HTML5

The newest iteration of hypertext markup language is HTML5 (often cited without a space before the digit). What follows is a barebones HTML5 page and some of the most common tags. Type this into a text editor and we'll upload it to the Huco server.

```
<!DOCTYPE html>
<html lang=en>

<head>
    <meta charset=utf-8>
    <title>My Webpage</title>
</head>

<body>
    <p>I'm at the <a href="http://www.ualberta.ca">University of Alberta</a>.</p>
</body>

</html>
```

A few important notes:

1. Here are two important concepts when it comes to markup languages:

a) being ***well-formed*** means that the containers nest properly inside one another (no quirky overlaps). The tags may or may not be legitimate, but the *structure* of the document is fine.

b) being ***valid*** means that the document also abides by the grammar set out in the Document Type Declaration (DTD) or the Schema (those two terms are roughly synonymous). The document contains only tags and attributes that are defined in the Schema and that there are no invalid nestings (it's illegal in HTML, for example, to put **body** inside **p**).

2. The file begins by declaring its Document Type Declaration (**DOCTYPE**). That statement defines the official "grammar" of this document as **html** (i.e., HTML5, as opposed to an earlier version of HTML). That particular grammar will be used when validating this document. (Visit the official W3 validator at <http://validator.w3.org> to try it out).

3. Notice that the document defines both a language (**en** = English) and a character set (**utf-8** = 8-bit Unicode).

4. Notice that in **attribute="value"** pairs, no space is allowed on either side of the **=** sign and that quotation marks are often optional. The quotation marks are not optional when the value contains spaces or characters that might be interpreted by the browser as demarcating characters. When in doubt, use quotation marks.

HTML5 Tags

meta-structural tags

html   the main container; all HTML content must reside inside **html**. Takes a **lang** attribute, which defines the language of the document.

head   container for meta-information about this page: its title, character set, links to CSS documents or to Javascript files, etc. (*Note: different from* **header**.)

body   the container for all "visible" content on the page.

*meta-tags inside **head***

link   contains an **href** attribute pointing to another document that should be imported (Cascading Style Sheets, Javascript, etc.) Must be inside **head**. (Not to be confused with **a**, which creates a hyperlink in the **body**.) You can have multiple **link** tags inside **head**. Documents will be imported in the order in which they are listed, which can be important for CSS since later rules can overwrite earlier rules.

meta   information such as character set or search keywords. Must be inside **head**.

content tags

<!--   beginning and ending comment markers. Comments are human-readable
-->   messages. You should feel free to add human-readable comments to your web pages: **<!-- this is an html comment -->**

a   anchor tag that create a hyperlink; takes an **href** (hypertext reference) attribute

article   contains content like forum posts, blog entries, newspaper articles, etc. Represents an independent, stand-alone chunk of content. (Compare **section**.)

aside   for extracted content (e.g., a pull quote in a magazine article)

em   for emphasized text. Default presentation is italics, but you can change that in the CSS.

footer   contains small print, links, logos, copyright data, authorship info, etc.; any info typically in a footer; interestingly, you can have more than one footer on a page.

form contains **input** elements used to gather information from users. Takes two attributes: 1) a **method** attribute (values are either **get** or **post**), which specifies how information will be delivered to the server; 2) an **action** attribute which specifies which URL to send the data to (if not specified, data is sent to this URL).

h1 *thru* **h6** headers; **h1** is the biggest, **h6** is the smallest (but you can change size and presentational qualities in the CSS). These are typically used for chapter titles or sub-headers in longer text passages.

header contains introductory info typically in a top banner: name of the site or organization, logos, perhaps **nav**, etc. (*Note: not to be confused with* **head**.)

img an image tag; takes a **src** attribute. Note: **img** is inline by default.

input a data collection element inside a **form**. The attribute value specifies the type of **input** element and notifies the browser how to draw the element. (Compare **select**.) In order to pass data to the server, input elements need to have **name** and/or **id** attributes.

label specifies a label for input tags; although you can add descriptive text to an input element without using a **label** tag, users can click on the **label** text to place the cursor inside the **input** element.

li list item; goes inside either **ol** (ordered list) or **ul** (unordered list).

nav designed to contain navigational elements (links, lists, etc.)

ol ordered list; contains **li** elements. The browser will auto-generate a numbered sequence, which can be re-defined via your CSS. (Compare **ul**.)

p text paragraph

section groups content chunks into subject areas. This might signify, for example, content to be repurposed or redistributed. (Compare **article**.)

select A pop-up (or "drop-down") menu. Items are defined with the **option** tag.

strong default presentation is bolded text, but you can change that in the CSS.

table generates a spreadsheet-like chart of rows and columns. Rows are specified with **tr** (table row) and individual cells are delimited with **td** (table data) tags.

ul unordered list; contains **li** elements. The browser will generate bullets or other signifiers as specified by your CSS. (Compare **ol**.)

legacy non-presentational grouping containers

div   a historical (pre-HTML5) non-presentational *block-level* container (i.e., default presentation is to be invisible). Typically seen with an **id** or **class** attribute.

span   a historical (pre-HTML5) non-presentational *inline* container (i.e., default presentation is to be invisible). Typically seen with an **id** or **class** attribute.

global attributes

Any HTML tag can take the following attributes:

class   defines an identifier that a group of elements may have (dot in CSS)

id   defines a unique identifier for that particular element (hash in CSS). Note: often accompanied by an identical **name** attribute, especially in **input** elements inside a **form**.