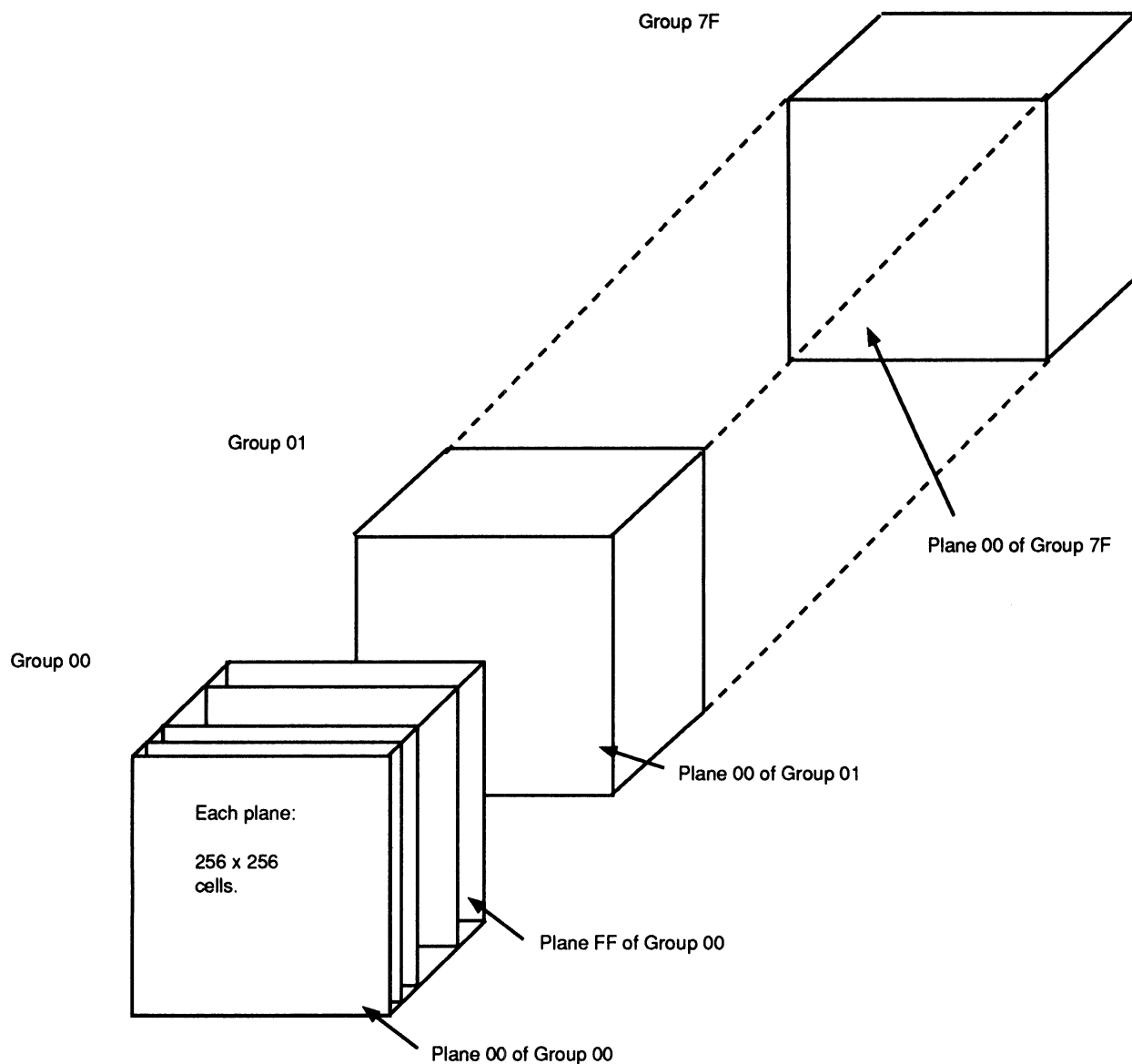
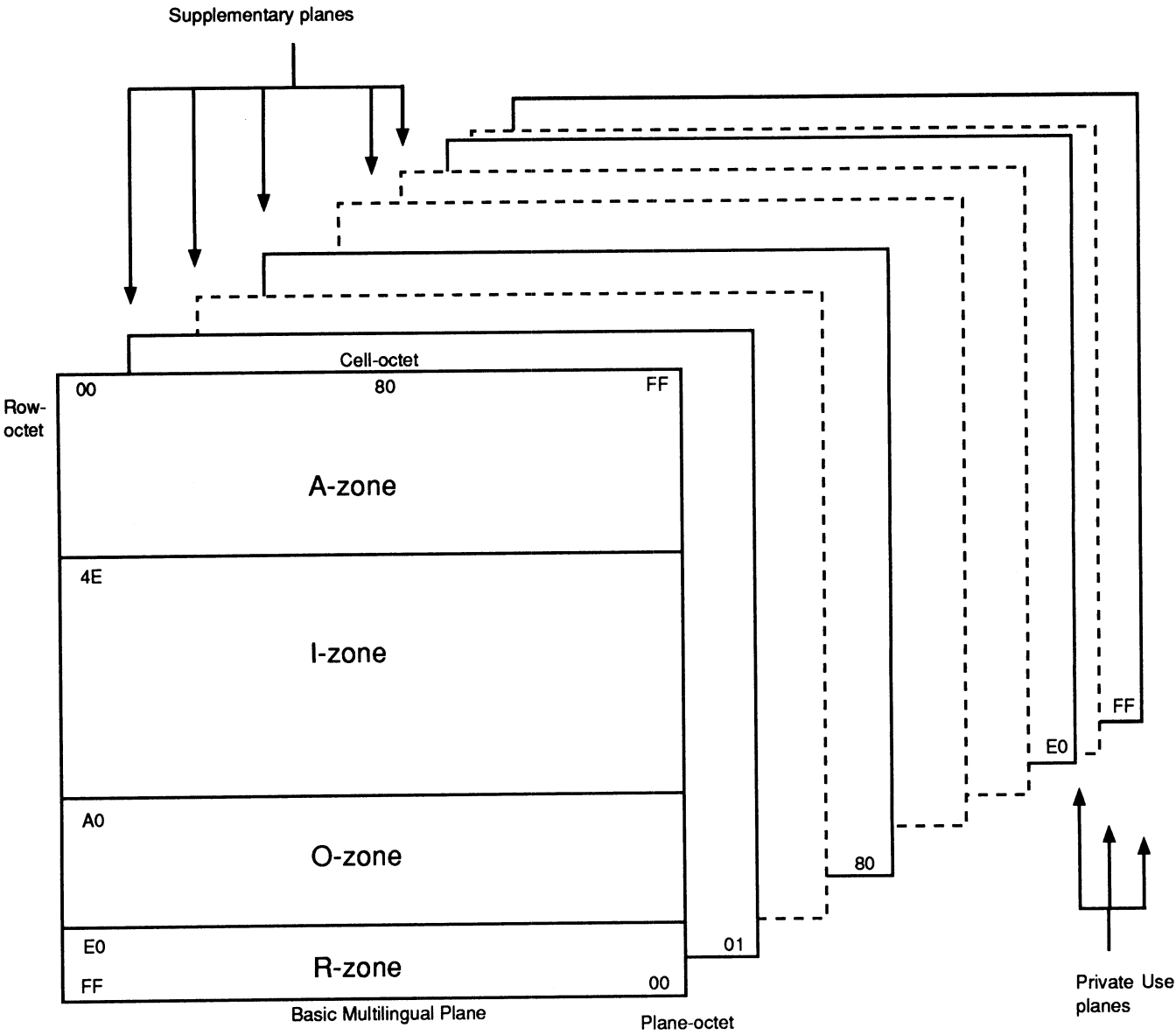


Figure 1 - Entire coding space of the Universal Multiple-Octet Coded Character Set



Labels A-zone, I-zone, O-zone, and R-zone are specified in clause 8.

Figure 2 - Group 00 of the Universal Multiple-Octet Coded Character Set

6.3 Octet order

The sequence of the octets that represent a character, and the most significant and least significant ends of it, shall be maintained as shown above. When serialised as octets, a more significant octet shall precede less significant octets. When not serialised as octets, the order of octets may be specified by agreement between sender and recipient (see 17.1 and annex F).

7 Special features of the UCS

The following characteristics apply to the entire coded character set.

1. The values of P-, and R-, and C-octets used for representing graphic characters shall be in the range 00 to FF. The values of G-octets used for representation of graphic characters shall be in the range 00 to 7F. On any plane, code positions FFFE and FFFF shall not be used.

NOTE - Code position FFFE is reserved for "signature" (see annex F). Code position FFFF can be used for internal processing uses requiring a numeric value that is guaranteed not to be a coded character such as in terminating tables, or signaling end-of-text. Since it is the largest two-octet value, it may also be used as the final value in binary or sequential searching index.

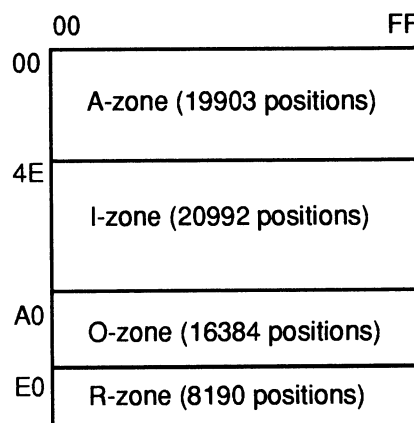
2. Code positions to which a character is not allocated, except for the positions reserved for Private Use characters, are reserved for future standardisation and shall not be used for any other purpose. Future editions of ISO/IEC 10646 will not allocate any characters to code positions reserved for Private Use characters.
3. The same graphic character shall not be allocated to more than one code position. There are graphic characters with similar shapes in the coded character set; they are used for different purposes and have different character names.
4. Compatibility characters are included in ISO/IEC 10646 primarily for compatibility with existing coded character sets to allow two-way code conversion without loss of information.

8 The Basic Multilingual Plane

Plane 00 of Group 00 shall be the Basic Multilingual Plane (BMP). The BMP can be used as a two-octet coded character set in which case it shall be called UCS-2 (see 14.1).

The Basic Multilingual Plane shall be divided into four zones:

A-zone: code positions 0000 to 4DFF
 I-zone: code positions 4E00 to 9FFF
 O-zone: code positions A000 to DFFF
 R-zone: code positions E000 to FFFD



Code positions 0000 to 001F in the BMP are reserved for control characters, and code position 007F is reserved for the character DELETE (see clause 16). Code positions 0080 to 009F are reserved.

In the Basic Multilingual Plane, the A-zone is used for alphabetic and syllabic scripts together with various symbols. The I-zone is used for Chinese/Japanese/Korean (CJK) unified ideographs (unified East Asian ideographs). The O-zone is reserved for future standardisation. The R-zone shall be used for the Restricted Use zone in the BMP which contains Private Use characters, presentation forms, and compatibility characters (see clause 10).

9 Other planes

Planes 01 to DF in Group 00 and planes 00 to FF in Groups 01 to 5F are reserved for future standardisation, and thus those code positions shall not be used for any other purpose.

10 The Restricted Use zone

Sets of graphic characters that are used in particular ways are provided in the Restricted Use zone. These sets include:

- a) Private Use characters,
- b) Presentation forms of characters,
- c) Compatibility characters (see item 4 in clause 7).

10.1 Private Use characters

Private Use characters are not restrained in any way by ISO/IEC 10646. Private Use characters can be used to provide user-defined characters. For example, this is a common requirement for users of ideographic scripts.

NOTE 1 - For meaningful interchange of Private Use characters, an agreement, independent of ISO/IEC 10646, is necessary between sender and recipient.

Private Use characters can be used for dynamically-redefinable characters (DRCS) applications.

NOTE 2 - For meaningful interchange of DRCS, an agreement, independent of ISO/IEC 10646 is necessary between sender and recipient. ISO/IEC 10646 does not specify the techniques for defining or setting up dynamically-redefinable characters.

10.2 Presentation forms of characters

Each presentation form of character provides an alternative form, for use in a particular context, to the nominal form of the character or sequence of characters from the other zones of graphic characters. The transformation from the nominal form to the presentation forms may involve substitution, superimposition, or combination.

The rules for the superimposition, choice of differently shaped characters, or combination into ligatures, or conjuncts — which are often of extreme complexity — are not specified in ISO/IEC 10646.

In general, presentation forms are not intended to be used as a substitute for the nominal forms of the graphic characters specified elsewhere within this coded character set. However, specific applications may encode these presentation forms instead of the nominal forms for specific reasons among which is compatibility with existing devices. The rules for searching, sorting and other processing operations on presentation forms are outside the scope of ISO/IEC 10646.

11 Private Use groups and planes

The code positions of 32 planes from Plane E0 to Plane FF of Group 00 shall be for Private Use.

The code positions of the 32 groups from Group 60 to Group 7F shall be for Private Use.

The contents of these code positions are not specified in ISO/IEC 10646 (see 10.1).

12 Revision and updating of the UCS

The revision and updating of this coded character set will be carried out by ISO/IEC JTC1/SC2.

NOTE - It is intended that in future editions of ISO/IEC 10646, the names and allocation of the characters in this edition will remain unchanged.

13 Subsets

ISO/IEC 10646 provides the specification of subsets of coded graphic characters for use in interchange, by originating devices and by receiving devices.

There are two alternatives for the specification of subsets; limited subset and selected subset. An adopted subset may comprise either of them, or a combination of the two.

13.1 Limited subset

A limited subset consists of a list of graphic characters in the specified subset. This specification allows applications and devices that were developed using other codes to interwork with this coded character set.

A claim of conformance referring to a limited subset shall list the graphic characters in the subset by the names of graphic characters or code positions as defined in ISO/IEC 10646.

13.2 Selected subset

A selected subset consists of a list of collections of graphic characters as defined in ISO/IEC 10646. The collections from which the selection may be made are listed in annex A of each part of ISO/IEC 10646. A selected subset shall always automatically include the Cells 20 to 7E of Row 00 of Plane 00 of Group 00.

A claim of conformance referring to a selected subset shall list the collections chosen as defined in ISO/IEC 10646.

14 Coded representation forms of the UCS

ISO/IEC 10646 provides two alternative forms of coded representation of characters.

NOTE - The characters from the ISO 646 IRV repertoire are coded by simple zero extensions to their coded representations in ISO 646 IRV. Therefore, their coded representations have the same integer values when represented as 8-bit, 16-bit, or 32-bit integers. For implementations sensitive to a zero valued octet (e.g. for use as a string terminator), use of 8-bit base-2 array data

type should be avoided as any zero valued octet may be interpreted incorrectly. Use of data types at least 16-bits wide is more suitable for UCS-2, and use of data types at least 32-bits wide is more suitable for UCS-4.

14.1 Two-octet BMP form

This coded representation form permits the use of characters from the Basic Multilingual Plane with each character represented by two octets.

Within a CC-data-element conforming to the two-octet BMP form, a character from the Basic Multilingual Plane shall be represented by two octets comprising the R-octet and the C-octet as specified in 6.2.

NOTE - A coded graphic character using the two-octet BMP form may be implemented by a 16-bit integer for processing.

14.2 Four-octet canonical form

The canonical form permits the use of all the characters of ISO/IEC 10646, with each character represented by four octets.

Within a CC-data-element conforming to the four-octet canonical form, every character shall be represented by four octets comprising the G-octet, the P-octet, the R-octet and the C-octet as specified in 6.2.

NOTE - A coded graphic character using the four-octet canonical form may be implemented by a 32-bit integer for processing.

15 Implementation levels

ISO/IEC 10646 specifies three levels of implementation. Combining characters are described in clause 23 and listed in annex B.

15.1 Implementation level 1

When implementation level 1 is used, a CC-data-element shall not contain coded representations of combining characters (see clause B.1) nor of characters from HANGUL JAMO block (see clause 24).

15.2 Implementation level 2

When implementation level 2 is used, a CC-data-element shall not contain coded representations of characters listed in clause B.2.

15.3 Implementation level 3

When implementation level 3 is used, a CC-data-element may contain coded representations of any characters.

16 Use of control functions with the UCS

This coded character set provides for use of control functions encoded according to ISO 2022, ISO/IEC 6429 or similarly structured standards for control functions, and standards derived from these. A set or subset of such coded control functions may be used in conjunction with this coded character set. These standards encode a control function as a sequence of one or more octets.

When a C0 control character of ISO/IEC 6429 is used with this coded character set, its coded representation as specified in ISO/IEC 6429 shall be padded to correspond with the number of octets in the adopted form (see clause 14). Thus, the least significant octet shall be the bit combination specified in ISO/IEC 6429, and the more significant octet(s) shall be zeros.

For example, the control character FORM FEED is represented by "000C" in the two-octet form, and "0000 000C" in the four-octet form.

For escape sequences, control sequences, and control strings (see ISO/IEC 6429) consisting of a coded control character followed by additional bit combinations in the range 20 to 7F, each bit combination shall be padded by octet(s) with value 00.

For example, the escape sequence "ESC 02/00 04/00" is represented by "001B 0020 0040" in the two-octet form, and "0000 001B 0000 0020 0000 0040" in the four-octet form.

When using a C1 control character of ISO/IEC 6429 with this coded character set, it shall be coded as ESC Fe sequence (see ISO/IEC 6429) padded as specified above.

For example, the control character PARTIAL LINE BACKWARD – PLU (08/12 in ISO/IEC 6429 representation) is represented by "001B 004C" in the two-octet form, and "0000 001B 0000 004C" in the four-octet form.

Code extension control functions for the ISO 2022 code extension techniques (such as designation escape sequence, single shift and locking shift) shall not be used with this coded character set.

17 Declaration of identification of features

17.1 Purpose and context of identification

CC-data-elements conforming to ISO/IEC 10646 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 ISO/IEC 10646 (including the form), the implementation level, and any subset of the coding space that have been adopted by the originator must also be available to the recipient. The route by which such identification is communicated to the recipient is outside the scope of ISO/IEC 10646.

However, some standards for interchange of coded information may permit, or require, that the coded representation of the identification applicable to the CC-data-element forms a part of the interchanged information. This clause specifies a coded representation for the identification of UCS with an implementation level and a subset of ISO/IEC 10646, and also of a C0 and a C1 set of control functions from ISO/IEC 6429 for use in conjunction with ISO/IEC 10646. Such coded representations provide all or part of an identification data element, which may be included in information interchange in accordance with the relevant standard.

If two or more of the identifications are present, the order of those identifications shall follow the order as specified in this clause.

NOTE - An alternative method of identification is described in annex M.

17.2 Identification of UCS coded representation form with implementation level

When the escape sequences from ISO 2022 are used, the identification of a coded representation form of UCS (see clause 14) and an implementation level (see clause 15) specified by ISO/IEC 10646 shall be by a designation sequence chosen from the following list:

ESC 02/05 02/15 04/00	UCS-2 with implementation level 1
ESC 02/05 02/15 04/01	UCS-4 with implementation level 1
ESC 02/05 02/15 04/03	UCS-2 with implementation level 2
ESC 02/05 02/15 04/04	UCS-4 with implementation level 2
ESC 02/05 02/15 04/05	UCS-2 with implementation level 3
ESC 02/05 02/15 04/06	UCS-4 with implementation level 3

If such an escape sequence appears within a CC-data-element conforming to ISO/IEC 10646, it

shall be padded in accordance with clause 16.

17.3 Identification of subsets of graphic characters

When the control sequences of ISO/IEC 6429 are used, the identification of subsets (see clause 13) specified by ISO/IEC 10646 shall be by a control sequence IDENTIFY UNIVERSAL CHARACTER SUBSET (IUCS) as shown below.

CSI Ps... 02/00 06/13

Ps... means that there can be any number of selective parameters. The parameters are to be taken from the subset collection numbers as shown in annex A of each part of ISO/IEC 10646. When there is more than one parameter, each parameter value is separated by an octet with value 03/11.

Parameter values are represented by digits where octet values 03/00 to 03/09 represent digits 0 to 9.

If such a control sequence appears within a CC-data-element conforming to ISO/IEC 10646, it shall be padded in accordance with clause 16.

17.4 Identification of control function set

When the escape sequences from ISO 2022 are used, the identification of each set of control functions (see clause 16) of ISO/IEC 6429 to be used in conjunction with ISO/IEC 10646 shall be an identifier sequence of the type shown below.

ESC 02/01 04/00	identifies the full C0 set of ISO/IEC 6429
ESC 02/02 04/03	identifies the full C1 set of ISO/IEC 6429

For a subset of C0 or C1 sets, the final octet F shall be obtained from the International Register of Coded Character Sets. The identifier sequences for these sets shall be:

ESC 02/01 F	identifies a C0 set
ESC 02/02 F	identifies a C1 set

If such an escape sequence appears within a CC-data-element conforming to ISO/IEC 10646, it shall be padded in accordance with clause 16.

17.5 Identification of return from UCS to ISO 2022

When the escape sequences from ISO 2022 are used, the identification of the return from UCS to the coding system of ISO 2022 shall be by the escape sequence ESC 02/05 04/00, padded in accordance with clause 16.

18 Structure of the code tables and lists

The clauses 25 and 26 set out the detailed code tables and the lists of character names for the graphic characters. Together, these specify graphic characters, their coded representation, and the character name for each character.

The graphic symbols are to be regarded as typical visual representations of the characters. ISO/IEC 10646 does not attempt to prescribe the exact shape of each character. The shape is affected by the design of the font employed, which is outside the scope of ISO/IEC 10646.

Graphic characters specified in ISO/IEC 10646 are uniquely identified by their names. This does not imply that the graphic symbols by which they are commonly imaged are always different. Examples of graphic characters with similar graphic symbols are LATIN CAPITAL LETTER A, GREEK CAPITAL LETTER ALPHA, and CYRILLIC CAPITAL LETTER A.

The meaning attributed to any character is not specified by ISO/IEC 10646; it may differ from country to country, or from one application to another.

For the alphabetic scripts, the general principle has been to arrange the characters within any row in approximate alphabetic sequence; where the script has capital and small letters, these are arranged in pairs. However, this general principle has been overridden in some cases. For example, for those scripts for which a relevant standard exists, the characters are allocated according to that standard. This arrangement within the code tables will aid conversion between the existing standards and this coded character set. In general, however, it is anticipated that conversion between this coded character set and any other coded character set will use a table lookup technique.

It is not intended, nor will it often be the case, that the characters needed by any one user will be found all grouped together in one part of the code table.

Furthermore, the user of any script will find that characters he needs may have been already coded earlier in this coded character set. This especially applies to the digits, to the symbols, and to the use of Latin letters in dual-script applications.

Therefore, in using this coded character set, the reader is advised to refer first to the block names list in clause 19 or an overview of the BMP in figure 3, and then to turn to the specific code table rows for the relevant script and for symbols and digits. In

addition, annex E contains an alphabetically sorted list of character names.

19 Block names

The following list contains the blocks defined in the BMP. The block names are used in providing for the specification of subsets (see annex A for subset collections).

Block name	from	to
BASIC LATIN	0020	007E
LATIN-1 SUPPLEMENT	00A0	00FF
LATIN EXTENDED-A	0100	017F
LATIN EXTENDED-B	0180	024F
IPA EXTENSIONS	0250	02AF
SPACING MODIFIER LETTERS	02B0	02FF
COMBINING DIACRITICAL MARKS	0300	036F
BASIC GREEK	0370	03CF
GREEK SYMBOLS AND COPTIC	03D0	03FF
CYRILLIC	0400	04FF
ARMENIAN	0530	058F
HEBREW EXTENDED-A	0590	05CF
BASIC HEBREW	05D0	05EA
HEBREW EXTENDED-B	05EB	05FF
BASIC ARABIC	0600	0652
ARABIC EXTENDED	0653	06FF
DEVANAGARI	0900	097F
BENGALI	0980	09FF
GURMUKHI	0A00	0A7F
GUJARATI	0A80	0AFF
ORIYA	0B00	0B7F
TAMIL	0B80	0BFF
TELUGU	0C00	0C7F
KANNADA	0C80	0CFF
MALAYALAM	0D00	0D7F
THAI	0E00	0E7F
LAO	0E80	0EFF
BASIC GEORGIAN	10D0	10FF
GEORGIAN EXTENDED	10A0	10CF
HANGUL JAMO	1100	11FF
LATIN EXTENDED ADDITIONAL	1E00	1EFF
GREEK EXTENDED	1F00	1FFF
GENERAL PUNCTUATION	2000	206F
SUPERSCRIPTS AND SUBSCRIPTS	2070	209F
CURRENCY SYMBOLS	20A0	20CF
COMBINING DIACRITICAL MARKS FOR SYMBOLS	20D0	20FF
LETTERLIKE SYMBOLS	2100	214F
NUMBER FORMS	2150	218F
ARROWS	2190	21FF
MATHEMATICAL OPERATORS	2200	22FF
MISCELLANEOUS TECHNICAL	2300	23FF
CONTROL PICTURES	2400	243F
OPTICAL CHARACTER RECOGNITION	2440	245F
ENCLOSED ALPHANUMERICS	2460	24FF
BOX DRAWING	2500	257F
BLOCK ELEMENTS	2580	259F
GEOMETRIC SHAPES	25A0	25FF
MISCELLANEOUS SYMBOLS	2600	26FF
DINGBATS	2700	27BF
CJK SYMBOLS AND PUNCTUATION	3000	303F
HIRAGANA	3040	309F
KATAKANA	30A0	30FF



BOPOMOFO	3100 - 312F
HANGUL COMPATIBILITY JAMO	3130 - 318F
CJK MISCELLANEOUS	3190 - 319F
ENCLOSED CJK LETTERS AND MONTHS	
	3200 - 32FF
CJK COMPATIBILITY	3300 - 33FF
HANGUL	3400 - 3D2D
HANGUL SUPPLEMENTARY-A	3D2E - 44B7
HANGUL SUPPLEMENTARY-B	44B8 - 4DFF
CJK UNIFIED IDEOGRAPHS	4E00 - 9FFF
PRIVATE USE AREA	E000 - F8FF
CJK COMPATIBILITY IDEOGRAPHS	F900 - FAFF
ALPHABETIC PRESENTATION FORMS	FB00 - FB4F
ARABIC PRESENTATION FORMS-A	FB50 - FDFF
COMBINING HALF MARKS	FE20 - FE2F
CJK COMPATIBILITY FORMS	FE30 - FE4F
SMALL FORM VARIANTS	FE50 - FE6F
ARABIC PRESENTATION FORMS-B	FE70 - FEFE
HALFWIDTH AND FULLWIDTH FORMS	FF00 - FFEF
SPECIALS	FFF0 - FFFF

20 Characters in bi-directional context

A class of left/right handed pairs of characters have special significance in the context of bi-directional text. In this context the terms LEFT or RIGHT in the character name are also intended to imply "opening" or "closing" forms of character shape, rather than a strict left-hand or right-hand form. These characters are listed below.

Code	Name
0028	LEFT PARENTHESIS
0029	RIGHT PARENTHESIS
005B	LEFT SQUARE BRACKET
005D	RIGHT SQUARE BRACKET
007B	LEFT CURLY BRACKET
007D	RIGHT CURLY BRACKET
2045	LEFT SQUARE BRACKET WITH QUILL
2046	RIGHT SQUARE BRACKET WITH QUILL
207D	SUPERSCRIP LEFT PARENTHESIS
207E	SUPERSCRIP RIGHT PARENTHESIS
208D	SUBSCRIPT LEFT PARENTHESIS
208E	SUBSCRIPT RIGHT PARENTHESIS
2329	LEFT-POINTING ANGLE BRACKET
232A	RIGHT-POINTING ANGLE BRACKET
3008	LEFT ANGLE BRACKET
3009	RIGHT ANGLE BRACKET
300A	LEFT DOUBLE ANGLE BRACKET
300B	RIGHT DOUBLE ANGLE BRACKET
300C	LEFT CORNER BRACKET
300D	RIGHT CORNER BRACKET
300E	LEFT WHITE CORNER BRACKET
300F	RIGHT WHITE CORNER BRACKET
3010	LEFT BLACK LENTICULAR BRACKET
3011	RIGHT BLACK LENTICULAR BRACKET
3014	LEFT TORTOISE SHELL BRACKET
3015	RIGHT TORTOISE SHELL BRACKET
3016	LEFT WHITE LENTICULAR BRACKET
3017	RIGHT WHITE LENTICULAR BRACKET
3018	LEFT WHITE TORTOISE SHELL BRACKET
3019	RIGHT WHITE TORTOISE SHELL BRACKET
301A	LEFT WHITE SQUARE BRACKET
301B	RIGHT WHITE SQUARE BRACKET

The interpretation and rendering of any of these

characters depend on the state of the SYMMETRIC SWAPPING (see D.2.2) and on the direction of the character being rendered that are in effect at the point in the CC-data-element where the coded representation of the character appears.

For example, if the SYMMETRIC SWAPPING is ACTIVATED and if the direction of the character is from right to left, the character shall be interpreted as if the term LEFT or RIGHT in its name had been replaced by the term RIGHT or LEFT, respectively.

NOTE - In the context of Arabic bi-directional text, a set of mathematical symbols may also have special significance (see annex C).

21 Special characters

There are some characters that do not have printable graphic symbols. These characters include space characters. They are

Code	Name
0020	SPACE
00A0	NO-BREAK SPACE
2000	EN QUAD
2001	EM QUAD
2002	EN SPACE
2003	EM SPACE
2004	THREE-PER-EM SPACE
2005	FOUR-PER-EM SPACE
2006	SIX-PER-EM SPACE
2007	FIGURE SPACE
2008	PUNCTUATION SPACE
2009	THIN SPACE
200A	HAIR SPACE
3000	IDEOGRAPHIC SPACE

Currency symbols in ISO/IEC 10646 do not necessarily identify the currency of a country. For example, YEN SIGN can be used for Japanese yen and Chinese yuan. Also, DOLLAR SIGN is used in numerous countries including the United States of America.

There is a special class of characters called Alternate Format Characters which are included for compatibility with some industry practices. These are described in annex D.

22 Order of characters

Usually, coded characters appear in a CC-data-element in logical order (logical or backing store order corresponds to the order in which characters are entered from the keyboard—after corrections such as insertions, deletions, and overtyping have taken place). This applies even when characters of different dominant direction are

mixed: left-to-right (Greek, Latin, Thai) with right-to-left (Arabic, Hebrew), or with vertical (Mongolian) script.

Some characters may not appear linearly in final rendered text. For example, the medial form of the short *i* in Devanagari is displayed before a character that it logically follows in the CC-data-element.

23 Combining characters

This clause specifies the use of combining characters. A list of combining characters is shown in clause B.1. A list of combining characters not allowed in implementation level 2 is shown in clause B.2.

NOTE - The names of many script-independent combining characters contain the word "COMBINING".

23.1 Order of combining characters

Coded representations of combining characters shall follow that of the graphic character with which they are associated (for example, coded representations of LATIN SMALL LETTER A followed by COMBINING TILDE represent a composite sequence for Latin "ã").

If a combining character is to be regarded as a composite sequence in its own right, it shall be coded as a composite sequence by association with the character SPACE. For example, grave accent can be composed as SPACE followed by COMBINING GRAVE ACCENT.

NOTE - Indic matras form a special category of combining characters, since the presentation can depend on more than one of the surrounding characters. Thus it might not be desirable to associate Indic matra with the character SPACE.

23.2 Appearance in code tables

Combining characters intended to be positioned relative to the associated character are depicted in the character code tables above, below, to the right of, to the left of, in, or through a dotted circle. In presentation, these characters are intended to be positioned relative to the preceding base character in some manner, and not to stand alone or function as base characters. This is the motivation for the term "combining". Diacritics are the principal class of combining characters used in European alphabets.

In the code tables for some scripts, such as Hebrew, Arabic, and the scripts of India and South-east Asia, combining characters are indicated in relation to dotted circles to show their position relative to the

base character. Many of these combining characters encode vowel letters; as such they are not generally referred to as "diacritical marks".

23.3 Multiple combining characters

There are instances where more than one combining character is applied to a single graphic character. ISO/IEC 10646 does not restrict the number of combining characters that can follow a base character. The following rules shall apply:

1. If the combining characters can interact in presentation (for example, a combining macron and a combining diaeresis), then the position of the combining characters in the resulting graphic display is determined by the order of the coded representation of the combining characters. The presentations of combining characters are to be positioned from the base character outward. For example, combining characters placed above a base character are stacked vertically, starting with the first encountered in the sequence of coded representations and continuing for as many marks above as are required by the coded combining characters following the coded base character. For combining characters placed below a base character, the situation is inverted, with the combining characters starting from the base character and stacking downward.

An example of multiple combining characters above the base character is found in Thai, where a consonant letter can have above it one of the vowels 0000 0E34 to 0000 0E37 and, above that, one of four tone marks 0000 0E48 to 0000 0E4B. The order of the coded representation is: base consonant, followed by vowel, followed by one tone mark.

2. Some specific combining characters override the default stacking behaviour by being positioned horizontally rather than stacking, or by forming a ligature with an adjacent combining character. When positioned horizontally, the order of coded representations is reflected by positioning in the dominant order of the script with which they are used. For example, horizontal accents in a left-to-right script are coded left-to-right.

Prominent characters that show such override behaviour are associated with specific scripts or alphabets. For example, the COMBINING GREEK KORONIS (0000 0343) requires that, together with a following acute or grave accent, they be rendered side-by-side above a letter, rather than the accent marks being stacked above the COMBINING GREEK KORONIS. The order of the coded representations is: the letter itself, followed by that of the breathing mark, followed by that of the accent

marks. Two Vietnamese tone marks which have the same graphic appearance as the Latin acute and grave accent marks do not stack above the three Vietnamese vowel letters which already contain the circumflex diacritic (â, ê, ô). Instead, they form ligatures with the circumflex component of the vowel letters.

3. If the combining characters do not interact in presentation (for example, when one combining character is above a graphic character and another is below), the resultant graphic symbol from the base character and combining characters in different orders may appear the same. For example, the coded representations of LATIN SMALL LETTER A, followed by COMBINING CARON, followed by COMBINING OGONEK may result in the same graphic symbol as the coded representations of LATIN SMALL LETTER A, followed by COMBINING OGONEK, followed by COMBINING CARON.

Combining characters in Hebrew or Arabic scripts do not normally interact. Therefore, the sequence of their coded representations in a composite sequence does not affect its graphic symbol. The rules for forming the combined graphic symbol are beyond the scope of ISO/IEC 10646.

NOTE - Where combining characters are used for the generation of composite sequences in implementation level 3, this facility may be used to provide an alternative coded representation of text. For example, in implementation level 3 the French word "là" may be represented by the characters LATIN SMALL LETTER L followed by LATIN SMALL LETTER A WITH GRAVE, or may be represented by the characters LATIN SMALL LETTER L followed by LATIN SMALL LETTER A followed by COMBINING GRAVE ACCENT.

23.4 Collections containing combining characters

In some collections of characters listed in annex A, such as collections 14 (BASIC ARABIC) or 25 (THAI), both combining characters and non-combining characters are included.

When implementation level 1 or 2 is adopted, a CC-data-element shall not contain the coded representations of combining characters listed in clause B.1 or B.2 respectively, even though the adopted subset may include them.

Other collections of characters listed in annex A comprise only combining characters, for example

collection 7 (COMBINING DIACRITICAL MARKS). Such a collection shall not be included in the adopted subset when implementation level 1 or 2 is adopted.

24 Hangul syllable composition method

In rendering, a sequence of Hangul Jamo (from HANGUL JAMO block: 1100 to 11FF) are displayed as a series of syllable blocks. Jamo can be classified into three classes: Choseong (syllable-initial character), Jungseong (syllable-peak character), and Jongseong (syllable-final character). A complete syllable block is composed of a Choseong and a Jungseong, and optionally a Jongseong.

An incomplete syllable is a string of one or more characters which does not constitute a complete syllable (for example, a Choseong alone, a Jungseong alone, a Jongseong alone, or a Jungseong followed by a Jongseong). An incomplete syllable which starts with a Jungseong or a Jongseong must be preceded by a CHOSEONG FILLER (0000 115F). An incomplete syllable composed of a Choseong alone must be followed by a JUNGSEONG FILLER (0000 1160).

Hangul Jamo are conjoining characters since they do not require non-combining characters for the syllable composition method. The implementation level 3 shall be used for the Hangul syllable composition method.

25 Code tables and lists of character names

An overview of the Basic Multilingual Plane is shown in figure 3. Detailed code tables and lists of character names for the Basic Multilingual Plane are shown on the following pages.

Guidelines to be used for constructing names of characters are given in annex K for information. In some cases, a name of a character is followed by additional explanatory statements not part of the name. These statements are in parentheses and not in capital letters except for the initials of the word, where required.

Row-octet

00	Basic Latin		Latin-1 Supplement	
01	Latin Extended-A		Latin Extended-B	
02	Latin Extended-B	IPA Extensions	Spacing Modifier Letters	
03	Combining Diacritical Marks		Basic Greek	Greek Symbols and Coptic
04	Cyrillic			
05	Armenian		Hebrew (Basic and Extended)	
06	Basic Arabic		Arabic Extended	
09	Devanagari		Bengali	
0A	Gurmukhi		Gujarati	
0B	Oriya		Tamil	
0C	Telugu		Kannada	
0D	Malayalam			
0E	Thai		Lao	
10			Georgian	
11	Hangul Jamo			
1E	Latin Extended Additional			
1F	Greek Extended			
20	General Punctuation	Super-/Subscripts	Currency Symbols	Comb. Diacritical Marks for Symbols
21	Letterlike Symbols	Number Forms	Arrows	
22	Mathematical Operators			
23	Miscellaneous Technical			
24	Control Pictures	O.C.R.	Enclosed Alphanumerics	
25	Box Drawing	Block Elements	Geometric Shapes	
26	Miscellaneous Symbols			
27	Dingbats			
30	CJK Symbols And Punctuation	Hiragana	Katakana	
31	Bopomofo	Hangul Compatibility Jamo	CJK Miscellaneous	
32	Enclosed CJK Letters and Months			
33	CJK Compatibility			
34	Hangul			
3D				
3E	Hangul Supplementary-A			
44				
45	Hangul Supplementary-B			
4D				
4E	CJK Unified Ideographs			
9F				
A0				
DF				
E0	Private Use Area			
F8				
F9	CJK Compatibility Ideographs			
FA				
FB	Alphabetic Presentation Forms			
FC	Arabic Presentation Forms-A			
FD				
FE	Comb. Half Marks	CJK Compat. Forms	Small Form Variants	Arabic Presentation Forms-B
FF	Halfwidth and Fullwidth Forms			Specials

 = reserved for future standardisation
 = not graphic characters

Figure 3 - Overview of the Basic Multilingual Plane

TABLE 1 - Row 00: BASIC LATIN

	000	001	002	003	004	005	006	007
0			SP	0	@	P	`	p
1			!	1	A	Q	a	q
2			"	2	B	R	b	r
3			#	3	C	S	c	s
4			\$	4	D	T	d	t
5			%	5	E	U	e	u
6			&	6	F	V	f	v
7			'	7	G	W	g	w
8			(	8	H	X	h	x
9			)	9	I	Y	i	y
A			*	:	J	Z	j	z
B			+	;	K	[	k	{
C			,	<	L	\	l	
D			-	=	M	]	m	}
E			.	>	N	^	n	~
F			/	?	O	_	o	

G = 00
P = 00

TABLE 1 - Row 00: BASIC LATIN

dec	hex	Name	dec	hex	Name
032	20	SPACE	123	7B	LEFT CURLY BRACKET
033	21	EXCLAMATION MARK	124	7C	VERTICAL LINE
034	22	QUOTATION MARK	125	7D	RIGHT CURLY BRACKET
035	23	NUMBER SIGN	126	7E	TILDE
036	24	DOLLAR SIGN			
037	25	PERCENT SIGN			
038	26	AMPERSAND			
039	27	APOSTROPHE			
040	28	LEFT PARENTHESIS			
041	29	RIGHT PARENTHESIS			
042	2A	ASTERISK			
043	2B	PLUS SIGN			
044	2C	COMMA			
045	2D	HYPHEN-MINUS			
046	2E	FULL STOP			
047	2F	SOLIDUS			
048	30	DIGIT ZERO			
049	31	DIGIT ONE			
050	32	DIGIT TWO			
051	33	DIGIT THREE			
052	34	DIGIT FOUR			
053	35	DIGIT FIVE			
054	36	DIGIT SIX			
055	37	DIGIT SEVEN			
056	38	DIGIT EIGHT			
057	39	DIGIT NINE			
058	3A	COLON			
059	3B	SEMICOLON			
060	3C	LESS-THAN SIGN			
061	3D	EQUALS SIGN			
062	3E	GREATER-THAN SIGN			
063	3F	QUESTION MARK			
064	40	COMMERCIAL AT			
065	41	LATIN CAPITAL LETTER A			
066	42	LATIN CAPITAL LETTER B			
067	43	LATIN CAPITAL LETTER C			
068	44	LATIN CAPITAL LETTER D			
069	45	LATIN CAPITAL LETTER E			
070	46	LATIN CAPITAL LETTER F			
071	47	LATIN CAPITAL LETTER G			
072	48	LATIN CAPITAL LETTER H			
073	49	LATIN CAPITAL LETTER I			
074	4A	LATIN CAPITAL LETTER J			
075	4B	LATIN CAPITAL LETTER K			
076	4C	LATIN CAPITAL LETTER L			
077	4D	LATIN CAPITAL LETTER M			
078	4E	LATIN CAPITAL LETTER N			
079	4F	LATIN CAPITAL LETTER O			
080	50	LATIN CAPITAL LETTER P			
081	51	LATIN CAPITAL LETTER Q			
082	52	LATIN CAPITAL LETTER R			
083	53	LATIN CAPITAL LETTER S			
084	54	LATIN CAPITAL LETTER T			
085	55	LATIN CAPITAL LETTER U			
086	56	LATIN CAPITAL LETTER V			
087	57	LATIN CAPITAL LETTER W			
088	58	LATIN CAPITAL LETTER X			
089	59	LATIN CAPITAL LETTER Y			
090	5A	LATIN CAPITAL LETTER Z			
091	5B	LEFT SQUARE BRACKET			
092	5C	REVERSE SOLIDUS			
093	5D	RIGHT SQUARE BRACKET			
094	5E	CIRCUMFLEX ACCENT			
095	5F	LOW LINE			
096	60	GRAVE ACCENT			
097	61	LATIN SMALL LETTER A			
098	62	LATIN SMALL LETTER B			
099	63	LATIN SMALL LETTER C			
100	64	LATIN SMALL LETTER D			
101	65	LATIN SMALL LETTER E			
102	66	LATIN SMALL LETTER F			
103	67	LATIN SMALL LETTER G			
104	68	LATIN SMALL LETTER H			
105	69	LATIN SMALL LETTER I			
106	6A	LATIN SMALL LETTER J			
107	6B	LATIN SMALL LETTER K			
108	6C	LATIN SMALL LETTER L			
109	6D	LATIN SMALL LETTER M			
110	6E	LATIN SMALL LETTER N			
111	6F	LATIN SMALL LETTER O			
112	70	LATIN SMALL LETTER P			
113	71	LATIN SMALL LETTER Q			
114	72	LATIN SMALL LETTER R			
115	73	LATIN SMALL LETTER S			
116	74	LATIN SMALL LETTER T			
117	75	LATIN SMALL LETTER U			
118	76	LATIN SMALL LETTER V			
119	77	LATIN SMALL LETTER W			
120	78	LATIN SMALL LETTER X			
121	79	LATIN SMALL LETTER Y			
122	7A	LATIN SMALL LETTER Z			

TABLE 2 - Row 00: LATIN-1 SUPPLEMENT

	008	009	00A	00B	00C	00D	00E	00F
0				°	À	Đ	à	đ
	128	144	160	176	192	208	224	240
1			¡	±	Á	Ñ	á	ñ
	129	145	161	177	193	209	225	241
2			¢	²	Â	Ò	â	ò
	130	146	162	178	194	210	226	242
3			£	³	Ã	Ó	ã	ó
	131	147	163	179	195	211	227	243
4			¤	´	Ä	Ô	ä	ô
	132	148	164	180	196	212	228	244
5			¥	µ	Å	Õ	å	õ
	133	149	165	181	197	213	229	245
6			¦	¶	Æ	Ö	æ	ö
	134	150	166	182	198	214	230	246
7			§	·	Ç	×	ç	÷
	135	151	167	183	199	215	231	247
8			¨	¸	È	Ø	è	ø
	136	152	168	184	200	216	232	248
9			©	¹	É	Ù	é	ù
	137	153	169	185	201	217	233	249
A			ª	º	Ê	Ú	ê	ú
	138	154	170	186	202	218	234	250
B			«	»	Ë	Û	ë	û
	139	155	171	187	203	219	235	251
C			¬	¼	Ì	Ü	ì	ü
	140	156	172	188	204	220	236	252
D				½	Í	Ý	í	ý
	141	157	173	189	205	221	237	253
E			®	¾	Î	Þ	î	þ
	142	158	174	190	206	222	238	254
F			-	¿	Ï	ß	ï	ÿ
	143	159	175	191	207	223	239	255

G = 00

P = 00

TABLE 2 – Row 00: LATIN-1 SUPPLEMENT

dec	hex	Name	dec	hex	Name
160	A0	NO-BREAK SPACE	251	FB	LATIN SMALL LETTER U WITH CIRCUMFLEX
161	A1	INVERTED EXCLAMATION MARK	252	FC	LATIN SMALL LETTER U WITH DIAERESIS
162	A2	CENT SIGN	253	FD	LATIN SMALL LETTER Y WITH ACUTE
163	A3	POUND SIGN	254	FE	LATIN SMALL LETTER THORN (Icelandic)
164	A4	CURRENCY SIGN	255	FF	LATIN SMALL LETTER Y WITH DIAERESIS
165	A5	YEN SIGN			
166	A6	BROKEN BAR			
167	A7	SECTION SIGN			
168	A8	DIAERESIS			
169	A9	COPYRIGHT SIGN			
170	AA	FEMININE ORDINAL INDICATOR			
171	AB	LEFT-POINTING DOUBLE ANGLE QUOTATION MARK			
172	AC	NOT SIGN			
173	AD	SOFT HYPHEN			
174	AE	REGISTERED SIGN			
175	AF	MACRON			
176	B0	DEGREE SIGN			
177	B1	PLUS-MINUS SIGN			
178	B2	SUPERSCRIP TWO			
179	B3	SUPERSCRIP THREE			
180	B4	ACUTE ACCENT			
181	B5	MICRO SIGN			
182	B6	PILCROW SIGN			
183	B7	MIDDLE DOT			
184	B8	CEDILLA			
185	B9	SUPERSCRIP ONE			
186	BA	MASCULINE ORDINAL INDICATOR			
187	BB	RIGHT-POINTING DOUBLE ANGLE QUOTATION MARK			
188	BC	VULGAR FRACTION ONE QUARTER			
189	BD	VULGAR FRACTION ONE HALF			
190	BE	VULGAR FRACTION THREE QUARTERS			
191	BF	INVERTED QUESTION MARK			
192	C0	LATIN CAPITAL LETTER A WITH GRAVE			
193	C1	LATIN CAPITAL LETTER A WITH ACUTE			
194	C2	LATIN CAPITAL LETTER A WITH CIRCUMFLEX			
195	C3	LATIN CAPITAL LETTER A WITH TILDE			
196	C4	LATIN CAPITAL LETTER A WITH DIAERESIS			
197	C5	LATIN CAPITAL LETTER A WITH RING ABOVE			
198	C6	LATIN CAPITAL LIGATURE AE			
199	C7	LATIN CAPITAL LETTER C WITH CEDILLA			
200	C8	LATIN CAPITAL LETTER E WITH GRAVE			
201	C9	LATIN CAPITAL LETTER E WITH ACUTE			
202	CA	LATIN CAPITAL LETTER E WITH CIRCUMFLEX			
203	CB	LATIN CAPITAL LETTER E WITH DIAERESIS			
204	CC	LATIN CAPITAL LETTER I WITH GRAVE			
205	CD	LATIN CAPITAL LETTER I WITH ACUTE			
206	CE	LATIN CAPITAL LETTER I WITH CIRCUMFLEX			
207	CF	LATIN CAPITAL LETTER I WITH DIAERESIS			
208	D0	LATIN CAPITAL LETTER ETH (Icelandic)			
209	D1	LATIN CAPITAL LETTER N WITH TILDE			
210	D2	LATIN CAPITAL LETTER O WITH GRAVE			
211	D3	LATIN CAPITAL LETTER O WITH ACUTE			
212	D4	LATIN CAPITAL LETTER O WITH CIRCUMFLEX			
213	D5	LATIN CAPITAL LETTER O WITH TILDE			
214	D6	LATIN CAPITAL LETTER O WITH DIAERESIS			
215	D7	MULTIPLICATION SIGN			
216	D8	LATIN CAPITAL LETTER O WITH STROKE			
217	D9	LATIN CAPITAL LETTER U WITH GRAVE			
218	DA	LATIN CAPITAL LETTER U WITH ACUTE			
219	DB	LATIN CAPITAL LETTER U WITH CIRCUMFLEX			
220	DC	LATIN CAPITAL LETTER U WITH DIAERESIS			
221	DD	LATIN CAPITAL LETTER Y WITH ACUTE			
222	DE	LATIN CAPITAL LETTER THORN (Icelandic)			
223	DF	LATIN SMALL LETTER SHARP S (German)			
224	E0	LATIN SMALL LETTER A WITH GRAVE			
225	E1	LATIN SMALL LETTER A WITH ACUTE			
226	E2	LATIN SMALL LETTER A WITH CIRCUMFLEX			
227	E3	LATIN SMALL LETTER A WITH TILDE			
228	E4	LATIN SMALL LETTER A WITH DIAERESIS			
229	E5	LATIN SMALL LETTER A WITH RING ABOVE			
230	E6	LATIN SMALL LIGATURE AE			
231	E7	LATIN SMALL LETTER C WITH CEDILLA			
232	E8	LATIN SMALL LETTER E WITH GRAVE			
233	E9	LATIN SMALL LETTER E WITH ACUTE			
234	EA	LATIN SMALL LETTER E WITH CIRCUMFLEX			
235	EB	LATIN SMALL LETTER E WITH DIAERESIS			
236	EC	LATIN SMALL LETTER I WITH GRAVE			
237	ED	LATIN SMALL LETTER I WITH ACUTE			
238	EE	LATIN SMALL LETTER I WITH CIRCUMFLEX			
239	EF	LATIN SMALL LETTER I WITH DIAERESIS			
240	F0	LATIN SMALL LETTER ETH (Icelandic)			
241	F1	LATIN SMALL LETTER N WITH TILDE			
242	F2	LATIN SMALL LETTER O WITH GRAVE			
243	F3	LATIN SMALL LETTER O WITH ACUTE			
244	F4	LATIN SMALL LETTER O WITH CIRCUMFLEX			
245	F5	LATIN SMALL LETTER O WITH TILDE			
246	F6	LATIN SMALL LETTER O WITH DIAERESIS			
247	F7	DIVISION SIGN			
248	F8	LATIN SMALL LETTER O WITH STROKE			
249	F9	LATIN SMALL LETTER U WITH GRAVE			
250	FA	LATIN SMALL LETTER U WITH ACUTE			

TABLE 3 - Row 01: LATIN EXTENDED-A

	010	011	012	013	014	015	016	017
0	Ā	Ð	Ġ	İ	ı	Ō	Š	Ů
	000	016	032	048	064	080	096	112
1	ā	đ	ġ	ı	Ł	ó	š	ů
	001	017	033	049	065	081	097	113
2	Ǻ	Ē	Ģ	IJ	ł	Œ	Ț	Ț
	002	018	034	050	066	082	098	114
3	ǻ	ē	ġ	ij	Ń	œ	ț	ț
	003	019	035	051	067	083	099	115
4	Ą	Ĕ	Ĥ	Ĵ	ń	Ŕ	Ť	Ŵ
	004	020	036	052	068	084	100	116
5	ą	ĕ	ĥ	ĵ	Ņ	ŕ	ť	ŵ
	005	021	037	053	069	085	101	117
6	Ć	Ė	Ħ	Ķ	ņ	Ŗ	Ŧ	Ŷ
	006	022	038	054	070	086	102	118
7	ć	ė	ħ	ķ	Ň	ŗ	ţ	ŷ
	007	023	039	055	071	087	103	119
8	Ĉ	Ę	Ĩ	κ	ň	Ř	Ů	Ỳ
	008	024	040	056	072	088	104	120
9	ĉ	ę	ĩ	Ł	'n	ř	ů	Ž
	009	025	041	057	073	089	105	121
A	Ċ	Ė	Ī	Í	Ŋ	Ś	Ū	ž
	010	026	042	058	074	090	106	122
B	ċ	ė	ī	Ľ	ŋ	ś	ū	Ž
	011	027	043	059	075	091	107	123
C	Č	Ĝ	Ĭ	Ľ	Ō	Ŝ	Ů	ž
	012	028	044	060	076	092	108	124
D	č	ĝ	ĭ	Ľ	ō	ŝ	ů	Ž
	013	029	045	061	077	093	109	125
E	Ď	Ĝ	Ĭ	Ĭ	Ŏ	Ş	Ű	ž
	014	030	046	062	078	094	110	126
F	ď	ğ	ị	Ľ	ő	ş	ű	f
	015	031	047	063	079	095	111	127

G = 00

P = 00

TABLE 3 – Row 01: LATIN EXTENDED-A

dec	hex	Name	dec	hex	Name
000	00	LATIN CAPITAL LETTER A WITH MACRON	091	5B	LATIN SMALL LETTER S WITH ACUTE
001	01	LATIN SMALL LETTER A WITH MACRON	092	5C	LATIN CAPITAL LETTER S WITH CIRCUMFLEX
002	02	LATIN CAPITAL LETTER A WITH BREVE	093	5D	LATIN SMALL LETTER S WITH CIRCUMFLEX
003	03	LATIN SMALL LETTER A WITH BREVE	094	5E	LATIN CAPITAL LETTER S WITH CEDILLA
004	04	LATIN CAPITAL LETTER A WITH OGONEK	095	5F	LATIN SMALL LETTER S WITH CEDILLA
005	05	LATIN SMALL LETTER A WITH OGONEK	096	60	LATIN CAPITAL LETTER S WITH CARON
006	06	LATIN CAPITAL LETTER C WITH ACUTE	097	61	LATIN SMALL LETTER S WITH CARON
007	07	LATIN SMALL LETTER C WITH ACUTE	098	62	LATIN CAPITAL LETTER T WITH CEDILLA
008	08	LATIN CAPITAL LETTER C WITH CIRCUMFLEX	099	63	LATIN SMALL LETTER T WITH CEDILLA
009	09	LATIN SMALL LETTER C WITH CIRCUMFLEX	100	64	LATIN CAPITAL LETTER T WITH CARON
010	0A	LATIN CAPITAL LETTER C WITH DOT ABOVE	101	65	LATIN SMALL LETTER T WITH CARON
011	0B	LATIN SMALL LETTER C WITH DOT ABOVE	102	66	LATIN CAPITAL LETTER T WITH STROKE
012	0C	LATIN CAPITAL LETTER C WITH CARON	103	67	LATIN SMALL LETTER T WITH STROKE
013	0D	LATIN SMALL LETTER C WITH CARON	104	68	LATIN CAPITAL LETTER U WITH TILDE
014	0E	LATIN CAPITAL LETTER D WITH CARON	105	69	LATIN SMALL LETTER U WITH TILDE
015	0F	LATIN SMALL LETTER D WITH CARON	106	6A	LATIN CAPITAL LETTER U WITH MACRON
016	10	LATIN CAPITAL LETTER D WITH STROKE	107	6B	LATIN SMALL LETTER U WITH MACRON
017	11	LATIN SMALL LETTER D WITH STROKE	108	6C	LATIN CAPITAL LETTER U WITH BREVE
018	12	LATIN CAPITAL LETTER E WITH MACRON	109	6D	LATIN SMALL LETTER U WITH BREVE
019	13	LATIN SMALL LETTER E WITH MACRON	110	6E	LATIN CAPITAL LETTER U WITH RING ABOVE
020	14	LATIN CAPITAL LETTER E WITH BREVE	111	6F	LATIN SMALL LETTER U WITH RING ABOVE
021	15	LATIN SMALL LETTER E WITH BREVE	112	70	LATIN CAPITAL LETTER U WITH DOUBLE ACUTE
022	16	LATIN CAPITAL LETTER E WITH DOT ABOVE	113	71	LATIN SMALL LETTER U WITH DOUBLE ACUTE
023	17	LATIN SMALL LETTER E WITH DOT ABOVE	114	72	LATIN CAPITAL LETTER U WITH OGONEK
024	18	LATIN CAPITAL LETTER E WITH OGONEK	115	73	LATIN SMALL LETTER U WITH OGONEK
025	19	LATIN SMALL LETTER E WITH OGONEK	116	74	LATIN CAPITAL LETTER W WITH CIRCUMFLEX
026	1A	LATIN CAPITAL LETTER E WITH CARON	117	75	LATIN SMALL LETTER W WITH CIRCUMFLEX
027	1B	LATIN SMALL LETTER E WITH CARON	118	76	LATIN CAPITAL LETTER Y WITH CIRCUMFLEX
028	1C	LATIN CAPITAL LETTER G WITH CIRCUMFLEX	119	77	LATIN SMALL LETTER Y WITH CIRCUMFLEX
029	1D	LATIN SMALL LETTER G WITH CIRCUMFLEX	120	78	LATIN CAPITAL LETTER Y WITH DIAERESIS
030	1E	LATIN CAPITAL LETTER G WITH BREVE	121	79	LATIN CAPITAL LETTER Z WITH ACUTE
031	1F	LATIN SMALL LETTER G WITH BREVE	122	7A	LATIN SMALL LETTER Z WITH ACUTE
032	20	LATIN CAPITAL LETTER G WITH DOT ABOVE	123	7B	LATIN CAPITAL LETTER Z WITH DOT ABOVE
033	21	LATIN SMALL LETTER G WITH DOT ABOVE	124	7C	LATIN SMALL LETTER Z WITH DOT ABOVE
034	22	LATIN CAPITAL LETTER G WITH CEDILLA	125	7D	LATIN CAPITAL LETTER Z WITH CARON
035	23	LATIN SMALL LETTER G WITH CEDILLA	126	7E	LATIN SMALL LETTER Z WITH CARON
036	24	LATIN CAPITAL LETTER H WITH CIRCUMFLEX	127	7F	LATIN SMALL LETTER LONG S
037	25	LATIN SMALL LETTER H WITH CIRCUMFLEX			
038	26	LATIN CAPITAL LETTER H WITH STROKE			
039	27	LATIN SMALL LETTER H WITH STROKE			
040	28	LATIN CAPITAL LETTER I WITH TILDE			
041	29	LATIN SMALL LETTER I WITH TILDE			
042	2A	LATIN CAPITAL LETTER I WITH MACRON			
043	2B	LATIN SMALL LETTER I WITH MACRON			
044	2C	LATIN CAPITAL LETTER I WITH BREVE			
045	2D	LATIN SMALL LETTER I WITH BREVE			
046	2E	LATIN CAPITAL LETTER I WITH OGONEK			
047	2F	LATIN SMALL LETTER I WITH OGONEK			
048	30	LATIN CAPITAL LETTER I WITH DOT ABOVE			
049	31	LATIN SMALL LETTER DOTLESS I			
050	32	LATIN CAPITAL LIGATURE IJ			
051	33	LATIN SMALL LIGATURE IJ			
052	34	LATIN CAPITAL LETTER J WITH CIRCUMFLEX			
053	35	LATIN SMALL LETTER J WITH CIRCUMFLEX			
054	36	LATIN CAPITAL LETTER K WITH CEDILLA			
055	37	LATIN SMALL LETTER K WITH CEDILLA			
056	38	LATIN SMALL LETTER KRA (Greenlandic)			
057	39	LATIN CAPITAL LETTER L WITH ACUTE			
058	3A	LATIN SMALL LETTER L WITH ACUTE			
059	3B	LATIN CAPITAL LETTER L WITH CEDILLA			
060	3C	LATIN SMALL LETTER L WITH CEDILLA			
061	3D	LATIN CAPITAL LETTER L WITH CARON			
062	3E	LATIN SMALL LETTER L WITH CARON			
063	3F	LATIN CAPITAL LETTER L WITH MIDDLE DOT			
064	40	LATIN SMALL LETTER L WITH MIDDLE DOT			
065	41	LATIN CAPITAL LETTER L WITH STROKE			
066	42	LATIN SMALL LETTER L WITH STROKE			
067	43	LATIN CAPITAL LETTER N WITH ACUTE			
068	44	LATIN SMALL LETTER N WITH ACUTE			
069	45	LATIN CAPITAL LETTER N WITH CEDILLA			
070	46	LATIN SMALL LETTER N WITH CEDILLA			
071	47	LATIN CAPITAL LETTER N WITH CARON			
072	48	LATIN SMALL LETTER N WITH CARON			
073	49	LATIN SMALL LETTER N PRECEDED BY APOSTROPHE			
074	4A	LATIN CAPITAL LETTER ENG (Sami)			
075	4B	LATIN SMALL LETTER ENG (Sami)			
076	4C	LATIN CAPITAL LETTER O WITH MACRON			
077	4D	LATIN SMALL LETTER O WITH MACRON			
078	4E	LATIN CAPITAL LETTER O WITH BREVE			
079	4F	LATIN SMALL LETTER O WITH BREVE			
080	50	LATIN CAPITAL LETTER O WITH DOUBLE ACUTE			
081	51	LATIN SMALL LETTER O WITH DOUBLE ACUTE			
082	52	LATIN CAPITAL LIGATURE OE			
083	53	LATIN SMALL LIGATURE OE			
084	54	LATIN CAPITAL LETTER R WITH ACUTE			
085	55	LATIN SMALL LETTER R WITH ACUTE			
086	56	LATIN CAPITAL LETTER R WITH CEDILLA			
087	57	LATIN SMALL LETTER R WITH CEDILLA			
088	58	LATIN CAPITAL LETTER R WITH CARON			
089	59	LATIN SMALL LETTER R WITH CARON			
090	5A	LATIN CAPITAL LETTER S WITH ACUTE			

TABLE 4 – Row 01: LATIN EXTENDED-B

	018	019	01A	01B	01C	01D	01E	01F
0	ƒ 128	ε 144	Ɔ 160	ƣ 176	 192	ı 208	Ā 224	ĵ 240
1	Ɓ 129	Ƒ 145	σ 161	Ʈ 177	 193	Ǿ 209	ā 225	DZ 241
2	Ɓ 130	ƒ 146	Ɔ 162	Ʈ 178	≠ 194	ǿ 210	Æ 226	Dz 242
3	ƒ 131	Ƒ 147	σ 163	ƣ 179	! 195	Ǿ 211	æ 227	dz 243
4	ƒ 132	Ƒ 148	Ɔ 164	ƣ 180	DŽ 196	ǿ 212	Ƒ 228	Ć 244
5	ƒ 133	h 149	ƒ 165	Z 181	Dž 197	Ū 213	g 229	ǵ 245
6	Ɔ 134	l 150	Ƒ 166	z 182	dž 198	ū 214	Ǧ 230	
7	Ɔ 135	Ƒ 151	2 167	3 183	LJ 199	Ū 215	ǧ 231	
8	Ɔ 136	K 152	2 168	Ɔ 184	Lj 200	ú 216	Ǩ 232	
9	Ɔ 137	k 153	Σ 169	Ɔ 185	lj 201	Ū 217	ǩ 233	
A	Ɔ 138	Ƒ 154	Ƒ 170	Ƒ 186	NJ 202	ǿ 218	Ɔ 234	Ǻ 250
B	Ɔ 139	λ 155	Ƒ 171	2 187	Nj 203	Ū 219	Ɔ 235	ǻ 251
C	Ɔ 140	Ƒ 156	Ƒ 172	Ƒ 188	nj 204	ù 220	Ɔ 236	Æ 252
D	Ɔ 141	Ƒ 157	Ƒ 173	Ƒ 189	Ǻ 205	ə 221	Ɔ 237	æ 253
E	Ɔ 142	η 158	Ƒ 174	Ƒ 190	ǻ 206	Ā 222	Ǻ 238	Ø 254
F	Ɔ 143	θ 159	Ƒ 175	Ƒ 191	Ǻ 207	ā 223	Ǻ 239	ø 255

G = 00
P = 00

TABLE 4 – Row 01: LATIN EXTENDED-B

dec	hex	Name	dec	hex	Name
128	80	LATIN SMALL LETTER B WITH STROKE	219	DB	LATIN CAPITAL LETTER U WITH DIAERESIS AND GRAVE
129	81	LATIN CAPITAL LETTER B WITH HOOK	220	DC	LATIN SMALL LETTER U WITH DIAERESIS AND GRAVE
130	82	LATIN CAPITAL LETTER B WITH TOPBAR	221	DD	LATIN SMALL LETTER TURNED E
131	83	LATIN SMALL LETTER B WITH TOPBAR	222	DE	LATIN CAPITAL LETTER A WITH DIAERESIS AND MACRON
132	84	LATIN CAPITAL LETTER TONE SIX	223	DF	LATIN SMALL LETTER A WITH DIAERESIS AND MACRON
133	85	LATIN SMALL LETTER TONE SIX	224	E0	LATIN CAPITAL LETTER A WITH DOT ABOVE AND MACRON
134	86	LATIN CAPITAL LETTER OPEN O	225	E1	LATIN SMALL LETTER A WITH DOT ABOVE AND MACRON
135	87	LATIN CAPITAL LETTER C WITH HOOK	226	E2	LATIN CAPITAL LIGATURE AE WITH MACRON
136	88	LATIN SMALL LETTER C WITH HOOK	227	E3	LATIN SMALL LIGATURE AE WITH MACRON
137	89	LATIN CAPITAL LETTER AFRICAN D	228	E4	LATIN CAPITAL LETTER G WITH STROKE
138	8A	LATIN CAPITAL LETTER D WITH HOOK	229	E5	LATIN SMALL LETTER G WITH STROKE
139	8B	LATIN CAPITAL LETTER D WITH TOPBAR	230	E6	LATIN CAPITAL LETTER G WITH CARON
140	8C	LATIN SMALL LETTER D WITH TOPBAR	231	E7	LATIN SMALL LETTER G WITH CARON
141	8D	LATIN SMALL LETTER TURNED DELTA	232	E8	LATIN CAPITAL LETTER K WITH CARON
142	8E	LATIN CAPITAL LETTER REVERSED E	233	E9	LATIN SMALL LETTER K WITH CARON
143	8F	LATIN CAPITAL LETTER SCHWA	234	EA	LATIN CAPITAL LETTER O WITH OGONEK
144	90	LATIN CAPITAL LETTER OPEN E	235	EB	LATIN SMALL LETTER O WITH OGONEK
145	91	LATIN CAPITAL LETTER F WITH HOOK	236	EC	LATIN CAPITAL LETTER O WITH OGONEK AND MACRON
146	92	LATIN SMALL LETTER F WITH HOOK	237	ED	LATIN SMALL LETTER O WITH OGONEK AND MACRON
147	93	LATIN CAPITAL LETTER G WITH HOOK	238	EE	LATIN CAPITAL LETTER EZH WITH CARON
148	94	LATIN CAPITAL LETTER GAMMA	239	EF	LATIN SMALL LETTER EZH WITH CARON
149	95	LATIN SMALL LETTER HV	240	F0	LATIN SMALL LETTER J WITH CARON
150	96	LATIN CAPITAL LETTER IOTA	241	F1	LATIN CAPITAL LETTER DZ
151	97	LATIN CAPITAL LETTER I WITH STROKE	242	F2	LATIN CAPITAL LETTER D WITH SMALL LETTER Z
152	98	LATIN CAPITAL LETTER K WITH HOOK	243	F3	LATIN SMALL LETTER DZ
153	99	LATIN SMALL LETTER K WITH HOOK	244	F4	LATIN CAPITAL LETTER G WITH ACUTE
154	9A	LATIN SMALL LETTER L WITH BAR	245	F5	LATIN SMALL LETTER G WITH ACUTE
155	9B	LATIN SMALL LETTER LAMBDA WITH STROKE	246	F6	(This position shall not be used)
156	9C	LATIN CAPITAL LETTER TURNED M	247	F7	(This position shall not be used)
157	9D	LATIN CAPITAL LETTER N WITH LEFT HOOK	248	F8	(This position shall not be used)
158	9E	LATIN SMALL LETTER N WITH LONG RIGHT TILDE	249	F9	(This position shall not be used)
159	9F	LATIN CAPITAL LETTER O WITH MIDDLE TILDE	250	FA	LATIN CAPITAL LETTER A WITH RING ABOVE AND ACUTE
160	A0	LATIN CAPITAL LETTER O WITH HORN	251	FB	LATIN SMALL LETTER A WITH RING ABOVE AND ACUTE
161	A1	LATIN SMALL LETTER O WITH HORN	252	FC	LATIN CAPITAL LIGATURE AE WITH ACUTE
162	A2	LATIN CAPITAL LETTER OI	253	FD	LATIN SMALL LIGATURE AE WITH ACUTE
163	A3	LATIN SMALL LETTER OI	254	FE	LATIN CAPITAL LETTER O WITH STROKE AND ACUTE
164	A4	LATIN CAPITAL LETTER P WITH HOOK	255	FF	LATIN SMALL LETTER O WITH STROKE AND ACUTE
165	A5	LATIN SMALL LETTER P WITH HOOK			
166	A6	LATIN LETTER YR			
167	A7	LATIN CAPITAL LETTER TONE TWO			
168	A8	LATIN SMALL LETTER TONE TWO			
169	A9	LATIN CAPITAL LETTER ESH			
170	AA	LATIN LETTER REVERSED ESH LOOP			
171	AB	LATIN SMALL LETTER T WITH PALATAL HOOK			
172	AC	LATIN CAPITAL LETTER T WITH HOOK			
173	AD	LATIN SMALL LETTER T WITH HOOK			
174	AE	LATIN CAPITAL LETTER T WITH RETROFLEX HOOK			
175	AF	LATIN CAPITAL LETTER U WITH HORN			
176	B0	LATIN SMALL LETTER U WITH HORN			
177	B1	LATIN CAPITAL LETTER UPSILON			
178	B2	LATIN CAPITAL LETTER V WITH HOOK			
179	B3	LATIN CAPITAL LETTER Y WITH HOOK			
180	B4	LATIN SMALL LETTER Y WITH HOOK			
181	B5	LATIN CAPITAL LETTER Z WITH STROKE			
182	B6	LATIN SMALL LETTER Z WITH STROKE			
183	B7	LATIN CAPITAL LETTER EZH			
184	B8	LATIN CAPITAL LETTER EZH REVERSED			
185	B9	LATIN SMALL LETTER EZH REVERSED			
186	BA	LATIN SMALL LETTER EZH WITH TAIL			
187	BB	LATIN LETTER TWO WITH STROKE			
188	BC	LATIN CAPITAL LETTER TONE FIVE			
189	BD	LATIN SMALL LETTER TONE FIVE			
190	BE	LATIN LETTER INVERTED GLOTTAL STOP WITH STROKE			
191	BF	LATIN LETTER WYNN			
192	C0	LATIN LETTER DENTAL CLICK			
193	C1	LATIN LETTER LATERAL CLICK			
194	C2	LATIN LETTER ALVEOLAR CLICK			
195	C3	LATIN LETTER RETROFLEX CLICK			
196	C4	LATIN CAPITAL LETTER DZ WITH CARON			
197	C5	LATIN CAPITAL LETTER D WITH SMALL LETTER Z WITH CARON			
198	C6	LATIN SMALL LETTER DZ WITH CARON			
199	C7	LATIN CAPITAL LETTER LJ			
200	C8	LATIN CAPITAL LETTER L WITH SMALL LETTER J			
201	C9	LATIN SMALL LETTER LJ			
202	CA	LATIN CAPITAL LETTER NJ			
203	CB	LATIN CAPITAL LETTER N WITH SMALL LETTER J			
204	CC	LATIN SMALL LETTER NJ			
205	CD	LATIN CAPITAL LETTER A WITH CARON			
206	CE	LATIN SMALL LETTER A WITH CARON			
207	CF	LATIN CAPITAL LETTER I WITH CARON			
208	D0	LATIN SMALL LETTER I WITH CARON			
209	D1	LATIN CAPITAL LETTER O WITH CARON			
210	D2	LATIN SMALL LETTER O WITH CARON			
211	D3	LATIN CAPITAL LETTER U WITH CARON			
212	D4	LATIN SMALL LETTER U WITH CARON			
213	D5	LATIN CAPITAL LETTER U WITH DIAERESIS AND MACRON			
214	D6	LATIN SMALL LETTER U WITH DIAERESIS AND MACRON			
215	D7	LATIN CAPITAL LETTER U WITH DIAERESIS AND ACUTE			
216	D8	LATIN SMALL LETTER U WITH DIAERESIS AND ACUTE			
217	D9	LATIN CAPITAL LETTER U WITH DIAERESIS AND CARON			
218	DA	LATIN SMALL LETTER U WITH DIAERESIS AND CARON			

Group 00

Plane 00

Row 01



Documents

Document Sharing Hub

23

TABLE 5 – Row 02: LATIN EXTENDED-B

	020	021	022	023	024
0	Ă	Ŕ			
	000	016	032	048	064
1	ă	ŕ			
	001	017	033	049	065
2	Â	Ŗ			
	002	018	034	050	066
3	â	ŗ			
	003	019	035	051	067
4	Ê	Û			
	004	020	036	052	068
5	ê	û			
	005	021	037	053	069
6	Ê	Û			
	006	022	038	054	070
7	ê	û			
	007	023	039	055	071
8	Ï				
	008	024	040	056	072
9	ï				
	009	025	041	057	073
A	Î				
	010	026	042	058	074
B	î				
	011	027	043	059	075
C	Õ				
	012	028	044	060	076
D	õ				
	013	029	045	061	077
E	Ô				
	014	030	046	062	078
F	ô				
	015	031	047	063	079

G = 00

P = 00

TABLE 5 - Row 02: LATIN EXTENDED-B

dec	hex	Name	dec	hex	Name
000	00	LATIN CAPITAL LETTER A WITH DOUBLE GRAVE			
001	01	LATIN SMALL LETTER A WITH DOUBLE GRAVE			
002	02	LATIN CAPITAL LETTER A WITH INVERTED BREVE			
003	03	LATIN SMALL LETTER A WITH INVERTED BREVE			
004	04	LATIN CAPITAL LETTER E WITH DOUBLE GRAVE			
005	05	LATIN SMALL LETTER E WITH DOUBLE GRAVE			
006	06	LATIN CAPITAL LETTER E WITH INVERTED BREVE			
007	07	LATIN SMALL LETTER E WITH INVERTED BREVE			
008	08	LATIN CAPITAL LETTER I WITH DOUBLE GRAVE			
009	09	LATIN SMALL LETTER I WITH DOUBLE GRAVE			
010	0A	LATIN CAPITAL LETTER I WITH INVERTED BREVE			
011	0B	LATIN SMALL LETTER I WITH INVERTED BREVE			
012	0C	LATIN CAPITAL LETTER O WITH DOUBLE GRAVE			
013	0D	LATIN SMALL LETTER O WITH DOUBLE GRAVE			
014	0E	LATIN CAPITAL LETTER O WITH INVERTED BREVE			
015	0F	LATIN SMALL LETTER O WITH INVERTED BREVE			
016	10	LATIN CAPITAL LETTER R WITH DOUBLE GRAVE			
017	11	LATIN SMALL LETTER R WITH DOUBLE GRAVE			
018	12	LATIN CAPITAL LETTER R WITH INVERTED BREVE			
019	13	LATIN SMALL LETTER R WITH INVERTED BREVE			
020	14	LATIN CAPITAL LETTER U WITH DOUBLE GRAVE			
021	15	LATIN SMALL LETTER U WITH DOUBLE GRAVE			
022	16	LATIN CAPITAL LETTER U WITH INVERTED BREVE			
023	17	LATIN SMALL LETTER U WITH INVERTED BREVE			
024	18	(This position shall not be used)			
025	19	(This position shall not be used)			
026	1A	(This position shall not be used)			
027	1B	(This position shall not be used)			
028	1C	(This position shall not be used)			
029	1D	(This position shall not be used)			
030	1E	(This position shall not be used)			
031	1F	(This position shall not be used)			
032	20	(This position shall not be used)			
033	21	(This position shall not be used)			
034	22	(This position shall not be used)			
035	23	(This position shall not be used)			
036	24	(This position shall not be used)			
037	25	(This position shall not be used)			
038	26	(This position shall not be used)			
039	27	(This position shall not be used)			
040	28	(This position shall not be used)			
041	29	(This position shall not be used)			
042	2A	(This position shall not be used)			
043	2B	(This position shall not be used)			
044	2C	(This position shall not be used)			
045	2D	(This position shall not be used)			
046	2E	(This position shall not be used)			
047	2F	(This position shall not be used)			
048	30	(This position shall not be used)			
049	31	(This position shall not be used)			
050	32	(This position shall not be used)			
051	33	(This position shall not be used)			
052	34	(This position shall not be used)			
053	35	(This position shall not be used)			
054	36	(This position shall not be used)			
055	37	(This position shall not be used)			
056	38	(This position shall not be used)			
057	39	(This position shall not be used)			
058	3A	(This position shall not be used)			
059	3B	(This position shall not be used)			
060	3C	(This position shall not be used)			
061	3D	(This position shall not be used)			
062	3E	(This position shall not be used)			
063	3F	(This position shall not be used)			
064	40	(This position shall not be used)			
065	41	(This position shall not be used)			
066	42	(This position shall not be used)			
067	43	(This position shall not be used)			
068	44	(This position shall not be used)			
069	45	(This position shall not be used)			
070	46	(This position shall not be used)			
071	47	(This position shall not be used)			
072	48	(This position shall not be used)			
073	49	(This position shall not be used)			
074	4A	(This position shall not be used)			
075	4B	(This position shall not be used)			
076	4C	(This position shall not be used)			
077	4D	(This position shall not be used)			
078	4E	(This position shall not be used)			
079	4F	(This position shall not be used)			

Group 00

Plane 00

TABLE 6 - Row 02: IPA EXTENSIONS

	025	026	027	028	029	02A
0	ɐ 080	ɟ 096	ɥ 112	ʀ 128	ʐ 144	ɖ 160
1	ɑ 081	ɡ 097	ɱ 113	ʁ 129	ʑ 145	ʢ 161
2	ɒ 082	ɢ 098	ɲ 114	ʂ 130	ʑ 146	ɕ 162
3	ɓ 083	ɣ 099	ɳ 115	ʃ 131	ʒ 147	ɗ 163
4	ɔ 084	ɹ 100	ɴ 116	ʄ 132	ʐ 148	ɖʑ 164
5	ɛ 085	ɥ 101	ɵ 117	ɹ 133	ɕ 149	ɗ 165
6	ɖ 086	ɦ 102	æ 118	ɭ 134	ɰ 150	ʈ 166
7	ɗ 087	ɧ 103	ω 119	ɰ 135	ɕ 151	ɧ 167
8	ə 088	ɨ 104	Φ 120	ʈ 136	ɕ 152	ʈɕ 168
9	ə 089	ɩ 105	ɰ 121	ɧ 137	ɕ 153	
A	æ 090	ɪ 106	ɰ 122	ʊ 138	ɕ 154	
B	ɛ 091	ɩ 107	ɰ 123	ʊ 139	ɕ 155	
C	ɛ 092	ɩ 108	ɰ 124	ʌ 140	ɧ 156	
D	ɛ 093	ɩ 109	ɰ 125	ʌ 141	ɩ 157	
E	ɛ 094	ɩ 110	ɰ 126	ʌ 142	ɩ 158	
F	ɩ 095	ɩ 111	ɰ 127	ʌ 143	ɩ 159	

G = 00

P = 00

TABLE 6 – Row 02: IPA EXTENSIONS

dec	hex	Name	dec	hex	Name
080	50	LATIN SMALL LETTER TURNED A	171	AB	(This position shall not be used)
081	51	LATIN SMALL LETTER ALPHA	172	AC	(This position shall not be used)
082	52	LATIN SMALL LETTER TURNED ALPHA	173	AD	(This position shall not be used)
083	53	LATIN SMALL LETTER B WITH HOOK	174	AE	(This position shall not be used)
084	54	LATIN SMALL LETTER OPEN O	175	AF	(This position shall not be used)
085	55	LATIN SMALL LETTER C WITH CURL			
086	56	LATIN SMALL LETTER D WITH TAIL			
087	57	LATIN SMALL LETTER D WITH HOOK			
088	58	LATIN SMALL LETTER REVERSED E			
089	59	LATIN SMALL LETTER SCHWA			
090	5A	LATIN SMALL LETTER SCHWA WITH HOOK			
091	5B	LATIN SMALL LETTER OPEN E			
092	5C	LATIN SMALL LETTER REVERSED OPEN E			
093	5D	LATIN SMALL LETTER REVERSED OPEN E WITH HOOK			
094	5E	LATIN SMALL LETTER CLOSED REVERSED OPEN E			
095	5F	LATIN SMALL LETTER DOTLESS J WITH STROKE			
096	60	LATIN SMALL LETTER G WITH HOOK			
097	61	LATIN SMALL LETTER SCRIPT G			
098	62	LATIN LETTER SMALL CAPITAL G			
099	63	LATIN SMALL LETTER GAMMA			
100	64	LATIN SMALL LETTER RAMS HORN			
101	65	LATIN SMALL LETTER TURNED H			
102	66	LATIN SMALL LETTER H WITH HOOK			
103	67	LATIN SMALL LETTER HENG WITH HOOK			
104	68	LATIN SMALL LETTER I WITH STROKE			
105	69	LATIN SMALL LETTER IOTA			
106	6A	LATIN LETTER SMALL CAPITAL I			
107	6B	LATIN SMALL LETTER L WITH MIDDLE TILDE			
108	6C	LATIN SMALL LETTER L WITH BELT			
109	6D	LATIN SMALL LETTER L WITH RETROFLEX HOOK			
110	6E	LATIN SMALL LETTER LEZH			
111	6F	LATIN SMALL LETTER TURNED M			
112	70	LATIN SMALL LETTER TURNED M WITH LONG LEG			
113	71	LATIN SMALL LETTER M WITH HOOK			
114	72	LATIN SMALL LETTER N WITH LEFT HOOK			
115	73	LATIN SMALL LETTER N WITH RETROFLEX HOOK			
116	74	LATIN LETTER SMALL CAPITAL N			
117	75	LATIN SMALL LETTER BARRED O			
118	76	LATIN LETTER SMALL CAPITAL OE			
119	77	LATIN SMALL LETTER CLOSED OMEGA			
120	78	LATIN SMALL LETTER PHI			
121	79	LATIN SMALL LETTER TURNED R			
122	7A	LATIN SMALL LETTER TURNED R WITH LONG LEG			
123	7B	LATIN SMALL LETTER TURNED R WITH HOOK			
124	7C	LATIN SMALL LETTER R WITH LONG LEG			
125	7D	LATIN SMALL LETTER R WITH TAIL			
126	7E	LATIN SMALL LETTER R WITH FISHHOOK			
127	7F	LATIN SMALL LETTER REVERSED R WITH FISHHOOK			
128	80	LATIN LETTER SMALL CAPITAL R			
129	81	LATIN LETTER SMALL CAPITAL INVERTED R			
130	82	LATIN SMALL LETTER S WITH HOOK			
131	83	LATIN SMALL LETTER ESH			
132	84	LATIN SMALL LETTER DOTLESS J WITH STROKE AND HOOK			
133	85	LATIN SMALL LETTER SQUAT REVERSED ESH			
134	86	LATIN SMALL LETTER ESH WITH CURL			
135	87	LATIN SMALL LETTER TURNED T			
136	88	LATIN SMALL LETTER T WITH RETROFLEX HOOK			
137	89	LATIN SMALL LETTER U BAR			
138	8A	LATIN SMALL LETTER UPSILON			
139	8B	LATIN SMALL LETTER V WITH HOOK			
140	8C	LATIN SMALL LETTER TURNED V			
141	8D	LATIN SMALL LETTER TURNED W			
142	8E	LATIN SMALL LETTER TURNED Y			
143	8F	LATIN LETTER SMALL CAPITAL Y			
144	90	LATIN SMALL LETTER Z WITH RETROFLEX HOOK			
145	91	LATIN SMALL LETTER Z WITH CURL			
146	92	LATIN SMALL LETTER EZH			
147	93	LATIN SMALL LETTER EZH WITH CURL			
148	94	LATIN LETTER GLOTTAL STOP			
149	95	LATIN LETTER PHARYNGEAL VOICED FRICATIVE			
150	96	LATIN LETTER INVERTED GLOTTAL STOP			
151	97	LATIN LETTER STRETCHED C			
152	98	LATIN LETTER BILABIAL CLICK			
153	99	LATIN LETTER SMALL CAPITAL B			
154	9A	LATIN SMALL LETTER CLOSED OPEN E			
155	9B	LATIN LETTER SMALL CAPITAL G WITH HOOK			
156	9C	LATIN LETTER SMALL CAPITAL H			
157	9D	LATIN SMALL LETTER J WITH CROSSED-TAIL			
158	9E	LATIN SMALL LETTER TURNED K			
159	9F	LATIN LETTER SMALL CAPITAL L			
160	A0	LATIN SMALL LETTER Q WITH HOOK			
161	A1	LATIN LETTER GLOTTAL STOP WITH STROKE			
162	A2	LATIN LETTER REVERSED GLOTTAL STOP WITH STROKE			
163	A3	LATIN SMALL LETTER DZ DIGRAPH			
164	A4	LATIN SMALL LETTER DEZH DIGRAPH			
165	A5	LATIN SMALL LETTER DZ DIGRAPH WITH CURL			
166	A6	LATIN SMALL LETTER TS DIGRAPH			
167	A7	LATIN SMALL LETTER TESH DIGRAPH			
168	A8	LATIN SMALL LETTER TC DIGRAPH WITH CURL			
169	A9	(This position shall not be used)			
170	AA	(This position shall not be used)			

TABLE 7 - Row 02: SPACING MODIFIER LETTERS

	02B	02C	02D	02E	02F
0	h 176	?	◌ 208	Y 224	
1	h 177	?	◌ 209	l 225	
2	j 178	<	◌ 210	s 226	
3	r 179	>	◌ 211	x 227	
4	I 180	^	◌ 212	ç 228	
5	l 181	v	◌ 213	7 229	
6	B 182	^	+ 214	7 230	
7	w 183	v	◌ 215	7 231	
8	y 184	l	◌ 216	7 232	
9	/ 185	-	◌ 217	7 233	
A	// 186	/	o 218		
B	‘ 187	\	◌ 219		
C	’ 188	l	~ 220		
D	€ 189	-	// 221		
E	◌ 190	\	~ 222		
F	◌ 191	/			

G = 00

P = 00

TABLE 7 - Row 02: SPACING MODIFIER LETTERS

dec	hex	Name	dec	hex	Name
176	B0	MODIFIER LETTER SMALL H			
177	B1	MODIFIER LETTER SMALL H WITH HOOK			
178	B2	MODIFIER LETTER SMALL J			
179	B3	MODIFIER LETTER SMALL R			
180	B4	MODIFIER LETTER SMALL TURNED R			
181	B5	MODIFIER LETTER SMALL TURNED R WITH HOOK			
182	B6	MODIFIER LETTER SMALL CAPITAL INVERTED R			
183	B7	MODIFIER LETTER SMALL W			
184	B8	MODIFIER LETTER SMALL Y			
185	B9	MODIFIER LETTER PRIME			
186	BA	MODIFIER LETTER DOUBLE PRIME			
187	BB	MODIFIER LETTER TURNED COMMA			
188	BC	MODIFIER LETTER APOSTROPHE			
189	BD	MODIFIER LETTER REVERSED COMMA			
190	BE	MODIFIER LETTER RIGHT HALF RING			
191	BF	MODIFIER LETTER LEFT HALF RING			
192	C0	MODIFIER LETTER GLOTTAL STOP			
193	C1	MODIFIER LETTER REVERSED GLOTTAL STOP			
194	C2	MODIFIER LETTER LEFT ARROWHEAD			
195	C3	MODIFIER LETTER RIGHT ARROWHEAD			
196	C4	MODIFIER LETTER UP ARROWHEAD			
197	C5	MODIFIER LETTER DOWN ARROWHEAD			
198	C6	MODIFIER LETTER CIRCUMFLEX ACCENT			
199	C7	CARON (Mandarin Chinese third tone)			
200	C8	MODIFIER LETTER VERTICAL LINE			
201	C9	MODIFIER LETTER MACRON (Mandarin Chinese first tone)			
202	CA	MODIFIER LETTER ACUTE ACCENT (Mandarin Chinese second tone)			
203	CB	MODIFIER LETTER GRAVE ACCENT (Mandarin Chinese fourth tone)			
204	CC	MODIFIER LETTER LOW VERTICAL LINE			
205	CD	MODIFIER LETTER LOW MACRON			
206	CE	MODIFIER LETTER LOW GRAVE ACCENT			
207	CF	MODIFIER LETTER LOW ACUTE ACCENT			
208	D0	MODIFIER LETTER TRIANGULAR COLON			
209	D1	MODIFIER LETTER HALF TRIANGULAR COLON			
210	D2	MODIFIER LETTER CENTRED RIGHT HALF RING			
211	D3	MODIFIER LETTER CENTRED LEFT HALF RING			
212	D4	MODIFIER LETTER UP TACK			
213	D5	MODIFIER LETTER DOWN TACK			
214	D6	MODIFIER LETTER PLUS SIGN			
215	D7	MODIFIER LETTER MINUS SIGN			
216	D8	BREVE			
217	D9	DOT ABOVE (Mandarin Chinese light tone)			
218	DA	RING ABOVE			
219	DB	OGONEK			
220	DC	SMALL TILDE			
221	DD	DOUBLE ACUTE ACCENT			
222	DE	MODIFIER LETTER RHOTIC HOOK			
223	DF	(This position shall not be used)			
224	E0	MODIFIER LETTER SMALL GAMMA			
225	E1	MODIFIER LETTER SMALL L			
226	E2	MODIFIER LETTER SMALL S			
227	E3	MODIFIER LETTER SMALL X			
228	E4	MODIFIER LETTER SMALL REVERSED GLOTTAL STOP			
229	E5	MODIFIER LETTER EXTRA-HIGH TONE BAR			
230	E6	MODIFIER LETTER HIGH TONE BAR			
231	E7	MODIFIER LETTER MID TONE BAR			
232	E8	MODIFIER LETTER LOW TONE BAR			
233	E9	MODIFIER LETTER EXTRA-LOW TONE BAR			
234	EA	(This position shall not be used)			
235	EB	(This position shall not be used)			
236	EC	(This position shall not be used)			
237	ED	(This position shall not be used)			
238	EE	(This position shall not be used)			
239	EF	(This position shall not be used)			
240	F0	(This position shall not be used)			
241	F1	(This position shall not be used)			
242	F2	(This position shall not be used)			
243	F3	(This position shall not be used)			
244	F4	(This position shall not be used)			
245	F5	(This position shall not be used)			
246	F6	(This position shall not be used)			
247	F7	(This position shall not be used)			
248	F8	(This position shall not be used)			
249	F9	(This position shall not be used)			
250	FA	(This position shall not be used)			
251	FB	(This position shall not be used)			
252	FC	(This position shall not be used)			
253	FD	(This position shall not be used)			
254	FE	(This position shall not be used)			
255	FF	(This position shall not be used)			

Group 00

Plane 00

Row 02



Documents

Document Sharing 29

TABLE 8 - Row 03: COMBINING DIACRITICAL MARKS

	030	031	032	033	034	035	036
0							
1							
2							
3							
4							
5							
6							
7							
8							
9							
A							
B							
C							
D							
E							
F							

G = 00

P = 00

TABLE 8 – Row 03: COMBINING DIACRITICAL MARKS

dec	hex	Name	dec	hex	Name
000	00	COMBINING GRAVE ACCENT (Varia)	091	5B	(This position shall not be used)
001	01	COMBINING ACUTE ACCENT (Oxia)	092	5C	(This position shall not be used)
002	02	COMBINING CIRCUMFLEX ACCENT	093	5D	(This position shall not be used)
003	03	COMBINING TILDE	094	5E	(This position shall not be used)
004	04	COMBINING MACRON	095	5F	(This position shall not be used)
005	05	COMBINING OVERLINE	096	60	COMBINING DOUBLE TILDE
006	06	COMBINING BREVE (Vrachy)	097	61	COMBINING DOUBLE INVERTED BREVE
007	07	COMBINING DOT ABOVE	098	62	(This position shall not be used)
008	08	COMBINING DIAERESIS (Dialytika)	099	63	(This position shall not be used)
009	09	COMBINING HOOK ABOVE	100	64	(This position shall not be used)
010	0A	COMBINING RING ABOVE	101	65	(This position shall not be used)
011	0B	COMBINING DOUBLE ACUTE ACCENT	102	66	(This position shall not be used)
012	0C	COMBINING CARON	103	67	(This position shall not be used)
013	0D	COMBINING VERTICAL LINE ABOVE (Tonos)	104	68	(This position shall not be used)
014	0E	COMBINING DOUBLE VERTICAL LINE ABOVE	105	69	(This position shall not be used)
015	0F	COMBINING DOUBLE GRAVE ACCENT	106	6A	(This position shall not be used)
016	10	COMBINING CANDRABINDU	107	6B	(This position shall not be used)
017	11	COMBINING INVERTED BREVE	108	6C	(This position shall not be used)
018	12	COMBINING TURNED COMMA ABOVE	109	6D	(This position shall not be used)
019	13	COMBINING COMMA ABOVE (Psili)	110	6E	(This position shall not be used)
020	14	COMBINING REVERSED COMMA ABOVE (Dasia)	111	6F	(This position shall not be used)
021	15	COMBINING COMMA ABOVE RIGHT			
022	16	COMBINING GRAVE ACCENT BELOW			
023	17	COMBINING ACUTE ACCENT BELOW			
024	18	COMBINING LEFT TACK BELOW			
025	19	COMBINING RIGHT TACK BELOW			
026	1A	COMBINING LEFT ANGLE ABOVE			
027	1B	COMBINING HORN			
028	1C	COMBINING LEFT HALF RING BELOW			
029	1D	COMBINING UP TACK BELOW			
030	1E	COMBINING DOWN TACK BELOW			
031	1F	COMBINING PLUS SIGN BELOW			
032	20	COMBINING MINUS SIGN BELOW			
033	21	COMBINING PALATALIZED HOOK BELOW			
034	22	COMBINING RETROFLEX HOOK BELOW			
035	23	COMBINING DOT BELOW			
036	24	COMBINING DIAERESIS BELOW			
037	25	COMBINING RING BELOW			
038	26	COMBINING COMMA BELOW			
039	27	COMBINING CEDILLA			
040	28	COMBINING OGONEK			
041	29	COMBINING VERTICAL LINE BELOW			
042	2A	COMBINING BRIDGE BELOW			
043	2B	COMBINING INVERTED DOUBLE ARCH BELOW			
044	2C	COMBINING CARON BELOW			
045	2D	COMBINING CIRCUMFLEX ACCENT BELOW			
046	2E	COMBINING BREVE BELOW			
047	2F	COMBINING INVERTED BREVE BELOW			
048	30	COMBINING TILDE BELOW			
049	31	COMBINING MACRON BELOW			
050	32	COMBINING LOW LINE			
051	33	COMBINING DOUBLE LOW LINE			
052	34	COMBINING TILDE OVERLAY			
053	35	COMBINING SHORT STROKE OVERLAY			
054	36	COMBINING LONG STROKE OVERLAY			
055	37	COMBINING SHORT SOLIDUS OVERLAY			
056	38	COMBINING LONG SOLIDUS OVERLAY			
057	39	COMBINING RIGHT HALF RING BELOW			
058	3A	COMBINING INVERTED BRIDGE BELOW			
059	3B	COMBINING SQUARE BELOW			
060	3C	COMBINING SEAGULL BELOW			
061	3D	COMBINING X ABOVE			
062	3E	COMBINING VERTICAL TILDE			
063	3F	COMBINING DOUBLE OVERLINE			
064	40	COMBINING GRAVE TONE MARK (Vietnamese)			
065	41	COMBINING ACUTE TONE MARK (Vietnamese)			
066	42	COMBINING GREEK PERISPOMENI			
067	43	COMBINING GREEK KORONIS			
068	44	COMBINING GREEK DIALYTIKA TONOS			
069	45	COMBINING GREEK YPOGEGRAMMENI			
070	46	(This position shall not be used)			
071	47	(This position shall not be used)			
072	48	(This position shall not be used)			
073	49	(This position shall not be used)			
074	4A	(This position shall not be used)			
075	4B	(This position shall not be used)			
076	4C	(This position shall not be used)			
077	4D	(This position shall not be used)			
078	4E	(This position shall not be used)			
079	4F	(This position shall not be used)			
080	50	(This position shall not be used)			
081	51	(This position shall not be used)			
082	52	(This position shall not be used)			
083	53	(This position shall not be used)			
084	54	(This position shall not be used)			
085	55	(This position shall not be used)			
086	56	(This position shall not be used)			
087	57	(This position shall not be used)			
088	58	(This position shall not be used)			
089	59	(This position shall not be used)			
090	5A	(This position shall not be used)			

Group 00

Plane 00

Row 03



Documents

Document Sharing Hub

TABLE 9 - Row 03: BASIC GREEK

	037	038	039	03A	03B	03C
0			ĩ	Π	ü	π
	118	128	144	160	176	192
1			A	P	α	ρ
	113	139	145	161	177	193
2			B		β	ς
	114	130	146	182	178	194
3			Γ	Σ	γ	σ
	115	131	147	163	179	195
4	'	'	Δ	T	δ	τ
	116	132	148	164	180	196
5	,	•	E	Υ	ε	υ
	117	133	149	165	181	197
6		'A	Z	Φ	ζ	φ
	118	134	150	166	182	198
7		•	H	X	η	χ
	119	135	151	167	183	199
8		'E	Θ	Ψ	θ	ψ
	120	136	152	168	184	200
9		'H	I	Ω	ι	ω
	121	137	153	169	185	201
A		'I	K	İ	κ	ĩ
	122	138	154	170	186	202
B			Λ	ÿ	λ	ü
	123	139	155	171	187	203
C		'O	M	á	μ	ó
	124	140	156	172	188	204
D			N	é	ν	ù
	125	141	157	173	189	205
E	;	'Υ	Ξ	ή	ξ	ώ
	126	142	158	174	190	206
F		'Ω	O	ì	ο	
	127	143	159	175	191	207

G = 00

P = 00

TABLE 9 – Row 03: BASIC GREEK

dec	hex	Name	dec	hex	Name
112	70	(This position shall not be used)	203	CB	GREEK SMALL LETTER UPSILON WITH DIALYTIKA
113	71	(This position shall not be used)	204	CC	GREEK SMALL LETTER OMICRON WITH TONOS
114	72	(This position shall not be used)	205	CD	GREEK SMALL LETTER UPSILON WITH TONOS
115	73	(This position shall not be used)	206	CE	GREEK SMALL LETTER OMEGA WITH TONOS
116	74	GREEK NUMERAL SIGN (Dexia keraia)	207	CF	(This position shall not be used)
117	75	GREEK LOWER NUMERAL SIGN (Aristeri keraia)			
118	76	(This position shall not be used)			
119	77	(This position shall not be used)			
120	78	(This position shall not be used)			
121	79	(This position shall not be used)			
122	7A	GREEK YPOGEGRAMMENI			
123	7B	(This position shall not be used)			
124	7C	(This position shall not be used)			
125	7D	(This position shall not be used)			
126	7E	GREEK QUESTION MARK (Erotimatiko)			
127	7F	(This position shall not be used)			
128	80	(This position shall not be used)			
129	81	(This position shall not be used)			
130	82	(This position shall not be used)			
131	83	(This position shall not be used)			
132	84	GREEK TONOS			
133	85	GREEK DIALYTIKA TONOS			
134	86	GREEK CAPITAL LETTER ALPHA WITH TONOS			
135	87	GREEK ANO TELEIA			
136	88	GREEK CAPITAL LETTER EPSILON WITH TONOS			
137	89	GREEK CAPITAL LETTER ETA WITH TONOS			
138	8A	GREEK CAPITAL LETTER IOTA WITH TONOS			
139	8B	(This position shall not be used)			
140	8C	GREEK CAPITAL LETTER OMICRON WITH TONOS			
141	8D	(This position shall not be used)			
142	8E	GREEK CAPITAL LETTER UPSILON WITH TONOS			
143	8F	GREEK CAPITAL LETTER OMEGA WITH TONOS			
144	90	GREEK SMALL LETTER IOTA WITH DIALYTIKA AND TONOS			
145	91	GREEK CAPITAL LETTER ALPHA			
146	92	GREEK CAPITAL LETTER BETA			
147	93	GREEK CAPITAL LETTER GAMMA			
148	94	GREEK CAPITAL LETTER DELTA			
149	95	GREEK CAPITAL LETTER EPSILON			
150	96	GREEK CAPITAL LETTER ZETA			
151	97	GREEK CAPITAL LETTER ETA			
152	98	GREEK CAPITAL LETTER THETA			
153	99	GREEK CAPITAL LETTER IOTA			
154	9A	GREEK CAPITAL LETTER KAPPA			
155	9B	GREEK CAPITAL LETTER LAMDA			
156	9C	GREEK CAPITAL LETTER MU			
157	9D	GREEK CAPITAL LETTER NU			
158	9E	GREEK CAPITAL LETTER XI			
159	9F	GREEK CAPITAL LETTER OMICRON			
160	A0	GREEK CAPITAL LETTER PI			
161	A1	GREEK CAPITAL LETTER RHO			
162	A2	(This position shall not be used)			
163	A3	GREEK CAPITAL LETTER SIGMA			
164	A4	GREEK CAPITAL LETTER TAU			
165	A5	GREEK CAPITAL LETTER UPSILON			
166	A6	GREEK CAPITAL LETTER PHI			
167	A7	GREEK CAPITAL LETTER CHI			
168	A8	GREEK CAPITAL LETTER PSI			
169	A9	GREEK CAPITAL LETTER OMEGA			
170	AA	GREEK CAPITAL LETTER IOTA WITH DIALYTIKA			
171	AB	GREEK CAPITAL LETTER UPSILON WITH DIALYTIKA			
172	AC	GREEK SMALL LETTER ALPHA WITH TONOS			
173	AD	GREEK SMALL LETTER EPSILON WITH TONOS			
174	AE	GREEK SMALL LETTER ETA WITH TONOS			
175	AF	GREEK SMALL LETTER IOTA WITH TONOS			
176	B0	GREEK SMALL LETTER UPSILON WITH DIALYTIKA AND TONOS			
177	B1	GREEK SMALL LETTER ALPHA			
178	B2	GREEK SMALL LETTER BETA			
179	B3	GREEK SMALL LETTER GAMMA			
180	B4	GREEK SMALL LETTER DELTA			
181	B5	GREEK SMALL LETTER EPSILON			
182	B6	GREEK SMALL LETTER ZETA			
183	B7	GREEK SMALL LETTER ETA			
184	B8	GREEK SMALL LETTER THETA			
185	B9	GREEK SMALL LETTER IOTA			
186	BA	GREEK SMALL LETTER KAPPA			
187	BB	GREEK SMALL LETTER LAMDA			
188	BC	GREEK SMALL LETTER MU			
189	BD	GREEK SMALL LETTER NU			
190	BE	GREEK SMALL LETTER XI			
191	BF	GREEK SMALL LETTER OMICRON			
192	C0	GREEK SMALL LETTER PI			
193	C1	GREEK SMALL LETTER RHO			
194	C2	GREEK SMALL LETTER FINAL SIGMA			
195	C3	GREEK SMALL LETTER SIGMA			
196	C4	GREEK SMALL LETTER TAU			
197	C5	GREEK SMALL LETTER UPSILON			
198	C6	GREEK SMALL LETTER PHI			
199	C7	GREEK SMALL LETTER CHI			
200	C8	GREEK SMALL LETTER PSI			
201	C9	GREEK SMALL LETTER OMEGA			
202	CA	GREEK SMALL LETTER IOTA WITH DIALYTIKA			

TABLE 10 - Row 03: GREEK SYMBOLS AND COPTIC

	03D	03E	03F
0	ϐ 208	ϑ 224	κ 240
1	ϑ 209		ρ 241
2	Υ 210	Ϡ 226	Ϸ 242
3	Ϛ 211	ϡ 227	Ϣ 243
4	ϛ 212	ϣ 228	
5	Ϝ 213	Ϥ 229	
6	ϝ 214	ϥ 230	
7		Ϧ 231	
8		ϧ 232	
9		Ϩ 233	
A	Ϫ 218	ϩ 234	
B		ϫ 235	
C	Ϭ 220	ϭ 236	
D		Ϯ 237	
E	ϯ 222	ϯ 238	
F		ϰ 239	

G = 00
P = 00

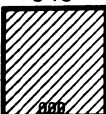
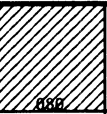
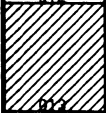
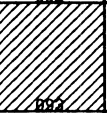
TABLE 10 – Row 03: GREEK SYMBOLS AND COPTIC

dec	hex	Name	dec	hex	Name
208	D0	GREEK BETA SYMBOL			
209	D1	GREEK THETA SYMBOL			
210	D2	GREEK UPSILON WITH HOOK SYMBOL			
211	D3	GREEK UPSILON WITH ACUTE AND HOOK SYMBOL			
212	D4	GREEK UPSILON WITH DIAERESIS AND HOOK SYMBOL			
213	D5	GREEK PHI SYMBOL			
214	D6	GREEK PI SYMBOL			
215	D7	(This position shall not be used)			
216	D8	(This position shall not be used)			
217	D9	(This position shall not be used)			
218	DA	GREEK LETTER STIGMA			
219	DB	(This position shall not be used)			
220	DC	GREEK LETTER DIGAMMA			
221	DD	(This position shall not be used)			
222	DE	GREEK LETTER KOPPA			
223	DF	(This position shall not be used)			
224	E0	GREEK LETTER SAMPI			
225	E1	(This position shall not be used)			
226	E2	COPTIC CAPITAL LETTER SHEI			
227	E3	COPTIC SMALL LETTER SHEI			
228	E4	COPTIC CAPITAL LETTER FEI			
229	E5	COPTIC SMALL LETTER FEI			
230	E6	COPTIC CAPITAL LETTER KHEI			
231	E7	COPTIC SMALL LETTER KHEI			
232	E8	COPTIC CAPITAL LETTER HORI			
233	E9	COPTIC SMALL LETTER HORI			
234	EA	COPTIC CAPITAL LETTER GANGIA			
235	EB	COPTIC SMALL LETTER GANGIA			
236	EC	COPTIC CAPITAL LETTER SHIMA			
237	ED	COPTIC SMALL LETTER SHIMA			
238	EE	COPTIC CAPITAL LETTER DEI			
239	EF	COPTIC SMALL LETTER DEI			
240	F0	GREEK KAPPA SYMBOL			
241	F1	GREEK RHO SYMBOL			
242	F2	GREEK LUNATE SIGMA SYMBOL			
243	F3	GREEK LETTER YOT			
244	F4	(This position shall not be used)			
245	F5	(This position shall not be used)			
246	F6	(This position shall not be used)			
247	F7	(This position shall not be used)			
248	F8	(This position shall not be used)			
249	F9	(This position shall not be used)			
250	FA	(This position shall not be used)			
251	FB	(This position shall not be used)			
252	FC	(This position shall not be used)			
253	FD	(This position shall not be used)			
254	FE	(This position shall not be used)			
255	FF	(This position shall not be used)			

Group 00

Plane 00

TABLE 11 – Row 04: CYRILLIC

	040	041	042	043	044	045	046	047
0		А	Р	а	р		Ѧ	Ѣ
1	Ё	Б	С	б	с	ё	Ѡ	ѣ
2	Ѥ	В	Т	в	т	ѥ	Ѣ	Ѧ
3	Ѓ	Г	У	г	у	ѓ	ѣ	Ѧ
4	Є	Д	Ф	д	ф	є	Ѓ	Ѣ
5	Ѕ	Е	Х	е	х	ѕ	Ѓ	Ѣ
6	Ї	Ж	Ц	ж	ц	ї	Ѓ	Ѣ
7	Ї	З	Ч	з	ч	ї	Ѓ	Ѣ
8	Ј	И	Ш	и	ш	ј	Ѓ	Ѣ
9	Љ	Й	Щ	й	щ	љ	Ѓ	Ѣ
A	Њ	К	Ѣ	к	ѣ	њ	Ѓ	Ѣ
B	Ѥ	Л	Ы	л	ы	ѥ	Ѓ	Ѣ
C	Ѓ	М	Ь	м	ь	Ѓ	Ѓ	Ѣ
D		Н	Э	н	э		Ѓ	Ѣ
E	Ѓ	О	Ю	о	ю	Ѓ	Ѓ	Ѣ
F	Ѣ	П	Я	п	я	ѣ	Ѓ	Ѣ

G = 00

P = 00

TABLE 11 - Row 04: CYRILLIC

dec	hex	Name	dec	hex	Name
000	00	(This position shall not be used)	091	5B	CYRILLIC SMALL LETTER TSHE (Serbocroatian)
001	01	CYRILLIC CAPITAL LETTER IO	092	5C	CYRILLIC SMALL LETTER KJE
002	02	CYRILLIC CAPITAL LETTER DJE (Serbocroatian)	093	5D	(This position shall not be used)
003	03	CYRILLIC CAPITAL LETTER GJE	094	5E	CYRILLIC SMALL LETTER SHORT U (Byelorussian)
004	04	CYRILLIC CAPITAL LETTER UKRAINIAN IE	095	5F	CYRILLIC SMALL LETTER DZHE
005	05	CYRILLIC CAPITAL LETTER DZE	096	60	CYRILLIC CAPITAL LETTER OMEGA
006	06	CYRILLIC CAPITAL LETTER BYELORUSSIAN-UKRAINIAN I	097	61	CYRILLIC SMALL LETTER OMEGA
007	07	CYRILLIC CAPITAL LETTER YI (Ukrainian)	098	62	CYRILLIC CAPITAL LETTER YAT
008	08	CYRILLIC CAPITAL LETTER JE	099	63	CYRILLIC SMALL LETTER YAT
009	09	CYRILLIC CAPITAL LETTER LJE	100	64	CYRILLIC CAPITAL LETTER IOTIFIED E
010	0A	CYRILLIC CAPITAL LETTER NJE	101	65	CYRILLIC SMALL LETTER IOTIFIED E
011	0B	CYRILLIC CAPITAL LETTER TSHE (Serbocroatian)	102	66	CYRILLIC CAPITAL LETTER LITTLE YUS
012	0C	CYRILLIC CAPITAL LETTER KJE	103	67	CYRILLIC SMALL LETTER LITTLE YUS
013	0D	(This position shall not be used)	104	68	CYRILLIC CAPITAL LETTER IOTIFIED LITTLE YUS
014	0E	CYRILLIC CAPITAL LETTER SHORT U (Byelorussian)	105	69	CYRILLIC SMALL LETTER IOTIFIED LITTLE YUS
015	0F	CYRILLIC CAPITAL LETTER DZHE	106	6A	CYRILLIC CAPITAL LETTER BIG YUS
016	10	CYRILLIC CAPITAL LETTER A	107	6B	CYRILLIC SMALL LETTER BIG YUS
017	11	CYRILLIC CAPITAL LETTER BE	108	6C	CYRILLIC CAPITAL LETTER IOTIFIED BIG YUS
018	12	CYRILLIC CAPITAL LETTER VE	109	6D	CYRILLIC SMALL LETTER IOTIFIED BIG YUS
019	13	CYRILLIC CAPITAL LETTER GHE	110	6E	CYRILLIC CAPITAL LETTER KSI
020	14	CYRILLIC CAPITAL LETTER DE	111	6F	CYRILLIC SMALL LETTER KSI
021	15	CYRILLIC CAPITAL LETTER IE	112	70	CYRILLIC CAPITAL LETTER PSI
022	16	CYRILLIC CAPITAL LETTER ZHE	113	71	CYRILLIC SMALL LETTER PSI
023	17	CYRILLIC CAPITAL LETTER ZE	114	72	CYRILLIC CAPITAL LETTER FITA
024	18	CYRILLIC CAPITAL LETTER I	115	73	CYRILLIC SMALL LETTER FITA
025	19	CYRILLIC CAPITAL LETTER SHORT I	116	74	CYRILLIC CAPITAL LETTER IZHITSA
026	1A	CYRILLIC CAPITAL LETTER KA	117	75	CYRILLIC SMALL LETTER IZHITSA
027	1B	CYRILLIC CAPITAL LETTER EL	118	76	CYRILLIC CAPITAL LETTER IZHITSA WITH DOUBLE GRAVE ACCENT
028	1C	CYRILLIC CAPITAL LETTER EM	119	77	CYRILLIC SMALL LETTER IZHITSA WITH DOUBLE GRAVE ACCENT
029	1D	CYRILLIC CAPITAL LETTER EN	120	78	CYRILLIC CAPITAL LETTER UK
030	1E	CYRILLIC CAPITAL LETTER O	121	79	CYRILLIC SMALL LETTER UK
031	1F	CYRILLIC CAPITAL LETTER PE	122	7A	CYRILLIC CAPITAL LETTER ROUND OMEGA
032	20	CYRILLIC CAPITAL LETTER ER	123	7B	CYRILLIC SMALL LETTER ROUND OMEGA
033	21	CYRILLIC CAPITAL LETTER ES	124	7C	CYRILLIC CAPITAL LETTER OMEGA WITH TITLO
034	22	CYRILLIC CAPITAL LETTER TE	125	7D	CYRILLIC SMALL LETTER OMEGA WITH TITLO
035	23	CYRILLIC CAPITAL LETTER U	126	7E	CYRILLIC CAPITAL LETTER OT
036	24	CYRILLIC CAPITAL LETTER EF	127	7F	CYRILLIC SMALL LETTER OT
037	25	CYRILLIC CAPITAL LETTER HA			
038	26	CYRILLIC CAPITAL LETTER TSE			
039	27	CYRILLIC CAPITAL LETTER CHE			
040	28	CYRILLIC CAPITAL LETTER SHA			
041	29	CYRILLIC CAPITAL LETTER SHCHA			
042	2A	CYRILLIC CAPITAL LETTER HARD SIGN			
043	2B	CYRILLIC CAPITAL LETTER YERU			
044	2C	CYRILLIC CAPITAL LETTER SOFT SIGN			
045	2D	CYRILLIC CAPITAL LETTER E			
046	2E	CYRILLIC CAPITAL LETTER YU			
047	2F	CYRILLIC CAPITAL LETTER YA			
048	30	CYRILLIC SMALL LETTER A			
049	31	CYRILLIC SMALL LETTER BE			
050	32	CYRILLIC SMALL LETTER VE			
051	33	CYRILLIC SMALL LETTER GHE			
052	34	CYRILLIC SMALL LETTER DE			
053	35	CYRILLIC SMALL LETTER IE			
054	36	CYRILLIC SMALL LETTER ZHE			
055	37	CYRILLIC SMALL LETTER ZE			
056	38	CYRILLIC SMALL LETTER I			
057	39	CYRILLIC SMALL LETTER SHORT I			
058	3A	CYRILLIC SMALL LETTER KA			
059	3B	CYRILLIC SMALL LETTER EL			
060	3C	CYRILLIC SMALL LETTER EM			
061	3D	CYRILLIC SMALL LETTER EN			
062	3E	CYRILLIC SMALL LETTER O			
063	3F	CYRILLIC SMALL LETTER PE			
064	40	CYRILLIC SMALL LETTER ER			
065	41	CYRILLIC SMALL LETTER ES			
066	42	CYRILLIC SMALL LETTER TE			
067	43	CYRILLIC SMALL LETTER U			
068	44	CYRILLIC SMALL LETTER EF			
069	45	CYRILLIC SMALL LETTER HA			
070	46	CYRILLIC SMALL LETTER TSE			
071	47	CYRILLIC SMALL LETTER CHE			
072	48	CYRILLIC SMALL LETTER SHA			
073	49	CYRILLIC SMALL LETTER SHCHA			
074	4A	CYRILLIC SMALL LETTER HARD SIGN			
075	4B	CYRILLIC SMALL LETTER YERU			
076	4C	CYRILLIC SMALL LETTER SOFT SIGN			
077	4D	CYRILLIC SMALL LETTER E			
078	4E	CYRILLIC SMALL LETTER YU			
079	4F	CYRILLIC SMALL LETTER YA			
080	50	(This position shall not be used)			
081	51	CYRILLIC SMALL LETTER IO			
082	52	CYRILLIC SMALL LETTER DJE (Serbocroatian)			
083	53	CYRILLIC SMALL LETTER GJE			
084	54	CYRILLIC SMALL LETTER UKRAINIAN IE			
085	55	CYRILLIC SMALL LETTER DZE			
086	56	CYRILLIC SMALL LETTER BYELORUSSIAN-UKRAINIAN I			
087	57	CYRILLIC SMALL LETTER YI (Ukrainian)			
088	58	CYRILLIC SMALL LETTER JE			
089	59	CYRILLIC SMALL LETTER LJE			
090	5A	CYRILLIC SMALL LETTER NJE			

Group 00

Plane 00

Row 04

TABLE 12 – Row 04: CYRILLIC

	048	049	04A	04B	04C	04D	04E	04F
0	Ѕ 128	Ґ 144	Ҝ 160	Ү 176	І 192	Ӑ 208	З 224	ӱ 240
1	ѕ 129	ґ 145	ҝ 161	ү 177	ӑ 193	ӓ 209	з 225	ӱ 241
2	Җ 130	Ғ 146	Ң 162	Ҳ 178	ӓ 194	Ӑ 210	Й 226	ӱ 242
3	Ү 131	ғ 147	ң 163	ҳ 179	ӑ 195	ӓ 211	й 227	ӱ 243
4	ӱ 132	Б 148	Н 164	Ц 180	ӑ 196	Æ 212	Й 228	ӱ 244
5	ӱ 133	Б 149	Н 165	Ц 181	ӑ 197	æ 213	й 229	ӱ 245
6	ӱ 134	Ж 150	П 166	Ч 182	ӑ 198	Ӑ 214	Ӑ 230	ӑ 246
7	ӑ 135	ж 151	п 167	ч 183	ӑ 199	ӓ 215	ӓ 231	ӑ 247
8	ӑ 136	З 152	Q 168	Ч 184	ӑ 200	ӓ 216	ӓ 232	ӑ 248
9	ӑ 137	з 153	Q 169	ч 185	ӑ 201	ӓ 217	ӓ 233	ӑ 249
A	ӑ 138	Қ 154	Ҹ 170	Һ 186	ӑ 202	Ӑ 218	Ӑ 234	ӑ 250
B	ӑ 139	қ 155	Ҹ 171	Һ 187	ӑ 203	ӓ 219	ӓ 235	ӑ 251
C	ӑ 140	К 156	Т 172	Ӑ 188	ӑ 204	ӑ 220	ӑ 236	ӑ 252
D	ӑ 141	к 157	т 173	Ӑ 189	ӑ 205	ӑ 221	ӑ 237	ӑ 253
E	ӑ 142	К 158	У 174	Ӑ 190	ӑ 206	Ӑ 222	ӑ 238	ӑ 254
F	ӑ 143	к 159	у 175	Ӑ 191	ӑ 207	Ӑ 223	ӑ 239	ӑ 255

G = 00
P = 00

TABLE 12 – Row 04: CYRILLIC

dec	hex	Name	dec	hex	Name
128	80	CYRILLIC CAPITAL LETTER KOPPA	219	DB	CYRILLIC SMALL LETTER SCHWA WITH DIAERESIS
129	81	CYRILLIC SMALL LETTER KOPPA	220	DC	CYRILLIC CAPITAL LETTER ZHE WITH DIAERESIS
130	82	CYRILLIC THOUSANDS SIGN	221	DD	CYRILLIC SMALL LETTER ZHE WITH DIAERESIS
131	83	COMBINING CYRILLIC TITLO	222	DE	CYRILLIC CAPITAL LETTER ZE WITH DIAERESIS
132	84	COMBINING CYRILLIC PALATALIZATION	223	DF	CYRILLIC SMALL LETTER ZE WITH DIAERESIS
133	85	COMBINING CYRILLIC DASIA PNEUMATA	224	E0	CYRILLIC CAPITAL LETTER ABKHASIAN DZE
134	86	COMBINING CYRILLIC PSILI PNEUMATA	225	E1	CYRILLIC SMALL LETTER ABKHASIAN DZE
135	87	(This position shall not be used)	226	E2	CYRILLIC CAPITAL LETTER I WITH MACRON
136	88	(This position shall not be used)	227	E3	CYRILLIC SMALL LETTER I WITH MACRON
137	89	(This position shall not be used)	228	E4	CYRILLIC CAPITAL LETTER I WITH DIAERESIS
138	8A	(This position shall not be used)	229	E5	CYRILLIC SMALL LETTER I WITH DIAERESIS
139	8B	(This position shall not be used)	230	E6	CYRILLIC CAPITAL LETTER O WITH DIAERESIS
140	8C	(This position shall not be used)	231	E7	CYRILLIC SMALL LETTER O WITH DIAERESIS
141	8D	(This position shall not be used)	232	E8	CYRILLIC CAPITAL LETTER BARRED O
142	8E	(This position shall not be used)	233	E9	CYRILLIC SMALL LETTER BARRED O
143	8F	(This position shall not be used)	234	EA	CYRILLIC CAPITAL LETTER BARRED O WITH DIAERESIS
144	90	CYRILLIC CAPITAL LETTER GHE WITH UPTURN	235	EB	CYRILLIC SMALL LETTER BARRED O WITH DIAERESIS
145	91	CYRILLIC SMALL LETTER GHE WITH UPTURN	236	EC	(This position shall not be used)
146	92	CYRILLIC CAPITAL LETTER GHE WITH STROKE	237	ED	(This position shall not be used)
147	93	CYRILLIC SMALL LETTER GHE WITH STROKE	238	EE	CYRILLIC CAPITAL LETTER U WITH MACRON
148	94	CYRILLIC CAPITAL LETTER GHE WITH MIDDLE HOOK	239	EF	CYRILLIC SMALL LETTER U WITH MACRON
149	95	CYRILLIC SMALL LETTER GHE WITH MIDDLE HOOK	240	F0	CYRILLIC CAPITAL LETTER U WITH DIAERESIS
150	96	CYRILLIC CAPITAL LETTER ZHE WITH DESCENDER	241	F1	CYRILLIC SMALL LETTER U WITH DIAERESIS
151	97	CYRILLIC SMALL LETTER ZHE WITH DESCENDER	242	F2	CYRILLIC CAPITAL LETTER U WITH DOUBLE ACUTE
152	98	CYRILLIC CAPITAL LETTER ZE WITH DESCENDER	243	F3	CYRILLIC SMALL LETTER U WITH DOUBLE ACUTE
153	99	CYRILLIC SMALL LETTER ZE WITH DESCENDER	244	F4	CYRILLIC CAPITAL LETTER CHE WITH DIAERESIS
154	9A	CYRILLIC CAPITAL LETTER KA WITH DESCENDER	245	F5	CYRILLIC SMALL LETTER CHE WITH DIAERESIS
155	9B	CYRILLIC SMALL LETTER KA WITH DESCENDER	246	F6	(This position shall not be used)
156	9C	CYRILLIC CAPITAL LETTER KA WITH VERTICAL STROKE	247	F7	(This position shall not be used)
157	9D	CYRILLIC SMALL LETTER KA WITH VERTICAL STROKE	248	F8	CYRILLIC CAPITAL LETTER YERU WITH DIAERESIS
158	9E	CYRILLIC CAPITAL LETTER KA WITH STROKE	249	F9	CYRILLIC SMALL LETTER YERU WITH DIAERESIS
159	9F	CYRILLIC SMALL LETTER KA WITH STROKE	250	FA	(This position shall not be used)
160	A0	CYRILLIC CAPITAL LETTER BASHKIR KA	251	FB	(This position shall not be used)
161	A1	CYRILLIC SMALL LETTER BASHKIR KA	252	FC	(This position shall not be used)
162	A2	CYRILLIC CAPITAL LETTER EN WITH DESCENDER	253	FD	(This position shall not be used)
163	A3	CYRILLIC SMALL LETTER EN WITH DESCENDER	254	FE	(This position shall not be used)
164	A4	CYRILLIC CAPITAL LIGATURE EN GHE	255	FF	(This position shall not be used)
165	A5	CYRILLIC SMALL LIGATURE EN GHE			
166	A6	CYRILLIC CAPITAL LETTER PE WITH MIDDLE HOOK (Abkhasian)			
167	A7	CYRILLIC SMALL LETTER PE WITH MIDDLE HOOK (Abkhasian)			
168	A8	CYRILLIC CAPITAL LETTER ABKHASIAN HA			
169	A9	CYRILLIC SMALL LETTER ABKHASIAN HA			
170	AA	CYRILLIC CAPITAL LETTER ES WITH DESCENDER			
171	AB	CYRILLIC SMALL LETTER ES WITH DESCENDER			
172	AC	CYRILLIC CAPITAL LETTER TE WITH DESCENDER			
173	AD	CYRILLIC SMALL LETTER TE WITH DESCENDER			
174	AE	CYRILLIC CAPITAL LETTER STRAIGHT U			
175	AF	CYRILLIC SMALL LETTER STRAIGHT U			
176	B0	CYRILLIC CAPITAL LETTER STRAIGHT U WITH STROKE			
177	B1	CYRILLIC SMALL LETTER STRAIGHT U WITH STROKE			
178	B2	CYRILLIC CAPITAL LETTER HA WITH DESCENDER			
179	B3	CYRILLIC SMALL LETTER HA WITH DESCENDER			
180	B4	CYRILLIC CAPITAL LIGATURE TE TSE (Abkhasian)			
181	B5	CYRILLIC SMALL LIGATURE TE TSE (Abkhasian)			
182	B6	CYRILLIC CAPITAL LETTER CHE WITH DESCENDER			
183	B7	CYRILLIC SMALL LETTER CHE WITH DESCENDER			
184	B8	CYRILLIC CAPITAL LETTER CHE WITH VERTICAL STROKE			
185	B9	CYRILLIC SMALL LETTER CHE WITH VERTICAL STROKE			
186	BA	CYRILLIC CAPITAL LETTER SHHA			
187	BB	CYRILLIC SMALL LETTER SHHA			
188	BC	CYRILLIC CAPITAL LETTER ABKHASIAN CHE			
189	BD	CYRILLIC SMALL LETTER ABKHASIAN CHE			
190	BE	CYRILLIC CAPITAL LETTER ABKHASIAN CHE WITH DESCENDER			
191	BF	CYRILLIC SMALL LETTER ABKHASIAN CHE WITH DESCENDER			
192	C0	CYRILLIC LETTER PALOCHKA			
193	C1	CYRILLIC CAPITAL LETTER ZHE WITH BREVE			
194	C2	CYRILLIC SMALL LETTER ZHE WITH BREVE			
195	C3	CYRILLIC CAPITAL LETTER KA WITH HOOK			
196	C4	CYRILLIC SMALL LETTER KA WITH HOOK			
197	C5	(This position shall not be used)			
198	C6	(This position shall not be used)			
199	C7	CYRILLIC CAPITAL LETTER EN WITH HOOK			
200	C8	CYRILLIC SMALL LETTER EN WITH HOOK			
201	C9	(This position shall not be used)			
202	CA	(This position shall not be used)			
203	CB	CYRILLIC CAPITAL LETTER KHAKASSIAN CHE			
204	CC	CYRILLIC SMALL LETTER KHAKASSIAN CHE			
205	CD	(This position shall not be used)			
206	CE	(This position shall not be used)			
207	CF	(This position shall not be used)			
208	D0	CYRILLIC CAPITAL LETTER A WITH BREVE			
209	D1	CYRILLIC SMALL LETTER A WITH BREVE			
210	D2	CYRILLIC CAPITAL LETTER A WITH DIAERESIS			
211	D3	CYRILLIC SMALL LETTER A WITH DIAERESIS			
212	D4	CYRILLIC CAPITAL LIGATURE A IE			
213	D5	CYRILLIC SMALL LIGATURE A IE			
214	D6	CYRILLIC CAPITAL LETTER IE WITH BREVE			
215	D7	CYRILLIC SMALL LETTER IE WITH BREVE			
216	D8	CYRILLIC CAPITAL LETTER SCHWA			
217	D9	CYRILLIC SMALL LETTER SCHWA			
218	DA	CYRILLIC CAPITAL LETTER SCHWA WITH DIAERESIS			

TABLE 13 – Row 05: ARMENIAN

	050	051	052	053	054	055	056	057	058
0					Հ	Ր		հ	ր
1				Ա	Զ	Յ	ա	ճ	ց
2				Բ	Ղ	Ի	բ	ղ	լ
3				Գ	Ճ	Փ	գ	ճ	փ
4				Դ	Մ	Ք	դ	մ	ֆ
5				Ե	Յ	Օ	ե	յ	օ
6				Զ	Ն	Ֆ	զ	ն	ֆ
7				Է	Շ		է	շ	և
8				Ը	Ո		ը	ո	
9				Թ	Չ	Շ	թ	չ	:
A				Ժ	Պ	՝	ժ	պ	
B				Ի	Ջ	՛	ի	ջ	
C				Լ	Ռ	՜	լ	ռ	
D				Խ	Ս	՝	խ	ս	
E				Ծ	Վ	՞	ծ	վ	
F				Կ	Տ	՟	կ	տ	

G = 00

P = 00

TABLE 13 – Row 05: ARMENIAN

dec	hex	Name	dec	hex	Name
000	00	(This position shall not be used)	091	5B	ARMENIAN EMPHASIS MARK
001	01	(This position shall not be used)	092	5C	ARMENIAN EXCLAMATION MARK
002	02	(This position shall not be used)	093	5D	ARMENIAN COMMA
003	03	(This position shall not be used)	094	5E	ARMENIAN QUESTION MARK
004	04	(This position shall not be used)	095	5F	ARMENIAN ABBREVIATION MARK
005	05	(This position shall not be used)	096	60	(This position shall not be used)
006	06	(This position shall not be used)	097	61	ARMENIAN SMALL LETTER AYB
007	07	(This position shall not be used)	098	62	ARMENIAN SMALL LETTER BEN
008	08	(This position shall not be used)	099	63	ARMENIAN SMALL LETTER GIM
009	09	(This position shall not be used)	100	64	ARMENIAN SMALL LETTER DA
010	0A	(This position shall not be used)	101	65	ARMENIAN SMALL LETTER ECH
011	0B	(This position shall not be used)	102	66	ARMENIAN SMALL LETTER ZA
012	0C	(This position shall not be used)	103	67	ARMENIAN SMALL LETTER EH
013	0D	(This position shall not be used)	104	68	ARMENIAN SMALL LETTER ET
014	0E	(This position shall not be used)	105	69	ARMENIAN SMALL LETTER TO
015	0F	(This position shall not be used)	106	6A	ARMENIAN SMALL LETTER ZHE
016	10	(This position shall not be used)	107	6B	ARMENIAN SMALL LETTER INI
017	11	(This position shall not be used)	108	6C	ARMENIAN SMALL LETTER LIWN
018	12	(This position shall not be used)	109	6D	ARMENIAN SMALL LETTER XEH
019	13	(This position shall not be used)	110	6E	ARMENIAN SMALL LETTER CA
020	14	(This position shall not be used)	111	6F	ARMENIAN SMALL LETTER KEN
021	15	(This position shall not be used)	112	70	ARMENIAN SMALL LETTER HO
022	16	(This position shall not be used)	113	71	ARMENIAN SMALL LETTER JA
023	17	(This position shall not be used)	114	72	ARMENIAN SMALL LETTER GHAD
024	18	(This position shall not be used)	115	73	ARMENIAN SMALL LETTER CHEH
025	19	(This position shall not be used)	116	74	ARMENIAN SMALL LETTER MEN
026	1A	(This position shall not be used)	117	75	ARMENIAN SMALL LETTER YI
027	1B	(This position shall not be used)	118	76	ARMENIAN SMALL LETTER NOW
028	1C	(This position shall not be used)	119	77	ARMENIAN SMALL LETTER SHA
029	1D	(This position shall not be used)	120	78	ARMENIAN SMALL LETTER VO
030	1E	(This position shall not be used)	121	79	ARMENIAN SMALL LETTER CHA
031	1F	(This position shall not be used)	122	7A	ARMENIAN SMALL LETTER PEH
032	20	(This position shall not be used)	123	7B	ARMENIAN SMALL LETTER JHEH
033	21	(This position shall not be used)	124	7C	ARMENIAN SMALL LETTER RA
034	22	(This position shall not be used)	125	7D	ARMENIAN SMALL LETTER SEH
035	23	(This position shall not be used)	126	7E	ARMENIAN SMALL LETTER VEW
036	24	(This position shall not be used)	127	7F	ARMENIAN SMALL LETTER TIWN
037	25	(This position shall not be used)	128	80	ARMENIAN SMALL LETTER REH
038	26	(This position shall not be used)	129	81	ARMENIAN SMALL LETTER CO
039	27	(This position shall not be used)	130	82	ARMENIAN SMALL LETTER YIWN
040	28	(This position shall not be used)	131	83	ARMENIAN SMALL LETTER PIWR
041	29	(This position shall not be used)	132	84	ARMENIAN SMALL LETTER KEH
042	2A	(This position shall not be used)	133	85	ARMENIAN SMALL LETTER OH
043	2B	(This position shall not be used)	134	86	ARMENIAN SMALL LETTER FEH
044	2C	(This position shall not be used)	135	87	ARMENIAN SMALL LIGATURE ECH YIWN
045	2D	(This position shall not be used)	136	88	(This position shall not be used)
046	2E	(This position shall not be used)	137	89	ARMENIAN FULL STOP
047	2F	(This position shall not be used)	138	8A	(This position shall not be used)
048	30	(This position shall not be used)	139	8B	(This position shall not be used)
049	31	ARMENIAN CAPITAL LETTER AYB	140	8C	(This position shall not be used)
050	32	ARMENIAN CAPITAL LETTER BEN	141	8D	(This position shall not be used)
051	33	ARMENIAN CAPITAL LETTER GIM	142	8E	(This position shall not be used)
052	34	ARMENIAN CAPITAL LETTER DA	143	8F	(This position shall not be used)
053	35	ARMENIAN CAPITAL LETTER ECH			
054	36	ARMENIAN CAPITAL LETTER ZA			
055	37	ARMENIAN CAPITAL LETTER EH			
056	38	ARMENIAN CAPITAL LETTER ET			
057	39	ARMENIAN CAPITAL LETTER TO			
058	3A	ARMENIAN CAPITAL LETTER ZHE			
059	3B	ARMENIAN CAPITAL LETTER INI			
060	3C	ARMENIAN CAPITAL LETTER LIWN			
061	3D	ARMENIAN CAPITAL LETTER XEH			
062	3E	ARMENIAN CAPITAL LETTER CA			
063	3F	ARMENIAN CAPITAL LETTER KEN			
064	40	ARMENIAN CAPITAL LETTER HO			
065	41	ARMENIAN CAPITAL LETTER JA			
066	42	ARMENIAN CAPITAL LETTER GHAD			
067	43	ARMENIAN CAPITAL LETTER CHEH			
068	44	ARMENIAN CAPITAL LETTER MEN			
069	45	ARMENIAN CAPITAL LETTER YI			
070	46	ARMENIAN CAPITAL LETTER NOW			
071	47	ARMENIAN CAPITAL LETTER SHA			
072	48	ARMENIAN CAPITAL LETTER VO			
073	49	ARMENIAN CAPITAL LETTER CHA			
074	4A	ARMENIAN CAPITAL LETTER PEH			
075	4B	ARMENIAN CAPITAL LETTER JHEH			
076	4C	ARMENIAN CAPITAL LETTER RA			
077	4D	ARMENIAN CAPITAL LETTER SEH			
078	4E	ARMENIAN CAPITAL LETTER VEW			
079	4F	ARMENIAN CAPITAL LETTER TIWN			
080	50	ARMENIAN CAPITAL LETTER REH			
081	51	ARMENIAN CAPITAL LETTER CO			
082	52	ARMENIAN CAPITAL LETTER YIWN			
083	53	ARMENIAN CAPITAL LETTER PIWR			
084	54	ARMENIAN CAPITAL LETTER KEH			
085	55	ARMENIAN CAPITAL LETTER OH			
086	56	ARMENIAN CAPITAL LETTER FEH			
087	57	(This position shall not be used)			
088	58	(This position shall not be used)			
089	59	ARMENIAN MODIFIER LETTER LEFT HALF RING			
090	5A	ARMENIAN APOSTROPHE			

Group 00

Plane 00

Row 05



Documents

Document Sharing Hub

TABLE 14 – Row 05: HEBREW (Basic and Extended)

	059	05A	05B	05C	05D	05E	05F
0			וּ	וּ	א	נ	וּ
1			וּ	וּ	ב	ס	י
2			וּ	וּ	ג	ע	י
3			וּ	וּ	ד	ף	'
4			וּ		ה	פ	"
5			וּ		ו	ץ	
6			וּ		ז	צ	
7			וּ		ח	ק	
8			וּ		ט	ר	
9			וּ		י	ש	
A					ך	ת	
B			וּ		כ		
C			וּ		ל		
D			וּ		ם		
E			וּ		מ		
F			וּ		ן		

G = 00

P = 00

TABLE 14 – Row 05: HEBREW (Basic and Extended)

dec	hex	Name	dec	hex	Name
144	90	(This position shall not be used)	235	EB	(This position shall not be used)
145	91	(This position shall not be used)	236	EC	(This position shall not be used)
146	92	(This position shall not be used)	237	ED	(This position shall not be used)
147	93	(This position shall not be used)	238	EE	(This position shall not be used)
148	94	(This position shall not be used)	239	EF	(This position shall not be used)
149	95	(This position shall not be used)	240	F0	HEBREW LIGATURE YIDDISH DOUBLE VAV
150	96	(This position shall not be used)	241	F1	HEBREW LIGATURE YIDDISH VAV YOD
151	97	(This position shall not be used)	242	F2	HEBREW LIGATURE YIDDISH DOUBLE YOD
152	98	(This position shall not be used)	243	F3	HEBREW PUNCTUATION GERESH
153	99	(This position shall not be used)	244	F4	HEBREW PUNCTUATION GERSHAYIM
154	9A	(This position shall not be used)	245	F5	(This position shall not be used)
155	9B	(This position shall not be used)	246	F6	(This position shall not be used)
156	9C	(This position shall not be used)	247	F7	(This position shall not be used)
157	9D	(This position shall not be used)	248	F8	(This position shall not be used)
158	9E	(This position shall not be used)	249	F9	(This position shall not be used)
159	9F	(This position shall not be used)	250	FA	(This position shall not be used)
160	A0	(This position shall not be used)	251	FB	(This position shall not be used)
161	A1	(This position shall not be used)	252	FC	(This position shall not be used)
162	A2	(This position shall not be used)	253	FD	(This position shall not be used)
163	A3	(This position shall not be used)	254	FE	(This position shall not be used)
164	A4	(This position shall not be used)	255	FF	(This position shall not be used)
165	A5	(This position shall not be used)			
166	A6	(This position shall not be used)			
167	A7	(This position shall not be used)			
168	A8	(This position shall not be used)			
169	A9	(This position shall not be used)			
170	AA	(This position shall not be used)			
171	AB	(This position shall not be used)			
172	AC	(This position shall not be used)			
173	AD	(This position shall not be used)			
174	AE	(This position shall not be used)			
175	AF	(This position shall not be used)			
176	B0	HEBREW POINT SHEVA			
177	B1	HEBREW POINT HATAF SEGOL			
178	B2	HEBREW POINT HATAF PATAH			
179	B3	HEBREW POINT HATAF QAMATS			
180	B4	HEBREW POINT HIRIQ			
181	B5	HEBREW POINT TSERE			
182	B6	HEBREW POINT SEGOL			
183	B7	HEBREW POINT PATAH			
184	B8	HEBREW POINT QAMATS			
185	B9	HEBREW POINT HOLAM			
186	BA	(This position shall not be used)			
187	BB	HEBREW POINT QUBUTS			
188	BC	HEBREW POINT DAGESH OR MAPIQ			
189	BD	HEBREW POINT METEG			
190	BE	HEBREW PUNCTUATION MAQAF			
191	BF	HEBREW POINT RAFE			
192	C0	HEBREW PUNCTUATION PASEQ			
193	C1	HEBREW POINT SHIN DOT			
194	C2	HEBREW POINT SIN DOT			
195	C3	HEBREW PUNCTUATION SOF PASUQ			
196	C4	(This position shall not be used)			
197	C5	(This position shall not be used)			
198	C6	(This position shall not be used)			
199	C7	(This position shall not be used)			
200	C8	(This position shall not be used)			
201	C9	(This position shall not be used)			
202	CA	(This position shall not be used)			
203	CB	(This position shall not be used)			
204	CC	(This position shall not be used)			
205	CD	(This position shall not be used)			
206	CE	(This position shall not be used)			
207	CF	(This position shall not be used)			
208	D0	HEBREW LETTER ALEF			
209	D1	HEBREW LETTER BET			
210	D2	HEBREW LETTER GIMEL			
211	D3	HEBREW LETTER DALET			
212	D4	HEBREW LETTER HE			
213	D5	HEBREW LETTER VAV			
214	D6	HEBREW LETTER ZAYIN			
215	D7	HEBREW LETTER HET			
216	D8	HEBREW LETTER TET			
217	D9	HEBREW LETTER YOD			
218	DA	HEBREW LETTER FINAL KAF			
219	DB	HEBREW LETTER KAF			
220	DC	HEBREW LETTER LAMED			
221	DD	HEBREW LETTER FINAL MEM			
222	DE	HEBREW LETTER MEM			
223	DF	HEBREW LETTER FINAL NUN			
224	E0	HEBREW LETTER NUN			
225	E1	HEBREW LETTER SAMEKH			
226	E2	HEBREW LETTER AYIN			
227	E3	HEBREW LETTER FINAL PE			
228	E4	HEBREW LETTER PE			
229	E5	HEBREW LETTER FINAL TSADI			
230	E6	HEBREW LETTER TSADI			
231	E7	HEBREW LETTER QOF			
232	E8	HEBREW LETTER RESH			
233	E9	HEBREW LETTER SHIN			
234	EA	HEBREW LETTER TAV			

Group 00

Plane 00

Row 05



Documents

Document Sharing Hub

TABLE 15 - Row 06: ARABIC

	060	061	062	063	064	065	066	067
0				ذ	ـ	َ	٠	٠
1			ء	ر	ف	ّ	١	أ
2			آ	ز	ق	ّ	٢	أ
3			أ	س	ك		٣	ل
4			ؤ	ش	ل		٤	ء
5			إ	ص	م		٥	أ
6			ئ	ض	ن		٦	و
7			ا	ط	ه		٧	ؤ
8			ب	ظ	و		٨	ع
9			ة	ع	ي		٩	ط
A			ت	غ	ي		%	ث
B		؛	ث		ّ		ر	ب
C	،		ج		ّ		و	ت
D			ح		ّ		*	ت
E			خ		ّ			پ
F		؟	د		ّ			ت

G = 00

P = 00

TABLE 15 – Row 06: ARABIC

dec	hex	Name	dec	hex	Name
000	00	(This position shall not be used)	091	5B	(This position shall not be used)
001	01	(This position shall not be used)	092	5C	(This position shall not be used)
002	02	(This position shall not be used)	093	5D	(This position shall not be used)
003	03	(This position shall not be used)	094	5E	(This position shall not be used)
004	04	(This position shall not be used)	095	5F	(This position shall not be used)
005	05	(This position shall not be used)	096	60	ARABIC-INDIC DIGIT ZERO
006	06	(This position shall not be used)	097	61	ARABIC-INDIC DIGIT ONE
007	07	(This position shall not be used)	098	62	ARABIC-INDIC DIGIT TWO
008	08	(This position shall not be used)	099	63	ARABIC-INDIC DIGIT THREE
009	09	(This position shall not be used)	100	64	ARABIC-INDIC DIGIT FOUR
010	0A	(This position shall not be used)	101	65	ARABIC-INDIC DIGIT FIVE
011	0B	(This position shall not be used)	102	66	ARABIC-INDIC DIGIT SIX
012	0C	ARABIC COMMA	103	67	ARABIC-INDIC DIGIT SEVEN
013	0D	(This position shall not be used)	104	68	ARABIC-INDIC DIGIT EIGHT
014	0E	(This position shall not be used)	105	69	ARABIC-INDIC DIGIT NINE
015	0F	(This position shall not be used)	106	6A	ARABIC PERCENT SIGN
016	10	(This position shall not be used)	107	6B	ARABIC DECIMAL SEPARATOR
017	11	(This position shall not be used)	108	6C	ARABIC THOUSANDS SEPARATOR
018	12	(This position shall not be used)	109	6D	ARABIC FIVE POINTED STAR
019	13	(This position shall not be used)	110	6E	(This position shall not be used)
020	14	(This position shall not be used)	111	6F	(This position shall not be used)
021	15	(This position shall not be used)	112	70	ARABIC LETTER SUPERScript ALEF
022	16	(This position shall not be used)	113	71	ARABIC LETTER ALEF WASLA
023	17	(This position shall not be used)	114	72	ARABIC LETTER ALEF WITH WAVY HAMZA ABOVE
024	18	(This position shall not be used)	115	73	ARABIC LETTER ALEF WITH WAVY HAMZA BELOW
025	19	(This position shall not be used)	116	74	ARABIC LETTER HIGH HAMZA
026	1A	(This position shall not be used)	117	75	ARABIC LETTER HIGH HAMZA ALEF
027	1B	ARABIC SEMICOLON	118	76	ARABIC LETTER HIGH HAMZA WAW
028	1C	(This position shall not be used)	119	77	ARABIC LETTER U WITH HAMZA ABOVE
029	1D	(This position shall not be used)	120	78	ARABIC LETTER HIGH HAMZA YEH
030	1E	(This position shall not be used)	121	79	ARABIC LETTER TTEH
031	1F	ARABIC QUESTION MARK	122	7A	ARABIC LETTER TTEHEH
032	20	(This position shall not be used)	123	7B	ARABIC LETTER BEEH
033	21	ARABIC LETTER HAMZA	124	7C	ARABIC LETTER TEH WITH RING
034	22	ARABIC LETTER ALEF WITH MADDA ABOVE	125	7D	ARABIC LETTER TEH WITH THREE DOTS ABOVE DOWNWARDS
035	23	ARABIC LETTER ALEF WITH HAMZA ABOVE	126	7E	ARABIC LETTER PEH
036	24	ARABIC LETTER WAW WITH HAMZA ABOVE	127	7F	ARABIC LETTER TEHEH
037	25	ARABIC LETTER ALEF WITH HAMZA BELOW			
038	26	ARABIC LETTER YEH WITH HAMZA ABOVE			
039	27	ARABIC LETTER ALEF			
040	28	ARABIC LETTER BEH			
041	29	ARABIC LETTER TEH MARBUTA			
042	2A	ARABIC LETTER TEH			
043	2B	ARABIC LETTER THEH			
044	2C	ARABIC LETTER JEEM			
045	2D	ARABIC LETTER HAH			
046	2E	ARABIC LETTER KHAH			
047	2F	ARABIC LETTER DAL			
048	30	ARABIC LETTER THAL			
049	31	ARABIC LETTER REH			
050	32	ARABIC LETTER ZAIN			
051	33	ARABIC LETTER SEEN			
052	34	ARABIC LETTER SHEEN			
053	35	ARABIC LETTER SAD			
054	36	ARABIC LETTER DAD			
055	37	ARABIC LETTER TAH			
056	38	ARABIC LETTER ZAH			
057	39	ARABIC LETTER AIN			
058	3A	ARABIC LETTER GHAIN			
059	3B	(This position shall not be used)			
060	3C	(This position shall not be used)			
061	3D	(This position shall not be used)			
062	3E	(This position shall not be used)			
063	3F	(This position shall not be used)			
064	40	ARABIC TATWEEL			
065	41	ARABIC LETTER FEH			
066	42	ARABIC LETTER QAF			
067	43	ARABIC LETTER KAF			
068	44	ARABIC LETTER LAM			
069	45	ARABIC LETTER MEEM			
070	46	ARABIC LETTER NOON			
071	47	ARABIC LETTER HEH			
072	48	ARABIC LETTER WAW			
073	49	ARABIC LETTER ALEF MAKSURA			
074	4A	ARABIC LETTER YEH			
075	4B	ARABIC FATHATAN			
076	4C	ARABIC DAMMATAN			
077	4D	ARABIC KASRATAN			
078	4E	ARABIC FATHA			
079	4F	ARABIC DAMMA			
080	50	ARABIC KASRA			
081	51	ARABIC SHADDA			
082	52	ARABIC SUKUN			
083	53	(This position shall not be used)			
084	54	(This position shall not be used)			
085	55	(This position shall not be used)			
086	56	(This position shall not be used)			
087	57	(This position shall not be used)			
088	58	(This position shall not be used)			
089	59	(This position shall not be used)			
090	5A	(This position shall not be used)			

Group 00

Plane 00

Row 06



Documents

Document Sharing Hub

TABLE 16 - Row 06: ARABIC

	068	069	06A	06B	06C	06D	06E	06F
0	پ 128	ڈ 144	ش 160	گ 176	ہ 192	ی 208	۰ 224	۰ 240
1	ع 129	ڑ 145	ف 161	گ 177	ہ 193	پی 209	۰ 225	۱ 241
2	خ 130	ز 146	ب 162	گی 178	ء 194	ے 210	۰ 226	۲ 242
3	ج 131	د 147	فب 163	گی 179	تہ 195	ئے 211	۰ 227	۳ 243
4	ج 132	ر 148	قا 164	گی 180	و 196	— 212	۰ 228	۴ 244
5	خ 133	ر 149	پ 165	ن 181	و 197	ہ 213	و 229	۵ 245
6	ج 134	ن 150	قا 166	ن 182	و 198	۰ 214	ے 230	۶ 246
7	ج 135	ر 151	ف 167	ش 183	و 199	۰ 215	۰ 231	۷ 247
8	ڈ 136	ر 152	ق 168	۰ 184	و 200	۰ 216	ن 232	۸ 248
9	د 137	ر 153	ک 169	۰ 185	و 201	لا 217	۰ 233	۹ 249
A	د 138	نیں 154	ک 170	ن 186	ق 202	۰ 218	۰ 234	۰ 250
B	ڈ 139	پیں 155	م 171	ن 187	و 203	۰ 219	۰ 235	۰ 251
C	ڈ 140	شیں 156	ن 172	ن 188	ی 204	س 220	۰ 236	۰ 252
D	د 141	صیں 157	ا 173	ش 189	ی 205	۰ 221	۰ 237	۰ 253
E	ڈ 142	صیں 158	ک 174	ہ 190	ی 206	۰ 222	۰ 238	۰ 254
F	ڈ 143	ظ 159	گ 175	۰ 191	۰ 207	۰ 223	۰ 239	۰ 255

G = 00

P = 00