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

08. Dezember 2016

XML und Datenbanken
— 8. Übungsblatt: XPath —

Teil a) muss nicht abgegeben werden. Sie sollten sich allerdings zu Hause mit diesen Fragen beschäftigen, um den Stoff der Vorlesung zu wiederholen.

Teil b) ist eine Hausaufgabe. Bitte schicken Sie eine Lösung per EMail an den Dozenten (mit "`xml16`" in der Betreff-Zeile, möglichst bis 08.12.2016, 16:00). Schicken Sie bitte keine leeren EMails nur mit Anhang, diese landen im Spam-Ordner. Nennen Sie die Datei(en) bitte "Nachname_Vorname", ggf. mit weiterem Suffix.

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

- Was ist der effektive boolesche Wert einer Sequenz? Wie ist er definiert? Wo wird der effektive boolesche Wert gebraucht?
- Was ist "Atomisierung"? Wo wird das angewendet? Mit welcher Funktion können Sie es explizit aufrufen?
- Was ist der Haupt-Unterschied zwischen allgemeinen Vergleichsoperatoren wie "`=`" und den Wert-Vergleichsoperatoren wie "`eq`"?
- Sie stellen mit XPath Anfragen an ein Dokument, das nicht validiert wurde. Welches Problem ergibt sich dann bei einer Anfrage wie "`//RESULT[@POINTS gt 8]`" Welchen Typ hat der Attributwert hier? Wie können Sie das Problem lösen?
- Zeigen Sie an einem einfachen Beispiel, dass "`=`" die Transitivität verletzt.

Hausaufgabe

b) Verwenden Sie die folgende Beispiel-Datei für XPath-Anfragen:

[http://www.informatik.uni-halle.de/~brass/xml16/cd.xml]

Die zugehörige XML Schema Definition finden Sie hier:

[http://www.informatik.uni-halle.de/~brass/xml16/cd.xsd]

Die Daten-Datei hat die gleiche Struktur wie die von Blatt 7, enthält aber deutlich mehr Daten. In der Schema-Datei ist ein Schlüssel korrigiert: Die Opus-Nummer ist natürlich nur innerhalb eines Komponisten eindeutig. Bei Bedarf gibt es auch eine DTD:

[http://www.informatik.uni-halle.de/~brass/xml16/cd.dtd]

Falls Sie Ihre Anfragen zusätzlich auch mit SQL probieren wollen, finden Sie hier ein Skript, das die Tabellen anlegt und füllt:

[http://www.informatik.uni-halle.de/~brass/xml16/cd.sql]

Schreiben Sie die folgenden Anfragen in XPath, und probieren Sie diese auch mit einem System aus! Beschreiben Sie in Ihrer Hausaufgabe kurz, mit welchem System Sie die Abfragen ausprobiert haben, und ob es dabei irgendwelche Schwierigkeiten gab (und ggf., wie Sie diese gelöst haben). Wie in der Vorlesung erläutert, wäre eine Möglichkeit, die BaseX XML Datenbank herunterzuladen:

[http://basex.org/home/]

Es gibt aber auch online demos wie z.B.

[http://www.freeformatter.com/xpath-tester.html]

Einfache XPath-Ausdrücke (XPath 1.0) können auch mit der in der Vorlesung genannten Methode mit einem XSLT-Stylesheet ausprobiert werden. Dazu finden Sie ein Beispiel-Stylesheet unter

[http://www.informatik.uni-halle.de/~brass/xml16/query.xsl]

Eine XML-Datei mit Verweis auf dieses Stylesheet steht hier:

[http://www.informatik.uni-halle.de/~brass/xml16/ex_query.xml]

Formulieren Sie die folgenden Anfragen in XPath:

- Geben Sie die Namen aller Solisten aus, die Violine spielen. Da ein Solist eventuell mehrere Instrumente spielt (z.B. "Violine" und "Viola"), steht das Instrument im Element `solist_aufnahme`, das in `solist` geschachtelt ist.
- Geben Sie alle CDs aus (Element-Knoten), die mehr als 3 Einzel-CDs enthalten (also mindestens den Wert 4 im Attribut `anz_cds` haben).
- Welche CDs (Attribut `name`) enthalten Stücke von "Wolfgang Amadeus Mozart"?
- Selektieren Sie alle Komponisten, bei denen keine Opus-Nummern verwendet werden, d.h. keines ihrer Stücke hat das Attribut "`opus`".