Question 1: XML Storage + RDF

An RDF document is actually an XML document and therefore, we can derive a DTD for it.

Consider the following fragment of RDF instance (abbreviated syntax) in Listing 1 and answer the following questions:

a) Give an XML DTD of which this RDF fragment is a valid instance.
b) Draw the DTD graph and provide a relational database schema to store the RDF fragment, using the inlining algorithm.

--- Listing 1. Fragment of an RDF instance---------------------------------------------------------------------

<ex:university rdf:ID="EPFL">
  <ex:name>EPFL</ex:name>
  <ex:departments>
    <rdf:Bag>
      <rdf:li rdf:resource="#IC"/>
      <rdf:li rdf:resource="#Physics"/>
    </rdf:Bag>
  </ex:departments>
</ex:university>

<ex:department rdf:ID="IC"/>
<ex:department rdf:ID="Physics"/>

--- End of listing 1. -----------------------------------------------------------------------------------------------------

SOLUTION:

a) Give an XML DTD of which this RDF document is a valid instance

<! ELEMENT ex:university (ex:name, ex:departments)>
<! ATTLIST ex:university rdf:ID CDATA #REQUIRED>
<! ELEMENT ex:name(#PCDATA)>
<! ELEMENT ex:departments(rdf:Bag)>
<! ELEMENT rdf:Bag (rdf:li)+>
<! ELEMENT rdf:li EMPTY>
<! ATTLIST rdf:li rdf:resource CDATA #REQUIRED >
<! ELEMENT ex:department EMPTY>
<! ATTLIST ex:department rdf:ID CDATA #REQUIRED>
b) Draw the DTD graph and provide a relational database schema to store the RDF fragment, using the inlining algorithm

**DTD graph**

```
  ex:name ← ex:university
    rdf:ID
  ex:departments
      ex:department
        rdf:Bag
          +
          rdf:li
            rdf:ID
            rdf:ID
            rdf:resource
```

**Top nodes**
ex:university: not reachable from any other node
ex:department: not reachable from any other node
rdf:li: child of + edge

**Relational schema**
ex_university (universityID, university_rdfid, root_type, university_name)
  (see Rule 4, 5, inlining algorithm, slide 19)

rdf_li(rdf_li_ID, root_type, parent_rdfli, rdfli_resource)
  (see Rule 6, slide 20)

ex_department(deptID, root_type, department_rdfid)
  (see Rule 6, slide 20)
Question 2: Graph Databases

Given is the following data graph

Data graph D

1.) Determine the data guide DG for the data graph D.
2.) Determine the non-deterministic schema graph NG constructed by using language equivalence.
3.) Identify all subsumption relations that exist among the two schema graphs DG and NG.
4.) Determine the simulation relationships which characterize the subsumption relations among the two schema graphs.

SOLUTION:

1.) Determine the data guide for the data graph D.
2.) Determine the non-deterministic schema graph based on language equivalence.

3.) Which subsumption relations exist among the two schema graphs (from questions 1. and 2.)?

Simulation Equivalence:

DG < DL
DL < DG

4.) Determine the relationships characterizing subsumption relations between two schema graphs.

<table>
<thead>
<tr>
<th>R1</th>
<th>(DG&lt; DL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL</td>
<td>DG</td>
</tr>
<tr>
<td>(d1)</td>
<td>(d1)</td>
</tr>
<tr>
<td>(d2, d3, d4)</td>
<td>(d2, d3, d4)</td>
</tr>
<tr>
<td>(d5, d6)</td>
<td>(d5, d6)</td>
</tr>
<tr>
<td>d7</td>
<td>(d7, (d5, d6, d7))</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>R2</th>
<th>(DL&lt; DG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL</td>
<td>DG</td>
</tr>
<tr>
<td>d1</td>
<td>(d1)</td>
</tr>
<tr>
<td>(d2, d3, d4)</td>
<td>(d2, d3, d4)</td>
</tr>
<tr>
<td>(d5, d6, d7)</td>
<td>(d5, d6, d7)</td>
</tr>
<tr>
<td>(d5, d6)</td>
<td>(d5, d6)</td>
</tr>
<tr>
<td>d7</td>
<td>(d7)</td>
</tr>
</tbody>
</table>