

Pflichtmodul Informationssysteme (SS 2014)

Prof. Dr. Jens Teubner

Leitung der Übungen: Marcel Preuß, Sebastian Breß, Martin Schwitalla, Karolina Hilkens

Übungsblatt Nr. 12

Ausgabe: 25.06.2014

Abgabe: 02.07.2014

Gegeben sei das XML-Dokument `dilbert.xml`:

```
<?xml version='1.0' encoding='iso-8859-1'?>
<strip copyright='United Feature Syndicate' year='2000'>
  <prolog>
    <series href='http://www.dilbert.com'>Dilbert</series>
    <author>Scott Adams</author>
    <characters>
      <character id='phb'>The Pointy-Haired Boss</character>
      <character id='dilbert'>Dilbert</character>
      <character id='wally'>Wally</character>
    </characters>
  </prolog>
  <panels length='3'>
    <panel no='1'>
      <scene visible='phb dilbert'>
        Pointy-Haired Boss and Dilbert sitting at table.
      </scene>
      <bubbles>
        <bubble speaker='phb'>
          We have a gigantic database full of customer behavior information.
        </bubble>
      </bubbles>
    </panel>
    <panel no='2'>
      <scene visible='dilbert'> Dilbert, looking enthusiastic. </scene>
      <bubbles>
        <bubble speaker='dilbert'>
          Excellent. We can use non-linear math and data mining technology
          to optimize our retail channels!
        </bubble>
      </bubbles>
    </panel>
    <panel no='3'>
      <scene visible='phb dilbert'>
        Pointy-Haired Boss and Dilbert sitting at table. </scene>
      <bubbles>
        <bubble speaker='phb'>
          If that's the same thing as SPAM, we're having a good meeting here.
        </bubble>
      </bubbles>
    </panel>
  </panels>
</strip>
```

Aufgabe 1 (DTD)

DTDs sind eine Möglichkeit, die Baumstruktur eines XML-Dokuments einzuschränken. Sie beschreiben die zulässige Struktur des Dokuments als *reguläre Baum-Grammatik*. Dabei gibt ein Satz von Regeln der Form

$$\textit{Tagname} \rightarrow \textit{RegulärerAusdruck}$$

zum Beispiel vor, welche Abfolge von Kindknoten innerhalb eines Elements mit Namen *Tagname* zulässig sind.

1. Machen Sie sich mit den Grundkonzepten Element- und Attributdeklaration von DTDs vertraut. Wir empfehlen dazu folgende Ressourcen im Internet:

<http://de.wikipedia.org/wiki/Dokumenttypdefinition>

<http://www.w3schools.com/dtd/>

2. Geben Sie eine DTD für `dilbert.xml` an.

Aufgabe 2 (XPath Anfragen)

1. Welche Elemente werden durch die folgenden XPath-Anfragen an das gegebene XML-Dokument ausgewählt?

(a) `/child::strip/child::panels/child::panel[attribute::no = "1"]/descendant-or-self::*`

(b) `/child::strip/child::panels/child::panel/preceding::panel`

(c) `/child::strip/child::prolog/child::characters/following::*`

2. Geben Sie zum gegebenen XML-Dokument XPath-Anfragen an, die die folgenden Fragen beantworten.

(a) Geben Sie die Namen aller Figuren im Dokument aus.

(b) Geben Sie den gesamten Text in einem "verärgerten" Ton gesprochen aus (d.h., der gesamte Text soll in Großbuchstaben ausgegeben werden).

Aufgabe 3 (Komplexe SQL-Anfragen)

1. Welche Präsidenten wurden in einem Jahr geboren, in dem ein anderer Präsident ins Amt eingeführt wurde?

Es sollen die Namen der Präsidenten, die in dem Jahr geboren wurden, die Namen der Präsidenten, die ins Amt eingeführt wurden, und die Jahre angegeben werden.

2. Um einem drohenden Rückgang der Bevölkerungszahl entgegenzuwirken, will die amerikanische Regierung eine Kampagne ins Leben rufen, die dafür werben soll, mehr Kinder zu bekommen. Auf der Suche nach einem Namen für die Kampagne hat man sich auf die folgenden Regeln geeinigt:

- Die Kampagne soll nach einem amerikanischen Präsidenten benannt werden.
- Namensgeber soll derjenige Präsident werden, der die meisten Kinder hatte.
- Um allen Präsidenten die gleichen Chancen zu geben, soll die Anzahl der Kinder jeweils durch das Alter des Präsidenten bei seinem Tod geteilt werden.
- Es kommen deshalb nur Präsidenten in Frage, die schon gestorben sind.
- Weiterhin sollen nur Präsidenten berücksichtigt werden, die nicht mehrfach verheiratet waren.

Entwickeln Sie eine (möglicherweise genestete) SQL-Anfrage, mit deren Hilfe sich der Namensgeber für die Kampagne bestimmen lässt. Es sollen die in Frage kommenden Präsidentennamen nebst Kinderanzahl/Todesalter-Quotient ausgegeben werden und die Ausgabe soll nach dem Quotienten sortiert werden.