

## Exercise Sheet 2

# DTD validation, Namespaces

### Exercise 1: Validation of XML with a DTD

Given is the following DTD:

```
<!ELEMENT movies (Movie+) >
<!ELEMENT Movie
    ( title, year, _director, (comment | newcomment)+)>
<!ATTLIST Movie id ID #REQUIRED>
<!ELEMENT title (#PCDATA)>
<!ELEMENT year (#PCDATA) >
<!ELEMENT _director (#PCDATA)>
<!ATTLIST _director name CDATA #IMPLIED>
<!ELEMENT comment (#PCDATA)>
<!ELEMENT newcomment (#PCDATA)>
<!ATTLIST comment lang CDATA #IMPLIED>
```

1.1. Correct the document from Exercise Sheet 1, Exercise 2.1, to conform to this DTD (use oXygen or Eclipse).

1.2. What is the difference between CDATA and PCDATA?

1.3. Provide four reasons for using elements instead of attributes.

Hint: a simple DTD tutorial can be found at <http://www.w3schools.com/dtd/default.asp>

### Exercise 2: Namespaces

2.1. Correct the mistakes in the following (not well-formed) XML document:

```
<?xml version="1.0"?>
<?DOCTYPE "eth" SYSTEM "eth.dtd"?>
<eth xmlns="http://www.ethz.ch"
    xmlns:xmldb="http://www.dbis.ethz.ch"
    date="11.11.2006"
    xmldb:date="12.11.2006">
  <date>13.11.2006</date>
  <president number="1">Empty</president>
  <Rektor>Name 2</rektor>
</Doc>
```

2.2. What are the namespaces of each attribute and element?

2.3. Try to define the external DTD **eth.dtd** so that the document is valid (use oXygen or Eclipse). Does DTD actually handle namespaces?

2.4. (Bonus) Take a look at exercise2-2-4.xml (from last week's exercise). Try to change the namespace for the html tag and check for validity. What error message did you expect? Why instead does oXygen want you to use the XHTML namespace, although there is nothing wrong with using a tag having local name "html" in another namespace? How can you get oXygen to output the expected error message?