**HTML Basics #5: Symbols and Charset**

Let’s consider this scenario, you are making an online tutorial about HTML, and you want the browser to display an HTML tag, <p>HTML</p>, so this is what you did:

<!DOCTYPE html>

<html>

<head>

<meta charset="utf-8" />

<title>My HTML Tutorial</title>

</head>

<body>

<h1>HTML Tutorial</h1>

<p> <p>HTML</p> </p>

</body>

</html>

But then you realize, the <p>HTML</p> tag will be rendered instead of displayed. How can you solve this problem? How can you display HTML tags in HTML documents?

Table of Contents

* [HTML Entities](https://www.ericsdevblog.com/index.php/2021/05/html-basics-5-symbols-and-charset/#HTML_Entities)
* [Symbols and Emojis](https://www.ericsdevblog.com/index.php/2021/05/html-basics-5-symbols-and-charset/#Symbols_and_Emojis)
  + [Some Common Symbols](https://www.ericsdevblog.com/index.php/2021/05/html-basics-5-symbols-and-charset/#Some_Common_Symbols)
  + [Some Mathematical Symbols](https://www.ericsdevblog.com/index.php/2021/05/html-basics-5-symbols-and-charset/#Some_Mathematical_Symbols)
  + [Some Common Emoji Symbols](https://www.ericsdevblog.com/index.php/2021/05/html-basics-5-symbols-and-charset/#Some_Common_Emoji_Symbols)
* [HTML Charset](https://www.ericsdevblog.com/index.php/2021/05/html-basics-5-symbols-and-charset/#HTML_Charset)
  + [The ASCII Character Set](https://www.ericsdevblog.com/index.php/2021/05/html-basics-5-symbols-and-charset/#The_ASCII_Character_Set)
  + [The ANSI Character Set (Windows-1252)](https://www.ericsdevblog.com/index.php/2021/05/html-basics-5-symbols-and-charset/#The_ANSI_Character_Set_Windows-1252)
  + [The ISO-8859-1 Character Set](https://www.ericsdevblog.com/index.php/2021/05/html-basics-5-symbols-and-charset/#The_ISO-8859-1_Character_Set)
  + [The UTF-8 Character Set](https://www.ericsdevblog.com/index.php/2021/05/html-basics-5-symbols-and-charset/#The_UTF-8_Character_Set)

**HTML ENTITIES**

Some characters in HTML are reserved, and to display them, we must replace them with HTML entities. An HTML entity has the format &entity\_name; or &#entity\_number;. One commonly used entity is the non-breaking space &nbsp;. Remember we talked about paragraphs (<p>) and how these two paragraphs are the same?

<p>This is a paragraph.</p>

<p>This is a

paragraph.</p>

Now, this example leaves us with a new problem, what if we want multiple spaces between two words? The answer is the HTML entity &nbsp;:

<p>This is a paragraph.</p>

<p>This&nbsp;&nbsp;&nbsp;&nbsp;&nbsp;is a paragraph.</p>

Text

Description automatically generated with medium confidence

**SYMBOLS AND EMOJIS**

**Some Common Symbols**

| **Char** | **Number** | **Entity** |
| --- | --- | --- |
| © | &#169; | &copy; |
| ® | &#174; | &reg; |
| € | &#8364; | &euro; |
| ™ | &#8482; | &trade; |
| ← | &#8592; | &larr; |
| ↑ | &#8593; | &uarr; |
| → | &#8594; | &rarr; |
| ↓ | &#8595; | &darr; |
| ♠ | &#9824; | &spades; |
| ♣ | &#9827; | &clubs; |
| ♥ | &#9829; | &hearts; |
| ♦ | &#9830; | &diams; |

**Some Mathematical Symbols**

| **Char** | **Number** | **Entity** |
| --- | --- | --- |
| ∀ | &#8704; | &forall; |
| ∂ | &#8706; | &part; |
| ∃ | &#8707; | &exist; |
| ∅ | &#8709; | &empty; |
| ∇ | &#8711; | &nabla; |
| ∈ | &#8712; | &isin; |
| ∉ | &#8713; | &notin; |
| ∋ | &#8715; | &ni; |
| ∏ | &#8719; | &prod; |
| ∑ | &#8721; | &sum; |

**Some Common Emoji Symbols**

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

**HTML CHARSET**

To make sure that these entities are displayed correctly in the browser, we need to specify the character encoding standard (Charset) that is used in the web page:

<meta charset="utf-8" />

There are many character sets that we can use, but HTML5 encourages us to use UTF-8, which covers almost all the characters and symbols in the world.

**The ASCII Character Set**

ASCII uses the values from 0 to 31 (and 127) for control characters.

ASCII uses the values from 32 to 126 for letters, digits, and symbols.

ASCII does not use the values from 128 to 255.

**The ANSI Character Set (Windows-1252)**

ANSI is identical to ASCII for the values from 0 to 127.

ANSI has a proprietary set of characters for the values from 128 to 159.

ANSI is identical to UTF-8 for the values from 160 to 255.

**The ISO-8859-1 Character Set**

ISO-8859-1 is identical to ASCII for the values from 0 to 127.

ISO-8859-1 does not use the values from 128 to 159.

ISO-8859-1 is identical to UTF-8 for the values from 160 to 255.

**The UTF-8 Character Set**

UTF-8 is identical to ASCII for the values from 0 to 127.

UTF-8 does not use the values from 128 to 159.

UTF-8 is identical to both ANSI and 8859-1 for the values from 160 to 255.

UTF-8 continues from the value 256 with more than 10 000 different characters.