This chapter describes the XML **Character Entities**. Before we understand the Character Entities, let us first understand what an XML entity is.

As put by [W3 Consortium](http://www.w3.org) the definition of entity is as follows:

> The document entity serves as the root of the entity tree and a starting-point for an XML processor.

This means, entities are the placeholders in XML. These can be declared in the document prolog or in a DTD. There are different types of entities and this chapter will discuss Character Entity.

Both, the HTML and the XML, have some symbols reserved for their use, which cannot be used as content in XML code. For example, `<` and `>` signs are used for opening and closing XML tags. To display these special characters, the character entities are used.

There are few special characters or symbols which are not available to be typed directly from keyboard. Character Entities can be used to display those symbols/special characters also.

### Types of Character Entities

There are three types of character entities:

- Predefined Character Entities
- Numbered Character Entities
- Named Character Entities

### Predefined Character Entities

They are introduced to avoid the ambiguity while using some symbols. For example, an ambiguity is observed when less than `<` or greater than `>` symbol is used with the angle tag `<>`. Character entities are basically used to delimit tags in XML. Following is a list of pre-defined character entities from XML specification. These can be used to express characters without ambiguity.

- Ampersand: `&amp;`
- Single quote: `apos;`
- Greater than: `&gt;`
- Less than: `&lt;`
- Double quote: `&quot;`

### Numeric Character Entities

The numeric reference is used to refer to a character entity. Numeric reference can either be in decimal or hexadecimal format. As there are thousands of numeric references available, these are a bit hard to remember. Numeric reference refers to the character by its number in the Unicode character set.

General syntax for decimal numeric reference is:

```xml
&# decimal number ;
```

General syntax for hexadecimal numeric reference is:
The following table lists some predefined character entities with their numeric values:

<table>
<thead>
<tr>
<th>Entity name</th>
<th>Character</th>
<th>Decimal reference</th>
<th>Hexadecimal reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>quot</td>
<td>&quot;</td>
<td>&amp;#34;</td>
<td>&amp;#x22;</td>
</tr>
<tr>
<td>amp</td>
<td>&amp;</td>
<td>&amp;#38;</td>
<td>&amp;#x26;</td>
</tr>
<tr>
<td>apos</td>
<td>'</td>
<td>&amp;#39;</td>
<td>&amp;#x27;</td>
</tr>
<tr>
<td>lt</td>
<td>&lt;</td>
<td>&amp;#60;</td>
<td>&amp;#x3C;</td>
</tr>
<tr>
<td>gt</td>
<td>&gt;</td>
<td>&amp;#62;</td>
<td>&amp;#x3E;</td>
</tr>
</tbody>
</table>

**Named Character Entity**

As it's hard to remember the numeric characters, the most preferred type of character entity is the named character entity. Here, each entity is identified with a name.

For example:

- 'Aacute' represents capital Á character with acute accent.
- 'ugrave' represents the small ù with grave accent.