

INFO-H-509 : Technologies XML

TP 2 - Schema languages for XML

Professor : Stijn Vansummeren
Teaching Assistant : Michaël Waumans
<http://cs.ulb.ac.be/public/teaching/infoh509>

XML Schema Validity

All documents required for completing the exercises below are available on the course's web page.

Exercise 1.1

The document `schedule.xml` describes a university's scheduling information of courses and exercise sessions. Is this document valid with respect to `schedule.xsd`? If not, make the necessary corrections to `schedule.xml` so that it becomes valid.

Exercise 1.2

The `schedule.xsd` schema is not very demanding with respect to `course` elements. The constraints that we would like to express are as follows :

- We refer to lectures and exercises collectively as *events*. Each event (be it lecture or exercise) must mention an attribute `on` (indicating the date of the event) and, optionally, a unique name defined by the `id` attribute. Each event must also be associated with a room.
- Lectures are described by `lecture` elements, which can contain multiple `topic` subelements. Topics within lectures can have `resource` subelements that mention URIs from which additional information can be fetched.
- Exercises are described by `exercises` elements, which can contain again multiple `topic` subelements. Here, however, the content of the `topic` subelements should refer to the ids of the lectures to which the exercises belong.

Proceed as follows to complete the schema :

- Complete the definition of `event` type, as a basis to describe courses and exercises. It will contain the definitions common to both types of events.
- Extend this type to obtain 2 types : 1 specialized for courses and 1 specialized for exercises. You may ignore the requirement that `topic` inside `exercise` elements should refer to a valid `lecture` id.
- Modify the type definition of `course`, so that courses and exercises sessions can be interlaced.

Exercise 1.3

Additional exercise : By the use of `unique`, `key` and `keyref` elements, ensure that lecture's resources are unique, a lecture refers to an existing room, and that an exercise refer to an existing lecture. Correct the document if necessary.

Exercise 1.4

Complete the following schema so that it uses the same namespace as the following XML document :

```
<message on="2010-03-23T17:55:00" xmlns="http://www.example.org/message">
  Hello!
</message>
```

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:complexType name="message">
    <xs:simpleContent>
      <xs:extension base="xs:string"
        <xs:attribute name="on" type="xs:dateTime" />
      </xs:extension>
    </xs:simpleContent>
  </xs:complexType>

  <xs:element name="message" type="message" />
</xs:schema>
```

Expressions régulières et DTD

Exercise 2.1

Consider the following regular expressions over the alphabet a, b, c, d :

1. $a(bb^?)?(c^+(a|d))$ 2. a^+bbccd 3. $c?b?(a^*|b|c)^+$ 4. $c^*(abcabc)$ 5. $(a^+b)^+(cd)^+$

Determine which regular expressions accept the following strings :

- A. $abcabc$ B. $abcd$ C. $abbccd$ D. $cabcabc$ E. $aaccd$ F. $abaaabcd$

Exercise 2.2

Complete the DTD `schedule.dtd` with the requirements of Exercise 1 : apply similar constraints to the `courses` element if possible.

Exercise 2.3

Compare DTD and XML Schema obtained in exercise 1 and 2.2. What are the main differences ?