

XML und Datenbanken — 9. Übungsblatt: XPath, Teil II —

Allgemeine Aufgabe

Notieren Sie sich eventuelle Verständnisfragen, so dass wir diese im nächsten Online-Treffen klären können. Da es erfahrungsgemäß sehr sill ist: Nehmen Sie sich bitte die Zeit und denken Sie bewusst über mögliche Fragen nach (wenigstens eine). Ich kann auch Studierende drannehmen, die sich nicht gemeldet haben. Ein Mal würde ich wohl verstehen, dass der Stoff so einfach war, dass Sie einfach keine Fragen haben. Wenn sich das wiederholt, müsste ich dann aber umgekehrt prüfungsähnliche Fragen stellen.

Hausaufgabe

Geben Sie die folgenden Aufgaben bis Montag, 19.12.2022, 16⁰⁰, über die Übungsplattform in StudIP ab. Schreiben Sie die Lösungen in eine `.txt`- oder `.xpath`- oder `.xq`-Datei. Es gibt 3 Punkte pro Anfrage.

Die Abgaben nur stichprobenartig kontrolliert. Wenn Ihre Abgabe nicht kontrolliert wurde, bekommen Sie die volle Punktzahl. Wenn Sie später wegen Plagiaten auffallen, oder bei einer Stichprobe eine fast gar nicht gelöste Aufgabe entdeckt wird, können auch alte Abgaben kontrolliert werden. Dann können auch rückwirkend Punkte abgezogen werden.

Sie benötigen 67% der Hausaufgabenpunkte und eine aktive Mitarbeit in den Übungen für die Studienleistung.

Die „Wiederholungsaufgaben“, also Teil h) und i), sind nicht abzugeben. Beschäftigen Sie sich aber bitte auch mit diesen Aufgaben. Sie müssen damit rechnen, dass Sie beim Online-Treffen gebeten werden, einen Teil des Vorlesungs-Stoffes zu wiederholen und insbesondere eine der Fragen zu beantworten.

Laden Sie die folgenden Dateien herunter, die Daten über eine CD-Sammlung für klassische Musik enthalten:

- Daten:
[<http://www.informatik.uni-halle.de/~brass/xml18/examples/cd.en/cd.xml>]
- XML Schema:
[<http://www.informatik.uni-halle.de/~brass/xml18/examples/cd.en/cd.xsd>]
- DTD:
[<http://www.informatik.uni-halle.de/~brass/xml18/examples/cd.en/cd.dtd>]

Schreiben Sie die folgenden Anfragen in XPath und testen Sie Ihre Anfragen mit einer XPath-Implementierung. 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). Sie können XPath auch mit Systemen ausprobieren, die XQuery unterstützen (XPath ist eine Teilmenge von XQuery):

- BaseX XML Datenbank:
[<http://basex.org/home/>]
- Eine weitere native XML Datenbank ist eXist-db:
[<http://exist-db.org/exist/apps/homepage/index.html>]
- Es gibt auch eine Implementierung von den Herausgebern der XQuery Spezifikation (Galax):
[<http://galax.sourceforge.net/>]
- Auch AltovaXML ist ein nützliches Programm mit Unterstützung vieler Technologien, die in dieser Vorlesung besprochen werden:
[<https://www.softpedia.com/get/Internet/Other-Internet-Related/AltovaXML.shtml>]
- Sie können XPath Anfragen aber auch mit folgender Webseite ausprobieren:
[<http://www.freeformatter.com/xpath-tester.html>]

Formulieren Sie folgende Anfragen in XPath:

- a) **(3 Punkte)** Gesucht sind alle Kommentarknoten im ganzen Dokument.
- b) **(3 Punkte)** Gesucht sind die Namen (Attributwert) von allen im Dokument vorkommenden Orchestern.
- c) **(3 Punkte)** Selektieren Sie alle Stücke (Elementknoten) von “Bach” in “E-dur”.
- d) **(3 Punkte)** 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.
- e) **(3 Punkte)** Geben Sie alle CDs aus (Element-Knoten), die mehr als 3 Einzel-CDs enthalten (also mindestens den Wert 4 im Attribut `anz_cds` haben).
- f) **(3 Punkte)** Welche CDs (Attribut `name`) enthalten Stücke von “Wolfgang Amadeus Mozart”?
- g) **(3 Punkte)** Selektieren Sie alle Komponisten, bei denen keine Opus-Nummern verwendet werden, d.h. keines ihrer Stücke hat das Attribut “opus”.

Wiederholungsaufgaben

Beschäftigen Sie sich mit diesen Aufgaben. Sie brauchen aber nichts abzugeben.

- h) Was würden Sie in einer mündlichen Prüfung auf folgende Fragen zu XPath antworten?
- XPath hat keine reservierten Worte, d.h. alle in XML erlaubten Bezeichner für Elementnamen können auch in XPath verwendet werden (ohne spezielle Kennzeichnung). Nennen Sie einige Beispiele, wie Schlüsselworte in XPath von normalen Elementnamen unterschieden werden können.
 - Was bedeutet der XPath-Ausdruck “`x-1`”?
 - Wie werden Variablen in XPath gekennzeichnet?
 - Wenn man in XPath eine Zeichenketten-Konstante mit `'` als Begrenzungszeichen schreibt, wie kann man dieses Zeichen als Daten der Zeichenkette einfügen? Warum funktioniert die XML Lösung mit “`'`” nicht in XPath? (Betrachten Sie dabei den Fall, dass der XPath-Ausdruck ein Attributwert in XML ist, wie z.B. bei einem XSLT-Stylesheet.)
 - Wie schreibt man einen Datumswert in XPath, z.B. den 24.12.2018?
 - Wie schreibt man boolesche Werte in XPath?
 - Wie kann man in XPath auf die drei wichtigsten Komponenten des Kontexts zugreifen (Kontext Element, Kontext Position und Kontext Größe)?
 - Was macht der Komma-Operator “`,`” in XPath? Vergleichen Sie es mit dem Vereinigungs-Operator “`|`” (den man auch “`union`” schreiben kann).
 - 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?
 - Geben Sie ein Beispiel, in dem gleichzeitig `$x = 1` und `$x != 1` gelten.
 - Zeigen Sie an einem einfachen Beispiel, dass “`=`” die Transitivität verletzen kann.
 - Wie kann man in XPath prüfen, ob zwei Knoten identisch sind?
 - Wie kann man abfragen, ob ein Knoten in der Dokument-Anordnung vor einem anderen Knoten steht?
 - Wie schreibt man den Allquantor und den Existenzquantor in XPath? Erklären Sie Syntax und Semantik. Warum ist es ein Problem, wenn die Bedingung einen Laufzeitfehler für einige der Werte erzeugt, über denen der Quantor läuft?

- Beschreiben Sie Syntax und Semantik von `for`-Ausdrücken in XPath.
- Normalerweise sind Schleifen typische Kennzeichen von imperativen Programmiersprachen. Vergleichen Sie `for`-Schleifen in Sprachen wie Java mit den `for`-Ausdrücken in XPath. Erklären Sie, warum XPath nicht imperativ ist. Wie könnten `for`-Ausdrücke von XPath auf parallelen Prozessoren ausgewertet werden? Warum sind typische Schleifen in Java viel schwieriger zu parallelisieren?
- Erklären Sie die Syntax und Semantik von `if`-Ausdrücken in Xpath. Warum gibt es keine Version ohne `else`? Was ist der Unterschied zu `if`-Statements in Java?

Für Interessierte

- i) Schauen Sie sich den Artikel „Implementing XQuery 1.0: The story of Galax“ von Mary Fernández und Jérôme Siméon an: [<https://dl.gi.de/handle/20.500.12116/28283>]