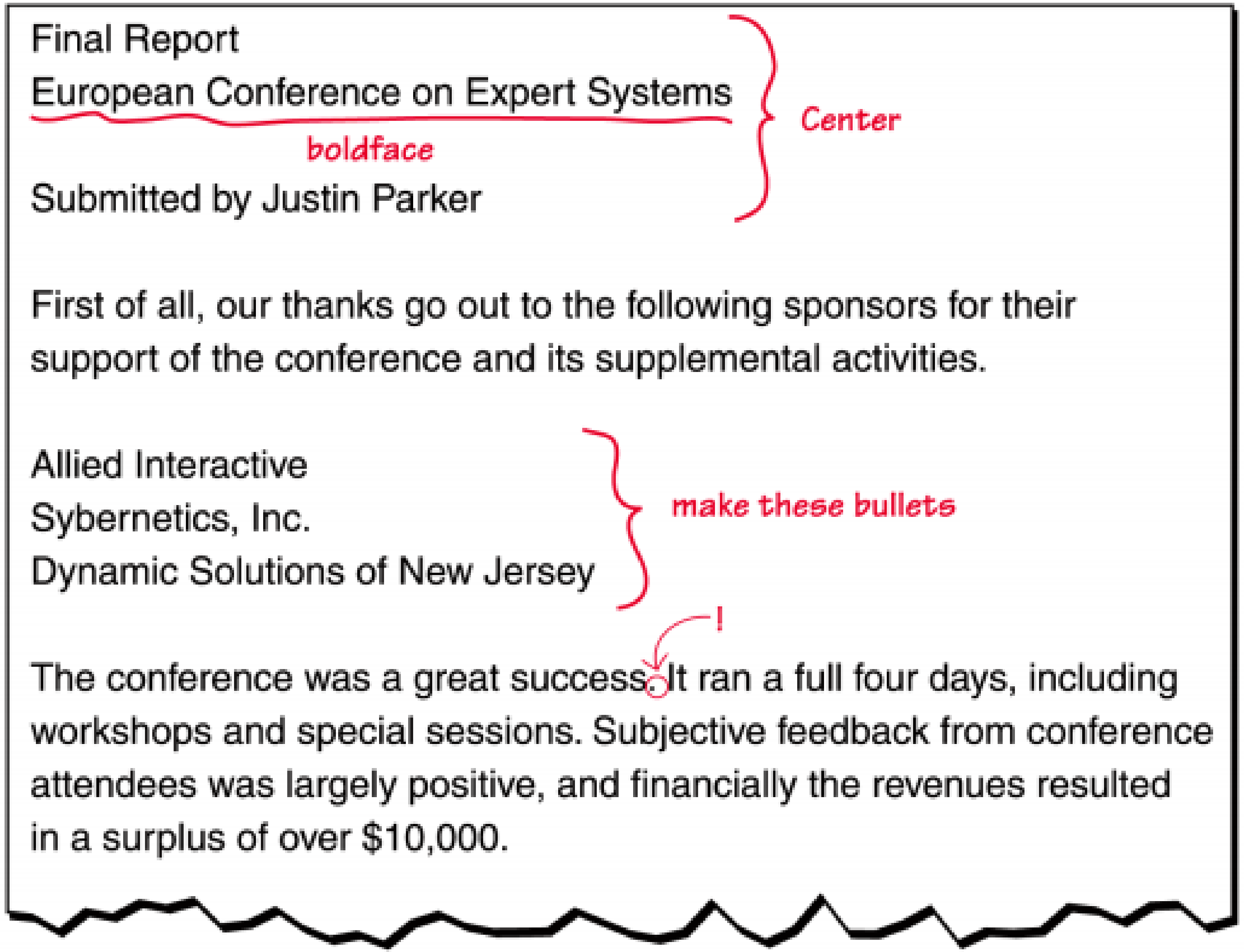
CSC 402 – Internet Technology

Recap

- Web Communication
- Web Service
- Web Service Vs. Data File Transfer
- Web Content
- Static and Dynamic Contents
- URL and its parts
- URI, URN, and URL

Markup



Final Report

European Conference on Expert Systems

boldface

Submitted by Justin Parker

Center

First of all, our thanks go out to the following sponsors for their support of the conference and its supplemental activities.

Allied Interactive

Sybernetics, Inc.

Dynamic Solutions of New Jersey

make these bullets

!

The conference was a great success. It ran a full four days, including workshops and special sessions. Subjective feedback from conference attendees was largely positive, and financially the revenues resulted in a surplus of over $10,000.

Markup Language

- A markup language is a set of characters or symbols that defines a document’s logical structure—that is, it specifies how a document should be printed or displayed.
- The browser wars began over **Dynamic Hypertext Markup Language (DHTML)**—a combination of various technologies, including HTML and JavaScript—which allows a Web page to change after it has been loaded by a browser.
  - Examples of DHTML functionality include the ability to position text and elements, change document background color, and create effects such as animation.
- Unfortunately for Netscape, the W3C adopted as the formal standard the version of DHTML in version IE v4, which prompted many loyal Netscape users to defect to Microsoft.
- At the time of this writing, browser market share (ref: https://www.w3counter.com/globalstats.php):
  - Chrome: 38.42%
  - Safari: 17.26%
  - Firefox: 8.64%
  - IE v11: 6.73%
- E.g.: LaTex, HTML, DHTML, WML, etc.

Markup Language

- Meta-markup language is a set of rules for creating a markup language.
    - SGML - Standard Generalized Markup Language
    - XML - eXtensible Markup Language
- Key Terminologies:
    - Tag
    - Element
    - Attributes

Markup Language – Terminologies

- Tag: A markup that constitutes an instruction to an interpreting program, and is not part of the text being marked up.
- In HTML, tags begins with "<" and ends with ">"
- They come in three flavors
    - Start-tags, for example <p>
    - End-tags, for example </p>
    - Empty-element tags, for example <br/>
- Element: If a document can be converted into a "tree"-like representation, as HTML can, then an element is a "node" in the "tree".
- In HTML, an element begins with a start-tag and ends with a matching end-tag or consists only of an empty-element tag.
- Characters between the start- and end-tags are the element's content
    - May include other elements, which are called child elements
    - An example of an element is <p>Hello, world.</p>

HTML

- HTML stands for Hyper Text Markup Language.
- HTML is based on an older language called **SGML**, which defines the data in a document independent of how the data will be displayed.
- SGML separates the data in a document from the way that data is formatted.
- Each element in an SGML document is marked according to its type, such as paragraphs, headings, and so on.
- Like SGML, HTML was originally designed as a way of defining the elements in a document independent of how they would appear.
- HTML was not intended to be used for designing the actual appearance of the pages in a Web browser, but the language gradually evolved to have this capability.

HTML

- A simple SGML application used to markup text documents into Web pages.
- Used only a small subset of SGML’s capabilities.
- HTML is not a programming language.
  - Specifies meaning and structure of content.
  - Not designed to specify presentation.
  - Created by Tim Berners-Lee in 1991.
  - Browser war created the proprietary HTML tags and incompatible implementations of HTML.
  - Intertwining of presentation layer and structure layer (e.g. font and center tags)
- **Note**: HTML is a markup language SGML is meta-markup language!

HTML

- HTML uses < > to denote **tags** or you can “mark” the logical elements of a document.
- Tags can provide:
  - Formatting information (e.g <i> for italic)
  - Meaning of text:
    - <h> means top-level heading.
    - <p> means paragraph.
    - <ul> <li> for unordered (bulleted) list.
  - Additional information to display (e.g. inline images using <img>).
    - <IMG ALIGN=LEFT SRC="http://somewebsite.com:80/eg/eg.1x1.4bit.COLOR.gif"> NEW RELEASE!
  - Tags can have tags inside (nesting supported)!

HTML

- Simple HTML:
-
```
<HTML>
        <HEAD>
        <TITLE>Hello Worl</TITLE>
        </HEAD>
        <BODY>
        Document body
        </BODY>
</HTML>
```
- Every HTML element is a box.
  - Explore boxes with developer tools: http://mr-pc.org/teaching/cisc1600/lec01.html
- Browser doesn’t care about indentations but Programmers do!

HTML

- Basic preparation tips:
  - Find documents that are similar to the ones you want to put online. Then use View Source to save the other document as a template.
  - Edit your document in one window
    - Use “Reload” option.
    - Be sure to view your document on multiple platforms like different Web browser,
  - Consider low-bandwidth users.
  - Use thumbnail images that link to larger images.
  - Give file sizes on links to large documents.
  - Minimize number of separate inline images in a document.
- Design Issues
  - People expect high quality graphics, but beware of glitz.
  - Don’t forget users using older equipment or slow modems.
  - Avoid designing for the machine on your desktop.
  - Beware of designing pages that only work with the latest version of browsers.
  - Avoid excessive use of harsh color and complex backgrounds.

Document Object Model (DOM)

- A web browser interprets your HTML and builds a model of the page, the DOM.
- The DOM is what is rendered to the screen.
- The DOM can be manipulated by CSS and JavaScript after it is built.
- Tasks that can be performed with DOM.
    - Navigate an XML document's structure, which is a tree stored in memory.
    - Report the information found at the nodes of the XML tree.
    - Add, delete, or modify elements in the XML document.
- DOM represents each node of the XML tree as an object with properties and behavior for processing the XML.
- The DOM originated as a specification to allow JavaScript scripts and Java programs to be portable among Web browsers.
- The DOM defines the objects and properties of all document elements, and the methods (interface) to access them.

DOM

- The DOM is separated into 3 different levels:
  - Core DOM – standard model for any structured document
  - XML DOM – standard model for XML documents
  - HTML DOM – standard model for HTML documents
- The HTML DOM defines the objects and properties of all HTML elements, and the methods (interface) to access them.
- The XML DOM is a standard for how to get, change, add, or delete XML elements.

XML

- eXtensible Markup Language.
- HTML is a HyperText Markup language.
  - Designed for a specific application, namely, presenting and linking hypertext documents.
- XML describes structure and content (“semantics”)
  - The presentation is defined separately from the structure and the content.
- XML supports internationalization (Unicode) and platform independence (an XML file is just a character file).
- No fixed set of tags
  - New tags can be added for new applications.
- An agreed upon set of tags can be used in many applications.
  - Namespaces facilitate uniform and coherent descriptions of data.
- For example, a namespace for address books determines whether to use <tel> or <phone>.
- Captures the document structure rather than the presentation information.

- An address as a book:

```
<addresses>
        <person>
                <name> Donald Duck</name>
                <tel> 01-987-1234 </tel>
                <email> donald@waltdisney.com </email>
        </person>
        <person>
                <name> Miki Mouse</name>
                <tel> 03-426-1142 </tel>
                <email>miki@waltdisney.com</email>
        </person>
</addresses>
```

HTML vs XML

```
<h2>Nonmonotonic Reasoning:
ContextDependent
Reasoning</h2>
        <i>by <b>V. Marek</b> and
        <b>M.
        Truszczynski</b></i><br>
Springer 1993<br>
ISBN 0387976892
```

```
<book>
    <title>Nonmonotonic Reasoning:
    ContextDependent
    Reasoning</title>
    <author>V. Marek and</author>
    <author>M. Truszczynski</author>
    <publisher>Springer</publisher>
    <year>1993</year>
    <ISBN>0387976892</ISBN>
</book>
```

HTML vs XML

- Style: HTML is a markup language for a specific purpose (display in browsers)Pre-defined tags.
- Pre-defined tags.
- Content and Style are in the same file.
- Uses server-side scripts.
- Unidirectional links.
- HTML: same intention, but has evolved into a presentation language.

- Semantics: XML is a framework for defining markup languages.
- User-defined tags.
- Content and Style are separated.
- Uses browser-side scripts.
- Multi-directional and powerful links.
- XML defines logical structure only,
- Optional user-defined DTD (Document Type Definition) to validate.

XHTML

- EXtensible Hyper Text Markup Language.
- An XML application for the structure and meaning of Web documents.
- XHTML keeps the vocab of HTML 4.01 but gains the stricter syntactic rules of XML.
- Offers more consistency in structure than HTML.
- Latest version: XHTML 1.1.

HTML5 VS. XHTML

- The good points of HTML5
  - It is the new standard endorsed by W3C.
  - It is supported by all new browsers and will eventually be widely supported.
  - It is backwards compatible with older HTML4 and XHTML.
  - It introduces some simplifications and new and exciting elements for audio and video.
- The bad points of HTML5
  - It abandons the stricter rules of XHTML.
  - Browsers do not enforce HTML rules.
  - Leads to bad web pages
  - Validating HTML5 documents is harder.