CS1115/CS5002
Web Development 1

Dr Derek Bridge

School of Computer Science & Information Technology
University College Cork

DOCTYPE

- The DOCTYPE tells the browser which version of HTML you are using
- If no DOCTYPE, a very old web page or the author of the page is incompetent
- HTML3, e.g.:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 3.2 Final//EN">
```

- HTML 4 has six different DOCTYPES, e.g.:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
    "http://www.w3.org/TR/html14/strict.dtd">
```

- HTML5 has simplified it:

```
<!DOCTYPE html>
```

The overall structure

- Everything other than DOCTYPE is enclosed between <html> and </html> and divided into <head> and <body>
- <head>:
    o Information *about* the document that is not considered part of the document content, e.g. <title>
- <body>:
    o The *content* of the document
- **Quick quiz:** In a browser, where are they displayed?

A very simple web page

```
<!DOCTYPE html>
<html>
    <head>
        <title>An example web page</title>
    </head>
    <body>
        <h1>My web page</h1>
        <p>
            This is a paragraph. Good, isn't it?
        </p>
        <p>
            This is another paragraph. Good, isn't it?
        </p>
    </body>
</html>
```

- Put this in a file whose **filename extension** is .html
- Save it on your server (in the correct folder with the correct permissions)
- Then people can use their browser to request to see it

Character references

- To include special characters in the content, use character references, e.g.:

<	<	>	>	&	&
Á	Á	á	á	€	€

- You can include them by name, by hexadecimal number or by decimal number, e.g. á á á
- A list: https://dev.w3.org/html5/html-author/charref

Lists of items

- An unordered list (usually shown with bullet points):

```
<ul>
<li>Sex</li>
<li>Drugs</li>
<li>Rock'n'Roll</li>
</ul>
```

- An ordered list (usually shown with numbers):

```
<ol>
<li>I came</li>
<li>I saw</li>
<li>I conquered</li>
</ol>
```

- Don't put lists inside paragraphs — in some sense, they *are* paragraphs

Headings and paragraphs

- Major headings:

```
<h1>My web page</h1>
```

- Paragraphs:

```
<p>
    This is a paragraph. Good, isn't it?
</p>
```

Information that browsers ignore

- The following usually have no effect on what gets displayed:
  o line breaks
  o tabs and multiple spaces ('whitespace')
  o empty <p> elements
  o comments

```
<!-- This is an example comment -->
```

- **Question:** Why nevertheless should we format our HTML nicely?

Markup

- **Markup:**
  - Annotations on a document
- **HTML** (HyperText Markup Language):
  - the markup that turns text documents into web pages
  - primarily, it indicates document **structure**

Nested elements

- Elements are often **nested**, e.g.:

```
<ol>
  <li>I came</li>
  <li>I saw</li>
  <li>I conquered</li>
</ol>
```

- Hence web pages have a **hierarchical structure**

HTML elements

- An HTML **element** comprises:
  - **content** (e.g. the text), plus
  - markup (the **tags**: start tag and end tag)
- element = start tag + content + end tag
  … except for **void elements**, which have start tag but no content and no end tag

Hierarchical structures: terminology

tree	node
root	leaf
parent	child
ancestor	descendant
sibling	

The hierarchical structure of HTML documents

```
<!DOCTYPE html>
<html>
  <head>
    <title>A simple document</title>
  </head>
  <body>
    <p>
      Some words.
    </p>
    <p>
      More words
      <em>and emphasised words</em>
      and final words.
    </p>
  </body>
</html>
```

document

html

head

body

title

p

p

A simple document

Some words.

More words

em

and final words.

and emphasised words