

Unicode

A standard character set, which represents most of the characters used in the world using a 16-bit encoding.

Note: The Unicode character set can be encoded using either UTF-16 or UTF-8. UTF-16 uses two bytes for every character. UTF-8 is able to store the most commonly used characters in western alphabets using a single byte, but it requires two bytes to encode accented characters and three bytes to encode symbols used in many non-European scripts.