

Internet Draft
<[draft-demchenko-koi8-ru-00.txt](#)>
Expires in six months

Yuri Demchenko
Kiev Polytechnic Institute
October 1997

Registration of a Ukrainian Cyrillic Character Set KOI8-RU
(as extension to Russian KOI8-R and ISO-IR-111)

Status of this Memo

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months. Internet-Drafts may be updated, replaced, or obsoleted by other documents at any time. It is not appropriate to use Internet-Drafts as reference material or to cite them other than as a "working draft" or "work in progress".

To learn the current status of any Internet-Draft, please check the `1id-abstracts.txt` listing contained in the Internet-Drafts Shadow Directories on `ds.internic.net` (US East Coast), `nic.nordu.net` (Europe), `ftp.isi.edu` (US West Coast), or `munnari.oz.au` (Pacific Rim).

Abstract

This document provides information about widely used in Ukrainian Internet community (including Ukraine and worldwide Ukrainian speaking community) character set for mail and news exchange as well as for presentation WWW information resources in Ukrainian language. 'KOI8-RU' (KOI8 Russian and Ukrainian) summarises and completes de-facto standard `koi8-u` accepted by all Ukrainian community in the Internet and unofficially published at many sites. KOI8-RU is compatible with KOI8-R ([RFC 1489](#)) in all Cyrillic Letters and completes it with four Ukrainian letters which locations are complaint with ISO-IR-111. KOI8-RU should be registered to support and facilitate general and cultural information content development and dissemination. Support of Ukrainian language in new software product is restrained by absent of officially registered and widely published de-facto used Ukrainian charset.

Introduction

This document provides information about widely

used in Ukrainian Internet community character set for mail and news exchange as well as for presentation WWW information resources in Ukrainian language.

Though the proposed character set "KOI8-RU" is not currently an international standard, there is large Internet user community

Expires in six months

[Page 1]

(including Ukraine and worldwide Ukrainian speaking community) supporting it.

"KOI8-RU" (KOI8 Russian and Ukrainian) summarises and completes de-facto standard koi8-u accepted by all Ukrainian community in the Internet and unofficially published at many sites (F.E., ftp://ftp.ua.net/pub/info/encodings/koi8-u/ukr_chars_in_koi8-u_and_others.txt; ftp://ftp.gu.kiev.ua/pub/koi8-u/ukr_chars_in_koi8-u_and_others.txt; <http://cad.ntu-kpi.kiev.ua/multiling/KOI8-U.html>).

Ukrainian language is the 20th among the world's languages (<http://www.isoc.org:8080/langues/iso639.htm>) and supported not only in Ukraine as national state but among Ukrainian communities over the world.

KOI8-RU should be registered to support and facilitate general and cultural information content development and dissemination. Support of Ukrainian language in new software product is restrained by absent of officially registered and widely published de-facto used Ukrainian charset.

One of the problem now is that all old codepages ISO-IR-111, ISO 8859-5 doesn't include new Ukr. letter GHE (with upturn). Now it's registered in UNICODE 2.0.14 as Cyrillic GHE with upturn (0490 - capital, 0491 - small). It is used in more than 25 ukrainian words and carry in some cases specific national features.

Compatibility with other character sets

This standard is unpublished but based on several published standards: first of all, [RFC1489](#) (it is fully complaint in all russian letters), ISO 8859-5, ISO-IR-111, UNICODE 2.0.14.

KOI8-RU is compatible with KOI8-R in all Cyrillic Letters and completes it with four Ukrainian (#164, #180 - ukr. ie, #166, #182 - ukr. i, #167, #183 - ukr. yi, #173, #189 - ukr. ghe with upturn) and one Byelorussian (#174, #190 - byelorussian short u) letters which locations are complaint with ISO-IR-111.

All FORMS except positions occupied by Ukrainian and Byelorussian letters and Bullets in positions #148, #149, #158 coincide with KOI8-R.

Positions #147, #150-153, #155-#157, #159 are used for important characters which are currently missing from ISO-IR-111.

Expires in six months

[Page 2]

KOI8-RU character set description

The description of all characters from the upper half of the table is compliance with ISO 10646 (Unicode). All Russian letters places have been left at their original KOI8-R places. Introduced new ukrainian letters occupy positions where they are used as standard-de-facto in Ukrainian language applications and newsgroups exchange accepted all Ukrainian language community.

<decimal> <hex-code> <Unicode> <description>

128	80	U2500	FORMS LIGHT HORIZONTAL
129	81	U2502	FORMS LIGHT VERTICAL
130	82	U250C	FORMS LIGHT DOWN AND RIGHT
131	83	U2510	FORMS LIGHT DOWN AND LEFT
132	84	U2514	FORMS LIGHT UP AND RIGHT
133	85	U2518	FORMS LIGHT UP AND LEFT
134	86	U251C	FORMS LIGHT VERTICAL AND RIGHT
135	87	U2524	FORMS LIGHT VERTICAL AND LEFT
136	88	U252C	FORMS LIGHT DOWN AND HORIZONTAL
137	89	U2534	FORMS LIGHT UP AND HORIZONTAL
138	8A	U253C	FORMS LIGHT VERTICAL AND HORIZONTAL
139	8B	U2580	UPPER HALF BLOCK
140	8C	U2584	LOWER HALF BLOCK
141	8D	U2588	FULL BLOCK
142	8E	U258C	LEFT HALF BLOCK
143	8F	U2590	RIGHT HALF BLOCK
144	90	U2591	LIGHT SHADE
145	91	U2592	MEDIUM SHADE
146	92	U2593	DARK SHADE
147	93	U201C	LEFT DOUBLE QUOTATION MARK
148	94	U25A0	BLACK SQUARE
149	95	U2219	BULLET OPERATOR
150	96	U201D	RIGHT DOUBLE QUOTATION MARK
151	97	U2014	EM DASH
152	98	U2116	NUMERO SIGN
153	99	U2122	TRADE MARK SIGN
154	9A	U00A0	NONBREAKING SPACE
155	9B	U00BB	RIGHT-POINTING DOUBLE ANGLE QUOTATION MARK
156	9C	U00AE	REGISTERED SIGN
157	9D	U00AB	LEFT-POINTING DOUBLE ANGLE QUOTATION MARK
158	9E	U00B7	MIDDLE DOT
159	9F	U00A4	CURRENCY SIGN
160	A0	U2550	FORMS DOUBLE HORIZONTAL
161	A1	U2551	FORMS DOUBLE VERTICAL

162	A2	U2552	FORMS DOWN SINGLE AND RIGHT DOUBLE
163	A3	U0451	CYRILLIC SMALL LETTER IO
164	A4	U0454	CYRILLIC SMALL LETTER UKRAINIAN IE
			UKR
165	A5	U2554	FORMS DOUBLE DOWN AND RIGHT

Expires in six months

[Page 3]

166	A6	U0456	CYRILLIC SMALL LETTER BELORUSSIAN- UKRAINIAN I	UKR
167	A7	U0457	CYRILLIC SMALL LETTER YI (UKRAINIAN)	UKR
168	A8	U2557	FORMS DOUBLE DOWN AND LEFT	
169	A9	U2558	FORMS UP SINGLE AND RIGHT DOUBLE	
170	AA	U2559	FORMS UP DOUBLE AND RIGHT SINGLE	
171	AB	U255A	FORMS DOUBLE UP AND RIGHT	
172	AC	U255B	FORMS UP SINGLE AND LEFT DOUBLE	
173	AD	U0491	CYRILLIC SMALL LETTER UKRAINIAN GHE (WITH UPTURN)	UKR
174	AE	U045E	CYRILLIC SMALL LETTER BELORUSSIAN SHORT U	BYEL
175	AF	U255E	FORMS VERTICAL SINGLE AND RIGHT DOUBLE	
176	*B0	U255F	FORMS VERTICAL DOUBLE AND RIGHT SINGLE	
177	B1	U2560	FORMS DOUBLE VERTICAL AND RIGHT	
178	B2	U2561	FORMS VERTICAL SINGLE AND LEFT DOUBLE	
179	B3	U0401	CYRILLIC CAPITAL LETTER IO	
180	B4	U0403	CYRILLIC CAPITAL LETTER UKRAINIAN	
IE				UKR
181	B5	U2563	FORMS DOUBLE VERTICAL AND LEFT	
182	B6	U0406	CYRILLIC CAPITAL LETTER BELORUSSIAN- UKRAINIAN I	UKR
183	B7	U0407	CYRILLIC CAPITAL LETTER YI (UKRAINIAN)	UKR
184	B8	U2566	FORMS DOUBLE DOWN AND HORIZONTAL	
185	B9	U2567	FORMS UP SINGLE AND HORIZONTAL DOUBLE	
186	BA	U2568	FORMS UP DOUBLE AND HORIZONTAL SINGLE	
187	BB	U2569	FORMS DOUBLE UP AND HORIZONTAL	
188	BC	U256A	FORMS VERTICAL SINGLE AND HORIZONTAL DOUBLE	
189	BD	U0490	CYRILLIC CAPITAL LETTER UKRAINIAN GHE (WITH UPTURN)	UKR
190	BE	U040E	CYRILLIC CAPITAL LETTER BELORUSSIAN SHORT U	BYEL
191	BF	U00A9	COPYRIGHT SIGN	
192	C0	U044E	CYRILLIC SMALL LETTER IU	
193	C1	U0430	CYRILLIC SMALL LETTER A	
194	C2	U0431	CYRILLIC SMALL LETTER BE	
195	C3	U0446	CYRILLIC SMALL LETTER TSE	
196	C4	U0434	CYRILLIC SMALL LETTER DE	
197	C5	U0435	CYRILLIC SMALL LETTER IE	
198	C6	U0444	CYRILLIC SMALL LETTER EF	
199	C7	U0433	CYRILLIC SMALL LETTER GE	
200	C8	U0445	CYRILLIC SMALL LETTER KHA	

201	C9	U0438	CYRILLIC SMALL LETTER II
202	CA	U0439	CYRILLIC SMALL LETTER SHORT II
203	CB	U043A	CYRILLIC SMALL LETTER KA
204	CC	U043B	CYRILLIC SMALL LETTER EL
205	CD	U043C	CYRILLIC SMALL LETTER EM
206	CE	U043D	CYRILLIC SMALL LETTER EN

Expires in six months

[Page 4]

207	CF	U043E	CYRILLIC SMALL LETTER O
208	D0	U043F	CYRILLIC SMALL LETTER PE
209	D1	U044F	CYRILLIC SMALL LETTER IA
210	D2	U0440	CYRILLIC SMALL LETTER ER
211	D3	U0441	CYRILLIC SMALL LETTER ES
212	D4	U0442	CYRILLIC SMALL LETTER TE
213	D5	U0443	CYRILLIC SMALL LETTER U
214	D6	U0436	CYRILLIC SMALL LETTER ZHE
215	D7	U0432	CYRILLIC SMALL LETTER VE
216	D8	U044C	CYRILLIC SMALL LETTER SOFT SIGN
217	D9	U044B	CYRILLIC SMALL LETTER YERI
218	DA	U0437	CYRILLIC SMALL LETTER ZE
219	DB	U0448	CYRILLIC SMALL LETTER SHA
220	DC	U044D	CYRILLIC SMALL LETTER REVERSED E
221	DD	U0449	CYRILLIC SMALL LETTER SHCHA
222	DE	U0447	CYRILLIC SMALL LETTER CHE
223	DF	U044A	CYRILLIC SMALL LETTER HARD SIGN
224	E0	U042E	CYRILLIC CAPITAL LETTER IU
225	E1	U0410	CYRILLIC CAPITAL LETTER A
226	E2	U0411	CYRILLIC CAPITAL LETTER BE
227	E3	U0426	CYRILLIC CAPITAL LETTER TSE
228	E4	U0414	CYRILLIC CAPITAL LETTER DE
229	E5	U0415	CYRILLIC CAPITAL LETTER IE
230	E6	U0424	CYRILLIC CAPITAL LETTER EF
231	E7	U0413	CYRILLIC CAPITAL LETTER GE
232	E8	U0425	CYRILLIC CAPITAL LETTER KHA
233	E9	U0418	CYRILLIC CAPITAL LETTER II
234	EA	U0419	CYRILLIC CAPITAL LETTER SHORT II
235	EB	U041A	CYRILLIC CAPITAL LETTER KA
236	EC	U041B	CYRILLIC CAPITAL LETTER EL
237	ED	U041C	CYRILLIC CAPITAL LETTER EM
238	EE	U041D	CYRILLIC CAPITAL LETTER EN
239	EF	U041E	CYRILLIC CAPITAL LETTER O
240	F0	U041F	CYRILLIC CAPITAL LETTER PE
241	F1	U042F	CYRILLIC CAPITAL LETTER IA
242	F2	U0420	CYRILLIC CAPITAL LETTER ER
243	F3	U0421	CYRILLIC CAPITAL LETTER ES
244	F4	U0422	CYRILLIC CAPITAL LETTER TE
245	F5	U0423	CYRILLIC CAPITAL LETTER U
246	F6	U0416	CYRILLIC CAPITAL LETTER ZHE
247	F7	U0412	CYRILLIC CAPITAL LETTER VE
248	F8	U042C	CYRILLIC CAPITAL LETTER SOFT SIGN
249	F9	U042B	CYRILLIC CAPITAL LETTER YERI
250	FA	U0417	CYRILLIC CAPITAL LETTER ZE
251	FB	U0428	CYRILLIC CAPITAL LETTER SHA
252	FC	U042D	CYRILLIC CAPITAL LETTER REVERSED E
253	FD	U0429	CYRILLIC CAPITAL LETTER SHCHA
254	FE	U0427	CYRILLIC CAPITAL LETTER CHE

255

FF

U042A

CYRILLIC CAPITAL LETTER HARD SIGN

Expires in six months

[Page 5]

Legend

UKR - New included Ukrainian letters

BYEL - New included Byelorussian letters

APPENDIX A

DIFFERENCE OF KOI8-RU from EXISTING KOI8-R and ISO-IR-111

KOI8-RU is compatible with KOI8-R in all Cyrillic Letters and completes it with Ukrainian letters UKRAINIAN IE #164, #180, CYRILLIC SMALL LETTER BELORUSSIAN-UKRAINIAN I #166, #182, UKRAINIAN YI #167, #183, UKRAINIAN GHE (WITH UPTURN) #173, #189, BELORUSSIAN SHORT U #174, #190.

Positions #147, #150 - #153, #155-#157, #159 are used for important characters which are currently missing from ISO-IR-111. In all other positions FORMS coincide with KOI8-R.

	147	93	U201C	LEFT DOUBLE QUOTATION MARK
	150	96	U201D	RIGHT DOUBLE QUOTATION MARK
	151	97	U2014	EM DASH
	152	98	U2116	NUMERO SIGN
	153	99	U2122	TRADE MARK SIGN
	155	9B	U00BB	RIGHT-POINTING DOUBLE ANGLE QUOTATION MARK
	156	9C	U00AE	REGISTERED SIGN
	157	9D	U00AB	LEFT-POINTING DOUBLE ANGLE QUOTATION MARK
	159	9F	U00A4	CURRENCY SIGN
IE	164	A4	U0454	CYRILLIC SMALL LETTER UKRAINIAN
				UKR
	166	A6	U0456	CYRILLIC SMALL LETTER BELORUSSIAN-UKRAINIAN I UKR
	167	A7	U0457	CYRILLIC SMALL LETTER YI
(UKRAINIAN)				UKR
	173	AD	U0491	CYRILLIC SMALL LETTER UKRAINIAN GHE (WITH UPTURN) UKR
	174	AE	U045E	CYRILLIC SMALL LETTER BELORUSSIAN SHORT U BYEL
IE	180	B4	U0403	CYRILLIC CAPITAL LETTER UKRAINIAN
				UKR
	182	B6	U0406	CYRILLIC CAPITAL LETTER BELORUSSIAN-

183	B7	U0407	UKRAINIAN I CYRILLIC CAPITAL LETTER YI (UKRAINIAN)	UKR UKR
189	BD	U0490	CYRILLIC CAPITAL LETTER UKRAINIAN (WITH UPTURN)	GHE UKR

Expires in six months

[Page 6]

190	BE	U040E	CYRILLIC CAPITAL LETTER BELORUSSIAN SHORT U	BYEL
191	BF	U00A9	COPYRIGHT SIGN	

KOI8-RU compatible with ISO-IR-111 in all Russian, Ukrainian and Belorussian letters but differs in positions of one additional Ukrainian letter GHE WITH UPTURN, non-specified in ISO-IR-111 positions #128-#159 are used for FORMS elements from KOI8-R and other important characters which are currently missing from ISO-IR-111 and KOI8-R.

128	80	U2500	FORMS LIGHT HORIZONTAL
129	81	U2502	FORMS LIGHT VERTICAL
130	82	U250C	FORMS LIGHT DOWN AND RIGHT
131	83	U2510	FORMS LIGHT DOWN AND LEFT
132	84	U2514	FORMS LIGHT UP AND RIGHT
133	85	U2518	FORMS LIGHT UP AND LEFT
134	86	U251C	FORMS LIGHT VERTICAL AND RIGHT
135	87	U2524	FORMS LIGHT VERTICAL AND LEFT
136	88	U252C	FORMS LIGHT DOWN AND HORIZONTAL
137	89	U2534	FORMS LIGHT UP AND HORIZONTAL
138	8A	U253C	FORMS LIGHT VERTICAL AND HORIZONTAL
139	8B	U2580	UPPER HALF BLOCK
140	8C	U2584	LOWER HALF BLOCK
141	8D	U2588	FULL BLOCK
142	8E	U258C	LEFT HALF BLOCK
143	8F	U2590	RIGHT HALF BLOCK
144	90	U2591	LIGHT SHADE
145	91	U2592	MEDIUM SHADE
146	92	U2593	DARK SHADE
147	93	U201C	LEFT DOUBLE QUOTATION MARK
148	94	U25A0	BLACK SQUARE
149	95	U2219	BULLET OPERATOR
150	96	U201D	RIGHT DOUBLE QUOTATION MARK
151	97	U2014	EM DASH
152	98	U2116	NUMERO SIGN
153	99	U2122	TRADE MARK SIGN
154	9A	U00A0	NONBREAKING SPACE
155	9B	U00BB	RIGHT-POINTING DOUBLE ANGLE QUOTATION MARK
156	9C	U00AE	REGISTERED SIGN
157	9D	U00AB	LEFT-POINTING DOUBLE ANGLE QUOTATION MARK
158	9E	U00B7	MIDDLE DOT
159	9F	U00A4	CURRENCY SIGN
160	A0	U2550	FORMS DOUBLE HORIZONTAL
161	A1	U2551	FORMS DOUBLE VERTICAL

162	A2	U2552	FORMS DOWN SINGLE AND RIGHT DOUBLE
165	A5	U2554	FORMS DOUBLE DOWN AND RIGHT
168	A8	U2557	FORMS DOUBLE DOWN AND LEFT
169	A9	U2558	FORMS UP SINGLE AND RIGHT DOUBLE
170	AA	U2559	FORMS UP DOUBLE AND RIGHT SINGLE
171	AB	U255A	FORMS DOUBLE UP AND RIGHT
172	AC	U255B	FORMS UP SINGLE AND LEFT DOUBLE
173	AD	U0491	CYRILLIC SMALL LETTER UKRAINIAN GHE (WITH UPTURN) UKR
175	AF	U255E	FORMS VERTICAL SINGLE AND RIGHT DOUBLE

Expires in six months

[Page 8]

176	B0	U255F	FORMS VERTICAL DOUBLE AND RIGHT SINGLE
177	B1	U2560	FORMS DOUBLE VERTICAL AND RIGHT
178	B2	U2561	FORMS VERTICAL SINGLE AND LEFT DOUBLE
181	B5	U2563	FORMS DOUBLE VERTICAL AND LEFT
184	B8	U2566	FORMS DOUBLE DOWN AND HORIZONTAL
185	B9	U2567	FORMS UP SINGLE AND HORIZONTAL DOUBLE
186	BA	U2568	FORMS UP DOUBLE AND HORIZONTAL SINGLE
187	BB	U2569	FORMS DOUBLE UP AND HORIZONTAL
188	BC	U256A	FORMS VERTICAL SINGLE AND HORIZONTAL DOUBLE
189	BD	U0490	CYRILLIC CAPITAL LETTER UKRAINIAN GHE (WITH UPTURN) UKR
191	BF	U00A9	COPYRIGHT SIGN

Security Considerations

Security issues are not discussed in this memo.

Acknowledgments

The motivation to write this document came from the work with Microsoft International Team and personally from Igor Sinitsyn. Also this work is based on large work provided by Dmitry Kohmanyuk and Igor Sviridov in setting up standard-de-facto koi8-r for Ukrainian

Internet community. Further acknowledgments go to all Ukrainian netizens who participated in discussion of proposed draft in ukr.nodes

and ukr.netnews and personally Andrew Stesin from Global Ukraine.

References

- [1] Chernov, A., "Registration of a Cyrillic Character Set", [RFC 1589](#), Network Working Group, July 1993.
- [2] UNICODE 2.0 CHARACTER DATABASE. - <ftp://unicode.org/pub/2.0-Update/UnicodeData-2.0.14.txt>
- [3] Ukrainian letters in koi8-u and other character sets ftp://ftp.ua.net/pub/info/encodings/koi8-u/ukr_chars_in_koi8-u_and_others.txt, June 1995.
- [4] ECMA-CYRILLIC. - <ftp://dkuug.dk/i18n/charmmaps.all/ECMA-CYRILLIC>

Expires in six months

[Page 9]

Author's Address

Yuri V. Demchenko
Computer Aided Design Department
National Technical University of Ukraine
"Kiev Polytechnic Institute"
Prospect Pobedy 37
Kiev-56, 252056
Ukraine

Tel: +380 44 441 1465

Fax: +380 44 274 5932

E-mail: demch@cad.ntu-kpi.kiev.ua