COMPILER CONSTRUCTION

Assignment 3



Instructor

Mam Tayyaba

Submitted by

NAMES	REGISTRATION NO
Areeba farooq	200901058

IST, Islamabad

BSCS-SECTION A

DATE:11Jan,2023

Overview of Parser:

Parsing an XML file involves analyzing the structure and content of an XML document and turning it into a more easily understandable format, such as a tree or a list of objects. This process is useful for extracting specific data from an XML file, as well as for manipulating and modifying the contents of the file. There are several tools and libraries available for parsing XML files, including DOM (Document Object Model) parsers, SAX (Simple API for XML) parsers, and XML Pull parsers. Each of these approaches has its own strengths and weaknesses, and the best choice will depend on the specific requirements of the task at hand. Regardless of the approach taken, proper handling of errors and edge cases is important to ensure that the XML file is parsed accurately and efficiently

In this Assignment I have parse xml file using python. The data which was extracted from file includes: 1. Book_Id

- 2. Author_Name
- 3. Title
- 4. Genre
- 5. Price
- 6. Publish_date
- 7. Description

We have use XML file as input and parse xml file using python

CODE:

```
@author: Areeba
import xml.etree.ElementTree as ET
from openpyxl import Workbook
#parsing
tree = ET.parse('compiler.xml')
root = tree.getroot()
# Create a new Excel workbook
workbook = Workbook()
# Get the active sheet
sheet = workbook.active
# Define the headers for the sheet
sheet.append(["Book_id", "Author_Name", "Title", "Genre", "Price", "Publish_date", "Description"])
for book in root.findall('book'):
    book_id = book.get('id')
    author_name = book.find('author').text
    title = book.find('title').text
genre = book.find('genre').text
    price = book.find('price').text
    publish_date = book.find('publish_date').text
    description = book.find('description').text
    print(f'Book_id: {book_id}')
    print(f'Author Name: {author_name}')
    print(f'Title: {title}')
   print(f'Genre: {genre}')
print(f'Price: {price}')
    print(f'Publish_date: {publish_date}')
    print(f'Description: {description}')
    sheet.append([book_id, author_name, title, genre, price, publish_date, description])
workbook.save("books.xlsx")
```

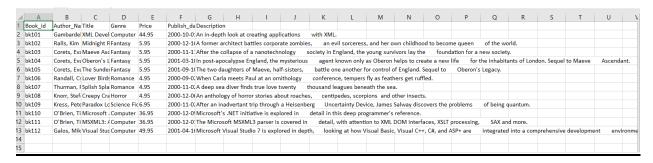
OUTPUT:

Book_id: bk101 Author Name: Gambardella, Matthew Title: XML Developer's Guide Genre: Computer Price: 44.95 Publish_date: 2000-10-01 Description: An in-depth look at creating applications with XML. Book_id: bk102 Author Name: Ralls, Kim Title: Midnight Rain Genre: Fantasy Price: 5.95 Publish date: 2000-12-16 Description: A former architect battles corporate zombies, an evil sorceress, and her own childhood to become queen of the world. Book id: bk103 Author Name: Corets, Eva Title: Maeve Ascendant Genre: Fantasy Price: 5.95 Publish_date: 2000-11-17 Description: After the collapse of a nanotechnology society in England, the young survivors lay the foundation for a new society. Book_id: bk104 Author Name: Corets, Eva Title: Oberon's Legacy Genre: Fantasy Price: 5.95 Publish_date: 2001-03-10 Description: In post-apocalypse England, the mysterious agent known only as Oberon helps to create a new life

for the inhabitants of London. Sequel to Maeve

```
Genre: Fantasy
Price: 5.95
Publish_date: 2001-03-10
Description: In post-apocalypse England, the mysterious
     agent known only as Oberon helps to create a new life
     for the inhabitants of London. Sequel to Maeve
     Ascendant.
Book id: bk105
Author Name: Corets, Eva
Title: The Sundered Grail
Genre: Fantasy
Price: 5.95
Publish_date: 2001-09-10
Description: The two daughters of Maeve, half-sisters,
     battle one another for control of England. Sequel to
     Oberon's Legacy.
Book id: bk106
Author Name: Randall, Cynthia
Title: Lover Birds
Genre: Romance
Price: 4.95
Publish_date: 2000-09-02
Description: When Carla meets Paul at an ornithology
     conference, tempers fly as feathers get ruffled.
Book id: bk107
Author Name: Thurman, Paula
Title: Splish Splash
Genre: Romance
Price: 4.95
Publish_date: 2000-11-02
Description: A deep sea diver finds true love twenty
     thousand leagues beneath the sea.
Book id: bk108
Author Name: Knorr, Stefan
Title: Creepy Crawlies
Genre: Horror
Price: 4.95
Publish_date: 2000-12-06
Description: An anthology of horror stories about roaches,
   Book id: bk109
   Author Name: Kress, Peter
   Title: Paradox Lost
   Genre: Science Fiction
   Price: 6.95
   Publish_date: 2000-11-02
   Description: After an inadvertant trip through a Heisenberg
         Uncertainty Device, James Salway discovers the problems
         of being quantum.
   Book_id: bk110
   Author Name: O'Brien, Tim
   Title: Microsoft .NET: The Programming Bible
   Genre: Computer
   Price: 36.95
   Publish_date: 2000-12-09
   Description: Microsoft's .NET initiative is explored in
         detail in this deep programmer's reference.
   Book_id: bk111
   Author Name: O'Brien, Tim
   Title: MSXML3: A Comprehensive Guide
   Genre: Computer
   Price: 36.95
   Publish date: 2000-12-01
   Description: The Microsoft MSXML3 parser is covered in
          detail, with attention to XML DOM interfaces, XSLT processing,
         SAX and more.
   Book id: bk112
   Author Name: Galos, Mike
   Title: Visual Studio 7: A Comprehensive Guide
   Genre: Computer
   Price: 49.95
   Publish_date: 2001-04-16
   Description: Microsoft Visual Studio 7 is explored in depth,
          looking at how Visual Basic, Visual C++, C#, and ASP+ are
          integrated into a comprehensive development
          environment.
```

After extracting the above information , I have created an excel file to record the data in the following table format



GitHub profile link:

https://github.com/AREEBA-FAR00Q2001/IST