

XML-Datenbanken
Wintersemester 2008/2009
2. Aufgabenblatt

Hinweis: Zur Diskussion der Übungsaufgaben in der Übung sind vorweg formlose Präsentationsfolien zu erstellen. Diese müssen im PDF-Format per Email an kupfer@ifis.cs.tu-bs.de mit dem Betreff "XMLDatenbanken" sowie der Gruppennummer bis spätestens Sonntag vor der Übung abgegeben werden. Achten Sie bitte stets darauf Ihren Namen auf dem Titelblatt zu vermerken.

Aufgabe 1: Erstellen Sie eine Hierarchie der in XML Schema verfügbaren Datentypen. Notieren Sie zusätzlich, wie diese Datentypen in einer DTD dargestellt werden können.

Aufgabe 2: Implementieren Sie ein Java Programm, das überprüft ob eine angegebene XML-Datei wohlgeformt ist und gegen eine gewählte XML-Schema-(.xsd)-Datei validiert werden kann. Verwenden Sie Ihr Programm, um die für Aufgabe 3 zu schreibenden Dateien zu korrigieren. Es empfiehlt sich auf vorhandene Implementierungen im JDK unter "javax.xml.validation.*" zurückzugreifen. Die Überprüfung läuft in 5 Schritten ab: Es sollte zunächst eine Instanz der SchemaFactory erstellt werden. Die xsd-Datei wird der Factory als neues Schema übergeben. Dann sollte ein Validator aus dem Schemaobjekt erzeugt werden, das zu prüfende Dokument z.B. als `javax.xml.transform.stream.StreamSource` geladen werden und schließlich mit der `validate`-Methode geprüft werden.

Aufgabe 3: Initialisierungsdateien (ini-Dateien) sind einfache Textdateien und bestehen aus Schlüssel-Werte-Paaren, vergleichbar mit einer Hashtable. Diese Dateien dienen der Speicherung von Programmkonfigurationen. Inzwischen werden sie u.a. durch XML abgelöst. Betrachten Sie den Ausschnitt aus einer solchen Datei `opera6.ini` auf der nächsten Seite.

Die Syntax gibt vor, dass pro Zeile ein Schlüssel und der zugehörige Wert, getrennt durch ein Gleichheitszeichen, definiert werden kann. Schlüssel und Wert sind jeweils beliebige Zeichenketten ohne Zeilenumbruch und Gleichheitszeichen. Einzelne Abschnitte werden durch eckige Klammern definiert und stellen jeweils eine eigene Hashtable dar.

- Warum wurde der Syntax in der Beispieldatei mit Komma "," und Pipe "|" zusätzliche Trennzeichen für Werte hinzugefügt?
- Vergleichen Sie die Mächtigkeit der Sprachen INI und XML.
- Definieren Sie einen eigenen XML-Dialekt um die Informationen dieser Datei darzustellen, indem Sie ein XML Schema erstellen.
- Entsteht dadurch neue Redundanz oder lässt sich Redundanz sogar reduzieren?
- Konvertieren Sie einzelne Beispiele aus der Datei nach XML. Können alle möglichen Zustände der ini-Datei durch Ihren XML-Dialekt dargestellt werden?

Ausschnitt aus einer opera6.ini Datei (Manche Zeilen mussten für diesen Ausdruck umgebrochen werden und sind mit " ➡" gekennzeichnet):

[User Prefs]

Plugin Path=/usr/lib/opera/plugins:/usr/X11R6/lib/Acrobat5/Browsers/intellinux:/usr/lib/RealPlayer8/
➡Plugins:/usr/lib/RealPlayer8:/usr/users/xmldb/.netscape/plugins:/usr/local/netscape/plugins
Button Set=/usr/share/opera/skin/standard_skin.zip
Has Restored Extensions=2

[Adv User Prefs]

HTTP Accept Language=en

[State]

Reading Plugins=0

[File Types Section Info]

Version=2

[File Types]

/=0,,,,*,|

...=0,,,,|

application/mime=2,,,,mim,mime,mhtml,|

application/x-pem-file=2,,,,pem,|

application/x-x509-ca-cert=2,,,,crt,ca,cer,509, der,|

application/x-x509-user-cert=2,,,,usr,|

application/x-pkcs12=2,,,,pfx,p12,|

application/x-javascript=2,,,,js,|

image/gif=2,,,,gif,|

image/jpeg=2,,,,jpeg,jpe,jpg,|

image/jpg=2,,,,jpg,|

image/pjpeg=2,,,,jpeg,|

image/bmp=2,,,,bmp,|

image/x-windows-bmp=2,,,,bmp,|

image/x-ms-bmp=2,,,,bmp,|

image/x-xbitmap=2,,,,xbm,|

image/xbm=2,,,,xbm,|

application/pdf=0,./usr/X11R6/lib/Acrobat5/Browsers/intellinux/nppdf.so,No Name,pdf,|Portable

➡Document Format

application/postscript=0,./usr/local/netscape/plugins/mozplugger.so,MozPlugger 1.3.0,ps,eps,ai,|

➡PostScript file

image/tiff=0,./usr/local/netscape/plugins/mozplugger.so,MozPlugger 1.3.0,tif,tiff,|TIFF image

application/x-pdf=6,./usr/local/netscape/plugins/mozplugger.so,MozPlugger 1.3.0,pdf,|PDF file

application/x-msdownload=0,,,,exe,bat,pif,com,scr,|

application/x-msdos-program=0,,,,exe,bat,pif,com,scr,|

application/internet-shortcut=2,,,,url,|

application/zip=0,,,,zip,|

application/x-zip-compressed=0,,,,zip,|

application/x-arj=0,,,,arj,|

application/x-rar-compressed=0,,,,rar,|