

Encoding ASCII into cf2022 colorForth tokens

EuroForth 2022 Sep 18

Cakes and Biscuits

Flour plus :

Name	Fat	Egg(s)	Raising Agent	Topping
Pancake	0.5	1	0	Maple Syrup
Biscuit	0.5	0	Soda	Chocolate
Cake	1	2	Soda	Icing
Bread	0	0	Yeast	Butter
Jaffa Cake	0.25	0.5	Soda	Chocolate

Encoding ASCII into cf2022 colorForth tokens

EuroForth 2022 Sep 18

Cakes and Biscuits

Computer plus :

Name	File	ASCII, Unicode	Raising Agent	Topping
C Source File	1	1, UTF-8	Pre-processor	Libraries
Word Docx	1	1, *	Templates	Styles
Forth Source	1	1, UTF-8	Forth	Source Database
PDF file	1	1, UTF-8	Forth-like	Zip, Signing
colorForth	0.01	0.01, UTF-8	Forth	Chocolate

Encoding ASCII into cf2022 colorForth tokens

EuroForth 2022 Sep 18

Layers



Choose your ingredients, add toppings.

Encoding ASCII into cf2022 colorForth tokens

EuroForth 2022 Sep 18

Layers

Win32Forth	Native Forth	colorForth	JaffaCake Forth
Application	Application	Application	Application
File System	File System	Blocks	Blocks
OS	OS	-	OS/VM
Hardware	Hardware	Hardware	Hardware

Encoding ASCII into cf2022 colorForth tokens

EuroForth 2022 Sep 18

```
; 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 '<<'
```

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

Encoding ASCII into cf2022 colorForth tokens

EuroForth 2022 Sep 18

```

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

```

Encoding ASCII into cf2022 colorForth tokens

EuroForth 2022 Sep 18

```

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

call showShannonFano

.forward:

; display as Shannon-Fano encoded token name
and _TOS_, byte -0x10 ; and _TOS_, 0xFFFFFFFF 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

```

Encoding ASCII into cf2022 colorForth tokens

EuroForth 2022 Sep 18

This simple change to the colorForth pre-parsed source format, and the corresponding editor display words, allows the colorForth system to continue to use 32 bit „colored“ tokens, and also support ASCII / UTF-8 characters.

ToDo: Add keyboard support to type ASCII / UTF-8 characters !

Encoding ASCII into cf2022 colorForth tokens

EuroForth 2022 Sep 18