

## Short Questions

1. The following fragment of HTML contains 6 errors (some appear more than once). Write the correct version such that it conforms to the XHTML standards

```
<HTML>
<head><title>My Web Page</title></head>
<body>
<h1>My Web Page</h2>
<img alt='My picture' src='mypic.jpg'>
<p>This page tells you a little about me & my interests!
<p>Here are my details:
<ul>
<li>Name: A. Student
<li>Occupation: student
<li>EMail: a.student@mail.cryst.bbk.ac.uk
</ul>
</body>
</HTML>
```

### Model answer:

```
<html>
<head><title>My Web Page</title></head>
<body>
<h1>My Web Page</h1>
<img alt='My picture' src='mypic.jpg' />
<p>This page tells you a little about me & my interests!</p>
<p>Here are my details:</p>
<ul>
<li>Name: A. Student</li>
<li>Occupation: student</li>
<li>EMail: a.student@mail.cryst.bbk.ac.uk</li>
</ul>
</body>
</html>
```

[1 mark for each error]

2. Explain the purpose each of the following fragments of Perl:

- a) 

```
use DBI;
$dbh = DBI->connect($dbsource, $username, $password);
($id, $ac, $sp, $seq) = $dbh->selectrow_array($sql);
```
- b) 

```
use CGI;
$cgi = new CGI;
print $cgi->header();
```
- c) 

```
use XML::DOM;
$parser = XML::DOM::Parser->new();
$doc = $parser->parsefile($file);
$name_tag = $doc->getElementsByTagName('name');
```

### Model answer

- a) Connects to a database specified by \$dbsource using the specified username and password; Performs the query specified in the variable \$sql – the selectrow\_array() function assumes that only one row will be returned by the query
- b) Creates a CGI object and prints the standard MIME header required by the web server to indicate the type of data (default HTML text)
- c) Creates an XML::DOM parser, parses an XML file and then extracts all the <name> tags and their contents into the \$name\_tag object.

[2 marks for each]

### **Long Question**

What is Remote Procedure Calling (RPC)? Outline the 4 main ways of achieving RPC giving examples of how they might be used and their pros and cons.

#### **Model answer**

RPC is accessing a program or subroutine on another computer as part of a program that you are running. You might want to do this: a) to distribute the load between computers, b) to access other people's methods, c) to access the latest data.

Four main ways of performing RPC:

- screen scraping
- simple cgi scripts
- custom code to work across networks
- standardized methods (e.g. CORBA, SOAP, XML-RPC)

RPC methods which work across the internet are often called “Web Services” (particularly SOAP and related protocols)

**[2 marks]**

**Screen scraping:** extracting content from a web page. **[1 mark]**

*Example:* any web server that provides a prediction (e.g. NNpredict for SS prediction) **[1 mark]**

*Disadvantages:* fragile – trying to interpret semantics from display-based markup.

Change of presentation breaks the screen scraper. **[1 mark]**

*Advantages:* easy to implement with Perl LWP module

'service provider' does not need to do anything other than provide their normal web-based form and CGI script. **[1 mark]**

**Simple CGI scripts (REST):** an extension of the screen scraping method, but relies on the service provider to provide a CGI script which is designed specifically for remote access. **[1 mark]**

*Advantages:* easy to implement with Perl LWP module.

Not fragile – data guaranteed to be returned in a simple format suitable for parsing and extracting semantic data.

Service provider just need to provide a modified CGI script. **[1 mark]**

*Disadvantages:* no standards for service or data description. **[1 mark]**

*Example:* NCBI 'Entrez programming utilities' **[1 mark]**

**Custom code:** generally implemented to distribute tasks around machines on a local network. **[1 mark]**

*Disadvantages:* complex **[1 mark]**

*Advantages:* fast – can be customized for specific purposes **[1 mark]**

*Examples:* low-level communication protocols such as HTTP, ssh, ftp, nfs, etc. **[1 mark]**

**Standardized methods:** high level protocols e.g. CORBA, XML-RPC and SOAP. **[1 mark]**

*Advantages* (of SOAP): information encoded in XML so is language independent. Transmission is over HTTP allowing access through firewalls. W3C standard with libraries to ease implementation in many languages. **[1 mark]**

*Disadvantages*: Slow **[1 mark]**

*Example*: Provide data servers and prediction services. **[1 mark]**