Prof. Dr. Stefan Brass
Institut für Informatik
MLU Halle-Wittenberg

06. Dezember 2018

XML und Datenbanken
— 7. Übungsblatt: XML Schema, DOM —

Bei Teil a) bis e) handelt es sich um Präsenzübungen, die nicht abgegeben werden müssen. Sie sollten über die Wiederholungsfragen a) allerdings vor der nächsten Übung selbst nachdenken. Teil f) und g) sind Hausaufgaben. Bitte schicken Sie eine Lösung per EMail an den Dozenten (mit "`xml18`" in der Betreff-Zeile, bis zum 12.12.2018). Schicken Sie bitte keine leeren EMails nur mit Anhang, diese landen im Spam-Ordner.

a) Wie würden Sie in einer mündlichen Prüfung auf folgende Fragen zu XML Schema antworten?

- Welche Möglichkeiten gibt es, um
  - Von einem "Simple Type" einen "Simple Type" abzuleiten?
  - Von einem "Simple Type" einen "Complex Type" abzuleiten?
  - Von einem "Complex Type" einen "Complex Type" abzuleiten?

  Geben Sie jeweils ein Beispiel.
- Vergleichen Sie abgeleitete Typen in XML Schema mit Unterklassen in Java.
- Angenommen, Sie schreiben ein Anwendungsprogramm, das XML Daten verarbeitet. Sie haben dafür ein Schema angegeben. Wenn Sie erlauben, das andere Leute Ihr Schema mit `include` benutzen, und erweitern, womit müssen Sie dann bei der Programmierung rechnen? Die erweiterten Schemata können natürlich keine Typen, Elemente oder Attribute umdefinieren, die Sie schon definiert haben. Sie können aber neue hinzufügen, auch Untertypen.
- Was ist eine "Substitution Group" bei Element-Deklarationen?
- Wozu dient "`xsi:type`"?
- Macht es einen Unterschied, ob Sie in einer Element-Deklaration keinen Typ definieren, also das `complexType`-Element weglassen, oder ein leeres `complexType`-Element hinschreiben?

Präsenzaufgaben

b) Beschaffen Sie sich das Beispiel einer Feuerwerksartikel-Datenbank:

- [http://users.informatik.uni-halle.de/~brass/xml18/examples/fw.xml]
- [http://users.informatik.uni-halle.de/~brass/xml18/examples/fw.xsd]

Die Daten sind:

```
<?xml version="1.0" encoding="UTF-8"?>
<Feuerwerksartikel>
    <Bombenrohr Herst="Diamond" Name="Silver Star">
        <Angebot Haendler="Röder" Preis="5.49"/>
        <Bew Jahr="2005">Goldpalme! Toll.</Bew>
        <Bew Jahr="2006">Na ja, geht so.</Bew>
    </Bombenrohr>
    <Batterie Herst="Weco" Name="Tanz der Vampire"
            Schuss="12" Hoehe="45" Dauer="30">
        <Angebot Haendler="Röder" Preis="8.50"/>
        <Angebot Haendler="Preisw-FW" Preis="7.69"/>
        <Bew Jahr="2006">Rote Blinker. Hübsch.</Bew>
    </Batterie>
</Feuerwerksartikel>
```

Es gibt viele verschiedene Arten von Feuerwerksartikeln, die unterschiedliche Attribute haben. "`Bombenrohr`" und "`Batterie`" sind nur zwei Beispiele, es gibt z.B. auch Fontänen, Sonnen, römische Lichter, Böller, u.s.w. In dem zugehörigen XML Schema ist ein Obertyp "`Artikeltyp`" definiert, von dem dann z.B. "`BombenrohrTyp`" und "`Batterietyp`" abgeleitet sind. Es ist außerdem eine Gruppe "`ArtikelArten`" als "`choice`-Verknüpfung der Elemente für die verschiedenen Artikelarten definiert.

Welche XML Schema Elemente wurden für die Typableitungen verwendet?

c) Erweitern Sie das Schema um einen Artikeltyp "`Sonne`" mit den Attributen

- "`Durchmesser`" und
- "`Dauer`".

d) Verwenden Sie Substitution Groups, um die Elemente für die verschiedenen Artikelarten einzuführen.

e) Probieren Sie eine Lösung mit `xsi:type` aus. Der Namespace ist

"`http://www.w3.org/2001/XMLSchema-instance`".

Hausaufgabe

f) Definieren Sie ein XML Schema für eine kleine Datenbank von Personen im Institut:

- Alle Personen haben Anrede (Herr/Frau), ggf. Titel, Vorname, Nachname, und optional eine EMail-Adresse.
- Studenten haben zusätzlich ein Studienfach, ein Semester, und eine Matrikelnummer. Falls sie schon einen Abschluss haben (z.B. Bachelor), soll dieser auch gespeichert werden.
- Professoren haben Fachgebiete. Dies können mehrere sein, z.B. "Datenbanken" und "Logische Programmierung".
- Zu Doktoranden soll eine Information über ihre Stelle/Finanzierung festgehalten werden (z.B. "Mitarbeiter", "Graduiertenstipendium") und ein Professor als Betreuer. Wenn Sie wollen, können Sie den Professoren zu diesem Zweck noch einen Schlüsselwert zuweisen, z.B. ein Kürzel. Sie können aber auch annehmen, dass Vorname und Nachname zusammen innerhalb des Instituts für Professoren eindeutig sind.

Da später noch andere Typen von Personen hinzukommen können (z.B. Sekretärinnen und Techniker), gestalten Sie Ihr Schema bitte möglichst leicht erweiterbar. Es soll ein äußeres Element "`Institut`" geben, darin direkt für jede Person ein Element.

g) Informieren Sie sich über die Verarbeitung von XML in Java mittels DOM, z.B. mit folgendem Tutorial:

[https://docs.oracle.com/javase/tutorial/jaxp/dom/index.html]

Weitere Tutorials finden Sie z.B. hier:

[https://www.tutorialspoint.com/java_xml/java_dom_parser.htm]
[https://www.mkyong.com/java/how-to-read-xml-file-in-java-dom-parser/]

Die DOM-Spezifikation stammt vom W3C:

[https://www.w3.org/TR/DOM-Level-3-Core/]
[https://www.w3.org/DOM/]

Schreiben Sie oder beschaffen Sie sich ein kleines Beispielprogramm, übersetzen Sie dies, und führen Sie es aus. Es soll mindestens die im Dokument befindlichen Elemente anzeigen (den Element-Namen). Sie können aber auch mehr bzw. alle Knoten ausgeben. Wenn die Baumstruktur z.B. durch Einrückung sichtbar wird, wäre dies ein Plus. Reichen Sie Ihr Programm per EMail ein. Wenn Sie ein Programm aus dem Internet verwenden, müssen Sie die Quelle nennen. Sie müssen es in der Übung erklären können. Achten Sie also gerade bei fremden Programmen darauf, dass Sie diese voll verstanden haben.