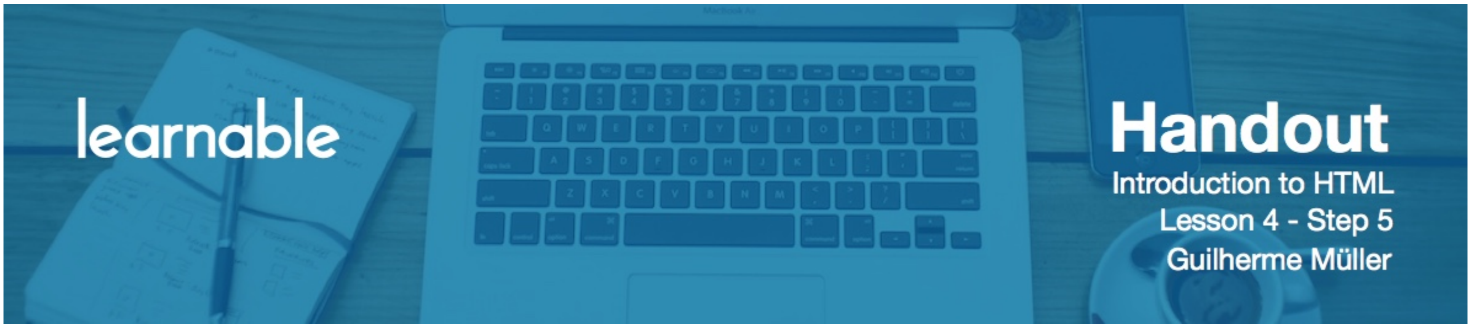

Handout

Introduction to HTML
Lesson 4 - Step 5
Guilherme Müller

Introduction

In this lesson we'll talk about the importance of page encoding and how to represent special characters in HTML using HTML entities. To prepare for the lesson open up the *"htmlentities-finish.html"* file or the *"htmlentities-begin.html"* if you want to test your own examples.

The importance of correct page encoding

In English, we use the regular 26-letter Roman alphabet. Accent marks or other special symbols are not used regularly, except for foreign words. This can hide some problems that might happen with file encoding. For example, if the `meta charset` tag is set to `UTF-8`, but the file is saved in another encoding, say, `ISO-8859-1`, which is a common encoding and usually called Latin1 or Western, we will have problems while rendering of some characters.

If you open the *"htmlentities-problem.html"* file you'll notice that most text is being shown without problems, but the special characters are replaced with question marks. This means that the meta tag encoding must match the file encoding.

When you save the file most text editors have tools to convert encoding to the desired one. When using Sublime Text it is very simple:

- Go to File menu
- Save with Encoding
- Choose the desired encoding

HTML entities

There is a way to represent special characters without typing them directly. Meet **HTML entities**. They start with an ampersand and end with a semicolon:

```
&aacute;
```

This will show the symbol regardless of problems with encoding. HTML entities can be used as well with other symbols like `©` for the copyright symbol or `&` for the ampersand itself:

```
&copy;
&amp;
```

There are many more examples on this page: http://dev.w3.org/html5/html-author/charref. If you open this page, you'll see that there are two other kinds of codes for each symbol which correspond to the universal character set. Think of it as an identification for a character inside the big table. The first uses decimal numbers and the other hexadecimal numbers.