

**ISO/IEC JTC 1/SC 2/WG 3**  
**7-bit and 8-bit codes and their extension**  
**SECRETARIAT : ELOT**

<b>DOC TYPE :</b>	Officer's Contribution
<b>TITLE :</b>	Proposal for Type 3 Technical Report, TR 15907, Information technology -- Revision of OCR-B standard (ISO 1073/II-1976)
<b>SOURCE :</b>	Project Editor
<b>PROJECT:</b>	---
<b>STATUS :</b>	At the 7th SC 2 plenary meeting held in Crete, Greece, 1997, SC 2 resolved to change the revision of ISO/IEC 1073-2 to preparation of Technical Report Type 2 (Resolution M07.14). This document has been prepared by Project Editor, Mr. K. I. Larsson. It is circulated to SC 2 members for discussion at the 9th SC 2 plenary meeting in Fukuoka, Japan, 1999-03-16/17.
<b>ACTION ID :</b>	ACT
<b>DUE DATE :</b>	----
<b>DISTRIBUTION :</b>	P, O and L Members of ISO/IEC JTC 1/SC 2 WG Conveners and Secretariats Secretariat, ISO/IEC JTC 1 ISO/IEC ITTF
<b>MEDIUM :</b>	Open
<b>NO OF PAGES :</b>	32

Contact 1: Secretariat ISO/IEC JTC 1/SC 2/WG 3 ELOT Mrs K.Velli (acting)  
Acharnon 313, 111 45 Kato Patissia, ATHENS – GREECE  
Tel: +30 1 21 20 307 Fax : +30 1 22 86 219 E-mail : [kkb@elot.gr](mailto:kkb@elot.gr)

Contact 2 : Convener ISO/IEC JTC 1/SC 2/WG 3 Mr E.Melagrakis  
Acharnon 313, 111 45 Kato Patissia, ATHENS – GREECE  
Tel: +30 1 21 20 301 Fax : +30 1 22 86 219 E-mail: [eam@elot.gr](mailto:eam@elot.gr)

## Contents

	Page
Foreword . . . . .	iii
Introduction . . . . .	iv
<b>1 Scope . . . . .</b>	<b>1</b>
<b>2 References . . . . .</b>	<b>1</b>
<b>3 Definitions . . . . .</b>	<b>1</b>
<b>4 Contents of report . . . . .</b>	<b>1</b>
<b>5 OCR application areas . . . . .</b>	<b>1</b>
<b>6 OCR glyph standards . . . . .</b>	<b>2</b>
<b>7 OCR-B characteristics . . . . .</b>	<b>2</b>
<b>8 OCR-B revision history . . . . .</b>	<b>3</b>
<b>9 Extension considerations . . . . .</b>	<b>3</b>
<b>Annex A: Editorially-revised text for OCR-B standard . . . . .</b>	<b>5</b>
<b>Annex B: Main extensions in ISO/IEC CD 1073-2.3 as compared to Annex A . . . . .</b>	<b>26</b>
<b>Annex C: OCR-B glyph repertoire extensions requested by National Bodies . . . . .</b>	<b>27</b>
<b>Annex D: Sample glyph definition according to ISO/IEC 9541-3 . . . . .</b>	<b>28</b>

© ISO/IEC 1998

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

ISO/IEC Copyright Office • Case Postale 56 • CH-1211 Genève 20 • Switzerland  
Printed in Switzerland

## Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

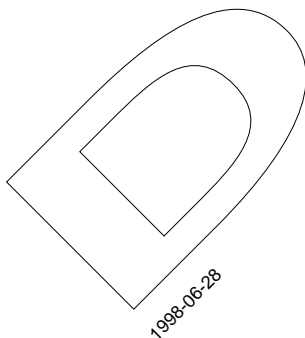
In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The main task of technical committees is to prepare International Standards. In exceptional circumstances a technical committee may propose the publication of a Technical Report of one of the following types :

- type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts:
- type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard:
- type 3, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example).

Technical reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

ISO/IEC TR 15907, which is a Technical Report of Type 3, was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 2, *Character sets and information coding*.

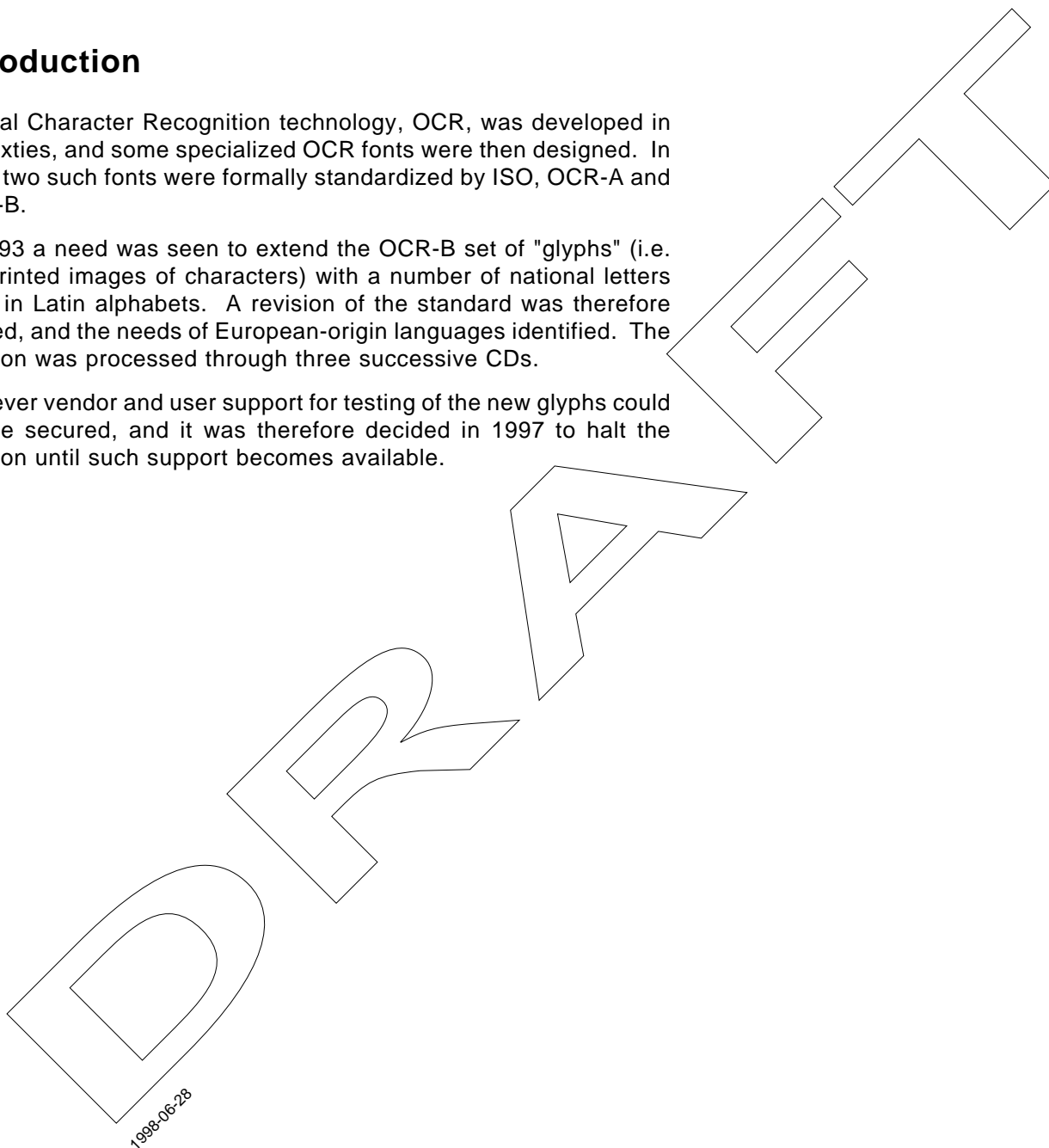


## Introduction

Optical Character Recognition technology, OCR, was developed in the sixties, and some specialized OCR fonts were then designed. In 1976 two such fonts were formally standardized by ISO, OCR-A and OCR-B.

In 1993 a need was seen to extend the OCR-B set of "glyphs" (i.e. the printed images of characters) with a number of national letters used in Latin alphabets. A revision of the standard was therefore started, and the needs of European-origin languages identified. The revision was processed through three successive CDs.

However vendor and user support for testing of the new glyphs could not be secured, and it was therefore decided in 1997 to halt the revision until such support becomes available.



# Information technology – Revision of OCR-B standard (ISO 1073/II-1976)

## 1 Scope

This Technical Report documents the work performed on a revision of the OCR-B font standard. That revision has been halted, and the work item changed to the production of this report; see revision history below.

## 2 References

ISO 1073/I-1976, *Alphanumeric character sets for optical recognition – Part I: Character set OCR-A – Shapes and dimensions of the printed image*

ISO 1073/II-1976, *Alphanumeric character sets for optical recognition – Part II: Character set OCR-B – Shapes and dimensions of the printed image*

ISO 1831-1980, *Printing specifications for optical character recognition*

ISO/IEC 9541-3:1994, *Information technology – Font information interchange – Part 3: Glyph shape representation*

Committee Draft ISO/IEC CD 1073-2.3, *Information technology – Alphanumeric character sets for optical recognition – Part 2: Character set OCR-B – Shapes and dimensions of the printed image* (1996-10-15)

## 3 Definitions

For definitions of terms used in the text of this report see Definitions in Annex A.

## 4 Contents of report

The text of this report does not cover Optical Character Recognition (OCR) technology as such. Neither is the related subject of bar coding covered.

The report however provides, in clauses 5 and 6, some information of a general nature on OCR application areas. The characteristics of the OCR-B font are described in clause 7, and the standard's

revision history in clause 8. Clause 9 contains some general considerations regarding extension of the font.

Annex A is an editorially revised text for the OCR-B standard, with a number of textual additions and clarifications as compared to the current standard. It however does not include the extensions to the glyph repertoire as proposed in the revision and contained in the latest (i.e. third) CD. The text in Annex A is therefore intended for use if it is decided to publish a new edition of the OCR-B standard, updated according to present editing rules, but without repertoire extensions. If, instead, the OCR-B revision is resumed to extend the standard, the latest CD should form the basis for such work.

The differences between Annex A and the latest CD are described in Annex B. Annex C contains a listing of glyph extensions as requested by National Bodies.

Annex D is a sample definition of an OCR-B glyph using the shape definition syntax of ISO/IEC 9541-3 (see clause 9).

## 5 OCR application areas

For the purpose of this report, it is suitable to distinguish between two differing OCR application areas, namely:

- reading of information in predefined fields of form-type documents
- reading of continuous-text documents

(Although in both areas not only printed but also hand-written text applications exist, the latter case is not considered in this report.)

OCR technology originally concentrated on the first of the areas, particularly within the financial sector. Already in the sixties a number of "Optical Readers" were marketed, capable of recognizing the glyphs of a few different fonts.

Applications in the second area were also in small-scale use in the sixties. The technology was greatly accelerated by the introduction of Personal Computers and the development of low-cost scanning equipment for such machines.

Forms-reading applications probably still constitute the largest volume of today's OCR processing. In general they place very high requirements on the reliability of glyph recognition and on the speed of processing. The hardware and software for these applications is generally tightly integrated, and marketed as packaged solutions.

Continuous-text reading applications generally accept lower reliability in glyph recognition and lower processing speed. In this case many independent hardware and software vendors exist, meaning that most software is designed to work with a number of different scanners.

For forms-type applications a specific font is usually specified. The recognition process therefore matches read images against shapes known in detail. In text document scanning, matching is instead to generic glyph shapes, and the process is "self-learning", sometimes referred to as Intelligent Optical Character Recognition, IOCR or ICR. This type of software usually permits operator interaction with the recognition process.

The OCR-B font is primarily intended for use in the forms-type application area. From a recognition point of view it is however also a very suitable font for applications in the document-text area, e.g. for long-time archiving.

## 6 OCR glyph standards

The development of the OCR technology in the sixties necessitated work on fonts suitable for machine-reading. Such work was started both by the computer industry and by standards organizations.

For the Latin alphabet separate work was carried out in the US and in Europe, resulting in two different glyph schemes. In 1976 these two were used to produce the two-part standard ISO 1073, and there named OCR-A and OCR-B, respectively.

This standard is complemented by ISO 1831, which covers printing quality requirements for the fonts as well as methods for conformance testing.

In 1994 an extension to the OCR-B font was proposed, to better accommodate European national letters. A revision of part II of the standard was therefore

started within ISO/IEC JTC 1/SC 2. It was however halted in 1997; see revision history below.

At the moment (June 1998) an interest exists within CEN to extend the glyph repertoire of OCR-B with a sign for the Euro currency. This may possibly result in the development of an EN (i.e. CEN standard) based on the OCR-B standard.

## 7 OCR-B characteristics

The OCR-B standard specifies a total of 121 glyphs. During the work started on the revision of the standard it has however been established that four out of these are unnecessary in present-day OCR, and should be deleted.

The font is specified for three sizes. Size I is the one most commonly used, e.g. in machine-readable passports and on some European banknotes. Its size corresponds roughly to 12 typographical points.

The standard defines two slightly differing versions of glyph shapes; the OCR-B standard therefore actually specifies two complete fonts.

The version originally intended for impact-type printers has the same thickness for all the lines – strokes – that the glyphs consist of. It is consequently termed "constant strokewidth". Also, the ends of most strokes are rounded. This version is the one usually used.

The second version was intended for more capable printing processes, and conforms to traditional typographical practices, with e.g. differing thickness of horizontal and vertical strokes. It is therefore termed "letterpress", and is useful as a font for typewriter-class applications. The ends of the strokes are squarely cut off, not rounded.

The constant-strokewidth version is primarily intended for use with a fixed horizontal spacing ("fixed pitch") of 2,54 mm, i.e. 10 characters per inch (10 cpi). The glyphs are however designed to permit a fixed spacing of 2,12 mm also, i.e. the 12 cpi more common in Europe. The letterpress version, although useable for fixed-pitch printing, is primarily intended for "proportional spacing" as normally used for text nowadays.

When output on medium- and high-quality printers the differences between the two versions can generally be distinguished by the human eye. For most OCR readers, on the other hand, the two font versions are in practice identical; differentiating between them certainly requires OCR readers with better than 300 "dots per inch" resolution. Considering inevitable printing defects and

tolerances, the two versions could therefore be seen as interchangeable from an OCR point of view.

The present OCR-B standard's glyph repertoire is illustrated in Annex A clause 16 at Size I, and also magnified from that four times (which does not correspond to any of the three sizes; this is for illustration only).

## 8 OCR-B revision history

In 1993 an extension was proposed in JTC 1/SC 17 to the standard for machine-readable passports ISO/IEC 7501-1, by the Turkish National Body. That standard does not permit any other letters than capital A–Z for writing names in the part of the passport intended for machine-reading (although they are permitted in the "Visual Inspection Zone" of the document).

The Turkish NB considered it necessary to have the possibility of representing names in a correct way for machine-reading also. This would avoid the ambiguity inherent in most transliteration schemes.

SC 17 at the time noted that the OCR-B standard, as referred to in ISO/IEC 7501, defines a very small repertoire of national letters, not completely covering e.g. the Turkish requirements. (SC 17 has also since concluded that, even if the OCR-B repertoire were extended with such letters, allowing them in passports would cause an unacceptable decrease in recognition performance, at least in installed equipment.)

As a consequence of the SC 17 discussions, the creation of a new part of ISO 1073 was originally proposed by the Turkish NB. In an SC 2/WG 3 meeting in 1994 it was however decided to instead start a revision of the OCR-B standard to extend its glyph repertoire, and also to editorially update it. This revision received work item number JTC 1.02.26.

The original planning was for a rather small extension to the repertoire, but successive NB comments caused a considerably enlargement of this planning. Therefore a number of CDs had to be produced; in July 1994, November 1995 and October 1996.

Since however necessary industry and user support could not be secured for finalizing the revision, in particular for testing the new glyphs proposed, SC 2 decided in July 1997 to halt the revision, and to produce this report instead.

Due partly to CEN requests for a Euro sign in OCR-B, the revision was again discussed in the February 1998 SC 2 Plenary. It was decided that the report should be finalized, even if some parallel work was started in CEN.

## 9 Extension considerations

The original OCR-B design concentrated on the digits 0-9, capital letters A-Z, currency signs and some punctuation marks. That core repertoire – which is the same as for OCR-A – was at the time considered sufficient for normal OCR applications.

Small letters and a few special national letters (like ÄÖÜ) were also designed, to make the font more useful in general office environments. These letters were however not intended for high-reliability OCR applications.

For the common size I, the "frame" within which the core repertoire glyphs all fit is approximately 1,76 mm wide by 2,66 mm high, although slight "over-shoots" occur for some glyphs. The special national letters in the original repertoire are designed to stay within that frame also.

The frame does not accommodate small letters with descenders (like j). Neither would capital letters with diacritical marks (like É) or small letters with ascenders and diacritical marks, as used in some Latin alphabets, stay within that frame.

Therefore extensions of the OCR-B glyph repertoire, as requested in the revision process, may be problematic for some existing OCR applications. Only if dimensional extensions to the recognition areas are possible could such extensions be permitted.

Specifically in the case of Machine-Readable Travel Documents according to ISO/IEC 7501 the defined area for OCR information could accommodate such increased size. It would however severely limit the tolerances needed both when printing such documents and when reading them.

For the current OCR-B repertoire, all glyphs are specified by scale 100:1 reference drawings (see Annex A 14.1). In the case a repertoire extension is decided, glyph specification by outline definitions according to ISO/IEC 9541-3 should be considered instead. Such a sample definition is illustrated in Annex D.

*(This page intentionally left blank)*

DRAFT

1998-06-28



## Annex A

### Editorially-revised text for new OCR-B standard

# Information technology – Alphanumeric character sets for optical recognition

## Part 2: Character set OCR-B – Shapes and dimensions of the printed image

### 1 Scope

This International Standard defines two sets of glyph images, designated OCR-A and OCR-B, and intended primarily for use in Optical Character Recognition (OCR) applications; but suitable also for visual, i.e. human, reading. It does not relate any coding scheme with these images (see clause 5).

**NOTE** – In the previous edition of this standard the term "character" was used not only in its strict sense, but also to mean the printed images used to represent characters visually. In this edition the term "glyph image" has been introduced for the latter meaning (except in the title of the standard, which has been kept unchanged).

This standard contains information on nominal dimensions for the glyph images. Tolerances, printing quality and other characteristics of the formats needed to satisfy interchange requirements are however covered in other International Standards (see clause 3).

The OCR-A set contains 69 glyph images comprising digits, capital letters and symbols. The OCR-A set is specified in part 1 of ISO/IEC 1073.

In this part 2 of the standard the OCR-B set is specified. The set contains 117 glyph images comprising digits, capital and small letters, diacritical marks, and symbols. It also contains a definition for SPACE.

The diacritical marks are designed for combination with small letters to produce composite glyph images complementing the basic images.

### 2 Conformance

A printing or OCR reading device is in conformance with this standard if it can generate/recognize, for either or both of the defined styles (see clause 6) and in one or more of the specified sizes (see

clause 7), all or part of the specified glyph image subsets (see clause 9).

A claim of conformance shall specify all the images in (each of) the style(s) and size(s) generated/recognized. Such a specification shall take the form of a reference to one of the subsets, a list of the images generated/recognized, or a combination of those (see also 9.4).

Additionally, a printing or OCR reading device must claim conformance to International Standard ISO 1831 (see clause 3).

Printed images produced by an OCR-B printing device are in conformance with this standard if their nominal shapes and dimensions are in accordance with their respective reference drawing(s) (see clause 14); with the claimed conformance to tolerances and printing quality factors specified in standard ISO 1831 considered.

### 3 Normative references

The following standards and other documents contain provisions which, through reference in this text, constitute provisions of this International Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this International Standard are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 1831-1980, *Printing specifications for optical character recognition*

*OCR-B character reference drawings and glyph definitions; see clause 14*

## 4 Definitions

For the purposes of this standard, the following definitions apply:

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

**4.2 coded character set:** A set of characters, defined by unambiguous rules that establish the character set and the relationship between the characters of the set and their coded representations.

**4.3 composite glyph image:** An image printed on paper or any other medium intended for OCR applications, obtained by superimposing two or more glyph images on the same area.

**4.4 glyph:** A recognizable abstract graphic symbol which is independent of any specific design.

**4.5 glyph image:** An image of a glyph, as obtained from a glyph representation printed on paper or any other medium intended for OCR applications.

**NOTE** – The definition above of "coded character set" differs slightly from definitions in other ISO/IEC standards, and the definition of "glyph image" is more limited. The definition of "composite glyph image" is specific to this standard (at the time of its publication).

## 5 Coding in OCR applications

This standard defines a set of glyph images, but does not specify corresponding characters, and relates no coding with the images. The images have been named as far as possible in the same way as the characters with corresponding glyphs in the ISO/IEC 10646-1:1993 standard, but this does not imply any formal association between the OCR-B glyph images and the characters of either ISO/IEC 10646 or any other standard for coded character sets.

Printing and/or OCR applications based on this standard must therefore define, through reference to other International Standards or otherwise, the set of glyph images which is available for printing and/or shall be recognized, and for each image the corresponding character and its coding.

## 6 OCR-B styles

The OCR-B glyph images are defined by this standard in two different styles.

The "constant-strokewidth" style is intended primarily for printer equipment in which the width of the strokes of the images is less controllable. This is for instance the case for some types of mechanical printers.

The "letterpress" style is intended for printing equipment which can reproduce fine details with high accuracy. For aesthetic reasons, the strokewidths of the letterpress images are varied deliberately, and the stroke endings are specially designed.

The shapes of the glyph images for the two styles are specified by reference drawings. The constant-strokewidth style drawings show the centrelines of the strokes and the contours of the ends of the strokes; the letterpress style drawings show the complete outlines of the images (see clause 14). The general dimensions and the centrelines of the images are the same for both styles.

## 7 OCR-B sizes

**7.1** Three sizes are specified for OCR-B glyph images in order to provide for use with a wide range of printing equipment possessing differing print quality characteristics.

**NOTE** – The metric and inch dimensions in this International Standard are rounded and therefore consistent but not exactly equal. Either system may be used but the two should not be intermixed.

**7.2** The letterpress font is specified in size I (the smallest) only. It provides the option of a variable pitch in printing as is usual with letterpress.

**7.3** The constant-strokewidth font is specified in three sizes: I, III and IV. Mechanisms using the constant-strokewidth font will usually maintain a fixed pitch.

**NOTE** – Size II which was originally in this standard has been deleted.

In fixed-pitch printing for OCR applications, the following minimum nominal pitches are recommended:

size I:	2,54 mm (0,100 in)
size III:	2,54 mm (0,100 in)
size IV:	3,63 mm (0,143 in)

**7.4** The centrelines for the three sizes are simply related by appropriate horizontal and vertical scale factors. The factors for size III and size IV referred to size I are:

for size III:	Vertical 1,333; horizontal 1,086
for size IV:	Vertical 1,500; horizontal 1,500

This scale relationship does not apply to the outline shapes since nominal strokewidth is not strictly proportional to centreline dimensions. The strokewidths for each size are shown in the reference drawings.

**7.5** The glyph image with the greatest height above the base line ("A" in figure 1) in each size is DIGIT EIGHT. The image with the greatest total height is SMALL LETTER J, because of its descender.

The centreline heights of the DIGIT EIGHT are:

- for size I: 2,40 mm (0,094 in)
- for size III: 3,20 mm (0,126 in)
- for size IV: 3,60 mm (0,142 in)

**7.6** The widest glyph image in each size (except for the alternative SMALL LETTER M) is DIGIT ZERO. Its centreline widths are:

- for size I: 1,40 mm (0,055 in)
- for size III: 1,52 mm (0,060 in)
- for size IV: 2,10 mm (0,083 in)

### 8 Typical dimensions of the nominal printed image

Typical dimensions for the nominal printed image of the letterpress font in size I are given in table 1. These dimensions are the heights above and below the horizontal base line of digits, capital and small letters, ascenders and descenders (see figure 1).

The shapes and dimensions of the constant-strokewidth glyph images are similar except that the stroke ends are rounded.

**NOTE** – It is recognized that some type-making and printing processes will not be able to produce sharp corners. Corners not given a specific radius should be as sharp as practicable. However, it is not necessary for OCR purposes that the radii of the corners of the nominal printed image be less than 0,08 mm (0,0035 in).

The dimensions in table 1 are for general information only. The values for individual glyph images are obtainable from the reference drawings.

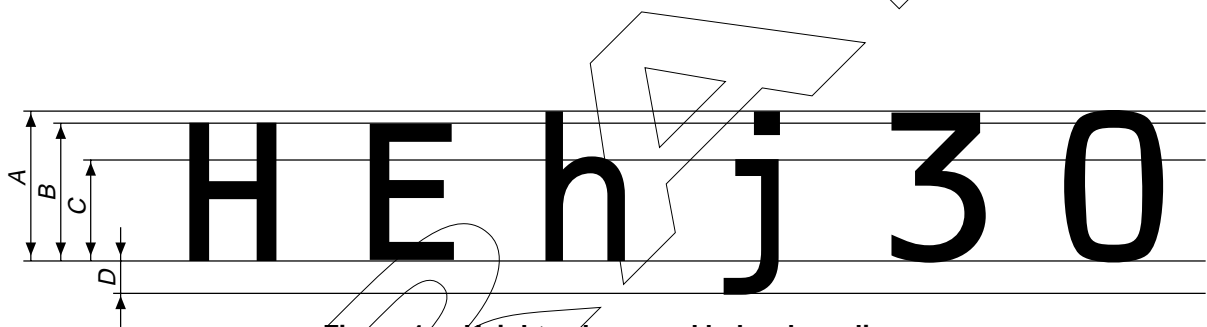


Figure 1 - Heights above and below base line

Table 1 - Typical dimensions A, B, C and D

Size	millimetres				inches			
	A	B	C	D	A	B	C	D
I	2,66	2,46	1,83	0,60	0,105	0,097	0,072	0,024

1998-06-28

## 9 OCR-B GLYPH IMAGE SET

The full set contains 116 glyph images and a definition for SPACE (see clause 13). Four subsets are defined:

### 9.1 Subset 1: Minimal alphanumeric subset

This subset applies to sizes I, III and IV in constant-strokewidth font and to size I in letterpress font. It contains 21 glyph images and SPACE:

0 1 2 3 4 5 6 7 8 9

C E N S T X Z

< + > | SPACE

### 9.2 Subset 2: Basic alphanumeric subset

This subset applies to sizes I and IV in constant-strokewidth font and to size I in letterpress font. It contains 25 glyph images in addition to subset 1, i.e. a total of 46 glyph images and SPACE:

0 1 2 3 4 5 6 7 8 9

A B C D E F G H I J K L M

N O P Q R S T U V W X Y Z

< + > \* - = / . , | SPACE

1998-06-28

**9.3 Subset 3: Extended alphanumeric subset**

This subset applies to sizes I and IV in constant-strokewidth font and to size I in letterpress font.

It contains 50 glyph images in addition to subset 2, i.e 96 glyphs in all (and SPACE); in particular those images corresponding to the characters listed in ISO/IEC 646 as unique, alternative, and International Reference Version.

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

**9.4 Subset 4: Options subset**

This subset applies to sizes I and IV in constant-strokewidth font, and to size I in letter-press font. It contains 8 capital national letters, 5 small national letters, 4 diacritical marks and 3 further glyph images.

Images from this subset shall be used only in conjunction with subset 3. A printing or OCR reading device may generate/recognize any of the images of this subset. The images generated/recognized by the device shall be listed in the claim of conformance (see clause 2).

À Á Â Ã Ä Å Æ Ç È É  
 à á â ã ä å æ ç è é  
 ' ^ " /  
 § Ÿ m

1998-06-28

## 10 Index table

**10.1** All glyph images are available in size I as constant-strokewidth font and as letterpress font.

Only the images of the minimal alphanumeric subset (subset 1) are available in size III as constant-strokewidth font.

All images are available in size IV as constant-strokewidth font, with the exception of VERTICAL LINE.

**10.2** In the following table each image is given with the number of its reference drawing(s) and the subset(s) in which it is comprised.

The drawings are identified as follows :

- L: for letterpress font, size I
- C: for the constant-strokewidth font, size I
- III: for the constant-strokewidth font, size III

**10.3** As stated in 14.6, the shapes for size IV are derived from those of size I for the constant-strokewidth font (designated by C).

**10.4** Application advice is given in the column "Remarks".

It is recommended that prospective users of this standard consult manufacturers before deciding on a particular set of glyph images.

**10.5** The ordering of the glyph images in the table does not imply any kind of coding scheme.

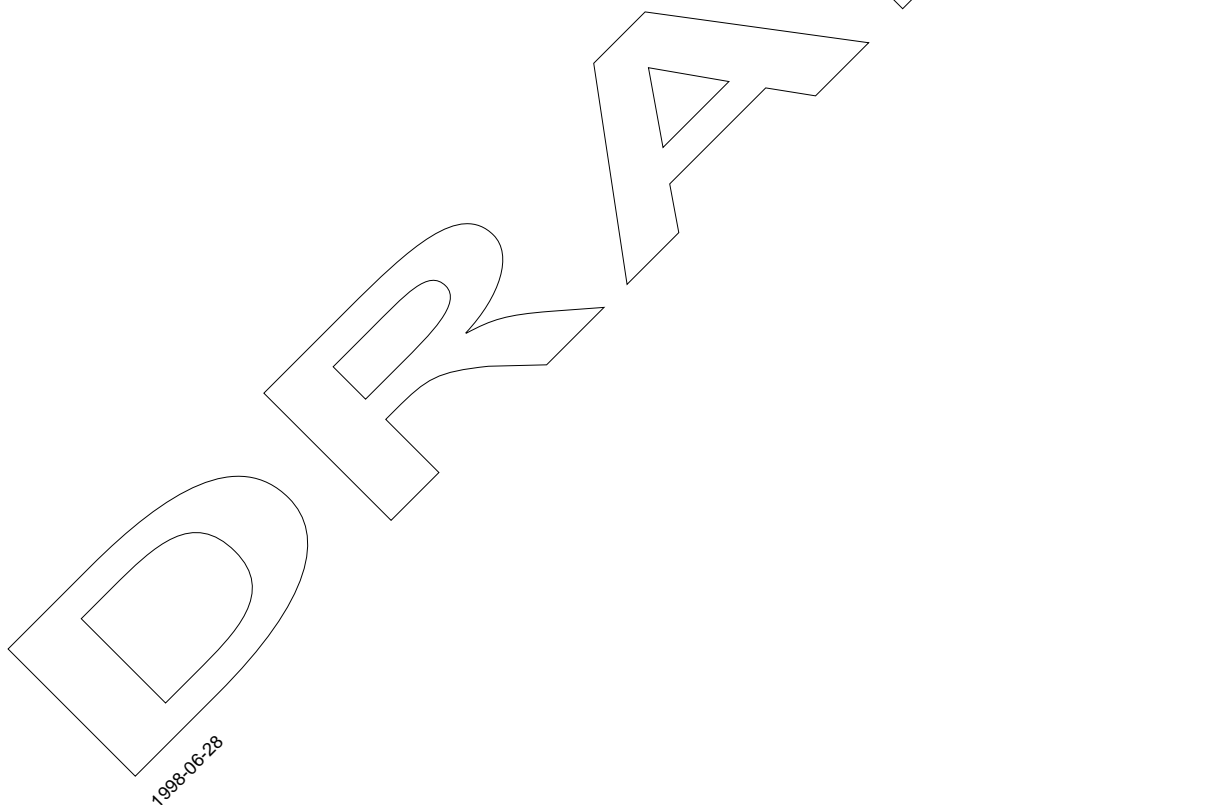


Table 2 – OCR-B glyph image set

Ref. No.	Shape	Drawing(s) No.	Name	Sets	Remarks
1	<b>1</b>	1 L, C, III	DIGIT ONE	1 2 3	
2	<b>2</b>	2 L, C, III	DIGIT TWO	1 2 3	
3	<b>3</b>	3 L, C, III	DIGIT THREE	1 2 3	
4	<b>4</b>	4 L, C, III	DIGIT FOUR	1 2 3	
5	<b>5</b>	5 L, C, III	DIGIT FIVE	1 2 3	
6	<b>6</b>	6 L, C, III	DIGIT SIX	1 2 3	
7	<b>7</b>	7 L, C, III	DIGIT SEVEN	1 2 3	
8	<b>8</b>	8 L, C, III	DIGIT EIGHT	1 2 3	
9	<b>9</b>	9 L, C, III	DIGIT NINE	1 2 3	
10	<b>0</b>	10 L, C, III	DIGIT ZERO	1 2 3	
11	<b>A</b>	11 L, C	LATIN CAPITAL LETTER A	2 3	
12	<b>B</b>	12 L, C	LATIN CAPITAL LETTER B	2 3	
13	<b>C</b>	13 L, C	LATIN CAPITAL LETTER C	1 2 3	
14	<b>D</b>	14 L, C	LATIN CAPITAL LETTER D	2 3	
15	<b>E</b>	15 L, C, III	LATIN CAPITAL LETTER E	1 2 3	

Table 2 (continued)

Ref. No.	Shape	Drawing(s) No.	Name	Sets	Remarks
16	<b>F</b>	16 L, C	LATIN CAPITAL LETTER F	2 3	
17	<b>G</b>	17 L, C	LATIN CAPITAL LETTER G	2 3	
18	<b>H</b>	18 L, C	LATIN CAPITAL LETTER H	2 3	
19	<b>I</b>	19 L, C	LATIN CAPITAL LETTER I	2 3	
20	<b>J</b>	20 L, C	LATIN CAPITAL LETTER J	2 3	
21	<b>K</b>	21 L, C	LATIN CAPITAL LETTER K	2 3	
22	<b>L</b>	22 L, C	LATIN CAPITAL LETTER L	2 3	
23	<b>M</b>	23 L, C	LATIN CAPITAL LETTER M	2 3	
24	<b>N</b>	24 L, C, III	LATIN CAPITAL LETTER N	1 2 3	
25	<b>O</b>	25 L, C	LATIN CAPITAL LETTER O	2 3	
26	<b>P</b>	26 L, C	LATIN CAPITAL LETTER P	2 3	
27	<b>Q</b>	27 L, C	LATIN CAPITAL LETTER Q	2 3	
28	<b>R</b>	28 L, C 1998-06-28	LATIN CAPITAL LETTER R	2 3	
29	<b>S</b>	29 L, C, III	LATIN CAPITAL LETTER S	1 2 3	
30	<b>T</b>	30 L, C, III	LATIN CAPITAL LETTER T	1 2 3	



Table 2 (continued)

Ref. No.	Shape	Drawing(s) No.	Name	Sets	Remarks
31	<b>U</b>	31 L, C	LATIN CAPITAL LETTER U	2 3	
32	<b>V</b>	32 L, C	LATIN CAPITAL LETTER V	2 3	
33	<b>W</b>	33 L, C	LATIN CAPITAL LETTER W	2 3	
34	<b>X</b>	34 L, C, III	LATIN CAPITAL LETTER X	1 2 3	
35	<b>Y</b>	35 L, C	LATIN CAPITAL LETTER Y	2 3	
36	<b>Z</b>	36 L, C, III	LATIN CAPITAL LETTER Z	1 2 3	
37	<b>a</b>	37 L, C	LATIN SMALL LETTER A	3	Smaller strokewidth; see clause 14
38	<b>b</b>	38 L, C	LATIN SMALL LETTER B	3	Smaller strokewidth; see clause 14
39	<b>c</b>	39 L, C	LATIN SMALL LETTER C	3	Smaller strokewidth; see clause 14
40	<b>d</b>	40 L, C	LATIN SMALL LETTER D	3	Smaller strokewidth; see clause 14
41	<b>e</b>	41 L, C	LATIN SMALL LETTER E	3	Smaller strokewidth; see clause 14
42	<b>f</b>	42 L, C	LATIN SMALL LETTER F	3	Smaller strokewidth; see clause 14
43	<b>g</b>	43 L, C 1998-06-28	LATIN SMALL LETTER G	3	Smaller strokewidth; see clause 14
44	<b>h</b>	44 L, C	LATIN SMALL LETTER H	3	Smaller strokewidth; see clause 14
45	<b>i</b>	45 L, C	LATIN SMALL LETTER I	3	Smaller strokewidth; see clause 14

Table 2 (continued)

Ref. No.	Shape	Drawing(s) No.	Name	Sets	Remarks
46	j	46 L, C	LATIN SMALL LETTER J	3	Smaller strokewidth; see clause 14
47	k	47 L, C	LATIN SMALL LETTER K	3	Smaller strokewidth; see clause 14
48	l	48 L, C	LATIN SMALL LETTER L	3	Smaller strokewidth; see clause 14
49	m	49 L, C	LATIN SMALL LETTER M	3	Smaller strokewidth; see clause 14
50	n	50 L, C	LATIN SMALL LETTER N	3	Smaller strokewidth; see clause 14
51	o	51 L, C	LATIN SMALL LETTER O	3	Smaller strokewidth; see clause 14
52	p	52 L, C	LATIN SMALL LETTER P	3	Smaller strokewidth; see clause 14
53	q	53 L, C	LATIN SMALL LETTER Q	3	Smaller strokewidth; see clause 14
54	r	54 L, C	LATIN SMALL LETTER R	3	Smaller strokewidth; see clause 14
55	s	55 L, C	LATIN SMALL LETTER S	3	Smaller strokewidth; see clause 14
56	t	56 L, C	LATIN SMALL LETTER T	3	Smaller strokewidth; see clause 14
57	u	57 L, C	LATIN SMALL LETTER U	3	Smaller strokewidth; see clause 14
58	v	58 L, C 1998-06-28	LATIN SMALL LETTER V	3	Smaller strokewidth; see clause 14
59	w	59 L, C	LATIN SMALL LETTER W	3	Smaller strokewidth; see clause 14
60	x	60 L, C	LATIN SMALL LETTER X	3	Smaller strokewidth; see clause 14

Table 2 (continued)

Ref. No.	Shape	Drawing(s) No.	Name	Sets	Remarks
61	y	61 L, C	LATIN SMALL LETTER Y	3	Smaller strokewidth; see clause 14
62	z	62 L, C	LATIN SMALL LETTER Z	3	Smaller strokewidth; see clause 14
63	*	63 L, C	ASTERISK	2 3	
64	+	64 L, C, III	PLUS SIGN	1 2 3	
65	-	65 L, C	HYPHEN - MINUS	2 3	
66	=	66 L, C	EQUALS SIGN	2 3	
67	/	67 L, C	SOLIDUS	2 3	
68	■	68 L, C	FULL STOP	2 3	
69	⁂	69 L, C	COMMA	2 3	Two vertical locations are specified, one of which projects below the base line for capital letters; see 14.5 and 14.8
70	⋮	70 L, C	COLON	3	
71	⋮	71 L, C	SEMICOLON	3	Two vertical locations are specified, one of which projects below the base line for capital letters; see 14.5 and 14.8
72	”	72 L, C	QUOTATION MARK	3	
73	’	73 L, C 1993-06-28	APOSTROPHE	3	
74	—	74 L, C	LOW LINE	3	Shall be used as a stand-alone character only, and shall not be printed under another character; see clause 12
75	?	75 L, C	QUESTION MARK	3	

Table 2 (continued)

Ref. No.	Shape	Drawing(s) No.	Name	Sets	Remarks
76	!	76 L, C	EXCLAMATION MARK	3	
77	(	77 L, C	LEFT PARENTHESIS	3	
78	)	78 L, C	RIGHT PARENTHESIS	3	
79	<	79 L, C, III	LESS-THAN SIGN	1 2 3	
80	>	80 L, C, III	GREATER-THAN SIGN	1 2 3	
81	[	81 L, C	LEFT SQUARE BRACKET	3	
82	]	82 L, C	RIGHT SQUARE BRACKET	3	
83	%	83 L, C	PERCENT SIGN	3	Smaller strokewidth; see clause 14
84	#	84 L, C	NUMBER SIGN	3	Smaller strokewidth; see clause 14
85	&	85 L, C	AMPERSAND	3	
86	@	86 L, C	COMMERCIAL AT	3	Smaller strokewidth; see clause 14
87	^	87 L, C	UP ARROWHEAD	3	
88	¤	88 L, C 1998-06-28	CURRENCY SIGN	3	
89	£	89 L, C	POUND SIGN	3	
90	\$	90 L, C	DOLLAR SIGN	3	

Table 2 (continued)

Ref. No.	Shape	Drawing(s) No.	Name	Sets	Remarks
91		91 L, C, III	VERTICAL LINE	1 2 3	
92					The glyph previously defined in this position (LONG VERTICAL MARK) has been unified with the preceding one (VERTICAL LINE)
93	\	93 L, C	REVERSE SOLIDUS	3	
94	Ä	94 L, C	LATIN CAPITAL LETTER A WITH DIAERESIS	4	
95	À	95 L, C	LATIN CAPITAL LETTER A WITH RING ABOVE	4	
96	Æ	96 L, C	LATIN CAPITAL LETTER AE	4	Shall also be used for LATIN CAPITAL LIGATURE AE
97	Ö	97 L, C	LATIN CAPITAL LETTER O WITH DIAERESIS	4	
98	Ø	98 L, C	LATIN CAPITAL LETTER O WITH STROKE	4	
99	Ü	99 L, C	LATIN CAPITAL LETTER U WITH DIAERESIS	4	
100	IJ	100 L, C	LATIN CAPITAL LIGATURE IJ	4	
101	Ñ	101 L, C	LATIN CAPITAL LETTER N WITH TILDE	4	
102	æ	102 L, C	LATIN SMALL LETTER A WITH RING ABOVE	4	Smaller strokewidth; see clause 14
103	æ	103 L, C 1998-06-28	LATIN SMALL LETTER AE	4	Shall also be used for LATIN SMALL LIGATURE AE Smaller strokewidth; see clause 14
104	ø	104 L, C	LATIN SMALL LETTER O WITH STROKE	4	Smaller strokewidth; see clause 14
105	ij	105 L, C	LATIN SMALL LIGATURE IJ	4	Smaller strokewidth; see clause 14

Table 2 (concluded)

Ref. No.	Shape	Drawing(s) No.	Name	Sets	Remarks
106	ß	106 L, C	LATIN SMALL LETTER SHARP S (German)	4	Smaller strokewidth; see clause 14
107	¨	107 L, C	DIAERESIS	4	For use see clauses 11 and 14
108	´	108 L, C	ACUTE ACCENT	4	For use see clauses 11 and 14
109	˘	109 L, C	GRAVE ACCENT	3	For use see clauses 11 and 14
110	ˆ	110 L, C	CIRCUMFLEX ACCENT	4	For use see clauses 11 and 14
111	˜	111 L, C	TILDE	3	For use see clauses 11 and 14
112	¸	112 L, C	CEDILLA	4	For use see clauses 11 and 14
113	{	113 L, C	LEFT CURLY BRACKET	3	
114	}	114 L, C	RIGHT CURLY BRACKET	3	
115	m	115 L, C	LATIN SMALL LETTER M (Alternative)	4	May be used in variable-pitch printing as a substitute for Ref. 49
116					The glyph previously defined in this position (CONTINUOUS UNDERLINE) has been deleted
117		No drawing	SPACE	1 2 3	SPACE is non-printing. For definition, see clause 13. Not all OCR readers will necessarily recognize SPACE
118	§	118 L, C 1998-06-28	SECTION SIGN	4	
119	¥	119 L, C	YEN SIGN	4	

**NOTE** – The glyphs previously defined with reference numbers 120 (CHARACTER ERASE) and 121 (GROUP ERASE) have been deleted.

## 11 Use of diacritical marks

**11.1** A number of diacritical marks are provided which have been designed and positioned in such a way that they can be combined with small letters to create national letters. These marks are:

DIAERESIS  
ACUTE ACCENT  
GRAVE ACCENT  
CIRCUMFLEX ACCENT  
TILDE  
CEDILLA

These diacritical marks may also be used free-standing.

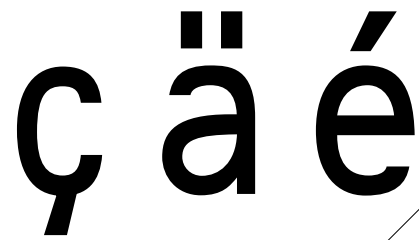
**11.2** The method for printing a composite glyph image, as well as the principle for deciding the coding of a recognized composite image, will be application-dependent, and is outside the scope of this standard. In particular, the naming of the diacritical marks in this standard, although corresponding to the names for free-standing marks in ISO/IEC 10646-1:1993, does not imply any specific method for combination.

### NOTES

1 For printing applications, a letter could be followed by a backspace operation followed by a diacritical mark; or the opposite sequence could be output; or the diacritical mark could be made non-spacing and followed by the letter; or the combined letter could be composed in the printing mechanism and output in a single operation. Likewise, for OCR applications a composite glyph image recognized in a reading operation could be represented in the system as a single code assigned to the corresponding character; as a "composite graphic character" (ISO/IEC 646:1991); as an "accented letter" (ISO/IEC 6937:1994); as a "composite sequence" (ISO/IEC 10646-1:1993); or by any other method defined in standards and/or specified for the particular application.

2 A national letter created by combining a diacritical mark with one of the letters a-z is not necessarily a variant of that specific letter. For instance, in the scripts of some languages the character ä is a separate letter of the alphabet, distinct from and unrelated to the letter a.

**11.3** The relative position of the diacritical mark and of the letter is obtained by superimposing the horizontal and vertical axes of the two glyph images concerned. Examples:



**Figure 2 – Examples of composite glyph images**

For OCR purposes, the superposition of a diacritical mark on the image of a basic letter must be done very accurately. Prospective users should consult manufacturers before planning inclusion of such letters in OCR glyph image sets.

**11.4** Subset 4 glyph images shall not be combined with diacritical marks to produce composite images. For the subset 3 glyph images the following restrictions apply:

- SMALL LETTER I may be combined only with CEDILLA;
- SMALL LETTER J shall not be combined with any diacritical mark;
- SMALL LETTER G and SMALL LETTER Y shall not be combined with CEDILLA;
- No letter shall be combined with multiple diacritical marks.

Apart from these rules, this International Standard imposes no restrictions on letter-mark combinations. Such restrictions should therefore be specified, as necessary, for any particular printing and/or OCR application based on OCR-B images; or restrictions could be implied by definition of a character set for the application, through reference to other International Standards or otherwise.

The validity of any specific combination of diacritical mark and letter, as well as of any free-standing diacritical mark, is application-dependent, and outside the scope of this standard.

**NOTE** – Not all letter-mark combinations permitted by the rules above will be valid national letters.

## 12 Use of the low line glyph

The glyph image LOW LINE shall be used in OCR applications free-standing only, and shall not be combined with (i.e. printed under) another image.

DH\_1925

Figure 3 – Example of use of LOW LINE

## 13 SPACE

The SPACE is an intentionally blank position in a line of printing. With constant-pitch printing, its nominal width is equal to the printing pitch (for example, 2,54 mm if the glyph images are printed 10 per 25,4 mm). With variable-pitch printing, its nominal width is equal to the largest glyph image width available.

## 14 Glyph image shape definition

### 14.1 Reference drawings

The shapes and dimensions of the OCR-B glyphs for both the letterpress and the constant-strokewidth fonts are specified by drawings for size I and III.

The characters are drawn at scale 100:1 on a 2 mm square grid. The total grid measures 280 mm × 380 mm. For the purpose of illustration, some of these drawings are reproduced in this standard at reduced size.

Grid readings should be made only from drawings on stable material. Photographic reproductions of drawings printed on paper are not satisfactory for this purpose – the dimensional stability of paper is not sufficient.

Points on the reference drawing can certainly be determined with an accuracy of half a square (10 µm at full size), and if desired one quarter of a square (5 µm at full size) should be possible.

### 14.2 Availability of duplicates and glyph shape representations

Original drawings on a stable base at 100:1 scale with the 280 mm × 380 mm grid exist in the following sets:

- OD1 Letterpress font, size I.
- OD2 Letterpress font, size I with the grid removed over approximately 2 mm around the glyph image outline. This set is particularly suitable for photographic reduction.
- OD3 Constant-strokewidth font, size I.
- OD4 Constant-strokewidth font, size III.

**NOTE** – Duplicates of the drawings could previously be ordered from ECMA and from the US National Bureau of Standards (later NIST). The drawings are at present (June 1998) archived by the Japanese Industrial Standards Committee. Future handling of this matter is under consideration.

### 14.3 Type dimensions

Attention is called to the fact that this standard specifies the shapes and dimensions of the nominal printed images. These dimensions are not necessarily the same as those used in the actual image generation process. Corrections for systematic effects occurring in the printing could be needed.

Details on dimensional tolerances etc. for the printed images are specified in standard ISO 1831.

### 14.4 Constant-strokewidth font, size I

**14.4.1** The nominal printed image of each glyph image is defined by its centreline and by its nominal strokewidth. The nominal strokewidth is:

0,35 mm (0,014 in) for most of the images

0,31 mm (0,012 in) for all small letters and the three images #, % and @.

The centreline and preferred line endings and corners are given in drawings marked "C". Pointers establish the vertical position (base line) and the orientation. Another pointer establishes the horizontal position for fixed-pitch printing.

The reference drawings for the COMMA and the SEMICOLON contain pointers to indicate two alternative vertical positions (base lines). Either position can be selected for the two glyph images for a specific font implementation, depending on the intended use of the font.

**14.4.2** A special effort should be made in type design and manufacturing to arrive at actual print that conforms as closely as possible to the given line endings and corners. This is especially important for the square corners of capital letters B and D.



**14.4.3** A pointer is provided to produce the most aesthetic spacing of glyph images in a line of printing. However, on printers having a significant horizontal spacing tolerance it is recommended to use the geometric centreline of the image instead of the line defined by the pointer where necessary to achieve an acceptable image separation.

#### 14.5 Constant-strokewidth font, size III

The nominal printed image of each glyph image is given by its centreline and by its nominal strokewidth. The nominal strokewidth is 0,38 mm (0,015 in). The 20 reference drawings for 0 1 2 3 4 5 6 7 8 9 C E N S T X Z < + > are marked "III" and include pointers. Subclauses 14.4.2 and 14.4.3 also apply.

#### 14.6 Constant-strokewidth font, size IV

The nominal printed image of each glyph image is given by its centreline and by its nominal strokewidth. The size IV centreline is derived from the corresponding size I centreline (see 14.4 and reference drawings marked "C") by a linear magnification of 1,5. For example, an image centreline width of 2,40 mm becomes  $1,5 \times 2,40 \text{ mm} = 3,60 \text{ mm}$  in size IV, and so on.

The nominal strokewidth is:

0,50 mm (0,020 in) for most of the glyph images.

0,44 mm (0,017 in) for all small letters and the three images #, % and @.

Preferred line endings and corners cannot be accurately arrived at by a 1,5 magnification since the ratio of nominal strokewidths for size IV and I is not exactly 1,5. However, given a 1,5 magnification of the size I drawing, the nominal size IV constant-strokewidth image can easily be constructed.

#### 14.7 Letterpress font, size I

The nominal printed image of each glyph image is drawn on a reference grid (see 14.1) to allow readings with any desired accuracy from drawings marked "L". Pointers establish the vertical position (base line), the orientation and the body width. A pointer also establishes the horizontal position for fixed-pitch printing:

The reference drawings for the COMMA and the SEMICOLON contain pointers to indicate two alternative vertical positions (base lines). Either position can be selected for the two glyph images for a specific implementation, depending on the intended use of the font.

The glyph images of the letterpress font are designed with minor strokewidth variations. However, strokewidths are always close to the nominal value of 0,35 mm (0,014 in) for digits and capital letters, and of 0,31 mm (0,012 in) for small letters and the three images #, % and @.

### 15 Printing the letterpress and constant-strokewidth fonts

In order to print the letterpress font and to achieve the most satisfactory appearance, the printing device should be able to print sharp corners and to keep the strokewidth variations under close control. These features are not required for printing the constant-strokewidth fonts, although a special effort should be made to produce sharp corners in the capital letters B and D.

There may well be printing equipment in which the accuracy of strokewidth control is intermediate between that required in letterpress quality and that provided by, for example, high-speed printers. It is at the discretion of the manufacturers of such printing equipment to design their type so that the printed images incorporate as many as practicable of the strokewidth variations which contribute to the aesthetically satisfactory appearance of the letterpress shapes.

Care should be taken that the printed image strokes are symmetrically distributed around the centrelines as specified in this document.

### 16 Illustration of OCR-B

Figure 4 shows the complete OCR-B set in size I at scales 4:1 and 1:1.

The drawings reproduced on the following pages show:

- DIGIT ONE, LATIN CAPITAL LETTER E, SECTION SIGN and YEN SIGN in size I as letterpress font and as constant-strokewidth font;
- DIGIT ONE and LATIN CAPITAL LETTER E in size III as constant-strokewidth font.

These reproductions of the original drawings are at approximately 70% of full scale.

0 1 2 3 4 5 6 7 8 9  
 A B C D E F G H I J K L M  
 N O P Q R S T U V W X Y Z  
 a b c d e f g h i j k l m  
 n o p q r s t u v w x y z  
 \* + - = / . , : ; " ' \_  
 ? ! ( ) < > [ ] % # & @ ^  
 ¤ £ \$ ¥ § | \ { } m  
 Ä Å Æ Ĩ Ñ Ö Ø Ù à æ ij ø ß  
 / \ ^ " ~  
 (SPACE)

0123456789  
 ABCDEFGHIJKLM  
 NOPQRSTUVWXYZ  
 abcdefghijklm  
 nopqrstuvwxyz  
 \*+ = / . , : ; " ' \_  
 ? ! ( ) < > [ ] % # & @ ^  
 ¤ £ \$ ¥ § | \ { } m  
 Ä Å Æ Ĩ Ñ Ö Ø Ù à æ ij ø ß  
 / \ ^ " ~

Figure 4 – Illustration of letterpress font size I, scales 4:1 and 1:1

1998-06-28

*Pages 23-41 from ISO 1073/II-1976 to be inserted here*

**DRAFT**

1998-06-28

*Text on this page to be included as Annex A (informative) to standard*

## Notes on the implementation of OCR-B

The design of the OCR-B font is based on fundamental aesthetic principles which, as far as feasible, correspond to the criteria emerging from the long development of our classic typography. One of the essential principles prescribes that in a letter design all vertical parts must be heavier than the horizontal parts. This is also true for so-called sans serif characters, that is for a design which at first sight has a thread-like appearance. This is precisely the case for OCR-B.

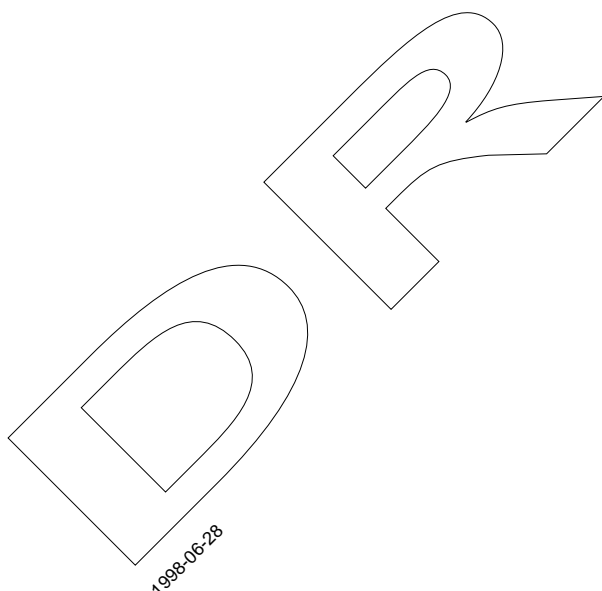
The OCR-B can be implemented in two clearly different forms. It can be used as a font with constant-strokewidth as well as a letterpress font. Type generation can be based on either implementation.

For printing devices like high-speed mechanical printers and similar machines, the centreline is the skeleton along which a stroke of prescribed width is placed. For engraving, it is recommended to use a tool the diameter of which is equal to the strokewidth. The resulting engraving is completely

thread-like, all strokes having an equal width. The aesthetic appearance as well as readability are partly diminished by this process.

For printing devices using other technologies, e.g. laser and ink-jet printers, the design of the font will depend on the resolution available in the device. Each font manufacturer is, of course, free to select either the constant-strokewidth or the letterpress font as a basis for the design.

The imaging resolution for some common types of printers will not permit reproduction of the finer details of the OCR-B letterpress font, like its varying strokewidth. It is however recommended that other properties of the letterpress font that can be achieved, notably the sharp angles at the end of the strokes, are implemented as far as possible. The font will then look less mechanical and bear more resemblance to the forms of traditional typography to which the human eye has been accustomed for centuries.



*Text on this page to be included as Annex B (informative) to standard*

**Main differences between ISO 1073/II-1976 and this first edition  
of this part 2 of ISO/IEC 1073**

*(To be prepared when standard text finalized; see also Annex B on following page)*

**DRAFT**

1998-06-28

## Annex B

### Main extensions in ISO/IEC CD 1073-2.3 as compared to Annex A (pages 5–25 of this document)

**B.1** Letter combinations with diacritical marks is allowed not only for small letters but also for capital letters, thereby providing glyph representations for all Latin letters in the present parts of ISO/IEC 8859. This is described in clause 1 and combination rules given in clauses 11 and 14.

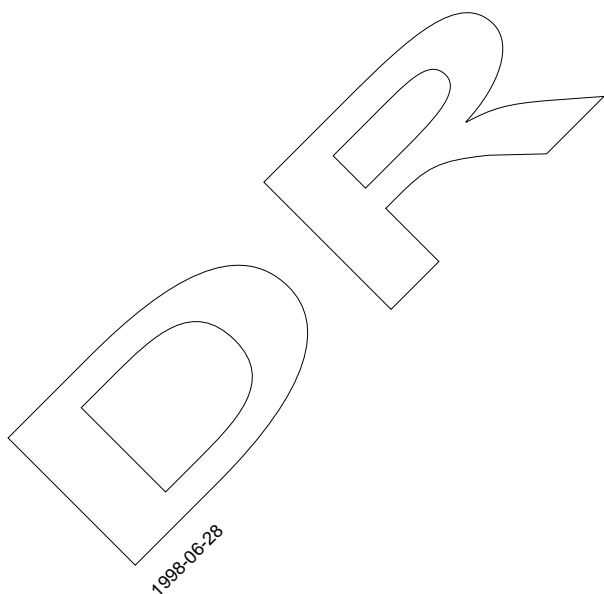
**B.2** The combination with capital letters increases the maximum height of OCR-B glyphs. This is specified in clause 8 (figure 1 and table 1).

**B.3** The OCR-B repertoire is extended with 40 new glyphs, defined in letterpress style only. Information relating to this extension is given in clauses 6, 7, 9, 10, 11 and 16 (figure 4). The requested new glyphs are listed in Annex C.

**B.4** The extension is intended to cover also the Greek capital alphabet, containing the ten Greek capital letters that have no corresponding glyphs in the Latin alphabet (gamma, delta, theta, lamda, xi, pi, sigma, phi, psi, omega). For the other "multiple-use glyphs" – like Greek capital alpha / latin A – rules are given in clauses 5 and 10 (table 2).

**B.5** Addition of new glyphs and the new combination rules for capital letters means that the precomposed letters ÅÄŒŮå in the original standard will have alternative combined representations. Rules for this situation are given in clauses 9 and 10 (table 2).

**B.6** The added glyphs are to be defined by outline shape definitions according to ISO/IEC 9541-3, not by reference drawings. That standard is therefore referenced in clause 3, and further information given in clauses 6, 8, and 14.



## Annex C

### OCR-B glyph repertoire extensions requested by National Bodies

In the processing of the OCR-B revision the following new glyphs have been requested (named here as the corresponding characters according to ISO/IEC 10646-1:1993):

#### Latin precomposed and other national letters:

LATIN CAPITAL LETTER A WITH OGONEK  
 LATIN CAPITAL LETTER D WITH STROKE  
 LATIN CAPITAL LETTER ETH  
 LATIN CAPITAL LETTER E WITH OGONEK  
 LATIN CAPITAL LETTER H WITH STROKE  
 LATIN CAPITAL LETTER L WITH STROKE  
 LATIN CAPITAL LETTER T WITH STROKE  
 LATIN CAPITAL LETTER THORN  
 LATIN CAPITAL LETTER ENG  
 LATIN CAPITAL LIGATURE OE  
 LATIN SMALL LETTER A WITH OGONEK  
 LATIN SMALL LETTER D WITH STROKE  
 LATIN SMALL LETTER G WITH CEDILLA  
 LATIN SMALL LETTER H WITH STROKE  
 LATIN SMALL LETTER DOTLESS I  
 LATIN SMALL LETTER J WITH CIRCUMFLEX ACCENT  
 LATIN SMALL LETTER L WITH STROKE  
 LATIN SMALL LETTER R WITH CEDILLA  
 LATIN SMALL LETTER T WITH STROKE  
 LATIN SMALL LETTER U WITH OGONEK  
 LATIN SMALL LETTER ETH  
 LATIN SMALL LETTER THORN  
 LATIN SMALL LETTER ENG  
 LATIN SMALL LIGATURE OE

**NOTE** – It is intended to unify the glyphs for CAPITAL LETTER D WITH STROKE and CAPITAL LETTER ETH.

#### Greek letters:

GREEK CAPITAL LETTER GAMMA  
 GREEK CAPITAL LETTER DELTA  
 GREEK CAPITAL LETTER THETA  
 GREEK CAPITAL LETTER LAMDA  
 GREEK CAPITAL LETTER XI  
 GREEK CAPITAL LETTER PI  
 GREEK CAPITAL LETTER SIGMA  
 GREEK CAPITAL LETTER PHI  
 GREEK CAPITAL LETTER PSI  
 GREEK CAPITAL LETTER OMEGA

#### Diacritical marks:

MACRON  
 BREVE  
 CARON  
 DOUBLE ACUTE ACCENT  
 DOT ABOVE  
 RING ABOVE  
 OGONEK

#### NOTES

1 Additional single and multiple diacritical marks specific to Vietnamese printing have been requested in ballot comments. These requests were not accommodated in the third CD, partly because of OCR-B design difficulties, and partly because no Vietnamese part of ISO/IEC 8859 existed (one of the purposes of the revision was to extend the OCR-B repertoire to cover existing Latin parts of ISO/IEC 8859).

2 Romanian COMMA BELOW has also been requested. This was considered impractical from an OCR design point of view. A note about possible use of the glyph for the CEDILLA, similar to the note in ISO/IEC 8859-2, should be included in the standard if the revision is resumed.

#### Other characters:

The EURO SIGN has been requested by CEN/TC304.

## Annex D

## Sample glyph definition according to ISO/IEC 9541-3

DEFINITION OF OCR-B CHARACTER "SMALL LIGATURE OE" (ISO/IEC 9541-3 SYNTAX)

```
84 135 594 0 rpe
0 78 vstem 174 78 vstem 348 78 vstem -5 69 hstem 207 69 hstem 397 69 hstem
```

```
0 hmoveto
-80 28 -60 90 vhcurveto 55 0 30 30 10 25 rrcurveto
10 -25 30 -30 55 0 rrcurveto 90 31 55 81 hvcurveto -78 hlineto
-40 -12 -27 -35 vhcurveto -35 -17 20 41 hvcurveto 82 vlineto 174 hlineto
118 vlineto 81 -28 60 -90 vhcurveto -55 0 -30 -24 -10 -27 rrcurveto
-10 27 -30 24 -55 0 rrcurveto -90 -28 -60 -81 hvcurveto
closepath
```

```
78 202 rmoveto
35 18 25 30 vhcurveto 30 18 -25 -35 hvcurveto -213 vlineto
-35 -18 -25 -30 vhcurveto -30 -18 25 35 hvcurveto
closepath
```

```
174 0 rmoveto
35 18 25 30 vhcurveto 30 18 -25 -35 hvcurveto -61 vlineto -96 hlineto
closepath
```

endglyph

THE DEFINITION ABOVE WILL PRODUCE THE FOLLOWING GLYPH:

