

Index

Symboles

- æ (U+00E6), 111, 360
' (apostrophe), 132
@ (arrobe), 134
* (astérisque), 133
/ (barre de fraction), 154
/ (barre oblique), 149
˘ (brève), 125
'\0' (caractère nul en C), 287
(croisillon), 132
⦿ (croix pointée), 154
° (degré), 136, 158
∂ (dérivée partielle), 159
(dièse), 132
\$ (dollar), 159
œ (e dans o), 41, 254, 279
€ (euro), 43, 45, 260
" (guillemet anglais), 132
' (guillemet-apostrophe), 133
» (guillemet chevron), 140
☞ (index vers la droite), 154
^a (indicateur ordinal féminin), 136
k (k barré), 64
λ lambda d'insertion (λ), 135
£ (livre), 159
̄ (macron), 129
μ (micro), 100
^w (modificateur), 124
– (moins), 152
∇ (nabla), 159
№ (numéro), 158
† (obèle), 134
§ (paragraphe), 135
& (perluète), 132
ϖ (pi dorique), 159
¶ (pied de mouche), 135
· (point), 131
• (point de coupure), 138
... (points de suspension), 151
· (point médian), 137, 154
• (puce), 138, 154
⋈ (quatre-points en croix), 154
¶ (révéler format), 147
ŀ (s long), 169
ß (s dur), 64, 92
□ (symbole monétaire), 39, 316
~ (tilde), 135
_ (tiret bas), 153, 197
— (tiret cadratin), 152
– (tiret demi-cadratin), 152
- (trait d'union), 152
- (trait d'union conditionnel), 138
- (trait d'union insécable), 139, 152
- (trait d'union-signe moins), 152

⌘ (transfini), 158

¨ (tréma), 128

ÿ, 127

A

accept-charset

sur balise form, 278

Accept-Charset

entête HTTP, 220

Accept-Encoding, 220

Accept-Language, 212, 220, 247

adresse IP, 220

afrikaans, 6, 92

Afrique, 124

ajouré, 158, 160

align, 248

allemand, 141

s dur, 92

alphabet phonétique international, 8, 123, 124

analyse contextuelle, 360

anglet, 142

Angström, 359

annotation interlinéaire, 166

ANSI, 37

jeu de caractères, 44

antiliant, 360

Apache, 207

API Voir alphabet phonétique international

apostrophe, 132

appel

de caractère, 254, 256, 257, 259

d'entité, 254, 256

arabe, 26, 71, 263, 362, 384

arobase Voir arrobe

arrobe, 134

ASCII, 37

astérisque, 133

attribut

align, 248

dir, 248

en XML, 251

lang, 245

xml:id, 253

xml:lang, 247, 252

azéri, 264

B

BabelMap, 31, 86, 95

balise, 273

<bdo>, 274

<body>, 248

, 86, 248

<div>, 248

, 365

<html>, 248

<link>, 248

<meta>, 204, 206, 207, 209

barre de fraction, 154

Base64, 195

BCP 47, 184, 245

bicaméral, 91

bidi Voir bidirectionnel

réflexion, 96, 266

bidirectionnel

algorithme, 263

caractère de commande, 273

feuilles de styles CSS, 273

passage, 271

réflexion, 70

XHTML, 274

BOM Voir indicateur d'ordre des octets

boutisme, 84, 85, 115, 170

brahmi, 63, 363

breton, 9, 64

C

C, 287

UTF-16, 287

C#, 287, 291

C++, 287

cadratin, 142, 145

caractère, 8, 64

abstrait, 64, 65

codage de, 200

codé, 37, 65

combinatoire, 65, 66

de base, 66, 72

de commande, 163

bidi, 273

ensemble C0, 259

ensemble C1, 40, 44, 164, 259

déconseillé, 169

- désuet, 169
 - en C#, 291
 - en C et C++, 287
 - en exposant, 263
 - en Java, 289
 - forme de sérialisation, 85
 - forme en mémoire, 81
 - propriété de, 86
 - spécial, 166
 - technique, 163
 - caron, 128
 - casse, 91, 305
 - correspondance, 92
 - de titre, 91
 - catalan, 137
 - catégorie générale, 88
 - ccTLD, 181, 221
 - Celsius, 158
 - CFE, 375
 - charset
 - HTTP, 42, 44
 - chasse, 11
 - nulle, 65
 - chiffre
 - arabe, 267
 - arabo-hindi, 95, 267
 - espace de même chasse, 145
 - européen, 267
 - hébreu, 95
 - séparateur, 144
 - chinois, 8, 13, 51, 62, 125, 172, 284, 286, 321
 - CJC, 13, 47, 62
 - ponctuation, 142
 - classe combinatoire canonique, 94
 - clavier, 25
 - virtuel, 26
 - CLDR, 113, 339
 - clone
 - caractère ASCII, 172
 - diacritique, 125, 134
 - cmap, 372
 - codage
 - centrifuge, 126
 - de caractères, 36
 - modal, 47
 - code page, 44
 - codet, 36
 - non attribué, 168
 - conformité, 113
 - conjointe, 363
 - Content-Encoding, 220
 - Content-Language, 220, 247
 - Content-Transfer-Encoding, 220
 - Content-Type, 220
 - contexte directionnel, 266
 - contraction, 111, 165
 - convertibilité, 72, 75
 - copte, 125
 - coréen
 - hangûl, 13
 - hanja, 13
 - corps Voir force
 - correspondance de casse, 92
 - coupure de lignes, 96, 139, 155, 367
 - en allemand, 138
 - en thaï, 98
 - coupure de mot, 133, 139
 - en allemand, 100, 138, 139
 - en anglais, 100, 138
 - en français, 100
 - en néerlandais, 139
 - en suédois, 139
 - courbe de Bézier, 373
 - CP 1252, 44
 - crénage, 376
 - croate, 91
 - croisillon, 132
 - CSS, 20, 243, 249
 - bidi, 273
 - polices, 358
 - pseudoclasse lang, 250
 - sélecteur, 249
- ## D
- danois, 9
 - date, 313
 - déclaration XML, 205
 - décomposition
 - canonique, 101
 - de compatibilité, 101
 - sans tête, 105
 - degré, 136
 - demi-cadratin, 142

demi-chasse, 46
demi-forme, 363
désunification, 158
dévanâgarî, 66, 72, 363
diacritique, 65, 66, 127
 clone, 125, 134
 couvrant, 129
 invisible bloquant, 165
DIB, 165
dièse, 132
dir Voir attribut dir
directionnalité, 95, 140, 158, 248
 faible, 267
 forte, 265
 globale, 267
 neutre, 266, 269
DNS, 220
<!DOCTYPE, 240
dollar, 159
domaine de tête, 220
 de code de pays, 221
 générique, 221
DTD, 240
DUCET (table), 110

E

e dans o, 41
EBCDIC, 46
Eclipse
 afficher de l'Unicode, 291
écriture, 7, 383
 arabe, 384
 chinoise, 13
 indicatif, 181
efficacité, 65, 69
élément de code, 36
empatement, 11
encoded-word, 197
entête HTTP
 Accept-Charset, 220
 Accept-Encoding, 220
 Accept-Language, 212, 220, 247
 Content-Encoding, 220
 Content-Language, 220
 Content-Transfer-Encoding, 220
 Content-Type, 220
 MIME-Version, 220

Subject, 220
Transfer-Encoding, 220
entête MIME, 189
 visualiser les, 189
entité
 &, 257
 ', 257
 €, 260
 >, 257
 <, 257
 , 148
 ", 257
 ‏, 257
  , 148
ESMTP, 189
espace, 142
 différents caractères, 144
 en français, 149
 fine, 144, 146
 idéographique, 143
 insécable, 41, 98, 143
 mot, 143
 sans chasse, 149
 sécable, 98
esperluette, 132
esprit
 doux, 126
 rude, 126
eszet Voir s dur
étiquette linguistique, 172
expansion, 111
expressions régulière, 295

F

foisonnement, 347
force, 11, 370
forme
 de sérialisation, 85
 en mémoire, 81
 normalisée, 101
formulaire HTML, 275
fractions, 263
fraktur Voir gothique
frontière de texte, 89
furigana, 167

G

GB 18030, 51
 gluon de mots, 146, 165, 262, 364
 glyf, 373
 glyphe, 11, 70, 369
 gothique, 12, 160
 GPOS, 126, 376, 382
 grand-boutiste, 170
 graphème, 9, 64
 grec, 125
 iota souscrit, 92
 gros-boutiste Voir boutisme
 GSUB, 126, 375, 382
 gTLD, 221, 225
 guillemet, 140
 anglais, 132
 chevron, 140
 d'Extrême-Orient, 142

H

haček, 128
 halant, 363
 hameçonnage, 226
 hangûl, 13
 hébreu, 71, 166
 heure, 313
 hinting Voir nuancement
 hiragana, 13, 46
 hirik, 166
 hmtx, 375
 homoglyphe, 227
 .htaccess, 209
 HTML, 167, 206, 245, 259
 espace fine insécable, 148
 formulaire, 275
 HTTP, 42, 44, 200, 206, 220, 276
 get, 276
 post, 276
 httpd.conf, 207

I

IANA, 183
 ICU, 292
 transformation, 299
 identificateurs de suites Unicode nommés, 127
 idéogramme, 7, 13, 62, 252

idéophonogramme, 7
 IDN, 224
 IETF, 182
 IIS, 208
 indicateur d'ordre des octets, 115, 147, 165,
 168, 170, 171, 205
 indicatif
 écriture, 181
 langue, 178
 pays, 180
 insécabilité, 124
 intégrale, 160
 internationalisation, 32, 284
 date, 312, 313
 extraction des données, 307
 heure, 313
 image, 349
 messages variables, 319
 montants, 316
 schéma XML, 251
 URI, 222

Internet Engineering Task Force Voir IETF
 interpunctum, 138

iota
 adscriit, 92
 souscrit, 92

IRI, 222, 229
 ISO 639, 178
 ISO 646, 39
 ISO 2022, 47
 ISO 3166, 180
 ISO 8859, 42
 ISO-8859-1, 204
 ISO 10646, 49
 ISO 14651, 109
 ISO 15924, 181
 italien, 141
 italique, 12, 160

J

J#, 292
 japonais, 13
 Java, 289
 échappement, 289
 message d'exception, 340
 propriété de caractère, 294
 UCS-2, 289

jawi, 6
 jeu de caractères, 36
 codés, 80
 JSP, 209
 justification, 369

K

katakana, 13, 46
 k barré, 64
 Kelvin, 157
 khmer, 149, 169, 262, 380
 KOI8-R, 47
 krakaziábry, 6

L

lang
 attribut, 245
 attribut HTML, 247
 pseudoclasse lang, 250
 langue, 245, 383
 rémanence du choix , 217
 lao, 98
 latin-1, 41
 lettres et nombres cerclés, 263
 lettres modificatives, 124
 liaison cursive, 362
 liant, 360
 ligature, 359
 lam alif, 264, 360, 376, 384
 touareg, 380
 lituanien, 128
 livre, 159
 locale, 303
 avec ICU, 303
 implicite, 305
 POSIX, 232
 sur .NET, 304
 localisation, 32, 253, 286
 LTag, 172

M

macron, 129
 malais, 6
 mathématique, 159
 MathML, 160

mécanisme de sérialisation, 84
 médiévales, lettres latines, 63
 meta, 204, 247 Voir balise <meta>
 méthode d'entrée, 28
 MIME, 189
 charset, 201
 MIME-Version , 220
 mode bizarreries, 241
 mojibaké, 6
 moment, 37
 mongol, 6
 MS Word
 espace fine insécable, 147
 espacement, 146
 trait d'union conditionnel, 139
 Multiviews, 214
 musique, 161, 262

N

nabla, 159
 nameprep, 224
 navigateur Internet, 44
 NDI, 224
 négociation de langue, 210
 neutralité culturelle, 348
 n'ko, 267
 nom de domaine, 220
 de tête, 181
 nomina sacra, 127
 non-caractère, 167, 168
 normalisation Unicode, 101, 297
 stabilité, 103
 notation, 7, 157
 nuancement, 374
 numéro, 158

O

obèle, 134
 occitan, 137
 octet, 80
 œil, 11, 41, 65, 372
 étranger, 356
 OpenType, 131, 371, 378
 écriture, 383
 langue, 383
 ordinal masculin, 136

ordre des mots, 318
 ordre logique, 66, 71, 264, 387
 ordre visuel, 368
 OTLS, 381

P

page de code, 44
 page HTML Voir page web
 page web, 201, 204, 209, 278
 paragraphe, 135
 parodie, 226
 passages bidi, 271
 patah, 166
 pays, 180
 perluète, 132
 persan, 362
 petit-boutiste Voir boutisme
 Phaïstos, 63
 phishing Voir hameçonnage
 PHP, 209
 pi dorique, 159
 pied-de-mouche, 135
 pinyin, 8, 28
 plan

- complémentaire
 - spécialisé, 60
- idéographique
 - complémentaire, 60
- multilingue
 - complémentaire, 61, 160, 254
 - de base, 56, 168, 169

 PMB Voir Plan multilingue de base
 PMC Voir Plan multilingue complémentaire
 point

- de code, 36
- de conduite, 151
- de suspension, 151
- médian, 137, 154

 polarité de stockage des octets Voir boutisme
 police, 250, 356, 357

- classification, 12
- de glyphes, 371
- de repli, 357
- de substitution, 357
- dite Unicode, 16
- incorporée, 21
- liée, 358

normalisation Unicode des diacritiques,
 378
 OpenType, 372
 TrueType, 372
 ponctuation, 131

- appariée, 140
- archaïque, 154

 position de code, 36
 prime, 133
 profil culturel Voir locale
 prologue XML, 240
 propriété de casse, 91
 protocole

- de haut niveau, 15, 93, 136, 160, 161
- Internet, 188

 pseudoclasse lang, 250
 puce, 138, 154
 punycode, 224

Q

quirks mode Voir mode bizarreries
 quoted-printable, 194, 198

R

réflexion bidi, 70, 96, 266
 rémanence du choix de langue, 217
 remplacement d'objet, 166
 rendu, 378
 réordonnancement, 368
 repérage, 105
 répertoire de caractères, 36

- abstrait, 80

 ResourceBundle, 307
 retour de chariot, 164
 RFC, 182
 RFC 821, 188
 RFC 1123, 221, 224
 RFC 1521, 193
 RFC 1738, 277
 RFC 1766, 184
 RFC 2045, 193
 RFC 2047, 190, 197
 RFC 2616, 200, 206
 RFC 2821, 188
 RFC 2822, 188, 197
 RFC 3066, 184 Voir RFC 4646

RFC 3454, 224
 RFC 3490, 3491 et 3492, 224
 RFC 3986, 222
 RFC 3987, 230
 RFC 4288, 189
 RFC 4646, 172, 181, 182
 RFC 4647, 184
 RFC 4648, 193, 195
 romain, 12
 romanisation, 8
 ronde, 160
 Ruby, 167, 252, 258, 262

S

schéma XML, 251, 252
 s dur, 64, 93
 seiset, 80, 81
 sélecteur CSS, 249, 359
 sélecteur de variante, 356
 sémantique, 70
 séquence équivalente, 73
 signe diacritique Voir diacritique
 signe moins, 134
 singleton, 105
 s long, 169
 slovène, 141
 SMTP, 188
 spéciaux (caractères), 166
 stabilité, 93, 104
 formes de normalisation, 103
 stringprep, 224
 Subject, 220
 suédois, 141
 suite mal formée, 82
 SVG, 245, 350
 symbole, 157
 hamiltonien, 160
 mathématique, 159
 mathématique alphanumérique, 160
 monétaire, 159
 musical, 161

T

table
 CFF, 375
 cmap, 372

glyf, 373
 GPOS, 376, 382
 GSUB, 375, 382
 hmtx, 375
 tchèque, 9
 texte
 brut, 15, 71, 258
 enrichi, 15, 71
 thaï, 98, 149, 286, 380
 tfinagh, 60, 361, 365
 touareg, 26, 264, 380
 tirect, 152
 TLD, 220
 touareg Voir tfinagh touareg
 trait d'union, 134, 152
 conditionnel, 41, 98, 138, 372
 insécable, 372
 signe moins, 372
 tramage, 370
 transcription, 7
 Transfer-Encoding, 220
 transfini, 158
 translittération, 8
 ICU, 299
 tri, 105, 325, 327, 329, 331, 333
 clé de, 109
 élément de, 109
 en Java, 325
 TrueType, 372
 tueur, 363
 turc, 141
 i sans point, 92, 383
 type directionnel Voir directionalité

U

U+00A0, 165
 U+034E, 165
 U+2060, 165
 U+FEFF, 147, 165, 170
 U+FFFE, 114, 168, 170
 U+FFFF, 114
 UAX, 96
 n° 13, 164
 n° 14, 96, 97, 98, 144, 146, 148, 335
 n° 29, 99, 335
 UCS-2, 81, 289
 UCS-4, 81, 85

umlaut, 166, 285
unicaméral, 91
UnicodeSet, 296
unification, 65, 72, 251
 ASCII, 158
Uniscribe, 357, 380
universalité, 68
URI, 222, 229, 276
URL, 149, 230
UTF-8, 82, 171, 204
 sauvegarder en, 206
UTF-16, 81, 85, 170
UTF-32, 83, 85

V

valeur scalaire, 36
vietnamien, 94, 378
 écriture annamite, 13
 quốc ngữ, 88
virâma, 363

W

wakiten, 253

Web, 200
Windows 1252, 44, 204
Windows Latin 1, 44

X

XHTML, 205, 240, 259
 bidi, 274
xml:id, 253
xml:lang, 252, 262
XML, 205, 245
 déclaration, 205
 id, 253
 lang, 252, 262
 normalisation des espaces, 143
 schéma, 251
 version 1.1, 259

Y

y tréma, 41

Z

zone à usage privé, 169
ZUP Voir Zone à usage privé

