

3S-6

情報処理学会第54回全国大会

SGML文書と差分情報とから指定の版を合成する実現方式
若島陸夫, 田中道夫
CALS技術研究組合(NCALS)

1. はじめに(Introduction)

1.1 全体の位置付け(Position in the Project)

日本国工業組合法による生産・調達・運用支援統合情報システム技術研究組合(略称=CALS技術研究組合, 英略称=NCALS)は, 次世代のデジタル化文書による生産・調達・運用の支援を研究している。その中でも, 筆者らは, 電子出版技術を研究対象とし, 業界を超越したデジタル化文書の交換を考慮したSGML文書型定義による様式を提案し, その実証実験を継続している[1], [2], [3]。

1.2 利用者要件(User Requirements)

- その電子出版の中でも, 技術文書では, 次の利用者要件が高い。
- a) 企業内及び業界内での特定の様式のデジタル文書から, 中立的な文書様式に変換して, 受信者へ送り文字及び映像を再生したい。
 - b) 文書内容の改定要望・技術改良を複数箇所から集め, その意見一覧を文書に添付したい。
 - c) 送付文書は, 最新版を省略時値として送るが, 受信者がその最新版を参照したい場合, 前の版に戻れるようにして欲しい。
 - d) 文書作成・編集は, その内部記述様式であるSGML-DTDの深い知識がなくてもできる高い利用者対機械間の界面を具備してほしい。

2. 改版情報の交換(Interchange of Version Information)

ここでは, 利用者要件の内, 異なる処理系間で改版情報を交換し, その改版情報を基に他の版も再生できる仕掛けを提案する。

2.1 改版制御の仕掛け(method for version control)

2.1.1 改版方法(methodology)

一般的に, 文書の改版は, 次のa)~c)のいずれかの方法又は組み合わせによって, 利用者に提示されている。

a) 版ごとの個別格納(no version-down)

個々の版を個別に格納し, 版の指定によって, 指定の版を再生する。検索・再生は, 煩であるが, 格納資源及び分散・遠隔格納の通信資源を消費する欠点がある。この方式は, 改版が稀にしか実施されない文書, 提示を主体とした文書などに適する。

例: HTML記述文書

b) 前向き改版(forward version-down)

初版からの変更履歴で書き換えながら, それより後の版を生成する方法である。版が重なるにつれて, 再生のための処理が増加する欠点がある。一方, 格納資源は, 初版と各版の間の差分情報とだけでよい利点がある。

新版文書 $\leftarrow$ (初版文書) + (差分1) ... + (差分 N)

c) 後ろ向き逆改版(backward version-down)

最新の版から最初の版に向かって, 一つ前の版に順次戻すことによって, 指定の版の文書を再生する方法である。すぐ, 前の版を再生する要件が高い応用に適する。再生したい版が初版に近いもの程, 逆改版に要する処理が増加する欠点がある。一方, 格納資源は, 最終版と各版の差分情報とに必要な資源だけでよい。

旧版文書 $\leftarrow$ (最新文書) - (差分 N) ... - (差分 1)

2.1.2 NCALSの改版伝達方式の選定(selected method in NCALS)

NCALSでは, 主に技術文書(仕様書, 技術連絡票, 製造資料, 運用手引き書など)を記述対象とすることから, 版ごとの個別格納方式と後ろ向き改版方式との混合利用によって, 前の版を自動合成する情報を交換することを選んだ。この方式は, 主要な改版の場合, 版を個別に格納し, 小変更の場合を後ろ向き改版情報によって, 自動合成できる基準で運用できる。このような, 改版運用ができれば, 大きく後ろ向きに逆改版する処理を圧縮できる。

なお, 主要改版と小改版との利用区別は, 対象とする文書種別, 業界, 運用方法の違いなどで偏位するので, 利用者側の決めに聞いておいた。

A Study of Version Control Methods for Electronic Technical Documents using SGML

Rick Wakatorl, Michio Tanaka

Nippon CALS Research partnership(NCALS)

c/o TIME24bldg. 24 Aomi(2), Kouto-Ku, Tokyo, JP135-73, JAPAN

2.2 逆改版情報の持ち方(version information in a document)

2.2.1 直前の発行情報にだけ、戻すための逆改版情報

NCALS-DTD適合文書では、直前の発行情報の段落内容(paragraph)まで戻す情報を変更命令書(chginsst)に削除(remove)及び挿入(insert)の順の組の並びによって、表現した箇所を示す。それによって、逆改版する場合、挿入された文面を削除し、削除された文面を挿入すればよいことを示す。

2.2.2 指定した版の文書へ戻すための逆改版情報

NCALS-DTD適合文書では、主要改版を別文書として扱うこととし、この逆改版の処理は、文書の探索機構に任せる。つまり、データベースには、各主要な版が格納されていることを前提とする。

2.3 全ての版及び発行情報を再生する場合の逆改版情報

全ての主要改版及び小改版を再生したい応用のために、NCALS-DTDでは、変更記録情報(chgrec)として、変更の並びを表形式で伝達する手段を用意している。この表に一定の様式を設定して、変更の都度、削除文面及び挿入文面を追記しておけば、それを後ろ向きで逆適用することによって、すべての版及び発行情報を生成できる。ただし、この方式の利用は、途中の変更状態(削除及び挿入)を正確に文書の前部(front/chgrec)へ記録しておく制約と、文書量が多くなる欠点がある。したがって、2.2.1及び2.2.2の組み合わせの方が短くなる。

これらのどれを応用するかは、利用者の属する業界で決めていただくつもりである。

3. 一般的な逆改版処理(reproduction of specified version)

一方、現在市販されているSGML処理系を概観すると、電子文書の改版は、特定処理系に埋め込まれた版情報によって、前向き改版、後ろ向き逆改版又は全ての版を保有しているかのどれかの形式に別れている。いずれをとっても、内部様式から(への)変換を必要としている。すなわち、それらの処理系は、同じ特定の様式で送信・受信が行なわれれば(同じ処理系を使えば)、改版制御を自在にできる半面、同じ処理系を期待できない他業界との間で、逆改版制御を自在にできない欠点をもつ。この異系統間の中継ぎの役が、NCALSが開発した文書型定義に込められている。

4. まとめ(conclusions)

NCALSでは、いつもは最新情報を提示し、指定された旧版も早く再生できる方式として、主要改版の場合に版、小改定の場合発行という概念を踏襲することによって、後ろ向き逆改版に適した文書型定義(DTD)を提案し、併せて、すべての変更履歴からどの版・発行の文書でも変更箇所と変更内容とを追跡できる文書中への表形式による表現様式とを提案した。

実際の処理系の逆改版のための内部処理では別な形態もありうるが、逆改版情報を受け渡す場合この様式へ変換し、受信者側は自己の処理系への変換をすることによって、異なるSGML処理系の間で、逆改版情報を含めて交換・処理・再生が可能となる。

NCALSでは、この思想に基付いた文書型定義(DTD)を公開し、更に、技術連絡指令(ECO)が外部処理を発火させることによって、この逆改版情報を付加したり、利用者側からの版の指定による逆改版文書の提示が可能かどうかなどについて、実証を続けている[4]。

なお、変更記録の情報の表の推奨様式は、次のURLによって、公開している。

<http://www.ncals.cif.or.jp/ncals/standards/chgrec.html>

謝辞 部分を含めた電子出版全体の研究活動に協力いただいている関係各位に感謝する。

参考文献(references)

- [1] 北村, 他: "NCALSのSGML文書データベースの開発方針と課題", 第53回全国大会論文集, 4-343, 情報処理学会, 1996
- [2] 若島, 他: "NCALSはん用文書型定義の開発方針と課題", 第53回全国大会論文集, 4-345, 情報処理学会, 1996
- [3] WAKATORI Rick, "Development of Geric document definition in NCALS", CALS EXPO '96
- [4] 北村: "SGML文書データベース間の変更管理共通方式の開発", 第54回全国大会論文集, 情報処理学会, 1997(予定)