

C16/+4/116/64 CHARACTER GENERATOR AND UNICODE

by Vladimir Lidovski

The character generator of the C16/116/+4 consists of 256 symbols, but only 153 are unique — 90 symbols are duplicated, two are triplicated and three even quadruplicated. 124 characters of the 153 are mapped to Unicode, the remaining 29 however are UNMAPPED! This may produce impossibility of using our well known Commodore symbols with modern and FUTURE computer systems. This problem has naturally appeared in the situation of conversion of Commodore texts to the other systems — some symbols become LOST. Without unicodes they will be lost FOREVER...

So we need to register them at the Unicode Consortium (www.unicode.org). The Unicode structure is not perfect. It's missed some obvious characters. They were sacrificed to this structure. For example, the chapter "Form and charts components" contains places only for 128 symbols — all these places are occupied, the chapter "Block elements" has 32 places and all are occupied, etc.

Eleven (73–75, 77, 78, 85, 86, 103, 118, 119, 120) of the 29 unregistered characters can simply get Unicode names. They are probably the victims of the mentioned structure. Their obvious names (prefixed by question mark) are presented in the next big chart.

Two characters (223, 233) have pictures very similar to some presented in Unicode. Their registration may not be necessary.

The remaining 16 (68–72, 76, 79, 80, 82, 84, 89, 92, 102, 104, 122, 222) require NAMES acceptable to ALL.

23 of the unmapped characters (68–80, 82, 84–86, 89, 103, 118–120, 122) can be placed in the extension to the above mentioned chapter "Form and charts components".


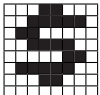

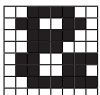

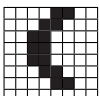

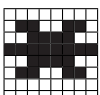


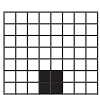
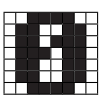
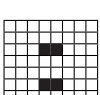

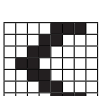

The location for the other six (92, 102, 104, 222, 223, 233) is more difficult to determine. It can be in the extensions of "Geometric shapes", or "Miscellaneous Symbols", or some other place.

The character ROM of C64 consists of 512 characters — 256 are the same as in C16/116/+4 and the other 256 are their inversion. The C64 has two character generators. The first 128 symbols parts of each are equal to first and second parts in the following big chart. The second 128 symbols parts contain inverted symbols of the first parts.

Commodore VIC20 character ROM structure is the same as C64, but this ROM contains THIN (light) characters which box drawings part is better represented in Unicode. This leaves only 21 unmapped characters.

CURRENT UNICODE MAP OF G₁₆/+4/116 CHARACTERS

0 0040		COMMERCIAL AT	1 0041		LATIN CAPITAL LETTER A
2 0042		LATIN CAPITAL LETTER B	3 0043		LATIN CAPITAL LETTER C
4 0044		LATIN CAPITAL LETTER D	5 0045		LATIN CAPITAL LETTER E
6 0046		LATIN CAPITAL LETTER F	7 0047		LATIN CAPITAL LETTER G
8 0048		LATIN CAPITAL LETTER H	9 0049		LATIN CAPITAL LETTER I
10 004A		LATIN CAPITAL LETTER J	11 004B		LATIN CAPITAL LETTER K
12 004C		LATIN CAPITAL LETTER L	13 004D		LATIN CAPITAL LETTER M
14 004E		LATIN CAPITAL LETTER N	15 004F		LATIN CAPITAL LETTER O
16 0050		LATIN CAPITAL LETTER P	17 0051		LATIN CAPITAL LETTER Q
18 0052		LATIN CAPITAL LETTER R	19 0053		LATIN CAPITAL LETTER S
20 0054		LATIN CAPITAL LETTER T	21 0055		LATIN CAPITAL LETTER U
22 0056		LATIN CAPITAL LETTER V	23 0057		LATIN CAPITAL LETTER W
24 0058		LATIN CAPITAL LETTER X	25 0059		LATIN CAPITAL LETTER Y
26 005A		LATIN CAPITAL LETTER Z	27 005B		LEFT SQUARE BRACKET
28 00A3		POUND SIGN	29 005D		RIGHT SQUARE BRACKET
30 2191		UPWARDS ARROW	31 2190		LEFTWARDS ARROW

32		SPACE	0020	33		EXCLAMATION MARK	0021
34		QUOTATION MARK	0022	35		NUMBER SIGN	0023
36		DOLLAR SIGN	0024	37		PERCENT SIGN	0025
38		AMPERSAND	0026	39		APOSTROPHE	0027
40		LEFT PARENTHESIS	0028	41		RIGHT PARENTHESIS	0029
42		ASTERISK	002A	43		PLUS SIGN	002B
44		COMMA	002C	45		HYPHEN-MINUS	002D
46		FULL STOP	002E	47		SOLIDUS	002F
48		DIGIT ZERO	0030	49		DIGIT ONE	0031
50		DIGIT TWO	0032	51		DIGIT THREE	0033
52		DIGIT FOUR	0034	53		DIGIT FIVE	0035
54		DIGIT SIX	0036	55		DIGIT SEVEN	0037
56		DIGIT EIGHT	0038	57		DIGIT NINE	0039
58		COLON	003A	59		SEMICOLON	003B
60		LESS-THAN SIGN	003C	61		EQUALS SIGN	003D
62		GREATER-THAN SIGN	003E	63		QUESTION MARK	003F

64 BOX DRAWINGS HEAVY HORIZONTAL
2501

66 BOX DRAWINGS HEAVY VERTICAL
2503

68 ?
NO!

70 ?
NO!

72 ?
NO!

74 ? BOX DRAWINGS HEAVY ARC UP AND RIGHT (similar to 2570)
NO!

76 ?
NO!

78 ? BOX DRAWINGS HEAVY DIAGONAL UPPER RIGHT TO LOWER LEFT (similar to 2571)
NO!

80 ?
NO!

82 ?
NO!

84 ?
NO!

86 ? BOX DRAWINGS HEAVY DIAGONAL CROSS (similar to 2573)
NO!

88 BLACK CLUB SUIT
2663

90 BLACK DIAMOND SUIT
2666

92 ?
NO!

94 GREEK SMALL LETTER PI
03C0

65 BLACK SPADE SUIT
2660

67 BOX DRAWINGS HEAVY HORIZONTAL
2501

69 ?
NO!

71 ?
NO!

73 ? BOX DRAWINGS HEAVY ARC DOWN AND LEFT (similar to 256E)
NO!

75 ? BOX DRAWINGS HEAVY ARC UP AND LEFT (similar to 256F)
NO!

77 ? BOX DRAWINGS HEAVY DIAGONAL UPPER LEFT TO LOWER RIGHT (similar to 2572)
NO!

79 ?
NO!

81 BLACK CIRCLE
25CF

83 BLACK HEART SUIT
2665

85 ? BOX DRAWINGS HEAVY ARC UP AND LEFT (similar to 256D)
NO!

87 WHITE CIRCLE
25CB

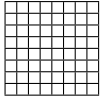
89 ?
NO!

91 BOX DRAWINGS HEAVY VERTICAL AND HORIZONTAL
254B

93 BOX DRAWINGS HEAVY VERTICAL
2503

95 BLACK UPPER RIGHT TRIANGLE
25E5

96
0020



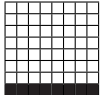
SPACE

98
2584



LOWER HALF BLOCK

100
2581



LOWER ONE EIGHTH BLOCK

102
NO!



?

104
NO!



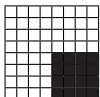
?

106
NO!



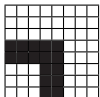
? RIGHT ONE QUARTER BLOCK

108
2597



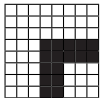
QUADRANT LOWER RIGHT

110
2513



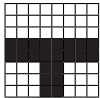
BOX DRAWINGS HEAVY DOWN AND LEFT

112
250F



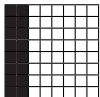
BOX DRAWINGS HEAVY DOWN AND RIGHT

114
2533



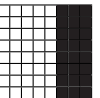
BOX DRAWINGS HEAVY DOWN AND HORIZONTAL

116
258E



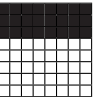
LEFT ONE QUARTER BLOCK

118
NO!



? RIGHT THREE EIGHTHS BLOCK

120
NO!



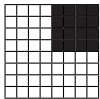
? UPPER THREE EIGHTHS BLOCK

122
NO!



?

124
259D



QUADRANT UPPER RIGHT

126
2598




QUADRANT UPPER LEFT

97
258C



LEFT HALF BLOCK

99
2594



UPPER ONE EIGHTH BLOCK

101
258E




LEFT ONE QUARTER BLOCK

103
NO!




? RIGHT ONE QUARTER BLOCK

105
25E4




BLACK UPPER LEFT TRIANGLE

107
2523




BOX DRAWINGS HEAVY VERTICAL AND RIGHT

109
2517




BOX DRAWINGS HEAVY UP AND RIGHT

111
2582




LOWER ONE QUARTER BLOCK

113
253B




BOX DRAWINGS HEAVY UP AND HORIZONTAL

115
252B




BOX DRAWINGS HEAVY VERTICAL AND LEFT

117
258D




LEFT THREE EIGHTHS BLOCK

119
NO!



? UPPER ONE QUARTER BLOCK

121
2583




LOWER THREE EIGHTHS BLOCK

123
2596




QUADRANT LOWER LEFT

125
251B



BOX DRAWINGS HEAVY UP AND LEFT

127
259A



QUADRANT UPPER LEFT AND LOWER RIGHT

128
0040  COMMERCIAL AT

130
0062  LATIN SMALL LETTER B

132
0064  LATIN SMALL LETTER D

134
0066  LATIN SMALL LETTER F

136
0068  LATIN SMALL LETTER H

138
006A  LATIN SMALL LETTER J

140
006C  LATIN SMALL LETTER L

142
006E  LATIN SMALL LETTER N

144
0070  LATIN SMALL LETTER P


146
0072  LATIN SMALL LETTER R

148
0074  LATIN SMALL LETTER T

150
0076  LATIN SMALL LETTER V

152
0078  LATIN SMALL LETTER X

154
007A  LATIN SMALL LETTER Z

156
00A3  POUND SIGN

158
2191  UPWARDS ARROW

129
0061  LATIN SMALL LETTER A

131
0063  LATIN SMALL LETTER C

133
0065  LATIN SMALL LETTER E

135
0067  LATIN SMALL LETTER G

137
0069  LATIN SMALL LETTER I

139
006B  LATIN SMALL LETTER K

141
006D  LATIN SMALL LETTER M

143
006F  LATIN SMALL LETTER O

145
0071  LATIN SMALL LETTER Q

147
0073  LATIN SMALL LETTER S

149
0075  LATIN SMALL LETTER U

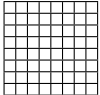


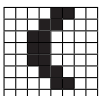

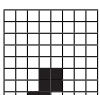




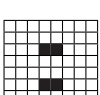
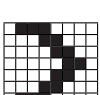

151
0077  LATIN SMALL LETTER W

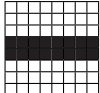
153
0079  LATIN SMALL LETTER Y

155
005B  LEFT SQUARE BRACKET

157
005D  RIGHT SQUARE BRACKET

159
2190  LEFTWARDS ARROW

160 0020		SPACE	161 0021		EXCLAMATION MARK
162 0022		QUOTATION MARK	163 0023		NUMBER SIGN
164 0024		DOLLAR SIGN	165 0025		PERCENT SIGN
166 0026		AMPERSAND	167 0027		APOSTROPHE
168 0028		LEFT PARENTHESIS	169 0029		RIGHT PARENTHESIS
170 002A		ASTERISK	171 002B		PLUS SIGN
172 002C		COMMA	173 002D		HYPHEN-MINUS
174 002E		FULL STOP	175 002F		SOLIDUS
176 0030		DIGIT ZERO	177 0031		DIGIT ONE
178 0032		DIGIT TWO	179 0033		DIGIT THREE
180 0034		DIGIT FOUR	181 0035		DIGIT FIVE
182 0036		DIGIT SIX	183 0037		DIGIT SEVEN
184 0038		DIGIT EIGHT	185 0039		DIGIT NINE
186 003A		COLON	187 003B		SEMICOLON
188 003C		LESS-THAN SIGN	189 003D		EQUALS SIGN
190 003E		GREATER-THAN SIGN	191 003F		QUESTION MARK

192  BOX DRAWINGS HEAVY HORIZONTAL
2501

194  LATIN CAPITAL LETTER B
0042

196  LATIN CAPITAL LETTER D
0044

198  LATIN CAPITAL LETTER F
0046

200  LATIN CAPITAL LETTER H
0048

202  LATIN CAPITAL LETTER J
004A

204  LATIN CAPITAL LETTER L
004C

206  LATIN CAPITAL LETTER N
004E

208  LATIN CAPITAL LETTER P
0050

210  LATIN CAPITAL LETTER R
0052

212  LATIN CAPITAL LETTER T
0054

214  LATIN CAPITAL LETTER V
0056

216  LATIN CAPITAL LETTER X
0058

218  LATIN CAPITAL LETTER Z
005A

220  ?
NO!

222  ?
NO!

193  LATIN CAPITAL LETTER A
0041

195  LATIN CAPITAL LETTER C
0043

197  LATIN CAPITAL LETTER E
0045

199  LATIN CAPITAL LETTER G
0047

201  LATIN CAPITAL LETTER I
0049

203  LATIN CAPITAL LETTER K
004B

205  LATIN CAPITAL LETTER M
004D

207  LATIN CAPITAL LETTER O
004F

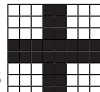
209  LATIN CAPITAL LETTER Q
0051

211  LATIN CAPITAL LETTER S
0053

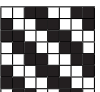
213  LATIN CAPITAL LETTER U
0055

215  LATIN CAPITAL LETTER W
0057

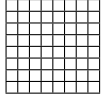
217  LATIN CAPITAL LETTER Y
0059

219  BOX DRAWINGS HEAVY VERTICAL AND HORIZONTAL
254B

221  BOX DRAWINGS HEAVY VERTICAL
2503

223  ? (similar to 25A7)
NO!

224
0020



SPACE

226
2584



LOWER HALF BLOCK

228
2581



LOWER ONE EIGHTH BLOCK

230
NO!



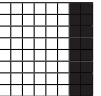
?

232
NO!



?

234
NO!



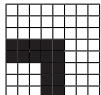
? RIGHT ONE QUARTER BLOCK

236
2597



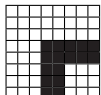
QUADRANT LOWER RIGHT

238
2513



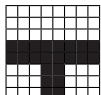
BOX DRAWINGS HEAVY DOWN AND LEFT

240
250F



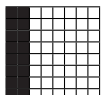
BOX DRAWINGS HEAVY DOWN AND RIGHT

242
2533



BOX DRAWINGS HEAVY DOWN AND HORIZONTAL

244
258E



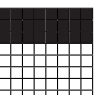
LEFT ONE QUARTER BLOCK

246
NO!



? RIGHT THREE EIGHTHS BLOCK

248
NO!



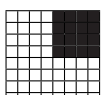
? UPPER THREE EIGHTH BLOCK

250
221A



SQUARE ROOT

252
259D



QUADRANT UPPER RIGHT

254
2598




QUADRANT UPPER LEFT

225
258C



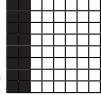
LEFT HALF BLOCK

227
2594



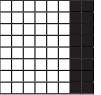
UPPER ONE EIGHTH BLOCK

229
258E



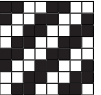
LEFT ONE QUARTER BLOCK

231
NO!



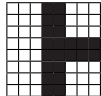
? RIGHT ONE QUARTER BLOCK

233
NO!



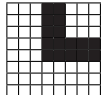
? (similar to 25A8)

235
2523



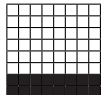
BOX DRAWINGS HEAVY VERTICAL AND RIGHT

237
2517



BOX DRAWINGS HEAVY UP AND RIGHT

239
2582



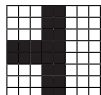
LOWER ONE QUARTER BLOCK

241
253B



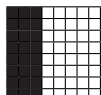
BOX DRAWINGS HEAVY UP AND HORIZONTAL

243
252B



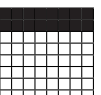
BOX DRAWINGS HEAVY VERTICAL AND LEFT

245
258D



LEFT THREE EIGHTHS BLOCK

247
NO!



? UPPER ONE QUARTER BLOCK

249
2583




LOWER THREE EIGHTHS BLOCK

251
2596



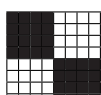
QUADRANT LOWER LEFT

253
251B



BOX DRAWINGS HEAVY UP AND LEFT

255
259A



QUADRANT UPPER LEFT AND LOWER RIGHT

ALL G=16/+4/116 UNICODES

0020	32 96 160 224	0021	33 161	0022	34 162
0023	35 163	0024	36 164	0025	37 165
0026	38 166	0027	39 167	0028	40 168
0029	41 169	002A	42 170	002B	43 171
002C	44 172	002D	45 173	002E	46 174
002F	47 175	0030	48 176	0031	49 177
0032	50 178	0033	51 179	0034	52 180
0035	53 181	0036	54 182	0037	55 183
0038	56 184	0039	57 185	003A	58 186
003B	59 187	003C	60 188	003D	61 189
003E	62 190	003F	63 191	0040	0 128
0041	1 193	0042	2 194	0043	3 195
0044	4 196	0045	5 197	0046	6 198
0047	7 199	0048	8 200	0049	9 201
004A	10 202	004B	11 203	004C	12 204
004D	13 205	004E	14 206	004F	15 207
0050	16 208	0051	17 209	0052	18 210
0053	19 211	0054	20 212	0055	21 213
0056	22 214	0057	23 215	0058	24 216
0059	25 217	005A	26 218	005B	27 155
005D	29 157	0061	129	0062	130
0063	131	0064	132	0065	133
0066	134	0067	135	0068	136
0069	137	006A	138	006B	139
006C	140	006D	141	006E	142
006F	143	0070	144	0071	145
0072	146	0073	147	0074	148
0075	149	0076	150	0077	151
0078	152	0079	153	007A	154
00A3	28 156	03C0	94	2190	31 159
2191	30 158	221A	250	2501	64 67 192
2503	66 93 221	250F	112 240	2513	110 238
2517	109 237	251B	125 253	2523	107 235
252B	115 243	2533	114 242	253B	113 241
254B	91 219	2581	100 228	2582	111 239
2583	121 249	2584	98 226	258C	97 225
258D	117 245	258E	101 116 229 244	2594	99 227
2596	123 251	2597	108 236	2598	126 254
259A	127 255	259D	124 252	25CB	87
25CF	81	25E4	105	25E5	95
2660	65	2663	88	2665	83
2666	90	NO!	102 230	NO!*	103 106 231 234
NO!	104 232	NO!*	118 246	NO!*	119 247
NO!*	120 248	NO!	122	NO!	222
NO!**	223	NO!**	233	NO!	68
NO!	69	NO!	70	NO!	71
NO!	72	NO!*	73	NO!*	74
NO!*	75	NO!	76	NO!*	77
NO!*	78	NO!	79	NO!	80
NO!	82	NO!	84	NO!*	85
NO!*	86	NO!	89	NO!	92 220

The asterisk (*) marks 11 symbols which MUST really be in Unicode even without influence of Commodore symbol table. Two asterisks (**) mark two symbols which have very similar matches in Unicode.

Edited by Csabo of LOD