

## Complementi di Basi di dati – Basi di dati II (secondo modulo)

Esame del 14 settembre 2009 – Compito A – Tempo a disposizione: 1 ora e 30 minuti (libri chiusi)

### Domanda 1 (10%)

Scrivere il DTD che consente di validare il seguente documento XML:

```
<?xml version="1.0" encoding="UTF-8"?><lettera numero="12">
Roma <data><giorno>14</giorno><mese>settembre</mese><anno>2009</anno></data>.
Si comunica che <cliente><nome>Mario</nome><cognome>Rossi</cognome></cliente> è il
fortunato vincitore della somma di <vincita>500</vincita> Euro. Per incassare la
vincita va sottoscritto un abbonamento alla rivista <rivista>Complementi di basi di
truffe</rivista> per <tempo>10</tempo> anni al prezzo in offerta speciale di
<prezzo>99</prezzo> Euro l'anno.
</lettera>
```

### Domanda 2 (35%)

Si vogliono esportare in formato XML il contenuto di una base di dati che memorizza informazioni relative gare di competizioni sportive e così strutturata:

- **Atleta**(Codice, Cognome, Nome, Età)
- **Competizione**(Codice, Nome, Organizzatore, Anno)
- **Gara**(Numero, Competizione, TipoGara, Giornata) con vincolo di integrità referenziale fra Competizione e la relazione Competizione
- **Classifica**(NumeroGara, Competizione, Atleta, Posizione) con vincoli di integrità referenziale fra gli attributi NumeroGara, Competizione e la relazione Tappa e fra Atleta e la relazione Ciclista

Scrivere un esempio di documento XML adatto a raccogliere dati di questo tipo e il relativo XML Schema.

### Domanda 3 (40%)

Considerare il seguente documento XML:

```
<vendite>
  <ricevuta numero="3244">
    <data><giorno>1</giorno><mese>9</mese><anno>2009</anno></data>
    <cliente>23456</cliente>
    <acquisti>
      <prodotto>123</prodotto><quantita>1</quantita><prezzo>3.35</prezzo>
      <prodotto>324</prodotto><quantita>2</quantita><prezzo>8.15</prezzo>
    </acquisti>
    <totale>11.50</totale>
  </ricevuta>
  <ricevuta numero="4321">
    <data><giorno>2</giorno><mese>9</mese><anno>2009</anno></data>
    <cliente>332211</cliente>
    <acquisti>
      <prodotto>254</prodotto><quantita>1</quantita><prezzo>4.05</prezzo>
      <prodotto>123</prodotto><quantita>2</quantita><prezzo>6.70</prezzo>
    </acquisti>
    <totale>10.75</totale>
  </ricevuta>
  <ricevuta numero="3254">
    <data><giorno>2</giorno><mese>9</mese><anno>2009</anno></data>
    <cliente>23456</cliente>
    <acquisti>
      <prodotto>324</prodotto><quantita>2</quantita><prezzo>8.15</prezzo>
      <prodotto>254</prodotto><quantita>2</quantita><prezzo>8.10</prezzo>
    </acquisti>
    <totale>16.25</totale>
  </ricevuta>
</vendite>
```

(continua sul retro del foglio)

Con riferimento a documenti XML di questo genere:

- (a) scrivere l'espressione XPATH che restituisce i codici dei clienti che, in un singolo acquisto, hanno comprato più di un prodotto spendendo complessivamente meno di 12 Euro.
- (b) scrivere il foglio di stile XSLT che restituisce il seguente file XML:

```
<statistica>
  <prodotto numero="123" quantita="3" incasso="10.05"/>
  <prodotto numero="324" quantita="3" incasso="16.30"/>
  <prodotto numero="254" quantita="4" incasso="12.15"/>
</statistica>
```

- (c) scrivere l'interrogazione XQuery che, per ogni data, restituisce un elemento vendite contenente i codici dei prodotti venduti in quel giorno, senza ripetizioni.
- (d) scrivere l'interrogazione XQuery che, per ogni prodotto, restituisce un elemento risultato contenente il codice del prodotto come attributo e un sottoelemento giorno per ogni data nella quale il prodotto è stato venduto almeno una volta contenete la data, la quantità complessiva venduta e l'incasso giornaliero per quel prodotto. Utilizzare allo scopo una funzione utente che, dato un codice di prodotto e una data, restituisce la quantità venduta e l'incasso giornaliero per quel prodotto.

**Domanda 4 (15%)**

Descrivere in maniera sintetica (mezza pagina) le varie fasi del processo di Knowledge Discovery con particolare riferimento agli obiettivi e alle tecniche del Data Mining,