

## Development of Internet Applications—Solutions

June 2005

1. (a) The value `janet` for the `spouse` attribute of the first `person` does not occur as a `pno` value in the document. The `spouse` attribute value of the second `person` contains two tokens, whereas only one is permitted. The `role` attribute of the third `person` uses a value that is not allowed by the DTD. (6 marks)
- (b) `/household/person[id(@spouse)/name='Mary']/name` (5 marks)
- (c) 

```
<xsl:template match="/household">
  <household>
    <xsl:for-each select="person">
      <xsl:choose>
        <xsl:when test="@role='adult'">
          <adult>
            <xsl:value-of select="name"/>
          </adult>
        </xsl:when>
        <xsl:when test="@role='child'">
          <child>
            <xsl:value-of select="name"/>
          </child>
        </xsl:when>
        <xsl:otherwise>
          <person>
            <xsl:value-of select="name"/>
          </person>
        </xsl:otherwise>
      </xsl:choose>
    </xsl:for-each>
  </household>
</xsl:template>
```

 (14 marks)
2. (a) TCP provides a reliable service by acknowledging the safe receipt of packets, detecting missing/corrupted/duplicated packets, and providing a method to resend missing/corrupted packets. UDP provides a connectionless, unreliable service, so it is faster than TCP. UDP adds only checksum and process-to-process addressing to IP. UDP is used for DNS and NFS. (7 marks)
- (b) Elements permit nested (and repeating) substructures, and are always ordered, even if the order is not significant for the application. Attributes can have only simple values, no ordering is implied by the order in which they appear, while an attribute name cannot repeat for a given element. Attributes tend to be used to provide additional information about elements, e.g., units of measurement in a periodic table document. Their values can also be restricted using types other than those allowed for elements when using a DTD. (6 marks)
- (c) i. `section[.//image]`  
ii. `..[name()='chapter']`

- iii. `preceding-sibling::*[1] — ../*[1]` is not equivalent if the context node is the first child  
(6 marks)
- (d) A strong validator for HTTP caching changes whenever the resource changes, whereas a weak validator changes only when the meaning of a resource changes. Changes could be made to an XML file that only alter the whitespace in the file, resulting in the same canonical representation as before the change. A weak validator could use this to detect that no meaningful change had occurred to the file. (6 marks)
3. (a) One might need to transform or produce an XML schema document using XSL in which case there may be a need to refer to the `element` element of XML schema and use the `element` element of XSL to output an element. Namespaces allow an XSL processor to distinguish between the two element names, by using distinct prefixes such as `xsd` and `xsl`. (5 marks)
- (b) In URIs special characters are escaped by using `%` followed by the ASCII value of the character in hexadecimal. In XML special characters are escaped using references to built-in entities, using the notation `&name;`, where `name` is the name of the built-in entity. (6 marks)
- (c) XML is seen as a universal, open, readable representation for software integration and data exchange. It enables integrated e-commerce through programs exchanging information. Historically, the usage of EDI was limited to large corporations because of high costs and complexity. Legacy data is data generated and used by processes that don't use the organisation's current technology. (10 marks)
- (d) Advantages of processing on client: user interactions can be managed on the client, reduces load on server (e.g. doing transformations on client, fewer round trips to server). Advantages of processing on server: don't have to worry about making code cross-browser compatible, don't have to expose all data to clients. (4 marks)
4. (a) Possibilities include: HTTP (HyperText Transfer Protocol), ftp (file transfer protocol), Network News Transfer Protocol (NNTP), Network News Reading Protocol (NNRP), Simple Mail Transfer Protocol (SMTP) and Post Office Protocol (POP3). (4 marks)
- (b) i. 

```
<!ATTLIST book
    isbn      CDATA      #REQUIRED
```

 (3 marks)
- ii. `//book[count(author) > 2]` (4 marks)
- iii. The expression `book[author/first-name][author/last-name]` can first be rewritten as `book[author/first-name][author]` because the DTD requires that every `author` has a `last-name`. This can then be rewritten as `book[author/first-name]` since every `book` with an `author` that has a `first-name` must have an `author`. (4 marks)
- (c) The use of many services will require contracts and billing: no standards yet exist in these areas. There is no way to deal gracefully with service version control: when an interface changes, the client code breaks. SOAP lacks authentication, authorization and encryption features. Trust relationships are difficult to manage when the parties involved may be unknown. Protocols which support atomic transactions are still under development. The protocols are seen as too low-level and tightly coupled: just like remote procedure calls (RPCs). Is UDDI necessary? If so, who will provide registries? There is too much focus on how to invoke services (syntax) and not enough on what web services actually provide and why they are provided (semantics). (10 marks)

5. (a) At the time MIME was proposed, the email protocol SMTP used a 7-bit ASCII format. This was inadequate for non-English and non-textual data. MIME allows non-US-ASCII message bodies, an extensible set of different formats for non-textual bodies, multi-part message bodies and non-US-ASCII textual header information. (5 marks)
- (b) The fragment first sets `elem` to refer to the element in the current XML document identified by `id="target1"`. It then creates a new `li` element node referred to by `node`, followed by a text node with value `Hello` referred to by `text`. The fourth line appends the text node to the `li` node, while the last line appends the new `li` element to the element referred to by `elem`. (10 marks)
- (c) The three approaches are: (1) scripting or programmatic approaches: e.g. common gateway interface (CGI) programs and Java servlets; (2) template approaches: e.g. server-side include (SSI) pages and Cold Fusion; and (3) hybrid approaches: e.g. Java server pages (JSP) and active server pages (ASP). They all mix content and presentation in the same code. (10 marks)

6. (a) 

```
<xsd:element name="recording">
  <xsd:complexType>
    <xsd:sequence>
      <xsd:element name="artist" type="xsd:string"/>
      <xsd:element name="track" maxOccurs="unbounded">
        <xsd:complexType>
          <xsd:sequence>
            <xsd:element name="title" type="xsd:string"/>
            <xsd:element name="length" minOccurs="0" type="xsd:timeDuration"/>
          </xsd:sequence>
        </xsd:complexType>
      </xsd:element>
    </xsd:sequence>
  </xsd:complexType>
</xsd:element>
```

(10 marks)

- (b) 

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

(10 marks)

- (c) In XML schema one can declare locally-scoped element names which allows the track element to have different content models in different contexts. This cannot be done using DTDs since every element declaration in a DTD is global. (5 marks)