

Klausur “Semistrukturierte Daten und XML”
Sommersemester 2008
Prof. Dr. Wolfgang May
16. Juli 2008, 10-12 Uhr
Bearbeitungszeit: 90 Minuten

Vorname:

Nachname:

Matrikelnummer:

Bei der Klausur sind **keine Hilfsmittel** (Skripten, Taschenrechner etc.) erlaubt. Handies müssen ausgeschaltet sein. Papier wird gestellt. Benutzen Sie nur die **ausgeteilten**, zusammengehefteten **Blätter** für Ihre Antworten. Schreiben Sie mit blauem/schwarzem Kugelschreiber, Füller etc.; Bleistift ist nicht erlaubt.

Zum **Bestehen** der Klausur sind **45** Punkte hinreichend.

- meine Note soll mit Matrikelnummer so bald wie möglich auf der Vorlesungs-Webseite veröffentlicht werden.
- meine Note soll nicht veröffentlicht werden; ich erfahre sie dann vom Prüfungsamt (bzw. für nicht von einem Prüfungsamt verwaltete Teilnehmer: beim Abholen des Scheins).

	Max. Punkte	Erreichte Punkte
Aufgabe 1 (Entwurf, DTD)	23	
Aufgabe 2 (XPath, XQuery)	22	
Aufgabe 3 (XQuery oder XSLT)	15	
Aufgabe 4 (XSLT)	20	
Aufgabe 5 (Konzeptuelles)	10	
Summe	90	

Note:

Die Klausur behandelt die Datenbank einer Fluggesellschaft: Flugstrecken, Flugplan, Flugbuchungen.

Gegeben sind die folgenden Dinge:

- Ein Rahmen (siehe letzte Seite, wird in Aufgabe 1 vervollständigt)
- Daten zu Flughäfen. Da diese prinzipiell von allen Fluggesellschaften benutzt werden, wird angenommen, dass alle Flughafendaten weltweit (mit eindeutigen Flughafencodes) von der IATA (International Air Transport Assoc.) vorgegeben sind.

Neben den bereits eingefügten Flughäfen soll die Datenbasis die folgenden fluglinienspezifischen Informationen enthalten:

- die Flugstrecken der Gesellschaft: z.B. Flugnummer "LH123" von "FRA" nach "LIS", Abflug 12h30min, Ankunft 14h30min und Flugnummer "LH124" von "FRA" nach "JFK", Abflug 8h40min, Ankunft 11h20min (alle Zeiten geben die lokale Zeit an). Sommer- und Winterzeit werden nicht berücksichtigt.
- Annahme: Kein Flug geht über Mitternacht, d.h. alle Flüge landen an demselben Tag, an dem sie gestartet sind.
- Annahme: jede Flugstrecke wird täglich geflogen; hierzu sind also keine speziellen Daten notwendig.
- Für jeden Tag werden die Buchungen gespeichert. Buchungen werden, z.B. von Online-Web-Services etc., als XML-Fragmente in HTTP-Messages verschickt. Der Inhalt einer solchen Nachricht sieht folgendermaßen aus:

```
<booking flight="LH123" firstname="John" name="Doe" nationality="uk">  
  <date year="2008" month="07" day="16"/>  
</booking>
```

- Annahme: Name+Vorname identifizieren eine Person eindeutig.

Aufgabe 1 (Entwurf, DTD [23 Punkte])

Überlegen Sie sich, wie Sie Flüge und Buchungen in der XML-Datenbasis ablegen möchten. Überlegen Sie sich, wie Sie am besten den Datentyp "Zeit" (für Anflugs- und Ankunftszeit) in XML repräsentieren, damit Sie darauf auch (zumindest ein bisschen) rechnen können.

1. Erweitern Sie das angegebene XML-Fragment entsprechend Ihren Überlegungen mit Beispieldaten zu Flügen und Buchungen. (8 P)
2. Geben Sie die DTD zu Ihrem Entwurf an. (15 P)

Aufgabe 2 (XPath, XQuery [22 Punkte])

Bearbeiten Sie die folgenden Aufgabenteile (1)–(4) entsprechend der von Ihnen entworfenen Datenbasis. Geben Sie zu jedem Aufgabenteil eine XPath- oder XQuery-Anfrage an, die das Ergebnis in dem jeweils angegebenen Format ausgibt (falls kein Format angegeben, beliebig).

1. Geben Sie einen XPath-Ausdruck an, der die Flugnummern aller Lufthansa-Flüge ausgibt, die *John Doe* im Juli 2008 gebucht hat. (3 P)
2. Geben Sie die Namen aller Flughäfen an, von denen man Albany (Code: ALB) mit genau einer Zwischenlandung erreichen kann. Die Ausgabe darf Duplikate enthalten (5 P).
3. Geben Sie die Namen aller Flughäfen an, von denen man alle portugiesischen Flughäfen (Landescode: "P") direkt erreichen kann. (5 P)
4. welche Passagiere fliegen am 16.7.2008 aus einem anderen Land in die USA, und am nächsten Tag mit einem Inlandsflug weiter? Geben Sie die Namen der Passagiere sowie die Flugnummern an (9 P).

Name:

MatNr.:

Aufgabe 3 (XQuery oder XSLT [15 Punkte])

Es soll eine HTML-Tabelle (Flugnummer, Start, Ziel) erstellt werden, die genau diejenigen Flüge enthält, die mindestens 8 Stunden dauern.

1. Diese Frage dient der Vorüberlegung und ist nicht in XQuery zu beantworten, sondern einfach mit einem Wert: Wieviel Uhr ist es in Frankfurt, wenn Flug LH123 in Lissabon landet? (2 P)
2. Geben Sie eine XQuery-Anfrage oder ein XSLT-Stylesheet an, das die gewünschte Tabelle erzeugt. (13 P)

Aufgabe 4 (XSLT [20 Punkte])

Jede Fluggesellschaft muss für ihre Flüge, die auf einem US-amerikanischen Flughafen landen, zwei Tage vorher einen Report übermitteln, der diese Flüge mit den Namen aller Passagiere gemäß folgender DTD enthält:

```
<ELEMENT airlineReport (name, flight+)>
  <ATTLIST airlineReport month CDATA #REQUIRED
                        day CDATA #REQUIRED
                        year CDATA #REQUIRED>
<ELEMENT name (#PCDATA)> <!-- name of the airline that submits the report -->
<ELEMENT flight (person+)>
  <ATTLIST flight code ID #REQUIRED>
<ELEMENT person (givenname, familyname, nationality)>
  <ELEMENT givenname (#PCDATA)>
  <ELEMENT familyname (#PCDATA)>
  <ELEMENT nationality (#PCDATA)>
```

1. Geben Sie ein kurzes charakteristisches XML-Fragment an, wie dieser Report für die Lufthansa aussieht. (3 P)
2. Schreiben Sie ein XSLT-Stylesheet, das z.B. mit

```
saxonXSL -s lufthansa.xml -xsl makeUSreport.xsl month=07 day=16 year=2008
```

aufgerufen wird und für den angegebenen Tag den entsprechenden Report der Lufthansa entsprechend Ihrer Datenbasis erstellt. (17 P)

Name:

MatNr.:

Aufgabe 5 (Konzeptuelles [10 Punkte])

In einer deklarativen Datenbank-Sprache kann man nicht nur Anfragen stellen, sondern auch Daten ändern, löschen und einfügen. Letzteres geschieht z.B. in SQL durch das Konstrukt

```
UPDATE table_name
SET attribute = new_value
WHERE condition .
```

Welche Informationen muss eine solche Sprache (denken Sie z.B. an eine Ergänzung zu XQuery) zum Einfügen, Löschen, oder Ändern von Werten in XML-Dokumenten ausdrücken können?

Wie würden Sie in einer solchen Sprache eine Flugplanänderung, z.B. die Ankunftszeit von LH123 auf 14:40 Uhr zu setzen, ausdrücken?

Name:

MatNr.:

Trennen Sie diese Seite ggf. ab, um damit die restlichen Aufgaben bearbeiten zu können (Kopien dieses Blattes sind verfügbar; versehen Sie diese bitte zuerst mit Ihrem Namen).

```
<lufthansa>
  <airport code="FRA" city="Frankfurt/Main" country="D" timezone="+1">
    <name>Rhein-Main-Flughafen Frankfurt</name>
  </airport>
  <airport code="LON" city="London" country="GB" timezone="0">
    <name>London Heathrow</name>
  </airport>
  <airport code="LIS" city="Lisbon" country="P" timezone="0">
    <name>Lisboa Portela</name>
  </airport>
  <airport code="JFK" city="New York" country="USA" timezone="-5">
    <name>John F. Kennedy Airport New York</name>
  </airport>
  <airport code="ALB" city="Albany" country="USA" timezone="-5">
    <name>Albany International Airport</name>
  </airport>
  <!-- weitere Flughäfen -->
```