Assignment#1:
CSV_BIN_XML CONVERTER

Due date

26th April, 23:55

Goal

In this assignment, you are asked to develop a command line tool to convert files between CSV, Binary, and XML file formats. In addition, to validate your XML file, a corresponding XSD file should be prepared. All codings must be in C programming language. Assignment will be done in groups of two.

The assignment is expected to help you practice basic file operations and understand the details of different file formats, as well as practicing C programming language.

Implementation Details and Requirements

A sample CSV file ("records.csv") is given to you to test your program. This file consists of some information about the customers of a bank.

- The details of the features are as follows:

Attribute Name	Description
name	Displays the name of a customer (assuming it as at most 20 characters)
surname	Displays the surname of a customer, at most 30 characters (assuming it as at most 30 characters)
gender	May have one of the values: "*M*" or "*F*"
occupancy	Displays the occupancy of a customer (assuming it as at most 30 characters)
level_of_education	May be one of the values "*PhD*", "*MSc*", "*BSc*", "*HS*", or "*PS*" where *PhD*: Doctor of Philosophy, *MSc*: Master of Science, *BSc*: Bachelor of Science, *HS*: High School, *PS*: Primary School
email	May be one of the email addresses ending with: "*@gmail.com*", "*@hotmail.com*", or "*@yahoo.com*"
bank_account_number	Displays the bank account number in a specific format (Starting with 4 digits and ending with 7 digits, separated by a hyphen), such as "*0123-4567890*"
IBAN	Displays the IBAN value in a specific format (Starting with TR and continuing with 20 digits separated by a space for every four characters such as "*TR12 3456 7890 1234 5678 90*")
account_type	Displays the bank account type which can be one of the types "deposit", "drawing", or "daily_deposit"
currency_unit	Displays the currency unit of the bank account that may be one of the following units: "€", "₺", or "$"
total_balance_available	Displays the available balances of the customer in the bank account
available_for_loan	May be one of the emojis: "😊", "☹"

- In the first part, you should read the given CSV file and convert it to a binary file.

- Your converter will then **produce an XML file** by pulling data **from the created binary file** (**not from CSV file directly**). XML file should be in the given format below:

```
<records>
     <row id="1">
          <customer_info>
               <name>Ahmet</name>
               <surname>Özdemir</surname>
               <gender>M</gender>
               <occupancy>Teacher</occupancy>
               <level_of_education>BSc</level_of_education>
               <email>ahmet_ozdemir35@gmail.com</email>
          </customer_info>
          <bank_account_info>
               <bank_account_number>0123-4567890</ bank_account_number>
               <IBAN>TR12 3456 7890 1234 5678 90</IBAN>
               <account_type>deposit</account_type>
               <total_balance_available currency_unit="₺" bigEnd_Version="196608">768</total_balance_available>
               <available_for_loan>☹</available_for_loan>
          <bank_account_info>
     </row>
                                                                  ...
</records>
```

- Please pay attention that the root element of the output XML file is the name of the output file. For each line that was read from the file, a row number is assigned as the “id” attribute starting from 1 and its value is increased by 1. The tags of the subelements under “row” in the XML file are found at the header of the “*records.csv*” file.

- The field “*total_balance_available*” is assumed as to be read in the Little Endian format. The corresponding Big Endian version of this field should be calculated and be written in XML file as a new attribute with the name "*bigEndVersion*". For example, the “*total_balance_available*” field is read in the Little Endian format as “768” in the above example. The converted value of “*total_balance_available*” to the Big Endian format is calculated as “196608” that should be the value of the attribute *bigEndVersion*.

- The tool takes command line arguments according to the formats you want to convert between them. A typical command line usage is as follows:

  `myConverter <input_file> <output_file> <type>`

- The first argument, `<input_file>` refers to the source file to be used for the conversion and the second one, `<output_file>`, refers to the target file, or XSD file. The last argument, `<type>` defines conversion type (1=CSV to BIN, 2=BIN to XML, 3=XML validation with XSD).

- A sample command line usage converting from binary file to XML is as follows:

  `myConverter records.csv records.dat 1`

- A sample command line usage converting from binary file to XML is as follows:

  `myConverter records.dat records.xml 2`

- You should also **create an XSD file** that will be used to validate your XML. XSD file should include all properties including patterns and restrictions. “validation.c” file is shared with you to

test your XSD file on the corresponding XML file. A sample command line usage for XML validation is as follows:

```
myConverter records.xml recors.xsd 3
```

Documentation

In this assignment, inline documentation is expected, as well as good coding practices such as consistent naming, proper usage of indentation and high readability of code.

Submission

- Name your source code file *xxx*.c and XSD file *xxx*.xsd, where *xxx* is your ***student id***. If you don't follow the naming rules, a penalty applies. (10 pts)
- Late submission is NOT ACCEPTED.

Honesty

Your submissions will be scanned among each other as well as the Internet repository **(including ChatGPT)**. Any assignments that are over the similarity threshold of a system for Detecting Software Similarity will get zero. We strongly encourage you not to submit your assignment rather than a dishonest submission.

Grading policy

- CSV file reading – 20%
- Binary file creation – 10%
- Binary file reading – 10%
- XML file creation – 20%
- Little Endian - Big Endian conversion – 10%
- XSD file preparation – 15%
- Inline documentation – 15%

For Questions

For any questions about the assignment please write under the topic "Assignment1 Questions" in Forum on the SAKAI platform. Before asking your question, please check carefully previous questions and answers, where similar questions that were asked by someone else were already answered. No private questions via email will be answered!!!

Good luck!!!

Read all the instructions carefully, if you find something UNCLEAR, please ask help to CLARIFY it!