
British Standard

Coded character sets for text communication

Part 1. General introduction

Jeux de caractères codés pour la transmission de texte
Partie 1. Généralités

Zeichensätze für Textkommunikation
Teil 1. Allgemeineinleitung

Contents

	Page		Page
National foreword		Annexes	
Committees responsible	Inside front cover Back cover	A. Method of identification	5
0. Introduction	1	B. Text communication concepts	8
1. Scope and field of application	1	C. Equipment conformance	11
2. References	2	D. Fall-back	12
3. Definitions	2	Figures	
4. Character repertoires and coded representations	3	1. Method of identification	6
5. The character SPACE	3	2. Numbering principle for alphabetic characters	7
6. Restrictions on the use of coded character sets	4	3. Illustration of the position of the coding interface (CI) in text communication	9
7. Conformance requirements for text communication	4	4. General structure of a text communication terminal	10

National foreword

This Part of BS 6692 has been prepared under the direction of the Office and Information Standards Committee. It is identical to ISO 6937/1-1983 'Information processing — Coded character sets for text communication — Part 1 : General introduction' published by the International Organization for Standardization (ISO).

At present this British Standard consists of two Parts:

- Part 1 General introduction
- Part 2 Specification for Latin alphabetic and non-alphabetic graphic characters

Part 2 is identical to ISO 6937/2-1983 'Information processing — Coded character sets for text communication — Part 2: Latin alphabetic and non-alphabetic graphic characters'. Other parts of the international standard are in preparation specifying the coded representation of other characters or control functions and these will be reviewed when available to see if further Parts of this British Standard should be prepared.

Terminology and conventions. The text of the international standard has been approved as suitable for publication as a British Standard without deviation. Some terminology and certain conventions are not identical with

those used in British Standards; attention is drawn especially to the following.

From clause 1 onwards, wherever the words 'International Standard' or 'part of ISO 6937' appear, referring to this standard, they should be read as 'British Standard' or 'Part of BS 6692'.

Cross-references

International standard	Corresponding British Standard
ISO 646-1983	BS 4730 : 1985 Specification for the UK 7-bit coded character set (Identical exercising national options)

The Technical Committee has reviewed the provisions of ISO 2022-1982 and ISO 6429-1983, to which reference is made in the text, and has decided that they are acceptable for use in conjunction with this standard.

Additional information. ISO 2022-1982, to which reference is made in the text, is currently being revised. Until that revision is published no decision will be made on its implementation as a British Standard. ISO 2022-1982 was not implemented as a British Standard. BS 4953, which was technically equivalent to ISO 2022-1973, has been withdrawn.

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

British Standard

Coded character sets for text communication

Part 1. General introduction

0 Introduction

This International Standard specifies repertoires of graphic characters and control functions, and their coded representations, for use in text communication. It applies to the communication of text in the form of binary-coded representations of graphic characters and control functions, using

- a) public communication networks;
- b) private communication networks;
- c) interchange media such as magnetic tapes and discs.

It is applicable to the communication of text at the coding interface (see 3.5 and annex B).

Although, in general, text (see 3.1) consists of characters and pictures, this International Standard applies only to text made up of characters.

This International Standard consists, at present, of three parts, as follows :

ISO 6937/1, General introduction.

ISO 6937/2, Latin alphabetic and non-alphabetic graphic characters.

ISO 6937/3, Control functions for page-image format (in preparation).

Other parts may be added later.

ISO 6937/2 and ISO 6937/3 may be used either in combination with one another or individually, but they should always be used in combination with this part of ISO 6937.

The specifications are based on the 7-bit coded character set specified in ISO 646, the 7-bit and 8-bit code extension techniques specified in ISO 2022, and the definitions of additional control functions given in ISO 6429.

ISO 6937/1, ISO 6937/2 and ISO 6937/3 have been developed in parallel with CCITT Recommendations S.61, *Character repertoire and coded character sets for the international teletex service* (Geneva, 1980) and S.100, *International information ex-*

change for interactive videotex (Geneva, 1980). The graphic character repertoire defined in ISO 6937/2 is a superset of the graphic character repertoires defined in CCITT Recommendations S.61 and S.100, and the control function repertoire to be defined in ISO 6937/3 is a superset of the control function repertoire defined in CCITT Recommendation S.61. However, the primary and supplementary sets of graphic characters specified in this International Standard are not identical to those defined in CCITT Recommendations S.61 and S.100 (see ISO 6937/2, annex F). To ensure compatible interworking between the public text communication services defined by the CCITT and terminal equipment communicating text conforming to this International Standard, special attention should be given to the differences between this International Standard and CCITT Recommendations S.61 and S.100.

1 Scope and field of application

This part of ISO 6937 constitutes a general introduction to coded character sets for text communication and, in particular,

- a) defines terms used in this and subsequent parts of ISO 6937;
- b) describes the general structure of the text communication code;
- c) specifies conformance requirements for the coded representation of communicated text at the coding interface.

In addition, it gives, in annexes,

- a) a description of the method used to identify each graphic character and control function defined in this International Standard (annex A);
- b) a description of some general concepts of text communication (annex B);
- c) suggestions for equipment conformance requirements (annex C);
- d) recommendations for fall-back implementations (annex D).

2 References

ISO 646, *Information processing — ISO 7-bit coded character set for information interchange*.

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

ISO 6429, *Information processing — ISO 7-bit and 8-bit coded character sets — Additional control functions for character-imaging devices*.

3 Definitions

For the purpose of this International Standard, the following definitions apply :

3.1 text : A representation of information for human comprehension that is intended for presentation in a two-dimensional form, for example printed on paper or displayed on a screen.

Text consists of symbols, phrases or sentences in natural or artificial languages, pictures, diagrams and tables.

NOTE — This International Standard applies only to text made up of characters.

3.2 text communication; communication of text : The transfer of text by means of telecommunication or by physical interchange of storage media.

NOTE — In the context of this International Standard, text communication is by means of binary-coded representations of characters.

3.3 terminal equipment (text communication) : Equipment that can exchange text by means of telecommunication or by physical interchange of storage media.

NOTE — The concept "terminal equipment" includes both individual terminals and terminal systems.

3.4 terminal system (text communication) : A complete system comprising a number of terminals and a central storage facility, with telecommunication between them.

NOTE — The coding arrangements on the communication lines within a terminal system may differ from those described in this International Standard, which are intended to be used between terminal systems.

3.5 coding interface : An interface through which character-coded text is passed between terminal equipment and communication media, or between a terminal system and communication media.

NOTE — The coding interface is the place at which the communication code is specified.

3.6 character¹⁾ : A member of a set of elements used for the organization, control or representation of text.

NOTE — Characters are classified in two categories : graphic characters and control functions.

3.7 bit combination¹⁾ : An ordered set of bits that represents a character or is used as a part of the representation of a character.

3.8 coded character set; code¹⁾ : A set of unambiguous rules that establishes a character set and the one-to-one relationship between the characters of the set and their bit combinations.

3.9 control function¹⁾ : An action that affects the recording, processing, transmission or interpretation of text and that has a coded representation consisting of one or more bit combinations.

3.10 control character¹⁾ : A control function, the coded representation of which consists of a single bit combination.

3.11 graphic character¹⁾ : A character, other than a control function, that has a visual representation normally hand-written, printed or displayed.

3.12 graphic character repertoire : A collection of graphic characters specified independently of coding, that each occupy a character position when imaged.

NOTE — A fall-back presentation of a graphic character of the repertoire may occupy more than one character position.

3.13 control function repertoire : A collection of control functions specified independently of coding.

3.14 escape sequence¹⁾ : A bit string that is used for control purposes in code extension procedures and that consists of two or more bit combinations. The first of these bit combinations represents the character ESCAPE.

NOTE — Formats and rules regarding the use of escape sequences are specified in ISO 2022.

3.15 control sequence : A bit string that represents a control function and that consists of two or more bit combinations.

In an 8-bit code, the first of these bit combinations represents the character CONTROL SEQUENCE INTRODUCER; in a 7-bit code, the CONTROL SEQUENCE INTRODUCER is represented by the first two bit combinations.

NOTE — Formats of control sequences are specified in ISO 6429. Control sequences are used, in particular, to represent control functions with parameters.

¹⁾ Definition based on the corresponding definition in ISO 2022.

3.16 to designate¹⁾ : To identify a set of characters that are to be represented, in some cases immediately and in others on the occurrence of a further control function, in a prescribed manner.

3.17 to invoke¹⁾ : To cause a designated set of characters to be represented by the prescribed bit combinations.

3.18 graphic symbol : The visual representation of a graphic character.

3.19 composite graphic symbol : A graphic symbol consisting of a combination of two or more other graphic symbols in a single character position, such as a diacritical mark and a basic letter, for example ä.

3.20 to present : To print on a document or display on a screen.

3.21 to perform : To carry out the operation specified by a control function.

3.22 fall-back : A technique whereby, if terminal equipment is unable to present a graphic character uniquely, or to perform a control function exactly, an approximation may be used.

4 Character repertoires and coded representations

4.1 Character repertoires for text communication consist of the character SPACE (see clause 5), and graphic characters and control functions defined in other parts of this International Standard (see clause 0).

Each graphic character or control function defined in this International Standard has a unique identifier in order to facilitate cross-references. The identifiers are assigned in a systematic way which is explained in annex A.

The character repertoires defined in this International Standard are coded using the 7-bit and 8-bit code extension techniques specified in ISO 2022.

4.2 Graphic characters and control functions other than SPACE are coded as elements of sets of up to 32 control functions and sets of up to 94 graphic characters.

4.2.1 There are two types of control function sets : “primary” sets and “supplementary” sets. A primary set of control functions is intended to be used as the C0 set of a 7-bit or 8-bit code, and the associated supplementary set is intended to be used as the C1 set.

The elements of a C0 set of control functions are represented by bit combinations in the range 0/0 to 1/15, both in a 7-bit code and in an 8-bit code.

In a 7-bit code, the elements of a C1 set of control functions are represented by escape sequences of the form ESC Fe, where Fe is a bit combination in the range 4/0 to 5/15.

In an 8-bit code, the elements of a C1 set of control functions are represented by single bit combinations in the range 8/0 to 9/15.

4.2.2 There are two types of graphic character sets : “primary” sets and “supplementary” sets. A primary set of graphic characters is intended to be used as the G0 set of a 7-bit or 8-bit code, and an associated supplementary set is intended to be used as a G1, G2 or G3 set.

The elements of a G0 set of graphic characters are represented by bit combinations in the range 2/1 to 7/14, both in a 7-bit code and in an 8-bit code.

In a 7-bit code, the elements of a G1, G2 or G3 set of graphic characters are also represented by bit combinations in the range 2/1 to 7/14, after invocation by an appropriate code extension function.

In an 8-bit code, the elements of a G1, G2 or G3 set of graphic characters are either represented by bit combinations in the range 10/1 to 15/14 or, as in the 7-bit code, represented by bit combinations in the range 2/1 to 7/14, after invocation by an appropriate code extension function.

The coded representations of the primary and supplementary sets of graphic characters and control functions are specified in the appropriate parts of this International Standard.

NOTES

1 More specific recommendations regarding the designations and the invocations of primary and supplementary sets of graphic characters and control functions are made in other parts of this International Standard.

2 In the notation of bit combinations in this International Standard, no distinction is made between 7-bit and 8-bit codes. The coded representation of the character SPACE, for example, is shown only as 2/0, not as 02/0 or 02/00.

5 The character SPACE

The definition of SPACE is as follows .

ID	ACRONYM	NAME and DEFINITION
SP01	SP	SPACE
		A character that is interpreted both as a graphic character and as a control function.
		As a graphic character, it has a visual representation consisting of the absence of a graphic symbol.
		As a control function, it acts as a format effector which causes the active position to be advanced one character position on the same line. ²⁾

¹⁾ Definition based on the corresponding definition in ISO 2022.

²⁾ Definitions of format effector and active position will be included in ISO 6937/3.