General Computing

The Unicode Standard
Worldwide Character Encoding
Version 1.0, Volume 1

The Unicode Consortium

The Unicode standard is a new international standard used to encode written characters for storage in computer files or transmission over communication lines. Whereas ASCII, the current North American standard, encodes the Latin alphabet, the Unicode standard addresses all of the characters used for written languages throughout the world. It also encodes mathematical and scientific symbols that were never provided for in the ASCII standard.

Based on a 16-bit coding architecture, the Unicode standard assigns each character or symbol a unique value. The simplicity and efficiency of this approach makes it easier to develop multilingual software and to exchange information internationally.

The Unicode Standard is the authorized description and guide to this new standard. It documents every aspect of the standard, including basic principles, code charts, and a discussion of implementation issues. Volume 1 covers the Latin, Cyrillic, Greek, Hebrew, and Arabic alphabets, and other alphabets used in countries across Europe, Africa, and the Indian subcontinent. Volume 2 is devoted to unified character codes for Chinese, Japanese, and Korean ideographs.

The Unicode Standard is an essential reference for computer programmers and software developers who deal with multilingual text.

The Unicode Consortium is a nonprofit organization founded to promote the use of the Unicode standard. Originating from an informal collaboration between engineering teams at Apple and Xerox, the consortium now includes other companies concerned with handling international text files, such as Adobe, Aldus, Borland, GO, IBM, Lotus, Metaphor, Microsoft, NeXT, Novell, Sun, and WordPerfect.

Cover design by Ned Williams

Addison-Wesley Publishing Company, Inc.