

SGML 文書の内容検証方式とその評価

今村 誠 森口 修 鈴木 克志 辻 秀一

E-mail: {imamura,mog,suzuki,htsuji}@isl.melco.co.jp

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

SGML では DTD により文書構造を業務に即して表現できるので、文書の SGML 化により業務システムでの文書データの活用が容易になる。しかし、検索や数値処理を行う業務システムでのデータ活用を前提とする SGML 文書交換では、文書交換のコミュニティ内で、文書構造の記述規約である DTD を共有するだけでなく、文書中の用語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の記述規約をも共有する必要がある。本稿では、企業間で交換する文書がコミュニティ内の文書内容記述規約に従っているという認定を与えるための文書内容検証方式を提案する。さらに、電子部品の部品カタログを対象とする文書内容検証ツールの開発と利用実験を通じて、文書内容検証合格の保証付き文書を流通させることにより、コミュニティ内での信頼性の高い文書情報の共有を促進できることを確認した。

A Method of Content Validation of SGML Documents and its Evaluation

Makoto IMAMURA Osamu MORIGUCHI Katsushi SUZUKI Hidekazu Tsuji

E-mail: {imamura, mog, suzuki, htsuji}@isl.melco.co.jp

Mitsubishi Electric Corporation
Information Technology R & D Center
Human Media Technology Dept.

SGML enables us to use document information effectively as data in application systems by DTD. In case we exchange documents among companies to use their data in application systems such as retrieval or calculation, not only DTD but also descriptive rules about document contents such as standard vocabularies, range of numerical value and unit of measure are necessary to share in a community of document exchange. In this paper, we propose a method of document validation to authorize documents to follow the descriptive rules in a community of document exchange. It promotes a reliable document exchange to circulate validated document to the content descriptive rules. Further we discuss the effectiveness of our method through development of content validation software and its experiments in electronic component industry field.

1. はじめに

WWW(World Wide Web)やJava等のネットワークインフラ技術の普及と 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。この要求にこたえるための文書形式が、CALS 標準の SGML(Standard Generalized Markup Language: ISO8879)や SGML のサブセットである XML(eXtensible Markup Language)である。

SGML や XML(以下では、SGML で代表する)は、表示や印刷だけを目的とした形式ではなく、プログラムが解釈できる形式でもある。例えば、文書中から必要な情報を機械的に抽出しデータベースに自動登録したり、文書の論理構造を組替えて別文書を生成する等の処理が可能になる。従って、文書を SGML 化することにより、文書情報を、ネットワークを介して業務システムから操作することが容易になる([1][2])。ネットワークを介した SGML 文書の業務システムでの活用案としては、以下がある。

(a) 製品販売企業と製品購入企業間の製品情報共有

製品販売企業各社が WWW 上で公開する SGML 形式の製品情報を収集し、製品購入企業のデータベースに自動登録し、販売企業各社の製品仕様を比較検索するシステム[3]。

(b) 営業部門と設計部門の製品仕様共有

製品顧客からの要求仕様を SGML 化した後、各々の営業所から設計部門へ伝送し、設計部門の設計支援システムのデータとして活用するシステム。従来の SGML の利用では、技術マニュアルをターゲットとするものが多く[4]、文書構造を規定する DTD(Document Type Definition)を共有することにより、企業間での文書交換が可能であった。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([5])。さらに、この文書内容制

約は、文書情報を交換/共有するコミュニティ内で、文書記述規約として共有される必要がある。例えば、CFI(CAD Framework Initiative)の ECIX(Electronic Component Information eXchange)プロジェクトでは、半導体部品のデータを電子的に配布する形式として SGML を採用し、文書の記述規約として、DTD だけでなく、カタログ中で使用する専門用語を辞書により規定している[6]。

本稿では、企業間で交換する文書がコミュニティ内の文書内容記述規約に従っているという認定を与えるための文書内容検証方式を提案する。文書内容検証合格の保証付き文書を流通させることにより、コミュニティ内での信頼性の高い文書情報の共有を促進できる。

以下、2章では、SGML による文書情報共有の特徴を整理した上で、文書内容検証の役割、必要機能、及び必要機能を実現するための要素技術の現状レベルを述べる。3章では、電子部品の分野を例にとり、文書内容検証方式の機能詳細と、その機能の実現方針を述べる。4章では、文書内容検証機能の実装について述べる。5章では、実現した方式の機能、処理性能、有用性、及び使いやすさを評価する。6章では、今後の課題について考察する。

2. 文書内容検証の課題

2.1 SGML による文書情報共有の特徴

SGML を従来のデータベースや HTML(HyperText Markup Language)による情報共有と比較すると、以下の特徴がある(表 1)。

- ・情報共有の目的としては、HTML が一般的な情報の共有を目的とするのに対して、SGML では、業務向け情報の共有を目的とする。
- ・情報共有する利用者の範囲としては、データベース(DB)の利用者がクローズであるのに対して、SGML では、DTD 共有により不特定多数の企業間でのオープンな文書情報共有である。
- ・共有される情報の質としては、SGML では、DB における数値やコードに代表される均質で整理さ

表 1 ネットワークによる情報共有を実現する上でのデータの形式の特徴

	DB	HTML	SGML
情報共有の目的	業務向け特定情報の共有	種々雑多な情報の共有	業務向け特定情報の共有
情報共有の範囲	クローズ	オープン	業務コミュニティ内でオープン
共有される情報の質	均質で整理されたプログラム解釈可能な情報	均質でない人間が解釈するテキスト情報	均質で整理されたプログラム解釈可能な情報と人間が解釈するためのテキスト情報の混在

れたプログラム解釈可能な情報だけでなく、人間が解釈するためのテキスト情報も混在している。

2.2 文書内容検証の役割

前節で述べたように SGML による文書情報共有では、従来の文書交換と異なりプログラムが解釈可能な情報を交換するので、従来クローズであったデータベース情報を業務特定のコミュニティ内でオープンに利用することになる。このような業務システムとのデータ連携を前提とするオープンな文書情報共有を実現するには、文書情報の利用形態を想定したそのコミュニティ内の規約に従った文書を交換しあうことが重要になる。このための機能が文書内容検証である。文書内容検証とは、流通する文書がコミュニティ内の規約に確かに従っているという認定を与える機能であり、文書内容検証合格の保証付き文書を流通させることにより、コミュニティ内での信頼性の高い文書情報の共有を促進できる。

従来の定型業務向けシステム(グループウェア等)では、利用者範囲が明確に定められており、かつ業務内容が定型的であり、ほとんど変化がないという特徴があった。そのため、利用するツールを固定することで、業務システムが許容するデータであることを入力時にチェックすることが可能であった(オーサリングツールによる文書内容の入力時チェック)。しかし、不特定の企業間、あるいは大企業部門間で共有する文書を作成する場合には、以下の理由によりオーサリングツールによる文書内容の入力時チェックは困難である。

- ・ あらかじめツールを固定することが困難。
- ・ 規約の改定結果をツールに反映させることを保証するのが困難。
- ・ 少々の違いを許容する、ゆるやかな規約に対応するのが困難。

2.3 設計業務と文書内容検証に必要な機能

本節では、電子部品の分野を例にとり、業務特定のコミュニティ内でのオープンなネットワークでの文書情報共有における業務機能の概略と、必要な文書内容検証について述べる。

2.3.1 設計業務の概略と部品情報の利用

設計業務の概略と設計業務における部品情報の利用について、プリント基板の設計を例にとり述

べる。プリント基板の設計は以下の4業務よりなる。

(1) 概要設計

既存類似部品の機能から部品を想定し、社内認定部品及び新規部品を調査する。新規部品に関しては最新の部品カタログや代理店から情報を収集する。

(2) EDA(Electronic Design Automation)ライブラリ作成

概要設計で選定した部品の回路シンボル、フットプリント等のライブラリを作成する。ライブラリの作成には、部品のデータブックの情報を利用する。

(3) プリント基板回路設計・実装設計

EDA ライブラリを元にして、基盤製造データや実装図、組み立て図、及びシルク図等の図面を作成する。

(4) プリント CAE(Computer Aided Engineering)解析

基盤設計図面と EDA ライブラリを元にして、回路解析、伝送線路解析、電気解析、熱解析、及び電磁波解析等の CAE 解析を行い、(3)で設計した図面を検証する。

2.3.2 文書内容検証に必要な機能

前項で述べたように、概略設計での部品選定と EDA ライブラリの作成時に部品カタログを参照する。部品カタログの内容検証機能として以下が必要になる。

(1) 文書構造の検証

部品カタログから部品の型名、仕様、及び機能等を機械的に自動抽出できるような標準化された文書構造を有しているかどうかを検証する。

(2) 文書内容の検証

自動抽出した部品仕様を表現するための部分類、部品パラメータ名称、及び部品パラメータ値の単位が部品メーカー間で標準化された用語で記述されているかどうかを検証する。

(3) EDAライブラリの検証

部品カタログ中に含まれる、または参照される EDA ライブラリが対応する CAD ツールで確かに利用できるものかどうかを検証する。

2.4 機能と要素技術の現状レベル

前項の必要機能に対応させて、機能の現状レベルについて述べる。但し、本稿では、検討範囲を SGML 文書の内容検証として、EDA ライブラリの検証については検討範囲外とする。

(1) 文書構造の検証

部品の型名、仕様、及び機能等を文書タグとして DTD により規定し、その DTD を満たすように SGML 形

式の部品カタログを作成することにより、部品の型名、仕様、及び機能等が機械的に自動抽出できるようになる。文書構造の検証は、SGML 文書が DTD に従っているかどうかを検証することであり、これはいわゆる SGML パーザの機能である。

(2) 文書内容の検証

SGML 文書のタグの内容を扱う要素技術としては、SGML 文書変換技術がある。SGML 文書変換の方式は、SGML パーザによる文書解析の結果得られる文書構造木をたどりながら、文書タグ毎に指定された変換命令を実行するイベント駆動アプローチと、SGML 文書の構造解析結果の木構造に対して、直接的な操作を指定する木構造操作アプローチとに大別できる。

イベント駆動アプローチをとる文書変換ツールの例として、OmniMark[8]と AEsop [9]がある。文書内容検証処理をイベント駆動アプローチにより実現すると以下の問題点がある。

- ・処理命令の実行順序が文書構造木をたどる順に依存するので、タグ間の値チェックなどの文書の部分構造間の制約のチェックを簡潔に記述することができない。
- ・従来ツールは、1入力文書を想定しているので、複数の文書間の制約をチェックすることができない。

木構造操作アプローチの文書変換の例として、ISO規格である DSSSL(Document Style Semantics and Specification) [ISO/IEC 96]の Standard Document Query Language (SDQL)がある。SDQL を用いれば、SGML 文書構造の操作に必要な基本関数とリスト処理機能を用いて、割付け処理で必要とされる複雑な変換処理を記述できる。しかし、内容検証処理の記述言語としては、SDQL が提供する基本関数はプリミティブなものに限定されているという問題点があった。

3. SGML 文書の内容検証方式の実現方式

本章では、電子機器メーカー、半導体メーカー、一般電子部品メーカー各社はじめ EDA ツールベンダーなど 43社が参加する電子機器・部品 CALS 実証実験プロジェクト (E-CALS) における電子部品カタログの情報共有の場合を例にとり、文書内容検証方式の機能と実現方針について述べる。

3.1 文書内容検証方式の機能

本節では、SGML 形式の電子部品カタログが E-CALS における電子部品カタログの記述規約である標準類

(記述規約標準類)を満たしているかどうかを検証するツールである文書内容検証ツールの運用形態、ねらい、及び機能について述べる。

3.1.1 電子部品カタログの記述規約標準類

E-CALS では、ECIX を参考にして、電子部品カタログ記述規約の標準を規定している。SGML 形式の部品カタログ記述に関する標準は、文書構造を規定する DTD である ECIS(E-CALS Component Information Standard) と、文書内容を規定する 3 つの用語辞書(部品分類辞書、部品パラメータ辞書、分類・パラメータ対応辞書)と内容検証規則からなる。ECIS は、ECIX の PCIS (Pinnacles Component Information Standard)1.2 版を参考にしている。ECIX との主要な差異は以下の 3 点である。

- ・部品カタログ中に部品分類を記述するタグを追加した。
- ・部品分類と部品パラメータの対応関係を規定する分類・パラメータ辞書を有する。
- ・部品カタログ中のタグの内容と辞書との対応関係を規定する内容検証規則を有する。

なお、電子部品カタログの記述規約標準類の詳細は、参考文献[9]に譲る。

3.1.2 文書内容検証ツールの運用形態

電子部品カタログの情報を共有するコミュニティは、部品情報を提供する電子部品メーカーを想定する情報提供側、部品情報を活用する電子機器メーカーを想定する情報活用側、及び部品情報に関する標準類を維持管理する情報センタを想定する情報流通側からなる。文書内容検証ツールの管理と提供は、カタログ記述規約標準類と共に情報流通側でなされることを想定する。文書内容検証ツールの運用手順フローを図 1 に示す。

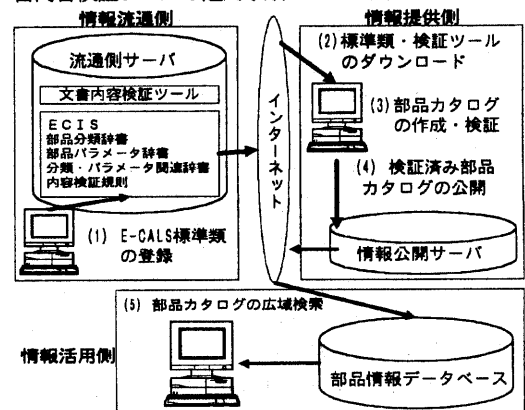


図 1 文書検証ツールの運用手順フロー

(1) E-CALS 標準類の登録

情報流通側の部品情報の維持管理者は、標準類が改定される毎に、最新の標準類を流通側の情報公開用サーバに登録する。

(2) 標準類・検証ツールのダウンロード

部品カタログを作成する提供側は、流通側サーバにアクセスし、標準類と文書内容検証ツールをダウンロードする。

(3) 部品カタログ作成・検証

(2)によりダウンロードした文書内容検証ツールを用いて、SGML形式の部品カタログが標準類に適合しているかどうかを検証する。検証の結果、部品カタログにエラーが含まれている場合は、検証結果メッセージを参考にしてエラー箇所を同定し、部品カタログを修正する。

(4) 検証済み部品カタログの公開

標準適合性検証に合格した部品カタログを情報提供側の情報公開サーバに登録する。

(5) 部品カタログの広域検索

情報活用側は、(4)により公開されている部品カタログをダウンロードし、その部品カタログから部品の型名、機能、仕様などを機械的に抽出し、情報活用側の部品情報データベースに登録する。

3.1.3 ねらい

部品情報の提供者が文書内容検証を行った上で部品カタログをインターネット上に公開することで、部品情報の活用者が必要とする情報が共通の形式で記述されていることが保証されるので、部品情報の会社間での共有をより進めることができる。

(1) 部品情報の提供側にとっての利点

・部品情報の提供者側は、本ツールの出力するエラーや警告メッセージを参照することにより、部品カタログが標準に適合しているかどうか、また、どの部分が適合しないかを知ることができる。したがって、部品カタログをインターネット上に公開することが容易になる。

(2) 部品情報の活用側にとっての利点

・電子部品情報の活用者は、公開された各社の文書内容検証済みの電子部品カタログ情報から部品の型名、機能、仕様などを機械的に抽出できるので、活

用側の部品データベースへの自動登録が可能となる。

・電子機器設計者は、上記の手順で自動登録された部品データベースを参照する検索ツールを使って、部品選択時に、E-CALS標準用語に基づく条件指定で複数部品メーカーの類似品を同時に比較検索したり、検索結果の各社の電子部品情報をダウンロードして利用できるようになる。

3.1.4 検証機能

本項では、文書内容検証ツールの機能について述べる。ツールの機能は、構造検証機能、内容検証機能、及び検証結果提示機能からなる。

(1) 構造検証機能

構造検証機能は、電子カタログが DTD を満たしているかどうかを検証する。

(2) 内容検証機能

内容検証機能は、内容検証規則に従って、電子カタログ中の文書内容が、部品分類辞書、部品パラメータ辞書、及び分類・パラメータ関連辞書に従っているかどうかを検証する。以下で、内容検証規則による検証内容について述べる。

(a) 部品分類に関する検証規則

① 分類コード検証規則

部品カタログ中の部品分類コードが部品分類辞書に登録されているかどうかを検証する。

② 分類名称検証規則

部品カタログ中の部品分類名称が部品分類辞書に登録されているかどうかを検証する。

③ 分類コード・名称整合性検証規則

部品カタログ中の部品分類コードと部品分類名称の組みが、部品分類辞書に登録されているものと同じかどうかを検証する。

(b) 部品パラメータに関する検証規則

④ 部品パラメータコード検証

部品カタログ中の部品パラメータコードが部品パラメータ辞書に登録されているかどうかを検証する。

⑤ 部品パラメータ名称検証

部品カタログ中の部品パラメータ名称が部品パラメータ辞書に登録されているかどうかを検証する。

⑥ 部品パラメータコード・名称整合性検証

部品カタログ中の部品パラメータコードとパラメータ名称の組みが部品パラメータ辞書に登録され

ているものと同じかどうかを検証する。

- ⑦ 部品パラメータコード・データ型整合性検証
部品カタログ中の部品パラメータの値が、部品パラメータ辞書のデータ型と一致するかどうかを検証する。
- ⑧ 部品パラメータコード・単位整合性検証
部品カタログ中の部品パラメータの値の単位が、部品パラメータ辞書の単位と一致するかどうかを検証する。
- (c) 部品分類とパラメータとの関係に関する検証規則
- ⑨ 必須パラメータ検証
部品分類・パラメータ関連辞書が規定する必須パラメータがすべて部品カタログ中に記載されているかどうかを検証する。
- ⑩ 任意パラメータ検証
部品カタログ中に記載される部品パラメータが部品分類・パラメータ関連辞書の必須または任意パラメータとして登録されているかどうかを検証する。

(3) 検証結果の提示機能

文書内容検証ツールでは、検証結果に関する以下の内容を含むメッセージを出力する。

- ・検証規則違反の種別(エラーまたは警告)
- ・検証規則違反の内容
- ・違反の検出された箇所
- ・検証に用いた辞書の名称及び版

3.2 実現方針

文書内容検証機能の実現方針は、検証仕様を表現する内容検証規則を簡潔に記述できる文書構造操作言語を導入し、その言語を解釈実行することにより内容検証機能を実現することである。標準適合性検証ツールの中心となる内容検証処理は、部品カタログの文書構造を操作する処理が中心である。そこで、SGML 文書の文書構造中の必要な部分構造の取り出し・比較等の基本演算式から構成される文書構造問い合わせ言語を導入した[5]。文書構造問い合わせ言語による文書内容検証処理の実現方針は、以下の通りである。

- (1) 検証用辞書(分類, パラメータ, 分類・パラメータ関連)を SGML により表現する。
- (2) 内容検証規則を、SGML 文書の文書構造中の必要な部分構造の取り出し・比較等の基本演算式から構成される構造化文書問い合わせ言語を用いて表現する。
- (3) 検証仕様記述を解釈実行することにより検証する。

本実現方式には、以下の特徴がある。

- ・従来ツールでは困難であった「SGML 文書の部分構造間の比較」や「複数文書にまたがる文書内容の制約関係の判定」といったグローバルな文書構造操作を記述することができるので、内容検証規則を簡潔に記述することができる。従って、検証用規則の改定時も、この問い合わせ言語による規則記述を修正するだけで済むので、直接 C 言語等のプログラム言語により実現した場合と比較して修正が容易である。
- ・内容検証処理は、部品カタログと辞書とを照合する処理がなされるが、辞書類も検証対象の文書形式と同じ SGML で記述するので、その照合処理を統一的な枠組みで記述できる。

なお、文書構造問い合わせ言語の詳細は、参考文献[5]に譲る。

4. 実装

4.1 開発環境と実行環境

SGML 文書内容検証ツールは、プログラミング言語 C++ で実装し、ソースコードは約 20K 行(コメントを除く)である。実行環境は、OS は Windows95、メモリ 32Mbyte 以上とした。

4.2 検証処理のフロー

SGML 文書内容検証ツールにおける検証処理フローを図 2 に示す。処理は、構造検証処理、内容検証処理、及び検証結果整形処理からなる。

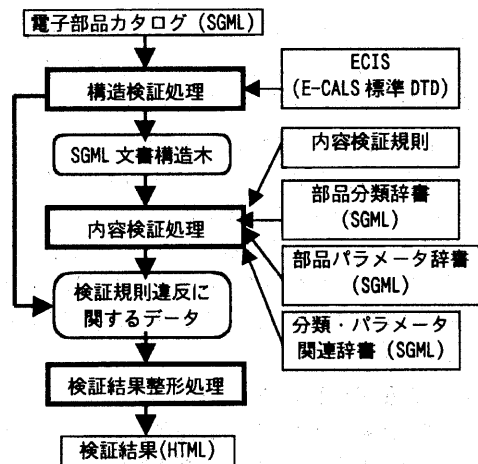


図 2 SGML 文書内容検証ツールの処理フロー

(1) 構造検証処理

検証対象の電子部品カタログを入力として、E-CALSの標準 DTD である ECIS を参照しながら、入力に対する SGML の文書構造木を生成する。入力の電子部品カタログが ECIS を満たしておらず文書構造木が生成できない場合には、そのエラー情報を検証規則違反に関するデータとして出力し、検証結果整形処理を行う。

(2) 内容検証処理

内容検証規則と用語辞書類を参照しながら、入力として与えられた SGML 文書構造木が内容検証規則を満たすかどうかをチェックし、その結果を検証規則違反に関するデータとして出力する。

(3) 検証結果整形処理

内容検証規則違反に関するデータを入力として、検証規則違反の内容とその違反箇所をハイパーリンクとして表現する HTML 文書を生成し、WEB ブラウザに表示する(図 3)。

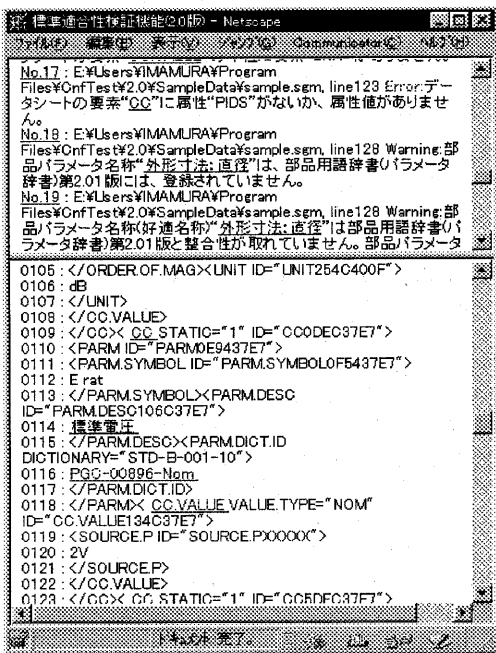


図 3 文書容検証エラーの出力画面例

5. 評価実験

5.1 実験環境

E-CALS プロジェクトにおいて、図 1の運用手順フ

ローに従って、文書内容検証ツールの利用実験を行った。本ツールは E-CALS プロジェクト参加の情報提供者側企業が流通側サーバからダウンロードして自由に利用できるようにした。

また、実験手順は、以下の作業(1)と作業(2)を順に行い、(2)でエラーが検出された場合は、(1)に戻って再編集した。

- (1) SGML エディタによる部品カタログ作成
- (2) 検証ツールによる部品カタログチェック

5.2 評価結果

以下、ツールの検証機能、ツールの処理性能、ツールを用いた部品カタログ作成時間、及び有用性と使いやすさについて述べる。

(1) ツールの検証機能

3.1.4 項で述べた検証機能の動作を確認した。以下に示すように、検証に合格した部品カタログは、部品カタログ中の部品情報を部品データベースに自動登録し、さらに E-CALS の標準用語により複数部品メーカーの部品仕様を比較検索できることが保証される。

- ・検証規則①,④,⑨,及び⑩により、部品カタログ中データが、部品分類に依存して決まる部品データベースのスキーマと対応していることを保証する。
- ・検証規則⑦により、部品カタログ中のデータが部品データベースに登録できることを保証する。
- ・検証規則②,③,⑤,⑥,及び⑧により、各社の部品カタログが E-CALS の標準用語により比較検索できることを保証する。

(2) ツールの処理性能

以下に、実証実験で用いられた電子部品カタログのサンプルでの検証処理時間(表示を含む)を示す。但し、実験環境 PC の CPU は PentiumPro(クロック 266Mhz)、メモリは 255Mbyte である。

表 2 文書内容検証ツールの検証処理時間

	ファイルサイズ	検証データ数	エラー数	検証時間
カタログ1	8 Kbyte	115	27	1.2 秒
カタログ2	41 Kbyte	504	69	2.3 秒
カタログ3	201Kbyte	2765	795	9.6 秒
カタログ4	652Kbyte	10246	1670	32.2 秒

また、検証用に用いた辞書のサイズは表 3に示す通りである。

表 3 検証用辞書のサイズ

	ファイルサイズ	語彙数
部品分類辞書	264Kbyte	1014
部品パラメータ辞書	931Kbyte	988
分類・パラメータ関連辞書	51Kbyte	1083(関連対の数)

(3) 部品カタログ作成時間

部品情報の提供者側企業における部品カタログの作成に要する時間に関する結果を表 4 に示す。

表 4 部品カタログ作成に関する時間

SGML エディタによる部品カタログの新規作成に要する平均時間(1件目)	204 分
SGML エディタによる部品カタログ作成の新規作成に要する平均時間(作成に慣れた後)	96 分
検証ツールによるエラー検出による部品カタログの再編集の平均回数	4.0 回
検証ツールによるエラー個所の修正に要する平均修正時間	39 分

(4) 有用性と使いやすさ

以下に、アンケート結果得られた結果を示す。括弧内の%は、肯定的な回答の割合である。

- ・記述規約標準類に準拠する部品カタログを作成する上で検証ツールのメッセージは参考になった(92%)。
- ・検証メッセージはほぼ理解できるものであった(「わかりやすい」、または「ほぼわかる」との回答が69%)。
- ・検証から結果提示に要する時間は、利用上気にならないレベルであった(92%)。

6. おわりに

本稿では、業務特定のコミュニティ内で、コミュニティ内の文書記述規約に確かに従っているという認定を与える機能である文書内容検証方式を提案した。そして、E-CALS の実証実験を通じて、文書内容検証合格の保証付き文書を流通させることにより、コミュニティ内での信頼性の高い文書情報の共有を促進できることを確認した。以下では、今後の課題について述べる。

(1) 標準の用語辞書を効率的に策定する枠組み

部品メーカー間で部品パラメータの測定条件が異なるために同じパラメータでも厳密には意味が異なっていたり、新機能をもった電子部品が出現するなどのため、標準類の辞書を策定することは手間のかかる作業である。そこで、標準となる辞書をより効率的に

策定する枠組みが望まれる。そのための支援ツールとしては、流通している SGML 形式の部品カタログから、部品パラメータや定義等を自動抽出し、標準用語の候補を整理するものが考えられる。

(2) 複合電子部品情報の検証機能

電子部品情報としては、今回の検証の主対象であった部品カタログのパラメータ情報だけでなく、人間が参照する PDF データ、電気設計用 EDA モデル、及び実装設計用 CAD データ等がある。SGML は、これらの複合的な電子部品情報を束ねる役割をになっもいる。タグ内容のテキスト情報だけでなく、これらの複合電子部品情報の適合性を検証する機能と連携することにより、より総合的な部品情報の検証が可能になる。

謝辞：文書内容検証ツールの実証実験に協力いただいた E-CALS の皆様に感謝いたします。電子機器・部品 CALS 実証実験プロジェクト(E-CALS)は、生産・調達・運用支援統合情報システム技術研究組合(NCALS)が、通商産業省の指導のもと情報処理振興事業協会(IPA)からの発注を受け推進する企業間高度電子商取引推進事業「CALC 実証実験」の一つである。

[参考文献]

- [1]石塚 英弘, 根岸 正光: 情報システム基盤技術としての SGML, 情報処理, Vol.37, No.3, pp207-212(1996)
- [2]村田 誠編著: XML入門, 日経新聞社(1998)
- [3]森口 修, 今村 誠, 鈴木 克志: 製品情報広域検索システムにおける検索方式, 情報処理学会デジタルドキュメント研究会資料 5-4, pp25-32(1997)
- [4]特井 浩, 小川 満広: SGML の概要, 日経バイト, 1995 年 8 月号, pp164-170(1995)
- [5]今村 誠, 森口 修, 鈴木 克志: SGML 文書内容検証方式の検討 情報処理学会 デジタルドキュメント研究会資料 6-4, pp25-32(1997)
- [6]Donald R. Cottrell: ECIX - Electronic Component Information Exchange, <http://www.cfi.org/ecix/toledo/index.html>, 1997
- [7]OmniMark Programmer's Guide Version 2, Exoterica Corporation (1993)
- [8]高橋 亨, 松本 正義: SGML 文書を対象とする文書構造操作言語の提案, 情報処理学会第 49 回全国大会, 3 分冊, pp265-266(1994).
- [9]今村 誠, 大師堂 清美: E-CALS における電子部品カタログ情報の標準適合性検証機能, CALS EXPO INTERNATIONAL 1997 論文集, T2-18(1997)