

**Aufgabe 1a)**

**[3 Punkte]**

Worin besteht der Unterschied zwischen well-formed und validem XML? Wie ändert sich ggf. diese Definition / dieser Unterschied, wenn Namensräume benutzt werden?

**Aufgabe 1b)**

**[4 Punkte]**

Beantworten/Bearbeiten Sie folgende Fragestellungen für die unten stehenden XML-Dateien. Ein Blick in die W3C-Recommendation für Namensräume hilft!<sup>1</sup>

- Ist die Deklaration der Namensräume korrekt?
- Ist deren Benutzung valide/korrekt, oder handelt es sich um ein Dokument, dass das Namensraumkonzept verletzt?
- Für den Fall, dass das Namensraumkonzept verletzt wurde: Korrigieren Sie das Dokument!

Markieren Sie in jedem der Dokumente, welcher Bestandteil (Element, Attribut, ggf. Text) zu welchem Namensraum gehört und dokumentieren/erläutern Sie nicht offensichtliche Zuordnungen **kurz**.

**XML-Datei #1:**

```
<?xml version="1.0" encoding="UTF-8"?>
<root
  xmlns:n1="test://mähdrescher"
  xmlns:n2="test://m%E4hdrescher"
  xmlns:n3="test:opaque_uri#">
  <n1:child n3:attr="attributwert"
    attr="attributwert"
    n1:attr="noch_ein_wert"/>
  <n2:child n3:attr="Hallo" n2:attr="Welt"/>
  <n3:child attr="Test"/>
</root>
```

**XML-Datei #2:**

```
<?xml version="1.0" encoding="UTF-8"?>
<root
  xmlns:n1="test://mähdrescher"
  xmlns:n2="test://mähdrescher"
  xmlns:n3="test:opaque_uri#">
  <n1:child n3:attr="attributwert" n2:attr="attributwert"
    n1:attr="noch_ein_wert"/>
  <n2:child n1:attr="Hallo, " attr="schöne"
    n2:attr="Welt"/>
  <child attr="Test"/>
</root>
```

<sup>1</sup> <http://www.w3.org/TR/REC-xml-names/>

DBLP (<http://dblp.uni-trier.de/>) ist ein händisch gepflegtes Online-Portal für wissenschaftliche Publikationen aus dem Bereich der Informatik/Computer Science. Machen Sie sich bitte kurz mit dem Online-Portal vertraut, indem sie nach „Schmidt-Thieme“ als Person suchen. Sie sollten ca. 77 Publikationen finden – viele davon zusammen mit Mitgliedern des ISMLL.

Neben dem HTML-basierten online-Portal bietet DBLP ebenfalls eine XML-basierte Datenschnittstelle an. Diese ist hier dokumentiert:

<http://dblp.uni-trier.de/xml/docu/dblpxmlreq.pdf>

Alle folgenden Aufgaben sind über die XML-Schnittstelle zu lösen!

**Aufgabe 2a)**

**[3 Punkte]**

Suchen Sie nach den Personen „Schmidt-Lars“ und „Schmidt-Thieme“ und geben Sie sie jeweiligen Ergebnisse als (einzelnes) XML-Dokument an.

**Aufgabe 2b)**

**[4 Punkte]**

Suchen Sie nach Publikationen von „Lars Schmidt-Thieme“. Argumentieren und diskutieren Sie **kurz** die XML-Dokumentmodellierung in Bezug auf:

- Warum ist „Lars Schmidt-Thieme“ als Attribut (`name`), und nicht als Element modelliert – und welche Auswirkungen hätte eine Änderung?
- Welche Auswirkungen hat es, wenn  
`<dblpkey>reference/rsh/MarinhoNSJHSS11</dblpkey>` in  
`<dblpkey ref="reference/rsh/MarinhoNSJHSS11"/>` geändert würde?

**Allgemeiner Bearbeitungshinweis:**

Geben Sie zu ihren Lösungen stets auch eine grobe Lösungsskizze mit an, damit ihre Ergebnisse nachvollziehbar sind!

Geben Sie die Lösung bis zum 13.06.2012, 8:00 Uhr ab, indem Sie sie entweder in Fach 45 einwerfen (links dreimalig als Buch getackert!), oder bevorzugt per E-Mail mit dem Betreff-prefix [is2] an [wi2@ismll.de](mailto:wi2@ismll.de) schicken. Sie erhalten eine Eingangsbestätigung bei E-Mailabgabe.

Aufgabe 3a)

[3 Punkte]

Erzeugen Sie eine **einzige** XML Schema Elementdefinition, um **alle** folgenden Dokumentfragmente abbilden zu können:

Dokumentfragment 1:

```
<pc manufacturer="dell" yr="2012">  
  <mainboard manufacturer="asus" formfactor="atx"></mainboard>  
  <ram manufacturer="infineon">1 GB</ram>  
  <ram manufacturer="kingston">2 GB</ram>  
</pc>
```

Dokumentfragment 2:

```
<pc manufacturer="apple" yr="2012">  
  <mainboard manufacturer="asus" formfactor="btx"/>  
  <ram manufacturer="corsair" bank="2">1 GB</ram>  
</pc>
```

Dokumentfragment 3:

```
<pc manufacturer="ibm" yr="2010">  
  <mainboard manufacturer="gigabyte" formfactor="atx"></mainboard>  
</pc>
```

Aufgabe 3b)

[3 Punkte]

Worin besteht der Unterschied zwischen `<xs:sequence>` und `<xs:choice>` bei Verwendung unterschiedlicher `min/maxOccurs`?

Benutzen Sie folgende alternative Modellierungsmöglichkeiten eines Namens, um die Unterschiede aufzuzeigen:

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="name1">
    <xs:complexType>
      <xs:sequence minOccurs="0" maxOccurs="unbounded">
        <xs:element name="vorname"/>
        <xs:element name="nachname"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>

  <xs:element name="name2">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="vorname" minOccurs="0" maxOccurs="unbounded"/>
        <xs:element name="nachname" minOccurs="0"
maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>

  <xs:element name="name3">
    <xs:complexType>
      <xs:choice minOccurs="0" maxOccurs="unbounded">
        <xs:element name="vorname"/>
        <xs:element name="nachname"/>
      </xs:choice>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

Wenn Modellierungen unterschiedliche Ergebnisse erzeugen können: Geben Sie ein Beispiel an, für die die eine Modellierung funktioniert, die andere jedoch nicht.

Wenn Modellierungen gleiche Ergebnisse erzeugen können: Welche Konsequenzen ergeben sich auf der einen, oder anderen, Modellierungsweise?

Aufgabe 3c&d)

[2+1 Punkte]

Gegeben ist folgendes XML Dokument:

```
<order xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:noNamespaceSchemaLocation="schema6.xsd">
  <Ware id="1" asin="3540011838">
    19,95 Euro
  </Ware>
  <Ware id="1" asin="3534144953" preis="29,90" einheit="Euro"/>
</order>
```

sowie folgendes Fragment eines XML Schemas

```
<schema xmlns="http://www.w3.org/2001/XMLSchema">
  <element name="order">
    <complexType>
      <!-- bitte hier die Definition ergänzen -->
    </complexType>
    <unique name="u">
      <selector xpath="./Ware"></selector>
      <field xpath="@id"></field>
    </unique>
  </element>
</schema>
```

- Ergänzen Sie das complexType-Element, sodass dieses eine möglichst präzise und valide Definition<sup>2</sup> im Sinne des Quelldokuments enthält.
- Ist für den hier im Ausschnitt dargestellten Fall das XML Dokument valide? Wenn nicht: wo liegt der Fehler, und was müsste geändert werden?

---

<sup>2</sup> Modellieren Sie mindestens die Kindelemente Ware, sowie deren Attribute!