Glossary

ANSI C

The first and only standard for the C programming language, which includes many of the locale mechanisms documented here.

character set

A set of alphabetic or other characters used to construct the words and other elementary units of a natural language or computer language. Character sets may consist of alphabets, ideographs, or other units.

code set

A list of unambiguous rules that establishes a character set and a one-to-one relationship between each character of the set and its bit representation. In short, a mapping between characters and computer code. ASCII is the most common code set; others are ISO 8859-1, JIS X0208, IBM DOS 437, and Unicode.

encoding scheme

A mechanism for combining multiple code sets into a unified character coding system. Examples are EUC, compound text, and ISO 2022.

EUC

An encoding scheme that allows up to four code sets to coexist in one data stream. EUC stands for Extended UNIX Code.

ideograph

Also called ideogram. A picture or symbol used in a system of writing to represent a thing or an idea but not a particular word or phrase for it. Often the symbol does not represent an object, but some thing or idea that the object is supposed to suggest. Ideographs do not comprise an alphabet.

internationalization

Designing and building software products that are adaptable to different written languages and regions, conforming to local requirements and customs, without engineering changes.

ISO and IEC

The International Standards Organization (ISO) and the International Electrotechnical Commission (IEC) together form a system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of international standards through technical committees established by ISO and IEC to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and nongovernmental, also take part in the work.

ISO divides topics of standardization into a number of areas, each assigned to a technical committee. One of these committees, the Open Systems Interconnection (OSI), was organized to deal with the complexity of communicating systems and the required protocol standards.

localization

Providing the necessary translations, format definitions, tables and other data to adapt an internationalized product to local requirements.

All the ISO 8859 code sets are listed in Table B-1 on page 67.

locale

The international environment of a computer program defining the localized behavior of that program at runtime. A locale name is comprised of language and territory, though often territory is dropped when not needed. Locale name is specified by the `LANG` environment variable.

locale category

One attribute of a locale, such as character type, date and time format, numeric and monetary conventions, collation sequences, or translated message.

multibyte character

Because 8-bit bytes can encode only 255 characters, many code sets require more than 8 bits, and contain characters that span two or more bytes, which are called multibyte characters. Since some encoding schemes contain ASCII, single-byte characters can be considered multibyte.

POSIX

An IEEE standard for the operating system interface, which includes most locale mechanisms documented here.

Background information for the POSIX standard follows.

/usr/group, a group of UNIX users in the U.S., established a committee with the objective of proposing a set of standards for application level interfaces. After publishing the 1984 /usr/group Standard, the group decided to seek IEEE status for the standard. (IEEE stands for the Institute of Electrical and Electronics Engineers.) In early 1984, the /usr/group Standards Committee closed its activities in its own name and its members were encouraged to become involved in the IEEE POSIX committee so that the work could become the basis for an official international standard.

The first externally visible result of this initiative was the publication of the IEEE Trial-Use Standard in March 1986, with the formal approval in August 1988 of IEEE Standard 1003.1-1988, a "Portable Operating System Interface for Computer Environments" (POSIX). POSIX was the first step toward a truly portable operating system.

Although originally planned to refer to the IEEE Standard 1003.1-1988, the name POSIX has come to refer to a family of related standards and parts of the International Standard ISO/IEC 9945 (see "UniForum (/usr/group)"). A preferred term, POSIX.1, emerged that did not confuse the 1003.1 standard with the POSIX family of standards. The 1003.1-1988 POSIX standard has also been accepted as an ISO standard, ISO IS9945/1 (1989).

UniForum (/usr/group)

The UniForum Technical Committee, formerly /usr/group, is an association of individuals, corporations, and institutions with an interest in the UNIX system. This organization provides input to POSIX and other standards committees and consortia to aid the development of independent industry-driven standards. UniForum members represent a cross section of the UNIX system community: over 450 members representing hardware manufacturers, vendors of operating systems and other software development tools, software designers, consultants, academics, authors, applications programmers, and others.

wide character

An integral data type large enough to hold any member of an extended character set. In program terms, it is an object of type `wchar_t`, which is an implementation-defined integral type defined in `<stddef.h>`. There is no direct relationship between display width and wide characters. Wide characters are sometimes called f2process code.

X/Open

A consortium that published the *X/Open Porting Guide* (XPG), which specified most locale mechanisms documented here. XPG Issue 3 incorporated all of POSIX. Background information for X/Open follows.

Founded in 1984, X/Open is a worldwide consortium of vendors attempting to adopt existing standards and adapt them into a consistent environment called the Common Applications Environment (CAE).

Many of the world's major hardware suppliers, including Sun, are now X/Open members. Most of X/Open's technical work is accomplished by people from its member companies. Each member company of X/Open has a contractual agreement with the consortium to provide compliant versions of the products specified.

X/Open publishes its specifications in the X/Open Portability Guide (XPG). The X/Open Portability Guide defines the interfaces identified as components of the Common Applications Environment. It contains an evolving portfolio of applications programming interfaces (APIs) that enhance portability of application programs at source code level. The interfaces are supported by an extensive set of conformance tests and the distinct trademark, the X/Open brand. The third version of the Portability Guide, XPG3, encompasses the IEEE POSIX.1 operating system interface.