                  Registration of a Cyrillic Character Set

Status of this Memo

   This memo provides information for the Internet community. It does
   not specify an Internet standard. Distribution of this memo is
   unlimited.

Introduction

   Though the proposed character set "koi8-r" is not currently an
   international standard, there is very large user community (including
   Relcom Net) supporting it. Factually, "koi8-r" is de-facto standard
   for Unix and global network applications in the former Soviet Union.
   This is the reason the Society of Unix User Groups (SUUG) believes
   "koi8-r" should be registered.

   MIME character set name: koi8-r

   Published specification: This standard is unpublished, but based on
   several published standards: GOST 19768-74 (old-koi8), ISO 6937/8
   (not registered) and variations: INIS-cyrillic, ISO 5427.

C language coding table from koi8-r to Unicode (ISO 10646) for the
upper half of the code table (the lower half of the table is fully
coinciding with ASCII character set)

```
   unsigned short koi82unicode[128] = {
      0x2500,0x2502,0x250c,0x2510,0x2514,0x2518,0x251c,0x2524,
      0x252c,0x2534,0x253c,0x2580,0x2584,0x2588,0x258c,0x2590,
      0x2591,0x2592,0x2593,0x2320,0x25a0,0x2219,0x221a,0x2248,
      0x2264,0x2265,0x00a0,0x2321,0x00b0,0x00b2,0x00b7,0x00f7,
      0x2550,0x2551,0x2552,0x0451,0x2553,0x2554,0x2555,0x2556,
      0x2557,0x2558,0x2559,0x255a,0x255b,0x255c,0x255d,0x255e,
      0x255f,0x2560,0x2561,0x0401,0x2562,0x2563,0x2564,0x2565,
      0x2566,0x2567,0x2568,0x2569,0x256a,0x256b,0x256c,0x00a9,
      0x044e,0x0430,0x0431,0x0446,0x0434,0x0435,0x0444,0x0433,
      0x0445,0x0438,0x0439,0x043a,0x043b,0x043c,0x043d,0x043e,
      0x043f,0x044f,0x0440,0x0441,0x0442,0x0443,0x0436,0x0432,
      0x044c,0x044b,0x0437,0x0448,0x044d,0x0449,0x0447,0x044a,
      0x042e,0x0410,0x0411,0x0426,0x0414,0x0415,0x0424,0x0413,
      0x0425,0x0418,0x0419,0x041a,0x041b,0x041c,0x041d,0x041e,
      0x041f,0x042f,0x0420,0x0421,0x0422,0x0423,0x0416,0x0412,
```

```
   0x042c,0x042b,0x0417,0x0428,0x042d,0x0429,0x0427,0x042a
};
```

   The description of all characters from the upper half of the table
   in compliance with ISO 10646 (Unicode)

   <hex-code> <description>

```
   80   FORMS LIGHT HORIZONTAL
   81   FORMS LIGHT VERTICAL
   82   FORMS LIGHT DOWN AND RIGHT
   83   FORMS LIGHT DOWN AND LEFT
   84   FORMS LIGHT UP AND RIGHT
   85   FORMS LIGHT UP AND LEFT
   86   FORMS LIGHT VERTICAL AND RIGHT
   87   FORMS LIGHT VERTICAL AND LEFT
   88   FORMS LIGHT DOWN AND HORIZONTAL
   89   FORMS LIGHT UP AND HORIZONTAL
   8A   FORMS LIGHT VERTICAL AND HORIZONTAL
   8B   UPPER HALF BLOCK
   8C   LOWER HALF BLOCK
   8D   FULL BLOCK
   8E   LEFT HALF BLOCK
   8F   RIGHT HALF BLOCK
   90   LIGHT SHADE
   91   MEDIUM SHADE
   92   DARK SHADE
   93   TOP HALF INTEGRAL
   94   BLACK SMALL SQUARE
   95   BULLET OPERATOR
   96   SQUARE ROOT
   97   ALMOST EQUAL TO
   98   LESS THAN OR EQUAL TO
   99   GREATER THAN OR EQUAL TO
   9A   NON-BREAKING SPACE
   9B   BOTTOM HALF INTEGRAL
   9C   DEGREE SIGN
   9D   SUPERSCRIPT DIGIT TWO
   9E   MIDDLE DOT
   9F   DIVISION SIGN
   A0   FORMS DOUBLE HORIZONTAL
   A1   FORMS DOUBLE VERTICAL
   A2   FORMS DOWN SINGLE AND RIGHT DOUBLE
   A3   CYRILLIC SMALL LETTER IO
   A4   FORMS DOWN DOUBLE AND RIGHT SINGLE
   A5   FORMS DOUBLE DOWN AND RIGHT
```

```
A6   FORMS DOWN SINGLE AND LEFT DOUBLE
A7   FORMS DOWN DOUBLE AND LEFT SINGLE
A8   FORMS DOUBLE DOWN AND LEFT
A9   FORMS UP SINGLE AND RIGHT DOUBLE
AA   FORMS UP DOUBLE AND RIGHT SINGLE
AB   FORMS DOUBLE UP AND RIGHT
AC   FORMS UP SINGLE AND LEFT DOUBLE
AD   FORMS UP DOUBLE AND LEFT SINGLE
AE   FORMS DOUBLE UP AND LEFT
AF   FORMS VERTICAL SINGLE AND RIGHT DOUBLE
B0   FORMS VERTICAL DOUBLE AND RIGHT SINGLE
B1   FORMS DOUBLE VERTICAL AND RIGHT
B2   FORMS VERTICAL SINGLE AND LEFT DOUBLE
B3   CYRILLIC CAPITAL LETTER IO
B4   FORMS VERTICAL DOUBLE AND LEFT SINGLE
B5   FORMS DOUBLE VERTICAL AND LEFT
B6   FORMS DOWN SINGLE AND HORIZONTAL DOUBLE
B7   FORMS DOWN DOUBLE AND HORIZONTAL SINGLE
B8   FORMS DOUBLE DOWN AND HORIZONTAL
B9   FORMS UP SINGLE AND HORIZONTAL DOUBLE
BA   FORMS UP DOUBLE AND HORIZONTAL SINGLE
BB   FORMS DOUBLE UP AND HORIZONTAL
BC   FORMS VERTICAL SINGLE AND HORIZONTAL DOUBLE
BD   FORMS VERTICAL DOUBLE AND HORIZONTAL SINGLE
BE   FORMS DOUBLE VERTICAL AND HORIZONTAL
BF   COPYRIGHT SIGN
C0   CYRILLIC SMALL LETTER IU
C1   CYRILLIC SMALL LETTER A
C2   CYRILLIC SMALL LETTER BE
C3   CYRILLIC SMALL LETTER TSE
C4   CYRILLIC SMALL LETTER DE
C5   CYRILLIC SMALL LETTER IE
C6   CYRILLIC SMALL LETTER EF
C7   CYRILLIC SMALL LETTER GE
C8   CYRILLIC SMALL LETTER KHA
C9   CYRILLIC SMALL LETTER II
CA   CYRILLIC SMALL LETTER SHORT II
CB   CYRILLIC SMALL LETTER KA
CC   CYRILLIC SMALL LETTER EL
CD   CYRILLIC SMALL LETTER EM
CE   CYRILLIC SMALL LETTER EN
CF   CYRILLIC SMALL LETTER O
D0   CYRILLIC SMALL LETTER PE
D1   CYRILLIC SMALL LETTER IA
D2   CYRILLIC SMALL LETTER ER
D3   CYRILLIC SMALL LETTER ES
D4   CYRILLIC SMALL LETTER TE
D5   CYRILLIC SMALL LETTER U
```

```
D6  CYRILLIC SMALL LETTER ZHE
D7  CYRILLIC SMALL LETTER VE
D8  CYRILLIC SMALL LETTER SOFT SIGN
D9  CYRILLIC SMALL LETTER YERI
DA  CYRILLIC SMALL LETTER ZE
DB  CYRILLIC SMALL LETTER SHA
DC  CYRILLIC SMALL LETTER REVERSED E
DD  CYRILLIC SMALL LETTER SHCHA
DE  CYRILLIC SMALL LETTER CHE
DF  CYRILLIC SMALL LETTER HARD SIGN
E0  CYRILLIC CAPITAL LETTER IU
E1  CYRILLIC CAPITAL LETTER A
E2  CYRILLIC CAPITAL LETTER BE
E3  CYRILLIC CAPITAL LETTER TSE
E4  CYRILLIC CAPITAL LETTER DE
E5  CYRILLIC CAPITAL LETTER IE
E6  CYRILLIC CAPITAL LETTER EF
E7  CYRILLIC CAPITAL LETTER GE
E8  CYRILLIC CAPITAL LETTER KHA
E9  CYRILLIC CAPITAL LETTER II
EA  CYRILLIC CAPITAL LETTER SHORT II
EB  CYRILLIC CAPITAL LETTER KA
EC  CYRILLIC CAPITAL LETTER EL
ED  CYRILLIC CAPITAL LETTER EM
EE  CYRILLIC CAPITAL LETTER EN
EF  CYRILLIC CAPITAL LETTER O
F0  CYRILLIC CAPITAL LETTER PE
F1  CYRILLIC CAPITAL LETTER IA
F2  CYRILLIC CAPITAL LETTER ER
F3  CYRILLIC CAPITAL LETTER ES
F4  CYRILLIC CAPITAL LETTER TE
F5  CYRILLIC CAPITAL LETTER U
F6  CYRILLIC CAPITAL LETTER ZHE
F7  CYRILLIC CAPITAL LETTER VE
F8  CYRILLIC CAPITAL LETTER SOFT SIGN
F9  CYRILLIC CAPITAL LETTER YERI
FA  CYRILLIC CAPITAL LETTER ZE
FB  CYRILLIC CAPITAL LETTER SHA
FC  CYRILLIC CAPITAL LETTER REVERSED E
FD  CYRILLIC CAPITAL LETTER SHCHA
FE  CYRILLIC CAPITAL LETTER CHE
FF  CYRILLIC CAPITAL LETTER HARD SIGN
```

Security Considerations

    Security issues are not discussed in this memo.

Author's Address

    Andrew A. Chernov
    RELCOM Development Team
    Moscow, Russia

    EMail: ache@astral.msk.su
    FIDOnet: 2:5020/23.34