

Proposer un contenu

Identifiant

Identifiant

Mot de passe

Mot de passe

☐ Connexion automatique

Se connecter

Pas de compte ? S'inscrire...

Bonjour les cornichons, Unicode vient de sortir en version 18

Posté par Ysabeau (site web personnel, Mastodon) le 21 septembre 2026 à 14:02. Édité par BAud, Benoît Sibaud et Pierre Jarillon. Modéré par patrick.g. Licence CC By-SA.

Étiquettes : écriture, unicode



Le standard Unicode a ajouté dans son bocal : 13 007 caractères. Il en contient donc aujourd'hui 172 808. Ce n'est pas une version révolutionnaire. Les modifications sont principalement liées à l'ajout des nouveaux systèmes d'écritures et caractères afin de pouvoir s'adapter à leur comportement, comme pour la [version précédente](#). Il n'y a donc pas de nouvel algorithme. Outre trois symboles monétaires, on voit arriver 11 328 idéogrammes d'écritures chinoises anciennes, le petit sceau et les sinogrammes Jurchen. Unicode 18 accueille aussi en son sein des chiffres proto-cunéiformes ou encore des symboles musicaux. Et bien sûr, des émojis.

Qu'écouter à la lecture de cette dépêche ? [Stéphane Bortzmeyer](#) suggère la chanson de Nino Ferrer « Les cornichons ». Vous avez le choix entre la version originale [chantée](#) ou une purement [instrumentale](#).



Sommaire

- [Des symboles monétaires](#)
 - [Des systèmes d'écritures](#)
 - [L'écriture petit sceau](#)
 - [L'écriture Jürchen](#)
 - [Les nombres proto-cunéiformes](#)
 - [Le bengali](#)
- [Aspect « technique »](#)
- [Des émojis et des symboles divers](#)
- [Remerciement et mots de la fin](#)

Selon votre navigateur, vous verrez, ou pas, les caractères s'afficher. Ils finiront très probablement par l'être un jour, de toute façon, raison pour laquelle ils figurent à côté de leur numéro Unicode.

Des symboles monétaires

Trois symboles monétaires ont été ajoutés. Dans sa [note de blog \(EN\)](#), le consortium Unicode les considère comme les caractères les plus attendus, à savoir :

- la rufyaa ou roupie maldivienne, MVR ou Rf, devise des Maldives, U+20C2, 𑖪 ;
- le D avec une double barre du dirham des Émirats arabes unies, AED, U+20C3, 𑄎 ;
- et le rial omanais, OMR, une boucle sur deux barres, monnaie officielle du sultanat d'Oman, U+20C4, 𑄌.

Des systèmes d'écritures

Les nouveautés d'Unicode 18 en matière d'écriture sont des anciennes, voire, des très anciennes écritures. Les références AEC signifient Avant l'Ère Commune, donc avec des périodes s'écrivant dans le sens inverse de la datation actuelle : du plus grand nombre au plus petit.

L'écriture petit sceau

L'écriture petit sceau ou *xiǎo zhuàn* en chinois est une [écriture sigillaire](#). En d'autres termes elle est basée sur les divers sceaux utilisés pour authentifier les différents actes et documents officiels mais aussi pour les particuliers et, évidemment, les artistes.

Elle fait partie des [six styles](#) de l'écriture chinoise¹ :

- le *grand sceau* ou *dà zhuàn* ;
- le *xiaozhuan* qui est une forme de l'écriture grand sceau ;
- le *petit sceau* qui nous occupe aujourd'hui ;
- le *lì shū*, écriture classique des documents administratifs des dynasties Qin et Han ;
- le *kǎi shū*, style régulier qui est la norme de l'écriture chinoise ;
- le *xíng shū* qui est une écriture semi-cursive.

L'écriture petit sceau a commencé à être utilisé sous la dynastie Qin, 221-207 AEC, et ce, pendant deux cents ans. C'est sous cette dynastie et sous l'impulsion du pouvoir impérial que l'écriture chinoise se normalisera. L'Unicode vient d'intégrer 11 328 caractères dans les blocs Unicode 3D000 à 3FC3F.

L'écriture Jürchen

Chinoise elle aussi, l'écriture [Jürchen](#) date de la dynastie Jin et du XII^e siècle. Elle comporte neuf cents caractères et n'est actuellement que partiellement déchiffrée.

Les Jürchen était un peuple du nord-est de la Chine. Le chef Jürchen Wanyan Aguda (1068-1123) devenu le premier empereur Jürchen en 1115 demande en 1119 à Wanyan Xiyin (?-1140) de travailler à un système d'écriture à la hauteur d'un empire. Ce sera donc l'écriture Jürchen qui occupe les blocs Unicode 18E00 à 1919F et 191A0 à 191DF pour les radicaux².

On pourrait écrire une histoire politique de la Chine à travers celle de son écriture.

Les nombres proto-cunéiformes

Le proto-cunéiforme ou cunéiforme archaïque est apparu en Mésopotamie vers 3 200 (voire, 3 400) AEC. C'est, probablement, le tout premier système d'écriture, tout du moins, c'est le système d'écriture le plus ancien dont on ait la trace. Il évoluera vers le cunéiforme aux alentours du deuxième millénaire. Ce dernier se répandra largement notamment grâce aux échanges commerciaux.

L'écriture se faisait sur des tablettes d'argile à l'aide de stylets, la forme des lettres rappelant celle des clous : une ligne avec une tête. Alors que le proto-cunéiforme était plus figuratif, le cunéiforme devient plus abstrait.

Le cunéiforme occupait déjà les blocs Unicode 12400 à 1247F. Les chiffres proto-cunéiforme, font leur entrée dans les blocs Unicode 12550 à 1268F.



Contrat de vente en proto-cunéiforme. La tablette est écrite en cases et se lit verticalement. Shuruppak, vers 2600 AEC. Musée du Louvre, photo Marie-Lan Nguyen.

Le bengali

Le Bengali est la langue nationale du Bangladesh. Il fait partie des vingt-deux langues officielles de l'Inde. C'est la langue des états du Bengale-occidental et de Tripura. Selon [l'Inalco](#) il y aurait dans le monde 250 millions de locuteurs et de locutrices de cette langue.

Il figurait déjà dans les blocs Unicode 0980 à 09FF. La version 18 vient d'ajouter un supplément : 11DF0 à 11DFF.

Aspect « technique »

Les modifications plus « techniques » concernent presque essentiellement la prise en compte de ces nouveaux systèmes d'écritures avec l'ajout de fichiers de données.

À cela se rajoutent des recommandations destinées aux implémentations pour rendre plus visibles les utilisations non conformes des sélecteurs de variation. Le langage de conformité de ces derniers a été mis à jour.

Des émojis et des symboles divers

Vous avez sans doute deviné au moins un des nouveaux émojis de cette version, on a donc :

- un cornichon, pickle, U+1FADD, 🥒, à ne pas confondre avec le concombre, U+1F952, 🥒 qui existait déjà ;
- une face fissurée, cracking face, qui ressemble assez à un genre de sablé rond avec des lézardes sur le pourtour, U+1FAEB, 🥞 ;
- un filet à papillon, net with handle, U+1FA8D, 🦋,
- une gomme, eraser, U+1FA8C, 🦋, impeccable en période de rentrée des classes ;
- une météorite, meteor, U+1FA8B, 🦋 ;
- et un papillon monarque, monarch butterfly, U+1FACC, 🦋. C'est d'ailleurs assez intéressant cet ajout car il existait déjà un papillon 🦋, U+1F98B.

Unicode 18 voit arriver des symboles musicaux supplémentaires : bloc 1D250 à 1D28F. On y trouvera notamment des symboles pour l'utilisation des talons et des orteils pour les pédales d'orgue, respectivement U+1D273 et U+1D274 ou encore, U+1D256, le symbole pour le bémol quart de ton arabe.

Les symboles divers et flèches se voient également gratifier de nouvelles entrées aux blocs 1DB00 à 1DBFF. Par exemple en U+1DB0A les deux lignes « plus grand que » et en 1DB0B les deux lignes « plus petit que ». Les deux symboles affichent deux lignes horizontales de tailles inégales.

Remerciement et mots de la fin

Cette dépêche n'aurait sans doute pas été rédigée si rapidement après la sortie d'Unicode 18 sans [Stéphane Bortzmeyer](#) qui m'a fait lire en avant-première sa note de blog au moment où j'allais lui demander de me tenir au courant. Tous mes remerciements. Évidemment, je vous suggère fortement de lire [sa note de blog](#) sur le sujet.

La version originale de l'image de tête est sur [OpenClipart \(EN\)](#), elle a été dessinée par [gubodup \(EN\)](#). Le bocal avait un [couvercle bleu \(EN\)](#) que l'auteur a repeint en [vert \(EN\)](#). Plus tard, [oksmith \(EN\)](#) a rempli le bocal de [cornichons \(EN\)](#), bocal que j'ai malencontreusement [renversé \(EN\)](#).

Et, bien sûr si vous avez la chanson de Nino Ferrer qui tourne en boucle dans votre crâne, je dénie toute responsabilité.

1. Selon les sources, on recense cinq ou six styles d'écriture chinoise. [↩](#)

2. Les radicaux sont les clés qui composent les idéogrammes chinois. [↩](#)

Aller plus loin

🇬🇧 [Note de blog annonçant la sortie d'Unicode 18](#) (2 clics)

🇬🇧 [Résumé des principaux changements apportés par Unicode 18](#) (2 clics)

🇬🇧 [Version 18 d'Unicode](#) (3 clics)

🇬🇧 [Tous les émojis Unicode](#) (3 clics)

🇬🇧 [Les nouveaux symboles musicaux \(pdf\)](#) (2 clics)

🇬🇧 [Les nouveaux symboles divers et flèches \(pdf\)](#) (2 clics)

(0 commentaire).

Markdown

EPUB

Envoyer un commentaire

Suivre le flux des commentaires

Note : les commentaires appartiennent à celles et ceux qui les ont postés. Nous n'en sommes pas responsables.

Revenir en haut de page

Derniers commentaires

- Re: La patrimonialis...
- réponse
- Re: Nombre de Frid...
- Re: La patrimonialis...
- Re: Encore une loi q...
- Nombre de Fridos v...
- Re: Plop fera la bulle.
- Re: Ho un voile !
- Re: XTEINK
- Re: Ho un voile !
- Re: Qui/quoi débran...

Étiquettes (tags) populaires

- intelligence_artificielle
- grands_modèles_de...
- merdification
- hugging_face
- jihad_butérien
- états-unis
- slop
- linux
- openai
- capitalisme
- startup_nation
- sortie_version

Sites amis

- Agenda du Libre
- April
- Éditions D-BookeR
- Éditions Diamond
- Éditions Eyrolles
- Éditions ENI
- En Vente Libre
- Framasoft
- La Quadrature du Net
- Lea-Linux
- Open Source Initiative
- Imprimerie Grafik Plus

À propos de LinuxFr.org

- Mentions légales
- Faire un don
- L'équipe de LinuxFr....
- Informations sur le s...
- Aide / Foire aux que...
- Suivi des suggestion...
- Wiki du site
- Règles de modération
- Statistiques
- API pour le dévelo...
- Code source du site
- Plan du site