به نام خدا

HTML

درس برنامه نویسی وب

مدرس: میترا عیسایی

دانشجوی دکتری مهندسی کامپیوتر-نرم افزار

مهر ۱۴۰۲

HTML Introduction

- HTML is the standard markup language for creating Web pages.

- HTML describes the structure of a Web page

HTML History

Year	Version
1989	Tim Berners-Lee invented www
1991	Tim Berners-Lee invented HTML
1993	Dave Raggett drafted HTML+
1995	HTML Working Group defined HTML 2.0
1997	W3C Recommendation: HTML 3.2
1999	W3C Recommendation: HTML 4.01
2000	W3C Recommendation: XHTML 1.0
2008	WHATWG HTML5 First Public Draft
2012	WHATWG HTML5 Living Standard
2014	W3C Recommendation: HTML5
2016	W3C Candidate Recommendation: HTML 5.1
2017	W3C Recommendation: HTML5.1 2nd Edition
2017	W3C Recommendation: HTML5.2

Learn HTML Using Notepad or TextEdit

- Web pages can be created and modified by using professional HTML editors.

- However, for learning HTML, it is recommend a simple text editor like Notepad (PC) or TextEdit (Mac).

- The use of a simple text editor is a good way to learn HTML.

How to View HTML Source

View HTML Source Code:

Click CTRL + U in an HTML page, **or** right-click on the page and select "View Page Source". This will open a new tab containing the HTML source code of the page.

Inspect an HTML Element:

Right-click on an element (or a blank area), and choose "Inspect" to see what elements are made up of (you will see both the HTML and the CSS).

You can also edit the HTML or CSS on-the-fly in the Elements or Styles panel that opens.

What is an HTML Element?

- An HTML **element** is defined by a start tag, some content, and an end tag:
- <tagname> Content goes here... </tagname>

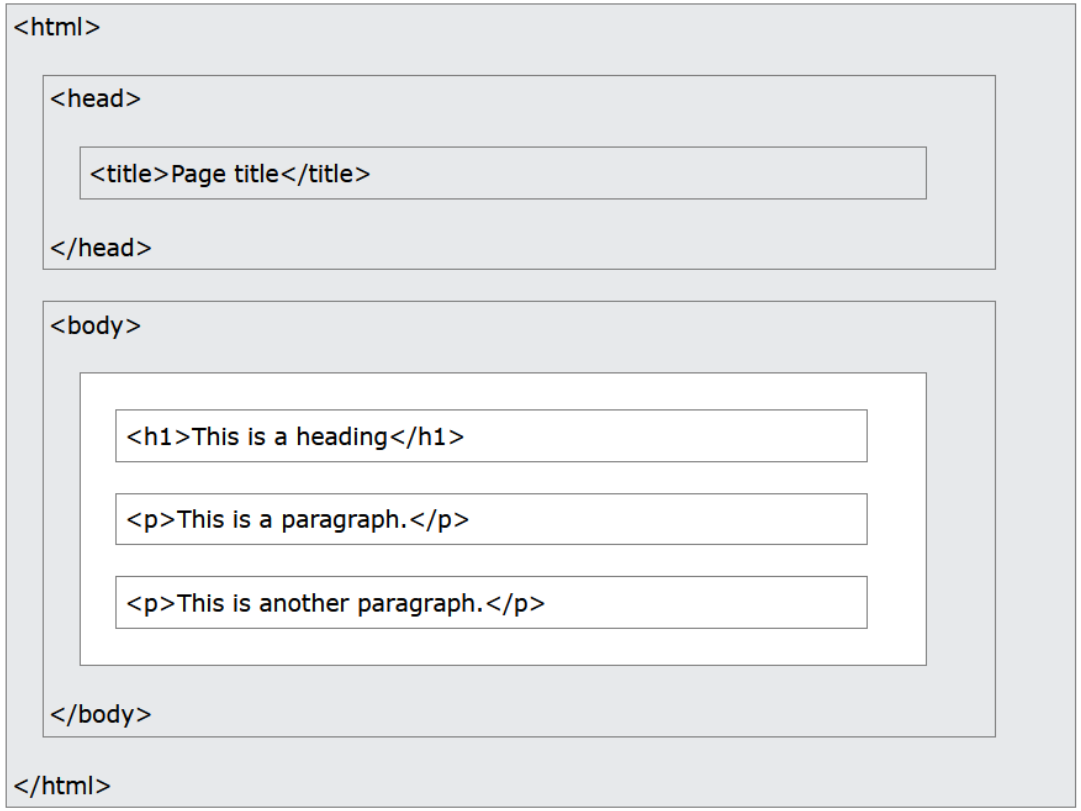
- **Empty elements:** Some HTML elements have no content (like the <br> element).

- Empty elements do not have an **end tag**!

Start tag	Element content	End tag
<h1>	My First Heading	</h1>
	My first paragraph.	
 	none	none

HTML Page Structure

- Below is a visualization of an HTML page structure:



Nested HTML Elements

Some elements can contain other elements

```
<!DOCTYPE html>
<html>
        <body>

                <h1>My First Heading</h1>
                <p>My first paragraph.</p>

        </body>
</html>
```

HTML Documents

All HTML documents must start with a document type declaration: <!DOCTYPE html>.

The <!DOCTYPE> declaration represents the document type, and helps browsers to display web pages correctly.

It must only appear once, at the top of the page (before any HTML tags).

<html>

The <html> element is the root element and it defines the whole HTML document.

It has a start tag <html> and an end tag </html>.

Then, inside the <html> element there is a <body> element:

<head>

- The <head> element is a container for metadata (data about data)
- between the <html> tag and the <body> tag.
- HTML metadata is data about the HTML document.
- Metadata is not displayed.
- Metadata typically define the document <title>, <style>, <meta>, <link>, <script>, and …

<title>

- The <title> tag defines the title of the document.
- it is shown in the browser's title bar.
- The contents of a page title is very important for search engine optimization (SEO)!
- displays a title for the page in search-engine results
- provides a title for the page when it is added to favorites

<title>

Here are some tips for creating good titles:

•Go for a longer, descriptive title (avoid one- or two-word titles)

•Search engines will display about 50-60 characters of the title, so try not to have titles longer than that.

•Do not use just a list of words as the title (this may reduce the page's position in search results)

- ✓ So, try to make the title as accurate and meaningful as possible!
- ✓ **Note:** You can NOT have more than one <title> element in an HTML document.

HTML Encoding (Character Sets)

- To display an HTML page correctly, a web browser must know which character set to use.

  This is specified in the <meta> tag:
  <meta charset="UTF-8">

- ASCII was the first character encoding standard. ASCII defined 128 different characters

- ISO-8859-1 was the default character set for HTML 4. This character set supported 256 different character codes.

- The HTML5 specification encourages web developers to use the UTF-8 character set, which covers almost all of the characters and symbols in the world!

HTML Attributes

- HTML attributes provide additional information about HTML elements.
- All HTML elements can have **attributes**
- Attributes are always specified in **the start tag**
- Attributes usually come in name/value pairs like: **name="value"**

HTML Attributes

- Google
- <img src="img_girl.jpg">
- <img src="img_girl.jpg" width="500" height="600">

Quote for Attribute Values

- The HTML standard does not require **quotes** around attribute values.
- However, W3C **recommends** quotes in HTML
- Double quotes around attribute values are the most common in HTML, but single quotes can also be used.
- In some situations, when the attribute value itself contains double quotes, it is necessary to use single quotes:

```
<p title='John "ShotGun" Nelson'>
```

The lang Attribute

the lang attribute inside the <html> tag to declare the language of the Web page.

This is meant to assist search engines and browsers.

```
<!DOCTYPE html>
<html lang="en">
        <body>
        /body>
</html>
```

The lang Attribute

Country codes can also be added to the language code in the lang attribute.

So, the first two characters define the language of the HTML page, and the last two characters define the country.

```
<!DOCTYPE html>
<html lang="en-US">
        <body>
        /body>
</html>
```

HTML style Attribute

- The style attribute specifies an inline style for an element, such as **color, font, size, and more**.
- The style attribute can be used on **any** HTML element.
- **However, it is not necessarily useful.**

- <*tagname* style="*property*:*value*;">
- This is a paragraph.

HTML Headings

- HTML headings are titles or subtitles that you want to display on a webpage.
- HTML headings are defined with the **<h1>** to **<h6>** tags.
    - <h1> defines the most important heading.
    - <h6> defines the least important heading.
- Search engines use the headings to index the structure and content of your web pages.
- Each HTML heading has a default size.
- Change the size for any heading with font-size property:

```
<h1 style="font-size:60px;">Heading 1</h1>
```

HTML Paragraphs

- A paragraph always starts on a new line
- browsers automatically add some **white space** (a margin) before and after a paragraph.
- The browser will automatically remove any extra spaces and lines when the page is displayed:
-
  This paragraph
  contains a lot of spaces
  in the source code,
  but the browser
  ignores it.

HTML Text Formatting

- HTML contains several elements for defining text with a special meaning.

HTML Formatting Elements

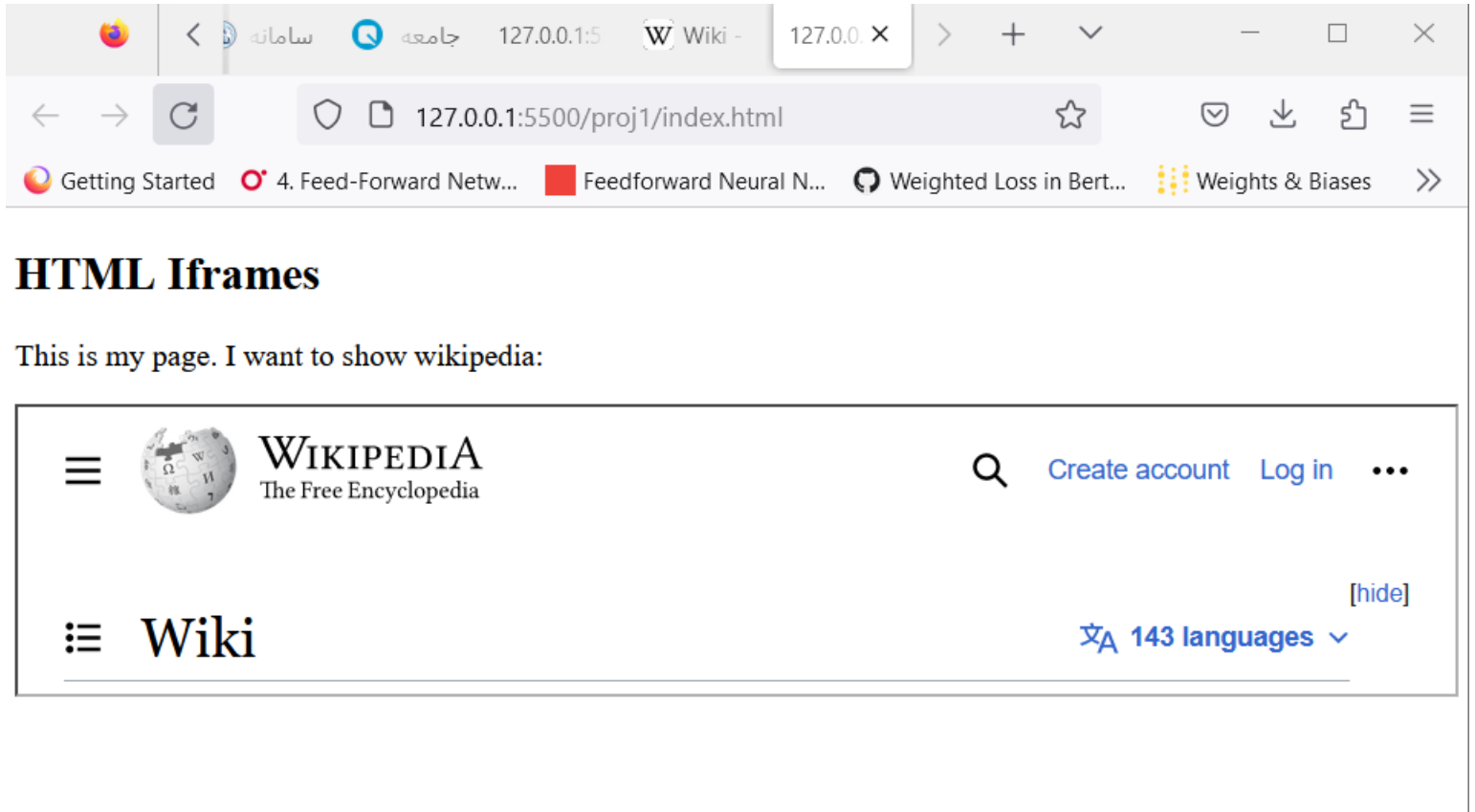
- <b> - Bold text
- - Important text
- <i> - Italic text
- - Emphasized text
- <mark> - Marked text
- <del> - Deleted text
- <ins> - Inserted text
- <sub> - Subscript text
- <sup> - Superscript text

HTML Computer Code Elements

- HTML contains several elements for defining user input and computer code.
- <kbd> defines keyboard input
- <samp> defines output from a computer program
- <code> defines a piece of computer code
- <var> defines a variable in a mathematical expression

HTML Iframes

- An HTML iframe is used to display a web page within a web page.



HTML Iframes

The HTML <iframe> tag specifies an inline frame. An inline frame is used to embed another **document** within the **current** HTML document.

```
<iframe src="url" title="description"></iframe>
```

Use the height and width attributes to specify the size of the iframe.

By default, an iframe has a border around it. **To remove** the border, use `frameborder="0"`

HTML Iframes

ClickJacking Attack!

ClickJacking is where a malicious user loads the site inside of some frame, while using the design of the site to try and get users to pass personal information that can be intercepted or collected.

```
<meta http-equiv="Content-Security-Policy" content="frame-ancestors 'self'">
```

HTML Responsive Web Design

- A responsive web design is about creating web pages that look good on **all devices**!
- A responsive web design will automatically adjust **for different screen sizes**.
- To create a responsive website, add the following <meta> tag to all your web pages:

```
<meta name="viewport" content="width=device-width, initial-scale=1.0">
```

- The width=device-width sets the width of the page to follow the **screen-width** of the device.
- The initial-scale=1.0 sets the initial zoom level when the page is **first loaded** by the browser.

HTML Caracter Entities

- Some characters are reserved in HTML. Reserved characters in HTML must be **replaced** with entities:

  &*entity_name*;

- If you use the less than **(<)** or greater than **(>)** signs in your HTML text, the browser might mix them with tags.

- To display a less than sign (<) we must write: **<**

- A commonly used HTML entity is the non-breaking space: ** **

HTML Caracter Entities

- Some Useful HTML Character Entities

- Entity names are case sensitive.

Result	Description	Name
	non-breaking space	
<	less than	&lt;
>	greater than	&gt;
&	ampersand	&amp;
"	double quotation mark	&quot;
'	single quotation mark	&apos;
¢	cent	&cent;
£	pound	&pound;
¥	yen	&yen;
€	euro	&euro;
©	copyright	&copy;
®	trademark	&reg;

HTML Caracter Entities

- Some Mathematical Entities

- Entity names are case sensitive.

Char	Number	Entity	Description
∀	&#8704;	&forall;	For all
∂	&#8706;	&part;	Partial differential
∃	&#8707;	&exist;	There exists
∅	&#8709;	&empty;	Empty sets
∇	&#8711;	&nabla;	Nabla
∈	&#8712;	&isin;	Element of
∉	&#8713;	&notin;	Not an element of
∋	&#8715;	&ni;	Contains as member
∏	&#8719;	&prod;	N-ary product
∑	&#8721;	&sum;	N-ary summation

Using Emojis in HTML

- Emojis look like images, or icons, but they are not.

Emoji	Value
	&#128507;
	&#128508;
	&#128509;
	&#128510;
	&#128511;
	&#128512;
	&#128513;
	&#128514;
	&#128515;
	&#128516;
	&#128517;

HTML Uniform Resource Locators

- A URL is another word for a web address.
- A URL can be composed of **words**, or an Internet Protocol **(IP) address.**
- A Uniform Resource Locator (URL) is used to address a document (or other data) on the web

HTML Uniform Resource Locators

- A web address follows these syntax rules:

scheme://prefix.domain:port/path/filename

- **scheme** - defines the **type** of Internet service (most common is **http or https**)
- **prefix** - defines a domain **prefix** (default for http is **www**)
- **domain** - defines the Internet **domain name** (like w3schools.com)
- **port** - defines the **port number** at the host (default for http is **80**)
- **path** - defines a **path** at the server (If omitted: the root directory of the site)
- **filename** - defines the name of a document or resource

HTML Uniform Resource Locators

- Common URL Schemes

Scheme	Short for	Used for
http	HyperText Transfer Protocol	Common web pages. Not encrypted
https	Secure HyperText Transfer Protocol	Secure web pages. Encrypted
ftp	File Transfer Protocol	Downloading or uploading files
file		A file on your computer

HTML Uniform Resource Locators

- **URL Encoding**
- URLs can only be sent over the Internet using the ASCII Character-set.
- If a URL contains characters outside the ASCII set, the URL has to be converted.
- URL encoding replaces non-ASCII characters with a "%" followed by hexadecimal digits.
- URLs cannot contain spaces. URL encoding normally replaces a space with a plus **(+) sign, or %20**.