

January 14, 2023

Compiler Construction

Assignment

3

Submitted by:

Syed Ali Raza

200901028

Submitted to:

Ma'am Reeda Saeed

Code

```
#Syed Ali Raza 200901028 CS-01-B

import xml.etree.ElementTree as ET
from openpyxl import Workbook
import os

#Reads input file
input_file = "compiler.xml"

#Creates tree to parse
tree = ET.parse(input_file)
root = tree.getroot()

#List of column names
column_names = ["Book Id", "Author Name", "Title", "Genre", "Price", "Publish Date",
"Description"]

#All required details are appended into a list
details = []
for child in root:
    details.append([child.get("id"), child.find("author").text,
child.find("title").text, child.find(
    "genre").text, child.find("price").text, child.find("publish_date").text,
child.find("description").text])

#Creating Workbook Class object
workbook = Workbook()
spreadsheet = workbook.active

#Writes column names
for j, value in enumerate(column_names):
    spreadsheet.cell(row=1, column=j+1).value = column_names[j]

#Inserts values into spreadsheet
for i, row in enumerate(details):
    for j, value in enumerate(row):
        spreadsheet.cell(row=i+2, column=j+1).value = value

#Creates new file
new_filename = os.path.abspath("./200901028_Assign_03.xlsx")
workbook.save(new_filename)
```

Explanation

Imported Modules:

1. xml.etree.ElementTree:

xml.etree.ElementTree is a library in python that is used to parse XML files. It allows easy traversal and extraction of data from the XML file.

2. openpyxl:

openpyxl is a Python library for reading and writing Excel files. It is used to create the Excel spreadsheet in this code.

3. os:

os is a Python library that provides a way of using operating system dependent functionality like reading or writing to files, starting new processes, killing processes etc. It is used in this code to work with the file system.

Functions used:

1. ET.parse(input_file):

This function is used to parse the input XML file. It takes the input file name as a parameter and returns an ElementTree object that allows for easy traversal and extraction of data from the XML file.

2. tree.getroot("id"):

This function gets the root element of the ElementTree object. In other words, it returns the top-most element of the XML file.

3. child.get():

This function is used to get the value of the "id" attribute of an XML element. In this case, it is used to get the "id" attribute of the book element.

4. child.find().text:

This function is used to find an element with a specific tag name and return its text content.

5. workbook.active:

This function returns the active worksheet in the workbook. In this code, it is used to set the active sheet for the Excel workbook.

Github Repository Link:

https://github.com/sydalirza/Parsing_XML_files