

COMP2405 Review

Pat Morin
COMP2405

Outline

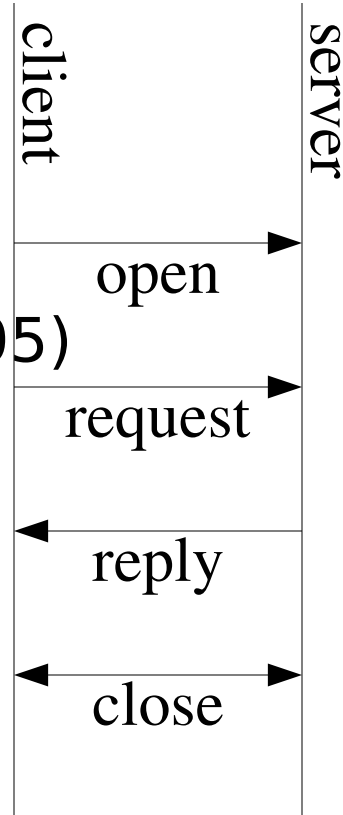
- Web technologies and types of applications
 - Web sites (HTTP, HTML, CSS, CRAP)
 - Server-side scripted (**CGI**, Perl, and PHP)
 - Client-side scripted (JavaScript, HTML DOM)
 - Both-side scripted (AJAX, XML, XMLHttpRequest)
- Other topics
 - History of the Internet
 - Cookies
 - HTTPS
 - JavaScript security

History of the Internet

- Internet timeline
 - 1970s until today
 - 1990s: The web
- The Internet != The Web
- Key players
 - Licklider, Roberts, Kleinrock, Fahlman, Morris, Berners-Lee

HTTP

- Basic HTTP transaction (4 steps)
- HTTP methods
 - GET, HEAD, PUT, POST
- HTTP response codes (200, 404, 405)
- HTTP Fields
 - Content-type
 - Content-length
 - Host



HTML: The Hypertext Markup Language

- History
 - Tim Berner-Lee and Robert Caillau at CERN (1989-1990)
 - Web launched in 1991
- Structure of an HTML Document
 - DOCTYPE, HEAD, BODY
- Logical versus Physical (Visual) Markup
 - Use HTML to describe logical structure of document
- Tags
 - Paragraphs, headings, forms, lists, tables, hyperlinks, targets, forms, form elements

CRAP: The Principles of Graphic Design

- Contrast
 - When two elements are not *exactly the same*, they should be *very different*
- Repetition
 - Styles and visual elements should be repeated across a page and across a web site
- Alignment
 - Our minds like things to be lined-up and evenly spaced
- Proximity
 - Proximity (closeness and distance) can be used to group related items together and separate unrelated items
- Be able to identify and fix graphic design

CSS: Cascading Style Sheets

- Advantages and disadvantages of CSS
 - Separating formatting from writing
- CSS Properties
 - Font and text properties, list properties
- CSS naming rules
 - class, id, and hierarchical naming
- CSS pseudo-elements
 - a:link, a:visited, a:hover, :before, :after and the content property
- CSS Units
 - Color, length

Layout with CSS

- The display property
 - block – P, H1, H2, UL, PRE, DIV
 - inline – A, EM, CODE, SPAN
- The CSS box model
 - margin, border, padding, content
- Layout properties
 - position – static, relative, or absolute
 - width, left, top, right, bottom
- Be able to recognize a layout from a CSS snippet
- Formatting lists

CGI: The Common Gateway Interface

- Specifies the interface between a web server and an application that serves web pages
- Use of environment variables and stdin
 - QUERY_STRING
 - Differences between GET and POST
- Output format
 - Content-length\n\ndata
- CGI and security
 - Client data can not be trusted
 - Be able to identify insecure code in CGI scripts

Perl: Practical Extraction and Report Language

- Literals
 - Strings – single quoted, double-quoted, and heredocs
 - Numbers
- Scalar variables - \$
- Arrays - @
- Hashes - %
- Contexts
 - array to scalar conversion

Perl (Cont'd)

- Subroutines
 - Defining a subroutine (with prototypes)
 - Calling a subroutine
- References
 - Taking a reference \
 - Dereferencing {}
 - Anonymous references {} (hash) and [] (array)
- Regular expressions
 - Operators * + ? {} [] ()
 - Special characters . ^ \$
 - Return values – scalar or array
 - Substitutions *s/pattern/replacement/*

JavaScript

- Data types
 - Numbers, strings, arrays, and objects
- Undefined is very well-defined
- Variables and constants
 - The var keyword and const keyword

JavaScript Functions

- Defined using the `function` keyword
- Can be anonymous
 - `var f = function () { ... }`
- Very flexible
 - Can have a variable number of arguments
 - accessed by the `arguments` variable
 - Can be called with more or fewer than the declared number of arguments

JavaScript Objects

- Object properties (instance variables) are accessed using the `.` or the `[]` operator
- Object syntax is very loose
 - Can add properties at any time simply by assigning to them
- Objects are created by calling the `new` operator on a constructor function
 - Constructor function modifies the object by accessing the `this` variable

JavaScript Advanced

- Getters and setters
 - Functions that look like properties
- Operators
 - delete, typeof, in, void, instanceof
- Class prototypes
- Using JavaScript to emulate Java
 - objects
 - instance variables
 - methods
 - private variables
 - inheritance
- Exception handling

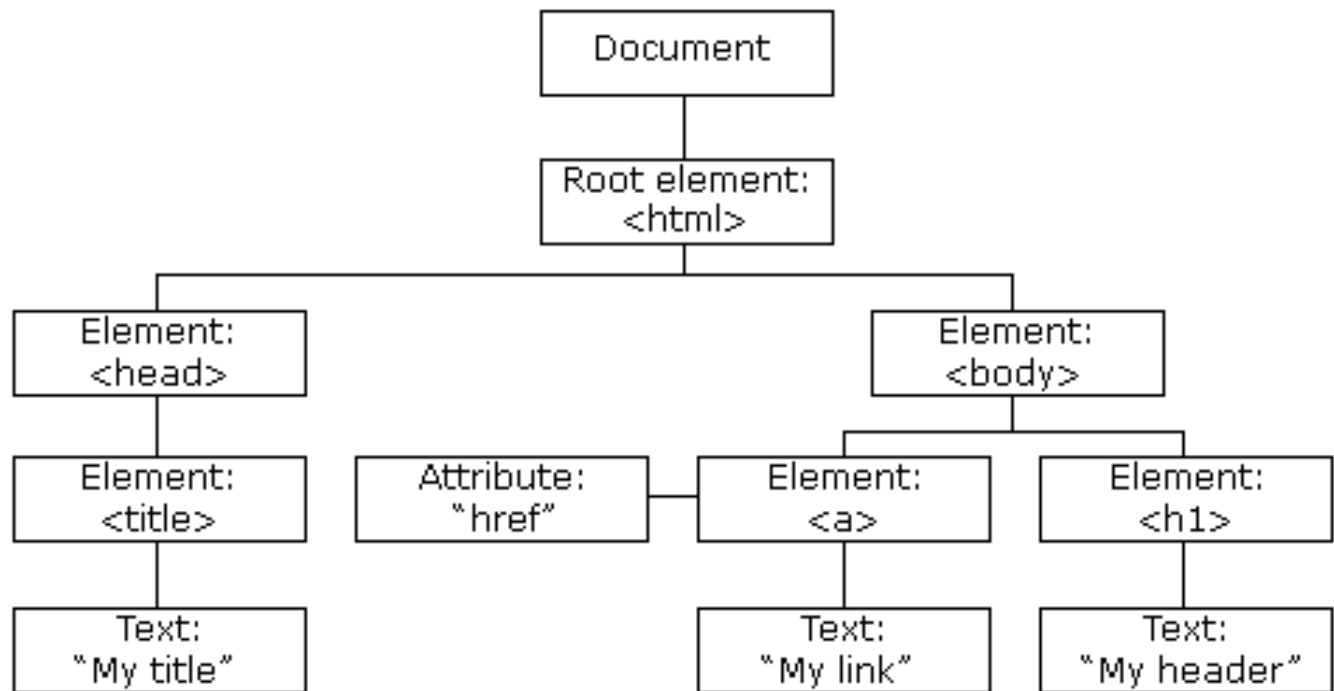
The HTML DOM

- The HTML Document Object Model specifies a mapping between parts of a document and objects

HTML DOM: The Document Object Model

- A document is represented by a *document tree*
 - This is the parse tree of the document
- Accessed by the global document object
 - `getElementById`, `getElementsByName`, `getElementsByTagName`
- A tree node `n` has children and can be modified
 - `n.childNodes`, `n.appendChild()`, `n.removeChild()`, `n.insertBefore()`
- Difference between text nodes and other nodes
 - `n.nodeValue` is either the text of the node or the tag

HTML DOM Example



Basic XML

- Basic XML is just like HTML but you can create your own tag names
- All documents have a root element
- All tags must be closed

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<note>
  <to>Tove</to>
  <from>Jani</from>
  <heading>Reminder</heading>
  <body>Don't forget me this weekend!</body>
</note>
```

Advanced XML

- XML is in a rapid state of development
- XML Document Type Definitions (DTD) specify the format of a valid XML document of a specific type
- XML can be formatted with CSS
- XHTML is an XML version of HTML
- In the future
 - XML Schemas will replace DTD
 - XSLT will replace formatting of XML with CSS

AJAX: Asynchronous JavaScript and XML

- AJAX combines server-side and client-side scripting
- Underpinned by the XMLHttpRequest
 - Client sends asynchronous request to server CGI script
 - Server replies with XML
- AJAX can be used to implement applications that interact with (save, store, retrieve, query) data on the server
- AJAX transactions take place without a page reload

AJAX Template

```
function sendRequest() {
    var xmlhttp = GetXmlHttpRequest();
    if (!xmlhttp) {
        return false;
    }
    xmlhttp.onreadystatechange = function() {
        if (xmlhttp.readyState == 4) {
            var xmlDoc =
                xmlhttp.responseXML.documentElement;
        }
    }
    var requestURI = xmlURI;
    xmlhttp.open("GET", requestURI, true);
    xmlhttp.send(null);
}
```

AJAX Summary

- AJAX combines everything
 - Part of an **HTML** page
 - Requires a **CGI** script on the server (can be written in **Perl**)
 - Requires **JavaScript** on the web page (client)
 - Sends an **HTTP** request and receives **XML** as the response text
 - The **XML** becomes a **DOM** document
- And it can be used to make very cool web applications

Other Topics

- Cookies
 - A small piece of data stored by the server on the client
 - Main application: keeping track of a user over several pages or visits (sessions)
- File upload
 - Done using an HTTP POST method
- HTTPS
 - An encapsulation of HTTP inside a cryptographically secure protocol
 - Uses public key encryption and symmetric encryption
 - Relies on the proper use of *certificates*
 - Prevents eavesdropping and man-in-the-middle attacks

JavaScript Security

- The same origin policy
 - JavaScript (should) only have access to information originating from the same origin as the current document
- JavaScript exploits
 - Popups
 - Chromeless windows
 - Cross-site scripting (also requires poorly-written CGI script)
 - Javascript port scanning

The Law

- Intellectual property
 - Trade secrets
 - Copyright
 - Trademarks
 - Patents
 - Industrial designs
- Privacy law
 - PIPEDA (federal)
 - FIPPA (provincial)
 - A web application may be under several privacy laws

Summary

- Exam Breakdown
 - 40% Pre mid-term material (HTML, CSS, CRAP, CGI, Perl)
 - 60% Post mid-term (Perl, JavaScript, AJAX, XML, DOM, Other)
- Same format as mid-term, but longer
- Study advice (beyond knowing the material)
 - Review these and all class notes just before the exam

What To Leave With

- What you should take out of this course
 - You now know how even the most sophisticated web applications work, from top to bottom
 - You now know something about security, vulnerabilities, and hacking
 - You will see regular expressions again