

# SGML/XML 文書の作成と表示に 関する取組み

---

岩垂 誠

富士通株式会社  
第一ソフトウェア事業部第四開発部

屋代 禎夫

株式会社富士通研究所  
情報サービス研究部

本稿では SGML/XML 文書を作成するためのエディタの実装形態と、その中で開発した、スタイルシート生成に関する実装例を紹介する。SGML/XML エディタは、DTD の定義に従った SGML 文書作成をするための専用 GUI を持ったものとし、さらにレイアウト表示と連携することにより文書作成時の内容確認作業の効率化を図っている。レイアウト表示では DSSSL スタイルシートを採用し、スタイルシート編集機能を持つ。ダイアログボックス入力によるスタイルシートエディタは、多様なスタイルの設定を簡便な操作で行うことを可能にしている。

## Approach to making and display of SGML/XML document

Makoto Iwadare

FUJITSU LIMITED  
DEVELOPMENT DEPT.IV

Sadao Yashiro

FUJITSU LABORATORIES LTD.  
INFO. SERVICE ARCHITECTURE LABO.

This paper describes an implementation about the SGML/XML document editor and the style-sheet editor which is integrated in the editor. The SGML/XML document editor has a specialized GUI(Graphical User Interface) to create SGML documents or XML documents which have DTD (ie. valid XML documents) easily. In addition, by communicating with a document browsing facility, this editor gives user enormous efficiency to check a document content. For a document style specification, that engine uses DSSSL style sheet, and provides the style sheet editing facility to specify document styles. This style sheet editing facility enables user to specify various kind of styles easily by form input method.

# 1 はじめに

CALS活動の活性化やWWW(World Wide Web)の普及に伴い、電子文書の形式として文書にマーク付けを行い構造化するためのメタ言語であるSGML(Standard Generalized Markup Language)やXML(eXtensible Markup Language)が注目されている。

SGML/XMLにより記述されるドキュメントは、文書のテキスト情報とその論理的な関係を構造化・データ化したもので、印刷・表示における情報からは切り離されていることが特徴であり、電算機によるドキュメントの管理・再利用を容易にする事を大きな利点としている。

SGML/XML文書作成機能の実現アプローチとしては、既存のワープロ・DTP・テキストエディタに対して構造化文書を作成するための機能を追加する方法と、SGML/XML文書作成のための専用エディタ機能を有する製品を開発する方法の2つに分けられる。

従来のワープロやDTPなどの既存の文書作成ツールを使った場合は、その文書作成ツールが本来もっている文書構造やユーザインタフェースと、作成したいSGML/XML文書の構造に食い違いがある場合は、意図した通りの正確なSGML/XML文書を作成することは容易ではなかった。

またSGML/XML文書作成のための専用エディタは、高機能/高価なこともあり、あまり普及しているとは言えない状況であった。

SGML/XML文書の作成と表示機能の実装に当たって、DTDに従った正確な文書作成が行え、かつ既存文書処理のレイアウト処理に影響されないことを目標とした。この条件を満たすために、既存文書処理ツールへの機能追加ではなく、SGML/XML文書作成のための専用エディタにより実装を行うことにした。

2節では、SGML/XML文書を作成できる専用のエディタの開発を行った背景、特徴、基本構造について紹介する。

3節では、レイアウト表示機能について簡単に特徴、基本構造について紹介する。

4節では、レイアウト表示の機能の中で実現したスタイル定義機能について、背景、特徴、基本構造について紹介する。

## 2 SGML/XML 文書エディタ

### 2.1 背景

既存製品としてのSGML文書の専用エディタは、その開発コンセプトにも依存するが操作性や文書の表示方法も様々である。大きく分けるとレイアウトされた状態でタグ名が見えないWYSIWYGタイプと、タグの構造や名前が見える非WYSIWYGタイプに別れる。WYSIWYGタイプのものは文書の作成前に予め文書要素に対して表示スタイルを設定するなど、文書作成前の事前の準備が必要になる。また、ワープロやDTPからの変換では、文書構造によっては文字データを含まないタグなどの作成が困難であったり、文書の書き方によってはDTDに沿わない文書でも作成できてしまうことがあるなど、SGML文書を作成するツールとして使うには不十分な面があった。

さらに、CALSで作られるSGML文書は数10Mbyteといった長大なものになることがあり、その場合は複数人による分担執筆となるため、複数のファイルから構成される文書を作成できることが要求される。

これらの課題に対して、正しいSGML文書を作成するための十分な機能を備えていること、コンパクトで軽快に使えること、文書作成前の準備などが最小限となること、分担執筆に使えることなどを目標にSGML文書を直接作成/編集するための文書エディタ(SGML-Editor)の開発を行った。

## 2.2 エディタの構成

SGML-Editor の機能構成を図-1 に示す。SGML-Editor は、SGML 文書編集とレイアウト表示の2つのコンポーネントからなっており、各機能について以下に説明する。

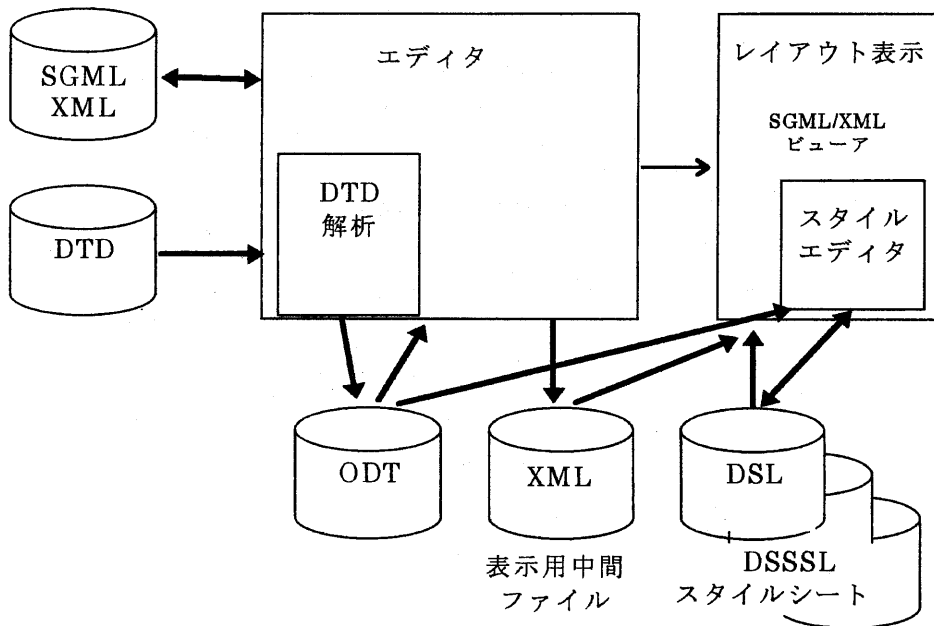


図-1 SGML エディタの構成

- エディタ  
SGML/XML 文書を階層的に表示し編集する
- レイアウト表示  
DSSSL(Document Style Semantics and Specification Language)スタイルシートを元にレイアウト結果を表示する。
- DTD 解析  
DTD ファイルを読み込み、独自の解析済み形式 (ODT) を作成する。
- スタイルエディタ  
DTD (ODT) を読み込み、各要素に対してスタイルの設定を行いスタイルシートを作成/更新する。

## 2.3 特徴

文書編集機能では、文書作成を行う前の準備作業を最小限にするためにオリジナルの DTD ファイルが直接扱えるようにし、分担執筆を支援する目的で部分文書編集機能を作成した。また、コンパクトで軽快に使えるという要件を満たすため、図-2に示すような非 WYSIWYG タイプのインターフェースとした。

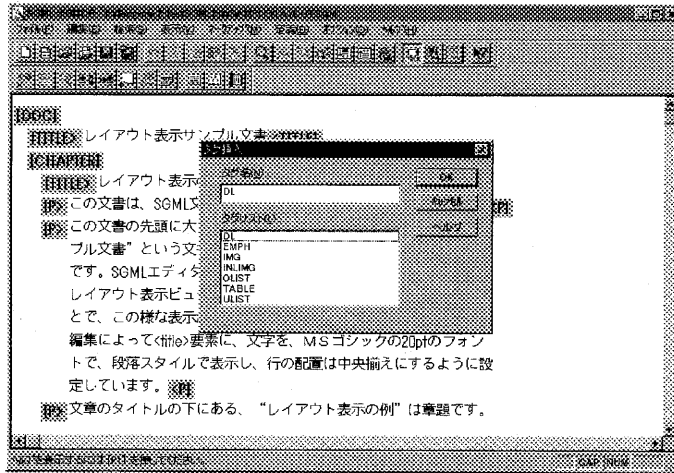


図-2 SGML/XML 文書編集画面

### (1) DTD ファイルの直接指定

DTD ファイルに特別な加工をすることなく、オリジナルの DTD を処理できるようにしたことで導入時の負担を軽減した。しかし DTD ファイルはテキスト形式であり、記述量に比例して構文解析時間が長くなるため、文書作成のための画面を表示するまでの時間が長くなるという問題がある。そこで解析済みの DTD 情報を DTD ファイルと対で保存しておくことで、2回目以降の文書編集時に DTD ファイルの解析処理を省略できるようにした。

### (2) 部分文書編集

SGML では、1つの文書を複数のファイルの集合として作成することが出来るため、長大な文書を分担して作成する場合、章や節毎に担当者を決めて各自担当の部分だけを作成し、これらの文書ファイルを集めて大きな文書を編成とすることができる。

こうした SGML 文書の特徴を生かした分担執筆を実現するために、部分文書作成機能を開発した。部分文書作成機能により章/節といった文書中の特定単位だけの文書ファイルを、DTD を特別に加工せずに、文書全体を編集する場合と同じ DTD を使って作成できるようにしている。

### (3) 図形データ編集

今回作成したエディタでは図形データを編集するための独自の機能は持たない。しかし、文書中から参照される外部形式データと処理するアプリケーションの関連付けに従って各図形データをそれぞれの図形アプリケーションと連携させることができる仕組みを作成した。

この連携機能では、SGML 文書作成で図形などの他メディアを組み込む場合に使われる次の2つの方法に対応している。

- 外部エンティティとして定義してエンティティ参照の形で組み込む
- タグの属性値としてファイル名を指定する

#### (4) レイアウト表示

図-2に示すような非 WYSIWYG タイプの場合、文書構造(タグの並びや階層)が分かりやすい反面、各タグに書かれた文字を中心に文章内容を確認する場合には読みにくいといった面を持っている。そこで図-3に示すような現在編集している SGML 文書ファイルの内容を簡易なレイアウト付けをして表示する機能を開発した。この開発には富士通研究所で開発された SGML/XML ビューア「HyBrick」の技術を利用している。

SGML 文書は体裁付けのための情報を含んでいないために、レイアウト表示には文書の表題、見出し、段落といった要素(タグ)毎にスタイル(体裁)を指定する必要がある。これらの設定は専用のダイアログから、対話的に行うことができるようにし、スタイル設定の作業が軽減されるようにしている。またこのレイアウト表示の機能は、文書編集の機能とは切り離し疎な連携を行うことで、レイアウト情報の作成が文書作成時に必須とならないようにした。

さらに、レイアウト情報を1つの DTD に対して複数作成することができるようにしたことで、文書内容全部の表示だけでなく確認したい対象(例えば章見出しだけや図のタイトルだけなど)だけを表示するような目的別のレイアウト表示が行えるようになった。

3 節以降ではこのレイアウト表示の機能について説明する。

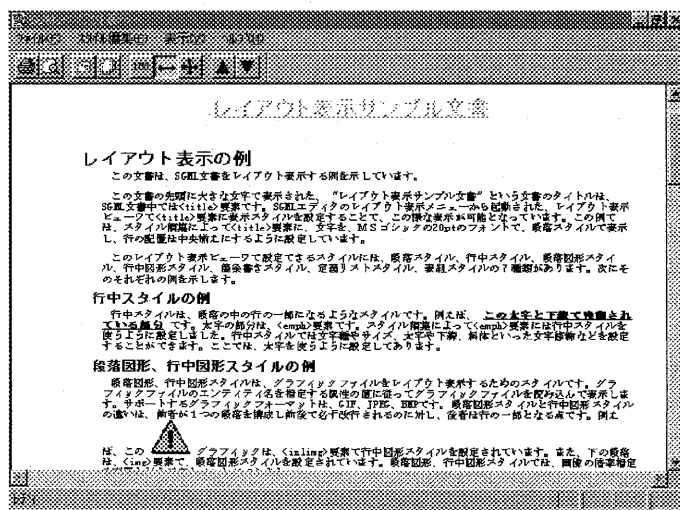


図-3 SGML/XML 文書レイアウト画面

### 3 レイアウト表示

レイアウト表示部は、編集された SGML/XML 文書の画面上での表示と印刷、スタイルシートの編集機能を持っている。レイアウト表示の構成は SGML、ODT、DSSSL 情報からレンダリングを行う表示部と、表示部からダイアログボックスの形で呼び出される DSSSL スタイルシートエディタ部の二つに大きく分かれている。

#### 3.1 表示部

表示部は SGML 文書と DSSSL 情報を受け取ると以下図-4の処理フローに従って処理を行う。

SGML エンジンでは SGML 文書をパースし論理構造に基づいたオブジェクトツリーを構成する。その SGML オブジェクトツリーを基に DSSSL エンジンでは与えられた DSSSL 情報から DSSSL フローオブジェクトツリーを作成する。画像デコーダは SGML 文書の参照している画像データを内部イメージ情報に復号化する。これらのエンジンの解析結果からレンダラは描画を行う。

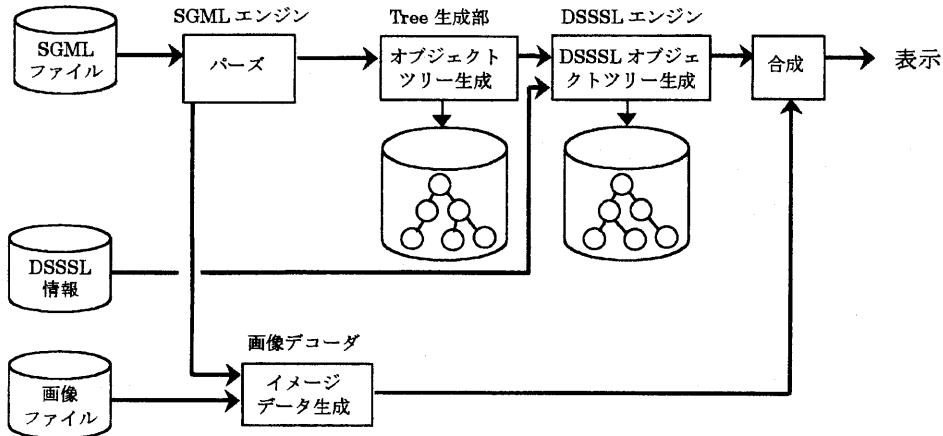


図-4 レイアウト表示処理フロー

### 3.2 SGML エディタとレイアウト表示の連携

レイアウト表示は SGML エディタのメニューから実行形式が呼び出される。SGML エディタは呼び出す際に、該当する文書の SGML ファイル、生成した DTD 情報、DSSSL ファイルの情報を引数で渡す。レイアウト表示はこの情報を基にレンダリングを行う。受け渡しに使う SGML ファイルは、エディタで正当性が保証されているため、レイアウト表示で改めて正当性を確認する必要はないことから、XML の well-formed な形式で行っている。

レイアウト表示が起動されている間の SGML エディタとの連携は、メッセージ送信により行われている。メッセージによる連携の機能は以下の 3 種類である。

#### 1)リロード(再表示)

SGML エディタにより文書に変更が加えられた時にユーザが指定する事により、エディタがレイアウト表示の再描画を指示する時に送信される。

#### 2)タグジャンプ

文書中の SGML エディタが指し示しているエレメント箇所の先頭になるようにスクロールしてレイアウト表示を行う場合に送信される。これは SGML エディタが渡す XML 形式の文書では、全てのタグに一意の ID を属性値で指定するようにしており、この ID でジャンプ先を指定することで実現している。

#### 3)終了

SGML エディタを終了する事をレイアウト表示に通知する。これを受けたレイアウト表示はその処理を終了し、ウィンドウを閉じる。

これらの連携により、SGML エディタによる文書の編集を妨げる事なくユーザは文書のレイアウトを確認する事ができる。

## 4 スタイルシート編集

スタイルシートの編集機能について、その必要性、構造、用意した文書スタイル、レイアウト表示との連携機能について説明する。スタイルシートを編集するユーザインタフェースは、帳票形式のダイアログボックスを用い、SGML/XML の表示・印刷に必要と思われる最小限の文書スタイルを用意した。

## 4.1 DSSSL 編集ツールの必要性

SGML/XML により記述されるドキュメントは、文書のテキスト情報とその論理的な関係を構造化・データ化したもので、印刷・表示における情報からは切り離されていることが特徴であり、電算機によるドキュメントの管理・再利用を容易にする事を大きな利点としている。そのため、SGML/XML を印刷・表示するアプリケーションではドキュメント情報の他に文書のレイアウトのための情報を必要とする。

DSSSL は SGML/XML ドキュメントの論理的な構造変更および表示・印刷におけるフォーマットを可能にするよう定められた国際規格(ISO-10179)であり、LISP の一種である Scheme に準拠した言語仕様により記述される。表示・印刷における様々な構成とレイアウト等のフォーマット処理は、フローオブジェクトと呼ばれるオブジェクト構造によりクラス化され、Scheme 言語処理系により管理される。

DSSSL を用いて SGML/XML ドキュメントをブラウジングするアプリケーションでは、SGML の各タグ構造に対し、フローオブジェクトによるレイアウト定義の関連付けを行う。同一のドキュメントの論理構造に対し同一の DSSSL を用いた場合、レイアウトの一致が保証される。

SGML/XML ドキュメントに対し、柔軟な表現手段と堅牢な一致性を保証する DSSSL であるが、反面 Scheme というプログラム言語を使用するためプログラミング技術を持たず、また、DSSSL フローオブジェクトが用意しているクラスとプロパティを知らないユーザが編集する事は困難であり、簡便な入力手段を求められていた。

## 4.2 スタイルシートエディタの構造

今回開発のスタイルシートエディタは入力方式として、MS-Windows 上の帳票入力(ダイアログボックス)を採用し、エディットボックスにパラメータを入力、リストから該当する項目を選択、等の単純な操作を行うだけで、目的とする SGML/XML ドキュメントについての DSSSL スタイルシートを自動生成、また再編集を行う事を目的とした。図-5 にその概念を示す。

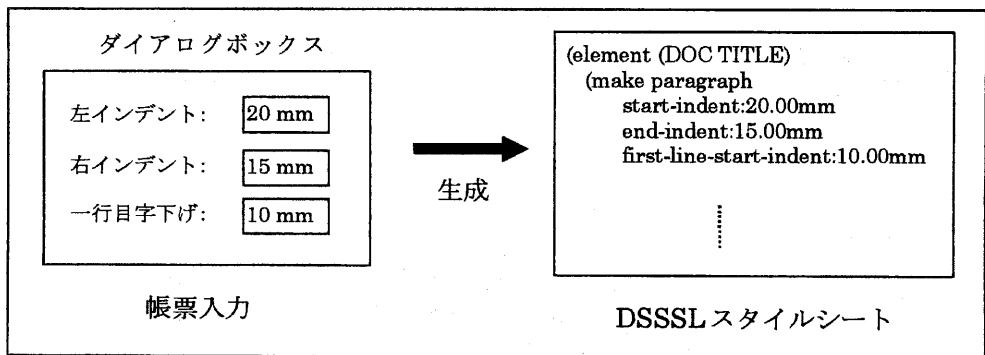


図-5 スタイルシートエディタ概念図

スタイルシートエディタでは、SGML/XML ドキュメントの要素の論理構造一つについて各々 DSSSL のフローオブジェクト定義を関連付ける方式を取っている。すなわち、タグ構造の階層が `<DOC><TITLE><P>~</P></TITLE></DOC>` となる様なドキュメント構造についてスタイルを定義したい場合は、この `<DOC><TITLE><P>` がスタイルシートエディタの管理する基本単位として、編集を行う。スタイルシートエディタではこれを要素の「適用部分」としている。

また、スタイルシートエディタでは、文書のレイアウトにおける情報をユーザが概念的に把握し易いように分類を行い、これを「スタイル定義」としてユーザの編集対象の一単位とする。各「スタイル定義」は一種類の「スタイル種類」を持つ。この「スタイル種類」は SGML/XML のレイ

アウト表示に必要なものを抽出・分類している。「スタイル種類」の一覧を表-1に示す。「スタイル種類」は先に説明した「要素の適用部分」を単独、あるいはスタイルの種類によっては複数を組み合わせた形で形成される。このスタイル定義の集合がスタイルシートのデータ全体を表す。この時のスタイルシートのデータ構造を図-6に示す。

それぞれの要素適用部分には設定すべきパラメータ群がその性質により分けられ、「書式」として定義される。この「書式」がスタイルシートエディタで編集を行う最小単位であり、その書式の種類ごとにダイアログボックス編集を行うよう設計されている。「書式」の一覧を表-2に示す。

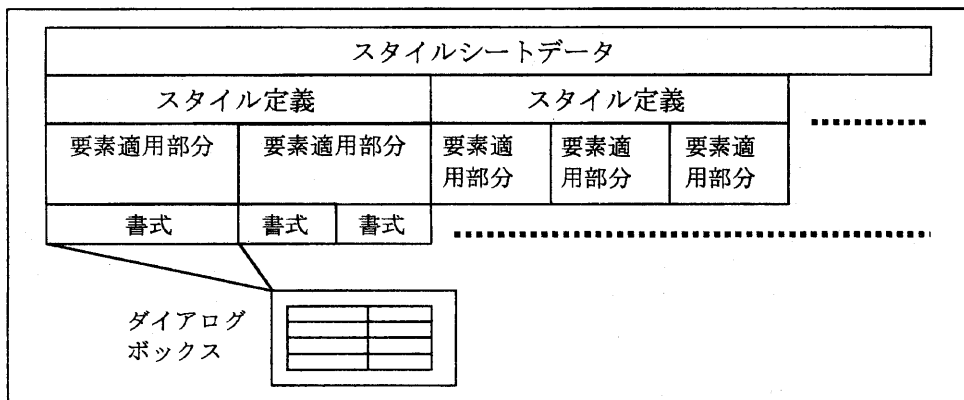


図-6 データ構造

表-1 スタイル種類

スタイル	用途
段落	段落で構成される文章
行中文字	文章中に含まれる特定の文字列を修飾する時など
段落図形	段落間に表示される図形
行中図形	文章などの中で表示される図形
箇条書き	箇条書き
定義リスト	項目とその説明文を列記する
表組	行と列からなる作表
段落内改行	段落内に改行を挿入する

表-2 書式種類

書式	設定する項目
段落	インデント、行間、段落間隔などの段落構成
文字	表示する文字のフォント情報
図形	表示する図形の情報
箇条書きヘッダ	箇条書きの各項目の先頭につける番号や記号
定義リスト配置	定義リストの項目・内容の位置情報
表枠線	表の枠線の太さ
表セル範囲	表の構成セルの行・列にまたがる範囲

### 4.3 レイアウト表示との連携

SGML/XMLビューアとスタイルシートエディタの関係を図-7に示す。スタイルシートエディタはSGML/XMLビューアの一つの機能モジュールとして実装されており、メニューから呼び出される。



SGML/XML ビューアは DSSSL スタイルシートを DTD、SGML ファイルと共に読み込み、レイアウト表示を行う。表示させたい文書と DTD のセットに対応する DSSSL ファイルがない場合には、デフォルトの DSSSL スタイルシートを生成し、SGML 文書をプレーンテキストで表示する。ユーザがスタイルシートを編集する場合は、SGML/XML ビューアからメニュー呼び出しでスタイルシートエディタを起動する。

起動するとスタイルシートエディタは、SGML/XML ビューアが使用しているスタイルシートの複製を一時ファイルとして持ち、編集を行う。ユーザがエディタ上での編集を終了すると、スタイルシートエディタはスタイルシートファイルを更新し、SGML/XML ビューアは更新されたファイルを再度読み直し、SGML 文書を再描画する。

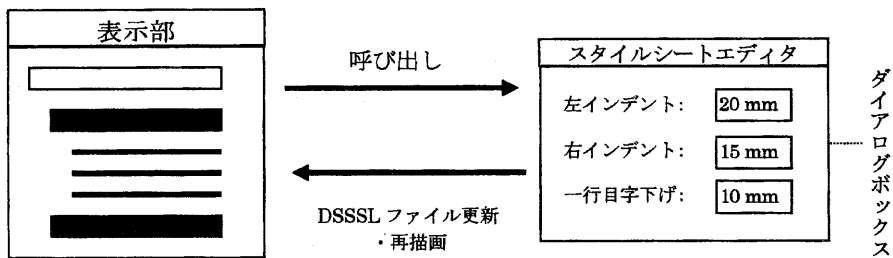


図-7 連携図

## 5 まとめ

本稿では SGML/XML 文書作成機能を実現する一アプローチを示した。SGML-Editor は SGML/XML 専用とすることで DTD に従った正確な文書作成と既存文書処理のレイアウト処理とは独立なスタイルシートを利用した表示と印刷を実現した。

スタイルシートに関しては、従来非常に煩雑であった DSSSL の記述を言語に不慣れなユーザーでも容易かつ簡便に行うことができるようにエディタを実装し、また SGML/XML ビューアとの連携により、SGML/XML ドキュメントに対してスタイルの編集と表示の整合性をビジュアルに、かつリアルタイムに把握する事ができ、文書のレイアウトに要する労力を大幅に軽減することが可能となった。

以上の結果、SGML 文書の作成ツールとして、多くの人に使っていただけるように、SGML 文書を作成するための基本的な編集機能を備え、かつ文書作成前の準備などが簡易にできる環境を提供した。しかし、SGML 規格を使った文書で記述できる文書の種類は多様であり、今後も表編集機能、図形の組み込み機能、リンク情報の設定機能など、さらに使いやすい機能の充実に取り組んでいきたい。

## 参考文献

- [1] 富士通 XML 推進チーム：はじめての XML，日経 B P 社，1997
- [2] 田村真：XML 入門，日本経済新聞社，1998
- [3] 門馬敦仁：XML の最新動向，情報処理
- [4] ISO/IEC 10179:1996  
Document Style Semantics and Specification Language(DSSSL)