

Локализация и интернационализация программного обеспечения

Инструментарий Современного Программиста

Иван Трепаков

NSU Sys.Pro

American Standard Code for Information Interchange

```
$ ascii -d
```

0 NUL	16 DLE	32	48 0	64 @	80 P	96 `	112 p
1 SOH	17 DC1	33 !	49 1	65 A	81 Q	97 a	113 q
2 STX	18 DC2	34 "	50 2	66 B	82 R	98 b	114 r
3 ETX	19 DC3	35 #	51 3	67 C	83 S	99 c	115 s
4 EOT	20 DC4	36 \$	52 4	68 D	84 T	100 d	116 t
5 ENQ	21 NAK	37 %	53 5	69 E	85 U	101 e	117 u
6 ACK	22 SYN	38 &	54 6	70 F	86 V	102 f	118 v
7 BEL	23 ETB	39 '	55 7	71 G	87 W	103 g	119 w
8 BS	24 CAN	40 (56 8	72 H	88 X	104 h	120 x
9 HT	25 EM	41)	57 9	73 I	89 Y	105 i	121 y
10 LF	26 SUB	42 *	58 :	74 J	90 Z	106 j	122 z
11 VT	27 ESC	43 +	59 ;	75 K	91 [107 k	123 {
12 FF	28 FS	44 ,	60 <	76 L	92 \	108 l	124
13 CR	29 GS	45 -	61 =	77 M	93]	109 m	125 }
14 SO	30 RS	46 .	62 >	78 N	94 ^	110 n	126 ~
15 SI	31 US	47 /	63 ?	79 O	95 _	111 o	127 DEL

ASCII

00:

10:

20: `!"#$%&'()*+,-./`

30: `0123456789:;<=>?`

40: `@ABCDEFGHIJKLMNO`

50: `PQRSTUVWXYZ[\]^_`

60: ``abcdefghijklmno`


70: `pqrstuvwxyz{|}~`

ASCII

00:	80:
10:	90:
20: <code>!"#\$%&'()*+,-./</code>	a0:
30: <code>0123456789:;<=>?</code>	b0:
40: <code>@ABCDEFGHIJKLMNO</code>	c0:
50: <code>PQRSTUVWXYZ[\]^_</code>	d0:
60: <code>`abcdefghijklmno</code>	e0:
70: <code>pqrstuvwxyz{ }~</code>	f0:

Расширения ASCII

ASCII

00: 80: АБВГДЕЖЗИЙКЛМНОП
10: 90: РСТУФХЦЧШЩЪЫЬЭЮЯ
20: ̀!"#\$%&'()*+,-./ а0: абвгдежзийклмноп
30: 0123456789:;<=>? б0:  | | | | | | | | | | | | | | | |
40: @ABCDEFGHIJKLMNO с0: | | | | | | | | | | | | | | | |
50: PQRSTUVWXYZ[\]^_ d0: | | | | | | | | | | | | | | | |
60: `abcdefghijklmno e0: рстуфхцчшщъыьэюя
70: pqrstuvwxyz{|}~ f0: ЁёЄєӢӓӮӹ••√№▪■

CP866

MS-DOS

Расширения ASCII

ASCII

00:
10:
20: _!"#\$%&'()*+,-./
30: 0123456789:;<=>?
40: @ABCDEFGHIJKLMNO
50: PQRSTUVWXYZ[\]^_
60: `abcdefghijklmnopqrstuvwxyz
70: pqrstuvwxyz{|}~

CP866

80: АБВГДЕЖЗИЙКЛМНОП
90: РСТУФХЦЧШЩЪЫЬЭЮЯ
a0: абвгдежзийклмноп
b0:
c0:
d0:
e0: рстуфхцчшщъыьэюя
f0: ЁёЄєӢӓӮӹ°•√№¤■

MS-DOS

KOI8-R

-| □ □ □ □ □ □ □
▒ ▒ ▒ ▒ ▒ ▒ ▒ ▒ ▒ ▒ ▒ ▒ ▒ ▒ ▒ ▒
=|| ёё г г г г г г г г
||| ЁЁ ||| ||| ||| ||| ||| |||
юабцдефгхийклмно
пярстужввызшэщчъ
ЮАБЦДЕФГХИЙКЛМНО
ПЯРСТУЖВВЫЗШЭЩЧЪ

Unix

Расширения ASCII

ASCII

CP866

KOI8-R

CP1251

00:	АБВГДЕЖЗИЙКЛМНОП
10:	РСТУФХЦЧШЩЪЫЬЭЮЯ
20:	а0: абвгдежзийклмноп
30:	б0:
40:	с0:
50:	д0:
60:	е0: рстуфхцчшщъыьэюя
70:	ф0: ЁёЄєİïŸŸ°•√№¤■

[illegible]

Ѓѓ,ѓ,„…†‡€%Љ<ЊЌЎ
ћ‘’“”•—™Ў>ЊЌћ
ЎЎЈѡГ|§Ё©€«¬-®Ї
°±Іігμ¶•ё№є»ЈЅѕі
АБВГДЕЖЗИЙКЛМНОП
РСТУФХЦЧШЩЪЫЬЭЮЯ
абвгдежзийклмно
рстуфхцчшщъыьэюя

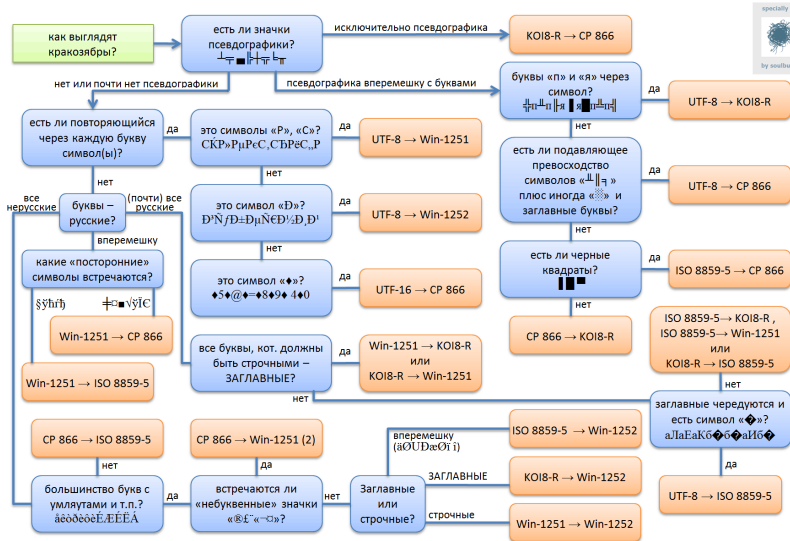
MS-DOS

Unix

Windows

опХБЕР ЛХП!

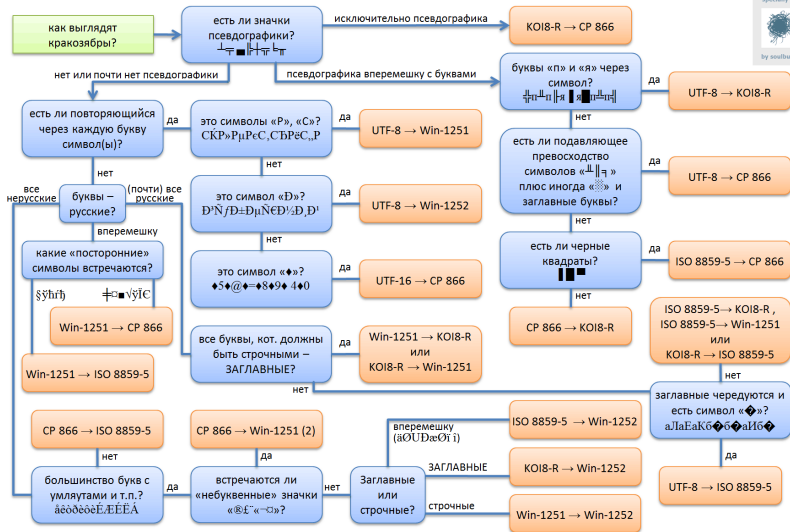
ОПХБЕР ЛХП!



Крокозябры

ОПХБЕР ЛХП!

```
$ echo "оПХБЕР ЛХП!" \
| iconv -t cp1251 \
| iconv -f koi8-r
Нояаен кyo!
```

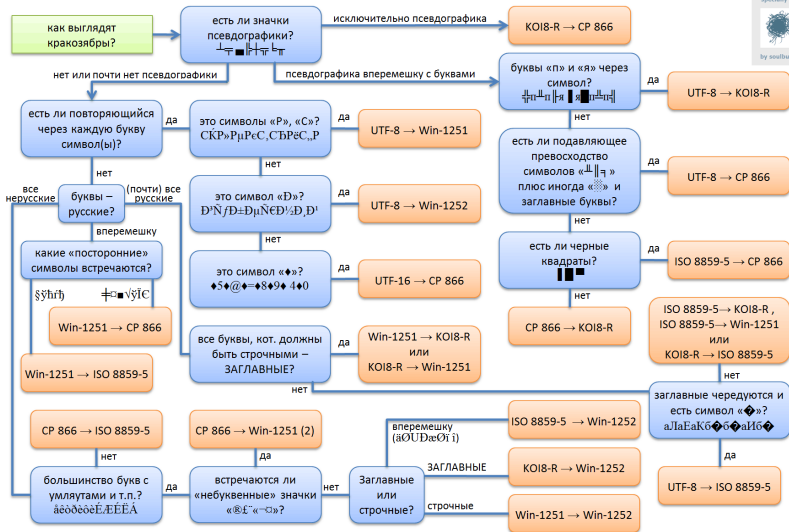


Крокозябры

оПХБЕР ЛХП!

```
$ echo "оПХБЕР ЛХП!" \  
| iconv -t cp1251 \  
| iconv -f koi8-r  
Ноуаеп кyo!
```

```
$ echo "оПХБЕР ЛХП!" \  
| iconv -t koi8-r \  
| iconv -f cp1251  
Привет мир!
```



Unicode Standard

- Универсальное кодирование символов
- Количество различных кодов 1 114 112
- **Unicode 16.0** (2024 год) использует лишь 154 998

Цели

- **Универсальность** Содержит все возможные символы современных и древних языков, технических текстов, диакритику и emoji
- **Эффективность** *Plain text* кодирование в одном из трех стандартных форматов: **UTF-32**, **UTF-16**, **UTF-8**
- **Однозначность** Каждый код *однозначно* соответствует единственному символу



Q&A