

Unicode Activity Report

1. Publication of Unicode 4.0.

The Unicode Consortium issued Unicode 4.0 in the spring of 2003. The standard was published in book form in September 2003. Since then, enhancements and corrections have been made, and Unicode 4.0.1 will be issued shortly.

Among the error corrections are substantial improvements to the Unihan database. The Mandarin readings have had extensive corrections made, and additions have been made to other fields.

The Unicode Technical Committee wishes to express its thanks to the IRG for all of the hard work done which has made Unicode 4.0 possible. We will be distributing complimentary copies of *The Unicode Standard, Version 4.0* to the delegations at this meeting.

2. Descriptions of characters in Extension B

We have prepared descriptions of approximately 17,000 of the ideographs in Extension B. Unfortunately, upon review of the data, we found an unacceptably large number of errors and it has been necessary to start the work over again.

Our plan is to use the "Character Description Language" described in IRG documents N985, N986, and N987 instead of ideographic description sequences for this work. This has the advantage of providing a more precise description and, in connection with the Wenlin computer program, allows for graphical editing. The CDL descriptions can also be turned into ideographic description sequences by computer.

The Unicode Technical Committee hopes that the IRG can review the CDL and comment on its appropriateness for this work. The UTC also feels that the IRG may find the CDL useful for its own work.

3. Work on the Basic Character Set

IRG document N988 contains the UTC's contribution on this subject.

The UTC is very concerned that vendors have urgent needs which need to be fulfilled as soon as possible, or they may start inventing their own solutions. The UTC hopes that the IRG can complete this work by the end of its next meeting.