

Annexe A (normative)

Collections de caractères graphiques pour les sous-ensembles

A.1 Collections des caractères graphiques codés

Les collections suivantes sont tirées du plan multilingue de base. Un astérisque (*) dans la colonne des positions indique que cette collection est fixe.

<u>Numéro et nom de collection</u>	<u>Positions</u>		
1 LATIN DE BASE	0020 - 007E *	23 KANNARA	0C80 - 0CFF 200C, 200D
2 SUPPLÉMENT LATIN-1	00A0 - 00FF *	24 MALAYALAM	0D00 - 0D7F 200C, 200D
3 LATIN ÉTENDU A	0100 - 017F *	25 THAÏ	0E00 - 0E7F
4 LATIN ÉTENDU B	0180 - 024F	26 LAO	0E80 - 0EFF
5 ALPHABET PHONÉTIQUE	0250 - 02AF	27 GÉORGIEN DE BASE	10D0 - 10FF
6 MODIFICATEURS DE LETTRE	02B0 - 02FF	28 GÉORGIEN ÉTENDU	10A0 - 10CF
7 SIGNES COMBINATOIRES	0300 - 036F	29 JAMOS HANGÛL	1100 - 11FF
8 GREC DE BASE	0370 - 03CF	30 LATIN ÉTENDU ADDITIONNEL	1E00 - 1EFF
9 COPTE & SYMBOLES GRECS	03D0 - 03FF	31 GREC ÉTENDU	1F00 - 1FFF
10 CYRILLIQUE	0400 - 04FF	32 PONCTUATION GÉNÉRALE	2000 - 206F
11 ARMÉNIEN	0530 - 058F	33 EXPOSANTS ET INDICES	2070 - 209F
12 HÉBREU DE BASE	05D0 - 05EA *	34 SYMBOLES MONÉTAIRES	20A0 - 20CF
13 HÉBREU ÉTENDU	0590 - 05CF 05EB - 05FF	35 SIGNES COMBINATOIRES POUR SYMBOLES	20D0 - 20FF
14 ARABE DE BASE	0600 - 065F	36 SYMBOLES DE TYPE LETTRE	2100 - 214F
15 ARABE ÉTENDU	0660 - 06FF	37 FORMES NUMÉRIQUES	2150 - 218F
16 DÉVANÂGARÎ	0900 - 097F 200C, 200D	38 FLÈCHES	2190 - 21FF
17 BENGALI	0980 - 09FF 200C, 200D	39 OPÉRATEURS MATHÉMATIQUES	2200 - 22FF
18 GOURMOUKHÎ	0A00 - 0A7F 200C, 200D	40 SIGNES TECHNIQUES DIVERS	2300 - 23FF
19 GOUDJARATI	0A80 - 0AFF 200C, 200D	41 PICTOGRAMMES DE COMMANDE	2400 - 243F
20 ORIYA	0B00 - 0B7F 200C, 200D	42 RECONNAISSANCE OPTIQUE DE CARACTÈRES	2440 - 245F
21 TAMOUL	0B80 - 0BFF 200C, 200D	43 ALPHANUMÉRIQUES CERCLÉS	2460 - 24FF
22 TÉLOUGOU	0C00 - 0C7F 200C, 200D	44 FILETS	2500 - 257F *
		45 PAVÉS	2580 - 259F
		46 FORMES GÉOMÉTRIQUES	25A0 - 25FF
		47 SYMBOLES DIVERS	2600 - 26FF
		48 CASSEAU	2700 - 27BF
		49 SYMBOLES ET PONCTUATION CJC	3000 - 303F

50 HIRAGANA	3040 - 309F	82 OGAM	1680 - 169F
51 KATAKANA	30A0 - 30FF	83 RUNES	16A0 - 16FF
52 BOPOMOFO	3100 - 312F 31A0 - 31BF	84 SINGHALAIS	0D80 - 0DFF
53 JAMOÛ DE COMPATIBILITÉ HANGÛL	3130 - 318F	85 SYRIAQUE	0700 - 074F
54 DIVERS CJC	3190 - 319F	86 THÂNA	0780 - 07BF
55 LETTRES ET MOIS CJC CERCLÉS	3200 - 32FF	87 BIRMAN DE BASE	1000 - 104F 200C, 200D
56 COMPATIBILITÉ CJC	3300 - 33FF	88 KHMER	1780 - 17FF 200C, 200D
57, 58, 59 (Ces numéros de collection ne doivent pas être utilisés, voir note 2.)		89 MONGOL	1800 - 18AF
60 IDÉOGRAMMES UNIFIÉS CJC	4E00 - 9FFF	90 BIRMAN ÉTENDU	1050 - 109F
61 ZONE À USAGE PRIVÉ	E000 - F8FF	91 TIBÉTAIN	0F00 - 0FFF
62 IDÉOGRAMMES DE COMPATIBILITÉ CJC	F900 - FAFF	Les collections suivantes rassemblent des caractères utilisés pour des formats optionnels ou des formats particuliers à un système d'écriture. Se reporter l'annexe F pour de plus amples infor- mations.	
63 FORMES DE PRÉSENTATION ALPHABÉTIQUES	FB00 - FB4F	200 INDICATEURS DE FRON- TIÈRE À CHASSE NULLE	200B - 200D FEFF
64 FORMES DE PRÉSENTATIONS ARABES A	FB50 - FDFE	201 SÉPARATEURS DE FORMATAGE	2028 - 2029
65 DEMI-SIGNES COMBINATOIRES	FE20 - FE2F	202 SIGNES DE FORMATAGE BIDIRECTIONNEL	200E - 200F
66 FORMES DE COMPATIBILITÉ CJC	FE30 - FE4F	203 ENCHÂSSEMENTS DE FOR- MATAGE BIDIRECTIONNEL	202A - 202E
67 PETITES VARIANTES DE FORMES	FE50 - FE6F	204 BOURRES HANGÛL	3164, FFA0
68 FORMES DE PRÉSENTATIONS ARABES B	FE70 - FEFE	205 SÉLECTEURS DE FORMAGE DE CARACTÈRES	206A - 206D
69 FORMES DE DEMI ET PLEINE CHASSE	FF00 - FFEF	206 SÉLECTEURS DE FORME NUMÉRIQUE	206E - 206F
70 CARACTÈRES SPÉCIAUX	FFF0 - FFFD	207 CARACTÈRES DE DESCRIPTION IDÉOGRAPHIQUE	2FF0 - 2FFF
71 SYLLABAIRE HANGÛL	AC00 - D7A3 *	Les numéros suivants désignent des collections qui constituent l'union de collections déclarées ci- dessus.	
72 TIBÉTAIN DE BASE	0F00 - 0FBF	250 CARACTÈRES DE FORMATAGE GÉNÉRAUX	Collections 200 - 203
73 ÉTHIOPIEN	1200 - 137F	251 CARACTÈRES DE FORMATAGE PROPRES À UNE ÉCRITURE	Collections 204 - 207
74 SYLLABAIRES AUTOCHTONES CANADIENS	1400 - 167F	Ci-dessous d'autres collections.	
75 CHÉROKÎ	13A0 - 13FF	270 CARACTÈRES COMBINATOIRES OU DIACRITIQUES	Caractères répertoriés dans l'annexe B.1
76 SYLLABAIRE YI	A000 - A48F		
77 CLÉS YI	A490 - A4CF		
78 CLÉS CHINOISES (K'ANG-HSI)	2F00 - 2FDF		
79 CLÉS CJC SUPPLÉMENTAIRES	2E80 - 2EFF		
80 COMBINAISONS BRAILLE	2800 - 28FF		
81 EXTENSION A AUX IDÉO- GRAMMES UNIFIÉS CJC	3400 - 4DBF FA1F, FA23		

271	CARACTÈRES COMBINATOIRES OU DIACRITIQUES B Caractères répertoriés dans l'annexe B.2		Copte	9
299	(Ce numéro de collection ne doit pas être utilisé, voir A.3.2.)		Cyrillique	10
300	PMB	0000 - D7FF E000 - FFFD	Demi chasse	69
301	PMB-AMD.7	voir A.3.1 *	Demi-signé	65
302	PMB DEUXIÈME ÉDITION	voir A.3.3 *	Dévanâgarî	16
			Devise	34
			Diacritique	7, 35, 65, 270, 271
			Enchâssement	203
			Éthiopien	73
			Exposant	33
			Filet	44
			Flèche	38
			Forme	37, 46, 63, 64, 67, 68, 205, 206
			Formatage	201, 202, 203, 250, 251
			Frontière	200
			Géométrie (forme)	46
			Géorgien	27, 28
400	PLANS À USAGE PRIVÉ G=00, P=0F, 10, & E0 - FF		Goudjarati	19
500	GROUPES À USAGE PRIVÉ G=60 - 7F		Gourmoukhî	18
			Grec	8, 9, 31
			Hangûl	29, 53, 71, 204
			Hébreu	12, 13
			Hiragana	50
			Idéogramme	60, 62, 81, 207
			Indicateur	200
			Indice	33
			Jamo	28, 53
			K'ang-hsi	78
			Kangxi	78
			Kannara	23
			Katakana	51
			Khmer	88
			Lao	26
			Latin	1, 2, 3, 4, 30
			Lettre	36, 55
			Malayalam	24
			Mathématique	39
			Modificateur	6
			Mois	55
			Mongol	89
			Monnaie	34
			Numérique	37
			Ogam	82
			Opérateur	39
			Oriya	20
			Pavé	45
			Petite variante	67
			Phonétique	5
			Pictogramme	41
			Pleine chasse	69
			PMB	300, 301, 302 (299)
			Ponctuation	32, 49
			Présentation	63, 64, 68
			Reconnaissance optique de caractères	42
			Rune	83
			Sélecteur	205, 206
			Singhalais	84
			Spécial (caractère)	70
			Syllabaire	71, 74, 76
			Symbole	9, 34, 35, 36, 47, 49.
			Syriaque	85
			Tamoul	21
			Technique	40
			Télougou	22

Les collections suivantes se situent en dehors du plan multilingue de base.

400 PLANS À USAGE PRIVÉ G=00,
P=0F, 10, & E0 - FF

500 GROUPES À USAGE PRIVÉ G=60 - 7F

NOTE 1 : L'utilisation des niveaux de mise en œuvre 1 et 2 restreint le répertoire de certaines collections de caractères (voir 24.4). Les collections suivantes comprennent des caractères combinatoires : 7, 10, 13 à 26, 35, 49, 50, 63, 65, 72, 84, 85, 86, 87, 88, 89, 90, et 91.

NOTE 2 : Les collections n° 57, 58, et 59 étaient décrites dans la première édition de cette norme internationale, elles sont maintenant supprimées.

NOTE 3 : On trouve ci-dessous, dans l'ordre alphabétique, les termes principaux ou mots-clés utilisés dans les noms de collection. Les numéros de collection de toutes les collections qui comprennent un mot-clé sont indiqués en regard de ce mot-clé. Ces termes ne fournissent pas une référence croisée complète vers toutes les collections ou des caractères qui partagent un certain attribut, comme le type d'écriture, se retrouvent. Bien que la plupart des termes identifient un attribut commun aux caractères de cette collection, certains caractères possédant cet attribut peuvent se retrouver dans d'autres collections dont les numéros n'apparaissent pas dans l'entrée correspondant à ce terme.

Alphabétique	63
Alphanumérique	43
API	5
Arabe	14, 15, 64, 68
Arménien	11
Autochtone canadien	74
Bengali	17
Bidirectionnel	202, 203
Birman	87, 90
Bopomofo	52
Bourre	204
Braille	80
Caractère combinatoire	v. diacritique
Casseau	48
Cerclé	43, 55
Chasse nulle	200
Chérokî	75
Chinois	78
CJC	49, 54, 55, 56, 60, 62, 66, 79, 81
Clés	77, 78, 79
Combinatoire (caractère)	v. diacritique
Commande	41
Compatibilité	53, 56, 62, 66

Thaï	25
Thâna	86
Tibétain	72, 91
Usage privé	61, 400, 500
Yi	76, 77

A.2 Blocs dans le PMB

Les blocs suivants sont définis dans le plan multilingue de base. Ils sont classés dans l'ordre de leur position de code.

Nom du bloc	de	à
LATIN DE BASE	0020	007E
SUPPLÉMENT LATIN-1	00A0	00FF
LATIN ÉTENDU A	0100	017F
LATIN ÉTENDU B	0180	024F
ALPHABET PHONÉTIQUE	0250	02AF
MODIFICATEURS DE LETTRES	02B0	02FF
SIGNES COMBINATOIRES	0300	036F
GREC ET COPTE	0370	03FF
CYRILLIQUE	0400	04FF
ARMÉNIEN	0530	058F
HÉBREU	0590	05FF
ARABE	0600	06FF
SYRIAQUE	0700	074F
THÂNA	0780	07BF
DÉVANÂGARÎ	0900	097F
BÉNGALI	0980	09FF
GOURMOUKHÎ	0A00	0A7F
GOUDJARATI	0A80	0AFF
ORIYA	0B00	0B7F
TAMOUL	0B80	0BFF
TÉLOUGOU	0C00	0C7F
KANNARA	0C80	0CFF
MALAYALAM	0D00	0D7F
SINGHALAIS	0D80	0DFF
THAÏ	0E00	0E7F
LAO	0E80	0EFF
TIBÉTAIN	0F00	0FFF
BIRMAN	1000	109F
GÉORGIEN	10A0	10FF
JAMOS HANGÛL	1100	11FF
ÉTHIOPIEN	1200	137F
CHÉROKÎ	13A0	13FF
SYLLABAIRES AUTOCHTONES		
CANADIENS	1400	167F
OGAM	1680	169F
RUNES	16A0	16FF
KHMER	1780	17FF
MONGOL	1800	18AF
LATIN ÉTENDU ADDITIONNEL	1E00	1EFF
GREC ÉTENDU	1F00	1FFF
PONCTUATION GÉNÉRALE	2000	206F
EXPOSANTS ET INDICES	2070	209F
SYMBOLES MONÉTAIRES	20A0	20CF
SIGNES COMBINATOIRES		
POUR SYMBOLES	20D0	20FF
SYMBOLES DE TYPE LETTRE	2100	214F
FORMES NUMÉRIQUES	2150	218F
FLÈCHES	2190	21FF
OPÉRATEURS MATHÉMATIQUES	2200	22FF
SIGNES TECHNIQUES DIVERS	2300	23FF
PICTOGRAMMES DE COMMANDE	2400	243F
RECONNAISSANCE OPTIQUE DE		
CARACTÈRES	2440	245F

ALPHANUMÉRIQUES CERCLÉS	2460	24FF
FILETS	2500	257F
PAVÉS	2580	259F
FORMES GÉOMÉTRIQUES	25A0	25FF
SYMBOLES DIVERS	2600	26FF
CASSEAU	2700	27BF
COMBINAISONS BRAILLE	2800	28FF
SUPPLÉMENT CLÉS CJC	2E80	2EFF
CLÉS CHINOISES (K'ANG-HSI)	2F00	2FDF
DESCRIPTION IDÉOGRAPHIQUE	2FF0	2FFF
SYMBOLES ET PONCTUATION CJC	3000	303F
HIRAGANA	3040	309F
KATAKANA	30A0	30FF
BOPOMOFO	3100	312F
JAMOS DE COMPATIBILITÉ HANGÛL	3130	318F
KANBUN (DIVERS CJC)	3190	319F
BOPOMOFO ÉTENDU	31A0	31BF
LETTRES ET MOIS CJC CERCLÉS	3200	32FF
COMPATIBILITÉ CJC	3300	33FF
EXTENSION A AUX IDÉOGRAMMES UNIFIÉS CJC	3400	4DBF
IDÉOGRAMMES UNIFIÉS CJC	4E00	9FFF
SYLLABAIRE YI DES MONTS-FRAIS	A000	A48F
CLÉS YI	A490	A4CF
SYLLABAIRE HANGÛL	AC00	D7A3
ZONE À USAGE PRIVÉ	E000	F8FF
IDÉOGRAMMES DE COMPATIBILITÉ CJC	F900	FAFF
FORMES DE PRÉSENTATION ALPHABÉTIQUES	FB00	FB4F
FORMES DE PRÉSENTATION ARABES A	FB50	FDFF
DEMI-SIGNES COMBINATOIRES	FE20	FE2F
FORMES DE COMPATIBILITÉ CJC	FE30	FE4F
PETITES VARIANTES DE FORME	FE50	FE6F
FORMES DE PRÉSENTATION ARABES B	FE70	FEFE
FORMES DE DEMI ET PLEINE CHASSE	FF00	FFEF
CARACTÈRES SPÉCIAUX	FFF0	FFFF

A.3 Collections fixes de l'ensemble du PMB

A.3.1 301 PMB-AMD.7

La collection 301 PMB-AMD.7 énumérée ci-dessous est une collection fixe (4.9). Elle est constituée uniquement des caractères codés qui faisaient partie du PMB de la première édition de cette norme internationale telle qu'amendée jusqu'à l'AMD.7. Par conséquent, le répertoire de cette collection n'est pas susceptible d'être modifié si de nouveaux caractères venaient à être ajoutés au PMB par des amendements ultérieurs.

NOTE : Le répertoire de la collection 300 PMB est, en revanche, susceptible d'être modifié si de nouveaux caractères étaient ajoutés au PMB à la suite d'amendements à cette norme internationale.

La collection 301 PMB-AMD.7 est définie par les intervalles suivants de positions de codes, précisés pour chaque rangée ou série de rangées contiguës.

Rangées Positions (cellules)

00	20-7E	A0-FF
01	00-F5	FA-FF

02	00-17 50-A8 B0-DE E0-E9
03	00-45 60-61 74-75 7A 7E 84-8A 8C 8E-A1 A3-CE D0-D6 DA DC DE E0 E2-F3
04	01-0C 0E-4F 51-5C 5E-86 90-C4 C7-C8 CB- CC D0-EB EE-F5 F8-F9
05	31-56 59-5F 61-87 89 91-A1 A3-B9 BB-C4 D0-EA F0-F4
06	0C 1B 1F 21-3A 40-52 60-6D 70-B7 BA-BE C0-CE D0-ED F0-F9
09	01-03 05-39 3C-4D 50-54 58-70 81-83 85-8C 8F-90 93-A8 AA-B0 B2 B6-B9 BC BE-C4 C7- C8 CB-CD D7 DC-DD DF-E3 E6-FA
0A	02 05-0A 0F-10 13-28 2A-30 32-33 35-36 38- 39 3C 3E-42 47-48 4B-4D 59-5C 5E 66-74 81-83 85-8B 8D 8F-91 93-A8 AA-B0 B2-B3 B5-B9 BC-C5 C7-C9 CB-CD D0 E0 E6-EF
0B	01-03 05-0C 0F-10 13-28 2A-30 32-33 36-39 3C-43 47-48 4B-4D 56-57 5C-5D 5F-61 66-70 82-83 85-8A 8E-90 92-25 99-9A 9C 9E-9F A3-A4 A8-AA AE-B5 B7-B9 BE-C2 C6-C8 CA-CD D7 E7-F2
0C	01-03 05-0C 0E-10 12-28 2A-33 35-39 3E-44 46-48 4A-4D 55-56 60-61 66-6F 82-83 85-8C 8E-90 92-A8 AA-B3 B5-B9 BE-C4 C6-C8 CA- CD D5-D6 DE E0-E1 E6-EF
0D	02-03 05-0C 0E-10 12-28 2A-39 3E-43 46-48 4A-4D 57 60-61 66-6F
0E	01-3A 3F-5B 81-82 84 87-88 8A 8D 94-97 99-9F A1-A3 A5 A7 AA-AB AD-B9 BB-BD C0-C4 C6 C8-CD D0-D9 DC-DD
0F	00-47 49-69 71-8B 90-95 97 99-AD B1-B7 B9
10	A0-C5 D0-F6 FB
11	00-59 5F-A2 A8-F9
1E	00-9B A0-F9
1F	00-15 18-1D 20-45 48-4D 50-57 59 5B 5D 5F-7D 80-B4 B6-C4 C6-D3 D6-DB DD-EF F2- F4 F6-FE
20	00-2E 30-46 6A-70 74-8E A0-AB D0-E1
21	00-38 53-82 90-EA
22	00-F1
23	00 02-7A
24	00-24 40-4A 60-EA
25	00-95 A0-EF
26	00-13 1A-6F
27	01-04 06-09 0C-27 29-4B 4D 4F-52 56 58-5E 61-67 76-94 98-AF B1-BE
30	00-37 3F 41-94 99-9E A1-FE
31	05-2C 31-8E 90-9F
32	00-1C 20-43 60-7B 7F-B0 C0-CB D0-FE
33	00-76 7B-DD E0-FE
4E-9F	4E00-9FA5
AC-D7	AC00-D7A3
E0-F8	E000-F8FF
F9-FA	F900-FA2D
FB	00-06 13-17 1E-36 38-3C 3E 40-41 43-44 46-B1 D3-FF
FC	00-FF
FD	00-3F 50-8F 92-C7 F0-FB
FE	20-23 30-44 49-52 54-66 68-6B 70-72 74 76- FC FF
FF	01-5E 61-BE C2-C7 CA-CF D2-D7 DA-DC E0-E6 E8-EE FD

A.3.2 299 PMB PREMIÈRE ÉDITION

Le numéro et le nom de collection

299 PMB PREMIÈRE ÉDITION

sont réservés et identifient la collection fixe constituée par tous les caractères codés qui faisaient partie du PMB dans sa première édition. Cette collection n'est plus conforme à la présente norme internationale.

NOTE : La spécification de la collection 299 PMB PREMIÈRE ÉDITION avait trait à la collection 301 PMD-AMD.7, à l'exception du remplacement des entrées correspondantes dans la liste ci-dessus par les entrées qui se trouvent ci-dessous.

<u>rangées</u>	<u>positions</u>
05	31-56 59-5F 61-87 89 B0-B9 BB-C3 D0-EA F0-F4
0F	[pas de positions]
1E	00-9A A0-F9
20	00-2E 30-46 6A-70 74-8E A0-AA D0-E1 AC-D7 [pas de positions]
ainsi que l'entrée suivante :	
34-4D	3400-4DFF
pour les intervalles de position de code des trois collections (57, 58, 59) de caractères codés; collections supprimées de cette norme internationale depuis sa première édition.	



A.3.3 302 PMB DEUXIÈME ÉDITION

La collection fixe 302 PMB DEUXIÈME ÉDITION comprend uniquement les caractères codés définis dans le PMB de cette seconde édition de l'ISO/CEI 10646-1. Par conséquent, le répertoire de cette collection n'est pas susceptible de modification si de nouveaux caractères devaient être ajoutés au PMB par des amendements ultérieurs.

La collection 302 PMB DEUXIÈME ÉDITION est définie par les intervalles suivants de positions de codes, précisés pour chaque rangée ou série de rangées contiguës.

Rangées	Positions (cellules)
00	20-7E A0-FF
01	00-FF
02	00-33 50-AD B0-EE
03	00-4E 60-62 74-75 7A 7E 84-8A 8C 8E-A1 A3-CE D0-D7 DA-F3
04	00-86 88-89 8C-C4 C7-C8 CB-CC D0-F5 F8-F9
05	31-56 59-5F 61-87 89-8A 91-A1 A3-B9 BB-C4 D0-EA F0-F4
06	0C 1B 1F 21-3A 40-55 60-6D 70-ED F0-FE
07	00-0D 0F-2C 30-4A 80-BF
09	01-03 05-39 3C-4D 50-54 58-70 81-83 85-8C 8F-90 93-A8 AA-B0 B2 B6-B9 BC BE-C4 C7-C8 CB-CD D7 DC-DD DF-E3 E6-FA
0A	02 05-0A 0F-10 13-28 2A-30 32-33 35-36 38-39 3C 3E-42 47-48 4B-4D 59-5C 5E 66-74 81-83 85-8B 8D 8F-91 93-A8 AA-B0 B2-B3 B5-B9 BC-C5 C7-C9 CB-CD D0 E0 E6-EF
0B	01-03 05-0C 0F-10 13-28 2A-30 32-33 36-39 3C-43 47-48 4B-4D 56-57 5C-5D 5F-61 66-70 82-83 85-8A 8E-90 92-25 99-9A 9C 9E-9F A3-A4 A8-AA AE-B5 B7-B9 BE-C2 C6-C8 CA-CD D7 E7-F2
0C	01-03 05-0C 0E-10 12-28 2A-33 35-39 3E-44 46-48 4A-4D 55-56 60-61 66-6F 82-83 85-8C 8E-90 92-A8 AA-B3 B5-B9 BE-C4 C6-C8 CA-CD D5-D6 DE E0-E1 E6-EF
0D	02-03 05-0C 0E-10 12-28 2A-39 3E-43 46-48 4A-4D 57 60-61 66-6F 82-83 85-96 9A-B1 B3-BB BD C0-C6 CA CF-D4 D6 D8-DF F2-F4
0E	01-3A 3F-5B 81-82 84 87-88 8A 8D 94-97 99-9F A1-A3 A5 A7 AA-AB AD-B9 BB-BD C0-C4 C6 C8-CD D0-D9 DC-DD

0F	00-47 49-6A 71-8B 90-97 99-BC BE-CC CF
10	00-21 23-27 29-2A 2C-32 36-39 40-59 A0-C5 D0-F6 FB
11	00-59 5F-A2 A8-F9
12	20-26 28-46 48 4A-4D 50-56 58 5A-5D 60-86 88 8A-8D 90-AE B0 B2-B5 B8-BE C0 C2-C5 C8-CE D0-D6 D8-EE F0-FF
13	00-0E 10 12-15 18-1E 20-46 48-5A 61-7C A0-F4
14-15	1401-15FF
16	00-76 80-9C A0-F0
17	80-DC E0-E9
18	00-0E 10-19 20-77 80-A9
1E	00-9B A0-F9
1F	00-15 18-1D 20-45 48-4D 50-57 59 5B 5D 5F-7D 80-B4 B6-C4 C6-D3 D6-DB DD-EF F2-F4 F6-FE
20	00-46 48-4D 6A-70 74-8E A0-AF D0-E3
21	00-3A 53-83 90-F3
22	00-F1
23	00-7B 7D-9A
24	00-26 40-4A 60-EA
25	00-95 A0-F7
26	00-13 19-71
27	01-04 06-09 0C-27 29-4B 4D 4F-52 56 58-5E 61-67 76-94 98-AF B1-BE
28	00-FF
2E	80-99 9B-F3
2F	00-D5 F0-FB
30	00-3A 3E-3F 41-94 99-9E A1-FE
31	05-2C 31-8E 90-B7
32	00-1C 20-43 60-7B 7F-B0 C0-CB D0-FE
33	00-76 7B-DD E0-FE
34-4D	3400-4DBF
4E-9F	4E00-9FA5
A0-A3	A000-A3FF
A4	00-8C 90-A1 A4-B3 B5-C0 C2-C4 C6
AC-D7	AC00-D7A3
E0-F8	E000-F8FF
F9-FA	F900-FA2D
FB	00-06 13-17 1D-36 38-3C 3E 40-41 43-44 46-B1 D3-FF
FC	00-FF
FD	00-3F 50-8F 92-C7 F0-FB
FE	20-23 30-44 49-52 54-66 68-6B 70-72 74 76-FC FF
FF	01-5E 61-BE C2-C7 CA-CF D2-D7 DA-DC E0-E6 E8-EE F9-FD

Annexe B (normative)

Liste des caractères combinatoires

B.1 Liste de tous les caractères combinatoires

Les caractères des collections de sous-ensembles suivantes sont des caractères combinatoires : SIGNES COMBINATOIRES (0300 à 036F), SIGNES COMBINATOIRES POUR SYMBOLES (20D0 à 20FF) et DEMI-SIGNES COMBINATOIRES (FE20 à FE2F). Les caractères suivants doivent également être considérés comme combinatoires.

0483	DIACRITIQUE CYRILLIQUE TITLO	05B8	POINT HÉBREU QAMATS
0484	DIACRITIQUE CYRILLIQUE DE PALATALISATION	05B9	POINT HÉBREU HOLAM
0485	DIACRITIQUE CYRILLIQUE DASIA PNEUMATA	05BB	POINT HÉBREU KOUBOUTS
0486	DIACRITIQUE CYRILLIQUE PSILI PNEUMATA	05BC	POINT HÉBREU DAGUËCH OU MAPIQ
0488	SIGNE CYRILLIQUE CENTAINE DE MILLIERS	05BD	POINT HÉBREU METEG
0489	SIGNE CYRILLIQUE MILLIONS	05BF	POINT HÉBREU RAFE
0591	ACCENT HÉBREU ATNAH	05C1	POINT HÉBREU CHIN
0592	ACCENT HÉBREU SEGOLTA	05C2	POINT HÉBREU SIN
0593	ACCENT HÉBREU CHALCHÉLET	05C4	SIGNE HÉBREU POINT SUPÉRIEUR
0594	ACCENT HÉBREU ZAQEF QATON	064B	FATHATAN ARABE
0595	ACCENT HÉBREU ZAQEF GADOL	064C	DAMMATAN ARABE
0596	ACCENT HÉBREU TARHA	064D	KASRATAN ARABE
0597	ACCENT HÉBREU RAVIA	064E	FATHA ARABE
0598	ACCENT HÉBREU ZARQA	064F	DAMMA ARABE
0599	ACCENT HÉBREU QADMA	0650	KASRA ARABE
059A	ACCENT HÉBREU YETIV	0651	CHADDA ARABE
059B	ACCENT HÉBREU TEVIR	0652	SOUKOUN ARABE
059C	ACCENT HÉBREU GERICH	0653	MADDA ARABE EN CHEF
059D	ACCENT HÉBREU GERICH MOUQDAM	0654	HAMZA ARABE EN CHEF
059E	ACCENT HÉBREU CHENÉ GRICHINE	0655	HAMZA ARABE SOUSCRIT
059F	ACCENT HÉBREU KARNÉ FARAH	0670	LETTRE ARABE ALIF EN CHEF
05A0	ACCENT HÉBREU TALCHAH	06D7	LIGATURE ARABE MINUSCULE QAF EN CHEF
05A1	ACCENT HÉBREU PAZER GADOL	06D8	LAM ALIF MAKSOURA
05A3	ACCENT HÉBREU CHOFAR HOLEKH	06D9	LETTRE ARABE MINUSCULE MÎM EN CHEF
05A4	ACCENT HÉBREU CHOFAR MEHOUPPAKH	06DA	INITIAL
05A5	ACCENT HÉBREU MÂRIKH	06DB	LETTRE ARABE MINUSCULE ALIF LAM EN CHEF
05A6	ACCENT HÉBREU TÊRÉ TÂMÉ	06DD	LETTRE ARABE MINUSCULE DJÎM EN CHEF
05A7	ACCENT HÉBREU DARGA	06DE	TROIS POINTS EN CHEF ARABE
05A8	ACCENT HÉBREU QADMA	06DF	LETTRE ARABE MINUSCULE SÎN EN CHEF
05A9	ACCENT HÉBREU TARSA	06E0	FIN DE AYAH ARABE
05AA	ACCENT HÉBREU YARÉAH BEN YOMO	06E1	DÉBUT DE ROUB EL HIZB ARABE
05AB	ACCENT HÉBREU OLEH	06E2	ZÉRO ARRONDI EN CHEF ARABE
05AC	ACCENT HÉBREU ILOUZ	06E3	ZÉRO RECTANGULAIRE EN CHEF ARABE
05AD	ACCENT HÉBREU DEHI	06E4	PARTIE HAUTE DE KHA' SANS POINT EN CHEF
05AE	ACCENT HÉBREU T̄SINOR	06E5	ARABE
05AF	CERCLE MASSORÉTIQUE	06E6	LETTRE ARABE MINUSCULE MÎM EN CHEF ISOLÉ
05B0	POINT HÉBREU CHEVA	06E7	LETTRE ARABE MINUSCULE SÎN SOUSCRIT
05B1	POINT HÉBREU HATAF SEGOL	06E8	LETTRE ARABE MINUSCULE MADDA EN CHEF
05B2	POINT HÉBREU HATAF PATAH	06E9	LETTRE ARABE MINUSCULE YA' EN CHEF
05B3	POINT HÉBREU HATAF QAMATS	06EA	LETTRE ARABE MINUSCULE NOÛN EN CHEF
05B4	POINT HÉBREU HIRIK	06EB	POINT CONTOUR SOUSCRIT ARABE
05B5	POINT HÉBREU TSÉRÉ	06EC	POINT CONTOUR EN CHEF ARABE
05B6	POINT HÉBREU SĒGOL	06ED	POINT PLEIN SOUSCRIT ARABE
05B7	POINT HÉBREU PATAH	0711	LETTRE ARABE MINUSCULE MÎM SOUSCRIT
		0730	LETTRE SYRIAQUE ÂLAPH EN CHEF
		0731	PETHÂHHÂ SYRIAQUE EN CHEF
		0732	PETHÂHHÂ SYRIAQUE SOUSCRIT
		0733	PETHÂHHÂ SYRIAQUE POINTÉ
		0734	ZEQÂPHÂ SYRIAQUE EN CHEF
		0735	ZEQÂPHÂ SYRIAQUE SOUSCRIT
		0736	ZEQÂPHÂ SYRIAQUE POINTILLÉ
		0737	REVÂSSÂ SYRIAQUE EN CHEF
		0738	REVÂSSÂ SYRIAQUE SOUSCRIT
			ZELÂMÂ SYRIAQUE POINTILLÉ HORIZONTAL

0739	ZELÂMÂ SYRIAQUE POINTILLÉ OBLIQUE	0983	SYMBOLE BENGALI VISARGA
073A	HHEVÂSSÂ SYRIAQUE EN CHEF	09BC	SYMBOLE BENGALI NOUKTA
073B	HHEVÂSSÂ SYRIAQUE SOUSCRIT	09BE	VOYELLE DIACRITIQUE BENGALI Â
073C	HHEVÂSSÂ/ESSÂSÂ SYRIAQUE POINTÉ	09BF	VOYELLE DIACRITIQUE BENGALI I
073D	'ESSÂSÂ SYRIAQUE EN CHEF	09C0	VOYELLE DIACRITIQUE BENGALI Î
073E	'ESSÂSÂ SYRIAQUE SOUSCRIT	09C1	VOYELLE DIACRITIQUE BENGALI OU
073F	REWÂHHÂ SYRIAQUE	09C2	VOYELLE DIACRITIQUE BENGALI OÛ
0740	POINT FÉMININ SYRIAQUE	09C3	VOYELLE DIACRITIQUE BENGALI R VOCALIQUE
0741	QOUCHCHÂI SYRIAQUE	09C4	VOYELLE DIACRITIQUE BENGALI RR VOCALIQUE
0742	ROUKKÂKH SYRIAQUE	09C7	VOYELLE DIACRITIQUE BENGALI É
0743	DEUX POINTS VERTICAUX SYRIAQUES EN CHEF	09C8	VOYELLE DIACRITIQUE BENGALI Ā
0744	DEUX POINTS VERTICAUX SYRIAQUES SOUSCRITS	09CB	VOYELLE DIACRITIQUE BENGALI Ō
0745	TROIS POINTS SYRIAQUES EN CHEF	09CC	VOYELLE DIACRITIQUE BENGALI AOU
0746	TROIS POINTS SYRIAQUES SOUSCRITS	09CD	SYMBOLE BENGALI VIRÂMA
0747	LIGNE OBLIQUE SYRIAQUE EN CHEF	09D7	SIGNE DE LONGUEUR BENGALI AOU
0748	LIGNE OBLIQUE SYRIAQUE SOUSCRITE	09E2	VOYELLE DIACRITIQUE BENGALI L VOCALIQUE
0749	MUSIQUE SYRIAQUE	09E3	VOYELLE DIACRITIQUE BENGALI LL VOCALIQUE
074A	BARREKH SYRIAQUE	0A02	SYMBOLE GOURMOUKHÎ BINDI
07A6	THÂNA ABAFILI	0A3C	SYMBOLE GOURMOUKHÎ NOUKTA
07A7	THÂNA ÂBÂFILI	0A3E	VOYELLE DIACRITIQUE GOURMOUKHÎ Â
07A8	THÂNA IBIFILI	0A3F	VOYELLE DIACRITIQUE GOURMOUKHÎ I
07A9	THÂNA ÎBÎFILI	0A40	VOYELLE DIACRITIQUE GOURMOUKHÎ Î
07AA	THÂNA OUBOUFILI	0A41	VOYELLE DIACRITIQUE GOURMOUKHÎ OU
07AB	THÂNA OÛBOÛFILI	0A42	VOYELLE DIACRITIQUE GOURMOUKHÎ OÛ
07AC	THÂNA ÉBÉFILI	0A47	VOYELLE DIACRITIQUE GOURMOUKHÎ É
07AD	THÂNA ÉBÉFILI	0A48	VOYELLE DIACRITIQUE GOURMOUKHÎ Ā
07AE	THÂNA OBOFILI	0A4B	VOYELLE DIACRITIQUE GOURMOUKHÎ Ō
07AF	THÂNA ÔBÔFILI	0A4C	VOYELLE DIACRITIQUE GOURMOUKHÎ AOU
07B0	THÂNA SOUKOUN	0A4D	SYMBOLE GOURMOUKHÎ VIRÂMA
0901	SYMBOLE DÉVANÂGARÎ TCHANDRABINDOU	0A70	TIPPI GOURMOUKHÎ
0902	SYMBOLE DÉVANÂGARÎ ANOUSVÂRA	0A71	ADDAK GOURMOUKHÎ
0903	SYMBOLE DÉVANÂGARÎ VISARGA	0A81	SYMBOLE GOUDJARATI TCHANDRABINDOU
093C	SYMBOLE DÉVANÂGARÎ NOUKTA	0A82	SYMBOLE GOUDJARATI ANOUSVÂRA
093E	VOYELLE DIACRITIQUE DÉVANÂGARÎ Â	0A83	SYMBOLE GOUDJARATI VISARGA
093F	VOYELLE DIACRITIQUE DÉVANÂGARÎ I	0ABC	SYMBOLE GOUDJARATI NOUKTA
0940	VOYELLE DIACRITIQUE DÉVANÂGARÎ Î	0ABE	VOYELLE DIACRITIQUE GOUDJARATI Â
0941	VOYELLE DIACRITIQUE DÉVANÂGARÎ OU	0ABF	VOYELLE DIACRITIQUE GOUDJARATI I
0942	VOYELLE DIACRITIQUE DÉVANÂGARÎ OÛ	0AC0	VOYELLE DIACRITIQUE GOUDJARATI Î
0943	VOYELLE DIACRITIQUE DÉVANÂGARÎ R VOCALIQUE	0AC1	VOYELLE DIACRITIQUE GOUDJARATI OU
0944	VOYELLE DIACRITIQUE DÉVANÂGARÎ RR VOCALIQUE	0AC2	VOYELLE DIACRITIQUE GOUDJARATI OÛ
0945	VOYELLE DIACRITIQUE DÉVANÂGARÎ É TCHANDRA	0AC3	VOYELLE DIACRITIQUE GOUDJARATI R VOCALIQUE
0946	VOYELLE DIACRITIQUE DÉVANÂGARÎ É BREF	0AC4	VOYELLE DIACRITIQUE GOUDJARATI RR VOCALIQUE
0947	VOYELLE DIACRITIQUE DÉVANÂGARÎ É	0AC5	VOYELLE DIACRITIQUE GOUDJARATI É TCHANDRA
0948	VOYELLE DIACRITIQUE DÉVANÂGARÎ Ā	0AC7	VOYELLE DIACRITIQUE GOUDJARATI É
0949	VOYELLE DIACRITIQUE DÉVANÂGARÎ O TCHANDRA	0AC8	VOYELLE DIACRITIQUE GOUDJARATI Ā
094A	VOYELLE DIACRITIQUE DÉVANÂGARÎ O BREF	0AC9	VOYELLE DIACRITIQUE GOUDJARATI O TCHANDRA
094B	VOYELLE DIACRITIQUE DÉVANÂGARÎ Ō	0ACB	VOYELLE DIACRITIQUE GOUDJARATI Ō
094C	VOYELLE DIACRITIQUE DÉVANÂGARÎ AOU	0ACC	VOYELLE DIACRITIQUE GOUDJARATI AOU
094D	SYMBOLE DÉVANÂGARÎ VIRÂMA	0ACD	SYMBOLE GOUDJARATI VIRÂMA
0951	SIGNE D'ACCENTUATION DÉVANÂGARÎ OUDATTA	0B01	SYMBOLE ORIYA TCHANDRABINDOU
0952	SIGNE D'ACCENTUATION DÉVANÂGARÎ ANOUDATTA	0B02	SYMBOLE ORIYA ANOUSVÂRA
0953	ACCENT GRAVE DÉVANÂGARÎ	0B03	SYMBOLE ORIYA VISARGA
0954	ACCENT AIGU DÉVANÂGARÎ	0B3C	SYMBOLE ORIYA NOUKTA
0962	VOYELLE DIACRITIQUE DÉVANÂGARÎ L VOCALIQUE	0B3E	VOYELLE DIACRITIQUE ORIYA Â
0963	VOYELLE DIACRITIQUE DÉVANÂGARÎ LL VOCALIQUE	0B3F	VOYELLE DIACRITIQUE ORIYA I
0981	SYMBOLE BENGALI TCHANDRABINDOU	0B40	VOYELLE DIACRITIQUE ORIYA Î
0982	SYMBOLE BENGALI ANOUSVÂRA	0B41	VOYELLE DIACRITIQUE ORIYA OU
		0B42	VOYELLE DIACRITIQUE ORIYA OÛ
		0B43	VOYELLE DIACRITIQUE ORIYA R VOCALIQUE
		0B47	VOYELLE DIACRITIQUE ORIYA É
		0B48	VOYELLE DIACRITIQUE ORIYA Ā
		0B4B	VOYELLE DIACRITIQUE ORIYA Ō

0B4C	VOYELLE DIACRITIQUE ORIYA AOU	0D43	VOYELLE DIACRITIQUE MALAYALAM R VOCALIQUE
0B4D	SYMBOLE ORIYA VIRÂMA	0D46	VOYELLE DIACRITIQUE MALAYALAM É
0B56	SIGNE DE LONGUEUR ORIYA Ā	0D47	VOYELLE DIACRITIQUE MALAYALAM Ê
0B57	SIGNE DE LONGUEUR ORIYA AOU	0D48	VOYELLE DIACRITIQUE MALAYALAM Ā
0B82	SYMBOLE TAMOUL ANOUSVÂRA	0D4A	VOYELLE DIACRITIQUE MALAYALAM O
0B83	SYMBOLE TAMOUL VISARGA	0D4B	VOYELLE DIACRITIQUE MALAYALAM Ô
0BBE	VOYELLE DIACRITIQUE TAMOUL Â	0D4C	VOYELLE DIACRITIQUE MALAYALAM AOU
0BBF	VOYELLE DIACRITIQUE TAMOUL I	0D4D	SYMBOLE MALAYALAM VIRÂMA
0BC0	VOYELLE DIACRITIQUE TAMOUL Î	0D57	SIGNE DE LONGUEUR MALAYALAM AOU
0BC1	VOYELLE DIACRITIQUE TAMOUL OU	0D82	SYMBOLE SINGHALAIS ANOUSVÂRAYA
0BC2	VOYELLE DIACRITIQUE TAMOUL OÛ	0D83	SYMBOLE SINGHALAIS VISARGAYA
0BC6	VOYELLE DIACRITIQUE TAMOUL É	0DCA	SYMBOLE SINGHALAIS AL-LAKOUNA
0BC7	VOYELLE DIACRITIQUE TAMOUL Ê	0DCF	VOYELLE DIACRITIQUE SINGHALAISE ÆELA-PILLA
0BC8	VOYELLE DIACRITIQUE TAMOUL Ā	0DD0	VOYELLE DIACRITIQUE SINGHALAISE KÉTTI ÆEDA-PILLA
0BCA	VOYELLE DIACRITIQUE TAMOUL O	0DD1	VOYELLE DIACRITIQUE SINGHALAISE DIGA ÆEDA-PILLA
0BCB	VOYELLE DIACRITIQUE TAMOUL Ô	0DD2	VOYELLE DIACRITIQUE SINGHALAISE KÉTTI IS-PILLA
0BCC	VOYELLE DIACRITIQUE TAMOUL AOU	0DD3	VOYELLE DIACRITIQUE SINGHALAISE DIGA IS-PILLA
0BCD	SYMBOLE TAMOUL VIRÂMA	0DD4	VOYELLE DIACRITIQUE SINGHALAISE KÉTTI PĀ-PILLA
0BD7	SIGNE DE LONGUEUR TAMOUL AOU	0DD6	VOYELLE DIACRITIQUE SINGHALAISE DIGA PĀ-PILLA
0C01	SYMBOLE TÉLOUGOU TCHANDRABINDOU	0DD8	VOYELLE DIACRITIQUE SINGHALAISE GÆTTA-PILLA
0C02	SYMBOLE TÉLOUGOU ANOUSVÂRA	0DD9	VOYELLE DIACRITIQUE SINGHALAISE KOMBOUVA
0C03	SYMBOLE TÉLOUGOU VISARGA	0DDA	VOYELLE DIACRITIQUE SINGHALAIS DIGA KOMBOUVA
0C3E	VOYELLE DIACRITIQUE TÉLOUGOU Â	0ddb	VOYELLE DIACRITIQUE SINGHALAISE KOMBOU DEKA
0C3F	VOYELLE DIACRITIQUE TÉLOUGOU I	0DDC	VOYELLE DIACRITIQUE SINGHALAIS KOMBOUVA HAA ÆELA-PILLA
0C40	VOYELLE DIACRITIQUE TÉLOUGOU Î	0DDD	VOYELLE DIACRITIQUE SINGHALAISE KOMBOUVA HAA DIGA ÆELA-PILLA
0C41	VOYELLE DIACRITIQUE TÉLOUGOU OU	0DDE	VOYELLE DIACRITIQUE SINGHALAISE KOMBOUVA HAA GAYANOUKITTA
0C42	VOYELLE DIACRITIQUE TÉLOUGOU OÛ	0DDF	VOYELLE DIACRITIQUE SINGHALAIS GAYANOUKITTA
0C43	VOYELLE DIACRITIQUE TÉLOUGOU R VOCALIQUE	0DF2	VOYELLE DIACRITIQUE SINGHALAISE DIGA GÆTTA-PILLA
0C44	VOYELLE DIACRITIQUE TÉLOUGOU RR VOCALIQUE	0DF3	VOYELLE DIACRITIQUE SINGHALAISE DIGA GAYANOUKITTA
0C46	VOYELLE DIACRITIQUE TÉLOUGOU E	0E31	LETTRE THĀĪ MAI HAN AKAT
0C47	VOYELLE DIACRITIQUE TÉLOUGOU Ê	0E34	LETTRE THĀĪ SARA I
0C48	VOYELLE DIACRITIQUE TÉLOUGOU Ā	0E35	LETTRE THĀĪ SARA Ī
0C4A	VOYELLE DIACRITIQUE TÉLOUGOU O	0E36	LETTRE THĀĪ SARA OUE
0C4B	VOYELLE DIACRITIQUE TÉLOUGOU Ô	0E37	LETTRE THĀĪ SARA OUE
0C4C	VOYELLE DIACRITIQUE TÉLOUGOU AOU	0E38	LETTRE THĀĪ SARA OU
0C4D	SYMBOLE TÉLOUGOU VIRÂMA	0E39	LETTRE THĀĪ SARA OÛ
0C55	SIGNE DE LONGUEUR TÉLOUGOU	0E3A	LETTRE THĀĪ PHINTHOU
0C56	SIGNE DE LONGUEUR TÉLOUGOU Ā	0E47	LETTRE THĀĪ MAĪTĀĪKHOU
0C82	SYMBOLE KANNARA ANOUSVÂRA	0E48	LETTRE THĀĪ MAĪ EK
0C83	SYMBOLE KANNARA VISARGA	0E49	LETTRE THĀĪ MAĪ THO
0CBE	VOYELLE DIACRITIQUE KANNARA Â	0E4A	LETTRE THĀĪ MAĪ TRI
0CBF	VOYELLE DIACRITIQUE KANNARA I	0E4B	LETTRE THĀĪ MAĪ TCHATTAWA
0CC0	VOYELLE DIACRITIQUE KANNARA Î	0E4C	LETTRE THĀĪ THANTHAKHAT
0CC1	VOYELLE DIACRITIQUE KANNARA OU	0E4D	LETTRE THĀĪ NIKHAHIT
0CC2	VOYELLE DIACRITIQUE KANNARA OÛ	0E4E	LETTRE THĀĪ YAMAKKAN
0CC3	VOYELLE DIACRITIQUE KANNARA R VOCALIQUE	0EB1	VOYELLE DIACRITIQUE LAOTIENNE MAĪ KAN
0CC4	VOYELLE DIACRITIQUE KANNARA RR VOCALIQUE	0EB4	VOYELLE DIACRITIQUE LAOTIENNE I
0CC6	VOYELLE DIACRITIQUE KANNARA É	0EB5	VOYELLE DIACRITIQUE LAOTIENNE Ī
0CC7	VOYELLE DIACRITIQUE KANNARA Ê	0EB6	VOYELLE DIACRITIQUE LAOTIENNE U
0CC8	VOYELLE DIACRITIQUE KANNARA Ā		
0CCA	VOYELLE DIACRITIQUE KANNARA O		
0CCB	VOYELLE DIACRITIQUE KANNARA Ô		
0CCC	VOYELLE DIACRITIQUE KANNARA AOU		
0CCD	SYMBOLE KANNARA VIRÂMA		
0CD5	SIGNE DE LONGUEUR KANNARA		
0CD6	SIGNE DE LONGUEUR KANNARA Ā		
0D02	SYMBOLE MALAYALAM ANOUSVÂRA		
0D03	SYMBOLE MALAYALAM VISARGA		
0D3E	VOYELLE DIACRITIQUE MALAYALAM Â		
0D3F	VOYELLE DIACRITIQUE MALAYALAM I		
0D40	VOYELLE DIACRITIQUE MALAYALAM Î		
0D41	VOYELLE DIACRITIQUE MALAYALAM OU		
0D42	VOYELLE DIACRITIQUE MALAYALAM OÛ		

0EB7	VOYELLE DIACRITIQUE LAOTIENNE UU	0FA8	LETTRE TIBÉTAINE MA SUBJOINTE
0EB8	VOYELLE DIACRITIQUE LAOTIENNE OU	0FA9	LETTRE TIBÉTAINE TSA SUBJOINTE
0EB9	VOYELLE DIACRITIQUE LAOTIENNE OÛ	0FAA	LETTRE TIBÉTAINE TSHA SUBJOINTE
0EBB	VOYELLE DIACRITIQUE LAOTIENNE MAÏ KON	0FAB	LETTRE TIBÉTAINE DZA SUBJOINTE
0EBC	DEMI-VOYELLE DIACRITIQUE LAOTIENNE LO	0FAC	LETTRE TIBÉTAINE DZHA SUBJOINTE
0EC8	MARQUE DE TON LAOTIEN MAÏ ÉK	0FAD	LETTRE TIBÉTAINE WA SUBJOINTE
0EC9	MARQUE DE TON LAOTIEN MAÏ THO	0FAE	LETTRE TIBÉTAINE ZHA SUBJOINTE
0ECA	MARQUE DE TON LAOTIEN MAÏ TI	0FAF	LETTRE TIBÉTAINE ZA SUBJOINTE
0ECB	MARQUE DE TON LAOTIEN MA TCHATAWA	0FB0	LETTRE TIBÉTAINE 'A SUBJOINTE
0ECC	SIGNE D'ANNULATION LAOTIEN	0FB1	LETTRE TIBÉTAINE YA SUBJOINTE
0ECD	SYMBOLE LAOTIEN NIGGAHITA	0FB2	LETTRE TIBÉTAINE RA SUBJOINTE
0F18	SIGNE TIBÉTAIN ASTROLOGIQUE 'KHYUD PA	0FB3	LETTRE TIBÉTAINE LA SUBJOINTE
0F19	SIGNE TIBÉTAIN ASTROLOGIQUE SDONG TSHUGS	0FB4	LETTRE TIBÉTAINE SHA SUBJOINTE
0F35	SIGNE TIBÉTAIN NGAS BZUNG ÑI ZLA	0FB5	LETTRE TIBÉTAINE SSA SUBJOINTE
0F37	SIGNE TIBÉTAIN NGAS BZUNG SGOR RTAGS	0FB6	LETTRE TIBÉTAINE SA SUBJOINTE
0F39	SIGNE TIBÉTAIN TSA 'PHRU	0FB7	LETTRE TIBÉTAINE HA SUBJOINTE
0F3E	SIGNE TIBÉTAIN YAR TSHES	0FB8	LETTRE TIBÉTAINE A SUBJOINTE
0F3F	SIGNE TIBÉTAIN MAR TSHES	0FB9	LETTRE TIBÉTAINE KSSA SUBJOINTE KSSA
0F71	VOYELLE DIACRITIQUE TIBÉTAINE Â	0FBA	LETTRE TIBÉTAINE WA SUBJOINTE À FORME FIXE
0F72	VOYELLE DIACRITIQUE TIBÉTAINE I	0FBB	LETTRE TIBÉTAINE SUBJOINTE YA À FORME FIXE
0F73	VOYELLE DIACRITIQUE TIBÉTAINE Î	0FBC	LETTRE TIBÉTAINE SUBJOINTE RA À FORME FIXE
0F74	VOYELLE DIACRITIQUE TIBÉTAINE OU	0FC6	SYMBOLE TIBÉTAIN PADMA GDAN
0F75	VOYELLE DIACRITIQUE TIBÉTAINE OÛ	102C	VOYELLE DIACRITIQUE BIRMANE Â
0F76	VOYELLE DIACRITIQUE TIBÉTAINE R VOCALIQUE	102D	VOYELLE DIACRITIQUE BIRMANE I
0F77	VOYELLE DIACRITIQUE TIBÉTAINE RR VOCALIQUE	102E	VOYELLE DIACRITIQUE BIRMANE Î
0F78	VOYELLE DIACRITIQUE TIBÉTAINE L VOCALIQUE	102F	VOYELLE DIACRITIQUE BIRMANE OU
0F79	VOYELLE DIACRITIQUE TIBÉTAINE LL VOCALIQUE	1030	VOYELLE DIACRITIQUE BIRMANE OÛ
0F7A	VOYELLE DIACRITIQUE TIBÉTAINE É	1031	VOYELLE DIACRITIQUE BIRMANE É
0F7B	VOYELLE DIACRITIQUE TIBÉTAINE Ê	1032	VOYELLE DIACRITIQUE BIRMANE AÏ
0F7C	VOYELLE DIACRITIQUE TIBÉTAINE O	1036	SYMBOLE BIRMAN ANOUSVÂRA
0F7D	VOYELLE DIACRITIQUE TIBÉTAINE Ô	1037	POINT SOUSCRIT BIRMAN
0F7E	SIGNE TIBÉTAIN RJES SO NGA RO	1038	VISARGA BIRMAN
0F7F	SIGNE TIBÉTAIN RNAM BCAD	1039	SYMBOLE BIRMAN VIRÂMA
0F80	VOYELLE DIACRITIQUE I RÉFLÉCHI	1056	VOYELLE DIACRITIQUE BIRMANE R VOCALIQUE
0F81	VOYELLE DIACRITIQUE I RÉFLÉCHI	1057	VOYELLE DIACRITIQUE BIRMANE RR VOCALIQUE
0F82	SIGNE TIBÉTAIN ÑI ZLA NÂ DA	1058	VOYELLE DIACRITIQUE BIRMANE L VOCALIQUE
0F83	SIGNE TIBÉTAIN SNA LDAN	1059	VOYELLE DIACRITIQUE BIRMANE LL VOCALIQUE
0F84	SIGNE TIBÉTAIN HALANTA	17B4	VOYELLE KHMÈRE INHÉRENTE A
0F86	SIGNE TIBÉTAIN LCI RTAGS	17B5	VOYELLE KHMÈRE INHÉRENTE Â
0F87	SIGNE TIBÉTAIN YANG RTAGS	17B6	VOYELLE DIACRITIQUE KHMÈRE Â
0F90	LETTRE TIBÉTAINE KA SUBJOINTE	17B7	VOYELLE DIACRITIQUE KHMÈRE I
0F91	LETTRE TIBÉTAINE KHA SUBJOINTE	17B8	VOYELLE DIACRITIQUE KHMÈRE Î
0F92	LETTRE TIBÉTAINE GA SUBJOINTE	17B9	VOYELLE DIACRITIQUE KHMÈRE Û
0F93	LETTRE TIBÉTAINE GHA SUBJOINTE	17BA	VOYELLE DIACRITIQUE KHMÈRE ÜÛ
0F94	LETTRE TIBÉTAINE NGA SUBJOINTE	17BB	VOYELLE DIACRITIQUE KHMÈRE OU
0F95	LETTRE TIBÉTAINE CA SUBJOINTE	17BC	VOYELLE DIACRITIQUE KHMÈRE OÛ
0F96	LETTRE TIBÉTAINE CHA SUBJOINTE	17BD	VOYELLE DIACRITIQUE KHMÈRE OUA
0F97	LETTRE TIBÉTAINE JA SUBJOINTE	17BE	VOYELLE DIACRITIQUE KHMÈRE OE
0F99	LETTRE TIBÉTAINE ÑASUBJOINTE	17BF	VOYELLE DIACRITIQUE KHMÈRE ÜË
0F9A	LETTRE TIBÉTAINE TTA SUBJOINTE	17C0	VOYELLE DIACRITIQUE KHMÈRE IÉ
0F9B	LETTRE TIBÉTAINE TTHA SUBJOINTE	17C1	VOYELLE DIACRITIQUE KHMÈRE É
0F9C	LETTRE TIBÉTAINE DDA SUBJOINTE	17C2	VOYELLE DIACRITIQUE KHMÈRE Ê
0F9D	LETTRE TIBÉTAINE DDHA SUBJOINTE	17C3	VOYELLE DIACRITIQUE KHMÈRE AÏ
0F9E	LETTRE TIBÉTAINE NNA SUBJOINTE	17C4	VOYELLE DIACRITIQUE KHMÈRE Ô
0F9F	LETTRE TIBÉTAINE TA SUBJOINTE	17C5	VOYELLE DIACRITIQUE KHMÈRE AOU
0FA0	LETTRE TIBÉTAINE THA SUBJOINTE	17C6	SIGNE KHMER NIKAHIT
0FA1	LETTRE TIBÉTAINE DA SUBJOINTE	17C7	SIGNE KHMER RÉAHMOUK
0FA2	LETTRE TIBÉTAINE DHA SUBJOINTE	17C8	SIGNE KHMER YOÛKALÉAPINTOU
0FA3	LETTRE TIBÉTAINE NA SUBJOINTE	17C9	SIGNE KHMER MOÛSIKATOAN
0FA4	LETTRE TIBÉTAINE PA SUBJOINTE	17CA	SIGNE KHMER TRÍSAP
0FA5	LETTRE TIBÉTAINE PHA SUBJOINTE	17CB	SIGNE KHMER BANTOTCH
0FA6	LETTRE TIBÉTAINE BA SUBJOINTE	17CC	SIGNE KHMER ROBAT
0FA7	LETTRE TIBÉTAINE BHA SUBJOINTE	17CD	SIGNE KHMER TOANDAKHIAT

17CE	SIGNE KHMER KAKABAT
17CF	SIGNE KHMER AHSDA
17D0	SIGNE KHMER SAÑÑOK SAÑÑĀ
17D1	SIGNE KHMER VIRIAM
17D2	SIGNE KHMER TCHOENG
17D3	SIGNE KHMER BATHAMASAT
18A9	LETTRE MONGOLE DAGALGA ALI GALI
302A	MARQUE IDÉOGRAPHIQUE DE TON FIXE
302B	MARQUE IDÉOGRAPHIQUE DE TON ASCENDANT
302C	MARQUE IDÉOGRAPHIQUE DE TON SORTANT
302D	MARQUE IDÉOGRAPHIQUE DE TON ENTRANT
302E	MARQUE DE TON HANGÛL POINT
302F	MARQUE DE TON HANGÛL DOUBLE POINT
3099	DIACRITIQUE KATAKANA-HIRAGANA SON VOISÉ
309A	DIACRITIQUE KATAKANA-HIRAGANA SON SEMI-VOISÉ
FB1E	POINT HÉBREU JUDÉO-ESPAGNOL VARIKA

iTeh STANDARD PREVIEW (standards.iteh.ai)

[ISO/IEC 10646-1:2000](https://standards.iteh.ai/catalog/standards/sist/23a1f63a-fe59-4266-ae0-8d25aead12e2/iso-iec-10646-1-2000)

<https://standards.iteh.ai/catalog/standards/sist/23a1f63a-fe59-4266-ae0-8d25aead12e2/iso-iec-10646-1-2000>

B.2 Liste des caractères interdits au niveau 2 de mise en œuvre

Les caractères des collections de sous-ensembles suivantes sont interdits au niveau 2 de mise en œuvre : SIGNES COMBINATOIRES (0300 à 036F), SIGNES COMBINATOIRES POUR SYMBOLES (20D0 à 20FF), JAMOS HANGÛL (1100 à 11FF) et DEMI-SIGNES COMBINATOIRES (FE20 à FE2F). Les caractères suivants sont également interdits :

NOTE : Cette liste est un sous-ensemble de la liste de l'article B.1, exception faite des JAMOS HANGÛL (voir 25.1).

0483	DIACRITIQUE CYRILLIQUE TITLO	05A7	ACCENT HÉBREU DARGA
0484	DIACRITIQUE CYRILLIQUE DE PALATALISATION	05A8	ACCENT HÉBREU QADMA
0485	DIACRITIQUE CYRILLIQUE DASIA PNEUMATA	05A9	ACCENT HÉBREU TARSA
0486	DIACRITIQUE CYRILLIQUE PSILI PNEUMATA	05AA	ACCENT HÉBREU YARÉAH BEN YOMO
0591	ACCENT HÉBREU ATNAH	05AB	ACCENT HÉBREU OLEH
0592	ACCENT HÉBREU SEGOLTA	05AC	ACCENT HÉBREU ILOUZ
0593	ACCENT HÉBREU CHALCHELET	05AD	ACCENT HÉBREU DEHI
0594	ACCENT HÉBREU ZAQEF QATON	05AE	ACCENT HÉBREU TSINOR
0595	ACCENT HÉBREU ZAQEF GADOL	05AF	CERCLE MASSORÉTIQUE
0596	ACCENT HÉBREU TARHA	05C4	SIGNE HÉBREU POINT SUPÉRIEUR
0597	ACCENT HÉBREU RAVIA	093C	SYMBOLE DÉVANÂGARÎ NOUKTA
0598	ACCENT HÉBREU ZARQA	0953	ACCENT GRAVE DÉVANÂGARÎ
0599	ACCENT HÉBREU QADMA	0954	ACCENT AIGU DÉVANÂGARÎ
059A	ACCENT HÉBREU YETIV	09BC	SYMBOLE BENGALI NOUKTA
059B	ACCENT HÉBREU TEVIR	09D7	SIGNE DE LONGUEUR BENGALI AOU
059C	ACCENT HÉBREU GERICH	0A3C	SYMBOLE GOURMOUKHÎ NOUKTA
059D	ACCENT HÉBREU GERICH MOUQDAM	0A70	TIPPI GOURMOUKHÎ
059E	ACCENT HÉBREU CHENÉ GRICHINE	0A71	ADDAK GOURMOUKHÎ
059F	ACCENT HÉBREU KARNÉ FARAH	0ABC	SYMBOLE GOUDJARATI NOUKTA
05A0	ACCENT HÉBREU TALCHAH	0B3C	SYMBOLE ORIYA NOUKTA
05A1	ACCENT HÉBREU PAZER GADOL	0B56	SIGNE DE LONGUEUR ORIYA AÏ
05A3	ACCENT HÉBREU CHOFAR HOLEKH	0B57	SIGNE DE LONGUEUR ORIYA AOU
05A4	ACCENT HÉBREU CHOFAR MEHOUPPAKH	0BD7	SIGNE DE LONGUEUR TAMOUL AOU
05A5	ACCENT HÉBREU MĀRIKH	0C55	SIGNE DE LONGUEUR TÉLOUGOU
05A6	ACCENT HÉBREU TERÉ TĀMÉ	0C56	SIGNE DE LONGUEUR TÉLOUGOU AÏ
		0CD5	SIGNE DE LONGUEUR KANNARA
		0CD6	SIGNE DE LONGUEUR KANNARA AÏ
		0D57	SIGNE DE LONGUEUR MALAYALAM AOU
		0F39	SIGNE TIBÉTAÏN TSA 'PHRU
		302A	MARQUE IDÉOGRAPHIQUE DE TON FIXE
		302B	MARQUE IDÉOGRAPHIQUE DE TON ASCENDANT
		302C	MARQUE IDÉOGRAPHIQUE DE TON SORTANT
		302D	MARQUE IDÉOGRAPHIQUE DE TON ENTRANT
		302E	MARQUE DE TON HANGÛL POINT
		302F	MARQUE DE TON HANGÛL DOUBLE POINT
		3099	DIACRITIQUE KATAKANA-HIRAGANA SON VOISÉ
		309A	DIACRITIQUE KATAKANA-HIRAGANA SON SEMI-VOISÉ

Annexe C (normative)

Format transformé pour 16 plans du groupe 00 (UTF-16)

UTF-16 fournit une représentation codée pour plus d'un million de caractères graphiques de l'UCS-4 en une forme compatible avec la forme à deux octets du PMB de l'UCS-2 (section 13.1). De la sorte, les caractères UCS-4 peuvent coexister avec des données de caractères codés conformes à l'UCS-2.

En UTF-16, chaque caractère graphique du répertoire de l'UCS-2 conserve sa représentation codée UCS-2. En outre, les représentations codées des caractères d'un bloc contigu formé de 16 plans du groupe 00 (1 048 576 positions) sont formées de paires de seizelets (4.32), où chacun des seizelets correspond à une cellule d'un bloc contigu de 8 rangées du PMB (2 048 positions). Ces positions sont réservées à cette forme de représentation codée et ne peuvent pas être utilisées à d'autres fins.

C.1 Définition d'UTF-16

On définit UTF-16 de la manière suivante :

1. La demi-zone haute d'indirection est constituée des 4 rangées D8 à DB du PMB, c'est-à-dire des 1 024 positions de la zone S à partir de D800 jusqu'à DBFF.
2. La demi-zone basse d'indirection est constituée des 4 rangées DC à DF du PMB, c'est-à-dire des 1 024 positions de la zone S de DC00 à DFFF.
3. Toutes les cellules des demi-zones haute et basse sont réservées de façon définitive à la forme de représentation codée UTF-16.
4. En UTF-16, tout caractère du PMB est représenté par son code UCS-2, telle que décrit dans le corps de cette Norme internationale.
5. En UTF-16, tout caractère du JUC dont la représentation codée UCS-4 se retrouve dans l'intervalle de 0001 0000 à 0010 FFFF est représenté par une suite de seizelets de la zone S, le premier étant un seizelet de la demi-zone haute, le second un seizelet la demi-zone basse.

L'algorithme de passage d'UCS-4 à UTF-16 pour ces caractères est décrit en C.3, l'algorithme assurant le passage inverse est décrit en C.4.

NOTE : Le standard Unicode, version 3.0, définit les formes suivantes d'UTF-16 :

- UTF-16 : ordre des octets (6.3) non précisé et signature (annexe H) optionnelle ;
- UTF-16BE : l'octet de poids fort précède l'octet de poids faible, tel que précisé en 6.2 ; sans signature ;
- UTF-16LE : l'octet de poids faible précède l'octet de poids fort ; sans signature.

C.2 Notation

1. Tous les nombres sont en notation hexadécimale.
2. Les seizelets UTF-16 sont séparés par des points-virgules.
3. Le symbole « % » indique l'opération modulo, par exemple ; $x \% y = x$ modulo y .
4. Le symbole « / » indique l'opération de division entière ; par exemple, $7 / 3 = 2$.
5. La hiérarchie des opérations est :
division entière > opération modulo > multiplication entière > addition entière.

C.3 De la forme UCS-4 à la forme UTF-16

UCS-4 (4-octets)	UTF-16, seizelets
$x = 0000\ 0000 \dots 0000\ FFFF$ (cf. Note)	$x \% 0001\ 0000$;
$x = 0001\ 0000 \dots 0010\ FFFF$	$y; z$;
où	$y = ((x - 0001\ 0000) / 400) + D800$ $z = ((x - 0001\ 0000) \% 400) + DC00$
$x = 0011\ 0000 \dots 7FFF\ FFFF$	(la transformation n'est pas définie)

NOTE : Les positions de 0000 D800 à 0000 DFFF sont réservées à la forme UTF-16 et ne concernent pas la forme UCS-4. Les valeurs 0000 FFFE et 0000 FFFF sont aussi exclues (voir article 8). La transformation de ces valeurs par UTF-16 n'est pas définie.

Exemple :

La séquence UCS-4 [0000 0048] [0000 0069]
[0001 0000] [0000 0021] [0000 0021]

représente « Hi<0001 0000>!! ».

Elle se transforme sous UTF-16 en :

[0048] [0069] [D800] [DC00] [0021] [0021]

Interprétée en UCS-2, cette suite vaut :

« Hi<seizet de demi-zone haute>
<seizet de demi-zone basse>!! »

C.4 De la forme UTF-16 à la forme UCS-4

UTF-16, seizets UCS-4 (4-octets)

x = 0000; ... D7FF;

x

x = E000; ... FFFF;

x

paire (x, y) telle que

x = D800; ... DBFF;

((x - D800) * 400

y = DC00; ... DFFF;

+ (y - DC00))

+ 0001 0000

Exemple :

La suite UTF-16

[0048] [0069] [D800] [DC00] [0021] [0021]

devient en UCS-4

[0000 0048] [0000 0069] [0001 0000]

[0000 0021] [0000 0021]

et représente « Hi<0001 0000>!! ».

C.5 Identification de l'UTF-16

Quand on utilise les séquences d'échappement de l'ISO/CEI 2022, il y a lieu d'identifier qu'il est fait recours à l'UTF-16 et à un niveau de mise en œuvre (cf. article 14) par une séquence de désignation choisie dans la liste suivante :

ESC 02/05 02/15 04/10

UTF-16 avec niveau de mise en œuvre 1

ESC 02/05 02/15 04/11

UTF-16 avec niveau de mise en œuvre 2

ESC 02/05 02/15 04/12

UTF-16 avec niveau de mise en œuvre 3

Si une telle séquence d'échappement apparaît dans un flux de données CC conforme à l'ISO/CEI 2022, elle ne devra comprendre que les combinaisons de bits indiquées ci-dessus.

Si une telle séquence d'échappement apparaît dans un flux de données CC conforme à l'ISO/CEI 10646, elle sera bourrée conformément à l'article 15.

Lors de l'utilisation de séquences d'échappement ISO 2022, l'identification d'un retour ou d'un transfert

de l'UTF-16 au système de codage ISO 2022 se fera conformément à l'article 16.5, ceci pour un retour ou un transfert depuis le JUC.

C.6 Seizets non appariés : interprétation par les dispositifs de réception

Selon C.1, un seizet non apparié (4.33) n'est pas conforme à l'UTF-16.

Si un dispositif de réception ayant adopté la forme UTF-16 reçoit un seizet non apparié à la suite d'erreurs :

- dans le dispositif de départ,
- ou dans l'échange entre dispositifs de départ et de réception,
- ou enfin dans le dispositif de réception lui-même,

le seizet non apparié sera interprété de la même manière qu'un caractère qui ne fait pas partie du sous-ensemble adopté tel qu'il a été identifié pour le dispositif (voir 2.3c).

NOTE : Puisque qu'un seizet de la demi-zone haute suivi d'un seizet de la demi-zone basse est une suite conforme à l'UTF-16, le seul type de suite syntaxiquement mal formée possible est un seizet non apparié.

Exemple :

Un dispositif qui ne gère que le répertoire latin de base et utilise des carreaux (à) pour l'affichage des caractères situés en dehors de ce répertoire afficherait :

« La lettre grecque Σ est la forme majuscule de la lettre σ. »

comme :

« La lettre grecque à est la forme majuscule de la lettre à. »

Par conséquent, un dispositif qui peut interpréter l'UTF-16 devra également afficher un carreau pour chaque seizet non apparié.

C.7 Dispositifs de réception : recommandations

Quand un dispositif de réception interprète des données CC conforme à UTF-16, les recommandations suivantes s'appliquent :

1. L'UTF-16 est conçu pour être compatible avec la forme à deux octets UCS-2 du PMB (13.1). Les demi-zones haute et basse sont situées dans deux régions séparées, auxquelles aucun caractère ne peut être affecté. La fonction de tout seizet (élément de deux octets) d'un flux UTF-16 est ainsi immédiatement identifiable par sa valeur, sans égard au contexte.

Par exemple, la suite UTF-16 valide [0048] [0069] [D800] [DC00] [0021] [0021] peut être interprétée, par un dispositif de réception ayant

seulement adopté l'UCS-2, comme la représentation codée de

« Hi<inconnu><inconnu>!! »

Cette forme de compatibilité est possible parce que les seizezets de la zone S sont interprétés soit selon l'UTF-16 par les dispositifs ayant adopté l'UTF-16, soit comme des caractères inconnus par les dispositifs n'ayant adopté que l'UCS-2. Un dispositif de départ peut ainsi transmettre de l'UTF-16 même si le récepteur l'interprète comme du texte UCS-2.

2. Les concepteurs de dispositifs peuvent retenir UTF-16 comme format de texte interne à des fins de traitement ou autres, en considérant deux questions fondamentales :

- Le dispositif interprète-t-il (c'est-à-dire traite-t-il selon la sémantique définie) un quelconque sous-ensemble des paires de seizezets des demi-zones haute et basse, en d'autres termes présente-t-il le caractère correspondant à la paire de seizezets ?
- Le dispositif garantit-il l'intégrité de chaque paire de seizezets de demi-zones haute et basse, c'est-à-dire ne les sépare jamais au cours d'opérations comme la troncature de chaînes, l'insertion, ou d'autres modifications de la suite de caractères codés ?

Ces considérations donnent lieu à quatre niveaux possibles de mise en œuvre :

- (U) Mise en œuvre UCS-2 :
Aucune interprétation de paires.
Aucune garantie d'intégrité des paires.

- (F) Mise en œuvre UTF-16 faible :
Interprétation d'un sous-ensemble non nul de paires.
Aucune garantie d'intégrité des paires.
- (I) Mise en œuvre UTF-16 informée :
Aucune interprétation de paires.
Garantie d'intégrité des paires.
- (C) Mise en œuvre UTF-16 complète :
Interprétation d'un sous-ensemble non nul de paires.
Garantie d'intégrité des paires.

Exemple :

La phrase suivante pourrait être affichée de quatre façons différentes, si les mises en œuvre faible et complète disposent toutes deux de fontes étrusques mais pas d'hiéroglyphes :

« La lettre grecque Σ correspond à <hiéroglypheHaut> <hiéroglypheBas> et à <étrusqueHaut> <étrusqueBas>. »

où <xxxHaut> et <xxxBas> représentent des seizezets, de la demi-zone haute ou basse respectivement, correspondant à un caractère du bloc xxx. Les quatre représentations sont :

U : « La lettre grecque Σ correspond à à et à à. »

F : « La lettre grecque Σ correspond à à et à S. »

I : « La lettre grecque Σ correspond à à et à à. »

C : « La lettre grecque Σ correspond à à et à S. »

où S indique ici la lettre ES dans la fonte étrusque.