

Les entités du HTML

*Introduction — Les 3 lettres réservées — Lettres avec accents ou cédilles
Ligatures — Ponctuation — Modificateurs
Symboles de calcul élémentaire — Mathématiques spécialisées
Caractères grecs — Caractères monétaires ou juridiques
Flèches — Lettres à caractère graphique — Autres*

Introduction

Le menu ci-dessus est un classement de la liste des entités HTML proposées par le consortium W3C dans [1], qu'on trouve aussi partiellement traduite en français dans [2].

Dans ces références, les caractères «spéciaux» sont donnés par leurs *entités* et par leurs *numéros unicodes*. Les *entités* sont des mots mnémotechniques qui s'emploient entre un & initial et un point-virgule final. Par exemple, l'entité pour «é» est «eacute» (*acute* après un «e» pour dire que ce «e» est modifié par un accent aigu), et on code **é** pour obtenir «é».

Les *numéros unicodes* sont les numéros de ces caractères dans les polices «unicode». Comme il s'agit d'une numérotation sur 2 octets, ces numéros peuvent aller jusqu'à 65000 (en décimal). On peut les employer pour former des caractères en les plaçant entre un &# initial et un point virgule final, ou bien, sous forme hexadécimale, entre un &#x et un point-virgule. Par exemple, un «à» (de numéro 224, ou e0 en hexa) s'obtient par **à** ou **à**.

Les entités ne sont pas toujours reconnues par les différents navigateurs, ou bien elles ne sont pas toujours correctement rendues. Par exemple, IE5-mac restitue beaucoup de caractères diacritiques «exotiques» comme le *scaron* «š» en plaçant le modificateur à côté de la lettre, au lieu de le mettre par dessus). Par la suite, nous utilisons des couleurs pour signaler divers problèmes :

- nous avons mis en rouge deux entités reconnues par Mozilla sur Linux, mais pas sur MacOS (ni par IE5 ou IE6, MacOS ou Windows).
- nous avons mis en magenta sombre les entités reconnues par Mozilla, mais pas (ou mal) par IE5-mac.
- nous avons mis en magenta intense les entités non reconnues par IE6 (Windows). A quelques exceptions près qui seront signalées le moment venu, les entités non reconnues par IE6 sont reconnues par IE5-mac.

Les trois lettres réservées du HTML

Rappelons que les 3 caractères <, > et & servent à former différents tags du HTML et ne peuvent pas apparaître en tant que telles à l'écran. Il faut les écrire avec leurs entités :

- **&** pour &
- **>** pour >
- **<** pour <

Lettres accentuées et cédilles

Le tableau ci-dessous rassemble les lettres accentuées et leurs entités. Nous n'avons pas rappelé les numéros unicode parce que la plupart de ces lettres (du moins, celles nécessaires au français) s'obtiennent directement à partir du clavier dans tout éditeur fonctionnant sous le charset 8859-1 ou 8859-15.

à	è	ì	ò	ù		À	È	Ì	Ò	Ù
agrave	egrave	igrave	ograve	ugrave		Agrave	Egrave	Igrave	Ograve	Igrave
á	é	í	ó	ú	ý	Á	É	Í	Ó	Ú
aacute	eacute	iacute	oacute	uacute	yacute	Aacute	Eacute	Iacute	Oacute	Uacute
â	ê	î	ô	û		Â	Ê	Î	Ô	Û
acirc	ecirc	icirc	ocirc	ucirc		Acirc	Ecirc	Icirc	Ocirc	Ucirc
ä	ë	ï	ö	ü	ÿ	Ä	Ë	Ï	Ö	Ü
auml	euml	iuml	ouml	uuml	yuml	Auml	Euml	Iuml	Ouml	Iuml
ã	õ			ñ	ÿ	Ã	Õ			Ñ
atilde	otilde			ntilde	Yuml	Atilde	Otilde			Ntilde
å	ø			š	ç	Å	Ø		Š	Ç
aring	oslash			šcaron	ccedil	Aring	Oslash		Šcaron	Ccedil

Caractères grecs

Le tableau ci-après rassemble les caractères grecs. Juste en dessous du caractère sur fond blanc, on lit l'entité correspondante, puis le numéro unicode décimal. Rappelons qu'on peut utiliser l'entité ou le numéro pour obtenir le caractère. Par exemple un alpha peut être codé au moyen de son entité (**α**) ou de son numéro unicode (**α**).

A	B	Γ	Δ	E	Z	H	Θ	I	K	Λ	M
Alpha	Beta	Gamma	Delta	Epsilon	Zeta	Eta	Theta	Iota	Kappa	Lambda	Mu
913	914	915	916	917	918	919	920	921	922	923	924
N	Ξ	O	Π	P	Σ	T	Υ	Φ	X	Ψ	Ω
Nu	Xi	Omicron	Pi	Rho	Sigma	Tau	Upsilon	Phi	Chi	Psi	Omega
925	926	927	928	929	931	932	933	934	935	936	937
α	β	γ	δ	ε	ζ	η	θ	ι	κ	λ	μ
alpha	beta	gamma	delta	epsilon	zeta	eta	theta	iota	kappa	lambda	mu
945	946	947	948	949	950	951	952	953	954	955	956
ν	ξ	ο	π	ρ	σ	τ	υ	φ	χ	ψ	ω
nu	xi	omicron	pi	rho	sigma	tau	upsilon	phi	chi	psi	omega
957	958	959	960	961	963	964	965	966	967	968	969

Il y a en plus une lettre ç (entité sigmafin ou code & #962;) pour les sigma en fin de mots, qui n'existe pas en majuscules.

Caractères monétaires ou juridiques

Le tableau ci-dessous donne le rendu de divers symboles, monétaires ou à caractère légal, avec les entités et numéros unicode correspondants et quelques commentaires.

Bien entendu, le dollar se trouve toujours directement sur le clavier (:-))

Caractère	Entité HTML	Numéro unicode	Description
€	euro	8364	euro
£	pound	163	livre sterling
¥	yen	165	yen
f	fnof	402	florin

¢	cent	162	cent US
¤	curren	164	symbole général de monnaie (euh ???)
©	copy	169	copyright
®	reg	174	enregistré (droits réservés)
™	trade	8482	Signe «marque déposée»

Modificateurs de caractères

Ces symboles correspondent à quelques uns des divers «accents» qu'on peut superposer aux lettres. Dans des langages typographiques comme TeX ou LaTeX, ces modificateurs peuvent s'appliquer à n'importe quelle lettre, c.à.d. qu'ils s'impriment par dessus le caractère modifié. Il semble qu'on ait prévu un mécanisme du même genre pour les polices unicode, le modificateur venant après le caractère modifié (c.à.d. un code comme `e´`; devant donner un é), mais ça ne fonctionne manifestement pas en cette mi-2003.

On notera que cette liste n'est pas cohérente avec les caractères disponibles. Par exemple, on n'y trouve ni accent grave, ni le petit rond de å. Inversement, il n'y a aucun code qui donne des lettres avec un «macron». Toutefois, tous ces caractères «manquants» sont accessibles par leurs numéros unicode dans la plage 300–36f (en hexa).

Caractère	Entité HTML	Numéro unicode	Description
¨	uml	168	tréma
´	acute	180	accent aigu
^	circ	710	accent circonflexe
~	tilde	732	petit tilda
—	macr	175	macron (en général, indique une voyelle allongée ou accentuée)
¸	cedil	184	cédille
—	oline	8254	surligné

Symboles de calcul élémentaires

Caractère	Entité HTML	Numéro unicode	Description
°	deg	176	degré
¹	sup1	185	puissance un
²	sup2	178	carré
³	sup3	179	cube
¼	frac14	188	fraction un-quart
½	frac12	189	fraction un-demi
¾	frac34	190	fraction trois-quarts
×	times	215	multiplié par
·	middot	183	point de produit (à mi-hauteur)
·	sdot	8901	point de produit (en principe légèrement plus gros que ·)
÷	divide	247	divisé par
/	frasl	8260	barre de fraction (plus petit que le slash / du clavier)
–	minus	8722	moins (quelle différence avec le signe - du clavier ???)

±	plusmn	177	plus-ou-moins
~	sim	8764	varie comme, environ (différent du tilda ~ &#126; ;))
≈	cong	8773	approximativement égal à
≈	asymp	8776	presque égal à
∝	prop	8733	proportionnel à
≠	ne	8800	différent de
≡	equiv	8801	identique à
≤	le	8804	inférieur ou égal à
≥	ge	8805	supérieur ou égal à
<	lt	60	plus petit que
>	gt	62	plus grand que
‰	permil	8240	pour mille (le % est au clavier!)
⊥	perp	8869	orthogonal
μ	micro	181	micron
'	prime	8242	minutes d'angle, pieds
"	Prime	8243	secondes d'angle, pouces

Symboles mathématiques spécialisés

Caractère	Entité HTML	Numéro unicode	Description
∀	forall	8704	quantificateur <i>pour tout</i>
∃	exist	8707	quantificateur <i>il existe</i>
∅	empty	8709	ensemble vide, nul, diamètre
∈	isin	8712	'élément de'
∉	notin	8713	non élément de
⊃	ni	8715	contient comme membre
∩	cap	8745	intersection
∪	cup	8746	union
⊂	sub	8834	sous-ensemble de
⊃	sup	8835	sur-ensemble de (notez : 'nsup', «non sur-ensemble de», n'existe pas)
⊄	nsup	8836	non sous-ensemble de
⊆	sube	8838	sous-ensemble ou identité
⊇	supe	8839	sur-ensemble ou identité
∂	part	8706	différentielle partielle
∇	nabla	8711	nabla
∏	prod	8719	n-produit, symbole «produit» différent du Pi majuscule, bien que le même glyphe puisse être utilisé
∑	sum	8721	n-somme différent du Sigma majuscule, bien que le même glyphe puisse être utilisé
*	lowast	8727	opérateur astérique
√	radic	8730	racine carrée, signe «radical»
∞	infin	8734	infini
∠	ang	8736	angle
∧	and	8743	ET logique

∨	or	8744	OU logique
∫	int	8747	intégrale
∴	there4	8756	implique
⊕	oplus	8853	plus encerclé, somme directe
⊗	otimes	8855	produit encerclé, produit vectoriel
∅	weierp	8472	ensemble de Weierstrass
ℑ	image	8465	partie imaginaire
℔	real	8476	partie réelle
ℵ	alefsym	8501	alef, cardinal infini. Ce n'est PAS la lettre hébraïque alef, (u+05D0) bien que le même glyphe puisse être utilisé pour les deux caractères
⟨	lang	9001	bra (plus grand que les signes < ou < (< ou ‹))
⟩	rang	9002	ket (plus grand que les signes > ou > (> ou ›))

Flèches

Le tableau ci-après montre les flèches qu'on peut obtenir avec les entités. On peut en obtenir quelques autres au moyen des numéros unicode, mais, même en unicode, les flèches en diagonale montante sont difficiles à trouver (:-))

Caractère	Entité HTML	Numéro unicode	Description
←	larr	8592	vers la gauche
↑	uarr	8593	vers le haut
→	rarr	8594	vers la droite
↓	darr	8595	vers le bas
↔	harr	8596	vers la gauche et la droite (bijection)
↵	crarr	8629	vers le bas, puis la gauche (retour chariot)
⇐	lArr	8656	double flèche gauche (impliqué par)
⇑	uArr	8657	double flèche montante
⇒	rArr	8658	double flèche droite (implique)
⇓	dArr	8659	double flèche descendante
⇔	hArr	8660	double flèche gauche-droite (équivalence)

Formes à caractère graphique

Le tableau ci-après montre les formes qu'on peut obtenir avec des entités. On peut en obtenir bien d'autres... si le navigateur le veut bien, au moyen des numéros unicode, par exemple toute une rangée de caractères «Dingbat» (Mozilla veut bien, mais ni IE5-mac, ni IE6...)

Caractère	Entité HTML	Numéro unicode	Description
¶	para	182	paragraphe
†	dagger	8224	daguer
‡	Dagger	8225	double daguer
§	sect	167	section
&	amp	38	esperluète
•	bull	8226	petit cercle noir

◇	loz	9674	losange
♠	spades	9824	pique
♣	clubs	9827	trèfle noir
♥	hearts	9829	cœur
♦	diams	9830	carreau
Γ	lceil	8968	euh...
⌋	rceil	8969	euh...
L	lfloor	8970	euh...
J	rfloor	8971	euh...
<	lang	9001	grande parenthèse anguleuse ouvrante
>	rang	9002	grande parenthèse anguleuse fermante

Les entités lceil, rceil, lfloor et rfloor ne sont reconnues ni par IE5-mac, ni par IE6-windows

Ligatures

Le tableau ci-après montre les quelques ligatures qu'on peut obtenir avec les entités. On peut en obtenir quelques autres au moyen des [numéros unicode](#) (enfin, mieux sous MacOS que sous Windows...)

Caractère	Entité HTML	Numéro unicode	Description
Æ	AElig	198	ligature A-E
æ	aelig	230	ligature a-e
Œ	OElig	338	ligature O-E
œ	oelig	339	ligature o-e
ß	szlig	223	ligature sz (allemand)

Ponctuation

Caractère	Entité HTML	Numéro unicode	Description
«	laquo	171	guillemet français ouvrant
»	raquo	187	guillemet français fermant
“	ldquo	8220	guillemet anglais ouvrant
”	rdquo	8221	guillemet anglais fermant
„	bdquo	8222	guillemet anglais bas (pour citation)
<	lsaquo	8249	petite parenthèse angulaire ouvrante
>	rsaquo	8250	petite parenthèse angulaire fermante (noter que › et › ne font pas encore officiellement partie du standard ISO)
"	quot	34	double quote
'	lsquo	8216	apostrophe ouvrante
'	rsquo	8217	apostrophe fermante
,	sbquo	8218	apostrophe basse
-			Signe - du clavier, ajouté pour que l'on puisse comparer avec les deux tirets ci-dessous
-	ndash	8211	tiret intermédiaire (portée de nombres)
—	mdash	8212	tiret long (dialogue)
...	hellip	8230	trois points

Les derniers éléments de ce tableau sont apparemment tous des espaces, mais à part le premier () bien connu, aucun ne fonctionne correctement : selon le navigateur, soit ils sont rendus comme , soit ils sont rendus avec une largeur nulle, comme s'ils n'existaient pas (à titre d'essai, dans les lignes ci-dessous, nous avons placé ces entités entre la légende et un point-virgule); ou encore ils ne sont pas reconnus du tout (IE6, lignes en magenta vif).

	nbsp	160	espace insécable ;
	ensp	8194	espace ;
	emsp	8195	espace ;
	thinsp	8201	espace ;
	zwnj	8204	zero width non-joiner;
	lrm	8206	left-to-right mark (???)
	rlm	8207	right-to-left mark (???)

Divers

Caractère	Entité HTML	Numéro unicode	Description
¡	iexcl	161	signe d'exclamation inversé
¿	quest	191	signe d'interrogation inversé
¦	brvbar	166	barre brisée (verticale)
§	sect	167	section
†	dagger	8224	dague
‡	Dagger	8225	double dague
ª	ordf	170	indicateur de genre féminin
º	ordm	186	indicateur de genre masculin
¬	not	172	non
	shy	173	hyphenation Ce caractère à l'intérieur d'un mot est une proposition de césure pour le logiciel de mise en page. Il n'est rendu par un tiret de césure que si le mot est coupé (à notre connaissance, aucun navigateur ne sait pratiquer de telles coupures).
Ð	ETH	208	Eth, islandais
Þ	THORN	222	THORN, islandais
ð	eth	240	eth, islandais
þ	thorn	254	thorn, islandais

Charles — Gérald

[Page précédente](#) — [Accueil des réponses](#)