INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

T.416

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(03/93)

TELEMATIC SERVICES

TERMINAL EQUIPMENTS AND PROTOCOLS FOR TELEMATIC SERVICES

INFORMATION TECHNOLOGY –

OPEN DOCUMENT ARCHITECTURE (ODA) AND INTERCHANGE FORMAT: CHARACTER CONTENT ARCHITECTURES

ITU-T Recommendation T.416

(Previously “CCITT Recommendation”)

Foreword

ITU (International Telecommunication Union) is the United Nations Specialized Agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the ITU. Some 179 member countries, 84 telecom operating entities, 145 scientific and industrial organizations and 38 international organizations participate in ITU-T which is the body which sets world telecommunications standards (Recommendations).

The approval of Recommendations by the members of ITU-T is covered by the procedure laid down in WTSC Resolution No. 1 (Helsinki, 1993). In addition, the World Telecommunication Standardization Conference (WTSC), which meets every four years, approves Recommendations submitted to it and establishes the study programme for the following period.

In some areas of information technology, which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC. The text of ITU-T Recommendation T.416 was approved by the WTSC (Helsinki, March 1-12, 1993). The identical text is also published as ISO/IEC International Standard 8613-6.

NOTE

1 As a consequence of a reform process within the International Telecommunication Union (ITU), the CCITT ceased to exist as of 28 February 1993. In its place, the ITU Telecommunication Standardization Sector (ITU-T) was created as of 1 March 1993. Similarly, in this reform process, the CCIR and the IFRB have been replaced by the Radiocommunication Sector.

In order not to delay publication of this Recommendation, no change has been made in the text to references containing the acronyms "CCITT, CCIR or IFRB" or their associated entities such as Plenary Assembly, Secretariat, etc. Future editions of this Recommendation will contain the proper terminology related to the new ITU structure.

2 In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Contents

Introduction

This ITU-T Recommendation | International Standard was prepared as a joint publication by ITU-T Study Group VII and ISO/IEC Joint Technical Committee 1.

At present, the ITU-T Set of Recommendations in the T.410 series | International Standard ISO 8613 consists of:

- Introduction and general principles;
- Document structures;
- Document profile;
- Open document interchange formats;
- Character content architectures;
- Raster graphics content architectures;
- Geometric graphics content architectures;
- Formal specification of the Open Document Architecture (FODA)

  (The formal specification is applicable to ISO/IEC 8613 only).

Further Recommendations | International Standards may be added to this series of ITU-T Recommendations | International Standards.

Development of this set of ITU-T Recommendations | International Standards was originally in parallel with ECMA-101 standard: *Open document architecture.*

This series of ITU-T Recommendations | International Standards is a new edition of the CCITT T.410-Series of Recommendations (1998) and ISO 8613 (1989).

Significant technical changes are the inclusion of the following amendments as agreed by ITU-T and ISO/IEC:

- Alternative Representation;
- Annex on use of MHS/MOTIS;
- Colour;
- Conformance Testing Annex;
- Document Application Profile, Proforma and Notation;
- Security;
- Streams;
- Styles;
- Tiled Raster Graphics.

In addition, a number of technical corrigenda have been applied to this ITU-T Recommendation | International Standard.

This ITU-T Recommendation | International Standard contains four annexes:

- Annex A (integral): SGML representation of character content-specific attributes for ODL;
- Annex B (non-integral): Summary of content architecture classes Summary of content architecture classes;
- Annex C (non-integral): Coded representations of control functions;
- Annex D (non-integral): Summary of object identifiers.

INTERNATIONAL STANDARD

ITU-T RECOMMENDATION

INFORMATION TECHNOLOGY – OPEN DOCUMENT ARCHITECTURE (ODA) AND INTERCHANGE FORMAT: CHARACTER CONTENT ARCHITECTURES

1 Scope

The purpose of the ITU-T Rec. T.410-Series | ISO/IEC 8613 is to facilitate the interchange of documents.

In the context of these Recommendations | International Standards, 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, raster graphic elements and geometric graphic elements, all potentially within one document.

NOTE – These Recommendations | International Standards are designed to allow for extensions, including hypermedia features, spreadsheets and additional types of content such as audio and video.

In addition to the content types defined in these Recommendations | International Standards, ODA also provides for arbitrary content types to be included in documents.

These Recommendations | International Standards apply to the interchange of documents by means of data communications or the exchange of storage media.

These Recommendations | International Standards provide 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.

These Recommendations | International Standards also provide for the interchange of ODA information structures used for the processing of interchanged documents.

This ITU-T Recommendation | International Standard:

- defines a character content architecture that can be used in conjunction with the document architecture defined in ITU-T Rec. T.412 | ISO/IEC 8613-2;
- defines the internal structure of content conforming to this character content architecture;
- defines those aspects of rendition applicable to the presentation of character content;
- defines the presentation and content portion attributes applicable to this character content architecture;
- describes a character content layout process which, together with the document processing model described in ITU-T Rec. T.412 | ISO/IEC 8613-2, determines the layout of character content in basic layout objects.

2 Normative references

The following ITU-T Recommendations and International Standards contain provisions which, through references, in this text, constitute provisions of this Recommendation | International Standard. At the time of publication, the editions indicated were valid. All Recommendations and Standards are subject to revision, and parties to agreements based on this Recommendation | International Standard are encouraged to investigate the possibility of applying the most recent edition of the Recommendations and Standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards. The ITU-T Secretariat maintains a list of currently valid ITU-T Recommendations.

2.1 Identical Recommendations | International Standards

- ITU-T Recommendation T.411 (1992) | ISO/IEC 8613-1:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Introduction and general principles*.
- ITU-T Recommendation T.412 (1992) | ISO/IEC 8613-2:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Document structures*.
- ITU-T Recommendation T.414 (1992) | ISO/IEC 8613-4:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Document profile*.
- ITU-T Recommendation T.415 (1992) | ISO/IEC 8613-5:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Open document interchange format*.
- ITU-T Recommendation T.417 (1992) | ISO/IEC 8613-7:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Raster graphics content architectures*.
- ITU-T Recommendation T.418 (1992) | ISO/IEC 8613-8:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Geometric graphics content architectures*.

2.2 Paired Recommendations | International Standards equivalent in technical content

- CCITT Recommendation X.208 (1988), *Specification of Abstract Syntax Notation One (ASN.1)*;

  ISO/IEC 8824:1990, *Information technology – Open Systems Interconnection – Specification of Abstract Syntax Notation One (ASN.1)*.

2.3 Additional references

- ISO 2022:1986, *Information processing – ISO 7-bit and 8-bit coded character sets – Code extension techniques.*
- ISO 2375:1985, *Data processing – Procedure for registration of escape sequences*.
- ISO/IEC 6429:1992, *Information technology – Control functions for coded character sets*.
- ISO 6937-1:1983, *Information processing – Coded character sets for text communication – Part 1: General introduction*.
- ISO 6937-2:1983, *Information processing – Coded character sets for text communication – Part 2: Latin alphabetic and non-alphabetic graphic characters*.
- ISO/IEC 7350:1991, *Information technology – Registration of repertoires of graphic characters from ISO 10367*.
- ISO/IEC 8613-10:1991, *Information technology – Text and office systems – Open Document Architecture (ODA) and interchange format – Part 10: Formal specifications*.
- ISO 8879:1986, *Information processing – Text and office systems – Standard Generalized Markup Language (SGML).*
- ISO/IEC 9541-1:1991, *Information technology – Font information interchange: Part 1 – Architecture*.

3 Definitions

For the purposes of this Recommendation | International Standard, the definitions given in ITU-T Rec. T.411 | ISO/IEC 8613-1 apply.

4 Abbreviations

BPH	Break Permitted Here
BS	Backspace
CR	Carriage Return
GCC	Graphic Character Composition
HPB	Character Position Backward
HPR	Character Position Relative
IGS	Identify Graphic Subrepertoire
JFY	No Justify
LF	Line Feed
NBH	No Break Here
PLD	Partial Line Up
PLU	Partial Line Down
PTX	Parallel Texts
SACS	Set Additional Character Separation
SCS	Set Character Spacing
SGR	Select Graphic Rendition
SHS	Select Character Spacing
SLS	Set Line Spacing
SOOS	Start Of Original String
SOS	Start Of String
SRCS	Set Reduced Character Separation
SRS	Start Reverse String
ST	String Terminator
SSW	Set Space Width
STAB	Selective Tabulation
SUB	Substitute Character
SVS	Select Line Spacing
VPB	Line Position Backward
VPR	Line Position Relative

5 Conventions

For the purpose of this Recommendation | International Standard, the conventions given in ITU-T Rec. T.411 | ISO/IEC 8613-1 apply.

6 General principles

6.1 Content architecture classes

This part of ITU-T Rec. T.410-Series | ISO/IEC 8613 defines three classes of character content architecture:

– 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.

6.2 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.

6.3 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 of the content layout and imaging processes.

All presentation attributes in this Specification are defaultable.

6.4 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 and, in the case of basic logical components, any basic logical component concatenated to this component (see 14.2.3).

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 ITU-T Rec. T.410-Series | ISO/IEC 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 of 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 or deleted as the result of a content layout process (see clause 14). The inserted 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 inserted character sequence.

6.5 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 12 and 13.1.17).

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.

6.6 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 14.2.1.3.2).

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.

6.7 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. If ODIF is used as the interchange format, the coding shall be performed for an 8-bit environment.

NOTE 1 – This is equivalent to presuming a code extension announcer value of ESC 2/0 4/7 for the C1 set.

If ODL is used in the interchange format, the coding should where possible be performed for an 8-bit environment. (The use of ODL is applicable to ISO/IEC 8613 only.)

NOTE 2 – If an ODL application requires coding for a 7-bit environment, this should be specified by the document application profile, preferably also stating the appropriate code extension announcer(s).

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

6.8 Internal structure

6.8.1 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 end of the content of a basic component implicitly terminates the last line.

6.8.2 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 ITU-T Rec. T.410-Series | ISO/IEC 8613.

In order to image content in this form, it is necessary to apply a content layout process (see clause 14) to the content which converts the processable content into formatted content (see 6.8.1) or into formatted processable content (see 6.8.3).

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 9 and 13). In addition, the character SPACE is regarded both as 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.

At the end of the content of a basic logical component to which another basic logical component is concatenated (see ITU-T Rec. T.412 | ISO/IEC 8613-2), 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.

6.8.3 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 as well as deleted logical content.

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 inserted within those delimiters.

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

At the end of the content of a basic logical component to which another basic logical component is concatenated (see ITU-T Rec. T.412 | ISO/IEC 8613-2), 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.

7 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/IEC 9541-1. This clause also caters for the positioning of characters pertaining to different fonts within the same basic layout object.

7.1 Basic concepts

7.1.1 Character fonts

In the context of this part of ITU-T Rec. T.410-Series | ISO/IEC 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/IEC 9541-1.

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 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 8.

7.1.2 Directions

In the context of this part of ITU-T Rec. T.410-Series | ISO/IEC 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).

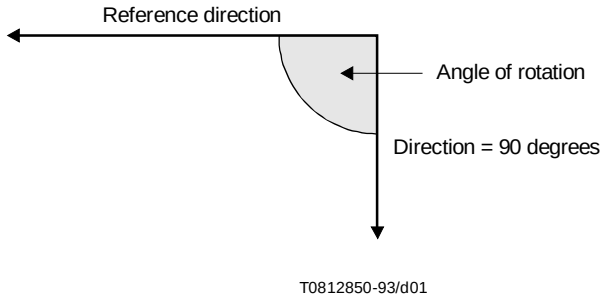


Figure 1 – Example of direction

The *character path* is the direction of progression of successive character images within a line box (defined in 7.1.7) 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 7.1.3) relative to the character path.

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

7.1.3 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).

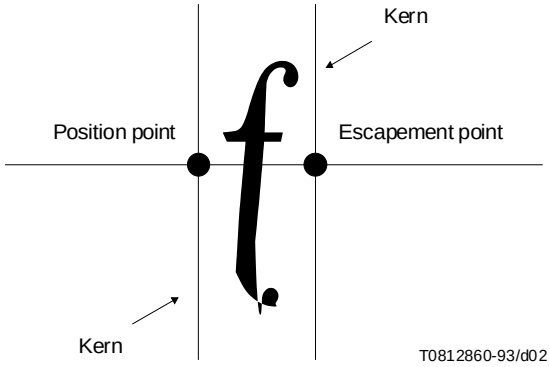
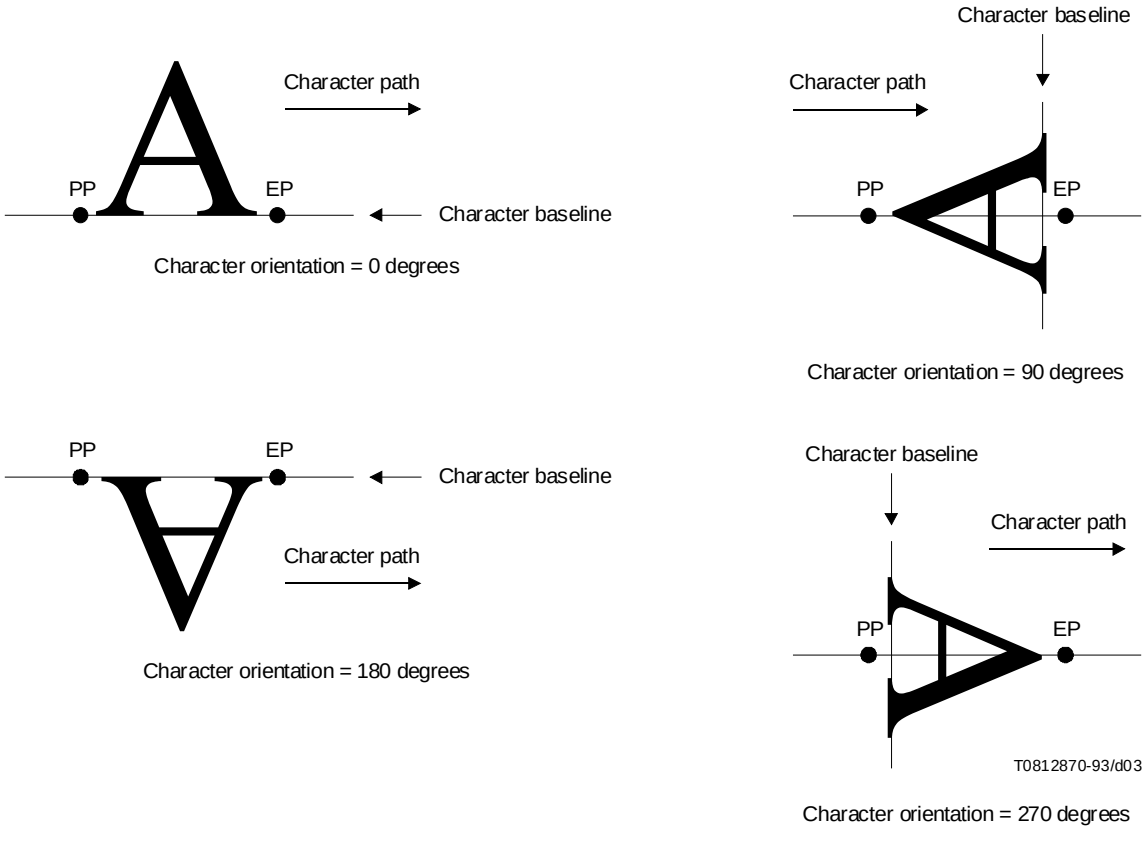


Figure 2 – Illustration of kerns



PP Position point
EP Escapement point

Figure 3 – Reference points for character image positioning

7.1.4 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 7.2.1.

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.

7.1.5 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 13) 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 7.2.1).

If the next character is a control function, this may cause the active position to move to another point within the positioning area.

7.1.6 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.

7.1.7 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 11). 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.

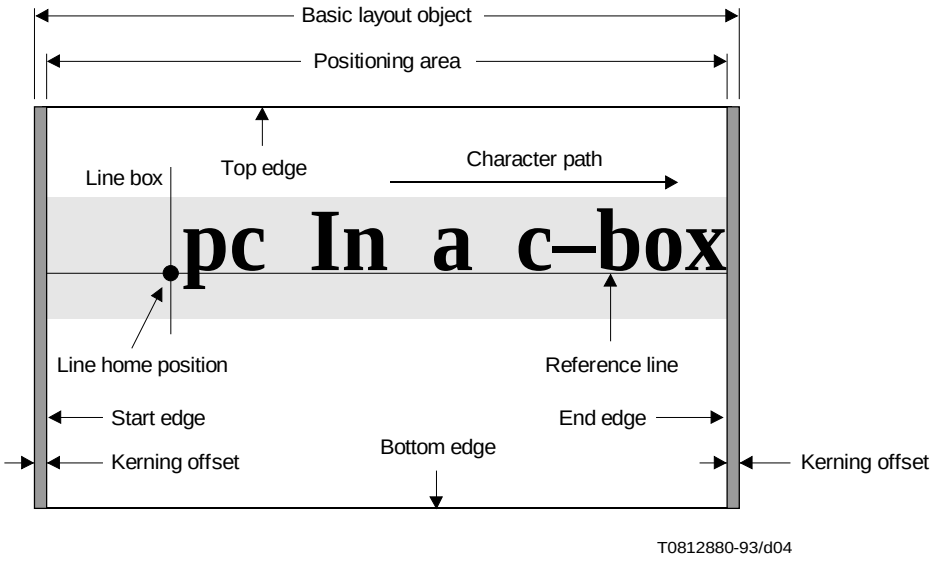


Figure 4 – Illustration of character positioning concepts

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 takes 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 do, line progression d270 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/IEC 9541-1.

If no graphic characters are present in a line box, the font that is used to determine the backward and forward extents shall be the last font that is invoked in the line box, if any, or, otherwise, the font that was invoked or was assumed to be invoked (see 8.2.3) before the line box was started.

7.2 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.

7.2.1 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 7.1.4). 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).

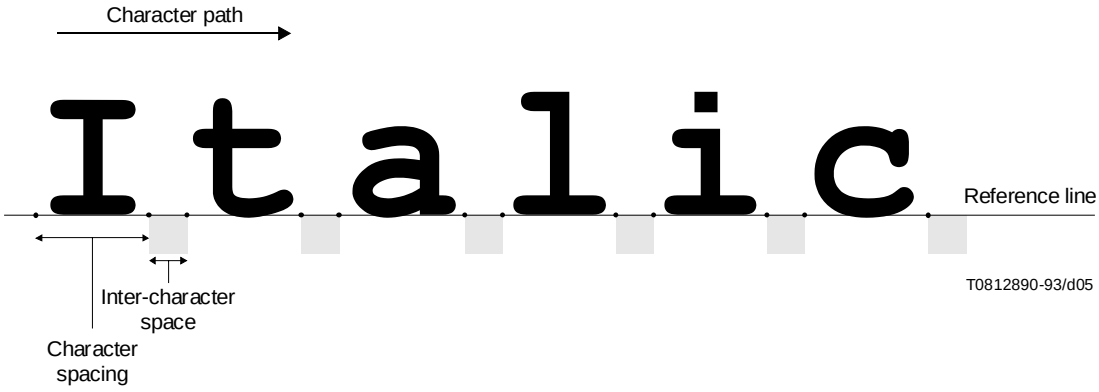


Figure 5 – Spacing for constant spacing font

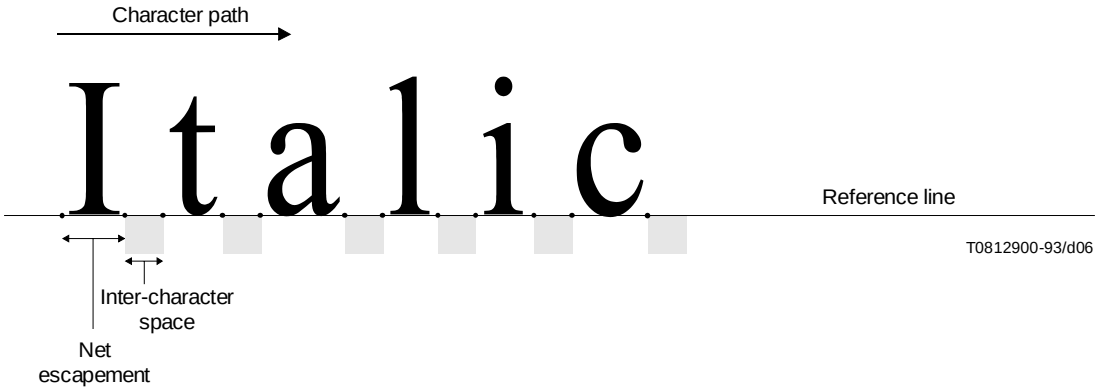


Figure 6 – Spacing for variable spacing font

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.

7.2.2 Alignment

The character images are positioned within a line box in accordance with the attribute “alignment” 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 made coincident with the end edge of the positioning area by appropriately setting the space width and/or the inter-character space.

7.2.3 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 equidistant 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.

7.2.4 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 shall 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.

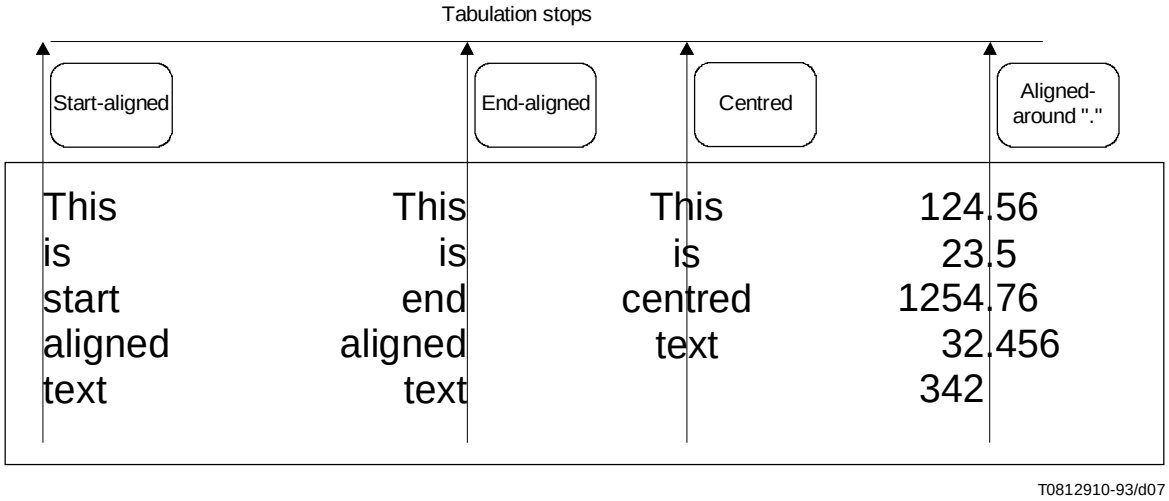


Figure 7 – Tabulation

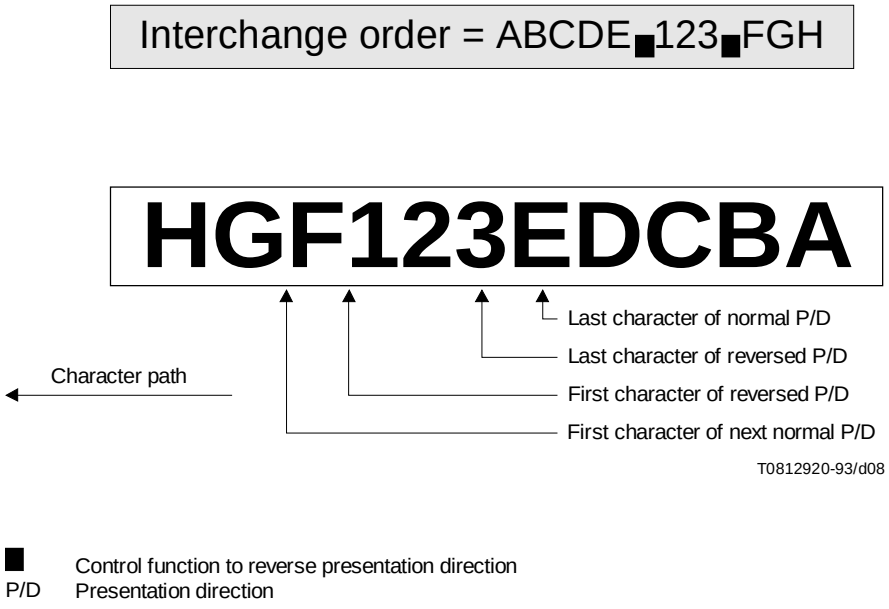
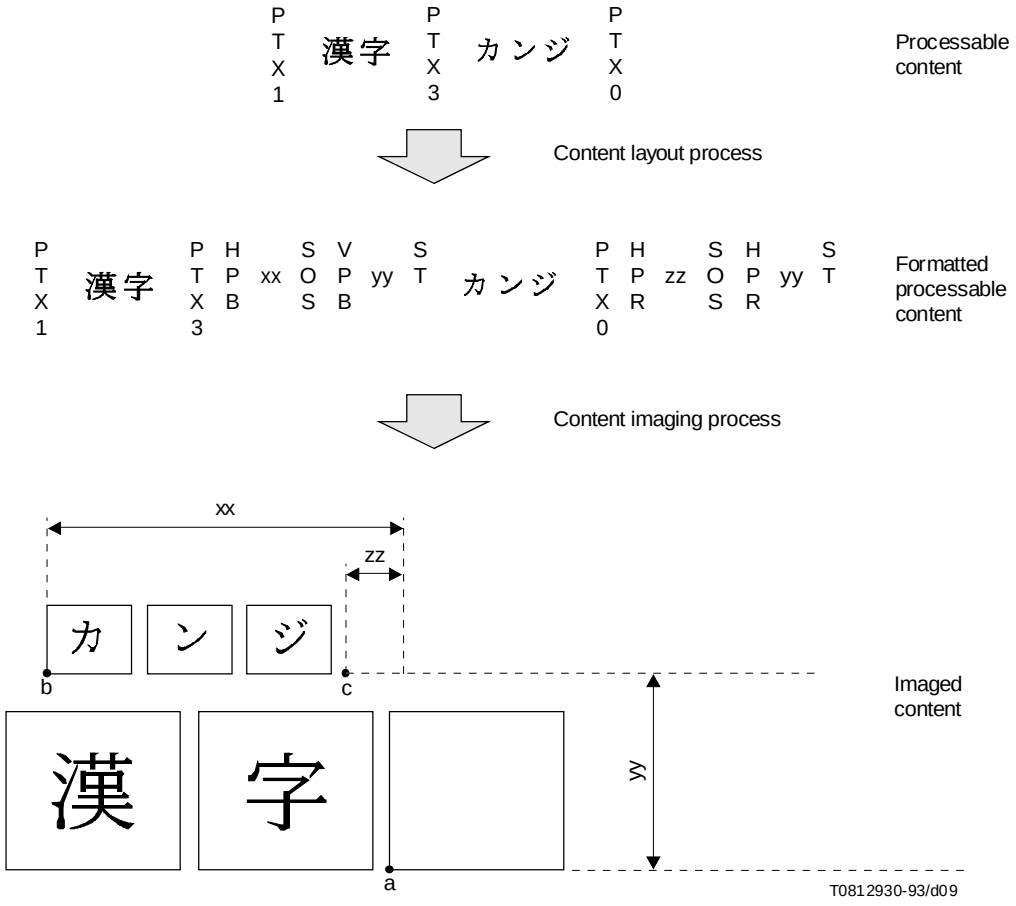


Figure 8 – Character ordering

7.2.5 Parallel annotation

Two character strings can be presented in parallel, by delimiting them such that the second string is used to indicate the pronunciation and interpretation of the first string. This feature is intended to be used for 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 centering 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.



HPB, HPR, VPB, VPR, SOS, ST, PTX — Control functions (see clause 13)
xx, zz, yy — Parameter values determined by the content layout process

The current reference point moves from “a” to “b” to “c” and then returns to “a”.

Figure 9 – An example of the specification of Japanese Ruby

7.2.6 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.

7.2.7 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.

7.2.8 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.

7.2.9 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:

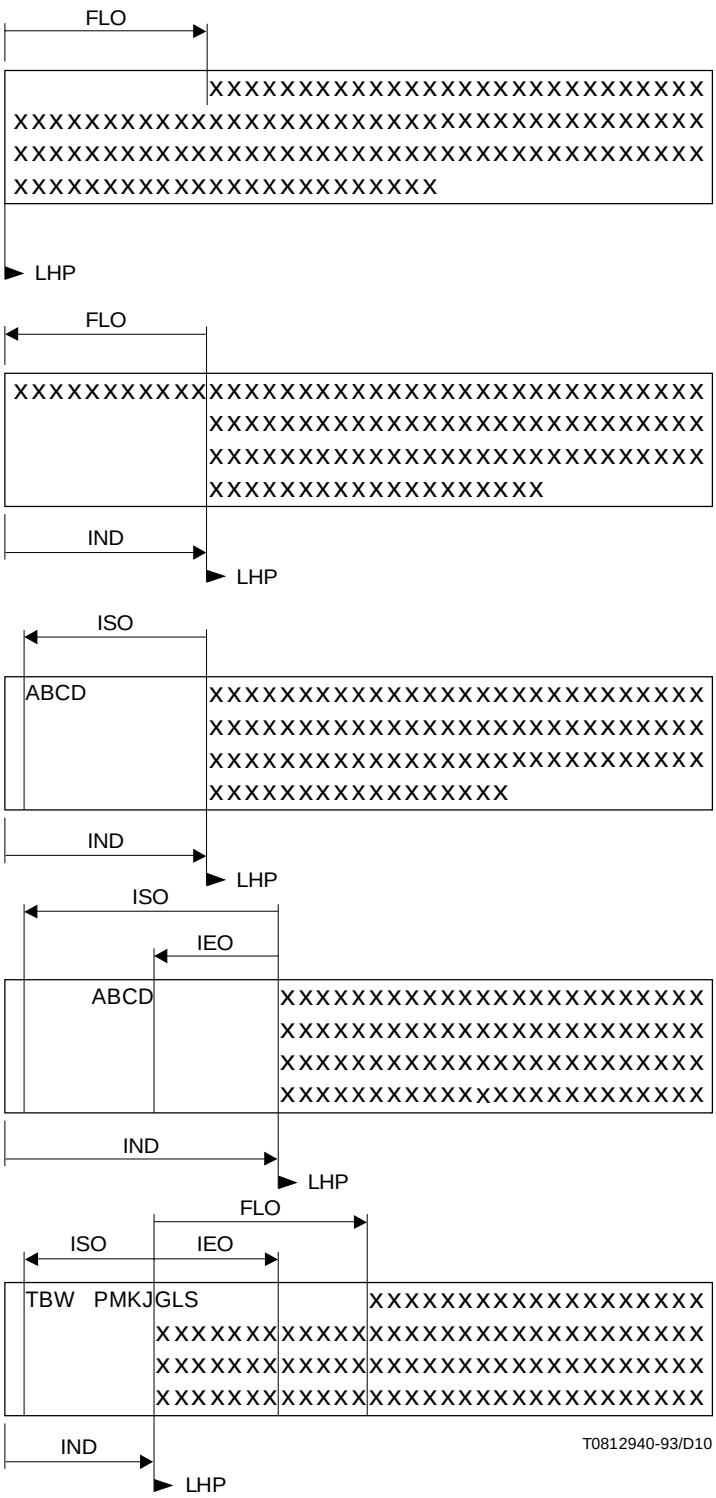
- *Start-aligned* – The position point of the first character image of the item identifier is placed at the start offset.
- *End-aligned* – The escapement point of the last character image of the item identifier is placed at the end offset.

7.3 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 9.2.2 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.



Example 10.1

First line offset:	positive
Identifier alignment:	no itemization
Identifier start offset:	zero
Identifier end offset:	zero
Indentation:	zero

Example 10.2

First line offset:	negative
Identifier alignment:	no itemization
Identifier start offset:	zero
Identifier end offset:	zero
Indentation:	positive

Example 10.3

First line offset:	zero
Identifier alignment:	start aligned
Identifier start offset:	negative
Identifier end offset:	zero
Indentation:	positive

Example 10.4

First line offset:	zero
Identifier alignment:	end aligned
Identifier start offset:	negative
Identifier end offset:	negative
Indentation:	positive

Example 10.5

First line offset:	positive
Identifier alignment:	start aligned
Identifier start offset:	negative
Identifier end offset:	positive
Indentation:	positive

FLO First line offset
ISO Identifier start offset
IEO Identifier end offset
LHP Line home position
IND Indentation

Figure 10 – Illustration of itemization, first line offset and indentation

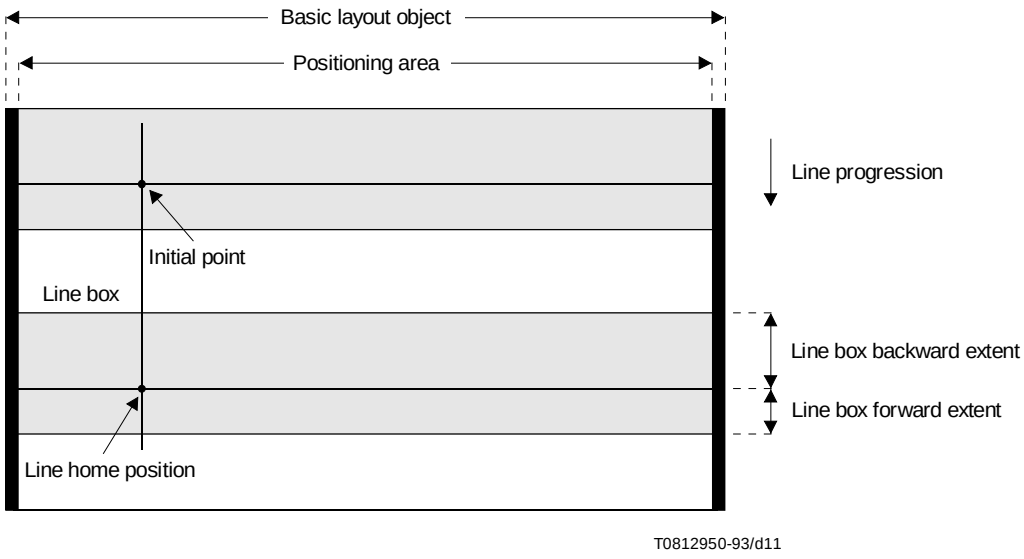


Figure 11 – Illustration of character positioning concepts

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 ITU-T Rec. T.410-Series | ISO/IEC 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.

8 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.

NOTES

1 Certain emphasis (e.g. ‘weight’, ‘posture’) and subscript/superscript rendition may be achieved by font selection.

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, ITU-T Rec. T.410-Series | ISO/IEC 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 8.1.6).

8.1 Emphasis

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

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

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 ITU-T Rec. T.410-Series | ISO/IEC 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), and `image inversion' may only be practicable by the use of appropriately designed cameo fonts.

8.1.1 Weight

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

Using rendition 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.

8.1.2 Posture

This feature provides for 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.

8.1.3 Underlining

This feature 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.

8.1.4 Blinking

This feature 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 others.

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 cycles per minute.

8.1.5 Image inversion

This feature provides for the colours of the graphic symbol and the line box to be exchanged.

Two renditions are provided:

- positive image;
- negative image.

If the character content is coloured, then the rendition ‘negative image’ causes the character foreground and character background colours to be exchanged.

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 .

8.1.6 Crossing-out

This feature 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.

8.1.7 Colour

This feature provides for characters, words or other elements of a character content portion to be presented in a different colour. More than one emphasis colour is permitted in a single character content portion. The character content architecture can specify colour specifications in RGB or CMY(K) colour spaces (see Figure 12). Colours may be specified in either direct mode or indexed mode.

When imaging text, there are two coloured elements for each character: the character image and the character image background (see Figure 13). For each character, the character image is imaged in a colour referred to as the *character foreground colour* and the character image background is imaged in a colour referred to as the character background colour.

The use of character background colours can be described by means of the following rules. The character background colour is applied to rectangular areas above and between the escapement and position points of successive character images in a line box termed *segments*.

Orthogonal to the character path, the extent of each line box delimits that segment with the character background colour.

For each line box, in the direction of the character path, the first segment begins at the line home position and the last segment ends at either the hard line terminator or the end of the line box. The specification of a new value for the character background colour or the emphasis "image inversion" (see 8.1.5) within the content terminates the current segment at the escapement point of the previous character and starts a new segment at the position point of the next character.

Usage of the control functions HPB (Character Position Backward) and HPR (Character Position Relative) terminates the current segment before moving the active position and starts a new segment at the new active position.

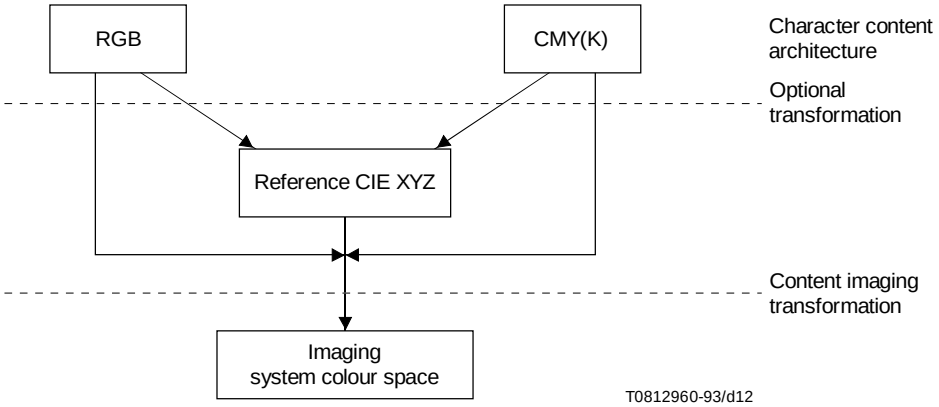


Figure 12 – Relationships between colour space

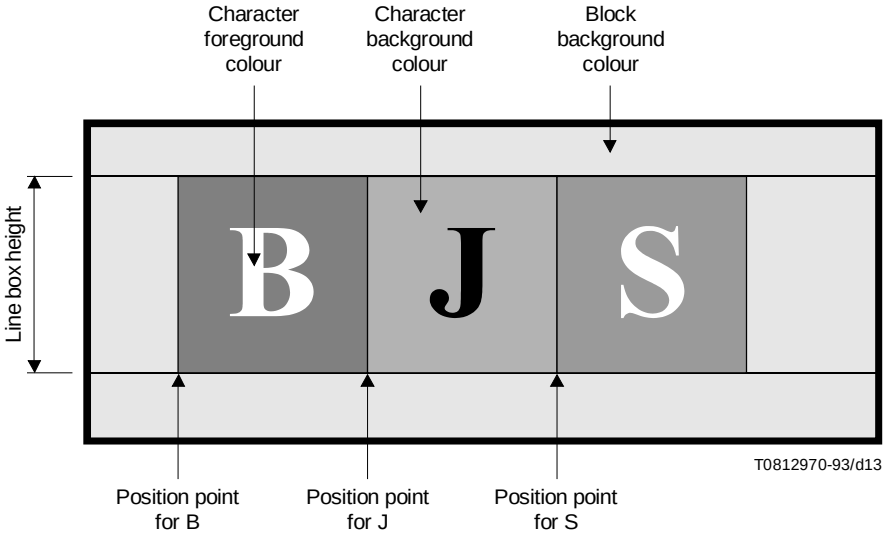


Figure 13 – Areas affected by colour emphasis

Initially, the values for the character foreground colour and the character background colour are those specified by the attributes “content foreground colour” and “colour background colour” of the basic object with the associated character content. These attributes are defined in ITU-T Rec. T.412 | ISO/IEC 8613-2. The two colours can independently be altered within the content information using the control function SGR (Select Graphic Rendition) with values 30 to 38 for the character foreground colour and values 40 to 48 for the character background colour. If negative image is in effect these two colours are exchanged.

8.2 Font selection

The capability for font selection is based upon:

- *font specification* – The required fonts and their characteristics shall be specified in the document profile (see ITU-T Rec. T.414 | ISO/IEC 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 ten 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).

8.2.1 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/IEC 9541-1. 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.

8.2.2 Font designation

The presentation attribute "Character fonts" enables a subset of up to ten 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.

NOTES

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

2 When selecting a suitable font, it is recommended to consider the specification of character spacing and line spacing for its dimension and the specification of emphasis for its design.

3 It is not recommended to use proportional fonts for the non-designated font in a layout process that generates a formatted document that is to be interchanged.

8.2.3 Font invocation

The control function SGR (Select Graphic Rendition) is used to invoke the primary font or one of the nine 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 font has 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 means of a code extension control function does not imply invocation of a font for that character set.

8.3 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).

8.4 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.

9 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.

For each presentation attribute, a default value is defined. This value is used in the defaulting mechanism as defined in ITU-T Rec. T.412 | ISO/IEC 8613-2.

This clause also defines values specific to character content architectures for the attribute “content architecture class” as defined in ITU-T Rec. T.412 | ISO/IEC 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	Logical attributes
Graphic character subrepertoire	
Graphic rendition	Indentation
Itemization	Proportional line spacing
Kerning offset	Orphan size
Line layout table	Widow size
Line progression	
Line spacing	
Pairwise kerning	

9.1 Shared presentation attributes

9.1.1 Alignment

CATEGORY: Shared

PERMISSIBLE VALUES:
‘start-aligned’
‘end-aligned’
‘centred’
‘justified’

DEFAULT VALUE: ‘start-aligned’

DEFINITION:

This attribute specifies the method of character alignment (see 7.2.2).

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 9.1.13).

9.1.2 Character fonts

CATEGORY: Shared

STRUCTURE: ten optional parameters:
primary font
first alternative font
second alternative font
third alternative font
fourth alternative font
fifth alternative font
sixth alternative font
seventh alternative font
eighth alternative font
ninth alternative font

Each parameter has two subparameters:
font size
font identifier

PERMISSIBLE VALUES:
font size: any positive integer
font identifier: any positive integer

DEFAULT VALUE: no fonts are designated

DEFINITION:

This attribute designates up to ten fonts which may be used within the basic component (see 7.1.2 and 8.2). These fonts are referred to as the primary font, the first alternative font, the second alternative font etc. The fonts designated shall be chosen from the fonts listed in the document profile (see ITU-T Rec. T.414 | ISO/IEC 8613-4).

The subparameter “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 subparameter “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).

9.1.3 Character orientation

CATEGORY: Shared

PERMISSIBLE VALUES: d0
d90
d180
d270

DEFAULT VALUE: d0

DEFINITION:

This attribute specifies the character orientation (see 7.1.3).

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

9.1.4 Character path

CATEGORY: Shared

PERMISSIBLE VALUES: d0
d90
d180

DEFAULT VALUE: d0

DEFINITION:

This attribute specifies the character path (see 7.1.2).

The character path cannot be altered within the content of a basic component. However, local changes of the relationship between the imaging order of the characters and the interchange order can be specified by the control function SRS (Start Reverse String).

9.1.5 Character spacing

CATEGORY: Shared

PERMISSIBLE VALUES: any positive integer

DEFAULT VALUE: the equivalent of 120 BMUs

DEFINITION:

This attribute specifies the character spacing which applies at the beginning of the basic component (see 7.1.4).

The value of this attribute is an integer specifying the distance in SMUs.

The character spacing value is only used while a constant spacing font is in use; it has no effect while a variable spacing font is in use.

It can be altered within the content of a basic component by means of the control functions SHS (Select Character Spacing) or SCS (Set Character Spacing).

9.1.6 Code extension announcers

CATEGORY: Shared

PERMISSIBLE VALUES: The value of this attribute consists of a string of escape sequences, in accordance with ISO 2022, to announce the use of code extension features. The permissible final characters of the escape sequences are: 4/1, 4/3, 4/4, 4/7, 4/9, 4/11, 4/12, 4/13, 4/14, 5/0, 5/3, 5/5, 5/10 and 5/11

DEFAULT VALUE: Escape sequences with final characters: 5/0, 5/5, 4/7, 4/9 and 4/11.

DEFINITION:

This attribute announces the code extension features allowed in the basic component.

The set of code extension features announced by this attribute cannot be altered within the content of a basic component.

The default value announces an 8-bit environment with an 8-bit code, the use of the G0 and G2 sets, the G2 set being invoked in columns 10 to 15, where the character sets may comprise 94 and/or 96 characters and each C1 control function is represented by a single bit combination from columns 08 and 09.

9.1.7 First line offset

CATEGORY: Shared

PERMISSIBLE VALUES: any integer

DEFAULT VALUE: 0

DEFINITION:

This attribute specifies an offset along the character path from the line home position, measured in SMUs (see 7.2.8).

The offset may be positive (in the direction of the character path), negative (in the direction opposite to the character path) or zero.

The position identified by application of this offset to the line home position is used instead of the line home position for the purposes of formatting and imaging the first line of the basic layout object in which the content of the basic logical component is laid out.

The value of the presentation attribute “indentation” (see 9.3.1) shall be set such that the resulting displacement of the line home position from the start edge of the positioning area is sufficient to enable overhanging characters to be imaged within the positioning area.

The first line offset cannot be altered within the content of a basic component.

9.1.8 Graphic character sets

CATEGORY: Shared

PERMISSIBLE VALUES: The value of this attribute consists of a string of escape sequences, in accordance with ISO 2022 and the register of ISO 2375, to designate one or more graphic character sets, and any locking shift functions needed to invoke these character sets.

DEFAULT VALUE: The escapade sequences and shift functions designating and invoking the primary character set of ISO 6937-2 as the G0 set and the supplementary character set of ISO 6937-2 as the G2 set in columns 10 to 15.

DEFINITION:

This attribute specifies the graphic character set(s) designated and/or invoked at the beginning of the basic component.

Other graphic character sets can be designated and/or invoked within the content of a basic component by means of the appropriate code extension escape sequences and shift functions.

9.1.9 Graphic character subrepertoire

CATEGORY: Shared

PERMISSIBLE VALUES: The value of this attribute is either 0 or the identifier of a subrepertoire assigned in the register of ISO/IEC 7350. The value 0 identifies the full repertoire of the graphic character sets that are designated at the beginning of the basic component.

DEFAULT VALUE: 0

DEFINITION:

This attribute identifies the subrepertoire of the graphic character repertoire of ISO 6937 used at the beginning of the basic component.

This attribute is only applicable if the graphic character sets of ISO 6937 are used.

Other graphic character subrepertoires can be invoked within the content of a basic component by means of the control function IGS (Identify Graphic Subrepertoire).

9.1.10 Graphic rendition

CATEGORY: Shared

PERMISSIBLE VALUES: A sequence of one or more integers corresponding to parameter values of the control function SGR (Select Graphic Rendition).

DEFAULT VALUE: 0

DEFINITION:

This attribute specifies the rendition parameters for font, underlining, etc. which apply at the beginning of the basic component (see 8.1). The SGR values 30 to 37 and 40 to 47 are not permitted in this attribute.

The graphic rendition can be altered within the content of a basic component by means of the control function SGR (Select Graphic Rendition).

NOTE – If more than one graphic rendition parameter is encoded, then it is the user's responsibility to ensure that they are consistent.

9.1.11 Itemization

CATEGORY:	Shared	
STRUCTURE:	One required parameter:	identifier alignment
	Two optional parameters:	identifier start offset, identifier end offset
PERMISSIBLE VALUES:	identifier alignment:	'no itemization' 'start-aligned' 'end-aligned'
	identifier start offset:	any integer value
	identifier end offset:	any integer value
DEFAULT VALUES:	identifier alignment:	'no itemization'
	identifier start offset:	the distance from the line home position to the start of the positioning area
	identifier end offset:	0

DEFINITION:

This attribute specifies the placement of an item identifier which may begin the basic component (see 7.2.9).

If the value of the parameter "identifier alignment" is 'no itemization', then no item identifier is present.

For other values of the parameter "identifier alignment", the parameters "identifier start offset" and "identifier end offset" specify offsets, in SMUs, from the line home position along the character path which identify, respectively, the start edge and end edge of a portion of the line box in which the item identifier will be formatted. These offsets may be positive (in the direction of the character path), negative (in the direction opposite to the character path) or zero.

The item identifier consists of all graphic characters preceding the first occurrence of the control function CR (Carriage Return) in the basic component. The content of the basic component following that CR shall be formatted as specified by the presentation attribute "first line offset".

Values of the parameter "identifier alignment" other than 'no itemization' specify the method of character alignment for the item identifier.

The value of the presentation attribute "indentation" (see 9.3.1) shall be set such that the resulting displacement of the line home position from the start edge of the positioning area is sufficient to enable the item identifier to be imaged within the positioning area.

Itemization cannot be altered within the content of a basic component.

9.1.12 Kerning offset

CATEGORY:	Shared	
STRUCTURE:	two parameters:	start edge offset end edge offset
PERMISSIBLE VALUE:	start edge offset: end edge offset:	any non-negative integer any non-negative integer
DEFAULT VALUES:	start edge offset: end edge offset:	0 0

DEFINITION:

This attribute specifies the kerning offset as a pair of integer values in SMUs (see 7.1.6). The parameter “start edge offset” specifies the distance from the edge of the basic layout object to the start edge of the positioning area. The parameter “end edge offset” specifies the distance from the edge of the basic layout object to the end edge of the positioning area.

The kerning offset cannot be altered within the content of a basic component.

9.1.13 Line layout table

CATEGORY:	Shared	
STRUCTURE:	A set of entries where each entry consists of three required parameters:	tab reference tab position alignment
	and one optional parameter:	alignment string
PERMISSIBLE VALUES:	tab reference:	A string of one to four decimal digits used as the reference parameter in the control function STAB (Selective Tabulation).
	tab position:	any non-negative integer
	alignment:	‘start-aligned’ ‘end-aligned’ ‘centred’ ‘aligned-around’
	alignment string:	Graphic characters from the set of graphic elements specified by the presentation attributes “graphic character sets” and “graphic character subrepertoire”.
DEFAULT VALUE:	The default value is such that no tabulation stop is defined.	

DEFINITION:

This attribute specifies the positions and types of a set of tabulation stops (see 7.2.3). The value of the parameter “tab position” specifies the distance in SMUs, in the direction of the character path, from the start edge of the positioning area to the tabulation stop.

If the value of the parameter “alignment” is ‘aligned-around’, then the parameter “alignment string” shall be specified otherwise this parameter shall not be specified.

When this presentation attribute specifies any tabulation stops, the presentation attribute “alignment” is assumed to have the value ‘start-aligned’ (see 9.1.1).

The tabulation stops cannot be altered within the content of a basic component.

9.1.14 Line progression

CATEGORY: Shared

PERMISSIBLE VALUES: d90
d270

DEFAULT VALUE: d270

DEFINITION:

This attribute specifies the line progression (see 7.1.2).

The line progression cannot be altered within the content of a basic component.

9.1.15 Line spacing

CATEGORY: Shared

PERMISSIBLE VALUES: any positive integer

DEFAULT VALUE: the equivalent of 200 BMUs

DEFINITION:

This attribute only applies when proportional line spacing is not to be done by the content layout process. In this case, it specifies the line spacing which applies at the beginning of the basic component (see 7.3).

The value of this attribute is an integer specifying the distance in SMUs.

The line spacing can be altered within the content of a basic component by means of the control functions SVS (Select Line Spacing) or SLS (Set Line Spacing).

9.1.16 Pairwise kerning

CATEGORY: Shared

PERMISSIBLE VALUES: ‘yes’
‘no’

DEFAULT VALUE: ‘no’

DEFINITION:

This attribute specifies whether pairwise kerning should be performed on the content during the formatting process (see 7.2.7).

The value ‘yes’ indicates that the formatting process should perform, if possible, pairwise kerning on the content.

The value ‘no’ specifies that pairwise kerning should not be performed on the content.

The specification of pairwise kerning cannot be altered within the content of a basic component.

9.2 Layout presentation attributes

9.2.1 Formatting indicator

CATEGORY: Layout

PERMISSIBLE VALUES: ‘yes’
‘no’

DEFAULT VALUE: ‘no’

DEFINITION:

This attribute specifies whether the content of a basic component has been formatted by a content layout process or not.

The value 'yes' indicates that the content of the basic layout component concerned contains layout control functions representing the effects of any use of the control function STAB (Selective Tabulation) or of the presentation attributes "alignment", "first line offset", "itemization" and/or "pairwise kerning" (see 14.2.1.3.1).

The specification of alignment cannot be altered within the content of a basic component.

NOTE – In interchange, a recipient can take advantage of this attribute only if he has a character font that is similar to that of the sender, i.e. a font that has the same width for each character as the sender's font.

9.2.2 Initial offset

CATEGORY: Layout

STRUCTURE: two parameters: horizontal coordinate
vertical coordinate

PERMISSIBLE VALUES: horizontal coordinate: any non-negative integer
vertical coordinate: any non-negative integer

DEFAULT VALUES: The default values of this attribute depend on character path, line progression and line spacing as defined in Table 2.

Table 2 – Default values of the presentation attribute "initial offset"

Character path (degrees)	Line progression (degrees)	Horizontal coordinate	Vertical coordinate
d0	d270 d90	0 0	S H-S
d90	d270 d90	S W-S	H H
d180	d270 d90	W W	H-S S
d270	d270 d90	W-S S	0 0

W Horizontal dimension of the basic component

H Vertical dimension of the basic component

S Backward extent of the first line box

DEFINITION:

This attribute specifies the position of the initial point (see 7.3).

The values of the parameters "horizontal coordinate" and "vertical coordinate" specify the horizontal and vertical co-ordinates, in SMUs, of the initial point relative to the top left corner of the basic component. The horizontal coordinate is measured positively from the vertical axis to the right and the vertical coordinate is measured positively from the horizontal axis downwards.

The position of the initial point cannot be altered within the content of a basic component.

9.3 Logical presentation attributes

9.3.1 Indentation

CATEGORY: Logical

PERMISSIBLE VALUES: any non-negative integer

DEFAULT VALUE: 0

DEFINITION:

This attribute specifies the distance, in the direction of character path, from the start edge of the positioning area to the initial point of the basic layout object in which the content of the basic logical component is laid out (see Figure 10). The distance is specified in SMUs.

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

9.3.2 Orphan size

CATEGORY: Logical

PERMISSIBLE VALUES: any positive integer

DEFAULT VALUE: 1

DEFINITION:

This attribute only applies when the content layout process would result in the basic logical object being laid out in two or more basic layout objects; for example, at a page or frame boundary.

The value of this attribute specifies the minimum number of lines of content that shall be placed in the first basic layout object.

If the number of lines remaining in the first basic layout object is less than the value of this attribute, all the content is laid out in subsequent layout objects.

9.3.3 Proportional line spacing

CATEGORY: Logical

PERMISSIBLE VALUES: ‘yes’
‘no’

DEFAULT VALUE: ‘no’

DEFINITION:

This attribute specifies how the content layout process is to determine the distance between the reference lines of two successive line boxes.

If the value is ‘yes’, the line spacing is variable and the content layout process takes into account the forward extent of the first line box and the backward extent of the second line box.

If the value is ‘no’, the line spacing does not depend on the content of the line boxes but is determined from values set by the presentation attribute “line spacing” or the control functions SLS (Set Line Spacing) and SVS (Select Line Spacing).

9.3.4 Widow size

CATEGORY: Logical

PERMISSIBLE VALUES: any positive integer

DEFAULT VALUE: 1

DEFINITION:

This attribute only applies when the content layout process would result in the basic logical object being laid out in two or more basic layout objects; for example, at a page or frame boundary.

The value of this attribute specifies the minimum number of lines of content that shall be placed in the last basic layout object.

If the number of lines in the last basic layout object is less than the value of this attribute, sufficient lines shall be moved from the end of the previous basic layout object to meet this requirement.

The widow size cannot be altered within the content of a basic component.

It is possible, e.g. for short basic layout objects, that, in satisfying this requirement and the requirements of the presentation attribute “orphan size”, the entire content may be removed from the first basic layout object.

9.4 Content architecture class attributes

9.4.1 Content architecture class

The value of the attribute “content architecture class” of a basic component description that conforms to this part of ITU-T Rec. T.410-Series | ISO/IEC 8613 is an ASN.1 object identifier with one of the following values:

{ 2 8 2 6 0 } for the formatted character content architecture class;

{ 2 8 2 6 1 } for the processable character content architecture class;

{ 2 8 2 6 2 } for the formatted processable character content architecture class;

9.5 Interactions between presentation attributes and layout directives

The attribute “concatenation” of the document architecture (see ITU-T Rec. T.412 | ISO/IEC 8613-2) is applicable to content portions belonging to the same character content architecture class. When concatenation is in effect, as a result of the layout directive “concatenation”, then, for the following presentation attributes

“alignment”

“character fonts”

“character orientation”

“character path”

“first line offset”

“indentation”

“itemization”

“kerning offset”

“line layout table”

“line progression”

“orphan size”

“pairwise kerning”

“widow size”

any value that is specified for the first component of the concatenated sequence applies to all components in the sequence. Thus, in the case of presentation attributes, the values specified for the first component override the values for the other components in the sequence.

10 Character content portion attributes

10.1 Common coding attributes

The value of the attribute “type of coding” of a content portion description that conforms to this part of ITU-T Rec. T.410-Series | ISO/IEC 8613 is an ASN.1 object identifier with the value { 2 8 3 6 0 }.

10.2 Other coding attributes

No other coding attributes are defined for content portions conforming to this part of ITU-T Rec. T.410-Series | ISO/IEC 8613.

10.3 Content information

For character content architectures, the value of this attribute is an octet string representing character coded information in accordance with this part of ITU-T Rec. T.410-Series | ISO/IEC 8613.

11 Formal definitions of character content architecture dependent data types

11.1 Introduction

This clause contains the formal definitions, in ASN.1 notation (defined in CCITT Rec. X.208 | ISO/IEC 8824), of the data types that are applicable to the character content architecture.

These data types are:

- the data type to represent character content architecture specific presentation attributes in basic components, presentation styles and default value lists;
- the data type to represent character content architecture coding attributes in content portions;
- the data type to represent non-basic values of the character content architecture presentation attributes in the document profile;
- the data type to represent non-basic values of the character content architecture coding attributes in the document profile;
- the data type to represent non-standard default values of the character content architecture presentation and coding attributes in the document profile.

11.2 Representation of presentation attributes

The data type “Character-Attributes” contains a set of subordinate data types that specify the character presentation attributes. Some of these subordinate data types are elementary but others are structured and are themselves made up of subordinate data types. The format of these data types is given below.

The subset of subordinate data types that may occur within a particular instance of the data type “Character-Attributes” depends upon a document application profile that is specified, using the rules specified in ITU-T Rec. T.411 | ISO/IEC 8613-1.Character-Presentation-Attributes { 2 8 1 6 2 }

```
DEFINITIONS ::= BEGIN

EXPORTS         Character-Attributes,
        One-Of-Four-Angles,
        One-Of-Two-Angles,
        Measure-Pair,
        Alignment,
        Layout-Table,Error! Bookmark not defined.
        Graphic-Rendition,
        formatting-Indicator,
        character-Fonts,
        itemization,
        kerning-Offset,
        proportional-Line-Spacing,
        pairwise-Kerning;

Character-Attributes                ::=     SET {
        character-path                  [0]     IMPLICIT One-Of-Four-Angles OPTIONAL,
        line-progression                [1]     IMPLICIT One-Of-Two-Angles OPTIONAL,
        character-orientation           [2]     IMPLICIT One-Of-Four-Angles OPTIONAL,
        initial-offset                  [3]     IMPLICIT Measure-Pair OPTIONAL,
        character-spacing               [6]     IMPLICIT INTEGER OPTIONAL,
        line-spacing                    [7]     IMPLICIT INTEGER OPTIONAL,
        alignment                               [8]     IMPLICIT Alignment OPTIONAL,
        line-layout-table               [9]     IMPLICIT Layout-Table OPTIONAL,
        graphic-rendition               [10]    IMPLICIT Graphic-Rendition OPTIONAL,
        formatting-indicator            [11]    IMPLICIT Formatting-Indicator OPTIONAL,
        character-fonts                 [12]    IMPLICIT Character-Fonts OPTIONAL,
        graphic-char-subrepertoire      [13]    IMPLICIT INTEGER OPTIONAL,
        itemization                     [14]    IMPLICIT Itemization OPTIONAL,
```

```
        widow-size                  [15]    IMPLICIT INTEGER OPTIONAL,
        orphan-size                 [16]    IMPLICIT INTEGER OPTIONAL,
        graphic-character-sets      [17]    IMPLICIT OCTET STRING OPTIONAL,
        indentation                 [19]    IMPLICIT INTEGER OPTIONAL,
        kerning-offset              [20]    IMPLICIT Kerning-Offset OPTIONAL,
        proportional-line-spacing   [21]    IMPLICIT Proportional-Line-Spacing OPTIONAL,
        pairwise-kerning            [22]    IMPLICIT Pairwise-Kerning OPTIONAL,
        first-line-offset           [23]    IMPLICIT INTEGER OPTIONAL,
        code-extension-announcers   [24]    IMPLICIT OCTET STRING OPTIONAL }

One-of-Four-Angles          ::=     INTEGER {   d0      (0),
                                                d90     (1),
                                                d180    (2),
                                                d270    (3) }

One-of-Two-Angles           ::=     INTEGER {   d90     (1),
                                                d270    (3) }

Measure-Pair                ::=     SEQUENCE {
        horizontal                          [0]     IMPLICIT INTEGER,
        vertical                        [1]     IMPLICIT INTEGER }

Alignment                   ::=     INTEGER {  start-aligned    (0),
                                               end-aligned      (1),
                                               centred          (2),
                                               justified        (3) }

Layout-Table                ::=     SET OF Tabulation-Stop

Tabulation-Stop             ::=     SET {
        tabulation-reference            [0]     IMPLICIT NumericString,
        tabulation-position             [1]     IMPLICIT INTEGER,
        alignment                               [2]     IMPLICIT INTEGER {
                                                            start-aligned   (0),
                                                            end-aligned     (1),
                                                            centred         (2),
                                                aligned-around (3) },
        alignment-character-string  [3] IMPLICIT OCTET STRING OPTIONAL }
                                        -- string of graphic characters
                                        -- from the set of graphic elements
                                        -- specified by the presentation
                                        -- attributes "graphic character
                                        -- sets" and "graphic character
                                        -- subrepertoire"

Graphic-Rendition           ::=     SET OF Graphic-Rendition-Aspect

Character-Fonts             ::=     SET {
        primary-font                    [0]     IMPLICIT Font-Type OPTIONAL,
        first-alternative-font          [1]     IMPLICIT Font-Type OPTIONAL,
        second-alternative-font         [2]     IMPLICIT Font-Type OPTIONAL,
        third-alternative-font          [3]     IMPLICIT Font-Type OPTIONAL,
        fourth-alternative-font         [4]     IMPLICIT Font-Type OPTIONAL,
        fifth-alternative-font          [5]     IMPLICIT Font-Type OPTIONAL,
        sixth-alternative-font          [6]     IMPLICIT Font-Type OPTIONAL,
        seventh-alternative-font        [7]     IMPLICIT Font-Type OPTIONAL,
        eighth-alternative-font         [8]     IMPLICIT Font-Type OPTIONAL,
        ninth-alternative-font          [9]     IMPLICIT Font-Type OPTIONAL }

Font-Type                   ::=     SET {
        font-size                       [0]     IMPLICIT INTEGER,
        font-identifier                 [1]     IMPLICIT INTEGER }

Graphic-Rendition-Aspect    ::=     INTEGER {
        cancel                          (0),
```

```
        increased-intensity                 (1),
        decreased-intensity                 (2),
        italicized                          (3),
        underlined                          (4),
        slowly-blinking                     (5),
        rapidly-blinking                    (6),
        negative-image                      (7),
        crossed-out                         (9),
        primary-font                        (10),
        first-alternative-font              (11),
        second-alternative-font             (12),
        third-alternative-font              (13),
        fourth-alternative-font             (14),
        fifth-alternative-font              (15),
        sixth-alternative-font              (16),
        seventh-alternative-font            (17),
        eighth-alternative-font             (18),
        ninth-alternative-font              (19),
        doubly-underlined                   (21),
        normal-intensity                    (22),
        not-italicized                      (23),
        not-underlined                      (24),
        steady                              (25),
        variable-spacing                    (26),
        positive-image                      (27),
        not-crossed-out                     (29),
        black-foreground                    (30),
        red-foreground                      (31),
        green-foreground                    (32),
        yellow-foreground                   (33),
        blue-foreground                     (34),
        magenta-foreground                  (35),
        cyan-foreground                     (36),
        white-foreground                    (37),
        select-char-foreground-colour       (38),
        black-background                    (40),
        red-background                      (41),
        green-background                    (42),
        yellow-background                   (43),
        blue-background                     (44),
        magenta-background                  (45),
        cyan-background                     (46),
        white-background                    (47),
        select-char-background-colour       (48),
        not-variable-spacing                (50) }

Formatting-Indicator                ::=     INTEGER { no (0), yes (1) }

Itemization                         ::=     SET {
        identifier-alignment                [0] IMPLICIT INTEGER {
                                            no-itemization (0),
                                            start-aligned (1),
                                            end-aligned (2) }
        identifier-start-offset             [1]     IMPLICIT INTEGER OPTIONAL,
        identifier-end-offset               [2]     IMPLICIT INTEGER OPTIONAL }

Kerning-Offset          ::=     SET {
        start-offset                        [0]     IMPLICIT INTEGER,
        end-offset                          [1]     IMPLICIT INTEGER }

Proportional-Line-Spacing           ::=     INTEGER { no (0), yes (1) }

Pairwise-Kerning                    ::=     INTEGER { no (0), yes (1) }

END
```

11.3 Representation of coding attributes

```
Character-Coding-Attributes { 2 8 1 6 3 }

DEFINITIONS ::= BEGIN

EXPORTS     Character-Coding-Attributes;
            Character-Coding-Attributes          ::=SET { }
                                                      -- no character coding attributes
                                                      -- are defined in this part of
                                                      -- ITU-T Rec. T.410-Series | ISO/IEC 8613

END
```

11.4 Representation of non-basic features and non-standard defaults

```
Character-Profile-Attributes { 2 8 1 6 4 }

DEFINITIONS      ::= BEGIN

EXPORTS          Character-Presentation-Feature,
                 Character-Coding-Attribute,
                 Character-Content-Defaults;

IMPORTS          Character-Attributes,
                 One-Of-Four-Angles,
                 One-Of-Two-Angles,
                 Measure-Pair,
                 Alignment,
                 Layout-Table,
                 Graphic-Rendition,
                 Formatting-Indicator,
                 Character-Fonts
                 Itemization,
                 Kerning-Offset,
                 Proportional-Line-Spacing,
                 Pairwise-Kerning
                 FROM Character-Presentation-Attributes;

                                                      -- see 11.2

Character-Presentation-Feature      ::=CHOICE {
        character-path                          [0]     IMPLICIT One-Of-Four-Angles,
        line-progression                [1]     IMPLICIT One-Of-Two-Angles,
        character-orientation           [2]     IMPLICIT One-Of-Four-Angles,
        initial-offset                  [3]     IMPLICIT Measure-Pair,
        character-spacing               [6]     IMPLICIT INTEGER,
        line-spacing                    [7]     IMPLICIT INTEGER,
        alignment                               [8]     IMPLICIT Alignment,
        line-layout-table               [9]     IMPLICIT Layout-Table,
        graphic-rendition               [10]    IMPLICIT Graphic-Rendition,
        formatting-indicator            [11]    IMPLICIT Formatting-Indicator,
        character-fonts                 [12]    IMPLICIT Character-Fonts,
        graphic-char-subrepertoire      [13]    IMPLICIT INTEGER,
        itemization                     [14]    IMPLICIT Itemization,
        widow-size                      [15]    IMPLICIT INTEGER,
        orphan-size                     [16]    IMPLICIT INTEGER,
        graphic-character-sets          [17]    IMPLICIT OCTET STRING,
        indentation                     [19]    IMPLICIT INTEGER,
        kerning-offset                          [20]    IMPLICIT Kerning-Offset,
        proportional-line-spacing       [21]    IMPLICIT Proportional-Line-Spacing,
        pairwise-kerning                [22]    IMPLICIT Pairwise-Kerning,
```

```
        first-line-offset                   [23]    IMPLICIT INTEGER,
        code-extension-announcers   [24]    IMPLICIT OCTET STRING }

Character-Coding-Attribute      ::=NULL
                                    --  no character coding attributes
                                    --  are defined in this part of
                                    --  ITU-T Rec. T.410-Series | ISO/IEC 8613

Character-Content-Defaults      ::=Character-Attributes

END
```

12 Graphic Characters

The set of graphic characters used in the content of a basic component, and their coded representations, are specified by the presentation attributes “Graphic character sets” and “Graphic character subrepertoire” and the associated control functions, that is, the code extension control functions and the control function Identify Graphic Subrepertoire (IGS) as follows:

- “graphic character sets” – This presentation attribute specifies the graphic character sets designated and/or invoked at the beginning of the basic component. The specification of the graphic character sets by this attribute implies the definition of a repertoire of graphic characters and of a unique coded representation for each character of that repertoire;
- “graphic character subrepertoire” – This presentation attribute can be used, when the graphic character sets designated are those of ISO 6937, to restrict the repertoire of graphic characters to a subset of the repertoire implied by the specification of the graphic character sets. Use of this attribute does not affect the coded representations of the graphic characters;
- code extension control functions – These control functions can be used within the content of a basic component to alter the designations and/or invocations of graphic character sets from those specified by the presentation attribute “graphic character sets”;
- Identify Graphic Subrepertoire (IGS) – This control function can be used within the content of a basic component, when the graphic character sets designated are those of ISO 6937, to alter the subrepertoire from that specified by the presentation attribute “graphic character subrepertoire”.

When the presentation attributes “graphic character sets” and “graphic character subrepertoire” are not specified for a basic component, their values are determined using the defaulting mechanism defined in ITU-T Rec. T.412 | ISO/IEC 8613-2, as for any other presentation attribute.

13 Definition of control functions and the character SPACE

Control functions are classified in the following four categories:

- Shared control functions, including code extension control functions, which are available to all classes of character content architecture;
- Layout control functions which are available only to formatted form and formatted processable form character content architectures;
- Logical control functions which are available only to processable form and formatted processable form character content architectures;
- Delimiters which are used to delimit graphic characters and/or shared control functions introduced as a result of a formatting process. The delimiters are available only to formatted processable form character content architectures.

The control functions in these four categories are defined in 13.1 to 13.4 respectively; subclause 13.5 defines the character SPACE. The allocation of control functions to categories is summarized in Table 3.

Control functions that have not been included explicitly in this clause are:

- the sequence introducers Escape (ESC) and Control Sequence Introducer (CSI) that are permitted to be used in the character content architecture when required to represent a graphic character or a control function;
- line terminators:
  1) A soft line terminator is represented by the control function Carriage Return (CR) then Line Feed (LF) in that order, where the CR and LF are enclosed between a pair of delimiters Start of String (SOS) and String Terminator (ST);
  2) A hard line terminator is represented by a control function Line Feed (LF) that immediately follows either a control function Carriage Return (CR) or another LF where the entire sequence of CR and LF(s) is NOT enclosed between a pair of delimiters Start of String (SOS) and String Terminator (ST).

Table 3 – Control functions

Shared control functions	CR Carriage Return GCC Graphic Character Composition IGS Identify Graphic Subrepertoire LF Line Feed PLD Partial Line Up PLU Partial Line Down SCS Set Character Spacing SGR Select Graphic Rendition SHS Select Character Spacing SLS Set Line Spacing SRS Start Reverse String STAB Selective Tabulation SUB Substitute Character SVS Select Line Spacing VPB Line Position Backward VPR Line Position Relative Code extension control functions
Layout control functions	BS Backspace HPB Character Position Backward HPR Character Position Relative JFY No Justify SACS Set Additional Character Separation SRCS Set Reduced Character Separation SSW Set SPACE Width
Logical control functions	BPH Break Permitted Here NBH No Break Here PTX Parallel Texts
Delimiters	SOOS (Start of Original String) SOS (Start of String) ST String Terminator

13.1 Shared control functions

13.1.1 Carriage Return (CR)

A control function which causes the active position to be moved to the line home position but not to be moved in the direction of line progression.

NOTE – CR is used in conjunction with the control function LF (Line Feed) to move the active position to the line home position at the beginning of a new line of text. It is also used to move the active position to the line home position. For example, after an item identifier (see 9.1.11). CR should not be used to cause character images to be superimposed.

13.1.2 Graphic Character Composition (GCC)

A control function with one optional parameter which specifies that two or more graphic characters are to be combined into one graphic symbol.

The value of the parameter is 0, 1 or 2.

GCC with the parameter value 0 indicates that the two following graphic characters are to be presented as a single symbol.

GCC with the parameter value 1 indicates the start and GCC with the parameter value 2 indicates the end of a string of graphic characters that are to be presented as a single symbol.

If no parameter value is specified the parameter value is assumed to be 0.

13.1.3 Identify Graphic Subrepertoire (IGS)

A control function with one optional selective parameter which is used to indicate that a subrepertoire of the graphic character repertoire of ISO 6937 is used in the subsequent text. All graphic character sets that are used to represent the indicated graphic character subrepertoire shall be explicitly designated, but need not be invoked, prior to the occurrence of IGS.

The identification of the graphic character subrepertoire may be changed at any point within a document and becomes effective immediately. No graphic characters other than those of the specified subrepertoire shall be used in the text following the occurrence of IGS. The effect of a graphic character subrepertoire identification ceases upon the next occurrence of

- another IGS;
- the end of the current basic object;
- the designation of any graphic character set.

If no parameter value is specified the parameter value is assumed to be 0.

A non-zero parameter value is the identifier assigned to a subrepertoire of the repertoire of ISO 6937 in accordance with the registration procedure specified in ISO/IEC 7350.

The parameter value zero identifies the entire repertoire of the currently designated graphic character sets.

In the absence of IGS, the subrepertoire identified by the presentation attribute “graphic character subrepertoire” applies or the entire repertoire of the currently designated graphic character sets otherwise.

NOTE – The use of IGS in document application profiles based on ISO/IEC 8613 is deprecated. IGS is included in this part of ITU-T Rec. T.410-Series | ISO/IEC 8613 only for compatibility with some existing applications such as those based upon Recommendation T.61 (1984).

13.1.4 Line Feed (LF)

A control function which causes the active position to be advanced in the direction of line progression but not to be moved in the direction of the character path. As a result, in the content layout process, the current line box is terminated and a new line box is started. The amount of movement is that specified by the most recent occurrence either of the control functions Set Line Spacing (SLS) or Select Line Spacing (SVS) if any, or otherwise by the presentation attribute “line spacing”.

LF is restricted to be used in the following cases:

- at the beginning of the content of a basic layout component;
- immediately following a control function Carriage Return (CR);
- immediately following another LF.

13.1.5 Partial Line Down (PLD)

A control function which causes either the start of subscript rendition or the end of superscript rendition of graphic characters.

When superscript rendition is in effect, it is terminated by PLD; otherwise, subscript rendition is initiated by PLD.

Any occurrence of PLD to start subscript rendition shall be followed by a control function PLU (Partial Line Up) in the same line before another PLD or control function LF (Line Feed) is used.

PLD does not affect the position of any (graphic) lines used to implement the graphic rendition “underlined”, “crossed-out”, or “doubly underlined” when such a graphic rendition is in effect prior to the occurrence of PLD.

NOTES

1 The implementation of the subscript rendition initiated by PLD may be accomplished with special character fonts and/or movement of the active position not exceeding a half line space.

2 The graphic rendition “underlined” may have been effected, prior to the occurrence of PLD, either by the control function Select Graphic Rendition (SGR) or by the non-spacing underline character (see ISO 6937).

13.1.6 Partial Line Up (PLU)

A control function which causes either the start of superscript rendition or the end of subscript rendition of graphic characters.

When subscript rendition is in effect, it is terminated by PLU; otherwise, superscript rendition is initiated by PLU.

Any occurrence of PLU to start superscript rendition shall be followed by a control function PLD (Partial Line Down) in the same line before another PLU or control function LF (Line Feed) is used.

PLU does not affect the position of any (graphic) lines used to implement the graphic rendition “underlined”, “crossed-out” or “doubly underlined” when such a graphic rendition is in effect prior to the occurrence of PLU.

NOTES

1 The implementation of the superscript rendition initiated by PLU may be accomplished with special character fonts and/or movement of the active position not exceeding a half line space.

2 The graphic rendition “underlined” may have been affected, prior to the occurrence of PLU, either by the control function SGR (Select Graphic Rendition) or by the non-spacing underline character (see ISO 6937).

13.1.7 Set Character Spacing (SCS)

A control function with one optional numeric parameter which specifies the character spacing to be applied to constant spacing fonts in subsequent text.

The specified character spacing takes effect immediately and remains in effect until it is changed by a subsequent occurrence of either SCS or the control function Select Character Spacing (SHS) in the current basic component.

The character spacing is expressed as an integral multiple of SMUs.

If no parameter value is specified the parameter value is assumed to be the equivalent of 120 BMUs.

13.1.8 Select Graphic Rendition (SGR)

A control function with zero, one or more selective parameters which specify one or more graphic rendition aspects for graphic characters and space characters in the subsequent text. The specified graphic rendition(s) takes effect immediately and remains in effect until a subsequent occurrence of SGR in the basic object.

The meaning of the parameter values is shown in Table 4.

If no parameter value is specified the parameter value is assumed to be 0.

Any graphic rendition aspect specified by an occurrence of SGR, apart from the exceptions specified below, is combined with the graphic rendition aspects that are in effect prior to that occurrence of SGR as a result of either an earlier occurrence of SGR or the presentation attribute “graphic rendition” of the current basic object.

Table 4 – Parameter values for SGR

Value	Meaning
0	Default rendition (implementation-defined); cancels the effect of any preceding occurrence of SGR and cancels the effect of the presentation attribute “graphic rendition”; invokes the primary font, character foreground and character background colour
1	Bold or increased intensity
2	Faint or decreased intensity
3	Italicized
4	Underlined
5	Slowly blinking
6	Rapidly blinking
7	Negative image
9	Crossed-out (characters still legible but marked as to be deleted)
10	Primary (default) font
11	First alternative font
12	Second alternative font
13	Third alternative font
14	Fourth alternative font
15	Fifth alternative font
16	Sixth alternative font
17	Seventh alternative font
18	Eighth alternative font
19	Ninth alternative font
21	Doubly underlined
22	Normal intensity (neither bold nor faint)
23	Not italicized
24	Not underlined (neither singly nor doubly)
25	Steady (not blinking)
26	Variable spacing
27	Positive image
29	Not crossed-out
30	Black foreground
31	Red foreground
32	Green foreground
33	Yellow foreground
34	Blue foreground
35	Magenta foreground
36	Cyan foreground
37	White foreground
38	Select character foreground colour
40	Black background
41	Red background
42	Green background
43	Yellow background
44	Blue background
45	Magenta background
46	Cyan background
47	White background
48	Select character background colour
50	Not variable spacing

When SGR is used to start underlining (singly or doubly), or crossing out, within the scope of subscript or superscript rendition (see 13.1.5 and 13.1.6), any lines used to implement such a graphic rendition are lowered or raised in order that the graphic rendition concerned applies to the subscript or superscript characters.

The default parameter value cannot be used in combination with any other parameter value.

When a designated font is invoked, any specification of weight or posture by an SGR parameter value is ignored. This applies to the parameter values 1, 2, 3, 22 and 23. These parameter values are enabled when a primary or alternative font is invoked for which no designation has been made.

NOTES

1 Several parameter values can be used in combination, in order to obtain, for example, underlined italics.

2 The use of parameter values 26 and 50 in document application profiles based on ISO/IEC 8613 is deprecated.

These values are included in this Specification only for compatibility with some existing applications such as those based upon Recommendation T.61 (1984).

The values 30 to 37 and 40 to 47 are interpreted as indexes into the colour table specified in the attribute "content colour table" applying to the object to which the content is associated. This attribute is defined in ITU-T Rec. T.412 | ISO/IEC 8613-2. The values are associated with indexes as follows:

SGR values	Entry in default colour table
30 and 40	index 1
31 and 41	index 2
32 and 42	index 3
33 and 43	index 4
34 and 44	index 5
35 and 45	index 6
36 and 46	index 7
37 and 47	index 0

The parameter values 38 and 48 are followed by a parameter substring used to select either the character foreground "colour value" or the character background "colour value".

A parameter substring for values 38 or 48 may be divided by one or more separators (03/10) into parameter elements, denoted as Pe. The format of such a parameter sub-string is indicated as:

Pe : P ...

Each parameter element consists of zero, one or more bit combinations from 03/00 to 03/09, representing the digits 0 to 9. An empty parameter element represents a default value for this parameter element. Empty parameter elements at the end of the parameter substring need not be included.

The first parameter element indicates a choice between:

0 implementation defined (only applicable for the character foreground colour)

1 transparent;

2 direct colour in RGB space;

3 direct colour in CMY space;

4 direct colour in CMYK space;

5 indexed colour.

If the first parameter has the value 0 or 1, there are no additional parameter elements.

If the first parameter element has the value 5, then there is a second parameter element specifying the index into the colour table given by the attribute "content colour table" applying to the object with which the content is associated.

If the first parameter element has the value 2, 3, or 4, the second parameter element specifies a colour space identifier referring to a colour space definition in the document profile.

If the first parameter element has the value 2, the parameter elements 3, 4, and 5, are three integers for red, green, and blue colour components. Parameter 6 has no meaning.

If the first parameter has the value 3, the parameter elements 3, 4, and 5 and three integers for cyan, magenta, and yellow colour components. Parameter 6 has no meaning.

If the first parameter has the value 4, the parameter elements 3, 4, 5, and 6, are four integers for cyan, magenta, yellow, and black colour components.

If the first parameter element has the value 2, 3, or 4, the parameter element 7 may be used to specify a tolerance value (an integer) and the parameter element 8 may be used to specify a colour space associated with the tolerance (0 for CIELUV, 1 for CIELAB).

NOTE 3 – The "colour space id" component will refer to the applicable colour space description in the document profile which may contain colour scaling data that describe the scale and offset to be applied to the specified colour components in the character content. Appropriate use of scaling and offsets may be required to map all colour values required into the integer encoding space provided. This may be particularly important if concatenated content requires the insertion of such SGR sequences by the content layout process.

13.1.9 Select Character Spacing (SHS)

A control function with one optional selective parameter which specifies the character spacing to be applied to constant spacing fonts in subsequent text.

The specified character spacing remains in effect until it is changed by a subsequent occurrence of either SHS or the control function Set Character Spacing (SCS) in the current basic object.

The meaning of the parameter value is:

0 120 BMUs;

1 100 BMUs;

2 80 BMUs;

3 200 BMUs;

4 400 BMUs.

If no parameter value is specified the parameter value is assumed to be 0.

13.1.10 Set Line Spacing (SLS)

A control function with one optional numeric parameter which specifies the line spacing for subsequent text.

The specified line spacing takes effect immediately and remains in effect until it is changed by a subsequent occurrence of either SLS or the control function Select Line Spacing (SVS) in the current basic component.

The line spacing is expressed as an integral multiple of SMUs.

If no parameter value is specified the parameter value is assumed to be the equivalent of 200 BMUs.

13.1.11 Start Reverse String (SRS)

A control function with one optional selective parameter which is used to indicate either the start or end of a string of graphic characters that are to be imaged in the direction opposite to that of the immediately preceding text (see 7.2.4).

SRS with parameter value 1 indicates the start of the string.

SRS with parameter value 0 indicates the end of the string.

If no parameter value is specified the parameter value is assumed to be 0.

Hard and soft line terminators shall not be used between SRS 1 and SRS 0.

Any occurrence of the control functions Partial Line Down (PLD), Partial Line Up (PLU), Line Position Backward (VPB) or Line Position Relative (VPR) within the string of characters delimited by SRS 1 and SRS 0 shall be matched by an occurrence of the opposite control function within the string.

Strings delimited by SRS may be nested.

13.1.12 Selective Tabulation (STAB)

A control function with one optional selective parameter which references a tabulation stop position in an associated “line layout table” (see 9.1.13).

This control function specifies the positioning of the subsequent text, until either the occurrence of another STAB or the end of the current line.

This text is to be positioned at the referenced tabulation stop and aligned in accordance with the properties specified for that tabulation stop.

If no parameter value is specified the parameter value is assumed to be a reference to the next tabulation stop position.

13.1.13 Substitute Character (SUB)

A control function which is used in the place of a character that has been found invalid or in error.

13.1.14 Select Line Spacing (SVS)

A control function with one optional selective parameter which specifies the line spacing for subsequent text. The specified line spacing takes effect immediately and remains in effect until it is changed by a subsequent occurrence of either SVS or the control function Set Line Spacing (SLS) in the current basic object.

The meaning of the parameter value is:

0 200 BMUs;

1 300 BMUs;

2 400 BMUs;

3 100 BMUs;

4 150 BMUs;

9 600 BMUs.

If no parameter value is specified the parameter value is assumed to be 0.

13.1.15 Line Position Backward (VPB)

A control function with one optional numeric parameter which causes the active position to be moved in the opposite direction to the line progression the number of SMUs specified by the parameter.

The combined effect of all occurrences of the control functions VPB and Line Position Relative (VPR) within a given line shall be such that the active position is returned to the reference line before the occurrence of a hard or soft line terminator.

If no parameter value is specified the parameter value is assumed to be the equivalent of 100 BMUs.

NOTE – The main purposes of VPB are to provide for positioning of parallel annotation and for explicit control for the positioning of superscripts.

13.1.16 Line Position Relative (VPR)

A control function with one optional numeric parameter which causes the active position to be moved in the direction of line progression the number of SMUs specified by the parameter.

The combined effect of all occurrences of the control functions Line Position Backward (VPB) and VPR within a given line shall be such that the active position is returned to the reference line before the occurrence of a hard or soft line terminator.

If no parameter value is specified the parameter value is assumed to be the equivalent of 100 BMUs.

NOTE - The main purposes of VPR are to provide for positioning of parallel annotation and for explicit control for the positioning of subscripts.

13.1.17 Code extension control functions

This is a category of control functions used for the designation and invocation of graphic character sets. They are defined in ISO 6429 and their use is defined in ISO 2022.

13.2 Layout control functions

13.2.1 Backspace (BS)

A control function that causes the active position to be moved, in the direction opposite to the character path, a distance specified by the most recent occurrence of either of the control functions Select Character Spacing (SHS) or Set Character Spacing (SCS), if any, or otherwise by the presentation attribute “character spacing”. The control function Set SPACE Width (SSW) has no effect on BS.

NOTE – BS allows for the positioning of item identifiers (see 7.2.9) on systems which do not implement the control function Character Position Backward (HPB). It is included in this Specification only for compatibility with Recommendation T.61 (1984). Its use in other document application profiles based upon ITU-T Rec. T.410-Series | ISO/IEC 8613 is deprecated. It should not be used to cause character images to be superimposed.

13.2.2 Character Position Backward (HPB)

A control function with one optional numeric parameter which causes the active position to be moved in the opposite direction to the character path the number of SMUs specified by the parameter.

If no parameter value is specified the parameter value is assumed to be the equivalent of 120 BMUs.

NOTE - The main purposes of HPB are to move the active position backwards from the line home position, and to provide for the positioning of parallel annotation.

13.2.3 Character Position Relative (HPR)

A control function with one optional numeric parameter which causes the active position to be moved in the direction of the character path the number of SMUs specified by the parameter.

If no parameter value is specified the parameter value is assumed to be the equivalent of 120 BMUs.

NOTES

1 Although HPR has a control effect similar to that of one or more space characters, it does not have the graphic equivalence of space characters. Therefore, HPR does not cause spaces to be imaged in accordance with the current graphic rendition, such as underlined, possibly specified by a preceding occurrence of the control function Select Graphic Rendition (SGR).

2 HPR also provides for the positioning of parallel annotation.

13.2.4 No Justify (JFY)

A control function with one optional selective parameter which is used at the beginning of a line to indicate that the line must not be justified. It has no effect on subsequent lines.

The meaning of the parameter value is:

0 no justification

If no parameter value is specified the parameter value is assumed to be 0.

13.2.5 Set Additional Character Separation (SACS)

A control function with one optional numeric parameter which specifies increased escapement between graphic characters in subsequent text i.e. a positive value for the inter-character space (see 7.2.1).

The specified value applies after the first subsequent graphic character and remains in effect until the next occurrence of SACS or a control function Set Reduced Character Separation (SRCS) or until it is reset to 0 by a subsequent occurrence of a hard or soft line terminator.

The parameter value is expressed in SMUs.

If no parameter value is specified the parameter value is assumed to be 0.

13.2.6 Set Reduced Character Separation (SRCS)

A control function with one optional numeric parameter which specifies reduced escapement between graphic characters in subsequent text i.e. a negative value for the inter-character space (see 7.2.1).

The specified value applies after the first subsequent graphic character and remains in effect until the next occurrence of SRCS or a control function Set Additional Character Separation (SACS) or until it is reset to 0 by a subsequent occurrence of a hard or soft line terminator.

The parameter value is expressed in SMUs.

If no parameter value is specified the parameter value is assumed to be 0.

13.2.7 Set SPACE Width (SSW)

A control function with one optional numeric parameter which specifies the character escapement associated with the character SPACE for subsequent text. The specified value takes effect immediately and remains in effect until it is changed by a subsequent occurrence of SSW or reset to the default value by a subsequent occurrence of a hard or soft line terminator.

The parameter value is expressed in SMUs.

If no parameter value is specified the parameter value is assumed to be equal to the character spacing if the current font has constant spacing. Otherwise it is determined by the font in use.

13.3 Logical control functions

13.3.1 Break Permitted Here (BPH)

A control function which indicates a point where a line break may occur when text is formatted (see 14.2.1.3.2).

13.3.2 No Break Here (NBH)

A control function which indicates a point where no line break may occur when text is formatted (see 14.2.1.3.2).

NOTE - The graphic character No Break SPace (NBSP) should be taken to be equivalent to SPACE followed by NBH.

13.3.3 Parallel Texts (PTX)

A control function with one optional selective parameter which delimits passages of text which are interchanged one after the other, but which are intended to be presented in parallel with one another (see 7.2.5).

The meaning of the parameter values are:

0 end of parallel texts

1 start of principal text

3 start of supplementary text (e.g. Japanese Ruby annotation).

If no parameter value is specified the parameter value is assumed to be 0.

PTX with parameter value 1 is the opening delimiter of the first (principal) of two passages of text intended to be presented in parallel with one another.

PTX with parameter value 3 is the closing delimiter of the first passage of text and the opening delimiter of the second (supplementary) passage intended to be presented in parallel with the first.

PTX with parameter value 0 indicates the end of the supplementary passage of text.

NOTE – Japanese Ruby permits the specification of exactly one supplementary passage of text.

13.4 Delimiters

13.4.1 Start Of Original String (SOOS)

A control function that acts as the opening delimiter of a string of graphic characters and/or control functions that is marked to facilitate its reintroduction by a subsequent editing or content layout process. The string is closed by either an intermediate control function Start of String (SOS) or the terminating delimiter control function String Terminator (ST), whichever occurs earlier in the content. If the string is closed by an SOS control function, the latter is interpreted as the opening delimiter of a string as defined in 13.4.2.

The string thus delimited may contain occurrences of graphic characters and control functions, in particular parts of a word subject to hyphenation, that have been removed from the content as a result of a formatting process (see 14.2.1.3).

13.4.2 Start Of String (SOS)

A control function that acts as the opening delimiter of a string of graphic characters and/or control functions that is marked to facilitate its removal by a subsequent content layout process. The string is closed by the terminating delimiter control function String Terminator (ST).

A string thus delimited may contain occurrences of graphic characters and control functions, in particular Carriage Return (CR), Line Feed (LF) and HYPHEN, introduced as a result of a formatting process (see 14.2.1.3).

13.4.3 String Terminator (ST)

A control function that acts as the terminating delimiter of a string opened by the delimiter control function Start of String (SOS) or Start of Original String (SOOS).

13.5 SPACE (SP)

A character with properties of both a graphic character and a logical control function.

As a control function, SP is significant to the content layout process. It acts as a word delimiter and indicates a potential line break point except when it is immediately followed by another SP or by an occurrence of the control function No Break Here (NBH) (see 13.3.2).

As a graphic character, SP causes the active position to be advanced without a graphic symbol to be imaged. However, any graphic renditions that are in effect, e.g. underlining, also apply to SP.

Any SP(s) that precede a line terminator after the last graphic character of a line, are ignored by the imaging process.

14 Content layout process

This clause describes a content layout process for basic logical objects associated with content architecture of type character.

Its purpose is to aid understanding of the semantics of the presentation attributes by describing the required results of such a process. However, it is not intended to specify any process that might be carried out in a particular implementation to achieve these results.

14.1 Introduction

14.1.1 Purpose

The content layout process defines a process of formatting and laying out character content into an allocated area. This area is referred to as the available area and is determined by the document layout process defined in ITU-T Rec. T.412 | ISO/IEC 8613-2.

The purpose of the content layout process is to convert content associated with basic logical components into content associated with basic layout objects. This might imply a transformation of the content from one form to another.

The content layout process results in the creation of basic layout object(s) into which the content is to be positioned.

The dimensions of the basic layout object(s) are returned to the document layout process which determines the precise position of that basic layout object within the available area.

14.1.2 Available area

The content layout process is constrained by the available area. That is, the maximum line length is constrained by the dimension of the available area in the direction of the character path.

During the layout of the content of a basic logical object into a basic layout object, the following cases can occur:

- the formatted or formatted processable content fits into the available area;
- the formatted or formatted processable content does not fit into the dimension of the available area in the direction of line progression. In this case, an additional or a new available area is required depending upon any constraints imposed by the document layout process;
- the formatted content does not fit into the dimension of the available area in the direction of the character path. In this case, a larger available area is required.

  NOTE – This case is most likely to occur when laying out formatted content associated with a basic logical component.

14.1.3 Presentation attributes

The content layout process is carried out taking into account the presentation attributes applying to the basic logical object with which the content is associated. The content layout process must also take into account any control functions that are embedded in the content.

The presentation attributes applying to the content layout process can be specified in the generic layout structure and presentation styles. The values of these presentation attributes are determined according to the defaulting rules specified in ITU-T Rec. T.412 | ISO/IEC 8613-2.

14.1.4 Character content architecture classes

The content layout process is described for basic logical objects associated with content that conforms to any of the three character content architecture classes (see 6.1) as follows:

- Processable form character content in which the content layout process provides for formatting of the content. The content layout process results in the output of content in formatted or formatted processable form depending upon the desired form of document.

- Formatted processable form character content in which the content layout process provides for reformatting of the content. This involves an initialization process which must be carried out on the content before the content layout process can be applied to that content.The content layout process results in the output of content in formatted or formatted processable form depending upon the desired form of document.
- Formatted form character content in which the content layout process has no effect on the content itself but still determines the dimensions of the basic layout object into which that content is to be positioned.

14.1.5 Use of delimiters

When formatted processable form content is created as a result of the content layout process, all shared control functions and graphic characters inserted into the content as a result of the content layout process are enclosed between the delimiter control functions Start of String (SOS) and String Terminator (ST), and control functions and graphic characters deleted from the content as a result of the layout process are either enclosed between the delimiter control functions Start of Original String (SOOS) and ST or between SOOS and an inserted string delimited by the control functions SOS and ST.

14.1.6 Layout of the content

For each of the three character content architecture classes, three cases of laying out the content of basic logical objects into basic layout objects are possible:

- single basic logical object to single basic layout object: the content of a single basic logical object can be laid out into a single basic layout object and is the only content associated with this basic layout object;
- single basic logical object to multiple basic layout objects: the content of a single basic logical object is split among two or more basic layout objects, i.e. the content portions associated with two or more basic layout objects are derived from a single basic logical object;
- multiple basic logical objects to single basic layout object: the content of two or more basic logical objects is laid out into a single basic layout object, i.e. the content portions of two or more basic logical objects are associated with a single basic layout object.

Multiple basic logical objects to multiple basic layout objects is also possible, but is not described explicitly since this is a combination of the last two cases above.

14.1.7 Layout sequence

In all cases, the same sequence of steps for laying out content associated with a basic logical object is executed as follows:

- initialization;
- determination of initial point;
- formatting of the content;
- identification of content portions;
- determination of basic layout object dimensions;
- determination of the value of the presentation attribute “initial offset”.

This also results in the creation of a basic layout object.

Although the sequence of steps is the same in all cases, the action performed at individual steps may vary.

14.1.8 Character sets

The value of the attribute “content information” is always interpreted as a sequence of characters from the character set(s) designated and invoked by the attribute “graphic character sets” and by corresponding control functions in the content. Character content specified in the value of the attribute “content generator” is interpreted as follows:

- Character content that is a string literal is interpreted in the same way as character content in the attribute “content information”.
- Character content that is generated by the application of string functions is interpreted as a sequence of characters from the character set(s) specified for those string functions. If these character sets are different from those used by any content preceding or following the content generated by application of the string functions, the character content layout process inserts appropriate designation and invocation control functions before and after the content generated by the application of the string functions.

14.2 Content layout process for processable content

14.2.1 Single basic logical object to single basic layout object

14.2.1.1 Initialization

In the case of processable form content, no initialization of the content is necessary.

14.2.1.2 Determination of initial point

The location of the initial point depends upon:

- the presentation attributes “character path” and “line progression” (determining the start edge and top edge of the basic layout object);
- the presentation attribute “kerning offset” (specifying the location of the start edge of the positioning area relative to the start edge of the basic layout object);
- the presentation attribute “indentation” (specifying the distance between the initial point and the start edge of the positioning area);
- the invocation of a font by the presentation attribute “graphic rendition” and the presentation attribute “character orientation” (specifying the minimum backward extent of the first line box);
- the presence of control functions Partial Line Up (PLU), Parallel Texts (PTX), Line Position Backward (VPB) and font invocation by Select Graphic Rendition (SGR) in the first line of characters to be imaged (modifying the backward extent of the first line box).

The position of the initial point relative to the start edge and top edge of the positioning area is determined such that:

- its distance from the top edge of the positioning area is equal to the backward extent of the first line box;
- its distance from the start edge of the positioning area is equal to the value specified by the presentation attribute “indentation”.

The position of the initial point relative to the upper left corner of the basic layout object can be determined only after the dimension of the basic layout object has been determined (see 14.2.1.5). This value is assigned to the presentation attribute “initial offset” and should always be specified explicitly in order to achieve the desired result by the content imaging process.

14.2.1.3 Formatting of the content

Formatting of the content involves:

- the positioning of character images within a line box (see 7.2);
- the determination of line breaks;
- the positioning of line boxes within the basic layout object (see 7.3).

It may involve the insertion of control functions and the assignment of presentation attribute values to the basic layout object.

14.2.1.3.1 Positioning of character images within a line box

There are two sets of operations that have an effect on the positioning of character images within a line box. One of these sets is related to the presentation attribute “formatting indicator”, the other is not.

The operations related to the attribute “formatting indicator” are:

- alignment (presentation attribute “alignment”);
- tabulation (presentation attribute “line layout table” and the control function STAB);
- first line offset (presentation attribute “first line offset”);
- itemization (presentation attribute “itemization”);
- pairwise kerning (presentation attribute “pairwise kerning”).

Alignment and tabulation are mutually exclusive but itemization and first line offset may be applied in combination with either of them.

These operations may be performed by either the content layout process or the content imaging process. However, they can be performed by the content layout process only if the presentation attribute “formatting indicator” and the control functions Character Position Backward (HPB), Character Position Relative (HPR), Set Additional Character Separation (SACS), Set Reduced Character Separation (SRCS) and Set SPACE Width (SSW) are available in the content architecture class concerned.

The result of performing these operations by the content layout process is the insertion of the above control functions in the content.

If the content layout process has performed all of the specified operations for a basic layout object and inserted all necessary control functions, the value of the presentation attribute “formatting indicator” is set to “yes”, otherwise it is set to “no”.

Whether or not the content layout process inserts these control functions, it always determines the allocation of characters to each line box and the dimensions of each line box.

NOTE – If the value of the presentation attribute “formatting indicator” is “no” or if the font substitution has been made, then the content imaging process should perform these operations.

The other set of operations which are not related to the presentation attribute “formatting indicator” are:

– character ordering (control function SRS);

– parallel annotation (control function PTX);

– graphic character composition (control function GCC).

14.2.1.3.1.1 Pairwise kerning

If the presentation attribute “pairwise kerning” specifies “yes” and the content layout process is capable of performing this function and the font used provides the necessary information, then certain combinations of character images are positioned closer to (or further apart from) each other than determined by their position and escapement points.

When pairwise kerning is performed by the content layout process, it will result in the insertion of control functions Character Position Backward (HPB) or Character Position Relative (HPR) between the two characters involved.

14.2.1.3.1.2 First line offset

The characters associated with the first line can be controlled to be laid out differently from the rest of the lines in this object.

The presentation attribute “first line offset” specifies if the first line has an overhang or indentation relative to the line home position.

When first line offset is performed by the content layout process, it will result in the insertion of a control function Character Position Backward (HPB) or Character Position Relative (HPR).

14.2.1.3.1.3 Itemization

The first line of a basic layout object may contain an item identifier. The position of the item identifier is controlled by the presentation attribute “itemization”.

When itemization is performed by the content layout process, it will result in the insertion of a control function Character Position Backward (HPB) or Character Position Relative (HPR) before the text of the item identifier.

14.2.1.3.1.4 Alignment

None of the alignments except “start-aligned“ can be performed until the dimensions of the basic layout object have been determined (see 14.2.1.5).

When alignment is performed by the content layout process, the line length for alignment is determined to be:

– for the first line, the distance between the line home position and the end edge of the positioning area minus the value of the presentation attribute “first line offset”;

– for all other lines, the distance between the line home position and the end edge of the positioning area.

The various values of the presentation attribute “alignment” are treated as follows:

– “start-aligned“ does not affect the output of the content layout process;

- "end-aligned" and "centred" result in the insertion of a control function Character Position Relative (HPR) either before the first graphic character of each line or after the Carriage Return (CR) delimiting the item identifier if the presentation attribute "itemization" specifies a value other than "no itemization";

- "justified" results in the insertion of zero, one or more control functions Set SPACE Width (SSW), Set Additional Character Separation (SACS) and/or Set Reduced Character Separation (SRCS) in each line which ends with a line terminator inserted by the content layout process. The precise usage of SACS, SRCS and SSW is implementation dependent.

  NOTE – The presentation attribute alignment" does not apply to the item identifier.

When the presentation attribute "alignment" has the value "justified", irrespective of whether the alignment is performed by the content layout process or the content imaging process, the control function No Justify (JFY) is inserted at the beginning of the last line of a character sequence in order to avoid justification of this line by the content imaging process.

14.2.1.3.1.5 Tabulation

When tabulation is performed by the content layout process, it results in the insertion of each occurrence of a control function Character Position Relative (HPR) or Character Position Backward (HPB) between each occurrence of the control function Selective Tabulation (STAB) and the first graphic character following it.

14.2.1.3.1.6 Parallel annotation

The occurrence of the control function Parallel Texts (PTX) in the content specifies that a string of characters is to be laid out as a parallel annotation to another specified string of characters. The result of the content layout process is as described in 7.2.5.

If the output of the content layout process is in formatted form, the positioning of the parallel annotation is achieved by the removal of all occurrences of PTX and the insertion of the appropriate control functions Character Position Relative (HPR), Character Position Backward (HPB), Line Position Relative (VPR) and Line Position Backward (VPB).

If the output of the content layout process is in formatted processable form, the positioning of the parallel annotation is achieved by inserting the control functions Character Position Relative (HPR), Character Position Backward (HPB), Line Position Relative (VPR) and Line Position Backward (VPB) and enclosing them by the delimiter control functions Start of String (SOS) and String Terminator (ST).

14.2.1.3.1.7 Character ordering

The occurrence of the control function Start Reverse String (SRS) in the content of a basic logical object controls the direction of imaging of the interchanged characters. The result of the content layout process is as described in 7.2.4.

14.2.1.3.1.8 Graphic character composition

The control function Graphic Character Composition (GCC) is used to combine two or more graphic characters into a single symbol, the width of which may be less than the sum of the widths of the component characters.

14.2.1.3.2 Insertion of line breaks

The formatting process commences by creating a line box. Each hard line terminator in the content causes the creation of a subsequent line box.

Additional line breaks, and hence additional line boxes, may be produced as a result of the formatting process. If the output of the formatting process is formatted processable form, the inserted line breaks are represented by CR/LF combinations delimited by the control function sequence SOS ... ST. This control function sequence may include other characters or control functions created as a result of the inserted line break (hyphens, for example). Optionally, characters deleted by such an algorithm are enclosed either in an SOOS-ST string or between an SOOS control function and the SOS-ST string.

If the output is in formatted form, the inserted line breaks consist of hard line terminators represented by the control function sequence CR LF. All occurrences of the control functions Break Permitted Here (BPH) and No Break Here (NBH) are deleted.

It is the intention of the formatting process that the number of characters between the inserted line breaks is the maximum possible for each line. The exact algorithm for inserting line breaks is implementation dependent and is not defined in UIT-T Rec. T.410-Series | ISO/IEC 8613. However, the line break algorithm shall conform to the following constraints:

a) a soft line break may be inserted

   – after a SP which is not immediately followed by another SP or the control function NBH (No Break Here);

   – after the control function BPH (Break Permitted Here);

   – at a point determined by an implementation or language dependent algorithm.

b) a line break is not permitted

   – when a subscripted rendition is active;

   – within a string with reversed presentation direction;

   – within a string of parallel annotation.

14.2.1.3.3 Positioning of line boxes

The first line box is positioned with its line home position at the initial point as described in 14.2.1.2.

Each line box is positioned with its line home position on the line from the initial point in the direction of line progression.

When proportional line spacing is not to be performed, the distance between the line home positions of two successive line boxes is equal to the current line spacing. The initial value of the current line spacing is the value of the presentation attribute “line spacing”. This value may be altered by occurrences of the control function Select Line Spacing (SVS) and Set Line Spacing (SLS).

When proportional line spacing is to be performed by the content layout process, the distance between the line home positions of two successive line boxes is evaluated by an implementation dependent algorithm. If the evaluated distance differs from the current value of line spacing, the control function Set Line Spacing (SLS) is inserted before the line terminator of the first line. The evaluated distance is inserted as the parameter of this control function and also becomes the current line spacing.

If the output of the content layout process is in formatted processable form, those occurrences of the control function Set Line Spacing (SLS) inserted by the content layout process are enclosed by the delimiters Start of String (SOS) and String Terminator (ST).

14.2.1.4 Identification of content portions

The content layout process shall also provide a value for the attribute “content portion identifier-layout” for each content portion associated with the layout structure.

14.2.1.5 Determination of object dimensions

The dimensions of the basic layout object created by the content layout process can only be determined after the formatting action is completed.

The dimension of the basic layout object in the direction of the line progression is such that this dimension is the minimum required to contain all line boxes that make up the formatted content allocated to it.

NOTE 1 – There may be only one line box created by the formatting process.

For the dimension of the basic layout object in the direction of the character path, two cases have to be distinguished.

– The formatting process can allocate all content to one single line box, taking into account all applicable presentation attributes. In this case, the dimension of the basic layout object in the direction of the character path is equal to the minimum dimension of that line box which can contain the formatted content.

  NOTE 2 – This dimension will usually be less than the dimension of the available area in the direction of the character path and therefore a value of “left-hand aligned“ or “centred“ for the attribute “block alignment” will have a visual effect.

– The formatting process cannot allocate all content to one single line box, taking into account all applicable presentation attributes. In this case, the dimension of the basic layout object in the direction of the character path is equal to the dimension of the available area in that direction. Full use of the value of this dimension should be made during the formatting process, i.e. the number of line boxes should be minimized, subject to the specifications of the presentation attributes and layout directives.

NOTE 3 – The formatting process is considered implementation dependent, i.e. different implementations may produce a different number of line boxes for the same character content.

The mapping to horizontal and vertical dimensions of the basic layout object is dependent upon the character path as shown in Table 5:

Table 5 – Character path and horizontal/vertical dimension of a basic layout object

Character path	Dimension in direction of character path	Dimension in direction orthogonal to character path
d0, d180 d90, d270	Horizontal dimension Vertical dimension	Vertical dimension Horizontal dimension

14.2.2 Single basic logical object to multiple basic layout objects

If the formatted content does not fit into the available area in the direction of line progression then an additional available area has to be obtained from the document layout process, e.g. in the case of a page boundary. In this case, the content of a single basic logical object is allocated to more than one basic layout object. The original content portion is split into several content portions, each corresponding to a different basic layout object.

There are two changes with respect to the description in 14.2.1:

– the presentation attributes of the second and subsequent basic layout objects are given values corresponding to the status of rendition (line spacing, character spacing, graphic renditions etc.) at the end of the previous basic layout object;

– the allocation of content to basic layout objects is performed such that the presentation attributes “orphan size” and “widow size” are fulfilled.

14.2.3 Multiple basic logical objects to single basic layout object

When concatenation results in the content associated with a more than one basic logical object being laid out in a single basic layout object (see 9.5), it may be necessary for the content layout process to insert control functions at the beginning of the second and subsequent basic logical objects so that the values of certain presentation attributes associated with those basic logical objects are applied. These control functions are:

– Select Character Spacing (SFS) or Set Character Spacing (SCS) for “character spacing”;

– Designation and/or invocation sequences for “graphic character sets”;

– Identify Graphic Subrepertoire (IGS) for “graphic character subrepertoire”;

– Select Graphic Rendition (SGR) for “graphic rendition”;

– Select Line Spacing (SVS) or Set Line Spacing (SLS) for “line spacing”.

If the output of the content layout process is in formatted processable form, then the control functions inserted by the content layout process are enclosed by the delimiters Start of String (SOS), Start of Original String (SOOS) and String Terminator (ST).

The presentation attribute “proportional line spacing” specified for the second or subsequent basic logical objects is interpreted by the content layout process as described in 14.2.1.

The other presentation attributes specified for the second and subsequent basic logical objects are ignored (see 9.5).

For logical and shared control functions other than those for which a corresponding presentation attribute exists that is not ignored when concatenation is in effect, the scope extends to the end of the concatenated sequence of basic logical objects. These control functions are: Partial Line Up (PLU) and Partial Line Down (PLD), Line Position Relative (VPR) and Line Position Backward (VPB), Start Reverse String (SRS), Parallel Texts (PTX), and Selective Tabulation (STAB). Accordingly, no insertion of such control functions at the start of a new basic logical object is necessary.

14.3 Content layout process for formatted processable content

For formatted processable form content, the initialization step of the content layout process shall

- remove all SOOS control functions and all SOS-ST control strings from the content;
- remove all layout control functions (BS, HPB, HPR, JFY, SACS, SRCS and SSW) from the content (see 13.2);
- combine all content portions associated with the same basic logical object into a single content portion in order to prevent unnecessary fragmentation of the document content which could otherwise occur;
- delete the content portion attribute “Content identifier-layout”, if present.

After initialization, the content is in processable form. The remaining steps of the content layout process are as described for processable content (see 14.2).

14.4 Content layout process for formatted content

For formatted content, the content layout process still has to determine the dimension of the basic layout object to be allocated.

The same steps of the content layout process are used as for processable form content. In this case, however, formatting only involves

- the positioning of character images within a line box as described in 7.2 and 14.2.1.3.1;
- the positioning of line boxes within a basic layout object as described in 7.3.

The dimension of the basic layout object in the direction of the character path is set equal to the minimum required to allow for the positioning within line boxes of all characters associated with the basic layout object.

15 Content imaging process

This clause describes a content imaging process for basic logical objects associated with content architecture of type character.

Its purpose is to aid understanding of the semantics of the presentation attributes and control functions by describing the required results of such a process. However, it is not intended to specify any process that might be carried out in a particular implementation to achieve these results.

15.1 Introduction

The content imaging process is only concerned with the layout structures, the presentation styles and the content of basic layout components conforming to this part of ITU-T Rec. T.410-Series | ISO/IEC 8613.

All logical control functions, if any, are ignored.

The content imaging process is only applicable to the formatted and formatted processable form character content architecture classes.

15.2 Content imaging process for formatted content

This subclause describes how the various shared and layout presentation attributes and shared and layout control functions influence the image of the contents.

Most shared presentation attributes and shared control functions serve the purpose of positioning and orienting character images along reference lines and of positioning and orienting these reference lines within the basic layout object.

Thus, the effects of most shared presentation attributes and shared control functions have already been described in the content layout process.

Most layout presentation attributes and control functions are related to positioning and, thus, have already been described as the result of the content layout process.

The following subclauses provide additional information relating to the content imaging process.

15.2.1 Determination of initial point

The active position for imaging is set on the initial point within the basic layout object. This information is derived from the presentation attribute “initial offset”.

15.2.2 Choosing character images

The following presentation attributes and control functions determine the character images to be chosen for imaging:

- “graphic character sets” and code extension announcers”, designation and invocation control functions;
- “character fonts” (together with the attribute “fonts list” in the document profile);
- “graphic rendition” and Select Graphic Rendition (SGR).

If the specified font is not available, the content imaging process may decide to substitute this font by making use of the font information available in the document profile attribute “fonts list”.

15.2.3 Formatting indicator

The presentation attribute formatting indicator” specifies whether first line offset, itemization, alignment, tabulation and pairwise kerning have already been performed by the content layout process or not.

If not, or if the result from the content layout process has been invalidated by a font substitution, then the content imaging process shall perform the task in the same way as described in the content layout process (see 14.2.1.3.1).

15.3 Content imaging process for formatted processable content

For content in formatted processable form, the only difference from the case of formatted form is that all logical control functions and the delimiter control functions Start of String (SOS) and String Terminator (ST) as well as any deleted text delimited by the delimiter control function Start of Original String (SOOS) and the next occurrence of one of the delimiter control functions SOS or ST are ignored.

The effect of shared and layout presentation attributes and shared and layout control functions is as described in 15.2.

16 Interactions between presentation attributes and control functions

This clause contains a summary of the interactions among and between presentation attributes and control functions defined in various places in this part of ITU-T Rec. T.410-Series | ISO/IEC 8613 as follows:

- Line Feed (LF) is restricted to be used in the following cases:
  1) at the beginning of the content of a basic layout component;
  2) immediately following a control function Carriage Return (CR);
  3) immediately following another LF.
- Break Permitted Here (BPH) or CR is not permitted in the following cases:
  1) when Partial Line Up (PLU) or Partial Line Down (PLD) is active;
  2) after Line Position Relative (VPR) and/or Line Position Backward (VPB) have moved the active position away from the reference line;
  3) between the control functions SRS 1 and SRS 0;
  4) between the control functions PTX 1 and PTX 0.

– Rendition aspects defined by some presentation attributes can be overridden by control functions embedded in content portions as shown in Table 6.

– When the presentation attribute “line layout table” specifies any tabulation stops, the presentation attribute “alignment” is assumed to have the value “start-aligned”.

Table 6 – Presentation attributes that can be overridden by control functions

Presentation attribute	Control function
Character spacing	SHS, SCS
Line spacing	SVS, SLS
Graphic rendition	SGR
Graphic character subrepertoire	IGS
Graphic character sets	Code extension control functions

17 Definition of character content architecture classes

This clause defines the three classes of character content architecture as described in 6.1, namely:

– A formatted form character content architecture which allows for document content to be presented (e.g. printed or displayed) as intended by the originator. Formatted form can be used in any basic component.

– A processable form character content architecture which allows for document content to be processed (e.g. edited or formatted). Processable form can be used in any basic logical component.

– A formatted processable form character content architecture which allows for document content to be processed and also to be presented as intended by the originator. Formatted processable form can be used in any basic component.

Subclauses 17.1, 17.2 and 17.3 specify the categories of presentation attributes and control functions that pertain to these content architecture classes. The individual presentation attributes and control functions are summarized in Tables 5 and 6.

The permissible values and the default values of the presentation attributes and the control functions parameter values are defined in clauses 9 and 13 respectively.

In order to aid the definition of a subset of a content architecture class for use in document application profiles (see ITU--T Rec. T.411 | ISO/IEC 8613-1), the presentation attributes and the control functions that are applicable to each content architecture class are listed in Annex B, together with their permissible values and default values.

17.1 Formatted character content architecture class

The following categories of presentation attributes and control functions pertain to the formatted character content architecture class:

– shared presentation attributes (see 9.1);

– layout presentation attributes (see 9.2);

– shared control functions (see 13.1);

– layout control functions (see 13.2).

17.2 Processable character content architecture class

The following categories of presentation attributes and control functions pertain to the processable character content architecture class:

- shared presentation attributes (see 9.1);
- logical presentation attributes (see 9.3);
- shared control functions (see 13.1);
- logical control functions (see 13.3).

17.3 Formatted processable character content architecture class

The following categories of presentation attributes and control functions pertain to the formatted processable character content architecture class:

- shared presentation attributes (see 9.1);
- layout presentation attributes (see 9.2);
- logical presentation attributes (see 9.3);
- shared control functions (see 13.1);
- layout control functions (see 13.2).
- logical control functions (see 13.3);
- delimiters (see 13.4).

Table 7 – Summary of presentation attributes

Presentation attribute	Character content architecture class		
	Formatted	Processable	Formatted processable
alignment	X	X	X
character fonts	X	X	X
character orientation	X	X	X
character path	X	X	X
character spacing	X	X	X
code extension announcers	X	X	X
first line offset	X		X
formatting indicator	X	X	X
graphic character sets	X	X	X
graphic character subrepertoire	X	X	X
graphic rendition	X	X	X
indentation		X	X
initial offset	X		X
itemization	X	X	X
kerning offset	X	X	X
line layout table	X	X	X
line progression	X	X	X
line spacing	X	X	X
orphan size		X	X
pairwise kerning	X	X	X
proportional line spacing		X	X
widow size		X	X

Table 8 – Summary of control functions

Control function		Character content architecture		
		Formatted	Processable	Formatted processable
BPH	Break Permitted Here		X	X
BS	Backspace	X		X
CR	Carriage Return	X	X	X
GCC	Graphic Character Composition	X	X	X
HPB	Character Position Relative	X		X
HPR	Character Position Backward	X		X
IGS	Identify Graphic Subrepertoire	X	X	X
JFY	No Justify	X		X
LF	Line Feed	X	X	X
NBH	No Break Here		X	X
PLD	Partial Line Down	X	X	X
PLU	Partial Line Up	X	X	X
PTX	Parallel Texts		X	X
SACS	Set Additional Character Spacing	X		X
SCS	Set Character Spacing	X	X	X
SGR	Select Graphic Rendition	X	X	X
SHS	Select Character Spacing	X	X	X
SLS	Set Line Spacing	X	X	X
SOOS	Start of Original String			X
SOS	Start of String			X
SP	SPACE	X	X	X
SRCS	Set Reduced Character Spacing	X		X
SRS	Start Reverse String	X	X	X
SSW	Set SPACE Width	X		X
ST	String Terminator			X
STAB	Selective Tabulation	X	X	X
SUB	Substitute	X	X	X
SVS	Selective Line Spacing	X	X	X
VPB	Line Position Backward	X	X	X
VPR	Line Position Relative	X	X	X
Code extension control functions		X	X	X

Annex A
SGML representation of character content-specific attributes for ODL

(This annex forms an integral part of this International Standard)

A.1 Introduction

This annex specifies a standardized SGML representation of attributes related to the character content architecture, for use with the Office Document Language (ODL) defined in ITU-T Rec. T.415 | ISO/IEC 8613-5. ODL is an SGML application conforming to ISO 8879.

The definitions of ISO 8879 apply to this annex.

A.2 Names and public identifiers

The following notation declarations include the public identifiers of the data content notations for content architecture classes defined in this Specification. The ODL content architecture class names follow the prefix ODA" in the notation names.

<!NOTATION ODAcf PUBLIC ISO/IEC 8613-6:1993//NOTATION
Character formatted content architecture//EN">

<!NOTATION ODAcfp PUBLIC ISO/IEC 8613-6:1993//NOTATION
Character formatted processable content architecture//EN">

<!NOTATION ODAcp PUBLIC ISO/IEC 8613-6:1993//NOTATION
Character processable content architecture//EN">

A.3 Representation of attribute values

Attribute values are represented in a clear text encoding, using the rules defined in this annex.

NOTE 1 – The content portions themselves are encoded according to the body of this part of ITU-T Rec. T.410-Series | ISO/IEC 8613.

The representations of the ODA attributes are presented in the form of SGML public text. In this form they can be referenced from a document, rather than be included within it.

The semantics of the attribute values are specified in the body of this part of ITU-T Rec. T.410-Series | ISO/IEC 8613.

The representation of attribute values is as specified in the body of this part of ITU-T Rec. T.410-Series | ISO/IEC 8613, except where a different representation is specified in the public text or elsewhere in this annex.

The default values specified in the public text are those defined in the body of this part of ITU-T Rec. T.410-Series | ISO/IEC 8613. If a different default value is wanted for an element (such as a non-standard default value specified in the document profile or in an object class description), the public text should not be referenced; instead, the definitions should be duplicated with the required changes made in the default values.

Attribute values are sequences of one or more parameters, separated by SGML separator characters. An omitted parameter is represented by the keyword: 00

A parameter is one of a number of primitive types: string, keyword, or integer. String parameters are delimited, and may contain separator characters. Other parameters are not delimited, and may not contain separator characters.

NOTE 2 – Most attribute values consist of a single parameter.

A.3.1 Constructed parameters

In this part of ITU-T Rec. T.410-Series | ISO/IEC 8613, a parameter is a constructed parameter if one or more of its permissible values is a group of two or more sub-parameters. The description of the attribute in the body of this part of ITU-T Rec. T.410-Series | ISO/IEC 8613 determines the number of sub-parameters, and whether any can be omitted. If more than one sub-parameter is present, they are separated from one another by commas. Successive commas denote an omitted sub-parameter, but they are required only if a succeeding sub-parameter is present.

NOTE - For an example of a constructed parameter, see the attribute line layout table" in the public text.

A.3.2 Parentheses

A constructed parameter is normally enclosed in parentheses but they can be omitted if no ambiguity would be created by doing so. If a sub-parameter is itself constructed, it shall be enclosed in parentheses.

An empty constructed parameter is represented by: ( )

A.3.3 Alternative representation

The value of a constructed parameter can optionally be represented as the name of a data entity that contains the actual constructed parameter.

NOTE - This technique can be used for long, constructed parameters that would otherwise cause the quantity limits of the concrete syntax to be exceeded.

A.3.4 String parameters

A string parameter could contain characters not permitted in an SGML name token, and is therefore delimited by SGML LIT or LITA delimiters.

A string parameter that is an escape sequence formulated in accordance with ISO 2022 is represented in the form used for the public text designating sequence" defined in ISO 8879.

NOTE – This is the clear text form commonly used in ISO standards.

A.3.5 Keyword parameters

Possible keyword values are defined by the body of this part of ITU-T Rec. T.410-Series | ISO/IEC 8613 for some parameters, and by this annex for others.

Lowercase letters in keyword parameters are treated as though they were uppercase.

For certain parameters whose permissible values constitute a set of keywords, fixed numeric values, or both keywords and fixed numeric values, the value is represented by choosing from a set of substitute keywords. These parameters are documented in comments in the public text, in the form:

parameter name: keyword keyword ...

with the keywords appearing in the same order as the permissible values that they represent appear within the body of this part of ISO 8613. For attributes that have only one parameter, the attribute name is the parameter name.

NOTE - For example: -- alignment: S E C J - means that in the alignment attribute, a value of "s" represents "start-aligned", a value of "e" represents "end- aligned", and so on.

A.3.6 Integer parameters

An integer is represented by a sequence of digits. If preceded by a hyphen, it represents a negative integer; otherwise, a positive integer.

Parameters whose permissible values constitute an enumerated set of quantities of degrees or SMUs are represented by the integer quantities alone, without the word "degrees" or "SMU".

A.4 Presentation attributes

A.4.1 Shared presentation attributes (format attribute-directives)

```
<! --     (C) International Organization for Standardization 1993
          Permission to copy in any form is granted for use with conforming SGML systems and applications as
          defined in ISO 8879, provided this notice is included in all copies.

-->

<! --     Public text entity. Typical invocation:

<! ENTITY % c-p-ad     PUBLIC "ISO/IEC 8613-6:1993//TEXT
                       Character Presentation format Attribute-Directives//EN">

<! ATTLIST (cf|cfp)       %c-p-a; %c-p-ad;>

-->
          calign          NAME          s             -- alignment: S E C J --
          cfonts          CDATA         #IMPLIED      -- character fonts --
          corient         NUMBER        0             -- character orientation: 0 90 180 270 --
          cpath           NUMBER        0             -- character path: 0 90 180 270 --
          ccharspc        NUMBER        #IMPLIED      -- character spacing --
```

```
        cannounc        CDATA           #IMPLIED        -- code extension announcer --
        cfloff          NMTOKEN         0               -- first line offset --
        cgrcsets        CDATA           #IMPLIED        -- graphic character sets --
        cgrcsub         NUMBER          0               -- graphic character subrepertoire --
        cgrrend         NUMBERS         0               -- graphic rendition --
        citem           NMTOKENS        #IMPLIED        -- itemization --
                        -- identifier alignment: N S E --
                        -- identifier start offset: integer --
                        -- identifier end offset: integer --
        ckernoff        NUMBERS         "0 0"           -- kerning offset
        lat             CDATA           " "             -- line layout table
                        -- tabulation reference: integer --
                        -- tabulation position: integer --
                        -- alignment: S E C A --
                        -- alignment character string: string --
                        -- Example: lat = "(1,600,s) (2,2400,a,".") (3,4800,e)" --
      clinepro          NUMBER          270             -- line progression: 90 270 --
      clinespc          NUMBER          #IMPLIED        -- line spacing --
      cpwkern           NAME            nk              -- pairwise kerning: K NK --
```

A.4.2 Layout presentation attributes (format attributes)

```
<! -- (C) International Organization for Standardization 1993
        Permission to copy in any form is granted for use with conforming SGML systems and applications as
        defined in ISO 8879, provided this notice is included in all copies.

 -->

<! -- Public text entity. Typical invocation:

<!ENTITY % c-p-a        PUBLIC ISO/IEC 8613-6: 1993//TEXT
                        Character Presentation Format Attributes//EN">

<!ATTLIST (cf|cfp) %c-p-a; %c-p-ad; >

-->
        cformat         NAME            nf              -- formatting indicator: F NF --
        cinitoff        NUMBERS         #IMPLIED        -- initial offset--
```

A.4.3 Logical presentation attributes (format directives)

```
<! -- (C) International Organization for Standardization 1993
        Permission to copy in any form is granted for use with conforming SGML systems and applications as
        defined in ISO 8879, provided this notice is included in all copies.

-->

<! -- Public text entity. Typical invocation:

<!ENTITY % c-p-d  PUBLIC ISO/IEC 8613-6: 1993//TEXT
                  Character Presentation Format Directives//EN">

<!ATTLIST (cfp|cp)      %c-p-d; >

-->
        cindent         NUMBER          0               -- indentation --
        corphan         NUMBER          1               -- orphan size --
        cpropls         NAME            np              -- proportional line spacing: P NP --
        cwidow          NUMBER          1               -- widow size --
```

A.5 Coding attributes

No character content coding attributes are defined for ODL.

Annex B
Summary of content architecture classes

(This annex does not form an integral part of this Recommendation | International Standard.)

This annex summarizes the presentation attributes and control functions that apply to each of the three content architecture classes (formatted, processable and formatted processable) defined in clause 17, together with their permissible values and default values or assumed values in the case that a control function/ parameter is not specified.

The purpose of this annex is to facilitate the definition of a subset of a content architecture class for use in document application profiles (see ITU-T Rec. T.411 | ISO/IEC 8613-1).

B.1 Formatted character content architecture class

Table B.1 – Summary of permissible and default values for presentation attributes applicable to the formatted character content architecture class

Attribute	Permissible value(s)	Default value
alignment	Start-aligned, end-aligned, centred, justified	Start-aligned
character fonts		
font size	Any positive integer	None
font identifier	Any positive integer	None
character orientation	d0, d90, d180, d270	d0
character path	0, d90, d180, d270	d0
character spacing	Any positive integer	Equivalent of 120 BMUs
code extension announcers	Any string of escape sequences in accordance with ISO 2022	Escape sequences for the G0 and G2 sets
first line offset	Any integer	0
formatting indicator	No, yes	No
graphic character sets	The escape sequences to designate, and any required locking shift functions to invoke, one or more registered graphic character sets	The escape sequences to designate, and the locking shift functions to invoke, the graphic character sets of ISO 6937-2
graphic character subrepertoire	0 or the identifier of any registered subrepertoire of ISO 6937	0
graphic rendition	0, 1-7, 9, 10-19, 21-27, 29, 30-38, 40-48, 50	0
initial offset		
horizontal coordinate	Any non-negative integer	See Table 2
vertical coordinate	Any non-negative integer	See Table 2
itemization		
identifier alignment	No itemization, start-aligned, end-aligned	No itemisation
identifier start offset	Any integer	The distance from the start edge of the positioning area to the line home position
identifier end offset	Any integer	0
kerning offset		
start edge offset	Any non-negative integer	0
end edge offset	Any non-negative integer	0
line layout table		
tab reference	One to four decimal digits	)
tab position	Any non-negative integer	)
alignment	Start-aligned, end-aligned, centred, aligned-around	) No) Tabulationstops)Defined)
alignment string	Any characters	)
line progression	d90, d270	d270
line spacing	Any positive integer	Equivalent of 200 BMUs
pairwise kerning	Yes, no	No

Table B.2 – Summary of permissible and assumed values for control functions applicable to the formatted character content architecture class

Attribute	Permissible value(s)	Default value
BS	Not applicable	Not applicable
CR	Not applicable	Not applicable
GCC	0, 1, 2	0
HPB	Any positive integer	Equivalent of 120 BMUs
HPR	Any positive integer	Equivalent of 120 BMUs
IGS	0 or the identifier of any registered subrepertoire of ISO 6937	0
JFY	0	0
LF	Not applicable	Not applicable
PLD	Not applicable	Not applicable
PLU	Not applicable	Not applicable
SACS	Any positive integer	0
SCS	Any positive integer	Equivalent of 120 BMUs
SGR	0, 1-7, 9, 10-19, 21-27, 29, 30-38, 40-48, 50	0
SHS	0, 1, 2, 3	0
SLS	Any positive integer	Equivalent of 200 BMUs
SP	Not applicable	Not applicable
SRCS	Any positive integer	0
SRS	0, 1	0
SSW	Any positive integer	None
STAB	Any	None
SUB	Not applicable	Not applicable
SVS	0, 1, 2, 3, 4, 9	0
VPB	Any positive integer	Equivalent of 100 BMUs
VPR	Any positive integer	Equivalent of 100 BMUs

In addition, any code extension control functions defined in ISO 2022, within the scope of the value of the attribute “code extension announcers”, is permitted.

B.2 Processable character content architecture class

Table B.3 – Summary of permissible and default values for presentation attributes applicable to the processable character content architecture class

Attribute	Permissible value(s)	Default value
alignment	Start-aligned, end-aligned, centred, justified	Start-aligned,
character fonts	Any positive integer	None
font size	Any positive integer	None
font identifier	d0, d90, d180, d270	d0
character orientation	d0, d90, d180, d270	d0
character path	Any positive integer	Equivalent of 120 BMUs
character spacing	Any string of escape	Escape sequences for the G0 and
code extension announcers	Sequences in accordance with ISO 2022	G2 sets
first line offset	Any integer	0
formatting indicator	No, yes	No
graphic character sets	The escape sequences to designate, and any required locking shift functions to invoke, one or more registered graphic character sets	The escape sequences to designate, and the locking shift functions to invoke, the graphic character sets of ISO 6937-2
graphic character subrepertoire	0 or the identifier of any registered subrepertoire of ISO 6937	0
graphic rendition	0, 1-7, 9, 10-19, 21-27, 29, 30-38, 40-48, 50	0
indentation	Any non-negative integer	0
itemization		
identifier alignment	No itemization, start-aligned, end-aligned	No itemization
identifier start offset	Any integer	The distance from the start edge of the positioning area to the line home position
identifier end offset	Any integer	0
kerning offset		
start edge offset	Any non-negative integer	0
end edge offset	Any non-negative integer	0
line layout table		
tab reference	One to four decimal digits	None
tab position	Any non-negative integer	None
alignment	Start-aligned, end-aligned, centred, aligned-around	Start-aligned
alignment string	Any characters	None
line progression	d90, d270	d270
line spacing	Any positive integer	Equivalent of 200 BMUs
orphan size	Any positive integer	1
pairwise kerning	Yes, no	No
proportional line spacing	Yes, no	No
widow size	Any positive integer	1

Table B.4 – Summary of permissible and assumed values for control functions applicable to the processable character content architecture class

Attribute	Permissible value(s)	Default value
BPH	Not applicable	Not applicable
CR	Not applicable	Not applicable
GCC	0, 1, 2	0
IGS	0 or the identifier of any registered subrepertoire of ISO 6937	0
LF	Not applicable	Not applicable
NBH	Not applicable	Not applicable
PLD	Not applicable	Not applicable
PLU	Not applicable	Not applicable
PTX	Not applicable	Not applicable
SCS	Any positive integer	Equivalent of 120 BMUs
SGR	0, 1-7, 9, 10-19, 21-27, 29, 30-38, 40-48, 50	0
SHS	0, 1, 2, 3	0
SLS	Any positive integer	Equivalent of 200 BMUs
SP	Not applicable	Not applicable
SRS	0, 1	0
STAB	Any	None
SUB	Not applicable	Not applicable
SVS	0, 1, 2, 3, 4, 9	0
VPB	Any positive integer	Equivalent of 100 BMUs
VPR	Any positive integer	Equivalent of 100 BMUs

In addition, any code extension control functions defined in ISO 2022, within the scope of the value of the attribute “code extension announcers”, is permitted.

B.3 Formatted processable character content architecture class

Table B.5 – Summary of permissible and default values for presentation attributes applicable to the formatted processable character content architecture classs

Attribute	Permissible value(s)	Default value
alignment	Start-aligned, end-aligned, centred, justified	Start-aligned
character fonts		
font size	Any positive integer	None
font identifier	Any positive integer	None
character orientation	d0, d90, d180, d270	d0
character path	d0, d90, d180, d270	d0
character spacing	Any positive integer	Equivalent of 120 BMUs
code extension announcers	Any string of escape sequences in accordance with ISO 2022	Escape sequences for the G0 and G2 sets
first line offset	Any integer	0
formatting indicator	No, yes	No
graphic character sets	The escape sequences to designate, and any required locking shift functions to invoke, one or more registered graphic character sets	The escape sequences to designate, and the locking shift functions to invoke, the graphic character sets of ISO 6937-2
graphic character subrepertoire	0 or the identifier of any registered subrepertoire of ISO 6937	0
graphic rendition	0, 1-7, 9, 10-19, 21-27, 29, 30-38, 40-48, 50	0
initial offset		
horizontal coordinate	Any non-negative integer	See Table 2
vertical coordinate	Any non-negative integer	See Table 2
indentation	Any positive integer	0
itemisation		
identifier alignment	No itemisation, start-aligned, end-aligned	No itemisation
identifier start offset	Any integer	The distance from the start edge of the positioning area to the line home position
identifier end offset	Any integer	0
kerning offset		
start edge offset	Any non-negative integer	0
end edge offset	Any non-negative integer	0
line layout table		
tab reference	One to four decimal digits	None
tab position	Any non-negative integer	None
alignment	start-aligned; end- aligned, centred, aligned-around	start aligned
alignment string	Any characters	None
line progression	d90, d270	d270
line spacing	Any positive integer	Equivalent of 200 BMUs
orphan size	Any positive integer	1
pairwise kerning	Yes, no	No
proportional line spacing	Yes, no	No
widow size	Any positive integer	1

Table B.6 – Summary of permissible and assumed values for control functions applicable to the formatted processable character content architecture class

Control function	Permissible value(s)	Assumed value
BPH	Not applicable	Not applicable
BS	Not applicable	Not applicable
CR	Not applicable	Not applicable
GCC	0, 1, 2	0
HPB	Any positive integer	Equivalent of 120 BMUs
HPR	Any positive integer	Equivalent of 120 BMUs
IGS	0 or the identifier of any registered subrepertoire of ISO 6937	0
JFY	0	0
LF	Not applicable	Not applicable
NBH	Not applicable	Not applicable
PLD	Not applicable	Not applicable
PLU	Not applicable	Not applicable
PTX	0, 1, 3	0
SACS	Any positive integer	0
SCS	Any positive integer	Equivalent of 120 BMUs
SGR	0, 1-7, 9, 10-19, 21-27, 29, 30-38, 40-48, 50	0
SHS	0, 1, 2, 3	0
SLS	Any positive integer	Equivalent of 200 BMUs
SOOS	Not applicable	Not applicable
SOS	Not applicable	Not applicable
SP	Not applicable	Not applicable
SRCS	Any positive integer	0
SRS	0, 1	0
SSW	Any positive integer	None
ST	Not applicable	Not applicable
STAB	Any	None
SUB	Not applicable	Not applicable
SVS	0, 1, 2, 3, 4, 9	0
VPB	Any positive integer	Equivalent of 100 BMUs
VPR	Any positive integer	Equivalent of 100 BMUs

In addition, any code extension control functions defined in ISO 2022, within the scope of the value of the attribute "code extension announcers", is permitted.

Annex C
Coded representations of control functions

(This annex does not form an integral part of this Recommendation | International Standard.)

Coded representations of control functions are defined in ISO 6429. A summary of the coded representations of the control functions defined in this Specification is given in Table C.1.

Table C.1 – Coded representations of control functions

Control function		Coded representation
BPH	Break Permitted Here	08/02
BS	Backspace	00/08
CR	Carriage Return	00/13
HPB	Character Position Backward	CSI Pn 06/10
HPR	Character Position Relative	CSI Pn 06/01
GCC	Graphic Character Composition	CSI Ps 02/00 05/15
IGS	Identify Graphic Subrepertoire	CSI Ps 02/00 04/13
JFY	No Justify	CSI 02/00 04/06
LF	Line Feed	00/10
NBH	No Break Here	08/03
PLD	Partial Line Down	08/11
PLU	Partial Line Up	08/12
PTX	Parallel Texts	CSI Ps 05/12
SCS	Set Character Spacing	CSI Pn 02/00 06/07
SGR	Select Graphic Rendition	CSI Ps... 06/13
SHS	Select Character Spacing	CSI Ps 02/00 04/11
SACS	Set Additional Character Spacing	CSI Pn 02/00 05/12
SLS	Set Line Spacing	CSI Pn 02/00 06/08
SRCS	Set Reduced Character Spacing	CSI Pn 02/00 06/06
SOOS	Start Of Original String	<currently undefined>
SOS	Start Of String	09/08
SSW	Set SPACE Width	CSI Pn 02/00 05/11
SP	Space	02/00
SRS	Start Reverse String	CSI Ps 05/11
ST	String Terminator	09/12
STAB	Selective Tabulation	CSI Ps 02/00 05/14
SUB	Substitute	01/10
SVS	Select Line Spacing	CSI Ps 02/00 04/12
VPB	Line Position Backwards	CSI Pn 06/11
VPR	Line Position Relative	CSI Pn 06/05

In Table C.1, CSI denotes the Control Sequence Introducer represented by the bit-combination 09/11 and Pn and Ps denote respectively numeric and selective parameter values represented by one or more bit combinations in the range 03/00 to 03/09.

The coded representation of a control function with parameters but with Pn or Ps omitted represents that control function with the default value of the parameter.

Annex D
Summary of object identifiers

(This annex does not form an integral part of this Recommendation | International Standard.)

Values of ASN.1 object identifiers are assigned in various clauses of this part of ITU-T Rec. T.410-Series | ISO/IEC 8613. These assignments are summarized in Table D.1.

Table D.1 – Summary of ASN.1 object identifiers

Object identifier value	Description	Clause
{ 2 8 1 6 2 }	Identifies module Character-Presentation-Attributes	11.2
{ 2 8 1 6 3 }	Identifies module Character-Coding-Attributes	11.3
{ 2 8 1 6 4 }	Identifies module Character-Profile-Attributes	11.4
{ 2 8 2 6 0 }	Identifies Formatted character content architecture class	9.4
{ 2 8 2 6 1 }	Identifies Processable character content architecture class	9.4
{ 2 8 2 6 2 }	Identifies Formatted-processable character content architecture class	9.4
{ 2 8 3 6 0 }	Identifies ISO 2022 type of coding	10.1