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

November 17, 2017

XML and Databases
— Exercise Sheet 5 —

You only have to submit the parts marked as “Homework Exercises”, i.e. Part d) and e). But please think about the questions in Part a) before the meeting! Send your homework solutions to the instructor via EMail: `brass@informatik.uni-halle.de` (with “`xml17`” in the subject line). The official deadline is November 23, 10:00 (before the lecture time).

Not all submitted homeworks will be corrected, but all homework exercises will be discussed in class. If you should have questions about your homework, please ask! A precondition for getting credit for this course is that you submit solutions to two thirds of the homework sheets. Obviously wrong or very incomplete submissions do not count.

Repetition Questions

a) Answer the following questions about XML schema:

- How does one specify a “mixed content model” in XML schema, i.e. define an element “`p`” that contains text with the possibility to force line breaks with “`<br/>`” and to insert links with “`<a href="...">...</a>`”?
- Does it make a difference whether you write an element declaration without a `complexType` child or with an empty `complexType` child?
- What is a “substitution group” in element declarations? For which purpose can it be used?
- What is the purpose of the attribute “`xsi:type`” in an XML data file?
- How does one specify keys in XML schema? What is the difference between `xs:key` and `xs:unique`? What is the purpose of `xs:selector` and `xs:field`?
- What are the main constructs that can be used in the XPath subset for the `xpath`-attribute of `xs:selector` and `xs:field`? Name an important restriction.
- Compare keys in a relational database with keys in XML schema. Which aspect of keys in XML Schema is much simpler in the relational model?
- How can one define keys in XML schema that correspond to weak entities in the ER-Model, i.e. what would you do if an element is identified by an attribute of its own in combination with an attribute of its parent entity?
- How does one specify foreign keys in XML schema?

In-Class Exercises

b) Define an XML schema for a small database of persons in the computer science institute (employees and students). This is an example of “subclasses”:

- Persons have a form of address (Mr./Ms.), an optional title (e.g. “Prof. Dr.”), first name, last name, and optionally an email address.
- Students have all data items (attributes or child elements) of persons, but in addition have a student ID (“Matrikelnummer”) and a study program (e.g., “Bioinformatics”).
- Professors are a subclass of persons about which an office number (e.g. “313”) and a set of specialization areas must be stored, e.g. “Databases” and “Logic Programming”.

c) Have a look on at least one of the following XML schemas, check that you know the constructs used in them, and find out which schema style is used.

- XML schema for XHTML:

  [https://www.w3.org/TR/xhtml1-schema/]

- XML Schema for WSDL, the Web Service Description Language:

  [https://www.w3.org/TR/wsdl#A4.1]

- XML Schema for SEPA direct debit:

  [http://www.sepazahlungsverkehr.de/spezifikationen/pain.008.002.02.xsd]

  Companies or clubs that want to collect money directly from customer/member bank accounts must use this format to send the transaction data to their bank. A specification is given here:

  [http://www.ebics.de/index.php?id=77]

Homework Exercises

d) Define an XML schema for your nutrition supplements database from Homework 1 and 2 that includes keys and foreign keys. In particular, all keys of the given relational database must be enforced. Think about removing IDs you added for the DTD version. You don't have to do that, but you must ensure that when valid XML data files are mapped back to the relational tables, all keys and foreign keys will be satisfied. Submit the modified XML data file and the XML schema.

- `Supplement(Name, Producer)`

Supplement	
Name	Producer
Centrum	Lederle
One A Day	Bayer

- `Vitamin(Substance, Daily_Value, Unit)`

Vitamin		
Substance	Daily_Value	Unit
Vitamin A	5000	IU
Vitamin C	60	mg
Biotin	300	mcg
Iron	18	mg

- `Contents(Name` $\rightarrow$ `Supplement, Substance` $\rightarrow$ `Vitamin, Quantity)`

Contents		
Name	Substance	Quantity
Centrum	Vitamin A	5000
One A Day	Vitamin A	5000
Centrum	Vitamin C	60
One A Day	Vitamin C	60
Centrum	Biotin	30
Centrum	Iron	18
One A Day	Iron	27

e) Please read the following page about Sitemaps for Google formatted as XML:

[https://support.google.com/webmasters/answer/183668?hl=en&ref_topic=4581190]

More information is available here:

[http://www.sitemaps.org/protocol.html]

The XML schema for sitemaps can be downloaded from this address:

[http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd]

Develop a sitemap (XML file) with at least three URIs that can be validated according to this schema. For instance, you could create a sitemap for the homepage of this course and some of its subpages. You can also use your own web pages as an example.

f*) This exercise is optional. If you like, you could look at the Google extension to sitemaps for Image Sitemaps (and use it in your example):

[https://support.google.com/webmasters/answer/178636]

An XML schema file for this is available here:

[https://www.google.com/schemas/sitemap-image/1.1/sitemap-image.xsd]