

Préface

Lorsque Patrick m'a demandé de préfacier son livre, j'étais quelque peu tiraillé car je craignais fort de souffrir à cause de mon français un peu rouillé !

Mais je connais Patrick depuis des années et je savais que son travail serait intéressant et bien ficelé. Il a une longue expérience du développement dans le domaine de l'internationalisation, notamment grâce à son rôle de premier plan dans la création de l'un des premiers navigateurs Internet multilingues. En même temps, il a la bosse de la communication, ayant donné plusieurs conférences bien structurées et écrit plusieurs articles fort clairs sur Unicode et l'internationalisation. En plus de ses penchants et de son enthousiasme, il montre aussi un intérêt marqué — et a connu du succès — pour la promotion de langues et d'écritures africaines comme le touareg et le tifinaghe.

Ce que j'ai trouvé a dépassé mes attentes. En plus de traiter des questions importantes du standard Unicode lui-même, ce livre est une mine d'informations sur les aspects pratiques de l'utilisation d'Unicode et de l'internationalisation des logiciels en général. Notamment, le livre commence par un résumé utile des principaux concepts et enjeux du standard Unicode et d'autres standards d'internationalisation apparentés, résumé rédigé dans un langage bien plus fluide que les standards eux-mêmes qui sont pour le moins arides !

Mais il va beaucoup plus loin. Cet ouvrage dans son ensemble constitue un complément précieux de ces standards : des descriptions solides de l'utilisation d'Unicode dans divers programmes, mais aussi de l'interface entre Unicode et le monde des polices de caractères, des concepts et problèmes relatifs à l'internationalisation des logiciels, et bien plus encore. Et, bien sûr, une attention particulière aux caractères et aux problèmes d'importance pour les francophones. En outre, le texte est si clairement écrit que même mon français peu assuré s'est avéré à la hauteur...

Unicode est une des pierres angulaires du logiciel moderne ; sa compréhension est importante autant pour les programmeurs que pour les utilisateurs. Sans fanfare, Unicode a complètement transformé les fondements du logiciel et des communications au cours de la dernière décennie. Dès que vous lisez ou écrivez quelque chose sur un ordinateur, vous utilisez Unicode. Dès que vous effectuez une recherche avec Google, Yahoo!, MSN, Wikipédia ou bien d'autres sites, vous utilisez Unicode.

Nous avons entrepris Unicode avec un but simple : unifier les centaines de manières de coder des caractères qui existaient en les remplaçant par un standard unique et universel. Les codages de caractères préexistants étaient à la fois incomplets et incohérents : à chaque fois que du texte devait passer d'un programme ou d'un système à un autre, il y avait un sérieux risque de corruption. La gestion de codages particuliers était codée en dur dans les programmes, rendant le développement, les tests et la maintenance de ces versions internationales difficiles, coûteux, voire cauchemardesques. En fin de compte les lancements de produit dans les marchés étrangers étaient à la fois onéreux et tardifs, au grand dam autant des éditeurs que de leurs clients. Les pays en développement souffraient particulièrement : le soutien des petits marchés était tout simplement impossible.

La situation a changé radicalement avec Unicode. Dorénavant, les programmes n'ont plus qu'à utiliser une seule représentation pour tout le texte, représentation qui gère toutes les langues du monde. Un programme peut facilement être structuré de façon à être rapidement déployé dans plusieurs langues. Ainsi, de multiples versions linguistiques d'un logiciel peuvent maintenant être développées presque simultanément et à coût marginal grandement réduit, même dans le cas de programmes extrêmement complexes comme Microsoft Office ou OpenOffice. Des compagnies comme Google peuvent traiter toutes les langues du monde avec une seule représentation uniforme. Les utilisateurs y gagnent énormément quant à la variété de programmes disponibles et à leur qualité.

À mesure qu'Unicode et les logiciels le mettant en œuvre ont pris de l'ampleur, l'ensemble est inévitablement devenu plus complexe. On fait face à une jungle de sigles et de concepts déconcertants : NDI (IDN), HTTP, XML, HTML, CSS, ICU, OpenType, BIDI, PMB (BMP), IOO (BOM), CLDR, FNC (NFC), UCA, UTF, LSC (ZWF), etc.

Ce livre trace un sentier dans cette jungle. Je ne connais pas d'autre ressource moderne, ni en anglais ni en français, regroupant une telle gamme de sujets utiles, et les expliquant de façon si claire et si accessible à un large public. J'espère que vous apprécierez cet ouvrage autant que moi.

Mark Davis, Ph. D.
Président et cofondateur du Consortium Unicode