Minimum Word Set and Comparison

Paul Bennett, HIDECS

Some discussion has been bubbling in the background since before last year's EuroForth. The problem is seen that, for deeply embedded controls, some words in the CORE are not suitable candidates for inclusion in the system.

If we are to allow a claim of compliance to the standard for such deeply embedded systems we may need to look at a sensible BASE-LINE set of words that we should be able to count on to be there for a system that may connect to the embedded controller via a communications port only, without being over-burdened by words that are only sensible in the desk-top realm.

Hence, trying to establish what words might be included in the Baseline set began via email discussions between a few of us who are more into the small controller markets. This can include 8-bit, 16-bit, 18-bit, 21-bit, and 32-bit processors.

So, the set suggested below is probably richer than an absolute minimalist set would be (that extreme is saved for the Umbilical Only System and may be as little as XC@, XC! and XCALL).

The following 71 words are suggested at this stage. Words for inclusion:

Memory Access:-

HERE C@ C! @ , C, ALIGN ALIGNED CELL

CELLS CALIGN CALIGNED CHAR+ CHARS ! ' [']

Arithmetic:-

+ - 2* * / 2/ MOD */MOD UM*

Logic:

0= 0< 0<> = < >

AND OR XOR LSHIFT RSHIFT

Stack:

DUP SWAP OVER DROP >R R> R@ ROT

Flow Control:

IF ELSE THEN BEGIN WHILE UNTIL DO

I J LOOP EXECUTE

Defining:

: ; CREATE DOES> CONSTANT VARIABLE

I/O:

KEY KEY? CR EMIT

Tools: (\ .S

Exception Management: CATCH THROW

Special Words: TBD.

Some special words will have to be added for controllers and FLASH and IO handling.

About The Exception Management Choice:
It has been felt that the Catch & Throw mechanism is the most flexible way of managing exceptions that may arise for different reasons.
The standard gives a range of values pre-defined for specific error types. Minimal action would be the actions of ABORT (-1 or any non-zero code), see Standard reference 6.1.0680.

Some final Words:

It is seen that this smaller set of words, defined as a Baseline Set, would allow simpler starter systems for those who may be confused by the rich dictionary of some of the offerings.
In the deeply embedded controller world it would also assist with preferring a fully certified product for such deeply embedded controllers, while also being able to claim full adherence to a standard.

The observant will realise that IMMEDIATE is not in the list. The need for this rather depends on implementation details and whether or not the IMMEDIATE flag is required in the word's header or not.
Other issues of STATE also come into this category and there have been lengthy discussions on both factors.
Please start a new thread on these specific topics rather than discuss those issues in this one.

Minimum Word Set Comparison									
Type		**Forth 2012 reference**	**Ting**	**Brinkhoff**	**Plichota**	**B16**	**B16 Small**	**GA-F18**	**Proposed**
Mem	C@	6.1.0870	1	1	1	0	1	0	1
	C!	6.1.0850	1	1	1	0	0	0	1
	@	6.1.0650	1	1	1	0	1	1	1
	!	6.1.0010	1	1	1	0	0	1	1
Math	+	6.1.0120	1	1	1	1	1	1	1
Flow	IF	6.1.1700	1	1	1	0	0	1	0
	XOR	6.1.2490	1	0	1	1	1	0	1
Stack	DUP	6.1.1290	1	0	1	1	1	1	1

	SWAP	6.1.2260	1	0	1	0	0	0	1
	DROP	6.1.1260	1	0	1	1	1	1	1
	>R	6.1.0580	1	0	1	1	1	1	1
	R>	6.1.2060	1	0	1	1	1	1	1
Defining	:	6.1.0450	1	0	1	0	0	0	1
	;	6.1.0450	1	0	1	0	0	1	1
	*	6.1.0090	1	0	1	0	0	0	0
	DO	6.1.1240	1	0	1	0	0	0	0
	EXIT	6.1.1380	1	1	0	0	0	0	1
I/O	IN		1	1	0	0	0	1	0
	OUT		1	1	0	0	0	0	0
Other	-	6.1.0160	1	0	0	0	0	1	1
	AND	6.1.0720	1	0	0	1	1	1	1
	OR	6.1.1980	1	0	0	1	1	1	1
	OVER	6.1.1990	1	0	0	1	1	1	1
	R@	6.1.2070	1	0	0	1	0	0	1
	BEGIN	6.1.0760	1	0	0	0	0	0	1
	AGAIN	6.2.0700	1	0	0	0	0	0	1
	EXECUTE	6.1.1370	1	0	0	1	1	1	1
	CONSTANT	6.1.0950	1	0	0	0	0	0	1
	VARIABLE	6.1.2410	1	0	0	0	0	0	1
	/	6.1.0230	1	0	0	0	0	0	0
	MOD	6.1.1890	1	0	0	0	0	0	0
	INVERT	6.1.1720	1	0	0	0	0	0	0
	ELSE	6.1.01310	1	0	0	0	0	0	0
	THEN	6.1.2270	1	0	0	0	0	0	0
	WHILE	6.1.2430	1	0	0	0	0	0	0
	REPEAT	6.1.2120	1	0	0	0	0	0	0
	LOOP	6.1.1800	1	0	0	0	0	0	0
Logic	NAND		0	1	1	0	0	0	0
	LIT		0	0	1	1	1	0	0
Control	0BRANCH		0	1	0	0	0	0	1
other	DODOES		0	1	0	0	0	0	1
	2/	6.1.0330	0	0	0	0	1	1	1
	UM*	6.1.2360	0	0	0	0	0	0	1
	KEY	6.1.0750	0	0	0	0	0	0	1
	KEY?	10.6.1.1755	0	0	0	0	0	0	1
	EMIT	6.1.1320	0	0	0	0	0	0	1
	CREATE	6.1.1000	0	0	0	0	0	0	1
	DOES>	6.1.1250	0	0	0	0	0	0	1
	,	6.1.0150	0	0	0	0	0	0	1

	C,	6.1.0860	0	0	0	0	0	0	1
	.S	15.6.1.0220	0	0	0	0	0	0	1
	2*	6.1.0320	0	0	0	0	0	1	0
	C@+		0	0	0	0	1	0	0
	C@.		0	0	0	0	1	0	0
	C!+		0	0	0	0	1	0	0
	C!.		0	0	0	0	1	0	0
	@.		0	0	0	0	1	0	0
	a		0	0	0	1	0	1	0
	@a		0	0	0	0	0	1	0
	A@		0	0	0	1	0	0	0
	A@+		0	0	0	1	0	0	0
	@p		0	0	0	0	0	1	0
	@b		0	0	0	0	0	1	0
	@+		0	0	0	0	1	1	0
	!p		0	0	0	0	0	1	0
	!+		0	0	0	0	1	1	0
	!.		0	0	0	0	1	0	0
	!b		0	0	0	0	0	1	0
	a!		0	0	0	1	0	1	0
	A!+		0	0	0	1	0	0	0
	b!		0	0	0	0	0	1	0
	AC!		0	0	0	1	0	0	0
	AC@+		0	0	0	1	0	0	0
	AC!+		0	0	0	1	0	0	0
	AC@+		0	0	0	1	0	0	0
	>A		0	0	0	1	0	0	0
	+C		0	0	0	0	1	0	0
	C2/		0	0	0	0	1	0	0
	+*		0	0	0	0	0	1	0
	*+		0	0	0	1	0	0	0
	+C		0	0	0	1	0	0	0
	/-		0	0	0	1	0	0	0
	R@+		0	0	0	1	0	0	0
	RC@		0	0	0	1	0	0	0
	RC@+		0	0	0	1	0	0	0
	NIP	6.2.1930	0	0	0	1	1	0	0
	-IF		0	0	0	0	0	1	0
	<name> ;		0	0	0	0	0	1	0
	<name>		0	0	0	0	0	1	0
	unext		0	0	0	0	0	1	0
	next		0	0	0	0	0	1	0

	RET		0	0	0	1	1	0	0
	JZ		0	0	0	1	1	0	0
	JNZ		0	0	0	1	1	0	0
	JC		0	0	0	1	1	0	0
	JNC		0	0	0	1	1	0	0
	GZ		0	0	0	1	1	0	0
	GNZ		0	0	0	1	1	0	0
	GC		0	0	0	1	1	0	0
	GNC		0	0	0	1	1	0	0
control	GOTO		0	0	0	1	1	0	0
	LITC		0	0	0	1	1	0	0
	data		0	0	0	0	0	1	0
	---u		0	0	0	0	0	1	0
	--l-		0	0	0	0	0	1	0
	--lu		0	0	0	0	0	1	0
	-d--		0	0	0	0	0	1	0
	-d-u		0	0	0	0	0	1	0
	-dl-		0	0	0	0	0	1	0
	-dlu		0	0	0	0	0	1	0
	r---		0	0	0	0	0	1	0
	r-l-		0	0	0	0	0	1	0
	r-lu		0	0	0	0	0	1	0
	rd--		0	0	0	0	0	1	0
	rdl-		0	0	0	0	0	1	0
	rdlu		0	0	0	0	0	1	0
x	PAUSE		0	0	0	0	0	0	1
	CALL		0	0	0	1	1	0	0
	JMP		0	0	0	1	1	0	0
	RET		0	0	0	1	1	0	0
	NOP		0	0	0	1	1	0	0
	COM		0	0	0	0	1	0	0
Total Base Word Set			**37**	**12**	**18**	**40**	**36**	**47**	**37**