Semantic Web

Assignment 3

Johannes Härtel   Iryna Dubrovska

Institute of Web Science and Technologies
Department of Computer Science
University of Koblenz-Landau

Some of the tasks may require you to do additional research extending the lecture. Please keep the citation rules in mind.

For all the assignment questions that require you to write a code, make sure to include the code in the answer sheet, along with a separate python file. Where screen shots are required, please add them in the answers directly and not as separate files.

Team Name: XXXX
Team Members: XXXX

1 XML and XML schema

Have a look at the following XML below.

```
1   <!DOCTYPE record [
2     <!ELEMENT record (artist |year |contributor)+>
3     <!ELEMENT artist (record)+>
4     <!ELEMENT year (#PCDATA)>
5     <!ELEMENT contributor (#PCDATA)>
6   ]>
7
8   <record>
9     <artist>
10       <record>text</record>
11     </artist>
12     <year>text</year>
13     <contributor>text</contributor>
14   </record>
15
```

1. Check the validity of the XML according to its schema. Is XML valid? Point out the issues and rewrite XML accordingly.

2. Translate the above DTD into an XML Schema. Please, submit your XML Schema in a separate file.

2 Python Programming

Write a program in Python that accesses the XML file and schema files from Task 1. Please submit your program in a single .py file named task3.py. You can use any XML library for this task.

Requirements:

a) Create two functions with the following signatures: `validateSchema(XMLfile, XMLSchema)`, `validateDTD(XMLfile, DTDSchema)`. The functions must validate the `XMLfile` against the schema files from task 1, and print "File Validated" if the XML file conforms to its DTD/XML Schema; otherwise, it must print "Your XML file does not conform to its DTD/XML Schema".

b) Create a third function that calls the appropriate function from task a) based on the file extension: the extension for DTD is .dtd, for XML schema the extension is .xsd.

c) Create a function that creates a modified XML file given the (valid) XML file from task 1 as the input. Modify the information about contributor and insert your own name there. Also add a new attribute to contributor with the name "gender" and give it some value. Write the result into a new XML document with the name "result.xml"

d) Validate the document "result.xml" against the XML schema from Task 1 using the functions you created in a) and b). You should get the message that "Your XML file does not conform with its Schema". Write a code to remove the attribute you created before and make your XML file valid again.

Important Notes

Submission

- Solutions have to be submitted to the OLAT repository Submission in the respected folder.

- The name of the group and the names of all participating students with matriculation numbers must be listed on each submission.

- Solution format: all solutions as *one* PDF document. Programming code has to be submitted as Python code to the OLAT repository. Upload *all* `.py` files of your program! Use `UTF-8` as the file encoding. *Other encodings will not be taken into account!*

- Check that your code compiles without errors.

- Make sure your code is formatted to be easy to read.
    - Make sure you code has consistent indentation.
    - Make sure you comment and document your code adequately in English.
    - Choose consistent and intuitive names for your identifiers.

- Do *not* use any accents, spaces or special characters in your filenames.

Acknowledgment

This pdfLaTeX template was adapted by Jun Sun and Iryna Dubrovska based on the LuaLaTeX version by Lukas Schmelzeisen.

LaTeX

Use `pdflatex assignment_X.tex` to build your PDF.