

INTERNATIONAL
STANDARDIZED
PROFILE

ISO/IEC
ISP
11181-1

First edition
1993-11-01

**Information technology — International
Standardized Profile FOD26 — Open
Document Format: Enhanced document
structure — Character, raster graphics and
geometric graphics content
architectures —**

Part 1:
Document Application Profile (DAP)

*Technologies de l'information — Profil normalisé international FOD26 —
Format de document ouvert: Structure de document avancé —
Architectures du contenu des caractères, des graphiques à raster et des
graphiques géométriques —
Partie 1: Profil du document d'application (DAP)*



Reference number
ISO/IEC ISP 11181-1:1993(E)

AVA **Documents**
Document Sharing Hub

Contents

1	Scope	1
2	Normative references	2
2.1	Paired Recommendations/International Standards equivalent in technical content	2
2.2	Additional references	3
3	Definitions	4
4	Relationship with other International Standardized Profiles	4
5	Conformance	5
5.1	Data stream conformance	5
5.2	Implementation conformance	5
6	Characteristics supported by this document application profile	5
6.1	Overview	6
6.1.1	General	6
6.1.2	Formatted form documents	6
6.1.3	Processable form documents	6
6.1.4	Formatted processable form documents	7
6.1.5	Generic-documents	7
6.1.6	Document profile	7
6.2	Logical characteristics	7
6.2.1	Introduction	7
6.2.2	Overview of the logical structure	8
6.2.3	Body part of the logical structure	8
6.2.3.1	DocumentLogicalRoot	8
6.2.3.2	Passage	8
6.2.3.3	NumberedSegment	9
6.2.3.4	Number	9
6.2.3.5	Paragraph	10
6.2.3.6	BodyText, BodyRaster and BodyGeometric	10
6.2.3.7	Footnote	10
6.2.3.8	FootnoteReference	11
6.2.3.9	FootnoteBody	11
6.2.3.10	FootnoteNumber	11
6.2.3.11	FootnoteText	12
6.2.4	Common content part of the logical structure	12
6.2.4.1	CommonContent	12

© ISO/IEC 1993

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

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

Printed in Switzerland

6.2.4.2	CommonText.....	12
6.2.4.3	PageNumber.....	12
6.2.4.4	CommonRaster.....	13
6.2.4.5	CommonGeometric	13
6.3	Layout characteristics.....	13
6.3.1	Overview of the layout characteristics	13
6.3.2	DocumentLayoutRoot.....	14
6.3.3	PageSet.....	14
6.3.4	Page characteristics	14
6.3.4.1	Page constituents	14
6.3.4.2	Page dimensions	15
6.3.4.3	Nominal page sizes.....	15
6.3.4.4	Page offset.....	16
6.3.4.5	Page layout characteristics.....	16
6.3.4.5.1	General characteristics.....	16
6.3.4.5.2	Page layout A	16
6.3.4.5.3	Page layout B	16
6.3.4.5.4	Page layout C	16
6.3.4.5.5	Page layout D	16
6.3.5	Body area characteristics	19
6.3.5.1	General characteristics.....	19
6.3.5.2	BasicBody	19
6.3.5.3	VariableCompositeBody.....	19
6.3.5.4	BasicFloat.....	23
6.3.5.5	SnakingColumns.....	23
6.3.5.6	SynchronizedColumns	25
6.3.5.7	ColumnVariable	26
6.3.5.8	ColumnFixed.....	26
6.3.5.9	FootnoteArea.....	26
6.3.6	Header and footer area characteristics.....	27
6.3.6.1	General characteristics.....	27
6.3.6.2	BasicHeader and BasicFooter	27
6.3.6.3	CompositeHeader and CompositeFooter	27
6.3.6.4	SourcedContentVariable	27
6.3.6.5	ArrangedContentVariable	28
6.3.6.6	SourcedContentFixed.....	28
6.3.6.7	ArrangedContentFixed.....	28
6.3.7	GenericBlock and SpecificBlock	28
6.4	Document layout characteristics	28
6.4.1	Flow controls.....	28
6.4.1.1	Allocation of content to page sets	29
6.4.1.2	Page breaks.....	29
6.4.1.3	Allocation of content to body areas.....	29

6.4.1.3.1	Layout of content into column areas	30
6.4.1.3.2	Layout of footnotes.....	30
6.4.1.4	Allocation of content to header-footer areas.....	31
6.4.2	Layout of the document content.....	31
6.4.2.1	Margins	31
6.4.2.2	Separation.....	31
6.4.2.3	Indivisibility	31
6.4.2.4	Same layout object	31
6.4.2.5	Concatenation.....	32
6.4.2.6	Block alignment	32
6.4.3	Layout controls applicable in the absence of a generic layout structure	32
6.5	Content layout and imaging characteristics.....	33
6.5.1	Character content.....	33
6.5.1.1	Introduction	33
6.5.1.2	Character content architecture class	34
6.5.1.3	Character repertoires.....	34
6.5.1.4	Code extension techniques.....	35
6.5.1.5	Line spacing	37
6.5.1.6	Character spacing	37
6.5.1.7	Character path and line progression	37
6.5.1.8	Character positioning controls	38
6.5.1.9	Character orientation	38
6.5.1.10	Emphasis	38
6.5.1.11	Tabulation	39
6.5.1.12	Indentation.....	39
6.5.1.13	Alignment.....	39
6.5.1.14	First line format	40
6.5.1.15	Widow and orphan sizes	40
6.5.1.16	Fonts.....	40
6.5.1.17	Reverse character strings	40
6.5.1.18	Kerning offset.....	41
6.5.1.19	Proportional line spacing.	41
6.5.1.20	Superscripts and subscripts	41
6.5.1.21	Line breaks	41
6.5.1.22	Substitution of characters.....	41
6.5.1.23	Initial point.....	41
6.5.1.24	Use of control functions.....	41
6.5.1.25	Formatting the content.....	42
6.5.2	Raster graphics content.....	42
6.5.2.1	Introduction	42
6.5.2.2	Raster graphics content architecture.....	42
6.5.2.3	Raster graphics encoding methods	42
6.5.2.4	Pel path and line progression	43
6.5.2.5	Clipping.....	43
6.5.2.6	Pel spacing	43
6.5.2.7	Spacing ratio.....	43

6.5.2.8 Image dimensions	43
6.5.3 Geometric graphics content	44
6.6 Miscellaneous features	44
6.6.1 Resource-documents	44
6.6.2 External documents	44
6.6.3 Border	44
6.6.4 Application comments	44
6.6.5 Alternative representation	45
6.6.6 Automatic numbering mechanisms	45
6.6.6.1 Page Numbering	45
6.6.6.2 Segment numbering	47
6.6.6.3 Footnote numbering	48
6.6.7 User readable comments	49
6.6.8 User visible name	49
6.7 Document management features	49
6.7.1 Document constituent information	49
6.7.1.1 Presence of document constituents	50
6.7.1.2 Resource-document information	50
6.7.1.3 External document information	50
6.7.2 Document identification information	50
6.7.2.1 Document application profile information	50
6.7.2.2 Document architecture class information	50
6.7.2.3 Content architecture classes information	50
6.7.2.4 Interchange format class information	50
6.7.2.5 ODA version information	50
6.7.2.6 Document reference	50
6.7.3 Document default information	51
6.7.4 Non-basic characteristics information	51
6.7.4.1 Profile character sets	51
6.7.4.2 Comments character sets	52
6.7.4.3 Alternative representation character sets	52
6.7.5 Fonts list	52
6.7.6 Document management attributes	52
7 Specification of constituent constraints	53
7.1 Introduction	53
7.1.1 Diagrams of relationships of logical constituents	53
7.1.2 Diagrams of relationships of layout constituents	55
7.1.3 Notation	56
7.2 Document profile constituent constraints	58
7.2.1 Macro definitions	58

7.2.2	Constituent constraints.....	62
7.2.2.1	DocumentProfile {	62
7.3	Logical constituent constraints	68
7.3.1	Macro definitions	68
7.3.2	Factor constraints	75
7.3.2.1	FACTOR ANY-LOGICAL {	75
7.3.3	Constituent constraints.....	75
7.3.3.1	DocumentLogicalRoot: ANY-LOGICAL {	75
7.3.3.2	Passage: ANY-LOGICAL {	76
7.3.3.3	NumberedSegment: ANY-LOGICAL {	77
7.3.3.4	Number: ANY-LOGICAL {	77
7.3.3.5	Paragraph: ANY-LOGICAL {	78
7.3.3.6	BodyText: ANY-LOGICAL {	78
7.3.3.7	BodyGeometric: ANY-LOGICAL {	79
7.3.3.8	BodyRaster: ANY-LOGICAL {	80
7.3.3.9	Footnote: ANY-LOGICAL {	80
7.3.3.10	FootnoteReference: ANY-LOGICAL {	81
7.3.3.11	FootnoteBody: ANY-LOGICAL {	82
7.3.3.12	FootnoteNumber: ANY-LOGICAL {	82
7.3.3.13	FootnoteText: ANY-LOGICAL {	83
7.3.3.14	CommonContent {	83
7.3.3.15	CommonText {	83
7.3.3.16	PageNumber {	84
7.3.3.17	CommonGeometric {	84
7.3.3.18	CommonRaster {	85
7.4	Layout constituent constraints	85
7.4.1	Macro definitions	85
7.4.2	Factor constraints	88
7.4.2.1	FACTOR ANY-LAYOUT {	88
7.4.2.2	FACTOR ANY-PAGE: ANY-LAYOUT {	88
7.4.2.3	FACTOR ANY-FRAME-FIXED: ANY-LAYOUT {	89
7.4.2.4	FACTOR ANY-FRAME-VARIABLE: ANY-LAYOUT {	90
7.4.2.5	FACTOR BLOCK {	90
7.4.3	Constituent constraints.....	92
7.4.3.1	DocumentLayoutRoot: ANY-LAYOUT {	92
7.4.3.2	PageSet: ANY-LAYOUT {	92
7.4.3.3	Page: ANY-PAGE {	93
7.4.3.4	RectoPage: ANY-PAGE {	93
7.4.3.5	VersoPage: ANY-PAGE {	94
7.4.3.6	BasicBody: ANY-FRAME-FIXED {	95
7.4.3.7	VariableCompositeBody: ANY-FRAME-FIXED {	95
7.4.3.8	BasicFloat: ANY-FRAME-VARIABLE {	96
7.4.3.9	SynchronizedColumns: ANY-FRAME-VARIABLE {	97
7.4.3.10	SnakingColumns: ANY-FRAME-VARIABLE {	99
7.4.3.11	ColumnVariable: ANY-FRAME-VARIABLE {	100

7.4.3.12 ColumnFixed: ANY-FRAME-VARIABLE {	102
7.4.3.13 FootnoteArea: ANY-FRAME-VARIABLE {	103
7.4.3.14 BasicHeader: ANY-FRAME-FIXED {	104
7.4.3.15 BasicFooter: ANY-FRAME-FIXED {	105
7.4.3.16 CompositeHeader: ANY-FRAME-FIXED {	105
7.4.3.17 CompositeFooter: ANY-FRAME-FIXED {	106
7.4.3.18 SourcedContentVariable: ANY-FRAME-VARIABLE {	106
7.4.3.19 ArrangedContentVariable: ANY-FRAME-VARIABLE {	108
7.4.3.20 SourcedContentFixed: ANY-FRAME-VARIABLE {	109
7.4.3.21 ArrangedContentFixed: ANY-FRAME-FIXED {	110
7.4.3.22 GenericBlock: BLOCK {	110
7.4.3.23 SpecificBlock: BLOCK {	111
7.5 Layout style constituent constraints	112
7.5.1 Macro definitions	112
7.5.2 FACTOR ANY-LAYOUT-STYLE {	112
7.5.3 Constituent constraints.....	112
7.5.3.1 L-Style1: ANY-LAYOUT-STYLE {	112
7.5.3.2 L-Style2: ANY-LAYOUT-STYLE {	113
7.5.3.3 L-Style3: ANY-LAYOUT-STYLE {	113
7.5.3.4 L-Style4: ANY-LAYOUT-STYLE {	114
7.5.3.5 L-Style5: ANY-LAYOUT-STYLE {	114
7.5.3.6 L-Style6: ANY-LAYOUT-STYLE {	115
7.5.3.7 L-Style7: ANY-LAYOUT-STYLE {	115
7.5.3.8 L-Style8: ANY-LAYOUT-STYLE {	115
7.5.3.9 L-Style9: ANY-LAYOUT-STYLE {	115
7.5.3.10 L-Style10: ANY-LAYOUT-STYLE {	116
7.5.3.11 L-Style11: ANY-LAYOUT-STYLE{	116
7.6 Presentation style constituent constraints	116
7.6.1 Macro definitions	116
7.6.2 Factor constraints	116
7.6.2.1 FACTOR ANY-PRESENTATION-STYLE {	116
7.6.3 Constituent constraints.....	117
7.6.3.1 P-Style1: ANY-PRESENTATION-STYLE {	117
7.6.3.2 P-Style2: ANY-PRESENTATION-STYLE {	118
7.6.3.3 P-Style3: ANY-PRESENTATION-STYLE {	118
7.6.3.4 P-Style4: ANY-PRESENTATION-STYLE {	119
7.7 Content portion constituent constraints	119
7.7.1 Macro definitions	119
7.7.2 Factor constraints	119
7.7.2.1 FACTOR ANY-CONTENT {	119
7.7.3 Constituent constraints.....	120
7.7.3.1 Character-content-portion: ANY-CONTENT {	120
7.7.3.2 Raster-graphics-content-portion: ANY-CONTENT {	121
7.7.3.3 Geometric-graphics-content-portion: ANY-CONTENT {	121

8	Interchange format	122
8.1	Interchange format class A	122
8.1.1	Interchange format	122
8.1.2	Document application profile object identifier	122
8.1.3	Encoding of application comments	122
8.1.4	Data lengths	122
8.2	Interchange format SDIF	122
8.2.1	Interchange format	122
8.2.2	Document application profile object identifier	123
8.2.3	Encoding of application comments	123
A.1	Amendments	124
A.1.1	Amendments to the base standard	124
A.1.2	Proposed changes to standards due to defects	124
A.2	Technical corrigenda	124
A.2.1	Technical corrigenda to this International Standardized Profile	124
A.3	Versions of the ODA International Standard	125
	Recommended Practices	126
B.1	Transfer methods for ODA	126
B.1.1	Conveyance of ODA over CCITT X.400-1984	126
B.1.2	Conveyance of ODA over FTAM	126
B.1.3	Conveyance of ODA over DTAM	126
B.1.4	Conveyance of ODA over flexible disks	127
B.2	Font reference	127
B.3	ISO/IEC 8632 (CGM) constraints for this document application profile	128
B.3.1	Delimiter elements	128
B.3.2	Metafile descriptor elements	129
B.3.3	Picture descriptor elements	130
B.3.4	Control elements	130
B.3.5	Graphical primitive elements	130
B.3.6	Attribute elements	131
B.3.7	External elements	133
B.4	Interoperability with SGML applications	133
	Bibliography	134

List of figures

Figure 1	Page layout type A	17
Figure 2	Page layout type B	17
Figure 3	Page layout type C	18
Figure 4	Page layout type D.....	18
Figure 5	Example of body area layout for page layout A.....	20
Figure 6	Example of body area layout for page layout B.....	21
Figure 7	Example of body area layout for page layout C and D.....	22
Figure 8	Examples of the layout of snaking columns.....	24
Figure 9	Example of synchronized column layout for page layout A.....	25
Figure 10	Specification of margins.....	32
Figure 11	Code extension features (basic case).....	36
Figure 12	Code extension features (all possible cases).....	36
Figure 13	The body part of the generic logical structure the passage and numbered segment levels.....	53
Figure 14	The body part of the generic logical structure - the paragraph level.....	54
Figure 15	The common part of the generic logical structure.....	54
Figure 16	The layout structure - document root and page sets.....	55
Figure 17	The layout structure - the page structure.....	55
Figure 18	The layout structure - the header and footer frame structure.....	56

List of tables

Table 1 Nominal page sizes 15

Table 2 List of number string identifiers 46

Table 3 Symbols to denote shift functions.....57

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1. In addition to developing International Standards, ISO/IEC JTC 1 has created a Special Group on Functional Standardization for the elaboration of International Standardized Profiles.

An International Standardized Profile is an internationally agreed, harmonized document which identifies a standard or group of standards, together with options and parameters, necessary to accomplish a function or set of functions.

Draft International Standardized Profiles are circulated to national bodies for voting. Publication as an International Standardized Profile requires approval by at least 75 % of the national bodies casting a vote.

International Standardized Profile ISO/IEC ISP 11181-1 was prepared with the collaboration of

- OSI Asia-Oceania Workshop (AOW);
- European Workshop for Open Systems (EWOS);
- OSE Implementors' Workshop (OIW);
- CCITT Study Group VIII.

ISO/IEC ISP 11181 consists of the following part under the general title *Information technology - International Standardized Profile FOD26 - Open Document Format: Enhanced document structure*

- *Character, raster graphics and geometric graphics content architectures:*

- *Part 1: Document Application Profile (DAP)*

Further parts of this International Standardized Profile may be developed.

Annex A forms an integral part of this part of ISO/IEC ISP 11181. Annexes B and C are for information only.

Introduction

This International Standardized Profile (ISP) is defined within the context of Functional Standardization, in accordance with the principles specified by ISO/IEC TR 10000, "Framework and Taxonomy of International Standardized Profiles". The context of Functional Standardization is one part of the overall field of Information Technology (IT) standardization activities, covering base standards, profiles and registration mechanisms. A Profile defines a combination of base standards that collectively perform a specific well-defined IT Function. Profiles standardize the use of options and other variations in the base standards, and provide a basis for the development of uniform, internationally recognized conformance test suites.

One of the most important roles for an ISP is to serve as the basis for the development (by organizations other than ISO and IEC) of internationally recognized tests and test centres. ISPs are produced not simply to "legitimize" a particular choice of base standards and options, but to promote real systems interoperability. The development and widespread acceptance of tests based on this and other ISPs is crucial to the successful realization of this goal.

Development of this document application profile has been done in close cooperation between several organizations. These include ODA expert groups within the Asia-Oceania Workshop (AOW), CCITT Study Group VIII, European Workshop for Open Systems (EWOS) and OSE Implementators' Workshop (OIW). The liaison between these organizations has occurred within the meetings of the Profile Alignment Group for ODA (PAGODA). These meetings have focused on the development of a single set of Internationally aligned ODA document application profiles. This ISP is harmonised among the three workshops and has been ratified by the three workshops' plenary assemblies.

The purpose of this International Standardized Profile (ISP) is to facilitate the interworking of applications interchanging documents based on ODA, CCITT Rec. T.410 series/ISO/IEC 8613. This International Standardized Profile is suitable for interchanging documents in formatted form, processable form or formatted processable form and has been defined in accordance with CCITT Rec. T.411/ISO 8613-1. The format of this International Standardized Profile is in accordance with the standardized proforma and notation defined in Addendum 1-Annex F of CCITT Rec. T.411/ISO 8613-1.

Information technology – International Standardized Profile FOD26 – Open Document Format: Enhanced document structure – Character, raster graphics and geometric graphics content architectures –

Part 1: Document Application Profile (DAP)

1 Scope

This International Standardized Profile specifies interchange formats for the transfer of structured documents between equipment designed for word or document processing. Such documents may contain character, raster graphics and geometric graphics content.

The documents that can be interchanged using this International Standardized Profile range from simple documents to highly structured technical reports, articles and typeset documents such as brochures. This International Standardized Profile provides a comprehensive level of features for the transfer of documents between these systems.

This International Standardized Profile allows documents to be interchanged in the following forms:

- formatted form;
- processable form;
- formatted processable form.

The architecture levels defined for these three forms have matching functionalities so that the interchange formats of a document are convertible from a processable form to any other form.

This International Standardized Profile is independent of the processes carried out in an end system to create, edit or reproduce documents. It is also independent of the means to transfer documents which, for example, may be by means of communication links or storage media.

A document structured in accordance with this International Standardized Profile is represented for interchange by one of two document application profiles. One document application profile uses the Office Document Interchange Format (ODIF), the other document application profile uses the Office Document Language (ODL), both as defined in CCITT Rec. T.415/ISO 8613-5.

When this document refers to *this International Standardized Profile*, it is referring to either of the document application profiles defined by this specification.

2 Normative references

The following documents contain provisions which, through reference in this text, constitute provisions of this International Standardized Profile. At the time of publication, the editions indicated were valid. All documents are subject to revision, and parties to agreements based on this International Standardized Profile are warned against automatically applying any more recent editions of the documents listed below, since the nature of references made by ISPs to such documents is that they may be specific to a particular edition. Members of IEC and ISO maintain registers of currently valid International Standards and ISPs, and CCITT maintains published editions of its current Recommendations.

2.1 Paired Recommendations/International Standards equivalent in technical content

- CCITT Recommendation T.411 (1988), *Open Document Architecture (ODA) and Interchange Format: Introduction and General Principles*.
ISO 8613-1: 1989, *Information processing - Text and office systems - Office Document Architecture (ODA) and interchange format - Part 1: Introduction and general principles*.
- CCITT Recommendation T.411 Annex F (1991), *Open Document Architecture (ODA) and Interchange Format: Introduction and General Principles - Annex F: Document Application Profile Proforma and Notation*.
ISO 8613-1 : 1989/Add.¹⁾ : *Information processing - Text and office systems - Office Document Architecture (ODA) and interchange format - Part 1: Introduction and general principles - Addendum 1 : Document Application Profile Proforma and Notation*.
- CCITT Recommendation T.412 (1988), *Open Document Architecture (ODA) and Interchange Format: Document Structures*.
ISO 8613-2: 1989, *Information processing - Text and office systems - Office Document Architecture (ODA) and interchange format - Part 2: Document structures*.
- CCITT Recommendation T.414 (1988), *Open Document Architecture (ODA) and Interchange Format: Document Profile*.
ISO 8613-4: 1989, *Information processing - Text and office systems - Office Document Architecture (ODA) and interchange format - Part 4: Document profile*.
- CCITT Recommendation T.415 (1988), *Open Document Architecture (ODA) and Interchange Format: Open Document Interchange Format (ODIF)*.
ISO 8613-5: 1989, *Information processing - Text and office systems - Office Document Architecture (ODA) and interchange format - Part 5: Office Document Interchange Format (ODIF)*.
- CCITT Recommendation T.416 (1988), *Open Document Architecture (ODA) and Interchange Format: Character Content Architectures*.
ISO 8613-6: 1989, *Information processing - Text and office systems - Office Document Architecture (ODA) and interchange format - Part 6: Character content architectures*.
- CCITT Recommendation T.417 (1988), *Open Document Architecture (ODA) and Interchange Format: Raster Graphics Content Architectures*.
ISO 8613-7: 1989, *Information processing - Text and office systems - Office Document Architecture (ODA) and interchange format - Part 7: Raster graphics content architectures*.
- CCITT Recommendation T.418 (1988), *Open Document Architecture (ODA) and Interchange Format: Geometric Graphics Content Architectures*.
ISO 8613-8: 1989, *Information processing - Text and office systems - Office Document Architecture (ODA) and interchange format - Part 8: Geometric graphics content architectures*.

1) To be published.

NOTE - In the titles of the various parts of ISO 8613 the term *Office* occurs. ISO has instructed that this term be changed to *Open* and the new edition of ISO/IEC 8613 (to be published) will use *Open* in place of *Office*. For this reason this International Standardized Profile uses the term *Open* in its title.

- 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)*.
- CCITT Recommendation X.209 (1988), *Specification of basic encoding rules for abstract syntax notation one (ASN.1)*.
ISO/IEC 8825: 1990, *Information technology - Open Systems Interconnection - Specification of Basic Encoding Rules for Abstract Syntax Notation One (ASN.1)*.
- CCITT Recommendation T.502¹⁾, *Document Application Profile PM-11 for the Interchange of Character Content Documents in Processable and Formatted Forms*.
ISO/IEC ISP 10610-1: 1993, *Information technology - International Standardized Profile FOD11 - Open Document Format: Simple document structure - Character content architecture only - Part 1: Document Application Profile (DAP)*.
- CCITT Recommendation T.506¹⁾, *Document Application Profile PM-36/FOD36 - Open Document Format: Extended document structure - Character, raster graphics and geometric graphics content architectures*.
ISO/IEC ISP 11182-1: 1993, *Information technology - International Standardized Profile FOD36 - Open Document Format: Extended document structure - Character, raster graphics and geometric graphics content architectures - Part 1: Document Application Profile (DAP)*.

2.2 Additional references

- CCITT Recommendation T.4 (1988), *Standardization of group 3 facsimile apparatus for document transmission*;
- CCITT Recommendation T.6 (1988), *Facsimile coding schemes and coding control functions for group 4 facsimile apparatus*;
- CCITT Recommendation T.400 (1988), *Introduction to Document Architecture, Transfer and Manipulation*.
- ISO/IEC 646 : 1991, *Information technology - ISO 7-bit coded character set for information interchange*.
- ISO 2022 : 1986, *Information processing - ISO 7-bit and 8-bit coded character sets - Code extension techniques*.
- ISO 2375: 1985, *Data processing - Procedure for registration of escape sequences*.
- ISO 6937-2 : 1983, *Information processing - Coded character sets for text communication - Part 2: Latin alphabetic and non-alphabetic graphic characters*.
- ISO 6937-2 : 1983/Add.1 : 1989, *Information processing - Coded character sets for text communication - Part 2: Latin alphabetic and non-alphabetic graphic characters - Addendum 1*.
- ISO/IEC 7350 : 1991, *Information technology - Registration of repertoires of graphic characters from ISO/IEC 10367*.
- ISO/IEC 8632 : 1992, *Information technology - Computer graphics - Metafile for the storage and transfer of picture description information*
- Part 1: Functional specification
- Part 3: Binary encoding.
- ISO 8859-1 : 1987, *Information processing - 8-bit single-byte coded graphic character sets - Part 1: Latin alphabet No 1*.
- ISO 8879 : 1986, *Information processing - Text and office systems - Standard Generalized Markup Language (SGML)*.

1) To be published.

- ISO 9069 : 1988, *Information processing - SGML support facilities - SGML Document Interchange Format (SDIF)*.
- ISO 9293 : 1987, *Information processing - Volume and file structure of flexible disk cartridges for information interchange*.
- ISO/IEC 9541 : 1991, *Information technology - Font information interchange*.
- ISO/IEC TR 10000-1: 1992, *Information technology - Framework and taxonomy of International Standardized Profiles - Part 1: Framework*.
- ISO/IEC TR 10000-2: 1992, *Information technology - Framework and taxonomy of International Standardized Profiles - Part 2: Taxonomy of ISO profiles*.

3 Definitions

For the purposes of this International Standardized Profile, the following definitions apply.

The definitions given in CCITT Rec. T.411/ISO 8613-1 are applicable to this International Standardized Profile.

Constituent constraint names: Each constituent that may be included in a document that conforms to this International Standardized Profile has been given a unique name which serves to associate that constituent with a constituent constraint defined in this International Standardized Profile.

The convention is that full names are used (i.e., no abbreviations are used), two or more words in a name are concatenated and each word begins with a capital. Examples of constituent constraint names used in this International Standardized Profile are BodyText, Footnote, RectoPage and ColumnFixed.

In clause 6 each constituent constraint provided by this International Standardized Profile is italicized once at the point in the text at which the purpose of that constituent constraint is defined. This also serves to identify all the constituent constraints provided by this International Standardized Profile.

The same constituent constraint names are also used in the technical specification in clause 7 so that there is a one-to-one correspondence between the use of these names in clauses 6 and 7.

Although the constituent constraint names relate to the purpose of the constituent constraints, the semantics of constituents shall not be implied from the names that are used. Also, these names do not appear in an interchanged document but a mechanism for associating constituents in a document with constituent constraints is provided (see 6.6.4). Thus in an application using this International Standardized Profile, the constituents may be known to the user by different names.

4 Relationship with other International Standardized Profiles

This International Standardized Profile belongs to a series of hierarchically related International Standardized Profiles which include FOD11 and FOD36.

The features supported by this International Standardized Profile are a superset of the features supported by the International Standardized Profile FOD11 and thus all data streams that are conformant to FOD11 are also conformant to this International Standardized Profile, apart from the document application profile identifier.

Also the features supported by this International Standardized Profile are a subset of the features supported by the International Standardized Profile FOD36 and thus all data streams conformant to this International Standardized Profile are also conformant to FOD36, apart from the document application profile identifier.

NOTE - This International Standardized Profile is identical to the CCITT Recommendation T.505 (PM-26) (to be published), except that PM-26 only defines the use of the ODIF interchange format.

5 Conformance

In order to conform to this International Standardized Profile, a data stream representing a document shall meet the requirements specified in 5.1.

This part of the International Standardized Profile does not define implementation support requirements. These requirements are defined in other parts that make use of this International Standardized Profile.

5.1 Data stream conformance

The following requirements apply to the encoding of data streams which conform to this International Standardized Profile:

- the data stream shall be encoded either in accordance with the ASN.1 encoding rules defined in CCITT Rec. X209/ISO/IEC 8825 or the SGML encoding rules defined in ISO 8879;
- the data stream shall be structured in accordance with the interchange formats defined in clause 8;
- the document, as represented by the data stream after resolution of any external references, shall be structured in accordance with one of the documents architecture classes as defined in 6.1 and shall contain all mandatory constituents specified for that class; other constituents may be included, provided that they are permitted for that class, as specified in clause 7;
- each constituent shall contain all those attributes specified as required for that constituent in this International Standardized Profile; other attributes may be specified provided that they are permitted for that constituent;
- the attribute values specified shall be within the range of permissible values specified in this International Standardized Profile;
- the encoded document shall be constructed in accordance with the abstract document architecture defined in CCITT Rec. T.412/ISO 8613-2;
- the document shall be structured in accordance with the characteristics and constraints specified in clause 6.

5.2 Implementation conformance

This subclause states the requirements for implementations claiming conformance to this International Standardized Profile.

A conforming receiving implementation shall be capable of receiving either any data streams conforming to this International Standardized Profile structured in accordance with ODIF or any data streams conforming to this International Standardized Profile structured in accordance with ODL or both of these. Receiving usually, but not always, involves recognizing and further processing the data stream elements. The explicit meaning of receiving is determined by another part of this International Standardized Profile.

A receiving system which claims conformance to this document application profile shall be capable of handling data streams which are conformant to document application profiles that are subordinate to this document application profile within the taxonomy described in clause 4, (i.e. FOD11).

6 Characteristics supported by this document application profile

This clause describes the characteristics of documents which may be represented by data streams conforming to this International Standardized Profile. This clause also describes how these characteristics are represented in terms of constituent constraints.

6.1 Overview

6.1.1 General

This International Standardized Profile supports the interchange of documents in the following forms:

- processable form, which facilitates the revision of a document by a recipient;

- formatted form, which facilitates the reproduction of a document as intended by the originator;
- formatted processable form, which facilitates the reproduction of a document as intended by the originator or facilitates the revision of a document by a recipient.

In addition this International Standardized Profile supports the interchange of:

- generic-documents;
- document profile.

The constituents that make up these five classes of data stream are defined in 6.1.2 to 6.1.6. Constituents defined as *required* shall occur in any data stream that conforms to this International Standardized Profile. Constituents listed as *optional* may or may not be present in the data stream depending on the requirements of the particular data stream.

The constituents that make up a complete document that is conformant to this International Standardized Profile include all those referenced and contained in resource-documents and external-documents, if any (see 6.6.1 and 6.6.2).

6.1.2 Formatted form documents

Required constituents:

- a document profile;
- layout object descriptions representing a specific layout structure.

Optional constituents:

- layout object class descriptions representing a factor generic layout structure;
- presentation styles;
- content portion descriptions.

6.1.3 Processable form documents

Required constituents:

- a document profile;
- logical object class descriptions representing a complete or partial generic logical structure;
- logical object descriptions representing a specific logical structure.

Optional constituents:

- layout object class descriptions representing a complete generic layout structure;
- layout styles;
- presentation styles;
- content portion descriptions.

In the case of processable form documents, when the generic layout structure is not present, additional restrictions are placed on the layout directives that may be included in layout styles. These restrictions are defined in 6.4.3.

6.1.4 Formatted processable form documents

Required constituents:

- a document profile;
- logical object class descriptions representing a complete or partial generic logical structure;
- logical object descriptions representing a specific logical structure;
- layout object class descriptions representing a complete generic layout structure;
- layout object descriptions representing a specific layout structure.

Optional constituents:

- layout styles;
- presentation styles;
- content portion descriptions.

6.1.5 Generic-documents

A generic-document consists of one of the following sets of constituents:

- a)
 - a document profile;
 - logical object class descriptions which represent a complete or partial generic logical structure;
 - layout styles whose presence is optional;
 - presentation styles whose presence is optional;
 - generic content portions whose presence is optional.
- b)
 - a document profile;
 - layout object class descriptions which represent a complete generic layout structure or factor set;
 - presentation styles whose presence is optional;
 - generic content portions whose presence is optional.
- c)
 - a document profile;
 - logical object class descriptions which represent a complete or partial generic logical structure;
 - layout object class descriptions which represent a complete generic layout structure;
 - layout styles whose presence is optional;
 - presentation styles whose presence is optional;
 - generic content portions whose presence is optional.

6.1.6 Document profile

This form of document contains a document profile only.

6.2 Logical characteristics

6.2.1 Introduction

This subclause defines the constituent constraints for logical components which are provided by this International Standardized Profile to represent the characteristics of documents containing logical component descriptions.

Different constituent constraints may be used to represent and distinguish parts of a document that have different logical characteristics. This subclause describes the general characteristics and typical uses of the constituent constraints that are provided.

The descriptions of the logical characteristics represented by each of the constituent constraints is provided for guidance only. It is the responsibility of the user to determine how a document is to be represented using the constituents provided. Adherence to these guidelines can enhance the mutual understanding of a document by an originator and a recipient.

6.2.2 Overview of the logical structure

From the logical point of view, the document consists of two parts, namely a body part and a common part.

The body part represents the main content of a document and is intended to be reproduced in the body area of the pages that make up the document.

The common part represents common content that is to be placed in reserved header and footer areas on each page of a document. Header and footer content are independently optional and so may be included in an interchanged document only if required.

6.2.3 Body part of the logical structure

6.2.3.1 DocumentLogicalRoot

DocumentLogicalRoot is a constituent constraint representing the top level in the document logical structure. Its immediate subordinates consist of a sequence of one or more constituent constraints of the type *Passage*.

The automatic numbering schemes that apply to constituent constraints of the types *NumberedSegment* and *Footnote* may be initialised on the *DocumentLogicalRoot*.

6.2.3.2 Passage

Passage is a constituent constraint that represents the first level of logical subdivision of a document. It may be used to indicate a logical grouping of subordinate parts of a document that are to be regarded as an entity for reading or that have common layout and presentation characteristics.

Passages are typically used to represent:

- the contents to be placed on the title page of a report;
- the front matter in the table of contents or foreword;
- the main matter of the document;
- the back matter, consisting of appendices, glossary or index.

The automatic numbering schemes that apply to subordinate constituent constraints of the types *NumberedSegment* and *Footnote* may be initialised on a *Passage*.

The immediate subordinates of a *Passage* consist of an optional arbitrary ordered sequence of one or more of the following constituent constraint types:

- *Paragraph*;
- *BodyGeometric*;
- *BodyRaster*;
- *BodyText*.

These may be optionally followed by one or more constituent constraints of the type *NumberedSegment*.

A *Passage* shall have at least one of the above constituent constraint types as a subordinate.

A document may contain several different class definitions of the type *Passage*, each of which defines the common characteristics of sets of *Passages* within the document such as their allowed subordinates or layout properties. For example, a class of *Passages* may be defined which always begin on a new page set.

6.2.3.3 NumberedSegment

NumberedSegment is a constituent constraint that represents a logical subdivision within a document that is distinguished by an identifier. The logical subdivision may be a subdivision of a Passage or another higher level *NumberedSegment*. The logical subdivision may also have some common layout characteristics.

The automatic numbering schemes that apply to subordinate constituent constraints of the types *NumberedSegment* and *Footnote* may be initialised on a *Passage*.

The immediate subordinates of a *NumberedSegment* consist of the constituent constraint *Number*, whose presence is mandatory and serves to carry the identifier of the *NumberedSegment*. This is followed by an optional arbitrary ordered sequence of one or more of the following constituent constraints:

- *Paragraph*;
- *BodyGeometric*;
- *BodyRaster*;
- *BodyText*.

These are optionally followed by a sequence of one or more constituent constraints of the type *NumberedSegment*. Hence a document may contain any number of nested levels of the *NumberedSegment* constituents.

A *NumberedSegment* is typically used to represent entities such as chapters, sections, nested sub-sections and appendices which contain an identifier that serves to distinguish that entity for human comprehension.

A document may contain any number of different class definitions of *NumberedSegment* which define the common characteristics of sets of *NumberedSegments*, such as their allowed subordinates and layout properties.

Class definitions of *NumberedSegments* may be recursively defined. In this case only one call of *NumberedSegment* may be specified, and the <simple-expr> construction in the macro *USENUMBERSTRINGS* in the bindings attribute of this class shall use the optional *ORD* construction only. If the recursive class definitions are used for *NumberedSegment*, the following constraints shall also apply. For all levels which reference recursively defined classes:

- numbering format shall be the same;
- no initial value other than 1 or re-initialisation of the numbering is possible;
- it is not possible to continue the numbering across *Passages*.

If class definitions are not recursive in this way, there shall be one and only one class definition for *NumberedSegments* corresponding to each level of numbering within each *Passage*. Class definitions may be shared between *NumberedSegments* belonging to different *Passages* but they shall then be used at the same level.

6.2.3.4 Number

Number is a constituent constraint that represents the identifier of a *NumberedSegment* to which it is subordinate. This identifier allows the *NumberedSegment* to be distinguished within the document for machine processing or human comprehension.

A *Number* defines a basic logical constituent which contains a content generator which, when evaluated, produces the identifier referred to above. This evaluation takes place during the layout process.

The identifiers are structured and consist of a sequence of one or more numerals that allow *NumberedSegments* at the same or different levels in a document structure to be uniquely distinguished. The numerals may be represented by Arabic or Roman numerals or by their alphabetic equivalent in lower or upper case characters (the number 1 is represented by "A", etc.). Each numeral in an identifier is distinguished by means of separator characters such as spaces and full stops (the character "period"); a typical example is "6.2.3.4".

NOTE - The separator may be an empty string.

Further details of the structure and generation of the identifiers are given in 6.6.6.2.

6.2.3.5 Paragraph

Paragraph is a constituent constraint that is a subdivision of a *Passage* or *NumberedSegment*. It is typically used to represent the grouping of parts of a document that deals with a single theme or topic. These parts may consist of character, raster graphics and geometric graphics content.

The immediate subordinates of a Paragraph consist of an arbitrary ordered sequence of one or more of the following constituent constraints:

- BodyText;
- BodyRaster;
- BodyGeometric;
- Footnote.

Content from any subordinate basic text objects within a paragraph may be run on one from another (that is, to continue on the same line) by use of the Concatenation feature (see 6.4.2.5). Alternatively, content from subordinates of a paragraph may be separated one from another to give white space between them, using Separation (see 6.4.2.2). This may be used to give an effect similar to that achieved with empty lines of text. Use of empty text lines to achieve white space between areas of text or other content may lead to unintended blank areas adjacent to the leading edge of layout objects (e.g., at page breaks) whereas the use of Separation avoids this.

Constituents of the type BodyText may be concatenated to form a continuous stream of character content which is laid out as a single unit. Sequences of constituents of the types BodyText and Footnote may be concatenated to represent a stream of character content with embedded footnotes. Multiple embedded footnotes, which may be consecutive without intervening text, may be included in the content.

Another typical use of a Paragraph is to represent a group of document parts that have common layout characteristics. An example is a graphical illustration with associated text which is to be laid out in a particular frame.

6.2.3.6 BodyText, BodyRaster and BodyGeometric

BodyText, *BodyRaster* and *BodyGeometric* are constituent constraints which represent the lowest level of logical subdivision of a document. These constituent constraints are subdivisions of Passages, NumberedSegments and Paragraphs. They allow the layout and presentation requirements of different parts of a document to be specified.

These are basic logical constituents that directly refer to content portions that contain character, raster graphics and geometric graphics content respectively. BodyText in the specific logical structure shall refer to one or more content portions each containing processable, formatted or formatted processable character content. BodyRaster and BodyGeometric shall only refer to a single content portion containing formatted processable raster graphics content or formatted processable geometric graphics content respectively.

Constituents of these types in the generic logical structure may refer to generic content. This provides the means of defining common content within the body part of a document.

6.2.3.7 Footnote

Footnote is a constituent constraint that is a subdivision of a Paragraph and is used to represent footnotes within a document.

A footnote is an amount of content that is logically associated with a particular part of the document body but which is intended to be read and laid out separately from its associated part of the document. Typically, a footnote consists of a footnote identifier, which is embedded within the document body, and the footnote itself, which is laid out elsewhere.

A Footnote defines a composite logical constituent whose immediate subordinates consist of the constituent constraint FootnoteReference, which represents the footnote identifier, followed by the constituent constraint FootnoteBody, which represents the footnote itself. Both of these subordinates are mandatory.

6.2.3.8 FootnoteReference

FootnoteReference defines a constituent constraint that is used to represent a footnote reference within the body of a document.

FootnoteReference defines a basic logical constituent that contains a content generator which when evaluated produces a character string which constitutes the footnote reference referred to above.

The generated character string consists of a label with optional prefix and suffix character strings. The label is used to uniquely identify a particular footnote and may consist of a number which is represented in the form of Arabic or Roman numerals or by an alphabetic equivalent. The number may be automatically generated so that its value is incremented for each successive footnote. Alternatively, the label may consist of a user defined character string.

In a sequence of footnotes, automatic and user defined labels may be freely mixed (giving for example the sequence 1,2*,3,4). If the label consists of a user-defined character string, the automatically generated number sequence is not incremented.

An example of a footnote reference is '(2)' where '(' and ')' are user defined prefix and suffix strings respectively and '2' is the automatically generated label. Another example is 'note⁵' where '5' is the label and 'note' is a prefix string which also contains the control function PLU to enable the label to be represented in the form of a superscript. In this case, a suffix string containing the control function PLD would be required to cause the superscripting to be cancelled before the following text.

The format of the content generator referred to above is described in 6.6.6.3.

6.2.3.9 FootnoteBody

FootnoteBody is a constituent constraint which represents the content of a footnote.

FootnoteBody defines a composite logical constituent whose subordinates consist of a *FootnoteNumber* constituent, which is mandatory and represents the footnote identifier, followed by a sequence of one or more *FootnoteText* constituents which represents the footnote content. The identifier referred to above is identical to the corresponding footnote identifier which is embedded in the content of the document body and represented by the constituent constraint *FootnoteReference*.

The constituents subordinate to *FootnoteBody* are intended to be laid out separately from the other parts of the document content. When a generic layout structure is specified for the document, these constituents are constrained to be laid out in a *FootnoteArea* frame (see 6.3.5.9).

6.2.3.10 FootnoteNumber

FootnoteNumber is a constituent constraint that represents the footnote identifier within the footnote body.

This identifier is identical to the content associated with the constituent constraint *FootnoteReference* but is intended to be laid out so that it immediately precedes the content of the footnote body.

FootnoteNumber defines a basic logical constituent that contains a content generator which when evaluated produces the identifier referenced above. The format of this content generator is the same as the content generator that may be specified for the constituent constraint *FootnoteReference*.

It is required to specify the attribute "layout category" with the value 'Footnote' for this constituent. This along with the attribute "permitted categories" with the value 'Footnote' on the footnote frame will ensure that this constituent is laid out in a *FootnoteArea* frame when generic layout structure is specified within the document.

6.2.3.11 FootnoteText

FootnoteText is a constituent constraint that is used to represent the footnote content. It is the lowest logical subdivision of a *FootnoteBody*.

FootnoteText defines a basic logical constituent that references one or more content portions each containing processable, formatted or formatted processable character content.

It is required to specify the attribute "layout category" with the value 'Footnote' for this constituent. This along with the attribute "permitted categories" with the value 'Footnote' on the footnote frame will ensure that this constituent is laid out in a *FootnoteArea* frame when generic layout structure is specified within the document.

6.2.4 Common content part of the logical structure

6.2.4.1 CommonContent

CommonContent is a constituent constraint that represents common content that is to be laid out in the header and footer areas of the pages of a document. Common content consists of any combination of character, raster graphics and geometric graphics content.

Any number of constituent constraints of the type *CommonContent* may be contained in a document. *CommonContent* is a composite logical object class whose immediate subordinates consist of an arbitrary ordered sequence of one or more of the following constituent constraints:

- *CommonText*;
- *PageNumber*;
- *CommonRaster*;
- *CommonGeometric*.

When the generic layout structure is present, constituents of the type *CommonContent* and their associated subordinate constituents are constrained to be laid out in frames representing header or footer areas using the logical source mechanism (see 6.3.6).

6.2.4.2 CommonText

CommonText is a constituent constraint that represents the common character content that is to be laid out in the header or footer area of a document. For example, header or footer content that appears on each page in a sequence of pages may be represented by this constituent constraint.

CommonText is a constituent constraint for a basic logical object class that references one or more content portions each containing character content in a processable, formatted or formatted processable form.

6.2.4.3 PageNumber

PageNumber is a constituent constraint that represents the common character content that is to be laid out in the header or footer area of a document. This constituent constraint is specifically used when it is required to present a header or footer content which contains an automatically generated page number.

PageNumber is a basic logical object class that contains a content generator. This content generator contains a reference to a page number which is automatically evaluated when the document is laid out. This provides the means of representing the page numbers that are displayed on the consecutive pages of a document.

Each page number consists of a single number which may be represented in the form of Arabic or Roman numerals or in its alphabetic equivalent. Page numbering schemes may start at 0 or any value greater than 0, at the document root or pageset levels.

The format of the content generators is defined in 6.6.6.1.

6.2.4.4 CommonRaster

CommonRaster is a constituent constraint that represents the common raster graphics content that is to be laid out in the header or footer area of a document. For example, this constraint may be used to represent a logo which is to be laid out on each page of a document.

CommonRaster is a basic logical object class which references a single content portion containing formatted processable raster graphics content.

6.2.4.5 CommonGeometric

CommonGeometric is a constituent constraint that represents the common geometric graphics content that is to be laid out in the header or footer area of a document. For example, this constraint may be used to represent a graphical icon which is to be laid out on each page of a document.

CommonGeometric is a constituent constraint for a basic logical object class which references a single content portion containing formatted processable geometric graphics content.

6.3 Layout characteristics

This subclause defines the constituent constraints for layout components which are provided by this International Standardized Profile to represent the characteristics of documents.

Different constituent constraints may be used to represent and distinguish parts of a document that have different layout characteristics. This subclause describes the general characteristics and typical uses of the constituent constraints that are provided.

The descriptions of the layout characteristics represented by each of the constituent constraints is provided for guidance only. It is the responsibility of the user to determine how a document is to be represented using the constituents provided. Adherence to these guidelines can enhance the mutual understanding of a document by an originator and a recipient.

6.3.1 Overview of the layout characteristics

The document structure allows the document content to be laid out and presented in one or more page sets. Each page set may be used for different parts of the document, for example, the title page, foreword, table of contents, document body and appendices.

Each page set consists of a series of pages. In general, each page may be sub-divided into three areas: the body area, which is used to layout the document body; and the header and footer areas, which may be used to layout the common content.

Four page layout types are provided by this International Standardized Profile. Each page layout type specifies how the body, header and footer areas are positioned within each page and how the content may be presented within each of those areas. These four types are referred to as page layouts A, B, C and D and are illustrated in Figures 1, 2, 3 and 4 respectively.

It is intended that all applications which use this International Standardized Profile shall support page layout A, whereas support for the other three page layouts may be specified as optional.

Page layout A is used when the character content is to be laid out horizontally (from left to right or from right to left) and from top to bottom within the body area. This layout is typically used for documents written in Latin based, Hebrew and Arabic languages.

Page layout B is used when the character content is to be laid out vertically (bottom to top or top to bottom) and from left to right within the body area. This layout is typically used for documents written in Latin based, Hebrew and Arabic languages in which it is required to layout the content in landscape orientation within the body area of the page.

Page layouts C and D are used when the character content is to be laid out vertically and from right to left within the body area. These layouts are typically used in documents written in languages which use ideograms, such as Japanese and Chinese characters.

The body area may be further sub-divided into areas composed of single and multiple columns and an area may be reserved for footnotes. In addition, the header and footer areas may be sub-divided to allow the representation of different content types.

6.3.2 DocumentLayoutRoot

DocumentLayoutRoot is a constituent constraint that represents the top level in the document layout structure. Its immediate subordinates consist of a sequence of one or more constituents of the type *PageSet*. The numbering schemes for pages may be initialised on this constituent constraint.

6.3.3 PageSet

PageSet is a constituent constraint that represents a grouping of pages within a document. A *PageSet* is typically used to represent a part of a document that has different layout requirements from other parts of a document. Also, a *PageSet* may correspond to a part of a document that has a certain logical significance, for example, a *PageSet* might represent the front matter in a document or an individual chapter.

Only one level of *PageSet* is allowed in a document. However, a document may contain any number of class definitions of the type *PageSet* which may be used, for example, to provide a choice of alternative layouts for different parts of a document or to specify the exact layout requirements for each successive part of a document.

The immediate subordinates of a *PageSet* consist of a combination of constituent constraints of the types *Page*, *RectoPage* and *VersoPage*, as described in 6.3.4.1.

6.3.4 Page characteristics

6.3.4.1 Page constituents

Three constituent constraints are provided to represent the pages within a document, namely *Page*, *RectoPage* and *VersoPage*.

The only difference in the characteristics of these types of constituent constraints concerns the values that may be specified for the parameter "side of sheet" in the attribute "medium type". In the case of *Page*, this parameter may be specified with the value 'recto', 'verso' or 'unspecified'. In the case of *RectoPage*, this parameter may be specified with the value 'recto' or 'unspecified'; in the case of *VersoPage*, this parameter may be specified with the value 'verso' or 'unspecified'. The values 'recto' and 'verso' of the parameter "side of sheet" of the attribute "medium type" are non-basic.

The pages that make up a page set consist of an optional initial page which is represented by the constituent constraint *Page* and which is optionally followed by either:

- A sequence of pages represented by the constituent constraint *Page*. All pages in this sequence shall have the same layout characteristics but these characteristics may differ from those of the initial page.
- A sequence of pages which are intended to be laid out alternatively on the recto and verso (or on the verso and recto) sides of the presentation medium and are represented by the constituent constraints *RectoPage* and *VersoPage* respectively. All pages in this sequence shall have the same layout characteristics but these characteristics may differ from those of the initial page.

Pages having the same layout characteristics are pages that have the same page layout (see 6.3.4.5) and for which the body area, header area (if present) and footer area (if present) have the same dimensions and positions within the page (see 6.3.4.3). Pages having the same layout characteristics do not necessarily have the same position on the presentation medium (see 6.3.4.4).

A page set shall contain at least one page.

An initial page is typically used at the beginning of a document or of a section within a document. It may be used, for example, for a title page whose layout requirements differ from the following pages.

The following restrictions also apply to the pages within a page set:

- all the pages shall have the same dimensions and orientation (see 6.3.4.2);
- all pages shall be laid out on the same size of presentation medium (see 6.3.4.3).

6.3.4.2 Page dimensions

The dimensions of the pages may be specified as any value (in BMUs) that is equivalent to or less than ISO A3 or ANSI B paper sizes in portrait or landscape orientation. The dimensions may be specified in portrait or landscape orientation. Japanese page sizes B4 and B5 are also supported but the dimensions of these pages lie within the range of dimensions given above.

Dimensions equivalent to or less than the common assured reproduction area of ISO A4 and ANSI A in portrait or landscape orientation are basic values. Larger page sizes are non-basic and their use shall be indicated in the document profile.

Any default page dimensions may be specified in the document profile subject to the maximum dimensions defined above.

NOTE - The size termed *North American Letter (NAL)* in CCITT Rec. T.410 series ISO/IEC 8613 clause 7 is called *ANSI A* in this International Standardized Profile to be consistent with the other references to ANSI standard paper sizes.

6.3.4.3 Nominal page sizes

The nominal page sizes that may be specified are listed in Table 1. These may be specified in portrait or landscape orientation. All values of nominal page size are non-basic and hence all values used in a document shall be indicated in the document profile.

Any value of nominal page size defined in Table 1, subject to the restrictions specified above, may be specified as the default value in the document profile.

Table 1 also includes the recommended assured reproduction area (ARA). Information loss may occur when a document is reproduced if the dimensions of constituent constraints of the type page exceed the ARA for the specified nominal page size.

Table 1 - Nominal page sizes

Page type	Size in inches or millimetres	Size in BMUs	ARA in BMUs
ISO A5	148mm x 210mm	7 015 x 9 920	not defined
ISO A4	210mm x 297mm	9 920 x 14 030	9 240 x 13 200
ISO A3	297mm x 420mm	14 030 x 19 840	13 200 x 18 480
ANSI legal	8,5 in. x 14 in.	10 200 x 16 800	9 240 x 15 480
ANSI A	8,5 in. x 11 in.	10 200 x 13 200	9 240 x 12 400
ANSI B	11 in. x 17 in.	13 200 x 20 400	12 744 x 19 656
Japan-legal	257mm x 364mm	12 141 x 17 196	11 200 x 15 300
Japan-letter	182mm x 257mm	8 598 x 12 141	7 600 x 10 200

6.3.4.4 Page offset

The page offset is the distance of the position of the left and top edges of the page relative to the left and top edges respectively of the presentation medium on which each page is reproduced. Any value of page offset may be specified provided that no part of the page area lies outside the area of the nominal page. Also, page offsets specified for the initial, recto and verso pages within a given page set may differ. The default page offset may be specified in the document profile.

6.3.4.5 Page layout characteristics

6.3.4.5.1 General characteristics

Each page in a document may be subdivided into three rectangular areas, as follows:

- a body area which is reserved for content that belongs to the body part of the document (see 6.3.5);
- a header area which is reserved for common header content (see 6.3.6);
- a footer area which is reserved for common footer content (see 6.3.6).

The body area is mandatory and shall occur on every page in a document. The header and footer areas are both optional.

Also, these three areas shall be entirely contained within the page area and shall not overlap.

Four types of page layout are defined:

6.3.4.5.2 Page layout A

For page layout A the header and footer areas are placed above and below the body area respectively. The layout paths in the header, body and footer areas are specified as 270°. This type of layout is illustrated in Figure 1.

6.3.4.5.3 Page layout B

For page layout B the header and footer areas are placed above and below the body area respectively. The layout path in the body area is specified as 0°; in the header and footer areas the layout paths are specified as 270°. This type of layout is illustrated in Figure 2.

6.3.4.5.4 Page layout C

For page layout C the header and footer areas are placed above and below the body area respectively. The layout path in the body area is specified as 180°; in the header and footer areas, the layout paths are specified as 270°. This type of layout is illustrated in Figure 3.

6.3.4.5.5 Page layout D

For page layout D the header and footer areas are placed to the right and left of the body area respectively. The layout paths in the header, body and footer areas are all specified as 180°. This type of layout is illustrated in Figure 4.

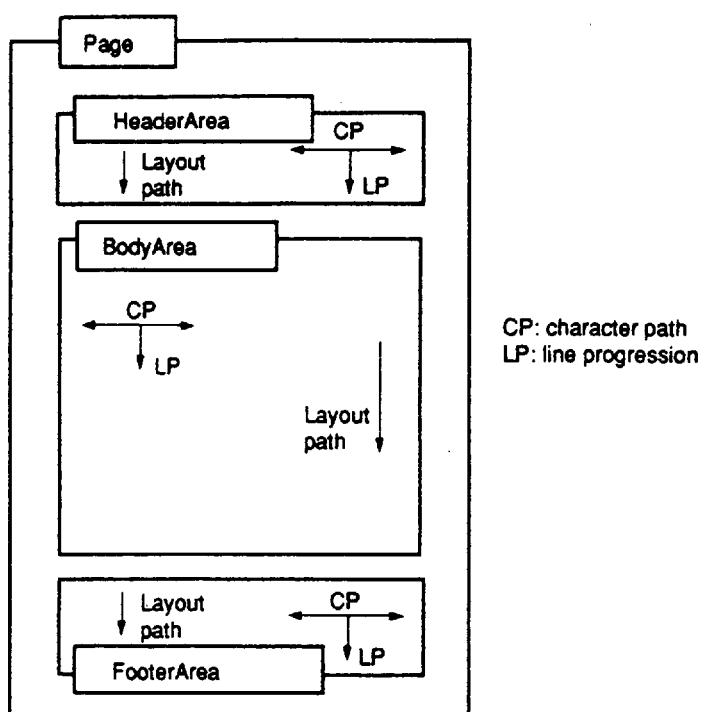


Figure 1 - Page layout type A

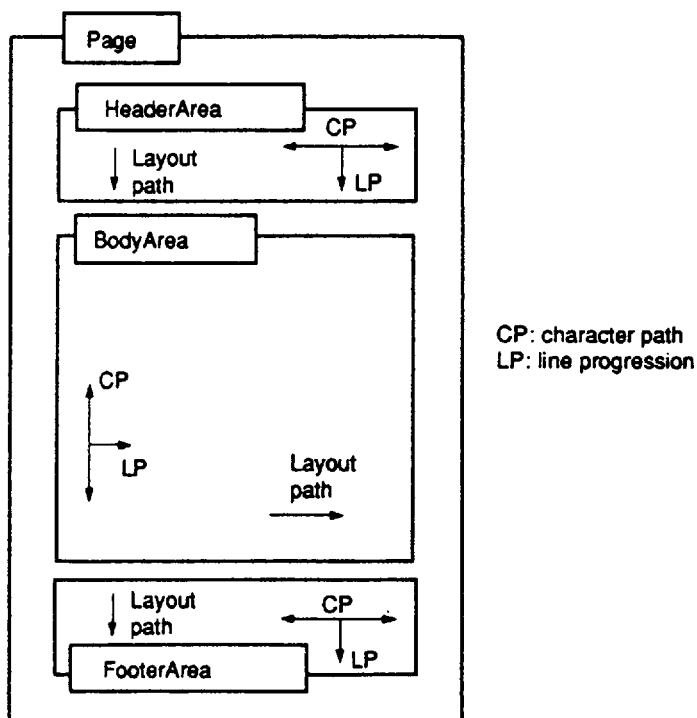


Figure 2 - Page layout type B

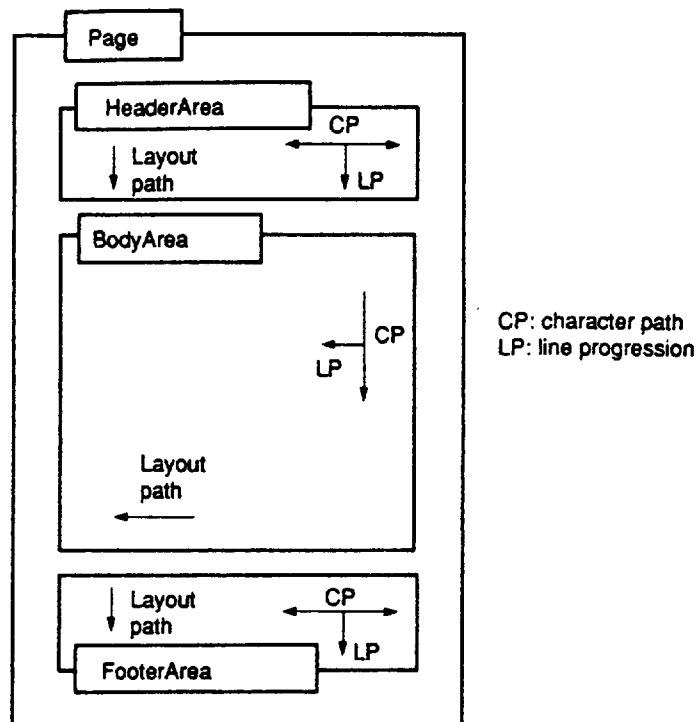


Figure 3 - Page layout type C

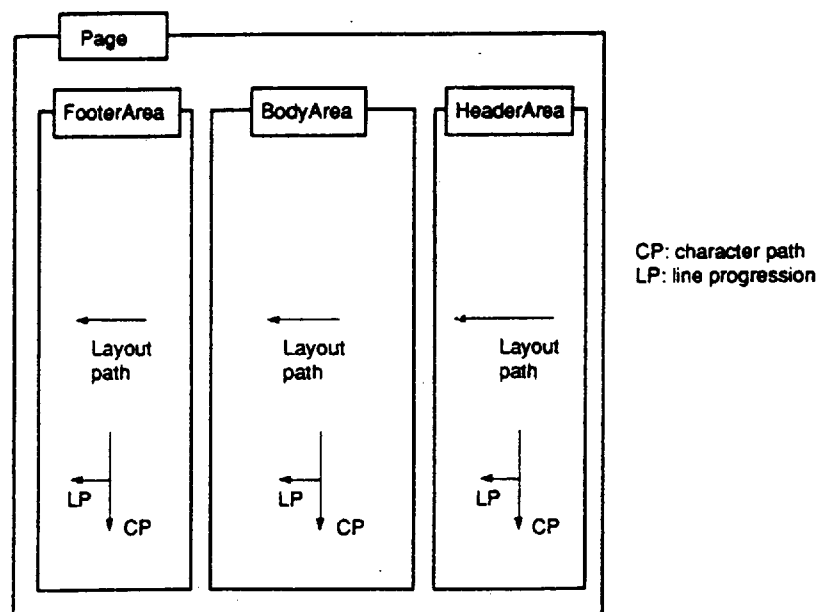


Figure 4 - Page layout type D

6.3.5 Body area characteristics

6.3.5.1 General characteristics

The body area is the area within a page where the main matter of the document, that is the *body* part of the document, is laid out.

The body area may consist of a single frame into which the content is directly laid out. In this case, the body area is represented by a BasicBody frame.

Alternatively, the body area may be subdivided into different rectangular areas to provide for different layout requirements. In this case, the body frame is represented by a VariableCompositeBody frame.

6.3.5.2 BasicBody

BasicBody is a constituent constraint which defines a lowest level frame into which content is directly laid out.

The position and dimensions of this frame are fixed. The layout path specified depends upon the page layout type being used (see 6.3.4.5).

6.3.5.3 VariableCompositeBody

VariableCompositeBody is a constituent constraint that defines a composite frame which contains one or more subordinate variably positioned frames. These subordinates are not pre-determined and are automatically adjusted during the layout process to accommodate the content that is allocated to them. For example, to provide for combinations of single or multiple column layout and the layout of footnotes. A VariableCompositeBody frame has a fixed position and fixed dimensions. The layout path specified for this frame depends upon the page layout type being used (see 6.3.4.5).

The immediate subordinates of VariableCompositeBody frames consist of an arbitrary ordered sequence of one or more frames of the following constituent constraints:

- BasicFloat;
- SnakingColumns;
- SynchronizedColumns

It may also contain a single frame of the type FootnoteArea.

The subordinate frames are all variably positioned and have variable dimensions.

Thus the relative positions of these frames in the body area may vary and depend upon the positions of other frames (if any) that are placed in the same VariableCompositeBody frame.

The layout path for VariableCompositeBody frames may be specified as 270°, 0° or 180°. This determines the page layout type used in the case where VariableCompositeBody represents the entire body area.

Frames of the type BasicFloat, SnakingColumns and SynchronizedColumns are laid out along the layout path specified (in normal positioning fill order). FootnoteArea frames are laid out in the same direction as the body area layout path, but reverse fill order is used.

These frames are constrained to have the same layout path as the VariableCompositeBody frame to which they are subordinate. An exception to this rule is made for frames of the type SnakingColumns (see subclause below).

Figures 5, 6 and 7 provide illustrations of the layout of frames within a VariableCompositeBody frame for the various page layout types.

A choice of subordinate frames of the types listed above may be specified for a VariableCompositeBody frame. Different frame types may be selected using various layout directives (see 6.4) and hence the layout characteristics of the body areas within a page set may change from page to page within a page set.

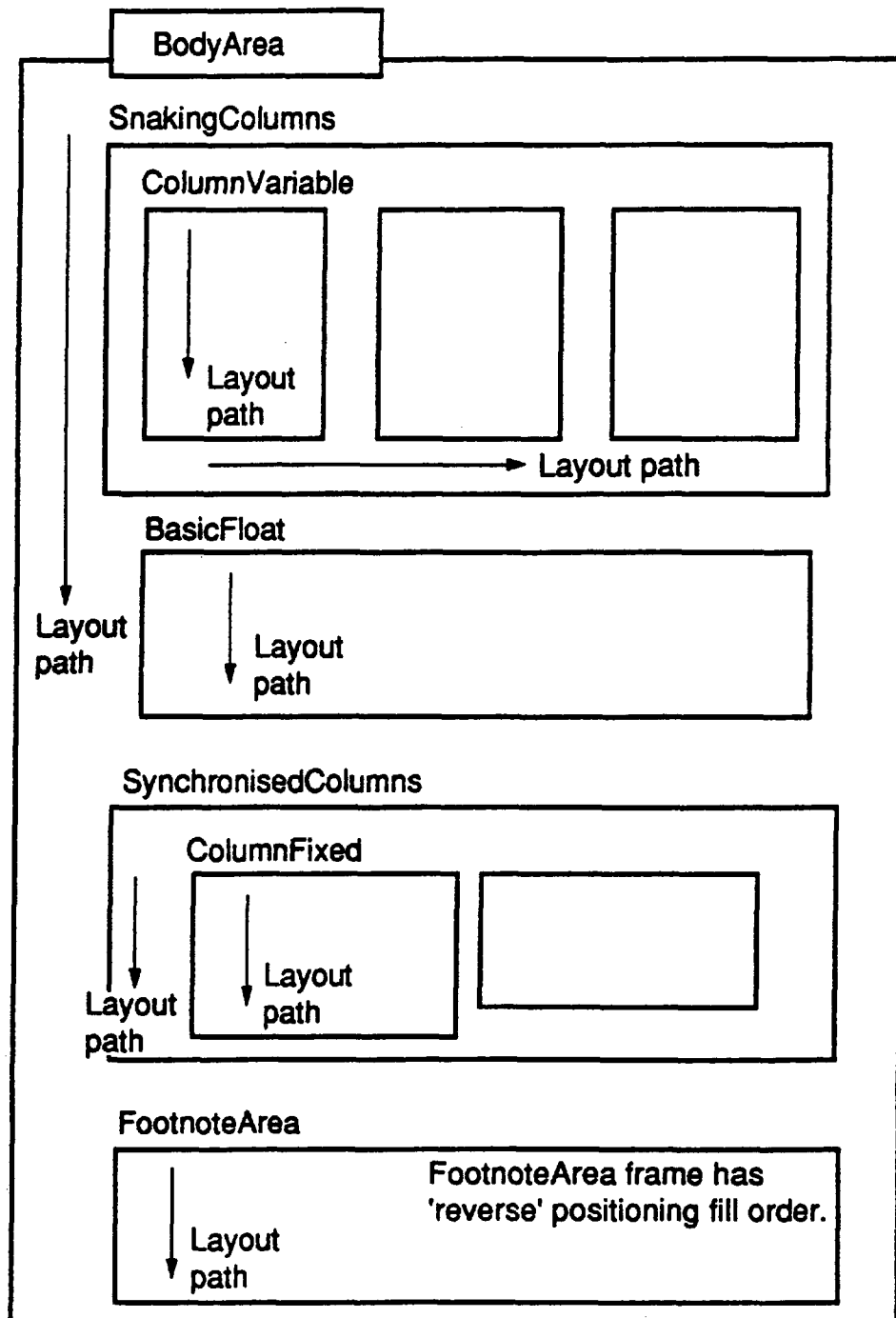


Figure 5 - Example of body area layout for page layout A

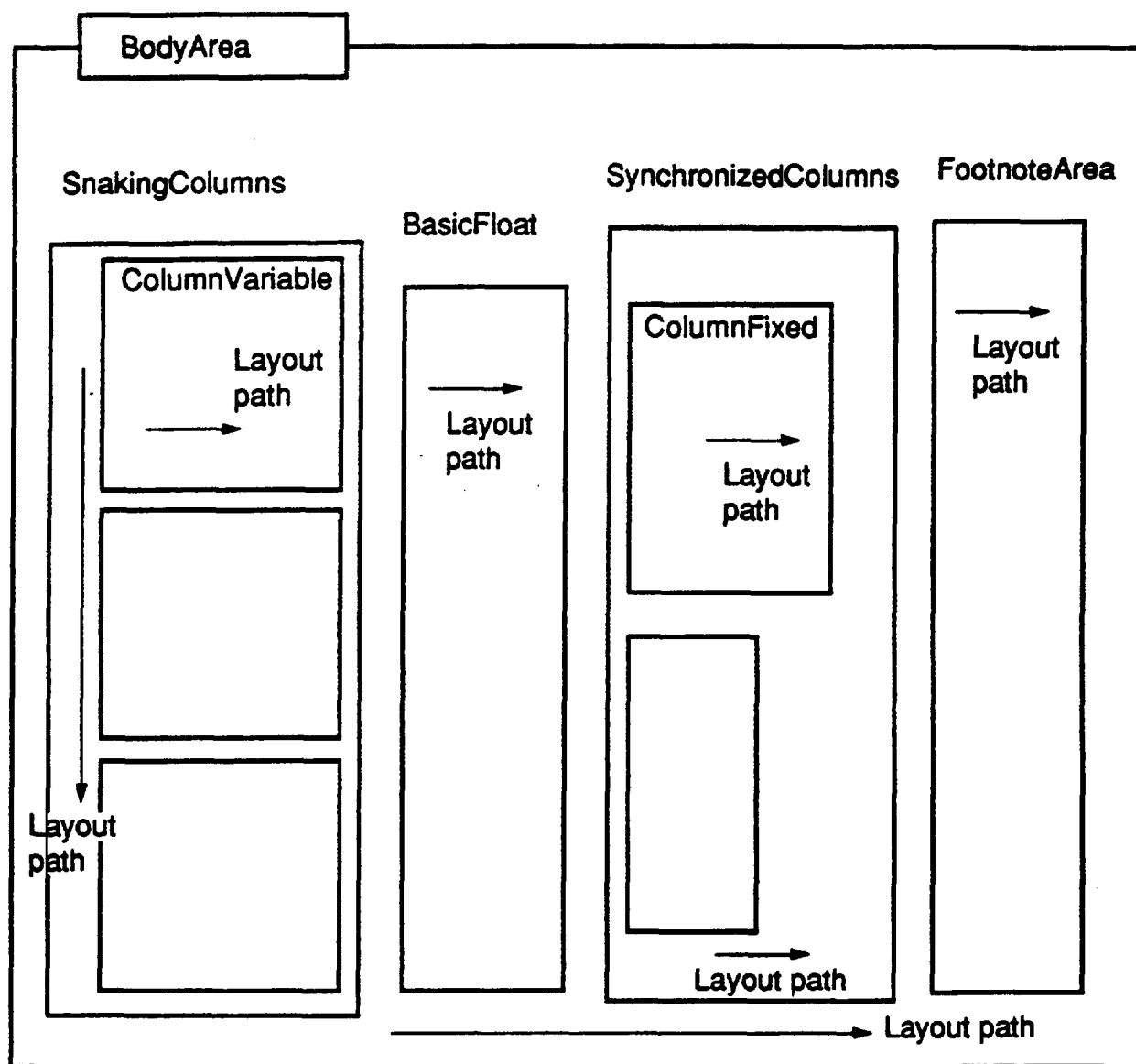


Figure 6 - Example of body area layout for page layout B

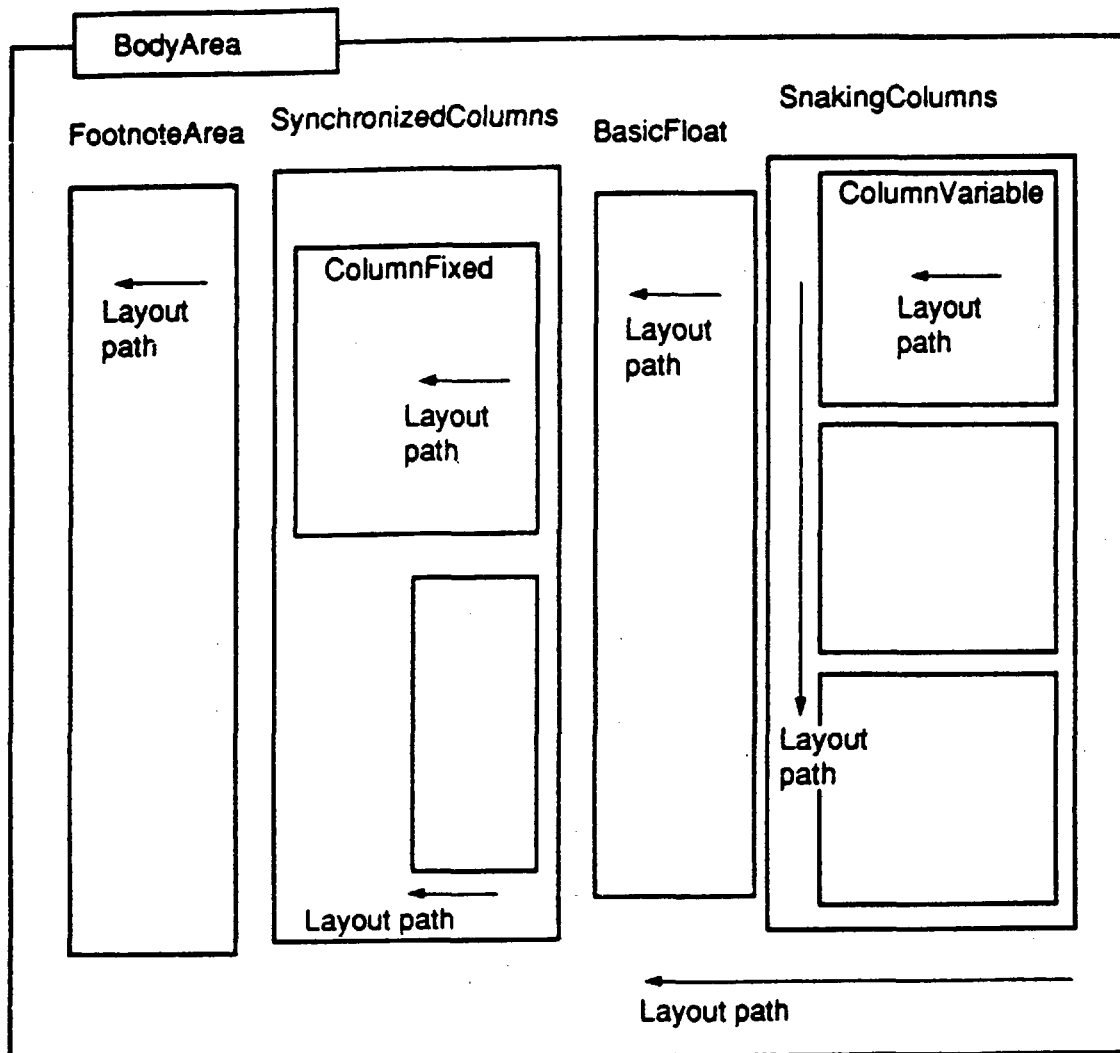


Figure 7 - Example of body area layout for page layouts C and D

6.3.5.4 BasicFloat

BasicFloat is a constituent constraint that defines a lowest level frame that is used to represent a single column area within a body area. A single column area is typically used to layout content in the form of a single column. This is a variably positioned frame.

The dimension of this frame in the direction orthogonal to the layout path of the body area is fixed or defaults to the maximum value allowed within the body area.

The dimension of this frame in the direction parallel to the layout path of the body area is specified by sub-parameter "Rule-B". This dimension is therefore automatically adjusted during the layout process to be the minimum required to contain all the content allocated to the frame.

The layout path specified for this frame is the same as that specified for the body area. Content shall only be laid out in this frame in the direction of the layout path specified.

6.3.5.5 SnakingColumns

SnakingColumns is a constituent constraint that defines a composite frame that represents a snaking columns area within a body area. A snaking columns area is typically used for the layout of one or more columns of content in which the content is allowed to flow freely from one column to the next.

This frame is variably positioned. Its immediate subordinates consist of one or more frames of the type *ColumnVariable*. Examples of the layout of *SnakingColumns* frames are given in Figure 8.

The dimension of a *SnakingColumns* frame in the direction orthogonal to the layout path of the body area is fixed or defaults to the maximum value allowed within the body area.

The dimension of this frame in the direction parallel to the layout path of the body area is specified by sub-parameter "Rule-B". This dimension is therefore automatically adjusted to accommodate the subordinate frames which are laid out in it.

The layout path for a *SnakingColumns* frame may be specified as 0° or 180° in the case of page layout A, 90° or 270° in the case of page layout B, and 270° in the cases of page layouts C and D.

The attribute "balance" may be specified for a *SnakingColumns* frame to indicate that two or more of the subordinate *ColumnVariable* frames are to be approximately equal in length in the vertical dimension in the case of page layout A and are to be approximately equal in length in the horizontal dimension in the cases of page layouts B, C and D (see note). Note that the term *approximately equal* in the context of the attribute "balance" means that the leading edges of the layout objects being balanced are aligned, as closely as possible, to a line orthogonal to the layout path for the objects.

The attribute "balance" may be ignored when the subordinate *ColumnVariable* frames have unequal widths.

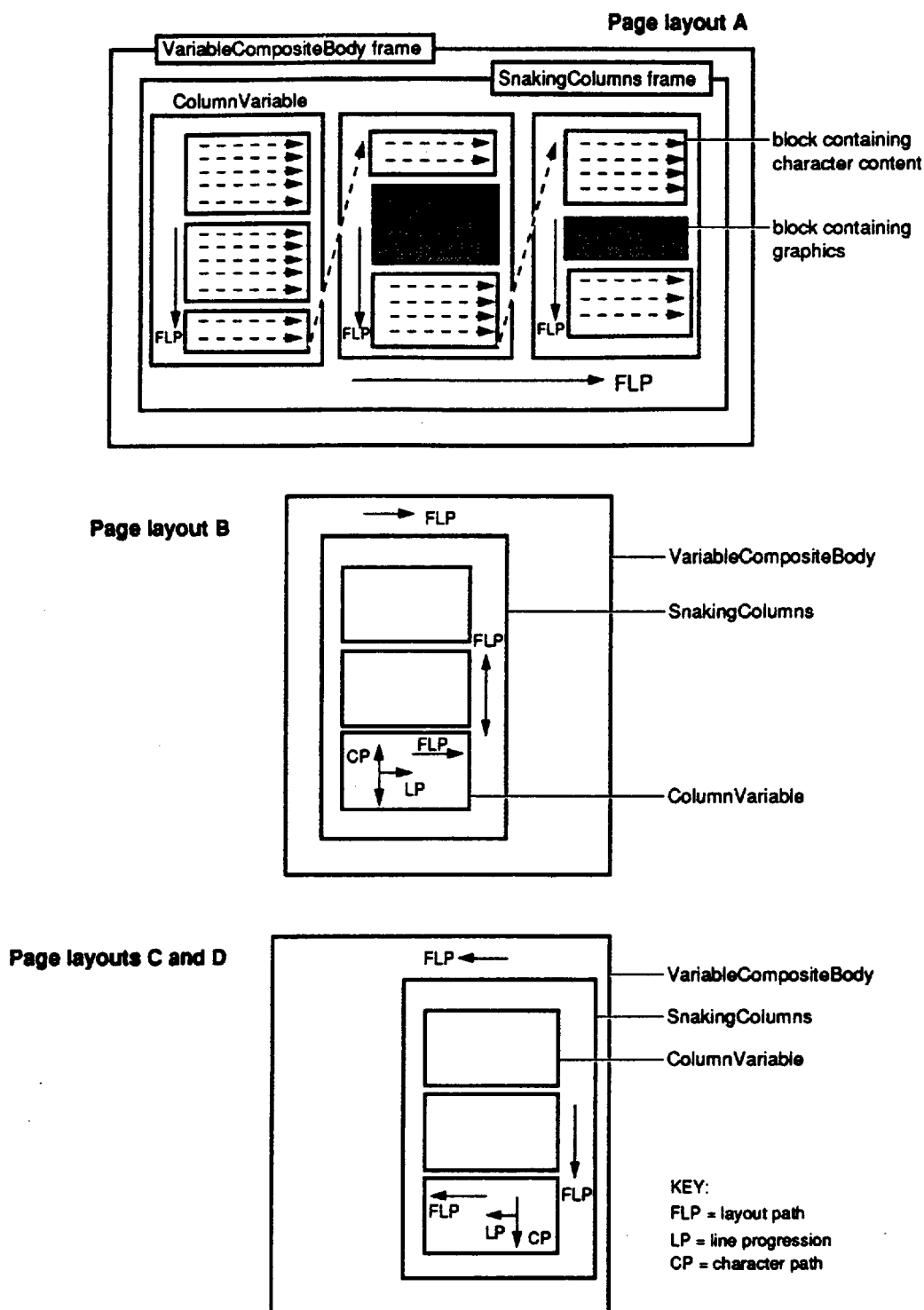


Figure 8 - Examples of the layout of snaking columns

6.3.5.6 SynchronizedColumns

SynchronizedColumns is a constituent constraint that defines a composite frame that represents a synchronized columns area within a body area. A synchronized columns area is typically used to represent one or more columns of content such that the content laid out in each column belongs to different layout streams. Thus content laid out in one column is not allowed to flow into the next column.

This type of column layout is typically used when it is required to layout separate amounts of content in parallel with one another such that they are aligned. Examples are the synchronized layout of content belonging to different languages and the layout of a figure in parallel with some text. An example is shown in Figure 9.

With regard to positioning and dimensioning, *SynchronizedColumns* frames have the same characteristics as *SnakingColumns* frames.

The immediate subordinates of a *SynchronizedColumns* frame consist of any number of frames of the type *ColumnFixed*.

The layout path for a *SynchronizedColumns* frame is 270° for page layout A, 0° for page layout B and 180° for page layouts C and D.

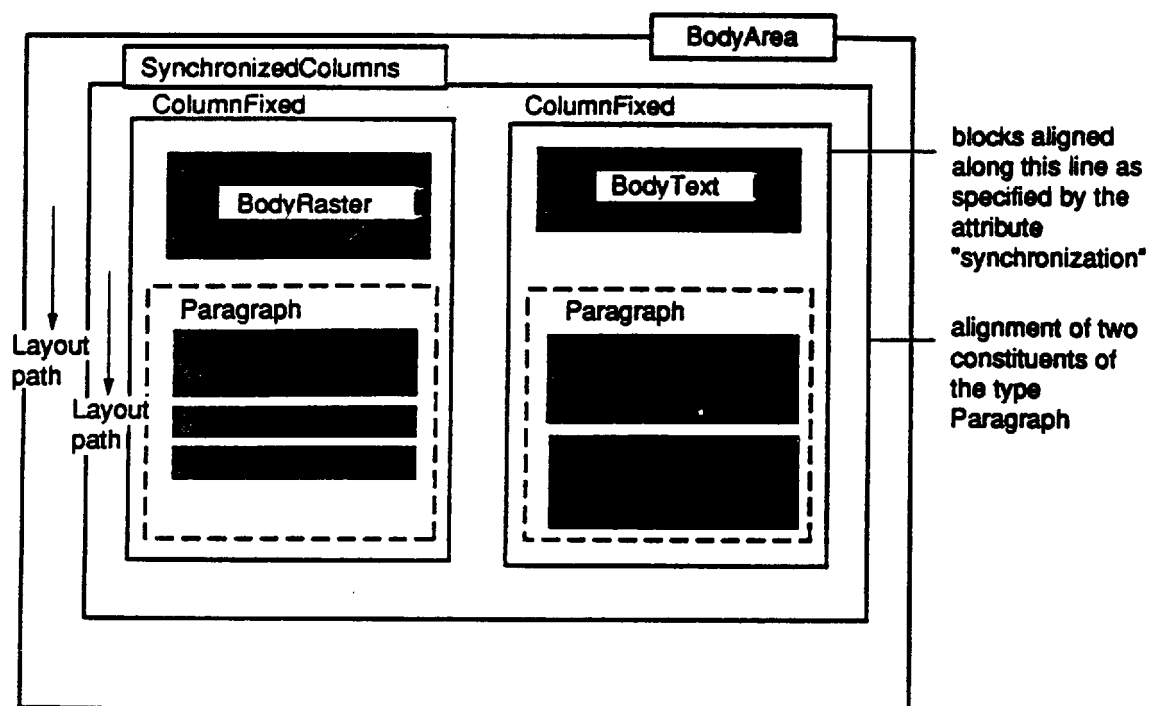


Figure 9 - Example of synchronized column layout for page layout A

6.3.5.7 ColumnVariable

ColumnVariable is a constituent constraint that defines a lowest level frame that is used to represent a column of content within a *SnakingColumns* frame. This is a frame which is variably positioned.

The dimension of this frame in the direction parallel to the layout path of the superior *SnakingColumns* frame (that is, the column width) is fixed. The dimensions of different instances of *ColumnVariable* frames within a given *SnakingColumns* frame may differ to allow columns of different widths to be specified.

The dimension in the direction orthogonal to the layout path of the superior frame (that is, the column length) may be specified by the sub-parameters

"Rule-B" or "maximum-size".

The layout path for *ColumnVariable* frames is 270° in the case of page layout A, 0° in page layout B and 180° in page layouts C and D.

All *ColumnVariable* frames subordinate to the same *SnakingColumns* frame shall have the same category name; different names may be used for *ColumnVariable* frames laid out in different *SnakingColumns* frames.

6.3.5.8 ColumnFixed

ColumnFixed is a constituent constraint that defines a lowest level frame that is used to represent a column of content within a *SynchronizedColumns* frame. This is a frame which has a fixed position.

The dimension of this frame in the direction orthogonal to the layout path of the superior *SynchronizedColumns* frame (that is, the column width) may be specified by either of the sub-parameters "fixed dimension" or "maximum size" in all page layout types. This dimension may differ for different instances of *ColumnFixed* frames within a given *SynchronizedColumns* frame to allow columns of different widths to be specified. However, the widths shall be specified such that the columns do not overlap.

The dimension of this frame in the direction parallel to the layout path of the superior frame (that is, the column length) may be specified by sub-parameters "Rule-B" or "maximum-size" in the cases of page layouts A and B. In the cases of page layouts C and D, this dimension shall only be specified as 'maximum-size'.

The *ColumnFixed* frames subordinate to a given *SynchronizedColumns* frame shall have different category names.

The layout path for *ColumnFixed* frames shall be equal to the layout path of the superior *SynchronizedColumns* frame.

The content laid out in different *ColumnFixed* frames within the same *SynchronizedColumns* frame may be specified synchronized by using the attribute "synchronization".

The sub-parameter "maximum size" shall only be specified for the right most *ColumnFixed* frames (relative to the page co-ordinate system) laid out in a *SynchronizedColumns* frame, to prevent overlapping of the frames.

6.3.5.9 FootnoteArea

FootnoteArea is a constituent constraint that defines a lowest level frame that is used to represent a footnote area within a body area. A footnote area is typically used for the layout of footnotes.

Frames of this type are variably positioned with a positioning fill order specified as 'reverse'. Hence this frame is positioned adjacent to the leading edge of the *VariableCompositeBody* frame.

The dimension of *FootnoteArea* frames in the direction orthogonal to the layout path of its superior frame is specified by one of the sub-parameters "fixed dimension" or "maximum size". In the direction of the layout path, the dimension is specified by sub-parameter "Rule-B" which means that this dimension is automatically adjusted to contain all the content that is allocated to it.

The layout path for *FootnoteArea* frames is the same as that specified for the body area.

The content that may be laid out in this frame is limited to the content that is associated with basic logical objects which are subordinates of the composite logical object *FootnoteBody*. To achieve this, the attribute "permitted categories" of this frame shall specify the value 'Footnote', the same name required on the basic logical objects for footnotes (see 6.2.3.10 and 6.2.3.11).

6.3.6 Header and footer area characteristics

6.3.6.1 General characteristics

The header and footer areas consist of either basic areas or composite areas.

A basic header or footer area is an area into which the content is directly laid out. This type of area is represented by a constituent constraint of the type *BasicHeader* or *BasicFooter* respectively.

A composite header or footer area is an area which is subdivided into separate sourced content and arranged content areas to provide greater versatility with regard to the layout of the content. This type of area is represented by a constituent constraint of the type *CompositeHeader* or *CompositeFooter* respectively.

In the case of basic header or footer areas, the content allocated to these areas is derived from the common part of the logical structure of a document. In the case of composite header or footer areas, the content may again be derived from the common part of the logical structure of a document but the content may also be derived from common content specified in the generic layout structure.

6.3.6.2 BasicHeader and BasicFooter

BasicHeader and *Basic Footer* are constituent constraints that define lowest level frames that represent areas within a page that are reserved for common content.

These types of frame have fixed positions and dimensions. The positioning of these frames within a page and the layout paths that may be specified for them depends upon the page layout type used (see 6.3.4.5)

The content that is laid out in these frames is derived, using the logical source mechanism, from the content associated with the composite logical object classes of the type *CommonContent*.

6.3.6.3 CompositeHeader and CompositeFooter

CompositeHeader and *CompositeFooter* are constituent constraints that define composite frames that represent areas within a page that are reserved for common content.

These types of frame have fixed positions and dimensions. The positioning of these frames within a page and the layout path that may be specified for them depends upon the page layout type used (see 6.3.4.5)

The subordinates of these frames may consist of either:

- a) any number and combination of variably positioned frames of the types *SourcedContentVariable* and *ArrangedContentVariable*.

or

- b) any number and combination of fixed positioned frames of the types *SourcedContentFixed* and *ArrangedContentFixed*.

In case b), the subordinate frames may overlap without restriction.

6.3.6.4 SourcedContentVariable

A *SourcedContentVariable* frame is a constituent constraint that defines a lowest level frame that represents a region within a header footer area that contains common content derived from the generic logical structure. This frame is variably positioned and its layout path is the same as that of the containing header or footer area.

The dimension of this frame in the direction orthogonal to the layout path of the superior frame is specified by one of the sub-parameters "fixed dimension" or "maximum-size". The dimensions of the frame in the direction parallel to the layout path of the superior frame is specified by one of the sub-parameters "fixed dimension" or "Rule-B".

This frame is required to specify the attribute "logical source" which indicates the particular instance of the constituent constraint *CommonContent* which contains the content to be laid out in this frame.

Typically, this frame is used for the positioning of content which is generated during the layout process, such as a character sequence containing a page number.

6.3.6.5 *ArrangedContentVariable*

An *ArrangedContentVariable* frame is a constituent constraint that defines a lowest level frame that represents a region within a header or footer area that contains pre-defined common content contained in the generic layout structure.

This frame is variably positioned and its dimensions are fixed.

This frame references one or more blocks of type *GenericBlock* (see 6.3.7) which contain the content to be laid out in this frame. Thus, this frame is typically used when it is required to layout pre-determined common content.

6.3.6.6 *SourcedContentFixed*

A *SourcedContentFixed* frame is a constituent constraint that defines a lowest level frame that represents a region within a header footer area that contains common content derived from the generic logical structure. This frame has a fixed position and dimensions and its layout path is equal to that of the containing header or footer area.

This frame is required to specify the attribute "logical source" which indicates the particular instance of the constituent constraint *CommonContent* which contains the content to be laid out in this frame.

Thus, as in the case of *SourcedContentVariable* frames, this frame is used for the positioning of content which is generated during the layout process, such as a character sequence containing a page number.

6.3.6.7 *ArrangedContentFixed*

An *ArrangedContentFixed* frame is a constituent constraint that defines a lowest level frame that represents a region within a header or footer area that contains pre-defined common content derived from the generic layout structure.

The position and dimensions of this frame are fixed. This frame references one or blocks of type *GenericBlock* (see 6.3.7) which contain the content to be laid out in this frame. Thus this frame is typically used when it is required to layout common content at pre-determined positions in the header or footer areas.

6.3.7 *GenericBlock* and *SpecificBlock*

Two types of constituent constraints of the object type 'block' are defined, namely *GenericBlock* and *SpecificBlock*. Object classes of the type *GenericBlock* may occur in the generic layout structure referenced by the attribute "generator for subordinates" of object classes of the types *ArrangedContentVariable* and *ArrangedContentFixed*. When the layout process is performed to produce a document in formatted processable form, equivalent blocks may occur in the specific layout structure. Objects of this type are restricted to occur within the header and footer areas of the page.

Objects of the type *SpecificBlock* shall only occur in the specific layout structure. They are created during the document layout process and result from the layout of basic logical objects into lowest level frames that constitute the body, header and footer areas.

6.4 Document layout characteristics

Mechanisms for controlling the allocation of logical constituents to various areas in the layout structure are defined in 6.4.1. Mechanisms for controlling the layout of the content within the allocated areas are defined in 6.4.2.

These mechanisms relate to documents for which a generic layout structure is specified. When a generic layout structure is not present, then these mechanisms are restricted as described in 6.4.3.

6.4.1 Flow controls

Various mechanisms are provided to control the allocation of constituent constraints representing the body parts of the logical structure of a document to pages sets, pages and body areas. These are described in 6.4.1.1, 6.4.1.2 and 6.4.1.3. The mechanisms for controlling the layout of the common parts of a document are described in 6.4.1.4.

6.4.1.1 Allocation of content to page sets

Two methods of allocating the constituent constraints associated with the *body* part of the document to page sets are provided:

- Layout in a nominated page set;
- Starting a new page set.

The first method provides the ability to specify that a part of a document is to be laid out entirely within a specified page set. This may be specified for constituent constraints of the types *Passage*, *NumberedSegment* and *Paragraph* using the attribute "layout object class" which specifies the object class identifier of the required class of page set.

The second method provides the ability to specify that the logical objects derived from a particular constituent constraint in a document and all subsequent parts of a document are to be laid out starting at the beginning of a new page set. This may be specified for the following constituent constraints:

- *Passage*;
- *NumberedSegment*;
- *Paragraph*;
- *Number*;
- *BodyText*;
- *BodyRaster*;
- *BodyGeometric*.

This is achieved using the attribute "new layout object" which specifies the object class identifier of the required class of page set.

6.4.1.2 Page breaks

This provides the ability to specify that the logical objects derived from a particular constituent constraint in a document and all subsequent parts of a document are to be laid out starting at the beginning of a new page. The page specified shall belong to the page set in which the logical objects from the immediately preceding constituent constraint is laid out. The specification of a page break shall not be used to layout part of a document in a new page set. If a new page set is required, then this shall be explicitly specified as described in 6.4.1.1.

This may be specified for the following constituent constraints:

- *Passage*;
- *NumberedSegment*;
- *Paragraph*;
- *Number*;
- *BodyText*;
- *BodyRaster*;
- *BodyGeometric*.

This is achieved using the attribute "new layout object". This attribute may specify the value 'page' indicating that the constituent constraint is to be laid out starting on the next available page which may be of any class. Alternatively, the attribute may specify that the constituent constraint is to be laid out starting on a page of a particular class; this is achieved by specifying the object class identifier of the required page class.

6.4.1.3 Allocation of content to body areas

If the page to which the content is allocated contains a basic body area, then the content is laid out in sequential order in that body area in the form of a single column.

If the page contains a composite body area, then the content is allocated to single, snaking and synchronized columns areas and footnote areas as described below.

6.4.1.3.1 Layout of content into column areas

When laying out content into a composite body area having more than one subordinate frame class (excluding FootnoteArea frame classes), it is necessary to indicate which of the column areas is to be used.

Logical objects of the types Paragraph, Number, FootnoteReference, BodyText, BodyRaster and BodyGeometric may be specified to be laid out in instances of one or more single columns area, snaking columns area or synchronized columns area.

This is done by giving each such basic logical component a value of the attribute "layout category" which corresponds to the value of the attribute "permitted categories" which applies to the lowest level of frame in which the content is to be laid out.

Note that any basic logical objects in the specific logical structure to which this attribute does not apply will be laid out only in a lowest level frame which has the implicit value of the attribute "permitted categories".

The use of layout categories ensures that if there is insufficient area on one page to lay out all of the content allocated to a particular type of area, then the laying out of the content will automatically continue in the same type of area in a succeeding page. Thus content is allowed to flow freely from one page to the next when the type of layout used at the end of one page is the same as that at the beginning of the succeeding page.

It is necessary to ensure the correct use of the mechanism for the layout of independent layout streams. In the absence of additional layout directives, content may be placed in available space within an earlier frame with appropriate values for the attribute "permitted categories". If this is not intended, it may be prevented by the use of the attribute "new layout object".

The attribute "new layout object" may be applied to logical components of the types NumberedSegment, Paragraph, Number, FootnoteReference, BodyText, BodyRaster and BodyGeometric, whenever a change in column layout is required.

The attribute "new layout object" may specify the identifier of the frame class that represents the single, snaking or synchronized column area required. In the case of single or synchronised column areas, the attribute "new layout object" may indicate the category name corresponding to the frame class of the single column area or of any of the columns within the synchronised column area that is required.

When layout occurs in a snaking columns area, column breaks may be indicated by using the attribute "new layout object". This attribute may specify the identifier or the category name of the frame corresponding to the column in which the layout is to continue. However only the use of a category name will ensure that a single column break is always obtained, irrespective of the frame class actually used.

When the layout is to occur in a synchronized columns area, category names are used to control the particular columns that are to be used to layout the logical entities. Each column within a synchronized columns area shall have a different permitted category and each basic logical entity to be laid out in this particular area shall have a category name corresponding to a name allocated to one of the columns. The logical entities allocated to different columns may be aligned using the attribute "synchronization".

6.4.1.3.2 Layout of footnotes

The basic logical objects derived from constituent constraints that represent the content belonging to a footnote (i.e FootnoteNumber and FootnoteText) are constrained to be laid out in a footnote area which is represented by a FootnoteArea frame (see 6.3.5.9).

This constraint is specified by means of category names. That is, the logical constituents of the types FootnoteNumber and FootnoteText and layout constituents of the type FootnoteArea are all required to have the category name with the value 'Footnote'.

More than one footnote may be placed in a footnote area within a given body area. In this case the content belonging to the footnotes are laid out sequentially in the footnote area in accordance with their reading order.

If the content belonging to a footnote cannot all be accommodated in the footnote area on one page, then the content may freely flow into the next footnote area on the subsequent page. Alternatively, it is possible to specify that a footnote is to be laid out entirely within a particular footnote area. This is achieved using the attribute "indivisibility".

6.4.1.4 Allocation of content to header-footer areas

A header or footer area may be basic or composite (see 6.3.6.1). In the case of a basic area, the frame representing that area specifies the attribute "logical source" which indicates the particular instance of the constituent constraint of the type CommonContent that is to be laid out in that area. The basic logical constituents subordinate to CommonContent are then laid out in accordance with their sequential order.

In the case of a composite header or footer area (see 6.3.6.3), the area is divided into one or more separate areas, each of which is represented by a lowest level frame. The content allocated to the separate areas may be derived from one of two sources. That is, the content may be pre-defined and represented by one or more blocks which are directly associated with the lowest level frame. Alternatively, the lowest level frame may specify the attribute "logical source" which, as above, indicates the particular logical object of the type CommonContent that is to be laid out in that frame.

6.4.2 Layout of the document content

Various constraints may be specified to control the layout of the content into the body, header and footer areas. These constraints are described below.

6.4.2.1 Margins

The margins are the minimum distances, or offsets, between a part of the document content and the edge of the particular area in which that content is laid out. The margins define the maximum extents of the available area into which the content shall be positioned.

Margins may be specified for any constituent constraint representing a basic logical object; different margin values may be specified for different constituent constraints without restriction.

Four margins may be independently specified for each constituent constraint, namely:

- trailing edge margin;
- leading edge margin;
- right hand edge margin;
- left hand edge margin.

These margins are defined in relationship to the layout path specified for the frame in which the content is to be laid out in (see Figure 10).

Any combination of the above margins may be specified for a particular constituent constraint. These margins are specified by the attribute "offset". Any value may be specified in units of BMUs. If a particular margin is not specified, then it is assumed to be 0 BMUs.

6.4.2.2 Separation

Leading separations is the minimum distance between one basic logical object and the next one, if any, when they are laid out; trailing separation is the minimum distance between one basic logical object and the previous one, if any, when they are laid out. Both may be specified for basic logical components of any constituent constraint types. These distances are specified in BMUs by the attribute "separation". If no value is specified, then the minimum distance is assumed to be 0 BMUs.

6.4.2.3 Indivisibility

Indivisibility provides the means to specify whether or not a basic or composite logical object derived from a constituent constraint is allowed to be split over more than one page or over more than one area within a page. It may be specified for constituent constraints of the types Passage, NumberedSegment, Paragraph, Footnote, Number, FootnoteReference and BodyText. The attribute "indivisibility" is used to specify this feature.

6.4.2.4 Same layout object

Same layout object provides the means to specify that the start of the content associated with a logical object and the end of the content associated with the previous logical object are to be laid out within a single layout object. This may be specified for logical objects of the types NumberedSegment, Paragraph, Number, Footnote, FootnoteReference, BodyText, BodyRaster and BodyGeometric. The attribute "same layout object" is used to specify this feature.

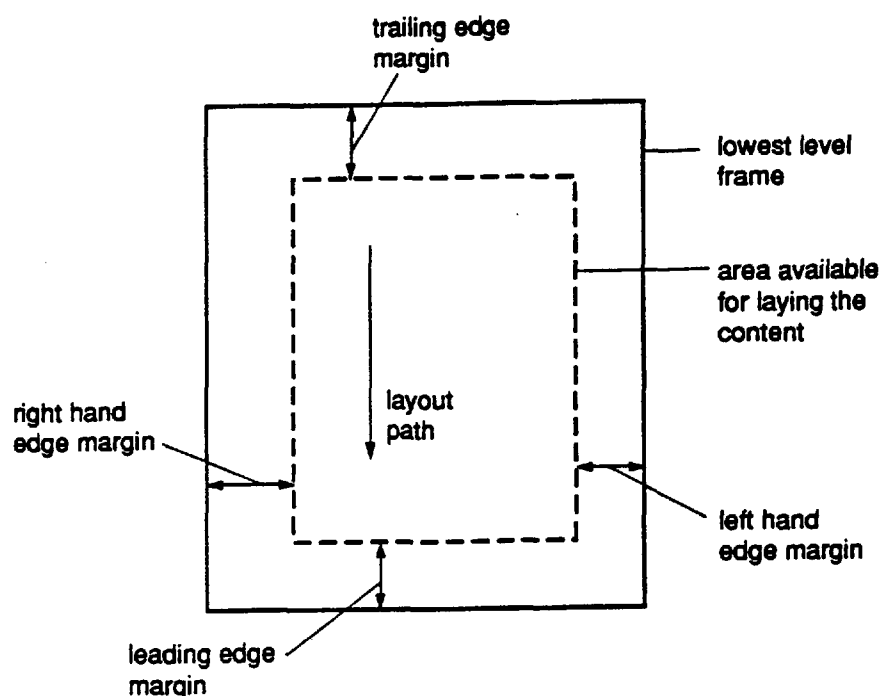


Figure 10 - Specification of margins

6.4.2.5 Concatenation

Concatenation provides the means to specify that the content associated with a basic logical object derived from a constituent constraint and the content associated with the basic logical object derived from a previous constituent constraint are to be regarded as an unbroken stream of content. This may be specified for constituent constraints of the types *BodyText*, *Number*, *FootnoteReference*, *FootnoteText*, *CommonText* and *PageNumber*. The attribute "concatenation" is used to specify this feature.

6.4.2.6 Block alignment

Block alignment allows the content associated with a basic logical object to be specified as 'left aligned', 'right aligned' or 'centred' within the area in which that content is laid out. Left aligned means that the content is laid out adjacent to the left hand edge margin. Right aligned means that the content is laid out adjacent to the right hand edge margin and centred means that the content is laid out midway between the left and right margins.

This feature may only be specified using the attribute "block alignment" for constituent constraints of the types *Number*, *FootnoteReference*, *FootnoteText*, *PageNumber*, *FootnoteNumber*, *BodyText* and *CommonText*, and when they contain formatted character content, *BodyRaster*, *BodyGeometric*, *CommonRaster* and *CommonGeometric*.

6.4.3 Layout controls applicable in the absence of a generic layout structure

In processable form documents the generic layout structure is optional. If the generic layout structure is omitted, then it is the responsibility of the receiver to define an appropriate layout structure. No limitations are placed on the layout structure that is used.

When a generic layout structure is not specified within a processable form document, then restrictions are placed on the layout control functions described in 6.4.1 and 6.4.2 that may be specified within the document. These restrictions are indicated below.

- It is not possible to specify that certain logical parts of a document are to be allocated to a given page set or that a part of a document is to be laid out starting in a new page set, as defined in 6.4.1.1.
- It is possible to specify page breaks as defined in 6.4.1.2 but it is only possible to indicate that the layout shall begin on a new page. It is not possible to specify a particular page class.
- The logical parts of the document that are intended to be laid out in the body area and in the header/footer areas of each page may be distinguished from each other by means of application comments specified for them (see 6.6.4). An exception is that it is not possible to distinguish whether a particular portion of the common content is to be placed in a header or footer area (or both).
- It is not possible to indicate the type of layout area to be used to layout each logical constituent in the body part of a document. That is, it is not possible to indicate whether single column or multiple column areas are to be used (see 6.4.1.3.1). This shall be decided by the receiver.
- Footnotes within the body part of a document may be distinguished by use of the attribute "application comments". Footnotes are intended to be read and laid out separately from the other logical constituents of the body part (see 6.4.1.3.2). However, it is the responsibility of the receiver to decide how footnotes are laid out.
- Margins, separation, indivisibility, same layout object, concatenation, and block alignment, as defined in 6.4.2 may all be specified. Only one restriction applies. Indivisibility (see 6.4.2.3) may be assumed to specify that a logical object derived from the constituent constraint is not to split over more than one page but indivisibility shall not be specified for other types of layout areas such as single or multiple column areas.

6.5 Content layout and imaging characteristics

A document may contain character, raster graphics and geometric graphics content.

The content architectures that may be specified using the attribute "content architecture class" are formatted character, processable character, formatted processable character, formatted processable raster graphics and formatted processable geometric graphics. Any of these may be specified as the default in the document profile.

6.5.1 Character content

6.5.1.1 Introduction

This subclause defines the features that are applicable to the character content contained in a document and the presentation attributes and control functions that may be used to specify these features. These features may apply to basic logical and layout components unless otherwise indicated.

The default values for the following features may be specified in the document profile:

- graphic character sets;
- graphic character subrepertoire;
- code extension announcers;
- line spacing;
- character spacing;
- character path;
- line progression;
- character orientation;
- graphic rendition, including the parameters values:

default rendition, increased intensity (bold), italicized, underlined, crossed-out, primary font, 1st alternative font, 2nd alternative font, 3rd alternative font, 4th alternative font, 5th alternative font, 6th alternative font, 7th alternative font,