

20.3. `html.entities` — Definitions of HTML general entities

Source code: <Lib/html/entities.py>

This module defines four dictionaries, `html5`, `name2codepoint`, `codepoint2name`, and `entitydefs`.

`html.entities.html5`

A dictionary that maps HTML5 named character references [1] to the equivalent Unicode character(s), e.g. `html5['gt;'] == '>'`. Note that the trailing semicolon is included in the name (e.g. `'gt;'`), however some of the names are accepted by the standard even without the semicolon: in this case the name is present with and without the `';'`. See also `html.unescape()`.

New in version 3.3.

`html.entities.entitydefs`

A dictionary mapping XHTML 1.0 entity definitions to their replacement text in ISO Latin-1.

`html.entities.name2codepoint`

A dictionary that maps HTML entity names to the Unicode code points.

`html.entities.codepoint2name`

A dictionary that maps Unicode code points to HTML entity names.

Footnotes

[1] See <https://www.w3.org/TR/html5/syntax.html#named-character-references>