

Differences_Between_XHTML_and_HTML

XHTML strictly conforms to the XML language syntax rules, while HTML does not. Since HTML does not require strict adherence to the coding rules of an XML language, there may be an inconsistent display of content on some browsers. XHTML's strict enforcement of XML coding rules ensures consistent presentation of the content on a wide range of browsers and devices. It also ensures that all documents are both well formed and strictly conforming XML, so you can use standard XML tools to manage, transform, manipulate, and parse them.

HTML is case-insensitive in tag names. XHTML is case-sensitive for all element and attribute names. All element names in XHTML must be lowercase. In the HTML column, some end tag can be missing, but XHTML all elements must be closed

Quotation marks do not surround the attribute value of the href attribute in the HTML code. Attribute values in XHTML must always be enclosed in double or single quotes.

All empty tags in XHTML must be "self-closing"; for example,
 has a closing / inside the tag to indicate that there is no corresponding </br> tag. This applies to br, hr, input, etc.

Attribute-value pairs cannot be minimized on XHTML as they can be in HTML. In HTML, when an attribute may take only one value, it is permitted to omit the value.

JP

Character reference overview

A numeric character reference refers to a character by its Universal Character Set/Unicode code point, and uses the format

```
&#nnnn;
```

or

```
&#xhhhh;
```

where nnnn is the code point in decimal form, and hhhh is the code point in hexadecimal form. The x must be lowercase in XML documents. The nnnn or hhhh may be any number of digits and may include leading zeros. The hhhh may mix uppercase and lowercase, though uppercase is the usual style.

In contrast, a character entity reference refers to a character by the name of an entity which has the desired character as its replacement text. The entity must either be predefined (built-in to the markup language) or explicitly declared in a Document Type Definition (DTD). The format is the same as for any entity reference:

```
&name;
```

where name is the name of the entity. The semicolon is required.

[edit] Predefined entities in XML

The XML specification does not use the term "character entity" or "character entity reference". The XML specification defines five "predefined entities" representing special characters, and requires that all XML processors honor them. The entities can be explicitly declared in a DTD, as well, but if this is done, the replacement text must be the same as the built-in definitions. XML also allows other named entities of any

Differences_Between_XHTML_and_HTML

size to be defined on a per-document basis.

The table below lists the five XML predefined entities. The "Name" column mentions the entity's name. The "Character" column shows the character, if it is renderable. In order to render the character, the format &name; is used; for example, & renders as &. The "Unicode code point" column cites the character via standard UCS/Unicode "U+" notation, which shows the character's code point in hexadecimal. The decimal equivalent of the code point is then shown in parentheses. The "Standard" column indicates the first version of XML that includes the entity. The "Description" column cites the character via its canonical UCS/Unicode name, in English. Name Character Unicode code point Standard Description quot " U+0022 (34) XML 1.0 (double) quotation mark amp & U+0026 (38) XML 1.0 ampersand apos ' U+0027 (39) XML 1.0 apostrophe (= apostrophe-quote) lt < U+003C (60) XML 1.0 less-than sign gt > U+003E (62) XML 1.0 greater-than sign

[edit] Character entities in HTML

The HTML 4 DTD explicitly declares 252 character entities. HTML processors must honor the HTML DTD's declarations, even if the DTD is not mentioned in the HTML document. HTML does not allow other named entities to be defined.

HTML document authors who have been exposed to XML and XHTML often overlook the fact that the apos entity is not defined in HTML. rsquo is the best alternative in this case (but it is not the same character), or authors can use a numeric character reference instead (' or ' which always designate the character using the numeric value of its Unicode code point, independently of the actual document encoding).

In the table below, the HTML built-in character entities are listed. The columns are as in the XML entity table, above, except "Standard" column indicates the first version of HTML that includes the entity. The version is one of the major releases of the HTML spec: 2.0, 3.2, or 4.0. HTML 4.01 didn't introduce any new entities. Name Character Unicode code point Standard DTD[1] Old ISO subset[2] Description[3] quot " U+0022 (34) HTML 2.0 HTMLspecial ISO num quotation mark (= APL quote) amp & U+0026 (38) HTML 2.0 HTMLspecial ISO num ampersand apos ' U+0027 (39) XHTML 1.0 HTMLspecial ISO num apostrophe (= apostrophe-quote); see below lt < U+003C (60) HTML 2.0 HTMLspecial ISO num less-than sign gt > U+003E (62) HTML 2.0 HTMLspecial ISO num greater-than sign nbsp U+00A0 (160) HTML 3.2 HTMLlat1 ISO num no-break space (= non-breaking space)[4] iexcl ¡ U+00A1 (161) HTML 3.2 HTMLlat1 ISO num inverted exclamation mark cent ¢ U+00A2 (162) HTML 3.2 HTMLlat1 ISO num cent sign pound £ U+00A3 (163) HTML 3.2 HTMLlat1 ISO num pound sign curren ¤ U+00A4 (164) HTML 3.2 HTMLlat1 ISO num currency sign yen ¥ U+00A5 (165) HTML 3.2 HTMLlat1 ISO num yen sign (= yuan sign) brvbar ¦ U+00A6 (166) HTML 3.2 HTMLlat1 ISO num broken bar (= broken vertical bar) sect § U+00A7 (167) HTML 3.2 HTMLlat1 ISO num section sign uml ¨ U+00A8 (168) HTML 3.2 HTMLlat1 ISO dia diaeresis (= spacing diaeresis); see German umlaut copy © U+00A9 (169) HTML 3.2 HTMLlat1 ISO num copyright sign ordf ª U+00AA (170) HTML 3.2 HTMLlat1 ISO num feminine ordinal indicator laquo « U+00AB (171) HTML 3.2 HTMLlat1 ISO num left-pointing double angle quotation mark (= left pointing guillemet) not ¬ U+00AC (172) HTML 3.2 HTMLlat1 ISO num not sign shy U+00AD (173) HTML 3.2 HTMLlat1 ISO num soft hyphen (= discretionary hyphen) reg ® U+00AE (174) HTML 3.2 HTMLlat1 ISO num registered sign (= registered trade mark sign) macr ¯ U+00AF (175) HTML 3.2 HTMLlat1 ISO dia macron (= spacing macron = overline = APL overbar) deg ° U+00B0 (176) HTML 3.2 HTMLlat1 ISO num degree sign plusmn ± U+00B1 (177) HTML 3.2 HTMLlat1 ISO num plus-minus sign (= plus-or-minus sign) sup2 ² U+00B2 (178) HTML 3.2 HTMLlat1 ISO num superscript two (= superscript digit two = squared) sup3 ³ U+00B3 (179) HTML 3.2 HTMLlat1 ISO num superscript three (= superscript digit three = cubed) acute ´ U+00B4 (180) HTML 3.2 HTMLlat1 ISO dia acute accent (= spacing acute) micro µ U+00B5 (181) HTML 3.2 HTMLlat1 ISO num micro sign para ¶ U+00B6 (182) HTML 3.2 HTMLlat1 ISO num pilcrow sign (= paragraph sign) middot · U+00B7 (183) HTML 3.2 HTMLlat1 ISO num middle dot (= Georgian comma = Greek middle dot) cedil ¸ U+00B8 (184) HTML 3.2 HTMLlat1 ISO dia cedilla (= spacing cedilla) sup1 ¹ U+00B9 (185) HTML 3.2

Differences_Between_XHTML_and_HTML

HTMLlat1 ISONum superscript one (= superscript digit one) ordm ° U+00BA (186) HTML 3.2 HTMLlat1 ISONum masculine ordinal indicator raquo » U+00BB (187) HTML 3.2 HTMLlat1 ISONum right-pointing double angle quotation mark (= right pointing guillemet) frac14 ¼ U+00BC (188) HTML 3.2 HTMLlat1 ISONum vulgar fraction one quarter (= fraction one quarter) frac12 ½ U+00BD (189) HTML 3.2 HTMLlat1 ISONum vulgar fraction one half (= fraction one half) frac34 ¾ U+00BE (190) HTML 3.2 HTMLlat1 ISONum vulgar fraction three quarters (= fraction three quarters) iquest ¿ U+00BF (191) HTML 3.2 HTMLlat1 ISONum inverted question mark (= turned question mark) Agrave À U+00C0 (192) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter A with grave (= Latin capital letter A grave) Aacute Á U+00C1 (193) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter A with acute Acirc Â U+00C2 (194) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter A with circumflex Atilde Ã U+00C3 (195) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter A with tilde Auml Ä U+00C4 (196) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter A with diaeresis Aring Å U+00C5 (197) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter A with ring above (= Latin capital letter A ring) Aelig Æ U+00C6 (198) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter AE (= Latin capital ligature AE) Ccedil Ç U+00C7 (199) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter C with cedilla Egrave È U+00C8 (200) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter E with grave Eacute É U+00C9 (201) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter E with acute Ecirc Ê U+00CA (202) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter E with circumflex Euml Ë U+00CB (203) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter E with diaeresis Igrave Ì U+00CC (204) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter I with grave Iacute Í U+00CD (205) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter I with acute Icirc Î U+00CE (206) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter I with circumflex Iuml Ï U+00CF (207) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter I with diaeresis ETH Ð U+00D0 (208) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter ETH Ñ U+00D1 (209) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter N with tilde Ograve Ò U+00D2 (210) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter O with grave Oacute Ó U+00D3 (211) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter O with acute Ocirc Ô U+00D4 (212) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter O with circumflex Otilde Õ U+00D5 (213) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter O with tilde Ouml Ö U+00D6 (214) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter O with diaeresis times × U+00D7 (215) HTML 3.2 HTMLlat1 ISONum multiplication sign Oslash Ø U+00D8 (216) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter O with stroke (= Latin capital letter O slash) Ugrave Ù U+00D9 (217) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter U with grave Uacute Ú U+00DA (218) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter U with acute Ucirc Û U+00DB (219) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter U with circumflex Uuml Ü U+00DC (220) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter U with diaeresis Yacute Ý U+00DD (221) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter Y with acute THORN Þ U+00DE (222) HTML 2.0 HTMLlat1 ISolat1 Latin capital letter THORN szlig ß U+00DF (223) HTML 2.0 HTMLlat1 ISolat1 Latin small letter sharp s (= ess-zed); see German Eszett agrave à U+00E0 (224) HTML 2.0 HTMLlat1 ISolat1 Latin small letter a with grave aacute á U+00E1 (225) HTML 2.0 HTMLlat1 ISolat1 Latin small letter a with acute acirc â U+00E2 (226) HTML 2.0 HTMLlat1 ISolat1 Latin small letter a with circumflex atilde ã U+00E3 (227) HTML 2.0 HTMLlat1 ISolat1 Latin small letter a with tilde auml ä U+00E4 (228) HTML 2.0 HTMLlat1 ISolat1 Latin small letter a with diaeresis aring å U+00E5 (229) HTML 2.0 HTMLlat1 ISolat1 Latin small letter a with ring above aelig æ U+00E6 (230) HTML 2.0 HTMLlat1 ISolat1 Latin small letter ae (= Latin small ligature ae) ccedil ç U+00E7 (231) HTML 2.0 HTMLlat1 ISolat1 Latin small letter c with cedilla egrave è U+00E8 (232) HTML 2.0 HTMLlat1 ISolat1 Latin small letter e with grave eacute é U+00E9 (233) HTML 2.0 HTMLlat1 ISolat1 Latin small letter e with acute ecirc ê U+00EA (234) HTML 2.0 HTMLlat1 ISolat1 Latin small letter e with circumflex euml ë U+00EB (235) HTML 2.0 HTMLlat1 ISolat1 Latin small letter e with diaeresis igrave ì U+00EC (236) HTML 2.0 HTMLlat1 ISolat1 Latin small letter i with grave iacute í U+00ED (237) HTML 2.0 HTMLlat1 ISolat1 Latin small letter i with acute icirc î U+00EE (238) HTML 2.0 HTMLlat1 ISolat1 Latin small letter i with circumflex iuml ï U+00EF (239) HTML 2.0 HTMLlat1 ISolat1 Latin small letter i with diaeresis eth ð U+00F0 (240) HTML 2.0 HTMLlat1 ISolat1 Latin small letter eth ntilde ñ U+00F1 (241) HTML 2.0 HTMLlat1 ISolat1 Latin small letter n with tilde ograve ò U+00F2 (242) HTML 2.0 HTMLlat1 ISolat1 Latin small letter o with grave oacute ó U+00F3 (243) HTML 2.0 HTMLlat1 ISolat1 Latin small letter o with acute ocirc ô U+00F4 (244) HTML 2.0 HTMLlat1 ISolat1 Latin small letter o with circumflex

Differences_Between_XHTML_and_HTML

otilde ð U+00F5 (245) HTML 2.0 HTMLlat1 ISolat1 Latin small letter o with tilde ouml ö U+00F6 (246) HTML 2.0 HTMLlat1 ISolat1 Latin small letter o with diaeresis divide ÷ U+00F7 (247) HTML 3.2 HTMLlat1 ISONum division sign oslash ø U+00F8 (248) HTML 2.0 HTMLlat1 ISolat1 Latin small letter o with stroke (= Latin small letter o slash) ugrave ù U+00F9 (249) HTML 2.0 HTMLlat1 ISolat1 Latin small letter u with grave uacute ú U+00FA (250) HTML 2.0 HTMLlat1 ISolat1 Latin small letter u with acute ucirc û U+00FB (251) HTML 2.0 HTMLlat1 ISolat1 Latin small letter u with circumflex uuml ü U+00FC (252) HTML 2.0 HTMLlat1 ISolat1 Latin small letter u with diaeresis yacute ý U+00FD (253) HTML 2.0 HTMLlat1 ISolat1 Latin small letter y with acute thorn þ U+00FE (254) HTML 2.0 HTMLlat1 ISolat1 Latin small letter thorn yuml ÿ U+00FF (255) HTML 2.0 HTMLlat1 ISolat1 Latin small letter y with diaeresis OElig ? U+0152 (338) HTML 4.0 HTMLspecial ISolat2 Latin capital ligature oe[5] oelig ? U+0153 (339) HTML 4.0 HTMLspecial ISolat2 Latin small ligature oe[6] Scaron ? U+0160 (352) HTML 4.0 HTMLspecial ISolat2 Latin capital letter s with caron scaron ? U+0161 (353) HTML 4.0 HTMLspecial ISolat2 Latin small letter s with caron Yuml ? U+0178 (376) HTML 4.0 HTMLspecial ISolat2 Latin capital letter y with diaeresis fnof ? U+0192 (402) HTML 4.0 HTMLsymbol ISOtech Latin small letter f with hook (= function = florin) circ ? U+02C6 (710) HTML 4.0 HTMLspecial ISOpub modifier letter circumflex accent tilde ? U+02DC (732) HTML 4.0 HTMLspecial ISODia small tilde Alpha ? U+0391 (913) HTML 4.0 HTMLsymbol Greek capital letter Alpha Beta ? U+0392 (914) HTML 4.0 HTMLsymbol Greek capital letter Beta Gamma ? U+0393 (915) HTML 4.0 HTMLsymbol ISOgrk3 Greek capital letter Gamma Delta ? U+0394 (916) HTML 4.0 HTMLsymbol ISOgrk3 Greek capital letter Delta Epsilon ? U+0395 (917) HTML 4.0 HTMLsymbol Greek capital letter Epsilon Zeta ? U+0396 (918) HTML 4.0 HTMLsymbol Greek capital letter Zeta Eta ? U+0397 (919) HTML 4.0 HTMLsymbol Greek capital letter Eta Theta ? U+0398 (920) HTML 4.0 HTMLsymbol ISOgrk3 Greek capital letter Theta Iota ? U+0399 (921) HTML 4.0 HTMLsymbol Greek capital letter Iota Kappa ? U+039A (922) HTML 4.0 HTMLsymbol Greek capital letter Kappa Lambda ? U+039B (923) HTML 4.0 HTMLsymbol ISOgrk3 Greek capital letter Lambda Mu ? U+039C (924) HTML 4.0 HTMLsymbol Greek capital letter Mu Nu ? U+039D (925) HTML 4.0 HTMLsymbol Greek capital letter Nu Xi ? U+039E (926) HTML 4.0 HTMLsymbol ISOgrk3 Greek capital letter Xi Omicron ? U+039F (927) HTML 4.0 HTMLsymbol Greek capital letter Omicron Pi ? U+03A0 (928) HTML 4.0 HTMLsymbol Greek capital letter Pi Rho ? U+03A1 (929) HTML 4.0 HTMLsymbol Greek capital letter Rho Sigma ? U+03A3 (931) HTML 4.0 HTMLsymbol ISOgrk3 Greek capital letter Sigma Tau ? U+03A4 (932) HTML 4.0 HTMLsymbol Greek capital letter Tau Upsilon ? U+03A5 (933) HTML 4.0 HTMLsymbol ISOgrk3 Greek capital letter Upsilon Phi ? U+03A6 (934) HTML 4.0 HTMLsymbol ISOgrk3 Greek capital letter Phi Chi ? U+03A7 (935) HTML 4.0 HTMLsymbol Greek capital letter Chi Psi ? U+03A8 (936) HTML 4.0 HTMLsymbol ISOgrk3 Greek capital letter Psi Omega ? U+03A9 (937) HTML 4.0 HTMLsymbol ISOgrk3 Greek capital letter Omega alpha ? U+03B1 (945) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter alpha beta ? U+03B2 (946) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter beta gamma ? U+03B3 (947) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter gamma delta ? U+03B4 (948) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter delta epsilon ? U+03B5 (949) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter epsilon zeta ? U+03B6 (950) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter zeta eta ? U+03B7 (951) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter eta theta ? U+03B8 (952) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter theta iota ? U+03B9 (953) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter iota kappa ? U+03BA (954) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter kappa lambda ? U+03BB (955) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter lambda mu ? U+03BC (956) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter mu nu ? U+03BD (957) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter nu xi ? U+03BE (958) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter xi omicron ? U+03BF (959) HTML 4.0 HTMLsymbol NEW Greek small letter omicron pi ? U+03C0 (960) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter pi rho ? U+03C1 (961) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter rho sigma ? U+03C2 (962) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter final sigma sigma ? U+03C3 (963) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter sigma tau ? U+03C4 (964) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter tau upsilon ? U+03C5 (965) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter upsilon phi ? U+03C6 (966) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter phi chi ? U+03C7 (967) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter chi psi ?

Differences_Between_XHTML_and_HTML

U+03C8 (968) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter psi omega ? U+03C9 (969) HTML 4.0 HTMLsymbol ISOgrk3 Greek small letter omega thetasym ? U+03D1 (977) HTML 4.0 HTMLsymbol NEW Greek theta symbol upsih ? U+03D2 (978) HTML 4.0 HTMLsymbol NEW Greek Upsilon with hook symbol piv ? U+03D6 (982) HTML 4.0 HTMLsymbol ISOgrk3 Greek pi symbol ensp ? U+2002 (8194) HTML 4.0 HTMLspecial ISOpub en space[7] emsp ? U+2003 (8195) HTML 4.0 HTMLspecial ISOpub em space[8] thinsp ? U+2009 (8201) HTML 4.0 HTMLspecial ISOpub thin space[9] zwnj U+200C (8204) HTML 4.0 HTMLspecial NEW RFC 2070 zero-width non-joiner zwj U+200D (8205) HTML 4.0 HTMLspecial NEW RFC 2070 zero-width joiner lrm U+200E (8206) HTML 4.0 HTMLspecial NEW RFC 2070 left-to-right mark rlm U+200F (8207) HTML 4.0 HTMLspecial NEW RFC 2070 right-to-left mark ndash ? U+2013 (8211) HTML 4.0 HTMLspecial ISOpub en dash mdash ? U+2014 (8212) HTML 4.0 HTMLspecial ISOpub em dash lsquo ? U+2018 (8216) HTML 4.0 HTMLspecial ISO num left single quotation mark rsquo ? U+2019 (8217) HTML 4.0 HTMLspecial ISO num right single quotation mark sbquo ? U+201A (8218) HTML 4.0 HTMLspecial NEW single low-9 quotation mark ldquo ? U+201C (8220) HTML 4.0 HTMLspecial ISO num left double quotation mark rdquo ? U+201D (8221) HTML 4.0 HTMLspecial ISO num right double quotation mark bdquo ? U+201E (8222) HTML 4.0 HTMLspecial NEW double low-9 quotation mark dagger ? U+2020 (8224) HTML 4.0 HTMLspecial ISOpub dagger Dagger ? U+2021 (8225) HTML 4.0 HTMLspecial ISOpub double dagger bull ? U+2022 (8226) HTML 4.0 HTMLspecial ISOpub bullet (= black small circle)[10] hellip ? U+2026 (8230) HTML 4.0 HTMLsymbol ISOpub horizontal ellipsis (= three dot leader) permil ? U+2030 (8240) HTML 4.0 HTMLspecial ISOtech per mille sign prime ? U+2032 (8242) HTML 4.0 HTMLsymbol ISOtech prime (= minutes = feet) Prime ? U+2033 (8243) HTML 4.0 HTMLsymbol ISOtech double prime (= seconds = inches) Isaquo ? U+2039 (8249) HTML 4.0 HTMLspecial ISO proposed single left-pointing angle quotation mark[11] rsaquo ? U+203A (8250) HTML 4.0 HTMLspecial ISO proposed single right-pointing angle quotation mark[12] oline ? U+203E (8254) HTML 4.0 HTMLsymbol NEW overline (= spacing overscore) frasl ? U+2044 (8260) HTML 4.0 HTMLsymbol NEW fraction slash (= solidus) euro ? U+20AC (8364) HTML 4.0 HTMLspecial NEW euro sign image ? U+2111 (8465) HTML 4.0 HTMLsymbol ISOamso black-letter capital I (= imaginary part) weierp ? U+2118 (8472) HTML 4.0 HTMLsymbol ISOamso script capital P (= power set = Weierstrass p) real ? U+211C (8476) HTML 4.0 HTMLsymbol ISOamso black-letter capital R (= real part symbol) trade ? U+2122 (8482) HTML 4.0 HTMLsymbol ISO num trademark sign alefsym ? U+2135 (8501) HTML 4.0 HTMLsymbol NEW alef symbol (= first transfinite cardinal)[13] larr ? U+2190 (8592) HTML 4.0 HTMLsymbol ISO num leftwards arrow uarr ? U+2191 (8593) HTML 4.0 HTMLsymbol ISO num upwards arrow rarr ? U+2192 (8594) HTML 4.0 HTMLsymbol ISO num rightwards arrow darr ? U+2193 (8595) HTML 4.0 HTMLsymbol ISO num downwards arrow harr ? U+2194 (8596) HTML 4.0 HTMLsymbol ISOamsa left right arrow crarr ? U+21B5 (8629) HTML 4.0 HTMLsymbol NEW downwards arrow with corner leftwards (= carriage return) lArr ? U+21D0 (8656) HTML 4.0 HTMLsymbol ISOtech leftwards double arrow[14] uArr ? U+21D1 (8657) HTML 4.0 HTMLsymbol ISOamsa upwards double arrow rArr ? U+21D2 (8658) HTML 4.0 HTMLsymbol ISO num rightwards double arrow[15] dArr ? U+21D3 (8659) HTML 4.0 HTMLsymbol ISOamsa downwards double arrow hArr ? U+21D4 (8660) HTML 4.0 HTMLsymbol ISOamsa left right double arrow forall ? U+2200 (8704) HTML 4.0 HTMLsymbol ISOtech for all part ? U+2202 (8706) HTML 4.0 HTMLsymbol ISOtech partial differential exist ? U+2203 (8707) HTML 4.0 HTMLsymbol ISOtech there exists empty ? U+2205 (8709) HTML 4.0 HTMLsymbol ISOamso empty set (= null set = diameter) nabla ? U+2207 (8711) HTML 4.0 HTMLsymbol ISOtech nabla (= backward difference) isin ? U+2208 (8712) HTML 4.0 HTMLsymbol ISOtech element of notin ? U+2209 (8713) HTML 4.0 HTMLsymbol ISOtech not an element of ni ? U+220B (8715) HTML 4.0 HTMLsymbol ISOtech contains as member prod ? U+220F (8719) HTML 4.0 HTMLsymbol ISOamsb n-ary product (= product sign)[16] sum ? U+2211 (8721) HTML 4.0 HTMLsymbol ISOamsb n-ary summation[17] minus ? U+2212 (8722) HTML 4.0 HTMLsymbol ISOtech minus sign lowast ? U+2217 (8727) HTML 4.0 HTMLsymbol ISOtech asterisk operator radic ? U+221A (8730) HTML 4.0 HTMLsymbol ISOtech square root (= radical sign) prop ? U+221D (8733) HTML 4.0 HTMLsymbol ISOtech proportional to infin ? U+221E (8734) HTML 4.0 HTMLsymbol ISOtech infinity ang ? U+2220 (8736) HTML 4.0 HTMLsymbol ISOamso angle and ? U+2227 (8743) HTML 4.0 HTMLsymbol ISOtech logical and (= wedge) or ? U+2228 (8744) HTML 4.0 HTMLsymbol ISOtech logical or (= vee) cap ? U+2229 (8745) HTML 4.0 HTMLsymbol ISOtech

Differences_Between_XHTML_and_HTML

intersection (= cap) cup ? U+222A (8746) HTML 4.0 HTMLsymbol ISOtech union (= cup) int ? U+222B (8747) HTML 4.0 HTMLsymbol ISOtech integral there4 ? U+2234 (8756) HTML 4.0 HTMLsymbol ISOtech therefore sim ? U+223C (8764) HTML 4.0 HTMLsymbol ISOtech tilde operator (= varies with = similar to)[18] cong ? U+2245 (8773) HTML 4.0 HTMLsymbol ISOtech congruent to asymp ? U+2248 (8776) HTML 4.0 HTMLsymbol ISOamsr almost equal to (= asymptotic to) ne ? U+2260 (8800) HTML 4.0 HTMLsymbol ISOtech not equal to equiv ? U+2261 (8801) HTML 4.0 HTMLsymbol ISOtech identical to; sometimes used for 'equivalent to' le ? U+2264 (8804) HTML 4.0 HTMLsymbol ISOtech less-than or equal to ge ? U+2265 (8805) HTML 4.0 HTMLsymbol ISOtech greater-than or equal to sub ? U+2282 (8834) HTML 4.0 HTMLsymbol ISOtech subset of sup ? U+2283 (8835) HTML 4.0 HTMLsymbol ISOtech superset of[19] nsub ? U+2284 (8836) HTML 4.0 HTMLsymbol ISOamsn not a subset of sube ? U+2286 (8838) HTML 4.0 HTMLsymbol ISOtech subset of or equal to supe ? U+2287 (8839) HTML 4.0 HTMLsymbol ISOtech superset of or equal to oplus ? U+2295 (8853) HTML 4.0 HTMLsymbol ISOamsb circled plus (= direct sum) otimes ? U+2297 (8855) HTML 4.0 HTMLsymbol ISOamsb circled times (= vector product) perp ? U+22A5 (8869) HTML 4.0 HTMLsymbol ISOtech up tack (= orthogonal to = perpendicular) sdot ? U+22C5 (8901) HTML 4.0 HTMLsymbol ISOamsb dot operator[20] lceil ? U+2308 (8968) HTML 4.0 HTMLsymbol ISOamsc left ceiling (= APL upstile) rceil ? U+2309 (8969) HTML 4.0 HTMLsymbol ISOamsc right ceiling lfloor ? U+230A (8970) HTML 4.0 HTMLsymbol ISOamsc left floor (= APL downstile) rfloor ? U+230B (8971) HTML 4.0 HTMLsymbol ISOamsc right floor lang ? U+2329 (9001) HTML 4.0 HTMLsymbol ISOtech left-pointing angle bracket (= bra)[21] rang ? U+232A (9002) HTML 4.0 HTMLsymbol ISOtech right-pointing angle bracket (= ket)[22] loz ? U+25CA (9674) HTML 4.0 HTMLsymbol ISOpub lozenge spades ? U+2660 (9824) HTML 4.0 HTMLsymbol ISOpub black spade suit[23] clubs ? U+2663 (9827) HTML 4.0 HTMLsymbol ISOpub black club suit (= shamrock)[24] hearts ? U+2665 (9829) HTML 4.0 HTMLsymbol ISOpub black heart suit (= valentine)[25] diams ? U+2666 (9830) HTML 4.0 HTMLsymbol ISOpub black diamond suit[26]

Notes:

* ^ DTD: the full public DTD name (where the character entity name is defined) is actually ma

HTMLlat1

maps to:

- o PUBLIC "-//W3C//ENTITIES Latin 1//EN//HTML" in HTML (the DTD is implicitly defined
- o PUBLIC "-//W3C//ENTITIES Latin 1 for XHTML//EN" "<http://www.w3.org/TR/xhtml1/DTD/x>

HTMLsymbol

maps to:

- o PUBLIC "-//W3C//ENTITIES Symbols//EN//HTML" in HTML (the DTD is implicitly defined
- o PUBLIC "-//W3C//ENTITIES Symbols for XHTML//EN" "<http://www.w3.org/TR/xhtml1/DTD/x>

HTMLspecial

maps to:

- o PUBLIC "-//W3C//ENTITIES Special//EN//HTML" in HTML (the DTD is implicitly defined
- o PUBLIC "-//W3C//ENTITIES Special for XHTML//EN" "<http://www.w3.org/TR/xhtml1/DTD/x>

* ^ Old ISO subset: these are old (documented) character subsets used in legacy encodings bef
* ^ Description: the standard ISO 10646 and Unicode character name is displayed first for eac
* ^ spaces: a blue background has been used in order to display each space's width.
* ^ ISO proposed: these characters have been standardized in ISO 10646 after the release of H
* ^ ligature: this is a standard misnomer as this is a separate character in some languages.
* ^ black: here it seems to mean filled as opposed to hollow.
* ^ alefsym: 'alef symbol' is NOT the same as U+05D0 'Hebrew letter alef', although the same
* ^ lArr: ISO 10646 does not say that 'leftwards double arrow' is the same as the 'is implied
* ^ rArr: ISO 10646 does not say that 'rightwards double arrow' is the 'implies' character bu
* ^ prod: 'n-ary product' is NOT the same character as U+03A0 'Greek capital letter Pi' thoug
* ^ sum: 'n-ary summation' is NOT the same character as U+03A3 'Greek capital letter Sigma' t

Differences_Between_XHTML_and_HTML

- * ^ sim: 'tilde operator' is NOT the same character as U+007E 'tilde', although the same glyph
- * ^ sup: note that nsup, U+2283 'not a superset of', is not covered by the Symbol font encoding
- * ^ sdot: 'dot operator' is NOT the same character as U+00B7 'middle dot'.
- * ^ lang: 'left-pointing angle bracket' is NOT the same character as U+003C 'less than' or U+003E 'greater than'
- * ^ rang: 'right-pointing angle bracket' is NOT the same character as U+003E 'greater than' or U+003C 'less than'

[edit] Entities representing special characters in XHTML

The XHTML DTDs explicitly declare 253 entities (including the 5 predefined entities of XML 1.0) whose expansion is a single character, which can therefore be informally referred to as "character entities". These (with the exception of the `'` entity) have the same names and represent the same characters as the 252 character entities in HTML. Also, by virtue of being XML, XHTML documents may reference the predefined `'` entity, which is not one of the 252 character entities in HTML. Additional entities of any size may be defined on a per-document basis. However, the usability of entity references in XHTML is affected by how the document is being processed:

- * If the document is read by a conforming HTML processor, then only the 252 HTML character entities are recognized.
- * If the document is read by an XML parser that does not or cannot read external entities, then the five built-in XML entities are recognized.
- * If the document is read by an XML parser that does read external entities, then the five built-in XML entities and any external entities are recognized.

Because of the special `'` case mentioned above, only `"`, `&`, `<`, and `>` will work in all processing situations.