Advanced HTML

Web Architecture and Information Management [./]
Spring 2009 — INFO 190-02 (CCN 42509)

Erik Wilde, UC Berkeley School of Information

2009-02-02

Advanced HTML ; Erik Wilde ; UC Berkeley School of In... http://localhost/drectures/web-spring09/html-advanced

Contents

Abstract (2)

This lecture covers linking in general and in header information, and a more general view of HTML layout based on the box model used by browsers. The concept of *frames* is introduced, which can be used in a combination of framesets and pages, or as inline frames. Finally, *image maps* are introduced as a way of how images can be turned not only into links, but into a set of various linked areas overlayed over the image.

Advanced HTML ; Erik Wilde ; UC Berkeley School of In... http://localhost/drectures/web-spring09/html-advanced

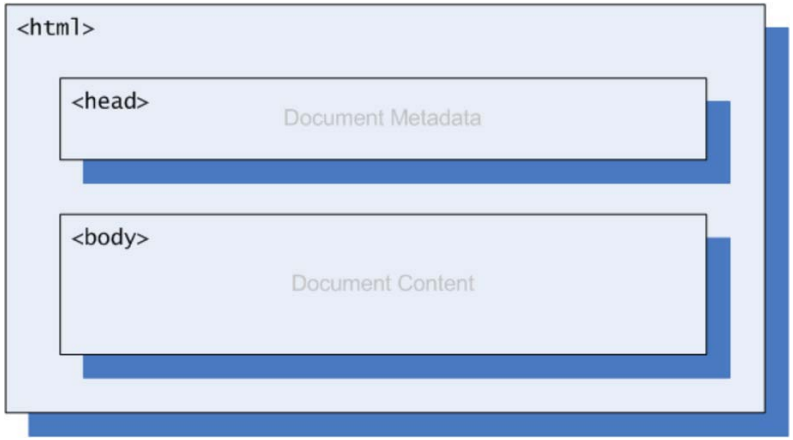
Header Information

Other Links (4)

- Links using a are the most important links on the Web
  - href points to the link target
  - most of the time, the link anchor is text or an image
- HTML has many more element linking to other resources
  - q/blockquote [http://www.w3.org/TR/REC-html40/struct/text.html#h-9.2.2] point to the source of quotations
  - img [http://www.w3.org/TR/REC-html40/struct/objects.html#h-13.2] specifies an image and embeds this image into the page
  - form [http://www.w3.org/TR/REC-html40/interact/forms.html#h-17.3] points to a URI to which the contents of a form are submitted
  - object [http://www.w3.org/TR/REC-html40/struct/objects.html#h-13.3] embeds an object in a web page (such as a Flash app)
  - frame [http://www.w3.org/TR/REC-html40/present/frames.html#h-16.2.2] loads a Web page into a frame
  - iframe [http://www.w3.org/TR/REC-html40/present/frames.html#h-16.5] embeds a Web page into a Web page
  - link [http://www.w3.org/TR/REC-html40/struct/links.html#h-12.3] connects a page to ancillary resources
  - script [http://www.w3.org/TR/REC-html40/interact/scripts.html#h-18.2.1] specifies the location of scripting code
  - this list is not complete (but close)

Web Architecture and Information Management [./]
Spring 2009 — INFO 190-02 (CCN 42509)
2009-02-02

HTML Document Structure (5)



Advanced HTML ; Erik Wilde ; UC Berkeley School of In... http://localhost/drectures/web-spring09/html-advanced

E. Wilde: Advanced HTML Header Information

HTML Document Type (6)

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN">
```

- HTML pages have to declare their document format
  - browsers/clients should know which version of HTML they are dealing with
  - HTML uses the same element names in different version
- A *Document Type Declaration* [http://www.w3.org/TR/REC-html40/struct/global.html#h-7.2] officially declares the document type
- HTML has three different document types:
  1. *Transitional* for backwards-compatibility
  2. *Strict* for HTML as it should be (more restricted than "Transitional")
  3. *Frameset* for using Frames [Frames (1)]
- The *public identifier* specifies HTML version information [http://www.w3.org/TR/REC-html40/struct/global.html#h-7.2]

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN">

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN">

<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Frameset//EN">
```

Web Architecture and Information Management [./]
Spring 2009 — INFO 190-02 (CCN 42509)
2009-02-02

Document Metadata (7)

- All document content is specified in the HTML body
  - this is what a browser renders in the document window
  - rendering may need additional information such as scripts and style
- Information *about* the document is contained in its head [http://www.w3.org/TR/html4/struct/global.html#h-7.4.1]

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN">
<html>
 <head>
  <meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
  <title>HTML Syntax and Structure</title>
  <link href="simple.css" rel="stylesheet" type="text/css">
 </head>
 <body>
  <h1>HTML Syntax and Structure</h1>
```


Advanced HTML ; Erik Wilde ; UC Berkeley School of In... http://localhost/drectures/web-spring09/html-advanced

E. Wilde: Advanced HTML Header Information

Essential Metadata (8)

- Page title [http://www.w3.org/TR/html4/struct/global.html#h-7.4.2]s are used in various places
  - in the title bar of the browser window and/or the browser tab
  - when creating bookmarks
  - in the results of search engines
- Style information can be embedded in style [http://www.w3.org/TR/html4/present/styles.html#edef-STYLE] (not reusable)
- HTML pages can link [http://www.w3.org/TR/html4/struct/links.html#edef-LINK] to external resources
  - a number of *link types* [http://www.w3.org/TR/html4/types.html#type-links] define relationships
  - some relationships are in widespread use but are not standardized (e.g., "`icon`")
  - one possible link type is "`stylesheet`" for pointing to external styles
- External styles need three pieces of information
  - href specifies the URI of the external stylesheet
  - rel specifies the link type "`stylesheet`"
  - type specifies the type of the stylesheet as a media type (`text/css`)

Web Architecture and Information Management [./] 2009-02-02
Spring 2009 — INFO 190-02 (CCN 42509)

Additional Metadata (9)

- base [http://www.w3.org/TR/html4/struct/links.html#edef-BASE] sets the base URI for all relative URIs
  - can be useful if the page contains many references to a different site

```
<base href="http://en.wikipedia.org/wiki/">

<a href="HTML" title="Wikipedia: HTML">HTML</a>
```

[http://en.wikipedia.org/wiki/HTML]

- meta [http://www.w3.org/TR/html4/struct/global.html#h-7.4.4.2] specifies general metadata for a page
  - keywords and description from the early days of search engines, largely ignored these days
  - additional metadata schemes have been defined (*Dublin Core* [http://dublincore.org/documents/dcq-html/])
  - metadata is a wide field and depends on usage and users

Advanced HTML ; Erik Wilde ; UC Berkeley School of In... http://localhost/drectures/web-spring09/html-advanced

Creating Content

All-Purpose Elements (11)

- HTML elements are supposed to convey structural semantics
  - lists [http://www.w3.org/TR/html4/struct/lists.html] are available in various flavors (ul, ol, dl)
  - various phrase markup elements [http://www.w3.org/TR/html4/struct/text.html#h-9.2.1] are available (em, strong, dfn, code, samp, kbd, var, cite, abbr, acronym)
  - various levels of headings [http://www.w3.org/TR/html4/struct/global.html#h-7.5.5] can be used (h1-h6)
- HTML content should represent structural information
  - not all content can be mapped to HTML elements
  - in many cases HTML elements are available (there are even diff elements [http://www.w3.org/TR/html4/struct/text.html#h-9.4])
- HTML also has all-purpose elements [http://www.w3.org/TR/html4/struct/global.html#h-7.5.4]
  - these elements have no semantics and are just containers
  - span is used as an inline container
  - div is used as a block container
  - all-purpose elements should only be used if no HTML element is available

Web Architecture and Information Management [./] 2009-02-02
Spring 2009 — INFO 190-02 (CCN 42509)

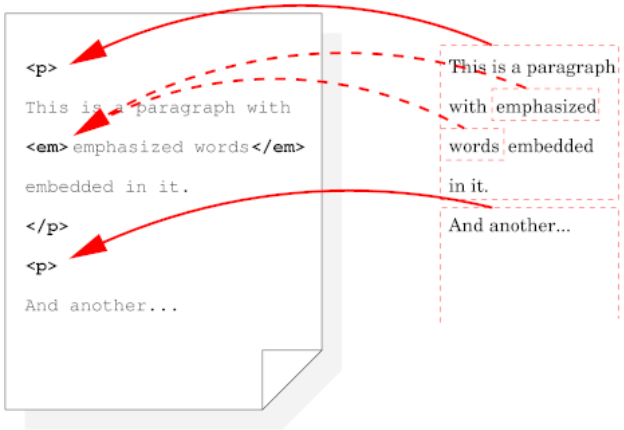
Advanced HTML ; Erik Wilde ; UC Berkeley School of In... http://localhost/drectures/web-spring09/html-advanced

Retain Content Structures (12)

- HTML should represent content structures
  - CSS [Cascading Style Sheets (CSS)] can be used to tweak the formatting (if required)
- Rich content should be mapped to rich Web pages
  - use HTML elements if available
  - augment HTML elements with CSS classes [Cascading Style Sheets (CSS); Use Classes & Containers (1)] for more specific semantics
  - use Microformats [Semantic Web and Microformats; Microformats (1)] for capturing more semantics
- HTML is just one possible representation of a resource
  - the data model of resources should not be limited by HTML
  - richer representations may become available in the future
  - why are there so few Web pages with tel: URIs?

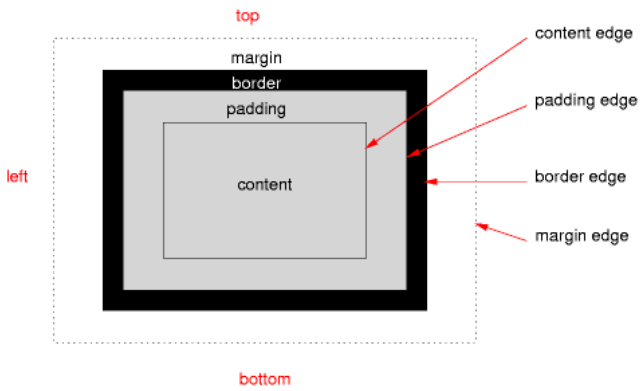
HTML/CSS Box Model

Structure and Layout (14)



Web Architecture and Information Management [./] 2009-02-02
Spring 2009 — INFO 190-02 (CCN 42509)

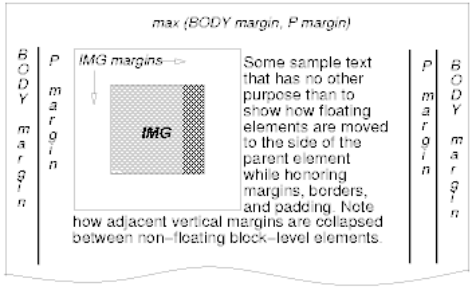
Box Structure (15)



Advanced HTML ; Erik Wilde ; UC Berkeley School of In... http://localhost/drectures/web-spring09/html-advanced

E. Wilde: Advanced HTML HTML/CSS Box Model

Floating Boxes Layout (16)



Floating Boxes Markup (17)

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN">
<html>
 <head>
  <title>Float example</title>
  <style type="text/css">
  body { font-size : 300% ; }
  body, p, img { margin : 1em ; }
  img { float : left ; width : 50% ; }
  </style>
 </head>
 <body>
  <p>
  <img src="../img/float-box.png" alt="This image will illustrate
floats"/>
  Some sample text that has no other purpose than to show how
floating elements are moved to the side of the parent element while
honoring margins, borders, and padding. Note how adjacent vertical
margins are collapsed between non-floating block-level elements.</p>
 </body>
</html>
```


Frames

Combining Documents in the Browser (19)

- HTML pages usually are one document loaded by the browser
- Frames were created to be able to combine documents
- Frames were created when server-side frameworks were primitive
  - building site navigation with frames is rather simple (HTML only)
  - building site navigation without frames is harder (server support required)
- More modern applications combine content differently
  - assembled on the server side and delivered as one document
  - assembled in the browser via scripting and one logical document

Problems with Frames (20)

- Frame-based sites are hard to use from the Web point of view
  - it is hard or impossible to link to pages
  - user's have a hard time creating bookmarks
- Search engines have problems pointing users to results
  - pointing to the frameset might not even contain the result
  - pointing to individual pages can result in unusable pages
- Printing frame-based pages usually does not work very well
  - most browsers support print functionality for one frame at a time
- Frames are not considered good practice anymore
  - they can still be useful for internal or limited audiences
  - they can be useful for rapid prototyping
  - pages for a general Web audience whould not use frames

Advanced HTML ; Erik Wilde ; UC Berkeley School of In... http://localhost/drectures/web-spring09/html-advanced

Regular Frames

Framesets and Frames (22)

- HTML pages can be HTML content or framesets [http://www.w3.org/TR/REC-html40/present/frames.html#h-16.1]
- For framesets, the page *only* defines a frameset "skeleton"
  - the frameset [http://www.w3.org/TR/REC-html40/present/frames.html#h-16.2.1] described the structure of the page
  - individual frame [http://www.w3.org/TR/REC-html40/present/frames.html#h-16.2.2]s point to actual HTML content
- The browser retrieves the frameset and all frame contents
  - rendering a frameset results in a *compound document*
- Links in the frameset can load content into individual frames
  - a frame's name identifies a frame with a name
  - a'a target instructs the browser to load content into that frame

Frameset and Frame Content (23)

```
<html>
 <head>
  <title>Lecture Browser</title>
 </head>
 <frameset cols="20%, 80%">
  <frame src="lectures-toc.html"/>
  <frame name="slides" src="../intro"/>
 </frameset>
</html>
```

```
<html>
 <head>
  <title>Lecture Table of Contents</title>
 </head>
 <body>
  <ul>
   <li><a href="../intro" target="slides">Introduction</a> (<a
href="../2009-01-21-intro.pdf" target="slides">PDF</a>)</li>
   <li><a href="../setup" target="slides">Setup</a> (<a
href="../2009-01-26-setup.pdf" target="slides">PDF</a>)</li>
   <li><a href="../html" target="slides">HTML</a> (<a
href="../2009-01-28-html.pdf" target="slides">PDF</a>)</li>
   <li><a href="../html-advanced" target="slides">Advanced HTML</a>
(<a href="../2009-02-02-html-advanced.pdf" target="slides">PDF</a>)
</li>
   <li><a href="../css" target="slides">CSS</a> (<a
href="../2009-02-04-css.pdf" target="slides">PDF</a>)</li>
  </ul>
 </body>
</html>
```


IFrames

Embedding HTML into HTML (25)

- *Inline frames* [http://www.w3.org/TR/REC-html40/present/frames.html#h-16.5] embed HTML pages in HTML pages
  - Regular Frames [Regular Frames (1)] segment the window and then load HTML into the parts
  - iframe is a box somewhere in an HTML page and contains an HTML page
- IFrames have the same usability/accessibility issues as frames
  - printing is a problem (scrolled content in scrolled content)
  - navigation is a problem (complex navigation in the context of one page)
  - addressing (bookmarks/search) is a problem

Advanced HTML ; Erik Wilde ; UC Berkeley School of In... http://localhost/drectures/web-spring09/html-advanced

IFrame Example (26)

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN">
<html>
 <head>
  <title>IFrames Lectures</title>
 </head>
 <body style="height : 100%">
  <table style="width : 100% ; height : 100% ; ">
   <tr style="height : 50%">
    <td><iframe src="../intro" width="100%"
height="300">Intro</iframe></td>
    <td><iframe src="../setup" width="100%"
height="300">Setup</iframe></td>
   </tr>
   <tr style="height : 50%">
    <td><iframe src="../html" width="100%" height="300">HTML</iframe>
</td>
    <td><iframe src="../html-advanced" width="100%"
height="300">Advanced HTML</iframe></td>
   </tr>
  </table>
 </body>
</html>
```


Image Maps

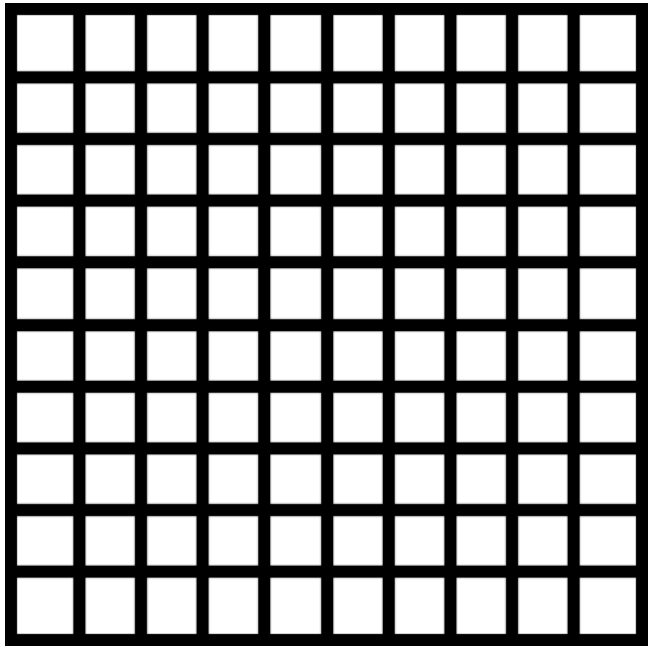
Clickable Images (28)

- Links on Web pages are often text or images
  - almost anything can be turned into a link by wrapping it in a
  - if an image turns into a link, any part of the image can be clicked
- Images may convey structural information
  - an organizational chart of a company
  - a world map with countries or regions
  - photographs with marked parts (buildings, persons, …)
- *HTML image maps* [] turn an image into a structured link
  - the older Server-Side Image Maps [Server-Side Image Maps (1)] are simpler for the browser
  - the newer Client-Side Image Maps [Client-Side Image Maps (1)] are much more user-friendly

Web Architecture and Information Management [./] 2009-02-02
Spring 2009 — INFO 190-02 (CCN 42509)

Server-Side Image Maps (29)

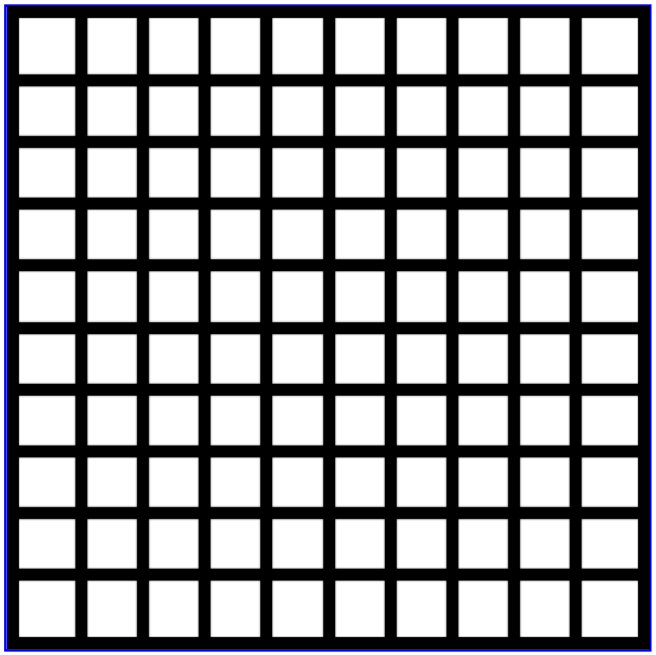
```
<a href="server-side-program"><img src="grid.png" ismap="ismap"/></a>
```



[server-side-program]

Client-Side Image Maps (30)

[square22]
[square36]



```
<img
src="grid.png" usemap="#gridmap">
<map name="gridmap">
        <a title="Square 2/2" href="square22" shape="rect"
coords="60,60,96,96"/>
        <a title="Square 3/6" href="square36" shape="rect"
coords="106,244,142,282"/>
</map>
```

- Various shapes are supported
  - rect for rectangles (x_1,y_1,x_2,y_2)
  - circle for circles (x,y,radius)
  - poly for polygons (x_1,y_1,…,x_n,y_n)