

INTERNATIONAL STANDARD

ISO/IEC
4873

Third edition
1991-12-15

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

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



Reference number
ISO/IEC 4873:1991(E)



Documents
Document Sharing Hub

Contents

	Page
1 Scope	1
2 Conformance and implementation	1
2.1 Conformance	1
2.1.1 Conformance of information interchange	1
2.1.2 Conformance of devices	1
2.2 Implementation	2
3 Normative references	2
4 Definitions	2
4.1 active position	2
4.2 bit combination	2
4.3 byte 2	
4.4 character	2
4.5 character position	2
4.6 coded-character-data-element (CC-data-element)	2
4.7 coded character set	3
4.8 code extension	3
4.9 code table	3
4.10 control character	3
4.11 control function	3
4.12 device	3
4.13 escape sequence	3
4.14 Final Byte	3
4.15 graphic character	3
4.16 graphic symbol	3
4.17 repertoire	3
4.18 user 3	
5 Notation, code table and names	3
5.1 Notation	3
5.2 Code table	4
5.3 Names	4
6 Structure of the 8-bit code	4
6.1 Elements of the 8-bit code	4
6.2 Identification of the elements of the 8-bit code	5

© ISO/IEC 1991

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

ISO/IEC Copyright Office • Case postale 56 • CH-1211 Genève 20 • Switzerland

Printed in Switzerland

6.3 Invocation	5
6.3.1 C0 set	
6.3.2 Character SPACE	5
6.3.3 G0 set	5
6.3.4 Character DELETE	5
6.3.5 C1 set	
6.3.6 G1 set	5
6.3.7 G2 set	5
6.3.8 G3 set	5
7 Specification of the characters of the 8-bit code	5
7.1 C0 set	6
7.2 Character ESCAPE	6
7.3 Character SPACE	6
7.4 G0 set	6
7.5 Character DELETE	9
7.6 C1 set	9
7.7 G1 set	9
7.8 G2 set	9
7.9 G3 set	10
7.10 Summary of the specification of the 8-bit code	10
8 Levels	10
8.1 Level 1	10
8.2 Level 2	10
8.3 Level 3	11
9 Version of the 8-bit code	11
9.1 Contents of a version	11
9.2 Unique coding of characters	11
10 Identification of version and level	11
10.1 Purpose and context of identification	11
10.2 Identification of level	12
10.3 Identification of a version	12
10.4 Switching from one version to another	12
10.5 Switching from one level to another	12
Annexes	
A Restrictions applicable to the C0 and C1 sets	16
B Shift functions	17
C Composite graphic characters	18

D	Use of bit combinations 00/14 and 00/15	19
E	Main differences between the 2nd edition (1986) and the present (third) edition of this International Standard	19

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. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

International Standard ISO/IEC 4873 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*.

This third edition cancels and replaces the second edition (ISO 4873 : 1986), which has been technically revised.

Annex A forms an integral part of this International Standard. Annexes B, C, D and E are for information only.

This page intentionally left blank

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

1 Scope

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

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

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

This code is primarily intended for general information interchange within an 8-bit environment among data processing systems and associated equipment, and within data communication systems. The need for graphic characters and control functions in data processing has also been taken into account.

The code includes the 10 digits as well as the 52 small and capital letters of the basic Latin alphabet and may include accented letters, special Latin letters and/or the letters of one or several non-Latin alphabet(s).

2 Conformance and implementation

2.1 Conformance

2.1.1 Conformance of information interchange

A coded-character-data-element (CC-data-element) within coded information for interchange is in conformance with a version of this International Standard if all the coded representations of characters within that CC-data-element conform to the requirements of clause 9.

A claim of conformance shall identify the version adopted.

2.1.2 Conformance of devices

A device is in conformance with this International Standard if it conforms to the requirements of 2.1.2.1, and either or both of 2.1.2.2 and 2.1.2.3. A claim of conformance shall identify the document which contains the description specified in 2.1.2.1, and shall identify the version adopted.

2.1.2.1 Device description

A device that conforms to this International Standard shall be the subject of a description that identifies the means by which the user may supply characters to the device, or may recognize them when they are made available to him, as specified respectively in 2.1.2.2 and 2.1.2.3.

2.1.2.2 Originating devices

An originating device shall allow its user to supply any sequence of characters from the version adopted, and shall be capable of transmitting their coded representations within a CC-data-element.

2.1.2.3 Receiving devices

A receiving device shall be capable of receiving and interpreting any coded representations of characters that are within a CC-data-element, and that conform to 2.1.1, and shall make the corresponding characters available to its user in such a way that the user can identify them from among those of the version adopted, and can distinguish them from each other.

2.2 Implementation

The use of this code requires definitions of its implementation in various media. For example, these could include punched tapes, punched cards, magnetic and optical media and transmission channels, thus permitting interchange of data to take place either indirectly by means of an intermediate recording in a physical medium, or by local connection of various units (such as input and output devices and computers) or by means of data transmission equipment.

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

3 Normative references

The following standards 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 listed below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO/IEC 646: 1991 *Information technology - ISO 7-bit coded character set for information interchange*.

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

ISO/IEC 6429:¹⁾ *Information technology - Control functions for 7-bit and 8-bit coded character sets*.

ISO/IEC 10367: 1991 *Information technology - Standardized coded graphic character sets for use in 8-bit codes*.

ISO International Register of Coded Character Sets to be Used with Escape Sequences (ISO 2375).

4 Definitions

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

4.1 active position : The character position which is to image the graphic symbol representing the next graphic character or relative to which the next control function is to be executed.

NOTE - In general, the active position is indicated in a display by a cursor.

4.2 bit combination : An ordered set of bits used for the representation of characters.

4.3 byte : A bit string that is operated upon as a unit.

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

4.5 character position : The portion of a display that is imaging or is capable of imaging a graphic symbol.

4.6 coded-character-data-element (CC-data-element) : An element of interchanged information that is specified to consist of a sequence of coded representations of characters, in accordance with one or more identified standards for coded character sets.

1) To be published.

NOTES

1. In a communication environment according to the Reference Model for Open Systems Interconnection (ISO 7498), a CC-data-element will form all or part of the information that corresponds to the Presentation-Protocol-Data-Unit (PPDU) defined in that International Standard.

2. When information interchange is accomplished by means of interchangeable media, a CC-data-element will form all or part of the information that corresponds to the user data, and not that recorded during formatting and initialization.

4.7 coded character set; code : A set of unambiguous rules that establishes a character set and the one-to-one relationship between the characters of the set and their bit combinations.

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

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

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

4.11 control function : An action that affects the recording, processing, transmission, or interpretation of data, and that has a coded representation consisting of one or more bit combinations.

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

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

4.13 escape sequence : A string of bit combinations that is used for control purposes in code extension procedures. The first of these bit combinations represents the control function ESCAPE.

4.14 Final Byte : The bit combination that terminates an escape sequence or a control sequence.

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

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

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

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

NOTES

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

2. The characters, as supplied by the user or made available to him, may be in the form of codes local to the device, or of non-conventional visible representations, provided that 2.1.2 above is satisfied.

5 Notation, code table and names

5.1 Notation

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

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