

# Office document architecture (ODA) and interchange format for text and office systems —

## Part 6: Character content architectures

# Committees responsible for this British Standard

The preparation of this British Standard was entrusted by the Information Systems Technology Standards Policy Committee (IST/-) to Technical Committee IST/18, upon which the following bodies were represented:

- British Computer Society
- British Facsimile Industry Consultative Committee
- British Printing Industries Federation
- British Telecommunications plc
- Computing Services Association
- Department of Trade and Industry (Information Technology Division)
- EEA (The Electronics and Business Equipment Association)
- HM Treasury (Central Computer and Telecommunications Agency)
- Her Majesty's Stationery Office
- Joint Network Team
- Ministry of Defence
- National Computer Users' Forum
- National Computing Centre Ltd.
- SGML Users' Group
- Telecommunication Engineering and Manufacturing Association
- University of London
- User Standards Forum for Information Technology (Institute of Data Processing Management)

This British Standard, having been prepared under the direction of the Information Systems Technology Standards Policy Committee, was published under the authority of the Board of BSI and comes into effect on 31 July 1990

© BSI 01-2000

The following BSI references relate to the work on this standard:  
Committee reference IST/18  
Draft for comment 86/67036 DC

ISBN 0 580 18340 8

## Amendments issued since publication

| Amd. No. | Date of issue | Comments |
|----------|---------------|----------|
|          |               |          |
|          |               |          |
|          |               |          |
|          |               |          |

# Contents

|                                                                                          | Page               |
|------------------------------------------------------------------------------------------|--------------------|
| Committees responsible                                                                   | Inside front cover |
| National foreword                                                                        | ii                 |
| 1 Scope                                                                                  | 1                  |
| 2 Normative references                                                                   | 1                  |
| 3 Definitions                                                                            | 2                  |
| 4 General principles                                                                     | 2                  |
| 5 Character positioning                                                                  | 4                  |
| 6 Character imaging                                                                      | 14                 |
| 7 Definition of character presentation attributes                                        | 18                 |
| 8 Character content portion attributes                                                   | 27                 |
| 9 Formal definitions of character content architecture dependent data types              | 27                 |
| 10 Graphic characters                                                                    | 32                 |
| 11 Definition of control functions and the character SPACE                               | 32                 |
| 12 Content layout process                                                                | 40                 |
| 13 Content imaging process                                                               | 46                 |
| 14 Interactions between presentation attributes and control functions                    | 47                 |
| 15 Definition of character content architecture classes                                  | 48                 |
| Annex A (informative) Summary of content architecture classes                            | 51                 |
| Annex B (informative) Character content architecture levels                              | 56                 |
| Annex C (informative) Coded representations of control functions                         | 69                 |
| Annex D (informative) Summary of object identifiers                                      | 70                 |
| Annex E (normative) SGML representation of character content specific attributes for ODL | 70                 |
| Figure 1 — Example of direction                                                          | 5                  |
| Figure 2 — Illustration of kerns                                                         | 6                  |
| Figure 3 — Reference points for character image positioning                              | 6                  |
| Figure 4 — Illustration of character positioning concepts                                | 8                  |
| Figure 5 — Spacing for constant spacing font                                             | 9                  |
| Figure 6 — Spacing for variable spacing font                                             | 9                  |
| Figure 7 — Tabulation                                                                    | 10                 |
| Figure 8 — Character ordering                                                            | 11                 |
| Figure 9 — An example of the specification of Japanese Ruby                              | 12                 |
| Figure 10 — Illustration of itemization, first line offset and indentation               | 16                 |
| Figure 11 — Illustration of line box positioning concepts                                | 14                 |
| Table 1 — Character presentation attributes                                              | 18                 |
| Table 2 — Default values of the presentation attribute “initial offset”                  | 25                 |
| Table 3 — Control functions                                                              | 33                 |
| Table 4 — Parameter values for SGR                                                       | 36                 |
| Table 5 — Summary of presentation attributes                                             | 49                 |
| Table 6 — Summary of control functions                                                   | 50                 |
| Publications referred to                                                                 | Inside back cover  |

## National foreword

This Part of BS 7299 has been prepared under the direction of the Information Systems Technology Standards Policy Committee. It is the UK implementation of, and is identical with, ISO 8613-6:1989 “*Information processing — Text and office systems — Office Document Architecture (ODA) and interchange format — Part 6: Character content architectures*”, published by the International Organization for Standardization (ISO).

The following list includes all the Parts of this British Standard.

- *Part 1: Introduction and general principles;*
- *Part 2: Document structures;*
- *Part 4: Document profile;*
- *Part 5: Office document interchange format (ODIF);*
- *Part 6: Character content architectures;*
- *Part 7: Raster graphics content architectures;*
- *Part 8: Geometric graphics content architectures.*

NOTE To maintain alignment with the numbering of the Parts of ISO 8613, there is at present no Part 3 of this British Standard.

The UK actively participated in the development of the International Standard through BSI Technical Committee IST/18 and fully approved publication of the International Standard.

There are no plans at present to supplement this British Standard with any additional guidance, and any comments or enquiries on the International Standard or its application should be addressed to the Secretary of IST/18 at the BSI headquarters address.

**Reasons for implementation.** This Part of BS 7299 implements the International Standard for the following reasons.

- a) Because information technology equipment, services and systems tend to be, and generally need to be, designed, purchased, and used in a world-wide context, BSI policy is to agree, wherever possible, standards at international level and not to prepare unique national standards. At the time of its approval, the text of the International Standard was considered to be the best that could be achieved, taking into account both the needs of the “market place” and the “state of the art” in this subject area, and the international standardization process.
- b) The Commission of the European Communities has mandated the European Committee for Standardization (CEN) to produce European Standards (ENs) (or, where this cannot be achieved, Harmonization Documents (HDs) or Draft European Standards (ENVs) in this subject area for application throughout the European Community by 1992. This British Standard anticipates that ISO 8613-6:1989 will be adopted without change as the Reference Document for the EN (or HD or ENV) and, therefore, fulfils in advance the UK obligation under CEN/CENELEC Common Rules to implement nationally each agreed European Standard.
- c) Publication of British Standards corresponding to the International Standards in which IST/- resources were invested will ensure that the BSI Catalogue becomes, and then remains, a comprehensive first source of information for UK users of IT equipment and services about the availability of International Standards on IT subjects. In turn, this will facilitate UK use of, and access to, those standards.

**Cross-references.** The cross-references relating to all Parts of this British Standard are given in the national foreword to Part 1.

A British Standard does not purport to include all the necessary provisions of a contract. Users of British Standards are responsible for their correct application.

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

### Summary of pages

This document comprises a front cover, an inside front cover, pages i to iv, pages 1 to 74, an inside back cover and a back cover.

This standard has been updated (see copyright date) and may have had amendments incorporated. This will be indicated in the amendment table on the inside front cover.



## 1 Scope

The purpose of ISO 8613 is to facilitate the interchange of documents.

In the context of ISO 8613, documents are considered to be items such as memoranda, letters, invoices, forms and reports, which may include pictures and tabular material. The content elements used within the documents may include graphic characters, geometric graphic elements and raster graphic elements, all potentially within one document.

NOTE ISO 8613 is designed to allow for extensions, including typographical features, colour, spreadsheets and additional types of content such as sound.

ISO 8613 applies to the interchange of documents by means of data communications or the exchange of storage media.

It provides for the interchange of documents for either or both of the following purposes:

- to allow presentation as intended by the originator;
- to allow processing such as editing and reformatting.

The composition of a document in interchange can take several forms:

- formatted form, allowing presentation of the document;
- processable form, allowing processing of the document;
- formatted processable form, allowing both presentation and processing.

ISO 8613 also provides for the interchange of ODA information structures used for the processing of interchanged documents.

Furthermore, ISO 8613 allows for the interchange of documents containing one or more different types of content such as character text, images, graphics and sound.

This part of ISO 8613

- defines character content architectures that can be used in conjunction with the document architecture defined in ISO 8613-2;
- defines the internal structure of content conforming to these character content architectures;
- defines those aspects of positioning and imaging applicable to presentation of these character content architectures in a basic layout object;
- defines the presentation attributes and control functions applicable to these character content architectures;
- describes a content layout process which, together with the document layout process described in ISO 8613-2, determines the layout of character content in basic layout objects and the dimensions of these basic layout objects.

## 2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO 8613. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO 8613 are encouraged to investigate the possibility of applying the most recent editions of the standards listed below. Members of ISO and IEC maintain registers of currently valid International Standards.

ISO 646:1983, *Information processing — 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 6429:1983, *Information processing — ISO 7-bit and 8-bit coded character sets — Additional control functions for character-imaging devices*.

ISO 6937:1983, *Information processing — Coded character sets for text communication*.

ISO 7350:1984, *Text communication — Registration of graphic character subrepertoires*.

ISO 8824:1987, *Information processing systems-Open Systems Interconnection — Specification of Abstract Syntax Notation One (ASN.1)*.

ISO 8879:1986, *Information processing — Text and office systems — Standardized Generalized Markup Language (SGML)*.

ISO 9541-5: *Information processing — Font and character information interchange — Part 5: Font attributes and character model*<sup>1)</sup>.

### **3 Definitions**

For the purpose of this part of ISO 8613, the definitions given in ISO 8613-1 apply.

## **4 General principles**

### **4.1 Classes**

Three classes of character content architecture are distinguished:

- a character content architecture for formatted content which allows for document content to be presented (e.g. printed or displayed) as intended by the originator. Formatted content can be used in any basic component;
- a character content architecture for processable content which allows for document content to be processed (e.g. edited or formatted). Processable content can be used in any basic logical component;
- a character content architecture for formatted processable content which allows for document content to be processed and also to be presented as intended by the originator. Formatted processable content can be used in any basic component.

### **4.2 Content**

The content of a basic component that conforms to a character content architecture is a character string. This character string is formed by concatenating the character strings in the content portions of the basic component.

The content character string consists of a combination of graphic characters, control functions and space characters.

### **4.3 Presentation attributes**

Presentation attributes are applicable to basic logical and layout components. They contain information that specifies the initial conditions relating to the layout, the imaging and the selection of graphic characters of the content of these basic components. Some of these conditions can be changed by control functions contained within the content.

Presentation attributes are classified as follows:

- logical presentation attributes which can be associated with processable and formatted processable character content. These attributes take effect during the content layout process but are ignored during the content imaging process;
- layout presentation attributes which can be associated with formatted and formatted processable character content. These attributes take effect during the content imaging process. They are generated either by a content layout process or by a process that creates or edits the formatted or formatted processable content;
- shared presentation attributes which can be associated with all character content architecture classes. These attributes take effect during either or both the content layout and imaging processes.

**NOTE** Presentation attributes can be applied to the content of a basic component in one of three ways. They can be specified directly in an object or object class description or they can be specified by means of a presentation style associated with the object or object class description. Alternatively, they can be indirectly applied to an object or object class by means of a default value list (see ISO 8613-2).

### **4.4 Control functions**

Control functions with zero or more parameters may specify information relating to the layout or imaging of subsequent graphic characters. A control function can also be used to extend or replace the set of graphic characters being used. The scope of all control functions is limited to the basic component in which they occur.

---

<sup>1)</sup> To be published.

Classification of control functions is similar to that of presentation attributes:

- logical control functions which can be used in processable and formatted processable character content. These control functions take effect during the content layout process but are ignored during the content imaging process;
- layout control functions which can be used in formatted and formatted processable character content. These control functions take effect during the content imaging process. They are generated by the content layout process. Alternatively, they may be inserted by a process (not described in ISO 8613) that creates or edits the formatted or formatted processable content;
- shared control functions which can be used in all character content architecture classes. These control functions take effect during either or both the content layout and imaging processes.

In addition, formatted processable content may contain control functions known as delimiters. These delimiters are used to indicate a string of one or more graphic characters and/or control functions that have been inserted as the result of a content layout process (see clause 12). The delimited graphic characters and/or control functions take effect only during the content imaging process. The delimiters take effect during the content layout process by deleting them and the enclosed character sequence.

#### 4.5 Graphic characters

The set of graphic characters used in the content of a basic component, and their coded representation, are specified by presentation attributes and code extension control functions (see clause 10 and 11.1.17).

Any set or sets of graphic characters may be used in the content of basic components, subject to the restrictions associated with the particular content architecture in use and subject to proper designation and invocation in accordance with ISO 2022.

Any non-spacing characters included in a graphic character set are not to be used in isolation but only in combination with spacing characters.

#### 4.6 Space characters

The character SPACE (SP) is considered both as a logical control function and as a graphic character. As a graphic character, it has a graphical representation consisting of the absence of a graphic symbol. As a control function, it indicates a potential line break point (see 12.2.1.3.2).

NOTE NBSP (No Break SPace) and any fixed-width space characters, such as “digit space”, “em space” and “en space” are regarded as graphic characters i.e. are not regarded as line break points.

#### 4.7 Coding of content information

The coded representation of the content information within a content portion is in accordance with the rules specified in ISO 2022.

Coded representations of control functions are defined in ISO 6429 and are summarised in Annex C.

#### 4.8 Internal structure

##### 4.8.1 Formatted content

Formatted content is content for which all the necessary information relating to the layout and imaging of that content has been specified. Content in this form is intended to be imaged as specified and is not intended to be revised by an editing process or to be reformatted.

The content of a basic component conforming to a formatted character content architecture consists of one or more lines of characters. Each pair of successive lines is separated by a hard line terminator. The last (or only) line may or may not be terminated by a hard line terminator; the end of the content of a basic component implicitly terminates the last line.

##### 4.8.2 Processable content

Processable content is content which has not been laid out. Content in this form is suitable for revision by an editing process.

NOTE The editing process is implementation dependent and is not described in ISO 8613.

In order to image content in this form, it is necessary to apply a content layout process (see clause 12) to the content which converts the processable content into formatted content (see 4.8.1) or into formatted processable content (see 4.8.3).

To assist the processing (i.e. editing or layout processes) of processable content, a number of logical presentation attributes and control functions have been defined (see clauses 7 and 11). In addition, the character SPACE is regarded as both a graphic character and as a control function that indicates where a line break may occur when the content is laid out.

The content of a basic component conforming to a processable character content architecture consists of one or more sequences of characters. Each pair of successive character sequences is separated by a hard line terminator control function. The last (or only) character sequence may or may not be terminated by a hard line terminator.

If the hard line terminator is omitted at the end of the content of a basic logical component to which another basic logical component is concatenated (see ISO 8613-2), then the last character sequence continues into the content of the next basic logical component. In all other cases, the end of the content of the basic logical component implicitly terminates the last character sequence.

The division into character sequences represents the internal structure of the processable content of a basic logical component. Each character sequence is anonymous, in that no name or identifier is associated with it, and no relationship exists among character sequences except that of sequence.

#### **4.8.3 Formatted processable content**

Formatted processable content is content that is structured such that it contains both the formatted content and the processable content as subsets. It is identical in structure to the processable content, except that it may contain additional control functions and graphic characters that have been added as a result of the content layout process. It is identical in structure to the formatted content, except that it may contain logical control functions and delimiters.

Thus, formatted processable content can be converted to processable content by deleting (or ignoring) all layout control functions, all occurrences of the delimiters and all control functions and characters within those delimiters.

Alternatively, formatted processable content can be converted to formatted content by deleting (or ignoring) all logical control functions and the delimiters but retaining the control functions and characters within the delimiters.

**NOTE** The conversion of formatted processable content to processable content is a reversible process (providing the same layout constraints are applicable to the content layout process) but converting formatted processable content to formatted content is irreversible.

The formatted view of a basic component conforming to a formatted processable character content architecture consists of one or more lines of characters. Each pair of successive lines is separated by either a hard or soft line terminator. The last (or only) line may or may not be terminated by a hard line terminator; the end of the content of a basic layout component implicitly terminates the last line.

The processable view of a basic component conforming to a formatted processable character content architecture consists of one or more sequences of characters. Each pair of successive character sequences is separated by a hard line terminator. The last (or only) character sequence may or may not be terminated by a hard line terminator.

If the hard line terminator is omitted at the end of the content of a basic logical component to which another basic logical component is concatenated (see ISO 8613-2), then the last character sequence continues into the content of the next basic logical component. In all other cases, the end of the content of the basic logical component implicitly terminates the last character sequence.

Soft line terminators are used as separators between lines within a character sequence. The division into character sequences represents the internal structure of the content of a basic logical component. Each character sequence is anonymous, in that no name or identifier is associated with it, and no relationship exists among character sequences except that of sequence.

## **5 Character positioning**

This clause specifies how characters are to be positioned within a basic layout object. The intention is to aid understanding of the presentation attributes and control functions that relate to character positioning.

This clause provides for the positioning of any font that is defined in accordance with ISO 9541-5. This clause also caters for the positioning of characters pertaining to different fonts within the same basic layout object.

## 5.1 Basic concepts

### 5.1.1 Character fonts

In the context of this part of ISO 8613, the term *graphic character* is used in its abstract sense; that is, this term refers to a member of a set of graphic symbols used for the representation of information. The term character image is then used to refer to the rendition of a graphic character on a presentation medium.

A *font* is a set of character images, normally with a common design and size. A set of font attributes is associated with the font as a whole and a set of character attributes is associated with each individual character. These attributes are defined in ISO 9541-5.

The main purpose of the font attributes is for the recipient to identify the font used by the originator and, in case the specified font is not available, the font and character attributes serve as a guidance for the recipient to find an appropriate substitute font among those available.

Further information concerning the designation and invocation of different fonts within a basic object is given in clause 6.

### 5.1.2 Directions

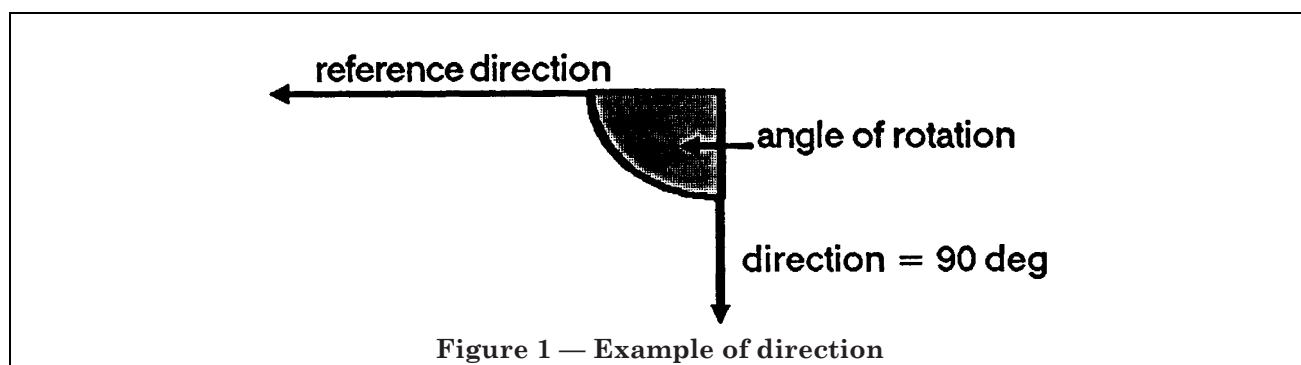
In the context of this part of ISO 8613, all directions are expressed as counter clockwise angles of rotation (in degrees) relative to a specified reference direction (an example is given in Figure 1).

The *character path* is the direction of progression of successive character images within a line box (defined in 5.1.7) and is expressed as a direction relative to the horizontal direction of the layout object (see Figure 4).

The *line progression* is the direction of progression of successive line boxes within the basic layout object and is expressed as a direction relative to the character path (see Figure 11).

The *character orientation* is the direction of the character baseline (defined in 5.1.3) relative to the character path.

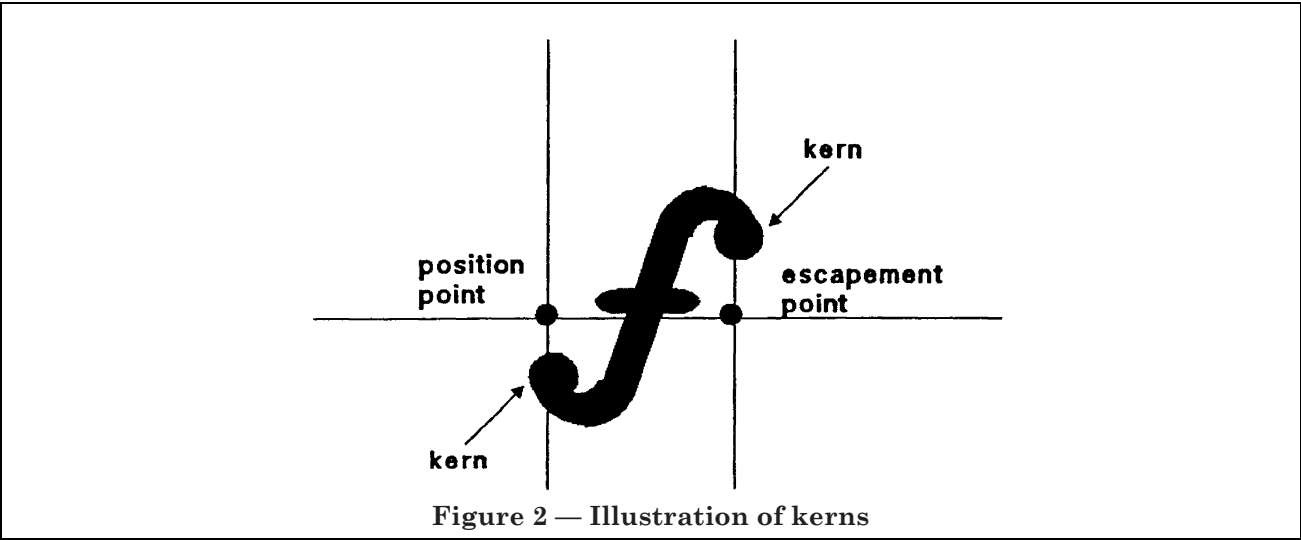
Only one value for the character path, line progression and character orientation may be specified for a basic component.



### 5.1.3 Character image model

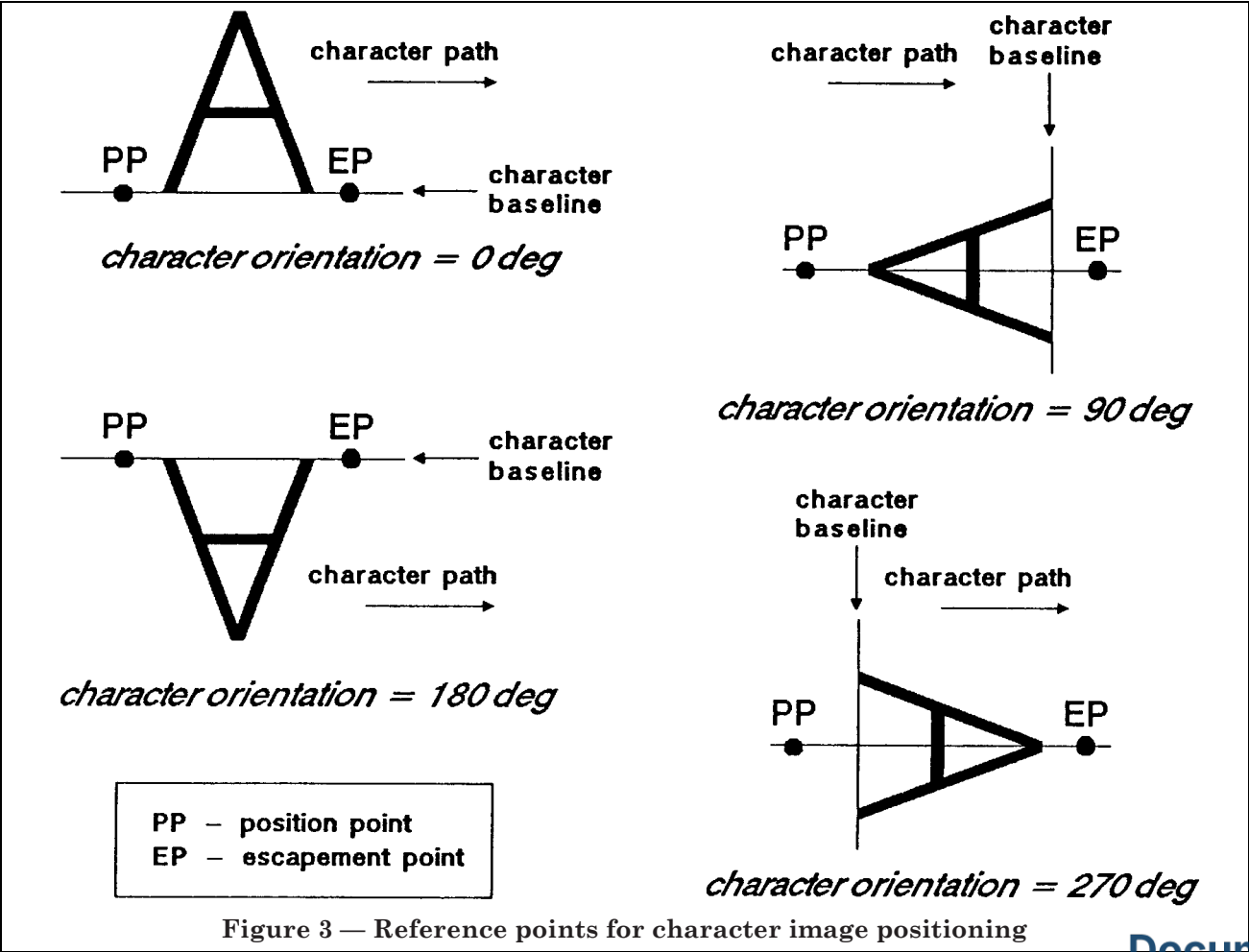
The *position point* is a reference point associated with a character image (see Figure 2). It is used for the positioning of the character image within a line box. The *escapement point* is a reference point associated with a character image (see Figure 2). It is used for the positioning of the next character image.

The *character baseline* is an imaginary line across a character image, for the purpose of defining the character orientation. The character baseline is a horizontal line when the character image is in its intended viewing orientation (see Figure 3).



A position point and escapement point shall be defined for each character orientation which is intended to be used (see Figure 3); i.e. “writing modes” corresponding to the required character orientations shall be defined in the font description, or fall-backs shall be defined in document application profiles.

A *kern* is that part of the character image that extends beyond its position and escapement points (see Figure 2).



#### 5.1.4 Character spacing

The concept of character spacing is only applicable when a constant spacing font is selected. It is used (in conjunction with the inter-character space) to determine the distance between character images within a line box as defined in 5.2.1.

The character spacing is equal to the distance between the position points of successive character images when the inter-character space is zero.

The character spacing is independent of the distance between the position point and escapement point of character images.

#### 5.1.5 Active position

The *active position* is an abstraction of an imaging device concept such as a cursor. This concept is used in the definitions of control functions (see clause 11) where a sequential method of processing a character string is assumed.

The active position indicates the point, within the positioning area of a basic layout object, at which the action specified by the next character (graphic character or control function) is to be effected.

If the next character is a graphic character, its character image is positioned with the position point at the active position and the active position is advanced in the direction of the character path by the amount of spacing defined in 5.2.1. If the next character is a control function, this may cause the active position to move to another point within the positioning area.

#### 5.1.6 Positioning area

A *positioning area* is a rectangular area, wholly contained within a basic layout object, within which position points and escapement points are to be positioned (see Figure 4). Kerns of character images are permitted to extend beyond the positioning area but are not permitted to extend beyond the edges of the basic layout object.

Two of the edges of the positioning area are referred to as the *start edge* and *end edge* (see Figure 4). The start edge and the end edge are defined such that the direction from the start edge to the end edge is in the direction of the character path.

The other two edges of the positioning area are referred to as the *top edge* and *bottom edge*. The top edge and the bottom edge are defined such that the direction from the top edge to the bottom edge is in the direction of line progression.

The start and end edges are indented from the corresponding edges of the basic layout object by a distance referred to as the *kerning offset* (see Figure 4). The kerning offset specified depends upon the fonts of the characters to be positioned in the positioning area. Its value is chosen such that no part of any character image with kerns will extend beyond the boundary of the basic layout object when sequences of character images are positioned within it.

#### 5.1.7 Line boxes

Within the positioning area, a sequence of character images is positioned within an area called a *line box* (see Figure 11). Each line box is a rectangular area that extends from the start edge to the end edge of the positioning area.

Each line box contains a reference point called the *line home position* (see Figure 4). This point is used for positioning the line box within the basic layout object. It also serves as the active position for the first graphic character or control function in each line.

Each line box contains an imaginary line called a *reference line* (see Figure 4). The reference line passes through the line home position in the direction of the character path. It extends from the start edge to the end edge within the line box and is used for the alignment of character images.

The length of the line box is equal to the distance between the start and end edges. The width (or height) of a line box is equal to the sum of the line box forward extent and the line box backward extent (see Figure 4). The line box backward extent is the distance between the reference line and the edge of the line box in the direction opposite to the direction of line progression. The line box forward extent is the distance between the reference line and the edge of the line box in the direction of line progression.

The values of the forward and backward extents depend on the maximum extents (measured perpendicular to the reference line) of the character fonts used in the line box. Determination of the forward and backward extents take into account any displacements of character images perpendicular to the character path, e.g. for subscripts, superscripts and parallel annotation.

NOTE As an example, for a Latin font with character path  $0^\circ$ , line progression  $270^\circ$  and when a single font is used in the line box, then the forward and backward extents are equal to the maximum right and maximum left extents for that font as defined in ISO 9541-5.

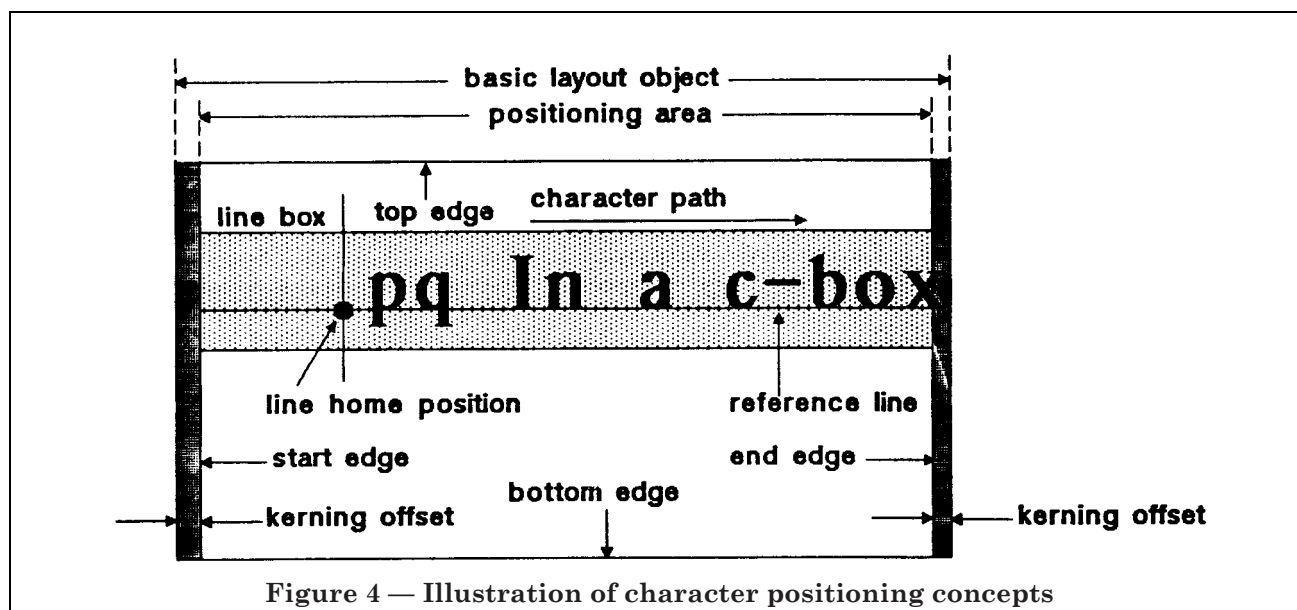


Figure 4 — Illustration of character positioning concepts

## 5.2 Positioning of character images within a line box

Successive character images are positioned within a line box in the direction of the character path.

The position points of the character images are aligned on the reference line unless the characters are imaged as subscripts, superscripts or parallel annotation.

There are several factors which affect the positioning of character images along the reference line:

- spacing between characters;
- alignment;
- tabulation;
- character ordering;
- parallel annotation;
- subscript/superscript;
- pairwise kerning;
- first line offset;
- itemization.

### 5.2.1 Spacing between characters

The *inter-character space* is an additional amount of spacing between the position points of successive character images, in the direction of the character path (see 5.1.4). A negative value indicates a reduction in the spacing between successive character images.

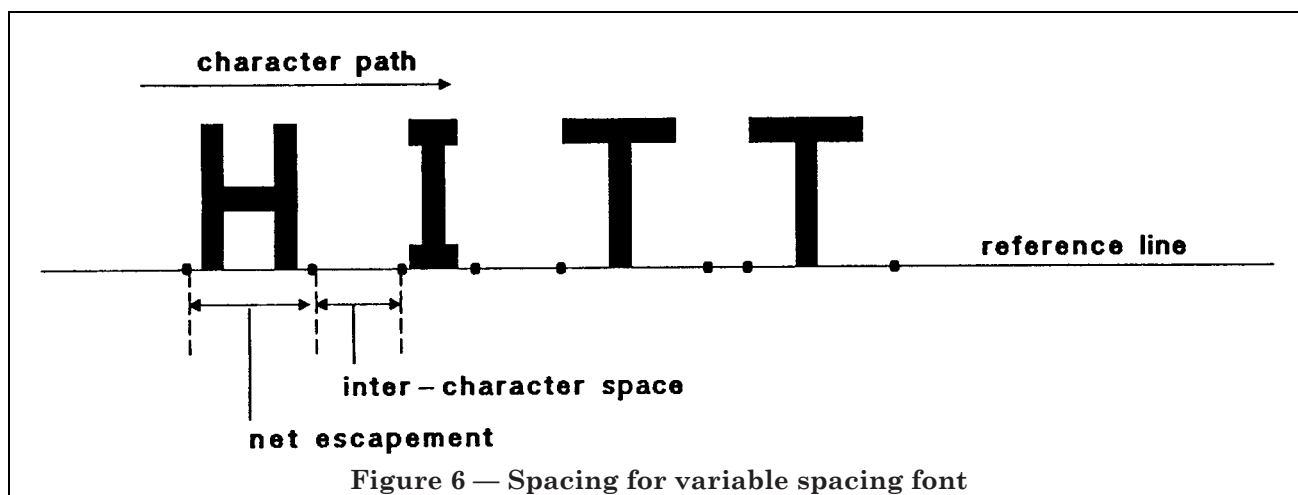
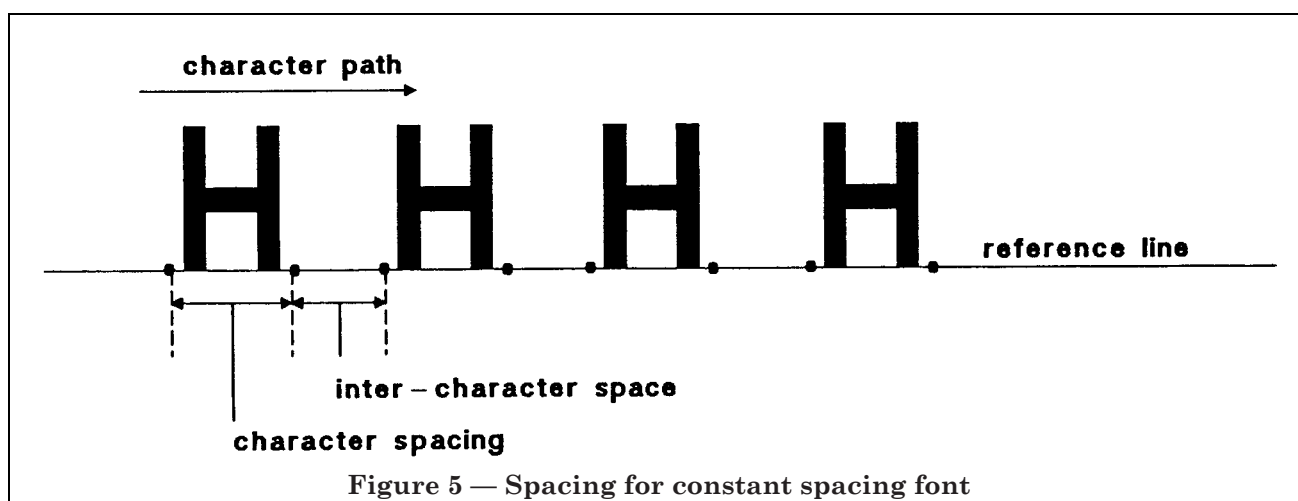
The distance between the position points of successive character images may be constant or variable depending upon the font as follows:

- For fonts with *constant spacing*, the distance between the position points of successive character images is independent of the characters and is the sum of the character spacing (as specified by presentation attributes and control functions) and the inter-character space (see Figure 5).

— For fonts with *variable spacing*, the distance between the position points of successive character images is dependent upon the character i.e. normally the distance between the position point and the escapement point of a character, and is the sum of the net escapement of the character (as specified by the font) and the inter-character space (see Figure 6).

The space width, i.e. the width of the SP (Space) character image, shall be determined as follows:

- for any SP that follows a soft line terminator and precedes the first graphic character of a line, or precedes a soft line terminator and follows the last graphic character of a line, the width is equal to zero;
- in a constant spacing font, the default width equals the character spacing;
- in a variable spacing font, the width is implicitly defined by the font;
- for all fonts, the width may be specified by a control function.



### 5.2.2 Alignment

The character images are positioned within a line box in accordance with the alignment attribute as follows:

- *start-aligned*: the position point of the first character image is placed at the line home position;
- *end aligned*: the escapement point of the last character is placed at the end edge of the positioning area;
- *centred*: the distance from the line home position to the position point of the first character image is approximately equal to the distance from the escapement point of the last character image to the end edge of the positioning area;

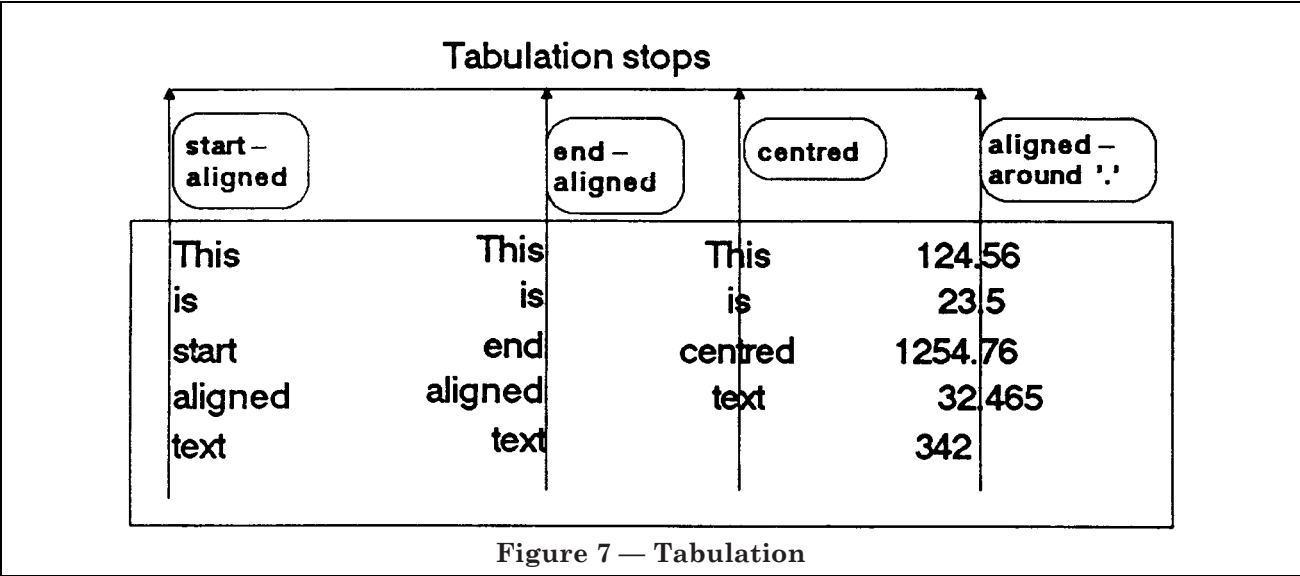
— *justified*: the position point of the first character image is placed at the line home position and the escapement point of the last character image is coincident with the end edge of the positioning area by appropriately setting the space width and/or the inter-character space.

5.2.3 Tabulation

The position of character images along a reference line can be controlled by means of a set of *tabulation stops*. Each tabulation stop specifies a point along a reference line relative to the start edge of the positioning area.

A string of character images can be placed at a tabulation stop by means of a control function embedded in the text. The string may be start-aligned, end-aligned, centred or aligned around one or more specified characters within that string as follows (see Figure 7):

- *start-aligned*: the position point of the first character image of the string is placed at the tabulation stop;
- *end-aligned*: the escapement point of the last character image of the string is placed at the tabulation stop;
- *centred*: the string is placed such that the position point of the first character image and the escapement point of the last character image of the string are approximately equi-distant from the tabulation stop;
- *aligned around*: the position point of the first character image of the first instance of the specified group of characters in that string is positioned at the tabulation stop. If the specified group of characters does not appear in the text associated with that tabulation stop, then the alignment defaults to end-aligned as defined above.



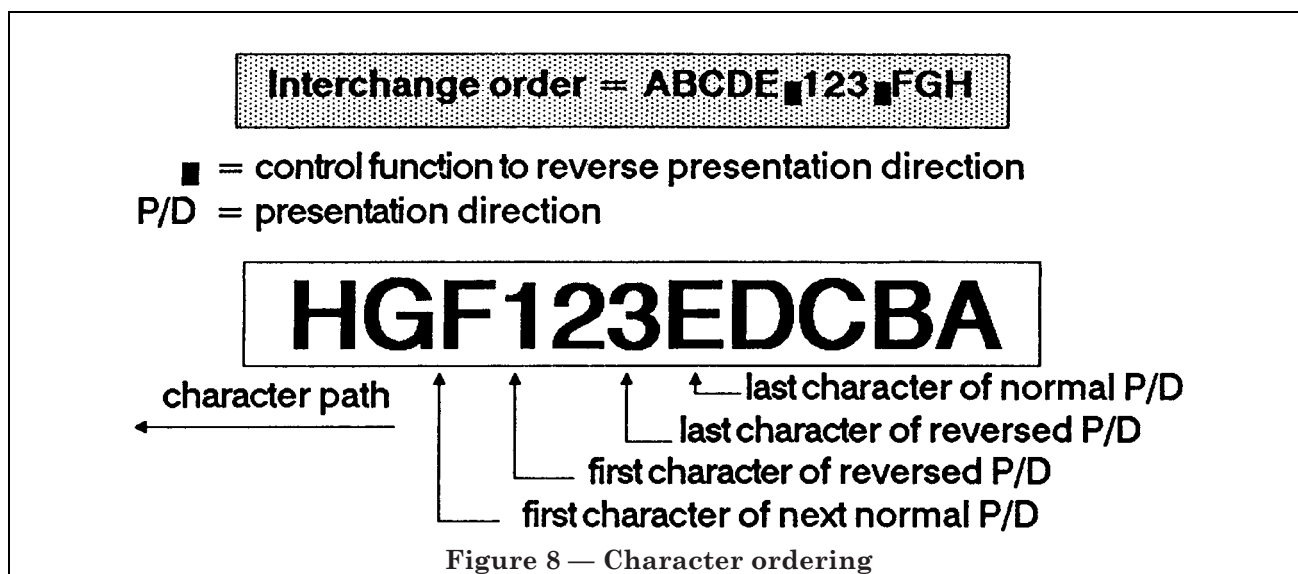
5.2.4 Character ordering

Within the content of a basic object, the interchange order is always the reading order of the language used. In the case of certain languages e.g. Arabic and Hebrew, where the alphanumeric text is read from right to left and the numeric text is read from left to right, the interchanged stream must indicate the change in presentation direction at the appropriate point(s).

This is necessary since control functions in character content architectures are defined to operate sequentially according to their position in the character stream.

When a string of characters with reversed presentation direction is embedded in the text with normal presentation direction, the image of the last character of the string with reversed presentation direction is positioned adjacent to the image of the last character of the preceding string with normal presentation direction (see Figure 8).

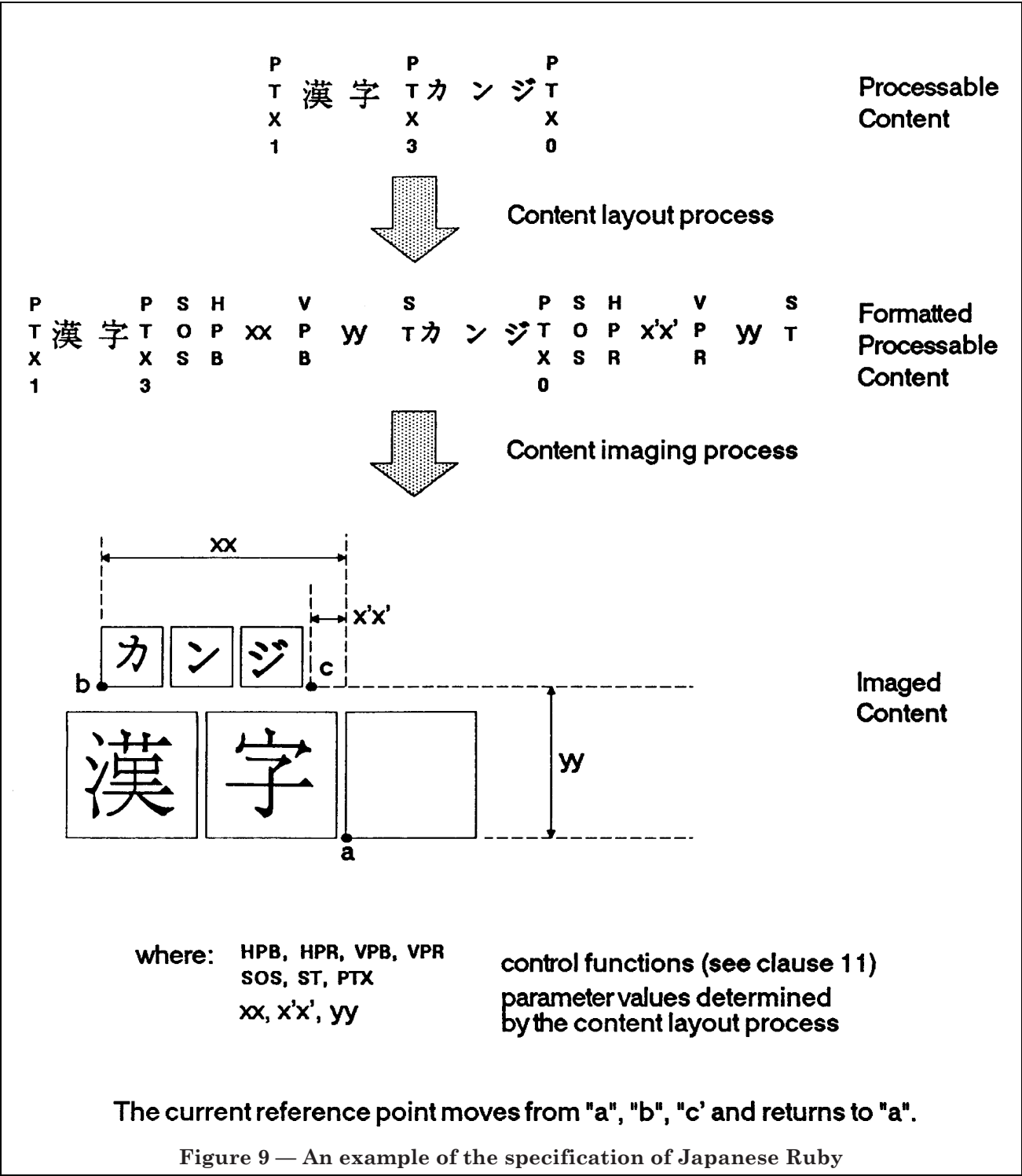
NOTE In Figure 8, the terms “first” and “last” are used in relation to the interchange order and the terms “normal” and “reversed” in relation to the direction of the character path.



#### 5.2.5 Parallel annotation

Two sequential character strings may be delimited such that the second string is used to indicate the pronunciation and interpretation of the first. It is intended to be used in the Japanese language to provide pronunciation and interpretation information (Ruby) in the form of Kana character(s) for one or more Kanji characters. This is indicated in the formatted text by the Kana character(s) being centred either above or to the right of the Kanji character(s), for character image orientations orthogonal and parallel to the character path respectively (see Figure 9).

Where centring would result in Kana characters being positioned outside the positioning area, then the Kana character string is positioned such that it is start-aligned or end-aligned with the edge of the available area.



### 5.2.6 Subscript/superscript

Subscript rendition allows for the active position to be displaced from the reference line in the direction of line progression.

Superscript rendition allows for the active position to be displaced from the reference line in the direction opposite to that of line progression.

The combined effect of all subscript/superscript renditions within a line box shall be such that the active position is returned to the reference line before the occurrence of a hard or soft line terminator.

### 5.2.7 Pairwise kerning

Pairwise kerning allows for the moving of the active position from that defined by the preceding character. The distance and direction depend both on the character being imaged and the preceding character.

In the case of a constant spacing font, pairwise kerning is ignored.

In the case of a variable spacing font, the actual distance between the escapement point of one character image and the position point of the next character image is modified by the kerning information as defined in the character attributes of the font.

### 5.2.8 First line offset

First line offset allows for character imaging of the first line of a basic component to start at a position displaced from the line home position.

The offset is either in the direction of the character path (producing first line indentation) or in the direction opposite to the character path (producing overhang) as illustrated in Figure 10.

### 5.2.9 Itemization

Itemization allows for imaging of an *item identifier* on the first line of a basic component in positions which are not directly constrained by the line home position and the first line offset (see Figure 10).

An item identifier is a string of characters that precedes and is separated from the remainder of the first line of a basic component.

A start offset and an end offset are defined relative to the line home position. These offsets determine the location of the item identifier and the separation between the item identifier and the line home position.

The item identifier may be positioned on the first line in accordance with the item identifier alignment attribute as follows:

- *start-aligned*: the position point of the first character image of the item identifier is placed at the start offset;
- *end aligned*: the escapement point of the last character image of the item identifier is placed at the end offset.

## 5.3 Positioning of line boxes within a basic layout object

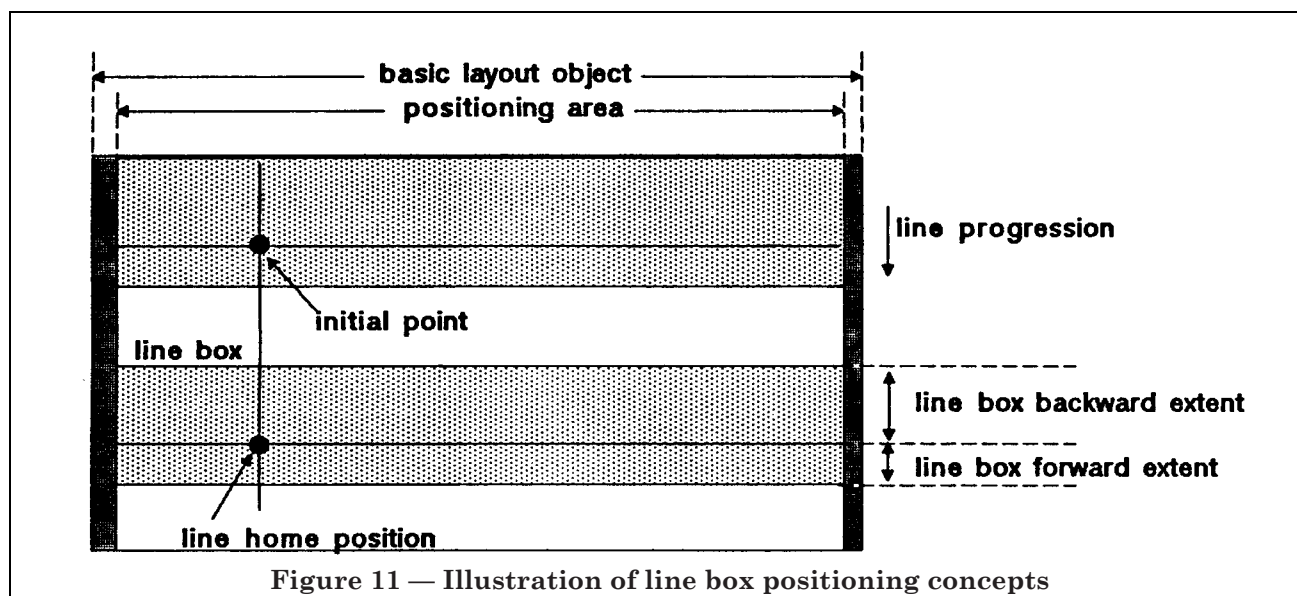
For positioning of line boxes in a basic layout object, the area of that object is independent of any adjoining areas. No part of the image shall be permitted to extend beyond the boundaries of the basic layout object.

The *initial point* is the point relative to which all line boxes are positioned within the basic layout object (see 7.2.2 and Figure 11).

The line home position of the first line box is at the initial point of the basic layout object. Subsequent line home positions are located on a line through the initial point in the direction of line progression.

The distance between the line home positions of two successive line boxes shall be determined as follows:

- when proportional line spacing is to be performed, the distance between the reference lines of two successive line boxes is evaluated by an implementation dependent algorithm not defined in ISO 8613;
- when proportional line spacing is not to be performed, the distance between the reference lines of two successive line boxes is equal to the current line spacing as specified by presentation attributes and control functions.



## 6 Character imaging

Four groups of specifications may apply to the imaging of graphic character elements in a basic object, namely those relating to

- emphasis;
- font selection;
- subscript and superscript;
- character combinations.

These groups of specifications are defined below.

NOTE 1 Emphasis (by “weight”, “posture” etc.) and subscript/superscript rendition may be achieved by font selection.

NOTE 2 Document application profiles may define additional restrictions on the use of the character features defined in this clause. In addition, an implementation may replace an imaging feature by an alternative fall-back feature. However, ISO 8613 does not define preferred fall-back features. It should also be noted that use of a fall-back feature may cause incorrect or misleading information to be conveyed to the user. It is particularly recommended that a fall-back is not used when the feature “crossed-out” is specified (see 6.1.6).

### 6.1 Emphasis

Portions of text may be visually differentiated or emphasized in character imaging. Six methods of emphasis are provided:

- weight;
- posture;
- underlining;
- blinking;
- image inversion;
- crossing-out.

All of these may be controlled by means of rendition selection using the presentation attribute “graphic rendition” and the control function SGR (Select Graphic Rendition).

The emphasis takes effect at the active position of the line box where the control function initiates it and ends at the active position of the line box where the control function specifies an end to the emphasis or at the end of the object.

NOTE Not all forms of emphasis allowed by ISO 8613 are applicable to all presentation devices. For example, in printed text, “blinking” may not be able to be represented, a second colour may be an available alternative to “faint” (decreased intensity), but “image inversion” may only be practicable by the use of appropriately designed cameo fonts.

### 6.1.1 *Weight*

This provides emphasis or de-emphasis by means of varying the contrast or intensity of the character image.

Using renditions selection, three weights are provided:

- faint (decreased intensity);
- normal intensity (neither faint nor bold);
- bold (increased intensity).

Only one of these may be in effect at any one point in the character stream, so that invoking one resets the others.

### 6.1.2 *Posture*

This causes a change of posture between an upright font and an italic font.

Using rendition selection, two renditions are provided:

- not italicized;
- italicized.

Only one of these renditions may be in effect at any one point in the character stream, so that invoking one resets the other.

### 6.1.3 *Underlining*

This provides for underlining character images. Three renditions are provided:

- not underlined;
- underlined;
- doubly underlined.

Only one of these renditions may be in effect at any one point in the character stream, so that invoking one resets the others.

NOTE For writing systems other than those that use a horizontal writing direction, underlining may be replaced by a suitable emphasis.

### 6.1.4 *Blinking*

This provides for flashing of the graphic symbol in the line box ON and OFF. Three renditions are provided:

- steady (not blinking);
- slowly blinking;
- rapidly blinking.

Only one of these renditions may be in effect at any one point in the character stream, so that invoking one resets the other.

NOTE Slow blinking should be taken as less than 150 ON/OFF cycles per minute and rapid blinking should be taken as more than 150 per minute.

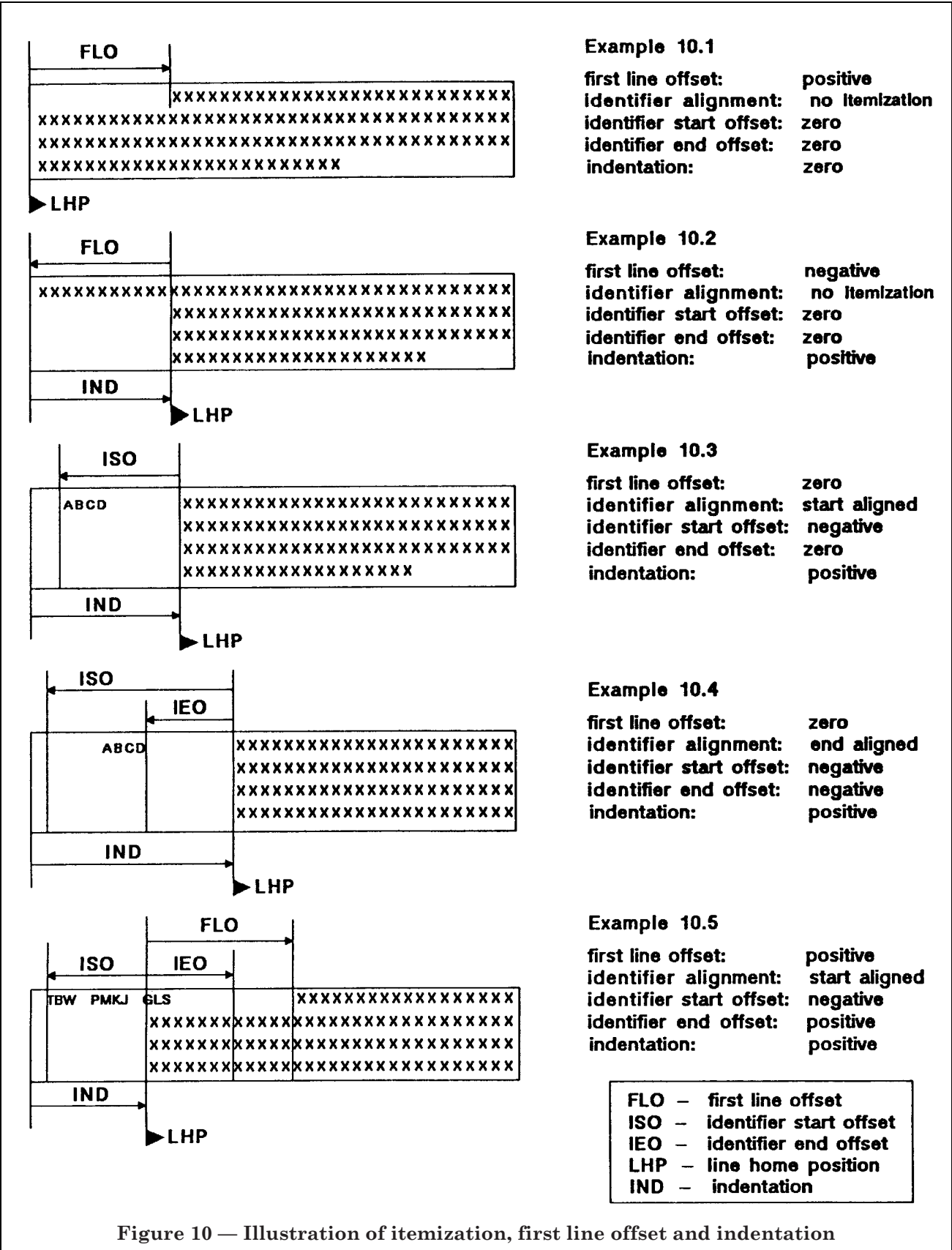
### 6.1.5 *Image inversion*

This causes the colours of the graphic symbol and the line box to be exchanged. Two renditions are provided:

- positive image;
- negative image.

Only one of these renditions may be in effect at any one point in the character stream, so that invoking one resets the other.

NOTE Simple inversion of the pels within the line box area may not be sufficient for inverting the character image.



### 6.1.6 *Crossing-out*

This allows characters to be marked for deletion. The method of crossing out is not defined but the characters should be legible. Two renditions are provided:

- not crossed-out;
- crossed-out.

Only one of these renditions may be in effect at any one point in the character stream, so that invoking one resets the other.

## 6.2 Font selection

The capability for font selection is based upon

- font specification: The required fonts and their characteristics must be specified in the document profile (see ISO 8613-4). There is no limit to the number of fonts that can be specified;
- font designation: A subset of these fonts can be designated for use within a basic component by means of the presentation attribute “character fonts”. This subset is limited to a maximum of 10 fonts;
- font invocation: Any of the fonts within this subset can be invoked by means of the presentation attribute “graphic rendition” and/or control function SGR (Select Graphic Rendition).

### 6.2.1 *Font specification*

Within the document profile, there is an attribute “fonts list”. This specifies, for each font used in the document

- a font identifier, represented by an integer, unique within the document;
- a font description consisting of a set of font attributes specifying properties that are common to all characters of the font, e.g. weight, posture, constant or variable spacing, as well as properties that are unique to the individual characters.

The font description includes a structured font name as defined in ISO 9541. The other font attributes allow a recipient to select a suitable substitute font when the font specified by the structured font name is not available.

**NOTE** It is necessary to consider that font designs are typically proprietary, and that, in addition, the recipient may have a limited repertoire. Thus, there can be no assurance that a particular font selected by the originator is actually available at the recipient's location, unless the originator has prior knowledge of the facilities available to the recipient. Since this would not generally be true in open interchange, it is necessary to specify the font and character characteristics in a manner that facilitates the selection of a near-equivalent alternative, from the repertoire of fonts available to the recipient.

### 6.2.2 *Font designation*

The presentation attribute “character fonts” enables a subset of up to 10 fonts from the set of fonts specified in the attribute “fonts list” in the document profile to be designated for use within the basic component.

Each font and its size is specified together with the method of invoking the font, i.e. as the primary or one of the alternative fonts.

If no font is designated for the primary font or for a particular alternative font, it is left to the recipient to select a suitable font for the primary or alternative font concerned.

**NOTE** The selection of a substitution font may depend upon the characteristics and capabilities of the presentation device.

### 6.2.3 *Font invocation*

The control function SGR (Select Graphic Rendition) is used to invoke the primary font or one of the 9 alternative fonts. Such an invocation can be performed anywhere within the content of a basic component.

For the initial conditions at the beginning of the basic component, the font indicated by the presentation attribute “graphic rendition” is invoked. In the absence of this attribute, the primary font is assumed to be invoked. If no fonts have been specified, the font used is implementation dependent.

When a designated font is invoked, any specification of weight or posture by an SGR parameter value, or by the presentation attribute “graphic rendition”, is ignored.

Invocation of a character set by way of code extension does not imply invocation of a font for that character set.

6.3 Subscript and superscript

The control functions PLD (Partial Line Down) and PLU (Partial Line Up) provide for characters to be imaged as subscript and superscript respectively.

Finer control of subscript and superscript rendition is provided by the control functions VPB (Line Position Backward) and VPR (Line Position Relative).

6.4 Character combinations

Two or more characters can be imaged as if they were a single symbol.

The control function GCC (Graphic Character Composition) within the content identifies the string of characters to be imaged as a single symbol.

7 Definition of character presentation attributes

Presentation attributes are applicable to basic logical and layout components. They specify the initial conditions at the start of the rendition of the content of that basic component. The content architecture associated with each basic component can have the means for changing certain of these presentation attributes by means of control functions embedded in the content; such capabilities are indicated in the definition of each presentation attribute.

The following categories of presentation attributes are defined:

- shared attributes which are available to all character content architecture classes;
- layout attributes which are available to formatted and formatted processable form character content architecture classes;
- logical attributes which are available to processable and formatted processable form character content architecture classes.

These attributes are listed in Table 1.

For each presentation attribute, a default value is defined. This value is used in the defaulting mechanism as defined in ISO 8613-2.

This clause also defines values for the content architecture attributes specific to character content architectures. These attributes are defined in ISO 8613-2.

Table 1 — Character presentation attributes

| Shared attributes               | Layout attributes    |
|---------------------------------|----------------------|
| Alignment                       | Formatting indicator |
| Character fonts                 | Initial offset       |
| Character orientation           |                      |
| Character path                  |                      |
| Character spacing               |                      |
| Code extension announcers       |                      |
| First line offset               |                      |
| Graphic character sets          |                      |
| Graphic character subrepertoire |                      |
| Graphic rendition               |                      |
| Itemization                     |                      |
| Kerning offset                  |                      |
| Line layout table               |                      |
| Line progression                |                      |
| Line spacing                    |                      |
| Pairwise kerning                |                      |

## 7.1 Shared presentation attributes

### 7.1.1 Alignment

|                     |                                                              |
|---------------------|--------------------------------------------------------------|
| CATEGORY:           | Shared                                                       |
| PERMISSIBLE VALUES: | “start-aligned”<br>“end-aligned”<br>“centred”<br>“justified” |
| DEFAULT VALUE:      | “start-aligned”                                              |
| DEFINITION:         |                                                              |

This attribute specifies the method of character alignment (see 5.2.2).

The character alignment cannot be altered within the content of a basic component. Application of the value “justified” may be suppressed by an occurrence of the control function JFY (No Justify).

This attribute can be overridden by the presentation attribute “line layout table” when any tabulation stops are specified (see 7.1.13).

### 7.1.2 Character fonts

|                     |                                                                          |
|---------------------|--------------------------------------------------------------------------|
| CATEGORY:           | Shared                                                                   |
| STRUCTURE:          | 10 pairs of two parameters: font size<br>font identifier                 |
| PERMISSIBLE VALUES: | font size: any positive integer<br>font identifier: any positive integer |
| DEFAULT VALUES:     | font size: {none is defined}<br>font identifier: {none is defined}       |
| DEFINITION:         |                                                                          |

This attribute designates up to 10 fonts which may be used within the basic component (see 5.1.2 and 6.2). These fonts are referred to as the primary font, the first alternative font, the second alternative font etc. The fonts designated must be chosen from the fonts listed in the document profile (see ISO 8613-4).

The parameter “font size” specifies the size of the font i.e. the height of the character image; its value is an integer representing the size of the font in SMUs.

The parameter “font identifier” is an integer equal to the font identifier associated with the font in the document profile attribute “fonts list”.

One of the designated fonts may be invoked at the start of the presentation of the content associated with a basic component by means of the presentation attribute “graphic rendition”, otherwise the primary font is assumed to be invoked. Also, fonts may be invoked within the content by means of the control function SGR (Select Graphic Rendition).

### 7.1.3 Character orientation

|                     |                     |
|---------------------|---------------------|
| CATEGORY:           | Shared              |
| PERMISSIBLE VALUES: | 0°, 90°, 180°, 270° |
| DEFAULT VALUE:      | 0°                  |
| DEFINITION:         |                     |

This attribute specifies the character orientation (see 5.1.3).

The character orientation cannot be altered within the content of a basic component.