

XML 文書変換アプリケーションの開発

寺門 正善 安井 浩之 松山 実†

武蔵工業大学‡

1 はじめに

近年，汎用的なデータ記述言語である XML(Extensible Markup Language)はインターネット上で交換可能なさまざまなデータ記述で使われており，XML 文書の変換機構の必要性が高まっている．^[1]

XML 文書の変換言語として W3C(World Wide Web Consortium) より勧告されている XSL(Extensible Stylesheet Language)は，文法が複雑であり高度な記法知識を要求される．また，変換後の文書の妥当性も保証されていない．

本報告では，シンプルな文法の変換言語を XML 上で提案するとともに，XML Schema に対応して変換後の文書の妥当性を保証する XML 文書変換アプリケーションの開発について述べる．

2 XML 文書変換言語：X2ML

2-1 X2ML

本稿で提案する X2ML(language for eXchange of XML)とは，XML 文書の変換方法を記述する言語(XML 文書変換言語)である．2つの文書の対応するルールを XML で記述し，同内容異構造文書(情報は同じだが XML Schema が異なる文書)間の双方向変換が効果的に実現できる．X2ML では変換手続きを記述するのではなく，対応ルールを記述する．この X2ML で記述した変換ルール文書を X2ML 文書とする．

2-2 X2ML 文書の仕様

例 1 のように，ルート要素は x2ml:rules で，子供に 1 つ以上の x2ml:pattern 要素を持つ．また，x2ml:pattern 要素は，子供に x2ml:from と x2ml:to 要素を 1 つずつ持つ．

例1 X2ML文書インスタンス

```
<?xml version="1.0" encoding="UTF-8">
<x2ml:rules xmlns:x2ml="
"http://www.tera.bounceme.net/x2ml/">
  <x2ml:pattern>
    <x2ml:from> <!-- from-pattern 1 --> </x2ml:from>
    <x2ml:to> <!-- to-pattern 1 --> </x2ml:to>
  </x2ml:pattern>
  ...
  <x2ml:pattern>
    <x2ml:from> <!-- from-pattern n --> </x2ml:from>
    <x2ml:to> <!-- to-pattern n --> </x2ml:to>
  </x2ml:pattern>
</x2ml:rules>
```

from-pattern には変換元文書の一部にマッチするパターンを記述する．変換元文書によって異なるテキスト要素などの部分構造は変数としてマッチさせる．to-pattern には変換先文書を構築するためのパターンを記述する．

3 XML 文書変換アプリケーション

ここで開発した XML 文書変換アプリケーションは，X2ML 文書の自動生成^[2]を行う X2ML ジェネレータと，X2ML 文書の処理系である X2ML プロセッサから成る．

3-1 X2ML ジェネレータ(図 1)

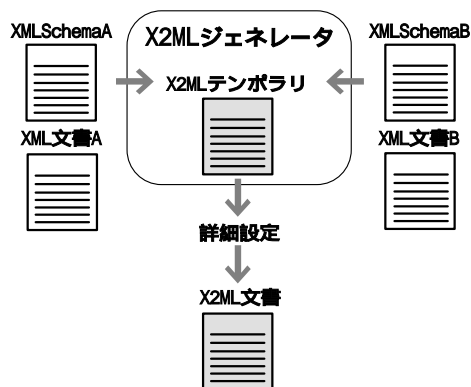


図1 X2MLジェネレータ

まず，XML 文書 A の XML Schema A と XML 文書 B の XML Schema B から X2ML テンポラリを自動生成する．次に，X2ML テンポラリに変数の対応などの詳細設定を行う(図 2)．最終的に例 2 のようなサンプル X2ML 文書を生成する．

Development of Application for XML Document Exchange

† Masayoshi Terakado, Hiroyuki Yasui, Minoru Matuyama

‡ Musashi Institute of Technology

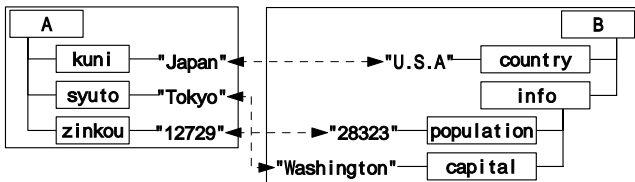


図2 サンプル文書の変数対応

例2 サンプルX2ML文書

```

<?xml version="1.0" encoding="UTF-8"?>
<x2ml:rules xmlns:x2ml=
"http://www.tera.bounceme.net/x2ml/">
  <x2ml:pattern>
    <x2ml:from>
      <A xmlns:xsi=
"http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="A.xsd">
        <kuni><x2ml:temp var="1"/></kuni>
        <syuto><x2ml:temp var="2"/></syuto>
        <zinkou><x2ml:temp var="3"/></zinkou>
      </A>
    </x2ml:from>
    <x2ml:to>
      <B xmlns:xsi=
"http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="B.xsd">
        <country><x2ml:temp var="1"/></country>
        <info>
          <population><x2ml:temp var="3"/></population>
          <capital><x2ml:temp var="2"/></capital>
        </info>
      </B>
    </x2ml:to>
  </x2ml:pattern>
</x2ml:rules>

```

3-2 X2ML プロセッサ(図3)

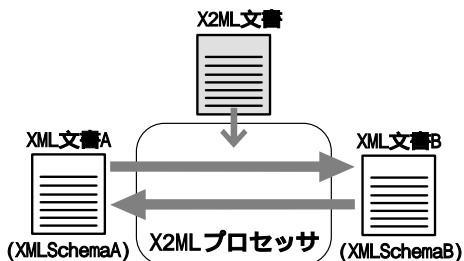


図3 X2MLプロセッサ

例2のX2ML文書によりXML文書AとBの間で双方向変換結果が得られる。(図4)

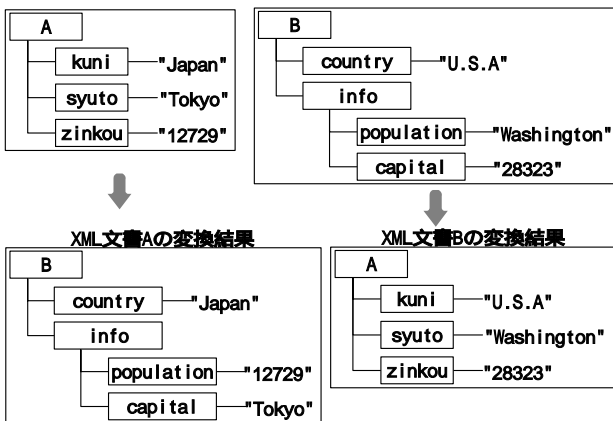


図4 サンプル文書の双方向変換結果

4 XSL との違い

4-1 生成文書の妥当性

X2ML 文書は XML Schema をもとに自動生成するので、X2ML 文書を用いての変換では、生成文書は必然的に妥当性が保証される。

4-2 変換方向

「パターンと手続き」で変換指定をする¹³⁾XSL に対し、X2ML では「パターンとパターン」で変換指定をするので、1つのX2ML文書で双方向変換が実現できる。

5 おわりに

5-1 実装

双方向変換を行う文書は、同内容異構造文書に限るという制限のもとで、XML Schema に対応して変換後の文書の妥当性を保証する、シンプルなX2MLによるXML文書変換アプリケーションを実装することができた。

5-2 今後の課題

現状では一部のXML Schemaにしか対応していないため、XSL との違いをより明確にさせるために、XML Schema に対するさらなる柔軟性を持つアプリケーションに拡張する必要がある。

また、変換によって欠落するデータが存在する場合、逆変換によって復元不可能であるため、可逆変換を可能とする拡張を検討する必要がある。

参考文献

[1] 丸山 宏, 田村 健人, 浦本 直彦 : XML と Java によるアプリケーション開発, ピアソン・エデュケーション, pp.33 ~ 56, 2000
 [2] 溝田 裕 : XML PRESS vol.1(JAXP/Xerces 実践活用術), 技術評論社, pp.97 ~ 145, 1999
 [3] Erick T. Ray : Learning XML, O'REILLY, pp.194 ~ 255, 2001