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Published specification: This standard is unpublished, but based on several published standards: GOST 19768-74 (old-koi8), ISO 6937/8 (not registered) and variations: INIS-cyrillic, ISO 5427.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

```c
unsigned short koi82unicode[128] = {
  0x2500, 0x2502, 0x250c, 0x2510, 0x2514, 0x2518, 0x251c, 0x2524,
  0x252c, 0x2534, 0x253c, 0x2540, 0x2544, 0x2548, 0x254c, 0x2550,
  0x2554, 0x2558, 0x255c, 0x2560, 0x2564, 0x2568, 0x256c, 0x2570,
  0x2574, 0x2578, 0x257c, 0x2580, 0x2584, 0x2588, 0x258c, 0x2590,
  0x2591, 0x2593, 0x2597, 0x2599, 0x259b, 0x259d, 0x259f, 0x25a1,
  0x25a3, 0x25a7, 0x25a9, 0x25ac, 0x25ad, 0x25b0, 0x25b2, 0x25b4,
  0x25b6, 0x25b8, 0x25ba, 0x25bc, 0x25be, 0x25c0, 0x25c2, 0x25c4,
  0x25c6, 0x25c8, 0x25ca, 0x25cc, 0x25ce, 0x25d0, 0x25d2, 0x25d4,
  0x25d6, 0x25d8, 0x25db, 0x25dc, 0x25e0, 0x25e2, 0x25e4, 0x25e6,
  0x25e8, 0x25ea, 0x25ec, 0x25ee, 0x25f0, 0x25f2, 0x25f4, 0x25f6,
  0x25f8, 0x25fa, 0x25fc, 0x25fe,
};
```
The description of all characters from the upper half of the table in compliance with ISO 10646 (Unicode)

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