Section A: Networks

(Answer at least TWO questions from this section.)

1. (a) TCP/IP protocols are organised into five conceptual layers.
   i. Layer 1: Physical - this layer corresponds to basic network hardware.
ii. Layer 2: Data-link or Network Interface - these protocols specify how to transmit data in frames over a network.

iii. Layer 3: Network - these protocols specify the format of packets sent across an internet and the mechanisms for forwarding packets through routers.

iv. Layer 4: Transport - these protocols specify how to ensure reliable transfer of data.

v. Layer 5: Application - each protocol specifies how one application uses an internet.

(b)

i. Voltage between +5 and +15 volts represents a 0 bit (SPACE)

ii. Voltage between -5 and -15 volts represents a 1 bit (MARK)

iii. Least-significant bit transmitted first

iv. Bits are recognised by timing

(c)

Although the two circuits carry data independently, it is possible for them to share a single ground wire.

(d) i. Amplitude shift modulation
ii. Phase shift modulation

(e) A technology that uses a small part of the electromagnetic spectrum is known as baseband technology, e.g. POTS having a bandwidth of 4,000 Hz. This limits data modems to 56 kbps. Higher throughput requires a larger bandwidth and the term broadband technology is used to characterise such technologies.

Broadband technology can be offered by telephone companies and cable companies by increasing the available bandwidth in the local loop (the wires between the exchange and the customer’s premises).

(f) The most popular offering is Asymmetric Digital Subscriber Line (ADSL) where the 1.1 MHz spectrum on the local loop is divided into 256 independent channels of 4,312.5 Hz each, as illustrated.
2. (a) i. The SOURCE PORT field identifies the port number used by the sender of the segment.

ii. The DESTINATION PORT identifies the port number of the process receiving this segment.

iii. The SEQUENCE NUMBER field identifies the first octet of the outgoing data. The receiver uses this to re-order segments arriving out of order and to compute an acknowledgement number.

iv. The ACKNOWLEDGEMENT NUMBER field identifies the sequence number of the incoming data that is expected next.

v. The WINDOW field identifies how much buffer space is available for incoming data.

(5 marks)

(b) TCP copes with the loss of packets using a technique called retransmission. When TCP data arrives, an acknowledgement is sent back to the sender. When TCP data is sent, a timer is started. If the timer expires before an acknowledgement arrives, TCP retransmits the data. This is illustrated below.

The host on the left is sending data, and the host on the right is receiving data.
TCP must be ready to retransmit any packet that is lost on either connection. (3 marks)

c) The three-way handshake is also used to terminate a connection, as illustrated below.

In this example, host 1 terminates the connection by transmitting a segment with the FIN flag set containing optional data. Host 2 acknowledges this (the FIN flag also consumes one byte of sequence space) and sets its own FIN flag. The third and last segment contains host 1’s acknowledgement of host 2’s FIN flag. (3 marks)

d) i. 21 FTP File transfer
    ii. 23 Telnet Remote login
    iii. 25 SMTP E-mail
    iv. 69 TFTP Trivial file transfer protocol
    v. 79 Finger Lookup information about a user
    vi. 7 HTTP World wide web
    vii. 110 POP-3 Remote e-mail access
    viii. 119 NNTP Usenet news

    (2 marks)

(e) UDP provides an end-to-end service that allows an application to send and receive individual messages, each of which travels in a separate datagram. An application can choose to restrict communication to one other application or communicate with multiple applications.

TCP provides a completely reliable (no data duplication or loss), connection-oriented, full-duplex stream transport service that allows two applications to form a connection, send data in either direction, and then terminate the connection. Each TCP connection is started reliably and terminated gracefully, with all data being delivered before the termination occurs. (5 marks)

(f) i. 7 Echo Server
    ii. 53 Domain Name System (DNS)
    iii. 69 Trivial File Transfer Protocol (TFTP)
    iv. 137 NetBIOS Name Service (WINS?)
v. 161 Simple Network Management Protocol (SNMP)
vi. Maybe a mention of RPC (not really :-)

(g) The computers at one site use local IP addresses which are private (or non-routable), and which are not valid global addresses. The three reserved ranges are 10.0.0.0/8, 172.16.0.0/12, and 192.168.0.0/16. Any datagram destined for the Internet has its private IP address replaced by the global IP address; any inbound datagram undergoes the reverse process.

<table>
<thead>
<tr>
<th>External socket</th>
<th>Internal socket</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCP 212.159.43.118:4001</td>
<td>192.168.0.1:2001 to 201.99.66.33</td>
</tr>
<tr>
<td>TCP 212.159.43.118:4002</td>
<td>192.168.0.1:2002 to 202.99.66.33</td>
</tr>
<tr>
<td>TCP 212.159.43.118:4003</td>
<td>192.168.0.2:2001 to 201.99.66.33</td>
</tr>
</tbody>
</table>

Similarly for UDP.

3. (a) • It is an application program that becomes a client temporarily when remote access is needed, but performs other computation locally.
  • It is invoked by a user and executes for one session.
  • It runs locally on the user’s computer.
  • It actively initiates contact with a server (CONNECT primitive).
  • It can access multiple services as needed.

(b) • It is a special-purpose program dedicated to providing one service.
  • It is invoked automatically when a system boots, and continues to execute through many sessions.
  • It runs on a shared computer.
  • It waits passively for contact from arbitrary remote clients (LISTEN primitive).
  • It accepts contact from arbitrary clients, but offers a single service.

(c) The translation of a domain name into an IP address is called name resolution and the software to perform the translation is called a resolver. Each resolver is configured with the address of a local domain name server. It sends a DNS request message and waits for the server to send a DNS reply. Most resolvers are configured to use UDP.

When an incoming request specifies a name for which that server is an authority, the server replies directly from its own database. Otherwise, the server temporarily becomes a client of another name server (usually a root server). When the second server returns a result, the original server sends a copy back to the resolver. It caches the result because it is highly likely that it will receive the same request within a short period of time.

Caching reduces the load on the root servers. In addition, the root servers are replicated around the world (there are currently 13) and a DNS server uses whichever root server is most responsive. In practice, the geographically closest root server usually responds first.
<table>
<thead>
<tr>
<th>Domain Name</th>
<th>Assigned To</th>
</tr>
</thead>
<tbody>
<tr>
<td>com</td>
<td>Commercial organization</td>
</tr>
<tr>
<td>edu</td>
<td>Educational institution</td>
</tr>
<tr>
<td>gov</td>
<td>Government organization</td>
</tr>
<tr>
<td>mil</td>
<td>Military group</td>
</tr>
<tr>
<td>net</td>
<td>Major network support center</td>
</tr>
<tr>
<td>org</td>
<td>Organization other than those above</td>
</tr>
<tr>
<td>arpa</td>
<td>Temporary ARPA domain (still used)</td>
</tr>
<tr>
<td>int</td>
<td>International organization</td>
</tr>
<tr>
<td>country code</td>
<td>A country</td>
</tr>
</tbody>
</table>

(d)

(e)  
   i. Static. The author determines the content at the time the document is written. Each request for the document results in exactly the same response.
   ii. Dynamic. When a request arrives, the web server runs an application program that creates the document. Each request for the document can result in different responses.
   iii. Active. When a request arrives, the web server returns a copy of a program that the browser runs locally. The contents can change as long as the user continues to run the program.

(f) State information is passed to a browser in the form of a cookie. Each cookie consists of a name/value pair. The name contains the name of the web site and the value is a small string that the browser stores. When it contacts the web site again, the browser inserts the cookie in the request.

(g) Any computer language can be used to write a CGI script, although the ability to read standard input and environment variables helps. Given the question, PHP might be suitable as it can run MySQL queries as well as generate HTML documents.

(h)  
   i. GET requests an item from the server. The server returns a header containing status information, followed by a blank line, followed by the item.
   ii. HEAD requests status information about an item. The server returns just the status information.
   iii. POST sends data to the server. The server appends the data to a specified item.
   iv. PUT sends data to the server. The server uses the data to replace a specified item.

©Birkbeck College 2010
COIY063H7
(Page 7 of 10)
Section B: Web technologies

(Answer at least TWO and at most THREE questions from this section.)

4. (a) A root element is required, end tags are required, elements must be properly nested, case is significant (so start and end tags must match), attribute values must be enclosed in quotes, and entity references, other than those to built-in entities, must be declared. (6 marks)

(b) Parameters are used in XSLT to pass external values into a stylesheet processing application. Parameters are declared by using an xsl:param element which appears as a child of the xsl:stylesheet element. The name attribute is used to give a name to the parameter. Values of parameters are referenced within the stylesheet by prefixing the name of the parameter with a dollar sign. Using parameters means that one stylesheet can be used where many would have been necessary without parameters. (6 marks)

(c) When async is set to false, the document load is synchronous which means that the Javascript interpreter waits for the document to load before continuing execution. When async is set to true, the document load is asynchronous which means that the interpreter does not wait for the document to load before continuing execution. As a result callback functions must be written, so that the application can react appropriately once the document has loaded. (5 marks)

(d) Ajax stands for Asynchronous Javascript and XML. Ajax uses Javascript and asynchronous calls back to a web server in order to provide responsive and dynamic web applications. (3 marks)

5. (a) A namespace is declared by including an attribute with name xmlns:prefix, where prefix is some user-defined string, in the start tag of an element in an XML document. The value of the attribute gives the name of the namespace. Elements are declared to be in the namespace by prefixing their names with prefix:. (4 marks)

(b) Line 1: XML declaration, declares file to conform to XML syntax, states version of XML used, character encoding used and that file depends on other files. Line 2: Document type declaration, declares document to be of type book. Line 3: Entity declaration, declares entity named HTML-chapter and states that its value is found by retrieving file html.xml. Line 5: Entity reference, results in reference being replaced by entity value. (8 marks)

(c) The variable xmlDoc is instantiated to an empty DOM document object. The load method is used to load the contents of an XML file into the DOM object. Another variable (xslDoc) is used to load an XSL file as a DOM document. The processor variable is instantiated to the XSLTProcessor object. The importStylesheet method is used to associate the stylesheet in variable xslDoc with the XSLT processor. The transformToDoc method of the processor object is used to transform the XML document in xmlDoc to an HTML document in variable htmlDoc by applying the stylesheet associated with the processor. Finally, the saveXML method
of the DOM object in \texttt{htmlDoc} is used to produce serialised XML which is produced on the standard output. The effect of this with PHP is for the web server to return the output to the client. (8 marks)

6. (a) A mixed content model is one in which both elements and character data can appear. In a mixed content model declaration, the character data (\texttt{#PCDATA} must appear first, all elements must be part of an alternation (|), and the whole model have the closure operator (*) applied to it. The effect of this is that any sequence comprising any number of the elements interspersed with character data is allowed in the documents valid with respect to the DTD. (7 marks)

(b) i. \texttt{//li} (1 mark)
   ii. \texttt{//ol[.//ul]} (2 marks)
   iii. \texttt{//ol[count(li)=1]} (3 marks)
   iv. \texttt{//ol/li[not(ol) and not(u1)]]} (4 marks)

(c) SSI instructions are embedded in HTML comments. When the server retrieves a page containing SSI instructions, it interprets them and replaces the comments by the output of the instruction(s). (3 marks)

7. (a) \texttt{<!ELEMENT photo ((title)?, (person)*, image) >}
\texttt{<!ELEMENT image EMPTY >}
\texttt{<!ATTLIST image}
\texttt{   ref CDATA #REQUIRED >}

(b) \texttt{<xsl:template match="/album">}
\texttt{ <html>}
\texttt{   <body>}
\texttt{     <h1>My photos</h1>}
\texttt{     <table>}
\texttt{       <tr>
\texttt{         <xsl:for-each select="photo">}
\texttt{           <th><xsl:value-of select="title"/></th>
\texttt{         </xsl:for-each>}
\texttt{       </tr>}
\texttt{       <tr>
\texttt{         <xsl:for-each select="photo">}
\texttt{           <td><img src="{image/@ref}"/></td>
\texttt{         </xsl:for-each>}
\texttt{       </tr>}
\texttt{       <tr>
\texttt{         <xsl:for-each select="photo">}
\texttt{           <td align="center">}
\texttt{             <xsl:for-each select="person">}
\texttt{               <xsl:value-of select="."/>
\texttt{             </xsl:for-each>}
\texttt{           </td>}
\texttt{         </xsl:for-each>}
\texttt{   </table>}
\texttt{   </html>}
\texttt{ </xsl:template>}

©Birkbeck College 2010
COIY063H7
(Page 9 of 10)
</td>
</xsl:for-each>
</tr>
</table>
</body>
</html>
</xsl:template>