School of Computer Science & IT, RMIT
CPT374/COSC2467 Document Markup Languages
SP4, 2015
Assignment One
Due date: **11:59pm Sunday 10th January 2016**
Submit via: Blackboard (not Weblearn)**- *more on this later***
This is an individual assignment that contributes **20%** of the total assessment

Introduction

The objective of this assignment is to test your understanding of XML documents, DTD, XML Schema, and how XML Schema provides a rich grammatical structure for XML documents that overcomes the limitations of DTD.

Plagiarism

All assignments will be checked with plagiarism-detection software; any student found to have plagiarised would be subject to disciplinary action. Plagiarism includes

- submitting code that is not your own or submitting text that is not your own
- allowing others to copy your work via email, printouts, social media etc.
- posting assignment questions (in full or partial) on external technical forums

All plagiarism will be penalised, there are no exceptions and no excuses. You have been warned.

General Requirements

This section contains information about the general requirements that your assignment must meet. Please read all requirements carefully before you start.

1. You must include descriptive comments where appropriate.
2. Please ensure that your submission obeys the file naming rules specified in the tasks below. File names are case sensitive, i.e. if it is specified that the file named as your student number s1234567, then 1234567.zip is exactly the file name you must submit; S1234567.zip, s1234567.zip, 1234567.ZIP, and anything else but 1234567.zip will be rejected.
3. Please validate the XML document against DTD and Schema files; the validation errors will translate to heavy penalty.
4. A penalty of 10% per day of the total marks will apply for each day late, including weekend. After five days, you will receive zero marks. For further information, please see http://www.rmit.edu.au/compsci/cgi#latesubmit.

The DML blackboard (login through myRMIT https://my.rmit.edu.au) will contain further announcements and a list of frequently asked questions. You are expected to check the discussion board on daily basis.

Submission of files

We will use **Blackboard**'s assignment submission feature (*this is not the same as Weblearn*), more instructions on this will follow at a later stage. An announcement will be made over the next few weeks.

Task 1 [10 marks]- *XML and DTD*

Note: Some of the specifications in Task 1 are open to interpretation(s), this is intentional. You will have to make assumptions while creating XML and DTD. The exercise will give you a good experience of XML serialization, DTD (and its limitations).

In order to improve the quality and efficiency of health care in Australia, the government has decided to record patient medical data on individual Medicare Cards. The data will be encrypted and will only be accessible to health care professionals. In order to make the data as portable as possible, patient health care information will be stored as an XML document. The Government has not finally decided on all the information that should be stored, but it must include at least:

1. The name of the patient
2. The gender of the patient
3. Date of birth (*think of an efficient way to serialize this*)
4. Blood type
5. Any known allergie**s**
6. Any known adverse drug reaction**s**
7. Any chronic condition**s**
8. Language**s** spoken
9. Information about each visit to a medical practitioner, including:

    (a) the type of practitioner, and associated information
    (b) the date of the visit
    (c) the patient's symptoms
    (d) the diagnosis
    (e) the prescription (if any)

10. Information about any hospital admissions, including:

    (a) date of admission and discharge
    (b) reason for admission
    (c) procedures performed

Your task is to write the following-

1. Sample XML document depicting information of two patients   (5 marks)

2. XML DTD – you must connect XML document to this external DTD (5 marks)

Name your documents as **Patient.xml** and **Patient.dtd**. You will be marked on the following criteria-
- You must use a data-centric approach.
- Element and attribute names must have meaningful semantics.
- You must include descriptive comments where appropriate, including detailed comments describing the model and why you have chosen to model the data as you have.
- **Patient.xml** must validate against **Patient.dtd**. Validation errors would lead to heavy penalty.

→

Task 2 [10 marks] – *XSD*

The **DatabaseInventory.dtd** file is an external DTD, existing in a separate file from the XML file (**DatabaseInventory.xml**) referencing it. Both of these files have been supplied as a zipped archive called **SupportingFiles.zip**.

DatabaseInventory.dtd

```
<!ELEMENT DatabaseInventory (DatabaseName+)>
<!ELEMENT DatabaseName     (GlobalDatabaseName
                            , OracleSID
                            , DatabaseDomain
                            , Administrator+
                            , DatabaseAttributes
                            , Comments)
>
<!ELEMENT GlobalDatabaseName (#PCDATA)>
<!ELEMENT OracleSID          (#PCDATA)>
<!ELEMENT DatabaseDomain     (#PCDATA)>
<!ELEMENT Administrator      (#PCDATA)>
<!ELEMENT DatabaseAttributes   EMPTY>
<!ELEMENT Comments           (#PCDATA)>

<!ATTLIST Administrator    EmailAlias CDATA
   #REQUIRED>
<!ATTLIST Administrator    Extension CDATA #IMPLIED>
<!ATTLIST DatabaseAttributes      Type(Production|Development|Testing)
   #REQUIRED>
<!ATTLIST DatabaseAttributes      Version
                           (7|8|8i|9i) "9i">
```

DatabaseInventory.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE DatabaseInventory SYSTEM "DatabaseInventory.dtd">
<DatabaseInventory>
 <DatabaseName>
   <GlobalDatabaseName>production.iDevelopment.info</GlobalDatabaseName>
   <OracleSID>Production</OracleSID>
   <DatabaseDomain>iDevelopment.info</DatabaseDomain>
     <Administrator EmailAlias="jhunter"
       Extension="6007"> Jeffrey Hunter
   </Administrator>
   <DatabaseAttributes Type="Production"
     Version="9i"/>
   <Comments>
     The following database should be considered the most stable for
       up-to-date data. The backup strategy includes running the
       database in Archive Log Mode and performing nightly backup ……
   </Comments>
 </DatabaseName>

  <!--
  other elements have been omitted here, kindly look at the supplied file
  -->
 </DatabaseInventory>
```

Your task is to create an XML Schema **`DatabaseInventory.xsd`** from the given
`DatabaseInventory.dtd` and `DatabaseInventory.xml`.

Your XSD must also fulfil the following requirements:

    a. There are at most three Administrator per database;
    b. The element Comments is less than 100 characters in length;
    c. The attribute Extension is 4 digits in length beginning with "6";
    d. The value of element OracleSID must be one from Production, Development, or
       Testing.

       (Note: *this will be exactly the same as attribute DatabaseAttributes, please design your
       schema to maximize reuse*.)

You must include descriptive comments where appropriate, including detailed   comments
describing the model.

Your XSD must be neatly indented.

Make sure you remove the DOCTYPE declaration and   provide a link from the XML
document to the XML Schema.

BONUS Task for coding enthusiasts [5 marks] – *this is extra and is not a part of the total
marks for the assignment. Attempt if you have some extra time.*

Your task is to create an XML Schema (**`DatabaseInventoryBonus.xsd`**) from task 2
again; only this time you must use a Schema pattern as outlined below-

Your XSD must use one of the three main design patterns
(0 – Salami slice, 1 - Venetian blinds, 2 - Garden of Eden), which allows reuse of both
elements and types.

Please use the last digit of your student **ID % 3** to select the design pattern you will use
in your submission. For example, if your student number is s1234567, then you will use 1 -
Venetian blinds (7 mod 3 = 1) in this assignment. A penalty of 50% of the total marks will
apply for using the wrong design pattern.

A reading document '**ReadForBonusTask.pdf**' has been supplied as a part of the
SupportingFiles.zip archive. The document talks about the use of Patterns in a Schema
document. This also makes for a good reading-
http://www.xfront.com/GlobalVersusLocal.html

The bonus task marks will be added to your examination score.

Submission

Before you submit anything, please read through the assignment specifications again carefully. Check that you have followed all instructions. Also check that you have attempted all parts of all questions.

Feel free to use any development environment you like in your assignments. However, it's your responsibility to ensure all your files have no mistakes. You must validate your XML instance documents against your XML schema documents (XSD/DTD).

The following 4 files must be submitted as <u>ONE single zipped file</u> named as your student number (for example, 1234567.zip if your student ID is s1234567, 1234567M.zip if your student ID is s1234567M) via Blackboard's assignment submission system:

1.   `Patient.xml` (Make sure that it has been correctly linked to your DTD)
2.   `Patient.dtd`
3.   `DatabaseInventory.xml` (Make sure that it has been correctly linked to your XSD)
4.   `DatabaseInventory.xsd`

   The following are optional- submit these if you have attempted the bonus task-

5.   `DatabaseInventoryBonus.xml` (Make sure that it has been correctly linked to your Bonus XSD)

6.   `DatabaseInventoryBonus.xsd`

The single zipped archive must be submitted via **Blackboard's assignment submission system**- more instructions on this to come later. Stay tuned!

What Happens Next

We will discuss the assignment during collaborate chats- if you cannot attend, please go through the recordings.

You will receive your results by Sunday 24/01/2016.