XML Programming Assignment

Due: 25th November, 2014 11:59pm.

You can do this assignment in groups of two.

Data Description

Data consists of the following entities and their information:

- **Airports.** Each airport has a airport name, 3 letter code, city, and country. E.g., <OHare Airport, ORD, Chicago, USA>.
- **Airlines.** Each airline has a name and a code (e.g., <United Airlines, UA>).
- **Flights**. Each airline has a set of flights, where each flight has a number, and an ordered-sequence of airports (given by their codes) it visits. With each airport in the sequence, there is also an arrival and departure time.
  E.g., United Airlines has two flights: <70, [SFO, ORD, JFK, CDG]> and <1, [SFO, HKG, IGI]>; here [, , ,] denotes an ordered list of items. Arrival/departure times are not shown for simplicity.

Key Constraints

You need to incorporate the following data constraints using keys and foreign keys:

- Flight number is unique within an airline. So, there is only one flight with a number 370 in UA.
- Each flight only visits an airport once.
- Each airport-code in the flight information must be a valid code (i.e., for the UA370 above, ORD, JFK and CDG must exist in the set of airports).

TO DO ITEMS

1. **Create an XML schema** for the above data application. Incorporate the above key constraints. There are many possible ways of creating a scheme (e.g., "city" could be an attribute or a sub-element of "airport") -- make your own choices, and do the below two tasks accordingly.
2. **Create an XML document** for the given data instance (see next page). Make sure the document conforms to the schema you define above.
3. **Write XML queries** for the following queries (your queries should output an appropriate XML document):
   a. For each airport, give the set of airlines that have a flight departing from the airport between 12pm and 4pm.
   b. Find pairs of airports that have a non-stop connection from two different airlines.

XML Processor, Instructions, and Example

See http://www.cs.sunysb.edu/~hgupta/532/hws/xml-instructions.html

Submission Instructions

You will email a .zip or .tar file with the following items:

1. An XML schema document with a .xsd extension.
2. An XML document with an .xml extension. This must be validated against the XML document.
3. A query file (with the above two queries) in a .txt file.
4. An optional README file. Herein, include any explanation if your program is not working fully – so, that we can appropriately award partial credit.

We should be able to just load your .xsd and .xml files, and paste your queries one-by-one and see the results.

Data Instance

Airports

- <Charles De Gaulle, CDG, Paris, France>
- <Heathrow, LHR, London, England>
- <John F Kennedy, JFK, New York, USA>
- <O'Hare, ORD, Chicago, USA>
- <San Francisco International, SFO, San Francisco, USA>
- <Logan, BOS, Boston, USA>
- <Hong Kong International, HKG, Hong Kong, China>
- <Indira Gandhi International, IGI, New Delhi, India>

Airlines

- <United Airlines, UA>
- <American Airlines, AA>
- <Delta Airlines, DAL>
- <Air France, AF>

Flights (The below data is completely made-up)

Here, each flight is denoted as <number, [sequence of airports]>. The airports with *'s have a departure time of 2pm; use arbitrary arrival/departure times for the rest.

UA Flights

<70, [SFO, ORD, JFK*, CDG]>
<1, [SFO, HKG*, IGI]>
<50, [LHR, JFK, ORD]>

AA Flights

<356, [CDG, LHR*, BOS]>
<275, [LHR, JFK, SFO]>

Delta Flights

<55, [BOS, ORD, SFO]>
<240, [ORD, BOS, LHR*, CDG]>
<23, [JFK, LHR, IGI]>

Air France Flights

<42, [CDG, BOS*, ORD]>