

# 計算機システム実験 K-11: XPath の利用

天笠俊之

2005 年 12 月 9 日

## 概要

XPath (XML Path Language) は, XML 文書の特定の部位を指示するための言語であり, XSLT や XQuery を始め, 極めて多くの仕様で用いられている. 今回は XPath の基礎を学習する.

## 1 準備

XPath の実習では, Apache プロジェクトによって開発が進められている XSLT プロセッサ Xalan-Java<sup>1</sup> を用いる. 最新版のパッケージは /home/lab/Denjo/amagasa/jikken/xalan に置いてあるが, 実習で必要なのは JAR ファイルだけなのに対してファイルの容量が大きいので (前回用いた Xerces も含めて) 特に必要な場合以外は JAR ファイル以外はコピーしないことを勧める.

必要な JAR ファイルは /home/lab/Denjo/amagasa/jikken/xalan にまとめておいてあるので, ここを直接参照するか, 各自のホームディレクトリ上にコピーした上で CLASSPATH を更新する.

/home/lab/Denjo/amagasa のファイルをそのまま参照する場合:

```
% setenv CLASSPATH ./home/lab/Denjo/amagasa/lib/java/resolver.jar:\
/home/lab/Denjo/amagasa/lib/java/xml-apis.jar:/home/lab/Denjo/amagasa/lib/java/xercesImpl.jar:\
/home/lab/Denjo/amagasa/lib/java/xalan.jar:/home/lab/Denjo/amagasa/lib/java/xalansamples.jar:\
/home/lab/Denjo/amagasa/lib/java/xercesSamples.jar:/home/lab/Denjo/amagasa/lib/java/serializer.jar
```

各自のホームディレクトリにコピーする場合:

```
% mkdir ~/lib
% cp -r ~/amagasa/lib/java ~/lib
% setenv CLASSPATH .:$HOME/lib/java/resolver.jar:$HOME/lib/java/xml-apis.jar:\
$HOME/lib/java/xercesImpl.jar:$HOME/lib/java/xalan.jar:$HOME/lib/java/xalansamples.jar:\
$HOME/lib/java/xercesSamples.jar:$HOME/lib/java/serializer.jar
```

なお, 環境変数の設定は .cshrc 等に入れておくと良い.

XPath の演習には, Xalan に付属している ApplyXPath クラスを用いる. これは, 引数に XML ファイルと XPath 式を取り, XPath 式に対応する部分を表示するプログラムである. 使い方は以下の通り:

```
% java ApplyXPath book.xml '//title'
Loading classes, parsing book.xml, and setting up serializer
Querying DOM using //title
<output>
<title>Database Management Systems</title>
<title>Readings in Database Systems</title>
<title>Foundations of Databases</title>
<title>Data on the Web: From Relations to Semistructured Data and
```

<sup>1</sup><http://xml.apache.org/xalan-j/>

```
XML</title>
</output>
%
```

[] や \* はコマンドシェルによって解釈されてしまう可能性があるので、XPath 式は必ずシングルクォートで括弧することに注意する。

## 2 実習課題

配布資料を読み、XPath について理解した上で、次に対応する XPath 式を考え、ApplyXPath で出力を確かめよ。なお、入力には先週の演習で作成した book.xml を使うこと。また、配布資料に含まれない XPath の機能については、<http://www.w3.org/TR/1999/REC-xpath-19991116>(和訳:<http://www.doranekeo.org/xml/xpath10/19991116/>)を適宜参照して欲しい。

### 2.1 単純な XPath 式

- Book 要素
- Book 要素の子の authors 要素
- Book 要素の全ての子要素
- authors 要素もしくは title 要素

### 2.2 軸を用いた表現

- authors 要素の全ての子孫要素
- title 要素の弟要素
- publisher 要素の兄弟要素

### 2.3 述語を含む表現

- 2 番目の person 要素
- 最後の person 要素
- 著者の firstname が “ Michael ” であるような本のタイトル
- 1900 年代に出版された label が “ Abi95 ” の本の著者の苗字 (familyname)
- タイトルに “XML” を含む本のタイトル, 出版社, 出版された年  
ヒント: contains() 関数を用いる。

上記 3, 4 番目の課題について、独自の内容でデータを作成した場合は、存在するデータに各自で適宜読み変えること。

## 3 課題の報告について

今回修得した内容を確認するための簡単なクイズを電子メールで送信しておいたので、空欄を埋めた上で、実験時間終了までに担当教官まで返送すること。

なお、不明な点や疑問点は、担当教官または TA に直接、またはメールで質問して欲しい。

amagasa@cs.tsukuba.ac.jp (担当教官)  
kido@kde.cs.tsukuba.ac.jp (TA)