

# 分散システム

## 第三回レポート課題

### 課題 1

#### 【課題内容】

講義資料(その2)ページ5に記載されているコードを参考に構文解析プログラムを作成し、第二回レポートで作成した HTML ファイルおよび XML ファイルが構文解析できるかどうか確かめよ。ただし、

- 構文解析に用いたソースコード
- 実行結果(スクリーンダンプにより)

を示すこと。また、実行結果は実行日時が分かるようにすること。

#### 【解析したソースコード】

##### *itemlist.html*

```
<html>
<head>
<title>Item List</title>
<meta http-equiv="Content-Type" content="text/html; charset=shift_jis">
</head>
<body>
<h2>Item List</h2>
<table border="3" width="300" cellspacing="0" cellpadding="5" bordercolor="#333333">
  <tr bgcolor="#CCCCCC">
    <th>Item Name
    <th>Value
  </tr>
  <tr>
    <td>File
    <td>¥120
  </tr>
  <tr>
    <td>Note
    <td>¥136
  </tr>
  <tr>
    <td>Pen
    <td>¥96
  </tr>
</table>
</body>
</html>
```

***itemlist.xml***

```
<?xml version="1.0" encoding="iso-8859-1"?>
<!DOCTYPE ItemList SYSTEM "itemlist.dtd">

<itemlist>
  <item>
    <name>File</name>
    <value>¥120</value>
  </item>
  <item>
    <name>Note</name>
    <value>¥136</value>
  </item>
  <item>
    <name>Pen</name>
    <value>¥96</value>
  </item>
</itemlist>
```

***itemlist.dtd***

```
<!ELEMENT itemlist (item)>
<!ELEMENT item (name, value)>
<!ELEMENT name (#PCDATA)>
<!ELEMENT value (#PCDATA)>
```

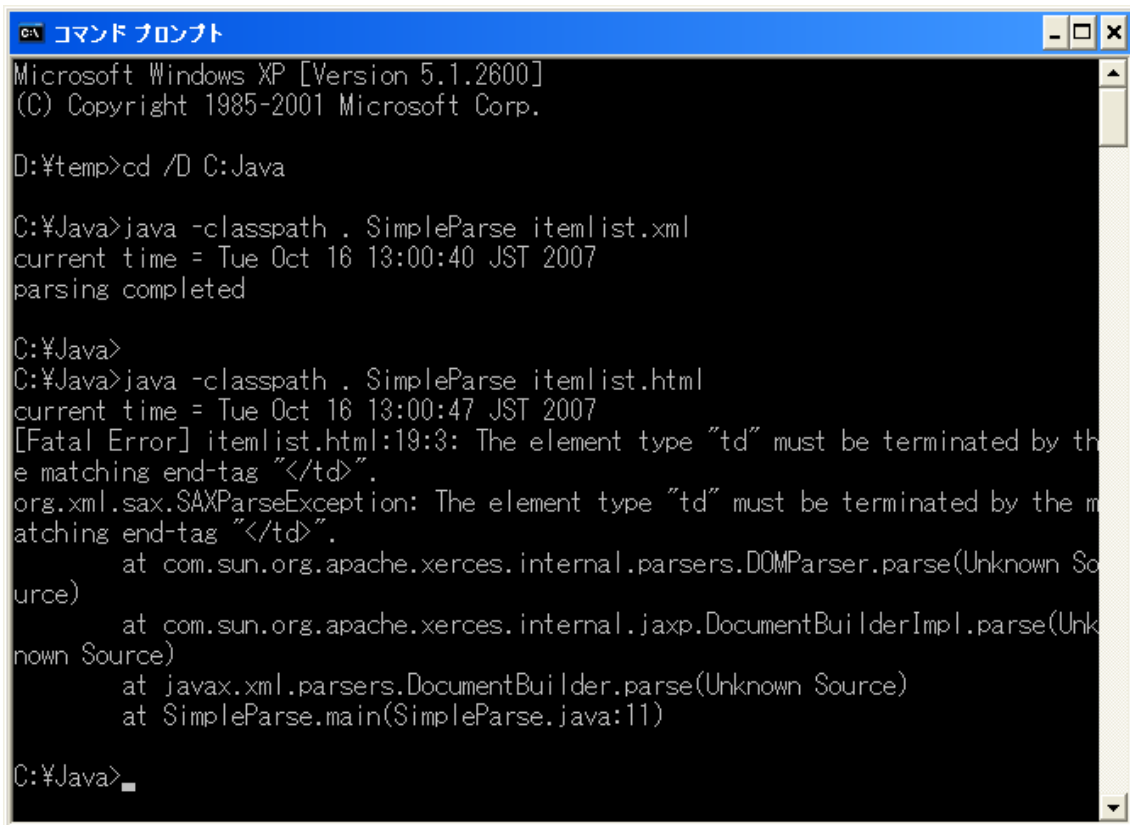
【実装したソースコード】

***SimpleParse.java***

```
import org.w3c.dom.Document;
import javax.xml.parsers.*;

public class SimpleParse {
    public static void main ( String[] args ) {
        System.out.println("current time = " + new java.util.Date());
        try {
            DocumentBuilderFactory factory =
                DocumentBuilderFactory.newInstance();
            DocumentBuilder parser = factory.newDocumentBuilder();
            Document doc = parser.parse(args[0]);
            System.out.println("parsing completed");
        } catch ( Exception e ) {
            e.printStackTrace();
        }
    }
}
```

【実行結果】



```
Microsoft Windows XP [Version 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

D:¥temp>cd /D C:Java

C:¥Java>java -classpath . SimpleParse itemlist.xml
current time = Tue Oct 16 13:00:40 JST 2007
parsing completed

C:¥Java>
C:¥Java>java -classpath . SimpleParse itemlist.html
current time = Tue Oct 16 13:00:47 JST 2007
[Fatal Error] itemlist.html:19:3: The element type "td" must be terminated by the
e matching end-tag "</td>".
org.xml.sax.SAXParseException: The element type "td" must be terminated by the m
atching end-tag "</td>".
    at com.sun.org.apache.xerces.internal.parsers.DOMParser.parse(Unknown Sou
urce)
    at com.sun.org.apache.xerces.internal.jaxp.DocumentBuilderImpl.parse(Unk
nown Source)
    at javax.xml.parsers.DocumentBuilder.parse(Unknown Source)
    at SimpleParse.main(SimpleParse.java:11)

C:¥Java>
```

【考察】

XML ファイルの構文解析を行うと、"parsing completed"という文字列が表示され、構文解析が成功していることが分かる。

HTML ファイルの構文解析を行うと、エラーが表示され、構文解析が失敗していることが分かる。エラーを見る限りでは、"td"タグの終了タグ"</td>"が明記されていないことが原因であると思われる。