Birkbeck
(University of London)

MSc and MRes Examination for Internal Students

MSc in Advanced Information Systems
MRes in Computer Science
MSc in E-Business

School of Computer Science and Information Systems

Development of Internet Applications (COIY032P)

Date of examination: Friday 3 June 2005
Duration of paper: 14:30–16:30

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

1. Consider the following XML document:

```
<?xml version="1.0"?>
<!DOCTYPE household [
<!ELEMENT household (person+)>
<!ELEMENT person   (name)>
<!ELEMENT name     (#PCDATA)>
<!ATTLIST  person
  pno      ID            #IMPLIED
  role     (adult|child)  #IMPLIED
  spouse   IDREF         #IMPLIED
  siblings IDREFS        #IMPLIED>
]>
<household>
  <person pno="john" spouse="janet"><name>John</name></person>
  <person role="adult" spouse="john jim"><name>Janet</name></person>
  <person pno="jim" role="dependent" siblings="john"><name/></person>
</household>
```

(a) Give all of the reasons why the above document is not valid. (6 marks)

(b) Given a document valid with respect to the same internal DTD, write an XPath expression that returns the `name` element of the `person` who has the person with `name` `Mary` as a `spouse`. (5 marks)

(c) Write an XSLT template rule that will take as input a valid `household` document and transform it to an output document as follows.

- The document element of the output is also `household`.
- Each `person` element in the input is transformed as follows.
    - If the `role` of the `person` is `adult`, then an `adult` element is output.
    - If the `role` of the `person` is `child`, then a `child` element is output.
    - If no `role` is specified for the `person`, then a `person` element is output.

    In each of these 3 cases the value of the corresponding `name` element in the input becomes the value of the `adult`, `child` or `person` element in the output.

So the first `person` above would be transformed to `<person>John</person>`, while the second would be transformed to `<adult>Janet</adult>`. (14 marks)

2. (a) Describe the differences between the TCP and UDP protocols, stating why UDP might be used in preference to TCP. (7 marks)

(b) XML allows information to be represented using either attributes or elements. Explain the factors to be taken into account when deciding which to use when designing a document type. (6 marks)

(c) For each of the following 3 XPath location steps which use the *full* syntax, write down an equivalent location step using the *abbreviated* syntax wherever possible:

 i. `child::section[descendent-or-self::image]`
 ii. `parent::chapter`
 iii. `preceding-sibling::*[position()=1]`

If abbreviated syntax cannot be used, explain why this is the case. (6 marks)

(d) *Canonical XML* is a representation of XML that, among other things, removes all irrelevant whitespace characters. Explain the relevance of canonical XML to the notion of a *weak validator*, in contrast to a *strong validator*, for resources used in HTTP caching. (6 marks)

3. (a) Both XML schema and XSL use elements named `element` and `attribute`. Use this fact to motivate the need for namespaces in XML. (5 marks)

(b) Certain characters are used with a special meaning in both URIs and XML. Explain the mechanism used for escaping the special meaning of these characters in the syntax for URIs and the syntax for XML. (6 marks)

(c) Why is XML seen as a better platform on which to build e-commerce solutions than historical EDI approaches? What is meant by *legacy data* in this context? (10 marks)

(d) Explain the relative advantages and disadvantages of processing dynamic Web content on the client versus the server. (4 marks)

4. (a) Give the full names for 4 application layer protocols used on the Internet. (4 marks)

(b) Consider the following DTD fragment:

```
<!ELEMENT book      (author*, title?, isbn, price)>
<!ELEMENT author    (first-name?, last-name)>
```

   i. Assume that the information about the `isbn` of a `book` is to be represented as an attribute rather than an element. Write down an appropriate attribute declaration for an `isbn` attribute. (3 marks)

   ii. Write an XPath expression to find `books` that have more than 2 `authors`. (4 marks)

   iii. Explain why the relative XPath expression

   `book[author/first-name][author/last-name]`

   is equivalent to the expression `book[author/first-name]` on documents valid with respect to the above DTD fragment. (4 marks)

(c) Explain some of the problems with web services that will need to be overcome if they are to be adopted more widely. (10 marks)

5. (a) Give the reasons for the introduction of MIME types on the Internet. (5 marks)

   (b) Consider the following fragment of JavaScript code:

```
var elem = document.getElementById("target1");
var node = document.createElement("li");
var text = document.createTextNode("Hello");
node.appendChild(text);
elem.appendChild(node);
```

   Assuming that the `document` object has been instantiated as a DOM document, explain, line by line, what the above fragment of code does. (10 marks)

   (c) Distinguish among three different approaches to developing server-side processing technologies. Name two technologies that follow each of the approaches. What is a disadvantage of all these approaches. (10 marks)

6. Consider an application where information about music recordings is represented in an XML file.

   (a) A `recording` element has the following content model. It comprises an `artist` element followed by one or more `track` elements, where each `track` element contains a `title` element followed by an optional `length` element. Write down an XML schema definition for the `recording` element that captures these requirements. (10 marks)

   (b) Now produce a template rule for an XSL stylesheet that will match a `recording` element (as defined in Part (a)) and will output an HMTL table row for each track. Each row should contain the track number, the track title and the artist. (10 marks)

   (c) Now consider the following change in requirements for the content model of a `recording` element (as defined in Part (a)): it comprises either

   i. an `artist` element followed by one or more `track` elements, where each `track` element contains a `title` element, or

   ii. one or more `track` elements, where each `track` element contains both a `title` element and an `artist` element.

   Can the new requirements for `track` elements be satisfied by a modified XML schema definition for `recording` elements? Can they be satisfied by using a DTD? There is no need to write down either an XML schema or a DTD, but you should explain your answers. (5 marks)