

WWW ブラウザによる SGML 文書入力方式について

1 Y-2

今村 誠 森口 修 鈴木 克志

三菱電機（株）情報技術総合研究所 音声・言語インタフェース技術部

1. はじめに

CALS(Commerce At Light Speed)やEC(Electronic Commerce)の進展に伴い、業務向けの文書情報を不特定多数の企業間ですばやく効率的に共有したいという要求が増している。この要求に応えるための文書形式が SGML(Standard Generalized Markup Language)である。SGML の特徴は、文書型定義(DTD:Document Type Definition)によって、業務に応じて文書の論理構造を厳密に規定できるので、文書情報の業務システムとのデータ連携が容易になる点にある。SGML 形式の製品情報をネットワークにより企業間で共有することにより、「製品販売企業と製品設計企業間での設計情報の共有」や「WWW(World Wide Web)上の製品情報の広域収集・比較検索」等の応用システムの開発が容易になる。

しかし、業務システムとのデータ連携を前提とした文書情報交換では、交換する SGML 文書が、単に DTD による文書構造制約を満足するだけでなく、文書中の専門用語/固有名詞や数値の単位や範囲といった業務アプリケーションに応じて決まる文書内容制約を満足する必要がある[1]。

本稿では、業務システム連携に必要な文書内容制約の入力時チェック機能を特長とする WWW ブラウザによる SGML 文書入力方式について述べる。

2. WWW ブラウザによる SGML 文書入力の課題

2.1 要求機能

SGML エディタが十分に普及していない現段階では、不特定多数の企業による SGML 文書交換を促進するには、WWW ブラウザからの SGML 文書入力機能を提供するのが望ましい。データ連携を前提とする SGML 文書作成支援ツールの要求機能を以下に示す。

Content Test of SGML Documents
 Makoto Imamura, Osamu Moriguchi, Katsushi Suzuki
 Mitsubishi Electric Corporation, Information
 Technology R & D CENTER, Human Media Technology Dept
 5-1-1 Ofuna, Kamakura, Kanagawa 247, JAPAN

(1) 入力時の文書内容チェック機能

業務アプリケーションに応じた文書内容制約を満足するかどうかをチェックする機能。また、その制約を満足しない場合にガイダンスやヘルプ情報を提示する機能。文書入力者のための機能である。

(2) DTD や文書内容制約の保守機能

DTD の変更や文書内容制約の修正/追加を容易にするための機能。システム管理者向け機能である。

2.2 既存技術の現状レベル

SGML 文書を作成支援するための既存技術として、「INCONTEXT(XSOFT 社)等の汎用 SGML エディタ」と「HTML による SGML 文書入力」[2]がある。現状の SGML エディタでは、DTD に対する正当性をチェックする機能はあるが、入力時の文書内容チェック機能をもたないという問題点がある。また、[2]では、入力枠をもつ HTML 文書と、その入力枠中の文字列を SGML 文書に写像する CGI(Common Gateway Interface)関数を DTD 毎に Perl 等のプログラムにより記述するので、DTD の変更や文書内容制約の修正/追加の際には直接ソースプログラムを修正する必要があり、保守性が低いという問題点があった。

3. 文書交換による SGML 入力用 HTML の生成方式

本稿では、SGML 文書変換技術を用いて、2.1の(1)と(2)を共に満たす SGML 文書入力方式を提案する。

3.1 実現方式

図1に基づいて、本方式の実現方針を述べる。

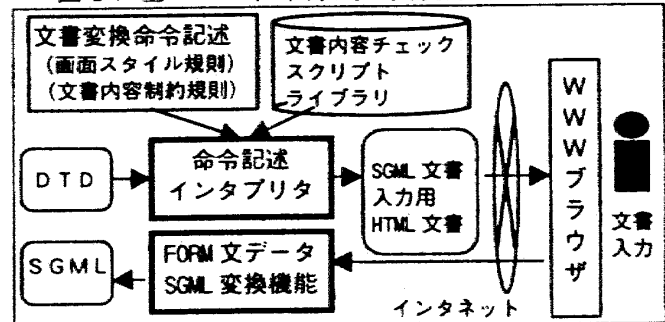


図1 SGML 入力用 HTML 生成方式の概念図

処理は、入力対象とする DTD にエレメントの繰り返し

しやORを含まないような決定的なDTDとそうでない非決定的なDTDとで異なる。以下では、決定的なDTDの場合の処理を示す。非決定的なDTDの場合については、3.3に述べる。

- ① 文書タグ毎に、入力枠の表示スタイルを指定する画面スタイル規則と入力内容の文書内容制約規則とを記述する文書変換命令記述を導入する。
- ② 命令記述インタプリタは、DTDと文書変換命令記述を入力として、DTD上に出現する文書タグ毎に①の文書変換命令記述を実行することにより、SGML文書入力用のHTML文書を生成する。但し、文書内容チェック機能はJavaScriptプログラムにより実現する。このプログラムは、文書内容チェックスクリプトライブラリから文書内容制約規則に対応するひな形を参照して生成する。
- ③ FORM文データSGML変換機能は、SGML文書編集者が②のHTML文書の入力枠に入力した文書内容をFORM文データとして受け取り、SGML文書を生成する。但し、命令記述インタプリタは、HTMLの入力枠とSGML文書タグとの対応関係を②のHTML文書中にhiddenタグの内容としてもたせることにより、FORM文データSGML変換機能は入力対象のDTDに依存しないようにする。

本方式では、DTDの変更や文書内容制約の修正/追加の際には文書変換命令記述だけを修正するだけで済むので、従来技術[2]の問題点であった保守性を向上させることができた(従来は、入力用HTML文書とFORM文データSGML変換機能の両者を修正)。

3.2 文書変換命令記述

3.1①で述べたように文書変換命令記述では、文書タグ毎に画面スタイル規則と文書内容制約規則を指定する。

(1) 画面スタイル規則

文書タグ毎に、以下を指定する。

- ・ 入力枠生成に必要なHTMLのinputタグの属性値(入力枠のタイプやサイズを指定する)
- ・ inputタグの前に挿入するHTML文(入力枠の名前やインデント等の情報を指定する)
- ・ inputタグの後に挿入するHTML文

(2) 文書内容制約規則

文書内容制約規則の種類を以下に示す。

- ① 初期値/強制値設定規則
- ② 単項目チェック規則
- ③ 項目間チェック規則(複数の文書タグ内容間の制約をチェック)
- ④ 文書入力をガイドするためのヘルプ規則

個々の規則毎に、その規則の起動のタイミング、起動条件、表示メッセージ、メッセージ表示先等を指定する。図2は、「サイズ」タグと「タイプ」タグ間の制約を表現する項目間チェック規則の例である。

```

サイズ [
  check.on = (onchange,
    if($<タイプ> == "A" | $self < 10 |,
      "タイプがAの時にはサイズは10以上でなければなりません",
      parent.GUIDE)
  ]

```

図2 文書内容制約規則の記述例

3.3 非決定的なDTDの場合の処理

以下の二つの方式がある。

- ① 入力用HTML文書生成の前に、非決定的な部分をすべてユーザに選択させることにより、3.1の場合に帰着させる。
- ② DTDから文書構造を受理するオートマトンを生成し、命令記述インタプリタがそのオートマトンをトレースし、非決定的性がある場合には、対話的にユーザに選択させることで、文書構造を決める。

4. おわりに

文書内容制約の入力時チェック機能と文書変換命令記述による保守性の向上を特徴とするSGML文書入力方式について述べた。今後の課題は、文書変換命令の記述能力を拡張し、「データベースとのデータ連携命令」や「文書タグ毎のアクセス権制御命令」を記述できるようにすることにより、SGML応用システム構築支援のためのSGML文書処理方式を確立することである。

参考文献

- [1] 今村 誠 他：SGML文書の内容検証方式の検討，情処学会デジタルドキュメント研究会6-4, P25~32(1997)。
- [2] 譚 澤華 他：SGMLデータベースを活用した質問回答システム，第54回情処学会全国大会, 3-309~3-310(1997)。