Birkbeck
(University of London)

MSc and MRes Examination for Internal Students
MSc in Advanced Information Systems
MSc in Web Information Management
MRes in Computer Science
MSc in E-Business

School of Computer Science and Information Systems

Development of Internet Applications (COIY032P)

Date of examination: Wednesday 24 May 2006
Duration of paper: 14:30–16:30

There are six questions on this paper. Candidates should attempt any FOUR of them. Calculators are not permitted.

1. (a) Explain some of the difficulties applications may have in consuming (parsing and interpreting) HTML. (4 marks)

   (b) By means of examples, demonstrate how the use of appropriate markup might promote the development of useful applications in future. (6 marks)

   (c) Suppose that we want to transform a source XHTML document into a target XHTML document as follows. The `title` of the source is to become an `h1` heading at the start of the target. This should be followed in the target by a `table`, where each row (`tr` element) contains information about each anchor (`a`) element in the source. Two table cells (`td` elements) are used to record the information: the first contains the value of the `href` attribute; the second contains the contents of the anchor element itself. Write XSLT rules to achieve this transformation (do not worry about the `stylesheet` element itself). (15 marks)

2. (a) Explain why certain start and end tags can be omitted in HTML. (4 marks)

   (b) Explain two specific, key differences between SGML and XML. (4 marks)

   (c) Assume that we have an XML file representing information about CD recordings. The file contains elements with name `CD`, each of which has a `publisher` attribute and a number of `performance` elements as children. Write an XPath expression that will return all `CD` elements that have a `publisher` value of "Decca" and more than 3 `performance` elements as children. (7 marks)

   (d) Explain what the `transformNode` method available in Internet Explorer allows one to do. (5 marks)

   (e) Explain or draw a diagram to show how a browser uses DNS lookup to help retrieving a web page based on its URL. (5 marks)

3. (a) Discuss why the namespace mechanism for XML as proposed by the W3C has caused much confusion. (6 marks)

   (b) Explain by means of an example how a location step of the form $x[y][z]$ in an XPath expression can give rise to a different answer to a location step of the form $x[y \text{ and } z]$ for particular instantiations of x, y and z. (In other words, you should find a way to substitute XPath expressions for x, y and z that makes the above true.) (6 marks)

   (c) Describe the purpose of the document object model (DOM) as defined by the W3C. (8 marks)

   (d) List the principal responsibilities of the email protocols POP, IMAP and SMTP. (5 marks)

4. Consider the following text appearing at the beginning of an XML file

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE plist PUBLIC "-//Apple Computer//DTD PLIST 1.0//EN"
                       "http://www.apple.com/DTDs/PropertyList-1.0.dtd">
```

along with the following DTD available from the URL given above:

```
<!ENTITY % plistObject
         "(array | date | dict | integer | string | true | false )" >
<!ELEMENT plist %plistObject;>
<!ATTLIST plist version CDATA "1.0" >
<!ELEMENT array (%plistObject;)*>
<!ELEMENT dict (key, %plistObject;)*>
<!ELEMENT key (#PCDATA)>
<!ELEMENT string (#PCDATA)>
<!ELEMENT date (#PCDATA)>
<!ELEMENT true EMPTY>
<!ELEMENT false EMPTY>
<!ELEMENT integer (#PCDATA)>
```

(a) Give the names of the constructs used at the beginning of the XML file and explain the purpose and format of each. (7 marks)

(b) Explain the purpose of the first three declarations in the DTD. (7 marks)

(c) Consider the following fragment of XML:

```
<key>Playlists</key>
<array>
  <dict>
    <key>Name</key><string>Library</string>
    <key>Master</key><true/>
    <key>Playlist Items</key>
    <array>
      <dict>
        <key>Track ID</key><integer>56</integer>
      </dict>
      <dict>
        <key>Track ID</key><integer>47</integer>
      </dict>
    </array>
  </dict>
</array>
```

Explain, for each type of element above, whether or not the above fragment forms valid content for a `dict` element. (11 marks)

5. (a) Consider the following fragment of XML:

```
<book xmlns="http://www.xxx.com">
   <author xmlns="http://www.yyy.com">
      <name>John Coetzee</name>
   </author>
   <year>1999</year>
</book>
```

For each of the elements `book`, `author`, `name` and `year`, state which namespace it is in. (4 marks)

(b) Describe 3 ways in which XSL provides a more powerful stylesheet mechanism for XML than CSS does. (3 marks)

(c) Name 5 URL schemes. (5 marks)

(d) Explain in what ways an HTTP client can alter the default caching behaviour provided in HTTP/1.1. (8 marks)

(e) Give the disadvantages of using CGI as a server-side processing technology. (5 marks)

6. (a) Consider the use of XML to represent information about the results of a particular tennis tournament. The application is described as follows. A tournament is a sequence of matches. Each match is between two players, each of whom has a name and a ranking which can be any integer greater than or equal to 1. For each player the result of the match is either "won" or "lost". The score of a match is a sequence of between 3 and 5 sets (inclusive), each of which has a value such as 6-4, i.e., one digit followed by a hyphen followed by another digit (we ignore sets involving more than 9 games, as well as tie-breaks in this simplification). Using the XML schema definition language, write down the declarations for the `player` and `set` elements only. (13 marks)

(b) Consider the following fragment of JavaScript code:

```
var elem = document.getElementById("target");
elem.parentNode.removeChild(elem);
```

Assuming that the `document` object has been instantiated as a DOM document, explain what the above fragment of code does. (4 marks)

(c) Contrast between the TCP and IP protocols in terms of the functionality they each provide. (8 marks)

