

ODA文書処理システムの試作(3)

1X-3

— 割付け処理 —

林 直樹、齋藤和雄、石田真美、村田 真
富士ゼロックス(株) システム技術研究所

1. はじめに

オフィス文書の国際規格としてODA(Office Document Architecture)がある。ODA文書は章や節といった論理構造とページや枠(フレーム)などの割付け構造を有しており、あるODA文書の割付け構造はその論理構造と割付けに対する種々の制約に基づいて生成せねばならない。この処理を割付け処理と呼ぶ。この割付け処理はこれらの種々の属性が相互に影響しあうため、複雑な処理を必要とする。今回、このODA文書割付け処理の実験システムをSmalltalk-80上に実装したので報告する。

2. 実装規格範囲

ODA規格の割付け処理に関する属性のうち、synchronization(同調), generator for subordinates(従属に対する生成子)のAGGなどの一部の属性を除くものを実装した。

現段階でフルセットの属性をサポートしていないとは言え、この実装範囲の属性のみでも

- ①複数の段組
- ②ヘッダ、フッタ
- ③自動章番号、頁番号、図番号付け
- ④脚注
- ⑤2か国語の文書の同調

などの特徴を有した文書の割付けを行うことが可能である。

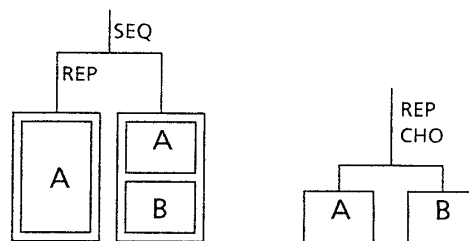
また、今回のシステム作成の目的は文書割付け処理のアルゴリズムを確立することにあるので、内容割付けについては規格に完全には従わず、かなり省略して行っている。

3. 割付け処理におけるバックトラック

ODA文書の割付け処理において欠かせない処理要素は幾つか挙げられるが、その内でもバックトラックは特に重要である。CHO, OPT, REPを含む構造式を持った共通割付け構造には構造選択の可能性が複数存在するので、ある時点で割付けが失敗した場合は、別の構造が生成できるところまで溯って、別の新しい構造を生成しなければならない。ここでは、このバックトラックの処理がかなり複雑になることを述べる。

このバックトラックの処理が難しい点は、すでに割付けられた構造を破棄する際に直接の失敗する原因となった割付けオブジェクトのみならず、それに関連した構造も同時に破棄し、それに伴って割

付け処理を再開する論理オブジェクトが変更される可能性を常に含んでいる点にある。さらに、layout category(割付け類別)が導入された場合にごく一般的に発生する割付け順序と論理順序の逆転、各種割付け指示属性の指定を考慮した処理が必要となる。



ページクラスα ページクラスβ

(a) 共通割付け構造

(b) 共通論理構造

図1. バックトラックの可能性を持つ文書構造

これを簡単な例を用いて説明する。与えられた共通割付け構造、共通論理構造が図1に示すようなものであるとする。ここで、A, Bはlayout categoryである。これらの共通構造から、特定論理構造はAまたはBのlayout categoryを持った基本論理オブジェクトが不規則に並んだもの、特定割付け構造はAが前半に、Bが後半に集められる構造となる。AとBの関係を論文などに見られる本文と参考文献の関係と考えると分かり易い。

この時、AあるいはBの長さは一般的には不定であるため、割付け処理は特定論理構造を辿り、Bの論理オブジェクトが出現した時点でページクラスαのREPの評価が終わったものと判断し、ページクラスβの生成を行う。割付けがさらに進行し、Aが1ページ目から溢れ、さらに2ページ目の枠より長くなった時、割付けは失敗し、バックトラック機構が起動される。その結果、2ページ目が破棄され、ページクラスαのREPが再評価されて、ページクラスαがもう1ページ生成される。このとき、破棄された2ページ目に割付けられていた論理オブジェクトは再び割付けが行われることになる。

この例の場合は、A, Bのlayout categoryを全く独立に割付ければこのような複雑なバックトラックは発生しないのであるが、permitted categoriesとしてA, Bの両方を許しているような枠に割付けるこ

とを考えてみれば、一般的に正しい割付け結果が得られないことは明らかであり、このようなバックトラック機構は必須である。

4. 割付け処理のSmalltalk-80への実装

4.1 構成

今回のSmalltalk-80上での実装では、前節で述べた点を考慮した割付け処理のための機能オブジェクトクラスLayouterを導入した。

このクラスのインスタンスaLayouterとSmalltalk-80上にマッピングされたODA文書との関係を図2に示す。aLayouterは、与えられたODA文書、現在の割付け対象である論理オブジェクト、そして、その論理オブジェクトが割付けられるであろう割付けオブジェクトを知っている(図2中の矢印①, ②, ③に各々相当する)。

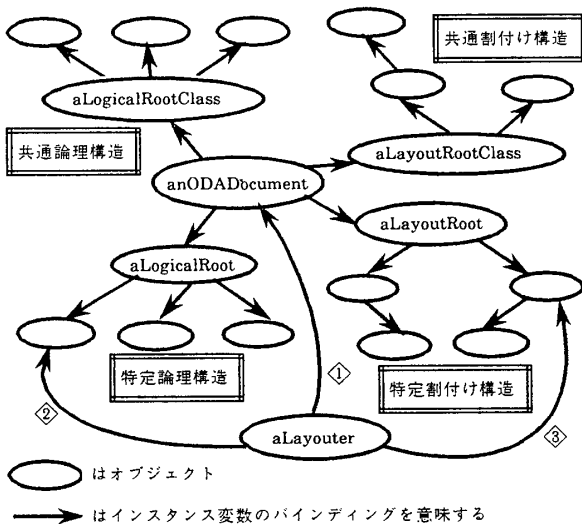


図2. オブジェクトaLayouterと文書構造

このクラスの機能は大きく5つの部分に分けられる。すなわち、指令部、初期化部、割付け構造生成部、内容割付け部、およびバックトラック部である。

[指令部]

他の各部分呼び出ししながら割付けを進める。

[初期化部]

割付け処理の前処理を行う部分で、前回の割付け結果の破棄などを行う。

[割付け構造生成部]

論理構造を辿りながら、論理側に指定されている割付けオブジェクトクラスなどの割付け指示属性に適合した割付け構造を生成する。

[内容割付け部]

割付けられるべき基本論理オブジェクトとその内容が中に割付けられる最下位枠オブジェクトが決定した時に呼び出される。内容割付けの結果に影響を与えるconcatenation(連結), layout path(割付け方位)などの属性に関してはここで評価される。内容割付けの結果、割付けられた内容部を有した区(ブロック)が生成され、現在の最下位枠に接続される。

[バックトラック部]

内容割付けに失敗した時、または論理オブジェクトに指定されている属性に適合する割付け構造が生成できなかった時などはバックトラック部が呼び出される。そして、新しい構造を生成できる可能性をもった割付けオブジェクトまで割付け結果を溯り、不必要となった割付けオブジェクトを破棄する。さらに、今破棄された割付けオブジェクトに割付けられていた論理オブジェクトの中から、次に割付けを再開すべき論理オブジェクトを決定する。

4.2 処理結果

図3にこの割付け処理システムによって割付けられた文書例を示す。文書例の入力及び出力には、参考文献3)で述べたツールを用いている。

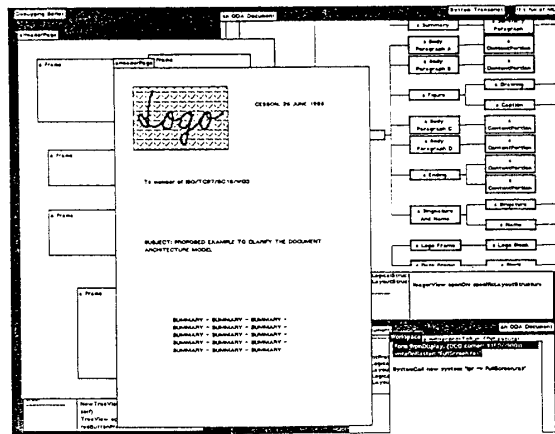


図3. 割付け処理結果を含む画面イメージ

5. おわりに

割付け処理において大きな問題となる点に、割付け指示属性間の優先順位がある。割付けは行ってみなければ失敗するかどうか分からないため、複数の属性が指定されている場合は可能性として組合せの数だけ試し、最終的にそれらの属性はすべて無視せねばならないといったことも考えられる。その場合にはかなり莫大な時間が無駄になることが容易に予想され、ヒューリスティクスあるいはバックトラックの際に過去の履歴を消さずに再利用するといったことが必要になるだろう。

謝辞

今回のシステム作成にあたって多大なるご助言を頂いたゼロックス社のMartin Cooper氏に深謝いたします。

参考文献

- 1) ISO/IS 8613 : Information Processing - Text and Office Systems - Office Document Architecture (ODA) and Interchange Format (March 1988).
- 2) 石田真美、村田 真 : ODA文書処理システムの試作(1)、情報処理学会第37回全国大会(1988).
- 3) 鈴木克明、中津山恒、村田 真 : ODA文書処理システムの試作(2)、情報処理学会第37回全国大会(1988).