

Universal Multiple-Octet Coded Character Set UCS

ISO/IEC JTC1/SC2/WG2 IRG N991

Date: 2003-11-12

Doc. Type:	Member body contribution
Title:	China's Basic Subset Proposal
Source:	China
Status:	Input to IRG
Action:	
Distribution:	IRG Members and Ideographic Experts
Reference:	
No. of pages:	1
Medium:	Electronic

The China's proposed character list of Basic Subset of CJK Unified Ideograph is derived from the being developed ***Standardized Hanzi list*** 規範漢字表(National Language Committee, Ministry of Education 教育部國家語委), with references of statistics of Hanzi using frequency according to the ***70 Million Contemporary Han Language Balance Corpus*** and the ***316 Million Hanzi Dynamically Circulating Corpus*** (Beijing Language and Culture University 北京語言大學) . There are 6944 characters in this list. The list may be modified according to other IRG members' contribution to ensure that simplified Hanzi and their unsimplified counterparts are in the Basic Subset of CJK Unified Ideograph.

Attachment: character list.