

HTML 5

This lecture is based on materials from:



HTML5: The Missing Manual
The Book That Should Have Been in the Box
Matthew MacDonald
O'Reilly Media

History



HTML 1 included 18 simple elements

HTML 4.1 had lots of additional features:

- form-based file upload
- tables
- image maps
- internationalization

Problem: Document were not well structured
Easy to create bad markup → undefined behaviour
Hard to process/mine automatically



Goal: Consistent Rendering

- XHTML

- Philosophy: Eliminate markup errors by enforcing strict syntax rules.
- Theoretically pure
- It failed, miserably!
- Why?



Goal: Consistent Rendering

- HTML5
 - Philosophy: Developers will mess up.
Standardize how to handle errors in markup
 - Pragmatic



Principles

- Don't break the web
- Pave the Cowpaths
- Be practical



New Features

- Semantic elements
- Enhanced web form widgets
- Audio and video support
- Canvas
- Drag and drop
- Local data storage
- Offline applications
- Server events
- Geolocation



Browser Support

- Differ among browsers and versions
- <http://caniuse.com>
- Runtime feature detection with *modernizr* Javascript library



Document Structure

```
<!DOCTYPE html>
```

```
<html lang="eng">
```

```
  <head>
```

```
    <title>The page's title</title>
```

```
    <meta charset="utf-8">
```

```
    <link href="mystylesheet.css" rel="stylesheet">
```

```
    <script src="myscript.js"> </script>
```

```
  </head>
```

```
  <body>
```

```
    <p lang="fr">Ceci est un paragraphe.</p>
```

```
  </body>
```

```
</html>
```




Syntax

- Reflects the reality in the web
- `<html>`, `<head>`, and `<body>` are optional
- Ignores capitalization
 - `<P>This is a fineexample</p>`
- OK to omit closing slash from void element
 - Both `
` and `
` are ok
- Attribute values don't need quotation marks
- Value-less attributes are allowed
 - `<input type="checkbox" checked>`



Good Style

- Include `<html>`, `<head>`, and `<body>` tags
- Use lowercase tags
- Use quotation marks around attribute values



Validation

- Checks that the document conforms to the specification
- Identifies
 - Missing mandatory elements (e.g., <title>)
 - Start tag without matching end tag
 - Nesting errors
 - Tags with missing attributes
 - Elements or content in the wrong place
- <http://validator.w3.org>



New Elements

- Semantic elements for structure
 - `<article>`, `<aside>`, `<figcaption>`, `<figure>`, `<footer>`,
`<header>`, `<hgroup>`, `<nav>`, `<section>`, `<details>`,
`<summary>`
- Semantic elements for text
 - `<mark>`, `<time>`, `<wbr>`
- Web forms
 - `<datalist>`, `<keygen>`, `<meter>`, `<progress>`,
`<command>`, `<menu>`, `<output>`
- Audio and video
 - `<audio>`, `<video>`, `<source>`
- Canvas
 - `<canvas>`



Cascading Style Sheets (CSS)



Structure vs. Presentation

- HTML specifies document structure
- CSS specifies presentation

Bad Idea: Mixing Presentation and Document Structure



Date	9498.30	11-21
London	1092.50	-1.94
New York	1092.50	-1.94

Quote delayed by 20 minutes.

```
<TABLE BORDER=1 CELLS=3 COLSPAN=1>
<TR>
<TD>
<TD>
<TD>
</TD>
</TR>
<TR>
<TD>
<TD>
<TD>
</TD>
</TR>
<TR>
<TD>
<TD>
<TD>
</TD>
</TR>
</TABLE>
<P>Quote delayed by 20 minutes.</P>
```

```
<TABLE>
<TR>
<TD>
<TD>
<TD>
</TD>
</TR>
<TR>
<TD>
<TD>
<TD>
</TD>
</TR>
<TR>
<TD>
<TD>
<TD>
</TD>
</TR>
</TABLE>
<P>Quote delayed by 20 minutes.</P>
```

2,400 HTML characters to describe 60 characters of content



Problems



Cascading Style Sheets (CSS)

- Separate structure from presentation
- “Simple” mechanism to attach style to structured documents
 - fonts, colours, spacing, ...



CSS Advantages

- Precise control over presentation
- Simplify site maintenance
- Faster downloads
- Media-specific rendering



CSS Language

stylesheet: ruleset*

ruleset: selector '{' [declaration ';']* '}'

declaration: property ':' expr ['! important']?

```
p {  
  font-family: sans-serif;  
  color: red;  
}
```



Selectors

- Type E
- Universal *
- Grouping E,G,F
- Attribute [foo="hi"]
- ID #myID or E#myID
- Class .myClass
- Pseudo-element E:pseudo-element
- Contextual
 - Descendent E F
 - Child E > F
 - Adjacent E + F



Available Formatting

- Font
- Text
- Background
- Display
- Box
- Positioning
- Animation

Display property

property	description
display	sets the type of CSS box model an element is displayed with

- values: inline | block | list-item | inline-block | table | inline-table | table-row-group | table-header-group | table-footer-group | table-row | table-column-group | table-column | table-cell | table-caption | none | inherit

```
<ul id="topmenu">  
  <li>Item 1</li>  
  <li>Item 2</li>  
  <li>Item 3</li>  
</ul>
```

HTML

```
#topmenu li {  
  display: inline;  
  border: 2px solid gray;  
  margin-right: 1em;  
}
```

CSS

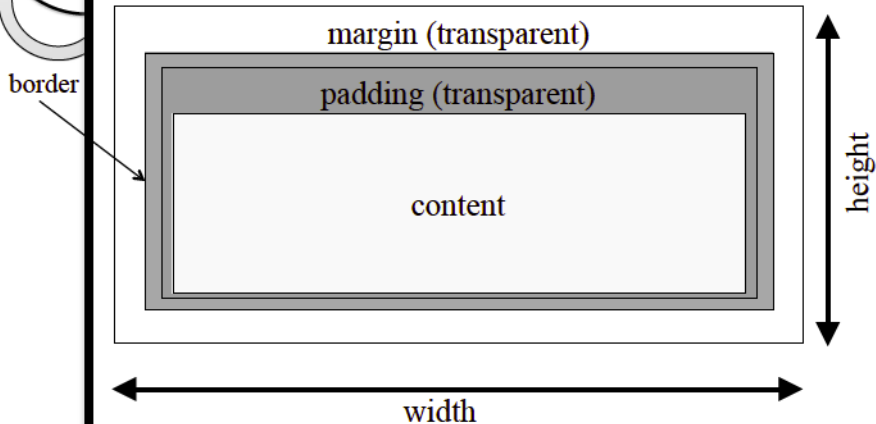
Item 1

Item 2

Item 3

output

Box Formatting Model



Firebug: CSS debugging

- Firefox add-on (<http://getfirebug.com/>)
- Inspect XHTML/CSS of any page
- Change styles dynamically

