Birkbeck
(University of London)

School of Computer Science and Information Systems

Development of Internet Applications (COIY032P)

Exam Solutions 2007

1. (a) The W3C Internationalization Activity tries to ensure that Web formats and protocols are usable worldwide in all languages and in all writing systems. Unicode is designed to be a universal system for encoding all the characters in all the world's languages. UTF-8 is the form of Unicode used in XML. It uses 1 to 4 bytes to represent each character. (8 marks)

   (b) A mixed content model is one that allow both elements and text to appear. For example, the declaration

   ```
   <!ELEMENT a (#PCDATA | b | c)* >
   ```

   states that element `a` can have both text nodes and elements `b` and `c` as children. Mixed content models are indicated in XSDL by including the attribute `mixed="true"` in a declaration. (6 marks)

   (c) The document is well-formed and has been found to satisfy all the constraints in an associated DTD or XML schema definition. (2 marks)

   (d) (i) `//award[year='2000']/title`.
       (ii) `//award[title='Amsterdam'][year='1998']/author`. (9 marks)

2. (a)
   ```
   <!ELEMENT stock      ((new-car | used-car)*) >
   <!ELEMENT new-car    (model, price) >
   <!ELEMENT used-car   (model, mileage, price) >
   <!ELEMENT model      (#PCDATA) >
   <!ELEMENT mileage    (#PCDATA) >
   <!ELEMENT price      (#PCDATA) >
   <!ELEMENT condition (#PCDATA) >
   <!ATTLIST price
     currency   CDATA            #REQUIRED>
   <!ATTLIST used-car
     condition (good|excellent)  #IMPLIED>
   ```
   With XSDL, `price` could be defined as a decimal number and mileage as as integer. Also `condition` could be declared as an element and still be restricted to one of two values. (13 marks)

   (b) URI is the more general term which includes any URL or URN. URLs and URNs do not overlap, since every URN starts with `urn:`. (5 marks)

   (c) When applied to a particular node in a DOM tree, `getElementByTagName` finds all descendant nodes whose name matches that given as an argument to `getElementByTagName`. As an alternative, we would have to use the method `childNodes` to retrieve the collection of children, and check their names using `nodeName`. This would have to be done repeatedly for each descendant of the given node. (7 marks)

3. (a)
```
<xsd:element name="CD">
  <xsd:complexType>
    <xsd:sequence>
      <xsd:element name="composer"     type="xsd:string"/>
      <xsd:element name="performance"  maxOccurs="unbounded">
        <xsd:complexType>
          <xsd:sequence>
            <xsd:element name="composition"  type="xsd:string"/>
            <xsd:sequence minOccurs="0">
              <xsd:element name="orchestra"  type="xsd:string"/>
              <xsd:element name="conductor"  type="xsd:string"/>
            </xsd:sequence>
          </xsd:sequence>
        </xsd:complexType>
      </xsd:element>
      <xsd:element name="length" type="xsd:timeDuration" minOccurs="0"/>
    </xsd:sequence>
  </xsd:complexType>
</xsd:element>
```

(16 marks)

(b)
```
num1 = window.prompt("Enter first number", "0");
num2 = window.prompt("Enter second number", "0");
if (num1 < num2)
      document.write(num2, " is larger")
else if (num1 > num2)
      document.write(num1, " is larger")
else
      document.write("The numbers are equal")
```

(9 marks)

4. (a)
```
<tree>
  <node>
    <node>
      <node>a</node>
      <node>a</node>
    </node>
    <node>
      <node>a</node>
      <node>a</node>
    </node>
  </node>
</tree>
```

(4 marks)

(b) The syntax relies on special attributes whose names start with `xmlns:`. The rest of the name gives the so-called namespace prefix. The value of the attribute gives the namespace URI. The scope of a namespace declaration is the element containing the declaration and all its descendants. (6 marks)

(c) Persistent connections send multiple request/response interactions over a single HTTP connection. They were introduced in HTTP/1.1 because of poor performance of HTTP/1.0 due to its use of separate connection for each request/response. The advantages of persistent connections include: saves CPU time and memory because of fewer TCP connections and fewer TCP control blocks, requests and responses can be pipelined (see below), network load is reduced because fewer packets are sent, and there is greater

tolerance to HTTP version inconsistencies since errors can be reported to the client without the penalty of closing the TCP connection. In pipelining a client can make multiple requests without waiting for each response from the server. This results in better utilisation of the connection. A consequence of pipelining is that the server must send responses back in the same order as requests were made (since HTTP is stateless). (15 marks)

5. (a) The `descendant` axis consists of nodes whose start tag follows the start tag of c and whose end tag precedes the end tag of c. The `following` axis consists of nodes whose start tag follows the end tag of c. (6 marks)

   (b) Exactly 5 digits, optionally followed by a hyphen and four digits. (4 marks)

   (c) A session is a sequence of related interactions between a client and a server. HTTP is a stateless protocol so there is no way of storing current state of a session, e.g., contents of shopping cart. Three solutions are to use URL rewriting, hidden form fields or cookies. URL rewriting typically tracks a session identifier in the URLs of links in the pages shown to a user. Using hidden form fields, the session identifier, or the whole session state, is placed in hidden form fields. A cookie is a small data item stored on a client but managed by a server. Users can thus be tracked independently of sessions. Cookies are controlled by HTTP headers. A server can set a cookie, e.g. session identifier, as part of an HTTP response. This is stored by the client and returned as a header in subsequent requests to the same server. (15 marks)

6. (a) Examples include SVG, MathML, RDF, SMIL, WML and many others. (2 marks)

   (b)
```
<xsl:template match="/">
  <html>
    <body>
      <xsl:apply-templates select="//title"/>
    </body>
  </html>
</xsl:template>

<xsl:template match="channel/title">
    <h1><xsl:value-of select="."/></h1>
</xsl:template>

<xsl:template match="image/title"/>

<xsl:template match="item/title">
  <p>
    <xsl:value-of select="."/>:
    <xsl:value-of select="../description"/>
  </p>
</xsl:template>
```
(14 marks)

   (c) Content negotiation is used between clients and servers to determine which representation of a resource to use. This is useful because clients may have varying capabilities, users may different preferences and resources may exist in different variants according to language, quality or encoding. (9 marks)