

General Instruction: After the second lecture of big data course, you should be able to answer the following questions.

1. What are database and database management system?
2. Why we need use a database management system to manage data, instead of file system?
3. What are three important features of data models?
4. What is the join operation between two relational tables?
5. What is DTD of XML document?
6. See an example of DTD and two XML documents. Check whether XML document 1 and 2 satisfy this DTD?

```
<!ELEMENT document
  (title*,subjectID,subjectname,prerequisite?,
   classes,assessment,syllabus,textbooks*)>
<!ELEMENT prerequisite (subjectID,subjectname)>
<!ELEMENT textbooks (author,booktitle)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT subjectID (#PCDATA)>
<!ELEMENT subjectname (#PCDATA)>
<!ELEMENT classes (#PCDATA)>
<!ELEMENT assessment (#PCDATA)>
<!ATTLIST assessment assessment_type (exam |
assignment) #IMPLIED>
<!ELEMENT syllabus (#PCDATA)>
<!ELEMENT author (#PCDATA)>
<!ELEMENT booktitle (#PCDATA)>
```

7. Note that indexes are relative to the parent. Consider the following data:

```
<x> <y/> <y/></x>
<x> <y/> <y/></x>
```

Answer the following XPath:

- (1) `x/y[1]`, (2) `x/y[position() = 1]`, (3) `(x/y)[2]`, (4) `x[1]/y[2]`

XML document 1:

```
<document>
  <title>Subjects available in Mechanical
Engineering.</title>
  <subjectID>2.303</subjectID>
  <subjectname>Fluid Mechanics</subjectname>
  <classes>4 hours per week </classes>
  <assessment>tutorial assignments and one 2hr exam at
end of course.</assessment>
  <syllabus>
    Fluid statics.
  </syllabus>
  <textbooks>
    <author>Foobar</author>
    <booktitle>The Study of Fluid
Mechanics</booktitle>
  </textbooks>
</document>
```

XML document 2:

```
<document>
  <subjectID>3.309</subjectID>
  <subjectname>Fluid Mechanics</subjectname>
  <prerequisite>
    <subjectID>1.001</subjectID>
    <subjectname>Mathematics</subjectname>
  </prerequisite>
  <classes>4 hours per week
</classes>
  <assessment assessment_type="exam">tutorial
assignments and one 2hr exam at end of course.
</assessment>
  <syllabus>
    Fluid statics.
  </syllabus>
  <textbooks>
    <author>Foobar</author>
    <booktitle>The Study of Fluid
Mechanics</booktitle>
  </textbooks>
</document>
```