

Inventio Software

Specialist Instrumentation Design and Programming

Encoding ASCII into cf2022 colorForth tokens

2022 Apr 21 b

```
; ASCII / UTF8 support. If the first Shannon-Fano encoded letter is a 4 bit NULL,  
; display the next 24 bits as three ASCII characters.  
; $03e3c009 is displayed as 'x'
```

```
Sandbox >< any old ASCII <=> 510  
test your code here ;  
  
SC† yrg*  
|<f>|j <↑↓>  
ab -mc+  
x. i
```

```
82b60c6a 0008f800 jg7 Sandb  
3ca00000 0008f804 ox  
03e3c009 0008f808 ><  
56980009 0008f80c m any  
3a600009 0008f810 m old  
5849dc0b 0008f814 y IX ASCII  
03c3d3e9 0008f818 ) <=>  
24810003 0008f81c ) test  
99e61004 0008f820 )↑ your  
91e04004 0008f824 )A code  
c8828004 0008f828 ) code  
f0000004 0008f82c ) here  
00000000 0008f830 ;  
00000000 0008f834  
00000000 0008f838  
00000000 0008f83c  
  
jg7 <m@>)m Vm <:y IX >)) □$)↑ )A )□ )((  
  
pyfi gcr l  
aoeu htns  
qkxd bnwv  
dump 9*
```

Inventio Software

Specialist Instrumentation Design and Programming

Appendix A cf2022.nasm code

```
lowercase:      ; display a white text word in normal lower-case letters
               call white
showSF_EDI_:   ; ( -- ) \ display a Shanon-Fano encoded token pointed to by edi in the current colour
               _DUP_
               mov _TOS_, [ ( edi * 4 ) - 0x04 ] ; fetch the next token - drops through to showShannonFano

showShannonFano: ; ( token -- ) \ display the Shannon-Fano encoded token on TOS
               ; ASCII / UTF8 support. If the first Shannon-Fano encoded letter is a 4 bit NULL,
               ; display the next 24 bits as three ASCII characters.
               mov _SCRATCH_, _TOS_ ; save the token value
               and _SCRATCH_, 0xF0000000
               cmp _SCRATCH_, 0x00000000
               jnz .forward
               ; display as three ASCII characters
               mov _SCRATCH_, _TOS_

               mov _TOS_, _SCRATCH_
               shr _TOS_, 20
               and _TOS_, 0x000000FF
               jz .null_terminator
               _DUP_
               call emit_

               mov _TOS_, _SCRATCH_
               shr _TOS_, 12
               and _TOS_, 0x000000FF
               jz .null_terminator
               _DUP_
               call emit_

               mov _TOS_, _SCRATCH_
               shr _TOS_, 4
               and _TOS_, 0x000000FF
               jz .null_terminator
               _DUP_
               call emit_

               ; arrive here if an ASCII character is an ASCII NULL, or if all three have been emitted
               .null_terminator:
               call space_ ; display a space character at the end of the word
               _DROP_
               ret

.forward:

               ; display as Shannon-Fano encoded token name
               and _TOS_, byte -0x10 ; and _TOS_, 0xFFFFFFFF0 ignore token colour when displaying the letters

lowercasePrimitive: ; ( token -- ) \ display the given Shanon-Fano encoded word in the current colour
               call unpack
               jz lowercasePrimitiveEnd
               call emitSF_
               jmp lowercasePrimitive
lowercasePrimitiveEnd:
               call space_
               _DROP_
               _DROP_
               ret
```