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 ‘><’





# 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\_, 0xFFFFFFF0 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