UI-OSF Application Platform Profile for Japanese Environment

Version 1.1

May 21,1993

UI-OSF Japanese Localization Group

Forward

This document is an Application Platform Profile that specifies options, parameters and extensions to assure portability of applications which support Japanese on operating systems conforming to ISO POSIX standards (ISO/IEC 9945-1 and ISO/IEC DIS 9945-2), such as UNIX operating system and OSF/1. The Japanese locale specified by this document will be registered in X/Open's locale registry.

NOTE:
This document was translated from the master edition which is written in Japanese. And this document was made as an abridged translation. If there are some ambiguous things in this English edition and any clarifications are needed, the Japanese edition must be referred.

Table of Contents

1. Preface This document is an Application Platform Profile specified by UI-OSF Japanese Localization Group (UOJLG), in order to increase portability and interoperability of applications that support Japanese language. UOJLG discloses this profile openly, imposing no condition on reference to, copy, distribution and implementation of this profile.

This profile is specified to be consistent with ISO POSIX standards, X/Open Portability Guide Issue 4(XPG4) and Draft Japanese National Profile for POSIX. This profile may be updated when the furture revisions of these standards and documents.

1.1 Scope This profile specifies options, parameters and extensions to assure portability of applications which support Japanese on operating systems conforming to ISO POSIX standards (ISO/IEC 9945-1 and ISO/IEC DIS 9945-2), such as UNIX operating system and OSF/1. It is intended to used by implementors of systems that support Japanese language.

This profile comprises four major components:

(1) Environment variables related with Japanese language
(2) Object names related with Japanese language handling
(3) Japanese locale definition
(4) Extensions to ISO POSIX standards required to support Japanese language

1.2 Normative reference The International, National and industry standards referred by this profile is listed as follows.

(1) International standards
- ISO/IEC 9945-1:1990 Information technology – Portable Operation System Interface (POSIX) – Part 1: System Application Program Interface (API)
- ISO/IEC DIS 9945-2 Information technology – Portable Operation System Interface (POSIX) – Part 2: Shell and Utilities
- ISO/IEC 9899:1990 Information technology – Programming Languages – C
- ISO/IEC 2022:1986 Information technology – ISO 7-bit and 8-bit coded character set – Code extension techniques
- ISO/IEC 646:1991 Information processing – ISO 7-bit coded character set for information interchange
- ISO 6429:1993 Control functions for 7-bit and 8-bit coded character sets
- ISO 639:1988 Codes for the representation of names of languages
- ISO 3166:1988 Codes for the representation of names of countries
- ISO 4217:1987 Codes for the Representation of Currencies and Funds

(2) Japanese Industrial Standards
- JIS X 0201: Code for Information Interchange
- JIS X 0208: Code of Japanese Graphic Character Set for Information Interchange
- JIS X 0212: Code of the supplementary Japanese graphic character set for information interchange

(3) Industry standards
- X/Open Portability Guide Issue 4(XPG4)
  - X/Open CAE Specification, System Interface Definitions, Issue 4
  - X/Open CAE Specification, System Interfaces and Headers, Issue 4
  - X/Open CAE Specification, Commands and Utilities, Issue 4
- UI-OSF Common Japanese EUC

(4) Draft standards and related documents
- ISO/IEC JTC1 SC22/WG15 Draft Japanese National Profile for POSIX Version 2.2
- ISO/IEC JTC1 SC22/WG14 N 239 Programming Language C – AMENDMENT 1

1.3 Conformance

1.3.1 Requirements Implementation of operating system conform to ISO POSIX standards and this profile shall satisfy ,in addition to the requirements of ISO POSIX standards, all of the following criteria.

(1) Implementation of operating system shall provide all interfaces (function, objects, object names and values) specified by ISO POSIX Standards, X/Open Portability Guide Issue 4 and this profile, except noted us optional features.

(2) Implementation of operating system may provide functions, features, symbols or locale definitions not required by ISO POSIX standards, X/Open Portability Guide Issue 4 or this profile, unless they do not require changes of applications using mandatory features of ISO POSIX standards and this profile.

1.3.2 Optional feature This profile specifies interfaces of optional feature, which should be provided by implementation of operating system conforming to this profile. It is implementation defined whether the optional features are provided.

2. Terminology For the purposes of this document, the following definitions apply:

2.1 Implementation defined An indication that the implementation provider shall define and document the requirements for correct program constructs and correct data of a value or behavior.

2.2 May An indication of an optional feature or behavior of the implementation that is not required by this part of this document, although there is no prohibition against providing it.

2.3 Shall An indication of a requirement on the implementation.

2.4 Should

(1) With respect to implementations, an indication of an implementation recommendation, but not a requirement.

(2) With respect to applications, an indication of recommended programming practice for application.

2.5 Undefined An indication that this document imposes no portability requirements on applications for erroneous program construction, erroneous data, or use of an indeterminate value.

Implementations may specify the result of using that value or causing that behavior.

2.6 unspecified An indication that this document imposes no portability requirements on applications for correct program construction, correct data.

Implementations may specify the result of using that value or causing that behavior.

3. Environment variable

3.1 TZ The TZ enviroment variable contains timezone name and the time difference from the Universal Coordinated Time. In the case of Japanese Standard Time, "JST" is used for timezone name and the time difference is nine hours advanced from UTC.

4. Object name

4.1 Name of coded character sets The object names of coded character sets is as follows. These object name are used for codeset field of locale name or arguments of iconv utility or functions.

eucJP UI-OSF Common Japanese EUC

SJIS Shift JIS or MS Kanji

ISO-2022-JP Coded representation of the character sets ISO 646 IRV or JIS X 0201, JIS X 0208, and JIS X 0212 using the designation sequence to G0 specified by IS0 2022.

JIS7 same as ISO-2022-JP

Note SJIS is also referred to as MS Kanji because the name is often the source of misunderstanding that it is specified by Japan Industry Standard (JIS).

Additional names for implementation defined coded character sets may be provided by an implementation.

4.2 Japanese locale Name

4.2.1 Locale name format The locale name format, according to X/Open Portability Guide Issue 4 is as follows. The fields enclosed by brackets may be omitted.

```
language [_territory] [.codeset] [@modifier]
```

language Two letter symbol of a language specified by ISO 639.

territory Two letter symbol of a country specified by ISO 3166.

codeset Symbolic name of coded character set specified by this profile. Symbolic names for coded character sets not specified by this profile are implementation defined.

modifier An implementation defined symbols used to distinguish detail difference of locale definition. This profile may specify symbols to use this filed in future.

4.2.2 Japanese locale Name The symbol for the Japanese locale specified by this profile is

```
ja_JP [.codeset]
```

The default value of codeset is eucJP. If an implementation allows users to set or modify the default value, and it is set to a value other than eucJP, the value may be used as default instead of eucJP. It is implementation defined whether an implementation provide locales for other codesets than eucJP.

4.3 Symbolic name of characters This profile has no requirement on symbolic name of characters used in charmap and localedef file. The following clauses describes the convention used only for the purpose of specification.

4.3.1 POSIX Portable characters The symbolic name used in POSIX standard is used.

4.3.2 C1 control character Symbol name specified by ISO 6429 is used.

4.3.3 JIS X 0201 Katakana characters

symbol	code point	symbol	code point
		<kana-TA>	12/0
<kana-full-stop>	10/1	<kana-CHI>	12/1
<kana-opening-bracket>	10/2	<kana-TSU>	12/2
<kana-closing-bracket>	10/3	<kana-TE>	12/3
<kana-comma>	10/4	<kana-TO>	12/4
<kana-conjunctive>	10/5	<kana-NA>	12/5
<kana-WO>	10/6	<kana-NI>	12/6
<kana-a>	10/7	<kana-NU>	12/7
<kana-i>	10/8	<kana-NE>	12/8
<kana-u>	10/9	<kana-NO>	12/9
<kana-e>	10/10	<kana-HA>	12/10
<kana-o>	10/11	<kana-HI>	12/11
<kana-ya>	10/12	<kana-FU>	12/12
<kana-yu>	10/13	<kana-HE>	12/13
<kana-yo>	10/14	<kana-HO>	12/14
<kana-tsu>	10/15	<kana-MA>	12/15
<kana-prolonged-sound>	11/0	<kana-MI>	13/0
<kana-A>	11/1	<kana-MU>	13/1
<kana-I>	11/2	<kana-ME>	13/2
<kana-U>	11/3	<kana-MO>	13/3
<kana-E>	11/4	<kana-YA>	13/4
<kana-O>	11/5	<kana-YU>	13/5
<kana-KA>	11/6	<kana-YO>	13/6
<kana-KI>	11/7	<kana-RA>	13/7

<kana-KU>	11/8	<kana-RI>	13/8
<kana-KE>	11/9	<kana-RU>	13/9
<kana-KO>	11/10	<kana-RE>	13/10
<kana-SA>	11/11	<kana-RO>	13/11
<kana-SHI>	11/12	<kana-WA>	13/12
<kana-SU>	11/13	<kana-N>	13/13
<kana-SE>	11/14	<kana-voiced-sound>	13/14
<kana-SO>	11/15	<kana-semivoiced-sound>	13/15

4.3.4 JIS X 0208 This document uses the symbolic name "<jxxyy>" for JIS X 0208 characters, where xx is a decimal value of row number, and yy is a decimal value of column number.

4.3.5 JIS X 0212 This document uses the symbolic name "<Jxxyy>" for JIS X 0212 characters, where xx is a decimal value of row number, and yy is a decimal value of column number.

5. Japanese charmap Conforming implementation shall include all characters belonging to ISO 646 IRV or JIS X 0201 7 bit code of Roman character, and JIS X 0208 in a charmap file used in conjunction with Japanese locale specified by this profile. Implementation should include all characters belonging to C1 control specified by ISO 6429, katakana characters of JIS X 0201, and characters of JIS X 0212. Implementation may include implementation defined user- and vendor-defined characters in charmap.

6. Japanese locale definition Implementation shall provide at least `ja_JP.eucJP` locale. The `ja_JP` locale is defined as follows.

6.1 LC_CTYPE category Conforming implementation shall provide all of the following character classification classes and character translation. Conforming implementation may provide implementation defined additional character classes. The meanings of the following character classes are based on POSIX.2 along with applied interpretations and extentions for Japanese locale. The detailed specification of LC_CTYPE category is in Annex B.1.

Implementation of C1 control specified by ISO 6429, Katakana characters specified by JIS X0201 and Supplementary Kanji specified by JIS X 0212 is optional, but if an implementation supports characters in the character sets listed above, such characters shall be classified as specified in this profile.

6.1.1 upper The upper class shall contain upper case letter of Latin, Greek and Cyrillic scripts including upper case letters with diacritical marks. The upper class may contain implementation defined characters which can be classified as upper case letters of Western alphabets.

6.1.2 lower The lower class shall contain lower case letter of Latin, Greek and Cyrillic scripts including lower case letters with diacritical marks. The lower class may contain implementation defined characters which can be classified as lower case letters of Western alphabets.

6.1.3 alpha The alpha class shall contain all characters in the upper or lower classes.

6.1.4 digit The digit class shall only contain 0,1,2,3,4,5,6,7,8 and 9 specified by ISO 646 IRV or JIS X 0201.

6.1.5 xdigit The xdigit class shall contain all of the following characters specified by ISO 646 IRV or JIS X 0201.

(1) 0,1,2,3,4,5,6,7,8 and 9 (3/0-3/9)

(2) A,B,C,D,E and F (4/1-4/6)

(3) a,b,c,d,e and f (6/1-6/6)

6.1.6 space The space class shall contain all of the following characters.

(1) SP,FF,LF,CR,HT and VT specified by ISO 646 IRV or JIS X 0201.

(2) Ideographic space <j0101> specified by JIS X 0208.

6.1.7 blank The blank class shall contain SP and HT specified by ISO 646 IRV or JIS X 0201.

6.1.8 cntrl The cntrl class shall contain all of the following characters:

(1) C0 control characters specified by ISO 6429.

(2) C1 control characters specified by ISO 6429.
(3) DEL character specified by ISO 646 IRV or JIS X 0201.

The cntrl class may exclude SS2 and SS3 in an implementation recognizing them as part of graphic characters, if the coded character set uses them as specified in ISO 2022. The cntrl class may contain implementation defined characters which can be classified as control character.

6.1.9 punct The punct class shall contain all characters specified by ISO 646 IRV or JIS X 0201,JIS X 0208 and JIS X 0212, except those classified as cntrl, alpha, digit, jdigit, jkanji,jhira or jkata. The punct class may contain implementation defined characters that can be classified as special symbols.

6.1.10 graph The graph class shall contain all character belonging to print class excluding characters which belong to space class.

6.1.11 print The print class shall contain all of the following characters:

(1) all characters specified by ISO 646 or JIS X 0201 7bit Alphanumeric Code excluding characters which belongs to cntrl class
(2) all characters specified by JIS X 0201 Graphic Character Set for Katakana
(3) all characters specified by JIS X 0208
(4) all characters specified by JIS X 0212
(5) all implementation defined characters which are printable.

6.1.12 ascii The ascii class shall contain all characters specified by ISO 646 IRV or JIS X 0201 7bit Alphanumeric Code.

6.1.13 line The line class shall contain all of Ruled Line Elements specified by JIS X 0208. The line class may contain implementation defined characters which can be classified as Ruled line or Box character.

6.1.14 jdigit The jdigit class shall contain digit 1 to 9 specified by JIS X 0208.

6.1.15 paren The paren class shall contain all the following characters:

(1) '(',')','[',']','{' and '}'
(2) row 1 column 13 to row 1 column 59 of JIS X 0208
(3) 2/2 and 2/3 of JIS X 0201 7bit Katakana Code

The paren class may contain implementation defined characters which are used in pair.

6.1.16 jparen The jparen class shall contain all the following characters:

(1) row 1 column 13 to row 1 column 59 of JIS X 0208
(2) 2/2 and 2/3 of JIS X 0201 7bit Katakana Code

The jparen class may contain implementation defined characters which are used in pair.

6.1.17 jisx0201 The jisx0201 class shall contain all characters specified by JIS X 0201. The jisx0201 class may contain implementation defined characters assigned in a code area for JIS X 0201.

6.1.18 jisx0201r The jisx0201r class shall contain all characters specified by JIS X 0201 7bit code for Katakana. The jisx0201r class may contain implementation defined characters assigned in a code area for JIS X 0201 7bit Katakana Code.

6.1.19 jisx0208 The jisx0208 class shall contain all characters specified by JIS X 0208. The jisx0208 class may contain implementation defined characters assigned in a code area for JIS X 0208.

6.1.20 jisx0212 The jisx0212 class shall contain all characters specified by JIS X 0212. The jisx0212 class may contain implementation defined characters assigned in a code area for JIS X 0212.

6.1.21 udc The udc class shall contain all implementation defined characters which are classified as User Defined Characters.

6.1.22 vdc The vdc class shall contain all implementation defined characters which are classified as Vendor Defined Characters.

6.1.23 gaiji The gaiji class shall contain all characters which belong to the udc class and the vdc class.

6.1.24 jhira The jhira class shall contain all of the following characters:

(1) Hiragana characters specified by JIS X 0208

(2) Voiced sound symbol , Semi-voiced sound symbol, Prolonged sound symbol, Hiragana iteration mark and Hiragana iteration mark for voiced sound specified by JIS X 0208

The jhira class may contain implementation defined characters which can be classified in Hiragana or special symbols for Hiragana.

6.1.25 jkata The jkata class shall contain all of the following characters:

(1) Katakana characters specified by JIS X 0208

(2) Voiced sound symbol , Semi-voiced sound symbol, Prolonged sound symbol, Katakana iteration mark and Katakana iteration mark for voiced sound specified by JIS X 0208

(3) Katakana, Voiced sound symbol , Semi-voiced sound symbol and Prolonged sound symbol specified by JIS X 0201

The jkata class may contain implementation defined characters which can be classified in Katakana or special symbols for Katakana.

6.1.26 jhankana The jhankana class shall contain all of the following characters:

(1) Katakana characters specified by JIS X 0201

(2) Voiced sound symbol, Semi-voiced sound symbol and iteration mark specified by JIS X 0201

The jhankana class may contain implementation defined characters which are assigned in a code area of JIS X 0201 7bit Katakana Code and can be classified as Katakana or Special symbol for Katakana.

6.1.27 jkanji The jkanji class shall contain all of the following characters:

(1) Kanji characters belongs to jisx0208 class

(2) Kanji characters belongs to jisx0212 class

(3) Iteration marks(<j0124>, <j0125>) and Kanji zero(<j0127>) specified by JIS X 0208

The jkanji class may contain implementation defined characters which can be classified as Kanji or special symbol for kanji.

6.1.28 jspace The jspace shall contain Ideographic space (<j0101>) specified by JIS X 0208.

6.1.29 tolower The tolower class shall specify character translation from upper class characters to corresponding lower class characters.

6.1.30 toupper The toupper class shall specify character translation from lower class characters to corresponding upper class characters.

6.2 LC_COLLATE category The recommended collation rule for Japanese characters is follows. Conforming implementation may provide other rules for Japanese character collation. However if an implementation provide other rules, it is recommended that it should provide an implementation defined means to allow user to switch the implementation defined rules and the recommended rule.

(1) ISO 646 IRV or JIS X 0201 7bit Alphanumeric Code

(2) ISO 6429 C1 control character

(3) JIS X 0201 Graphic character set for Katakana

(4) JIS X 0208 including implementation defined characters assigned in a code area of JIS X 0208

(5) JIS X 0212 including implementation defined characters assigned in a code area of JIS X 0212

(6) others

Japanese characters are classified into the above categories. Collation sequence among the above groups of characters is from lower number of category to upper number. Within a group of characters, collation sequence is code point order.

6.3 LC_MESSAGES category LC_MESSAGES category shall contain the following definitions:

```
yesexpr "^[<y><Y><j0389><j0357>]"
noexpr "^[<n><N><j0378><j0346>]"
```

Conforming implementation may modify the strings with affirmative or nagative reply as far as they satisfy the above definitions.

6.4 LC_MONETARY LC_MONETARY category shall contain the following definitions.

```
int_curr_symbol         "<J><P><Y><space>"
currency_symbol         "<yen-sign>"
mon_decimal_point       ""
mon_thousands_sep       "<comma>"
mon_grouping            3
positive_sign           ""
negative_sign           "<hyphen>"
int_frac_digits         0
frac_digits             0
p_cs_precedes           1
p_sep_by_space          0
n_cs_precedes           1
n_sep_by_space          0
p_sign_posn             1
n_sign_posn             4
```

Conforming implementation may replace definitions for positive/negative sign (positive_sign, negative_sign, p_cs_precedes, p_sep_by_space, n_cs_precedes, n_sep_by_space, p_sign_posn and n_sign_posn) with implementation defined value, though the above definitions are recommended by this profile.

6.5 LC_NUMERIC category LC_NUMERIC category shall contain the following definitions.

```
decimal_point           "<period>"
thousands_sep           "<comma>"
grouping 3
```

6.6 LC_TIME category LC_TIME shall contain the following definitions:

```
abday   "<j3892>";"<j2378>";"<j1848>";"<j3169>";\
        "<j4458>";"<j2266>";"<j3758>"
day     "<j3892><j4543><j3892>";"<j2378><j4543><j3892>";\
        "<j1848><j4543><j3892>";"<j3169><j4543><j3892>";\
        "<j4458><j4543><j3892>";"<j2266><j4543><j3892>";\
        "<j3758><j4543><j3892>"
abmon   "<space><one><j2378>";"<space><two><j2378>";"<space><three><j2378>";\
        "<space><four><j2378>";"<space><five><j2378>";"<space><six><j2378>";\
        "<space><seven><j2378>";"<space><eight><j2378>";"<space><nine><j2378>";\
        "<one><zero><j2378>";"<one><one><j2378>";\
        "<one><two><j2378>"
mon     "<one><j2378>";"<two><j2378>";"<three><j2378>";\
        "<four><j2378>";"<five><j2378>";"<six><j2378>";\
        "<seven><j2378>";"<eight><j2378>";"<nine><j2378>";\
        "<one><zero><j2378>";"<one><one><j2378>";\
        "<one><two><j2378>"
d_t_fmt "%Y<j3915>%m<j2378>%d<j3892><space>%H<j2794>%M<j4212>%S<j4135>"
d_fmt   "%Y<j3915>%m<j2378>%d<j3892>"
t_fmt   "%H<j2794>%M<j4212>%S<j4135>"
```

```
am_pm   "<j2465><j3316>";"<j2465><j2469>"
t_fmt_ampm      "%p%I<j2794>%M<j4212>%S<j4135>"
era     "+:2:1990/01/01:+*:<j4231><j3214>:%EC%Ey<j3915>";\
        "+:1:1989/01/08:1989/12/31:<j4231><j3214>:%EC<j2421><j3915>";\
        "+:2:1927/01/01:1989/01/07:<j3028><j4734>:%EC%Ey<j3915>"
era_d_fmt       "%EY%m<j2378>%d<j3892>"
era_d_t_fmt     "%EY%m<j2378>%d<j3892><space>%H<j2794>%M<j4212>%S<j4135>"
```

7. Character code conversion function Conforming Implementation shall provide all of the following character code conversion rules via iconv utility and functions specified by X/Open Portability Guide Issue 4.

source coded character set	destination coded character set
eucJP	ISO-2022-JP
eucJP	JIS7
eucJP	SJIS
ISO-2022-JP	eucJP
JIS7	eucJP
SJIS	eucJP

Conforming implementation may provide implementation defined character code conversion rules.

8. Optional features Conforming implementation should provide the following optional features.

8.1 Wide character translation functions Wide character translation functions translate a wide character to another by referring to a translation rules which is connected to current locale by implementation defined means. The implementation which provide these functions should provide the following symbols and translation rules. Conforming implementation may provide additional implementation defined rules.

Symbol	rule
tojhira	Translate a Katakana character defined by JIS X 0208 to corresponding Hiragana character defined by JIS X 0208.
tojkata	Translate a Hiragana character defined by JIS X 0208 to corresponding Katakana character defined by JIS X 0208.
tojisx0201	Translate a character defined by JIS X 0208 to corresponding one defined by JIS X 0201.
tojisx0208	Translate a character defined by JIS X 0201 to corresponding one defined by JIS X 0208.

8.1.1 wctrans()

8.1.1.1 Synopsis

```
#include <wchar.h>
wctrans_t wctrans(const char *tname)
```

8.1.1.2 Description wctrans() function tests if the translation rule specified by tname is available in the current locale. If available, the wctrans() function returns a value which will be used by towctrans() function for specifying the rule.

tname is a symbol corresponding to a translation rule from a wide character in a class to one in another class. The wctrans() function returns a value of wctrans_t which is used for an argument of towctrans() function. The wctrans() function determines the value of wctrans_t by referring to the current locale of LC_CTYPE category. The return value of wctrans() function can be used until the locale of LC_CTYPE category is changed by the call of setlocale() function.

8.1.1.3 Returns If corresponding translation rule with the value of tname is available in the current locale, wctrans() function returns a value corresponding to tname. If the rule is not available, the wctrans() function returns (wctrans_t)0.

8.1.2 towctrans()

8.1.2.1 Synopsis

```
#include <wchar.h>
wint_t towctrans(wint_t c, wctrans_t wc_trans)
```

8.1.2.2 Description The towctrans() function translate a wide character stored in c to a corresponding wide character by referring to the translation rule pointed to by wc_trans. If the translation rule pointed to by wc_trans is not available in the current locale assigned in LC_CTYPE category, the behavior of towctrans() function is undefined.

8.1.2.3 Returns If a wide character pointed to by c exists in the translation rule pointed to by wc_trans, the towctrans() function returns the corresponding wide character. If not exists, the towctrans() function returns c as is.

Annex A Charmap for Japanese

This normative annex specifies charmap using eucJP as codeset. This charmap contains JIS standard characters specified by JIS X 0201, JIS X 0208 and JIS X 0212, and C1 control characters specified by ISO 6429. Conforming implementation may remove characters belonging to C1 control, JIS X 0201 Katakana Graphic character set or JIS X 0212. Conforming implementation may include implementation defined characters in the charmap.

It is recommended to use Yen-sign specified by JIS X 0201 as the character pointed to by <yen-sign>. If a conforming implementation uses ISO 646 IRV which does not have yen-sign, instead of JIS X 0201, as the POSIX portable character set, the implementation may assign Yen sign specified by JIS X 0208 or other implementation defined character to the character pointed to by <yen-sign>.

Any comments in the following example are only for better understanding, they are not included into the specification.

```
#
#
#       charmap for eucJP
#
#
<code_set_name>                 eucJP
<mb_cur_max>                    3
<mb_cur_min>                    1
#
#       CHARMAP
#
CHARMAP
#
#       Portable Character Set
#
<NUL>                           \x00
<SOH>                           \x01
<STX>                           \x02
<ETX>                           \x03
<EOT>                           \x04
<ENQ>                           \x05
<ACK>                           \x06
<alert>                         \x07
<BEL>                           \x07
<backspace>                     \x08
<tab>                           \x09
<newline>                       \x0a
<vertical-tab>                  \x0b
<form-feed>                     \x0c
<carriage-return>               \x0d
<SO>                            \x0e
<SI>                            \x0f
<DLE>                           \x10
<DC1>                           \x11
```

<DC2>	\x12
<DC3>	\x13
<DC4>	\x14
<NAK>	\x15
<SYN>	\x16
<ETB>	\x17
<CAN>	\x18
<EM>	\x19
<SUB>	\x1a
<ESC>	\x1b
<IS4>	\x1c
<IS3>	\x1d
<IS2>	\x1e
<IS1>	\x1f
<space>	\x20
<SP>	\x20
<exclamation-mark>	\x21
<quotation-mark>	\x22
<number-sign>	\x23
<dollar-sign>	\x24
<percent-sign>	\x25
<ampersand>	\x26
<apostrophe>	\x27
<left-parenthesis>	\x28
<right-parenthesis>	\x29
<asterisk>	\x2a
<plus-sign>	\x2b
<comma>	\x2c
<hyphen>	\x2d
<hyphen-minus>	\x2d
<period>	\x2e
<full-stop>	\x2e
<slash>	\x2f
<solidus>	\x2f
<zero>	\x30
<one>	\x31
<two>	\x32
<three>	\x33
<four>	\x34
<five>	\x35
<six>	\x36
<seven>	\x37
<eight>	\x38
<nine>	\x39
<colon>	\x3a

```
<semicolon>                  \x3b
<less-than-sign>             \x3c
<equals-sign>                \x3d
<greater-than-sign>          \x3e
<question-mark>              \x3f
<commercial-at>              \x40
<A>                          \x41
<B>                          \x42
<C>                          \x43
<D>                          \x44
<E>                          \x45
<F>                          \x46
<G>                          \x47
<H>                          \x48
<I>                          \x49
<J>                          \x4a
<K>                          \x4b
<L>                          \x4c
<M>                          \x4d
<N>                          \x4e
<O>                          \x4f
<P>                          \x50
<Q>                          \x51
<R>                          \x52
<S>                          \x53
<T>                          \x54
<U>                          \x55
<V>                          \x56
<W>                          \x57
<X>                          \x58
<Y>                          \x59
<Z>                          \x5a
<left-square-bracket>        \x5b
<backslash>                  \x5c
<reverse-solidus>            \x5c
<right-square-bracket>       \x5d
<circumflex>                 \x5e
<circumflex-accent>          \x5e
<underscore>                 \x5f
<low-line>                   \x5f
<grave-accent>               \x60
<a>                          \x61
<b>                          \x62
<c>                          \x63
<d>                          \x64
```

```
<e>                             \x65
<f>                             \x66
<g>                             \x67
<h>                             \x68
<i>                             \x69
<j>                             \x6a
<k>                             \x6b
<l>                             \x6c
<m>                             \x6d
<n>                             \x6e
<o>                             \x6f
<p>                             \x70
<q>                             \x71
<r>                             \x72
<s>                             \x73
<t>                             \x74
<u>                             \x75
<v>                             \x76
<w>                             \x77
<x>                             \x78
<y>                             \x79
<z>                             \x7a
<left-brace>                    \x7b
<left-curly-bracket>            \x7b
<vertical-line>                 \x7c
<right-brace>                   \x7d
<right-curly-bracket>           \x7d
<tilde>                         \x7e
<DEL>                           \x7f

# End Portable Character Set
#
# Additional Japanese symbols
#
<yen-sign>                      \x5c
<overline>                      \x7e
# End of additional Japanese symbols
#
# C1 Control Character Set
#
# Note: SS2 and SS3 are not included because they are used as the first
# bytes of multibyte characters.
#
<BPH>                           \x82
<NBH>                           \x83
```

```
<NEL>                           \x85
<SSA>                           \x86
<ESA>                           \x87
<HTS>                           \x88
<HTJ>                           \x89
<VTS>                           \x8a
<PLD>                           \x8b
<PLU>                           \x8c
<RI>                            \x8d
<DCS>                           \x90
<PU1>                           \x91
<PU2>                           \x92
<STS>                           \x93
<CCH>                           \x94
<MW>                            \x95
<SPA>                           \x96
<EPA>                           \x97
<SOS>                           \x98
<SCI>                           \x9a
<CSI>                           \x9b
<ST>                            \x9c
<OSC>                           \x9d
<PM>                            \x9e
<APC>                           \x9f
#       End C1 Control
#
#       JIS X 0201 Katakana Characters
#
<kana-full-stop>                \x8e\xa1
<kana-opening-bracket>          \x8e\xa2
<kana-closing-bracket>          \x8e\xa3
<kana-comma>                    \x8e\xa4
<kana-conjunctive>              \x8e\xa5
<kana-WO>                       \x8e\xa6
<kana-a>                        \x8e\xa7
<kana-i>                        \x8e\xa8
<kana-u>                        \x8e\xa9
<kana-e>                        \x8e\xaa
<kana-o>                        \x8e\xab
<kana-ya>                       \x8e\xac
<kana-yu>                       \x8e\xad
<kana-yo>                       \x8e\xae
<kana-tsu>                      \x8e\xaf
<kana-prolonged-sound>          \x8e\xb0
<kana-A>                        \x8e\xb1
```

```
<kana-I>                    \x8e\xb2
<kana-U>                    \x8e\xb3
<kana-E>                    \x8e\xb4
<kana-O>                    \x8e\xb5
<kana-KA>                   \x8e\xb6
<kana-KI>                   \x8e\xb7
<kana-KU>                   \x8e\xb8
<kana-KE>                   \x8e\xb9
<kana-KO>                   \x8e\xba
<kana-SA>                   \x8e\xbb
<kana-SHI>                  \x8e\xbc
<kana-SU>                   \x8e\xbd
<kana-SE>                   \x8e\xbe
<kana-SO>                   \x8e\xbf
<kana-TA>                   \x8e\xc0
<kana-CHI>                  \x8e\xc1
<kana-TSU>                  \x8e\xc2
<kana-TE>                   \x8e\xc3
<kana-TO>                   \x8e\xc4
<kana-NA>                   \x8e\xc5
<kana-NI>                   \x8e\xc6
<kana-NU>                   \x8e\xc7
<kana-NE>                   \x8e\xc8
<kana-NO>                   \x8e\xc9
<kana-HA>                   \x8e\xca
<kana-HI>                   \x8e\xcb
<kana-FU>                   \x8e\xcc
<kana-HE>                   \x8e\xcd
<kana-HO>                   \x8e\xce
<kana-MA>                   \x8e\xcf
<kana-MI>                   \x8e\xd0
<kana-MU>                   \x8e\xd1
<kana-ME>                   \x8e\xd2
<kana-MO>                   \x8e\xd3
<kana-YA>                   \x8e\xd4
<kana-YU>                   \x8e\xd5
<kana-YO>                   \x8e\xd6
<kana-RA>                   \x8e\xd7
<kana-RI>                   \x8e\xd8
<kana-RU>                   \x8e\xd9
<kana-RE>                   \x8e\xda
<kana-RO>                   \x8e\xdb
<kana-WA>                   \x8e\xdc
<kana-N>                    \x8e\xdd
<kana-voiced-sound>         \x8e\xde
```

```
<kana-semivoiced-sound>         \x8e\xdf
#       End JIS X 0201 Katakana
#
#       JIS X 0208 Characters
#
<j0101>...<j0194>               \xa1\xa1
<j0201>...<j0294>               \xa2\xa1
<j0301>...<j0394>               \xa3\xa1
<j0401>...<j0494>               \xa4\xa1
<j0501>...<j0594>               \xa5\xa1
<j0601>...<j0694>               \xa6\xa1
<j0701>...<j0794>               \xa7\xa1
<j0801>...<j0894>               \xa8\xa1
<j0901>...<j0994>               \xa9\xa1
<j1001>...<j1094>               \xaa\xa1
<j1101>...<j1194>               \xab\xa1
<j1201>...<j1294>               \xac\xa1
<j1301>...<j1394>               \xad\xa1
<j1401>...<j1494>               \xae\xa1
<j1501>...<j1594>               \xaf\xa1
<j1601>...<j1694>               \xb0\xa1
<j1701>...<j1794>               \xb1\xa1
<j1801>...<j1894>               \xb2\xa1
<j1901>...<j1994>               \xb3\xa1
<j2001>...<j2094>               \xb4\xa1
<j2101>...<j2194>               \xb5\xa1
<j2201>...<j2294>               \xb6\xa1
<j2301>...<j2394>               \xb7\xa1
<j2401>...<j2494>               \xb8\xa1
<j2501>...<j2594>               \xb9\xa1
<j2601>...<j2694>               \xba\xa1
<j2701>...<j2794>               \xbb\xa1
<j2801>...<j2894>               \xbc\xa1
<j2901>...<j2994>               \xbd\xa1
<j3001>...<j3094>               \xbe\xa1
<j3101>...<j3194>               \xbf\xa1
<j3201>...<j3294>               \xc0\xa1
<j3301>...<j3394>               \xc1\xa1
<j3401>...<j3494>               \xc2\xa1
<j3501>...<j3594>               \xc3\xa1
<j3601>...<j3694>               \xc4\xa1
<j3701>...<j3794>               \xc5\xa1
<j3801>...<j3894>               \xc6\xa1
<j3901>...<j3994>               \xc7\xa1
<j4001>...<j4094>               \xc8\xa1
```

```
<j4101>...<j4194>              \xc9\xa1
<j4201>...<j4294>              \xca\xa1
<j4301>...<j4394>              \xcb\xa1
<j4401>...<j4494>              \xcc\xa1
<j4501>...<j4594>              \xcd\xa1
<j4601>...<j4694>              \xce\xa1
<j4701>...<j4794>              \xcf\xa1
<j4801>...<j4894>              \xd0\xa1
<j4901>...<j4994>              \xd1\xa1
<j5001>...<j5094>              \xd2\xa1
<j5101>...<j5194>              \xd3\xa1
<j5201>...<j5294>              \xd4\xa1
<j5301>...<j5394>              \xd5\xa1
<j5401>...<j5494>              \xd6\xa1
<j5501>...<j5594>              \xd7\xa1
<j5601>...<j5694>              \xd8\xa1
<j5701>...<j5794>              \xd9\xa1
<j5801>...<j5894>              \xda\xa1
<j5901>...<j5994>              \xdb\xa1
<j6001>...<j6094>              \xdc\xa1
<j6101>...<j6194>              \xdd\xa1
<j6201>...<j6294>              \xde\xa1
<j6301>...<j6394>              \xdf\xa1
<j6401>...<j6494>              \xe0\xa1
<j6501>...<j6594>              \xe1\xa1
<j6601>...<j6694>              \xe2\xa1
<j6701>...<j6794>              \xe3\xa1
<j6801>...<j6894>              \xe4\xa1
<j6901>...<j6994>              \xe5\xa1
<j7001>...<j7094>              \xe6\xa1
<j7101>...<j7194>              \xe7\xa1
<j7201>...<j7294>              \xe8\xa1
<j7301>...<j7394>              \xe9\xa1
<j7401>...<j7494>              \xea\xa1
<j7501>...<j7594>              \xeb\xa1
<j7601>...<j7694>              \xec\xa1
<j7701>...<j7794>              \xed\xa1
<j7801>...<j7894>              \xee\xa1
<j7901>...<j7994>              \xef\xa1
<j8001>...<j8094>              \xf0\xa1
<j8101>...<j8194>              \xf1\xa1
<j8201>...<j8294>              \xf2\xa1
<j8301>...<j8394>              \xf3\xa1
<j8401>...<j8494>              \xf4\xa1
<j8501>...<j8594>              \xf5\xa1
```

```
<j8601>...<j8694>               \xf6\xa1
<j8701>...<j8794>               \xf7\xa1
<j8801>...<j8894>               \xf8\xa1
<j8901>...<j8994>               \xf9\xa1
<j9001>...<j9094>               \xfa\xa1
<j9101>...<j9194>               \xfb\xa1
<j9201>...<j9294>               \xfc\xa1
<j9301>...<j9394>               \xfd\xa1
<j9401>...<j9494>               \xfe\xa1
#       End JIS X 0208
#
#       JIS X 0212 Characters
#
<J0101>...<J0194>               \x8f\xa1\xa1
<J0201>...<J0294>               \x8f\xa2\xa1
<J0301>...<J0394>               \x8f\xa3\xa1
<J0401>...<J0494>               \x8f\xa4\xa1
<J0501>...<J0594>               \x8f\xa5\xa1
<J0601>...<J0694>               \x8f\xa6\xa1
<J0701>...<J0794>               \x8f\xa7\xa1
<J0801>...<J0894>               \x8f\xa8\xa1
<J0901>...<J0994>               \x8f\xa9\xa1
<J1001>...<J1094>               \x8f\xaa\xa1
<J1101>...<J1194>               \x8f\xab\xa1
<J1201>...<J1294>               \x8f\xac\xa1
<J1301>...<J1394>               \x8f\xad\xa1
<J1401>...<J1494>               \x8f\xae\xa1
<J1501>...<J1594>               \x8f\xaf\xa1
<J1601>...<J1694>               \x8f\xb0\xa1
<J1701>...<J1794>               \x8f\xb1\xa1
<J1801>...<J1894>               \x8f\xb2\xa1
<J1901>...<J1994>               \x8f\xb3\xa1
<J2001>...<J2094>               \x8f\xb4\xa1
<J2101>...<J2194>               \x8f\xb5\xa1
<J2201>...<J2294>               \x8f\xb6\xa1
<J2301>...<J2394>               \x8f\xb7\xa1
<J2401>...<J2494>               \x8f\xb8\xa1
<J2501>...<J2594>               \x8f\xb9\xa1
<J2601>...<J2694>               \x8f\xba\xa1
<J2701>...<J2794>               \x8f\xbb\xa1
<J2801>...<J2894>               \x8f\xbc\xa1
<J2901>...<J2994>               \x8f\xbd\xa1
<J3001>...<J3094>               \x8f\xbe\xa1
<J3101>...<J3194>               \x8f\xbf\xa1
<J3201>...<J3294>               \x8f\xc0\xa1
```

<J3301>...<J3394>	\x8f\xc1\xa1
<J3401>...<J3494>	\x8f\xc2\xa1
<J3501>...<J3594>	\x8f\xc3\xa1
<J3601>...<J3694>	\x8f\xc4\xa1
<J3701>...<J3794>	\x8f\xc5\xa1
<J3801>...<J3894>	\x8f\xc6\xa1
<J3901>...<J3994>	\x8f\xc7\xa1
<J4001>...<J4094>	\x8f\xc8\xa1
<J4101>...<J4194>	\x8f\xc9\xa1
<J4201>...<J4294>	\x8f\xca\xa1
<J4301>...<J4394>	\x8f\xcb\xa1
<J4401>...<J4494>	\x8f\xcc\xa1
<J4501>...<J4594>	\x8f\xcd\xa1
<J4601>...<J4694>	\x8f\xce\xa1
<J4701>...<J4794>	\x8f\xcf\xa1
<J4801>...<J4894>	\x8f\xd0\xa1
<J4901>...<J4994>	\x8f\xd1\xa1
<J5001>...<J5094>	\x8f\xd2\xa1
<J5101>...<J5194>	\x8f\xd3\xa1
<J5201>...<J5294>	\x8f\xd4\xa1
<J5301>...<J5394>	\x8f\xd5\xa1
<J5401>...<J5494>	\x8f\xd6\xa1
<J5501>...<J5594>	\x8f\xd7\xa1
<J5601>...<J5694>	\x8f\xd8\xa1
<J5701>...<J5794>	\x8f\xd9\xa1
<J5801>...<J5894>	\x8f\xda\xa1
<J5901>...<J5994>	\x8f\xdb\xa1
<J6001>...<J6094>	\x8f\xdc\xa1
<J6101>...<J6194>	\x8f\xdd\xa1
<J6201>...<J6294>	\x8f\xde\xa1
<J6301>...<J6394>	\x8f\xdf\xa1
<J6401>...<J6494>	\x8f\xe0\xa1
<J6501>...<J6594>	\x8f\xe1\xa1
<J6601>...<J6694>	\x8f\xe2\xa1
<J6701>...<J6794>	\x8f\xe3\xa1
<J6801>...<J6894>	\x8f\xe4\xa1
<J6901>...<J6994>	\x8f\xe5\xa1
<J7001>...<J7094>	\x8f\xe6\xa1
<J7101>...<J7194>	\x8f\xe7\xa1
<J7201>...<J7294>	\x8f\xe8\xa1
<J7301>...<J7394>	\x8f\xe9\xa1
<J7401>...<J7494>	\x8f\xea\xa1
<J7501>...<J7594>	\x8f\xeb\xa1
<J7601>...<J7694>	\x8f\xec\xa1
<J7701>...<J7794>	\x8f\xed\xa1

```
<J7801>...<J7894>               \x8f\xee\xa1
<J7901>...<J7994>               \x8f\xef\xa1
<J8001>...<J8094>               \x8f\xf0\xa1
<J8101>...<J8194>               \x8f\xf1\xa1
<J8201>...<J8294>               \x8f\xf2\xa1
<J8301>...<J8394>               \x8f\xf3\xa1
<J8401>...<J8494>               \x8f\xf4\xa1
<J8501>...<J8594>               \x8f\xf5\xa1
<J8601>...<J8694>               \x8f\xf6\xa1
<J8701>...<J8794>               \x8f\xf7\xa1
<J8801>...<J8894>               \x8f\xf8\xa1
<J8901>...<J8994>               \x8f\xf9\xa1
<J9001>...<J9094>               \x8f\xfa\xa1
<J9101>...<J9194>               \x8f\xfb\xa1
<J9201>...<J9294>               \x8f\xfc\xa1
<J9301>...<J9394>               \x8f\xfd\xa1
<J9401>...<J9494>               \x8f\xfe\xa1
#       End JIS X 0212
#
END CHARMAP
```

Annex B Example of Japanese Locale Definition

This Annex gives an example locale definition for eucJP encoding which is confoming to this profile. Implementations may modify this example as far as the corresponding definition in the body of this document allows.

Any comments in the following example are only for better understanding, they are not included into the specification.

B.1 LC_CTYPE

```
#
#       LC_CTYPE
#
LC_CTYPE
#
# upper class:
#       Uppercase alphabets in portable character set,
#       Roman uppercase letters in JIS X 0208,
#       Greek uppercase letters in JIS X 0208,
#       Russian uppercase letters in JIS X 0208, and
#       Uppercase letters in JIS X 0212.
#       Uppercase letters in udc or vdc classes may be added.

upper   <A>;<B>;<C>;<D>;<E>;<F>;<G>;<H>;<I>;<J>;<K>;<L>;<M>;\
        <N>;<O>;<P>;<Q>;<R>;<S>;<T>;<U>;<V>;<W>;<X>;<Y>;<Z>;\
        <j0333>;...;<j0358>;\
        <j0601>;...;<j0624>;\
        <j0701>;...;<j0733>;\
        <J0665>;...;<J0669>;\
        <J0671>;\
        <J0673>;\
        <J0674>;\
        <J0676>;\
        <J0734>;...;<J0746>;\
        <J0901>;\
        <J0902>;\
        <J0904>;\
        <J0906>;\
        <J0908>;\
        <J0909>;\
        <J0911>;...;<J0913>;\
        <J0915>;\
        <J0916>;\
        <J1001>;...;<J1024>;\
        <J1026>;...;<J1087>
#
# lower class:
#       Lowercase alphabets in portable character set,
#       Roman lowercase letters in JIS X 0208,
```

```
#       Greek lowercase letters in JIS X 0208,
#       Russian lowercase letters in JIS X 0208, and
#       Lowercase letters in JIS X 0212.
#       Lowercase letters in udc or vdc classes may be added

lower   <a>;<b>;<c>;<d>;<e>;<f>;<g>;<h>;<i>;<j>;<k>;<l>;<m>;\
        <n>;<o>;<p>;<q>;<r>;<s>;<t>;<u>;<v>;<w>;<x>;<y>;<z>;\
        <j0365>;...;<j0390>;\
        <j0633>;...;<j0656>;\
        <j0749>;...;<j0781>;\
        <J0681>;...;<J0692>;\
        <J0782>;...;<J0794>;\
        <J0933>;...;<J0948>;\
        <J1101>;...;<J1127>;\
        <J1129>;...;<J1135>;\
        <J1137>;...;<J1187>
#
# alpha class (default):
# It includes, by default, all the characters defined in the upper class
# and the lower class.
#
#
# digit class
#
digit   <zero>;<one>;<two>;<three>;<four>;\
        <five>;<six>;<seven>;<eight>;<nine>
#
# space class:
#       Space characters defined in ISO DIS 9945-2 "POSIX" locale
#       Space in JIS X 0208
#

space   <tab>;<newline>;<vertical-tab>;<form-feed>;\
        <carriage-return>;<space>;\
        <j0101>
#
# cntrl class:
#       C0 and C1 control characters as per ISO 6429.
#       SS2 and SS3 may be excluded if an accompanying charmap uses them
#       as single shifts to invoke graphic characters (as in EUC).
#       Control characters in udc or vdc may be added.

cntrl   <alert>;<backspace>;<tab>;<newline>;<vertical-tab>;\
        <form-feed>;<carriage-return>;\
        <NUL>;<SOH>;<STX>;<ETX>;<EOT>;<ENQ>;<ACK>;<SO>;\
```

```
        <SI>;<DLE>;<DC1>;<DC2>;<DC3>;<DC4>;<NAK>;<SYN>;\
        <ETB>;<CAN>;<EM>;<SUB>;<ESC>;<IS4>;<IS3>;<IS2>;\
        <IS1>;<DEL>;\
        <BPH>;<NBH>;<NEL>;<SSA>;<ESA>;<HTS>;<HTJ>;\
        <VTS>;<PLD>;<PLU>;<RI>;<SS2>;<SS3>;\
        <DCS>;<PU1>;<PU2>;<STS>;<CCH>;<MW>;<SPA>;<EPA>;\
        <SOS>;<SCI>;<CSI>;<ST>;<OSC>;<PM>;<APC>
#
# punct class:
# Special characters in udc or vdc may be added as long as they
# do not belong to classes cntrl, alpha, digit, jkanji, jhira,
# jkata or jdigit.
#
punct   <exclamation-mark>;<quotation-mark>;<number-sign>;\
        <dollar-sign>;<percent-sign>;<ampersand>;<apostrophe>;\
        <left-parenthesis>;<right-parenthesis>;<asterisk>;\
        <plus-sign>;<comma>;<hyphen>;<period>;<slash>;\
        <colon>;<semicolon>;<less-than-sign>;<equals-sign>;\
        <greater-than-sign>;<question-mark>;\
        <commercial-at>;\
        <left-square-bracket>;<backslash>;<yen-sign>;<right-square-bracket>;\
        <circumflex>;<underscore>;\
        <grave-accent>;\
        <left-curly-bracket>;<vertical-line>;<right-curly-bracket>;<tilde>;\
        <overline>;\
        <kana-full-stop>;<kana-opening-bracket>;<kana-closing-bracket>;\
        <kana-comma>;<kana-conjunctive>;\
        <j0102>;...;<j0110>;\
        <j0113>;...;<j0118>;\
        <j0123>;\
        <j0126>;\
        <j0129>;...;<j0194>;\
        <j0201>;...;<j0214>;\
        <j0226>;...;<j0233>;\
        <j0242>;...;<j0248>;\
        <j0260>;...;<j0274>;\
        <j0282>;...;<j0289>;\
        <j0294>;\
        <j0801>;...;<j0832>;\
        <J0215>;...;<J0225>;\
        <J0234>;...;<J0236>;\
        <J0275>;...;<J0281>

#
# graph class:
```

```
#       upper, lower, alpha, digit, xdigit
#       JIS X 0201 graphic characters
#       JIS X 0208 graphic characters
#       JIS X 0212 graphic characters
#       Graphic characters in udc or vdc classes may be added.

graph   <exclamation-mark>;<quotation-mark>;<number-sign>;\
        <dollar-sign>;<percent-sign>;<ampersand>;<apostrophe>;\
        <left-parenthesis>;<right-parenthesis>;<asterisk>;\
        <plus-sign>;<comma>;<hyphen>;<period>;<slash>;\
        <zero>;<one>;<two>;<three>;<four>;\
        <five>;<six>;<seven>;<eight>;<nine>;\
        <colon>;<semicolon>;<less-than-sign>;<equals-sign>;\
        <greater-than-sign>;<question-mark>;\
        <commercial-at>;\
        <A>;<B>;<C>;<D>;<E>;<F>;<G>;<H>;<I>;<J>;<K>;<L>;<M>;\
        <N>;<O>;<P>;<Q>;<R>;<S>;<T>;<U>;<V>;<W>;<X>;<Y>;<Z>;\
        <left-square-bracket>;<backslash>;<yen-sign>;<right-square-bracket>;\
        <circumflex>;<underscore>;\
        <grave-accent>;\
        <a>;<b>;<c>;<d>;<e>;<f>;<g>;<h>;<i>;<j>;<k>;<l>;<m>;\
        <n>;<o>;<p>;<q>;<r>;<s>;<t>;<u>;<v>;<w>;<x>;<y>;<z>;\
        <left-curly-bracket>;<vertical-line>;<right-curly-bracket>;<tilde>;\
        <overline>;\
        <kana-full-stop>;<kana-opening-bracket>;<kana-closing-bracket>;\
        <kana-comma>;<kana-conjunctive>;\
        <kana-WO>;...;<kana-tsu>;\
        <kana-prolonged-sound>;\
        <kana-A>;...;<kana-N>;\
        <kana-voiced-sound>;<kana-semivoiced-sound>;\
        <j0102>;...;<j0194>;\
        <j0201>;...;<j0214>;\
        <j0226>;...;<j0233>;\
        <j0242>;...;<j0248>;\
        <j0260>;...;<j0274>;\
        <j0282>;...;<j0289>;\
        <j0294>;\
        <j0316>;...;<j0325>;\
        <j0333>;...;<j0358>;\
        <j0365>;...;<j0390>;\
        <j0401>;...;<j0483>;\
        <j0501>;...;<j0586>;\
        <j0601>;...;<j0624>;\
        <j0633>;...;<j0656>;\
        <j0701>;...;<j0733>;\
```

```
<j0749>;...;<j0781>;\
<j0801>;...;<j0832>;\
<j1601>;...;<j1694>;\
<j1701>;...;<j1794>;\
<j1801>;...;<j1894>;\
<j1901>;...;<j1994>;\
<j2001>;...;<j2094>;\
<j2101>;...;<j2194>;\
<j2201>;...;<j2294>;\
<j2301>;...;<j2394>;\
<j2401>;...;<j2494>;\
<j2501>;...;<j2594>;\
<j2601>;...;<j2694>;\
<j2701>;...;<j2794>;\
<j2801>;...;<j2894>;\
<j2901>;...;<j2994>;\
<j3001>;...;<j3094>;\
<j3101>;...;<j3194>;\
<j3201>;...;<j3294>;\
<j3301>;...;<j3394>;\
<j3401>;...;<j3494>;\
<j3501>;...;<j3594>;\
<j3601>;...;<j3694>;\
<j3701>;...;<j3794>;\
<j3801>;...;<j3894>;\
<j3901>;...;<j3994>;\
<j4001>;...;<j4094>;\
<j4101>;...;<j4194>;\
<j4201>;...;<j4294>;\
<j4301>;...;<j4394>;\
<j4401>;...;<j4494>;\
<j4501>;...;<j4594>;\
<j4601>;...;<j4694>;\
<j4701>;...;<j4751>;\
<j4801>;...;<j4894>;\
<j4901>;...;<j4994>;\
<j5001>;...;<j5094>;\
<j5101>;...;<j5194>;\
<j5201>;...;<j5294>;\
<j5301>;...;<j5394>;\
<j5401>;...;<j5494>;\
<j5501>;...;<j5594>;\
<j5601>;...;<j5694>;\
<j5701>;...;<j5794>;\
<j5801>;...;<j5894>;\
```

<j5901>;...;<j5994>;\
<j6001>;...;<j6094>;\
<j6101>;...;<j6194>;\
<j6201>;...;<j6294>;\
<j6301>;...;<j6394>;\
<j6401>;...;<j6494>;\
<j6501>;...;<j6594>;\
<j6601>;...;<j6694>;\
<j6701>;...;<j6794>;\
<j6801>;...;<j6894>;\
<j6901>;...;<j6994>;\
<j7001>;...;<j7094>;\
<j7101>;...;<j7194>;\
<j7201>;...;<j7294>;\
<j7301>;...;<j7394>;\
<j7401>;...;<j7494>;\
<j7501>;...;<j7594>;\
<j7601>;...;<j7694>;\
<j7701>;...;<j7794>;\
<j7801>;...;<j7894>;\
<j7901>;...;<j7994>;\
<j8001>;...;<j8094>;\
<j8101>;...;<j8194>;\
<j8201>;...;<j8294>;\
<j8301>;...;<j8394>;\
<j8401>;...;<j8406>;\
<J0215>;...;<J0225>;\
<J0234>;...;<J0236>;\
<J0275>;...;<J0281>;\
<J0665>;...;<J0669>;\
<J0671>;\
<J0673>;\
<J0674>;\
<J0676>;\
<J0681>;...;<J0692>;\
<J0734>;...;<J0746>;\
<J0782>;...;<J0794>;\
<J0901>;\
<J0902>;\
<J0904>;\
<J0906>;\
<J0908>;\
<J0909>;\
<J0911>;...;<J0913>;\
<J0915>;\

```
<J0916>;\
<J0933>;...;<J0948>;\
<J1001>;...;<J1024>;\
<J1026>;...;<J1087>;\
<J1101>;...;<J1127>;\
<J1129>;...;<J1135>;\
<J1137>;...;<J1187>;\
<J1601>;...;<J1694>;\
<J1701>;...;<J1794>;\
<J1801>;...;<J1894>;\
<J1901>;...;<J1994>;\
<J2001>;...;<J2094>;\
<J2101>;...;<J2194>;\
<J2201>;...;<J2294>;\
<J2301>;...;<J2394>;\
<J2401>;...;<J2494>;\
<J2501>;...;<J2594>;\
<J2601>;...;<J2694>;\
<J2701>;...;<J2794>;\
<J2801>;...;<J2894>;\
<J2901>;...;<J2994>;\
<J3001>;...;<J3094>;\
<J3101>;...;<J3194>;\
<J3201>;...;<J3294>;\
<J3301>;...;<J3394>;\
<J3401>;...;<J3494>;\
<J3501>;...;<J3594>;\
<J3601>;...;<J3694>;\
<J3701>;...;<J3794>;\
<J3801>;...;<J3894>;\
<J3901>;...;<J3994>;\
<J4001>;...;<J4094>;\
<J4101>;...;<J4194>;\
<J4201>;...;<J4294>;\
<J4301>;...;<J4394>;\
<J4401>;...;<J4494>;\
<J4501>;...;<J4594>;\
<J4601>;...;<J4694>;\
<J4701>;...;<J4794>;\
<J4801>;...;<J4894>;\
<J4901>;...;<J4994>;\
<J5001>;...;<J5094>;\
<J5101>;...;<J5194>;\
<J5201>;...;<J5294>;\
<J5301>;...;<J5394>;\
```

```
        <J5401>;...;<J5494>;\
        <J5501>;...;<J5594>;\
        <J5601>;...;<J5694>;\
        <J5701>;...;<J5794>;\
        <J5801>;...;<J5894>;\
        <J5901>;...;<J5994>;\
        <J6001>;...;<J6094>;\
        <J6101>;...;<J6194>;\
        <J6201>;...;<J6294>;\
        <J6301>;...;<J6394>;\
        <J6401>;...;<J6494>;\
        <J6501>;...;<J6594>;\
        <J6601>;...;<J6694>;\
        <J6701>;...;<J6794>;\
        <J6801>;...;<J6894>;\
        <J6901>;...;<J6994>;\
        <J7001>;...;<J7094>;\
        <J7101>;...;<J7194>;\
        <J7201>;...;<J7294>;\
        <J7301>;...;<J7394>;\
        <J7401>;...;<J7494>;\
        <J7501>;...;<J7594>;\
        <J7601>;...;<J7694>;\
        <J7701>;...;<J7767>
#
# print class:
#       <space>, <j0101>,
#       upper, lower, alpha, digit, xdigit
#       JIS X 0201 printable characters
#       JIS X 0208 printable characters
#       JIS X 0212 printable characters
#       Pritable characters in udc or vdc classes may be added.

print   <space>;\
        <exclamation-mark>;<quotation-mark>;<number-sign>;\
        <dollar-sign>;<percent-sign>;<ampersand>;<apostrophe>;\
        <left-parenthesis>;<right-parenthesis>;<asterisk>;\
        <plus-sign>;<comma>;<hyphen>;<period>;<slash>;\
        <zero>;<one>;<two>;<three>;<four>;\
        <five>;<six>;<seven>;<eight>;<nine>;\
        <colon>;<semicolon>;<less-than-sign>;<equals-sign>;\
        <greater-than-sign>;<question-mark>;\
        <commercial-at>;\
        <A>;<B>;<C>;<D>;<E>;<F>;<G>;<H>;<I>;<J>;<K>;<L>;<M>;\
        <N>;<O>;<P>;<Q>;<R>;<S>;<T>;<U>;<V>;<W>;<X>;<Y>;<Z>;\
```

```
<left-square-bracket>;<backslash>;<yen-sign>;<right-square-bracket>;\
<circumflex>;<underscore>;\
<grave-accent>;\
<a>;<b>;<c>;<d>;<e>;<f>;<g>;<h>;<i>;<j>;<k>;<l>;<m>;\
<n>;<o>;<p>;<q>;<r>;<s>;<t>;<u>;<v>;<w>;<x>;<y>;<z>;\
<left-curly-bracket>;<vertical-line>;<right-curly-bracket>;<tilde>;\
<overline>;\
<kana-full-stop>;<kana-opening-bracket>;<kana-closing-bracket>;\
<kana-comma>;<kana-conjunctive>;\
<kana-WO>;...;<kana-tsu>;\
<kana-prolonged-sound>;\
<kana-A>;...;<kana-N>;\
<kana-voiced-sound>;<kana-semivoiced-sound>;\
<j0101>;...;<j0194>;\
<j0201>;...;<j0214>;\
<j0226>;...;<j0233>;\
<j0242>;...;<j0248>;\
<j0260>;...;<j0274>;\
<j0282>;...;<j0289>;\
<j0294>;\
<j0316>;...;<j0325>;\
<j0333>;...;<j0358>;\
<j0365>;...;<j0390>;\
<j0401>;...;<j0483>;\
<j0501>;...;<j0586>;\
<j0601>;...;<j0624>;\
<j0633>;...;<j0656>;\
<j0701>;...;<j0733>;\
<j0749>;...;<j0781>;\
<j0801>;...;<j0832>;\
<j1601>;...;<j1694>;\
<j1701>;...;<j1794>;\
<j1801>;...;<j1894>;\
<j1901>;...;<j1994>;\
<j2001>;...;<j2094>;\
<j2101>;...;<j2194>;\
<j2201>;...;<j2294>;\
<j2301>;...;<j2394>;\
<j2401>;...;<j2494>;\
<j2501>;...;<j2594>;\
<j2601>;...;<j2694>;\
<j2701>;...;<j2794>;\
<j2801>;...;<j2894>;\
<j2901>;...;<j2994>;\
<j3001>;...;<j3094>;\
```

```
<j3101>;...;<j3194>;\
<j3201>;...;<j3294>;\
<j3301>;...;<j3394>;\
<j3401>;...;<j3494>;\
<j3501>;...;<j3594>;\
<j3601>;...;<j3694>;\
<j3701>;...;<j3794>;\
<j3801>;...;<j3894>;\
<j3901>;...;<j3994>;\
<j4001>;...;<j4094>;\
<j4101>;...;<j4194>;\
<j4201>;...;<j4294>;\
<j4301>;...;<j4394>;\
<j4401>;...;<j4494>;\
<j4501>;...;<j4594>;\
<j4601>;...;<j4694>;\
<j4701>;...;<j4751>;\
<j4801>;...;<j4894>;\
<j4901>;...;<j4994>;\
<j5001>;...;<j5094>;\
<j5101>;...;<j5194>;\
<j5201>;...;<j5294>;\
<j5301>;...;<j5394>;\
<j5401>;...;<j5494>;\
<j5501>;...;<j5594>;\
<j5601>;...;<j5694>;\
<j5701>;...;<j5794>;\
<j5801>;...;<j5894>;\
<j5901>;...;<j5994>;\
<j6001>;...;<j6094>;\
<j6101>;...;<j6194>;\
<j6201>;...;<j6294>;\
<j6301>;...;<j6394>;\
<j6401>;...;<j6494>;\
<j6501>;...;<j6594>;\
<j6601>;...;<j6694>;\
<j6701>;...;<j6794>;\
<j6801>;...;<j6894>;\
<j6901>;...;<j6994>;\
<j7001>;...;<j7094>;\
<j7101>;...;<j7194>;\
<j7201>;...;<j7294>;\
<j7301>;...;<j7394>;\
<j7401>;...;<j7494>;\
<j7501>;...;<j7594>;\
```

```
<j7601>;...;<j7694>;\
<j7701>;...;<j7794>;\
<j7801>;...;<j7894>;\
<j7901>;...;<j7994>;\
<j8001>;...;<j8094>;\
<j8101>;...;<j8194>;\
<j8201>;...;<j8294>;\
<j8301>;...;<j8394>;\
<j8401>;...;<j8406>;\
<J0215>;...;<J0225>;\
<J0234>;...;<J0236>;\
<J0275>;...;<J0281>;\
<J0665>;...;<J0669>;\
<J0671>;\
<J0673>;\
<J0674>;\
<J0676>;\
<J0681>;...;<J0692>;\
<J0734>;...;<J0746>;\
<J0782>;...;<J0794>;\
<J0901>;\
<J0902>;\
<J0904>;\
<J0906>;\
<J0908>;\
<J0909>;\
<J0911>;...;<J0913>;\
<J0915>;\
<J0916>;\
<J0933>;...;<J0948>;\
<J1001>;...;<J1024>;\
<J1026>;...;<J1087>;\
<J1101>;...;<J1127>;\
<J1129>;...;<J1135>;\
<J1137>;...;<J1187>;\
<J1601>;...;<J1694>;\
<J1701>;...;<J1794>;\
<J1801>;...;<J1894>;\
<J1901>;...;<J1994>;\
<J2001>;...;<J2094>;\
<J2101>;...;<J2194>;\
<J2201>;...;<J2294>;\
<J2301>;...;<J2394>;\
<J2401>;...;<J2494>;\
<J2501>;...;<J2594>;\
```

<J2601>;...;<J2694>;\
<J2701>;...;<J2794>;\
<J2801>;...;<J2894>;\
<J2901>;...;<J2994>;\
<J3001>;...;<J3094>;\
<J3101>;...;<J3194>;\
<J3201>;...;<J3294>;\
<J3301>;...;<J3394>;\
<J3401>;...;<J3494>;\
<J3501>;...;<J3594>;\
<J3601>;...;<J3694>;\
<J3701>;...;<J3794>;\
<J3801>;...;<J3894>;\
<J3901>;...;<J3994>;\
<J4001>;...;<J4094>;\
<J4101>;...;<J4194>;\
<J4201>;...;<J4294>;\
<J4301>;...;<J4394>;\
<J4401>;...;<J4494>;\
<J4501>;...;<J4594>;\
<J4601>;...;<J4694>;\
<J4701>;...;<J4794>;\
<J4801>;...;<J4894>;\
<J4901>;...;<J4994>;\
<J5001>;...;<J5094>;\
<J5101>;...;<J5194>;\
<J5201>;...;<J5294>;\
<J5301>;...;<J5394>;\
<J5401>;...;<J5494>;\
<J5501>;...;<J5594>;\
<J5601>;...;<J5694>;\
<J5701>;...;<J5794>;\
<J5801>;...;<J5894>;\
<J5901>;...;<J5994>;\
<J6001>;...;<J6094>;\
<J6101>;...;<J6194>;\
<J6201>;...;<J6294>;\
<J6301>;...;<J6394>;\
<J6401>;...;<J6494>;\
<J6501>;...;<J6594>;\
<J6601>;...;<J6694>;\
<J6701>;...;<J6794>;\
<J6801>;...;<J6894>;\
<J6901>;...;<J6994>;\
<J7001>;...;<J7094>;\

```
        <J7101>;...;<J7194>;\
        <J7201>;...;<J7294>;\
        <J7301>;...;<J7394>;\
        <J7401>;...;<J7494>;\
        <J7501>;...;<J7594>;\
        <J7601>;...;<J7694>;\
        <J7701>;...;<J7767>
#
# xdigit class
#
xdigit  <zero>;<one>;<two>;<three>;<four>;\
        <five>;<six>;<seven>;<eight>;<nine>;\
        <A>;<B>;<C>;<D>;<E>;<F>;\
        <a>;<b>;<c>;<d>;<e>;<f>
#
# blank class
#
blank   <space>;<tab>
#
# Non-standard character classes specific to the ja_JP locale
#

charclass       ascii;line;jdigit;paren;jparen;jisx0201;jisx0201r;\
                jisx0208;jisx0212;udc;vdc;gaiji;jhira;jkata;jhankana;\
                jkanji;jspace
#
# ascii class: characters for which isascii() returns true
#       C0 control characters
#       SPACE and DELETE characters in ASCII (or JIS X 0201 Roman)
#       Graphic characters in ASCII (or JIS X 0201 Roman)

ascii   <NUL>;<SOH>;<STX>;<ETX>;<EOT>;<ENQ>;<ACK>;\
        <alert>;<BEL>;<backspace>;<tab>;<newline>;\
        <vertical-tab>;<form-feed>;<carriage-return>;\
        <SO>;<SI>;<DLE>;<DC1>;<DC2>;<DC3>;<DC4>;<NAK>;<SYN>;\
        <ETB>;<CAN>;<EM>;<SUB>;<ESC>;<IS4>;<IS3>;<IS2>;<IS1>;\
        <space>;\
        <exclamation-mark>;<quotation-mark>;<number-sign>;\
        <dollar-sign>;<percent-sign>;<ampersand>;<apostrophe>;\
        <left-parenthesis>;<right-parenthesis>;<asterisk>;\
        <plus-sign>;<comma>;<hyphen>;<period>;<slash>;\
        <zero>;<one>;<two>;<three>;<four>;\
        <five>;<six>;<seven>;<eight>;<nine>;\
        <colon>;<semicolon>;<less-than-sign>;<equals-sign>;\
        <greater-than-sign>;<question-mark>;\
```

```
        <commercial-at>;\
        <A>;<B>;<C>;<D>;<E>;<F>;<G>;<H>;<I>;<J>;<K>;<L>;<M>;\
        <N>;<O>;<P>;<Q>;<R>;<S>;<T>;<U>;<V>;<W>;<X>;<Y>;<Z>;\
        <left-square-bracket>;<backslash>;<right-square-bracket>;\
        <circumflex>;<underscore>;\
        <grave-accent>;\
        <a>;<b>;<c>;<d>;<e>;<f>;<g>;<h>;<i>;<j>;<k>;<l>;<m>;\
        <n>;<o>;<p>;<q>;<r>;<s>;<t>;<u>;<v>;<w>;<x>;<y>;<z>;\
        <left-curly-bracket>;<vertical-line>;<right-curly-bracket>;<tilde>;\
        <DEL>
#
# line class:
# The line drawing characters in JIS X 0208
# Line drawing characters in udc or vdc classes may be added.
#

line    <j0801>;...;<j0832>
#
# jdigit class: The digit characters in JIS X 0208

jdigit  <j0316>;...;<j0325>
#
# paren class:
# Parentheses and paired symbols in JIS X 0201 and JIS X 0208.
# Parentheses or paired symbols in udc or vdc classes may be added.
#

paren   <left-parenthesis>;<right-parenthesis>;\
        <left-square-bracket>;<right-square-bracket>;\
        <left-curly-bracket>;<right-curly-bracket>;\
        <kana-opening-bracket>;<kana-closing-bracket>;\
        <j0138>;...;<j0159>
#
# jparen class:
# The kana bracket characters in JIS X 0201 and the parentheses in JIS X 0208.
# Parentheses or paired symbols in udc or vdc classes may be added.
#

jparen  <kana-opening-bracket>;<kana-closing-bracket>;\
        <j0138>;...;<j0159>
#
# jisx0201 class:
# All the printable characters in JIS X 0201.
# Printable characters in udc or vdc classes with their code points
# in undefined area of JIS X 0201 may be added.
```

```
#
jisx0201 <alert>;<backspace>;<tab>;<newline>;<vertical-tab>;\
        <form-feed>;<carriage-return>;\
        <NUL>;<SOH>;<STX>;<ETX>;<EOT>;<ENQ>;<ACK>;<SO>;\
        <SI>;<DLE>;<DC1>;<DC2>;<DC3>;<DC4>;<NAK>;<SYN>;\
        <ETB>;<CAN>;<EM>;<SUB>;<ESC>;<IS4>;<IS3>;<IS2>;\
        <IS1>;\
        <space>;\
        <exclamation-mark>;<quotation-mark>;<number-sign>;\
        <dollar-sign>;<percent-sign>;<ampersand>;<apostrophe>;\
        <left-parenthesis>;<right-parenthesis>;<asterisk>;\
        <plus-sign>;<comma>;<hyphen>;<period>;<slash>;\
        <zero>;<one>;<two>;<three>;<four>;\
        <five>;<six>;<seven>;<eight>;<nine>;\
        <colon>;<semicolon>;<less-than-sign>;<equals-sign>;\
        <greater-than-sign>;<question-mark>;\
        <commercial-at>;\
        <A>;<B>;<C>;<D>;<E>;<F>;<G>;<H>;<I>;<J>;<K>;<L>;<M>;\
        <N>;<O>;<P>;<Q>;<R>;<S>;<T>;<U>;<V>;<W>;<X>;<Y>;<Z>;\
        <left-square-bracket>;<backslash>;<right-square-bracket>;\
        <circumflex>;<underscore>;\
        <grave-accent>;\
        <a>;<b>;<c>;<d>;<e>;<f>;<g>;<h>;<i>;<j>;<k>;<l>;<m>;\
        <n>;<o>;<p>;<q>;<r>;<s>;<t>;<u>;<v>;<w>;<x>;<y>;<z>;\
        <left-curly-bracket>;<vertical-line>;<right-curly-bracket>;<tilde>;\
        <DEL>;\
        <kana-full-stop>;<kana-opening-bracket>;<kana-closing-bracket>;\
        <kana-comma>;<kana-conjunctive>;\
        <kana-WO>;...;<kana-tsu>;\
        <kana-prolonged-sound>;\
        <kana-A>;...;<kana-N>;\
        <kana-voiced-sound>;<kana-semivoiced-sound>
#
# jisx0201r class:
# All the printable characters in the right hand side of JIS X 0201.
# Printable characters in udc or vdc classes with their code points
# in undefined area of JIS X 0201 right hand side may be added.
#
jisx0201r <kana-full-stop>;<kana-opening-bracket>;<kana-closing-bracket>;\
        <kana-comma>;<kana-conjunctive>;\
        <kana-WO>;...;<kana-tsu>;\
        <kana-prolonged-sound>;\
        <kana-A>;...;<kana-N>;\
        <kana-voiced-sound>;<kana-semivoiced-sound>
#
```

```
# jisx0208 class:
# All the printable characters in JIS X 0208.
# Printable characters in udc or vdc classes whose code points are in
# the undefined area of JIS X 0208 may be added.
#

jisx0208 <j0101>;...;<j0194>;\
        <j0201>;...;<j0214>;\
        <j0226>;...;<j0233>;\
        <j0242>;...;<j0248>;\
        <j0260>;...;<j0274>;\
        <j0282>;...;<j0289>;\
        <j0294>;\
        <j0316>;...;<j0325>;\
        <j0333>;...;<j0358>;\
        <j0365>;...;<j0390>;\
        <j0401>;...;<j0483>;\
        <j0501>;...;<j0586>;\
        <j0601>;...;<j0624>;\
        <j0633>;...;<j0656>;\
        <j0701>;...;<j0733>;\
        <j0749>;...;<j0781>;\
        <j0801>;...;<j0832>;\
        <j1601>;...;<j1694>;\
        <j1701>;...;<j1794>;\
        <j1801>;...;<j1894>;\
        <j1901>;...;<j1994>;\
        <j2001>;...;<j2094>;\
        <j2101>;...;<j2194>;\
        <j2201>;...;<j2294>;\
        <j2301>;...;<j2394>;\
        <j2401>;...;<j2494>;\
        <j2501>;...;<j2594>;\
        <j2601>;...;<j2694>;\
        <j2701>;...;<j2794>;\
        <j2801>;...;<j2894>;\
        <j2901>;...;<j2994>;\
        <j3001>;...;<j3094>;\
        <j3101>;...;<j3194>;\
        <j3201>;...;<j3294>;\
        <j3301>;...;<j3394>;\
        <j3401>;...;<j3494>;\
        <j3501>;...;<j3594>;\
        <j3601>;...;<j3694>;\
        <j3701>;...;<j3794>;\
```

```
<j3801>;...;<j3894>;\
<j3901>;...;<j3994>;\
<j4001>;...;<j4094>;\
<j4101>;...;<j4194>;\
<j4201>;...;<j4294>;\
<j4301>;...;<j4394>;\
<j4401>;...;<j4494>;\
<j4501>;...;<j4594>;\
<j4601>;...;<j4694>;\
<j4701>;...;<j4751>;\
<j4801>;...;<j4894>;\
<j4901>;...;<j4994>;\
<j5001>;...;<j5094>;\
<j5101>;...;<j5194>;\
<j5201>;...;<j5294>;\
<j5301>;...;<j5394>;\
<j5401>;...;<j5494>;\
<j5501>;...;<j5594>;\
<j5601>;...;<j5694>;\
<j5701>;...;<j5794>;\
<j5801>;...;<j5894>;\
<j5901>;...;<j5994>;\
<j6001>;...;<j6094>;\
<j6101>;...;<j6194>;\
<j6201>;...;<j6294>;\
<j6301>;...;<j6394>;\
<j6401>;...;<j6494>;\
<j6501>;...;<j6594>;\
<j6601>;...;<j6694>;\
<j6701>;...;<j6794>;\
<j6801>;...;<j6894>;\
<j6901>;...;<j6994>;\
<j7001>;...;<j7094>;\
<j7101>;...;<j7194>;\
<j7201>;...;<j7294>;\
<j7301>;...;<j7394>;\
<j7401>;...;<j7494>;\
<j7501>;...;<j7594>;\
<j7601>;...;<j7694>;\
<j7701>;...;<j7794>;\
<j7801>;...;<j7894>;\
<j7901>;...;<j7994>;\
<j8001>;...;<j8094>;\
<j8101>;...;<j8194>;\
<j8201>;...;<j8294>;\
```

```
        <j8301>;...;<j8394>;\
        <j8401>;...;<j8406>
#
# jisx0212 class:
# All the printable characters in JIS X 0212.
# Printable characters in udc or vdc classes whose code points are in
# the undefined area of JIS X 0212 may be added.
#

jisx0212 <J0215>;...;<J0225>;\
        <J0234>;...;<J0236>;\
        <J0275>;...;<J0281>;\
        <J0665>;...;<J0669>;\
        <J0671>;\
        <J0673>;\
        <J0674>;\
        <J0676>;\
        <J0681>;...;<J0692>;\
        <J0734>;...;<J0746>;\
        <J0782>;...;<J0794>;\
        <J0901>;\
        <J0902>;\
        <J0904>;\
        <J0906>;\
        <J0908>;\
        <J0909>;\
        <J0911>;...;<J0913>;\
        <J0915>;\
        <J0916>;\
        <J0933>;...;<J0948>;\
        <J1001>;...;<J1024>;\
        <J1026>;...;<J1087>;\
        <J1101>;...;<J1127>;\
        <J1129>;...;<J1135>;\
        <J1137>;...;<J1187>;\
        <J1601>;...;<J1694>;\
        <J1701>;...;<J1794>;\
        <J1801>;...;<J1894>;\
        <J1901>;...;<J1994>;\
        <J2001>;...;<J2094>;\
        <J2101>;...;<J2194>;\
        <J2201>;...;<J2294>;\
        <J2301>;...;<J2394>;\
        <J2401>;...;<J2494>;\
        <J2501>;...;<J2594>;\
```

<J2601>;...;<J2694>;\
<J2701>;...;<J2794>;\
<J2801>;...;<J2894>;\
<J2901>;...;<J2994>;\
<J3001>;...;<J3094>;\
<J3101>;...;<J3194>;\
<J3201>;...;<J3294>;\
<J3301>;...;<J3394>;\
<J3401>;...;<J3494>;\
<J3501>;...;<J3594>;\
<J3601>;...;<J3694>;\
<J3701>;...;<J3794>;\
<J3801>;...;<J3894>;\
<J3901>;...;<J3994>;\
<J4001>;...;<J4094>;\
<J4101>;...;<J4194>;\
<J4201>;...;<J4294>;\
<J4301>;...;<J4394>;\
<J4401>;...;<J4494>;\
<J4501>;...;<J4594>;\
<J4601>;...;<J4694>;\
<J4701>;...;<J4794>;\
<J4801>;...;<J4894>;\
<J4901>;...;<J4994>;\
<J5001>;...;<J5094>;\
<J5101>;...;<J5194>;\
<J5201>;...;<J5294>;\
<J5301>;...;<J5394>;\
<J5401>;...;<J5494>;\
<J5501>;...;<J5594>;\
<J5601>;...;<J5694>;\
<J5701>;...;<J5794>;\
<J5801>;...;<J5894>;\
<J5901>;...;<J5994>;\
<J6001>;...;<J6094>;\
<J6101>;...;<J6194>;\
<J6201>;...;<J6294>;\
<J6301>;...;<J6394>;\
<J6401>;...;<J6494>;\
<J6501>;...;<J6594>;\
<J6601>;...;<J6694>;\
<J6701>;...;<J6794>;\
<J6801>;...;<J6894>;\
<J6901>;...;<J6994>;\
<J7001>;...;<J7094>;\

```
        <J7101>;...;<J7194>;\
        <J7201>;...;<J7294>;\
        <J7301>;...;<J7394>;\
        <J7401>;...;<J7494>;\
        <J7501>;...;<J7594>;\
        <J7601>;...;<J7694>;\
        <J7701>;...;<J7767>
#
# udc class: user defined characters
#

udc
#
# vdc class: vender defined characters
#

vdc
#
# gaiji class: udc or vdc
#

gaiji
#
# jhira class:
# The Hiragana characters in JIS X 0208.
# Hiragana characters in udc or vdc classes may be added.
#

jhira  <j0401>;...;<j0483>;\
       <j0111>;<j0112>;\
       <j0121>;<j0122>;<j0128>
#
# jkata class:
# The Katakana characters in JIS X 0208 JIS X 0201.
# The voiced, semivoiced and prolonged sound marks in JIS X 0208
# and JIS X 0201.
# The Katakana iteration marks in JIS X 0208.
# Katakana characters in udc or vdc classes may be added.
#

jkata  <kana-WO>;<kana-a>;<kana-i>;<kana-u>;<kana-e>;<kana-o>;\
       <kana-ya>;<kana-yu>;<kana-yo>;<kana-tsu>;<kana-prolonged-sound>;\
       <kana-A>;<kana-I>;<kana-U>;<kana-E>;<kana-O>;\
       <kana-KA>;<kana-KI>;<kana-KU>;<kana-KE>;<kana-KO>;\
       <kana-SA>;<kana-SHI>;<kana-SU>;<kana-SE>;<kana-SO>;\
```

```
       <kana-TA>;<kana-CHI>;<kana-TSU>;<kana-TE>;<kana-TO>;\
       <kana-NA>;<kana-NI>;<kana-NU>;<kana-NE>;<kana-NO>;\
       <kana-HA>;<kana-HI>;<kana-FU>;<kana-HE>;<kana-HO>;\
       <kana-MA>;<kana-MI>;<kana-MU>;<kana-ME>;<kana-MO>;\
       <kana-YA>;<kana-YU>;<kana-YO>;<kana-RA>;<kana-RI>;\
       <kana-RU>;<kana-RE>;<kana-RO>;<kana-WA>;<kana-N>;\
       <kana-voiced-sound>;<kana-semivoiced-sound>;\
       <j0501>;...;<j0586>;\
       <j0111>;<j0112>;\
       <j0119>;<j0120>;<j0128>
#
# jhankana class:
# The Katakana characters in JIS X 0201.
# The voiced, semivoiced and prolonged sound marks in JIS X 0201.
# Katakana characters, Katakana symbols in JIS X 0201, or udc/vdc
# in undefined area of JIS X 0201 may be added.
#

jhankana  <kana-WO>;<kana-a>;<kana-i>;<kana-u>;<kana-e>;<kana-o>;\
       <kana-ya>;<kana-yu>;<kana-yo>;<kana-tsu>;<kana-prolonged-sound>;\
       <kana-A>;<kana-I>;<kana-U>;<kana-E>;<kana-O>;\
       <kana-KA>;<kana-KI>;<kana-KU>;<kana-KE>;<kana-KO>;\
       <kana-SA>;<kana-SHI>;<kana-SU>;<kana-SE>;<kana-SO>;\
       <kana-TA>;<kana-CHI>;<kana-TSU>;<kana-TE>;<kana-TO>;\
       <kana-NA>;<kana-NI>;<kana-NU>;<kana-NE>;<kana-NO>;\
       <kana-HA>;<kana-HI>;<kana-FU>;<kana-HE>;<kana-HO>;\
       <kana-MA>;<kana-MI>;<kana-MU>;<kana-ME>;<kana-MO>;\
       <kana-YA>;<kana-YU>;<kana-YO>;<kana-RA>;<kana-RI>;\
       <kana-RU>;<kana-RE>;<kana-RO>;<kana-WA>;<kana-N>;\
       <kana-voiced-sound>;<kana-semivoiced-sound>
#
# jkanji class: Kanji (Ideograms)
# Kanji in JIS X 0208 and JIS X 0212.
# Kanji Iteration mark in JIS X 0208.
# Han-numeral zero in JIS X 0208.
# Kanji in udc or vdc classes may be added.
#

jkanji  <j1601>;...;<j1694>;\
        <j1701>;...;<j1794>;\
        <j1801>;...;<j1894>;\
        <j1901>;...;<j1994>;\
        <j2001>;...;<j2094>;\
        <j2101>;...;<j2194>;\
        <j2201>;...;<j2294>;\
```

```
<j2301>;...;<j2394>;\
<j2401>;...;<j2494>;\
<j2501>;...;<j2594>;\
<j2601>;...;<j2694>;\
<j2701>;...;<j2794>;\
<j2801>;...;<j2894>;\
<j2901>;...;<j2994>;\
<j3001>;...;<j3094>;\
<j3101>;...;<j3194>;\
<j3201>;...;<j3294>;\
<j3301>;...;<j3394>;\
<j3401>;...;<j3494>;\
<j3501>;...;<j3594>;\
<j3601>;...;<j3694>;\
<j3701>;...;<j3794>;\
<j3801>;...;<j3894>;\
<j3901>;...;<j3994>;\
<j4001>;...;<j4094>;\
<j4101>;...;<j4194>;\
<j4201>;...;<j4294>;\
<j4301>;...;<j4394>;\
<j4401>;...;<j4494>;\
<j4501>;...;<j4594>;\
<j4601>;...;<j4694>;\
<j4701>;...;<j4751>;\
<j4801>;...;<j4894>;\
<j4901>;...;<j4994>;\
<j5001>;...;<j5094>;\
<j5101>;...;<j5194>;\
<j5201>;...;<j5294>;\
<j5301>;...;<j5394>;\
<j5401>;...;<j5494>;\
<j5501>;...;<j5594>;\
<j5601>;...;<j5694>;\
<j5701>;...;<j5794>;\
<j5801>;...;<j5894>;\
<j5901>;...;<j5994>;\
<j6001>;...;<j6094>;\
<j6101>;...;<j6194>;\
<j6201>;...;<j6294>;\
<j6301>;...;<j6394>;\
<j6401>;...;<j6494>;\
<j6501>;...;<j6594>;\
<j6601>;...;<j6694>;\
<j6701>;...;<j6794>;\
```

<j6801>;...;<j6894>;\
<j6901>;...;<j6994>;\
<j7001>;...;<j7094>;\
<j7101>;...;<j7194>;\
<j7201>;...;<j7294>;\
<j7301>;...;<j7394>;\
<j7401>;...;<j7494>;\
<j7501>;...;<j7594>;\
<j7601>;...;<j7694>;\
<j7701>;...;<j7794>;\
<j7801>;...;<j7894>;\
<j7901>;...;<j7994>;\
<j8001>;...;<j8094>;\
<j8101>;...;<j8194>;\
<j8201>;...;<j8294>;\
<j8301>;...;<j8394>;\
<j8401>;...;<j8406>;\
<J1601>;...;<J1694>;\
<J1701>;...;<J1794>;\
<J1801>;...;<J1894>;\
<J1901>;...;<J1994>;\
<J2001>;...;<J2094>;\
<J2101>;...;<J2194>;\
<J2201>;...;<J2294>;\
<J2301>;...;<J2394>;\
<J2401>;...;<J2494>;\
<J2501>;...;<J2594>;\
<J2601>;...;<J2694>;\
<J2701>;...;<J2794>;\
<J2801>;...;<J2894>;\
<J2901>;...;<J2994>;\
<J3001>;...;<J3094>;\
<J3101>;...;<J3194>;\
<J3201>;...;<J3294>;\
<J3301>;...;<J3394>;\
<J3401>;...;<J3494>;\
<J3501>;...;<J3594>;\
<J3601>;...;<J3694>;\
<J3701>;...;<J3794>;\
<J3801>;...;<J3894>;\
<J3901>;...;<J3994>;\
<J4001>;...;<J4094>;\
<J4101>;...;<J4194>;\
<J4201>;...;<J4294>;\
<J4301>;...;<J4394>;\

```
        <J4401>;...;<J4494>;\
        <J4501>;...;<J4594>;\
        <J4601>;...;<J4694>;\
        <J4701>;...;<J4794>;\
        <J4801>;...;<J4894>;\
        <J4901>;...;<J4994>;\
        <J5001>;...;<J5094>;\
        <J5101>;...;<J5194>;\
        <J5201>;...;<J5294>;\
        <J5301>;...;<J5394>;\
        <J5401>;...;<J5494>;\
        <J5501>;...;<J5594>;\
        <J5601>;...;<J5694>;\
        <J5701>;...;<J5794>;\
        <J5801>;...;<J5894>;\
        <J5901>;...;<J5994>;\
        <J6001>;...;<J6094>;\
        <J6101>;...;<J6194>;\
        <J6201>;...;<J6294>;\
        <J6301>;...;<J6394>;\
        <J6401>;...;<J6494>;\
        <J6501>;...;<J6594>;\
        <J6601>;...;<J6694>;\
        <J6701>;...;<J6794>;\
        <J6801>;...;<J6894>;\
        <J6901>;...;<J6994>;\
        <J7001>;...;<J7094>;\
        <J7101>;...;<J7194>;\
        <J7201>;...;<J7294>;\
        <J7301>;...;<J7394>;\
        <J7401>;...;<J7494>;\
        <J7501>;...;<J7594>;\
        <J7601>;...;<J7694>;\
        <J7701>;...;<J7767>;\
        <j0124>;<j0125>;<j0127>
#
# jspace class: The space character in JIS X 0208
#

jspace  <j0101>
#
# toupper and tolower:  also handle Roman, Greek and Russian
#                       characters in JIS X 0208 and JIS X 0212

toupper (<a>,<A>);(<b>,<B>);(<c>,<C>);(<d>,<D>);(<e>,<E>);\
```

```
(<f>,<F>);(<g>,<G>);(<h>,<H>);(<i>,<I>);(<j>,<J>);\
(<k>,<K>);(<l>,<L>);(<m>,<M>);(<n>,<N>);(<o>,<O>);\
(<p>,<P>);(<q>,<Q>);(<r>,<R>);(<s>,<S>);(<t>,<T>);\
(<u>,<U>);(<v>,<V>);(<w>,<W>);(<x>,<X>);(<y>,<Y>);\
(<z>,<Z>);\
(<j0365>,<j0333>);(<j0366>,<j0334>);(<j0367>,<j0335>);\
(<j0368>,<j0336>);(<j0369>,<j0337>);(<j0370>,<j0338>);\
(<j0371>,<j0339>);(<j0372>,<j0340>);(<j0373>,<j0341>);\
(<j0374>,<j0342>);(<j0375>,<j0343>);(<j0376>,<j0344>);\
(<j0377>,<j0345>);(<j0378>,<j0346>);(<j0379>,<j0347>);\
(<j0380>,<j0348>);(<j0381>,<j0349>);(<j0382>,<j0350>);\
(<j0383>,<j0351>);(<j0384>,<j0352>);(<j0385>,<j0353>);\
(<j0386>,<j0354>);(<j0387>,<j0355>);(<j0388>,<j0356>);\
(<j0389>,<j0357>);(<j0390>,<j0358>);\
(<j0633>,<j0601>);(<j0634>,<j0602>);(<j0635>,<j0603>);\
(<j0636>,<j0604>);(<j0637>,<j0605>);(<j0638>,<j0606>);\
(<j0639>,<j0607>);(<j0640>,<j0608>);(<j0641>,<j0609>);\
(<j0642>,<j0610>);(<j0643>,<j0611>);(<j0644>,<j0612>);\
(<j0645>,<j0613>);(<j0646>,<j0614>);(<j0647>,<j0615>);\
(<j0648>,<j0616>);(<j0649>,<j0617>);(<j0650>,<j0618>);\
(<j0651>,<j0619>);(<j0652>,<j0620>);(<j0653>,<j0621>);\
(<j0654>,<j0622>);(<j0655>,<j0623>);(<j0656>,<j0624>);\
(<j0749>,<j0701>);(<j0750>,<j0702>);(<j0751>,<j0703>);\
(<j0752>,<j0704>);(<j0753>,<j0705>);(<j0754>,<j0706>);\
(<j0755>,<j0707>);(<j0756>,<j0708>);(<j0757>,<j0709>);\
(<j0758>,<j0710>);(<j0759>,<j0711>);(<j0760>,<j0712>);\
(<j0761>,<j0713>);(<j0762>,<j0714>);(<j0763>,<j0715>);\
(<j0764>,<j0716>);(<j0765>,<j0717>);(<j0766>,<j0718>);\
(<j0767>,<j0719>);(<j0768>,<j0720>);(<j0769>,<j0721>);\
(<j0770>,<j0722>);(<j0771>,<j0723>);(<j0772>,<j0724>);\
(<j0773>,<j0725>);(<j0774>,<j0726>);(<j0775>,<j0727>);\
(<j0776>,<j0728>);(<j0777>,<j0729>);(<j0778>,<j0730>);\
(<j0779>,<j0731>);(<j0780>,<j0732>);(<j0781>,<j0733>);\
(<J0681>,<J0665>);(<J0682>,<J0666>);(<J0683>,<J0667>);\
(<J0684>,<J0668>);(<J0685>,<J0669>);(<J0687>,<J0671>);\
(<J0689>,<J0673>);(<J0690>,<J0674>);(<J0692>,<J0676>);\
(<J0782>,<J0734>);(<J0783>,<J0735>);(<J0784>,<J0736>);\
(<J0785>,<J0737>);(<J0786>,<J0738>);(<J0787>,<J0739>);\
(<J0788>,<J0740>);(<J0789>,<J0741>);(<J0790>,<J0742>);\
(<J0791>,<J0743>);(<J0792>,<J0744>);(<J0793>,<J0745>);\
(<J0794>,<J0746>);(<J0933>,<J0901>);(<J0934>,<J0902>);\
(<J0936>,<J0904>);(<J0938>,<J0906>);(<J0940>,<J0908>);\
(<J0941>,<J0909>);(<J0943>,<J0911>);(<J0944>,<J0912>);\
(<J0945>,<J0913>);(<J0947>,<J0915>);(<J0948>,<J0916>);\
(<J1101>,<J1001>);(<J1102>,<J1002>);(<J1103>,<J1003>);\
```

```
        (<J1104>,<J1004>);(<J1105>,<J1005>);(<J1106>,<J1006>);\
        (<J1107>,<J1007>);(<J1108>,<J1008>);(<J1109>,<J1009>);\
        (<J1110>,<J1010>);(<J1111>,<J1011>);(<J1112>,<J1012>);\
        (<J1113>,<J1013>);(<J1114>,<J1014>);(<J1115>,<J1015>);\
        (<J1116>,<J1016>);(<J1117>,<J1017>);(<J1118>,<J1018>);\
        (<J1119>,<J1019>);(<J1120>,<J1020>);(<J1121>,<J1021>);\
        (<J1122>,<J1022>);(<J1123>,<J1023>);(<J1124>,<J1024>);\
        (<J1126>,<J1026>);(<J1127>,<J1027>);(<J1129>,<J1029>);\
        (<J1130>,<J1030>);(<J1131>,<J1031>);(<J1132>,<J1032>);\
        (<J1133>,<J1033>);(<J1134>,<J1034>);(<J1135>,<J1035>);\
        (<J1137>,<J1037>);(<J1138>,<J1038>);(<J1139>,<J1039>);\
        (<J1140>,<J1040>);(<J1141>,<J1041>);(<J1142>,<J1042>);\
        (<J1143>,<J1043>);(<J1144>,<J1044>);(<J1145>,<J1045>);\
        (<J1146>,<J1046>);(<J1147>,<J1047>);(<J1148>,<J1048>);\
        (<J1149>,<J1049>);(<J1150>,<J1050>);(<J1151>,<J1051>);\
        (<J1152>,<J1052>);(<J1153>,<J1053>);(<J1154>,<J1054>);\
        (<J1155>,<J1055>);(<J1156>,<J1056>);(<J1157>,<J1057>);\
        (<J1158>,<J1058>);(<J1159>,<J1059>);(<J1160>,<J1060>);\
        (<J1161>,<J1061>);(<J1162>,<J1062>);(<J1163>,<J1063>);\
        (<J1164>,<J1064>);(<J1165>,<J1065>);(<J1166>,<J1066>);\
        (<J1167>,<J1067>);(<J1168>,<J1068>);(<J1169>,<J1069>);\
        (<J1170>,<J1070>);(<J1171>,<J1071>);(<J1172>,<J1072>);\
        (<J1173>,<J1073>);(<J1174>,<J1074>);(<J1175>,<J1075>);\
        (<J1176>,<J1076>);(<J1177>,<J1077>);(<J1178>,<J1078>);\
        (<J1179>,<J1079>);(<J1180>,<J1080>);(<J1181>,<J1081>);\
        (<J1182>,<J1082>);(<J1183>,<J1083>);(<J1184>,<J1084>);\
        (<J1185>,<J1085>);(<J1186>,<J1086>);(<J1187>,<J1087>)
tolower (<A>,<a>);(<B>,<b>);(<C>,<c>);(<D>,<d>);(<E>,<e>);\
        (<F>,<f>);(<G>,<g>);(<H>,<h>);(<I>,<i>);(<J>,<j>);\
        (<K>,<k>);(<L>,<l>);(<M>,<m>);(<N>,<n>);(<O>,<o>);\
        (<P>,<p>);(<Q>,<q>);(<R>,<r>);(<S>,<s>);(<T>,<t>);\
        (<U>,<u>);(<V>,<v>);(<W>,<w>);(<X>,<x>);(<Y>,<y>);\
        (<Z>,<z>);\
        (<j0333>,<j0365>);(<j0334>,<j0366>);(<j0335>,<j0367>);\
        (<j0336>,<j0368>);(<j0337>,<j0369>);(<j0338>,<j0370>);\
        (<j0339>,<j0371>);(<j0340>,<j0372>);(<j0341>,<j0373>);\
        (<j0342>,<j0374>);(<j0343>,<j0375>);(<j0344>,<j0376>);\
        (<j0345>,<j0377>);(<j0346>,<j0378>);(<j0347>,<j0379>);\
        (<j0348>,<j0380>);(<j0349>,<j0381>);(<j0350>,<j0382>);\
        (<j0351>,<j0383>);(<j0352>,<j0384>);(<j0353>,<j0385>);\
        (<j0354>,<j0386>);(<j0355>,<j0387>);(<j0356>,<j0388>);\
        (<j0357>,<j0389>);(<j0358>,<j0390>);\
        (<j0601>,<j0633>);(<j0602>,<j0634>);(<j0603>,<j0635>);\
        (<j0604>,<j0636>);(<j0605>,<j0637>);(<j0606>,<j0638>);\
        (<j0607>,<j0639>);(<j0608>,<j0640>);(<j0609>,<j0641>);\
```

```
(<j0610>,<j0642>);(<j0611>,<j0643>);(<j0612>,<j0644>);\
(<j0613>,<j0645>);(<j0614>,<j0646>);(<j0615>,<j0647>);\
(<j0616>,<j0648>);(<j0617>,<j0649>);(<j0618>,<j0650>);\
(<j0619>,<j0651>);(<j0620>,<j0652>);(<j0621>,<j0653>);\
(<j0622>,<j0654>);(<j0623>,<j0655>);(<j0624>,<j0656>);\
(<j0701>,<j0749>);(<j0702>,<j0750>);(<j0703>,<j0751>);\
(<j0704>,<j0752>);(<j0705>,<j0753>);(<j0706>,<j0754>);\
(<j0707>,<j0755>);(<j0708>,<j0756>);(<j0709>,<j0757>);\
(<j0710>,<j0758>);(<j0711>,<j0759>);(<j0712>,<j0760>);\
(<j0713>,<j0761>);(<j0714>,<j0762>);(<j0715>,<j0763>);\
(<j0716>,<j0764>);(<j0717>,<j0765>);(<j0718>,<j0766>);\
(<j0719>,<j0767>);(<j0720>,<j0768>);(<j0721>,<j0769>);\
(<j0722>,<j0770>);(<j0723>,<j0771>);(<j0724>,<j0772>);\
(<j0725>,<j0773>);(<j0726>,<j0774>);(<j0727>,<j0775>);\
(<j0728>,<j0776>);(<j0729>,<j0777>);(<j0730>,<j0778>);\
(<j0731>,<j0779>);(<j0732>,<j0780>);(<j0733>,<j0781>);\
(<J0665>,<J0681>);(<J0666>,<J0682>);(<J0667>,<J0683>);\
(<J0668>,<J0684>);(<J0669>,<J0685>);(<J0671>,<J0687>);\
(<J0673>,<J0689>);(<J0674>,<J0690>);(<J0676>,<J0692>);\
(<J0734>,<J0782>);(<J0735>,<J0783>);(<J0736>,<J0784>);\
(<J0737>,<J0785>);(<J0738>,<J0786>);(<J0739>,<J0787>);\
(<J0740>,<J0788>);(<J0741>,<J0789>);(<J0742>,<J0790>);\
(<J0743>,<J0791>);(<J0744>,<J0792>);(<J0745>,<J0793>);\
(<J0746>,<J0794>);(<J0901>,<J0933>);(<J0902>,<J0934>);\
(<J0904>,<J0936>);(<J0906>,<J0938>);(<J0908>,<J0940>);\
(<J0909>,<J0941>);(<J0911>,<J0943>);(<J0912>,<J0944>);\
(<J0913>,<J0945>);(<J0915>,<J0947>);(<J0916>,<J0948>);\
(<J1001>,<J1101>);(<J1002>,<J1102>);(<J1003>,<J1103>);\
(<J1004>,<J1104>);(<J1005>,<J1105>);(<J1006>,<J1106>);\
(<J1007>,<J1107>);(<J1008>,<J1108>);(<J1009>,<J1109>);\
(<J1010>,<J1110>);(<J1011>,<J1111>);(<J1012>,<J1112>);\
(<J1013>,<J1113>);(<J1014>,<J1114>);(<J1015>,<J1115>);\
(<J1016>,<J1116>);(<J1017>,<J1117>);(<J1018>,<J1118>);\
(<J1019>,<J1119>);(<J1020>,<J1120>);(<J1021>,<J1121>);\
(<J1022>,<J1122>);(<J1023>,<J1123>);(<J1024>,<J1124>);\
(<J1026>,<J1126>);(<J1027>,<J1127>);(<J1029>,<J1129>);\
(<J1030>,<J1130>);(<J1031>,<J1131>);(<J1032>,<J1132>);\
(<J1033>,<J1133>);(<J1034>,<J1134>);(<J1035>,<J1135>);\
(<J1037>,<J1137>);(<J1038>,<J1138>);(<J1039>,<J1139>);\
(<J1040>,<J1140>);(<J1041>,<J1141>);(<J1042>,<J1142>);\
(<J1043>,<J1143>);(<J1044>,<J1144>);(<J1045>,<J1145>);\
(<J1046>,<J1146>);(<J1047>,<J1147>);(<J1048>,<J1148>);\
(<J1049>,<J1149>);(<J1050>,<J1150>);(<J1051>,<J1151>);\
(<J1052>,<J1152>);(<J1053>,<J1153>);(<J1054>,<J1154>);\
(<J1055>,<J1155>);(<J1056>,<J1156>);(<J1057>,<J1157>);\
```

```
        (<J1058>,<J1158>);(<J1059>,<J1159>);(<J1060>,<J1160>);\
        (<J1061>,<J1161>);(<J1062>,<J1162>);(<J1063>,<J1163>);\
        (<J1064>,<J1164>);(<J1065>,<J1165>);(<J1066>,<J1166>);\
        (<J1067>,<J1167>);(<J1068>,<J1168>);(<J1069>,<J1169>);\
        (<J1070>,<J1170>);(<J1071>,<J1171>);(<J1072>,<J1172>);\
        (<J1073>,<J1173>);(<J1074>,<J1174>);(<J1075>,<J1175>);\
        (<J1076>,<J1176>);(<J1077>,<J1177>);(<J1078>,<J1178>);\
        (<J1079>,<J1179>);(<J1080>,<J1180>);(<J1081>,<J1181>);\
        (<J1082>,<J1182>);(<J1083>,<J1183>);(<J1084>,<J1184>);\
        (<J1085>,<J1185>);(<J1086>,<J1186>);(<J1087>,<J1187>)
END LC_CTYPE
```

B.2 LC_COLLATE

```
#
#       LC_COLLATE
#
LC_COLLATE

order_start     forward
#
#       ISO 646 IRV  or  JIS X 0201 Roman
#
<NUL>
<SOH>
<STX>
<ETX>
<EOT>
<ENQ>
<ACK>
<alert>
<backspace>
<tab>
<newline>
<vertical-tab>
<form-feed>
<carriage-return>
<SO>
<SI>
<DLE>
<DC1>
<DC2>
<DC3>
<DC4>
<NAK>
<SYN>
<ETB>
```

<CAN>
<EM>
<SUB>
<ESC>
<IS4>
<IS3>
<IS2>
<IS1>
<space>
<exclamation-mark>
<quotation-mark>
<number-sign>
<dollar-sign>
<percent-sign>
<ampersand>
<apostrophe>
<left-parenthesis>
<right-parenthesis>
<asterisk>
<plus-sign>
<comma>
<hyphen>
<period>
<slash>
<zero>
<one>
<two>
<three>
<four>
<five>
<six>
<seven>
<eight>
<nine>
<colon>
<semicolon>
<less-than-sign>
<equals-sign>
<greater-than-sign>
<question-mark>
<commercial-at>
<A>
<B>
<C>
<D>

<E>
<F>
<G>
<H>
<I>
<J>
<K>
<L>
<M>
<N>
<O>
<P>
<Q>
<R>
<S>
<T>
<U>
<V>
<W>
<X>
<Y>
<Z>
<left-square-bracket>
<backslash>
<right-square-bracket>
<circumflex>
<underscore>
<grave-accent>
<a>
<b>
<c>
<d>
<e>
<f>
<g>
<h>
<i>
<j>
<k>
<l>
<m>
<n>
<o>
<p>
<q>

<r>
<s>
<t>
<u>
<v>
<w>
<x>
<y>
<z>
<left-curly-bracket>
<vertical-line>
<right-curly-bracket>
<tilde>
<DEL>
#
C1 control
#
<BPH>
<NBH>
<NEL>
<SSA>
<ESA>
<HTS>
<HTJ>
<VTS>
<PLD>
<PLU>
<RI>
<SS2>
<SS3>
<DCS>
<PU1>
<PU2>
<STS>
<CCH>
<MW>
<SPA>
<EPA>
<SOS>
<SCI>
<CSI>
<ST>
<OSC>
<PM>
<APC>

```
#
#       JIS X 0201 Katakana
#
<kana-full-stop>
<kana-opening-bracket>
<kana-closing-bracket>
<kana-comma>
<kana-conjunctive>
<kana-WO>
<kana-a>
<kana-i>
<kana-u>
<kana-e>
<kana-o>
<kana-ya>
<kana-yu>
<kana-yo>
<kana-tsu>
<kana-prolonged-sound>
<kana-A>
<kana-I>
<kana-U>
<kana-E>
<kana-O>
<kana-KA>
<kana-KI>
<kana-KU>
<kana-KE>
<kana-KO>
<kana-SA>
<kana-SHI>
<kana-SU>
<kana-SE>
<kana-SO>
<kana-TA>
<kana-CHI>
<kana-TSU>
<kana-TE>
<kana-TO>
<kana-NA>
<kana-NI>
<kana-NU>
<kana-NE>
<kana-NO>
<kana-HA>
```

```
<kana-HI>
<kana-FU>
<kana-HE>
<kana-HO>
<kana-MA>
<kana-MI>
<kana-MU>
<kana-ME>
<kana-MO>
<kana-YA>
<kana-YU>
<kana-YO>
<kana-RA>
<kana-RI>
<kana-RU>
<kana-RE>
<kana-RO>
<kana-WA>
<kana-N>
<kana-voiced-sound>
<kana-semivoiced-sound>
#
#       JIS X 0208
#
<j0101>
...
<j0194>
<j0201>
...
<j0214>
<j0226>
...
<j0233>
<j0242>
...
<j0248>
<j0260>
...
<j0274>
<j0282>
...
<j0289>
<j0294>
<j0316>
...
```

<j0325>
<j0333>
...
<j0358>
<j0365>
...
<j0390>
<j0401>
...
<j0483>
<j0501>
...
<j0586>
<j0601>
...
<j0624>
<j0633>
...
<j0656>
<j0701>
...
<j0733>
<j0749>
...
<j0781>
<j0801>
...
<j0832>
<j1601>
...
<j1694>
<j1701>
...
<j1794>
<j1801>
...
<j1894>
<j1901>
...
<j1994>
<j2001>
...
<j2094>
<j2101>
...

<j2194>
<j2201>
...
<j2294>
<j2301>
...
<j2394>
<j2401>
...
<j2494>
<j2501>
...
<j2594>
<j2601>
...
<j2694>
<j2701>
...
<j2794>
<j2801>
...
<j2894>
<j2901>
...
<j2994>
<j3001>
...
<j3094>
<j3101>
...
<j3194>
<j3201>
...
<j3294>
<j3301>
...
<j3394>
<j3401>
...
<j3494>
<j3501>
...
<j3594>
<j3601>
...

<j3694>
<j3701>
...
<j3794>
<j3801>
...
<j3894>
<j3901>
...
<j3994>
<j4001>
...
<j4094>
<j4101>
...
<j4194>
<j4201>
...
<j4294>
<j4301>
...
<j4394>
<j4401>
...
<j4494>
<j4501>
...
<j4594>
<j4601>
...
<j4694>
<j4701>
...
<j4751>
<j4801>
...
<j4894>
<j4901>
...
<j4994>
<j5001>
...
<j5094>
<j5101>
...

<j5194>
<j5201>
...
<j5294>
<j5301>
...
<j5394>
<j5401>
...
<j5494>
<j5501>
...
<j5594>
<j5601>
...
<j5694>
<j5701>
...
<j5794>
<j5801>
...
<j5894>
<j5901>
...
<j5994>
<j6001>
...
<j6094>
<j6101>
...
<j6194>
<j6201>
...
<j6294>
<j6301>
...
<j6394>
<j6401>
...
<j6494>
<j6501>
...
<j6594>
<j6601>
...

<j6694>
<j6701>
...
<j6794>
<j6801>
...
<j6894>
<j6901>
...
<j6994>
<j7001>
...
<j7094>
<j7101>
...
<j7194>
<j7201>
...
<j7294>
<j7301>
...
<j7394>
<j7401>
...
<j7494>
<j7501>
...
<j7594>
<j7601>
...
<j7694>
<j7701>
...
<j7794>
<j7801>
...
<j7894>
<j7901>
...
<j7994>
<j8001>
...
<j8094>
<j8101>
...

```
<j8194>
<j8201>
...
<j8294>
<j8301>
...
<j8394>
<j8401>
...
<j8406>
#
#       JIS X 0212
#
<J0215>
...
<J0225>
<J0234>
...
<J0236>
<J0275>
...
<J0281>
<J0665>
...
<J0669>
<J0671>
<J0673>
<J0674>
<J0676>
<J0681>
...
<J0692>
<J0734>
...
<J0746>
<J0782>
...
<J0794>
<J0901>
<J0902>
<J0904>
<J0906>
<J0908>
<J0909>
<J0911>
```

...
<J0913>
<J0915>
<J0916>
<J0933>
...
<J0948>
<J1001>
...
<J1024>
<J1026>
...
<J1087>
<J1101>
...
<J1127>
<J1129>
...
<J1135>
<J1137>
...
<J1187>
<J1601>
...
<J1694>
<J1701>
...
<J1794>
<J1801>
...
<J1894>
<J1901>
...
<J1994>
<J2001>
...
<J2094>
<J2101>
...
<J2194>
<J2201>
...
<J2294>
<J2301>
...

<J2394>
<J2401>
...
<J2494>
<J2501>
...
<J2594>
<J2601>
...
<J2694>
<J2701>
...
<J2794>
<J2801>
...
<J2894>
<J2901>
...
<J2994>
<J3001>
...
<J3094>
<J3101>
...
<J3194>
<J3201>
...
<J3294>
<J3301>
...
<J3394>
<J3401>
...
<J3494>
<J3501>
...
<J3594>
<J3601>
...
<J3694>
<J3701>
...
<J3794>
<J3801>
...

<J3894>
<J3901>
...
<J3994>
<J4001>
...
<J4094>
<J4101>
...
<J4194>
<J4201>
...
<J4294>
<J4301>
...
<J4394>
<J4401>
...
<J4494>
<J4501>
...
<J4594>
<J4601>
...
<J4694>
<J4701>
...
<J4794>
<J4801>
...
<J4894>
<J4901>
...
<J4994>
<J5001>
...
<J5094>
<J5101>
...
<J5194>
<J5201>
...
<J5294>
<J5301>
...

<J5394>
<J5401>
...
<J5494>
<J5501>
...
<J5594>
<J5601>
...
<J5694>
<J5701>
...
<J5794>
<J5801>
...
<J5894>
<J5901>
...
<J5994>
<J6001>
...
<J6094>
<J6101>
...
<J6194>
<J6201>
...
<J6294>
<J6301>
...
<J6394>
<J6401>
...
<J6494>
<J6501>
...
<J6594>
<J6601>
...
<J6694>
<J6701>
...
<J6794>
<J6801>
...

```
<J6894>
<J6901>
...
<J6994>
<J7001>
...
<J7094>
<J7101>
...
<J7194>
<J7201>
...
<J7294>
<J7301>
...
<J7394>
<J7401>
...
<J7494>
<J7501>
...
<J7594>
<J7601>
...
<J7694>
<J7701>
...
<J7767>
#
#       Undefined
#
UNDEFINED
order_end

#
END LC_COLLATE
```

B.3 LC_MESSAGES

```
#
#       LC_MESSAGES
#
LC_MESSAGES

yesexpr "^[<y><Y><j0389><j0357>]"
noexpr  "^[<n><N><j0378><j0346>]"
```

```
#
END LC_MESSAGES
```

B.4 LC_MONETARY

```
#
#       LC_MONETARY
#
LC_MONETARY

int_curr_symbol         "<J><P><Y><space>"
currency_symbol         "<yen-sign>"
mon_decimal_point       ""
mon_thousands_sep       "<comma>"
mon_grouping            3
positive_sign           ""
negative_sign           "<hyphen>"
int_frac_digits         0
frac_digits             0
p_cs_precedes           1
p_sep_by_space          0
n_cs_precedes           1
n_sep_by_space          0
p_sign_posn             1
n_sign_posn             4

#
END LC_MONETARY
```

B.5 LC_NUMERIC

```
#
#       LC_NUMERIC
#
LC_NUMERIC

decimal_point   "<period>"
thousands_sep   "<comma>"
grouping        3

#
END LC_NUMERIC
```

B.6 LC_TIME

```
#
#       LC_TIME
#
LC_TIME
# abday: abbreviated weekday names
# abday is defined as the first letters of Japanese weekday names in Kanji,
```

```
# such as Nichi, Getsu, and Ka.
abday           "<j3892>";"<j2378>";"<j1848>";"<j3169>";\
                "<j4458>";"<j2266>";"<j3758>"

# day: full weekday names
# day is defined as full names of Japanese weekday names in Kanji, such as
# Nichiyoubi, Getsuyoubi, and Kayoubi.
day             "<j3892><j4543><j3892>";"<j2378><j4543><j3892>";\
                "<j1848><j4543><j3892>";"<j3169><j4543><j3892>";\
                "<j4458><j4543><j3892>";"<j2266><j4543><j3892>";\
                "<j3758><j4543><j3892>"

# abmon: abbreviated month names
# abmon is defined as two columns digit of month number, and Gatsu
# in Kanji.  If month number is less than ten, leading space is padded.
# The names are " 1Gatsu", " 2Gatsu", ..., "12Gatsu".
abmon           "<space><one><j2378>";"<space><two><j2378>";"<space><three><j2378>";\
                "<space><four><j2378>";"<space><five><j2378>";"<space><six><j2378>";\
                "<space><seven><j2378>";"<space><eight><j2378>";"<space><nine><j2378>";\
                "<one><zero><j2378>";"<one><one><j2378>";\
                "<one><two><j2378>"

# mon: full month names
# mon is defined as digit of month number, and Gatsu in Kanji.  No space
# is padded.  The names are "1Gatsu", "2Gatsu", ..., "12Gatsu".
mon             "<one><j2378>";"<two><j2378>";"<three><j2378>";\
                "<four><j2378>";"<five><j2378>";"<six><j2378>";\
                "<seven><j2378>";"<eight><j2378>";"<nine><j2378>";\
                "<one><zero><j2378>";"<one><one><j2378>";\
                "<one><two><j2378>"

# d_t_fmt: date and time format
# The format is defined as:
#       "%YNen%mGatsu%dNichi %HJi%MFun%SByou"
# which will be formatted as, for example,
#       "1993Nen02Gatsu06Nichi 08Ji59Fun07Byou"
d_t_fmt         "%Y<j3915>%m<j2378>%d<j3892><space>%H<j2794>%M<j4212>%S<j4135>"

# d_fmt: date format
# The format is defined as:
#       "%YNen%mGatsu%dNichi"
# which will be formatted as, for example,
#       "1993Nen02Gatsu06Nichi"
d_fmt           "%Y<j3915>%m<j2378>%d<j3892>"
```

```
# t_fmt: time format
# The format is defined as:
#       "%HJi%MFun%SByou"
# which will be formatted as, for example,
#       "08Ji59Fun07Byou"
t_fmt           "%H<j2794>%M<j4212>%S<j4135>"

# am_pm: ante meridiem (AM) and post meridiem (PM) strings
# "Gozen" for AM, "Gogo" for PM
am_pm           "<j2465><j3316>";"<j2465><j2469>"

# t_fmt_ampm: time format using am_pm
# The format is defined as:
#       "%p%IJi%MFun%SByou"
# which will be formatted as, for example,
#       "Gozen08Ji59Fun07Byou"
t_fmt_ampm      "%p%I<j2794>%M<j4212>%S<j4135>"

# era: year count and format for era
# era (Gengou) is defined as follows:
#       from 1990-01-01:                "Heisei%EyNen" (1990 is Heisei 2 Nen.)
#       from 1989-01-08 to 1989-12-31:  "HeiseiGannen"
#       from 1927-01-01 to 1989-01-07:  "Showa%EyNen" (1927 is Showa 2 Nen.)
era             "+:2:1990/01/01:+*:<j4231><j3214>:%EC%Ey<j3915>";\
                "+:1:1989/01/08:1989/12/31:<j4231><j3214>:%EC<j2421><j3915>";\
                "+:2:1927/01/01:1989/01/07:<j3028><j4734>:%EC%Ey<j3915>"

# era_d_fmt: date format using era
# The format is defined as:
#       "%EY%mGatsu%dNichi"
# which will be formatted as, for example,
#       "Heisei5Nen02Gatsu06Nichi"
era_d_fmt       "%EY%m<j2378>%d<j3892>"

# era_d_t_fmt: date and time format using era
# The format is defined as:
#       "%EY%mGatsu%dNichi %HJi%MFun%SByou"
# which will be formatted as, for example,
#       "Heisei5Nen02Gatsu06Nichi 08Ji59Fun07Byou"
era_d_t_fmt     "%EY%m<j2378>%d<j3892><space>%H<j2794>%M<j4212>%S<j4135>"

# era_t_fmt and alt_digits are not defined here.
#
END LC_TIME
```

Annex C Definition and Notes of Japanese EUC

The material for UI-OSF-USLP Joint Announcement Definition and Notes of Japanese EUC December 10, 1991

C.1 Definition of Japanese EUC

C.1.1 Scope

This document describes the definition of an Advanced Japanese EUC (AJEC) which is agreed to be a common Japanese character encoding among the Open Software Foundation, Inc. (OSF), UNIX International, Inc. (UI), and UNIX System Laboratories Pacific, Ltd. (USLP).

To deal with a new Japanese character set standard "Japanese Supplementary Character set Standard (JIS X 0212)" issued during 1990 in a consistent manner in the industry through the three organizations, this document evolves the definition of Japanese character sets assignment to the country specific part of EUC, without changing the EUC framework itself.

The definition refers to the International Standards (ISO standards) and the Japanese National Standards (JIS standards) relating to code for information interchange, as of November 1991.

This document also defines common usage of unassigned code points of the JIS standards in the AJEC, but not specify the implementation of fonts for the AJEC.

C.1.2 Definition of Japanese EUC

The assignment of graphic character sets, their code set width (a number of bytes of each character in the code set) and display width information for the supplementary code sets in EUC is defined by the AJEC as follows:

code set	number of bytes	display width	character set
0	1	1	ANS X 3.4 (ASCII)
1	2	2	JIS X 0208
2	1	1	JIS X 0201 Katakana
3	2	2	JIS X 0212

The number of bytes does not include that of "single shift" bytes.

The display width is meaningful for character devices where ASCII graphic characters occupy a single column while other characters occupy single or multiple columns.

C.1.3 Usage of Unassigned Code Points in JIS Code Tables

To avoid conflict and collision in the use of unassigned code points in the JIS X 0208 and JIS X 0212 code tables (so called "Empty Positions" in the JIS Rationale), the AJEC applies the following rules:

(1) As a subset of the unassigned code points of the JIS standards, the "Common Free Area" is defined in the AJEC as follows.
- The zone from row 85 to row 94 inclusive, out of the JIS X 0208unassigned code points.
- The zone from row 78 to row 94 inclusive, out of the JIS X 0212 unassigned code points.

(2) User and/or vendor defined characters can be assigned to the code points in the "Common Free Area."

C.1.4 Transparency of Codes

All code points of the AJEC including those unassigned are transparent. That is, all code points are treated as valid characters and kept unchanged during processing.

1. EUC: Extended UNIX Code. See "A. Definition of EUC" in the attached Notes and Rationale for more information.

C.2 Note and Rationale

C.2.1 Definition of EUC

EUC (Extended UNIX Code) is an encoding framework which was based on Japanese UNIX internal code introduced by the Japanese UNIX System Advisory Committee in the proposal entitled "Proposal to AT&T for Japanese Capability on UNIX System V" in 1985, and enhanced by AT&T Bell Laboratories for world-wide usage.

The code extension scheme is defined in compliance with the ISO 2022 code extension technique which can handle both single and multibyte characters concurrently. The scheme is designed to be applicable to every coded character set in each country or region in the world, as long as it is defined based on ISO 2022.

As allowed in ISO 2022, EUC can deal with a maximum of four character sets at the same time, without using locking shift. It permanently assigns the US ASCII character set to the "primary code set" (G0) and allows the assignment of country or user specific character set(s) to three "supplementary code sets" (G1−G3). The encoding scheme is defined as follows:

code set *1	code representation				associated standard \ character set
	1st byte	2nd byte	3rd byte	4th byte	
C0 set (C0) *2	000XXXXX	-	-	-	ISO 646
C1 set (C1) *3	100XXXXX	-	-	-	(ISO 6429)
SPACE character (SP)	00100000	-	-	-	
DELETE character (DEL)	01111111	-	-	-	
G0 set (G0) *4	0XXXXXXX	-	-	-	ASCII (ISO 646 variant)
G1 set (G1) *5	1XXXXXXX	-	-	-	country specific *6
	1XXXXXXX	1XXXXXXX	-	-	
	1XXXXXXX	1XXXXXXX	1XXXXXXX	-	
G2 set (G2) *5	SS2	1XXXXXXX	-	-	country specific *6
	SS2	1XXXXXXX	1XXXXXXX	-	
	SS2	1XXXXXXX	1XXXXXXX	1XXXXXXX	
G3 set (G3) *5	SS3	1XXXXXXX	-	-	country specific *6
	SS3	1XXXXXXX	1XXXXXXX	-	
	SS3	1XXXXXXX	1XXXXXXX	1XXXXXXX	

*1 The name of each code set is based on ISO 2022 (JIS X 0202).
*2 000XXXXX can have values in 00000000-00011111 inclusive.
*3 100XXXXX can have values in 10000000-10011111 inclusive.
*4 0XXXXXXX can have values in 00100001-01111110 inclusive.
*5 1XXXXXXX can have values in 10100000-11111111 inclusive.
*6 Multibyte character set of more than three bytes can be defined in the same way.

C.2.2 Control Sequences that Specifies Japanese EUC

As described above, this document is based on the latest standard regulations adopted as of November 1991. The control sequences that define the Japanese EUC code assignment are shown below.

From the viewpoint of ISO 2022 (JIS X 0202), the Japanese EUC encoding scheme is interpreted in such a way that the code set assignment is fixed and the following control sequences, except for the use of SS2 and SS3, are omitted under agreement between the interchanging parties.

C.2.2.1 Announce Sequences

Extension Facilities utilized	Escape Sequence
Invoke G0 as GL by LS0	ESC 02/0 05/0
Invoke G1 as GR by LS1R	ESC 02/0 05/3
Invoke G2 by SS2	ESC 02/0 05/10
Invoke G3 by SS3	ESC 02/0 05/11
Use C1	ESC 02/0 04/7

C.2.2.2 Designation of Graphic Character Sets

Designation	Escape Sequence	Character Set
G0	ESC 02/8 04/2 or ESC 02/8 04/10	ASCII (ANS X3.4-1968) or JIS X 0201-1976 Roman
G1	ESC 02/6 04/0 ESC 02/4 02/9 04/2	JIS X 0208-1990 Kanji
G2	ESC 02/10 04/9	JIS X 0201-1976 Katakana
G3	ESC 02/4 02/11 04/4	JIS X 0212-1990 Supplemental Kanji

Remarks: On the designation of G1

(1) The escape sequence that designates JIS X 0208-1990 is shown in accord with ISO 2022:1986 (JIS X 0202-1991) where a revision sequence for registered character sets has recently been introduced. JIS X 0208-1990 is regarded as a revised version of the previously registered JIS X 0208-1983. Therefore, strictly speaking, it is necessary to add a revision sequence (ESC 02/6 04/0).

On systems that are based on ISO 2022:1982 (JIS X 0202-1984), that is an older version of ISO 2022:1986 (JIS X 0202-1991), this revision sequence is not defined. The corresponding escape sequence of JIS X 0208-1990 might be:

ESC 02/4 02/9 04/2,

which is the same as the escape sequence of JIS X 0208-1983. In such systems, there is no way to distinguish them using the escape sequence.

(2) This document does not specify the exact version (year of issue) of standards on character sets. Therefore, on systems that designate JIS X 0208-1983 or JIS X 0208-1978, the corresponding escape sequences are as follows:

Designate JIS X 0208-1983 as G1	ESC 02/4 02/9 04/2
Designate JIS X 0208-1978 as G1	ESC 02/4 02/9 04/0

C.2.2.3 Invocation of Graphic Character Sets

Invocation	Escape sequence
G0	LS0 (00/15)
G1	LS1R (ESC 07/14)
G2	SS2 (08/14)
G3	SS3 (08/15)

C.2.3 Code Conversion

This section provides several important notes for code conversion between Japanese EUC and other codes.

C.2.3.1 JIS 7-bit code

In this section, "JIS 7-bit code" means an encoding scheme that does both of the following:

- designates basic character set ASCII (JIS X 0201 Roman character) as G0 and JIS X 0201 Katakana as G1 by default,
- switches G0 to ASCII, JIS X 0208 or JIS X 0212 back and forth by using appropriate escape sequences.

The underlying aim of this scheme is to support systems based on an ISO 2022 previous to 1982 (JIS X 0202 previous to 1984) as well to handle several Japanese character sets.

The following should be taken into consideration when converting codes between this JIS 7-bit code and Japanese EUC.

(1) Switching to JIS X 0208-1990 (so called "Kanji-In" sequence)

On systems that strictly comply with JIS X 0208-1990 and ISO 2022:1986 (JIS X 0202-1991), the sequence for designating G0 is as follows:

(1-1) ESC 2/6 4/0 ESC 2/4 2/8 4/2

But on systems that use the older version of ISO 2022 (JIS X 0202) and JIS X 0208 (1978 or 1983 version),

- a revision number sequence, such as ESC 2/6 4/0, may not be supported,
- only one of JIS X 0208-1978 or JIS X 0208-1983 may be supported,
- sometimes ESC 2/4 F may be used instead of ESC 2/4 2/8 F to designate a multibyte character set as G0,
- furthermore, vendor specific "Kanji-In" sequence may sometimes be used.

Because of the above reasons, it is conceivable that the following escape sequences may appear. Therefore, it is desirable to consider these escape sequences while converting codes in either direction.

(1-2) ESC 2/4 2/8 4/2	designation of JIS X 0208-1983 or 1990
(1-3) ESC 2/4 2/8 4/0	designation of JIS X 0208-1978
(1-4) ESC 2/4 4/2	designation of JIS X 0208-1983 or 1990
(1-5) ESC 2/4 4/0	designation of JIS X 0208-1978
(1-6) ESC 2/6 4/0 ESC 2/4 4/2	designation of JIS X 0208-1990
(1-7) Vendor Specific Sequence	for all of JIS X 0208-1978, 1983 and 1990

(2) Switching to JIS X 0212-1990

Taking account of the various escape sequences above, the following are conceivable for switching to JIS X 0212-1990:

(2-1) ESC 2/4 2/8 4/4	conforming to an ESC 2/4 2/8 F type
(2-2) ESC 2/4 4/4	conforming to an ESC 2/4 F type (obsolete)
(2-3) Vendor Specific Sequence	for JIS X 0212-1990

(3) Switching to ASCII or JIS X 0201 Roman character (so called "Kanji-Out" sequence)

The following variations need to be considered:

(3-1) ESC 2/8 4/2	ISO 646 American version (ASCII; ANS X3.4-1968)
(3-2) ESC 2/8 4/10	ISO 646 Japanese version (JIS X 0201 Roman)
(3-3) ESC 2/8 4/0	ISO 646 IRV
(3-4) Vendor Specific Sequence	for ASCII/JIS X 0201

(4) Switching to JIS X 0201 Katakana

When JIS X 0201 Katakana is designated as the G1 set, it is usually invoked by shift out (SO; 0/14).

(4-1) SO (0/14)	switch to G1 set (Katakana)

In this case, the return to G0 (ASCII/JIS X 0201 Roman character) is done with:

(3-5) SI (0/15)	switch to G0 set

and it becomes necessary to handle SI/SO in converting codes.

On the other hand, some systems designate and invoke JIS X 0201 Katakana as well as other code sets as the G0 set rather than as the G1 set. In this case, the following sequences should be also taken into consideration.

(4-2) ESC 2/8 4/9	designate JIS X 0201-1976 Katakana to G0
(4-3) Vendor Specific Sequence	for JIS X 0201-1976 Katakana

C.2.3.2 JIS 8-bit code

In this section, "JIS 8-bit code" means an encoding scheme that extends the above "JIS 7-bit code" to 8-bit. This encoding scheme designates the basic character set ASCII (JIS X 0201 Roman character) as G0 and invokes it as GL, designates JIS X 0201 Katakana as G1 and invokes it as GR, and switches G0 among ASCII, JIS X 0208 and JIS X 0212 using the appropriate escape sequence. The same care should

be taken here as that of JIS 7-bit code. One difference with the JIS 7-bit code concerns the conversion of JIS X 0201 Katakana.

(4-4) JIS X 0201 Katakana	By default, this character set is designated as G1 and invoked as GR (assuming that ESC 02/9 04/9 is issued). But this control sequence can be omitted under agreement between the interchanging parties. Therefore, JIS X 0201 Katakana is represented as a single byte 8-bit character.

C.2.3.3 Shifted JIS code (MS Kanji code)

As of November 1991, no widely accepted method is available for allocating the JIS X 0212 Supplemental Kanji character set to the framework of Shifted JIS code. It may be impossible to map the JIS X 0212 characters into a Shifted JIS code.

C.2.3.4 Transparency of user/vendor-defined characters in Common Free Area

The transparency of characters in JIS unassigned area (including user/vendor-defined characters) is not guaranteed after they have been converted to codes other than Japanese EUC.

But it is desirable to keep mapping relationship of code points during code conversion.

C.2.4 Free Area

The "Common Free Area" can be used to assign user/vendor-defined characters. This document does not specify the use of this area.

The Rationale for JIS X 0208 and JIS X 0212 describes two types of unassigned area. One is the "(area that strongly bears the nature of) reserved area" which the standard reserves for future extension. Another is the "(area that strongly bears the nature of) free area" which is the part for temporary and/or local use. Although the Rationale for JIS itself does not constitute a part of JIS standard, this AJEC document has defined the "Common Free Area" based on the intention of the JIS Rationale. This is to avoid the unnecessary collision as far as possible between user/vendor-defined characters and future standardized characters that will be added to JIS.

The Common Free Area is selected from zones that are left unassigned in a whole row in JIS with the following criteria:

- Exclude those zones that have a great possibility that vendor-defined characters have been assigned already, and may cause confusion if they are newly assigned as Common Free Area (9th through 15th rows inclusive, in JIS X 0208).
- Include those zones that are treated as "Free Area" in the Rationale on JIS (85th through 94th rows inclusive, in both of JIS X 0208 and JIS X 0212).
- Include those zones that form the contiguous area of maximum size in the rest of the unassigned code points, in order to secure as many user/vendor-defined characters as possible (74th through 84th rows inclusive, in JIS X 0212).

When using these code points, the following policies are recommended.

- Use user/vendor-defined characters in the order from 94th to 85th rows in both JIS X 0208 and JIS X 0212.
- Limit the use of JIS X 0212 "Reserved Area" (78th through 84th rows inclusive) to the case where all the code points of other "Free Area" have been used up for user/vendor-defined characters.

C.2.5 Problems in Handling Fonts

In general, fonts are implemented independently of character encoding by presentation devices such as terminals and printers or by their support subsystems. Although this document refers to the latest JIS standards available, it does not specify the implementation of fonts or graphical representation of characters "glyphs", so as not to exclude the use of existing (not the latest) various Japanese presentation devices. The following are points to be noted in applying the Japanese EUC defined in this document to actual presentation devices.

(1) On devices that have no JIS X 0212 fonts, characters of JIS X 0212 are not displayed nor printed.

(2) In the strict sense, JIS X 0208 has three versions: 1978, 1983, and 1990. And there are some differences among those versions. Presentation (display or print) of characters of JIS X 0208 depends on actual fonts implemented on the devices.

(3) Presentation of user/vendor-defined characters depends on the characters being defined on the devices. Presentation of these characters is not guaranteed on devices where these user/vendor-defined characters are not defined.

(4) Some presentation devices implement only either one of ANS X3.4 or JIS X 0201. Even on devices which implement both of them, application programs may not be able to specify which one to present.