

**Implementation of
EN ISP 11182-1:1994**

**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) —**

**(Regional variant of
ISO/IEC ISP 11182-1:1993**

The European Standard EN ISP 11182-1:1994 has the status of a
British Standard

UDC 681.3:621.39:651.51.7

Committees responsible for this British Standard

The preparation of this British Standard was entrusted to Technical Committee IST/18, Document processing and related communication, upon which the following bodies were represented:

- British Computer Society
- British Telecommunications plc
- Digital Equipment Co. Ltd.
- HM Treasury (Central Computer and Telecommunications Agency)
- IBM United Kingdom Ltd.
- International Computers Limited
- Joint Information Systems Committee of the Higher Education Funding Council
- Kodak Limited
- Level 7 Ltd.
- National Computing Centre Ltd.
- National Health Services
- Rank Xerox Ltd.
- SEMA Group Systems Ltd.

This British Standard, having been prepared under the direction of the Information Systems Technology Assembly, was published under the authority of the Standards Board and comes into effect on 15 July 1995

© BSI 03-2000

The following BSI references relate to the work on this standard:
Committee reference IST/18
Draft for comment 91/68284 DC

ISBN 0 580 24055 X

Amendments issued since publication

Amd. No.	Date	Comments

Contents

	Page
Committees responsible	Inside front cover
National foreword	ii
Foreword	2
Foreword	xiii
Text of ISO/IEC ISP 11182-1	1

National foreword

This British Standard reproduces verbatim EN ISP 11182-1:1994 and implements it as the UK national standard. This British Standard supersedes DD ENV 41510:1991 which is withdrawn.

This British Standard is published under the direction of the Information Systems Technology Assembly, whose Technical Committee IST/18 has the responsibility to:

- aid enquirers to understand the text;
- present to the responsible European committee any enquiries on interpretation, or proposals for change, and keep UK interests informed;
- monitor related international and European developments and promulgate them in the UK.

NOTE International and European Standards, as well as overseas standards, are available from Customer Services, BSI, 389 Chiswick High Road, London W4 4AL.

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

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

Summary of pages

This document comprises a front cover, an inside front cover, pages i and ii, the EN ISP title page, page 2, the ISO/IEC ISP title page, pages ii to xiv, pages 1 to 205 and a back cover.

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

ICS 35.100

Descriptors: Data processing, text processing, information interchange, documentation, documents, logical structure, layout, formats, profiles.

English version

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)

(Regional variant of ISO/IEC ISP 11182-1:1993)

Technologies de l'information — Profil
normalisé international FOD36 — 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)
(Variante régionale de
l'ISO/IEC ISP 11182-1:1993)

Informationstechnik — Internationale
Profilnorm FOD36 — Format für
Bürodokumente — Komplexe
Dokumentstruktur — Inhaltsarchitekturen für
Schriftzeichen, Rasterbilder und
Geometriegrafiken
Teil 1: Dokumentanwendungsprofil (DAP)
(regionale Version von
ISO/IEC ISP 11182-1:1993)

This European Standard was approved by CEN on 1994-12-05. CEN members are bound to comply with the CEN/CENELEC Internal Regulations which stipulate the conditions for giving this European Standard the status of a national standard without any alteration.

Up-to-date lists and bibliographical references concerning such national standards may be obtained on application to the Central Secretariat or to any CEN member.

This European Standard exists in three official versions (English, French, German). A version in any other language made by translation under the responsibility of a CEN member into its own language and notified to the Central Secretariat has the same status as the official versions.

CEN members are the national standards bodies of Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and United Kingdom.

CEN

European Committee for Standardization
Comité Européen de Normalisation
Europäisches Komitee für Normung

Central Secretariat: rue de Stassart 36, B-1050 Brussels

© 1994 Copyright reserved to CEN members

Foreword

The Technical Board decided to submit the International Standard *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)* (Regional variant of ISO/IEC ISP 11182-1:1993) for adoption by the Unique Acceptance Procedure (UAP), in accordance with clause 4.6 of the CEN/CENELEC Internal Regulations Part 2 of April 1990.

The result of the Unique Acceptance Procedure was positive.

For the time being this standard exists in the English version only.

This document is the first part of a multi-part Functional Standard identified in M-IT-02 as *F/OD36 — Extended Document Format — Character, raster graphics and geometric graphics content architectures*.

M-IT-02, *CEN/CENELEC/ETSI Memorandum on taxonomy of profiles and directory of functional standards*.

This European Standard supersedes ENV 41510:1990 which shall be withdrawn.

This European Standard shall be given the status of a National Standard, either by publication of an identical text or by endorsement, at the latest by June 1995, and conflicting national standards shall be withdrawn at the latest by June 1995.

According to the CEN/CENELEC Internal Regulations, the following countries are bound to implement this European Standard: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and United Kingdom.

Endorsement notice

The text of the International Standard ISO/IEC ISP 11182-1:1993 was approved by CEN as a European Standard with agreed regional variations as given below. Section and line numbers correspond to the equivalent numbers in the International Standard.

1 Scope

Replace the last two paragraphs by:

“A document structured in accordance with this International Standard Profile is represented for interchange by a document application profile which uses the Open Document Interchange Format (ODIF) as defined in CCITT Rec. T.415/ISO 8613-5.”

5.1 Data stream conformance

Replace 5.1, first bullet with:

“the data stream shall be encoded in accordance with the ASN.1 encoding rules defined in CCITT Rec. X.209/ISO/IEC 8825

5.2 Implementation conformance

Replace 5.2, first paragraph, first sentence with:

“A conforming receiving implementation shall be capable of receiving any data streams conforming to this International Standardized Profile structured in accordance with ODIF.”

8 Interchange format

Replace 8, first paragraph with:

“For conformance to this International Standardized Profile, the interchange format class A shall be used.”

8.2 Interchange format SDIF

8.2 and its sub-clauses do not apply.

INTERNATIONAL
STANDARDIZED
PROFILE

ISO/IEC
ISP
11182-1

First edition
1993-11-01

**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)

*Technologies de l'information — Profil normalisé international FOD36 —
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 11182-1:1993(E)

AVA Documents
Document Sharing Hub

Contents

	Page
Foreword	xiii
Introduction	1
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	4
5.1 Data stream conformance	4
5.2 Implementation conformance	5
6 Characteristics supported by this document application profile	5
6.1 Overview	5
6.1.1 General	5
6.1.2 Formatted form documents	5
6.1.3 Processable form documents	6
6.1.4 Formatted processable form documents	6
6.1.5 Generic documents	6
6.1.6 Document profile	7
6.2 Logical characteristics	7
6.2.1 Introduction	7
6.2.2 Overview of the logical structure	7
6.2.3 Body part of the logical structure	7
6.2.3.1 DocumentLogicalRoot	7
6.2.3.2 Passage	7
6.2.3.3 NumberedSegment	8
6.2.3.4 Number	9
6.2.3.5 Paragraph	9
6.2.3.6 BodyText, BodyRaster and BodyGeometric	9
6.2.3.7 Footnote	10
6.2.3.8 FootnoteReference	10
6.2.3.9 FootnoteBody	11
6.2.3.10 FootnoteNumber	11
6.2.3.11 FootnoteText	11
6.2.3.12 Title	11
6.2.3.13 Phrase	12
6.2.3.14 Constituents that provide for a general referencing mechanism	12
6.2.3.14.1 Reference	12
6.2.3.14.2 ReferencedContent	13
6.2.3.15 Constituents representing illustrations	13
6.2.3.15.1 Introduction	13
6.2.3.15.2 Figure	13
6.2.3.15.3 Artwork	14
6.2.3.15.4 Form	14
6.2.3.15.5 Caption	14
6.2.3.15.6 Description	14
6.2.3.15.7 EntryGroup	15
6.2.3.16 Constituents used to represent tables	15
6.2.3.16.1 Introduction	15

	Page	
6.2.3.16.2	Table	15
6.2.3.16.3	Row	15
6.2.3.16.4	TableComponent	15
6.2.3.16.5	RowComponent	16
6.2.3.16.6	EntryElement	16
6.2.3.16.7	EntryText, EntryRaster and EntryGeometric	16
6.2.3.17	Constituents representing lists	16
6.2.3.17.1	Introduction	16
6.2.3.17.2	NumberedList	16
6.2.3.17.3	UnNumberedList	17
6.2.3.17.4	DefinitionList	17
6.2.3.17.5	ListItem	17
6.2.3.17.6	ListTerm	17
6.2.4	Common content part of the logical structure	17
6.2.4.1	CommonContent	17
6.2.4.2	CommonText	18
6.2.4.3	PageNumber	18
6.2.4.4	CommonRaster	18
6.2.4.5	CommonGeometric	18
6.2.4.6	CommonReference	18
6.2.4.7	CurrentInstance	18
6.2.4.8	TableNumber	19
6.2.4.9	CommonNumber	19
6.3	Layout characteristics	19
6.3.1	Overview of the layout characteristics	19
6.3.2	DocumentLayoutRoot	20
6.3.3	PageSet	20
6.3.4	Page characteristics	20
6.3.4.1	Page constituents	20
6.3.4.2	Page dimensions	20
6.3.4.3	Nominal page sizes	21
6.3.4.4	Page offset	21
6.3.4.5	Page layout characteristics	21
6.3.4.5.1	General characteristics	21
6.3.4.5.2	Body area layout	21
6.3.4.5.3	Header/footer area layout	22
6.3.4.5.4	Page layout	22
6.3.5	Body area characteristics	24
6.3.5.1	General characteristics	24
6.3.5.2	BasicBody	24
6.3.5.3	VariableCompositeBody	24
6.3.5.4	BasicFloat	25
6.3.5.5	SnakingColumns	29
6.3.5.6	SynchronizedColumns	30
6.3.5.7	ColumnVariable	31
6.3.5.8	ColumnFixed	31
6.3.5.9	FootnoteArea	32
6.3.5.10	FixedCompositeBody	32

	Page
6.3.5.11 CompositeFloat	33
6.3.5.12 CompositeFixtureVariable	33
6.3.5.13 CompositeFixtureFixed	34
6.3.5.14 CompositeArtwork	34
6.3.5.15 BasicFixture	34
6.3.5.16 FormArea	35
6.3.5.17 EntryGroupArea	36
6.3.5.18 FormEntryArea	36
6.3.5.19 BasicColumn	36
6.3.5.20 CompositeColumnVariable	37
6.3.5.21 CompositeColumnFixed	37
6.3.5.22 CompositeCommon	37
6.3.5.23 Constituents used for laying out tables	38
6.3.5.23.1 Overview	38
6.3.5.23.2 TableArea	40
6.3.5.23.3 TableHeader	40
6.3.5.23.4 TableLabel	41
6.3.5.23.5 CompositeTableLabel	41
6.3.5.23.6 LabelComponent	41
6.3.5.23.7 TableLabelContent	42
6.3.5.23.8 RowArea	42
6.3.5.23.9 SubRowGroup	42
6.3.5.23.10 SubRow	42
6.3.5.23.11 Cell	42
6.3.6 Header and footer area characteristics	43
6.3.6.1 General characteristics	43
6.3.6.2 BasicHeader and BasicFooter	43
6.3.6.3 CompositeHeader and CompositeFooter	43
6.3.6.4 SourcedContentVariable	43
6.3.6.5 ArrangedContentVariable	44
6.3.6.6 SourcedContentFixed	44
6.3.6.7 ArrangedContentFixed	45
6.3.7 GenericBlock and SpecificBlock	45
6.4 Document layout characteristics	45
6.4.1 Flow controls	45
6.4.1.1 Allocation of content to page sets	45
6.4.1.2 Page breaks	46
6.4.1.3 Allocation of content to body areas	46
6.4.1.3.1 General mechanism for laying out content in a composite body area	46
6.4.1.3.2 Layout into BasicFloat and BasicFixture frames	47
6.4.1.3.3 Layout in SnakingColumns frames	47
6.4.1.3.4 Layout in SynchronizedColumns frames	48
6.4.1.3.5 Layout in CompositeFloat frames	48
6.4.1.3.6 Layout of figures	48
6.4.1.3.7 Layout of tables	49
6.4.1.3.8 Layout of forms	50
6.4.1.3.9 Layout of footnotes	50
6.4.1.4 Allocation of content to header and footer areas	51

		Page
6.4.2	Layout of document content	51
6.4.2.1	Margins	51
6.4.2.2	Separation	51
6.4.2.3	Indivisibility	51
6.4.2.4	Same layout object	52
6.4.2.5	Concatenation	52
6.4.2.6	Block alignment	52
6.4.3	Layout controls applicable in the absence of a generic layout structure	53
6.5	Content layout and imaging characteristics	53
6.5.1	Character content	53
6.5.1.1	Introduction	53
6.5.1.2	Character content architecture class	54
6.5.1.3	Character repertoires	54
6.5.1.4	Code extension techniques	55
6.5.1.5	Line spacing	57
6.5.1.6	Character spacing	57
6.5.1.7	Character path and line progression	57
6.5.1.8	Character positioning controls	58
6.5.1.9	Character orientation	58
6.5.1.10	Emphasis	58
6.5.1.11	Tabulation	59
6.5.1.12	Indentation	59
6.5.1.13	Alignment	59
6.5.1.14	First line format	60
6.5.1.15	Widow and orphan sizes	60
6.5.1.16	Fonts	60
6.5.1.17	Reverse character strings	61
6.5.1.18	Kerning offset	61
6.5.1.19	Proportional line spacing	61
6.5.1.20	Superscripts and subscripts	61
6.5.1.21	Line breaks	61
6.5.1.22	Substitution of characters	61
6.5.1.23	Initial point	61
6.5.1.24	Use of control functions	61
6.5.1.25	Formatting the content	62
6.5.1.26	Graphic character composition	62
6.5.1.27	Parallel text	62
6.5.1.28	Pairwise kerning	62
6.5.2	Raster graphics content	63
6.5.2.1	Introduction	63
6.5.2.2	Raster graphics content architecture	63
6.5.2.3	Raster graphics encoding methods	63
6.5.2.4	Pel path and line progression	63
6.5.2.5	Clipping	64
6.5.2.6	Pel spacing	64
6.5.2.7	Spacing ratio	64
6.5.2.8	Image dimensions	64

	Page
6.5.3 Geometric graphics content	64
6.6 Miscellaneous features	64
6.6.1 Resource-documents	64
6.6.2 External-documents	65
6.6.3 Border	65
6.6.4 Application comments	65
6.6.5 Alternative representation	65
6.6.6 Automatic numbering and referencing mechanisms	67
6.6.6.1 Introduction	67
6.6.6.2 Bindings	67
6.6.6.2.1 Binding “prefixes-<n>”	68
6.6.6.2.2 Binding “suffixes-<n>”	68
6.6.6.2.3 Binding “numberstring-<n>”	68
6.6.6.2.4 Binding “numbers-<n>”	68
6.6.6.2.5 Binding “separator-<n>”	68
6.6.6.2.6 Binding “string-<n>”	68
6.6.6.2.7 Binding “PGnum”	68
6.6.6.2.8 Binding “fnotenumber”	68
6.6.6.2.9 Binding “fnotestring”	68
6.6.6.3 Numbering of segments	68
6.6.6.4 Numbering of figures	70
6.6.6.5 Numbering of lists	70
6.6.6.6 Numbering of tables	70
6.6.6.7 Footnote numbering	70
6.6.6.8 Page numbering	71
6.6.6.9 Referenced content	71
6.6.6.10 Common references	72
6.6.6.11 Current instance references	72
6.6.6.12 Common number references	72
6.6.7 User readable comments	73
6.6.8 User visible name	73
6.6.9 Unit scaling	73
6.7 Document management features	73
6.7.1 Document constituent information	73
6.7.1.1 Presence of document constituents	73
6.7.1.2 Resource-document information	74
6.7.1.3 External-document information	74
6.7.2 Document identification information	74
6.7.2.1 Document application profile information	74
6.7.2.2 Document architecture class information	74
6.7.2.3 Content architecture classes information	74
6.7.2.4 Interchange format class information	74
6.7.2.5 ODA version information	74
6.7.2.6 Document reference	74
6.7.3 Document default information	74
6.7.4 Non-basic characteristics information	74
6.7.4.1 Profile character sets	75
6.7.4.2 Comments character sets	75

	Page
6.7.4.3	Alternative representation character sets 76
6.7.5	Fonts list 76
6.7.6	Document management attributes 76
7	Specification of constituent constraints 76
7.1	Introduction 76
7.1.1	Diagrams of relationships of logical constituents 77
7.1.2	Diagrams of relationships of layout constituents 82
7.1.3	Notation 88
7.2	Document profile constituent constraints 89
7.2.1	Macro definitions 89
7.2.2	Constituent constraints 95
7.2.2.1	DocumentProfile 95
7.3	Logical constituent constraints 100
7.3.1	Macro definitions 100
7.3.2	Factor constraints 118
7.3.2.1	FACTOR ANY-LOGICAL 118
7.3.2.2	FACTOR COMP-LOGICAL 118
7.3.2.3	FACTOR BASIC-LOGICAL 119
7.3.2.4	FACTOR ANY-COMMON 119
7.3.3	Constituent constraints 119
7.3.3.1	DocumentLogicalRoot 119
7.3.3.2	Passage 120
7.3.3.3	NumberedSegment 120
7.3.3.4	Number 121
7.3.3.5	Title 121
7.3.3.6	Caption 122
7.3.3.7	Paragraph 122
7.3.3.8	Phrase 122
7.3.3.9	Footnote 123
7.3.3.10	FootnoteNumber 123
7.3.3.11	FootnoteReference 124
7.3.3.12	FootnoteBody 124
7.3.3.13	FootnoteText 124
7.3.3.14	Figure 125
7.3.3.15	BodyText 125
7.3.3.16	Reference 125
7.3.3.17	ReferencedContent 126
7.3.3.18	BodyRaster 126
7.3.3.19	BodyGeometric 127
7.3.3.20	CommonContent 127
7.3.3.21	CommonText 127
7.3.3.22	CommonReference 128
7.3.3.23	CommonNumber 128
7.3.3.24	CurrentInstance 128
7.3.3.25	CommonRaster 128
7.3.3.26	CommonGeometric 129
7.3.3.27	PageNumber 129
7.3.3.28	TableNumber 129

	Page
7.3.3.29 Description	129
7.3.3.30 Artwork	130
7.3.3.31 NumberedList	130
7.3.3.32 UnNumberedList	131
7.3.3.33 DefinitionList	131
7.3.3.34 ListItem	131
7.3.3.35 ListTerm	132
7.3.3.36 Table	132
7.3.3.37 Row	132
7.3.3.38 TableComponent	133
7.3.3.39 RowComponent	133
7.3.3.40 Form	133
7.3.3.41 EntryElement	134
7.3.3.42 EntryGroup	134
7.3.3.43 EntryText	134
7.3.3.44 EntryRaster	135
7.3.3.45 EntryGeometric	135
7.4 Layout constituent constraints	136
7.4.1 Macro definitions	136
7.4.2 Factor constraints	140
7.4.2.1 FACTOR ANY-LAYOUT	140
7.4.2.2 FACTOR ANY-PAGE	141
7.4.2.3 FACTOR ANY-FRAME-FIXED	141
7.4.2.4 FACTOR ANY-FRAME-VARIABLE	142
7.4.3 Constituent constraints	143
7.4.3.1 DocumentLayoutRoot	143
7.4.3.2 PageSet	143
7.4.3.3 Page	144
7.4.3.4 RectoPage	144
7.4.3.5 VersoPage	145
7.4.3.6 CompositeHeader	145
7.4.3.7 CompositeFooter	146
7.4.3.8 FixedCompositeBody	146
7.4.3.9 VariableCompositeBody	147
7.4.3.10 ColumnFixed	148
7.4.3.11 ColumnVariable	149
7.4.3.12 SnakingColumns	150
7.4.3.13 SynchronizedColumns	151
7.4.3.14 BasicFloat	152
7.4.3.15 CompositeFloat	153
7.4.3.16 BasicColumn	154
7.4.3.17 FootnoteArea	156
7.4.3.18 ArrangedContentFixed	157
7.4.3.19 ArrangedContentVariable	157
7.4.3.20 SourcedContentFixed	158
7.4.3.21 SourcedContentVariable	159
7.4.3.22 CompositeFixtureVariable	161
7.4.3.23 CompositeFixtureFixed	163

	Page
7.4.3.24 BasicFixture	164
7.4.3.25 CompositeColumnFixed	165
7.4.3.26 CompositeColumnVariable	166
7.4.3.27 CompositeCommon	167
7.4.3.28 CompositeArtwork	168
7.4.3.29 BasicHeader	169
7.4.3.30 BasicFooter	169
7.4.3.31 BasicBody	170
7.4.3.32 GenericBlock	170
7.4.3.33 SpecificBlock	172
7.4.3.34 FormArea	173
7.4.3.35 EntryGroupArea	173
7.4.3.36 TableArea	174
7.4.3.37 TableHeader	175
7.4.3.38 TableLabel	176
7.4.3.39 CompositeTableLabel	177
7.4.3.40 LabelComponent	178
7.4.3.41 RowArea	179
7.4.3.42 Cell	180
7.4.3.43 SubRowGroup	181
7.4.3.44 SubRow	182
7.4.3.45 TableLabelContent	183
7.4.3.46 FormEntryArea	183
7.5 Layout style constituent constraints	184
7.5.1 Macro definitions	184
7.5.2 Factor constraints	184
7.5.3 Constituent constraints	184
7.5.3.1 LStyle1	184
7.5.3.2 LStyle2	185
7.5.3.3 LStyle3	186
7.5.3.4 LStyle4	186
7.5.3.5 LStyle5	186
7.5.3.6 LStyle6	187
7.5.3.7 LStyle7	187
7.5.3.8 LStyle8	187
7.5.3.9 LStyle9	187
7.5.3.10 LStyle10	188
7.5.3.11 LStyle11	188
7.5.3.12 LStyle12	189
7.5.3.13 LStyleT1	189
7.5.3.14 LStyleT2	189
7.5.3.15 LStyleT3	190
7.5.3.16 LStyleT4	190
7.5.3.17 LStyleT5	190
7.5.3.18 LStyleT6	190
7.5.3.19 LStyleT7	191
7.5.3.20 LStyleT8	191
7.5.3.21 LStyleT9	191

		Page
7.6	Presentation style constituent constraints	191
7.6.1	Macro definitions	191
7.6.2	Factor constraints	191
7.6.3	Constituent constraints	192
7.6.3.1	PStyle1	192
7.6.3.2	PStyle2	192
7.6.3.3	PStyle3	193
7.6.3.4	PStyle4	193
7.7	Content portion constituent constraints	194
7.7.1	Macro definitions	194
7.7.2	Factor constraints	194
7.7.3	Constituent constraints	194
7.7.3.1	Character-content-portion	194
7.7.3.2	Raster-graphics-content-portion	196
7.7.3.3	Geometric-graphics-content-portion	196
8	Interchange format	196
8.1	Interchange format class A	196
8.1.1	Interchange format	196
8.1.2	Document application profile object identifier	196
8.1.3	Encoding of application comments	196
8.1.4	Data lengths	197
8.2	Interchange format SDIF	197
8.2.1	Interchange format	197
8.2.2	Document application profile object identifier	197
8.2.3	Encoding of application comments	197
Annex A	Amendments and technical corrigenda	198
A.1	Amendments	198
A.1.1	Amendments to the base standard	198
A.1.2	Proposed changes to standards due to defects	198
A.2	Technical corrigenda	198
A.2.1	Technical corrigenda to this International Standardized Profile	198
A.3	Versions of the ODA International Standard	198
Annex B	Recommended practices	199
B.1	Transfer methods for ODA	199
B.1.1	Conveyance of ODA over CCITT X.400-1984	199
B.1.2	Conveyance of ODA over FTAM	199
B.1.3	Conveyance of ODA over DTAM	199
B.1.4	Conveyance of ODA over flexible disks	199
B.2	Font reference	200
B.3	ISO/IEC 8632 (CGM) constraints for this document application profile	201
B.3.1	Delimiter elements	201
B.3.2	Metafile descriptor elements	201
B.3.3	Picture descriptor elements	202
B.3.4	Control elements	202
B.3.5	Graphical primitive elements	202
B.3.6	Attribute elements	203
B.3.7	External elements	204

	Page
B.4 Interoperability with SGML applications	204
Annex C Bibliography	204
Figure 1 — Body layout types A, B and C with header and footer above and below the body area	22
Figure 2 — Header and footer frame layouts corresponding to Figure 1	23
Figure 3 — Body layout types A, B and C with header and footer to the right and left of the body area	23
Figure 4 — Header and footer frame layouts corresponding to Figure 3	24
Figure 5 — Example of body area layout for body layout type A	26
Figure 6 — Example of body area layout for body layout type B	27
Figure 7 — Example of body area layout for body layout type C	28
Figure 8 — Example of the layout of a snaking columns frame	29
Figure 9 — Another example of the layout of a snaking columns frame	30
Figure 10 — Example of the layout of a synchronized column	31
Figure 11 — Example of a CompositeFloat frame	33
Figure 12 — Example of a CompositeFixtureVariable frame	35
Figure 13 — Example of the layout of a form	36
Figure 14 — Example of a <i>simple</i> table	38
Figure 15 — Example of a <i>complex</i> table	39
Figure 16 — Frames used to represent table labels	40
Figure 17 — Frames used to represent table rows	41
Figure 18 — Specification of margins	52
Figure 19 — Code extension features (basic case)	56
Figure 20 — Code extension features (all possible cases)	57
Figure 21 — DocumentLogicalRoot, 1st level	77
Figure 22 — DocumentLogicalRoot, 2nd level	77
Figure 23 — Title, Paragraph and ListTerm	78
Figure 24 — Phrase, Caption and Description	78
Figure 25 — Footnote	79
Figure 26 — Figure	79
Figure 27 — Form	80
Figure 28 — Table	80
Figure 29 — Reference	81
Figure 30 — NumberedList	81
Figure 31 — UnNumberedList	81
Figure 32 — DefinitionList	82
Figure 33 — CommonContent	82
Figure 34 — DocumentLayoutRoot	82
Figure 35 — Page, RectoPage and VersoPage	83
Figure 36 — CompositeHeader, CompositeFooter and CompositeCommon	83
Figure 37 — FixedCompositeBody	84
Figure 38 — VariableCompositeBody	84
Figure 39 — SnakingColumns	85
Figure 40 — SynchronizedColumns	85
Figure 41 — CompositeFloat	85
Figure 42 — CompositeColumnVariable and CompositeColumnFixed	86
Figure 43 — CompositeFixtureVariable and CompositeFixtureFixed	86
Figure 44 — FormArea and EntryGroup	87

	Page
Figure 45 — TableArea	87
Table 1 — Nominal page sizes	21
Table 2 — Layout of figure	48
Table 3 — Layout of tables	49
Table 4 — Layout of forms	50
Table 5 — List of number string identifiers for logical constituent constraints	66
Table 6 — List of number string identifiers for layout constituent constraints	67
Table 7 — Symbols to denote shift functions	88

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 11182-1 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 11182 consists of the following part under the general title *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).*

Further parts of this International Standardized Profile may be developed.

Annex A forms an integral part of this part of ISO/IEC ISP 11182. Annex B and Annex 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 rules 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 Implementors’ 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 is to facilitate the interworking of applications interchanging documents based on ODA, CCITT Rec. T.410 series | ISO 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 Amendment 1 — Annex F of CCITT Rec. T.411 | ISO 8613-1.

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 International Standardized Profile.

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 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: A Document Application Profile Proforma and Notation*.
- ISO 8613-1 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 Notation¹⁾*.
- 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 Architecture*.
- 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*.
- 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)*.

¹⁾ To be published.

ISO/IEC 8825:1990, *Information technology — Open Systems Interconnection — Specification of Basic Encoding Rules for Abstract Syntax Notation One (ASN.1)*.

— CCITT Recommendation T.502, *Document Application Profile PM-11 for the Interchange of Character Content Documents in Processable and Formatted Forms*²⁾.

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.505, *Document Application Profile PM-26 for the Interchange of Mixed Content Documents in Processable and Formatted Forms*²⁾.

ISO/IEC ISP 11181-1:1993, *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)*.

NOTE In the titles of the various parts of ISO 8613 the term “Office” occurs. ISO has decided that this shall 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.

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)*.

— 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*.

²⁾ To be published. The title of this recommendation may be changed.

- ISO/IEC TR 10000-2:1992, *Information technology — Framework and taxonomy of International Standardized Profiles — Part 2: Taxonomy of OSI Profiles*³⁾.
- CCITT Recommendation T.506, *Document Application Profile PM-36 for the Interchange of Mixed Content Documents in Processable and Formatted Forms*⁴⁾.

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 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 FOD26.

The features supported by this International Standardized Profile are a superset of the features supported by the International Standardized Profiles FOD11 and FOD26 and thus all data streams that are conformant to FOD11 and FOD26 are also conformant to this International Standardized Profile apart from the object identifier of the document application profile.

NOTE This International Standardized Profile (FOD36) is identical to the CCITT Recommendation T.506 (PM-36)(to be published), except that PM-36 only defines the use of the Office Document Interchange Format (ODIF).

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.X.209 | 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;

³⁾ Taxonomy of this document is based on the first edition of ISO/IEC TR 10000-2 (published 1990). In the present edition of ISO/IEC TR 10000-2 (published 1992), the profile identifiers of FOD11, FOD26 and FOD36 were changed to FOD011, FOD026 and FOD036.

⁴⁾ To be published. The title of this recommendation may be changed.

- 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 Office Document Interchange Format (ODIF), or any data streams conforming to this International Standardized Profile structured in accordance with Office Document Language (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 and FOD26.

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-document;
- 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 logical constituent constraints 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 an arbitrary ordered sequence of one or more constituent constraints of the types *Passage* and *NumberedSegment*.

The automatic numbering schemes that apply to constituent constraints of the types *NumberedSegment*, *Figure*, *Table*, *NumberedList* and *Footnote* may be initialized on the *DocumentLogicalRoot*.

6.2.3.2 Passage

Passage is a constituent constraint that represents a subdivision of a document that corresponds to a logical grouping of subordinate parts of a document. This grouping may be regarded as a logical entity for reading, or it may represent parts of a document 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, consisting of the table of contents or foreword;
- the main matter of the document;
- the back matter, consisting of appendices, glossary or index;
- an illustration with associated text which is inserted as a distinct entity within a section of a document.

The automatic numbering schemes that apply to subordinate constituent constraints of the types *NumberedSegment*, *Figure*, *Table*, *NumberedList* and *Footnote* may be initialized on a *Passage*.

The immediate subordinates of a *Passage* consist of an optional constituent constraint of the type *Title* which is followed by an arbitrary ordered sequence of one or more of the following constituent constraints:

- *Paragraph*;
- *BodyGeometric*;
- *BodyRaster*;
- *BodyText*;
- *Figure*;

- Table;
- NumberedList;
- UnNumberedList;
- DefinitionList;
- NumberedSegment;
- Passage.

A Passage shall have at least one of the above constituent constraints as a subordinate.

Therefore, a Passage may itself contain one or more subordinates of the type Passage, and thus Passages may be nested to any number of levels. This allows logical entities within a document to be described in terms of their component logical entities. Also, a NumberedSegment may contain one or more subordinate Passages which allows different logical entities within a NumberedSegment to be distinguished.

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 begins on a new page set.

6.2.3.3 *NumberedSegment*

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

The automatic numbering schemes that apply to the subordinate constituent constraints NumberedSegment, Figure, Table, NumberedList and Footnote may be initialized on any logical object or logical object class, typically on Passage or a NumberedSegment.

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 optionally by a constituent constraint Title which in turn is followed by an optional arbitrary ordered sequence of one or more of the following constituent constraints:

- Paragraph;
- BodyGeometric;
- BodyRaster;
- BodyText;
- Figure;
- Table;
- NumberedList;
- UnNumberedList;
- DefinitionList;
- Passage;
- NumberedSegment.

The above indicates that a document may contain any number of nested levels of the constituent constraint NumberedSegment.

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 class 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 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-initialization of the numbering is possible;
- it is not possible to continue the numbering across Passages.

6.2.3.4 *Number*

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

A *Number* is a basic logical constituent constraint 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.

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*;
- *Phrase*;
- *Reference*.

The character content associated with a *Paragraph* may be concatenated to form a continuous stream of character content which may contain single or multiple references to footnotes or other parts of the document, and may be laid out as single unit. Entities within this content which have particular logical significance or presentation features may be distinguished using the constituent constraint *Phrase*.

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.

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 act as carriers for the document content and may be specified as subordinates of constituent constraints:

- *Passage*;
- *NumberedSegment*;
- *Paragraph*;
- *Title*;
- *ListTerm*;
- *Artwork*;
- *Phrase*;

- Reference;
- UnNumberedList.

In addition, BodyText may be specified as a subordinate to Phrase, Caption and Description. These constituent constraints allow the layout and presentation requirements of different parts of the content of a document to be specified.

These are basic logical constituent constraints that directly refer to content portions that contain character, raster graphics and geometric graphics content respectively. BodyText shall refer to one or more content portions which may contain either 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.

Constituent constraints 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.

Where the superior constituent constraint referenced is subordinate to a FootnoteBody, it is required to specify one of the layout category names for this constituent constraint, "Footnote" or "Footnote-<n>". This along with a "permitted categories" attribute of the same name on the footnote frame will ensure that a logical object from this constituent constraint is laid out in a FootnoteArea frame when generic layout structure is specified within the document.

6.2.3.7 Footnote

Footnote is a constituent constraint that is used to represent footnotes within a document. This constituent constraint may be specified as a subordinate to:

- Paragraph;
- Title;
- ListTerm;
- Phrase;
- Caption;
- Description.

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

FootnoteReference is 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 numbers 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 6.6.6.7.

6.2.3.9 *FootnoteBody*

FootnoteBody is a constituent constraint which represents the content of a footnote. The content consists of a stream of character content which may contain embedded references to other parts of the document. *FootnoteBody* is a composite logical constituent constraint whose subordinates consist of the constituent constraint *FootnoteNumber*, which is mandatory and represents the footnote identifier, followed by a sequence of one or more constituent constraints of the type *FootnoteText* and *Reference* which represent 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 constituent constraints 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 constituent constraints 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 is a basic logical constituent constraint 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 one of the layout category names for this constituent constraint, “Footnote” or “Footnote-<n>”. This along with a “permitted categories” attribute of the same name on the footnote frame will ensure that a logical object from this constituent constraint 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 is a basic logical constituent constraint that references one or more content portions each containing processable, formatted or formatted processable character content.

It is required to specify one of the layout category names for this constituent constraint, “Footnote” or “Footnote-<n>”. This along with a “permitted categories” attribute of the same name on the footnote frame will ensure that a logical object from this constituent constraint is laid out in a *FootnoteArea* frame when generic layout structure is specified within the document.

6.2.3.12 *Title*

Title is a constituent constraint that is used to represent the title, heading or name of the Passage or NumberedSegment to which it is an immediate subordinate. This constituent constraint may consist of character, raster graphics and geometric graphics content.

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

- *BodyText*;
- *BodyRaster*;
- *BodyGeometric*;
- *Phrase*;
- *Reference*;
- *Footnote*.

The character content associated with a Title may be concatenated to form a continuous stream of character content which may contain single or multiple references to footnotes or other parts of the document, and may be laid out as single unit. Entities within this content which have particular logical significance or presentation features may be distinguished using the constituent constraint Phrase. 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 Concatentation (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, whereas the use of Separation avoids this.

6.2.3.13 Phrase

Phrase is a constituent constraint that is used to group together an amount of character content that represents a single logical entity that needs to be distinguished for some purpose. That is, the content represented by a Phrase may have a particular logical significance, or it may have certain layout or presentation requirements. The character content may contain embedded footnotes and references to other parts of the document content. A typical example is a quotation that is to be reproduced in italics.

A Phrase may be used as subdivision of constituent constraints of the types Paragraph, Title, Caption, Description, Artwork and ListItem. Also, a Phrase may be subordinate to another Phrase and, therefore, constituent constraints of the type Phrase may be nested.

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

- BodyText;
- Phrase;
- Reference;
- Footnote.

The character content associated with a Phrase may be concatenated to form a continuous stream of character content which may contain single or multiple references to footnotes or other parts of the document, and may be laid out as single unit. Alternatively, the character content may contain hard line terminators, which will cause parts of the content to be separated when laid out.

6.2.3.14 Constituents that provide for a general referencing mechanism

This subclause defines the constituent constraints that are provided to support a general referencing mechanism in a document.

6.2.3.14.1 Reference

A *Reference* is a constituent constraint which represents a reference consisting of character content that is derived either fully or partially from another part of the document. This constituent constraint provides a general cross-referencing mechanism in a document.

This constituent constraint may be specified as a subordinate to constituent constraints of the types:

- Paragraph;
- Title;
- ListTerm;
- Phrase;
- Caption;
- Description;
- FootnoteBody.

It is a composite constituent constraint whose immediate subordinates may consist of an ordered sequence of constituent constraints of the types BodyText, ReferencedContent and BodyText.

The general format of the content associated with a constituent constraint of the type Reference is:

[<prefix-string>]<reference-string>[<suffix-string>]

The prefix and suffix strings are optional and, if required, are represented by constituent constraints of the type BodyText. The reference string is represented by the constituent constraint, ReferencedContent.

A reference string may, for example, contain references to identifiers such as a number that distinguishes a chapter or section, a table, a figure, a footnote, an item in a numbered list or a page number. Each reference may contain multiple, concatenated references to different parts of a document; a typical example is the reference “see Table 3 in chapter 2 on page 4” where the values “3”, “2” and “4” are derived automatically from the appropriate table, chapter and page in the document.

6.2.3.14.2 ReferencedContent

ReferencedContent is a constituent constraint that represents a character string that contains a single reference to content in other parts of the document (see **6.2.3.14.1**).

It is a basic logical constituent constraint that is an immediate subordinate to the constituent constraint *Reference*. It contains a content generator which, when evaluated, produces the character string containing the referenced content.

A sequence of two or more constituent constraints of this type may be used to represent a composite reference string such as “see Table 2 in section 3.1 beginning on page 6”, where the strings “2”, “3.1” and “6” are automatically generated by referring to number strings attached to particular parts of the document.

The format of this content generator and its evaluation is described in **6.6.6.9**.

Where the superior constituent constraint referenced is subordinate to a *FootnoteBody*, it is required to specify one of the layout category names for this constituent constraint, “Footnote” or “footnote-*<n>*”. This along with a “permitted categories” attribute of the same name on the footnote frame will ensure that a logical object from this constituent constraint is laid out in a *FootnoteArea* frame when generic layout structure is specified within the document.

6.2.3.15 Constituents representing illustrations

6.2.3.15.1 Introduction

This International Standardized Profile supports the representation of illustrations or figures consisting of artwork or simple forms.

Artwork typically consists of a diagram or figure which is an image composed of a single content type or which is formed by overlaying two or more separate images consisting of character, raster graphics or geometric graphics content.

A form typically consists of a collection of logical entities, each of which has a certain logical significance. Each logical entity may be further subdivided into subordinate logical entities. A typical example is an order form consisting of a reference number, information about the originator, including the name, address and telephone number, a list of items required and their expected delivery dates.

The entities which make up a form are intended to be laid out in a designated area which is subdivided into areas that are specially reserved for each particular type of entity. This designated area is specified by the frame *FormArea* when the generic layout structure is present in the document.

Optionally, an illustration may also contain an identifier which may be used to distinguish the illustration from other parts of the document, a caption which may be used to identify the purpose of the illustration, and an amount of associated descriptive text.

Further information concerning the layout of illustrations consisting of artwork or forms is given in **6.4.1.3.6**.

The constituent constraints used to represent illustrations are defined below.

6.2.3.15.2 Figure

Figure is a constituent constraint that is typically used to represent an illustration. Such an illustration may consist of artwork or a form as described in **6.2.3.15.1**.

A *Figure* is a composite logical constituent constraint which may be specified as a subordinate to a *Passage* or *NumberedSegment*.

The subordinates of a *Figure* may consist of a sequence, in any order, of the following constituent constraints:

- either *Artwork* or *Form*;
- *Number*;
- *Caption*;

— Description.

A constituent constraint of the type Artwork or Form shall always be present as a subordinate to the constituent constraint Figure. Constituent constraints of the types Number, Caption and Description are independently optional. The constituent constraints may occur in any order, except that a Number and a Caption shall be in this order.

6.2.3.15.3 Artwork

Artwork is a constituent constraint that is typically used to represent a graphical image within an illustration or figure. This may be a simple image that is represented by character, raster graphics or geometric graphics content, or it may be a composite image consisting of a combination of these content types.

This constituent constraint shall only occur as a subordinate to a constituent constraint of the type Figure. The immediate subordinates of Artwork consist of an arbitrary ordered sequence of one or more of the following constituent constraints:

- Phrase;
- BodyRaster;
- BodyGeometric.

The constituent constraint Phrase is used to enable the image to contain character content which contains references to footnotes or other parts of the document.

6.2.3.15.4 Form

Form is a constituent constraint that is used to represent a simple form within an illustration or figure. It is a composite logical constituent constraint which shall only occur as a subordinate to a constituent constraint of the type Figure.

A Form consists of an arbitrary order sequence of basic entries and composite entries. Basic entries, which may consist of character, raster graphics or geometric graphics content, are represented by subordinate constituent constraints of the type EntryElement. Composite entries are represented by subordinate constituent constraints of the type EntryGroup.

Constituent constraints of the type EntryGroup may be further subdivided into a set of one or more basic or composite entries. Thus a composite entry may be nested to any number of levels such that each level may consist of a set of subordinate basic and composite entries.

This International Standardized Profile does not define a mechanism for defining the semantics associated with each basic or composite entry. This can be achieved by individual applications by making use of the parameter “external data” in the attribute “application comments” (see 6.6.4).

6.2.3.15.5 Caption

Caption is a constituent constraint that typically represents a title or header that is associated with an illustration or figure. This constituent constraint represents character content that may contain embedded references to footnotes and to other content within the document.

A Caption is a composite logical constituent constraint which shall only occur as a subordinate to a constituent constraint of the type Figure.

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

- BodyText;
- Reference;
- Phrase;
- Footnote.

The above constituent constraints may be concatenated to form a continuous stream of character content which is to be laid out as a single unit.

6.2.3.15.6 Description

Description is a constituent constraint that typically represents some general supplementary information that forms part of an illustration that contains a figure or form.

This constituent constraint is a composite logical constituent constraint which shall only occur as a subordinate to a constituent constraint of the type Figure.

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

- BodyText;
- Reference;
- Phrase;
- Footnote.

The above constituent constraints may be concatenated to form a continuous stream of character content which is to be laid out as a single unit.

6.2.3.15.7 EntryGroup

EntryGroup is a constituent constraint that represents a composite logical entry within a form.

It is a composite logical constituent constraint which shall only occur as a subordinate to a constituent constraint of the types Form or EntryGroup.

The immediate subordinates of a EntryGroup consist of an arbitrary ordered sequence of one or more constituent constraints of the types EntryElement and EntryGroup. Therefore, the subordinates to EntryGroup may be nested to any number of levels.

NOTE Constituent constraint EntryElement is defined in 6.2.3.16.6.

6.2.3.16 Constituents used to represent tables

6.2.3.16.1 Introduction

For the purpose of this International Standardized Profile, a table is a logical entity that consists of an ordered sequence of elements, called cells, that are arranged into a two dimensional array of rows and columns.

Each row may consist of a simple row which contains a sequence of one or more cells. Alternatively, a row may consist of a composite row which contains a single cell followed by a sequence of one or more subrows, each of which contains a sequence of cells.

The subclauses below define the logical constituents used to represent tables. Figure 28 illustrates the structural relationships between the constituents used to represent a table. Subclause 6.4.1.3.7 describes how tables are intended to be laid out.

6.2.3.16.2 Table

Table is a logical constituent constraint that represents a table as a whole. This constituent constraint may be specified as a subordinate to constituent constraints Passage and NumberedSegment.

The immediate subordinates of this constituent constraint consist of a sequence of constituent constraints Row. Each row may or may not have the same characteristics with regard to its sub-structure.

6.2.3.16.3 Row

Row is a constituent constraint that is a subordinate of the constituent constraint Table and represents a row of entries in a table. This may consist of a sequence of entries, or it may consist of a single entry followed by a sequence of subrows of entries.

In order to represent these two cases, the immediate subordinates of a Row may consist of either:

- a sequence of constituent constraints EntryElement, or;
- a single constituent constraint EntryElement, followed by a single constituent constraint of the type TableComponent.

6.2.3.16.4 TableComponent

TableComponent is a constituent constraint that is a subordinate of the constituent constraint, Row, and represents a sequence of one or more subrows of entries within a row of a table.

The immediate subordinates of this constituent constraint consist of a sequence of one or more constituent constraints of the type RowComponent. Each RowComponent shall have the same characteristics with regard to its sub-structure.

6.2.3.16.5 *RowComponent*

RowComponent is a constituent constraint that is a subordinate of the constituent constraint *TableComponent* and represents a sub-row of entries in a row of entries in a table.

The immediate subordinates of this constituent constraint consist of a sequence of one or more components of the type *EntryElement*. Each *EntryElement* may or may not have the same characteristics with regard to its sub-structure.

6.2.3.16.6 *EntryElement*

EntryElement is a constituent constraint that represents a single entry in a form or table. It is a sub-division of constituent constraints of the types *Table* and *Form*. In the case of *Table*, it is specified as a subordinate to a *Row* or a *RowComponent* and represents a single entry in a table. In the case of *Form*, it is specified as a subordinate of a *Form* itself or of an *EntryGroup*.

Each entry in a table or form may consist of character, raster graphics or geometric graphics content and, hence, *EntryElement* has a single immediate subordinate constituent constraint of the type *EntryText*, *EntryRaster* or *EntryGeometric*.

6.2.3.16.7 *EntryText*, *EntryRaster* and *EntryGeometric*

EntryText, *EntryRaster* and *EntryGeometric* are constituent constraints which represent content that is to be entered into tables and forms. These constituent constraints may be specified as subordinates of the constituent constraint *EntryElement* and allow the layout and presentation requirements for the content allocated to tables and forms to be specified.

These are basic logical constituent constraints that directly refer to content portions that contain character, raster graphics and geometric graphics content respectively. *EntryText* shall refer to one or more content portions which may contain either processable, formatted or formatted processable character content.

EntryRaster and *EntryGeometric* shall only refer to a single content portion containing formatted processable raster graphics content or formatted processable geometric graphics content respectively.

Constituent constraints of these types in the generic logical structure may refer to generic content. This provides the means of defining common content within tables and forms.

6.2.3.17 *Constituents representing lists*

6.2.3.17.1 *Introduction*

This International Standardized Profile supports the representation of three types of lists, as follows:

- numbered lists consisting of ordered lists of items, each of which is preceded by an identifier such as an alphabetic character or numeral;
- unnumbered lists consisting of unordered lists of items, each of which may optionally be preceded by a separator such as a hyphen, bullet or small circle;
- definition lists consisting of lists of ordered pairs of items such as a term and its corresponding definition.

Each type of list may be nested without restriction, and one particular type of list may be composed of lists of other types. For example, an item in a numbered list can consist of a subordinate numbered list, unnumbered list or definition list.

The constituent constraints that may be used to represent these list types are defined below.

6.2.3.17.2 *NumberedList*

A *NumberedList* is a constituent constraint that is used to represent a list of items, each of which is preceded by an identifier that serves to distinguish that item.

This constituent constraint may be specified as an immediate subordinate to a *Passage*, *NumberedSegment*, or *ListItem*.

The immediate subordinates of a *NumberedList* consist of the constituent constraint *Number* which contains a content generator that generates the identifier corresponding to each item in the list, followed by the constituent constraint, *ListItem*. This pair of constituent constraints may be repeated without limitation.

Further information concerning the numbering of items in a list is contained in 6.6.6.5.

6.2.3.17.3 *UnNumberedList*

UnNumberedList is a constituent constraint that is used to represent a list of items, each of which may be preceded by an optional separator consisting of character, raster graphics or geometric graphics content. This constituent constraint may be specified as an immediate subordinate to a Passage, NumberedSegment, or ListItem.

The immediate subordinates of an *UnNumberedList* consist of a separator, which is represented by a constituent constraint of the type BodyText, BodyRaster or BodyGeometric, followed by the constituent constraint ListItem. This pair of constituent constraints may be repeated without limitation.

6.2.3.17.4 *DefinitionList*

DefinitionList is a constituent constraint that represents a sequence of ordered pairs of items.

This constituent constraint may be specified as an immediate subordinate to a Passage, NumberedSegment, or ListItem.

The immediate subordinates of this item consist of the constituent constraint ListTerm, followed by the constituent constraint ListItem. This pair of constituent constraints may be repeated without limitation.

6.2.3.17.5 *ListItem*

ListItem is a constituent constraint that represents an item within a NumberedList, UnNumberedList or DefinitionList. That is, this constituent constraint represents the second element of each pair of elements within a numbered, unnumbered or definition list.

The immediate subordinates of this constituent constraint may consist of a sequence of one or more constituent constraints Phrase, or one of the constituent constraints NumberedList, UnNumberedList or DefinitionList.

Thus this constituent constraint represents an amount of character content that may contain embedded references to footnote and other parts of the document, or it represents a subordinate list of items.

6.2.3.17.6 *ListTerm*

ListTerm is a constituent constraint that represents a term element within a DefinitionList. A term element is the first item of each pair of items that constitutes a definition list.

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

- BodyText;
- BodyRaster;
- BodyGeometric;
- Phrase;
- Reference;
- Footnote.

Thus this constituent constraint represents an amount of character, raster graphics and geometric graphics content. Character content may contain embedded phrases, references and footnotes.

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 laid out in the body, header, or footer area 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 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;
- CommonReference;

- *CurrentInstance*;
- *TableNumber*;
- *CommonNumber*.

When the generic layout structure is present, constituent constraints of the type *CommonContent* and their associated subordinate constituent constraints are constrained to be laid out in a specified frame within a body, header or footer area 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 a specified area within a page.

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 a specified area in a page. This constituent constraint is specifically used when it is required to represent 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. For example, 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. The page number that is generated may have a prefix or suffix character string.

The format of the content generators is defined in 6.6.6.8.

6.2.4.4 *CommonRaster*

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

CommonRaster is a constituent constraint for 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 a specified area within a page. For example, this constituent 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.

6.2.4.6 *CommonReference*

CommonReference is a constituent constraint that represents the common character content that is to be laid out in a specified area in a page and which represents a character string that contains references to other parts of the document. Such a reference may consist of a reference to a number that relates to a segment, table, figure, footnote or page number.

CommonReference is a constituent constraint for a basic logical object class that contains a content generator which, when evaluated, produces a character string containing the references. The format of this reference string is defined in 6.6.6.10.

6.2.4.7 *CurrentInstance*

CurrentInstance is a constituent constraint that represents the common character content that is to be laid out in a specified area in a page. This constituent constraint is specifically used when it is required to refer to a character string that is attached to any logical constituent constraint in the document or any of the layout constituent constraints *DocumentLayoutRoot*, *PageSet*, *RectoPage*, *VersoPage* or *Page*. The number string may represent, for example, the title of a sub-section or table that is contained elsewhere in the document.

CurrentInstance is a constituent constraint for a basic logical object class that contains a content generator which, when evaluated, produces a copy of the character string associated with a specified constituent. The format of this reference is defined in **6.6.6.11**.

6.2.4.8 TableNumber

TableNumber is a constituent constraint that represents the common character content that is to be laid out in a specified area in a page. This constituent constraint is specifically used when it is required to represent a table number which is to be placed within the header area of a table.

TableNumber is a constituent constraint for a basic logical object class that contains a content generator which, when evaluated, generates the required table number. The format of the content generator is defined in **6.6.6.6**.

6.2.4.9 CommonNumber

CommonNumber is a constituent constraint that represents the common character content that is to be laid out in a specified area in a page. This constituent constraint is specifically used when it is required to refer to character content consisting of a number string which can, for example, represent the number of a subsection or a current page.

CommonNumber is a constituent constraint for a basic logical object class that contains a content generator which, when evaluated, generates the reference to the required number. The format of this reference is defined in **6.6.6.12**.

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.

It should be noted that, in the case of FOD36, common content may also be laid out in the body area, as well as the header/footer area.

Three body layout types are provided by this International Standardized Profile. Each body layout type specifies how the body is positioned within each page, and how the content may be presented within the body. These are referred to as body layouts A, B, and C, and are defined in **6.3.4.5**.

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

Body 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 contents written in Latin based, Hebrew, Arabic, and Japanese (most cases) languages.

Body 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 contents written in Latin based, Hebrew, Arabic, and Japanese (most cases) languages in which it is required to layout the content in landscape orientation within the body area of the page.

Body layout C is used when the character content is to be laid out vertically and from right to left within the body area. This layout may be typically used for contents written in languages which use ideograms, such as Japanese and Chinese characters.

The body, header and footer areas may be further sub-divided into areas to support a wide range of different layout requirements. These features are described further in **6.3.5** and **6.3.6**.

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 constituent constraints of the type, *PageSet*. The numbering schemes for pages may be initialized 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*, the value of this parameter may be specified as “recto”, “verso” or “unspecified”. In the case of *RectoPage*, the value of this parameter may be specified as “recto” or “unspecified”; in the case of *VersoPage*, the value of this parameter may be specified as “verso” or “unspecified”. The values “recto” and “verso” of the “side of sheet” parameter of the “medium type” attribute are non-basic.

The pages that make up a page set consist of an arbitrary sequence of one or more of the constituent constraints *Page*, *RectoPage* and *VersoPage*.

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

- all the pages shall have the same dimensions, but may differ in orientation (see **6.3.4.2**);
- all the pages are to be laid out on the same size of presentation medium (see **6.3.4.3**);
- all the pages instantiated from a given page class shall have the same layout characteristics. That is, for a given page class, there is not a choice of layout characteristics. However, the layout characteristics of pages in a page set may or may not be the same.

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

6.3.4.2 *Page dimensions*

The dimensions of the pages may be specified as any value (in SMUs) that is equivalent to or less than ISO A0 or ANSI E paper sizes. The dimensions may be specified in portrait or landscape orientation. Japanese page sizes B4 and B5 are also supported, but the dimension 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 8613 (e.g., in CCITT Rec. T.412 | ISO 8613-2 clause 7) is called “ANSI A” in this International Standardized Profile to be consistent with the other reference 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 millimeters	Size in BMUs	ARA in BMUs
ISO A5	148 mm × 210 mm	7 015 × 9 920	not defined
ISO A4	210 mm × 297 mm	9 920 × 14 030	9 240 × 13 200
ISO A3	297 mm × 420 mm	14 030 × 19 840	13 200 × 18 480
ISO A2	420 mm × 594 mm	19 840 × 28 060	18 898 × 27 118
ISO A1	594 mm × 841 mm	28 060 × 39 680	26 173 × 37 843
ISO A0	841 mm × 1 189 mm	39 680 × 56 120	37 843 × 54 283
ANSI legal	8,5 in. × 14 in.	10 200 × 16 800	9 240 × 15 480
ANSI A	8,5 in. × 11 in.	10 200 × 13 200	9 240 × 12 400
ANSI B	11 in. × 17 in.	13 200 × 20 400	12 744 × 19 656
ANSI C	17 in. × 22 in.	20 400 × 26 400	19 500 × 25 800
ANSI D	22 in. × 34 in.	26 400 × 40 800	25 800 × 39 600
ANSI E	34 in. × 44 in.	40 800 × 52 800	39 600 × 52 200
ANSI F	28 in. × 40 in.	33 600 × 48 000	32 400 × 47 400
Japan-legal	257 mm × 364 mm	12 141 × 17 196	11 200 × 15 300
Japan-letter	182 mm × 257 mm	8 598 × 12 141	7 600 × 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.

6.3.4.5.2 Body area layout

Three types of layout of body area are defined:

- Body layout type A. In this case, the layout path in the body area is specified as 270 degrees.
- Body layout type B. In this case, the layout path in the body area is specified as 0 degrees.

— Body layout type C. In this case, the layout path in the body area is specified as 180 degrees.

6.3.4.5.3 Header/footer area layout

Four types of layout of header/footer area are defined:

- H/F layout A1. In this case, the layout path in the header and footer area is 270 degrees. If the header or footer area is composite, the layout paths in the lowest frames are 270 degrees;
- H/F layout A2. In this case, the layout path in the header and footer area is 0 degrees. If the header or footer area is composite, the layout paths in the lowest frames are 270 degrees;
- H/F layout B1. In this case, the layout path in the header and footer area is 180 degrees. If the header or footer area is composite, the layout paths in the lowest frames are 180 degrees;
- H/F layout B2. In this case, the layout path in the header and footer area is 270 degrees. If the header or footer area is composite, the layout paths in the lowest frames are 180 degrees.

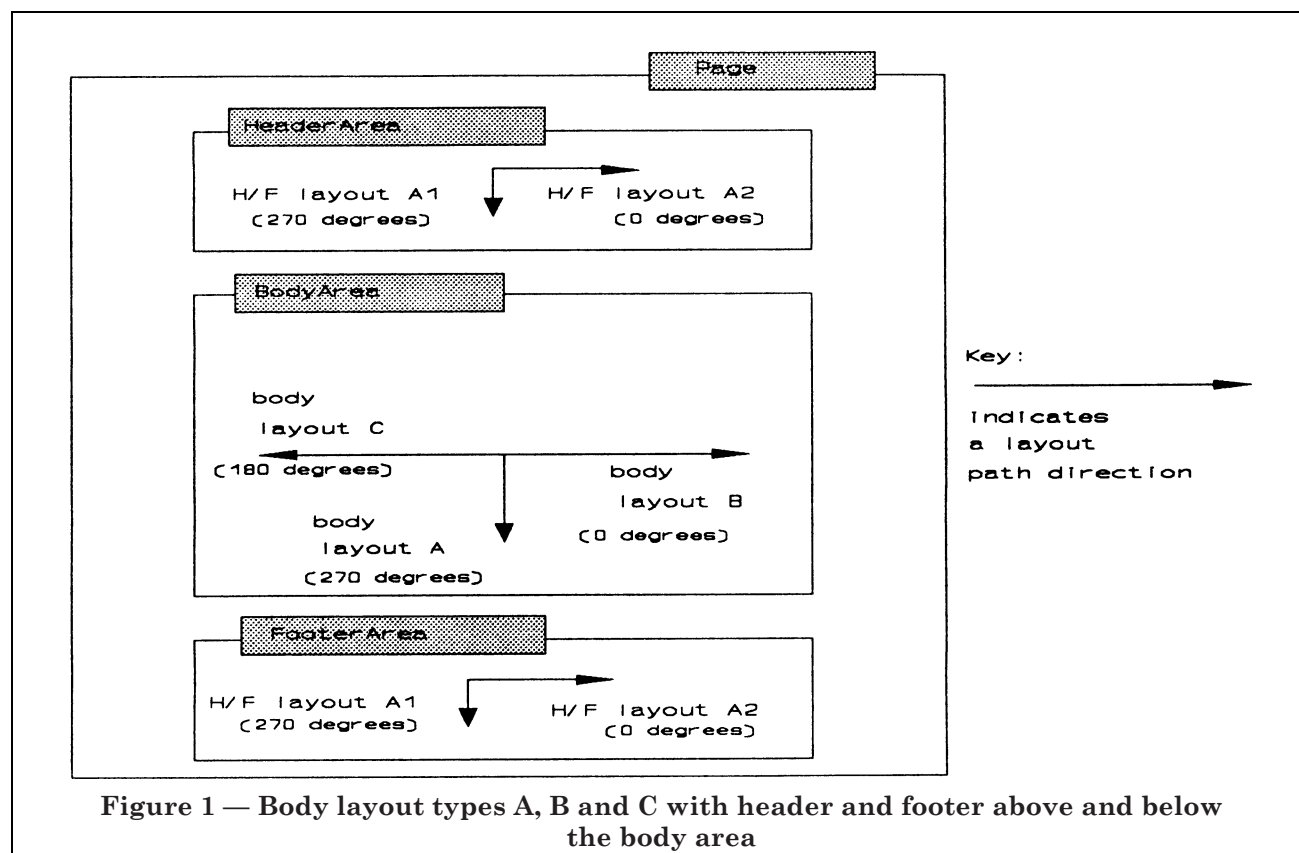
6.3.4.5.4 Page layout

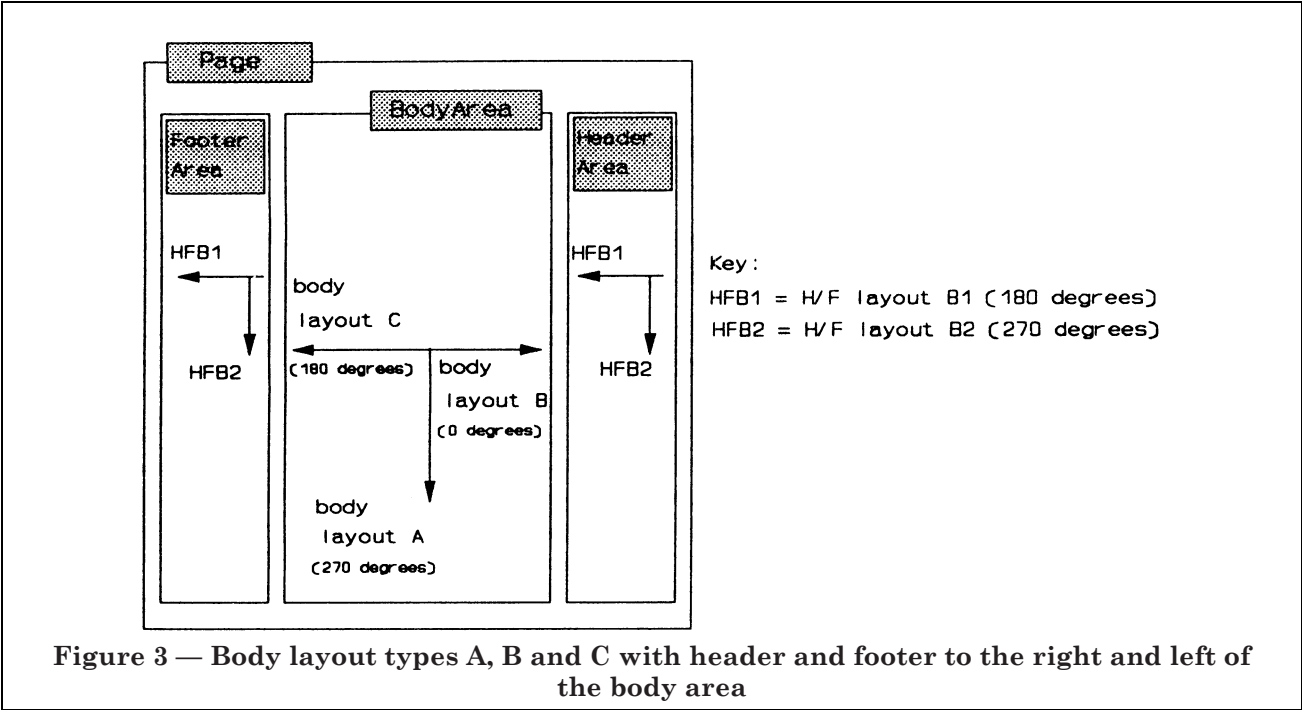
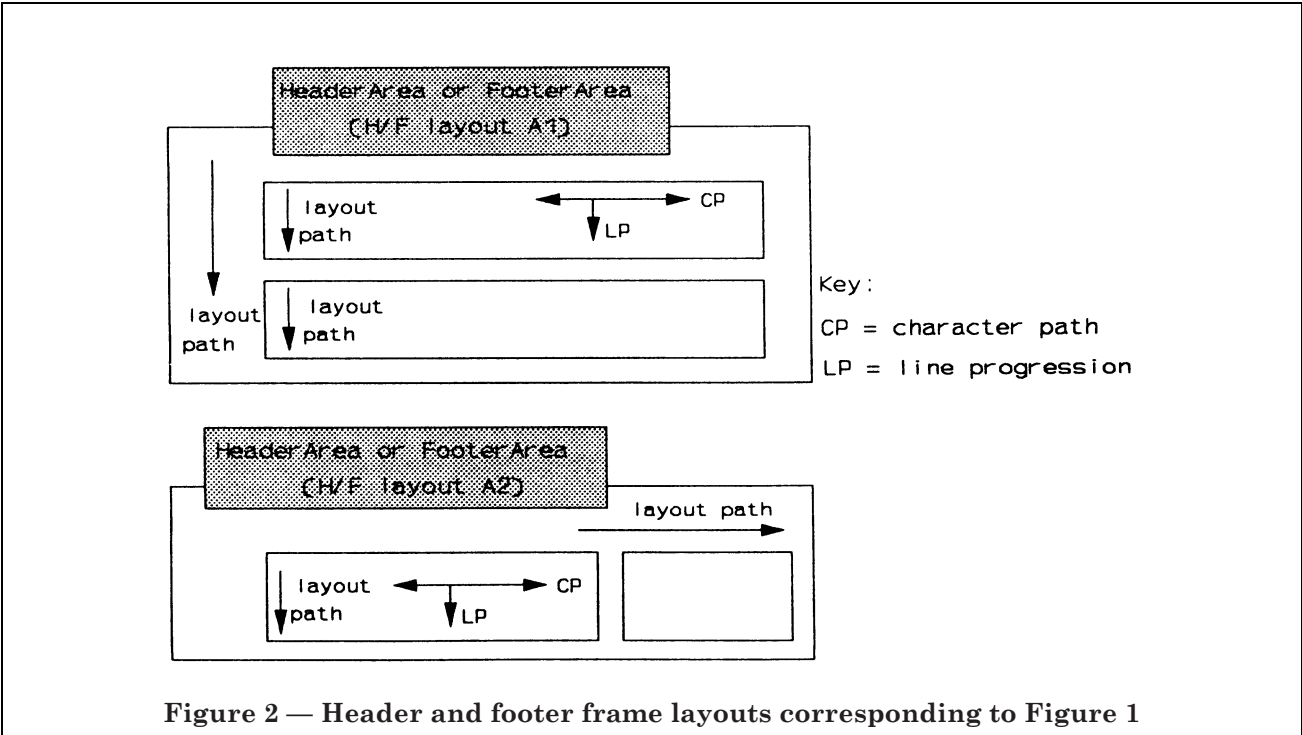
Page layout type is determined by a combination of the body layout type and the H/F layout type. Any combination is permitted. However, it is intended that all applications which use this International Standardized Profile shall support the combinations of body layout A and H/F layout A1 and A2, whereas support for other combinations may be specified as optional.

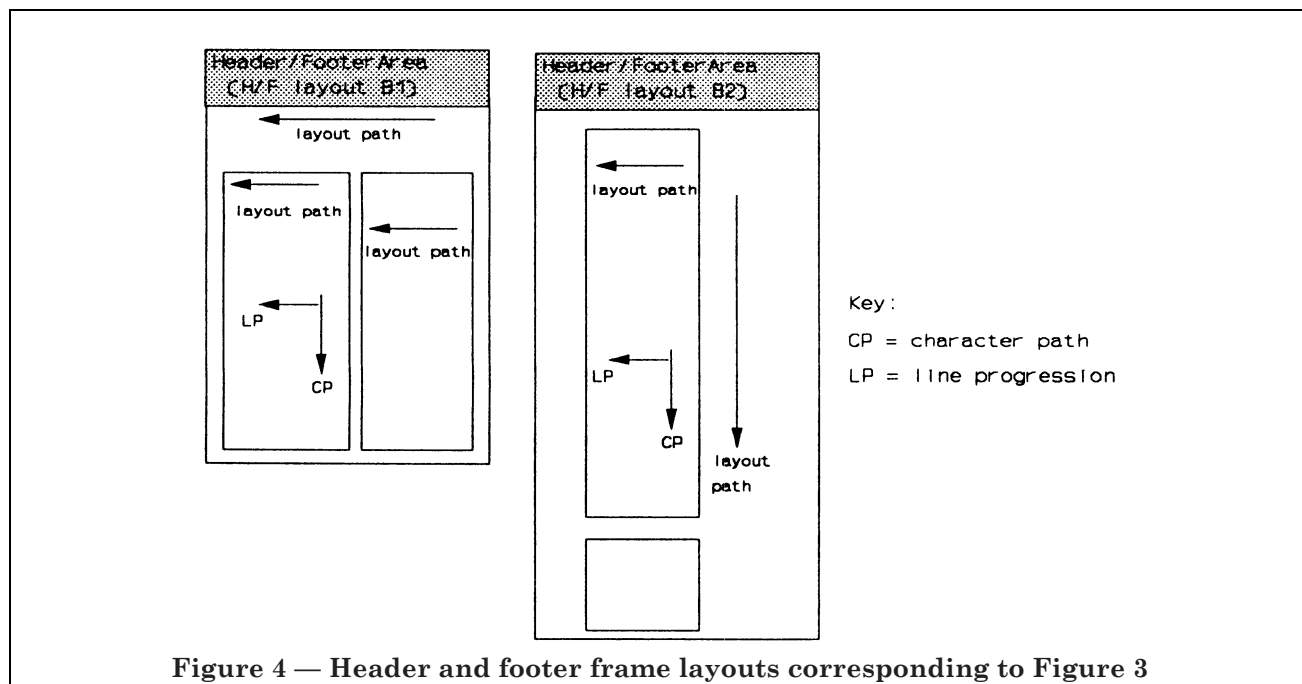
NOTE The combinations of body layout A and H/F layout A1, B and A1, C and A1, and C and B1 are equivalent to page layouts A, B, C and D respectively in FOD26.

The header and footer of H/F layout A1 or A2 are laid out above and below the body area. Figure 1 illustrates this case, and Figure 2 illustrates H/F layout type A1 and A2 corresponding to this case.

The header and footer of H/F layout B1 or B2 are laid out to the right and left of the body area. Figure 3 illustrates this case, and Figure 4 illustrates H/F layout type B1 and B2 corresponding to this case.







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 layout path specified in the body area determines the body layout type being used.

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 or FixedCompositeBody frame.

The subordinate areas within a VariableCompositeBody frame are represented by variably positioned frames. Thus the subordinate areas are not pre-determined and are automatically adjusted during the layout process to accommodate the content that is allocated to them.

When a FixedCompositeBody frame is used, the subordinate areas are represented by fixed positioned frames, and hence the body area layout can be precisely specified.

However, in order to provide both areas whose layout is fixed and areas whose layout is variable within a single body area, it is possible to specify one or more VariableCompositeBody frames as subordinates to a FixedCompositeBody frame. In this case, the layout paths for each of the VariableCompositeBody frames may or may not be the same. This allows, for example, text belonging to different languages to be laid out on the same page.

6.3.5.2 BasicBody

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

The position and dimensions of this frame are fixed. The layout path specified depends upon the body 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 represents either the entire body area or a part of it, and which contains one or more subordinate variably positioned frames. A VariableCompositeBody frame has a fixed position and fixed dimensions. The layout path specified for this frame depends upon the layout type used (see 6.3.1).

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

- BasicFloat;
- SnakingColumns;
- SynchronizedColumns;
- FootnoteArea;
- CompositeFloat;
- CompositeFixtureVariable;
- TableArea;
- ArrangedContentVariable;
- SourcedContentVariable.

The subordinate frames are all variably positioned and have variable dimensions except for ArrangedContentVariable frames. 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 degrees. This determines the body layout type used in the case where VariableCompositeBody represents the entire body area (see 6.3.4.5).

All immediate subordinate frames 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.

Also frames are constrained to have the same layout path as the VariableCompositeBody frame to which they are subordinate. However, exceptions to this rule are frames of the types CompositeFixtureVariable, CompositeFloat, SnakingColumns and SourcedContentVariable (see appropriate subclause below).

Figure 5, Figure 6 and Figure 7 provide illustrations of the layout of frames within a VariableCompositeBody frame for the various body 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, therefore, the layout characteristics of the body areas within a page set may change from page to page within a page set.

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.

This is a variably positioned frame which may be specified as a subordinate to frames of the types VariableCompositeBody, CompositeColumnVariable, CompositeColumnFixed, CompositeFixtureVariable and CompositeFixtureFixed.

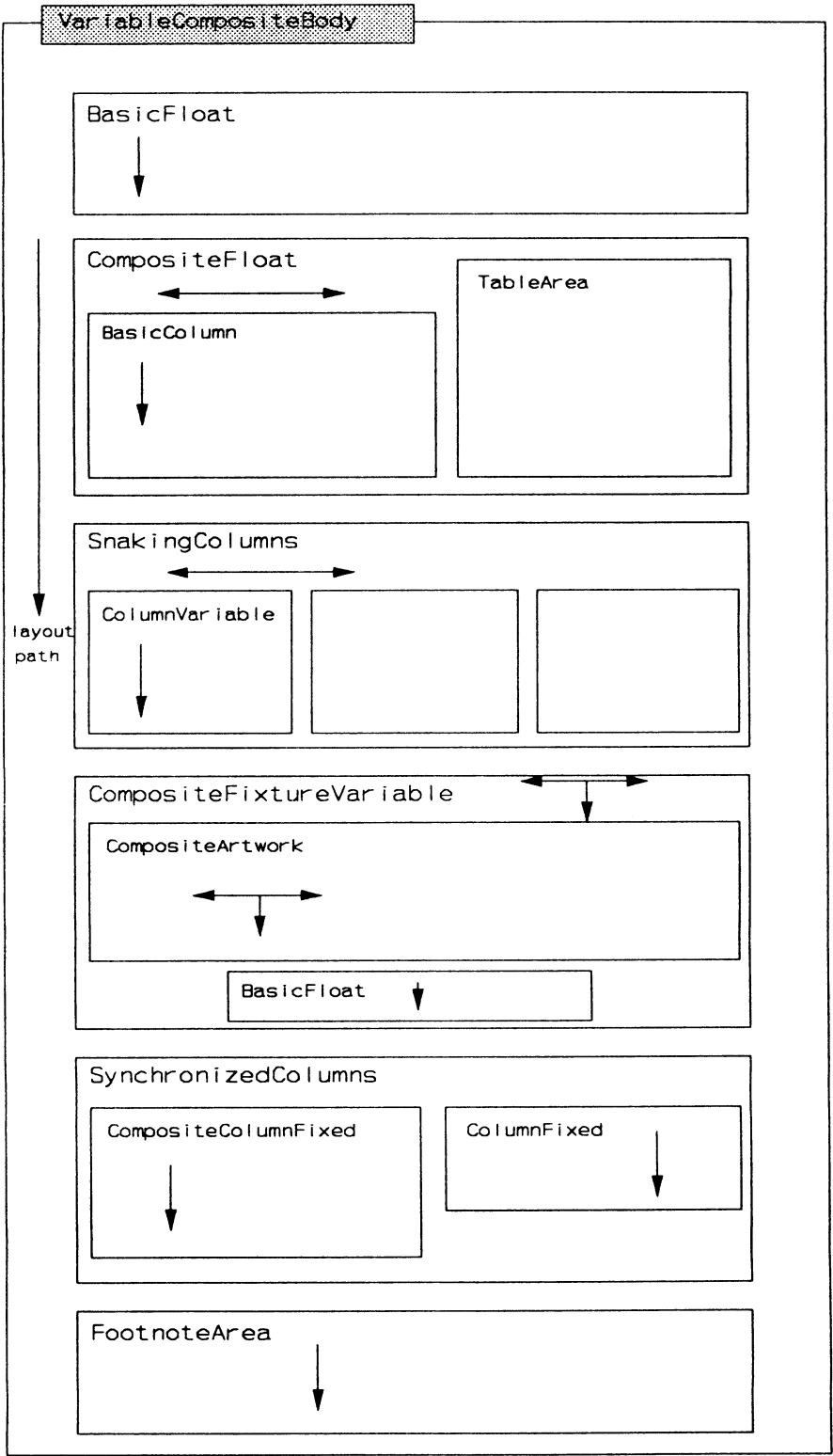


Figure 5 — Example of body area layout for body layout type A

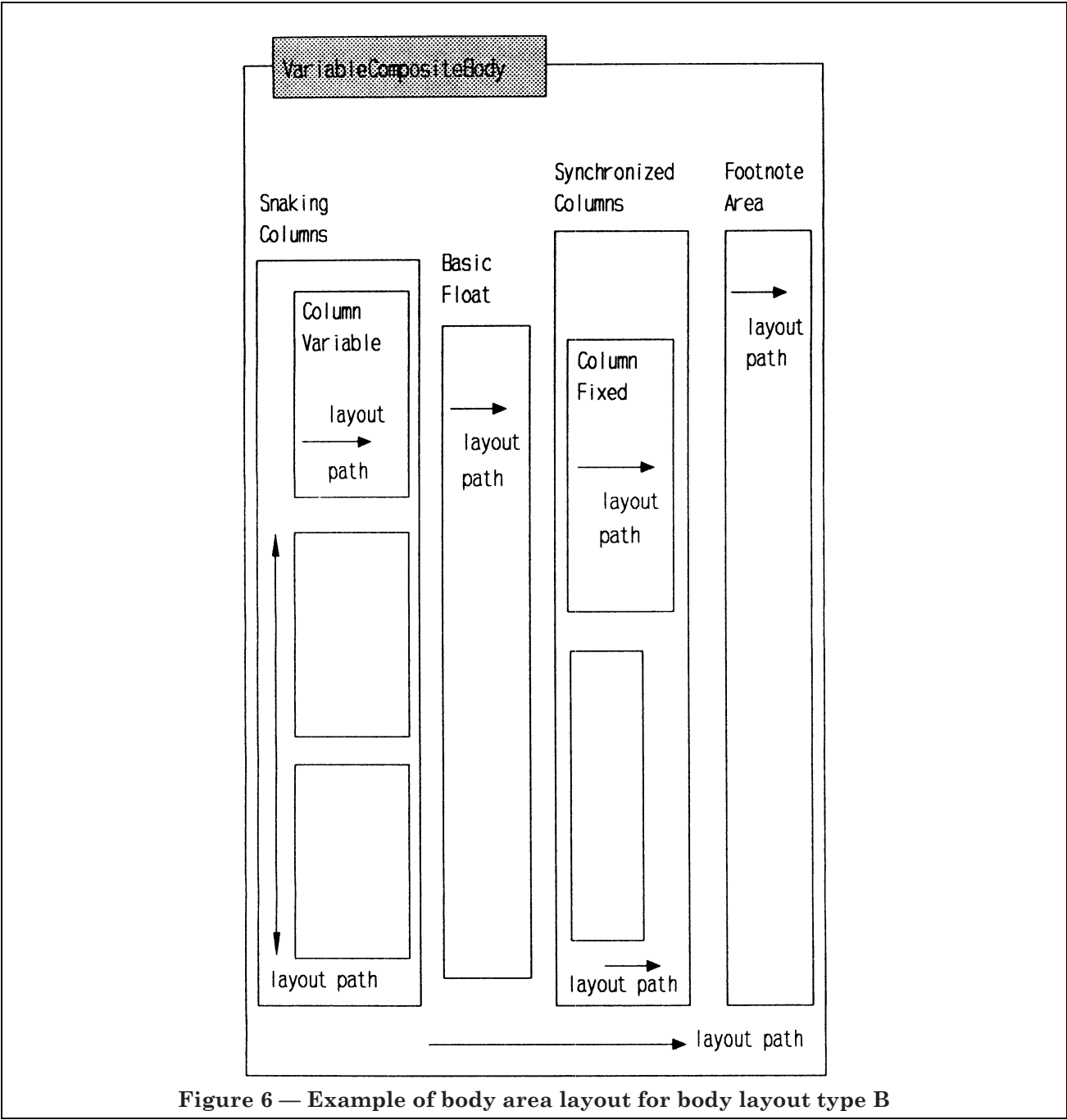


Figure 6 — Example of body area layout for body layout type B

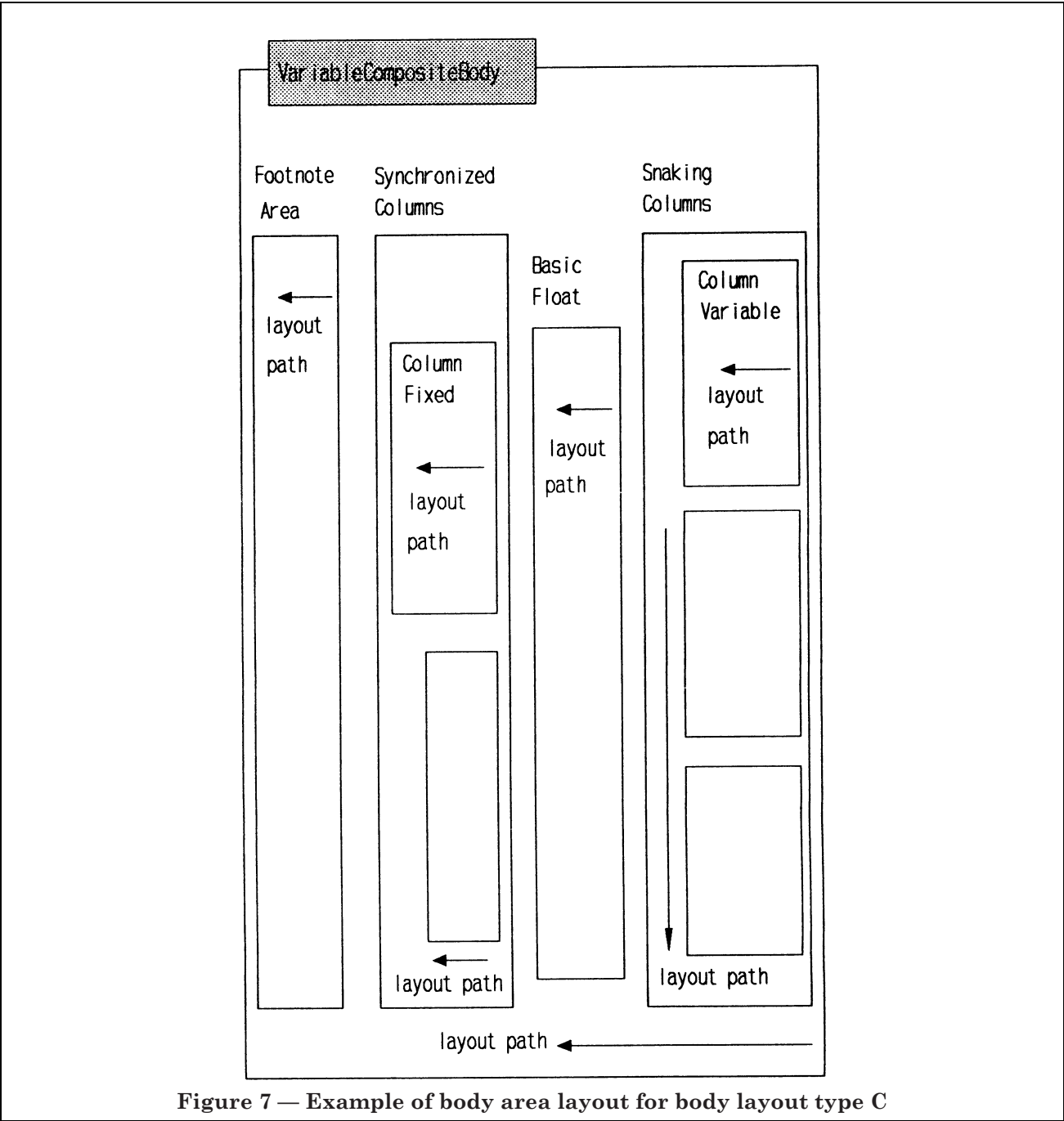


Figure 7 — Example of body area layout for body layout type C

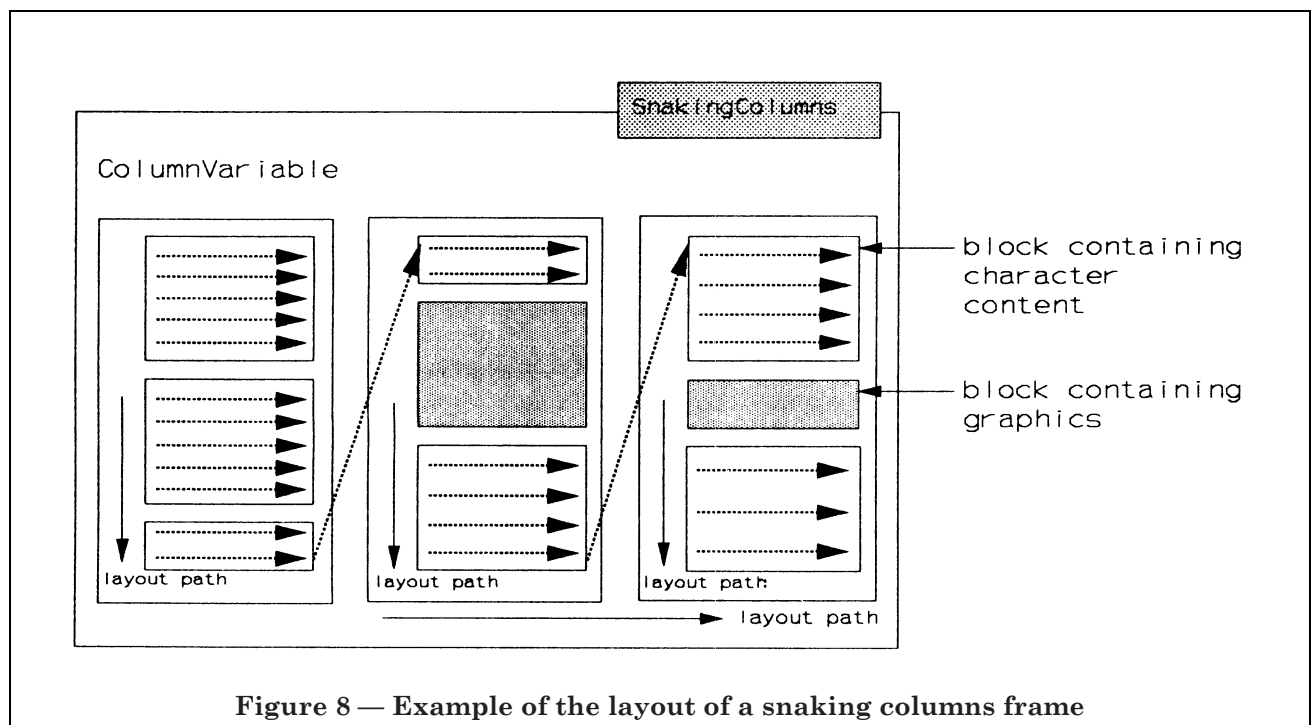
The dimension of this frame in the direction orthogonal to the layout path of the body area is fixed or defaults to the maximum value allowed within the body area.

The dimension of this frame in the direction parallel to the layout path of the body area is specified by the 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 its superior frame. Content shall only be laid out in this frame in the direction of the layout path specified.

6.3.5.5 SnakingColumns

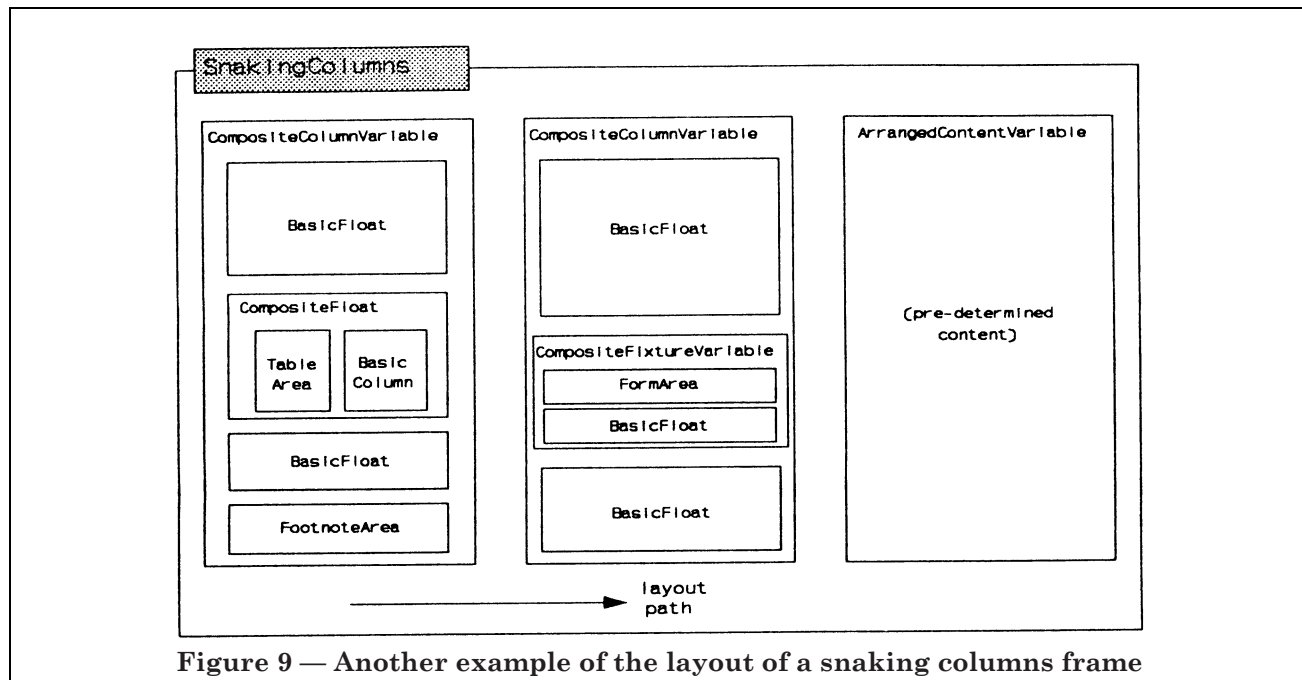
SnakingColumns is a constituent constraint that defines a composite frame that represents a snaking columns area within a body area. A snaking columns area is typically used for the layout of one or more columns of content in which the content is allowed to flow freely from one column to the next. Examples are shown in Figure 8 and Figure 9.



This is a variably positioned frame which may only be specified as a subordinate to a *VariableCompositeBody* frame.

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

- *ColumnVariable*;
- *CompositeColumnVariable*;
- *ArrangedContentVariable*;
- *SourcedContentVariable*.



The dimension of the 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 the 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 degrees in the case of body layout A, 90 or 270 degrees in the case of body layout B, and 270 degrees in the case of body layout C.

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 body layout A and approximately equal in length in the horizontal dimension in the cases of body layouts B and C. Note that “approximately equal” in the context of the “balance” attribute 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.

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 10.

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

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

- ColumnFixed;
- CompositeColumnFixed;
- ArrangedContentFixed;
- SourcedContentFixed.