WEB ENGINEERING

Monday 21st March 2016 - Morning
Answer any THREE questions out of FIVE. All questions carry equal marks.
Time: THREE hours.

Answer any Section A questions you attempt in Answer Book A
Answer any Section B questions you attempt in Answer Book B

The marks given in brackets are indicative of the weight given to each part of the question.

Calculators are NOT allowed in this examination.

The figures referenced in this question paper are provided in a separate booklet.

Section A
Answer Section A questions in Answer Book A

A1. Web developers face a number of technical choices when developing a website.

a) A solution stack is a complete set of software components required to run an application.
   i) LAMP is a well-known solution stack for dynamic web sites. Give the name of every component in this stack, and explain their role within the stack.
      (4 marks)
   ii) Give an example of an alternative to LAMP
       (1 mark)

b) Server-side scripting can be performed using a variety of programming languages.
   i) Give the names of FOUR different server-side scripting languages.
      (2 marks)
   ii) Give THREE distinct factors which may lead a web developer to choose (or avoid) a particular server-side scripting language when embarking on a brand new project. Clearly explain the impact (risk/cost or benefit) of each factor on the project.
      (3 marks)
iii) Are there any different factors that must be considered when choosing a client-side scripting language? Justify your answer with a clear explanation.

(2 marks)

c) When developing complex web applications, it is often advised to reuse existing code (and supporting tools). These additional software components are typically distributed as libraries or structured as frameworks. An alternative to using these well-established solutions would be writing all the code yourself.

i) Give the name of a well-known web application framework (indicate the language it is written in)

(1 mark)

ii) Explain, with appropriate justification, when it is preferable to reuse external code, and when it is preferable to write your own.

Give at least TWO arguments in favour of each approach.

(4 marks)

d) Your company intends to launch a database driven website. A senior manager has suggested the company should buy its own server and host the website itself, using its existing broadband connection. Write a short presentation (no more than 10 slides) analysing this proposal.

The presentation should:
- consider one or more alternatives
- compare and contrast available options, considering a range of business and technical factors
- conclude with a recommendation

(8 marks)
A2. Since August 2014, several United States-based Healthcare providers and Health Insurance companies have had serious data breaches, which may affect in total almost 100 million users. Please read the following extract from a Website article on the topic and then answer the questions below:

Health Care Info Used to Hack Phone NFC
A survey done by tax advisory firm KPMG reports that 81 percent of health care executives say their organization has been compromised by a data breach in the last two years. And the effects of those breaches can be seen reaching into pockets, wallets and phones — all at the same time.

After polling 223 chief information officers, chief technology officers, chief security officers and chief compliance officers at health care providers and health plans, KPMG “found the number of attacks increasing, with 13 percent saying they are targeted by external hack attempts about once a day and another 12 percent seeing about two or more attacks per week. More concerning, 16 percent of healthcare organizations said they cannot detect in real-time if their systems are compromised,” as stated in a press release.

Consumer Affairs reports just a few of the recent attacks on the vulnerability of health care organizations. In August of 2014, Community Health Systems admitted that Chinese hackers had breached their network and made off with data on more than 4.5 million patients. The following February, the Anthem health-insurance network admitted that hackers had stolen up to 80 million medical records dating back to 2004.

In March, Premera Blue Cross admitted to a breach compromising 11 million medical and financial records dating back to 2002. CareFirst Blue Cross/Blue Shield admitted to a hacking in May. In mid-July, the UCLA Health System admitted that 4.5 million patient records were at risk from a hacking UCLA had discovered two months earlier, Consumer Affairs reports.

Malware, software designed to disrupt or gain access to private computer systems, is the most frequently reported line of attack on health care organizations during the past 12 to 24 months, according to 65 percent respondents in KPMG’s survey. Botnet attacks, where computers are hijacked to issue spam or attack other systems, and “internal” attack vectors, such as employees compromising security, were cited by 26 percent of respondents.

According to the survey, the areas with the greatest vulnerabilities within an organization include external attackers (65 percent), sharing data with third parties (48 percent), employee breaches (35 percent), wireless computing (35 percent) and inadequate firewalls (27 percent).


a)  
   i)  For each of these three attack vectors on the business listed in the article, outline a suitable countermeasure that could be used to minimise risk or impact of:
      - Malware within the corporate network;
      - Denial of service attack by botnets; and
      - Internal attacks from employees.

      (3 marks)

   ii) State and explain THREE wired-network security risks that could result in disclosure of user data. (N.B.: when answering this question, you should not list risks that are solely client- or server-based).

      (6 marks)
iii) Wireless networking (e.g. 802.11n) is often used for convenience. State and explain ONE additional security risk introduced by the use of wireless networks, and provide a suitable countermeasure.

(2 marks)

iv) Customers may wish to access services using mobile devices (e.g. mobile phones, tablets), where there is a real risk of losing the device. Provide TWO countermeasures that could be used to minimise risk or impact of a lost device.

(2 marks)

b) i) One approach to reducing risk after a serious compromise is to disable all user accounts. With a suitable example, describe ONE positive consequence of this action and, ONE negative consequence of this action. (N.B.: this may be considered from either the corporate or user perspective).

(2 marks)

ii) After disabling all user accounts, the company might ask the user to re-activate them after proving their identity.

Explain why each of the following three normal methods of account verification might not be safe:

a. Sending an activation code to the user’s email address.

b. Writing a letter to the user’s home address asking them to call a telephone number to get a new password.

c. Asking the user questions about their family (e.g. ages of siblings) over instant messaging.

(3 marks)

iii) Taking into account the weaknesses of the methods discussed in part c) ii), propose (with appropriate justification) a robust method of verifying users’ identities that offers a good balance between user account safety and cost effectiveness.

(3 marks)

c) The article reports that one in six providers cannot detect in real time if their systems are compromised.

In no more than 300 words, explain what benefits and drawbacks real-time monitoring may offer to the management of security risks and – with appropriate justification – recommend whether it is worthwhile for Healthcare Providers to use real-time monitoring.

(4 marks)
Section B
Answer Section B questions in Answer Book B

B3.  

a) Data can be stored in child elements or in attributes as demonstrated by the example in Figure 3.1. Briefly state 5 factors to be considered when using attributes instead of child elements.

(5 marks)

b) The XML document in Figure 3.2 contains a number of errors when validated against the DTD (document data type) in Figure 3.3. Identify all the errors, and provide a solution for each.

(5 marks)

c) An XML schema can be written as a DTD or as XSD (xml schema document). State the benefits and drawbacks of each approach. Convert the DTD in Figure 3.3 to an XML schema.

(5 marks)

d) Referring to XML document in Figure 3.4, write a DTD which enforces the following constraints on cbcatalogue:

- Sequence of elements is as shown in the XML document (i.e. cbcatalogue is a container of subject elements).
- subject element must be present zero or more times.
- Each subject element contains a subject_title and zero or more books.
- The two attributes in book are mandatory.
- A cover image for each book that captures attributes of height, width and location of image
- book_url is optional.
- book_url must have a page attribute.
- For each book, exactly one book_title and one Publisher element must be present
- For each book, there must be one or more authors.
- An author has a name and optionally a short biography and an email.
- A name consists of both a first_name and a last_name.
- For each book, there must be one or more review.

(10 marks)
B4. a) 

i) Briefly explain how XPath expressions can manipulate an XML document.

(2 marks)

ii) With reference to the XML file shown in Figure 4.1, write an XPath expression to show all details of all products supplied by Coste.

(3 marks)

b) The XML document in Figure 4.2 (list of journals and articles) needs to be displayed as a web page as shown in Figure 4.3. Using the HTML template provided in Figure 4.4, in your answer book provide the missing code. Use the section marked <!--TO BE COMPLETED --> to accomplish this. You may assume that the journal.css file exists and no style code needs to be written.

(5 marks)

c) Modify the HTML template further so that the XML file is rendered as shown in Figure 4.5. The journal title and the publisher information are to be active links to the respective websites. In your answer book write the code to do this.

(5 marks)

d) Modify the HTML template further so that the XML file is rendered as shown in Figure 4.6. The editor email address is an active email link. In your answer book write the code to do this.

(5 marks)

e) The last stage in rendering the XML file is to display the details of the articles within each journal as shown in Figure 4.7. In your answer book write the code to do this.

(5 marks)
B5.  

a) Define the term API and state the purpose it serves.  

(4 marks)

b) Companies such as eBay™ and Google™ provide public APIs.  

State FOUR benefits of APIs.  

(3 marks)

c) Identify the stages and issues when creating an API from a producer’s viewpoint.  

(4 marks)

d) Data can be returned as XML or JSON by an API call. As a developer, give two distinct examples where each of these is most suitable, based on the type of device in use during the API call.  

(4 marks)

e) As a Web Engineer, identify at least FIVE major challenges in the design and development of an image-hosting site, where users can upload their images to a central server and the images can be retrieved via a web link or an API.  

Initially you are to focus on two aspects only:

- the ability to upload (write) an image to the server
- the ability to query for an image.

For each of the challenges identified, provide an outline solution including the technology to be used. Your answer should be in the form of a slide presentation to the client (5 slides with bullet points).  

(10 marks)

NOTE: Your answers should focus on the various technologies and challenges in using those technologies. No credit will be given for “text book” answers that follow the Software Development Life Cycle Model (SDLC).
option 1:

<booking reference="20120521" date="12/05/2012">
  <agent>M.Miles</agent>
  <client_name>Jane Doe</client_name>
  <client_email>jdoe@jmail.com</client_email>
  <advice>No changes allowed</advice>
</booking>

option 2:

<booking reference="20120521">
  <date>12/05/2012</date>
  <agent>M.Miles</agent>
  <client_name>Jane Doe</client_name>
  <client_email>jdoe@jmail.com</client_email>
  <advice>No changes allowed</advice>
</booking>

option 3:

<booking reference="20120521">
  <date>
    <day>12</day>
    <month>05</month>
    <year>2012</year>
  </date>
  <agent>M.Miles</agent>
  <client_name>
    <first_name>Jane</first_name>
    <last_name>Doe</last_name>
  </client_name>
  <client_email>jdoe@jmail.com</client_email>
  <advice>No changes allowed</advice>
</booking>
Figure 3.2

```xml
<?xml version="1.0"?>
<!DOCTYPE cv SYSTEM "cv.dtd">
<cv>
  <preface> Personal Data </preface>
  <preface> Name: Jane Doe </preface>
  <qualification> BSc </qualification>
  <experience> Charity work </experience>
  <qualification> MSc </qualification>
  <hobbies> running </hobbies>
  <experience> Charity work </experience>
  <qualification> MSc </qualification>
  <referees> Joe Blogs, BSc </referees>
  <!-- <hobbies> kite flying </hobbies> -->
</cv>
```

Figure 3.3

```xml
<!ELEMENT cv (preface, (qualification | experience)+, hobbies?, referee*)>
<!ELEMENT preface (#PCDATA)>
<!ELEMENT qualification (#PCDATA)>
<!ELEMENT experience (#PCDATA)>
<!ELEMENT hobbies (#PCDATA)>
<!ELEMENT referee (#PCDATA)>
```
Figure 3.4
<?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>
<ReadyMeals>
  <Food>
    <lunch supplier="Coste" id="1">
      <price>3.50</price>
      <quantity>20</quantity>
      <calories>1500</calories>
    </lunch>
    <dinner supplier="Daas" id="2">
      <price>4.50</price>
      <quantity>50</quantity>
      <calories>2000</calories>
    </dinner>
  </Food>
  <Drink>
    <juice supplier="Coste" id="3">
      <price>1.50</price>
      <quantity>60</quantity>
      <calories>180</calories>
    </juice>
  </Drink>
</ReadyMeals>

Figure 4.1
<?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>
<xml-stylesheet type="text/xsl" href="Journal.xsl"/>
<JournalCatalogue>
  <Journal ISBN="10104556" Issue_no="24">
    <Journal_title>Nature</Journal_title>
    <Journal_url page="http://www.nature.com"/>
    <Publisher>Wiley</Publisher>
    <Publisher_url page="http://www.wiley.com"/>
    <Editor>Bob Jones</Editor>
    <Editor_email>bjones@bcs.ac.uk</Editor_email>
    <Article>
      <Article_title>Go green</Article_title>
      <Abstract>Global warming and eco systems</Abstract>
      <Author>
        <Name>
          <Last_name>Simpson</Last_name>
          <First_name>Bart</First_name>
        </Name>
        <Author_email>simpsb@eco.com</Author_email>
        <Author_url page="www.eco.com"/>
        <Occupation>Academic</Occupation>
      </Author>
      <Author>
        <Name>
          <Last_name>Simpson</Last_name>
          <First_name>Lisa</First_name>
        </Name>
        <Author_email>simpsl@eco.com</Author_email>
        <Author_url page="www.eco.com"/>
        <Occupation>Scientist</Occupation>
      </Author>
    </Article>
    <Article>
      <Article_title>Forestation</Article_title>
      <Abstract>Flood Management</Abstract>
      <Author>
        <Name>
          <Last_name>Hobs</Last_name>
          <First_name>Harold</First_name>
        </Name>
        <Author_email>hhobs@eco.com</Author_email>
        <Author_url page="www.eco.com"/>
        <Occupation>Engineer</Occupation>
      </Author>
    </Article>
  </Journal>
  <Journal ISBN="10104557" Issue_no="7">
    <Journal_title>Wired</Journal_title>
    <Journal_url page="http://www.wired.com"/>
    <Publisher>New Riders</Publisher>
    <Publisher_url page="http://www.newriders.com"/>
    <Editor>Bob Jones</Editor>
    <Editor_email>bjones@bcs.ac.uk</Editor_email>
    <Article>
      <Article_title>Convergence</Article_title>
      <Abstract>One device as a container for all other devices</Abstract>
      <Author>
        <Name>
          <Last_name>Bates</Last_name>
          <First_name>Bill</First_name>
        </Name>
        <Author_email>bates@wires.com</Author_email>
        <Occupation>Academic</Occupation>
      </Author>
    </Article>
    <Article>
      <Article_title>USB3</Article_title>
      <Abstract>New standards</Abstract>
      <Author>
        <Name>
          <Last_name>Jones</Last_name>
          <First_name>Joe</First_name>
        </Name>
        <Author_email>jones@wires.com</Author_email>
        <Occupation>Engineer</Occupation>
      </Author>
    </Article>
  </Journal>
</JournalCatalogue>
Figure 4.3

```xml
<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0">
  <xsl:template match="/">
    <html>
      <head>
        <title>Catalogue of Journals</title>
      </head>
      <body>
        <h1>Catalogue of Journals - updated January 2016</h1>
        <table border="1">
          <tr bgcolor="#9acd32">
            <th align="left">Title</th>
            <th align="left">Journal Details</th>
            <th align="left">Editor</th>
            <th align="left">Publisher</th>
          </tr>
          <tr>!-- TO BE COMPLETED -->
          <tr>!--</tr>
        </table>
      </body>
    </html>
  </xsl:template>
</xsl:stylesheet>
```

Figure 4.4
## Catalogue of Journals - updated January 2016

<table>
<thead>
<tr>
<th>Title</th>
<th>Journal Details</th>
<th>Editor</th>
<th>Publisher</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><a href="mailto:bJones@bcs.ac.uk">bJones@bcs.ac.uk</a></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="mailto:bJones@bcs.ac.uk">bJones@bcs.ac.uk</a></td>
<td></td>
</tr>
</tbody>
</table>

### Article Titles

**Go green**
- Global warming and eco systems

**Forestation**
- Flood Management

**Convergence**
- One device as a container for all other devices

**USB3**
- New standards

1. Bart Simpson  
   Academic

2. Lisa Simpson  
   Scientist

1. Harold Hobs   
   Engineer

1. Bill Bates   
   Academic

1. Joe Jones     
   Engineer