

Chapitre 4

Propriétés des caractères

Avertissement

Dans la mesure du possible, le consortium Unicode a vérifié le contenu de tous les tableaux de propriétés de caractère. Toutefois, le consortium Unicode ne garantit pas que les tableaux imprimés dans cet ouvrage ou sur le disque optique qui l'accompagne sont corrects dans leurs moindres détails, il n'est donc pas responsable ni des erreurs qui auraient pu se glisser dans ces tableaux de propriétés de caractères ni des erreurs qui pourraient survenir dans les logiciels qui mettraient en œuvre ces tableaux. *Le contenu de tous les tableaux de ce chapitre peut être remplacé ou amplifié par les données tenues à jour sur le site internet du consortium.*

Ce chapitre décrit les propriétés de caractère définies par le standard Unicode et précise le lien entre les caractères du standard et certaines de ces propriétés. La liste complète de toutes les propriétés Unicode est reprise dans la *Base de données de caractères Unicode* présente sur le disque optique.

Bien que le consortium Unicode s'efforce, autant que se peut, de minimiser les changements aux fichiers de propriétés de caractère, il arrive parfois qu'il faille modifier ces propriétés. Quand cela se produit, on revoit tous les fichiers appropriés de la Base de données de caractères Unicode. Les fichiers rectifiés sont par la suite affichés sur le site internet du standard Unicode et forment une mise à jour du standard.

Propriétés normatives et informatives

Tel qu'indiqué au chapitre 3, *Conformité*, le standard Unicode définit à la fois des propriétés normatives et informatives.

Propriétés normatives. Le terme normatif signifie que les mises en œuvre qui prétendent être conformes au standard Unicode (à une certaine version de celui-ci) et qui utilisent une propriété particulière doivent se plier aux spécifications du standard dans l'utilisation de cette propriété. C'est ainsi que le *type de caractère bidirectionnel* est nécessaire pour être conforme quand on affiche du texte bidirectionnel, comme de l'arabe ou de l'hébreu. *Normatif* ne veut pas dire, dans le cadre des propriétés de caractère, que la valeur de cette propriété ne changera jamais. Il se peut qu'à l'avenir des corrections ou des ajouts apportés au standard nécessitent des modifications mineures aux valeurs normatives, bien que le comité technique Unicode s'évertue à réduire au minimum de tels changements.

Propriétés informatives. Si une propriété de caractère est simplement qualifiée d'informatives, une mise en œuvre conforme peut alors choisir d'utiliser ou même de changer ces valeurs selon ses besoins tout en restant conforme au standard. L'emploi de ces valeurs est

cependant fortement recommandé. Certaines mises en œuvre peuvent décider d'ignorer les valeurs qui ne sont pas normatives et de les remplacer par d'autres. S'ils le désirent, les concepteurs peuvent mettre au point un protocole afin de transmettre cette information.

Références normatives. D'autres documents sont libres de considérer normatives certaines propriétés Unicode bien qu'elles soient informatives pour Unicode.

Les propriétés de caractère normatives du standard Unicode sont indiquées au Tableau 4-1, les propriétés informatives au Tableau 4-2. Ces mêmes tableaux précisent également où ces propriétés sont décrites et où trouver la liste des caractères ayant cette propriété.

Tableau 4-1. Propriétés normatives des caractères

Propriété	Description	Liste
Casse	Chapitre 4 (§ 4.1)	Chapitre 15 ^a
Classes combinatoires	Chapitre 3 (§ 3.10)	Tableau 4-3
Correspondance de casses	Chapitre 4	Disque optique
Décomposition (canonique et de compatibilité)	Chapitre 3 (§ 3.6)	Chapitre 15
Directionnalité	Chapitre 3 (§ 3.12)	Disque optique
Jamos jointifs (U+1100-U+11FF)	Chapitre 3 (§ 3.11)	Chapitre 15
Miroir	Chapitre 3 (D10, L4) et chapitre 4 (§ 4.7)	Tableau 4-9
Nom abrégé de jamo	Chapitre 3 (§ 3.11)	Tableau 4-4
Noms des caractères Unicode	Chapitre 15	Chapitre 15
Propriétés des caractères spéciaux	Chapitre 14	Section 3.9
Seizet d'indirection	Chapitre 3 (§ 3.7)	Section 13.4
Usage privé	Chapitre 3 (D12)	Section 13.5
Valeur numérique	Chapitre 4 (§ 4.8)	Tableau 4-6

a. Cette propriété apparaît comme annotation sous les caractères quand la casse n'est pas évidente.

Tableau 4-2. Propriétés informatives des caractères

Propriété	Description	Liste
Chasse extrême-orientale	Chapitre 10 (§ 10.3)	Disque, UTR n°11
Coupsures de ligne	Chapitre 14	Chapitre 13 (§ 13.1 et § 13.2)
Espaces	Chapitre 6	Tableau 6-1
Lettres (alphabétiques et idéographiques)	Chapitre 4	Disque optique
Propriété mathématique	Chapitre 4	Chapitre 4 (§ 4.9)
Tirets	Chapitre 7	Tableau 7-2

Propriétés implicites. Il est nécessaire que les mises en œuvre disposent de propriétés précises pour tous les points de code, même pour ceux qui ne sont pas affectés à des caractères. À cette fin, le standard Unicode attribue des propriétés implicites aux intervalles de points de code non affectés. Voir Tableau 3-8 pour un exemple.

Cohérence des propriétés. Le standard Unicode est le produit de nombreux compromis. Il doit concilier l'uniformité de traitement des caractères similaires et la compatibilité envers des pratiques courantes utilisées pour des caractères hérités de codages préexistants. Ce tour d'équilibriste entraîne, on le comprend, quelques anomalies dans les propriétés de caractère. Ainsi, certaines paires de caractères auraient-elles pu être considérées comme canoniquement équivalentes, mais on les a laissées distinctes pour des raisons de compatibilité avec des normes qui les considéraient comme telles. C'est le cas du U+00B5 μ SYMBOLE MICRO (cf. U+03BC μ LETTRE MINUSCULE GRECQUE MU) ainsi que de plusieurs jamos coréens.

En outre, certains caractères ont pu avoir des propriétés différentes de celles qui leur auraient été attribuées par ce standard, on a cependant conservé ces propriétés pour des raisons de compatibilité avec leur utilisation actuelle. C'est le cas pour les signes de voisement japonais demi-chasse (U+FF9E ◌ MARQUE KATAKANA DE SON VOISÉ DEMI-CHASSE et FF9F ◌ MARQUE KATAKANA DE SON SEMI-VOISÉ DEMI-CHASSE) qu'on aurait pu classer plus adéquatement parmi les signes combinatoires avec chasse. Il en va de même pour les jamos hangûl jointifs que l'on aurait analysé de manière plus satisfaisante comme composés d'un caractère de base initial suivi d'un caractère combinatoire médian et d'un autre caractère combinatoire final cette fois. Pour plus d'efficacité et d'uniformité dans les algorithmes, il est permis aux mises en œuvre de réanalyser de la sorte ces propriétés de caractère pour autant que le résultat de cette nouvelle interprétation n'entrent pas en conflit avec les propriétés normatives du standard.

4.1 Casse – normatif

La casse est une propriété normative associée aux caractères de certains alphabets qui considèrent certains caractères comme des variantes d'une même lettre. On appelle ces variantes, qui peuvent sensiblement différer tant par leur dessin que leur taille, la lettre *majuscule* (également nommée *capitale* ou lettre de *haut de casse*) et la lettre *minuscule* (ou de *bas de casse*). Les majuscules sont, en règle générale, plus grandes que les minuscules.

L'introduction, pour des raisons de compatibilité, de certains caractères composites (tels que U+01F1 DZ LETTRE MAJUSCULE LATINE DZ) entraîne l'emploi d'une troisième casse, appelée *casse de titre*, lorsque le premier caractère d'un mot doit être en majuscule. U+01F2 DZ LETTRE MAJUSCULE LATINE D AVEC LETTRE MINUSCULE Z constitue un exemple de ce type de caractère. Les trois formes de casses sont donc MAJUSCULE, Casse De Titre et minuscule.

Pour les écritures qui connaissent la distinction de casse (latin, grec, cyrillique, arménien et ancien géorgien), le nom du caractère Unicode indique habituellement sa casse. Ceci ne s'applique qu'à ces cinq écritures. Le nom des caractères majuscules comporte d'ordinaire le qualificatif *majuscule*, alors que le nom des caractères minuscules inclut le mot *minuscule*. Le terme minuscule qui pourrait apparaître dans le nom de caractères appartenant à d'autres écritures n'a aucun rapport avec la casse. Bien que l'ancien système d'écriture géorgien comprenne des paires majuscules-minuscules, celles-ci sont rarement utilisées en géorgien moderne. Voir la section 8.5, *Géorgien*.

Correspondance de casses. *La correspondance de casse implicite des minuscules* relie un caractère minuscule à son caractère majuscule correspondant. Le disque optique contient Les tableaux de correspondance Unicode, ceux-ci sont informatifs. Le fichier SpecialCasing.txt précise les exceptions aux règles habituelles de correspondance de casses. Pour de plus amples informations sur la correspondance de casses, veuillez vous référer à la section 5.18, *Correspondance de casses*, ainsi qu'au rapport technique Unicode n° 21, « Case Mappings » sur le cédérom ou, pour une version tenue à jour, sur le site internet du consortium.

4.2 Classes combinatoires – normatif

Chaque caractère combinatoire appartient à une *classe combinatoire* canonique normative. Cette classe sert, lors de l'algorithme de mise en ordre canonique, à établir quels caractères interagissent au niveau typographique et par là permet d'ordonner canoniquement les suites de caractères combinatoires. Pour les besoins de la mise en ordre canonique, les caractères combinatoires de la classe *zéro* se comportent comme des caractères de base. Les caractères combinatoires n'appartenant pas à la classe *zéro* participent au réordonnement afin d'établir la forme canonique des suites de caractères (cf. la section 3.10, *Mise en ordre canonique*, pour une description de l'algorithme).

Le tableau 4-3 présente la liste des caractères combinatoires et de leur classe combinatoire canonique. La plupart des caractères combinatoires ne chassent pas. Les caractères combinatoires de classe zéro qui chassent sont indiqués comme tels.

Tableau 4-3. Classes combinatoires

Avec et sans chasse		(Classe 0)
U+07A6	0	THÂNA ABAFILI
U+07A7	0	THÂNA ÂBÂFILI
U+07A8	0	THÂNA IBIFILI
U+07A9	0	THÂNA ÎBÎFILI
U+07AA	0	THÂNA OUBOUFILI
U+07AB	0	THÂNA OÛBOÛFILI
U+07AC	0	THÂNA ÉBÉFILI
U+07AD	0	THÂNA ÊBÊFILI
U+07AE	0	THÂNA OBOFILI
U+07AF	0	THÂNA ÔBÔFILI
U+07B0	0	THÂNA SOUKOUN
U+0901	0	SYMBOLE DÉVANÂGARÎ TCHANDRABINDOU
U+0902	0	SYMBOLE DÉVANÂGARÎ ANOUSVÂRA
U+0903	0	SYMBOLE DÉVANÂGARÎ VISARGA
U+093E	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ Â
U+0940	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ Î
U+0941	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ OU
U+0942	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ OÛ
U+0943	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ R VOCALIQUE
U+0944	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ RR VOCALIQUE
U+0945	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ É TCHANDRA
U+0946	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ É BREF
U+0947	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ Ê
U+0948	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ AÏ
U+0949	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ O TCHANDRA
U+094A	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ O BREF
U+094B	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ Ô
U+094C	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ AOU
U+0962	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ L VOCALIQUE
U+0963	0	VOYELLE DIACRITIQUE DÉVANÂGARÎ LL VOCALIQUE
U+0981	0	SYMBOLE BENGALI TCHANDRABINDOU
U+0982	0	SYMBOLE BENGALI ANOUSVÂRA
U+0983	0	SYMBOLE BENGALI VISARGA
U+09BE	0	VOYELLE DIACRITIQUE BENGALI Â
U+09C0	0	VOYELLE DIACRITIQUE BENGALI Î
U+09C1	0	VOYELLE DIACRITIQUE BENGALI OU
U+09C2	0	VOYELLE DIACRITIQUE BENGALI OÛ
U+09C3	0	VOYELLE DIACRITIQUE BENGALI R VOCALIQUE
U+09C4	0	VOYELLE DIACRITIQUE BENGALI RR VOCALIQUE
U+09CB	0	VOYELLE DIACRITIQUE BENGALI Ô
U+09CC	0	VOYELLE DIACRITIQUE BENGALI AOU
U+09D7	0	SIGNE DE LONGUEUR BENGALI AOU
U+09E2	0	VOYELLE DIACRITIQUE BENGALI L VOCALIQUE
U+09E3	0	VOYELLE DIACRITIQUE BENGALI LL VOCALIQUE
U+0A02	0	SYMBOLE GOURMOUKHÎ BINDI
U+0A3E	0	VOYELLE DIACRITIQUE GOURMOUKHÎ Â
U+0A40	0	VOYELLE DIACRITIQUE GOURMOUKHÎ Î

Tableau 4-3. Classes combinatoires

U+0A41	0	VOYELLE DIACRITIQUE GOURMOUKHÎ OU
U+0A42	0	VOYELLE DIACRITIQUE GOURMOUKHÎ OÛ
U+0A47	0	VOYELLE DIACRITIQUE GOURMOUKHÎ Ê
U+0A48	0	VOYELLE DIACRITIQUE GOURMOUKHÎ Ä
U+0A4B	0	VOYELLE DIACRITIQUE GOURMOUKHÎ Ô
U+0A4C	0	VOYELLE DIACRITIQUE GOURMOUKHÎ AOU
U+0A70	0	TIPPI GOURMOUKHÎ
U+0A71	0	ADDAK GOURMOUKHÎ
U+0A81	0	SYMBOLE GOUDJARATI TCHANDRABINDOU
U+0A82	0	SYMBOLE GOUDJARATI ANOUSVÂRA
U+0A83	0	SYMBOLE GOUDJARATI VISARGA
U+0ABE	0	VOYELLE DIACRITIQUE GOUDJARATI Â
U+0AC0	0	VOYELLE DIACRITIQUE GOUDJARATI Î
U+0AC1	0	VOYELLE DIACRITIQUE GOUDJARATI OU
U+0AC2	0	VOYELLE DIACRITIQUE GOUDJARATI OÛ
U+0AC3	0	VOYELLE DIACRITIQUE GOUDJARATI R VOCALIQUE
U+0AC4	0	VOYELLE DIACRITIQUE GOUDJARATI RR VOCALIQUE
U+0AC5	0	VOYELLE DIACRITIQUE GOUDJARATI É TCHANDRA
U+0AC7	0	VOYELLE DIACRITIQUE GOUDJARATI Ê
U+0AC8	0	VOYELLE DIACRITIQUE GOUDJARATI Ä
U+0AC9	0	VOYELLE DIACRITIQUE GOUDJARATI O TCHANDRA
U+0ACB	0	VOYELLE DIACRITIQUE GOUDJARATI Ô
U+0ACC	0	VOYELLE DIACRITIQUE GOUDJARATI AOU
U+0B01	0	SYMBOLE ORIYA TCHANDRABINDOU
U+0B02	0	SYMBOLE ORIYA ANOUSVÂRA
U+0B03	0	SYMBOLE ORIYA VISARGA
U+0B3E	0	VOYELLE DIACRITIQUE ORIYA Â
U+0B3F	0	VOYELLE DIACRITIQUE ORIYA I
U+0B40	0	VOYELLE DIACRITIQUE ORIYA Î
U+0B41	0	VOYELLE DIACRITIQUE ORIYA OU
U+0B42	0	VOYELLE DIACRITIQUE ORIYA OÛ
U+0B43	0	VOYELLE DIACRITIQUE ORIYA R VOCALIQUE
U+0B56	0	SIGNE DE LONGUEUR ORIYA Ä
U+0B57	0	SIGNE DE LONGUEUR ORIYA AOU
U+0B82	0	SYMBOLE TAMOUL ANOUSVÂRA
U+0B83	0	SYMBOLE TAMOUL VISARGA
U+0BBE	0	VOYELLE DIACRITIQUE TAMOULE Â
U+0BBF	0	VOYELLE DIACRITIQUE TAMOULE I
U+0BC0	0	VOYELLE DIACRITIQUE TAMOULE Î
U+0BC1	0	VOYELLE DIACRITIQUE TAMOULE OU
U+0BC2	0	VOYELLE DIACRITIQUE TAMOULE OÛ
U+0BD7	0	SIGNE DE LONGUEUR TAMOUL AOU
U+0C01	0	SYMBOLE TÉLOUGOU TCHANDRABINDOU
U+0C02	0	SYMBOLE TÉLOUGOU ANOUSVÂRA
U+0C03	0	SYMBOLE TÉLOUGOU VISARGA
U+0C3E	0	VOYELLE DIACRITIQUE TÉLOUGOU Â
U+0C3F	0	VOYELLE DIACRITIQUE TÉLOUGOU I
U+0C40	0	VOYELLE DIACRITIQUE TÉLOUGOU Î
U+0C41	0	VOYELLE DIACRITIQUE TÉLOUGOU OU
U+0C42	0	VOYELLE DIACRITIQUE TÉLOUGOU OÛ
U+0C43	0	VOYELLE DIACRITIQUE TÉLOUGOU R VOCALIQUE
U+0C44	0	VOYELLE DIACRITIQUE TÉLOUGOU RR VOCALIQUE
U+0C46	0	VOYELLE DIACRITIQUE TÉLOUGOU E

Tableau 4-3. Classes combinatoires

U+0C47	0	VOYELLE DIACRITIQUE TÉLOUGOU Ê
U+0C48	0	VOYELLE DIACRITIQUE TÉLOUGOU ĀĪ
U+0C4A	0	VOYELLE DIACRITIQUE TÉLOUGOU O
U+0C4B	0	VOYELLE DIACRITIQUE TÉLOUGOU Ô
U+0C4C	0	VOYELLE DIACRITIQUE TÉLOUGOU AOU
U+0C82	0	SYMBOLE KANNARA ANOUSVĀRA
U+0C83	0	SYMBOLE KANNARA VISARGA
U+0CBE	0	VOYELLE DIACRITIQUE KANNARA Ā
U+0CBF	0	VOYELLE DIACRITIQUE KANNARA I
U+0CC1	0	VOYELLE DIACRITIQUE KANNARA OU
U+0CC2	0	VOYELLE DIACRITIQUE KANNARA OŪ
U+0CC3	0	VOYELLE DIACRITIQUE KANNARA R VOCALIQUE
U+0CC4	0	VOYELLE DIACRITIQUE KANNARA RR VOCALIQUE
U+0CC6	0	VOYELLE DIACRITIQUE KANNARA É
U+0CCC	0	VOYELLE DIACRITIQUE KANNARA AOU
U+0CD5	0	SIGNE DE LONGUEUR KANNARA
U+0CD6	0	SIGNE DE LONGUEUR KANNARA ĀĪ
U+0D02	0	SYMBOLE MALAYALAM ANOUSVĀRA
U+0D03	0	SYMBOLE MALAYALAM VISARGA
U+0D3E	0	VOYELLE DIACRITIQUE MALAYALAM Ā
U+0D3F	0	VOYELLE DIACRITIQUE MALAYALAM I
U+0D40	0	VOYELLE DIACRITIQUE MALAYALAM Î
U+0D41	0	VOYELLE DIACRITIQUE MALAYALAM OU
U+0D42	0	VOYELLE DIACRITIQUE MALAYALAM OŪ
U+0D43	0	VOYELLE DIACRITIQUE MALAYALAM R VOCALIQUE
U+0D57	0	SIGNE DE LONGUEUR MALAYALAM AOU
U+0D82	0	SYMBOLE SINGHALAIS ANOUSVĀRAYA
U+0D83	0	SYMBOLE SINGHALAIS VISARGAYA
U+0DCF	0	VOYELLE DIACRITIQUE SINGHALAISE ÆLA-PILLA
U+0DD0	0	VOYELLE DIACRITIQUE SINGHALAISE KĒTTI ÆDA-PILLA
U+0DD1	0	VOYELLE DIACRITIQUE SINGHALAISE DIGA ÆDA-PILLA
U+0DD2	0	VOYELLE DIACRITIQUE SINGHALAISE KĒTTI IS-PILLA
U+0DD3	0	VOYELLE DIACRITIQUE SINGHALAISE DIGA IS-PILLA
U+0DD4	0	VOYELLE DIACRITIQUE SINGHALAISE KĒTTI PĀ-PILLA
U+0DD6	0	VOYELLE DIACRITIQUE SINGHALAISE DIGA PĀ-PILLA
U+0DD8	0	VOYELLE DIACRITIQUE SINGHALAISE GĀETTA-PILLA
U+0DDF	0	VOYELLE DIACRITIQUE SINGHALAIS GAYANOUKITTA
U+0DF2	0	VOYELLE DIACRITIQUE SINGHALAISE DIGA GĀETTA-PILLA
U+0DF3	0	VOYELLE DIACRITIQUE SINGHALAISE DIGA GAYANOUKITTA
U+0E31	0	LETTRE THAÏE MAI HAN AKAT
U+0E34	0	LETTRE THAÏE SARA I
U+0E35	0	LETTRE THAÏE SARA Î
U+0E36	0	LETTRE THAÏE SARA OUE
U+0E37	0	LETTRE THAÏE SARA OUE
U+0E47	0	LETTRE THAÏE MAÏTAÏKHOU
U+0E4C	0	LETTRE THAÏE THANTHAKHAT
U+0E4D	0	LETTRE THAÏE NIKHAHIT
U+0E4E	0	LETTRE THAÏE YAMAKKAN
U+0EB1	0	VOYELLE DIACRITIQUE LAOTIENNE MAÏ KAN
U+0EB4	0	VOYELLE DIACRITIQUE LAOTIENNE I
U+0EB5	0	VOYELLE DIACRITIQUE LAOTIENNE Î
U+0EB6	0	VOYELLE DIACRITIQUE LAOTIENNE U
U+0EB7	0	VOYELLE DIACRITIQUE LAOTIENNE UU

Tableau 4-3. Classes combinatoires

U+0EBB	0	VOYELLE DIACRITIQUE LAOTIENNE MAĪ KON
U+0EBC	0	DEMI-VOYELLE DIACRITIQUE LAOTIENNE LO
U+0ECC	0	SIGNE D'ANNULATION LAOTIEN
U+0ECD	0	SYMBOLE LAOTIEN NIGGAHITA
U+0F3E	0	SIGNE TIBÉTAIN YAR TSHES
U+0F3F	0	SIGNE TIBÉTAIN MAR TSHES
U+0F73	0	VOYELLE DIACRITIQUE TIBÉTAINE Î
U+0F75	0	VOYELLE DIACRITIQUE TIBÉTAINE OÛ
U+0F76	0	VOYELLE DIACRITIQUE TIBÉTAINE R VOCALIQUE
U+0F77	0	VOYELLE DIACRITIQUE TIBÉTAINE RR VOCALIQUE
U+0F78	0	VOYELLE DIACRITIQUE TIBÉTAINE L VOCALIQUE
U+0F79	0	VOYELLE DIACRITIQUE TIBÉTAINE LL VOCALIQUE
U+0F7E	0	SIGNE TIBÉTAIN RJES SO NGA RO
U+0F7F	0	SIGNE TIBÉTAIN RNAM BCAD (nam tché)
U+0F81	0	VOYELLE DIACRITIQUE Î RÉFLÉCHI
U+102C	0	VOYELLE DIACRITIQUE BIRMANE Â
U+102D	0	VOYELLE DIACRITIQUE BIRMANE I
U+102E	0	VOYELLE DIACRITIQUE BIRMANE Î
U+102F	0	VOYELLE DIACRITIQUE BIRMANE OU
U+1030	0	VOYELLE DIACRITIQUE BIRMANE OÛ
U+1032	0	VOYELLE DIACRITIQUE BIRMANE AĪ
U+1036	0	SYMBOLE BIRMAN ANOUSVĀRA
U+1038	0	VISARGA BIRMAN
U+1056	0	VOYELLE DIACRITIQUE BIRMANE R VOCALIQUE
U+1057	0	VOYELLE DIACRITIQUE BIRMANE RR VOCALIQUE
U+1058	0	VOYELLE DIACRITIQUE BIRMANE L VOCALIQUE
U+1059	0	VOYELLE DIACRITIQUE BIRMANE LL VOCALIQUE
U+17B4	0	VOYELLE IMPLICITE KHMÈRE A
U+17B5	0	VOYELLE IMPLICITE KHMÈRE Â
U+17B6	0	VOYELLE DIACRITIQUE KHMÈRE Â
U+17B7	0	VOYELLE DIACRITIQUE KHMÈRE I
U+17B8	0	VOYELLE DIACRITIQUE KHMÈRE Î
U+17B9	0	VOYELLE DIACRITIQUE KHMÈRE Û
U+17BA	0	VOYELLE DIACRITIQUE KHMÈRE ÜÛ
U+17BB	0	VOYELLE DIACRITIQUE KHMÈRE OU
U+17BC	0	VOYELLE DIACRITIQUE KHMÈRE OÛ
U+17BD	0	VOYELLE DIACRITIQUE KHMÈRE OUA
U+17C6	0	SIGNE KHMER NIKAHIT
U+17C7	0	SIGNE KHMER RÉAHMOUK
U+17C8	0	SIGNE KHMER YOÛKALÉAPINTOU
U+17C9	0	SIGNE KHMER MOÛSIKETOEN
U+17CA	0	SIGNE KHMER TRĪSAP
U+17CB	0	SIGNE KHMER BANTOTCH
U+17CC	0	SIGNE KHMER ROBAT
U+17CD	0	SIGNE KHMER TOANDAKHIAT
U+17CE	0	SIGNE KHMER KAKABAT
U+17CF	0	SIGNE KHMER AHSDA
U+17D0	0	SIGNE KHMER SAÑÑOK SAÑÑA
U+17D1	0	SIGNE KHMER VIRIAM
U+17D3	0	SIGNE KHMER BATHAMASAT
Fendues		(Classe 0)
U+09CB	0	VOYELLE DIACRITIQUE BENGALI Ô
U+09CC	0	VOYELLE DIACRITIQUE BENGALI AOU

Tableau 4-3. Classes combinatoires

U+0B48	0	VOYELLE DIACRITIQUE ORIYA ĀĪ
U+0B4B	0	VOYELLE DIACRITIQUE ORIYA Ō
U+0B4C	0	VOYELLE DIACRITIQUE ORIYA AOU
U+0BCA	0	VOYELLE DIACRITIQUE TAMOULE O
U+0BCB	0	VOYELLE DIACRITIQUE TAMOULE Ō
U+0BCC	0	VOYELLE DIACRITIQUE TAMOULE AOU
U+0CC0	0	VOYELLE DIACRITIQUE KANNARA Ī
U+0CC7	0	VOYELLE DIACRITIQUE KANNARA Ê
U+0CC8	0	VOYELLE DIACRITIQUE KANNARA ĀĪ
U+0CCA	0	VOYELLE DIACRITIQUE KANNARA O
U+0CCB	0	VOYELLE DIACRITIQUE KANNARA Ō
U+0D4A	0	VOYELLE DIACRITIQUE MALAYALAM O
U+0D4B	0	VOYELLE DIACRITIQUE MALAYALAM Ō
U+0D4C	0	VOYELLE DIACRITIQUE MALAYALAM AOU
U+0DDC	0	VOYELLE DIACRITIQUE SINGHALAIS KOMBOUVA HAA ÆLA-PILLA
U+0DDD	0	VOYELLE DIACRITIQUE SINGHALAISE KOMBOUVA HAA DIGA ÆLA-PILLA
U+0DDE	0	VOYELLE DIACRITIQUE SINGHALAISE KOMBOUVA HAA GAYANOUKITTA
U+17BF	0	VOYELLE DIACRITIQUE KHMÈRE ÜŒ
U+17C0	0	VOYELLE DIACRITIQUE KHMÈRE IĒ
U+17C4	0	VOYELLE DIACRITIQUE KHMÈRE Ō
U+17C5	0	VOYELLE DIACRITIQUE KHMÈRE AOU
Antéposées		(Classe 0)
U+093F	0	VOYELLE DIACRITIQUE DÉVANĀGARĪ I
U+09BF	0	VOYELLE DIACRITIQUE BENGALI I
U+09C7	0	VOYELLE DIACRITIQUE BENGALI É
U+09C8	0	VOYELLE DIACRITIQUE BENGALI ĀĪ
U+0A3F	0	VOYELLE DIACRITIQUE GOURMOUKHĪ I
U+0ABF	0	VOYELLE DIACRITIQUE GOUDJARATI I
U+0B47	0	VOYELLE DIACRITIQUE ORIYA É
U+0BC6	0	VOYELLE DIACRITIQUE TAMOULE É
U+0BC7	0	VOYELLE DIACRITIQUE TAMOULE Ê
U+0BC8	0	VOYELLE DIACRITIQUE TAMOULE ĀĪ
U+0D46	0	VOYELLE DIACRITIQUE MALAYALAM É
U+0D47	0	VOYELLE DIACRITIQUE MALAYALAM Ê
U+0D48	0	VOYELLE DIACRITIQUE MALAYALAM ĀĪ
U+0DD9	0	VOYELLE DIACRITIQUE SINGHALAISE KOMBOUVA
U+0DDA	0	VOYELLE DIACRITIQUE SINGHALAIS DIGA KOMBOUVA
U+0ddb	0	VOYELLE DIACRITIQUE SINGHALAISE KOMBOU DEKA
U+1031	0	VOYELLE DIACRITIQUE BIRMANE É
U+17BE	0	VOYELLE DIACRITIQUE KHMÈRE OE
U+17C1	0	VOYELLE DIACRITIQUE KHMÈRE É
U+17C2	0	VOYELLE DIACRITIQUE KHMÈRE È
U+17C3	0	VOYELLE DIACRITIQUE KHMÈRE ĀĪ
Lettres tibétaines subjointes		(Classe 0)
U+0F90	0	LETTRE TIBÉTAINE KA SUBJOINTE
U+0F91	0	LETTRE TIBÉTAINE KHA SUBJOINTE
U+0F92	0	LETTRE TIBÉTAINE GA SUBJOINTE
U+0F93	0	LETTRE TIBÉTAINE GHA SUBJOINTE
U+0F94	0	LETTRE TIBÉTAINE NGA SUBJOINTE

Tableau 4-3. Classes combinatoires

U+0F95	0	LETTRE TIBÉTAINE CA SUBJOINTE
U+0F96	0	LETTRE TIBÉTAINE CHA SUBJOINTE
U+0F97	0	LETTRE TIBÉTAINE JA SUBJOINTE
U+0F99	0	LETTRE TIBÉTAINE ÑA SUBJOINTE
U+0F9A	0	LETTRE TIBÉTAINE TTA SUBJOINTE
U+0F9B	0	LETTRE TIBÉTAINE TTHA SUBJOINTE
U+0F9C	0	LETTRE TIBÉTAINE DDA SUBJOINTE
U+0F9D	0	LETTRE TIBÉTAINE DDHA SUBJOINTE
U+0F9E	0	LETTRE TIBÉTAINE NNA SUBJOINTE
U+0F9F	0	LETTRE TIBÉTAINE TA SUBJOINTE
U+0FA0	0	LETTRE TIBÉTAINE THA SUBJOINTE
U+0FA1	0	LETTRE TIBÉTAINE DA SUBJOINTE
U+0FA2	0	LETTRE TIBÉTAINE DHA SUBJOINTE
U+0FA3	0	LETTRE TIBÉTAINE NA SUBJOINTE
U+0FA4	0	LETTRE TIBÉTAINE PA SUBJOINTE
U+0FA5	0	LETTRE TIBÉTAINE PHA SUBJOINTE
U+0FA6	0	LETTRE TIBÉTAINE BA SUBJOINTE
U+0FA7	0	LETTRE TIBÉTAINE BHA SUBJOINTE
U+0FA8	0	LETTRE TIBÉTAINE MA SUBJOINTE
U+0FA9	0	LETTRE TIBÉTAINE TSA SUBJOINTE
U+0FAA	0	LETTRE TIBÉTAINE TSHA SUBJOINTE
U+0FAB	0	LETTRE TIBÉTAINE DZA SUBJOINTE
U+0FAC	0	LETTRE TIBÉTAINE DZHA SUBJOINTE
U+0FAD	0	LETTRE TIBÉTAINE WA SUBJOINTE
U+0FAE	0	LETTRE TIBÉTAINE ZHA SUBJOINTE
U+0FAF	0	LETTRE TIBÉTAINE ZA SUBJOINTE
U+0FB0	0	LETTRE TIBÉTAINE 'A SUBJOINTE
U+0FB1	0	LETTRE TIBÉTAINE YA SUBJOINTE
U+0FB2	0	LETTRE TIBÉTAINE RA SUBJOINTE
U+0FB3	0	LETTRE TIBÉTAINE LA SUBJOINTE
U+0FB4	0	LETTRE TIBÉTAINE SHA SUBJOINTE
U+0FB5	0	LETTRE TIBÉTAINE SSA SUBJOINTE
U+0FB6	0	LETTRE TIBÉTAINE SA SUBJOINTE
U+0FB7	0	LETTRE TIBÉTAINE HA SUBJOINTE
U+0FB8	0	LETTRE TIBÉTAINE A SUBJOINTE
U+0FB9	0	LETTRE TIBÉTAINE KSSA SUBJOINTE KSSA
U+0FBA	0	LETTRE TIBÉTAINE WA SUBJOINTE À FORME FIXE
U+0FBB	0	LETTRE TIBÉTAINE SUBJOINTE YA À FORME FIXE
U+0FBC	0	LETTRE TIBÉTAINE SUBJOINTE RA À FORME FIXE
Englobantes		(Classe 0)
U+0488	0	SIGNE CYRILLIQUE CENTAINE DE MILLIERS
U+0489	0	SIGNE CYRILLIQUE MILLIONS
U+06DD	0	FIN DE AYAH ARABE
U+06DE	0	DÉBUT DE ROUB EL HIZB ARABE
U+20DD	0	DIACRITIQUE CERCLE ENGLOBANT
U+20DE	0	DIACRITIQUE CARRÉ ENGLOBANT
U+20DF	0	DIACRITIQUE LOSANGE ENGLOBANT
U+20E0	0	DIACRITIQUE CERCLE ENGLOBANT ET BARRE OBLIQUE INVERSÉE
U+20E2	0	DIACRITIQUE ÉCRAN ENGLOBANT
U+20E3	0	DIACRITIQUE TOUCHE ENGLOBANTE

Tableau 4-3. Classes combinatoires

Couvrantes/ Intérieures		(Classe 1)
U+0334	1	DIACRITIQUE TILDE COUVRANT
U+0335	1	DIACRITIQUE BARRE COURTE COUVRANTE
U+0336	1	DIACRITIQUE BARRE LONGUE COUVRANTE
U+0337	1	DIACRITIQUE BARRE OBLIQUE COURTE COUVRANTE
U+0338	1	DIACRITIQUE BARRE OBLIQUE LONGUE COUVRANTE
U+20D2	1	DIACRITIQUE LIGNE VERTICALE LONGUE COUVRANTE
U+20D3	1	DIACRITIQUE LIGNE VERTICALE COURTE COUVRANTE
U+20D8	1	DIACRITIQUE ROND COUVRANT
U+20D9	1	DIACRITIQUE ROND EN SENS NÉGATIF COUVRANT
U+20DA	1	DIACRITIQUE ROND EN SENS POSITIF COUVRANT
Nouktas		(Classe 7)
U+093C	7	SYMBOLE DÉVANÂGARÎ NOUKTA
U+09BC	7	SYMBOLE BENGALI NOUKTA
U+0A3C	7	SYMBOLE GOURMOUKHÎ NOUKTA
U+0ABC	7	SYMBOLE GOUDJARATI NOUKTA
U+0B3C	7	SYMBOLE ORIYA NOUKTA
U+1037	7	POINT SOUSCRIT BIRMAN
Signes de voisement kana		(Classe 8)
U+3099	8	DIACRITIQUE KATAKANA-HIRAGANA SON VOISÉ
U+309A	8	DIACRITIQUE KATAKANA-HIRAGANA SON SEMI-VOISÉ
Viramas		(Classe 9)
U+094D	9	SYMBOLE DÉVANÂGARÎ VIRÂMA
U+09CD	9	SYMBOLE BENGALI VIRÂMA
U+0A4D	9	SYMBOLE GOURMOUKHÎ VIRÂMA
U+0ACD	9	SYMBOLE GOUDJARATI VIRÂMA
U+0B4D	9	SYMBOLE ORIYA VIRÂMA
U+0BCD	9	SYMBOLE TAMOUL VIRÂMA
U+0C4D	9	SYMBOLE TÉLOUGOU VIRÂMA
U+0CCD	9	SYMBOLE KANNARA VIRÂMA
U+0D4D	9	SYMBOLE MALAYALAM VIRÂMA
U+0DCA	9	SYMBOLE SINGHALAIS AL-LAKOUNA
U+0E3A	9	LETTRE THAÏE PHINTHOU
U+0F84	9	SIGNE TIBÉTAÏN HALANTA
U+1039	9	SYMBOLE BIRMAN VIRÂMA
U+17D2	9	SIGNE KHMER TCHOENG
Classes à position fixe		(Classes 10..199)
U+05B0	10	POINT HÉBREU CHEVA
U+05B1	11	POINT HÉBREU HATAF SEGOL
U+05B2	12	POINT HÉBREU HATAF PATAH
U+05B3	13	POINT HÉBREU HATAF QAMATS
U+05B4	14	POINT HÉBREU HIRIK
U+05B5	15	POINT HÉBREU TSÉRÈ
U+05B6	16	POINT HÉBREU SÈGOL
U+05B7	17	POINT HÉBREU PATAH
U+05B8	18	POINT HÉBREU QAMATS
U+05B9	19	POINT HÉBREU HOLAM
U+05BB	20	POINT HÉBREU KOUBOUTS
U+05BC	21	POINT HÉBREU DAGUÈCH OU MAPIQ

Tableau 4-3. Classes combinatoires

U+05BD	22	POINT HÉBREU METEG
U+05BF	23	POINT HÉBREU RAFE
U+05C1	24	POINT HÉBREU CHIN
U+05C2	25	POINT HÉBREU SIN
U+FB1E	26	POINT HÉBREU JUDÉO-ESPAGNOL VARIKA
U+064B	27	FATHATAN ARABE
U+064C	28	DAMMATAN ARABE
U+064D	29	KASRATAN ARABE
U+064E	30	FATHA ARABE
U+064F	31	DAMMA ARABE
U+0650	32	KASRA ARABE
U+0651	33	CHADDA ARABE
U+0652	34	SOUKOUN ARABE
U+0670	35	LETTRE ARABE ALIF EN CHEF
U+0711	36	LETTRE SYRIQUE ÂLAPH EN CHEF
U+0C55	84	SIGNE DE LONGUEUR TÉLOUGOU
U+0C56	91	SIGNE DE LONGUEUR TÉLOUGOU Ā
U+0E38	103	LETTRE THAÏE SARA OU
U+0E39	103	LETTRE THAÏE SARA OÙ
U+0E48	107	LETTRE THAÏE MAÏ EK
U+0E49	107	LETTRE THAÏE MAÏ THO
U+0E4A	107	LETTRE THAÏE MAÏ TRI
U+0E4B	107	LETTRE THAÏE MAÏ TCHATTAWA
U+0EB8	118	VOYELLE DIACRITIQUE LAOTIENNE OU
U+0EB9	118	VOYELLE DIACRITIQUE LAOTIENNE OÙ
U+0EC8	122	MARQUE DE TON LAOTIEN MAÏ ĒK
U+0EC9	122	MARQUE DE TON LAOTIEN MAÏ THO
U+0ECA	122	MARQUE DE TON LAOTIEN MAÏ TI
U+0ECB	122	MARQUE DE TON LAOTIEN MA TCHATAWA
U+0F71	129	VOYELLE DIACRITIQUE TIBÉTAINE Â
U+0F72	130	VOYELLE DIACRITIQUE TIBÉTAINE I
U+0F7A	130	VOYELLE DIACRITIQUE TIBÉTAINE É
U+0F7B	130	VOYELLE DIACRITIQUE TIBÉTAINE Ê
U+0F7C	130	VOYELLE DIACRITIQUE TIBÉTAINE O
U+0F7D	130	VOYELLE DIACRITIQUE TIBÉTAINE Ô
U+0F80	130	VOYELLE DIACRITIQUE I RÉFLÉCHI
U+0F74	132	VOYELLE DIACRITIQUE TIBÉTAINE OU
Jointes en bas à gauche		(Classe 200)
Jointes en bas		(Classe 202)
U+0321	202	DIACRITIQUE HAMEÇON PALATAL SOUSCRIT
U+0322	202	DIACRITIQUE HAMEÇON RÉTROFLEXE SOUSCRIT
U+0327	202	DIACRITIQUE CÉDILLE
U+0328	202	DIACRITIQUE OGONEK
Jointes en bas à droite		(Classe 204)
Antéposées et jointes à gauche		(Classe 208)
Jointes à droite		(Classe 210)
Jointes en haut à gauche		(Classe 212)

Tableau 4-3. Classes combinatoires

Jointes en haut		(Classe 214)
Jointes en haut à droite		(Classe 216)
U+031B	216	DIACRITIQUE CORNU
U+0F39	216	SIGNE TIBÉTAIN TSA 'PHRU
Souscrites à gauche		(Classe 218)
U+302A	218	MARQUE IDÉOGRAPHIQUE DE TON FIXE
Souscrites		(Classe 220)
U+0316	220	DIACRITIQUE ACCENT GRAVE SOUSCRIT
U+0317	220	DIACRITIQUE ACCENT AIGU SOUSCRIT
U+0318	220	DIACRITIQUE TAQUET GAUCHE SOUSCRIT
U+0319	220	DIACRITIQUE TAQUET DROIT SOUSCRIT
U+031C	220	DIACRITIQUE DEMI-ROND GAUCHE SOUSCRIT
U+031D	220	DIACRITIQUE TAQUET HAUT SOUSCRIT
U+031E	220	DIACRITIQUE TAQUET BAS SOUSCRIT
U+031F	220	DIACRITIQUE SIGNE PLUS SOUSCRIT
U+0320	220	DIACRITIQUE SIGNE MOINS SOUSCRIT
U+0323	220	DIACRITIQUE POINT SOUSCRIT
U+0324	220	DIACRITIQUE TRÉMA SOUSCRIT
U+0325	220	DIACRITIQUE ROND SOUSCRIT
U+0326	220	DIACRITIQUE VIRGULE SOUSCRITE
U+0329	220	DIACRITIQUE LIGNE VERTICALE SOUSCRITE
U+032A	220	DIACRITIQUE PONTET SOUSCRIT
U+032B	220	DIACRITIQUE OMÉGA SOUSCRIT
U+032C	220	DIACRITIQUE CARON SOUSCRIT
U+032D	220	DIACRITIQUE ACCENT CIRCONFLEXE SOUSCRIT
U+032E	220	DIACRITIQUE BRÈVE SOUSCRITE
U+032F	220	DIACRITIQUE BRÈVE RENVERSÉE SOUSCRITE
U+0330	220	DIACRITIQUE TILDE SOUSCRIT
U+0331	220	DIACRITIQUE MACRON SOUSCRIT
U+0332	220	DIACRITIQUE TRAIT SOUSCRIT
U+0333	220	DIACRITIQUE DOUBLE TRAIT SOUSCRIT
U+0339	220	DIACRITIQUE DEMI-ROND À DROITE SOUSCRIT
U+033A	220	DIACRITIQUE PONTET RENVERSÉ SOUSCRIT
U+033B	220	DIACRITIQUE CARRÉ SOUSCRIT
U+033C	220	DIACRITIQUE MOUETTE SOUSCRITE
U+0347	220	DIACRITIQUE SIGNE ÉGAL SOUSCRIT
U+0348	220	DIACRITIQUE DOUBLE LIGNE VERTICALE SOUSCRITE
U+0349	220	DIACRITIQUE ANGLE GAUCHE SOUSCRIT
U+034D	220	DIACRITIQUE FLÈCHE GAUCHE ET DROITE SOUSCRITE
U+034E	220	DIACRITIQUE FLÈCHE VERS LE HAUT SOUSCRITE
U+0591	220	ACCENT HÉBREU ATNAH
U+0596	220	ACCENT HÉBREU TARHA
U+059B	220	ACCENT HÉBREU TEVIR
U+05A3	220	ACCENT HÉBREU CHOFAR HOLEKH
U+05A4	220	ACCENT HÉBREU CHOFAR MEHOUPPAKH
U+05A5	220	ACCENT HÉBREU MÂRIKH
U+05A6	220	ACCENT HÉBREU TERÉ TÂMÉ
U+05A7	220	ACCENT HÉBREU DARGA
U+05AA	220	ACCENT HÉBREU YARÉAH BEN YOMO
U+0655	220	HAMZA ARABE SOUSCRIT

Tableau 4-3. Classes combinatoires

U+06E3	220	LETTRE ARABE MINUSCULE SÎN SOUSCRIT
U+06EA	220	POINT CONTOUR SOUSCRIT ARABE
U+06ED	220	LETTRE ARABE MINUSCULE MÎM SOUSCRIT
U+0731	220	PETHÂHHÂ SYRIAQUE SOUSCRIT
U+0734	220	ZEQÂPHÂ SYRIAQUE SOUSCRIT
U+0737	220	REVÂSSÂ SYRIAQUE SOUSCRIT
U+0738	220	ZELÂMÂ SYRIAQUE POINTILLÉ HORIZONTAL
U+0739	220	ZELÂMÂ SYRIAQUE POINTILLÉ OBLIQUE
U+073B	220	HHEVÂSSÂ SYRIAQUE SOUSCRIT
U+073C	220	HHEVÂSSÂ/’ESSÂSÂ SYRIAQUE POINTÉ
U+073E	220	’ESSÂSÂ SYRIAQUE SOUSCRIT
U+0742	220	ROUKKÂKH SYRIAQUE
U+0744	220	DEUX POINTS VERTICAUX SYRIAQUES SOUSCRITS
U+0746	220	TROIS POINTS SYRIAQUES SOUSCRITS
U+0748	220	LIGNE OBLIQUE SYRIAQUE SOUSCRITE
U+0952	220	SIGNE D’ACCENTUATION DÉVANÂGARÎ ANOUDATTA
U+0F18	220	SIGNE TIBÉTAIN ASTROLOGIQUE ‘KHYUD PA
U+0F19	220	SIGNE TIBÉTAIN ASTROLOGIQUE SDONG TSHUGS
U+0F35	220	SIGNE TIBÉTAIN NGAS BZUNG ÑI ZLA
U+0F37	220	SIGNE TIBÉTAIN NGAS BZUNG SGOR RTAGS
U+0FC6	220	SYMBOLE TIBÉTAIN PADMA GDAN
Souscrites à droite		(Classe 222)
U+059A	222	ACCENT HÉBREU YETIV
U+05AD	222	ACCENT HÉBREU DEHI
U+302D	222	MARQUE IDÉOGRAPHIQUE DE TON ENTRANT
Adscrites à gauche		(Classe 224)
U+302E	224	MARQUE DE TON HANGÛL POINT
U+302F	224	MARQUE DE TON HANGÛL DOUBLE POINT
Adscrites à droite		(Classe 226)
Suscrites à gauche		(Classe 228)
U+05AE	228	ACCENT HÉBREU TSINOR
U+18A9	228	LETTRE MONGOLE DAGALGA ALI GALI
U+302B	228	MARQUE IDÉOGRAPHIQUE DE TON ASCENDANT
Suscrites		(Classe 230)
U+0300	230	DIACRITIQUE ACCENT GRAVE
U+0301	230	DIACRITIQUE ACCENT AIGU
U+0302	230	DIACRITIQUE ACCENT CIRCONFLEXE
U+0303	230	DIACRITIQUE TILDE
U+0304	230	DIACRITIQUE MACRON
U+0305	230	DIACRITIQUE TIRET HAUT
U+0306	230	DIACRITIQUE BRÈVE
U+0307	230	DIACRITIQUE POINT EN CHEF
U+0308	230	DIACRITIQUE TRÉMA
U+0309	230	DIACRITIQUE CROCHET EN CHEF
U+030A	230	DIACRITIQUE ROND EN CHEF
U+030B	230	DIACRITIQUE DOUBLE ACCENT AIGU
U+030C	230	DIACRITIQUE CARON
U+030D	230	DIACRITIQUE LIGNE VERTICALE EN CHEF
U+030E	230	DIACRITIQUE DOUBLE LIGNE VERTICALE EN CHEF
U+030F	230	DIACRITIQUE DOUBLE ACCENT GRAVE

Tableau 4-3. Classes combinatoires

U+0310	230	DIACRITIQUE TCHANDRABINDOU
U+0311	230	DIACRITIQUE BRÈVE RENVERSÉE
U+0312	230	DIACRITIQUE VIRGULE CULBUTÉE EN CHEF
U+0313	230	DIACRITIQUE VIRGULE EN CHEF
U+0314	230	DIACRITIQUE VIRGULE RÉFLÉCHIE EN CHEF
U+033D	230	DIACRITIQUE X EN CHEF
U+033E	230	DIACRITIQUE TILDE VERTICAL
U+033F	230	DIACRITIQUE DOUBLE TRAIT EN CHEF
U+0340	230	DIACRITIQUE MARQUE DE TON GRAVE
U+0341	230	DIACRITIQUE MARQUE DE TON AIGU
U+0342	230	DIACRITIQUE GREC ACCENT CIRCONFLEXE
U+0343	230	DIACRITIQUE GREC CORONIS
U+0344	230	DIACRITIQUE GREC DIALYTIKA TONOS
U+0346	230	DIACRITIQUE PONTET EN CHEF
U+034A	230	DIACRITIQUE TILDE BARRÉ EN CHEF
U+034B	230	DIACRITIQUE HOMOTHÉTIQUE EN CHEF
U+034C	230	DIACRITIQUE PRESQUE ÉGAL À EN CHEF
U+0483	230	DIACRITIQUE CYRILLIQUE TITLO
U+0484	230	DIACRITIQUE CYRILLIQUE DE PALATALISATION
U+0485	230	DIACRITIQUE CYRILLIQUE DASIA PNEUMATA
U+0486	230	DIACRITIQUE CYRILLIQUE PSILI PNEUMATA
U+0592	230	ACCENT HÉBREU SEGOLTA
U+0593	230	ACCENT HÉBREU CHALCHELET
U+0594	230	ACCENT HÉBREU ZAQEF QATON
U+0595	230	ACCENT HÉBREU ZAQEF GADOL
U+0597	230	ACCENT HÉBREU RAVIA
U+0598	230	ACCENT HÉBREU ZARQA
U+0599	230	ACCENT HÉBREU QADMA
U+059C	230	ACCENT HÉBREU GERICH
U+059D	230	ACCENT HÉBREU GERICH MOUQDAM
U+059E	230	ACCENT HÉBREU CHENÉ GRICHINE
U+059F	230	ACCENT HÉBREU KARNÉ FARAH
U+05A0	230	ACCENT HÉBREU TALCHAH
U+05A1	230	ACCENT HÉBREU PAZER GADOL
U+05A8	230	ACCENT HÉBREU QADMA
U+05A9	230	ACCENT HÉBREU TARSA
U+05AB	230	ACCENT HÉBREU OLEH
U+05AC	230	ACCENT HÉBREU ILOUZ
U+05AF	230	CERCLE MASSORÉTIQUE
U+05C4	230	SIGNE HÉBREU POINT SUPÉRIEUR
U+0653	230	MADDA ARABE EN CHEF
U+0654	230	HAMZA ARABE EN CHEF
U+06D6	230	LIGATURE ARABE MINUSCULE SAD EN CHEF LAM ALIF MAKSOURA
U+06D7	230	LIGATURE ARABE MINUSCULE QAF EN CHEF LAM ALIF MAKSOURA
U+06D8	230	LETTRE ARABE MINUSCULE MĪM EN CHEF INITIAL
U+06D9	230	LETTRE ARABE MINUSCULE ALIF LAM EN CHEF
U+06DA	230	LETTRE ARABE MINUSCULE DJĪM EN CHEF
U+06DB	230	TROIS POINTS EN CHEF ARABES
U+06DC	230	LETTRE ARABE MINUSCULE SĪN EN CHEF
U+06DF	230	ZÉRO ARRONDI EN CHEF ARABE
U+06E0	230	ZÉRO RECTANGULAIRE EN CHEF ARABE
U+06E1	230	PARTIE HAUTE DE KHA' SANS POINT EN CHEF ARABE
U+06E2	230	LETTRE ARABE MINUSCULE MĪM EN CHEF ISOLÉ

Tableau 4-3. Classes combinatoires

U+06E4	230	LETTRE ARABE MINUSCULE MADDA EN CHEF
U+06E7	230	LETTRE ARABE MINUSCULE YA' EN CHEF
U+06E8	230	LETTRE ARABE MINUSCULE NOÛN EN CHEF
U+06EB	230	POINT CONTOUR EN CHEF ARABE
U+06EC	230	POINT PLEIN SOUSCRIT ARABE
U+0730	230	PETHÂHHÂ SYRIAQUE EN CHEF
U+0732	230	PETHÂHHÂ SYRIAQUE POINTÉ
U+0733	230	ZEQÂPHÂ SYRIAQUE EN CHEF
U+0735	230	ZEQÂPHÂ SYRIAQUE POINTILLÉ
U+0736	230	REVÂSSÂ SYRIAQUE EN CHEF
U+073A	230	HHEVÂSSÂ SYRIAQUE EN CHEF
U+073D	230	'ESSÂSÂ SYRIAQUE EN CHEF
U+073F	230	REWÂHHÂ SYRIAQUE
U+0740	230	POINT FÉMININ SYRIAQUE
U+0741	230	QOUCHCHÂI SYRIAQUE
U+0743	230	DEUX POINTS VERTICAUX SYRIAQUES EN CHEF
U+0745	230	TROIS POINTS SYRIAQUES EN CHEF
U+0747	230	LIGNE OBLIQUE SYRIAQUE EN CHEF
U+0749	230	MUSIQUE SYRIAQUE
U+074A	230	BARREKH SYRIAQUE
U+0951	230	SIGNE D'ACCENTUATION DÉVANÂGARÎ OUDATTA
U+0953	230	ACCENT GRAVE DÉVANÂGARÎ
U+0954	230	ACCENT AIGU DÉVANÂGARÎ
U+0F82	230	SIGNE TIBÉTAIN ÑI ZLA NÂ DA
U+0F83	230	SIGNE TIBÉTAIN SNA LDAN
U+0F86	230	SIGNE TIBÉTAIN LCI RTAGS
U+0F87	230	SIGNE TIBÉTAIN YANG RTAGS
U+20D0	230	DIACRITIQUE VECTEUR VERS LA GAUCHE EN CHEF
U+20D1	230	DIACRITIQUE VECTEUR VERS LA DROITE EN CHEF
U+20D4	230	DIACRITIQUE FLÈCHE EN SENS POSITIF EN CHEF
U+20D5	230	DIACRITIQUE FLÈCHE EN SENS NÉGATIF EN CHEF
U+20D6	230	DIACRITIQUE FLÈCHE VERS LA GAUCHE EN CHEF
U+20D7	230	DIACRITIQUE FLÈCHE VERS LA DROITE EN CHEF
U+20DB	230	DIACRITIQUE TROIS POINTS EN CHEF
U+20DC	230	DIACRITIQUE QUATRE POINTS EN CHEF
U+20E1	230	DIACRITIQUE FLÈCHE GAUCHE ET DROITE EN CHEF
U+FE20	230	DIACRITIQUE MOITIÉ GAUCHE DE LIGATURE
U+FE21	230	DIACRITIQUE MOITIÉ DROITE DE LIGATURE
U+FE22	230	DIACRITIQUE MOITIÉ GAUCHE DE TILDE DOUBLE
U+FE23	230	DIACRITIQUE MOITIÉ DROITE DE TILDE DOUBLE
Suscrites à droite		(Classe 232)
U+0315	232	DIACRITIQUE VIRGULE EN CHEF À DROITE
U+031A	232	DIACRITIQUE ANGLE GAUCHE EN CHEF
U+302C	232	MARQUE IDÉOGRAPHIQUE DE TON SORTANT
Doubles souscrites		(Classe 233)
U+0362	233	DIACRITIQUE DOUBLE FLÈCHE VERS LA DROITE SOUSCRITE
Doubles suscrites		(Classe 234)
U+0360	234	DIACRITIQUE DOUBLE TILDE
U+0361	234	DIACRITIQUE DOUBLE BRÈVE RENVERSÉE

Tableau 4-3. Classes combinatoires

Iota souscrit		(Classe 240)
U+0345	240	DIACRITIQUE GREC IOTA SOUSCRIT

4.3 Directionnalité – normatif

On interprète le comportement directionnel conformément à l'algorithme bidirectionnel Unicode (voir la *section 3.12, Comportement bidirectionnel*). À cet effet, tous les caractères du standard Unicode possèdent un type *directionnel normatif*. Les types directionnels gauche-à-droite et droite-à-gauche sont qualifiés de *forts*, on appelle les caractères dotés de ces types des caractères directionnels forts (ou encore à *forte directionnalité*). Les types gauche-à-droite comprennent la plupart des caractères alphabétiques et syllabiques ainsi que les caractères idéographiques hans. Les types droite-à-gauche comprennent l'arabe, l'hébreu, le syriaque et le thâna, ainsi que la majorité de la ponctuation associée à ces écritures. L'algorithme bidirectionnel Unicode utilise de surcroît des types de caractères dits *faibles* et *neutres*.

Pour plus d'informations sur les types directionnels des caractères Unicode, veuillez consulter la *Base de données des caractères Unicode* sur le cédérom ainsi que la section 3.12 susmentionnée.

4.4 Noms abrégés des jamos – normatif

Le *nom abrégé des jamos* est une propriété normative des caractères jamos jointifs. Ces noms abrégés, repris au *Tableau 4-4*, servent à établir les noms des lettres isolées¹ issues de la décomposition des syllabes hangûl.

Tableau 4-4. Noms abrégés des jamos

Valeur	Nom abrégé	Glyphe	Nom ISO 10646
U+1100	K	ㄱ	HANGÛL TCH'ÔSONG KIYOK
U+1101	KK	ㄲ	HANGÛL TCH'ÔSONG SSANGKIYOK
U+1102	N	ㄴ	HANGÛL TCH'ÔSONG NIÛN
U+1103	T	ㄷ	HANGÛL TCH'ÔSONG TIKÛT
U+1104	TT	ㄸ	HANGÛL TCH'ÔSONG SSANGTIKÛT
U+1105	L	ㄹ	HANGÛL TCH'ÔSONG RIÛL
U+1106	M	ㅁ	HANGÛL TCH'ÔSONG MIÛM
U+1107	P	ㅂ	HANGÛL TCH'ÔSONG PIÛP
U+1108	PP	ㅃ	HANGÛL TCH'ÔSONG SSANGPIÛP
U+1109	S	ㅅ	HANGÛL TCH'ÔSONG SIÛS
U+110A	SS	ㅆ	HANGÛL TCH'ÔSONG SSANGSIÛS
U+110B		ㅇ	HANGÛL TCH'ÔSONG IÛNG
U+110C	TCH	ㅈ	HANGÛL TCH'ÔSONG TCHIÛTCH
U+110D	TCHTCH	ㅊ	HANGÛL TCH'ÔSONG SSANGTCHIÛTCH
U+110E	TCH'	ㅉ	HANGÛL TCH'ÔSONG TCH'IÛTCH'
U+110F	K'	ㅋ	HANGÛL TCH'ÔSONG K'IÛK'
U+1110	T'	ㅌ	HANGÛL TCH'ÔSONG T'IÛT'
U+1111	P'	ㅍ	HANGÛL TCH'ÔSONG P'IÛP'
U+1112	H'	ㅎ	HANGÛL TCH'ÔSONG HIÛH

¹ C'est le sens de jamo : consonne ou voyelle qui forment une syllabe hangûl.

Valeur	Nom abrégé	Glyphe	Nom ISO 10646
U+1161	A	ㅏ	HANGŪL DJOUNGSONG A
U+1162	Ē	ㅑ	HANGŪL DJOUNGSONG Ē
U+1163	YA	ㅓ	HANGŪL DJOUNGSONG YA
U+1164	YĒ	ㅕ	HANGŪL DJOUNGSONG YĒ
U+1165	O	ㅗ	HANGŪL DJOUNGSONG O
U+1166	É	ㅛ	HANGŪL DJOUNGSONG É
U+1167	YO	ㅜ	HANGŪL DJOUNGSONG YO
U+1168	YÉ	ㅠ	HANGŪL DJOUNGSONG YÉ
U+1169	Ô	ㅡ	HANGŪL DJOUNGSONG Ô
U+116A	WA	ㅜ	HANGŪL DJOUNGSONG WA
U+116B	WĒ	ㅝ	HANGŪL DJOUNGSONG WĒ
U+116C	EU	ㅞ	HANGŪL DJOUNGSONG EU
U+116D	YÔ	ㅟ	HANGŪL DJOUNGSONG YÔ
U+116E	OU	ㅠ	HANGŪL DJOUNGSONG OU
U+116F	WO	ㅡ	HANGŪL DJOUNGSONG WO
U+1170	WĒ	ㅢ	HANGŪL DJOUNGSONG WĒ
U+1171	WI	ㅣ	HANGŪL DJOUNGSONG WI
U+1172	YOU	ㅤ	HANGŪL DJOUNGSONG YOU
U+1173	Û	ㅥ	HANGŪL DJOUNGSONG Û
U+1174	ÛI	ㅦ	HANGŪL DJOUNGSONG ÛI
U+1175	I	ㅧ	HANGŪL DJOUNGSONG I
U+11A8	K	ㄱ	HANGŪL DJŎNGSONG KIYOK
U+11A9	KK	ㄲ	HANGŪL DJŎNGSONG SSANGKIYOK
U+11AA	KS	ㅋ	HANGŪL DJŎNGSONG KIYOK-SIÔS
U+11AB	N	ㄴ	HANGŪL DJŎNGSONG NIÛN
U+11AC	NTCH	ㄴ	HANGŪL DJŎNGSONG NIÛN-TCHIÛTCH
U+11AD	NH	ㄴ	HANGŪL DJŎNGSONG NIÛN-HIÛH
U+11AE	T	ㄷ	HANGŪL DJŎNGSONG TIKÛT
U+11AF	L	ㄹ	HANGŪL DJŎNGSONG RIÛL
U+11B0	LK	ㄹ	HANGŪL DJŎNGSONG RIÛL-KIYOK
U+11B1	LM	ㄹ	HANGŪL DJŎNGSONG RIÛL-MIÛM
U+11B2	LP	ㄹ	HANGŪL DJŎNGSONG RIÛL-PIÛP
U+11B3	LS	ㄹ	HANGŪL DJŎNGSONG RIÛL-SIÔS
U+11B4	LT'	ㄹ	HANGŪL DJŎNGSONG RIÛL-T'IÛT'
U+11B5	LP'	ㄹ	HANGŪL DJŎNGSONG RIÛL-P'IÛP'
U+11B6	LH	ㄹ	HANGŪL DJŎNGSONG RIÛL-HIÛH
U+11B7	M	ㅁ	HANGŪL DJŎNGSONG MIÛM
U+11B8	P	ㅂ	HANGŪL DJŎNGSONG PIÛP
U+11B9	PS	ㅂ	HANGŪL DJŎNGSONG PIÛP-SIÔS
U+11BA	S	ㅅ	HANGŪL DJŎNGSONG SIÔS
U+11BB	SS	ㅆ	HANGŪL DJŎNGSONG SSANGSIÔS
U+11BC	NG	ㅇ	HANGŪL DJŎNGSONG IÛNG
U+11BD	TCH	ㅈ	HANGŪL DJŎNGSONG TCHIÛTCH
U+11BE	TCH'	ㅊ	HANGŪL DJŎNGSONG TCH'IÛTCH'
U+11BF	K'	ㅋ	HANGŪL DJŎNGSONG K'IÛK'
U+11C0	T'	ㅌ	HANGŪL DJŎNGSONG T'IÛT'
U+11C1	P'	ㅍ	HANGŪL DJŎNGSONG P'IÛP'
U+11C2	H	ㅎ	HANGŪL DJŎNGSONG HIÛH

4.5 Catégorie générale – normatif

La *Base de données de caractères Unicode* fournie sur le disque définit une catégorie générale valable pour tous les caractères Unicode. Cette catégorie générale opère une partition au sein des caractères et les divise en plusieurs grandes classes, telles que les lettres, les signes de ponctuation et les symboles ainsi que d'autres sous-classes pour chacune de ces grandes classes.

On attribue à chaque caractère Unicode une valeur de catégorie générale. Chaque valeur est définie à l'aide d'une abréviation à deux lettres. La première lettre précise la grande classe, la seconde la sous-classe de cette grande classe. Pour chaque classe, la sous-classe « autre » rassemble simplement les caractères restants de la grande classe. Ainsi, la sous-classe « No » (Nombre, autres) comprend les caractères de la classe Nombre qui ne sont ni des chiffres décimaux ni des lettres. Ces caractères peuvent avoir très peu de traits en commun si ce n'est d'appartenir à la même grande classe.

On considère Zs, Zl, et Zp comme des caractères de formatage, mais leur appartenance à la classe Z (séparateur) a préséance sur leur appartenance à la classe Cf, puisque la catégorie générale n'accepte qu'une seule valeur par caractère.

La catégorie générale est normative car certaines de ses valeurs sont directement corrélées à d'autres propriétés normatives des caractères Unicode (état numérique, état combinatoire ou d'autres propriétés particulières). Le *Tableau 4-5* énumère les différentes valeurs de catégorie générale accompagnées d'une brève description de celles-ci.

Tableau 4-5. Catégorie générale

Normatives

Lu	=	Lettre, majuscule
Li	=	Lettre, minuscule
Lt	=	Lettre, casse de titre
Mn	=	Marque, à chasse nulle
Mc	=	Marque, combinatoire avec chasse
Me	=	Marque, englobante
Nd	=	Nombre, chiffre décimal
Nl	=	Nombre, lettre
No	=	Nombre, autre
Zs	=	Séparateur, espace
Zl	=	Séparateur, lignes
Zp	=	Séparateur, paragraphes
Cc	=	Autre, commande
Cf	=	Autre, formatage
Cs	=	Autre, seizet d'indirection
Co	=	Autre, usage privé
Cn	=	Autre, non affecté
Lm	=	Lettre, modificateur
Lo	=	Lettre, autre
Pc	=	Ponctuation, connecteur
Pd	=	Ponctuation, tiret
Ps	=	Ponctuation, début
Pe	=	Ponctuation, fin

Pi	=	Ponctuation, guillemet initial (selon l'utilisation, peut se comporter comme Ps ou Pe)
Pf	=	Ponctuation, guillemet final (selon l'utilisation, peut se comporter comme Ps ou Pe)
Po	=	Ponctuation, autre
Sm	=	Symbole, mathématique
Sc	=	Symbole, devise monétaire
Sk	=	Symbole, modificateur
So	=	Symbole, autre

On utilise souvent la catégorie générale d'un caractère Unicode pour repérer plus facilement les frontières de texte, voir la section 5.15, *Repérage des frontières d'élément textuel*. Elle sert également fréquemment à établir si une chaîne est un identificateur valable pour un langage informatique ou peut servir de balisage, voir la section 5.16, *Identificateurs*. Cette propriété permet également la mise en œuvre de fonctions communes comme `isLetter()`, `isUpperCase()` et ainsi de suite.

4.6 Valeur numérique – normatif

La *valeur numérique* est une propriété normative des caractères numériques (ceux qui représentent des nombres). Ce groupe comprend les fractions, les indices et les exposants, les chiffres romains, les numérateurs monétaires, les nombres cerclés et les chiffres propres à des écritures. Dans de nombreux systèmes numériques traditionnels, les lettres peuvent servir à représenter des nombres. C'est le cas des lettres grecques et hébraïques ainsi que des lettres latines utilisées comme numéros de paragraphe (II.A.1.b). Ces cas particuliers de lettres numériques ne font pas partie des nombres considérés ici.

Les *chiffres décimaux* forment une sous-catégorie importante des nombres, elle comprend tous les chiffres qui peuvent former des nombres à base décimale. On y retrouve les chiffres propres à différentes écritures, mais pas de lettres numériques comme les chiffres romains, ni les indices ou les exposants. Les numériques autres que les chiffres décimaux peuvent servir dans les expressions numériques, mais il appartient à l'utilisateur d'établir ces utilisations spécialisées.

Le standard Unicode attribue des numéros de caractères distincts aux chiffres utilisés dans les différents systèmes d'écriture ; que ce soit, par exemple, les chiffres arabes, les idéogrammes numériques ou les chiffres des langues de l'Inde. Pour les règles d'affectation de noms, consultez l'introduction de la section 9.2, *Arabe*.

Le *Tableau 4-6* fournit la valeur numérique des caractères numériques d'Unicode. Certains idéogrammes CJC ont également une valeur numérique, ils ne sont pas repris dans le *Tableau 4-6*, mais font l'objet d'un examen dans le tableau suivant.

Tableau 4-6. Propriétés numériques

Valeur	Décimal	Nombre	Nom
U+0030	✓	0	CHIFFRE ZÉRO
U+0031	✓	1	CHIFFRE UN
U+0032	✓	2	CHIFFRE DEUX
U+0033	✓	3	CHIFFRE TROIS
U+0034	✓	4	CHIFFRE QUATRE

Tableau 4-6. Propriétés numériques

Valeur	Décimal	Nombre	Nom
U+0035	✓	5	CHIFFRE CINQ
U+0036	✓	6	CHIFFRE SIX
U+0037	✓	7	CHIFFRE SEPT
U+0038	✓	8	CHIFFRE HUIT
U+0039	✓	9	CHIFFRE NEUF
U+00B2		2	EXPOSANT DEUX
U+00B3		3	EXPOSANT TROIS
U+00B9		1	EXPOSANT UN
U+00BC		1/4	FRACTION UN QUART
U+00BD		1/2	FRACTION UN DEMI
U+00BE		3/4	FRACTION TROIS QUARTS
U+0660	✓	0	CHIFFRE ARABE-HINDI ZÉRO
U+0661	✓	1	CHIFFRE ARABE-HINDI UN
U+0662	✓	2	CHIFFRE ARABE-HINDI DEUX
U+0663	✓	3	CHIFFRE ARABE-HINDI TROIS
U+0664	✓	4	CHIFFRE ARABE-HINDI QUATRE
U+0665	✓	5	CHIFFRE ARABE-HINDI CINQ
U+0666	✓	6	CHIFFRE ARABE-HINDI SIX
U+0667	✓	7	CHIFFRE ARABE-HINDI SEPT
U+0668	✓	8	CHIFFRE ARABE-HINDI HUIT
U+0669	✓	9	CHIFFRE ARABE-HINDI NEUF
U+06F0	✓	0	CHIFFRE ARABE-HINDI ÉTENDU ZÉRO
U+06F1	✓	1	CHIFFRE ARABE-HINDI ÉTENDU UN
U+06F2	✓	2	CHIFFRE ARABE-HINDI ÉTENDU DEUX
U+06F3	✓	3	CHIFFRE ARABE-HINDI ÉTENDU TROIS
U+06F4	✓	4	CHIFFRE ARABE-HINDI ÉTENDU QUATRE
U+06F5	✓	5	CHIFFRE ARABE-HINDI ÉTENDU CINQ
U+06F6	✓	6	CHIFFRE ARABE-HINDI ÉTENDU SIX
U+06F7	✓	7	CHIFFRE ARABE-HINDI ÉTENDU SEPT
U+06F8	✓	8	CHIFFRE ARABE-HINDI ÉTENDU HUIT
U+06F9	✓	9	CHIFFRE ARABE-HINDI ÉTENDU NEUF
U+0966	✓	0	CHIFFRE DÉVANÂGARÎ ZÉRO
U+0967	✓	1	CHIFFRE DÉVANÂGARÎ UN
U+0968	✓	2	CHIFFRE DÉVANÂGARÎ DEUX
U+0969	✓	3	CHIFFRE DÉVANÂGARÎ TROIS
U+096A	✓	4	CHIFFRE DÉVANÂGARÎ QUATRE
U+096B	✓	5	CHIFFRE DÉVANÂGARÎ CINQ
U+096C	✓	6	CHIFFRE DÉVANÂGARÎ SIX
U+096D	✓	7	CHIFFRE DÉVANÂGARÎ SEPT
U+096E	✓	8	CHIFFRE DÉVANÂGARÎ HUIT
U+096F	✓	9	CHIFFRE DÉVANÂGARÎ NEUF
U+09E6	✓	0	CHIFFRE BENGALI ZÉRO
U+09E7	✓	1	CHIFFRE BENGALI UN
U+09E8	✓	2	CHIFFRE BENGALI DEUX
U+09E9	✓	3	CHIFFRE BENGALI TROIS
U+09EA	✓	4	CHIFFRE BENGALI QUATRE
U+09EB	✓	5	CHIFFRE BENGALI CINQ
U+09EC	✓	6	CHIFFRE BENGALI SIX
U+09ED	✓	7	CHIFFRE BENGALI SEPT
U+09EE	✓	8	CHIFFRE BENGALI HUIT
U+09EF	✓	9	CHIFFRE BENGALI NEUF
U+09F4		1	NUMÉRATEUR MONÉTAIRE BENGALI UN

Tableau 4-6. Propriétés numériques

Valeur	Décimal	Nombre	Nom
U+09F5		2	NUMÉRATEUR MONÉTAIRE BENGALI DEUX
U+09F6		3	NUMÉRATEUR MONÉTAIRE BENGALI TROIS
U+09F7		4	NUMÉRATEUR MONÉTAIRE BENGALI QUATRE
U+09F8		-	NUMÉRATEUR MONÉTAIRE BENGALI DÉNOMINATEUR MOINS UN
U+09F9		16	DÉNOMINATEUR MONÉTAIRE BENGALI SEIZE
U+0A66	✓	0	CHIFFRE GOURMOUKHÎ ZÉRO
U+0A67	✓	1	CHIFFRE GOURMOUKHÎ UN
U+0A68	✓	2	CHIFFRE GOURMOUKHÎ DEUX
U+0A69	✓	3	CHIFFRE GOURMOUKHÎ TROIS
U+0A6A	✓	4	CHIFFRE GOURMOUKHÎ QUATRE
U+0A6B	✓	5	CHIFFRE GOURMOUKHÎ CINQ
U+0A6C	✓	6	CHIFFRE GOURMOUKHÎ SIX
U+0A6D	✓	7	CHIFFRE GOURMOUKHÎ SEPT
U+0A6E	✓	8	CHIFFRE GOURMOUKHÎ HUIT
U+0A6F	✓	9	CHIFFRE GOURMOUKHÎ NEUF
U+0AE6	✓	0	CHIFFRE GOUDJARATI ZÉRO
U+0AE7	✓	1	CHIFFRE GOUDJARATI UN
U+0AE8	✓	2	CHIFFRE GOUDJARATI DEUX
U+0AE9	✓	3	CHIFFRE GOUDJARATI TROIS
U+0AEA	✓	4	CHIFFRE GOUDJARATI QUATRE
U+0AEB	✓	5	CHIFFRE GOUDJARATI CINQ
U+0AEC	✓	6	CHIFFRE GOUDJARATI SIX
U+0AED	✓	7	CHIFFRE GOUDJARATI SEPT
U+0AEE	✓	8	CHIFFRE GOUDJARATI HUIT
U+0AEF	✓	9	CHIFFRE GOUDJARATI NEUF
U+0B66	✓	0	CHIFFRE ORIYA ZÉRO
U+0B67	✓	1	CHIFFRE ORIYA UN
U+0B68	✓	2	CHIFFRE ORIYA DEUX
U+0B69	✓	3	CHIFFRE ORIYA TROIS
U+0B6A	✓	4	CHIFFRE ORIYA QUATRE
U+0B6B	✓	5	CHIFFRE ORIYA CINQ
U+0B6C	✓	6	CHIFFRE ORIYA SIX
U+0B6D	✓	7	CHIFFRE ORIYA SEPT
U+0B6E	✓	8	CHIFFRE ORIYA HUIT
U+0B6F	✓	9	CHIFFRE ORIYA NEUF
U+0BE7	✓	1	CHIFFRE TAMOUL UN
U+0BE8	✓	2	CHIFFRE TAMOUL DEUX
U+0BE9	✓	3	CHIFFRE TAMOUL TROIS
U+0BEA	✓	4	CHIFFRE TAMOUL QUATRE
U+0BEB	✓	5	CHIFFRE TAMOUL CINQ
U+0BEC	✓	6	CHIFFRE TAMOUL SIX
U+0BED	✓	7	CHIFFRE TAMOUL SEPT
U+0BEE	✓	8	CHIFFRE TAMOUL HUIT
U+0BEF	✓	9	CHIFFRE TAMOUL NEUF
U+0BF0		10	NOMBRE TAMOUL DIX
U+0BF1		100	NOMBRE TAMOUL CENT
U+0BF2		1 000	NOMBRE TAMOUL MILLE
U+0C66	✓	0	CHIFFRE TÉLOUGOU ZÉRO
U+0C67	✓	1	CHIFFRE TÉLOUGOU UN
U+0C68	✓	2	CHIFFRE TÉLOUGOU DEUX
U+0C69	✓	3	CHIFFRE TÉLOUGOU TROIS

Tableau 4-6. Propriétés numériques

Valeur	Décimal	Nombre	Nom
U+0C6A	✓	4	CHIFFRE TÉLOUGOU QUATRE
U+0C6B	✓	5	CHIFFRE TÉLOUGOU CINQ
U+0C6C	✓	6	CHIFFRE TÉLOUGOU SIX
U+0C6D	✓	7	CHIFFRE TÉLOUGOU SEPT
U+0C6E	✓	8	CHIFFRE TÉLOUGOU HUIT
U+0C6F	✓	9	CHIFFRE TÉLOUGOU NEUF
U+0CE6	✓	0	CHIFFRE KANNARA ZÉRO
U+0CE7	✓	1	CHIFFRE KANNARA UN
U+0CE8	✓	2	CHIFFRE KANNARA DEUX
U+0CE9	✓	3	CHIFFRE KANNARA TROIS
U+0CEA	✓	4	CHIFFRE KANNARA QUATRE
U+0CEB	✓	5	CHIFFRE KANNARA CINQ
U+0CEC	✓	6	CHIFFRE KANNARA SIX
U+0CED	✓	7	CHIFFRE KANNARA SEPT
U+0CEE	✓	8	CHIFFRE KANNARA HUIT
U+0CEF	✓	9	CHIFFRE KANNARA NEUF
U+0D66	✓	0	CHIFFRE MALAYALAM ZÉRO
U+0D67	✓	1	CHIFFRE MALAYALAM UN
U+0D68	✓	2	CHIFFRE MALAYALAM DEUX
U+0D69	✓	3	CHIFFRE MALAYALAM TROIS
U+0D6A	✓	4	CHIFFRE MALAYALAM QUATRE
U+0D6B	✓	5	CHIFFRE MALAYALAM CINQ
U+0D6C	✓	6	CHIFFRE MALAYALAM SIX
U+0D6D	✓	7	CHIFFRE MALAYALAM SEPT
U+0D6E	✓	8	CHIFFRE MALAYALAM HUIT
U+0D6F	✓	9	CHIFFRE MALAYALAM NEUF
U+0E50	✓	0	CHIFFRE THAÏ ZÉRO
U+0E51	✓	1	CHIFFRE THAÏ UN
U+0E52	✓	2	CHIFFRE THAÏ DEUX
U+0E53	✓	3	CHIFFRE THAÏ TROIS
U+0E54	✓	4	CHIFFRE THAÏ QUATRE
U+0E55	✓	5	CHIFFRE THAÏ CINQ
U+0E56	✓	6	CHIFFRE THAÏ SIX
U+0E57	✓	7	CHIFFRE THAÏ SEPT
U+0E58	✓	8	CHIFFRE THAÏ HUIT
U+0E59	✓	9	CHIFFRE THAÏ NEUF
U+0ED0	✓	0	CHIFFRE LAOTIEN ZÉRO
U+0ED1	✓	1	CHIFFRE LAOTIEN UN
U+0ED2	✓	2	CHIFFRE LAOTIEN DEUX
U+0ED3	✓	3	CHIFFRE LAOTIEN TROIS
U+0ED4	✓	4	CHIFFRE LAOTIEN QUATRE
U+0ED5	✓	5	CHIFFRE LAOTIEN CINQ
U+0ED6	✓	6	CHIFFRE LAOTIEN SIX
U+0ED7	✓	7	CHIFFRE LAOTIEN SEPT
U+0ED8	✓	8	CHIFFRE LAOTIEN HUIT
U+0ED9	✓	9	CHIFFRE LAOTIEN NEUF
U+0F20	✓	0	CHIFFRE TIBÉTAIN ZÉRO
U+0F21	✓	1	CHIFFRE TIBÉTAIN UN
U+0F22	✓	2	CHIFFRE TIBÉTAIN DEUX
U+0F23	✓	3	CHIFFRE TIBÉTAIN TROIS
U+0F24	✓	4	CHIFFRE TIBÉTAIN QUATRE
U+0F25	✓	5	CHIFFRE TIBÉTAIN CINQ

Tableau 4-6. Propriétés numériques

Valeur	Décimal	Nombre	Nom
U+0F26	✓	6	CHIFFRE TIBÉTAIN SIX
U+0F27	✓	7	CHIFFRE TIBÉTAIN SEPT
U+0F28	✓	8	CHIFFRE TIBÉTAIN HUIT
U+0F29	✓	9	CHIFFRE TIBÉTAIN NEUF
U+0F2A		1/2	CHIFFRE TIBÉTAIN UN MOINS UN DEMI
U+0F2B		3/2	CHIFFRE TIBÉTAIN DEUX MOINS UN DEMI
U+0F2C		5/2	CHIFFRE TIBÉTAIN TROIS MOINS UN DEMI
U+0F2D		7/2	CHIFFRE TIBÉTAIN QUATRE MOINS UN DEMI
U+0F2E		9/2	CHIFFRE TIBÉTAIN CINQ MOINS UN DEMI
U+0F2F		11/2	CHIFFRE TIBÉTAIN SIX MOINS UN DEMI
U+0F30		13/2	CHIFFRE TIBÉTAIN SEPT MOINS UN DEMI
U+0F31		15/2	CHIFFRE TIBÉTAIN HUIT MOINS UN DEMI
U+0F32		17/2	CHIFFRE TIBÉTAIN NEUF MOINS UN DEMI
U+0F33		-1/2	CHIFFRE TIBÉTAIN ZÉRO MOINS UN DEMI
U+1040	✓	0	CHIFFRE BIRMAN ZÉRO
U+1041	✓	1	CHIFFRE BIRMAN UN
U+1042	✓	2	CHIFFRE BIRMAN DEUX
U+1043	✓	3	CHIFFRE BIRMAN TROIS
U+1044	✓	4	CHIFFRE BIRMAN QUATRE
U+1045	✓	5	CHIFFRE BIRMAN CINQ
U+1046	✓	6	CHIFFRE BIRMAN SIX
U+1047	✓	7	CHIFFRE BIRMAN SEPT
U+1048	✓	8	CHIFFRE BIRMAN HUIT
U+1049	✓	9	CHIFFRE BIRMAN NEUF
U+1369	✓	1	CHIFFRE ÉTHIOPIEN UN
U+136A	✓	2	CHIFFRE ÉTHIOPIEN DEUX
U+136B	✓	3	CHIFFRE ÉTHIOPIEN TROIS
U+136C	✓	4	CHIFFRE ÉTHIOPIEN QUATRE
U+136D	✓	5	CHIFFRE ÉTHIOPIEN CINQ
U+136E	✓	6	CHIFFRE ÉTHIOPIEN SIX
U+136F	✓	7	CHIFFRE ÉTHIOPIEN SEPT
U+1370	✓	8	CHIFFRE ÉTHIOPIEN HUIT
U+1371	✓	9	CHIFFRE ÉTHIOPIEN NEUF
U+1372		10	NOMBRE ÉTHIOPIEN DIX
U+1373		20	NOMBRE ÉTHIOPIEN VINGT
U+1374		30	NOMBRE ÉTHIOPIEN TRENTE
U+1375		40	NOMBRE ÉTHIOPIEN QUARANTE
U+1376		50	NOMBRE ÉTHIOPIEN CINQUANTE
U+1377		60	NOMBRE ÉTHIOPIEN SOIXANTE
U+1378		70	NOMBRE ÉTHIOPIEN SOIXANTE-DIX
U+1379		80	NOMBRE ÉTHIOPIEN QUATRE-VINGTS
U+137A		90	NOMBRE ÉTHIOPIEN QUATRE-VINGT-DIX
U+137B		100	NOMBRE ÉTHIOPIEN CENT
U+137C		10 000	NOMBRE ÉTHIOPIEN DIX MILLE
U+16EE		17	SYMBOLE RUNIQUE ARLAUG
U+16EF		18	SYMBOLE RUNIQUE TVIMADUR
U+16F0		19	SYMBOLE RUNIQUE BELGTHOR
U+17E0	✓	0	CHIFFRE KHMER ZÉRO
U+17E1	✓	1	CHIFFRE KHMER UN
U+17E2	✓	2	CHIFFRE KHMER DEUX
U+17E3	✓	3	CHIFFRE KHMER TROIS
U+17E4	✓	4	CHIFFRE KHMER QUATRE

Tableau 4-6. Propriétés numériques

Valeur	Décimal	Nombre	Nom
U+17E5	✓	5	CHIFFRE KHMER CINQ
U+17E6	✓	6	CHIFFRE KHMER SIX
U+17E7	✓	7	CHIFFRE KHMER SEPT
U+17E8	✓	8	CHIFFRE KHMER HUIT
U+17E9	✓	9	CHIFFRE KHMER NEUF
U+1810	✓	0	CHIFFRE MONGOL ZÉRO
U+1811	✓	1	CHIFFRE MONGOL UN
U+1812	✓	2	CHIFFRE MONGOL DEUX
U+1813	✓	3	CHIFFRE MONGOL TROIS
U+1814	✓	4	CHIFFRE MONGOL QUATRE
U+1815	✓	5	CHIFFRE MONGOL CINQ
U+1816	✓	6	CHIFFRE MONGOL SIX
U+1817	✓	7	CHIFFRE MONGOL SEPT
U+1818	✓	8	CHIFFRE MONGOL HUIT
U+1819	✓	9	CHIFFRE MONGOL NEUF
U+2070		0	EXPOSANT ZÉRO
U+2074		4	EXPOSANT QUATRE
U+2075		5	EXPOSANT CINQ
U+2076		6	EXPOSANT SIX
U+2077		7	EXPOSANT SEPT
U+2078		8	EXPOSANT HUIT
U+2079		9	EXPOSANT NEUF
U+2080		0	INDICE ZÉRO
U+2081		1	INDICE UN
U+2082		2	INDICE DEUX
U+2083		3	INDICE TROIS
U+2084		4	INDICE QUATRE
U+2085		5	INDICE CINQ
U+2086		6	INDICE SIX
U+2087		7	INDICE SEPT
U+2088		8	INDICE HUIT
U+2089		9	INDICE NEUF
U+2153		1/3	FRACTION UN TIERS
U+2154		2/3	FRACTION DEUX TIERS
U+2155		1/5	FRACTION UN CINQUIÈME
U+2156		2/5	FRACTION DEUX CINQUIÈMES
U+2157		3/5	FRACTION TROIS CINQUIÈMES
U+2158		4/5	FRACTION QUATRE CINQUIÈMES
U+2159		1/6	FRACTION UN SIXIÈME
U+215A		5/6	FRACTION CINQ SIXIÈMES
U+215B		1/8	FRACTION UN HUITIÈME
U+215C		3/8	FRACTION TROIS HUITIÈMES
U+215D		5/8	FRACTION CINQ HUITIÈMES
U+215E		7/8	FRACTION SEPT HUITIÈMES
U+215F		1	NUMÉRATEUR FRACTIONNAIRE UN
U+2160		1	CHIFFRE ROMAIN UN
U+2161		2	CHIFFRE ROMAIN DEUX
U+2162		3	CHIFFRE ROMAIN TROIS
U+2163		4	CHIFFRE ROMAIN QUATRE
U+2164		5	CHIFFRE ROMAIN CINQ
U+2165		6	CHIFFRE ROMAIN SIX
U+2166		7	CHIFFRE ROMAIN SEPT

Tableau 4-6. Propriétés numériques

Valeur	Décimal	Nombre	Nom
U+2167		8	CHIFFRE ROMAIN HUIT
U+2168		9	CHIFFRE ROMAIN NEUF
U+2169		10	CHIFFRE ROMAIN DIX
U+216A		11	CHIFFRE ROMAIN ONZE
U+216B		12	CHIFFRE ROMAIN DOUZE
U+216C		50	CHIFFRE ROMAIN CINQUANTE
U+216D		100	CHIFFRE ROMAIN CENT
U+216E		500	CHIFFRE ROMAIN CINQ CENTS
U+216F		1 000	CHIFFRE ROMAIN MILLE
U+2170		1	CHIFFRE ROMAIN MINUSCULE UN
U+2171		2	CHIFFRE ROMAIN MINUSCULE DEUX
U+2172		3	CHIFFRE ROMAIN MINUSCULE TROIS
U+2173		4	CHIFFRE ROMAIN MINUSCULE QUATRE
U+2174		5	CHIFFRE ROMAIN MINUSCULE CINQ
U+2175		6	CHIFFRE ROMAIN MINUSCULE SIX
U+2176		7	CHIFFRE ROMAIN MINUSCULE SEPT
U+2177		8	CHIFFRE ROMAIN MINUSCULE HUIT
U+2178		9	CHIFFRE ROMAIN MINUSCULE NEUF
U+2179		10	CHIFFRE ROMAIN MINUSCULE DIX
U+217A		11	CHIFFRE ROMAIN MINUSCULE ONZE
U+217B		12	CHIFFRE ROMAIN MINUSCULE DOUZE
U+217C		50	CHIFFRE ROMAIN MINUSCULE CINQUANTE
U+217D		100	CHIFFRE ROMAIN MINUSCULE CENT
U+217E		500	CHIFFRE ROMAIN MINUSCULE CINQ CENTS
U+217F		1 000	CHIFFRE ROMAIN MINUSCULE MILLE
U+2180		1 000	CHIFFRE ROMAIN MILLE C D
U+2181		5 000	CHIFFRE ROMAIN CINQ MILLE
U+2182		10 000	CHIFFRE ROMAIN DIX MILLE
U+2183		-	CHIFFRE ROMAIN CENT CULBUTÉ
U+2460		1	CHIFFRE UN CERCLÉ
U+2461		2	CHIFFRE DEUX CERCLÉ
U+2462		3	CHIFFRE TROIS CERCLÉ
U+2463		4	CHIFFRE QUATRE CERCLÉ
U+2464		5	CHIFFRE CINQ CERCLÉ
U+2465		6	CHIFFRE SIX CERCLÉ
U+2466		7	CHIFFRE SEPT CERCLÉ
U+2467		8	CHIFFRE HUIT CERCLÉ
U+2468		9	CHIFFRE NEUF CERCLÉ
U+2469		10	NOMBRE DIX CERCLÉ
U+246A		11	NOMBRE ONZE CERCLÉ
U+246B		12	NOMBRE DOUZE CERCLÉ
U+246C		13	NOMBRE TREIZE CERCLÉ
U+246D		14	NOMBRE QUATORZE CERCLÉ
U+246E		15	NOMBRE QUINZE CERCLÉ
U+246F		16	NOMBRE SEIZE CERCLÉ
U+2470		17	NOMBRE DIX-SEPT CERCLÉ
U+2471		18	NOMBRE DIX-HUIT CERCLÉ
U+2472		19	NOMBRE DIX-NEUF CERCLÉ
U+2473		20	NOMBRE VINGT CERCLÉ
U+2474		1	CHIFFRE UN ENTRE PARENTHÈSES
U+2475		2	CHIFFRE DEUX ENTRE PARENTHÈSES
U+2476		3	CHIFFRE TROIS ENTRE PARENTHÈSES

Tableau 4-6. Propriétés numériques

Valeur	Décimal	Nombre	Nom
U+2477		4	CHIFFRE QUATRE ENTRE PARENTHÈSES
U+2478		5	CHIFFRE CINQ ENTRE PARENTHÈSES
U+2479		6	CHIFFRE SIX ENTRE PARENTHÈSES
U+247A		7	CHIFFRE SEPT ENTRE PARENTHÈSES
U+247B		8	CHIFFRE HUIT ENTRE PARENTHÈSES
U+247C		9	CHIFFRE NEUF ENTRE PARENTHÈSES
U+247D		10	NOMBRE DIX ENTRE PARENTHÈSES
U+247E		11	NOMBRE ONZE ENTRE PARENTHÈSES
U+247F		12	NOMBRE DOUZE ENTRE PARENTHÈSES
U+2480		13	NOMBRE TREIZE ENTRE PARENTHÈSES
U+2481		14	NOMBRE QUATORZE ENTRE PARENTHÈSES
U+2482		15	NOMBRE QUINZE ENTRE PARENTHÈSES
U+2483		16	NOMBRE SEIZE ENTRE PARENTHÈSES
U+2484		17	NOMBRE DIX-SEPT ENTRE PARENTHÈSES
U+2485		18	NOMBRE DIX-HUIT ENTRE PARENTHÈSES
U+2486		19	NOMBRE DIX-NEUF ENTRE PARENTHÈSES
U+2487		20	NOMBRE VINGT ENTRE PARENTHÈSES
U+2488		1	CHIFFRE UN POINT
U+2489		2	CHIFFRE DEUX POINT
U+248A		3	CHIFFRE TROIS POINT
U+248B		4	CHIFFRE QUATRE POINT
U+248C		5	CHIFFRE CINQ POINT
U+248D		6	CHIFFRE SIX POINT
U+248E		7	CHIFFRE SEPT POINT
U+248F		8	CHIFFRE HUIT POINT
U+2490		9	CHIFFRE NEUF POINT
U+2491		10	NOMBRE DIX POINT
U+2492		11	NOMBRE ONZE POINT
U+2493		12	NOMBRE DOUZE POINT
U+2494		13	NOMBRE TREIZE POINT
U+2495		14	NOMBRE QUATORZE POINT
U+2496		15	NOMBRE QUINZE POINT
U+2497		16	NOMBRE SEIZE POINT
U+2498		17	NOMBRE DIX-SEPT POINT
U+2499		18	NOMBRE DIX-HUIT POINT
U+249A		19	NOMBRE DIX-NEUF POINT
U+249B		20	NOMBRE VINGT POINT
U+24EA		0	CHIFFRE ZÉRO CERCLÉ
U+2776		1	VIGNETTE CHIFFRE UN NÉGATIF CERCLÉ
U+2777		2	VIGNETTE CHIFFRE DEUX NÉGATIF CERCLÉ
U+2778		3	VIGNETTE CHIFFRE TROIS NÉGATIF CERCLÉ
U+2779		4	VIGNETTE CHIFFRE QUATRE NÉGATIF CERCLÉ
U+277A		5	VIGNETTE CHIFFRE CINQ NÉGATIF CERCLÉ
U+277B		6	VIGNETTE CHIFFRE SIX NÉGATIF CERCLÉ
U+277C		7	VIGNETTE CHIFFRE SEPT NÉGATIF CERCLÉ
U+277D		8	VIGNETTE CHIFFRE HUIT NÉGATIF CERCLÉ
U+277E		9	VIGNETTE CHIFFRE NEUF NÉGATIF CERCLÉ
U+277F		10	VIGNETTE NOMBRE DIX NÉGATIF CERCLÉ
U+2780		1	VIGNETTE CHIFFRE UN SANS EMPATTEMENT CERCLÉ
U+2781		2	VIGNETTE CHIFFRE DEUX SANS EMPATTEMENT CERCLÉ
U+2782		3	VIGNETTE CHIFFRE TROIS SANS EMPATTEMENT CERCLÉ
U+2783		4	VIGNETTE CHIFFRE QUATRE SANS EMPATTEMENT CERCLÉ

Tableau 4-6. Propriétés numériques

Valeur	Décimal	Nombre	Nom
U+2784		5	VIGNETTE CHIFFRE CINQ SANS EMPATTEMENT CERCLÉ
U+2785		6	VIGNETTE CHIFFRE SIX SANS EMPATTEMENT CERCLÉ
U+2786		7	VIGNETTE CHIFFRE SEPT SANS EMPATTEMENT CERCLÉ
U+2787		8	VIGNETTE CHIFFRE HUIT SANS EMPATTEMENT CERCLÉ
U+2788		9	VIGNETTE CHIFFRE NEUF SANS EMPATTEMENT CERCLÉ
U+2789		10	VIGNETTE NOMBRE DIX SANS EMPATTEMENT CERCLÉ
U+278A		1	VIGNETTE CHIFFRE UN SANS EMPATTEMENT NÉGATIF CERCLÉ
U+278B		2	VIGNETTE CHIFFRE DEUX SANS EMPATTEMENT NÉGATIF CERCLÉ
U+278C		3	VIGNETTE CHIFFRE TROIS SANS EMPATTEMENT NÉGATIF CERCLÉ
U+278D		4	VIGNETTE CHIFFRE QUATRE SANS EMPATTEMENT NÉGATIF CERCLÉ
U+278E		5	VIGNETTE CHIFFRE CINQ SANS EMPATTEMENT NÉGATIF CERCLÉ
U+278F		6	VIGNETTE CHIFFRE SIX SANS EMPATTEMENT NÉGATIF CERCLÉ
U+2790		7	VIGNETTE CHIFFRE SEPT SANS EMPATTEMENT NÉGATIF CERCLÉ
U+2791		8	VIGNETTE CHIFFRE HUIT SANS EMPATTEMENT NÉGATIF CERCLÉ
U+2792		9	VIGNETTE CHIFFRE NEUF SANS EMPATTEMENT NÉGATIF CERCLÉ
U+2793		10	VIGNETTE NOMBRE DIX SANS EMPATTEMENT NÉGATIF CERCLÉ
U+3007		0	ZÉRO IDÉOGRAPHIQUE
U+3021		1	CHIFFRE UN HANGZHOU
U+3022		2	CHIFFRE DEUX HANGZHOU
U+3023		3	CHIFFRE TROIS HANGZHOU
U+3024		4	CHIFFRE QUATRE HANGZHOU
U+3025		5	CHIFFRE CINQ HANGZHOU
U+3026		6	CHIFFRE SIX HANGZHOU
U+3027		7	CHIFFRE SEPT HANGZHOU
U+3028		8	CHIFFRE HUIT HANGZHOU
U+3029		9	CHIFFRE NEUF HANGZHOU
U+3038		10	NOMBRE DIX HANGZHOU
U+3039		20	NOMBRE VINGT HANGZHOU
U+303A		30	NOMBRE TRENTE HANGZHOU
U+3192		1	MARQUE D'ANNOTATION IDÉOGRAPHIQUE UN
U+3193		2	MARQUE D'ANNOTATION IDÉOGRAPHIQUE DEUX
U+3194		3	MARQUE D'ANNOTATION IDÉOGRAPHIQUE TROIS
U+3195		4	MARQUE D'ANNOTATION IDÉOGRAPHIQUE QUATRE
U+3220		1	IDÉOGRAMME UN ENTRE PARENTHÈSES
U+3221		2	IDÉOGRAMME DEUX ENTRE PARENTHÈSES
U+3222		3	IDÉOGRAMME TROIS ENTRE PARENTHÈSES
U+3223		4	IDÉOGRAMME QUATRE ENTRE PARENTHÈSES
U+3224		5	IDÉOGRAMME CINQ ENTRE PARENTHÈSES
U+3225		6	IDÉOGRAMME SIX ENTRE PARENTHÈSES
U+3226		7	IDÉOGRAMME SEPT ENTRE PARENTHÈSES
U+3227		8	IDÉOGRAMME HUIT ENTRE PARENTHÈSES
U+3228		9	IDÉOGRAMME NEUF ENTRE PARENTHÈSES

Tableau 4-6. Propriétés numériques

Valeur	Décimal	Nombre	Nom
U+3229		10	IDÉOGRAMME DIX ENTRE PARENTHÈSES
U+3280		1	IDÉOGRAMME UN CERCLÉ
U+3281		2	IDÉOGRAMME DEUX CERCLÉ
U+3282		3	IDÉOGRAMME TROIS CERCLÉ
U+3283		4	IDÉOGRAMME QUATRE CERCLÉ
U+3284		5	IDÉOGRAMME CINQ CERCLÉ
U+3285		6	IDÉOGRAMME SIX CERCLÉ
U+3286		7	IDÉOGRAMME SEPT CERCLÉ
U+3287		8	IDÉOGRAMME HUIT CERCLÉ
U+3288		9	IDÉOGRAMME NEUF CERCLÉ
U+3289		10	IDÉOGRAMME DIX CERCLÉ
U+FF10	✓	0	CHIFFRE ZÉRO PLEINE CHASSE
U+FF11	✓	1	CHIFFRE UN PLEINE CHASSE
U+FF12	✓	2	CHIFFRE DEUX PLEINE CHASSE
U+FF13	✓	3	CHIFFRE TROIS PLEINE CHASSE
U+FF14	✓	4	CHIFFRE QUATRE PLEINE CHASSE
U+FF15	✓	5	CHIFFRE CINQ PLEINE CHASSE
U+FF16	✓	6	CHIFFRE SIX PLEINE CHASSE
U+FF17	✓	7	CHIFFRE SEPT PLEINE CHASSE
U+FF18	✓	8	CHIFFRE HUIT PLEINE CHASSE
U+FF19	✓	9	CHIFFRE NEUF PLEINE CHASSE

Les idéogrammes CJC issus des *Répertoire et classement unifiés*² peuvent également représenter des nombres. Le tableau 4-7 énumère les principaux idéogrammes numériques. Quand ces idéogrammes représentent des nombres en notation décimale, le zéro est représenté par U+3007 〇, sinon on utilise U+96F6 零.

Tableau 4-7. Principaux idéogrammes numériques

Valeur Unicode	Idéogramme	Nombre décimal
U+96F6	零	0
U+4E00	一	1
U+4E8C	二	2
U+4E09	三	3
U+56DB	四	4
U+4E94	五	5
U+516D	六	6
U+4E03	七	7
U+516B	八	8
U+4E5D	九	9
U+5341	十	10
U+767E	百	100
U+5343	千	1 000
U+4E07	万	10 000
U+5104	億	100 000 000 (10 000 x 10 000)
U+4EBF	亿	100 000 000 (10 000 x 10 000)

² Voir Annexe H, *Unification han*.

U+5146 兆 1 000 000 000 000 (10 000 x 10 000 x 10 000)

Les nombres idéographiques comptables sont fréquemment utilisés sur les chèques et autres documents financiers afin de réduire les risques d'équivoque ou de fraude. L'ensemble des nombres comptables varie quelque peu selon la langue (japonais, chinois, coréen). Le tableau 4-8 fournit une liste passablement complète des caractères comptables. Certains caractères sont des idéogrammes polysémiques « enrôlés » par le domaine financier, les autres ne s'utilisent que comme nombres comptables.

Tableau 4-8. Principaux idéogrammes numériques comptables

Nombre	Idéogrammes
1	U+58F9 壹, U+58F1 壹, U+5F0C ^a 弍
2	U+8CAE ^a 貳, U+8D30 ^a 貳, U+5F10 ^a 弍, U+5F0D ^a 弍
3	U+53C3 參, U+53C2 參, U+53C1 ^a 叁, U+5F0E ^a 弍
4	U+8086 肆
5	U+4F0D 伍
6	U+9678 陸, U+9646 陆
7	U+67D2 ^b 柒
8	U+634C 捌
9	U+7396 玖
10	U+62FE 拾
100	U+4F70 ^a 佰, U+964C 陌
1 000	U+4EDF 仟
10 000	U+842C 萬

a. Ces caractères servent uniquement de nombres comptables, ils n'ont pas d'autre sens.

b. Au Japon, U+67D2 柒 se prononce également *urusi* (*ouloussi*), c'est-à-dire « laque » et il constitue une variante du caractère standard pour « laque » 6F06 漆.

4.7 Caractères miroirs – normatif

L'attribut miroir (ou spéculaire) est une propriété normative des caractères comme les parenthèses dont l'œil est réfléchi dans les textes composés de droite à gauche. Ainsi, U+0028 PARENTHÈSE GAUCHE est interprété comme une *parenthèse ouvrante*; dans un contexte gauche-à-droite il s'affiche sous la forme d'une « (» alors que dans un contexte droite-à-gauche il apparaît sous la forme de son œil miroir «) ». Le tableau 4-9 énumère la liste des caractères spéculaires. Il convient de remarquer que l'effet miroir ne s'applique pas uniquement aux caractères appariés mais que tout caractère doté de la propriété miroir aura besoin de deux glyphes miroir. Cette exigence garantit que ce caractère s'affichera correctement dans un contexte bidirectionnel.

Tableau 4-9. Caractères miroirs

Valeur	Glyphe gauche-à-droite	Nom
U+0028	(PARENTHÈSE GAUCHE
U+0029)	PARENTHÈSE DROITE
U+003C	<	SIGNE INFÉRIEUR À
U+003E	>	SIGNE SUPÉRIEUR À
U+005B	[CROCHET GAUCHE

Tableau 4-9. Caractères miroirs

Valeur	Glyphe gauche-à-droite	Nom
U+005D]	CROCHET DROIT
U+007B	{	ACCOLADE GAUCHE
U+007D	}	ACCOLADE DROITE
U+00AB	«	GUILLEMET GAUCHE
U+00BB	»	GUILLEMET DROIT
U+2039	<	GUILLEMET SIMPLE VERS LA GAUCHE
U+203A	>	GUILLEMET SIMPLE VERS LA DROITE
U+2045	[CROCHET GAUCHE AVEC PIQUE
U+2046]	CROCHET DROIT AVEC PIQUE
U+207D	(EXPOSANT PARENTHÈSE GAUCHE
U+207E)	EXPOSANT PARENTHÈSE DROITE
U+208D	(INDICE PARENTHÈSE GAUCHE
U+208E)	INDICE PARENTHÈSE DROITE
U+2201	∅	COMPLÉMENT
U+2202	∂	DÉRIVÉE PARTIELLE
U+2203	∃	IL EXISTE
U+2204	∄	IL N'EXISTE PAS
U+2208	∈	APPARTIENT À
U+2209	∉	N'APPARTIENT PAS À
U+220A	∈	PETIT APPARTIENT À
U+220B	⊃	CONTIENT COMME ÉLÉMENT
U+220C	⊄	NE CONTIENT PAS COMME ÉLÉMENT
U+220D	⊃	PETIT CONTIENT COMME ÉLÉMENT
U+2211	∑	SOMMATION DE LA FAMILLE
U+2215	/	BARRE OBLIQUE DE DIVISION
U+2216	\	DIFFÉRENCE D'ENSEMBLE
U+221A	√	RACINE CARRÉE
U+221B	∛	RACINE CUBIQUE
U+221C	∜	RACINE QUATRIÈME
U+221D	∝	PROPORTIONNEL À
U+221F	⊥	ANGLE DROIT
U+2220	∠	ANGLE
U+2221	∠	ANGLE MESURÉ
U+2222	∠	ANGLE SPHÉRIQUE
U+2224	∤	N'EST PAS UN DIVISEUR DE
U+2226	∦	NON PARALLÈLE À
U+222B	∫	INTÉGRALE
U+222C	∫	INTÉGRALE DOUBLE
U+222D	∫	INTÉGRALE TRIPLE
U+222E	∮	INTÉGRALE DE CONTOUR
U+222F	∯	INTÉGRALE DE SURFACE
U+2230	∭	INTÉGRALE DE VOLUME
U+2231	∫	INTÉGRALE EN SENS NÉGATIF
U+2232	∮	INTÉGRALE DE CONTOUR EN SENS NÉGATIF
U+2233	∮	INTÉGRALE DE CONTOUR EN SENS POSITIF
U+2239	∴	EXCÈS
U+223B	∼	HOMOTHÉTIQUE
U+223C	∼	OPÉRATEUR TILDE
U+223D	∼	OPÉRATEUR TILDE RENVERSÉ
U+223E	∞	S COUCHÉ RENVERSÉ

Tableau 4-9. Caractères miroirs

Valeur	Glyphe gauche- à-droite	Nom
U+223F	≈	SINUSOÏDE
U+2240	∩	PRODUIT COURONNE
U+2241	∴	NON TILDE
U+2242	≈	MOINS TILDE
U+2243	∓	ASYMPTOTIQUEMENT ÉGAL À
U+2244	≠	NON ASYMPTOTIQUEMENT ÉGAL À
U+2245	∓	APPROXIMATIVEMENT ÉGAL À
U+2246	≠	APPROXIMATIVEMENT MAIS NON STRICTEMENT ÉGAL À
U+2247	≠	NI APPROXIMATIVEMENT NI STRICTEMENT ÉGAL À
U+2248	≈	PRESQUE ÉGAL À
U+2249	≠	NON PRESQUE ÉGAL À
U+224A	≈	PRESQUE ÉGAL OU ÉGAL À
U+224B	≈	TRIPLE TILDE
U+224C	∓	ENTIÈREMENT ÉGAL À
U+2252	∓	APPROXIMATIVEMENT ÉGAL À OU IMAGE DE
U+2253	∓	IMAGE DE OU APPROXIMATIVEMENT ÉGAL À
U+2254	∓	DEUX-POINTS ÉGAL
U+2255	∓	ÉGAL DEUX-POINTS
U+225F	∓	ÉGALITÉ EN DOUTE
U+2260	≠	NON ÉGAL À
U+2262	≠	NON IDENTIQUE À
U+2264	∩	PLUS PETIT OU ÉGAL À
U+2265	∪	PLUS GRAND OU ÉGAL À
U+2266	∩	PLUS PETIT QUE PAR-DESSUS ÉGAL À
U+2267	∪	PLUS GRAND QUE PAR-DESSUS ÉGAL À
U+2268	∩	PLUS PETIT MAIS NON ÉGAL À
U+2269	∪	PLUS GRAND MAIS NON ÉGAL À
U+226A	∩	BEAUCOUP PLUS PETIT QUE
U+226B	∪	BEAUCOUP PLUS GRAND QUE
U+226E	∩	PAS PLUS PETIT QUE
U+226F	∪	PAS PLUS GRAND QUE
U+2270	∩	NI PLUS PETIT NI ÉGAL À
U+2271	∪	NI PLUS GRAND NI ÉGAL À
U+2272	∩	PLUS PETIT OU ÉQUIVALENT À
U+2273	∪	PLUS GRAND OU ÉQUIVALENT À
U+2274	∩	NI PLUS PETIT NI ÉQUIVALENT À
U+2275	∪	NI PLUS GRAND NI ÉQUIVALENT À
U+2276	∩	PLUS PETIT OU PLUS GRAND QUE
U+2277	∪	PLUS GRAND OU PLUS PETIT QUE
U+2278	∩	NI PLUS PETIT NI PLUS GRAND QUE
U+2279	∪	NI PLUS GRAND NI PLUS PETIT QUE
U+227A	∩	PRÉCÈDE
U+227B	∪	SUIT
U+227C	∩	PRÉCÈDE OU EST ÉGAL À
U+227D	∪	SUIT OU EST ÉGAL À
U+227E	∩	PRÉCÈDE OU EST ÉQUIVALENT À
U+227F	∪	SUIT OU EST ÉQUIVALENT À
U+2280	∩	NE PRÉCÈDE PAS
U+2281	∪	NE SUIT PAS
U+2282	⊂	SOUS-ENSEMBLE DE

Tableau 4-9. Caractères miroirs

Valeur	Glyphe gauche- à-droite	Nom
U+2283	⊃	SUR-ENSEMBLE DE
U+2284	⊄	PAS UN SOUS-ENSEMBLE DE
U+2285	⊇	PAS UN SUR-ENSEMBLE DE
U+2286	⊆	SOUS-ENSEMBLE OU ÉGAL À
U+2287	⊇	SUR-ENSEMBLE OU ÉGAL À
U+2288	⊈	NI UN SOUS-ENSEMBLE NI ÉGAL À
U+2289	⊉	NI UN SUR-ENSEMBLE NI ÉGAL À
U+228A	⊈	SOUS-ENSEMBLE MAIS NON ÉGAL À
U+228B	⊉	SUR-ENSEMBLE MAIS NON ÉGAL À
U+228C	⊋	MULTIENSEMBLE
U+228F	⊐	IMAGE CARRÉE DE
U+2290	⊑	ORIGINAL CARRÉ DE
U+2291	⊐	IMAGE CARRÉE OU ÉGAL À
U+2292	⊑	ORIGINAL CARRÉ OU ÉGAL À
U+2298	⊘	BARRE OBLIQUE DE DIVISION CERCLÉE
U+22A2	⊢	TAQUET DROIT
U+22A3	⊣	TAQUET GAUCHE
U+22A6	⊤	ASSERTION
U+22A7	⊥	MODÉLISE
U+22A8	⊦	VRAI
U+22A9	⊧	OBLIGE
U+22AA	⊨	BARRE VERTICALE TRIPLE AVEC TOURNIQUET À DROITE
U+22AB	⊩	BARRE VERTICALE DOUBLE AVEC TOURNIQUET À DROITE DOUBLE
U+22AC	⊮	NE PROUVE PAS
U+22AD	⊯	PAS VRAI
U+22AE	⊰	N'OBLIGE PAS
U+22AF	⊱	NON-BARRE VERTICALE DOUBLE AVEC TOURNIQUET À DROITE DOUBLE
U+22B0	⋈	PRÉCÈDE SOUS RELATION
U+22B1	⋉	SUIT SOUS RELATION
U+22B2	⋊	SOUS-GROUPE NORMAL DE
U+22B3	⋋	CONTIENT COMME SOUS-GROUPE NORMAL
U+22B4	⋌	SOUS-GROUPE NORMAL OU ÉGAL À
U+22B5	⋍	CONTIENT COMME SOUS-GROUPE NORMAL OU ÉGAL À
U+22B6	⋎	ORIGINAL DE
U+22B7	⋏	IMAGE DE
U+22B8	⋐	MULTIJECTION
U+22BE	⊞	ANGLE DROIT AVEC ARC
U+22BF	⊟	TRIANGLE RECTANGLE
U+22C9	⊗	PRODUIT SEMI-DIRECT À GAUCHE DE FACTEUR NORMAL
U+22CA	⊘	PRODUIT SEMI-DIRECT À DROITE DE FACTEUR NORMAL
U+22CB	⊗	PRODUIT SEMI-DIRECT À GAUCHE
U+22CC	⊘	PRODUIT SEMI-DIRECT À DROITE
U+22CD	⋚	MOINS TILDE RENVERSÉ
U+22D0	⊕	SOUS-ENSEMBLE DOUBLE
U+22D1	⊖	SUR-ENSEMBLE DOUBLE
U+22D6	⋈	PLUS PETIT QUE POINTÉ
U+22D7	⋉	PLUS GRAND QUE POINTÉ
U+22D8	⋊	CONSIDÉRABLEMENT PLUS PETIT QUE
U+22D9	⋋	CONSIDÉRABLEMENT PLUS GRAND QUE
U+22DA	⋌	PLUS PETIT OU ÉGAL OU PLUS GRAND QUE

Tableau 4-9. Caractères miroirs

Valeur	Glyphe gauche-à-droite	Nom
U+22DB	≧	PLUS GRAND OU ÉGAL OU PLUS PETIT QUE
U+22DC	≦	ÉGAL OU PLUS PETIT QUE
U+22DD	≧	ÉGAL OU PLUS GRAND QUE
U+22DE	⋈	ÉGAL OU PRÉCÈDE
U+22DF	⋉	ÉGAL OU SUIV
U+22E0	⋊	NI PRÉCÉDANT NI ÉGAL À
U+22E1	⋋	NI SUIVANT NI ÉGAL À
U+22E2	⋌	NI IMAGE CARRÉE NI ÉGAL À
U+22E3	⋍	NI ORIGINAL CARRÉ NI ÉGAL À
U+22E4	⋎	IMAGE CARRÉ OU DIFFÉRENT DE
U+22E5	⋏	ORIGINAL CARRÉ OU DIFFÉRENT DE
U+22E6	⋐	PLUS PETIT MAIS NON ÉQUIVALENT À
U+22E7	⋑	PLUS GRAND MAIS NON ÉQUIVALENT À
U+22E8	⋒	PRÉCÉDANT MAIS NON ÉQUIVALENT À
U+22E9	⋓	SUIVANT MAIS NON ÉQUIVALENT À
U+22EA	⋔	PAS UN SOUS-GROUPE NORMAL DE
U+22EB	⋕	NE CONTIENT PAS COMME SOUS-GROUPE NORMAL
U+22EC	⋖	NI SOUS-GROUPE NORMAL NI ÉGAL À
U+22ED	⋗	NI ÉGAL NI CONTENANT COMME SOUS-GROUPE NORMAL
U+22F0	⋘	TROIS POINTS DIAGONAUX VERS LE COIN HAUT À DROITE
U+22F1	⋙	TROIS POINTS DIAGONAUX VERS LE COIN BAS À DROITE
U+2308	┌	PLAFOND À GAUCHE
U+2309	┐	PLAFOND À DROITE
U+230A	└	PLANCHER À GAUCHE
U+230B	┘	PLANCHER À DROITE
U+2320	⌈	MOITIÉ SUPÉRIEURE D'INTÉGRALE
U+2321	⌋	MOITIÉ INFÉRIEURE D'INTÉGRALE
U+2329	⟨	CROCHET VERS LA GAUCHE
U+232A	⟩	CROCHET VERS LA DROITE
U+3008	⟨	CROCHET ANGULAIRE À GAUCHE
U+3009	⟩	CROCHET ANGULAIRE À DROITE
U+300A	⟨⟨	DOUBLE CROCHET ANGULAIRE À GAUCHE
U+300B	⟩⟩	DOUBLE CROCHET ANGULAIRE À DROITE
U+300C	┌	CROCHET COIN GAUCHE
U+300D	┐	CROCHET COIN DROIT
U+300E	┌	CROCHET BLANC COIN GAUCHE
U+300F	┐	CROCHET BLANC COIN DROIT
U+3010	【	CROCHET NOIR LENTICULAIRE GAUCHE
U+3011	】	CROCHET NOIR LENTICULAIRE DROIT
U+3014	{	CROCHET GAUCHE EN ÉCAILLE
U+3015	}	CROCHET DROIT EN ÉCAILLE
U+3016	[CROCHET BLANC LENTICULAIRE GAUCHE
U+3017]	CROCHET BLANC LENTICULAIRE DROIT
U+3018	{	CROCHET BLANC GAUCHE EN ÉCAILLE
U+3019	}	CROCHET BLANC DROIT EN ÉCAILLE
U+301A	[CROCHET BLANC GAUCHE
U+301B]	CROCHET BLANC DROIT

4.8 Propriété mathématique – informatif

La *propriété mathématique* est une propriété informative des caractères qui peuvent servir d'opérateurs dans les formules mathématiques. La propriété mathématique peut s'avérer utile pour les algorithmes qui affichent du texte mathématique ou des formules. Cependant, un certain nombre de ces caractères peuvent s'employer de plusieurs façons et peuvent avoir une sémantique non mathématique. Ainsi, 002D - TRAIT D'UNION-SIGNE MOINS peut aussi servir de trait d'union — et non, dans ce cas, de signe moins. D'autres caractères (parmi lesquels certaines lettres, nombres, signes de ponctuation, espaces, flèches et symboles géométriques) s'utilisent également parfois dans des expressions mathématiques, mais leur identité dépend encore plus du contexte. Les caractères Unicode ci-dessous ont la propriété mathématique.

Caractères à propriété mathématique et membre de la catégorie générale Sm :

002B, 003C..003E, 007C, 007E, 00AC, 00B1, 00D7, 00F7, 2044,
207A..207C, 208A..208C, 2190..2194, 219A..219B, 21A0, 21A3, 21A6,
21AE, 21CE..21CF, 21D2, 21D4, 2200..22F1, 2308..230B, 2320..2321,
25B7, 25C1, 266F, FB29, FE62, FE64..FE66, FF0B, FF1C..FF1E, FF5C,
FF5E, FFE2, FFE9..FFEC

Caractères à propriété mathématique et membre d'autre catégorie générale que Sm :

0028..002A, 002D, 002F, 005B..005E, 007B, 007D, 2016, 2032..2034,
207D..207E, 208D..208E, 20D0..20DC, 20E1, 2329..232A, 300A..300B,
301A..301B, FE35..FE38, FE59..FE5C, FE61, FE63, FE68, FF08..FF0A,
FF0D, FF0F, FF3B..FF3E, FF5B, FF5D

4.9 Lettres et autres propriétés utiles

Le disque optique qui accompagne cet ouvrage contient des fichiers de données qui reprennent d'autres propriétés informatives des caractères Unicode utiles. La liste complète de ces propriétés se trouve dans les fichiers de données ; consultez plus particulièrement *PropList.txt*. Cette section se penche sur certaines de ces propriétés qui peuvent avoir une incidence sur des questions de mise en œuvre comme l'analyse des identificateurs (voir la section 5.16, *Identificateurs*).

Les langages informatiques décrivent souvent les identificateurs comme étant constitué de lettres, de syllabes, d'idéogrammes et de chiffres sans définir précisément ce qu'est une « lettre », une « syllabe », un « idéogramme » ou un « chiffre » et laissent implicitement ces définitions soit à un jeu de caractères normé ou une valeur de la « locale ». L'ampleur du standard Unicode et le grand nombre de systèmes d'écriture impliqués signifient qu'il n'est pas toujours aussi facile de départager ces notions, chose qui était relativement aisée pour des systèmes uniquement conçus pour les langues d'Europe occidentale ou le japonais. Plus particulièrement, alors que le standard Unicode comprend plusieurs « alphabets » et « syllabaires », il comprend également des systèmes d'écriture qui se situent quelque part entre les deux. C'est pourquoi le standard ne tente pas d'opposer les lettres aux syllabes au niveau des propriétés de caractère.

Lettre. Cette propriété informative s'applique aux caractères utilisés pour écrire des mots. Ce groupe comprend des caractères comme les lettres majuscules, les minuscules, les idéogrammes, les syllabes hangûl et les lettres modificatives avec chasse. Les caractères

combinatoires adoptent, en règle générale, la propriété « lettre » du caractère de base précédent. Ainsi, lors de la recherche des frontières de mot, ne sépare-t-on pas les caractères combinatoires de leur caractère de base.

Alphabétique. La propriété dite alphabétique est une propriété informative des unités de base qui forment les alphabets ou les syllabaires, que ces unités soient combinatoires ou non. Ce groupe comprend les digrammes, les variantes contextuelles, les ligatures de caractères alphabétiques, les variantes contextuelles de ces ligatures, les lettres modificatives, les symboles de type lettre qui sont des équivalents compatibles de lettres alphabétiques isolées, les caractères composites dont la décomposition canonique comprend un caractère de base alphabétique suivi d'au moins un caractère combinatoire ainsi que, finalement, divers éléments de lettre. Les caractères U+00AA ^a INDICATEUR ORDINAL FÉMININ et U+00BA ^o INDICATEUR, ORDINAL MASCULIN, notamment, sont simplement des formes abrégées qui utilisent une lettre latine et font donc partie des symboles alphabétiques.

Idéographique. La propriété idéographique est une propriété informative de l'ensemble des idéogrammes unifiés CJC (U+4E00..U+9FA5), du supplément A aux idéogrammes unifiés CJC (U+3400..U+4DB5), de l'ensemble des idéogrammes de compatibilités CJC (U+F900..U+FA2D), du ZÉRO IDÉOGRAPHIQUE (U+3007), de la MARQUE IDÉOGRAPHIQUE DE FERMETURE (U+3006) et des nombres à la façon de Hang-tcheou ou Hangzhou (U+3021..U+3029, U+3038..U+303A).

Stabilité des identificateurs. Unicode tente dans la mesure du possible de conserver les valeurs associées à la Catégorie Générale ; il se peut néanmoins que certaines valeurs changent au point d'affecter la définition des identificateurs dans des versions ultérieures. Il est donc possible donc que d'autres normes ou produits, lorsqu'ils passent à cette nouvelle version d'Unicode, doivent traiter des caractères qui ne font désormais plus partie de ID_Start (début de l'identificateur) ou de ID_Continue (suite de l'identificateur).

On peut remédier à cette difficulté en utilisant deux listes explicitement compatibles vers le bas : Supplément_Début_ID et Supplément_Suite_ID. Les spécifications de mise en œuvre des identificateurs doivent alors faire référence à l'union des propriétés Unicode correspondantes et de ces listes supplémentaires.