

SGML 文書の表組みレイアウト作成支援方式

4 V - 8

森口 修 今村 誠 鈴木 克志

三菱電機株式会社 情報技術総合研究所

1. はじめに

文書の電子化が日常的になったが、単に電子化しただけの文書では、画面に表示するか紙に印刷するかして人間が読むという使い道に限定される。

そこで、SGML(Standard Generalized Markup Language) / XML(Extensible Markup Language)という文書記述言語が注目されている。SGML/XMLでは、文書を人間が読むという使い方だけでなく、文書内に記述された特定の内容を抽出してデータとして扱うという使い方ができる。さらに SGML/XML は文書を定義するメタ言語でもあるため、業務に特化した文書を自由に定義することができる。

文書を「計算機がデータとして扱う」ことと「人間が読む」ことを両立させるため、SGML/XML には文書の論理構造を記述し、スタイル(割り付け)情報は特定のタグで SGML/XML 文書内に埋め込む方法[1]やスタイルシートと呼ぶ別の文書に記述する方法がある。文書の論理構造とは、文書を構成する論理要素間の階層関係である。計算機は文書の論理構造を解析して所望のデータを抽出する。文書のスタイル情報とは、表示位置や文字サイズなどの表示方法である。

スタイルシートでは、文書の論理要素毎に表示方法を定義する[2]ため、表組レイアウトを行なうスタイルシートを作成する場合、論理要素毎に表を構成する行・列やセルのスタイル指定が必要であり、さらにそのスタイル指定は表組み範囲内の論理要素間の関係を把握しないと不可能であるため熟練と手間を要するという問題があった。

これらの問題を解決するには、スタイルシートの作成を容易にするツールが必要となる。本論文では、スタイルシートの作成を支援するツールについて、特に「表組み」という複雑な表示方法に着目して述べる。

A Supporting Method of Layouting Table Format from SGML Documents

Osamu MORIGUCHI, Makoto IMAMURA, Katsushi SUZUKI
Mitsubishi Electric Corporation.

5-1-1 Ofuna, Kamakura, Kanagawa 247, JAPAN

2. 表組みパターンの分析

業務文書の形式は、表組みレイアウトとなっているものが多い。そこで、実際の業務文書の論理構造と表組みレイアウトとの関係を分析した結果、以下のように表組みパターンを分類した。

- 表組みパターン 1 (図1)
文書の論理構造との相関関係がなく、論理要素の出現順に表セルを埋めるパターン。
- 表組みパターン 2 (図2)
下位の論理要素が同一である上位の論理要素の集合を表組みするパターン。
- 表組みパターン 3 (図3)
下位の論理要素を持つ上位の論理要素がリストであるパターン。
- 表組みパターン 4 (図4)
下位の論理要素がリストであり、その上位の論理要素もリストであるパターン。

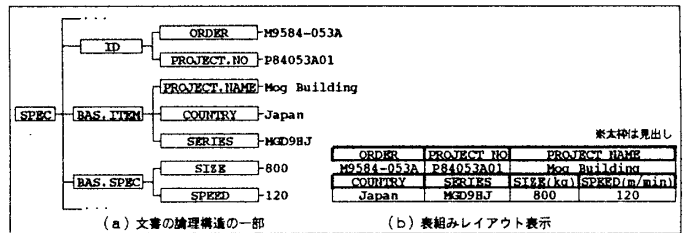


図1. 表組みパターン1

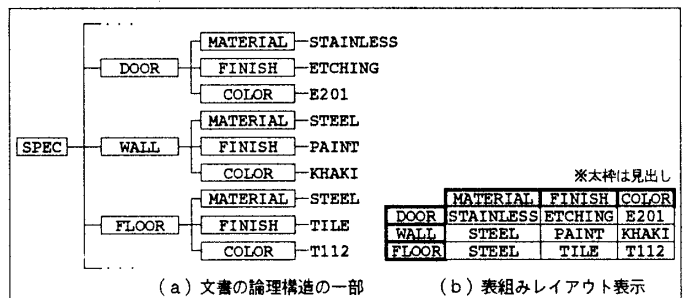


図2. 表組みパターン2

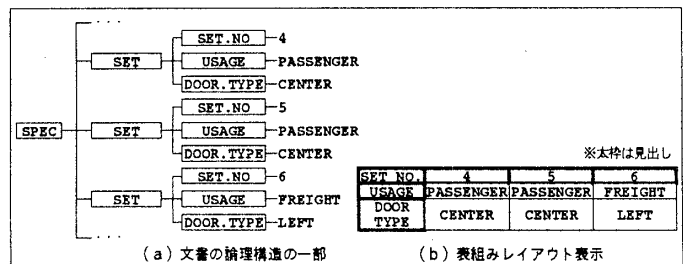


図3. 表組みパターン3

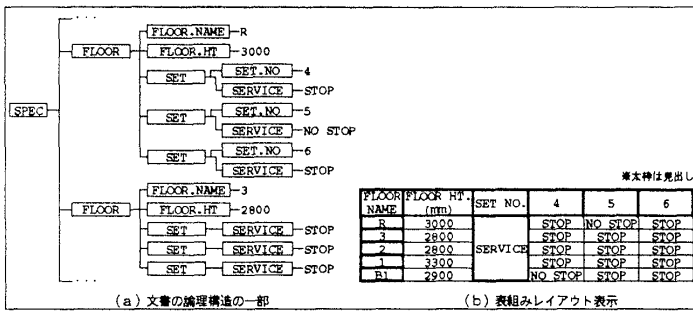


図4. 表組みパターン4

3. スタイルシート作成ツール

スタイルシート作成ツールでは、文書の論理構造に対して、表組みレイアウトを行なう範囲と2章で分類した表組みパターンを指定することにより、表組みレイアウト表示を行なうスタイルシートが作成できることを特徴とする。このうち、本論文では表組みパターン1について述べる。

図5は、図1(a)に示す文書の論理構造に対して、論理要素IDから論理要素SPEEDまでの範囲を表組みパターン1に指定した時のスタイルシート作成ツールの画面である。

スタイル表			
タグ名	L	見出し	表組み
白ID	2		TABI 設定
白ORDER	#	注文番号	設定
白PROJECT.NO	#	工事番号	設定
白BAS.ITEM	2		設定
白PROJECT.NAME	#	工事名	設定
白COUNTRY	#	国名	LP 設定
白SERIES	#	機種名	設定
白BAS.SPEC	2		設定
白SIZE	#	サイズ(kg)	設定
白SPEED	#	速度(m/min)	/TABI 設定

図5. スタイルシート作成ツール

図5で表組みパターンを指定するだけで表組みレイアウトのスタイル定義は完了であるが、さらに「設定」ボタンを押すと図6に示す表組みプロパティ設定画面が表示され、方向や列数などの表組みのカスタマイズができる。

以上のスタイル指定操作により作成したスタイルシートを用いると、図1(a)に示す文書を図1(b)に示す表組みレイアウトでの表示ができる。なお、図1(b)では論理要素との対応をわかりやすくするために英語の見出しを用いているが、図5

で日本語の見出しを指定すると、図1(b)の見出しは日本語となる。

表組みプロパティ		
表組みパターン 1	方向	<input checked="" type="radio"/> 水平 <input type="radio"/> 垂直
	列数	4
	枠線の太さ	1
	セル間スペース	0
	セル内マージン	1
	見出し背景色	SKYBLUE
	その他背景色	WHITE
	行番号	<input type="radio"/> あり <input checked="" type="radio"/> なし
	角見出し	
	キャプション	配置
サイズ		+1
親見出し	1	
親見出し	2	
親見出し	3	
		設定 取消

図6. 表組みプロパティの設定

4. まとめ

業務文書で用いられる表組み形式をパターン分けしたことにより、スタイルシート作成ツールでは、表組み範囲と表組みパターンを指定するという簡単な操作でSGML/XML文書の表組みレイアウト表示ができる。このとき論理要素毎にスタイルを指定する必要はない。

今後の課題について以下に述べる。

- (1) 多くの業務文書について適用し、本ツールの汎用性の確認と改良を行なう。
- (2) スタイルシート定義言語 XSL (Extensible Style Language)やフォーム定義言語 XFDL (Extensible Forms Description Language) などの W3C (World Wide Web Consortium) による標準提案に対応する。
- (3) スタイル変更の度にレイアウトを画面上で確認できるよう操作インタフェースを改良する。

参考文献

- [1] 石田,神谷,市山:「既存文書のレイアウト情報付き構造化とその利用」, 情処学会情報学基礎研究会報告 FI-96-44-4, pp.25-32,1996
- [2] 遠瀬,早川,並木,高橋:「文書構造を用いた文書レイアウト機能の設計と実現」, 情処学会第54回全国大会論文集 3S-3, pp.323-324,1997