Technical Report

ISO/IEC TR 2375

First edition
2024-07

Information technology – Registered escape sequences and coded character sets

Technologies de l'information — Séquences d'échappement et jeux de caractères codés enregistrés

Contents

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 2, *Coded character sets*.

This first edition cancels and replaces ISO/IEC 2375:2003, which has been technically revised.

The main changes are as follows:

— removal of the registration procedure specification due to the cancellation of the Registration Authority;

— inclusion of the registered data published by the former Registration Authority.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

ISO/IEC 2022 uses the escape sequence to control the code extension procedures. The International Register defined by ISO/IEC 2375 (previous edition of this document) provided the escape sequences and character sets.

Since no new escape sequences have been registered since 2004, and no new organization was found to take over the registration authority as of 2020, ISO/IEC JTC1 / SC2 decided to replace ISO/IEC 2375 with ISO/IEC TR 2375 (this document), which contains the data provided by the former registration authority, to ensure interoperability with archival digital data encoded by ISO/IEC 2022. This document and electronic attachments replace the International Register.

Information technology – Registered escape sequences and coded character sets

1 Scope

This document provides the escape sequences and coded character sets that were registered and published by the former registration authority in ISO/IEC 2375 (the previous edition of this document).

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
bit combination
ordered set of bits used for the representation of characters

3.2
byte
bit string that is operated upon as a unit

3.3
catalogue
list of the registrations and supplementary data, with reference to their locations

3.4
character
member of a set of elements used for the organization, control, or representation of data

3.5
coded character set
set of unambiguous rules that establishes a character set and the relationship between the characters of the set and their coded representation

3.6
code position
part of a code table identified by its column and row coordinates

3.7
code table
table showing the characters allocated to each bit combination in a code

3.8
combining character
member of an identified subset of the coded character set intended for combination either

a) with the preceding non-combining graphic character, or with a sequence of combining characters preceded by a non-combining character (as presented in ISO/IEC 10646), or

b) with the following non-combining graphic character, or with a sequence of combining characters followed by a non-combining character (as presented in ISO/IEC 6937)

3.9
combining sequence
sequence of graphic characters consisting of either

a) a non-combining character followed by one or more combining characters (as presented in ISO/IEC 10646), or

b) a non-combining character preceded by one or more combining characters (as presented in ISO/IEC 6937)

3.10
control function
action that affects the recording, processing, transmission, or interpretation of data, and that has a coded representation consisting of one or more bit-combinations

3.11
escape sequence
string of bit combinations that is used for control purposes in code extension procedures

Note 1 to entry: The first of these bit combinations represents the control function ESCAPE.

3.12
ESC F$_s$ sequence
escape sequence with the second bit combination in the range 6/0 to 7/14

Note 1 to entry: ESC Fs sequences are used for the standardized single control functions.

3.13
former registration authority
organization designated by ISO that ensured the maintenance of the registry of the escape sequence, character set, and the mapping tables defined by ISO/IEC 2375:2003 (the previous edition of this document)

3.14
graphic character
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

3.15
International Register
register of the coded character sets and the escape sequences in ISO/IEC 2022

Note 1 to entry: In this document, this International Register means the archive of the former International Register (from ISO/IEC 2375, now withdrawn and for which the registration mechanism is no longer available), which is now provided in electronic attachment in this document.

3.16
octet
ordered sequence of eight bits considered as a unit

3.17
repertoire
specified set of characters that are each represented by one or more bit-combinations of a coded character set

Note 1 to entry: A registration does not specify the repertoire of the sequences obtained by combining the characters (see A.3).

3.18
standard return
escape sequence to switch the coding system to the ISO/IEC 646 character set under the ISO/IEC 2022 coding system, i.e. "ESC 2/5 4/0"

Note 1 to entry: Standard return fits the DESIGNATE OTHER CODING SYSTEM specified in ISO/IEC 2022:1994, 15.4.

4   International Register

4.1   Purpose of the International Register

The International Register provides the catalogue of the registered character set standards with their registration numbers, the names of the character set standards, and their final bytes (for the definition of final bytes, see ISO/IEC 2022:1994, 4.14). The catalogue includes hyperlinks to the electronic document for each registration, which is preserved in the International Register.

Registration of the graphic character sets and the control character sets consists of three parts: a cover page, a code table, and a list of character names. Registration of the coding systems not in conformance with ISO/IEC 2022 has a cover page, and a description text of the coding system, but can lack both a code table and a list of character names.

4.2   Location of the International Register

The International Register can be found at

https://standards.iso.org/iso-iec/tr/2375/ed-1/en

4.3   Machine readability

The electronic data in the International Register were primarily produced for printing, while some data were based on scanned images which were image data and not machine-readable text. In the case that registration has the mapping table to ISO/IEC 10646, it is in machine-readable format (see an example in Annex E).

4.4   Catalogue

4.4.1   Structure of the catalogue

The catalogue classifies the registered character set standards into nine types under three main categories. The top category is determined by the types of the character defined by ISO/IEC 2022 the graphic character (see ISO/IEC 2022:1994, 4.15), the control character (see ISO/IEC 2022:1994, 4.9), and the coding systems not in conformance with ISO/IEC 2022 (see B.1).

a)   Graphic character sets

ISO/IEC 2022 specifies three types of coded graphic character sets (see ISO/IEC 2022:1994, 6.3.1):

—   "94-character set";

—   "96-character set";

—   "multiple-byte set".

The International Register has two subtypes under "94-character graphic character set", which are distinguished by the length of their intermediate bytes (see ISO/IEC 2022:1994, 4.14):

—   "94-character graphic character sets" with one intermediate byte;

—   "94-character graphic character sets" with two intermediate bytes.

Hence, the graphic character sets in the International Register fall into four types:

— 94-character graphic character sets with one intermediate byte;

— 94-character graphic character sets with two intermediate bytes;

— 96-character graphic character sets;

> NOTE 1   The escape sequences for this type use one intermediate byte.

— multiple byte graphic character sets.

> NOTE 2   The escape sequences for this type use two intermediate bytes.

b)   Control character sets

ISO/IEC 2022 specifies two types of "sets of control functions":

— the primary sets of coded control functions (C0) (see ISO/IEC 2022:1994, 6.4.1 and 6.4.4);

— the supplementary sets of coded control functions (C1) (see ISO/IEC 2022:1994, 6.4.2 and 6.4.4);

In addition to these two types, ISO/IEC 2022 specifies the "coded single additional control functions" (see ISO/IEC 2022:1994, 6.5). Hence, the International Register has three types of character sets for C0, C1, and single control functions. The former registration authority considered the criteria in Annex C for the registration of ESC Fs (standardized single control function, defined in ISO/IEC 2022:1994, 13.2.1).

— C0-Control Character Sets

— C1-Control Character Sets

— Single Control Functions

> NOTE   The escape sequences for all three of these types use no intermediate bytes.

c)   Coding Systems not in conformance with ISO/IEC 2022

The International Register includes some registrations of coding systems not in conformance with ISO/IEC 2022. The registrations are classified into two types according to whether or not the coding system can be switched to the ISO/IEC 2022 coding system by standard return:

— Coding systems with standard return;

— Coding systems without standard return.

NOTE   The escape sequences for these two types use two intermediate bytes.

4.4.2   Registration number

A registration number was assigned by the former registration authority to each application for registration in the form of a single positive integer (e.g. 1, 2, 4) or two positive integers concatenated by a hyphen (e.g. 8-1, 8-2) in the case that the registration consists of multiple sets (see B.3).

In the case that an application was withdrawn or rejected, the registration number was not reused for other applications. Consequently, there are some numbers in the International Register that were never registered (e.g. 210 to 225).

4.5   Contents in a registration

4.5.1   Copyright of the content

In some registrations, the code table or the list of character names were reproduced from copyrighted material. The registrations in the archive are reproduced from the former registration authority, which obtained the copyrights to reproduce them through the sponsoring authority.

4.5.2   Cover page

The cover page contains the following information:

—   the type of the character set (4.4);

—   the registration number (4.4.2);

—   the date of registration;

—   the escape sequences (4.5.4);

—   the name of the character set;

—   the description of the character set;

—   the sponsoring authority (4.5.5);

—   the origin of the character set (4.5.6).

In addition, the cover page can contain the following information:

—   the owner of origin (4.5.7);

—   the copyright owner (4.5.8);

—   the field of utilization.

Whether each information was provided by the sponsoring authority or the former registration authority is described in A.1.

4.5.3   Format of registration number on the cover page

On the cover page, the registration number is shown by three digits with zero padding. In the case that the registration number consists of two numbers, the first major number is shown by three digits with zero padding, but the second number is shown without zero padding.

4.5.4   Escape sequences

The escape sequences were assigned by the former registration authority. After the assignment of the escape sequence, it was never reallocated for another registration, even when the original registration was withdrawn.

4.5.5   Sponsoring authority

The sponsoring authority was the organization that submitted applications concerning the meanings of escape sequences to the former registration authority. The sponsoring authority obtained copyright permission from the copyright owner so that the former registration authority reproduced the publication that specified the coded character set in the International Register. If the registration was intended for a particular application, the sponsoring authority obtained the endorsement of the developer of that application to register the coded character set.

4.5.6   Origin

The origin is a name for a coded character standard of the registration (e.g. an ISO, IEC, or ITU standard, a national standard, or a character set defined by an organization).

4.5.7   Owner of origin

The owner of origin was the organization or individual responsible for the development of a coded character set in the registration. The owner of origin had ultimate authority over the content of its coded character set. Where a sponsoring authority was not the owner of origin, the owner of origin was given separately on the cover page.

4.5.8   Copyright owner

The copyright owner was the organization or individual holding the copyright for the publication that specified a coded character set in the registration. Where a sponsoring authority was not the copyright owner, the copyright owner was given separately on the cover page.

4.5.9   Code table

4.5.9.1   Purpose of the code table

As a description of the coded character set, the registration of a graphic character set or the control functions includes the code table. The format of the code table is determined by the type of the character set, the details are defined in A.1.2.2. The layouts are defined in Annex D.

4.5.9.2   Exceptional registration without code table

Registration of a coding system not in conformance with ISO/IEC 2022 can refer to external documents which define the registered coding system rather than include a code table.

4.5.10   List of character names

4.5.10.1   Purpose of the list of character names

As a description of the coded character set, registration of a graphic character set or the control functions includes the list of character names. The format of the list of character names is determined by the type of the character set; details are defined in A.1.2.3.

4.5.10.2   Registration without a list of character names

Registration of a multiple byte set can lack a list of character names because it contains too many characters to assign names adequately representing their nature.

NOTE   In the case of ISO/IEC 10646, some ideographic characters have the names assigned by their code positions.

Registration of a coding system not in conformance with ISO/IEC 2022 can refer to external documents which define the registered coding system rather than include a list of character names.

4.6   Format of reference to an existing registration

A reference to an existing registration in the International Register is made by using the prefix "ISO-IR" followed by a SPACE and the registration number.

EXAMPLES

      ISO-IR 16

identifies the particular version of ISO/IEC 646 for the Portuguese language registered on 1976-12-30, while

ISO-IR 48

identifies the set of control functions registered on 1981-07-15.

4.7 Exception for reference to international and national standards

Reference to an international or national standard in the International Register is made by using the identifier assigned by for example, ISO, IEC or ITU (for international standards) or the national body (for national standards). Some registrations were based on international or national standards. However, the registration covered only some of the standards on which the registration was based. Thus, the registration numbers are inappropriate for indicating the coded character set standard.

EXAMPLES

"ISO/IEC 8859-14" is preferred over "ISO-IR 199".

"JIS X 0208-1990" is preferred over "ISO-IR 168".

5   ISO/IEC supervisory body

5.1   Administrative responsibility

The ISO/IEC JTC 1 subcommittee concerned with coded character sets (covered in particular by ISO/IEC 646, ISO/IEC 2022, ISO/IEC 4873, ISO/IEC 6429, and ISO/IEC 10646) has administrative responsibility for this document and the content of the International Register.

NOTE   At the time of publication, subcommittee ISO/IEC JTC 1/SC 2, *Coded character sets*, has this responsibility.

5.2   Editorial maintenance

The contents of the International Register are transferred from the different locations hosted by the former registration authority. They can include some hyperlinks referring to the previous locations. The supervisory body is responsible for the editorial maintenance involved in correcting these unreachable hyperlinks when they are correctable.

5.3   Record of corrections

Some registrations have a record of correction (such as typographical errors or errors in character shapes) which were approved. The corrections were approved by the former registration authority, the sponsoring authority, the owner of origin, or the copyright owner. When a registration had a correction, the cover page included the date of the correction and the reason for the change. The corrected pages were also published.

5.4   Record of revisions

5.4.1   Revisions to the coded character set

As specified in ISO/IEC 2022:1994, 14.5, some registered character sets were revised in the form of the upwardly compatible versions. In the case that the revised character set standard was not identical to the registered coded character set, the revised character set standard was registered with new registration with a new escape sequence. The criteria for how the former registration authority determined whether two-character sets are identical or different are described in B.2.

5.4.2   Addition of or revision to the ISO/IEC 10646 mapping

In some registrations, the mapping table to ISO/IEC 10646 was added or revised after the registration of the original character set standard. Details of the mapping table is described in A.2.

5.5   Withdrawal

Some registrations were withdrawn by the sponsoring authority. The catalogue advises of the item's status as withdrawn status. Some withdrawn registrations include the withdrawal notice page before the cover page (e.g. ISO-IR 25) while others lack it altogether (e.g. ISO-IR 61).

8

Annex A
(informative)

Details of registrations in the International Register

A.1 Contents of a registration

A.1.1 Cover page

Each registration has a cover page.

A.1.1.1 Content provided by the sponsoring authority

The sponsoring authority provided the following elements on the cover page:

— "type" indicates the type of coded character set registration; the type is one of the three categories or one of the eight types from the list below.

    a) graphic coded character set

        1) 94-character graphic character set

        NOTE   In the catalogue, this type is divided into two subtypes according to the length of the Intermediate bytes allocated by the former registration authority (see 4.4.1). Since the intermediate bytes cannot be specified at the application stage, the type on the cover page is not divided.

        2) 96-character graphic character set

        3) multiple-octet graphic character set

    b) control functions

        1) C0-control character set

        2) C1-control character set

        3) single control function (see ISO/IEC 2022:1994, 6.5.1)

    c) coding systems not in conformance with ISO/IEC 2022

        1) coding system that uses the standard return

        2) coding system that does not use the standard return

— a short name for the coded character set

— a short description

— the sponsoring authority (4.5.5)

— the owner of origin of the character or coded character set (4.5.7)

— a general indication of the intended field of application

— the description states if any of the following conditions apply:

— a mapping to ISO/IEC 10646 is included (see A.2, and an example in Annex E);

— the coded character set is intended for use in combination with one or more other registered sets;

— the coded character set is intentionally a subset or a superset of one or more other registered sets (if the coded character set is a part of one or more standards, reference(s) to the standard(s) is included either in the short description or under "origin");

— the registration is for a coding system not in conformance with ISO/IEC 2022 and the ISO/IEC 2022 standard return (escape sequence ESC 2/5 4/0) applies;

— the referential information to obtain a publicly available document describing the coding system if it is not in conformance with ISO/IEC 2022 and the registration does not include the code table and list of character names;

— the registration can be subject to the modification after the registration (see 5.3 and 5.4);

— the registration is a revision to a previously registered standard; if so, then the identity of that coded character set is given (see 5.4).

A.1.1.2   Content provided by the former registration authority

The former registration authority provided the following elements on the cover page:

— the registration number;

— the date of registration;

— the allocated escape sequence(s);

— whenever registration was revised, the date of the revision and description of each change;

— when a sponsoring authority withdrew the registration, the date of the withdrawal (and the reason, when available);

— if the registration was for the revision of a standard that was previously registered, the new registration was identified on the cover page of the original registration and the original registration was noted on the cover page of the new registration (see 5.4).

A.1.2   Description of the coded character set

A.1.2.1   Components to describe the coded character set

The description of the coded character set contains both a code table and a list of character names. A description can be omitted from the registration for an ISO or ISO/IEC coded character set standard or a coding system not in conformance with ISO/IEC 2022 but documented in a publicly available document (see B.1). All other registration applications have the descriptions of the coded character set.

A.1.2.2   Code table

A.1.2.2.1   Graphic character set

For 94-character coded graphic character sets, the layout of the code table in D.1 is used; for 96-character coded graphic character sets, the layout of the code table in D.2 is used; for multiple-octet coded graphic character sets, the 16 × 16 or 24 × 24 layouts of the code table in D.3 are used.

NOTE   Some registrations use multiple code tables with other arrangements, as necessary. ISO-IR 169, for instance, uses code tables of 24 rows × 8 columns.

A.1.2.2.2   Control functions

For C0 sets, the layout of the table in D.4 is used. For C1 sets, the layout of the table in D.5 is used. For C1 sets, the two-character escape sequences of type ESC Fs (see ISO/IEC 2022:1994, 13.2.1) are listed for 7-bit coding.

A.1.2.2.3   Non-conformant graphic character sets

In this document, no layout of the code table is defined for coding systems not in conformance with ISO/IEC 2022 (see B.1). However, an example presented to the sponsoring authority is retained in D.6.

A.1.2.3   List of character names

A.1.2.3.1   Graphic character sets

For graphic character sets, the list of character names shows all the code positions in the code table and indicates the name of the character allocated to each code position as the name appears in the coded character set being registered. Combining characters are identified as such, by adding the text "(combining character)" immediately following the character name.

A.1.2.3.2   Control functions

For character sets of control functions, the list of character names shows the control functions of the set by indicating the name and definition for each code position in the code table as the name appears in the coded character set being registered. Following the list of character names, a registration lists the control functions of the set indicating the name and definition for each code position.

A.1.2.3.3   Unused positions

Unused positions are indicated in the list of character names by the presence of "(this position shall not be used)" in place of a character name. For a contiguous range of unused positions, the list can show the range of code positions as a single entry, where the code position shows the first code position in the range, the word "to", and the last code position, and the text for the character name is "(these positions shall not be used)".

A.1.2.3.4   Notes in the list of character names

When it is important for the understanding of a graphic character, a short note for the character is included in the list of character names.

A.2   Mapping to ISO/IEC 10646

A.2.1   Purpose of the mapping table

A mapping of the characters in the coded character set in the registration to ISO/IEC 10646 or an equivalent standard can be included in the registration as an option. If such a mapping is included, then the following clauses apply.

A.2.2   Identity of the coded character set

The mapping identifies the coded character set mapped to ISO/IEC 10646, for example, by name and registration number.

A.2.3   Date of creation

The mapping includes the date of creation.

A.2.4 Part and edition of mapped ISO/IEC 10646

The mapping identifies the applicable part and edition of ISO/IEC 10646 plus any amendments and corrigenda on which the mapping is based.

A.2.5 Format of the mapping table

The mapping to ISO/IEC 10646 is in machine-readable form.

A.2.6 Mapped entity

The mapping equates each character in the coded character set to exactly one of these alternatives:

— a single character in ISO/IEC 10646;

— a combining sequence in ISO/IEC 10646;

— no ISO/IEC 10646 character;

— optionally, when no equivalent character is in ISO/IEC 10646, a character in either the private use area or private use planes of ISO/IEC 10646.

A.2.7 Mapped entity

Unused code positions in the coded character set are not included in the mapping table.

For each character in the registration, the mapping contains the following two elements:

— the code position in the registered coded character set;

— the corresponding ISO/IEC 10646 code position or combining sequence.

A.2.8 Registration without mapping

If a mapping to ISO/IEC 10646 does not exist for a character in the registration, then the mapping makes this clear, for example, the word "none" replaces the ISO/IEC 10646 code position. As an alternative, the mapping can specify a code position in the private use area of the basic multilingual plane or one of the private use planes of ISO/IEC 10646. Since the use of private use areas is based on an agreement between the sender and receiver about the meaning of the code positions in the private use area, the use of private use areas is discouraged. Consequently, if a code position from the private use area or planes is specified, the mapping highlights this, for example, with the text "(Private use character)".

A.2.9 Combining character

If combining characters are included in the registration, the mapping documentation indicates whether the base character precedes or follows the corresponding combining character or characters in the combining sequence for the registered coded character set.

A.2.10 Alternate mapping

To make the mapping precise and unambiguous, the description can include alternate mappings and special situations considered by the implementer. Registered alternate mapping was designed to meet the following conditions:

— The mapping for each character is in a single line of text.

— Each line of text in an individual character mapping contains the elements specified in the same order. A character from the registration is expected to be clearly separated and distinguishable from the corresponding character or characters in ISO/IEC 10646. For example, registration code positions can be separated from the corresponding ISO/IEC 10646 code position or code positions by the character

tabulation control character of ISO/IEC 6429 (CHARACTER TABULATION (HT)). Code positions for combining sequences can be separated from each other by a COMMA.

— Code positions use hexadecimal (base 16) notation (the digits "0" through "9", and the Latin letters "A" through "F" (or "a" through "f").

— The machine-readable format of the mapping is limited to the repertoire for ISO/IEC 646 IRV and the three characters CHARACTER TABULATION (HT), CARRIAGE RETURN (CR), and LINE FEED (LF) from ISO/IEC 6429.

— Records are ordered by the code position of the character in the registration.

A.2.11 Supplementary information

After the mapping records, the mapping can include supplementary information for clarification, for example, when a special situation warrants an alternate mapping for a character.

A.2.12 Date of acceptance

The mapping includes the date that the mapping was accepted by the former registration authority.

A.2.13 Additional information

The former registration authority can have specified additional presentation and information guidelines for the mapping.

A.2.14 Example of a mapping table

Annex E shows an example of a mapping table.

A.3   Repertoire

For coded graphic character sets, the registration specifies only the characters of the set and their coded representations, as shown in the code table of the registration. It does not specify a repertoire of characters that can be obtained by combining the characters of the set, for example by means of backspace sequences or of combining sequences.

Annex B
(informative)

Coded character sets with special consideration

B.1 Coding systems not in conformance with ISO/IEC 2022

B.1.1 Acceptable registrations

A coding system not in conformance with ISO/IEC 2022 was registered only if

— the application identifies the publicly available document that describes the coded character set, or

— the application includes the code table and list of character names.

B.1.2 Code table and list of character names

If the registration does not include the code table and list of character names, the cover page indicates where a publicly available document describing the coding system can be obtained.

B.1.3 Escape sequence for the coding systems not in conformance with ISO/IEC 2022

Although the coding system is not in conformance with ISO/IEC 2022, the registered escape sequence for the coding system is in accordance with ISO/IEC 2022.

B.2 Identical sets

B.2.1 Avoidance of duplicated character sets in the International Register

If an application for registration contained a coded character set identical to an already-registered coded character set, it was refused to avoid the multiple registrations of the identical character sets.

B.2.2 Conditions to distinguish the character sets

Two coded character sets are deemed to be identical if

— both sets are of the same type, for example, both a C0 or both a C1 coded character set,

— the number of characters is the same,

— the names of the characters are the same according to the terminology of the former registration authority,

— the same code positions (values) are used for the same characters,

— the definitions of control characters are functionally equivalent (a more restricted definition is not considered equivalent),

— graphic characters have the same geometric shape apart from typographic variations between fonts.

B.3 Multiple registrations for the same application

Provided that identical registrations do not occur (see B.2), a single application (for example a programming language or a natural language) can have multiple coded character sets in the International Register.

Annex C
(informative)

Criteria for the allocation of ESC F_s sequences

C.1 Priority to control functions

ISO/IEC 2022 provides for a very limited number of ESC F_s sequences. Priority in the allocation of ESC F_s sequences was given to control functions used for general code extension purposes.

C.2 Priority to the broad applicability

Other candidates for ESC F_s representation are of a general nature with broad applicability. The action of such control functions is largely independent of the graphical or control character sets invoked at the time.

C.3 Independence of control functions

Registered control functions were designed to be logically independent of each other except when a control function was one half of a complementary pair, e.g. in an ON/OFF action.

C.4 Purpose of an ESC F_s sequence in the application

Any application for registration for an ESC F_s sequence included

— a complete definition of the control function with an indication of the overall environment in which it was to be used, and

— a justification for more efficient coding of the function via a registered ISO/IEC 2022 ESC F_s sequence.

Annex D
(informative)

Layout of code tables

D.1   94-character graphic character sets

Figure D.1 illustrates the layout of a code table for a 94-character graphic character set. The shaded positions correspond to code positions reserved in ISO/IEC 2022 for control characters. For registration of character sets not conforming to ISO/IEC 2022, such shading can be omitted.

b_5	0	0	0	0	1	1	1	1
b_6	0	0	1	1	0	0	1	1
b_7	0	1	0	1	0	1	0	1

b_4 b_3 b_2 b_1		0	1	2	3	4	5	6	7	HEX
0 0 0 0	0	▓	▓	▓						0
0 0 0 1	1	▓	▓							1
0 0 1 0	2	▓	▓							2
0 0 1 1	3	▓	▓							3
0 1 0 0	4	▓	▓							4
0 1 0 1	5	▓	▓							5
0 1 1 0	6	▓	▓							6
0 1 1 1	7	▓	▓							7
1 0 0 0	8	▓	▓							8
1 0 0 1	9	▓	▓							9
1 0 1 0	10	▓	▓							A
1 0 1 1	11	▓	▓							B
1 1 0 0	12	▓	▓							C
1 1 0 1	13	▓	▓							D
1 1 1 0	14	▓	▓							E
1 1 1 1	15	▓	▓						▓	F
		0	1	2	3	4	5	6	7	HEX

Figure D.1 — Layout of a code table for a 94-character graphic character set

D.2   96-character graphic character sets

Figure D.2 illustrates the layout of a code table for a 96-character graphic character set. The shaded positions correspond to code positions reserved in ISO/IEC 2022 for control characters. For registration of character sets not conforming to ISO/IEC 2022, such shading can be omitted.

					0	1	2	3	4	5	6	7		
b5					0	0	0	0	1	1	1	1		
b6					0	0	1	1	0	0	1	1		
b7					0	1	0	1	0	1	0	1		
b4	b3	b2	b1		0	1	2	3	4	5	6	7		HEX
0	0	0	0	0										0
0	0	0	1	1										1
0	0	1	0	2										2
0	0	1	1	3										3
0	1	0	0	4										4
0	1	0	1	5										5
0	1	1	0	6										6
0	1	1	1	7										7
1	0	0	0	8										8
1	0	0	1	9										9
1	0	1	0	10										A
1	0	1	1	11										B
1	1	0	0	12										C
1	1	0	1	13										D
1	1	1	0	14										E
1	1	1	1	15										F
					0	1	2	3	4	5	6	7		HEX

Figure D.2 — Layout of a code table for a 96-character graphic character set

D.3 Multi-octet graphic character sets

D.3.1 16 × 16 presentation format of multi-octet graphic character sets

Figure D.3 illustrates the layout of a code table for a multi-octet graphic character set, in 16 × 16 presentation format

b5																
b5	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
b6	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
b7	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
b8	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

b4 b3 b2 b1		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
0 0 0 0	0																	0
0 0 0 1	1																	1
0 0 1 0	2																	2
0 0 1 1	3																	3
0 1 0 0	4																	4
0 1 0 1	5																	5
0 1 1 0	6																	6
0 1 1 1	7																	7
1 0 0 0	8																	8
1 0 0 1	9																	9
1 0 1 0	10																	A
1 0 1 1	11																	B
1 1 0 0	12																	C
1 1 0 1	13																	D
1 1 1 0	14																	E
1 1 1 1	15																	F
		0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	HEX

Figure D.3 — 16 × 16 layout of a code table for a multi-octet graphic character set

D.3.2   24 × 24 presentation format of multi-octet graphic character sets

Figure D.4 illustrates the layout of a code table for a multi-octet graphic character set, in 24 × 24 presentation format.

b₇	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
b₆	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
b₅	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
b₄	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	1
b₃	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1	0
b₂	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0
b₁	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0

b₇ b₆ b₅ b₄ b₃ b₂ b₁		33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	
0 1 0 0 0 0 1	33																									21
0 1 0 0 0 1 0	34																									22
0 1 0 0 0 1 1	35																									23
0 1 0 0 1 0 0	36																									24
0 1 0 0 1 0 1	37																									25
0 1 0 0 1 1 0	38																									26
0 1 0 0 1 1 1	39																									27
0 1 0 1 0 0 0	40																									28
0 1 0 1 0 0 1	41																									29
0 1 0 1 0 1 0	42																									2A
0 1 0 1 0 1 1	43																									2B
0 1 0 1 1 0 0	44																									2C
0 1 0 1 1 0 1	45																									2D
0 1 0 1 1 1 0	46																									2E
0 1 0 1 1 1 1	47																									2F
0 1 1 0 0 0 0	48																									30
0 1 1 0 0 0 1	49																									31
0 1 1 0 0 1 0	50																									32
0 1 1 0 0 1 1	51																									33
0 1 1 0 1 0 0	52																									34
0 1 1 0 1 0 1	53																									35
0 1 1 0 1 1 0	54																									36
0 1 1 0 1 1 1	55																									37
0 1 1 1 0 0 0	56																									38
		21	22	23	24	25	26	27	28	29	2A	2B	2C	2D	2E	2F	30	31	32	33	34	35	36	37	38	HEX

Figure D.4 — 24 × 24 layout of a code table for a multi-octet graphic character set

D.4   C0 control function sets

Figure D.5 illustrates the layout of a code table for a C0 control function set, in 7-bit encoding and 8-bit encoding. The positions in the code tables in Figure D.5 correspond to code positions reserved in ISO/IEC 2022 for control characters.

b₄	b₃	b₂	b₁		00	01	HEX
0	0	0	0	00			0
0	0	0	1	01			1
0	0	1	0	02			2
0	0	1	1	03			3
0	1	0	0	04			4
0	1	0	1	05			5
0	1	1	0	06			6
0	1	1	1	07			7
1	0	0	0	08			8
1	0	0	1	09			9
1	0	1	0	10			A
1	0	1	1	11			B
1	1	0	0	12			C
1	1	0	1	13			D
1	1	1	0	14			E
1	1	1	1	15			F
					0	1	HEX

Figure D.5 — Layout of a code table for a C0 control function set

D.5   C1 control function sets

Figure D.6 illustrates the layout of a code table for a C1 control function set. The positions in the code tables in Figure D.6 correspond to code positions reserved in ISO/IEC 2022 for control characters.

b8	1	1
b7	0	0
b6	0	0
b5	0	1

b4	b3	b2	b1		08	09	
0	0	0	0	00			0
0	0	0	1	01			1
0	0	1	0	02			2
0	0	1	1	03			3
0	1	0	0	04			4
0	1	0	1	05			5
0	1	1	0	06			6
0	1	1	1	07			7
1	0	0	0	08			8
1	0	0	1	09			9
1	0	1	0	10			A
1	0	1	1	11			B
1	1	0	0	12			C
1	1	0	1	13			D
1	1	1	0	14			E
1	1	1	1	15			F
					8	9	HEX

Figure D.6 — Layout of a code table for a C1 control function set

D.6 Example of an 8-bit graphic character set not in conformance with ISO/IEC 2022

Registration of an 8-bit graphic character set in conformance with ISO/IEC 2022 consists of two separate entries in the International Register:

— registration of the graphic characters of the lower range as a 94-character graphic character set;

— registration of the graphic characters of the upper range as a 96-character graphic character set.

A registration of an 8-bit graphic character set not in conformance with ISO/IEC 2022 can consist of one part. Figure D.7 can be used for the full 8-bit set of a coding system not in conformance with ISO/IEC 2022.

NOTE   The shaded blocks in this example indicate general practice for such character sets. They do not imply conformance with ISO/IEC 2022.