

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации

Doc Type: Working Group document
Title: Function and property of character names
Source: Unicode Consortium and Ireland
Status: Liaison contribution
Date: 2001-04-05

With regard to the representation of standard character names in the Unicode Standard, in response to some concerns presented by Korean experts, the following should be noted:

The principle function of character names in the Unicode Standard is to serve as non-numeric identifiers for characters. Names therefore have three properties:

- They are unique
- They are fixed
- They are arbitrary (in practice).

With regard to arbitrariness, it should be noted that while we *do* try to make the names meaningful, once they are finally assigned, they *cannot* be changed even if we get more information later.

It may be appropriate to translate character names (into languages like Korean, for instance) to help local users locate the correct character code.

The purpose of the standard is interoperability. Interoperability is based on using the same codes for a given entity in data transfer. Translating names can help to increase interoperability, because it helps native users identify characters in the Unicode Standard. It is important that when standards are translated, that they reproduce the same technical specification as the original English version of the standard, so that implementations are interoperable.

France and Canada chose to translate character names in the French translation in the standard (there exists a French translation of both ISO/IEC 10646 and the Unicode Standard). Japan, on the other hand, chose not to do so, but instead relies on the English names as arbitrary strings. Because of this, it is especially important to reiterate that we do not change names once they are adopted into the standard.