Unicode Locale Status

Hideki Hiura (SUN)
Akio Kido (IBM)
Tetsuji Orita (IBM)
Kentaro Noji (IBM)

UTC #84 August, 2000
Coworking with Linux Internationalization Initiative (Li18nux) creating Unicode base locales for GNU C library (glibc 2.2) as an Open Source project.

- 65 languages (total 152 locales)
- UTF-8 base charmap
- Full Unicode 3.0 character repertoire
- Unicode conforming collation

The source format is POSIX plus some extension
Project status (cont.)

/ Version 0.9 (as of Aug 4, 2000)
  / About 100 locales are available.
  / Collation
    / Unicode 2.1.9 base
    / Additional characters from Unicode 2.1.9 to Unicode 3.0 are ordered by Unicode code point order (except CJK Ext-A)
    / CJK Unified + CJK Ext-A in IRG SuperCJK 10.0 ordering

/ Version 1.0 (plan: Oct. 2000)
  / All of 152 locales will be supported
  / Unicode 3.0 conforming collation
Availablity

Those locales will be licensed under IBM Public License (one of approved open source license).

http://www.opensource.org/licenses/

Those locales will be available through

Li18nux Web sites

IBM developerWorks Web site

Those locales intended to be bundled with some Open Source SW, e.g. glibc, and Linux distributions.
Requests to UTC

/ Please provide us Unicode 3.0 base collation data as soon as possible.
  / Early draft (when it become available)
  / Approved data (before Oct. 2000)
/ Please ask Unicode members to participate in the testing activity.
/ Please put a copy or link to the locale data on the Unicode consortium Web site.