

インタオペラブルマルチメディア実装規約の研究開発

5S-6 単純文書交換形式の文書構造

[INTAP第2専門委員会]

濱田 眞美(日本電気)

秦 務(松下電器)

1. はじめに

文字、図形、画像、音声などの混在する情報を、異機種の計算機システム間で相互に利用するニーズが高まってきている。

通商産業省工業技術院大型プロジェクト「電子計算機相互運用データベースシステム」の研究開発の一環として計算機システム間における相互運用性を確立するために必要な技術の研究開発が昭和60年度から開始された。このプロジェクトを実施している財団法人情報処理相互運用技術協会(INTAP)の第2専門委員会では、そのうちマルチメディア文書の相互運用性に関する技術の検討と実装規約の開発を行っている。

本稿は、この第2専門委員会の検討成果の内、マルチメディア実装規約の単純文書交換形式の文書構造について報告する。

2. 単純文書交換形式の設定目的

文書を扱う装置には多様な機能が要求される。しかしまた、手軽な価格帯での文書内容情報の交換/利用も要求される。INTAPでは、後者に主眼を置いた文書の実装規約として単純文書交換形式(AE.111n-J)を設定した[1]。

海外でも同じ目的で類似の実装規約を設定している。SPAGの文書交換形式Q/111、CCITTのT.502はAE.1111-Jとほぼ同等の機能を持つ。

3. 単純文書交換形式の規約設定の考え方

対象文書を単一カラムで構成される単純な階層を持つ文書とした。すなわち、論理構造は

- 複合論理要素を1階層とする

制限を加えた。また、割付け構造についても

- ページ集合を1階層とする
- 枠(frame)を1階層とする

制限を加えた。

文書体系クラスについては、

書式付き文書体系(FDA2)

処理可能文書体系(PDA5)

書式付き処理可能文書体系(FPDA2)

を使用可能とした。

属性値の設定に当たっては次の方針に従った。

- 現在実現可能なハードウェアが許す機能範囲とする
- 日本語の特性上必要な機能は標準機能とする

以上を基本的な考え方として、基本標準の文書応用プロフィール設定の原則に従い、単純文書交換形式実装規約を開発した。

4. 単純文書交換形式の文書構造の形式的記述

規定する全ての文書の構造を表現するために、文書構造レベルごとに形式的記述を行った。

特定構造 : 個々の文書の構造を規定

ジレック構造 : 文書の型紙の構造を規定

マジレック構造 : ジレック構造の型紙の構造を規定

これを理解し易くするために図式表現で解説した。

AE.111n-Jでは、文書構造が簡単であるために、マジレック論理構造はジレック論理構造に一致し、マジレック割付け構造はジレック割付け構造のページ集合以上を重ね合わせただけの構造とした。

5. 単純文書交換形式で記述できる文書

単純文書形式では、文書は柱などに使用するヘッダ領域、実際の文字/図形/画像などのメディア枠が配置される本文領域およびページ番号、脚書きな

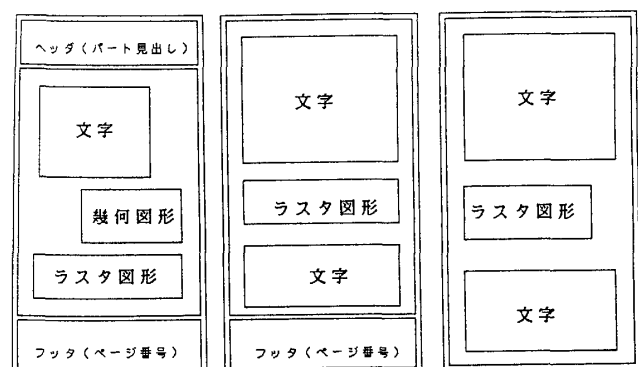


図1 AE.111n-Jのページ形式

どに使用されるフッタ領域の、3つの領域から成る(図1)。

5.1 単純文書交換形式で表せる文書の構造

(1) 論理構造

文書をパラグラフおよび図(基本論理要素)が任意順、任意回数並んだものとして取り扱う。パラグラフは、文字テキストから成る。図とは、ラスタ図形内容または幾何図形内容を持つものである。パラグラフおよび図を複数個まとめて、一つのパート(複合論理要素)として取り扱う。文書はパートの一次元的な集まりとなる(図2)。

ヘッダ・フッタは本文部分とは独立の複合論理構造を持つ。

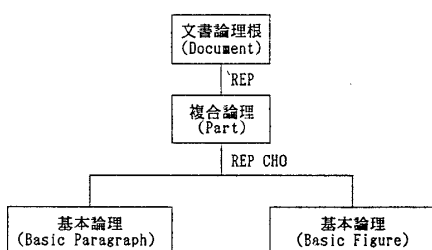


図2 AE.111n-J のメトリック論理構造

(2) 割付け構造

文書は複数のページ集合から成り、ページ集合は同一サイズのページから成る集合である。ページは省略可能なヘッダ枠と、文章の本文を含むボディ枠と、省略可能なフッタ枠から構成する(図3)。ページ集合は、省略可能な先頭ページの次に、以下のいずれかが続くものとした。

- (i) ページの中のヘッダ枠、フッタ枠、ボディ枠の割付け位置と、ページ番号を除いたヘッダ枠フッタ枠の内容が同一であるページが連続するもの
- (ii) 上記のようなページの2種類のものが交互に並ぶもの

ページのボディ枠には、文書の本文であるパラグラフまたは図を一次元的に上から下へ配置する。枠が横方向に並ぶこと及び互いに重なることは無い。

割付け構造であるヘッダ枠、フッタ枠から論理ソース属性を用いて、ヘッダ/フッタの内容を持つ論理構造を参照する。

5.2 単純文書交換形式の文書機能

本形式の文書は簡易な文書構造であるが、以下に挙げる高度な機能をも合わせ持つ。

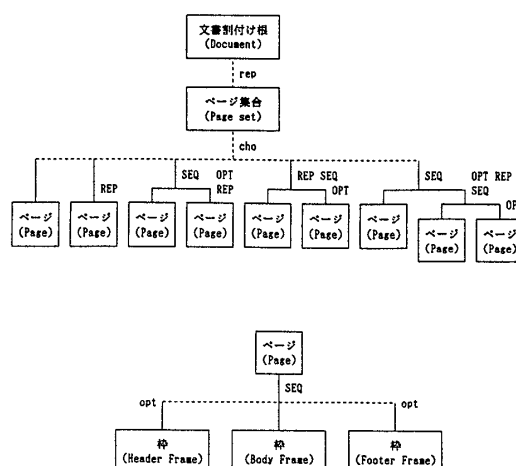


図3 AE.111n-J のメトリック割付け構造

- 紙面の裏表を意識したレイアウト： 綴じ代を見込んだ印刷位置を設定する
- 新割付け対象指示： 新規のページ形式に割付けられることを指示
- 非分割性指示： 文書の論理的単位が、ページ境界で分離しない様に配置する
- 同一割付け対象指示： 複数の論理内容を同一ページに割付ける
- ヘッダ/フッタ中のページ番号を自動生成する指示

6. おわりに

単純文書交換形式は構造に強い制約をした文書群であるが、表を記述する能力の付与など機能強化が要求されている。逆に、内容種別の組み合わせを減らして文書応用プロファイルを簡素化する要求もある。また、欧米とのデータ交換を可能とするためにまず単純文書交換形式を、欧米の文書プロファイルと整合させることが必要である。

INTAP第2専門委員会では、これらの課題を解決すべく、単純文書交換形式の検討を更に進める予定である。

7. 参考文献

- [1] インタポラブル マルチメディア 実装規約の研究開発—全体概要— 情報処理学会第37回全国大会予稿
- [2] // —中程度文書交換形式の文書構造— 情報処理学会第37回全国大会予稿
- [3] // —文字内容体系および日本語関機能について— 情報処理学会第37回全国大会予稿
- [4] // —ラスタ図形内容体系および幾何図形内容体系について— 情報処理学会第37回全国大会予稿
- [5] ISO TC97 ISO 8613 「Information processing Text and office systems - Office Document Architecture (ODA) and interchange format」 Part1,2,4 ~8