

# Table of Contents

<i>Foreword</i> .....	<i>xiii</i>
<i>Preface</i> .....	<i>xv</i>
<b>1. CJKV Information Processing Overview</b> .....	<b>1</b>
Multiple Writing Systems .....	2
Character Set Standards .....	6
Encoding Methods .....	8
Input Methods .....	11
Typography .....	13
Basic Concepts and Terminology .....	14
<b>2. Writing Systems</b> .....	<b>27</b>
Latin Characters and Transliteration .....	27
Zhuyin .....	40
Kana .....	42
Hangul .....	47
Chinese Characters .....	50
Non-Chinese Chinese Characters .....	61
<b>3. Character Set Standards</b> .....	<b>66</b>
Non-Coded Character Set Standards .....	67
Coded Character Set Standards .....	71
International Character Set Standards .....	120

Character Set Standard Oddities .....	130
Non-Coded Versus Coded Character Sets .....	132
Information Interchange Versus Professional Publishing .....	134
Advice to Developers .....	136
<b>4. Encoding Methods .....</b>	<b>138</b>
Locale-Independent Encoding Methods .....	140
Locale-Specific Encoding Methods .....	169
Comparing CJKV Encoding Methods .....	185
International Encoding Methods .....	186
Charset Designations .....	196
Code Pages .....	199
Code Conversion .....	202
Repairing Unreadable CJKV Text .....	209
Beware of Little and Big Endian Issues .....	214
Advice to Developers .....	214
<b>5. Input Methods .....</b>	<b>216</b>
Transliteration Techniques .....	217
Input Techniques .....	224
User Interface Concerns .....	237
Keyboard Arrays .....	238
Other Input Hardware .....	260
Input Method Software .....	262
<b>6. Font Formats .....</b>	<b>269</b>
Typeface Design Issues .....	270
Bitmapped Fonts .....	271
Outline Fonts .....	276
Ruby Fonts .....	308
Host-Based Versus Printer-Resident Fonts .....	309
Creating Your Own Fonts .....	321
External Character Handling .....	325
Advice to Developers .....	335
<b>7. Typography .....</b>	<b>336</b>
Rules, Rules, Rules... .....	337
Typographic Units and Measurements .....	338
Horizontal and Vertical Layout .....	342
Line Breaking and Word Wrapping .....	352

Character Spanning .....	356
Alternate Metrics .....	357
Kerning .....	363
Line Length Issues .....	365
Multilingual Text .....	366
Glyph Substitution .....	370
Annotations .....	372
Typographic Software .....	377
<b>8. Output Methods .....</b>	<b>387</b>
Where Can Fonts Live? .....	388
Printer Output .....	389
PostScript CJKV Printers .....	389
Computer Monitor Output .....	394
Other Printing Methods .....	398
The Role of Printer Drivers .....	399
Output Tips and Tricks .....	402
Advice to Developers .....	404
<b>9. Information Processing Techniques .....</b>	<b>406</b>
Language, Country, and Script Codes .....	407
Programming Languages .....	410
Code Conversion Algorithms .....	414
Java Programming Examples .....	422
Miscellaneous Algorithms .....	426
Byte Versus Character Handling .....	433
Character Sorting .....	440
Natural Language Processing .....	443
Regular Expressions .....	445
Search Engines .....	447
Code Processing Tools .....	448
<b>10. Operating Systems, Text Editors, and Word Processors .....</b>	<b>455</b>
Viewing CJKV Text on Non-CJKV Systems .....	456
Operating Systems .....	457
Hybrid Environments .....	468
Text Editors .....	472
Word Processors .....	479
Dedicated Word Processors .....	482

<b>11. Dictionaries and Dictionary Software</b> .....	<b>484</b>
Chinese Character Dictionary Indexes .....	484
Character Dictionaries .....	491
Other Useful Dictionaries .....	497
Dictionary Hardware .....	498
Dictionary Software .....	499
Machine Translation Software .....	507
Machine Translation Services .....	508
Learning Aids .....	509
<b>12. The Internet</b> .....	<b>511</b>
Email .....	512
News .....	517
FTP and Telnet .....	518
Network Domains .....	520
Getting Connected .....	522
Internet Software .....	523
<b>13. The World Wide Web</b> .....	<b>530</b>
Content Versus Presentation .....	530
Displaying Web Documents .....	533
Authoring HTML Documents .....	534
Authoring XML Documents .....	538
Authoring PDF Documents .....	539
Character References .....	541
CGI Programming Examples .....	542
Shall We Surf? .....	545
<b>A. Code Conversion Tables</b> .....	<b>547</b>
<b>B. Notation Conversion Table</b> .....	<b>551</b>
<b>C. Vendor Character Set Standards</b> .....	<b>554</b>
<b>D. Vendor Encoding Methods</b> .....	<b>605</b>
<b>E. GB 2312-80 Table</b> .....	<b>638</b>
<b>F. GB/T 12345-90 Table</b> .....	<b>653</b>
<b>G. CNS 11643-1992 Table</b> .....	<b>668</b>

---

<i>H. Big Ffive Table</i> .....	773
<i>I. Hong Kong GCCS Table</i> .....	804
<i>J. JIS X 0208:1997 Table</i> .....	814
<i>K. JIS X 0212-1990 Table</i> .....	828
<i>L. KS X 1001:1992 Table</i> .....	840
<i>M. KS X 1002:1991 Hanja Table</i> .....	856
<i>N. Hangul Reading Table</i> .....	862
<i>O. TCVN 6056:1995 Table</i> .....	876
<i>P. Code Table Indexes</i> .....	883
<i>Q. Character Lists and Mapping Tables</i> .....	897
<i>R. Chinese Character Lists</i> .....	941
<i>S. Single-Byte Code Tables</i> .....	965
<i>T. Software and Document Sources</i> .....	975
<i>U. Mailing Lists</i> .....	995
<i>V. Professional Organizations</i> .....	1006
<i>W. Perl Code Examples</i> .....	1008
<i>X. Glossary</i> .....	1025
<i>Bibliography</i> .....	1053
<i>Index</i> .....	1073