1. (a) A codeword is a fixed size unit of data comprising data and check bits. The Hamming distance between two codewords is the number of bit positions in which they differ. If two codewords are distance $d$ apart, it will require $d$ single-bit errors to convert one to the other. (7 marks)

(b) i. $<!ATTLIST article$
   
   type (journal | conference) #IMPLIED$>$

   (3 marks)

ii. $//article[@type='conference'][count(author)>3]$

   (4 marks)

iii. According to the DTD, every author must have a last-name, so the expression article[author/first-name] will match authors who have both a first-name and a last-name. (3 marks)

(c) XSL can process elements more than once, it can reorder the input, and it can output arbitrary text or elements. (3 marks)

2. (a) The document must refer to a DTD and some element or attribute in the document must violate the constraints placed on it in the DTD. (3 marks)

(b) Request headers include:

- Host: the domain name (and port) of the server; required in every request; allows server to differentiate requests for multiple hosts with same IP address
- User-Agent: information about the client program (type, version)
- Accept: formats acceptable to the client, given using MIME types

Response headers include:

- Server: name and version of server
- Content-Type: the (MIME) media type of the resource being returned
- Content-Length: size of message body in bytes
- Last-Modified: date and time when entity was last modified

(5 marks)

(c) There would need to be an input textbox in an HTML form. This would need an onkeyup attribute which would call a Javascript function on the page. This function would asynchronously send the string typed by the user so far to a program running on a server. This program would find matches for the string (from a database, e.g.) and would send these back to the client, which, using a callback function, would display these suggested matches below the textbox. The use could then select from among these. (5 marks)
(d) TCP uses a window mechanism to control the flow of data. When a connection is established, each end of the connection allocates a buffer to hold incoming data, and sends the size of the buffer to the other end. As data arrives, the receiver sends acknowledgements together with the amount of buffer space available, called a window advertisement. If the receiving application can read data as quickly as it arrives, the receiver will send a positive window advertisement with each acknowledgement. However, if the sender is faster than the receiver, e.g. has a faster CPU, incoming data will eventually fill the receiver’s buffer, causing the receiver to advertise a zero window. A sender that receives a zero window advertisement must stop sending until it receives a positive window advertisement. (7 marks)

3. (a) The namespaces are as follows:
- **product**: http://www.xxx.com
- **manufacturer**: http://www.xxx.com
- **name**: http://www.yyy.com
- **description**: http://www.zzz.com
- **price**: http://www.xxx.com

(b) When applied to a particular node in a DOM tree, `getElementByTagName` finds all descendant nodes whose name matches that given as its argument. Alternatively, we could use `evaluate` (Firefox) or `selectNodes` (IE) with an XPath expression as its argument. We could also 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. (5 marks)

(c) The completed table is given below.

<table>
<thead>
<tr>
<th>Event</th>
<th>Segment 1</th>
<th>Segment 2</th>
<th>Frame sent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bridge boots</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>A sends to B</td>
<td>A</td>
<td>–</td>
<td>both segments</td>
</tr>
<tr>
<td>B sends to A</td>
<td>A,B</td>
<td>–</td>
<td>segment 1 only</td>
</tr>
<tr>
<td>X broadcasts</td>
<td>A,B</td>
<td>X</td>
<td>both segments</td>
</tr>
<tr>
<td>Y sends to A</td>
<td>A,B</td>
<td>X,Y</td>
<td>both segments</td>
</tr>
<tr>
<td>Y sends to X</td>
<td>A,B</td>
<td>X,Y</td>
<td>segment 2 only</td>
</tr>
<tr>
<td>X sends to Z</td>
<td>A,B</td>
<td>X,Y</td>
<td>both segments</td>
</tr>
<tr>
<td>Z sends to X</td>
<td>A,B</td>
<td>X,Y,Z</td>
<td>segment 2 only</td>
</tr>
</tbody>
</table>

(d) The mask is 20 bits long, so the network part is 20 bits. The remaining 12 bits are for the host, so 4094 (4096 - 2) host addresses exist. (3 marks)
4. (a) URI is the more general term which includes any URL or URN. URLs and URNs do not overlap, since every URN starts with urn:. (4 marks)

(b) The DNS servers provide a mapping between domain names and IP addresses, such that when a request for a Web page is received, the browser can look up in the DNS server the IP address corresponding to the domain name of the requested page, and then download the requested page from that IP address. If all the DNS servers in the world were to crash at the same time, one would not be able to map between domain names and IP addresses. Therefore, the only way to access Web pages would be by using the IP address of the host server instead of the domain name. Since most of us do not know the IP addresses of the servers we access, this type of situation would make use of the Internet extremely inefficient, if not virtually impossible for most users. (5 marks)

(c) The page that displays the form looks like this:

```
<html>
<head> <title> Multiplier </title> </head>
<body>
<form action="action.php" method="post">
<p> Please enter first number: <input type="text" name="first"> </p>
<p> Please enter second number: <input type="text" name="second"> </p>
<input type="submit">
</form>
</body>
</html>
```

The PHP script that does the processing looks like this:

```
<html>
<head> <title> Multiplication </title> </head>
<body>
The product is <?PHP echo $first * $second; ?>
</body>
</html>
```

(7 marks)

(d) After stuffing, the output will be
A B esc esc C esc esc esc eot esc soh D (4 marks)

5. (a) SMTP: simple mail transfer protocol; POP3: post office protocol; IMAP4: internet mail access protocol. SMTP is used for sending email from a client to a server, and from server to server. POP and IMAP are used when a client wants to retrieve emails. (5 marks)

(b) i. The DTD is as follows. (7 marks)

```
<!ELEMENT recording (artist, track+)> 
<!ELEMENT artist (#PCDATA)> 
<!ELEMENT track (title, length?)> 
<!ELEMENT title (#PCDATA)> 
<!ELEMENT length (#PCDATA)>
```

©Birkbeck College 2012
COIY063H7
(Page 3 of 5)
ii. The XSLT rule is as follows. (8 marks)

```xml
<xsl:template match="recording">
  <table>
    <xsl:for-each select="track">
      <tr>
        <td><xsl:number/></td>
        <td><xsl:value-of select="title"/></td>
        <td><xsl:value-of select="../artist"/></td>
      </tr>
    </xsl:for-each>
  </table>
</xsl:template>
```

6. (a) General entities can be used in either DTDs or documents and are referenced using an initial &. Parameter entities can only be used in DTDs and are referenced using an initial %; both references are terminated by a ;. (5 marks)

(b) The first returns all sections if there is an image anywhere in the document, whereas the second returns all sections that have an image element as a descendant. (3 marks)

(c) Multiplexing refers to the combination of information streams from multiple sources for transmission over a shared medium. Frequency Division Multiplexing (FDM) uses a separate channel (i.e. carrier frequency) for each stream. Time Division Multiplexing (TDM) simply means transmitting an item from one source, then from another, and so on. (6 marks)

(d) The signal is a square wave with two values, high (H) and low (L). The pattern is LHLHLHLHLHLHLHLHLHLHLHLHHLLHHL. (3 marks)

(e) The official RFC 1738 way to do this is http://dns-name:port/file. (3 marks)
7. (a) The output will consist of

    <html><h1>Some conference</h1></html>

but this will be followed by the textual contents of all other elements in the document. The reason is that the stylesheet processor does not find any matching rule for the root node, and so automatically tries to find matching rules for its child, the proceedings element. Again, it finds no matching rule, so tries to finds rules for name and contents. It finds a rule for name and applies it. But for contents it carries on looking for rules for its children until it reaches text nodes, when it applies its default rule to output the corresponding text. (8 marks)

(b) Event attributes in HTML allow scripts to be executed in response to (user) events on a web page. Some events are onClick (triggered when the left mouse button is clicked), onLoad (triggered when the page is loaded), and onKeyUp (triggered when a key on the keyboard is released). (4 marks)

(c) The packets are routed as follows: (i) Interface 1 (ii) Interface 0 (iii) Router 2 (iv) Router 1 (v) Router 2 (5 marks)

(d) No. IP packets contain IP addresses, which specify a destination machine. Once such a packet arrived, how would the network handler know which process to give it to? UDP packets contain a destination port. This information is essential so they can be delivered to the correct process. (3 marks)