

J.J. Hoekstra

doc-version 001

© 2024

# wabiPi4 programmers-guide

for wabiPi4v2.12.1



**WARNING**  
a Raspberry running wabiPi4  
must be cooled adequately

## Copyright and license

(c) 2024 J.J. Hoekstra.

---

This **PROGRAMMERS-GUIDE** is licensed under the Creative Commons Attribution-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nd/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.



---

**WabiPi4** itself is copyright 2016–2024 (and all years in between) by J.J. Hoekstra.

WabiPi4 is licensed for free use under AGPL-3.0. See the source-code of wabiPi4 for the complete license text.

---

The transcendental floating point functions in wabiPi4 are based on (parts of) the NE10 library from ARM limited.

The NE10 library is licensed under the BSD-3 clause license. See the source-code of wabiPi4 for the complete license text.

---

Brands and names mentioned herein are the trademarks of their respective owners.

## Foreword

In 1983 I got my first home-computer, and I was blown away by it. No matter that it was slow, had limited memory and was riddled with bugs. What fun I had!

I still fondly remember the long nights – with an old television for a monitor; programming, debugging, playing – all the while struggling with a crappy cassette-recorder interface.

At that time the fastest, baddest, 'bestest' computer by far in the world was the CRAY-1. Needless to say that I fantasised about a home-computer with the unbridled power of a CRAY-1. This fantasy later expanded: the build-in programming language had to be Forth.

In 2016 I started a project to finally fulfil that old dream: create a system as enjoyable as my 80s home-computer – with a decent Forth build-in, and performance-wise close to a CRAY-1. WabiPi4 is a result of this project.

The performance of wabiPi4 exceeds my wildest dreams. Using the 'Sieve' benchmark from BYTE magazine, wabiPi4 is over 123 times faster than a CRAY-1. As only around 90 CRAY-1's were ever built, wabiPi4 is more powerful than the combined power of all CRAY-1's ever built...

### Dream fulfilled? I should think so!

Home computing in the early 80's focused on DIY programming. Graphics, sound-generation and limited interfacing were important building blocks of the fun. The best home-computers also had a build-in assembler. WabiPi4 stays true to this original focus.

WabiPi4 also follows another 80's tradition: like any home-computer from that time, it is 'blessed' with quirks and faults. For instance: most modern technologies are not supported. Things like USB-ports, internet or wired/wireless networking. And, like most early 80s home-computers, the manual, which you are reading now, is incomplete, unstructured and a constant work-in-progress.

It is unlikely that any of this will change a lot in the future. WabiPi4's raison d'être is the enjoyment of programming and playing around, like the early home-computers. But with more memory and very, very much faster. 'Vintage computing on a lot of steroids'.

WabiPi4 is what it is. Enjoy it and have fun, or move to other and better, systems. (like noForth-T, CiForth, iForth, eForth or Mcrisp)

Jeroen Hoekstra

**Content**

Copyright and license	2
Foreword	3
Introduction	6
WabiPi4 – it's all in the name	7
Cooling the wabiPi4 system	8
Operating-system	10
Starting, aborting and exiting wabiPi4	12
Limits of the wabiPi4 system	15
Dictionary	16
Stack manipulation	19
DO...LOOP & FOR...NEXT	21
MOVE & FILL	27
Store, fetch and values	28
Comparisons	31
User Input	32
User Output	33
String handling	35
Overflow protected string variables	37
Memory-map and statistics of wabiPi4 system	40
Memory allocation	42
Wordlists (a.k.a. vocabularies)	44
Stacks	46
Random numbers	48
System constants, variables and values	53
Time and time-measurement related words	55
Not so serious stuff	60
GPIO	61
Of bytes, bits and bobs...	65
Cyclic Redundancy Check ( a.k.a. CRC )	70
WAVE – sound generator	72
PIXEL	76
Compilation/Execution control	84
CATCH THROW	85
Optimisation in wabiPi4	86
Hash-table functionality	91
Real time clock module	94

ARM-v8 assembler	96
Lessons from one opcode...	106
Snippets of code – examples of assembly	107
Primitives using high level code	113
Example of a longer assembly routine	116
Thumb assembler	118
5CH code	120
Interrupts	121
Cache-handling	124
ARM-VideoCore mailbox	125
ARM-processor ID, status and info registers	133
ANSI-required documentation	135
Air Traffic Controller	141

# Introduction

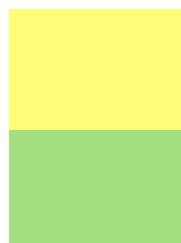
**As of Januari 2024 wabiPi4 contains 2453 definitions.** Compared to some other Forth system not a huge number. But a daunting number if you are a beginner. Fortunately most definitions are related to topics like home-computer functionality, the sound-system and the graphical system. And the assembler alone accounts for more than 550 definitions.  
 All useful and fun functionality, but not essential when learning Forth. Once you have learned a bit of Forth, there is a lot to explore within wabiPi4.

Call .STATS and you will get an overview of the different categories of definitions:

```
=====
904 words
493 inlinable primitives
165 primitives
50 values
0 2values
55 variables
5 2variables
241 constants
6 2constants
534 created with does>
0 deferred
=====
2453 definitions in total
```

**This manual presumes knowledge of Forth.** If you do not know how to program using Forth, there are excellent introductory books available. The classic books "**Starting Forth**" and "**Thinking Forth**" by Leo Brodie are a very good starting point. And Elisabeth Rather, the second Forth-programmer in history, has co-authored an excellent book: "**Forth Programmer's Handbook**".

Membership of a local Forth Interest Group is also a very good way of starting with Forth.



## WabiPi4 – it's all in the name



The name wabiPi4 comes from the Japanese concept '**wabi-sabi**'.

Translating these two words is the easy part:

**wabi:** "nothing is ever finished or perfect"  
**sabi:** "everything comes to an end"

Explaining the deeper meaning, and the role, of these two words in Japanese Culture is very complex though, if not impossible.

An explanation could start with something like: 'You can only be happy if you accept that something is never perfect and will not last forever'. The second phase is: "if nothing can be perfect, then at least let's make sure that the imperfect is as enjoyable as possible".

Apply this to wabiPi4 and you get something like this:

**wabiPi4 is unfinished and imperfect**

**but I tried hard to make it really enjoyable.**

# Cooling the wabiPi4 system

**The wabiPi4 system is based on a Raspberry 4b with 4 or 8 GB.** At the start of the project, a Raspberry Pi was the easiest and cheapest way of getting a fast processor with a large memory. The other hardware on the Raspberry (USB, network-connection, Bluetooth etc.) is irrelevant for wabiPi4.

## Minimum wabiPi4 system:

1. Raspberry 4b with 4/8 GB with an excellent cooling solution
2. Power supply unit 5v, 2.5A or better
3. serial to USB conversion cable
4. terminal program on a PC (OSX, Windows, Linux)
5. SD-card with wabiPi4

**A good power-supply** is critical for a Raspberry to run reliable. Overall power is important, but the stability of the delivered voltage is even more important. A short dip below 4.6 volt will result in a reset or crash. If this happens you could try to raise the voltage of the power-supply. Aim for 5.2–5.3 volt.

## Power-consumption of the pi4b

The actual power-consumption of a pi4b running wabiPi4 is between 4.5–5.0 Watt. A full workload or very complex sound-generation adds 1–2 Watt.

Not a lot, but enough to make good cooling essential!

## STEP 1: ADD COOLING – COOLING IS ESSENTIAL

A Raspberry running Linux does not need CPU-cooling. Linux constantly measures the temperature of the CPU, and lowers the CPU-frequency if it gets too hot. In addition Linux enables only those parts of the CPU which are in active use.

## WabiPi4 does the opposite

WabiPi4 overclocks the CPU, works 3 of the 4 cores pretty hard, activates most of the CPU, and at all times, even when waiting for user-input, generates high memory-traffic. And, to add insult to injury, wabiPi4 does not check the temperature of the CPU. And so:

---

## A Raspberry running wabiPi4 must be cooled!

---

Add a large cooling-element to it. Ideal are the cooling cases of Joy-It. Or a heat-pipe based cooling. Large (25x25mm or larger) cooling fins for SMD chips also work, but only just. The cheap 14x14mm cooling fins sold everywhere are inadequate.



Examples of good cooling solutions for the CPU of a Raspberry

## STEP 2: CONNECT THE REST

Connect the Raspberry via a serial-to-USB cable to a PC. On the PC a terminal program is used to communicate with the Raspberry. In the terminal-program choose **115200-8N1** as protocol for the setup for the connection, resulting in a speed of 115.2 Kb/s.

A monitor can be connected to the Raspberry using a HDMI cable. Unfortunately the HDMI connector on the Raspberry 4B is a micro-version. This port is mechanically rather fragile, take care!

The video-output is fixed @ 1024x768 pixels, a dream-resolution for a home-computer-user of the 80s. The colour-depth per pixel is 24 bits.

## STEP 3: LOAD WABIPI4 ON SD

Load the images in the top-directory of the sd-card, put the sd-card carefully in the Raspberry and you are done.

## OPTIONAL STEPS – BETTER DO THIS LATER

### Optional extra's

Included in wabiPi4 are I2C drivers for a 20x4 character LCD-screen, 384 kB eeprom and a RTC. Obviously the user can develop other drivers, for I2C or SPI or anything else. **Project Forth Works** on GitHub has some interesting examples for other I2C and SPI drivers.

# Operating-system

WabiPi4 is a bare-metal implementation of Forth. There is no operating system running in the background of WabiPi4.

No Linux, no Android, no RTOS, *no nuthink...*

After completion of the boot-process, three tasks are running in the background:

1. on **core1**: the graphics-system
2. on **core2**: the keyboard input and buffer system, with **core0** handling user-aborts
3. also on **core2**: the sound-system management

As **core1** and **core2** handle all background tasks, wabiPi4 on **core0** is free from interrupts, and is available all of the time. This is optimal for performance and exact timing.

The only interrupt possible for **core0** is a user-abort. This is when the user wants to interrupt a running program, for instance because the program is stuck in a loop, by pressing 'FnCtrlBackspace'.

Any other functionality normally provided by an operating system is either handled by wabiPi4, is provided/programmed by the user, or is not possible.

## ACCESS RIGHTS

An important goal during the development of wabiPi4 was to give the user the fullest possible access to the ARM-processor, in this case the Cortex-a72. In the 80s, experimenting with the processor was one of the most exiting aspects of home-computing. And so wabiPi4 *had* to support experimentation with the Cortex-a72.

The 'h' in 'software development' stands for 'happiness'...

In concreto: wabiPi4 does not manage, control or limit access-rights. On the contrary, the system runs in (secure) supervisor-mode, giving the most complete access possible to the CPU. And in addition, every available option giving even more access to the ARM-processor is enabled.

The only part of the system not open are those registers which are only exclusively accessible in unsecure modes, specifically HYP-mode related registers.

All other registers, around 99%, are accessible. If you want to write to a configuration-register of the CPU, you can. The system might crash, but you *can* try.

The hypermode, as an aside, is relevant for WabiPi4. One example is setting the **CNTVOFF** register. This CPU register contains an offset from the physical counter for the virtual system counter, and is only accessible in (unsecure) HYP-mode.

The other is self-hosted DEBUG. Self-hosted DEBUG-mode is enabled in WabiPi4 at the highest level (ie. invasive DEBUG-mode). Enabling self-hosted DEBUG for a secure mode can also only be done in HYP-mode. And as invasive DEBUG-mode must be enabled to get access to the secure CPU-cycle counter, wabiPi4 runs in invasive self-hosting DEBUG-mode.

There is no specific support in WabiPi4 for self-hosted debugging. But if you want to have a go, the registers are accessible.

As the system runs in 'supervisor' mode, it is, by the way, entirely possible to program a Hyper-mode management system with wabiPi4. That is left as an interesting little exercise for the user – but, maybe, not as an exercise for an absolute beginner.

## MULTITASKING

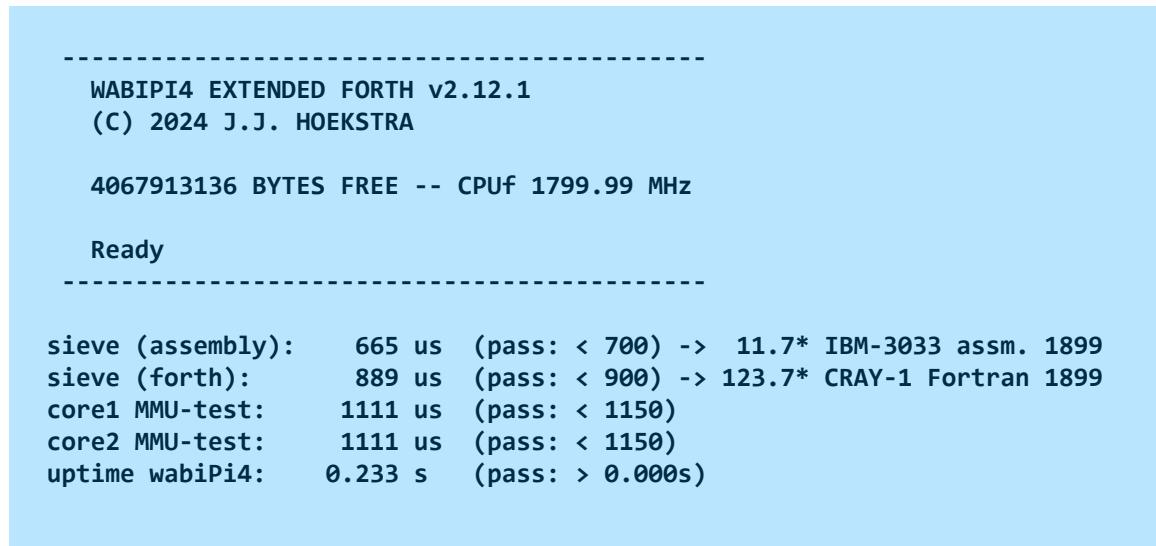
There is no multi-tasking. This has to be provided by the user. It is a tentative intention to add this functionality in future, but this is not guaranteed nor in the planning.

# Starting, aborting and exiting wabiPi4

## Starting wabiPi4

Wabi starts by switching the power 'on'.

After a successful start you will be greeted by a boot-screen:



The **bootscreen** gives information on the system and the results of 5 system sanity-checks.

The **first part** is inspired by the start-screen of my first computer, the 'ORIC-1' from 1983. The wabiPi4-variant shows the version number, a copyright notice, the amount of free memory and the actual present CPU-frequency.

The **second part** shows the results of 5 sanity-checks. If all 5 checks pass, there is a reasonable chance that the system is running as intended. No guarantees, but for the checks to run successfully, a lot of steps in the setup-process have to be executed correctly.

**Test 1** (sieve in assembly) – shows that core0 is running as intended. There is also a bit of unnecessary bragging, as it also shows that the sieve benchmark runs more than 11 times faster in assembly than the IBM-3033 (the 'big one') did in assembly. Finally the number 1899 proves that the sieve actually found the right amount of integers.

**Test 2** (sieve in forth) – shows that wabiPi4 is compiling, optimising and executing correctly. And again a bit of bragging as it shows that the sieve benchmark in wabiForth runs ~123 times faster than a CRAY-1 using optimised Fortran.

**Tests 3 & 4** (core1 and core2 MMU-test) – show that cores 1 and 2 are configured and running correctly.

**Test 5** (uptime) – shows that the clock-system is running. You might notice that the first time wabiForth starts, an uptime of around 0.232s is reported, although starting takes around 5 seconds. This is because starting the Raspberry Pi4 hardware takes 4.8 seconds. Much slower, by the way, than a Raspberry Pi3, which starts within a second.

## Reset of wabiPi4

**COLD** restarts wabiPi4 without resetting the CPU. Most system-settings are reset and any previous program is not present anymore. Memory is not cleared though by **COLD**. Using **DUMP** you can see the remnants of previous programs after using **COLD**. The settings for the compiler and the screen are not reset. This allows the programmer to experiment with the effect of these settings of the compiler on the Sieve benchmark.

A full reset of the CPU is done with **RELOAD**. **RELOAD** resets every part and every core of the system in accordance with the architectural specifications of ARM and Broadcom. It then reloads wabiPi4 from the SD-card. Any program and any data in memory is lost and all system settings are re-written.

Rebooting by using the boot-vector at address 0x0 (by typing '**0 execute**') restarts core 0. Core 1 and core 2 are left as they are. So this is not a full reset, but it does usually function. You can overwrite the vector at address 0x0 if you want to. A, rather theoretical, example where this might be useful is if you want to play with ARM64 code. A reset is needed to switch to ARM64, and this address points to the code to be executed after the reset. For more info see **CPU\_RMR64**.

## Exit of a running program

Pressing the keys 'fn', 'ctrl' and 'backspace' at the same time, on an Apple OSX computer using 'Terminal', will abort a running wabiPi4 program and return wabiPi4 to where commands can be typed in.

The memory is not cleared by this abort, giving you the chance to check values or restart the program or whatever else you want to do.

If 'fn', 'ctrl' and 'backspace' does not function then something is seriously wrong with wabiPi4. Or you are not using an Apple OSX computer, which is the same. In that case only a hard reset, by switching the system 'off' and 'on', will help. Any programs and data in the system will be lost.

As wabiForth is running in 'invasive self-debug mode', the 'debug-over-reboot' feature of the ARM-CPU, in rare cases, makes it necessary to switch off for 10–20 second or so. Just like old home-computers. This is not a bug, this is a 'feature'...

## Aborting a running program from within the program

Aborting a running program is done with the words **ABORT** or **ABORT"**. **ABORT** exits a running program immediately, giving control back to the user.

**ABORT"** takes a values from the stack. If the value is not zero, it prints the string between the quotes and then aborts the program.

## Exit of wabiPi4

is done by switching it 'off'. The classic word '**BYE**' is available but has a peculiar will of its own and thus its use is discouraged (try it...).

**COLD** ( -- ) restarts wabiPi4, resets most of the system-variables and reloads the tools-dictionary. The compiler optimisation-settings and screen-settings are not reset.

**RELOAD** ( -- ) restarts the Raspberry Pi, reloads the firmware of the Raspberry and reloads wabiPi4. (synonym: **REBOOT**)

**ABORT** ( -- ) immediately exits a running program and puts control back with the user. It does not clear memory, so the user can inspect items like variables etc. If an abort happens during compilation, **HERE** is put back to where it was before the compilation of the word started. Thus avoiding a memory leak.

**ABORT"** ( flag -- ) **ABORT"** is followed by a string ending with a ''''. If it finds a '**true**' flag on stack, it prints the string between the quotes and exits a running program. The message can be used by the user to signal where the abort occurred and what the reason for the abort was. **ABORT"** only functions inside a definition.

**BYE** ( -- ) has no useful functionality. Where would you go outside of wabiPi4? There is not even a bare emptiness...

## Limits of the wabiPi4 system

WabiPi4 is a 32 bit system and has access to 4 GB of memory in the system, also on a pi4 with 8 GB memory. The memory of a 4 GB pi4 is split into a lower 1 GB part and an upper 3 GB part. The dictionary of wabiForth is located in the lower part, the upper part is available for data.

<b>maximum size dictionary</b>	983 MB
<b>maximum size allotted memory block within the dictionary</b>	983 MB
<b>maximum size allocated memory block in data-part of memory</b>	2880 MB
<b>maximum size string</b>	strings have a 32 bit index – depending on where they are defined, the actual maximum size is 983 or 2880 MB.
<b>maximum size of a single word/definition</b>	32 MB per definition. The system might, but must not, become unstable if a definition is larger than 32MB. Utterly irrelevant in normal coding, but it could be relevant for automatically generated code.
<b>maximum number of definitions in the dictionary</b>	Limited by the size of the hash-table. Between 4753485 (worst case) and 5592373 (best case) definitions. Likely enough for most applications.
<b>maximum number of vocabularies in the dictionary</b>	Up to 255 word-lists can be defined in the system. Up to 8 can be put in the search-order-list at the same time.
<b>depth of stacks</b>	There are four stacks in wabiPi4: data, return, cpu-return and floats. Each stack is 2048 cells deep. No checks are done on under or over-runs.

## Hardware specific wabiPi4 limitations:

The words **2@** (two-fetch) and **2!** (two-store) take an address as argument. Contrary to wabiForth for the pi3b+, for wabiPi4 these addresses do not have to be 4-byte aligned.

### Assembler related limitations

Several opcodes in the ARMv8 assembler have specific limitations with regard to address-alignment. That is how the ARM-cooperation designed them. For specifics see the ARMv8 manuals from the ARM-cooperation.

# Dictionary

The **wabiPi4 dictionary** is a linked list of definitions. Listing words in the dictionary can be done using the ANSI standard **WORDS** or the following words:

**WORDS** ( -- ): lists all words in the dictionary and prints the total number of definitions in the dictionary. Hidden words are not shown. Prints the definitions sorted by wordlist in the search-order.

**LWORDS** ( -- ): prints out the most recent words. The number of words printed is contained in the value **#WORDSLW**. The default is 50. Prints the definitions sorted by wordlist in the search-order.

**AWORDS** (<"ccc"> -- ): prints all the words starting with the first character specified after **AWORDS**. For instance: typing '**AWORDS d**' will show all words starting with 'D' or 'd'. **AWORDS** prints all non-hidden words, irrespective of word-lists, and is case-insensitive.

**WORDSL** ( -- ): prints out a list of all words in the dictionary including the execution-token and the inline-length. A word can be inlined if the inline-length is larger than zero. For a word to be inlined, inlining must be switched 'on'. In addition the inline-length of the word must not larger than the length contained in the value **MAXINLINE**. The default behaviour is that all primitives are inlined.

As a side note: on the Raspberry 3b+, the maximum inline-length is set by default at 8 opcodes, as inlining with more than 8 opcodes never gives any speed-advantage with the processor used in de raspberry 3b+. (well almost, in some cases a marginal 1 CPU-cycle advantage is measurable, but at the cost of heavier use of the I-cache).

**WORDSCOMMA** ( -- ): prints a comma-delimited list of all words in the dictionary, including the execution-token, the inlining-length and the length of the name. This list can easily be imported into a spreadsheet.

**LWORDSL** ( -- ): combined functionality of **WORDSL** & **LWORDS**

**LWORDSCOMMA** ( -- ): equivalent to **WORDSCOMMA**

**WORD#** ( -- n ): returns the total number of definitions in the dictionary, including redefined definitions and user-definitions.

**#WORDSLW** ( -- n ): The value **#WORDSLW** contains the number of words to be printed by the word **LASTWORDS**. By default this is 50 but it can be changed by the user.

If you for instance enter:

**10 TO #WORDSLW**

**LASTWORDS** from then onwards will print the last 10 words.

## FINDING DEFINITIONS

### Printing the name of definition

If you want to know the execution token (**XT**) of a definition you use '**(TICK)**', a ANSI standard word. But what if you have an **XT** and you want to know which definition belongs to it? One way is to use **.XTNAME**.

More complicated is the case where you only have an address and you want to know whether it is part of a definition and, if yes, from what definition. For this you can use '**ADDR>XT**'. This word, for a given address, finds the corresponding **XT** or returns a 'zero' if the address is not part of any **XT**.

The words **.XTNAME** and **ADDR>XT** are handy when writing debugging-routines.

The following examples shows two definition:

```
.MYSELF prints the name of definition in which .MYSELF is
located.
.BYWHOM prints the calling definition.

: .MYSELF r14@ addr>xt .xtname space ;
: TST1 .myself ;

: .BYWHOM ( -- )
  r13@ 4+ @
  addr>xt dup ." called by: " .xtname
  space ." with XT: 0x" .hex ;

: WHO? cr .myself .bywhom ;
: TST2 cr .myself .bywhom who? ;
```

Execute **TST2** and you will see that it was called by **INTERPRET**, and **WHO?** will report it was called by **TST2**.

The definitions **.MYSELF** and **.BYWHOM** rely on **R13@** and **R14@**. Words which get the present value of R13, the CPU return-stack pointer, respectively R14, the CPU link register. They are explained elsewhere in this document.

**.XTNAME** ( **xt --** ): given a valid **XT**, prints the name of a definition. If there is a **0x0** on the stack, **.XTNAME** will print "not part of the dictionary".

**.XTNAMECAPS** ( **xt --** ): given a valid **XT**, prints the name of a definition in capital characters. If there is a **0x0** on the stack, **.XTNAMECAPS** will print "not part of the dictionary".

**ADDR>XT** ( **addr -- XT/0** ): for a given address, finds the corresponding valid **XT**, or returns 'false' if the address is not part of the dictionary. This word works well together with **.XTNAME** and **.XTNAMECAPS**.

## THE HEADER

Each definition in the dictionary has a header which contains the link to the next word, the name of the definition and essential meta-data. The header is key to correct functioning of the dictionary.

The header has a fixed length of 48 bytes. The name of the definition can contain a maximum of 32 characters. The first 4 bytes of each header are available to the user. Byte 5 and 6 are reserved for future use.

### header of a wabiPi4 definition

#### byte content of byte/word/string

byte	content of byte/word/string
0	free
1	free
2	free
3	free
4	reserved for future use
5	reserved for future use
6	object_id: 1=value, 2=variable, 3=constant, 4=variable, 5=constant, 6=deferred, 7=primitive, 8=created w/ DOES>, 9=wabiForth def, 10=inlinable primitive, 11=2value
7	hidden flag: 0=print, 1=do not print
8	link word - points to XT of next word
:	...
:	...
:	...
12	inlining length: 0=never inline - 1 or higher: number of opcodes to be inlined if inlining is active
13	compiling=-1 immediate=0
14	WID (wordlist identifier) - 0=Forth
15	length name - max=32 characters
16	1st char name
:	2n char name
:	...
:	...
47	last char name
48	1st opcode - address = XT
:	

## Stack manipulation

The ANSI standard defines 26 words concerning stack manipulation:

DUP	2DUP	FDUP
DROP	2DROP	FDROP
OVER	2OVER	FOVER
SWAP	2SWAP	FSWAP
ROT	2ROT	FROT
R>	2R>	
R@	2R@	
>R	2>R	
?DUP		
NIP		
PICK		
ROLL		
TUCK		



oh happiness...

In wabiPi4 these words function as described in the standard.

From the empty yellow cells in the table above, it is clear that some functionality is left out of the ANSI standard. For instance, there is no standard way of pushing a float on the return-stack, there is no **2NIP**, etc. At the time of the development of the ANSI-standard, there was a movement to limit the number of words in de standard.

Too many words was seen as too difficult to remember for Forth-programmers. And so some words, deemed less important, were left out. The unfortunate result is that a programmer now has to remember which words do NOT exist...

Fine if you are a Forth-purist, glory be! But less so if you do a lot of floating point calculations, or work a lot with doubles. In those cases you are relegated to self-developed and slow high-level definitions.

WabiPi4 adds the missing words, mostly as optimised primitives, and then some more.

Use of any one of these words renders a program non-standard, unless you provide an alternative high-level implementation of the offending word. But inclusion makes your program run faster. Sometimes hardly measurable, sometimes clearly noticeable.

**?2DUP** ( `dn -- double_zero / dn dn` ): conditional duplicate. If dn is unequal to zero -> duplicate the dn.

**2NIP** ( `dn dm -- dm` ): remove dn from stack.

**2PICK** ( `dn*x n -- dn*x dn` ): pick the nth double value from stack and put it on top.

**2ROLL** ( `du du-1...d0 u -- du-1...d0 du` ): Like the ANSI word **ROLL**, but for double values. U is a single. Programmed in hi-level Forth.

**FPICK** ( `fn*x n -- fn*x fn` ): Pick the n-th floating point value from the floating point stack and put it on top of the stack.

**3DUP ( n o p -- n o p n o p ):** copies the top 3 values on stack

**4DUP ( n o p q -- n o p q n o p q ):** copies the top 4 values on stack

**1DROP ( n o -- n ):** exactly the same function as **DROP**. But has the same opcode structure as **2DROP**, **3DROP** etc. This makes **1DROP** useful in timing measurements.

**-ROT ( n o p -- p n o ):** moves the top value to the 3rd position of the stack and moves the other two values up one position.

**2-ROT ( dn do dp -- dp dn do ):** moves the top double to the 3rd double position of the stack and moves the other two double values up one double position.

**F-ROT ( fn fo fp -- fp fn fo ):** moves the top float value to the 3rd position of the float-stack and moves the other two float values up one position.

2TUCK

RDROP ( FRDROP not needed )

#tbd: FTUCK, FR>, F>R, FR@,

# DO...LOOP & FOR...NEXT

## DO...LOOP

WabiPi4 follows the ANSI standard: **DO**, **LOOP**, **?DO**, **+LOOP**, the indexes '**I**' and '**J**' and **LEAVE**, **UNLOOP** and **EXIT** are available. In addition the non-ANSI word '**K**' is available.

### WabiPi4 specifics

WabiPi4 introduces **!DO (exclamation-DO)**. **!DO** does the opposite of **?DO**: if the two input-values are equal **!DO** will **ABORT**.

**!DO ( n m -- )**: same function as **DO**. But aborts on equal input-values.

This 'early abort' principle can be helpful in finding bugs early.

### Limits of DO...LOOP in wabiPi4

The distance of the jump back from **LOOP** to **DO** cannot exceed 32MB. For normal programming not an issue, but it could be an issue with algorithmically generated computer programs.

**LEAVE**, if used, must be located within the **DO...LOOP**. It cannot be part of a word called from the **DO...LOOP**.

### The DO...LOOP is highly optimised

The **DO...LOOP** is the most optimised part of the wabiPi4 system. Two CPU-registers are dedicated permanently to the **DO...LOOP**. This in itself makes looping a lot faster. But it also allows for additional optimisation using inlining and compounding. The following example compares moving a memory-block from source to a destination with the **DO...LOOP** and with **MOVE**. **MOVE** uses the fastest method (according to the ARM-corporation) in assembly. Obviously it is faster, but the **DO...LOOP** doesn't do too badly.

```
decimal
100000000 allocate drop constant tstsource
100000000 allocate drop constant tstdest

: fillsource 25000000 0 do      \ fill with random numbers
    rndm32 tstsource i 4 * + !
loop ;

: tst1 25000000 0 do
    tstsource i 4 * + @        \ get a word
    tstdest   i 4 * + !        \ store a word
loop ;

: tst2 tstsource tstdest 100000000 move ;

fillsource t[ tst1 ]t.
fillsource t[ tst2 ]t.
```

The **maximum bandwidth of the memory-system** of the Raspberry 4b is around 5 GB/s, and the GPU consumes ~1.3 GB/s of that. So **MOVE** comes close to the maximum.

method	typical execution times (microseconds)	relative to FASTMOVE	bandwidth (MB/s)
<b>DO...LOOP</b>	~134359	253%	~1488
<b>MOVE</b>	~52973	100%	~3775

The **DO...LOOP** is 2.5 times slower than **MOVE**. Slower, but still, pretty impressive. Consider that **MOVE** uses the fastest possible method in assembly. It is possible that you get ~30% faster times. A reset of the system will cure that silly thing...

Some specific reasons why the **DO...LOOP** is slower:

1. more opcodes are used for the move of the data
2. the forth-words between **DO** and **LOOP** manipulate the stack which consumes memory-bandwidth in itself.

#### What if the GPU is not active?

In that situation the whole of the memory-bandwidth is available for the move of data, and both routines are faster:

method	typical execution times (microseconds)	relative to FASTMOVE	bandwidth (MB/s)
<b>DO...LOOP</b>	~88956	258%	2248
<b>MOVE</b>	~38487	100%	5196

#### Memory-alignment

**MOVE**, @ and ! function with and without memory-alignment. But whereas the performance of the **DO...LOOP** is largely independent of alignment, the **MOVE** routine can slow down somewhat when data is moved from or to an unaligned address.

Please note that using:

```
create tstdest 100000000 allot
create tstsource 100000000 allot
```

to define the two memory-blocks would result in slower execution of the code.

There are two reasons for this:

1. this creates non-inlinable, and thus slower, words to put an address on stack. Normally not noticeable, but clearly so when you do 25 million loops with a **DO...LOOP**, where these two words each put a value on stack every loop.
2. **ALLOT** reserves space in the dictionary-area. The memory-attributes (especially related to caching of data) for this area are for mixed use, ie executables and data. The consequence is slower performance for data-only tasks. (see '**ALLOCATE**')

**DO...LOOP** can be combined with **AFT/THEN**. See **FOR...NEXT** for an example.  
 This is fun to try, but non-conformant with the ANSI-standard.  
 And so hinders exchange of source-code with other Forth-versions.

## **FOR...NEXT**

On small systems, the **FOR...NEXT** construct is faster than the **DO...LOOP**, uses less cells on the return-stack and is smaller and easier to implement. All important advantages.

### **Reality-check...**

On larger CPUs, like the Cortex-a72 in the Raspberry 4b, testing shows that there is no speed-advantage. And saving a few cells on the return-stack or a few words of memory is irrelevant if you have GigaBytes to work with.

### **FOR...NEXT – choices to be made**

As there is no speed-advantage to be gained, wabiPi4 provides the **FOR...NEXT** loop mainly for compatibility-reasons. Preferably with as many different Forth-implementations as is possible.

This is not an easy task; **FOR...NEXT** is not part of the ANSI-standard. The ANSI committee found that standardising **FOR...NEXT** was impossible, as the various existing implementations differed too much.

There are (at least) two ways of designing a **FOR...NEXT** loop:

#### 1. **PRE loop subtraction**

At the start of the loop, subtract 1 from the index and check if the end of the index is reached, if not perform a loop. This version does not perform a loop if the input is zero. The first index available is the input-value minus 1.

#### 2. **POST loop subtraction**

First perform a loop, then subtract 1 from the index and check if another loop has to be done. This version always performs at least 1 loop. The first index available is the input-value and this version does 1 loop more than the input-value.

**Please note** that for programming reasons the variants can be programmed in different ways but with the same functionality.

Other differences between different Forth-implementations are related to:

- the availability of index '**J**' for a **FOR...NEXT** loop contained in a **DO...LOOP** and visa versa
- combining of **FOR...NEXT** with **WHILE/ELSE/THEN**
- the availability of the words **AFT/THEN**
- the availability of **LEAVE**, **EXIT** and **UNLOOP**
- use of signed positive numbers or unsigned numbers for the index

Take all these differences into account and it is easy to see that the humble **FOR...NEXT** loop has greater complexity than obvious at first sight, and it is easy to understand why a universal standard solution eluded the ANSI-committee.

### The wabiPi4 **FOR...NEXT** implementation

WabiPi4 supports both variants of **FOR...NEXT**: **FOR\_pre** and **FOR\_post**.

The following features are available in wabiPi4 for **FOR...NEXT**:

- interactive choice of the two variants during compilation
- The index '**J**' is available in nested loops. This is true for **DO...LOOP** and **FOR...NEXT** in any combination
- **UNLOOP** and **EXIT** can be used

In addition:

- **AFT/THEN** and **LEAVE** are available in the **FOR\_post** variant
- **WHILE/ELSE/THEN** is available in the **FOR\_pre** variant

### Choosing the **FOR...NEXT** variant

Both variants are available in wabiPi4. Both principles have advantages and disadvantages. During compilation two words govern which variant is compiled:

**FOR\_post** ( -- ): immediate – The **FOR...NEXT** loop functions as in eForth: The index starts at the input-value, and a loop is always done at least once. **LEAVE** and **AFT/THEN** can be used, **WHILE/ELSE/THEN** not.

**FOR\_pre** ( -- ): immediate – The **FOR...NEXT** loop functions as in noForth. The index starts counting at the input-value minus 1, and the loop will not run on an input of zero. **WHILE/ELSE/THEN** is available. **LEAVE** and **AFT/THEN** are not.

Changing which **FOR...NEXT** variant has effect on the next **FOR...NEXT** loop(s) to be compiled. Existing loops, and loops still being compiled are not influenced.

### Compatibility with other Forth-implementations

The two variants together should allow compatibility with most Forth-implementations.

The **FOR\_post** variant is, for instance, compatible with eForth and iForth and many other implementations. The **FOR\_pre** variant is compatible with noForth.

**Please note:** **LEAVE** from **FOR...NEXT** is not always available in other Forth-implementations; do not use it if you want to be compatible with other implementations.

Example of the effect of switching the **FOR...NEXT** system:

```
FOR_pre
: for_v1 cr 5 for i . next ;

FOR_post
: for_v2 cr 5 for i . next ;

for_v1 ( prints: 4 3 2 1 0 )
for_v2 ( prints: 5 4 3 2 1 0 )
```

Example of **UNLOOP** and **EXIT** in **FOR...NEXT**

```
: tstf 10 for
    i dup . 5 = if
        unloop exit      \ exits the definition!
    then
    next ." end" ; \ end is never printed!

tstf ( prints: 10 9 8 7 6 5 )
```

Example of **LEAVE** in **FOR...NEXT**

```
FOR_post
: tstl 10 for
    i dup . 5 = if
        leave          \ exits the loop
    then
    next ." end" ; \ end is printed

tstl ( prints: 10 9 8 7 6 5 end )
```

### Use of **AFT & THEN**

As far as I can find, the word **AFT** was introduced by eForth. In wabiPi4 it is included for compatibility reasons. This word is one of those concepts which either grows slowly on you, or not at all.

It is a powerful word. You will find several interesting examples of the word in the book: "eForth Overview" by dr. C.H. Ting.

**AFT ( -- )**: immediate – exchanges a destination from R-stack with a destination pointing to just after **AFT**. And adds a new destination on R-stack pointing to **AFT** itself.

This manipulation of destinations allows the program to do something different during the first **FOR...NEXT** loop compared to the other loops.

Like this:

**AFT and THEN in FOR...NEXT:**

```

FOR
    {only executes during first loop}
AFT
    {executes during all other loops}
THEN
    {executes during all loops}
NEXT

```

**Please note:** in wabiPi4, **AFT/THEN** can also be used in the **DO...LOOP**. But ecod, a useful purpose eludes me...

**Example with FOR...AFT...THEN...NEXT**

This short example defines a word '.A' (DOT\_A) which prints as many A's as the value on stack. If the value on stack is zero, it will print zero A's, like it should. And that without any check if there is a zero as input.

```

FOR_post
: .A ( n -- ) for aft ." A" then next ;

```

**Example with FOR...NEXT and WHILE/ELSE/THEN**

This example shows the use of **WHILE/ELSE/THEN** in **FOR...NEXT**. **KEY\_TIMEOUT** waits for a keystroke and returns the keystroke. If it takes too long an error-message is returned.

Because of the speed of wabiPi4 a large input-value is needed. A value of a hundred million gives a time-out time of around 5 seconds. This large number is very energy-inefficient, as the CPU is working very hard while waiting. It also looks silly. In the **assembly-examples** chapter there is an example showing how, by mixing in one opcode with the assembler, a more energy efficient wait-loop can be implemented.

```

FOR_pre                                \ FOR_post does not support this
: KEY_TIMEOUT
  1000000000 for
    key? 0= while
  next
    ." time-out"
  else
    key emit unloop
  then ;

```

## MOVE & FILL

**MOVE**, **CMOVE**, **CMOVE>**, and **FILL** function as described in the ANSI standard.

In addition the following words have been defined:

**FASTMOVE** ( **orig**, **dest**, **no\_of\_words** -- ): Moves 32b words from origin to destination. On the pi3b+ this word is clearly faster than **MOVE**, thus the name. But on the pi4b, with a more efficient memory and store/load system, this is usually not the case. It is provided to be compatible with wabiForth for the pi3b+. Contrary to the pi3b+ version, there is no need to align the origin and destination addresses for **FASTMOVE**. But using unaligned addresses can result in a up to 25% slower move of data.

In the **DO...LOOP** chapter there is an example comparing the speed of **FASTMOVE** and **DO...LOOP**

**Please note:** in wabi v2.5.3 and earlier, **FASTMOVE** contains a bug which resulted in errors when moving up. From v2.6.0 onwards this bug has been resolved.

**FASTFILL** ( **addr**, **no\_of\_words**, **n** -- ): Fills the memory starting at address with the 32b word n, no\_of\_words times. **FASTFILL** on the Raspberry pi3b+ was clearly faster than **FILL**, but on the pi4b there is no difference. **FASTFILL** is provided here for compatibility with pi3b+ programs.

Addresses can be non-aligned at little or no loss of speed.

# Store, fetch and values

WabiPi4 provides all the standard words to store and fetch data from memory and variable. The standard words !, @, 2!, 2@, C!, C@, +! function as defined in the ANSI-standard. And naturally the standard words **VALUE**, **T0** and **+T0** are included.

## DOUBLE VALUES

The ANSI standard is lacking standards for double values, which seems a bit inconsequent. In wabiPi4 double values are available, and **T0** and **+T0** handle both **VALUE** and **2VALUE** correctly. Overall there is little or no speed advantage in comparison with double variables. **VALUES** and **2VALUES** are used to avoid having to write directly to memory, thus avoiding an important source of bugs, bugs which can be difficult to track. The word **M+T0** is available to add/subtract single numbers from a double value, it is faster than **+T0**.

**2VALUE <name>** compiling:( double -- ) execution:( -- double ): like **2VARIABLE**, stores double numbers. **T0** and **+T0** stay the same. They correctly handle both **VALUE** and **2VALUE**.

**M+T0 <name>** ( n -- ): adds a single to a double **VALUE**. Significantly faster than **+T0** for a **2VALUE**. So if you want to add a small number to a double, for instance to count something, use **M+T0**. Compared to **M+2!**, **M+T0** is also faster, around 25%.

## SPECIFIC STORE AND FETCH-WORDS

The following words are included for performance reasons. They can all easily be programmed in Forth, but as modern CPU's usually have specific opcodes for this functionality, it makes sense to include them as primitives.

### Halfword fetch/store

**H@ ( addr -- 16b )**: fetches a halfword (=16bit value) from memory.

**H! ( 16b addr -- )**: stores a halfword (=16b value) in memory

**SH@ ( addr -- 16b )**: fetches a signed halfword (= signed 16bit value) from memory and extends the sign-bit to 32b.

### Signed character fetch

**SC@ ( addr -- 16b )**: fetches a 8bit value from memory and extends the sign-bit to 32b.

### Add-stores

**C+! ( n addr -- )**: raises the 8bit value at address with 8bit value n – idea: W. Baden

**H+! ( n addr -- )**: raises the 16bit value at address with 16bit value n

**M+2! ( n addr -- )**: raises the double at address with the single signed value n

**D+2! ( d addr -- )**: raises the double signed value at address with the double signed value d

**Auto-indexing fetches**

**C@+ ( addr -- addr+1, c )**: fetches a 8bit value from address and raises the address with 1. This is equal to **COUNT** on most small systems with maximum string-lengths of 256 characters. Using **C@+** is an efficient way of going thru a character array.

**SC@+ ( addr -- addr+1, sc )**: fetches a 8bit value from address, extends the sign-bit to 32bit and raises the address with 1.

**H@+ ( addr -- addr+2, h )**: fetches a 16bit value from address and raises the address with 2.

**SH@+ ( addr -- addr+2, sh )**: fetches a 16bit value from address, extends the sign-bit to 32bit and raises the address with 2.

**@+ ( addr -- addr+4, n )**: fetches a value from address and raises the address with 4. This is the same as **COUNT** in wabiPi4. Using **@+** is an efficient way to go thru an array.

**2@+ ( addr -- addr+8, d )**: fetches a double value from address and raises the address with 8. Using **2@+** is an efficient way to go thru an array of doubles.

**Auto-indexing stores**

**C!+ ( c addr -- addr+1 )**: stores character **c** at **addr** and leaves **addr+1** on stack

**H!+ ( h addr -- addr+2 )**: stores halfword **h** at **addr** and leaves **addr+2** on stack

**!+ ( n addr -- addr+4 )**: stores value **n** at **addr** and leaves **addr+4** on stack

**2!+ ( d addr -- addr+8 )**: stores double **d** at **addr** and **addr+4**, and leaves **addr+8** on stack

**Overview of all available variants of fetch/store words:**

element	fetch	auto-indexing fetch	store	auto-indexing store	memory addition
char/byte	<b>C@</b>	<b>C@+</b>	<b>C!</b>	<b>C!+</b>	<b>C+!</b>
signed char/byte	<b>SC@</b>	<b>SC@+</b>			
halfword	<b>H@</b>	<b>H@+</b>	<b>H!</b>	<b>H!+</b>	<b>H+!</b>
signed halfword	<b>SH@</b>	<b>SH@+</b>			
word (cell)	<b>@</b>	<b>@+</b>	<b>!</b>	<b>!+</b>	<b>+!</b>
double	<b>2@</b>	<b>2@+</b>	<b>2!</b>	<b>2!+</b>	<b>D+2!</b>
mixed word and double					<b>M+2!</b>

## Some other words doing fetch and store

There are six words defined which do a very specific sort of fetch, namely 16, 24 and 32 bit BIG-ENDIAN fetch/store from and to an unaligned address. This can easily be programmed in Forth with **C@**, **C!** and **LSHIFT**. But as inlinable primitives they are 10-20 times faster. I found these words particularly useful with large scale simulations, where speed of execution is very relevant.

**16B@ ( un-addr -- u )**: fetches a 16bit unsigned number (ie. without extension of the sign-bit) from an unaligned address in BIG-endian fashion.

**24B@ ( un-addr -- u )**: fetches a 24bit unsigned number (ie. without extension of the sign-bit) from an unaligned address in BIG-endian fashion.

**32B@ ( un-addr -- u )**: fetches a 32bit number from an unaligned address in BIG-endian fashion.

**16B! ( 16b un-addr -- )**: stores a 16bit number BIG-endian in an unaligned address.

**24B! ( 24b un-addr -- )**: stores a 24bit number BIG-endian in an unaligned address.

**32B! ( 32b un-addr -- )**: stores a 32bit number BIG-endian in an unaligned address.

## FLOATING POINT FETCH AND STORE

There are also fetch/store words for floating point numbers. These will be discussed in future in the future chapter on floating point calculations and numbers.

These words are mostly already in the dictionary, you just have to find them... Hint: type '**awords f**'

## Comparisons

wabiPi4 contains all comparison operators from the ANSI standard. These are:  **$0<$ ,  $0>$ ,  $0=$ ,  $0<>$ ,  $<$ ,  $=$ ,  $<>$ ,  $>$ ,  $U<$ ,  $U>$ ,  $WITHIN$ ,  $D0<$ ,  $D0=$ ,  $D<$ ,  $D=$ ,  $DU<$ ,  $F0<$  and  $F0=$** . All function as specified in the standard.

The ANSI-standard operators form only a small subset of possible comparisons. For instance, the standard only describes 2 floating-point comparisons. Other floating-point comparison-operators have to be constructed by the programmer. This is at least less efficient than having these available in the dictionary. The more so as modern processors are especially efficient in this area.

So it is logical that wabiPi4 contains more operators than those contained in the standard.

All comparisons in wabiPi4 are inlinable primitives, and all are optimised for the Cortex-a72 processor in the Raspberry pi4b.

The following **additional comparison-operators** are contained:  
 $\leq$ ,  $\geq$ ,  $\leq\geq$ ,  $\geq\leq$ ,  $\odot\leq$ ,  $\odot\geq$ ,  $2^=$ ,  $EVEN?$ ,  $ODD?$ ,

**<< CHAPTER STILL TO BE COMPLETED >>**

## User Input

**User input** is via the serial line at **GPIO 14** and **15**, and technically is handled by the WAVE-loop running on core2. This loop polls the UART-hardware 44100 times per second for input. The advantage of this way of handling the input is that **core0** can run interrupt-free.

Received characters are put into a circular buffer with a size of 2 MB. This large buffer ensures that the system is reasonably resilient against buffer overruns. You can find definitions to manage the input-buffer below.

### Adherence to the ANSI standard:

The entry-system is the basis for the normal Forth user-input words. **KEY**, **KEY?**, **EXPECT**, **SPAN**, **>IN**, and **TIB** function as described in the ANSI-standard. **ACCEPT** works as described in the standard but has, as additional feature, the handling of 'esc'.

**The user can interrupt a running program** by pressing 'fn', 'ctrl' and 'backspace' at the same time. If this combination of keys is observed by **core2**, a fast interrupt request is send via de GIC (global interrupt controller) to **core0**, which then aborts the running program and puts the user back in control. This is very handy if for instance a program is stuck in an endless loop.

**ACCEPT** ( addr len\_max -- len\_string ): Accepts len\_max characters input from the keyboard. Entry ends by a 'return' or an 'esc'. On pressing 'esc' entry ends immediately and a 1 character string is returned with the value 27 as the first character.

The following additional definitions are available

**BACKSPACE** ( -- ): executes the sequence: **8 EMIT 32 EMIT 8 EMIT**

The following definitions are available to reset and get the status of the UART entry-buffer.

**UARTCLEAR** ( -- ): fills the UART entry-buffer with zeroes and resets the position of the read and write pointers

**CHARSINBUF** ( -- u ): returns the number of characters left in the input-buffer

**UARTRWRITE** ( -- addr ): variable – contains the position of the write-pointer of the UART entry-buffer. This variable is used by the entry-system and changing the value probably results in strange behaviour

**UARTRREAD** ( -- addr ): variable – contains the position of the read-pointer of the UART entry-buffer. This variable is used by the entry-system and changing the value likely results in strange behaviour

# User Output

## DOT ROUTINES

Numeric output is via the standard ANSI numeric output words: **.** (dot), **U.** (u-dot), **.R** (dot-R), **U.R** (u-dot-R), and the pictured number output words **<#** (less-number-sign), **#>** (number-sign-greater), **#** (number-sign), **#S** (number-sign-s), **HOLD**, **SIGN**. They all follow the standard.

In addition the following words are available. I use these so often that I included them in the permanent dictionary. Saves a bit of source code.

### Pictured numeric output

**-SIGN ( -- )**: use inside **<#** and **#>**. Places a 'minus\_sign' when the input number is negative, but does not place a sign when the number is zero or positive. This is the usual convention for most people.

### Variations of DOT words

**.BYTE ( c -- )**: prints the lower 8 bits of a value in the present **BASE**

**.SBYTE ( c -- )**: prints a signed byte. That is bit 7 is extended as sign-bit. For instance the value 130 is printed as -126, and the value 255 is printed as -1. Uses present **BASE**

**.BIT ( n -- )**: prints a zero if value equals zero, otherwise prints a one.

**.FLAG ( n -- )**: prints a zero if value equals zero, otherwise prints a minus one.

**.BIN ( n -- )**: prints in binary, independent of actual present **BASE**

**.8BIN ( n -- )**: prints the lower 8 bits of value n in binary, including leading zeroes.

**.16BIN ( n -- )**: prints the lower 16 bits of value n in binary, including leading zeroes.

**.32BIN ( n -- )**: prints the 32 bits of value n in binary, including leading zeroes.

**.64BIN ( d -- )**: prints the 64 bits of double value d in binary, including leading zeroes

**.64BIN\_ ( d -- )**: prints the 64 bits of double value d in binary, including leading zeroes, and without space after the printed number

**.HEX ( n -- )**: prints in hex, independent of actual present **BASE**

**.RHEX ( n r -- )**: prints in hex, right aligned with r characters

**.32HEX ( n -- )**: prints value n in hex, including leading zeroes.

**.64HEX ( n -- )**: prints double value d in hex, including leading zeroes.

**.64HEX\_ ( d -- ):** prints the 64 bits of double value d in hex, including leading zeroes, and without space after the printed number

**.DEC ( n -- ):** prints in decimals, independent of actual present BASE

**FIXED FLOAT printing**

**.1DEC ( n -- ):** prints value n in decimals, as if divided by 10. For instance the value 123 is printed as 12.3

**.2DEC ( n -- ):** prints value n in decimals, as if divided by 100. For instance the value 1234 is printed as 12.34

**.3DEC ( n -- ):** prints value n in decimals, as if divided by 1000. For instance the value 12345 is printed as 12.345

**.4DEC ( n -- ):** prints value n in decimals, as if divided by 10000. For instance the value 123456 is printed as 12.3456

## String handling

String handling is an important area of WabiPi4. Strings in wabiPi4 have a 32 bit index and thus are not limited to 255 characters. In addition, the handling of strings in wabiPi4 is fast, all primitives are written in optimised assembly.

The words **C@**, **C!**, **FILL**, **C**, (**c-comma**), **MOVE**, **CMOVE**, **CMOVE>**, **FILL**, **." (dot-quote)**, **S" (s-quote)**, **C" (c-quote)**, **SEARCH**, **COMPARE**, **/STRING (slash-string)**, **-TRAILING**, **CHAR**, **[CHAR]**, **CHAR+** and **CHARS** function as specified in the ANSI-standard, with the following exceptions:

- **SLITERAL** also works during interpretation
- the use of **DOUBLE-QUOTES** in the string definition words.

**In addition there is the ability to define and manipulate overflow protected string variables. See the next chapter for that topic.**

### SLITERAL

In wabiPi4 **SLITERAL** can also be used during interpretation. It is placed after a string defined with **S"..."** and is followed by a name. Calling the name during execution will put the address and length of the original string on stack, allowing direct typing or other string manipulations.

The ANSI standard explicitly disallows changing this string, however wabiPi4 does not guard against this. Complying with the standard is the responsibility of the user.

An example can be found below.

```
SLITERAL    interpreting: ( addr len ... <ccc> -- )
exec: ( -- addr len )
```

### DOUBLE-QUOTES

The words **."**, **S"** and **C"** (and **ABORT**) have an additional feature. In standard ANSI Forth the string contained in these words cannot contain the character **"'"** (double-quote). This character marks the end of the string, and so the first occurrence of a double-quote after these 4 words ends the string. In wabiPi4, but only in these 4 words, two double-quotes are interpreted as being one double-quote, which does not end the string.

Example of printing double quotes:

```
." ""Hello"""
will print:
"HELLO" (including the double quotes)
```

Example of the definition of a string containing double quotes:

```
S" cr ."" Hello "" 1 . cr 2 ." SLITERAL teststring
```

```
Once a string has a name it can be typed like this:
  teststring type
will print:
  cr ." Hello " 1 . cr 2 .
```

It can also be evaluated if it is valid wabiForth like this:

```
  teststring evaluate
will print:
  Hello 1
  2
```

As you see with the double double-quotes **EVALUATE** can also handle string-words correctly.

In addition the following definitions are available:

**>CAPS (char -- CHAR)**: converts a character in the range 'a'-'z' to 'A'-'Z'

**>LOWER (char -- CHAR)**: converts a character in the range 'A'-'Z' to 'a'-'z'

**\$>CAPS ( addr len -- )**: converts the string defined by addr and len to all capitals characters.

**\$>LOWER ( addr len -- )**: converts the string defined by addr and len to all lower case characters.

**STRING<> ( addr1 len1 addr2 len2 -- f/t )**: compares 2 strings and gives a true if the strings are not the same. Case-insensitive.

An example of the use of **\$>CAPS**.

```
s" a string without capitals" sliteral mystring
: show mystring type ;
: makecaps mystring $>caps ; \ << not ANSI compliant!!

show
makecaps
show
```

You will see that the string is first printed in lower case and then in upper case. The string is changed permanently. Using '**mystring \$>lower**' will change it back to all lower characters.

# Overflow protected string variables

**There are at least two problems with the strings** as defined in the ANSI standard. One is that the functionality seems limited compared to other programming languages. This is in reality not an huge problem, but anybody who has ever programmed in BASIC with MID\$ etc. will recognise this. The second problem is that there is no overflow-protection build in.

Using the ANSI standard words it is possible to create overflow-protected string-variables, and provide ample string-functionality. The following definitions are all done in hi-level wabiForth.

A **string-variable** consists structurally of 2 parts: a named definition, created in the dictionary, and a buffer where the actual string resides, created in the data-area of wabipi4. The data-area has ~2.8GB available space, and, consequently, a string-variable has a maximum length of ~2.8GB. If you consider that all works of Shakespeare taken together are around 5.5 million characters, it is clear that 2.8GB is a pretty long string.

These are the definitions:

## GENERAL FUNCTIONS

**\$VARIABLE ( max\_len \_\$VARIABLE\_ <name> )**: \$VARIABLE is used to define string-variables with a maximum length.

An example:

**1000 \$VARIABLE example\$**

defines a string variable named **example\$** with a maximum length of a 1000 characters, and allocates the buffer.

**A string-variable MUST be declared before it can be used in any of the following words.**

**GET\$ ( var\$ -- addr len )**: for a given string-variable, returns the address and length of the contained string.

**DEL\$ ( var\$ -- )**: deletes the given string-variable by setting the length to zero. The string itself is still visible with words like **DUMP** or **C@**. If you do not want that, use the word **CLEAN\$** after **DEL\$**.

**CLEAN\$ ( var\$ -- )**: fills the empty space of the given string-variable with zeroes. The string itself does not change, only the unused space in the buffer where the string resides.

**\$TYPE ( var\$ -- )**: prints the given string-variable. Only prints the first 2500 characters. Does the same as '**GET\$ 2500 MIN TYPE**'. Defining comparable words, if something different is wanted, is easy.

**ADR\$ ( var\$ -- addr ):** for a given string-variable returns the start address of the contained string.

**LEN\$ ( var\$ -- length ):** for a given string-variable, returns the length of the contained string.

**MAX\$ ( var\$ -- max ):** for a given string-variable, returns the maximum length the contained string can have.

## STRING MANIPULATION FUNCTIONS

These definitions change an existing string by adding to or deleting from a string variable.

**ADD\$ ( addr len var\$ -- ):** adds the string, defined by 'addr len' at the end of the string contained in the given string-variable.

If the available space in the buffer is shorter than the length of the string to be added, only the available space is filled. Thus ensuring that the maximum length is never overwritten. No error-messages are generated if this happens.

**CADD\$ ( char var\$ -- ):** adds a character at the end of the string contained in the given string-variable, provided there is space left in the buffer. This word can, for instance, be used to add ASCII-control characters to a string.

**OVERWRITE\$ ( addr len start var\$ -- ):** overwrites the string contained in the given sting-variable with the string defined by 'addr len' by overwriting from the point defined by 'start'. If the available space in the buffer is not enough to contain the whole of the string to be added, only the part which fits is written.

**INSERT\$ ( addr len start var\$ -- ):** inserts the given string into the string contained in the given string-variable. In any situation, the maximum length is kept.

**SET\$ ( addr len var\$ -- ):** sets the string in string-variable to the string defined by 'addr len'. Any existing string is removed from the string-variable.

**RSUB\$ ( len var\$ -- ):** removes 'len' number of characters from the right side of the string contained in the given string-variable. If 'len' is bigger than the length of the string contained in the string-variable, the whole string is deleted.

**LSUB\$ ( len var\$ -- ):** removes 'len' number of characters from the left side (ie the start) of the string contained in the given string-variable. If 'len' is bigger than the length of the string contained in the string-variable, the whole string is deleted.

**MIDSUB\$ ( start len var\$ -- ):** removes the part of the string contained in the given string-variable defined by the start and the length of the part contained. If the start-point lies outside of the string, nothing happens.

## STRING SELECTION DEFINITIONS:

These words select part of a string in a string-variable, but do not change the string. The programmer can use the returned addr and length to change the content of the string if so wanted.

**RSEL\$ ( len var\$ -- addr len )**: selects 'len' number of characters from the right side of the string contained in the given string-variable by returning the address and length of the selected part.

**LSEL\$ ( len var\$ -- addr len )**: selects 'len' number of characters from the left side of the string contained in the given string-variable by returning the address and length of the selected part.

**MIDSEL\$ ( start len var\$ -- addr len )**: selects 'len' number of characters from the starting point 'start' of the string contained in the given string-variable by returning the address and length of the selected part.

## EXAMPLES OF THE USE OF STRING VARIABLES

Adding a string to itself:

```
100 $VARIABLE STR$  \ creates an empty string-variable
s" This is example text..." str$ set$ \ add$ also functions

str$ $type
types: This is example text...

str$ get$ str$ add$ \ copies the string after itself
str$ $type
types: This is example text...This is example text...
```

# Memory-map and statistics of wabiPi4 system

**WabiPi4 supports a 4 or 8 GB Raspberry pi 4b.** For the 8 GB version only the lower 4 GB is accessible. Support for a 2 GB Raspberry pi 4b might be added in future. A 1 GB pi4b will not be supported.

Below is the system **memory map**.

item	start in memory	end	size
CPU exception vectors	0x00000000	0x00000020	32 B
SOC data-area	0x00000020	0x00008000	~32KB
wabiPi4 system	0x00008000	~0x00068514*	~385KB*
user-dictionary	~0x00068514*	0x3D800000	~983 MB
internal data and variables	0x3D800000	0x3E000000	8 MB
GPU-memory	0x3E000000	0x40000000	32 MB
PIXEL buffers	0x40000000	0x41000000	16 MB
task memory	0x41000000	0x43000000	32 MB**
hashtable	0x43000000	0x47000000	64 MB
data memory	0x47000000	0xFB000000	2880 MB
managed memory	0xFB000000	0xFC000000	16 MB**
SOC memory mapped registers	0xFC000000	0x00000000	64 MB

\* : depends on the version of the wabiPi4 system

\*\* : memory reserved but functionality not yet implemented

The following definitions return data on the free space in the main sections of the memory:

**UNUSED ( -- n )**: returns available memory in bytes available for new definitions in the dictionary – ANSI standard word

**MEMFREE ( -- u )**: returns the total amount of free memory in bytes available on the system

**DATAFREE ( -- u )**: returns the total amount of free data memory in bytes still available on the system. Both **ALLOCATE** and **TEMPALLOCATE** take their allocations from this pool.

**.AVSIZEWORD ( -- )**: prints the average number of opcodes (1 opcode is 4 bytes) in a word. The average is based on all words in the dictionary, system and user-defined.

It is mainly for informal purposes, it is not a reliable statistic. This is because allotted blocks of memory are also part of the dictionary. A slightly more accurate statistic can be reached by using **ALLOCATE** instead of **ALLOT** when defining blocks of memory. Allocated blocks of memory do not influence the calculation of **.AVSIZEWORD**.

The wabiPi4 system itself consists of 4 main parts:

1. **System-data:** a block of memory used by the system. It contains internal variables, constants and strings. It also contains the FORTH source-code which is loaded during the boot-process. This block starts at **0x8000** and stretches till the address contained in constant **STARTDICT**.

2. **Base-dictionary:** this is where the assembly code resides. Code for the following functionality is contained:

- the setup of the SOC/CPU
- initiation of the PIXEL graphic and WAVE sound system
- input system
- primitive Forth definitions
- interpreter
- optimising compiler

Parts 1 and 2 together form a primitive FORTH system able to run independently. The base-dictionary starts at the address contained in the constant **STARTDICT** and ends at **STARTHILEVEL**.

3. **HiLevel-dictionary.** This part of the dictionary is where wabiPi4 is defined in its final form. It adds useful functionality like drivers and tools. For instance the I2C drivers and most date and time related words are defined here. And the assembler is mostly defined here.

This part of the dictionary is compiled real-time during the booting-process using **LOADBOOT**. **LOADBOOT** loads the source code contained in System-data. Executing a **COLD** also reloads this dictionary. The hilevel-dictionary starts at the address contained in the constant **STARTHILEVEL** and ends at **ENDBOOT**. The source-code for the hi-level dictionary is contained in wabiPi4 in a readable format. It starts at the address returned by **SOURCEHILEVEL**.

4. **User-defined definitions.** Occupies the memory from **ENDBOOT** upto **HERE**.

The following words return statistics on the size of different parts of the wabiPi4 system.

**STARTDICT** ( -- u ): constant – returns address of first definition

**STARTHILEVEL** ( -- u ): constant – returns address of first definition of hi-level part of dictionary

**SOURCEHILEVEL** ( -- u ): constant – returns address of source-code of the hilevel dictionary

**ENDBOOT** ( -- u ): constant – returns address of first free memory-location after booting but before any definitions are added by the user.

**SZSYSDATA** ( -- u ): size of the internal block of memory.

**SZBASEDICT** ( -- u ): size of the base-dictionary.

**SZHILEVELDICT** ( -- u ): size of the hi-level dictionary.

**SZUSERDICT** ( -- u ): size of the user-part of the dictionary.

**SZSYSTOTAL** ( -- u ): size of total wabiPi4 system.

**SZDICT** ( -- u ): size of dictionary.

## Memory allocation

WabiPi4 largely follows the ANSI-standard with regard to memory allocation. The main difference to the standards are the words **FREE** and **RESIZE**.

Both **FREE** and **RESIZE** are CRC-protected, to reduce the chance of spurious data-erasure due to an error in a program. The words **FREE** and **RESIZE** will only take effect if you give the exact start-address of a block as input. Using a wrong address will result in an error-code on stack. A programmer can use the words as often as needed without fear of a memory leak.

### NON-STANDARD WORDS:

In addition to the standard words there are two words focused on temporary data. They are: **TEMPALLOCATE** and **FREETEMPDATA**.

**TEMPALLOCATE** is intended to allocate a buffer for data which is only needed on a temporary basis. **FREETEMPDATA** frees all memory allocated to temporary data.

The rationale is that with these words temporary and permanent buffers can be separated by the programmer. This makes it possible to clear the temporary buffers from memory once the contained data is no longer needed. This usually avoids complex and time-consuming garbage-collection.

**ALLOCATE** ( *u -- addr t/f* ): allocates a block of memory at least *u* bytes long in the data-area. The allocated block of memory will always be 64 byte aligned, to optimise cache-allocation. If the allocation was successful, a valid address and 'false' are returned. If there was not enough space, a true-flag will be returned with an invalid address.

Memory reserved with **ALLOCATE** can be faster than memory reserved with **ALLOT**. This is due to different cache-specifications for the memory. Memory in the data-area is flagged as 'EXECUTE-NEVER'. Because of this the default aggressive pre-loading of the instruction-cache is avoided.

For programs with a lot of data being accessed in a random or semi-random fashion, the difference in speed might be noticeable.

**ALLOCATE** works from high to low data-memory, and all empty data-memory space is available. A maximum of 2880 MB of memory can be allocated, in 1 block if so needed.

**FREE** ( *addr -- f/t* ): releases a block of memory previously defined by **ALLOCATE** or changed by **RESIZE**. **FREE** can be used on any block defined, but all blocks more recently allocated than the block of data being FREEed will also be FREEed.

**RESIZE** ( *addr - f/t* ): in wabiPi4 resizes a block of memory and moves the content to the new starting-position. If a block is reduced in size, data falling outside of the new size will be deleted without further warning.

Normally only the most recently defined block in memory can be resized. If you resize a block of memory which is not the lowest

(ie most recent), it will be the lowest after resizing. Words below the word being resized will be deleted without warning.

**TEMPALLOCATE** ( u -- addr/u, f/t ): allocates u bytes for temporary data. If u is larger than the available free memory, a true-flag is returned. If the allocation is successful, the address of the buffer and a 'false' flag are returned. The allocated block of memory will always be 64 byte aligned, to optimise cache-allocation.

**FREETEMPDATA** ( -- u ): frees all temporarily allocated buffer space in memory. Returns the actual number of bytes freed. This can be larger than the total size of the requested buffers due to the 64 bytes alignment of memory blocks.

**Example:**

```
: CHECK? abort" problem with memory-block" ;
```

**CHECK?** checks if a memory-block action was successful and aborts on an error.

```
1000000 ALLOCATE check? value MYBLOCK
```

Allocates a block of 1 million bytes and puts the address in value MYBLOCK.

```
20000000 myblock resize check? to myblock
```

Resizes the memory-block MYBLOCK from 1 to 20 million bytes

```
500000 myblock resize check? to myblock
```

Resize the memory-block again, but now to a smaller size

```
myblock FREE check?
```

Frees the memory occupied by the memory-block MYBLOCK

**#To Be Done:**

**BLCKSIZE?** ( addr - max\_size ) Returns the maximum size of the block at addr. If 'TRUE' is returned an error occurred.

**BLCKFILL** ( u addr - f/t ) Fill the memory block at addr with U. Return a flag with '**FALSE**' signifying succes and '**TRUE**' signifying an error.

```
BLCKCOPY ( addr u u addr u ), BLCK
```

## Wordlists (a.k.a. vocabularies)

The ANSI-standard introduced the more generic term '**WORDLIST**' for vocabularies.

WabiPi4 supports the following standard words: **WORDLIST**, **FORTH-WORDLIST**, **SEARCH-WORDLIST**, **FORTH**, **DEFINITIONS**, **ALSO**, **ONLY**, **PREVIOUS**, **ORDER**, **SET-ORDER**, **GET-ORDER**, **SET-CURRENT** and **GET-CURRENT**.

These words follow the ANSI-standard. Implementation-specific details are:

**PREVIOUS:** When an attempt is made to remove a word-list from an empty search-order list, PREVIOUS will be ignored and no error-condition occurs.

**ONLY:** According to the ANSI-standard the minimum search-order is implementation dependent. In wabiPi4 the minimum search-order includes all system-defined words. This is done by ensuring that FORTH is always searched at least once.

In addition the following definitions are available.

**VOCABULARY** ( <name> -- ) The word vocabulary defines a new vocabulary with the name <name> which follows the word

**VOCABULARY**. This word is not part of the ANSI standard, but in common use.

“It’s OK to figure out murder mysteries, but you shouldn’t need to figure out code. You should be able to read it.”  
– Steve McConnell

**HIDDEN** ( -- ) A definition with HIDDEN after it will not be printed with WORDS.

The combination of vocabularies and **HIDDEN** is very powerful. **HIDDEN** separates the functionality of 'do\_not\_print' from the coding functionality of a library. Using two vocabularies for this purpose can lead to very hard to find bugs, especially if the programs get larger.

**.VOC** ( -- ) Prints an overview of all vocabularies defined. It also prints vocabularies which are not in the search-order.

**Example:** Enter the following:

```
order
vocabulary NEWVOC
also newvoc
definitions
order
```

And you will see the following:

```
(0 0 0) order
definitions :
FORTH
search order:
FORTH ok
(0 0 0) vocabulary NEWVOC ok
(0 0 0) also newvoc ok
(0 0 0) definitions ok
```

```
(0 0 0) order
definitions :
    NEWVOC
search order:
    NEWVOC
    FORTH ok
```

**ORDER** shows the present search-order. Then **VOCABULARY** is used to make a new vocabulary with the name **NEJVOC**. **ALSO** **NEJVOC** adds **NEJVOC** to the top of the search-order list. **DEFINITIONS** ensures that new definition are added to the new vocabulary. Finally **ORDER** shows the final situation: 2 vocabularies and new definitions are compiled in **NEJVOC**.

**.VOCS** will show the two available vocabularies, including their wordlist-identifiers.

```
wid name
--- -----
1 NEWVOC
0 FORTH
--- -----
```

# Stacks

The ANSI standard contains 2 definitions giving information on a stack: **DEPTH & FDEPTH**. But there are several other related or common words with which the programmer can get information on the stacks and their status. Some of these are even mentioned and used in the ANSI standard. For instance **SP@** and **SP!** are used in the example sources for **CATCH** and **THROW** of the ANSI-standard, although they are not part of that standard.

This is the set of such words in wabiPi4:

**DEPTH ( -- n )**: puts depth of the data stack on stack as it was before the depth was put on stack

**FDEPTH ( -- n )**: puts the depth of the floating point stack on stack

**CPUDEPTH ( -- n )**: puts depth of the CPU return-stack on stack. WabiPi4 has two separate stacks for the return: the user-stack and the CPU-stack. This was done primarily to ensure efficient functionality of the branch-predictors of the CPU. As a side-effect, it also ensures that most user-mistakes with the return-stack do not result in a crash of the system.

**USDEPTH ( -- n )**: puts depth of the user return on stack. WabiPi4 has separated stacks for the return available to the user and for the CPU. **USDEPTH** returns the depth for the user-stack.

**SP@ ( -- addr )**: fetches the address pointed to by the stack pointer

**SP! ( addr -- )**: puts the address on stack in the data stack pointer

**RP@ ( -- addr )**: fetches the address pointed to by the return stack pointer

**RP! ( addr -- )**: puts the address on stack in the return stack pointer

**FP@ ( -- addr )**: fetches the address pointed to by the floats stack pointer

**FP! ( addr -- )**: puts the address on stack in the floats stack pointer

**R13@ ( -- addr )**: puts the content of CPU register 13 on stack. In the chapter '**Overview of Dictionary**' R13@ is used in an example.

**R13! ( addr -- )**: puts the address from the stack in CPU register 13

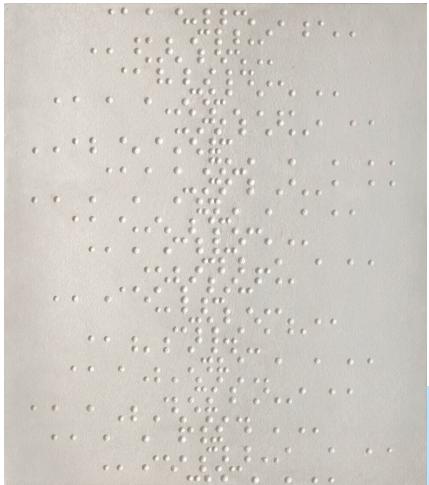
**R14@ ( -- addr )**: fetches the address contained in register 14 of the CPU. **This word can only be used in interpreting mode!** Register 14 is the link-register of the ARM-processor.

**[R14@] ( -- addr )**: fetches the address contained in register 14 of the CPU. **This word can only be used within a definition!** Register 14 is the link-register of the ARM-processor.

**R15@ ( -- addr ):** fetches the address contained in register 15 of the CPU. Register 15 is the program counter of the ARM-processor.

# Random numbers

Generating high quality random numbers reliably is surprisingly complex, and care has to be taken to avoid pitfalls. WabiPi4 has 3 different methods for the generation of random numbers. A very fast, deterministic generator of excellent quality called **RNDM**. And a slow non-deterministic method which generates true random numbers called **TRNG**.



GERHARD VON GRAEVENITZ

Weiße Struktur (Zufallsverteilung  
auf vertikaler Achse II)

1960

The wabiSystem also needs random numbers for system-management purposes, in some cases quit a lot of numbers. To avoid loosing the deterministic character of the main generator (**RNDM**), these numbers are generated by an independent 3rd generator called **SYSRNDM**.

**SYSRNDM** might be of use if a programmer needs a random number but wants to keep the state and order of the main **RNDM**-generator in tact.

**SYSRNDM** is a bit slower than the main method, and the quality of **RNDM**-generation is only fair.

These three methods should cover all needs, apart from serious cryptographic work.

## TRUE RANDOM NUMBERS

are generated by the iProc-RNG200 true random generator from Broadcom. The output cannot be predicted, there is no periodicity, and it is impossible to seed or reseed this generator. The generation of random numbers depends on the random jitter of ring-oscillators.

This True Random Number Generator (**TRNG**) conforms, according to Broadcom, to NIST SP800-90B, NIST SP800-22, FIPS 140-2, and BSI AIS-31. However, the wabiPi4-implementation has *NOT* been stress-tested.

The **TRNG** is slow. After a full reboot of the system, it takes 0.7s to generate the first random number. This delay guarantees that even the very first generated number is fully unpredictable. After the first number, a new 32bit random number is generated every ~106000c. The **TRNG** has a buffer to store up to three random-numbers for faster access. In case a number is available in the buffer, reading it takes around ~3350c. WabiPi4 handles this automatically.

Typically the **TRNG** is used to seed the other random generator once or periodically. The effect is fast generation of unpredictable random numbers.

### Sanity-check of TRNG

To comply with regulations governing random-number generation, such as BSI AIS-31, a hardware sanity-check of the generator must be performed before retrieving a random-value. WabiPi4 takes care of this.

If the sanity-check shows that an error within the **TRNG** generator was flagged by the hardware, a full reset of the TRNG-generator is performed and the **TRNG** error counter raised by 1. This count is available in the value **TRNGERRORCNT**.

Please note: the returned values are always unpredictable, also after a full reset of the **TRNG**. A reset of the TRNG does not reset seeds or anything comparable to that, as seeds do not form the basis of the random numbers generated.

## Definitions for the True Random Generator

**TRNG** ( *n -- m* ): returns a random number between 0 and *n* minus 1 in ~3350c. If the 3-value buffer is empty then it takes ~106000c to return a next random number.

A sanity-check is performed before a number is returned. In case the sanity-check fails, a hardware-reset of the **TRNG** is done to ensure correct functioning. In case of an error, **TRNGERRORCNT** is also raised by 1 to flag the failure.

**TRNG32** ( *-- u* ): Uses the same process as **TRNG**, but returns a 32bit random number.

**TRNGERRORCNT** ( *-- u* ): Value containing the number of times an error with the TRNG-generator was reported by the hardware since booting of the system. This should stay 0x0 at all times.

**RESETTRNG** ( *--* ): resets and (re)starts the true random generator (**TRNG**). It is called during start-up of the system. After a reset the generated numbers are still fully unpredictable. It takes 0.7 seconds after a reset before the first random number is available. There is no way of halting the TRNG-generator.

## THE MAIN RANDOM-METHOD

is based on the XOSHIRO-method as published by David Blackman and Sebastiano Vigna in 2018. The version used here is the 256 bit version with multiply-roll-multiply scrambling. It has a period of >1e77 and returns a random-number in 9c.

This method is intended as an efficient, reliable 'workhorse' function. It produces random values of the highest possible quality. Any combination of bits, both within a generated random value and from value to value, passes all known quality-tests for random-generators (incl. Diehard, Crush and BigCrush).

*Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin.*

—JOHN VON NEUMANN (1951)

how rare: the chance that you encounter the end of the solar-system in the next 24 hours, is larger than the chance that you will encounter such an unlikely sequence of numbers. And if you would encounter such an unlikely sequence, you would not notice it. There is simply no way of knowing that it happened.

The authors state that, based on their testing, this method can be used for any serious task. For any critical work you should test for suitability before use. It should also be noted that the method is not intended for serious cryptographic work.

The XOSHIRO-method has been put into public domain by David Blackman and Sebastiano Vigna, and can thus be used legally for any and all tasks.

The main words are **RNDM**, **TRNG** and **FRNDM**

**RNDM** ( n -- m ): returns a random number between 0 and n minus 1 in ~9c.

As the XOSHIRO generator resides permanently inside the NEON-unit of the ARM-processor it has no influence on – and is not slowed by – the state of the data-caches. In other words, dirty cache-lines do *not* slow the generator down, as would happen with memory based generators. This is especially beneficial in situations where data is read or written at a high rate and at random from or to a large block of memory.

**RNDM32** ( -- u ): returns a 32bit random number in ~9c. It uses the same generator as **RNDM**.

**RNDM** and **RNDM32** are not used by internal words or by the system. Their status only depends on the programmer.

**FRNDM** ( -- f ): returns a 32b floating point number between 0 and 1. This generator is also based on the fast method described above. As a first step a random 32b integer is generated, which is then converted to a floating number between 0 and 1. It returns a number in ~12c.

## RESEEDING THE RANDOM GENERATOR

The XOSHIRO algorithm uses 4 internal 64bit seeds to generate the random values, the sequence of random values is exclusively defined by these 4 seeds.

A quick way of setting seeds is using **RNDMSEED**. It takes a double value from stack and copies it into the 4 64 bit seeds and does 30 dummy runs to prime the random generator.

A more formal way to reseed the generator is with the word **RNDMTOSEED**. With this word the 4 seeds can be set to any value, and no dummy runs are done. If the present values of the seeds need to be preserved for later use, the word **GETSEEDX** copies the seeds to the stack.

**RNDMSEED** ( d -- ): uses a double value to load the 4 seeds and performs 30 dummy runs to prime the random generator. A quick and easy way of randomising to a new sequence. The double value cannot be zero.

**RNDMTOSEED** ( d\_s0 d\_s1 d\_s2 d\_s3 -- ) Copies the 4 doubles from the data-stack to the seeds in the XOSHIRO generator. Any 4 value can be used as long as at least 1 of the 4 seeds is non-zero.

If the seeds have a very low number of set bits, for instance if only 1 of the 4 numbers is non-zero, the first 10 to 20 random values generated might lack randomness.

**RNDMGETSEED** ( -- d\_s0 d\_s1 d\_s2 d\_s3 ): Copies the 4 64-bit seeds to the data-stack.

**.RNDMSEEDS** ( -- ): prints the 4 seeds.

The random generator can be reset to the initial state, as found after a reboot of the system, or randomised to a fully random new starting point.

**RNDMRESET ( -- )**: resets the seeds for the random generator to the situation as found after a reboot of the system.

**RANDOMIZE ( -- )**: sets the seeds of the random generator to four random values. The internal true random generator is the basis for this function, so predictability after **RANDOMIZE** is zero.

## JUMP GENERATION

In certain rare use-cases it is important to define 2 or more subsequences of random values which with 100% certainty do not overlap. For a linear method, like the XOSHIRO algorithm, it is possible to define a jump function. A jump function calculates what the value of the seeds would be after a certain number of random values have been generated, without actually generating all these numbers. The jump-function in wabiPi4 calculates how the seeds would be after the generation of  $2^{128}$  values. This allows for the creation of up to  $2^{128}$  non-overlapping subsequences.

The maximum rate of random value creation with wabiPi4 is around 150 million values/s. Generating  $2^{128}$  samples at this rate would take around  $7.2 \times 10^{22}$  years. The jump function saves time by performing this calculation in 5.8  $\mu$ s. A pretty nice example of a smart algorithm which is faster than brute-force calculation. In this case  $3.9 \times 10^{35}$  times faster.

**RNDMJUMP ( d\_s0 d\_s1 d\_s2 d\_s3 -- d\_s0 d\_s1 d\_s2 d\_s3 )**: calculates, based on the four 64 bit seeds as input, what the value of these seeds would be after generating  $2^{128}$  random samples. This allows for the definition of non-overlapping sequences of  $2^{128}$  samples each.

## THE SYSRNDM GENERATOR

This is a basic 32bit random generator with two 32bit seeds, following the **XorShift** principle of Marsaglia, and using the optimised symmetrical setup as published by professor Brent. The period is  $1.844 \times 10^{19}$ . If you generate random samples at the maximum speed rate, 120 million/sec, this generator returns to the starting point after 4871 years.

As stated previously, this generator is used by the system when it needs random-numbers. The consequence is that the exact state of this generator is unknown. And so experiments using this generator are not guaranteed to be repeatable.

If you need a quick random number though, while using the main RNDM generator for something big, like a massive simulation, this generator comes in very handy.

**SYSRNDM ( n -- rndm )**: returns a random number between 0 and n-1 in 10-13 cycles. This generator is memory-based and thus suffers

from performance-degradation if the amount of dirtying of cache-lines is high. Performance in those cases can be significantly slower.

**SYSRNDM32 ( -- rndm ):** returns a 32bit random number in 10-13 cycles.

The **sys-random generator** can be reset to the initial state, as found after a reboot of the system, or randomised to a new starting point.

**SYSRNDMRESET ( -- ):** resets the sys-random generator to the situation after a **BOOT**.

**SYSRANDOMIZE ( -- ):** loads the two seeds of the sys-random generator with values from the true random generator.

# System constants, variables and values

The wabiSystem depends on a range of variables, values, constants and tables for a correct flow of actions and in general to keep track of stuff. A lot of these are available to the programmer. For instance if you are interested in the internals of the system, or if you want to add new functionality.

The list here shows just some...

**LAST ( -- addr )**: variable, contains the address of the link-field of the most recent word in the dictionary.

**LATESTWORD ( -- addr )**: value, contains the address of the name of the most recent defined word. 'C!+ TYPE' prints this name.

**DEFINING? ( -- n )**: value, contains a flag which, if true, signifies that the system is compiling a definition. This is NOT the same as '**STATE**'. '**STATE**' can temporarily be put on hold with '[' whereas **DEFINING?** tracks if a colon was closed by a semi-colon or not.

**COLDCOUNT ( -- value )**: value, returns the number of times a COLD was executed by the system since the last reboot. It is used internally for the sanity-checks on the boot-screen.

**TRANS ( -- addr )**: constant, contains base-address of buffer for trans-area (a scratch-buffer which temporarily holds strings during compilation).

**TRANSNUM ( -- addr )**: constant, contains base-address of buffer for nummer-conversion.

**IBADDR ( -- addr )**: variable, contains the presently used input-buffer address. Usually points to the Text Input Buffer (TIB), but this is changed during the evaluation of text-files.

**LASTXTHERE ( -- addr )**: variable, contains the first location of the last XT compiled. Used during optimisation.

**WORD\_LIST\_COUNT ( -- addr )**: variable, tracks overall number of wordlists defined - cannot be larger than 255.

**SEARCH\_ORDER\_TABLE ( -- addr )**: constant, start-address of the table with the search\_order lists.

**FREENOCACHEMEM ( -- addr )**: constant, start-address of the unoccupied part of a uncached memory-block.

**NOCACHEDATA ( -- addr )**: variable, pointer into no\_cached memory block.

**FREECOHERENTMEM ( -- addr )**: constant, contains start-address of the unoccupied part of a coherent memory-block.

**COHERENTDATA ( -- addr )**: variable, pointer into coherent memory block.

**FREEUNSECUREMEM ( -- addr )**: constant, start-address of the free/available part of an unsecure memory-block. Unsecure is a term related to the ARM-processor. Programs requiring the HYPER-state of the ARM-processor can only run in, and read/write data from/to unsecure memory.

**UNSECUREDATA ( -- addr )**: variable, pointer into available unsecure memory. Some experimental data, for potential future functionality, resides here.

**UARTREC ( -- value )**: constant, start of UART receive buffer.

**UARTRWRITE ( -- addr )**: variable located in coherent memory-block, contains the location within the UART receive buffer where the next character received will be stored in the buffer.

**UARTRREAD ( -- addr )**: variable located in the coherent memory-block, contains the location within the UART receive buffer where the next character will be read from the buffer.

**UARTRECLEN ( -- value )**: constant, contains length of the UART receive buffer.

**GIC\_BASE ( -- value )**: constant, contains the base-address of the Global Interrupt Controller (GIC). See the ARM documentation for more details.

**FONT1 ( -- addr )**: constant, contains address of the font-table

## Time and time-measurement related words

WabiPi4 contains two 64bit counters. One, the micro-second counter, counts up with a frequency of 1 MHz. The other is the CPU-cycle counter, which counts up with a frequency of 1.8 GHz. Both start counting from zero after a full reboot of the system. The CPU-cycles counter starts counting as soon as the CPU is powered up, whereas the micro-second counter starts counting as soon as wabiPi4 is started.

The micro-second counter is intended as a general purpose time-measurement tool. The CPU-cycle counter is especially useful when optimising the performance of algorithms, as it counts the actual number of CPU-cycles needed to execute an individual routine.

### THE MICRO-SECOND COUNTER

is accessed with the word **DMCS**. **DMCS** returns the time in microseconds since the last reboot of the system, as a double. The underlying counter in fact counts at 5.6448 MHz to get better rounding-properties for **DMCS** (and also because 5.6448 MHz divided by 128 gives 44100, a number relevant for sound-generation). In theory it takes more than 103.000 year before this underlying counter wraps around to zero. But according to ARM, this counter only needs to be implemented in hardware with 56 bits. Which gives a shortest wrap\_to\_zero time of something like 400 years. Shorter than 103.000 year but still not a problem in most situations.

A related word is **MCS**, which does the same as **DMCS** but returns a 32bit value.

**DMCS ( -- d )**: returns the elapsed time in microseconds since the last reboot, as a double. It does so by dividing the underlying counter with 5.6448 to get at microseconds.

**MCS ( -- u )**: returns the elapsed time in microseconds since the last reboot, as an unsigned single. As **MCS** wraps to zero after 4295 seconds, care should be taken that this does not impede on interval calculations.

**MS? ( -- u )**: returns the elapsed time in milliseconds since the last reboot, as an unsigned single. **MS?** wraps to zero after 49.7 days.

**Please note the question mark!** **MS** is a ANSI-standard word which waits for a given number of milliseconds.

**MS ( n - )**: wait n milliseconds. A relic from the olden days, when computers struggled measuring periods shorter than milliseconds. But part of the ANSI-standard.

**CENTIS ( -- u )**: returns the elapsed time in centi-seconds since the last reboot, as an unsigned single. CENTIS wraps to zero after 497 days.

**SEC ( -- u )**: returns the time in seconds since the restart of the system, as an unsigned single. **SEC** wraps around to zero after ~136 years.

**DMCS**, **MCS**, **MS?**, **CENTIS** and **SEC** are non-inlinable primitives, and take ~12 cycles to put an answer on the stack.

## THE CPU-CYCLE COUNTER

is accessed with the word **CPUCYCLES**. It starts counting at 0x0 after power-up of the CPU and after **RELOAD**. The counter counts the actual clock-cycles used by CPU-core0. Which in concreto means that this counter counts up at 1.8 GHz. If a core is temporarily halted, for instance using WFE or WFI, this counter also stops counting. If the frequency of the CPU changes, this clock will count at that changed frequency.

It is important to realise that the measured number of cycles will vary greatly. This variability reflects reality. **CPUCYCLES** counts the actual number of cycles. The state of the different caches, the state of the pipelines and other aspects influence the number of cycles actually required.

If you want to measure the fastest possible time of a routine use '()' and ')'. If, on the other hand, you are interested in the actual time taken, use **CPUCYCLES**.

**CPUCYCLES ( -- d )**: puts the elapsed number of CPU-cycles since the last reboot on stack as a double. It wraps around to 0x0 in about 325 years.

Executing **CPUCYCLES** twice directly after each other will show a difference of (usually) 9 to 13 cycles. These 9 to 13 cycles are the time required to put the values on the datastack.

Subtracting 9 to 13 cycles from a measurement will give you the time taken by whatever is between two **CPUCYCLES**. Which makes clear that averaging is needed to get really reliable measurements.

## EXAMPLES:

```
: MINIMUM cpucycles cpucycles 2swap d- d. ;
: SHOWMIN 10 0 do cr minimum loop ;
```

Executing **SHOWMIN** shows ten times the difference from two consecutively executed **CPUCYCLES**. The most common time shown is 12 cycles. By running **MINIMUM** ten times you can sometimes see the effect of the cache-system and the branch-prediction system.

Contrary to the Cortex-a53 in the Raspberry Pi3b+, the timing of the Cortex-a72 in the Raspberry 4b is called 'deterministic' by the ARM-cooperation. That is the timing of individual tasks is predictable.

That this predictability is not always self-evident shows the following example:

```
: HOWLONG cpucycles 1 1 + drop cpucycles 2swap d- d. ;
```

Here '**1 1 + drop**' is added to the line of code in the '**MINIMUM**' definition above. The result printed can be anything from 9 cycles to around 1900 or so.

The interesting thing is that the most common timing returned is 9 cycles. Why this longer line of code usually runs faster than '**MINIMUM**' above is unknown to me.

## OTHER WORDS FOR TIME-MEASUREMENTS:

**.UPTIME ( -- )**: shows how long the system has been running since the last start/restart of wabiPi4. It uses the format `00h00m00s`. When programmed correctly, wabiPi4 is a very stable system, uptimes of thousands of hours are entirely possible. **.UPTIME** can track the uptime for at least ~400 years.

## Interval measurements

For measuring an time-interval the following words are available:

**T[ ( -- dmcs )**: puts the present time in micro-second ( $\mu s$ ) as a double on stack

**]T ( dmcs -- dmcs )**: calculates the difference in  $\mu s$  with the last **T[**

**]T. ( dmcs -- )**: calculates the difference in  $\mu s$  with the last **T[** and prints it in a formatted way

For instance:

```
: TST 100000 0 do i drop loop ;
: go t[ tst ]t. ;
go
```

prints "278 us" or something like that. In other words: putting I on stack and dropping it again one hundred thousand times takes around 278 micro-seconds. But it can also take 223 or 389  $\mu s$ . That depends on the state of the CPU.

**C[ ( -- dcpucycles )**: puts the present number of cpu-cycles as a double on stack

**]C ( dcpucycles -- dcpucycles )**: calculates the difference in cycles with the last **C[**

**]C. ( dcpucycles -- )**: calculates the difference in cycles with the last **C[** and prints it in a formatted way

For instance:

```
: TST 100000 0 do i drop loop ;
: go c[ tst ]c. ;
go
```

prints "501082 cycles" or something like that. (but it sometimes prints a number around 400000 or 700000 cycles)

If you experiment with **CPUCYCLES** you will see that measurements with **CPUCYCLES** have a high variability. That variability reflects the reality of modern CPUs, which are very fast on average but unpredictable for the performance of individual pieces of code.

### Measurement of number of cycles

If you want to measure pretty exact how long the execution of a word or short snippet of Forth-code takes at its fastest you can use the words `((` and `)`.

The measurements are done by running the Forth-code between (( and )) a million times and measuring the elapsed time to calculate the number of cycles per loop. As the code is called a million times, the cache-system, the branch predictor and the pipe-line system all function optimally. This ensures that the maximum performance of the CPU is measured, like measuring the top-speed of a car.

In normal software, especially when the rate of dirtying of cache-lines is high, the actual speed is slower, sometimes a lot slower.

Please note that the word or forth-code snippet being measured must be stack-neutral. And that (( and )) can only be used inside a compiled word.

```
(( ( -- mcs ): Initiates the measurement and 1 million loops
)) ( mcs -- ): takes the time from (( from stack, calculates the
number of cycles used and prints this in a formatted way.
```

For instance:

```
: TST 10 0 do loop ;
: HOWMANYCYCLES (( tst )) ;
```

Executing **HOWMANYCYCLES** prints something like:  
**17779 us -> 32.0 cycles**

The 17780 µs reported is the time (in micro-second) it takes to call the word **TST** a million times. The 32.0 cycles reported is the number of CPU-cycles it takes on average to call and execute the word **TST** 1 time.

In other words, calling **TST** one time takes on average ~17.8 nanoseconds. To make clear how short a time that is: light travels only 5.3 meters in 17.8 nanoseconds. In that very short time wabiPi4 jumps to a subroutine, does 10 loops and returns. Which I find truly awe-inspiring!

## Defining a timer

The following code-example defines a timer called **TIMERA** which measures microseconds. As many individual timers as needed can be set up this way. Note that the timing of the two different versions is very comparable.

```
2variable timera
: SETTIMERA ( -- ) dmcs timera 2! ; ( 17.0c )
: GETTIMERA ( -- d ) dmcs timera 2@ d- ; ( 17.7c )
```

or like this, using the Wabi-specific 2VALUE:

```
0 0 2value timera
: SETTIMERA ( -- ) dmcs to timera ; ( 17.7c )
: GETTIMERA ( -- d ) dmcs timera d- ; ( 17.0 )
```

**SETTIMERA ( -- ):** gets a baseline, which functionally is af if timer A was reset to zero.

**GETTIMERA ( -- d ):** returns the time in microseconds since timera was last reset, as a double.

The following code shows how this timer can be used. It measures how long it takes to do 10000 empty loops using do...loop:

```
: HOWLONG settimera 10000 0 do loop gettimera d. ;
```

At the time of writing it takes at least 11 microseconds ( i.e. 11 millionth of a second ) to do ten thousand empty do...loops with wabiPi4. Longer times are also possible.

## Pausing execution

Pausing execution for a certain time can be done using the following words:

**WAITMCS ( n -- ):** waits for n microseconds before continuing with execution. Pausing is done by constantly polling the elapsed time, and thus is reasonable accurate.

**WAIT ( -- ):** waits for 1 second before continuing execution

**BLINK ( -- ):** waits for 100 milliseconds before continuing execution. This is the time of a fast blink of the eye, although some sources give a longer duration like 330 milliseconds as the duration of a blink.

**10MS ( -- ):** waits for 10 milliseconds before continuing execution

**1MS ( -- ):** waits for 1 millisecond before continuing execution

**MS ( n -- ):** waits n milliseconds before continuing execution

## Not so serious stuff

The Cortex-A72 processor is a member of a long line of ARM-processors with an equally long history. One or two of those remnants of this long history are still visible. Features that were once useful but are now outdated. It is my intention to fully support these features, if only out of respect for all the brilliant people who developed these processors.

### **Scratch-byte**

In case the 4067920432 (~4 billion) bytes of free memory available in wabiPi4 are not enough, the CPU contains 1 whole extra byte of scratch-storage in the UART-section.

To make the use of this extra byte as convenient as possible, wabiPi4 contains 2 words: **C!SCRATCH** and **C@SCRATCH**. With these you can store a byte for later use, and fetch it again.

Contrary to the pi3b+, on the pi4b the access-time of this byte-field is not exceptionally long, namely ~15 cycles (around 220 cycles for the pi3b+). Normal memory is still at least 10 times faster. But if you really need that 1 more byte, you'll take all you can get...

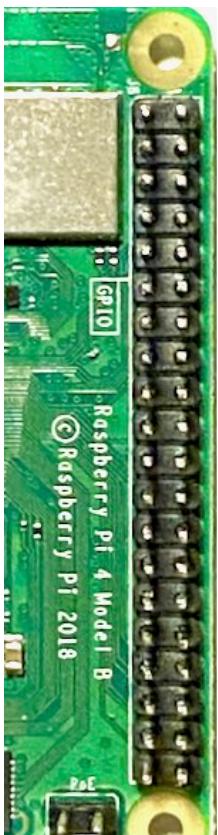
To the cynics who just have to notice that these two definitions consume 160 or so bytes of the available RAM; and that it would have been 160 times more efficient to just leave these words out; I can only say: 'wow, how boring you are...'.

**C!SCRATCH** ( c -- ): stores a character in the scratch-byte

**C@SCRATCH** ( -- c ): fetches a character from the scratch-byte

## GPIO

The Raspberry Pi4b board contains a 40 pin connector for IO (input/output). Twenty-six of these are IO pins available to the user. The usual term for these is GPIO (General Purpose In/Output), as they have more functions than just IO. For the most common actions wabiPi4 offers specific words. Full control of the GPIOs is possible via the CPU-registers.



The picture shows you that the GPIO-numbers are not the same as the pin-numbers. All wabiPi4 words always use the GPIO-numbers.

3.3v	1	2	5v
gpio2	3	4	5v
gpio3	5	6	0v
gpio4	7	8	gpio14
0v	9	10	gpio15
gpio17	11	12	gpio18
gpio27	13	14	0v
gpio22	15	16	gpio23
3.3v	17	18	gpio24
gpio10	19	20	0v
gpio9	21	22	gpio25
gpio11	23	24	gpio8
0v	25	26	gpio7
HAT id	27	28	HAT id
gpio5	29	30	0v
gpio6	31	32	gpio12
gpio13	33	34	0v
gpio19	35	36	gpio16
gpio26	37	38	gpio20
0v	39	40	gpio21

Pin 27 and 28 are used by the Raspberry for HAT-identification (HAT=Hardware Attached on Top), and are not available to the user; even if no HAT is present.

All other GPIO's can be in- or output. In addition most GPIOs support other functions. On the next page there is an overview.

Pins 2 and 4 are the 5.0v pins. Connecting these with any other pin, no matter how short, might (and probably will) destroy the CPU!

You can actually use one of these 2 pins to connect an external 5.0v power-supply. This way of supplying power is called back-feeding and it works fine. It is wise to cover the 5.0v pins if they are not in use.

**Use of GPIO-pins in wabiPi4:** wabiPi4 has default settings for 6 GPIO's. The UART1 on pin 8 and 10 is critical for communication with wabiPi4, the functionality should not be changed. The other pins (I2C and PWM) can be changed to other functions.

pin	GPIO	function
pin 8	gpio 14	UART1 – TX
pin 10	gpio 15	UART1 – RCV
pin 3	gpio 2	I2C – SDA1
pin 5	gpio 3	I2C – SCL1
pin 32	gpio 12	PWM – output synthesizer
pin 33	gpio 13	PWM – output synthesizer

**WabiPi4 contains definitions** for setting **ALT functions** of a GPIO, for setting and reading a GPIO and for the pull-up and pull-down resistors. Interrupts related to the GPIOs are not covered by **wabiPi4** itself. This might be added in future as functionality, but don't count on it...

**SETFUNCGPIO ( n gpio# -- ):** sets the alternate function of port gpio# to function n. Valid inputs are between 0 and 7. When 0 is the input, a pin as configured input, 1 configures a gpio as output. This is true for all GPIO's. The other options are pin-specific. See the ARM-documentation for full details on the other special functions.

Please be careful as the special functions-values have a strange order. These are the inputs for the different functions and ALT functions:

value for input	function
0	input
1	output
4	ALT0
5	ALT1
6	ALT2
7	ALT3
3	ALT4
2	ALT5

The Pi4b has a lot more I0-hardware on board than the pi3b+. Some of the alternate functions are rarely of practical use for most programmers. Examples like the liquid crystal bus, alternate databus and JTAG interface. An overview without these alternate functions makes more sense for most programmer.

#### Simplified alternate functions Raspberry Pi 4B

ALT5	ALT4	ALT3	ALTO	GPIO	pin	pin	GPIO	ALTO	ALT3	ALT4	ALT5
				3.3v	01	02	5.0v				
SDA3	CTS2	SPI3_MOSI	SDA1	gpio2	03	04	5.0v				
SCL3	RTS2	SPI3_SCLK	SCL1	gpio3	05	06	0v				
SDA3	TXD3	SPI4_CE0	GPCLK0	gpio4	07	08	gpio14	TXD0	SPI5_MOSI	CTS5	TXD1
				0v	09	10	gpio15	RXD0	SPI5_SCLK	RTS5	RXD1
RTS1	SPI1_CE1	RTS0		gpio17	11	12	gpio18	PCM_CLK	SPI6_CE0	SPI1_CE0	PWM0_0
SPI6_CE1				gpio27	13	14	0v				
SDA6				gpio22	15	16	gpio23				SCL6
				3.3v	17	18	gpio24				SPI3_CE1
SDA5	CTS4	BSCSL_MOSI	SPI0_MOSI	gpio10	19	20	0v				
SCL4	RXD4	BSCSL_SDA/MISO	SPI0_MISO	gpio9	21	22	gpio25				SPI4_CE1
SCL5	RTS4	BSCSL_CL/SCLK	SPI0_SCLK	gpio11	23	24	gpio8	SPI0_CE0	BSCSL_CE	TXD4	SDA4
				0v	25	26	gpio7	SPI0_CE1	SPI4_SCLK	RTS3	SCL4
				HAT id	27	28	HAT id				
SCL3	RXD3	SPI4_MISO	GPCLK1	gpio5	29	30	0v				
SDA4	CTS3	SPI4_MOSI	GPCLK2	gpio6	31	32	gpio12	PWM0_0	SPI5_CE0	TXD5	SDA5
SCL5	RXD5	SPI5_MISO	PWM0_1	gpio13	33	34	0v				
PWM0_1	SPI1_MISO	SPI6_MISO	PCM_FS	gpio19	35	36	gpio16	CTS0	SPI1_CE2	CTS1	
SPI5_CE1				gpio26	37	38	gpio20	PCM_DIN	SPI6_MOSI	SPI1_MOSI	GPCLK0
				0v	39	40	gpio21	PCM_DOUT	SPI6_SCLK	SPI1_SCLK	GPCLK1

## WABIPI4 WORDS TO SETUP THE GPIO FUNCTIONS

**SETGPOUT ( true/false gpio# -- )**: sets port gpio# to 'on' or 'off'. The valid range of port-numbers is 0–53. Other values are ignored. The maximum frequency which can be reached is ~19.5 MHz.

This word only has effect for GPIO-pins configured as output. But the value is stored. If you write 'true' as output to a pin, the pin will go high as soon as it is configured as output.

Please note that the ARM-processor can only source around 2mA per pin! This is much lower than for instance an Arduino. So in most cases you will need a buffer to drive a LED or other effectuators.

**GETGPIN ( gpio# -- 0/1 )**: gets the input-value at port gpio#. The maximum frequency with which ports can be read is ~75 MHz.

**GETGPIN** gets the actual value of the pin, also if the pin is configured as output. This can be handy when debugging a routine.

The inputs are all sampled, that is the Raspberry does several reads of a port and then a majority-vote determines what level to return. This is more reliable than reading a level once, but takes longer.

**SETPULLUD ( action gpio# -- )**: controls the use of pull-up/pull-down resistors setting the GPIO specified by gpio# to action. The pull-up/down feature only has effect for GPIO-pins configured as input. Valid gpio# are 0–53, valid input for action is 0–2. If input is invalid than **SETPULLUD** doesn't do anything.

The following table shows what input gives which action.

value for input	input
0	no resistors
1	pull-up
2	pull-down
3	reserved

Please note that the logic for pull-up and pull-down is reversed compared to the values for the Pi3!

It might be interesting to note that the pull-up/pull-down resistors can be used as high-impedance outputs. In that case the pin needs to be configured as INPUT. And then, by setting the pull-up/down resistors to high or low, the pin in acts high-impedance output with clearly defined levels.

**On the Raspberry Pi 4B** it is possible to set the drive-strength, slew-rate and hysteresis of the GPIOs. It is not a feature which is well documented.

Setting values is not done per GPIO, but per group of GPIOs. There are 3 groups. Group 0 contains GPIOs 0–27, group 1 contains GPIOs 28–45 and group 2 contains GPIOs 46–53. The settings are always the same for all GPIOs within a group.

There are three options available for each group:

- |                   |            |
|-------------------|------------|
| 1: drive-strength | values 0–7 |
| 2: slew-rate      | values 0–1 |
| 3: hysteresis     | values 0–1 |

**The drive strength** specifies at which current (measured in mA) the ports in a group are guaranteed to reach correct output-voltages. The values 0–7 correspond to 1–8mA. The default setting is 8mA.

**Please NOTE:** this sets the drive-strength, it does NOT set a limit on current.

The maximum current flowing is only limited by the internal and external resistance. The current should never be allowed to go higher than 8mA per GPIO.

**The hysteresis** can be 0=disabled or 1=enabled. This value only has effect for GPIOs used as input.

**The slew-rate** can be 0=limited (=slower) or 1=unlimited (=faster)

#### **Coding of the three options:**

The value for drive-strength is coded into bits 0–2, the hysteresis is bit 3, and the slew-rate is bit 4. The default value is hexadecimal 0x1F for group 0 and 1, and (for unknown reasons) 0x1E for group 2. 0x1F maximises drive-strength (8mA), enables hysteresis and sets unlimited slew-rate. 0x1E is the same but specifies 7mA instead of 8mA as the drive-strength.

**PADSPECS ( value group -- ):** controls the settings of the GPIO-PADS.

Bit 2:0 of value contain the drive-strength, bit 3 is the hysteresis enable-bit and bit 4 the slew-rate control bit. The value of any other bit is ignored.

Valid input for group are 0, 1 or 2. Values higher as 2 are interpreted as being group 2.

## Of bytes, bits and bobs...

Any Forth makes it easy to handle individual bytes and bits with words like **AND** and **OR**. However, assembly is sometimes much faster. Especially if there are specific opcodes for a task. WabiPi4 contains a set of primitive words to benefit from the ARM-opcodes for byte and bit handling. Some of these words are over 50 times faster than the corresponding Forth code.

### BYTE & BIT REVERSE & REFLECT

focus on handling specific situations. For instance the word **BYTEREV**. It reverses the order of the 4 bytes in a 32b word. Programmed in forth, the fastest routine takes 53 cycles to do the reversal. The ARM-processor has a specific opcode for this task. It can usually reverse the bytes in effectively 1 cycle.

**BYTEREV ( n -- n )**: reverses the order of the 4 bytes in a 32b value. Example: 0xAABBCCDD becomes 0xDDCCBBAA.

To see for yourself, try:

```
HEX AABBCCDD byterev u. \ note the U.
```

-> DDCCBBAA is printed

**REFLECT ( n -- reflected\_n )**: reverses the order of the 32 bits in the value on top. So the lowest bit becomes the highest and the lowest becomes the highest bit etc. This is not the same as **BYTEREV**!

To see for yourself, try:

```
HEX AABBCCDD reflect u. \ note the U.
```

-> BB33DD55 is printed

**CREFLECT ( char -- reflected\_char )**: reverses the order of the 8 bits of a character.

**4CREFLECT ( n -- reflected\_4chars )**: reverses the order of the 8 bits of each of the four characters within a 32b value. The 4 characters each stay on their position within the 32b value.

**HREFLECT ( 16b -- reflected\_16b )**: reverses the order of the 16 bits of a 16b value.

**2HREFLECT ( n -- reflected\_2\_16b )**: reverses the order of the 16 bits of each of the two 16b values within a 32b value. The 2 values stay on their position within the 32b value.

## BIT COUNTING

There are three words concerning the counting of bits: **POPCOUNT**, **CNTLZERO** and **CNTTZERO**. Short for population count, count\_leading\_zeroes and count\_trailing\_zeroes.

**POPCOUNT** simply counts the number of set bits in a value on stack. It is very handy to determine the hamming distance of two values. In that case the two values are first XORed with each other, a popcount of the result then gives the hamming-distance.

**POPCOUNT** ( n -- count ): for a given number counts the number of set bits. This is a non-inlinable primitive based on a NEON routine. Returns a count in 3 cycles.

**CNTLZERO** ( n -- count ): for a given number counts the number of zeroes till the first set bit. This is a inlinable primitive and it returns a value in 1 cycles (and sometimes in 0.5 or even 0.0 cycles).

**CNTTZERO** ( n -- count ): for a given number counts the number of zeroes after the last set bit. This is a inlinable primitive and it returns a value in 1 cycle.

## BIT ROTATE RIGHT AND ROTATE LEFT

Most processors offer some kind of bit-rotate for a value. A rotate right of a value means that all bits are shifted 1 or more places to the right and that the bits which are dropped from the right, are introduced again on the left.

Such a rotate forms, for instance, the core of a lot of cryptographic functions. It is easy to program in hi-level forth, but very much faster in assembly.

WabiPi4 supports the following definitions for bitwise rotates:

**BROR** ( n m -- n (rolled right by m bits )): rotates value n right with m bits. Rotates are 32b based.

**DBROR** ( d m -- d (rolled right by m bits )): rotates double d right with m bits. Rotates are 64b based.

**BROL** ( n m -- n (rolled left by m bits )): rotates value n left with m bits. Rotates are 32b based.

**DBROL** ( d m -- d (rolled left by m bits )): rotates double d left with m bits. Rotates are 64b based.

All bit-rotates are optimised primitives. The single rotate operations **BROR** and **BROL** are inlinable, and return a value in ~1.5 cycles. The double rotates return a value in ~7.5 cycles.

Please do not confuse these operations with **ROLL**, a ANSI standard definition where a single value n positions deep on the stack is rolled to the top of the stack.

## BIT FETCH & STORE

focus on setting/clearing/retrieving or counting bits. Usually at a given address or in a bit-array. For instance with the words **BIT@** and **BIT!**.

The word **FLAG@** gets the value of a bit from a given memory address and than converts this value of the bit (0 or 1) to a

flag-value (0 or -1). Very easy to do in Forth with **0<>**, but faster in assembly. There is no need for a **FLAG!** as **BIT!** handles both bits and flags correctly.

The word **NEGFLAG@** does the same as **FLAG@** but with a negative logic. So if the bit is zero, the flag is true, and if the bit is 1, the flag returned is false,

The words **1BIT!** and **0BIT!** do the same as **BIT!** but 3 cycles faster.

**BIT@** ( bit# addr -- 0/1 ): gets the value of the bit specified by bit# from the value contained in addr. Returns a 0 or 1.

**FLAG@** ( bit# addr -- false/true ): as **BIT@** but returns FALSE or TRUE.

**NEGFLAG@** ( bit# addr -- true/false ): as **FLAG@** but with reversed logic

**BIT!** ( 0/non-zero bit# addr -- ): stores a 0 (=clears) or 1 (=sets) in the bit specified by bit# from the value contained in addr. Also works with FALSE/TRUE as input.

**1BIT!** ( bit# addr -- ): sets the bit specified by bit# from the value contained in addr. Faster execution as '1 addr bit!'.

**0BIT!** ( bit# addr -- ): clears the bit specified by bit# from the value contained in addr. Faster execution as '0 addr bit!'.

### Comparing execution time

Using the following code-snippets you get an impression of how much faster the optimised primitives are.

```
: forthbit@ ( bit# addr -- 0/1 )
  @          \ ( bit# value )
  swap       \ ( value bit# )
  rshift     \ ( value>>bit# )
  1 and ;   \ ( 0/1 )

: timeforth (( 23 1000000 forthbit@ drop )) ;
: timeassembly (( 23 1000000 bit@ drop )) ;

timeforth
timeassembly
```

If you run the code you will see that a forth-based **BIT@** takes ~15 cycles and the primitive **BIT@** takes ~3 cycles.

For **FLAG@** the difference is a bit larger, 17 cycles in wabiPi4 and 3c as primitive. Avoiding the conversion saves some cycles.

## BIT-ARRAYS

The words **BIT[]@** and **BIT[]!** support the use of an array of bit-flags spanning across a range of memory. With these words the specified bit to fetch or store can be any number. For instance, if you specified bit 32 with these words, it would address the first bit in the address 4 bytes higher than the base-address. Bit 33 is the second bit in the base address+4 bytes. Bit 64 is the first bit of the value at base address+8 bytes. The full range of an unsigned number can be used, so a bit-array can contain up to  $2^{32}$  bits, corresponding to a bit-array of 128 MB of memory.

The words **FLAG[]@**, **NEGFLAG[]@**, **FLIPBIT[]**, **1BIT[]!** and **0BIT[]!** are also available.

**BIT[]@** ( *u addr -- 0/1* ): Gets the value of the *n*-th bit from the bit-array starting at base-address *addr*. The value *u* is a 32 bit unsigned value. The maximum size of a bit-array is thus 128MB. This word is a inlinable primitive and executes in 2.4 cycles, which is surprisingly fast (as are the following words)

**FLAG[]@** ( *u addr -- false/true* ): as **BIT[]@**, but returns FALSE or TRUE.

**NEGFLAG[]@** ( *u addr -- true/false* ): as **FLAG[]@** but with the opposite logic.

**BIT[]!** ( *0/non-zero u addr --* ): stores a 0 (=clears) or 1 (=sets) in the bit specified by *n* from the value contained in *addr*. The value *n* can be anything 32 bit unsigned value. Also works with FALSE/TRUE as input. This word is a non-inlinable primitive which takes 7 cycles to complete. As do the following two words.

**1BIT[]!** ( *u addr --* ): the same as **1BIT[]**, but for bit-arrays up to 128 MB long. Somewhat faster than '**1 u addr BIT[]!**'

**0BIT[]!** ( *u addr --* ): the same as **0BIT[]**, but for bit-arrays up to 128 MB long. Somewhat faster than '**0 u addr BIT[]!**'

**FLIPBIT[]** ( *u addr --* ): flips bit '*u*' in the bit-array starting at address from 0 to 1 or from 1 to 0. **FLIPBIT[]** can for instance be useful in error-correction routines.

## FIELD INSERT

Especially when setting up registers for peripherals it is very common that the programmer has to set/clear certain bits in a register without changing the other bits of the register. If it is an individual bit then the above mentioned words can be used. If it concerns a field of bits within an address, the word **FIELDINS** might be handy.

**FIELDINS** inserts a value in a bitfield in a value contained in an address.

For example:

```
17 5 8 myaddr FIELDINS
```

will insert the value 17 in the field starting at bit 8 with a length of 5 bits in the memory-location **myaddr**. The other bits in **myaddr** are not changed.

**FIELDINS** ( **u** **len** **start** **addr** **--** ): inserts the value **U** in the bit-field defined starting at bit **start** and with a length **len** at the memory-address **addr**. The other bits are left unchanged. Using a value of **0** resets the bits in the bit-field to **0**.

## **BOBS**

**BOBS** ( **? --** ): seriously, don't try this one...

## Cyclic Redundancy Check ( a.k.a. CRC )

The ARMv8 processor in the Wabi system has specific opcodes for the generation of CRCs. Generating CRCs this way is a lot faster than software-based algorithms.

Where relevant, the definitions can use either the polynomial **0x04C11DB7** (=ISO/IEEE) or **0x1EDC6F41** (as published by Castagnoli) for the generation of the CRC. The Castagnoli variant is denoted by an extra 'C' in the name. The Castagnoli version is especially good at longer messages. The IEEE version is in general very good and excels with shorter messages.

The following words are available:

**CRC8B** ( `CRC_previous, byte -- CRC_updated` ): Generates a new 32bit CRC based on the previous CRC and a byte as input.  
Using **CRC8B** on the 4 bytes of a word generates the same CRC as using **CRC32B** on that word.

**CRC8CB** ( `CRC_previous, byte -- CRC_updated` ): the same as **CRC8B** but using the Castagnoli polynomial.

**CRC32B** ( `CRC_previous, n -- CRC_updated` ): Based on the previous CRC value and a word as input generates a new 32bit CRC.

**CRC32CB** ( `CRC_previous, n -- CRC_updated` ): the same as **CRC32B** but using the Castagnoli polynomial.

**DATACRC32** ( `CRC_previous, addr_c, len -- CRC_updated` ): Based on a previous 32bit CRC and a data-set starting at `addr_c` and with a length of `len`, generates a 32bit CRC.

This word uses the data as is. This means that a string starting with a capital character results in a different CRC than the same string starting with a small character. If the generated CRC has to be independent of the case of the characters used, use **\$CRC32** instead of **DATACRC32**.

**DATACRC32C** ( `CRC_previous, addr_c, len -- CRC_updated` ): the same as **DATACRC32** but uses the Castagnoli polynomial.

**FASTCRC32[]** ( `CRC_previous, addr, len_words -- CRC_updated` ): does the same as **DATACRC32** but only on 32b words. Especially for short sections of data where a CRC is needed rapidly this word is faster. This is for instance relevant with some error-correction routines where millions of CRC-codes are generated while correcting errors.

**FASTCRC32C[]** ( `CRC_previous, addr, len_words -- CRC_updated` ): does the same as **FASTCRC32[]** but uses the Castagnoli polynomial.

**\$CRC32** ( `CRC_previous`, `addr_c`, `len` -- `CRC_updated` ): Based on a previous 32bit CRC and a string starting at `addr` and with a length of `len`, generates a case-independent 32bit CRC. **\$CRC32** is 4-5 times slower than **DATA\_CRC32**. This word works by converting all lower-case characters to upper-case characters before the generation of the CRC.

**\$CRC32C** ( `CRC_previous`, `addr_c`, `len` -- `CRC_updated` ): the same as **\$CRC32** but using the Castagnoli polynomial.

**CRCXMODEM** ( `byte_in` -- ): Generates a 16 bit CRC as used by the XMODEM-protocol.

Before generating a CRC for one XMODEM block, the variable **CRCXM** must be set to zero. After this the individual bytes of the XMODEM block are fed to this word. After all bytes of 1 block for the XMODEM protocol have been included, the variable **CRCXM** contains the 16bit CRC.

**CRCXM:** ( -- `addr` ): variable – used in combination with **CRCXMODEM**.

## WAVE – sound generator

### **\*\* THIS CHAPTER IS UNDER DEVELOPMENT \*\***

As the sound generator has a lot of options, defining these and describing these clearly and in detail is a challenge.

And as the pi4 has more hardware options than the pi3, there might be changes to the sound-system. The introduction of a 3rd channel specially for bass-generation is envisioned.

### **\*\* THIS CHAPTER IS UNDER DEVELOPMENT \*\***

**The Wabi-system contains a flexible sound-generation system.** Core 2 of the CPU is dedicated permanently to this. The system generates digital sound at 44.1 KHz and has 2 channels. These can either be used as left and right channel, or as 1 for sounds with a low pitch (ie. the 'bass'-channel), and 1 for sounds with a higher pitch. This depends on the implementation of the reconstruction-filters connected to the outputs of the Raspberry.

On the original home computers, sound-generators had a fixed number of oscillators, 1 or more noise-generators and a fixed number of registers to define frequencies, wave-forms, inter-connections and other such options.

The wabi-system functions in a different way. There is no fixed number of oscillators or other elements. The user has the possibility to define whatever number of oscillator and other elements is needed. Obviously within the technical limits of the ARM processor. An estimate is that around a 1200 elements can be defined at the same time. For instance using 100 oscillators in parallel to generate a sound is perfectly fine. Silly, obviously. But it will work.

**The way WAVE functions** is that the user defines sound-elements, and links these elements together. The defined elements then generate or manipulate the sound.

A sound-element is created by creating a control-block. Such a control-block is nothing more than 8 to 12 words of data describing what an individual sound-element should do. So for instance for a sinus-generator there is a word specifying what frequency the generator should generate.

All the control-blocks are put after each other in a list (called the WAVE-list). The WAVE-sound generator uses this information to generate the specified sound/sounds.

You surely will know the phrase "sheer endless possibilities". Well that is especially true here. A guesstimate of the number of possible different settings for WAVE is a 1 followed by 21250 zeroes (give or take a zero...).

**The WAVE sound-generator is a prototype** and testing will take a lot of time. But the system is already usable. And what is clear is that it is an exiting tool with a broad spectrum of functionality. As with any new tool, changes and additions can happen in the future.

A **quick-start example** is probably the easiest way to show the principles of WAVE. The code:

```
resetwv
440 ~sine constant osc1
osc1 osc1 ~output drop
startwv
```

will sound a 440 HZ sinus on both channels. If you want to stop the sound use STOPWV. STARTWV will start it again.

Five words are used here:

**RESETWV** ( -- ): resets the WAVE generator, empties the WAVE-list and stops all output. Is used before a new setup of WAVE is done.

**STARTWV** ( -- ): starts the WAVE generator. It is usually done after setup, but can also be done earlier to hear the effect of individual setup-steps.

**STOPWV** ( -- ): halts the WAVE generator but does not reset it. Executing STARTWV will restart it from where it was stopped.

**~SINE** ( F -- \*out ): adds a sine-oscillator to the WAVE-list. F specifies the wanted frequency. \*OUT is the address where the output of the oscillator is available. With ~SINE frequencies between 1 Hz – 20KHz can be generated. By putting the address (ie.: \*OUT) in a named constant, this output-address is available as input for other elements.

**~OUTPUT** ( ch0 ch1 -- \*out ): adds an output-element to the WAVE-list. At least 1 output element is needed to hear output. But the system itself will happily generate wave-data without an output module. You just don't hear anything.

## The oscillators:

**There are presently 11 different oscillators** available: 3 sawtooth oscillators, 2 block-wave oscillators, a sine-oscillator, 2 arbitrary wave generators and finally 3 noise-generators. With the exception of 2 noise-generators, all these generators are frequency based. That is, their main defining characteristic, next to the wave-form, is a base-frequency.

The 9 frequency-based oscillators have some properties in common: Each has a base frequency. They have one output each. The phase of an oscillator can be controlled realtime. The frequency can be fixed or be defined by an external source, for instance a sequencer or from wabiPi4. And finally all have technically a very wide range of frequencies which can be generated. Technically, the upper range is more than 1 MHz. And on the opposite side, the maximum period a wave can have, is around 27 hours.

In addition, some oscillators have specific functionality which are explained later.

## The sawtooth oscillators:

**There are 3 sawtooth oscillators.** The base saw, the reverse saw and the universal saw.

**The base sawtooth** produces a sawtooth formed wave, where the wave slowly rises from minimum to maximum and then rapidly descents to minimum again. In other words, the top of the wave is at the end of the wave.

**The reverse sawtooth** also produces a sawtooth formed wave but with a reversed profile: it rapidly ascend to maximum and then slowly descends to minimum. In other words, the top of the wave is at the start.

**The universal sawtooth** oscillator can produce a wave-form with the top at a user-specified position. This allows the sound to be more mellow, or sharper, or anything in-between.

**~BSAW ( F -- \*out ):** Creates a base sawtooth oscillator, oscillating at frequency F

**~RSAW ( F -- \*out ):** Creates a reversed sawtooth oscillator, oscillating at frequency F

**~USAW ( F top -- \*out ):** Creates a sawtooth oscillator, oscillating at frequency F and where the top is between 0 and 100% of the width of the wave. Top is a signed 32b number where lowest negative value indicates a top at the start of the wave and highest positive value indicates a top at the end of the wave. An easy way of getting correct numbers is by using '**~%**'. It will calculate the position using a percentage -> 50 ~% will put the top in the middle.

For example:

```
440 50 ~% ~usaw constant saw1 \ create sawtooth-oscillator
    saw1 saw1 ~output drop           \ create output element
```

produces a sawtooth oscillator with the top at 50% of the width of the wave, and an output element with the same saw-wave oscillator as input in both of its inputs.

## The block-wave oscillators:

There are two block wave oscillators:

**The base block oscillator** produces a block-form wave with a fixed 50/50% ratio. It produces a lot of nasty harmonics, and sounds rather unpleasant at any frequency and even worse at low frequencies. But the sound is very recognisable, as it was used a lot on early home-computers. So if you are working on a vintage game, this is the oscillator you would consider first. Listen to a few beeps of this oscillator and the early eighties spring back in mind forcefully.

**The universal block wave oscillator** produces a wave with a user-definable ratio. Any ratio between 0 and 100% can be chosen by the user.

**~BBLOCK ( F -- \*out ):** Creates a base block oscillator with a fixed 50/50 ratio and oscillating at frequency F

**~UBLOCK ( F ratio -- \*out ):** Creates a universal block oscillator with a user-definable ratio and oscillating at

frequency F. The ratio is signed integer number between minimum and maximum integer value. Zero gives a ratio of 50%. The '**~%**' word can be used (see '**~USAW**')

## The sine oscillator:

**There is 1 sine generator** as there is only 1 sine. It produces a very nice mellow sound, but some find it a bit boring.

**~SINUS ( F -- \*out )**: Creates a base sine oscillator, oscillating at frequency F

## The noise generators:

There are three noise generators. They produce slightly different sounding noise. In reality the difference between the sound produced is hardly audible.

**~FNOISE ( F -- \*out )**: produces a noise which changes value randomly at the given frequency F. Some base-frequency might be audible. This generator has the same common features as the above mentioned oscillators. Thus the frequency can also be controlled by an external source and phase-control is available. Why you would need phase-control for a noise-generator is unclear, but it is there for you to play with if you feel the need to do so.

**~BNOISE ( cycle -- \*out )**: produces a basic noise which randomly changes it's output from max-integer to min-integer or vv. at the specified interval of cycles. The sound of this generator is pretty harsh.

**~SNOISE ( cycle -- \*out )**: produces a slightly softer noise which changes it's output to a random value at the specified interval of cycles. This generator is comparable to the ~FNOISE generator, but with a different time-base.

# PIXEL

PIXEL is under development and more importantly, under consideration.

I have the feeling, subjective no doubt, that the present resolution of 1024x768 and the 24bit colours does not reflect the spirit of the kind of computing I want to do. So it might be that in future both the resolution and the amount of colours are reduced.

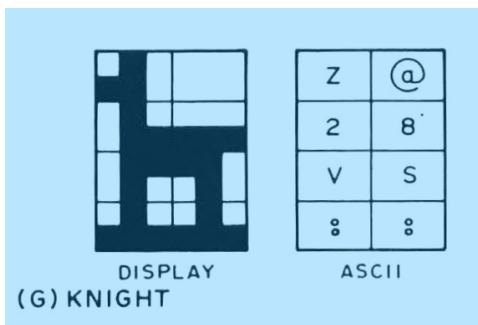
I have experimented with resolutions of 704\*528, 768\*576, 720\*540, 800\*600 and 848\*480. These are resolutions which can be shown on TVs and monitors very well. And the Raspberry Pi has the option of using a 256 color table. This would give 8-bit colours from the full 24 bit range. These options have a better vibe than the present, very bland, 1024\*758 24bit colour system. The performance of PIXEL would be higher, while at the same time reducing the needed memory-bandwidth. And one could playing with the colour-table for fast animation. Yeahh!!

## Overview:

The wabiPi4 system contains an experimental and very basic graphical system called PIXEL. Not as basic as the drawing of a knight for a chess-game on the left, but basic nonetheless. The setup of PIXEL tries to follow the way the graphical systems of the home-computers of the mid-eighties worked.

All the management tasks for PIXEL are provided by a dedicated core of the ARM-processor, the Raspberry-GPU is not used. The big advantage is that the user has full control over every pixel being drawn.

**PEEKing** and **POKEing** of pixels is very much a feature of wabiPi4. Developing the latest 3D-shooter is probably not feasible.



## THE PIXEL WINDOW-SYSTEM

WabiPi4 contains a basic windowing-system. Compared to modern standards, and even compared to Windows 3.1, it is very primitive. But it is fun to play with and it has a couple of unique features.

A **maximum of 8 windows** can be defined in wabiPi4. The order of visible windows is fixed, window 0 is the lowest and the other windows are visible in order of numbering.

**Window 0 is special.** It is always visible and it always has the same size as the screen and thus fills the background completely. Any other visible window floats on top of window 0.

A number of routines are specifically made for window 0 to make use of the fact that size and memory location of window 0 are fixed, and thus using window 0 can be faster.

It is intended that window 0 will be the standard output for system and error-messages but this is still under development.

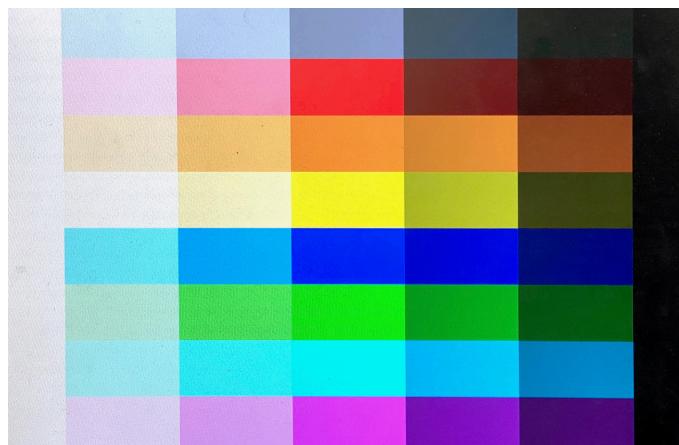
## COLOUR-SYSTEM:

WabiPi4 uses 24 of the available 32 bits for defining colours, in the normal RGB order. The 4th byte is the alpha-byte, which on the Raspberry does not function. WabiPi4 can used it to store information on the screen.

### Predefined colours:

The following colours have been pre-defined as CONSTANTS:

black	white					
gray	dgray	vdgray	lgray	vlgray		
red	dred	vdred	lred	vlred		
orange	dorange	vdorange	lorange	vlorange		
yellow	dyellow	vyellow	lyellow	vlyellow		
blue	dblue	vdblue	lblue	vlblue		
green	dgreen	vdgreen	lgreen	vlgreen		
cyan	dcyan	vdcyan	lcyan	vlcyan		
magenta	dmagenta	vdmagenta	lmagenta	vlmagenta		



In addition there is the very special color green, defined by my daughter Lize:  
`greenlo`

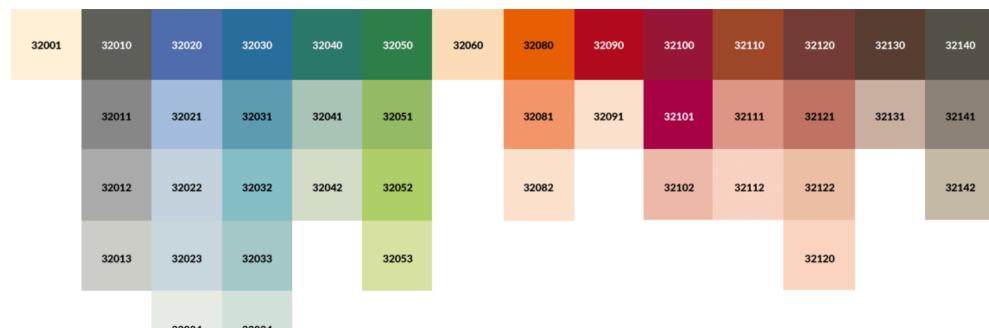
### The Corbusier colour palettes:

The 1931 and 1959 colour palettes of Le Corbusier have also been pre-defined as CONSTANTS. You need a good monitor to see them as they are intended. All colours match when combined.

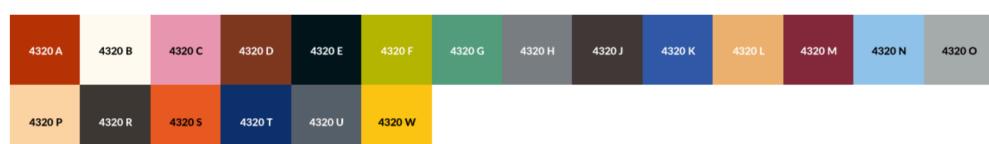
32001_blanc	32010_grisfon31
32011_gris31	32012_grismoy
32013_grisclr31	32020_bleouotr31
32021_outrmoy	32022_outrclr
32023_outrpal	32024_outrgris
32030_bleuceru31	32031_ceruvif
32032_cerumoy	32033_ceruclr
32034_cerupal	32041_vertanglclr
32042_vertanglpal	32050_vertfon
32051_vert31	32052_vertclr

32053_vertjauneclr	32060_ocre
32080_orange	32081_orangeclr
32082_orangepal	32090_rouverm31
32091_rosepal	32100_roucar
32101_rourub	32102_roseclr
32110_ocrerou	32111_ocreroumoy
32112_ocrerouclr	32120_tersnnbru31
32121_tersnnbrq	32122_tersnnclre31
32123_tersnnpal	32130_terombbru31
32131_ombbrucle	32140_ombnat31
32141_ombratmoy	32142_ombratclre
4320A_rouverm59	4320B_blancivo
4320C_rosevif	4320D_tersnnbru59
4320E_noirivo	4320F_vertolvif
4320G_vert59	4320H_gris59
4320J_terdombburu59	4320K_bleuoutr59
4320L_ocrejauneclr	4320M_lerubis
4320N_bleuceru59	4320O_grisclr59
4320P_tersnnclre59	4320R_ombnat59
4320S_orangevif	4320T_bleuoutrfon
4320U_grisfon59	4320W_jaune

The colour palette of 1931



The colour palette of 1959



## WORDS AVAILABLE FOR PIXEL

### Control of PIXEL:

**PIXELWFE ( =VALUE )**: the PIXEL-system functions by aggregating all windows into one screen-buffer, which is then copied to the actual screen memory. Once this is done, PIXEL waits a short period. The length of this waiting-period can be controlled by the user with the value **PIXELWFE**. The default value is 800 and gives a refresh-rate of around 77 frames/sec. (this is the rate with which the windows are updated internally, this is not the rate screen are send to the monitor)

The advantage of making the refresh-rate slower is that this makes more memory-bandwidth available to other applications. A

memory-intensive task will usually go a bit faster if the value in **PIXELWFE** is raised.

## DRAWING:

**WINPIXEL** ( *x y win# --* ): draws a pixel at *x*, *y* of window *win#* using the **INK** setting of *win#*. If *x,y* are out of bounds, nothing is changed. If *x,y* point to a section of a window outside of the screen, the pixel is drawn in the window.

**WIN0PIXEL** ( *x y --* ): draws a pixel at *x*, *y* in window 0 using the **INK**-setting of window 0. If *x* or *y* are out of bounds, no changes are made. This routine is upto ~3\* faster than the general pixel-draw routine **WINPIXEL**.

**WINLINE** ( *x0 y0 x1 y1 win# --* ): draws a line from *x0,y0* to *x1,y1* on window *win#* using the **INK** setting of the window. The routine draws 1 pixel if the length of the line is zero. This is an implementation of the Bresenham algorithm.

**HORLINE** ( *x y len win# --* ): draws a horizontal line starting at *x,y* with a length *len* using the **INK** setting. The routine does not draw anything if the *len=0*. Drawing a horizontal line is ~5 times faster than drawing a line with **WINLINE**.

**DRAW4x4** ( *addr x y win# --* ): draws a 4x4 box using the pattern stored at *addr*. The pattern is stored as 16 32bit color constants. The ranges of X and Y are from 0-255 and from 0-192 respectively. This is an optimised assembly routine, and as such upto 20x faster than an equivalent in wabiPi4.

**DRAW2x2** ( *x y win# --* ): draws a 2x2 box using the **INK** setting of the window. Range *x=0-511 y=0-383*.

**WIPE2x2** ( *x y win# --* ): draws a 2x2 box using the **CANVAS** setting of the window *win#*. Range *x=0-511, range y = 0-383*.

**DRAWBOX** ( *x y lenx leny win# --* ): draws a filled rectangle on window *win#*, using the **INK**-setting of window *win#*. The rectangle starts at *x,y* and has a width of *lenx* and a height of *leny*. It is valid for the box to be partly or completely outside of the window.

## WORDS TO MAKE AND RESIZE A WINDOW:

**MAKEWIN** ( *maxx maxy win# --* ): creates a new window with size specified by *maxx* and *maxy*. A window can be resize afterwards at will but cannot be bigger than *maxx* and *maxy*. A new window uses the present values for the variables **INK** and **CANVAS**. A newly defined window is not visible. This allows for flicker-free changes to the content.

**>WINSIZE** ( *x y win# --* ): sets the new size of a window. The size can not be larger than the maximum size set during definition of a new window. Changing the size of a window invalidates the content of a window but a **WINCLEAR** has to be performed by the programmer.

**WINGETSIZE** ( *win# -- x y* ): for window *win#* gets the size in pixels.

## WORDS RELATED TO THE HANDLING OF WINDOWS

**WINCHAR** ( char x y win# — ): writes character char at position x,y on window win# using the INK and CANVAS settings for the window.

**WINHOME** ( win# — ): puts the non-visible text-cursor at position 0,0 in window win#. The window is not cleared.

**WINTAB** ( position win# — ): moves the text-cursor on the present line to position by printing spaces. If the cursor is already past the position, nothing happens. Normal the ANSI word TAB is used, which calls WINTAB.

**WINCLEAR** ( win# — ): clears window win# using the CANVAS-setting of the window. Does not clear the border of a window.

**WINCURSOR** ( win# — x y ): for window win# returns the present position of the text-cursor.

**WINVISIBLE** ( n win# — ): sets the visible flag to n. If 'false', window win# is not visible. If flag is greater than 0, window win# is visible.

A wabiPi4 specific feature is that that this flag counts down at the refresh-rate. This means that ANY window is only temporarily visible with the exception of window 0. Setting this flag to -1 results in the maximum period of visibility, which is around 10 months since the setting of the flag. For most applications this is long enough.

**>WINRAND** ( pixel win# — ): sets the width of the border of a window. The maximum width is 32 pixels. The border is not part of the active window. For instance text never prints in the border and a WINCLEAR does not clear the border.

Whether you want to use a border is up to you and your taste. But a nice fat red border functions well as a warning!

**>WINORIG** ( x y win# — ): sets the origin of a window. X and Y are the location of the upper left corner of a window, and changing them moves the window. It is valid for a window to be partly or completely outside the screen. This does not change the content or functionality of a window.

**WINSCROLL** ( p win# — ): scrolls window win# up by p pixels, clears the area at the bottom with the canvas color of the window. If p is larger than the height of the window, the window is just cleared.

**>WINNOSCROLL** ( scroll-disable-flag win# — ): sets or clears the window scroll disable flag. The content of a window will not scroll up if the flag is set to 'true'. Default-value for a window is 'false'. Scrolling up normally happens when printing text in a window.

## WORDS TO SET AND GET INK AND CANVAS VALUES

**>WININK** ( ink win# — ): sets the INK setting for window win#. Any 24 bit number can be used.

**>WIN0INK** ( ink -- ): sets the **INK** setting for window 0 to the given ink. Any 24 bit number can be used. See the **PIXEL** intro text for the details of **INK**. This word is ~3\* faster then the general **>WININK**.

**WINGETINK** ( win# -- ink ): gets the present **INK** setting for window win#

**>WINCANVAS** ( canvas win# -- ): sets the **CANVAS** (=background color) setting for window#. **CANVAS** and **INK** have the same properties, they are 24bit numbers.

**WINGETCANVAS** ( win# -- canvas ): gets the present **CANVAS** setting for window win#.

## WORDS TO CONTROL WHERE TO PRINT

**WIN#>TASK#** ( win# task# -- ): sets the window on which a task prints. Presently only task 0 is active. But task 0 can be made to print on any of the 8 windows. The effect of the word is to direct the output of words like **EMIT** to a given window. A task can only print to 1 window at a time. By default task 0 prints to window 0.

**UART>TASK#** ( flag task# -- ): sets or clears the UART print-disable flag for a given task#. Presently only task 0 is active. Setting the flag to true results in the disabling of printing for a given task. Disabling the UART-printing can save time if you print a lot to the screen.

## WORDS RELATED TO THE HANDLING OF THE ALPHA-BYTE

**GETWIN0ALFA** ( x y - (false) or (addr value true)): (notice the spelling of the word ALFA!) gets the value of alpha of the pixel at x,y of window 0. If the x or y are out of bounds a false is returned. Otherwise the address of the pixel, the alpha-value and a true are returned. The address allows for fast writing of a new alpha-value without recalculating the address.

For an explanation of alpha and why that is available see the introduction text of **PIXEL**.

**SETWIN0ALFA** ( alpha x y -- ): sets the alpha-value of the pixel at x,y. If out of bounds, nothing is changed. The alpha-value is a 8 bit value.

## WINDOWS INFO-TABLE

**PIXEL** contains all control-data for each window in an array of 8\*132 bytes. For each of the 8 windows a list of values is stored. Not all items are presently functional.

Please take note that the info-table is mainly ment for getting data on a window. For changing of values usually it is better to use the specific words for that. Changing values directly might have unexpected results.

For instance changing the size of a window should be done with **>WINSIZE** and not by writing the new x and y values directly in this table. The same is true for **>WINRAND**.

variable	offset	comments
pxmaxx	0	maximum width window – window can be smaller but never larger
pxmaxy	4	maximum height window – window can be smaller but never larger
pxorigx	8	location window relative to screen – can be outside of screen
pxorigy	12	ditto
pxsizex	16	current size window in x-direction
pxsizey	20	ditto
pxcurx	24	position graphical cursor in x-direction
pxcury	28	ditto
pxbuffer	32	pointer to memory-buffer of window – filled during the creation of a window
pxsizebuf	36	size of the memory-buffer – filled during the creation of a window
pxbackgr	40	not in use
pxpitch	44	number of bytes per line of pixels of a window – allows fast fill of pxwopitch
pxmask	48	not in use
pxsizemask	52	not in use
pxtext	56	not in use
pxsizetxt	60	not in use
pxcurstxt	64	not in use
pxrand	68	size border around window – can be from 0 (default) to 32 pixels wide – the border is an area which will not be overwritten by graphical commands
pxink	72	ink for this window – start with off-white for win0 – other windows start with the value contained in the variable INK
pxcanvas	76	canvas color for this window – win0 starts with very dark cyan – other windows start with the value contained in the variable CANVAS
pxfont	80	pointer to default-font for a window – presently always point to the font NeoForth and has no actual function
pxspacingx	84	horizontal spacing between chars – presently defaults to 2 and has no actual function
pxspacingy	88	vertical spacing between chars – presently defaults to 2 and has no actual function

pxcharx	92	number of chars per line - calculated during the creation of a window
pxchary	96	number of lines in window - calculated during the creation of a window
pxvarout	100	varOUT for a window - is x_pos of the textcursor
pxvarrow	104	which row is text cursor
px_wobase	108	address 1st pixel to draw -> this value is filled during creation of the window
px_wohix	112	size x window minus 2*border - limits of drawing -> redo: see px_wobase
px_wohiy	116	size y window minus 2*border - limits of drawing -> redo: see px_wobase
pxflags	120	not in use
pxnoscroll	124	flag=true-> no scrolling
pxvisible	128	window is visible if flag <> 0 -this value counts down to zero at the screen-refresh rate, once zero is reached the window will become invisible

### WORD AVAILABLE FOR USING THE WINDOWS INFO-TABLE

**WINADR>** ( win# -- addr ): for a given window (0 - 7) returns the address where the relevant part of the info-table starts. Adding the offset from the table above to the returned address gives the actual address of a value.

# Compilation/Execution control

## COMPILATION/EXECUTION CONTROL IS EXPERIMENTAL AND MIGHT NOT FUNCTION

This section of wabiPi4 was developed with the support of Albert Nijhof and Willem Ouwerkerk.

The words **[IF]**, **[ELSE]**, **[THEN]**, **REFILL**, and **PARSE** function as specified in the standard. In addition **[IF]** is available.

**[IF:** [IF allows for conditional compilation based on categories.

**[IF\_VARIANT:** is a VALUE which contains the categorie to be executed or compiled.

As the example shows it can be used while interpreting. This example shows the effect of different category selections

```
char A to [if_variant
[IF AC]    1 [ELSE] 0 [THEN] .
[IF CA]    1 [ELSE] 0 [THEN] .
[IF B]     1 [ELSE] 0 [THEN] .
[IF BCEFD] 1 [ELSE] 0 [THEN] .
[IF 13%A]  1 [ELSE] 0 [THEN] .
[IF ]       1 [ELSE] 0 [THEN] .
```

And it can be used for conditional compilation:

```
char a to [if_variant
: test1 [if BCEFD] 1 [else] 0 [then] . ;

char c to [if_variant
: test2 [if BCEFD] 1 [else] 0 [then] . ;

test1 -> 0
test2 -> 1
```

As you can see in the examples above, the categories used are NOT case-sensitive.

## CATCH THROW

WabiPi4 implements **CATCH** and **THROW** as described in the ANSI standard.  
The examples in the ANSI documentation function as described.

**Please note:**

The implementation of both **CATCH** and **THROW** is experimental and might be changed or even removed in the future if long-term system-stability of wabiPi4 is not assured due to **CATCH** and **THROW**.

Contrary to some other Forth-implementations, **CATCH/THROW** is not used for **ABORT**. And there is no last resort **CATCH** defined in **QUIT**.

# Optimisation in wabiPi4

During the development of wabiPi4, performance was an important factor, and several design-decisions result in higher running speed. Some of these can be influenced by the user during programming, others are fixed.

## Optimisations implemented in wabiPi4

**WabiPi4 uses a mix of subroutine threading and native code.** At all times wabiPi4, and any Forth program compiled by the user, is running in ARMv8 assembly. This is the fastest way of running a Forth-program for the ARM Cortex-a72 processor.

**Separate CPU- and user-return-stack:** the return-stack used by the CPU and the return-stack available to the user are separated in wabiPi4.

The main reason is the efficiency of the branch-predictor. This branch-predictor is essential for the performance of the CPU. But it only functions well if the return-stack for the CPU is not changed by a user-program.

An added bonus is that the separate return-stack makes wabiPi4 more stable and resilient against mistakes. For instance if you pop a value from the return-stack by mistake, wabiPi4 will keep running without a problem.

For the programmer there are no consequences of this separation.

**Use of assembly** and full use of the 'leaf' concept of ARM. The 'leaf' concept of ARM is: routines which do not call other routines (called leaf-routine in ARM-terminology) do not need to save and restore their return-address on stack when called. And by dedicating two CPU-registers exclusively to leaf-routines, they also do not have to save and restore registers when called. This allows for extremely fast calling and returning from these routines.

Wherever possible, wabiPi4 primitives are written in optimised assembly using this leaf-principle. For instance all comparisons, most mathematical functions and all stack-manipulation words are written as leaf-routines.

More complex Forth-primitives usually do not benefit from or even cannot use the principle, but these are also written in assembly when useful to do so. Forth-words where speed is less relevant are defined in wabiPi4 to save space and thus raise cache-efficiency.

**Dedicated memory blocks** for several time-sensitive functions. Faster routines and algorithms can be developed by fixing where certain data is stored in memory. WabiPi4 extensively uses this option. This also optimises the use of the data-caches, by ensuring that critical datablocks always start at cache-line borders.

**Dedicated CPU-registers for the DO...LOOP.** Dedicated CPU-registers for I and the loop-limit ensure the fastest possible **DO...LOOP**. This allows the **DO...LOOP** to do a loop in 2 CPU cycles. And the index 'I' can be put on the stack in 1-2 CPU cycles.

**Inlining of Forth-primitives.** The use of the leaf-principle also allows for efficient inlining of Forth-primitives, resulting in an additional increase of speed. About a third of all words in the dictionary can be inlined.

Inlining avoids time-consuming calls to subroutines. And the calls that are still made are mostly to words where inlining does not bring speed-advantages. WabiPi4 inlines a lot. Most programs mainly consist of ARM-opcodes with a few subroutine-calls dispersed in-between the opcodes.

Values, variables, constants and literals are always inlined by wabiPi4. There simply is no benefit of not inlining these. The inlining of other words can be controlled by the user (see later).

**Pruning** of unnecessary opcodes. If during compilation wabiPi4 notices that certain opcodes are unnecessary, it will prune the superfluous ones. An example is if an opcode for DROP follows an opcode for DUP. In that case both can be deleted. A few other combinations are also pruned or optimised.

Another example is when a number is put on the stack. If the number is positive and smaller as  $2^{16}$ , two opcodes are generated, otherwise three opcodes are generated.

The pruning-routine is very conservative. In case of any doubt, it will not prune. It also ensures that pruning never crosses destinations of jumps, as that could lead to errors and possibly crashes. It also ensures that compounding and pruning do not clash.

It is maybe interesting to note that pruning is done at opcode-level, not at wabiPi4 word-level. The word DROP followed by the word DUP is not compounded away. But the resulting opcodes might be pruned away.

**Hash-Table based FIND:** The searching of words in the dictionary is based on a hashing technique. This speeds up the searching for words. For large dictionaries FINDing a word can be more than a 1000 times faster. This is relevant for the speed of compilation and for the speed of interpretation of source-code. The hash-table functionality is also available to the programmer for use within a program. For this see the separate chapter.

**Compounding of words:** There are many combinations of FORTH words where the assembly-code for the combination of two words is faster than the assembly-code for the two words separately. A few of these combinations are so common that they are part of the ANSI standard. An example is '1+'. The only reason why it exists is that it is faster than a separate '1' and '+'.

An examples which is not an ANSI standard is 'OVER +'. It takes 4 CPU cycles execution-time when compiled as two separate words, but only 1 CPU cycle when coded as a combination of both words. There are hundreds, possibly even thousands of these time-saving combinations. It would be difficult for a programmer to remember which words-combinations are available as compound words in the dictionary. Especially as the combinations are partly hardware specific.

To ensure that the programmer can program in a care-free way using standard ANSI Forth, whilst at the same time ensuring best possible performance, wabiPi4 compounds words where possible. As an example: if you put 'OVER +' in the source code, wabiPi4 will compound these two words into the much faster word 'OVER+'.

The compounding feature is based on the ANSI-standard. Combinations of words where a ANSI standard exists, like '1 +' are not compounded. WabiPi4 assumes that if you do not use the ANSI-standard word, this must be on purpose.

WabiPi4 will also not correct programming-errors like a **DUP** followed by a **DROP**. WabiPi4 assumes that if you do something seemingly so illogical, that undoubtedly you must have a good reason for doing it.

First final remark: additional combinations are added to wabiPi4 in a haphazard way. Compounding functions well, but as it is a complex function, it will take some time before most wrinkles (read: 'bugs') are ironed out.

Last final remark: the CPU in the pi4b, the Cortex-A72, also does a bit of pruning. It is called 'synthesis' and it is not well documented. But combinations like '**DUP DROP**' are faster on a Cortex-A72 than one would expect.

## User-options of wabiPi4 optimisation:

### USER OPTIONS INLINING

WabiPi4 can inline primitives during compilation of definitions. Whether an individual primitive is inlined or not depends on exactly two factors:

- the **inlining-value** available in the header of each primitive.
- the value contained in **MAXINLINE** at the time the primitive is included into a definition.

When the **inlining-value** is `0x0`, the primitive is never inlined. This is the case for all non-primitive definitions and some primitive definitions.

When the inlining value is not `0x0`, then the primitive is inlined if the **inlining-value** is equal to or smaller than the value in **MAXINLINE**.

All primitives have been timed with and without inlining, and for some primitives inlining does not bring a speed-advantage. These have an inlining-value of `0x0` in the header, and will therefore never be inlined.

Inlining can be influenced by the user with the immediate words **INLINE** and **NOINLINE**. These words can be used interactively during programming. This allows the user to test whether inlining has benefits for part of a program or not.

Inlining can also be influenced by the value **MAXINLINE**.

Inlining is most beneficial for short primitives. Some primitives are only 1 opcode long and execute in 1 or even less cycles. A branch to such a primitive and the branch back together take at least 3 cycles. This means that inlining reduces the execution time from 4c to 1c (or even to 0c!).

For the pi3, inlining a maximum of 8 opcodes is optimal. Inlining longer primitives does not make the program faster, indeed the program can even be slower. For the pi4, with a different ARM-processor, tests have shown that inlining is always beneficial.

The overall speed advantage of inlining can be quite large. In some cases code can run 300% faster. The sieve benchmark without inlining runs in ~2425us. With inlining it takes ~890us, a difference of ~270%.

The lowest improvement I have seen is a 0.2% faster read-speed with a driver I developed for flash-memory. Avoiding all optimisation reduced the size of the program by 6.5%, but memory is rarely an issue. And that is why inlining is switched on by default.

**Inlining is fully transparent to the user**, there are no risks associated with inlining. The only, small, disadvantage is that inlining increases the size of a program. To give you an indication: small programs can double in size, whereas larger programs rarely increase size by more than 15%.

**Inlining can be halted** temporarily with **HALTINLINE**, or forced temporarily with **FORCEINLINE**. **HALTINLINE** and **FORCEINLINE** remember the previous inline-settings, so that inlining can later be set back to the settings as they were by using **INLINEASBEFORE**.

**HALTINLINE** and **FORCEINLINE** are not nestable. So if you use **HALTINLINE** or **FORCEINLINE** twice after each other, only the last value will be remembered.

**INLINE** ( — ): Immediate word. Starts inlining from that point onwards by setting the value **MAXINLINE** to -1 for the pi4 and 8 for the pi3.

**NOINLINE** ( — ): Immediate word. Stops inlining from that point onwards by setting the value **MAXINLINE** to 0x0. Can be used within the source-code to stop inlining from that point of the source-code.

**HALTINLINE** ( — ): halts inlining temporarily and remembers the previous inlining-settings. Not nestable.

**FORCEINLINE** ( — ): starts inlining and remembers the previous inlining-settings. Not nestable.

**INLINEASBEFORE** ( — ): reinstates the previous inlining-setting before **HALTINLINE** or **FORCEINLINE** were executed.

**MAXINLINE** ( — max\_inline\_length ): value – defines the upper limit of the length of primitives which are inlined. For instance, if **MAXINLINE** contains the value 8, primitives with a length of 8 or less opcodes are inlined. Primitives with a length of 9 or longer are not inlined.

If **MAXINLINE** contains -1, inlining is done for all inlinable primitives. If **MAXINLINE** contains 0x0, inlining is switched off. For the pi3 a value of 8 is optimal, for the pi4 -1 is optimal.

## USER OPTIONS PRUNING

Pruning can be influenced by the user with the immediate words **PRUNE** and **NOPRUNE**. They can be used interactively during programming.

The overall speed advantage of pruning is usually negligible. But in some rare cases pruning can result in an increase of speed of ~10% or more for individual words.

Pruning is switched on as default after a reset or start of the system. Pruning is fully transparent to the user. As far as

presently is known there are no disadvantages or risks associated with pruning with the present setup.

**PRUNE** ( -- ): immediate - pruning is done from the location of **PRUNE** in the source.

**NOPRUNE** ( -- ): immediate - halts pruning from the location of **NOPRUNE** in the source.

## USER OPTIONS COMPOUNDING OF WORDS

Compounding is switched on as a default. It can be influenced by the user with the immediate words **COMPOUND** and **NOCOMPOUND**. These words can be used interactively during programming.

The overall speed advantage of compounding is usually small but noticeable. Expect anything between 0% (if you are unlucky) and 30% (if you are this lucky, consider playing the lotto...). speed advantage, depending on the sort of program you are writing and your programming-style.

Executing **.COMPOUNDTABLE** will give an overview of which combinations of words are combined into more efficient alternative definitions. The alternative words used for compounding are hidden and so not visible when listing words with **WORDS**. But, as with all hidden words, you can use them in a program if you want to.

**COMPOUND** ( -- ): immediate - compounding is done from the location of **COMPOUND** in the source.

**NOCOMPOUND** ( -- ): immediate - halts compounding from the location of **NOCOMPOUND** in the source.

**.COMPOUNDTABLE** ( -- ): lists the complete compound-table.

## Hash-table functionality

**Hash-table functionality:** Wabi-Forth uses a hash-table, also known as key-to-address-transformation, to speed up searching for definitions in the dictionary. In Forth, the word **FIND** is the standard word to find a word/definition in the dictionary. Classically **FIND** does this by checking one definition after the other till the correct one is found, or until it can be confirmed that the word does not exist in the dictionary. The disadvantage of this way of trying to find a definition is that it is slow. And it gets slower when the dictionary grows. In tests done with a library of 2420 words, using the classic way of trying to **FIND** a word, wabi-Forth on average needs ~25000 machine cycles to find a definition. Confirming that a given word does not exist in the dictionary takes ~49000 machine cycles.

A usual way of speeding up **FIND** is by logically splitting up the dictionary into several smaller dictionaries. This makes **FIND** faster, depending on into how many parts the dictionary is split. But **FIND** still gets slower as the dictionary grows.

The better way of speeding up the **FIND** is by using a hash-table. This is transparent to the user, it is just there and functions. Finding a 5 character word in a 2420 word dictionary with support of a hash table is around a 550 times faster than the classic way. And as the dictionary grows, the speed advantage grows as well. Finding the same 5 character word in a 10000 word dictionary is already around 2000 times faster.

Finding a definition fast is important, but speedily confirming that a word is not available in the dictionary is even more important. And here a hash-table really shines. The best case is confirming that a 1 character definition does not exist in the dictionary. In tests with a 2420 word dictionary, such a short definition can be confirmed to *not* exist in 30 ns. Which is around 850 times faster than using the classic way. This is the best case, the average improvement is closer to 400 times.

### MEASURING THE BENEFITS OF THE HASH-TABLE

The **benefit of the hash-table** can be big. The following example shows how big:

```
s" y" sliteral tsty
: tst0 (( tsty uncount find 2drop )) ; \ note: UNCOUNT !
```

If you run **tst0** twice, once with the hash-table and without, like this:  
**usehtb drop**  
**tst0**  
**nohtb**  
**tst0**  
**usehtb drop**

You will get the time it takes to confirm that the string "y" is *not* part of the dictionary, based on trying 1 million times. With the hash-table it is around 850\* faster than without the hash-table.

```
31113 us -> 56.0 cycles
26632375 us -> 47938.3 cycles
```

A longer string takes a bit longer, but is still very much faster with a hash-table than without. Try it!

```
s" Doemaarwat" sliteral tsf10
: tst1 (( tsf10 uncount find 2drop )) ;
usehtb drop
tst1
nohtb
tst1

prints:
48225 us -> 86.8 cycles ok
27956388 us -> 50321.5 cycles
```

The following example also clearly demonstrates the benefit of a hash-table. It measures the shortest possible time in total to:

- put 2 numbers on stack
- add these numbers
- drop the result

**TST0** uses normal wabiPi4, **TST1** interprets the needed steps with the word EVALUATE. **TST1** is run twice, once with hash-table on, and once with the hash-table off.

```
: tst0 (( 2 2 + drop )) ;
: tst1 (( s" 2 2 + drop" evaluate )) ;

tst0
tst1
nohtb
tst1 ( << grab a coffee, this takes a while... )
usehtb drop
```

You will see that the difference in running time is large, the running time of the second run of TST1 is around 71 second. The slowest way is 32 thousand times slower than the fastest way. Using the hash-table is ~165\* faster than not using the hash-table with around 2400 words in the dictionary.

## TECHNICAL SETUP OF HASH-TABLE

The hash-table in the wabiPi4 system is a two-level system, with an entry-pool of 838860 hash-table rows and a link-pool containing 4753513 hash-table rows. Each row contains 3 words. The first word is the link-field, the second is the CRC-field and the 3rd word contains the message. A new definition is always hashed into the entry-pool. If on putting a new definition into the entry-pool it is found that a row is already in use by another definition, this original row will be moved into the link-pool and the new definition is put into the entry-pool. The link-field ensures that the original row can still be found.

The implementation with entry-pool and link-pool ensures that even with a very large dictionary, containing millions of definitions, the average maximum number of links to be followed to find a word is ~6. And the CRC field ensures that almost all that needs to be done is the direct comparison of two 32bit values. Which is very fast.

**User functionality:** A special feature of the wabi-Forth hash-table is that the use by **FIND** can be switched off. The hash-table is then available for the programmer. This gives the best of two worlds: during the compilation of a program, the hash-table speeds up the compilation. And once compilation is ready, hash-table functionality is available to speed up user-programs if so wanted.

Using a hash table is an advanced topic which cannot be discussed here. But the following words are available for the user who wants to give it a try.

**USEHTB** ( -- n ): resets and starts the use of the hash-table by the wabi-system. Returns the number of words hashed.

**NOHTB** ( -- ): stops the use of the hash-table by wabi-system, and makes the hash table available for use by the user. Compilation still functions, but is slower.

**HTB\_CLEAR** ( -- ): clears the existing hash-table and switches the wabi-Forth system to the classic and slow FIND. The hash-table is then available to the programmer.

**HTB\_REDO** ( -- n ): re-fills the hash-table with the definitions from the dictionary and returns the number of definitions hashed. And switches the wabi-Forth system back to hash-table based fast FIND.

**HASHWORD** ( message, addr, len -- ): puts the 32bit message in the hash-table based on a hash of the data at the specified address and the length. What meaning the data has, und thus what meaning the message has, is up to the programmer. It could be a link to a field in a table, or point to a memory-block. Or whatever you need. If wabiPi4 is using the hash-table, it stores the XT of a word in the message-field.

**HTBFIND** ( addr\_of\_string -- message ): finds the message based on the hashing of the counted string at addr\_of\_string. Please note that the standard Forth word **FIND**, und thus **HTBFIND**, uses a counted string to find something.

## Real time clock module

The RTC related words only function if a RTC-module is connected to the Raspberry. WabiPi4 contains a driver for such real-time clock. The clock can be read and controlled using the following words:

### Date and time related functions

**RTCSET** ( hh mm ss dd mm yy -- ) sets the RTC-clock according to the data on the stack. Please be aware that the system, in accordance with a longstanding FORTH-tradition, does not perform any checks on the data. Making sure that the data is valid is your responsibility.

**RTCGETDATA** ( -- ): fills the following values with data from the real time clock.

**RTCYEAR** ( -- year ): returns the year in the format yyyy

**RTCMONTH** ( -- month ): returns the month

**RTCDAY** ( -- day): returns the day of the month

**RTCWEEKDAY** ( -- weekday ): returns the weekday as a number, where 1=Sunday, 2=Monday etc

**RTCHOURS** ( -- hours ): returns the hours in 24 hour format

**RTCMINUTES** ( -- minutes ): returns the minutes

**RTCSECONDS** ( -- seconds ): returns the seconds

The following printing-words are available:

**.WEEKDAY** ( 1-7 -- ): prints the name of a day as a 3 character abbreviation, where 1=sun, 2=mon etc.

**.MONTHNAME** ( 1-12 -- ): prints the name of a months as a 3 character abbreviation, where 1=jan, 2=feb etc.

**.DATE** ( -- ): prints the date in the format dd-mm-yyy

**.DATECAMEL** ( -- ): prints the date in the format DDmmYY

**.TIME** ( -- ): prints the time in the format hh:mm:ss

**.RTC** ( -- ): prints date, weekday and time

These printing words all always execute **RTCGETDATA** (see above) and so print the most up-to-date time and dates.

### OTHER DATE-RELATED WORDS

- **GETWEEKDAY** ( dd, mm, yyyy -- 1-7 ): returns the day of the week for a given date in the Gregorian calendar (which starts at October the 15th 1582). Where 1=Sunday, 2=Monday etc.
- **LEAPYEAR?** ( yyyy -- flag ): returns 'true' if the year on the stack is a leap-year, otherwise returns a 'false'.

## BATTERY PROTECTED MEMORY OF THE RTC-MODULE

The **real time clock module** contains 64 bytes of battery protected RAM, and these are available to the user. These 64 bytes are at locations 32 to 95. Positions 32 to 39 also double as power-down memory. If the RTC-module looses external power, the date and time where the battery took over the power-supply of the RTC-module is stored in these positions. This allows to monitor an external power-supply.

The following words are available to use and manage the 64 bytes of RAM:

**RTCCLEARRAM** ( -- ): clears all 64 bytes of RAM, including the Power-down bytes.

**RTCDUMPRAM** ( -- ): prints the 64 bytes of RAM in table form.

**RTCRAM!** ( char, 32-95 -- ): stores char n at position 32 to 95. When a position outside of the range is specified, no warning is given and no value is stored.

**RTCRAM@** ( 32-95 -- char | -1 ): fetches an unsigned char from position 32 to 95. If a position outside this range is specified than -1 is returned.

## ARM-v8 assembler

An **integrated ARMv8 assembler** was an important requirement for WabiForth. After a reboot or '**COLD**', the assembler is available directly. No need for messing around with vocabularies or loading of source-code.  
 You can use '**FORGET ASSEMBLER**' immediately after a reboot or '**COLD**' to remove the assembler from the dictionary. This saves a cool 0,0013% of total available memory.

**The assembler is reasonably complete.** Missing are NEON opcodes and literal pools. The NEON opcodes will be added in future, the literal pools not (but you can always use **PC**, as register).

And finally one has to use **MOV** for the shifts and rotate, but this does not limit functionality.

Any relevant additions to the assembler, especially functionality, will happen once more experience is gained with the present version. The focus is presently on testing and debugging.

On the following pages the available assembler words are listed.

### Why using the assembler

There are several reasons to use the assembler:

First: **the fun** you can have with an assembler. Being in total control, and being fully responsible for every aspect of a routine can be very satisfying.

Second: **speed**. WabiPi4 pushes the hardware pretty hard, with the assembler you can push it even harder.

Third reason: **help the optimiser**. The optimiser only handles a few common situations and is risk-averse to the absurd. For instance it will never optimise directly after an immediate word. Chances are pretty high that you can do better than the optimiser.

Fourth reason: **access to parts of the ARM-system** which are not covered by wabiPi4 functionality. For instance wabiPi4 has no build in support for the debug-functionality of the ARM-processor, although it is switched on. But if you wish to do so, this is available using the assembler.

### Forth words and opcodes can be mixed

Assembly code can be used in any of the defining words. A very important feature of the wabiPi4 assembler is that it is possible to mix Forth words and assembly code at will, and within any definition.

The following, very silly, example (which you do not have to understand and which does nothing useful) shows respectively a literal (10000), an opcode (**MOV**,) a **DUP**, an opcode (**MVN**,) and a **STAR**. It puts -49 on top of stack. It shows the principle of mixing Forth words and opcodes. (it also shows 2 ways of creating values in assembly)

```

: silly ( -- -49 )
    10000          \ << this puts 10000 on the stack
    [ top, 0 7 i#, mov, ] \ << this changes the 10000 to 7
    dup           \ << this duplicates the 7
    [ top, 0 6 i#, mvn, ] \ << this changes one 7 in -7
    *             \ << this multiplies -7 and 7
;

```

**Format and syntax:**

The assembler follows the established Forth-way of assembling by having the registers followed by the opcode.

An example:

```
ADD r1, r2, r3      \ add r2 and r3 with the result in r1
```

in the standard ARM-assembler notation.

Would become:

```
r1, r2, r3, ADD,
```

in the wabi assembler.

**Notice the comma's** after every word in the wabi-assembler. They are there to avoid confusion between Forth definitions and assembler definitions.

There are a few exceptions to the 'comma'-convention:

- destination labels end with a colon (fi: **label0:** )
- ordinary numbers are written normally (fi: **23** )
- the word **LBL[** has no comma as it manages the label-system and does not assemble anything

**Assembler words in wabiPi4 are not immediate**

This means that you have to put them between brackets within a new definition.

There are pro's and con's to having the words compiling or immediate. The big advantage of normal compiling words is that it is much easier to develop macro's. No need to resort to endless **POSTPONE** statements.

The disadvantage is that you have to put brackets ( '[' and ']' ) around blocks of assembly words.

## ASSEMBLER WORDS

**Normal definitions and primitives**

Opcodes are either part of a normal wabiPi4 definition or part of a primitive. The distinction is important. Primitives must follow certain rules to function correctly.

**CODE: defines a primitive**

In most Forths the colon, ':', is for defining Forth-words and the word **CODE:** (or something similar) is used for definitions using assembly.

WabiPi4 is more flexible, assembly/opcodes can be mixed with Forth words at any time and place. The function of **CODE:** is not to distinguish between Forth and assembly, but rather to define a primitive.

A primitive is a definition which does not call another definition. Because of this there is no need for the link-register (=r14) to be saved and restored. And that is the exact function of **CODE**: It starts a definition which does not save at the start of the definition, nor does it restore the link-register at the end of the definition. And exactly this allows a primitive to be inlined if wanted.

In the chapter "**Snippets of code – examples of assembly**" there are examples of inlinable primitives.

#### **Rules for a primitive**

- primitives must not call any other definitions
- *only* the registers **V**, and **W**, can be used as scratch-registers within a primitive. Registers **r0**, **r1**, **r2**, **r3** and **r4** cannot be used in a primitive. The registers with specific functions, **DTS**, and **TOP**, can be used within a primitive for their specific function.
- registers **V**, and **W**, can be used outside of a primitive, but their preservation is not assured after a call to another word.

WabiPi4 will only function correctly if these rules are followed. The bugs resulting from violating these rules can be dreadfully hard to find.

## **REGISTERS**

The ARM processor in ARM32 mode has 16 registers. Register 5 to 15 each have two names. The standard ARM-name, like "r5,", and the name with which they are known internally in the wabi system. These names are synonyms and can be mixed.

r0,	(scratch for non-primitives)
r1,	(scratch for non-primitives)
r2,	(scratch for non-primitives)
r3,	(scratch for non-primitives)
r4,	(internal scratch, do not use)
r5, = top,	(top of stack register)
r6, = fps,	(floating point stack pointer)
r7, = i,	(index of D0...LOOP and FOR...NEXT)
r8, = lim,	(limit of D0...LOOP)
r9, = dts,	(data stack pointer)
r10, = uss,	(user return stack pointer)
r11, = v,	(scratch, only use in a primitive)
r12, = w,	(scratch, only use in a primitive)
r13, = sp,	(CPU return-stack pointer)
r14, = lr,	(link register)
r15, = pc,	(program-counter)

## **CONDITIONAL EXECUTION**

Use in front of any opcode which can be executed conditionally. See the ARM documentation for more information.

eq,	
ne,	
cs, = hs,	
cc, = lo,	

```

mi,
pl,
vs,
vc,
hi,
ls,
ge,
lt,
gt,
le,
al, ( default, optional )

```

## OPCODES

### Data Processing Opcodes:

```

and,      ands,
eor,      eors,
sub,      subs,
rsb,      rsbs,
add,      adds,
adc,      adcs,
sbc,      sbcs,
rsc,      rscs,
orr,      orrs,
mov,      movs,
bic,      bics,
mvn,      mvns,

tst,
teq,
cmp,
cmn,

```

### Operand 2:

This is a very flexible feature of the ARM-processor.

Operand 2 can either be:

- register
- register with it's content shifted immediately
- register with it's content shifted defined by another reg
- immediate, a 8 bit value rolled right with 0, 2, 4, etc

I#,

Please note: the following 8 words are ONLY for use in operand 2. Use MOV, if you want to shift or rotate values.

```

lsl#,
lsr#,
asr#,
ror#,
lslr,
lsrr,
asrr,
rror,

```

### 16b immediate:

```

movt,
movw,

```

### various:

```

sel,
clz,

```

usad8,  
usada8,  
clrex,  
setendle,  
setendbe,

**Multiply:**

mul,	muls,		
mls,			
mla,	mlas,		
smull,	smulls,		
smlal,	smlals,		
umull,	umulls,		
umlal,	umlals,		
umaal,			
smllabb,	smlabt,	smlatb,	smlatt,
smlad,	smladx,		
smlalbb,	smlalbt,	smlaltb,	smlaltt,
smlald,	smlaldx,		
smlawb,	smlawt,		
smlsd,	smlsdx,		
smmla,	smmlar,		
smmls,	smmlsr,		
smmul,	smmulr,		
smuad,	smuadx,		
smulbb,	smulbt,	smultb,	smultt,
smulwb,	smulwt,		
smusd,	smusdx,		
smlsld,	smlsldx,		

**Load and Store:**

ldr,	str,
ldrb,	strb,
ldrh,	strh,
ldrsh,	
ldrsb,	
ldrd,	strd,

**Load and Store acquire/exclusive**

lda,	stl,
ldab,	stlb,
ldah,	stlh,
ldaeax,	stlex,
ldaeaxb,	stlexb,
ldaeaxh,	stlexh,
ldaeaxd,	stlexd,
ldrex,	strext,
ldrexrb,	strexb,
ldrexrh,	strext,
ldrexrd,	strexd,
ldrt,	strt,
ldrbt,	strbt,
ldrht,	strht,
ldrsht,	
ldrsbt,	

**Addressing modes:**

+],	-],
+]!,	-]!,

```
]+!,      ]-!,
i+],      i-],
i+]!,      i-!]!,
]i+!,      ]i-!,
```

**Block Load and Store:**

```
ldmed,
ldmib,
ldmfd,
ldmia,
ldmea,
ldmdb,
ldmfa,
ldmda,
```

```
stmfa,
stmb,
stmea,
stmia,
stmfd,
stmdb,
stmed,
stmda,
```

**Register-block definition:**

```
{,
r-r,
},
}!,
}^,
}!^,
```

**Processor state:**

```
cps,
cpsid,
cpsie,
```

**processor modes**  
 user,  
 fiq,  
 irq,  
 supervisor,  
 monitor,  
 abort,  
 hyp,  
 undefined,  
 system,

**a,i,f flags**  
 <a>, <ai>, <af>, <aif>, <i>, <if>, <f>,

**CoProcessor:**

```
mcr,
mcrr,
mrc,
mrrc,
```

**CoProcessor registers:**

```
p14, p15,
c0, c1, c2, c3, c4, c5, c6, c7, c8,
c9, c10, c11, c12, c13, c14, c15,
```

**Move to/from special registers:**

```

mrs,
msr, ( only reg to banked_reg - no immediate )

special registers:
spsr, cpsr, apsr, ( only for mrs, !! )
r8_usr, r9_usr, r10_usr, r11_usr, r12_usr, sp_usr, lr_usr,
r8_fiq, r9_fiq, r10_fiq, r11_fiq, r12_fiq, sp_fiq, lr_fiq,
lr_irq, sp_irq,
lr_svc, sp_svc,
lr_abt, sp_abt,
lr_und, sp_und,
lr_mon, sp_mon,
elr_hyp, sp_hyp,
spsr_fiq,
spsr_irq,
spsr_svc,
spsr_abt,
spsr_und,
spsr_mon,
spsr_hyp,

```

**External debugging:**

```

ldc,      Rn, 8bIMM, {index_mode} ldc, -> 8b immediate
          mandatory for every mode of opcode. Leave out p14 and
          c5.
stc,      ditto

```

**Division:**

```

sdiv,
udiv,

```

**Barriers:**

```

csdb, ( no options )
pssbb, ( no options )

dmb,
dsb,
isb,

```

**options for domain of memory barriers:**

```

sy, ( default, optional )
st,
ld,
ish,
ishst,
ishld,
nsh,
nshst,
nshld,
osh,
oshst,
oshld,

```

**Hints:**

```

sev,
sevl,
wfe,
wfi,
yield,
nop,

```

```

pldw,      for literal variant use pc,
pld,       ditto
pli,       ditto

```

**Exception handling:**

```

eret,
bkpt,      16b value before opcode mandatory – no condition
hvc,       ditto

smc,      4b value before opcode mandatory – condition allowed
svc,      24b value before opcode mandatory – condition allowed

rfe,      rfe!, \ rfe & refia are synonyms
rfeia,    rfeia!,
rfeda,    rfeda!,
rfedb,    rfedb!,
rfeib,    rfeib!, 

\ for all following: put processor mode before opcode
-> fi: hyp, srsia,
srs,      srs!, \ srs and srsia are synonyms
srsia,    srsia!,
srsda,    srsda!,
srsdb,    srsdb!,
srsib,    srsib!,

```

**Cyclic redundancy check:**

```

crc32w, = crc32,
crc32h,
crc32b,
crc32cw, = crc32c,
crc32ch,
crc32cb,

```

**Bitfield manipulation:**

```

bfc,
bfi,
sbfz,
ubfz,

```

**Extent (and add): use ror#, to specify which byte - 0 is default**

```

sxtb,
sxtab,
sxth,
sxtah,
uxtb,
uxtab,
uxth,
uxtah,

```

**SIMD extent (and add):**

```

sxtb16,
sxtab16,
uxtb16,
uxtab16,

```

**Saturated arithmetic:**

```

qadd,
qdadd,
qsub,

```

qdsb,

ssat,  
usat,  
ssat16,  
usat16,

**Packing and unpacking:**

pkhtb,  
pkhtb,

**Parallel additions and subtraction:**

sadd16, qadd16, shadd16, uadd16, uqadd16, uhadd16,  
sadd8, qadd8, shadd8, uadd8, uqadd8, uhadd8,  
saxs, qasx, shasx, uasx, uqasx, uhasx,  
ssax, qsax, shsax, usax, uqsax, uhsax,  
ssub16, qsub16, shsub16, usub16, uqsub16, uhsub16,  
ssub8, qsub8, shsub8, usub8, uqsub8, uhsub8,

**Byte and Bit Reverse:**

rev,  
rev16,  
rbit,  
revsh,

**Branch:**

**bx**, ( reg -- ): - branches to address in register  
**blx**, ( reg -- ): - linked branch to address in register - this  
is like a jump to a subroutine with the destination address in  
register

**Conditional Branch:**

Some branches have two different names for the same function.  
They are put next to each other in the following list

beq,  
bne,  
bcs, = bhs,  
bcc, = blo,  
bmi,  
bpl,  
bvs,  
bvc,  
bhi,  
bls,  
bge,  
blt,  
bgt,  
ble,  
b, = bal,

bleq,  
blne,  
blcs, = blhs,  
blcc, = bllo,  
blmi,  
blpl,  
blvs,  
blvc,  
blhi,  
bls,  
blge,  
blt,  
blgt,  
blle,

```
bl,    = blal,
```

## LABELS

The assembler contains a simple label system. It has a capacity of a maximum of 512 individual labels. There are 12 pre-defined labels, named **label0:** to **label11::**. The word **NEWLBL** lets you create more labels.

The word **LBL[** is used to initiate the labels.

**LBL[ ( -- )**: resets the label-addresses. Normally used once in a program. Multiple wabiForth definitions can share labels. This makes it possible to branch from one definition to a label inside another.

**NEWLBL ( <name> -- )**: using the name after the word NEWLBL, creates a new assembler-label.

### Pre-defined labels:

```
label0:  
label1:  
label2:  
label3:  
label4:  
label5:  
label6:  
label7:  
label8:  
label9:  
label10:  
label11:
```

## Lessons from one opcode...

The simplest definition with assembly imaginable is a definition containing 1 opcode. The example here just adds 15 to the top. This is enough to learn a few valuable lessons on assembly.

As base for comparison, the example in forth:  
`: f15+ 15 + ;`

The same but using assembly inside a normal definition. Note: there is no need to uses LBL[ as there are no labels.  
`: w15+ [ top, top, 0 15 i#, add, ] ;`

And using assembly as a primitive.

`code: c15+ [ top, top, 0 15 i#, add, ] ;`

We can compare the execution-times of the three definitions with the words (( and )). A suitable test-setup would be the following. We put a zero on top of the stack, add 15 to it with one of the test-routines above, and then drop the result.

`: tstf (( 0 f15+ drop )) ;  
: tstd (( 0 w15+ drop )) ;  
: tstdc (( 0 c15+ drop )) ;`

Running the three test-words gives the number of cycles execution-time for each of the three versions of our definition:  
`tstf 8334 us -> 15.0 cycles  
tstd 7779 us -> 14.0 cycles  
tstdc 2224 us -> 4.0 cycles`

Assembly is mostly used to enhance performance, and the increase in speed is clearly noticeable.

But there is one more we can do. Make an **inlinable** primitive assembly routine. In this specific case we only have to add the word '**inlinable**', with a 1 in front of it, after the definition of our primitive.

Like this:

`code: ic15+ [ top, top, 0 15 i#, add, ] ; 1 inlinable`

We measure the time in the same way:

`: tsti (( 0 ic15+ drop )) ;  
tsti 556 us -> 1.0 cycles`

A very impressive result! The routine is now 15 times faster than the original routine. The main reason is that inlining avoids the linked branch to and the return from the called definition.

### Lessons learned:

1. The rudimentary compiler of wabiPi4, coupled with the excellent Cortex-a72 processor, do a good job of creating and running efficient and fast code.
2. Assembly is faster than Forth.
3. At least for short assembly routines, developing inlinable primitives brings the biggest benefits with regard to speed.

## Snippets of code – examples of assembly

Violà: a chapter with an eclectic mix of definitions using assembly. There is no special order and it, for sure, is not enough to learn ARM-assembly. But it shows how certain specific cases are programmed. More examples will be added in a haphazard way.

This example shows how **conditional execution** saves opcodes and branches. With only 3 opcodes, lower ASCII characters are converted to capitals. Also note that in this case **CODE:** is not mandatory. The version with **CODE:** will be faster, but it does function with a **COLON**.

```
: 2caps ( char -- CHAR ) [
    v, top, 0 97 i#, sub,
    v, 0 26 i#, cmp,
    top, top, 0 32 i#, lo, sub,      \ << conditional execution
] ;
```

This example shows 2 different versions of **ROT** in assembly. The 4 opcode version is the one used in wabiPi4 as it is slightly faster. The examples shows the use of **normal and immediate indexed addressing** for Load and Store opcodes. It also shows the use of **CODE:** and **INLINABLE**.

```
code: ROT ( m n o -- n o m ) [
    v, top, mov,
    w, dts, ldr,
    top, dts, 4 i+], ldr,
    v, dts, str,
    w, dts, 4 i+], str,
] ; 5 inlinable

code: ROT ( m n o -- n o m ) [
    v, top, mov,
    w, dts, ldr,
    top, dts, 4 i+], ldr,
    dts, {, v, w, }, stmia,
] ; 4 inlinable
```

Here is shown how **-ROT** is coded in assembly. It shows the use of **multiple load and store**.

```
code: -ROT ( m n o -- o m n ) [
    w, top, mov,
    dts, {, top, v, }, ldmia,
    dts, {, v, w, }, stmia,
] ; 3 inlinable
```

The following is an example of **an optimising word**. It combines the function of the words ROT and DROP into 1 word to reduce execution time. ROT uses 5 opcodes and DROP uses 1 opcode. But combined into 1 word, ROTDROP, only 2 opcodes are needed. And

this combined word is 3 cycles faster than **ROT** and **DROP** separately.

If **CODE:** and **2 INLINABLE** are not used, the word is slower than **ROT** and **DROP**. It also shows the use of the 'post-increment with an immediate' addressing-mode.

```
code: rotdrop [
    v, dts, 4 ]i+!, ldr,
    v, dts, str,
] ; 2 inlinable
```

#### Comparing a conditional branch vs conditional execution.

Both examples perform the same trivial action, namely to check if the top is equal to 23. But one uses a conditional branch and the other uses conditional execution. On the Cortex-a72 the latter is the shorter but NOT the faster option. This in contrast to the Cortex-a53 processor of the pi3b+ where the second is much faster. This is a consequence of design-decisions of the ARM-corporation.

Example with **conditional branch** – 2.0 cycles execution time

```
code: 23=
    lbl[ [                                \ lvl[ resets the labels
        top, 0 23 i#, cmp,
        top, 0 movw,
        bne, label1:           \ branch on not-equal to label1:
        top, 0 0 i#, mvn,
        label1:
    ] ; 4 inlinable
```

Example with **conditional execution** – 2.0 cycles execution time

```
code: fast23= [
    top, 0 23 i#, cmp,
    top, 0 0 i#, ne, mov,
    top, 0 0 i#, eq, mvn,
] ; 3 inlinable
```

Example using **MOV**, instead of **LSL**, (and comparable)

```
code: shift16left ( n -- m<<16 ) [
    top, top, 16 lsl#, mov,
] ;
```

Counting leading or trailing zeroes.

This example shows how to use the CLZ and RBIT opcodes to quickly count the number of leading or trailing zeroes within a number. RBIT reverses the order of the 32 bits of a number.

```

code: cntlz ( n - leading_zeroes )
    [ top, top, clz, ] ; 1 inlinable

code: cnttz ( n -- trailing zeroes )
    [ w, top, rbit,
      top, w, clz,
    ] ; 2 inlinable

```

This example is not relevant for the RPi4b, as **2@ (twofetch)** on the rpi4b does not need memory to be aligned. But it still is a good example how to make a **2@ (twofetch)**.

```

code: 2@nonaligned ( addr -- n )
[ v, top, 4 i+]!, ldr,
  w, top, ldr,
  top, v, mov,
  w, dts, 4 i-]!, str,
] ; 4 inlinable

```

**Example of an assembly-macro.** The following example shows how to make a simple macro. On the ARM-processor getting an 32b immediate into a reg takes two opcodes, MOVW and MOVT. This macro takes a 32b value from stack and generates the two opcodes directly.

```

: ldv32, ( reg n -- ) \ no square brackets!!!
  2dup
  16 lshift 16 rshift movw,
  16 rshift movt, ;

```

### Please note that square brackets are not used in a macro!

On a ARMv8 processor **movw**, followed by **movt**, is the fastest way of getting an 32 bit immediate into a register. On the Cortex-a72, the processor in the RPi 4b, these two opcodes run in parallel, when put in the order MOVW directly followed by MOVT. And together take only 1 cycle execution-time.

**LDV32**, can be used as a normal opcode in the assembler. The following silly example shows how to put the number 123456789 into the top register.

```
: tld [ top, 123456789 ldv32, ] ;
```

If you enter:  
`0 tld .`

123456789 will be printed as **tld** changes the **0** into 123456789.

**Calling other definitions from assembly.** In wabiPi4, assembly and wabiPi4 definitions can be mixed at will. This obviates the need to call a wabiPi4 word from within assembly, you can just mix them. However, if you really want to, it is possible to call a Forth-word from assembly directly.

The easiest way is by using the macro defined above (**ldv32**),  
The following example shows how to call the word MYSQ, from  
assembly, with the help of the macro LDV32.

First lets define MYSQ, which simply squares the top of stack:

```
: mysq ( n -- n^2 )
    dup * ;
```

We can call this definition by getting the XT of MYSQ with TICK,  
putting the XT into the r0-register and than branch-linking to  
that address.

```
: sqcall ( n -- n^2 )
[ r0, ' mysq ldv32,           \ put xt of MYSQ in r0
  r0, blx,                   \ branch linked to address in r0
] ;                         \ << do not use inlinable
```

**NOTE:** this routine is not a primitive as it calls another Forth  
definition. And so **CODE:** and **INLINABLE** cannot be used!  
Also note that there is no need to use '['], as assembling is  
done between brackets.

The macro **LDV32**, itself can be part of inlinable code. Remember  
that it creates two opcodes.

Example of the use of **shifts within operand 2**. The word implements a 1-  
seed random-generator according to Marsaglia.

First the function is implemented using wabiPi4:

```
variable seed
2345 seed !
```

```
: forthrandom ( address_seed - rndm_val ) ( 37c )
    dup >r @
    dup 5 lshift xor
    dup 9 rshift xor
    dup 7 lshift xor
    dup r> ! ;
```

The same routine, but now assembly. The code for the Forth  
version is 31 opcodes long once compiled. The assembly-version  
is only 7 opcodes long (6 opcodes when inlining) and is 3.5  
times faster.

```
variable seed
2345 seed !
```

```
code: asmrandom ( address_seed - rndm_val ) ( 10c )
[ w, top, ldr,
  w, w, w, 5 lsl#, eor,
  w, w, w, 9 lsr#, eor,
  w, w, w, 7 lsl#, eor,
  w, top, str,
  top, w, mov, ]
; 6 inlinable
```

This example splits a word on top of the stack up into 3 bytes. It shows the use of UBFX, ie unsigned bit-extract. This is a very useful opcode when handling lots of data.

The word **SPLIT24** might for instance be useful when handling SPI-addresses, which are send to a peer as 3 bytes, hi, mid and low byte. The assembly routine **SPLIT24** takes about 3 cycles. In wabiPi4 it takes 28 cycles. When speed is essential, and it often is with SPI-data-transfers, assembly routines can really help.

Coded using assembly:

```
code: SPLIT24 ( w -- c_lo c_mid c_hi )
[
    w, top, 0 8 ubfx,
    w, dts, 4 i-]!, str,
    w, top, 8 8 ubfx,
    w, dts, 4 i-]!, str,
    top, top, 16 8 ubfx, ]
; 5 inlinable

: tst1 (( 2345678 split24 3drop )) ; \ -> 3.0c
```

The same coded in wabiPi4:

```
: split24w ( w -- c_lo c_mid b_hi )
>r
r@ 255 and
r@ 8 rshift 255 and
r> 16 rshift 255 and
;

: tst2 (( 2345678 split24w 3drop )) ; \ -> 28.0c
```

Indexing with a **shifted register added to an address** in a register:

**ldr r0, [r1, r2, lsl #2]**  
becomes  
**r0, r1, r2, 2 lsl#, +], ldr,**

This is useful when going thru an array. Put the base-address in r1, the index in r2 and the data gets loaded into r0.

Example of the **signed byte extend** opcode. It is used to make a primitive which gets the **IMMEDIATE** flag (a signed byte in the header) for a given XT and converts it to a valid 32b flag.

```
code: getimmediate ( xt -- flag ) [
    v, top, 35 i-], ldrb,
    top, v, sxtb, ]
; 2 inlinable
```

The code above shows the use of the **sign\_extend** opcode. But it is not the optimal code. Using **ldrsb**, a signed byte can be loaded directly:

```
code: getimmediate ( xt -- flag ) [
    v, top, 35 i-], ldrsb, ]
; 1 inlinable
```

**Adding and subtracting a value to/from a double number.** The example shows the use of **ADDS**, and **SUBS**, (instead of **ADD**, and **SUB**,) to set the carry. Also note the two values in front of **i#**, to construct the values 0, 1 and 4. The wabiPi4 assembler uses the two values separately rather than combine into the one resulting value. Any other value which fits in the 4\_8 immediate scheme can be used.

adding 1 to a double:

```
code: 1md+ ( d -- d+1 ) [
    w, dts, ldr,
    w, w, 0 1 i#, adds,
    w, dts, str,
    top, top, 0 0 i#, adc, ]
; 4 inlinable
```

subtracting 4 from a double:

```
code: 4md- ( d -- d-4 ) [
    w, dts, ldr,
    w, w, 0 4 i#, subs,
    w, dts, str,
    top, top, 0 0 i#, sbc, ]
; 4 inlinable
```

#### Reading data from a 64 bit system-register.

This example reads data from the **PMCCNTR** register. This is a 64 bit counter which counts the actual number of CPU-cycles. The data is put as a double on the stack. The function only needs 3 opcodes to get the double value, raise the stack-pointer by two cells and put the double value on the stack.

```
code: cycles ( -- no_of_cpu_cycles )
[ top, dts, 4 i-]!, str,           \ << this is in fact a dup
  p15, 0 w, top, c9, mrrc,
  w, dts, 4 i-]!, str, ]
; 3 inlinable
```

## Primitives using high level code

This chapter shows a principle that might be fun to play with, and sometimes could help you save a few CPU-cycles. However, it is clearly *not* an example of 'good programming practise'. See the remarks at the end of this chapter for a better way.

**Inlinable primitives normally are made using the assembler.** But now and then it is possible to make a definition with normal wabiPi4 words and still make a inlinable primitive.  
This seems a contradiction as a primitive is a word which does not call other words. Inlining sometimes makes this possible.

The principle is this: wabiPi4 by default inlines as much words as possible. And an inlined word is, per definition, not called anymore.

If a definition for a word has no calls at all to existing words in the dictionary, it is in fact a primitive. And primitives can be made inlinable.

With regard to speed, inlining a definition is only useful if the resulting definition is short, usually not longer than 8 opcodes. Definitions with more opcodes hardly, or not at all, benefit from inlining. This because of the efficient branch-prediction and pre-emptive branching of the ARM-processor. Furthermore, converting a definition into a primitive is only useful if the compiler was able to generate efficient code.

An example is the words **BOUNDS**, a common word although not in the ANSI-standard. It converts a starting address and length into the upper and lower bounds, ready for use by a DO...LOOP. It is defined thus:

```
: BOUNDS ( addr len -- upper lower ) over + swap ;
```

If you enter '**SEE bounds**' you can see that it compiles to 6 opcodes. Two opcodes for fastprolog and return (called 'fastnext') to the calling word, and four opcodes to handle the actual functionality.

**see bounds**

name:BOUNDS	defid:WORD	hddn:0	inl:0	imm:-1	wid:FORTH
1	000635DC	E92D4000	_@_		fastprolog lr only
2	000635E0	E599C000	___		copy 2nd to reg:w
3	000635E4	E08CB005	___		reg:v = top + reg:w
4	000635E8	E589B000	___		copy reg:v to 2nd
5	000635EC	E1A0500C	_P_		copy reg:w to top
6	000635F0	E8BD8000	___		fastnext lr only

From the listing it is clear that no calls to other words are made. So there is nothing which prevents the programmer to use CODE: thus:

**code: boundsc over + swap ; 4 inlinable**

**see boundsc**

<b>name:BOUNDS defid:CODE hddn:0 inl:4 imm:-1 wid:FORTH</b>	
1 00063624 E599C000	_____   copy 2nd to reg:w
2 00063628 E08CB005	_____   reg:v = top + reg:w
3 0006362C E589B000	_____   copy reg:v to 2nd
4 00063630 E1A0500C	__P_   copy reg:w to top
5 00063634 E12FFF1E	__/_   codenext

The wabiPi4 version runs in 8 cycles, the version using CODE: runs in 4 cycles. The internal version, of course an optimised inlinable primitive, also runs in 4 cycles. So in this specific example the Forth compiler found the optimal solution. Please note that this is an exception. An internal primitive is usually faster than a version coded in wabiPi4.

The word **>BIT** can also be coded in this way. >BIT returns a number with only the bit at the position n set. So 1 >BIT returns 1, 3 >BIT returns 8 etc. It is a word which might be used when creating a mask for isolating bits from a value.

**: >BIT ( n -- 2^n ) 1 swap lshift ;**

**see >bit**

<b>name:&gt;BIT defid:WORD hddn:0 inl:0 imm:-1 wid:FORTH</b>	
1 00063668 E92D4000	__@_   fastprolog lr only
2 0006366C E3A0C001	_____   reg:w =0x1
3 00063670 E1A0551C	__U_   LSHIFT: top = reg:w shifted left with top
4 00063674 E8BD8000	_____   fastnext lr only

As with the example with **BOUNDS**, there is nothing which prevents us from coding it thus:

**code: >BITc 1 swap lshift ; 2 inlinable**

<b>name:&gt;BITC defid:CODE hddn:0 inl:2 imm:-1 wid:FORTH</b>	
1 000636A8 E3A0C001	_____   reg:w =0x1
2 000636AC E1A0551C	__U_   LSHIFT: top = reg:w shifted left with top
3 000636B0 E12FFF1E	__/_   codenext

The words >BIT coded in normal wabiPi4 takes about 5c to run. The version using **CODE:** and **INLINABLE** runs in 1 cycle.

#### **Please note:**

The method above will only work if *inlining* is switched 'on'. Without inlining, hi-level Forth cannot create a primitive, and the primitives created using **CODE:** will not function anymore. And your program might very well crash. If inlining is switched off in your program, use **FORCEINLINE** and **INLINEASBEFORE** before and after the definition to avoid problems.

An **alternative way** to speed up things is to use IMMEDIATE and POSTPONED in a definition. This way results in a performance which is as fast as making inlinable primitives with the method above. And halting inlining will not stop your program from working.

An example:

```
code: >BITc 1 swap lshift ; 2 inlinable
```

Could also be written as:

```
: >BITp ( n -- 2^n )
postpone 1 postpone swap postpone lshift ; immediate
```

**This version of >BIT can only be used INSIDE a definition!**

This version does not depend on the correct setting of INLINING, but cannot be used outside a definition. The resulting code is as fast as the version with CODE:. And the code is ANSI-compatible!

## Example of a longer assembly routine

This is the **sieve-benchmark** as published in BYTE in 1983. It counts the number of primes upto 16k and does that 10 times. In 1983 the fastest computer in existence, the CRAY-1, using an optimizing Fortran compiler, did this in 0.11 seconds. The Raspberry Pi 4b, using wabiPi4, takes around 890 micro seconds, around 124 times faster than the CRAY-1. The assembler version shown below takes around 665 microseconds, around 165 times faster. It clearly shows the unbelievable progress computing has made.

This routine is not picked as an example because of its beauty. Rather because it shows a number of useful things when using the assembler.

For instance it shows mixing assembly and normal wabiPi4 words. It shows the use of brackets around blocks of assembly words, even if a block only consists of 1 opcode. It also shows the use of macro's. And how to save and restore registers in an assembly routine.

```
\ the assembler macro's ----

: prologmax
    r13, {, r0, r4, r-r, r6, r8, r-r, v, w, }!, stmfdf, ;
: nextmax
    r13, {, r0, r4, r-r, r6, r8, r-r, v, w, }!, ldmfd, ;
: pushdts ( reg -- )  \ pushes reg on top of stack
    top, dts, 4 i-]!, str,
    top, swap mov, ;

lbl[                                \ reset of labels

: asmsieve \ ( odd# -- #primes, us )
[ prologmax ]                      \ brackets needed!!

mcs swap
dup 8 +
allocate          \ array in high memory

[ top, dts, 4 ]i+!, ldr,      \ drop flag from allocate

r1, top, mov,           \ top in r1
top, dts, 4 ]i+!, ldr,      \ =drop

r6, top, mov,           \ top in r6
top, dts, 4 ]i+!, ldr,      \ =drop

r0, 10 movw,           \ r0=M -> 10 iterations

\ do 10 iterations
label1:
    r2, 0 movw,           \ zero for each iteration
    top, dts, 4 i-]!, str,
    top, r1, mov,          \ r1=start array

    w, r6, 2 lsr#, mov, \ w=r6 right shifted by 2
    w, w, 0 1 i#, add,
```

```

    top, dts, 4 i-]!, str,      \ =dup
    top, w, mov, ]          \ push content of w on stack

[ hex] 01010101 [decimal]
fastfill           \ flags in array

[ r3, 0 movw,
r4, 0 movw,

label3:
r7, r1, r3, +], ldrb,      \ get r3'th element
r7, 0 0 i#, cmp,
beq, label4:

r7, r3, r3, add,
r7, r7, 0 3 i#, add, \ prime is i+i+3
r8, r7, r3, add,          \ k=prime+i

label6:
r8, r6, cmp,              \ if k>odd#
bgt, label5:

r4, r1, r8, +], strb,      \ flags[k]=0
r8, r8, r7, add,
b, label6:                 \ repeat if k<=no of odd#

label5:
r2, r2, 0 1 i#, add, \ r2=count of primes

label4:
r3, r3, 0 1 i#, add,
r3, r6, cmp,
blt, label3:               \ end sieve-loop

r0, r0, 0 1 i#, subs,      \ 1 from r0=m
bne, label1: ]             \ end 10*-loop

mcs swap -
[ r2, pushdts ]           \ primes# on stack
swap
[ nextmax ] ;

Enter:
8190 asmsieve . .

```

and you will get something like 670 and 1899 printed. The first number is the time in microseconds to complete 10 counts. The second is the actual count of primes. It should be 1899...

## Thumb assembler

The CPU in the Raspberry pi4 is capable of running in thumb-execution-mode. It runs thumb-code very efficient and fast.

To be able to experiment with this, a THUMB assembler for wabiPi4 is under development. The assembler supports thumb for M0 processor of ARM. A draft version of the assembler can be found at the wabiForth GitHub site under 'examples'.

The actual handling of the switch to thumb-execution mode and back is handled by wabiPi4 when it encounters the word **THUMB:**.

**THUMB:** ( -- ): starts a definition using the thumb-opcode set. Please note that thumb cannot be mixed with forth words. A thumb primitive is ended with the normal ';' (**SEMICOLON**). WabiPi4 handles the switch to thumb and switching back to ARM32 mode after the word has finished.

### Experimental phase:

Everything with thumb is experimental and it might be that in future things are updated or changed. For instance presently the switch to and from thumb-mode incorporate two **ISB**'s (instruction barriers) to add extra stability. This works fine but consumes about 30 or so extra CPU-cycles. More efficient solutions might be found in future.

You have to **load the Thumb-assembler** before you can use thumb-opcodes.

Here a simple routine which adds the sum between 0 and n. So if n=10, the sum calculated is  $1+2+3+4+5+6+7+8+9+10=55$ .

```
lbl[ thumb: intsum ( n -- total(0-n) )
\ calcs the total of number 0 to n on stack
[
    r0, top, movs,      \ mov top into r0 - set flags
    beq, label1:        \ skip to exit on top=0
    top, 0 movs,
label0:
    top, top, r0, adds, \ add r0 to top
    r0, r0, 1 subs,    \ sub 1 from r0
    bne, label0:        \ jump on not zero to label0:
label1:
];
;
```

It shows the use of labels, including clearing the label-table; the use of THUMB:, the format of the opcodes and conditional branching.

If you do not have the thumb-assembler running, the following will write the opcodes as hex codes and function exactly the same.

```
hex
thumb: intsum ( n -- total(n-n) )
[ 0028 16bopcode,
  D003 16bopcode,
  2500 16bopcode,
  182D 16bopcode,
  1E40 16bopcode,
  D1FC 16bopcode, ]
; decimal
```

## 5CH code

Coding in assembly is necessary if you want the fastest routines possible. But assembly code can take up a lot of space in your source-code. It would be a nice feature that, once you are satisfied with a routine in assembly, you are able to make the source-code compacter. **5ch**-code does just that. It allows for opcodes to be put into source-code in a compact fashion.

**CAVE:** Only code without jumps to other words can safely be converted to 5ch-code. This is because jumps to other words are not always relocatable in wabiPi4. So definitions which use a mix of Forth and assembly could easily crash when converted into 5ch-code. Use SEE to check whether a definition contains any jumps to other words.

**5CHDUMP** ( addr, n -- ): converts n opcodes starting at addr into a 5ch code of 1 or more lines and prints these lines. These lines can then be copied and pasted into source-code.

**5C|** ( -- ): converts a string of 5ch-code into opcodes, and commas these at HERE and onwards in a cache-safe way. This word is usually used by creating lines of 5ch-code using 5CHDUMP.

The example here shows how the word fast23= looks when it is converted into 5CH code. It also shows that making it inlinable is still possible.

Original:

```
code: fast23= [
    top, 0 23 i#, cmp,
    top, 0 0 i#, ne, mov,
    top, 0 0 i#, eq, mvn,
] ; 3 inlinable
```

**5C|** version:

```
code: fast23= 5c| S|_k%FB`X!qp$s ; 3 inlinable
```

# Interrupts

The raspberry pi4b uses the **GIC-400** (global interrupt controller=GIC) from the ARM-corporation to aid in handling interrupts. The GIC-400 has the reputation of being complicated in use. But this is only limited to the initial setup, which is done by the wabiPi4 system.

Once setup and running, it is relatively easy to handle interrupts with it. Learning how is not trivial and takes a bit of perseverance, but entirely possible. I did! ARM has excellent documentation for the GIC-400 in which the functionality is described in detail.

The wabiPi4 system contains constants pointing to the addresses of the most important GIC-registers which control the use of the GIC-400 functionality. As wabiPi4 runs in (secure) supervisor operating-mode, all functionality of the GIC-400 is available to the user.

## THE INTERRUPTS OF THE WABIPI4 SYSTEM

There is only 1 interrupt active in the wabiPi4 system, and that is a user-request for an ABORT. If the user wants to abort a running program this can be done by pressing 'fn', 'ctrl' and 'backspace' at the same time. This is recognised by the input-handler running on core2. If recognised, core2 issues a Software Generated Interrupt (SGI) to the GIC with core0 as destination. The resulting interrupt is forwarded to the FIQ-handler of core0. The FIQ-handler checks what the origin is of the interrupt, and if it is a user-request for an ABORT, an ABORT is issued.

So core0 is running without any interruption until the user requests an interrupt.

This has several advantages:

- execution of any program is as fast as possible
- core0 is always available to perform time-critical tasks. For instance checking a GPIO-port or anything else which has to be checked
- measurements of time are as reliable as technically possible

But sometimes having a regular interrupt available can be beneficial. For instance if you want to use WFI (Wait For Interrupt) for an energy-conserving pause of execution. For this reason it is possible to switch on an interrupt, by default at 500 Hz. Presently, the only use of this interrupt is for WFI. But in future might be used to regularly call a wabiPi4 routine.

The words used for this functionality are **FIQDIRECT** or **FIQBYPASS** (the names have to do with terminology of the GIC-400 interrupt controller of the pi4). Calling **FIQBYPASS** switches on a 500 Hz interrupt, calling **FIQDIRECT** disables the interrupt.

**FIQCOUNTER** points to a variable which counts up with each call of the FIQhandler of core0. It can be used to check if interrupts are reaching the FIQhandler.

**FIQBYPASS** ( -- ): switches on a 500 Hz interrupt, handled by the FIQhandler of core0. It is a slow word and takes around 13700 cycles to complete.

**FIQDIRECT** ( -- ): switches the interrupt off again. It is a slow word and takes around 13700 cycles to complete. No interrupts of core0 is the default situation.

**FIQCOUNTER** ( -- addr ): points to an address which counts up with each call of the FIQhandler of core0.

**Example:**

Executing the following will print out the frequency of calls to the FIQhandler of core0. It returns 0 after a call to FIQDIRECT, and 500 after FIQBYPASS is called.

```
fiqcounter @ wait fiqcounter @ swap - .
```

## AN ENERGY-SAVING WAIT-LOOP USING WFI OR WFE

The ARM opcode **WFI** waits for a next interrupt (hence the name: Wait For Interrupt) to occur in a system (any interrupt is okay) and then simply continues with the next opcode. The interesting feature of **WFI** is that it switches off the CPU-core while waiting but retains memory and status.

Please note that it is irrelevant for the function of **WFI** why an interrupt was activated. **WFI** also doesn't influence or change the interrupt in any way.

**WFI** is a very powerful concept. It is this opcode (combined with very smart software) which allows an Apple Watch with an 2-core ARM-processor running @ 1800 MHz, to run for 20 hours on a small battery. **WFI** ensures that for most of the time the CPU actually isn't running at all.

In wabiPi4 **WFI** normally does not function. The reason is easy to understand: waiting for an interrupt takes forever on wabiPi4 as core0 runs interrupt-free. One solution is using **WFE**. **WFE** does exactly the same as **WFI**, only it waits for an event, rather than an interrupt.

What an event is, is defined by the programmer. In wabiPi4 there is a global event-stream defined with 44.1 KHz events. This is the basis for sound-synthesis in core2. And because it is global, it can be used to activate **WFE**.

The following routine does the same as the example in the **DO...LOOP - FOR...NEXT** chapter. But by adding **WFE** or **WFI** it becomes energy-efficient.

**A solution using WFE:**

```

FOR_PRE
: KEY_TIMEOUT
  441000 for
    wfe
    key? 0= while
  next
    ." time-out"
  else
    key emit unloop
  then
;

```

**A solution using WFI:**

Another possibility is to switch on interrupts for core0 with **FIQBYPASS** (see above for an explanation). Interrupts are switched off again using **FIQDIRECT**.

Please note that the interrupts must be switched on when using the word containing **WFI**, or the program will wait forever. A simple user-interrupt will alleviate the situation if this happens. The status of the interrupts during compilation is irrelevant.

```

FOR_PRE
: KEY_TIMEOUT fiqbypass
  5000 for
    wfi
    key? 0= while
  next
    ." time-out"
  else
    key emit unloop
  then
fiqdirect ;

```

This example switches on the interrupts for core0 and starts waiting for input. After input or time-out it switches off the interrupts for core0 again.

# Cache-handling

## CACHE-HANDLING

One of the specific characteristics of the Forth programming language is that programming extends the language itself, rather than create a separate program. For a compiling Forth, like wabiPi4, this means that the act of programming changes a running program. This can only be done reliably if the handling of the caches is done very carefully and very conservative. All cache-handling is taken care of by wabiPi4. The caches are initialised during the boot process. When words are added to the dictionary, or when words are removed from the dictionary using **FORGET**, wabiPi4 handles the coherency between the data and instruction-caches and handles the branch-predictor- and translation table caches.

## LIMITATIONS OF CACHE-SETUP

The setup of a cache-system is always a balancing-act between performance, functionality and reliability. Although the system-design of wabiPi4 usually focusses on performance, in case of the setup of the cache-system reliability was prime. Therefore it was decided to run wabiPi4 with full coherence between the different cores. As such this is a bit slower than without coherence. But the big advantage is that wabiPi4 is 'rock-solid'. For instance: adding words using **EVALUATE** to a running program from within that program is possible, without fear of **ABORTS** or crashes due to coherency-problems.

There is one user-word related to cache-handling:

**CLEANCACHE** ( -- ): The word **CLEANCACHE** makes cache invalidation easy. It flushes the data-cache, invalidates the data and instruction-caches, and invalidates the branch-predictor,

"There are only two hard things in computer science: cache invalidation and naming things"

Phil Karlton

translation table and look-aside buffers. The result is that all data in the memory is up-to-date, and all caches are filled afresh when data, opcodes or branches are accessed in memory.

The biggest benefit of the word is that **it creates a known starting-condition**. This is especially useful when doing performance measurements as it raises the consistency of measurements. And obviously if you want to experiment with the effect of caches on a program, it might have some value.

**CLEANCACHE** is a pretty slow word, it can easily take more than 90 thousand cycles to complete. So you do not want to use it in fast loops.

The other area where the word might be useful is when using external agents like the DMA-controllers. External agents act independent of the CPU and thus do not know whether data in the memory is valid or invalid. By using the word **CLEANCACHE** a known starting state of the caches is created.

## ARM-VideoCore mailbox

A **Raspberry Pi 4b has 2 main processing units**, the ARM-processor and the Video-Core. These two communicate with each other through a feature called 'MAILBOX'. This mailbox is not to be confused with the mailboxes used for communication between the 4 cores of the ARM-processor. Here we are dealing with the mailbox for communication between the ARM-processor and the Video-Core.

The **mailbox-system** technically supports communication in both directions between the ARM-processor and the video-core, using 16 channels. But on the Raspberry Pi only the ARM-processor initiates communication, and always on channel 8. The ARM-processor does a 'process-request', usually accompanied by some data. The VideoCore then reacts to these 'process-requests' by performing certain actions, or giving information back etc.

The **communication is done via TAGs**. For the Raspberry 4b there are 68 valid documented TAGs, and at least 10 undocumented ones. Each TAG has a specific function covering topics like firmware- and board-version, available DMA-channels, clocks, temperature and LEDs. In addition you can change video-settings and control the cursor.

Not all **TAGs** are created equal, most return useful information or do useful tasks, several do not. WabiPi4 supports all known **TAGs**, useful or not.

WabiPi4 does presently not support multiple **TAGs** in a single process-request.

**Documentation** from the Raspberry-organisation is extensive, but slightly confusing and not always in concordance with what a TAG actually seems to do. If you are interested, get your hands on it and study it. For help you can consult the help-forums of the Raspberry organisation.

The **tags have the tag-number as name**. These tag-numbers are the same as the numbers used in the Raspberry documentation.

**TAG00001** ( -- firmware\_revision ): returns firmware revision of board wabiPi4 is running on.

**TAG10001** ( -- board\_model ): returns model-number of board

**TAG10002** ( -- board\_revision ): returns board-revision code

**TAG10003** ( -- mac\_addr=6\*char ): returns mac-address of board as 6 bytes with the most significant on top of the stack

**TAG10004** ( -- serial\_number ): returns a 64bit serial number, the serial\_number is unique for each individual board

**TAG10005** ( -- base\_addr, size ): returns the base address and size of the memory below the video-core

**TAG10006** ( -- base\_addr, size ): returns the base address and size of the video-core memory

**TAG10007** ( -- 12\*(parent\_clock, clock) ): returns the clocks in the Raspberry with their parent clock. Not very useful, see for yourself.

**TAG50001** ( -- addr, len ): returns the address and length of the command-line. TYPE can be used to print the string.

**TAG60001** ( -- mask ): returns a mask specifying which DMA channels (0-15) are available and which are in use. The bits 15:0 each represent a DMA-channel. A 0x0 means that the corresponding channel is not available for use by the programmer.

**TAG20001** ( device\_id -- device\_id, power-state ): returns the power-state for the given device.

**Unique device\_IDs (according to the Raspberry documentation):**

0x00000000: SD Card
0x00000001: UART0
0x00000002: UART1
0x00000003: USB HCD
0x00000004: I2C0
0x00000005: I2C1
0x00000006: I2C2
0x00000007: SPI
0x00000008: CCP2TX
0x00000009: Unknown (RPi4)
0x0000000A: Unknown (RPi4)

This is one of the TAGs which suffers from incomplete documentation. For instance for many other device\_IDs it is returned that a powered or unpowered device exists with that device\_ID. Which seems logical as the pi4b has far more devices than the pi3b+. It is presently undocumented which of these unspecified devices is what.

**Power-state returned:**

bit 0: 0=off, 1=on
bit 1: 0=device exists, 1=device does not exist

The returned power-state is also documented incompletely. For instance: for the UART1 (the UART used for wabiPi4 communication) it is returned that it is unpowered. Also during actual use of the UART.

**TAG28001** ( device\_id, wanted\_state -- device\_id, returned\_state ): For the given device switch the power on or off. Returns the device\_id and the result.

**wanted state:**

bit 0: 0=off, 1=on
bit 1: 0=do not wait, 1=wait for the power to become stable

**returned state:**

bit 0: 0=off, 1=on
bit 1: 0=device exists, 1=device does not exist

**TAG20002** ( device\_id -- device\_id, enable\_waittime ): For the given device, returns the required wait-time in microseconds for the power-supply to be stable after enabling the power. If a 0x0 is returned it signifies that the device does not exist. Usually a time of 1 is returned, for UART0 a time of 4000 microseconds is returned.

**TAG30001** ( clock\_id -- clock\_id, returned\_state ): For the given clock\_ID returns the present state of that clock

**unique clock IDs:**

- 0x0000000000: reserved
- 0x0000000001: EMMC
- 0x0000000002: UART
- 0x0000000003: ARM
- 0x0000000004: CORE
- 0x0000000005: V3D
- 0x0000000006: H264
- 0x0000000007: ISP
- 0x0000000008: SDRAM
- 0x0000000009: PIXEL
- 0x000000000A: PWM
- 0x000000000B: HEVC
- 0x000000000C: EMMC2
- 0x000000000D: M2MC
- 0x000000000E: PIXEL\_BVB

**state:**

- bit 0: 0=off, 1=on
- bit 1: 0=clock exists, 1=clock does not exist

**TAG38001** ( *clock\_id*, *state* -- *clock\_id*, *returned\_state* ): For the given *clock\_ID* sets the state of that *clock*. The wanted state is 0=off/1=on, the returned state is as for **TAG30001**.

**TAG30002** ( *clock\_id* -- *clock\_id*, *rate* ): for the given *clock\_ID* returns the set rate, in Hz, for that *clock*. Returns 0x0 for a non-existent *clock*. The *clock\_rate* is also returned for disabled *clocks*.

The following 3 LED-related **TAGS** function differently than documented. This might be because I use an old version of the firmware. The description below is as close as I can get to the actual way the 3 TAGs work.

**TAG30041** ( -- *powerLED\_state*, *statusLED\_state* ): returns the state of the two onboard LEDs: red power-LED and green status-LED.

**pin\_state:**

- bit 0: 0=on, 1=off << please note the reverse logic

This is different from the documentation, which states that the pin-number of the LED is returned with the status. At least on the pi4b this TAG does not return the pin-numbers.

**TAG34041** ( *wanted\_state*, *pin-number* -- *pin-number*, *state* ): checks if the wanted state is possible for the two onboard LEDs: red power-LED and green status-LED.

This **TAG** function differers from the documentation. The states are

**pin\_state:**

- for 130=red power LED: 0=on, 1=off
- for 42=green status LED: 0=off, 1=on

**TAG38041** ( *wanted\_state*, *pin-number* -- 0x0, *state* ): sets the specified pin-number to the requested state. Returns a 0x0 and the resulting state. Please note that the logic for both LEDs differs!

**pin-number:**  
 42: green status LED  
 130: red power LED

**pin\_state:**  
 for 130=red power LED: 0=on, 1=off  
 for 42=green status LED: 0=off, 1=on

**TAG30047** (clock\_id -- clock\_id, measured\_rate\_Hz ): takes a measurement of the specified clock\_id and returns the measurement in Hz.

**TAG38002** ( clock\_id, wanted\_rate, turbo\_flag -- clock\_id, set\_rate ): For the specified clock\_id set the new rate to the wanted\_rate. Setting the turbo\_flag to 1 means that nothing else happens than the rate. Setting the flag to 0x0 means that the CPU can change other things like voltage as a result of the new rate. It is good to notice that internal limits and other implementation issues mean that the clock can stay unchanged. For instance setting the ARM-clock does not function. But changing the rate of the CORE-clock does function.

**TAG30004** ( clock\_id -- clock\_id, max\_rate ): for the given clock\_id, return the maximum rate allowed by the system.

**TAG30007** ( clock\_id -- clock\_id, min\_rate ): for the given clock\_id, return the minimum rate allowed by the system.

**TAG30009** ( 0x0 -- 0x0, turbo\_state ): return the state of the turbo\_flag. 0=non\_turbo/1=turbo

**TAG38009** ( 0x0, wanted\_turbo\_state -- 0x0, turbo\_state ): set the state of the turbo\_flag, where 0=non\_turbo/1=turbo. Again a TAG where the actual function and the documentation do not match. Under certain circumstances the TAG returns 0x8000000 instead of the turbo-state.

**TAG30003** ( voltage\_id -- voltage\_id, voltage ): for the given voltage\_id returns the set voltage in microvolt. A voltage of -1 is returned if a voltage\_id does not exist. For voltage\_id=0, 0x8000000 is returned.

#### Unique voltage IDs:

- 0x0: reserved
- 0x1: core
- 0x2: SDRAM\_C
- 0x3: SDRAM\_P
- 0x4: SDRAM\_I

**TAG38003** ( voltage\_id, wanted\_voltage -- voltage\_id, voltage ): for the given voltage\_Id set the voltage to the wanted voltage.

The wanted voltage can be specified in three ways, depending on the value:

- **value <= 16**: relative to platform-specific typical voltage, in 25mV steps -> for instance a value=10 would raise the voltage by 250mV
- **16 < value < 500000**: relative to platform-specific typical voltage, in microvolts -> for instance a value=100000 would raise the voltage by 0.1V above the default voltage
- **value >= 500000**: absolute voltage in microvolts. -> for instance a value of 1000000 woud set the voltage to 1.0V.

**TAG30005** ( voltage\_id -- voltage\_id, max\_voltage ): returns the maximum voltage in microvolt for the given voltage\_id.

**TAG30008** ( voltage\_id -- voltage\_id, min\_voltage ): returns the minimum voltage in microvolt for the given voltage\_id.

**TAG30006** ( 0x0 – 0x0, temperature ): return the actual temperature of the SOC. The returned temperature is in mili-degrees Celsius but is always rounded to whole degrees.

**TAG3000A** ( 0x0 – 0x0, max\_temperature ): returns the maximum temperature of the SOC. The returned temperature is in mili-degrees Celsius but is always rounded to whole degrees.  
Overclocking may be disabled above this temperature by the SOC.

**TAG3000C** ( size, alignment, flags -- handle ): Allocates contiguous block of memory on the GPU. Size and alignment are in bytes. A non-zero value for the handle signifies success.

#### Flags contain:

- **bit0: MEM\_FLAG\_DISCARDABLE**: if true then can be resized to 0x0 at any time. Intended for cached data.
- **bit1**: unused
- **bit2-3**:
  - 0x0: **MEM\_FLAG\_NORMAL** -> normal allocating memory - do not use from ARM
  - 0x1: **MEM\_FLAG\_DIRECT** -> 0xC alias -> uncached
  - 0x2: **MEM\_FLAG\_COHERENT** -> 0x8 alias -> non-allocating in L2-cache but coherent
  - 0x3: **MEM\_FLAG\_L1\_NONALLOCATING** -> allocating in L2, not allocating in L1
- **bit4: MEM\_FLAG\_ZERO** -> if true than initialise buffer to all zero's
- **bit5: MEM\_FLAG\_NO\_INIT** -> if true -> don't initialise, otherwise initialise to all one's
- **bit6: MEM\_FLAG\_HINT\_PERMALOCK** -> if true hint to MMU of VC that buffer is likely to be locked for long periods of time.
- **other bits**: unused

**TAG3000D** ( handle -- bus\_address ): lock buffer in place, and return a bus address. Must be done before memory can be accessed. Bus-address <> 0 signals success.

**TAG3000E** ( handle -- status ): unlock buffer. The contents is retained, but may be moved. The buffer must be locked before next use. A returned status = 0x0 signals success.

**TAG3000F** ( handle -- status ): release the memory buffer. A returned status = 0x0 signals success.

**TAG30010** ( bus\_addr, r0, r1, r2, r3, r4, r5 -- r0 ): calls the function at given bus\_address with the values moved to the registers r0-r5. It blocks until the call completes. The (GPU) instruction cache is flushed. Setting bit:0 of the bus\_address will suppress the instruction\_cache flush. This is safe if the buffer hasn't changed since last execution.

This must be one of the most exiting TAGs in existence! An extra processor, which can be called in parallel to the ARM-processor. And if I tell you that this processor is a vector processor which can process 16 data-fields in parallel, than maybe you understand my excitement. Disadvantage is that the architecture of this processor is not documented, but fortunately most of the specs have been reverse engineered.

**TAG30014** ( disp\_manx\_handle -- status:0=success, mem\_handle ): Gets the disp\_mem\_handle associated with a created disp\_manx resource. This resource can be locked and the memory directly written from the ARM to avoid having to copy image data into the GPU. This TAG basically allows a block of memory outside of the GPU to be part of the operations of the GPU. Thus avoiding the copy of a block of memory to the GPU a lot of times per second. The details of this TAG are somewhat nebulous.

**TAG30020** ( block\_number -- block\_number, addr, length ): Get EDID-block – returns the block\_number, an address and the length of the data at the address which is always 132 bytes. The first 4 bytes are the status, the rest are the 128 bytes of the EDID-block. Status=0x0 is success. Keep getting blocks (starting at 0x0) till the status<>0x0.

## SCREEN SETUP WITH TAGS

All TAGs related to the setup of a screen in the GPU (frame\_buffers etc.) have to be used in 1 operation. This is a major complication of the previous situation where one could use single TAGs till completion of the task. The main problem is that the error-reporting is much more unclear in the present way.

WabiPi4 handles the single operation with several TAGs for the screen setup for the user in assembly. It does presently not support combining several TAGs into 1 action from Forth directly. If you need that functionality, you have to program it yourself, which obviously is entirely possible.

**TAG40001** ( alignment\_bytes -- base\_addr, frame\_size ) allocates a frame buffer in the GPU memory. The size of the buffer is calculated by the GPU based on the size of the virtual and physical screens in pixel height and width, and the colour-depth (ie 1, 2, 4, 8, 16 or 32 bits). This is why the setup has to be done in 1 operation, as other TAGs are used for that information.

**TAG48001** ( -- ): release the frame-buffer

**TAG40002** ( 0=off/1=on -- 0=off/1=on ): blank the screen

**TAG40003** ( -- width, height ): get physical width and height of the display. Please note that the physical width and height can differ from the size of the screen send to the monitor.

**TAG44003** ( width, height -- width, height ): test if the requested width and height for the display are possible. If

possible, the requested width and height are returned. This TAG tests, but does not change anything.

**TAG48003** ( width, heigth -- width, heigth ): set physical width and height for the display.

**TAG40004** ( -- width, heigth ): get virtual width and height of the display. Please note that the virtual buffer size is the portion of buffer that is sent to the display device, not the resolution of the buffer itself. The virtual buffer size may be smaller than the allocated buffer size in order to implement efficient scrolling and panning.

**TAG44004** ( width, heigth -- width, heigth ): test virtual width and height of the display. If possible, the requested width and height are returned. This TAG tests, but does not change anything.

**TAG48004** ( width, heigth -- width, heigth ): set virtual width and height for the display.

**TAG40005** ( -- depth ): get depth in bits per pixel. Can be 1, 2, 4, 8, 16 or 32.

**TAG44005** ( depth -- depth ): test depth in bits per pixel. Can be 1, 2, 4, 8, 16 or 32. If the returned depth is the same as the input than the test was successful. This TAG does not change any settings.

**TAG48005** ( depth -- depth ): set depth in bits per pixel. Can be 1, 2, 4, 8, 16 or 32. The returned depth might be different from the requested depth.

**TAG40006** ( -- pixel-order ): get pixel-order. 0x0=BGR/0x1=RGB.

**TAG44006** ( pixel-order -- pixel-order ): test pixel-order. If the returned pixel-order is the same as the input than the test was successful. This TAG does not change any settings.

**TAG48006** ( pixel-order -- pixel-order ) set pixel-order. The returned pixel-order might differ from the requested pixel-order.

**TAG40007** ( -- alpha-state ): get alpha-state.

**Alpha-state:**

- 0x0: alpha channel enabled (0 = fully opaque)
- 0x1: alpha channel reversed (0 = fully transparent)
- 0x2: alpha channel ignored

**TAG44007** ( alpha-state -- alpha-state ): test alpha-state. If the returned alpha-state is the same as the requested state than the test was successful. This TAG does not change any settings.

**TAG48007** ( alpha-state -- alpha-state ): set the alpha-state. The returned alpha-state might differ from the requested alpha-state.

**TAG40008** ( -- pitch ): get pitch in bytes per line.

**TAG40009** ( -- x\_pixels, y\_pixels ): get virtual offset in pixels. Can be a positive or negative number.

**TAG44009** ( x\_pixels, y\_pixels -- x\_pixels, y\_pixels ): test virtual offset in pixels. Can be a positive or negative number.

If the returned values are the same as the requested values than the test was successful.

**TAG48009** ( *x\_pixels*, *y\_pixels* -- *x\_pixels*, *y\_pixels* ): set virtual offset in pixels. Can be a positive or negative number. The returned values can differ from the requested values. Useful for panning and scrolling.

**TAG4000A** ( -- *top*, *bottom*, *left*, *right* ): get overscan, all in pixels.

**TAG4400A** ( *top*, *bottom*, *left*, *right* -- *top*, *bottom*, *left*, *right* ): test overscan, all in pixels. If the returned values are equal to the requested values then the test was successful.

**TAG4800A** ( *top*, *bottom*, *left*, *right* -- *top*, *bottom*, *left*, *right* ): set overscan, all in pixels. The returned values can differ from the requested values.

**TAG4000B** ( -- *256\*32bit\_RGB\_value* ): get palet.

# ARM-processor ID, status and info registers

The ARM-processor contains a wide selection of registers which return information describing the specific features of processor or statuses of functional blocks of the processor. A selection of these registers are directly supported by wabiPi4.

[For more information see the ARMv8 documentation.](#)

**CPU\_PARWRITE** ( addr -- n ): for a given address returns the memory attributes for a write-action to that address. The address itself is not changed.

**CPU\_PARREAD** ( addr -- n ) ditto for a read-action from that address.

**CPU\_ID\_MMFR0** ( -- n ): returns content of CPU-register ID\_MMFR0 – describes memory-features

**CPU\_ID\_MMFR1** ( -- n ): ditto for ID\_MMFR1

**CPU\_ID\_MMFR2** ( -- n ): ditto for ID\_MMFR2

**CPU\_ID\_DFR0** ( -- n ): returns content of CPU-register ID\_DFR0 – which gives top level information about the debug system. The debug system is not directly supported by wabiPi4, but is enabled.

**CPU\_CTR** ( -- n ): returns content of CPU-register CTR = cache-type-register

**CPU\_L2CTLR** ( -- n ): returns content of L2 Control Register

**CPU\_CCSIDR** ( 0, 1 or 2 -- n ): returns content of the Cache Size ID Register. The input specifies the cache: 0x0=L1 data-cache, 0x1=L1 instruction-cache, 0x2=unified cache. Other input-values result in a return of true. See the ARM-documentation for interpretation of the returned values.

**CPU\_CBAR** ( -- n ): returns content of Configuration Base Address Register which is base-address of the GIC-400 registers. Returns 0xFF84000.

**CPU\_PMCR** ( -- n ): returns content of CPU-register PMCR = performance monitors control register.

**CPU\_PMCEID1** ( -- n ): returns content of CPU-register PMCEID0 – aka the Performance Monitors Common Event Identification register 0. It describes which performance monitors are available in the CPU.

**CPU\_PMCEID1** ( -- n ): ditto for PMCEID1

**CPU\_PCCFILTR** ( -- n ) returns content of CPU-register PCCFILTR

**CPU\_CNTFRQ** ( -- 5644800 ): returns content of CPU-register CNTFRQ which holds the actual frequency of the physical, virtual and core counters – returns 5644800. This register has a pure informational character, changing the content does *not* change the frequency.

**CPU\_CPSR** ( -- n ): returns content of CPU-register CPSR = current processor state register

**CPU\_SCTLR** ( -- n ): returns content of CPU-register SCTLR = system control register

**CPU\_SCR** ( -- n ): returns content of CPU-register SCR = security configuration register  
**CPU\_SDCR** ( -- n ): returns content of CPU-register SDCR = secure debug config registers

**CPU\_RMR32** ( -- ): warm-reset request register, resets the CPU and goes to Aarch32 mode.

**CPU\_RMR64** ( -- ): warm-reset request register, resets the CPU and goes to Aarch64 mode.

Theoretically, this word allows the development of a system running in Aarch64 mode. For this to function the programmer points the vector at address 0x0 to the start of the Aarch64 system. **This word is intended for experimentation and testing.** (You know: messing around with a processor and trying to get something to run. What larks!!)

**CPU\_FPSID** ( -- n ): returns content of CPU-register FPSID = describes floating point processor features

**CPU\_DFAR** ( -- n ): returns content of CPU-register DFAR = Data Fault Address Register

**CPU\_DFSR** ( -- n ): ditto for DFSR = Data Fault Status Register

**CPU\_MAIR0** ( -- n ): returns content of CPU-register MAIR0 = memory attribute indirection register 0. This register is the same as the PRRR register which is used by wabiPi4. See the ARM documentation for more details.

**CPU\_MAIR1** ( -- n ): returns content of CPU-register MAIR1 = memory attribute indirection register 1. IN wabiPi4 this register is set but not in actual use.

The **ID\_ISARn registers** of the ARM-processor describe the available features of the specific processor wabiPi4 is running on. See the ARM documentation for technical details.

**CPU\_ID\_ISAR0** ( -- n ): returns content of CPU-register ID\_ISAR0  
**CPU\_ID\_ISAR1** ( -- n ): ditto for ID\_ISAR1  
**CPU\_ID\_ISAR2** ( -- n ): ditto for ID\_ISAR2  
**CPU\_ID\_ISAR3** ( -- n ): ditto for ID\_ISAR3  
**CPU\_ID\_ISAR4** ( -- n ): ditto for ID\_ISAR4  
**CPU\_ID\_ISAR5** ( -- n ): ditto for ID\_ISAR5

# ANSI-required documentation

The **ANSI-standard** specifies which implementation defined items must be documented. See section 4.1.1 of the ANSI-standard.

REQUIRED CATEGORY OF DOCUMENTATION	IMPLEMENTATION
aligned address requirements <b>(3.1.3.3 Addresses)</b>	the system technically does not rely on alignment of addresses, except for floating point values and ARM-opcodes, but alignment is assured during normal operation.
behavior of <b>6.1.1320 EMIT</b> for non-graphic characters	<b>EMIT</b> handles 8 bit characters for the screen and 7 bits for the UART. Characters below 32 are handled in accordance with the ASCII standard. A bell sound is not sounded.
character editing of <b>6.1.0695 ACCEPT</b> and <b>6.2.1390 EXPECT</b>	<b>ACCEPT</b> follows the standard. In addition during entry, entry can be stopped by pressing 'ESC'. In that case a string with a length of 1 and 27 as first character is returned. <b>EXPECT</b> follows the standard.
character set (3.1.2 Character types, 6.1.1320 <b>EMIT</b> , 6.1.1750 <b>KEY</b> )	The system uses ASCII for characters below 128. For the screen only, character values between 128 and 255 are printed in accordance with the definition of the character in the character table. The user can change any value in the character table.
character-aligned address requirements <b>(3.1.3.3 Addresses)</b>	none
character-set-extensions matching characteristics <b>(3.4.2 Finding definition names)</b>	Character-set-extensions are not supported. Finding definitions is exclusively done using ASCII characters.
conditions under which control characters match a space delimiter <b>(3.4.1.1 Delimiters)</b>	The control-characters 9, 10, 11, 12 and 13 are always skipped/ignored when parsing.
format of the control-flow stack <b>(3.2.3.2 Control-flow stack)</b>	The return-stack is used as control-flow stack, addresses and optionally flags denoting certain kinds of control-flow are used.
conversion of digits larger than thirty-five <b>(3.2.1.2 Digit conversion)</b>	Trying to convert digits larger than <b>BASE</b> or larger than 35 results in an <b>ABORT</b> and error-message
display after input terminates in 6.1.0695 <b>ACCEPT</b> and 6.2.1390 <b>EXPECT</b>	<b>ACCEPT</b> and <b>EXPECT</b> show the characters as entered by the user. After termination of input the characters are kept as is.
exception abort sequence (as in 6.1.0680 <b>ABORT</b> )	<b>ABORT</b> " prints a message and then aborts. The user is put back into interpreting mode in <b>QUIT</b> .
input line terminator <b>(3.2.4.1 User input device)</b>	Input line terminator: <b>CR</b> , <b>LF</b> and <b>CR/LF</b> are all accepted during input. <b>LF</b> is used during output.

REQUIRED CATEGORY OF DOCUMENTATION	IMPLEMENTATION
maximum size of a counted string, in characters <b>(3.1.3.4 Counted strings, 6.1.2450 WORD)</b>	4294967295 characters.
maximum size of a parsed string <b>(3.4.1 Parsing)</b>	1024 characters.
maximum size of a definition name, in characters <b>(3.3.1.2 Definition names)</b>	32 characters.
maximum string length for <b>6.1.1345 ENVIRONMENT?, in characters</b>	1024 characters.
method of selecting <b>3.2.4.1 User input device</b>	only 1 device supported
method of selecting <b>3.2.4.2 User output device</b>	by default the UART and the screen are used in parallel for output. The UART can be switched on/off.
methods of dictionary compilation <b>(3.3 The Forth dictionary)</b>	Dictionary compilation method: subroutine threaded with inlining of most primitives, compounding and some pruning of useless opcodes.
number of bits in one address unit <b>(3.1.3.3 Addresses)</b>	32 bits
number representation and arithmetic <b>(3.2.1.1 Internal number representation)</b>	Numbers are internally represented with least significant byte first and can be 32 or 64 bits long.
ranges for <i>n</i> , <i>+n</i> , <i>u</i> , <i>d</i> , <i>+d</i> , and <i>ud</i> <b>(3.1.3 Single-cell types, 3.1.4 Cell-pair types)</b>	<i>n</i> : -2147483648 to 2147483647 <i>+n</i> : 0 to 2147483647 <i>u</i> : 0 to 4294967295 <i>d</i> : -9223372036854775808 to 9223372036854775807 <i>+d</i> : 0 to 9223372036854775807 <i>ud</i> : 0 to 18446744073709551615
read-only data-space regions <b>(3.3.3 Data space)</b>	All data-space regions have read/write characteristics
size of buffer at <b>6.1.2450 WORD (3.3.3.6 Other transient regions)</b>	1024 characters
size of one cell in address units <b>(3.1.3 Single-cell types)</b>	4
size of one character in address units <b>(3.1.2 Character types)</b>	1
size of the keyboard terminal input buffer <b>(3.3.3.5 Input buffers)</b>	1024 characters
size of the pictured numeric output string buffer <b>(3.3.3.6 Other transient regions)</b>	255 characters
size of the scratch area whose address is returned by <b>6.2.2000 PAD (3.3.3.6 Other transient regions)</b>	4095 characters

REQUIRED CATEGORY OF DOCUMENTATION	IMPLEMENTATION
system case-sensitivity characteristics (3.4.2 Finding definition names)	WabiPi4 is case-insensitive.
system prompt (3.4 The Forth text interpreter, 6.1.2050 QUIT)	<p>The depth of each of the three user-stacks (data, float, return) is printed after each linefeed in QUIT.</p> <p>like this: (0 0 0) ok</p> <p>For efficient debugging, negative numbers are possible for the depth of the stacks.</p>
type of division rounding (3.2.2.1 Integer division, 6.1.0100 */, 6.1.0110 */MOD, 6.1.0230 /, 6.1.0240 /MOD, 6.1.1890 MOD)	Division is symmetric with truncation as rounding-method
values of 6.1.2250 STATE when true	1
values returned after arithmetic overflow (3.2.2.2 Other integer operations)	After an overflow, usually the lower 32 bits of the calculation are returned.
Whether the current definition can be found after 6.1.1250 DOES> (6.1.0450 :)	A definition can only be found after successful execution of a semicolon.

## AMBIGUOUS CONDITIONS – SYSTEM ACTION

REQUIRED CATEGORY OF DOCUMENTATION	SYSTEM ACTION
a name is neither a valid definition name nor a valid number during text interpretation (3.4 The Forth text interpreter)	The system aborts with a message specifying which word was unknown.
a definition name exceeded the maximum length allowed (3.3.1.2 Definition names)	The system aborts with a message specifying which word was too long
addressing a region not listed in 3.3.3 Data Space	<p>All data-regions in the system are available to the user without exception.</p> <p>If non-existent memory is addressed, a hardware-generated exception occurs and a message is printed specifying the illegal address causing the exception.</p>
argument type incompatible with specified input parameter, e.g., passing a flag to a word expecting an n (3.1 Data types)	WabiPi4 does not check for type-compatibility of input-arguments with the following exception: for T0 & +T0 wabiPi4 checks if the name after T0/+T0 is a VALUE or a 2VALUE
attempting to obtain the execution token, (e.g., with 6.1.0070 ', 6.1.1550 FIND, etc.) of a definition with undefined interpretation semantics	Execution-tokens for any definition can be obtained.

REQUIRED CATEGORY OF DOCUMENTATION	SYSTEM ACTION
dividing by zero	The system aborts with a message specifying that a division by zero occurred.
insufficient data-stack space or return-stack space (stack overflow)	The system allows for a limited amount of overflow. No checks are done by the wabiPi4-system.
insufficient space for loop-control parameters	No checks are done by the system. The stacks have space for 2048 cells each.
insufficient space in the dictionary	No checks are done by the system. The maximum size of the dictionary is 983 MB. In addition there is 2880 MB data-space, addressed by <b>ALLOCATE</b> and <b>TEMPALLOCATE</b> . Both words do not allocate memory if the space is insufficient for the request and return a <b>true</b> as warning.
interpreting a word with undefined interpretation semantics	No checks are done by the system.
modifying the contents of the input buffer or a string literal (3.3.3.4 Text-literal regions, 3.3.3.5 Input buffers);	By design, the system has read/write access to any data-region.
overflow of a pictured numeric output string	This overflow will not happen under normal circumstances in the wabiPi4 system. No actual checks are presently done.
parsed string overflow	The system aborts with a message specifying the cause of the abort
producing a result out of range, e.g., multiplication (using *) results in a value too big to be represented by a single-cell integer	No checks are done by the system. The result is most likely the lower 32 bits of the answer, but this is not guaranteed.
reading from an empty data stack or return stack (stack underflow)	The system allows for (a limited amount of) negative depths and the system will continue functioning. No specific checks are done by the wabiPi4-system.
unexpected end of input buffer, resulting in an attempt to use a zero-length string as a name	The system aborts with a message specifying that a name was expected.

## SPECIFIC AMBIGUOUS CONDITIONS – SYSTEM ACTION

REQUIRED CATEGORY OF DOCUMENTATION	IMPLEMENTATION
>IN greater than size of input buffer (3.4.1 Parsing)	Under normal circumstances >IN cannot become larger than the buffer. No checks are done by the system
RECURSE appears after 6.1.1250 DOES>	The system will abort with a clarifying message.

REQUIRED CATEGORY OF DOCUMENTATION	IMPLEMENTATION
Argument input source different than current input source for 6.2.2148 <b>RESTORE-INPUT</b>	<b>RESTORE-INPUT</b> returns a true flag on failure, no <b>ABORT</b> occurs.
data space containing definitions is de-allocated (3.3.3.2 Contiguous regions)	No checks are done by the system.
data space read/write with incorrect alignment (3.3.3.1 Address alignment)	Address-alignment is mandatory for floating-point related words and for ARM-opcodes. The system will abort with an exception if problems with these occur. Address-alignment is technically not needed for other values. Normal functioning of the system ensures alignment automatically.
data-space pointer not properly aligned (6.1.0150 , , 6.1.0860 C ,)	<b>COMMA</b> can handle unaligned addresses. However ARM-opcodes must be aligned, so if <b>COMMA</b> is used to write opcodes on unaligned addresses, executing the resulting code will result in an unrecognised opcode with a hardware generated prefetch-exception. The system will abort and print a message which address caused the exception.
less than u+2 stack items (6.2.2030 <b>PICK</b> , 6.2.2150 <b>ROLL</b> )	No checks on this are done by the system. Items below the lowest stack-position can be <b>PICK</b> 'd or <b>ROLL</b> 'd.
loop-control parameters not available	No checks on this are done by the system.
most recent definition does not have a name	Using a zero length string for a name result in an abort with a clarifying message.
name not defined by 6.2.2405 <b>VALUE</b> used by 6.2.2295 <b>TO</b>	The system aborts with a message specifying which is the offending name.
name not found (6.1.0070 ', , 6.1.2033 <b>POSTPONE</b> , 6.1.2510 '[' ] )	The system aborts with a message specifying which is the offending name.
parameters are not of the same type (6.1.1240 <b>DO</b> , 6.2.0620 ? <b>DO</b> , 6.2.2440 <b>WITHIN</b> )	No checks on this are done by the system.
6.1.2033 <b>POSTPONE</b> or 6.2.2530 [ <b>COMPILE</b> ] applied to 6.2.2295 <b>TO</b>	In wabiPi4 <b>TO</b> can be postponed and the containing definition can be used to put a value into a <b>VALUE</b> . [ <b>COMPILE</b> ] is not implemented in wabiPi4.
string longer than a counted string returned by 6.1.2450 <b>WORD</b>	Technically and logically impossible in wabiPi4 as a counted string has a 32b index.
u greater than or equal to the number of bits in a cell ( 6.1.1805 <b>LSHIFT</b> , 6.1.2162 <b>RSHIFT</b> )	0x0 is returned, which is logically correct.
word not defined via 6.1.1000 <b>CREATE</b> (6.1.0550 > <b>BODY</b> , 6.1.1250 <b>DOES&gt;</b> )	No checks on this are done by the system. In wabiPi4 the words <b>CREATE</b> and <b>DOES&gt;</b> must be within the same definition.

REQUIRED CATEGORY OF DOCUMENTATION	IMPLEMENTATION
words improperly used outside 6.1.0490 <# and 6.1.0040 #> (6.1.0030 #, 6.1.0050 #S, 6.1.1670 HOLD, 6.1.2210 SIGN)	The words #, #S and HOLD can be used outside of <# and #> without problems occurring. No spurious filling of the buffer occurs, but there are stack-effects.

## OTHER SYSTEM DOCUMENTATION

REQUIRED CATEGORY OF DOCUMENTATION	IMPLEMENTATION
list of non-standard words using 6.2.2000 PAD (3.3.3.6 Other transient regions)	PAD is not used by any (standard or non-standard) word in the dictionary of wabiPi4.
operator's terminal facilities available	--
program data space available, in address units	1031385952 bytes
return stack space available, in cells	2048 cells – Wabipi4 technically separates the return-stack into 2 stacks, one for the CPU and one for the user, to optimize the branch-prediction system of the CPU. Both return-stacks are 2048 cells deep.
stack space available, in cells	2048 cells
system dictionary space required, in address units	380060 bytes for the overall system. For details see the chapter on the <b>memory-map</b> of the system.

# Air Traffic Controller

```

(*
Air Traffic Controller
Based on a game by Dave Mannering for Creative Computing 1978-1981
This original wabiPi4 version (C) 2022 - J.J. Hoekstra
Requires wabiPi4 2.5.0 or higher
28852 bytes 184 definitions ( ?? bytes noinline )
*)

forget atcshield : atcshield ; word# unused

\ constants ****
25 constant xdir
15 constant ydir
27 constant #plane \ 0=no_plane, 1=plane A, etc.
16 constant column

\ 0 constant pstat \ { state-engine plane (0/1/2/3) - 0=inactive - 1=proceed,
    2=waiting
\ 1 constant ppj \ prop/jet - (0/1)
\ 2 constant phgt \ height (0-15)
\ 3 constant porig \ origin (0-11) - 1 of 10 entry-points or 1 of 2 airfields
\ 4 constant pdest \ destination (0-11) - cannot be full rndm
\ 5 constant pxpos \ x (0-24) - the origin - ditto
\ 6 constant pypos \ y (0-14) - ditto - ditto
\ 7 constant pdir \ direction (0-7)
\ 10 constant pstrt \ aligned 16 bit number - start number 99*4 -> meer als 8 bits
\ 12 constant pxnew \ newx (0-24) - the destinaton position
\ 13 constant pynew \ newy (0-15) - ditto
\ 14 constant phnew \ new height (0-15) -ditto }

0 constant oexit
1 constant oxpos
2 constant oypos
3 constant odirc

12 constant origplane
4 constant origcol

\ values **** {
30 value #mins
#mins 4* value #steps

0 value escflag
0 value opd4plane
0 value opdracht
0 value getal
0 value microstep
0 value prop_phase
0 value presentstep
2 value screenwrite
5 value screenshow
0 value linesprinted
0 value dodemo

\ string literals ****
s" N NEE SES SWW NW" $literal $directions

\ arrays ****
create [planes] #plane column * allot
create [orig] origplane origcol * allot
create $input 16 allot
create $seed 16 allot
\ }

```

```

\ routines ****
: defwins  \
    \ window 2 - field A of radar
        790 496 2 makewin
        gray 2 >wincanvas
        2 wincls
        3 2 >winrand          \ grijze rand
translucent 2 >wincanvas
        2 wincls
        white 2 >winink
        true 2 winvisible      \ translucent

\ window 5 - field B of radar
    790 496 5 makewin
    gray 5 >wincanvas
    5 wincls
    3 5 >winrand          \ grijze rand
translucent 5 >wincanvas
    5 wincls
    white 5 >winink
    false 5 winvisible     \ translucent

    2 0 win#>task#

\ windows 1 - copyright etc
    false 1 winvisible

\ window 0 - achtergrond
    black 0 >wincanvas
    white 0 >winink
    0 wincls
    0 winhome

\ window 3 - lijst met active viegtuigen
    230 496 3 makewin
    794 0 3 >winorig
    vlyellow 3 >wincanvas
    3 wincls
    4 3 >winrand
translucent 3 >wincanvas
    3 wincls
    black 3 >winink
    true 3 winvisible

\ window 4 - orders
    610 264 4 makewin           \ was 790 264
    0 500 4 >winorig
    gray 4 >wincanvas
    4 wincls
    3 4 >winrand          \ grijze rand
translucent 4 >wincanvas
    4 wincls
    7 4 >winrand          \ marge voor duidelijker text
    white 4 >winink
    true 4 winvisible

\ window 6 - various
    410 264 6 makewin
    614 500 6 >winorig
    gray 6 >wincanvas
    6 wincls
    3 6 >winrand          \ grijze rand
translucent 6 >wincanvas
    6 wincls
    white 6 >winink

```

```

        true 6 winvisible
; \ }
: showscreen2                                \ { for double-buffering
    2 to screenshow 5 to screenwrite
    true 2 winvisible false 5 winvisible ;

: showscreen5                                \ for double-buffering
    5 to screenshow 2 to screenwrite
    true 5 winvisible false 2 winvisible ;

: switchscreens                                \ synced screen-flip
    begin scr_sync until                      \ flip if screen sync = 'true'
        screenwrite 2 = if showscreen2
        else showscreen5 then ;                \ }
: switchpropphase prop_phase 1 xor to prop_phase ;

0 value timera
: resettimer ( -- ) centis to timera ;
: gettimer ( -- n ) centis timera - ;

: win4 4 0 win#>task# ;
: win6 6 0 win#>task# ;
: win4emit win4 emit win6 ;
: win4bs win4 backspace win6 ;
: win4cr win4 cr win6 ;

: *[planes] ( p col -- addr ) swap column * [planes] + + ;
: plane@   ( p col -- n )   *[planes] sc@ ;
: plane!    ( n p col -- )   *[planes] c! ;
: 16plane@ ( p col -- n )   *[planes] h@ ;
: 16plane!  ( n p col -- )   *[planes] h! ;

: stat@    ( p -- dest ) 0 plane@ ;          \ { = pstat
: stat!    ( p -- dest ) 0 plane! ;

: prop@    ( p -- dir ) 1 plane@ ;           \ = ppj
: prop!    ( n p -- ) 1 plane! ;

: height@  ( p -- height ) 2 plane@ ;         \ = phgt
: height!  ( n p -- ) 2 plane! ;

: orig@    ( p -- dest ) 3 plane@ ;           \ = porig
: orig!    ( d p -- ) 3 plane! ;

: dest@    ( p -- dest ) 4 plane@ ;           \ = pdest
: dest!    ( d p -- ) 4 plane! ;

: xpos@    ( p -- dest ) 5 plane@ ;           \ = pxpos
: xpos!    ( d p -- ) 5 plane! ;

: ypos@    ( p -- dest ) 6 plane@ ;           \ = pypos
: ypos!    ( d p -- ) 6 plane! ;

: direc@   ( p -- dir ) 7 plane@ ;            \ = pdir
: direc!   ( n p -- ) 7 plane! ;

: xnew@    ( p -- dir ) 12 plane@ ;           \ = pxnew
: xnew!    ( n p -- ) 12 plane! ;

: ynew@    ( p -- dir ) 13 plane@ ;           \ = pynew
: ynew!    ( n p -- ) 13 plane! ;

: pstart@  ( p -- start ) 10 16plane@ ;       \ = pstrt
: pstart!  ( p -- start ) 10 16plane! ;

: newhght@ ( p -- height ) 14 plane@ ;         \ = phnew

```

```

: newhght! ( n p -- ) 14 plane! ;           \ }

: newxy@ ( plane -- x y ) >r r@ xnew@ r> ynew@ ;
: xy@ ( plane -- x y ) >r r@ xpos@ r> ypos@ ;

: planestat! opd4plane stat! ;

: planeactive? ( p -- t/f ) stat@ 0<> ;
: planemoving? ( p -- t/f )
    stat@ dup 0<> if
        2 5 within if
            false
        else
            true
        then
    else
        drop false
    then
;

: .plchar ( p -- ) 64 + emit space ;
: >orig origcol * [orig] + >r r@ 3 + c! r@ 2 + c! r@ 1 + c! r> c! ;
: clear[planes] [planes] #plane column * 0 fill ;

: fill[orig] \ {
\ cn x y dr #
  9 0 8 2 0 >orig
  8 5 0 4 1 >orig
  7 10 0 3 2 >orig
  6 18 0 4 3 >orig
  5 0 3 3 4 >orig
  4 11 14 7 5 >orig
  3 18 14 0 6 >orig
  2 24 14 7 7 >orig
  1 5 14 0 8 >orig
  0 24 8 6 9 >orig
  0 10 8 6 10 >orig
  0 14 4 7 11 >orig
  0 18 8 6 12 >orig
  0 5 8 0 13 >orig ; \ }

create movetbl \ {
  0 c, -1 c, \ N
  1 c, -1 c, \ NE
  1 c, 0 c, \ E
  1 c, 1 c, \ SE
  0 c, 1 c, \ S
  -1 c, 1 c, \ SW
  -1 c, 0 c, \ W
  -1 c, -1 c, \ NW }

: (xstep) ( dir -- step )
    dup 0 8 within if 2* movetbl + sc@
    else cr ." illegal value in (xstep): " . abort then ;
: (ystep) ( dir -- step )
    dup 0 8 within if 2* movetbl + 1+ sc@
    else cr ." illegal value in (ystep): " . abort then ;

: clear$seed $seed 5 32 fill ;
: fill$seed cr ." please enter a 5 character SEED: " $seed 5 accept drop ;
: $>seed clear$seed fill$seed $seed 5cval ;
: setseeds $>seed dup rndmseed ( rndm32 space . ) ;
: origtbl@ ( orig ocol -- n ) swap origcol * [orig] ++ sc@ ; \ test of sc@

\ {
: (coord2xy) ( x y -- sx sy )
    30 * 34 + swap 30 * 34 + swap ;
: (.line) ( x y x y -- )

```

```

green 0 >winink 0 winline white 0 >winink ;
: .alllines 5 0 do
    i xpos origtbl@      i ypos origtbl@      \ start line
    (coord2xy) 1+ swap 1+ swap
    9 i - xpos origtbl@  9 i - ypos origtbl@ \ end line
    (coord2xy) 1+ swap 1+ swap
    (.line) loop ;
: (.bbox) ( x y -- ) 4 - swap 4 - swap 11 11 0
    black 0 >winink drawbox white 0 >winink ; hidden
: (.bdot) ( x y -- ) (coord2xy) 2dup (.bbox) 3 3 0 drawbox ; hidden

: .alldots ydir 0 do xdir 0 do i j (.bdot) loop loop ; hidden
: .bluelines lblue 0 >winink 24 1 do
    805 i 20 * 19 + 208 0 horline loop ; hidden
: (.1char) ( char orig -- ) \ prints char airfields/navi-beacon
    >r r@ xpos origtbl@
    r> ypos origtbl@
    (coord2xy)
    7 - swap 3 - swap
    0 winchar ; hidden
: .4chars 130 10 (.1char) 131 11 (.1char) 132 12 (.1char) 132 13 (.1char) ; hidden
: .entries 10 0 do
    i 48 + i (.1char)
    loop ;
: .whitesheet vlyellow 0 >winink 798 0 228 494 0 drawbox .bluelines
    ; hidden
: .paintfield black 0 >winink 0 wincls
    .alllines .alldots .4chars .entries .whitesheet
    ; hidden
: .origdest ( o/d -- )
    dup 0 10 within if ._ else
        10 = if ." #" else ." %" then then
    ;
: .drctn ( d -- ) 2* $directions drop + 2 type
    ; \ }
: .plane ( number -- )                                \
: .planel ( number -- )                               \
    >r
    r@ 64 + emit                                     \ =plane-character without space
    r@      prop@
        0= if ." P"
        else ." J"
        then
    r> height@ .hex
    ;                                                 \ height-character }
: .planel ( number -- )                               \
    >r
    r@      .plane
    r@ orig@   .origdest
        ." ->"
    r@ dest@   .origdest
        space
    r@ xpos@   .2d_
        [char] .
    r@ ypos@   .2d_
        space
    r@ direc@  .drctn

    rdrop
    ;
: backtonormal                                     \
    0 0 win#>task# vdcyan 0 >wincanvas
    white 0 >winink rstwins cls
    ;
: drawsqrplane ( x y )                           \
    black screenwrite >winink
    8 - swap 10 - swap 19 17 screenwrite drawbox
    ;

```

```

        white screenwrite >winink
        ;
: (mcctx) ( x p - new_x )
    dup direc@ (xstep) >r
    prop@
    if
        r> microstep * +
    else
        prop_phase if
            r@ microstep *
            r> 30 * + 2 / +
        else
            r> microstep * 2 / +
    then then
    ;
: (mcsty) ( y p - new_y )
    dup direc@ (ystep) >r
    prop@
    if
        r> microstep * +
    else
        prop_phase if
            r@ microstep *
            r> 30 * + 2 / +
        else
            r> microstep * 2 / +
    then then ;
: (addmicrosteps) ( x y i -- new_x new_y )
    >r r@ (mcsty) swap r> (mcctx) swap
    ;
: (drawplane) ( x y plane# -- )
    >r
    2dup drawsqrplane
    2dup r@ 64 + -rot
    9 - swap 9 - swap
    screenwrite winchar
    9 - r> height@ 48 + -rot
    screenwrite winchar
    ;
: .planesfield
    screenwrite wincls
    black screenwrite >wincanvas
    #plane 1 do
        i planeactive? if
            i xy@ (coord2xy)
            i planemoving? if
                i (addmicrosteps)
            then
                i (drawplane)
            then
        loop
        translucent screenwrite >wincanvas
        ;
: myinput ( -- len )
    $input 4+ 8 accept dup $input !
    ;
: checkmin ( -- 0/value )
    0 $input >number 0<> if drop 0 else then
    ;
: input#min ( -- 18-99 ) ." How long should the game last? (18 - 99 min): "
    myinput 0= if 25 else checkmin then \ default of 25 min on return
    18 99 clip 4* to #steps
    ;
s" ALRHMP\$%" sliteral orders$
\ : makeopdracht ( char -- ) \ codes the tasks 1-9 with 0 for invalid
\     0 to opdracht

```

```

\ >caps
\ orders$ bounds do
\   dup i c@ = if
\     i orders$ drop - 1+ to opdracht
\     leave
\   then
\ loop drop
\ ;
: makeopdracht ( char -- )           \ codes the tasks 1-9 with 0 for invalid
  >caps orders$ instring 1+ to opdracht
  ;
: checknumber ( char -- f/t )        \ also sets getal
  >caps 48 -
  dup 0 6 within if
    to getal true
  else
    drop -1 to getal false
  then
  ;
\ ** state engine input-handler
\ 0 -> 0 chars - waiting for A-Z for plane - goto 1 - bs not possible
\ 1 -> 1 chars - waiting for "A, L, R" for opdracht -> goto 2 - bs -> goto 0
\           Hold, Maintain, Proceed, Status, approach_#, approach_%, -> goto 3
\ 2 -> 2 chars - waiting for 1-5 for opdracht -> goto 3 - bs -> goto 1
\ 3 -> 2/3 chars received - do opdracht, cr, goto 1 - bs -> goto 2

0 value is=
: input_handler ( key -- )
  dup 8 = if                                \ backspace!
  is= 1 = if
    win4bs
    0 to is=
  then
  is= 2 = if
    win4bs
    1 to is=
  then
  drop exit
  then
  dup 0 33 within if drop exit then       \ exit on all ctrl chars and space

  >r r@ win4emit
  is= 0 = if r> 64 - to opd4plane 1 to is= exit then
  is= 2 = if r> 48 - to getal      3 to is= exit then
  is= 1 = if r> makeopdracht
    opdracht 4 < if
      2 to is=                      \ if opdracht 1, 2, 3 -> 3 chars input
    else
      3 to is=                      \ if opdracht > 3 -> 2 chars input
    then exit
  then
  ;
: fromairfield ( p -- )
  >r 0 r@ stat!                            \
  40 rndm 10 < if                          \ =waiting at airfield
    11 r@ orig!                           \ 25/75%
  else
    10 r@ orig!                           \ orig=10
  then

  12 rndm r@ dest!                         \ dest 0-11 (inc airfield)
  \ #TBD: bias for busy routes

  0 r@ height!
  0 r> newheight!
  ;
: overflight ( p -- )                     \
  \ {

```

```

>r
10 rndm dup
    r@ orig!
    \ orig 0-9
9 swap - r> dest!
;
: toairfield ( p -- )
    >r
    10 rndm r@ orig!

4 rndm 1 < if
    11 r> dest!
else
    10 r> dest!
then
;
: makeorides ( plane -- )
    >r
    6 rndm dup

2 < if drop
    r> fromairfield
else
    4 > if
        r> overflight
    else
        r> toairfield
    then
then
;
: fillxyd ( plane -- )
    >r
    r@ orig@

dup oxpos origtbl@ r@ xpos!
dup oypos origtbl@ r@ ypos!
odirc origtbl@ r@ direc!
rdrop
;
0 value step*100 0 value rndmfact
: correctstart ( -- )
    1 pstart@
#plane 2 do
    dup
    i pstart@
    > if
        dup i pstart!
    else
        drop i pstart@
    then
loop drop ;
\ }

0 value tempstart 0 value temporig 0 value tempplane
: goback ( pl relevante_orig startt -- height ) \ {
    to tempstart
    to temporig
    to tempplane

1 tempplane 1- do
    temporig i orig@ = if
        tempstart
        i pstart@
        - 13 <
        if
            i height@
            1+ dup
            tempplane height!
\ negative do can have equal inputs
\ origs equal?
\ starttime of plane being checked
\ starttime of lower plane
\ calculate diff and check
\ if < 13 (=3 min)
\ if smaller -> get hight of lower plane
\ add 1 to it and ready

```

```

        tempplane newhght!
    then
        leave
    then
        -1 +loop
    ;
: checkheights ( -- )
    #plane 2 do
        i orig@ >r
        r@ 0 10 within if
            i
            r@
            i pstart@
            goback
            startttimes
            then rdrop
        then
    loop
    ;
: fillstarttime
    #steps 60 - 100 * 26 / to step*100
    step*100 95 100 */ to rndmfact
    0 1 pstart!

    #plane 1- 2 do
        i step*100 *
        rndmfact rndm
        2 rndm 0 = if - else + then
        100 / i pstart!
    loop
    step*100 #plane * 100 / 26 pstart!
    correctstart
        number
    checkheights
    ;
\ : testdata
\  20 6 stat!
\  40 8 stat!
\  30 22 stat!
\  40 19 stat!
\  11 19 dest! ;
: defplanes
    \ setseeds
    clear[planes]
    #plane 1 do      0 i stat!
        3 rndm 0= if 1 i prop! then
            6 i height!
            6 i newhght!
            i makeorides
            i fillxyd
    loop
    fillstarttime
    ( testdata ) ;
: gameinit
    rndmreset
    defwins
    fill[orig]
    defplanes
    .paintfield
    uartclear

    2 to screenwrite
    5 to screenshow

    6 0 win#>task#
    0 to presentstep
\ setup of double buffering
\ 6=other 4=entry screen

```

```

;
: .upcommingplanes          \ }
    2 to linesprinted
    3 winhome
    3 0 win#>task#
    true 0 uart>task#

    ." upcomming planes    " cr
    #plane 1 do
        i pstart@
        presentstep -
        1 3 within if           \ 1 5 -> planes next minute
            i space .planel cr
            1 +to linesprinted
        then
    loop
    linesprinted 2 = if
        1 spaces ." ---" 16 spaces
    else
        20 spaces
    then
    win6
    false 0 uart>task#
    ;
: .activeplanes             \ {
    3 0 win#>task#
    true 0 uart>task#

    2 +to linesprinted
    cr ." active planes      "
    #plane 1 do
        i stat@ 0 > if
            cr i space .planel
            1 +to linesprinted
        then
    loop

    linesprinted 24 < if
        25 linesprinted do
            cr 20 spaces loop
    then
    win6
    false 0 uart>task#
    ;
: ***** from here game logic *****
: getxstep ( p -- step ) direc@ 2* movetbl + sc@ ;
: getystep ( p -- step ) direc@ 2* movetbl + 1+ sc@ ;
: revdir ( dir -- dir+180graad ) 4 + 7 and ;
: (leftdir) ( dir -- dir+180graad ) 1 - 7 and ;
: (rightdir) ( dir -- dir+180graad ) 1 + 7 and ;
: leftdir ( p -- )
    >r r@ direc@ (leftdir) r> direc! ;
: rightdir ( p -- )
    >r r@ direc@ (rightdir) r> direc! ;
: (makenewxy) ( plane -- ) >r
    r@ getxstep r@ xpos@ +
    r@ xnew!                                \ new x
    r@ getystep r@ ypos@ +
    r> ynew!                                \ new y
    ;
: makenewxy ( plane -- ) >r
    prop_phase r@ prop@ + if
        plane
        r@ (makenewxy)

```

```

    then rdrop                                \ 0=> altijd prop met prop phase 0
    ;
: boundaries? ( x y -- f/t )                \ true if x, y is outside of boundaries
    0 ydir within swap 0 xdir within and 0=
    ;
: planeatexit ( p -- f/dest t )
    >r
    r@ dest@ dup                                \ 2*destination of plane
    xpos origtbl@ r@ xpos@ =                  \ compare x of plane and x of dest
    swap
    ypos origtbl@ r@ ypos@ =                  \ ditto for y
    and if                                         \ combine flags
        r> dest@ true
    else
        rdrop false
    then
    ;
: correctexit? ( p -- f/t )                 \ { check correct x, y and direction
    >r r@ planeatexit if
        odirc origtbl@ revdir
        orig
        r> direc@ =
        r> ( r_dir ) -> reversed entry direction of
    else
        rdrop false
    then
    ;
: deactivateplane ( p -- ) 0 swap stat! ;
: activateplane   ( p -- )                   \ checked
    airfield autohold=state=30
    >r r@ orig@ 10 < 2 + r@ stat!           \ for planes from orig to
                                                \ 1=active or 2=waiting at an
                                                \ airport

    r@ orig@ 10 <
    r@ dest@ 9 > and if
        dest > 9 -> autohold
        30 r@ stat!
    then
    rdrop
    ;
: adaptheight ( p -- )
    >r
    r@ planemoving? 0= if                    \ exit if plane does not move
        rdrop exit
    then

    r@ newheight@ r@ height@ -
    dup 0> if
        r@ height@ 1+ r> height! drop exit
    then
    0< if
        r@ height@ 1- r> height! exit
    then
    rdrop
    ;
: preplanes ( -- ) \ { new xy in destination active planes & activates planes
    #plane 1 do                                \ for all planes
        i stat@ 0= if
            i pstart@ presentstep = if
                ." activating plane "          \ debug
                i .plchar cr                \ debug
                i activateplane
            then
        then

        i planemoving? if                  \ only moving planes
            i makenewxy

```

```

    then

    i stat@ 3 = if                                \ for 3=permission to start
        prop_phase 0= if                           \ wait till prop-phase=0
            1 i stat!
            1 i height!
        then
    then

        i stat@ 4 = if                                \ for 4=permission to start
            3 i stat!                               \ goto 3=permission to start
        then
    loop
;
: updatexy ( p -- )                                \ { newxy of moving planes to xpos/ypos
    >r r@ newxy@ r@ ypos! r> xpos!
;
: field#? ( x y -- f/t ) 8 = swap 10 = and 0<>; 
: field%? ( x y -- f/t ) 4 = swap 14 = and 0<>; 

: abortlanding? ( p -- )
    >r r@ stat@ 40 =
    r@ height@ 0 = or if                         \ trying to land? (state=40 or height=0)
        1 r@ stat!
        1 r@ height!
        1 r@ newheight!
        ." plane " r@ .plchar ." aborting landing" cr
    then rdrop
;
: landed? ( p -- f\t )
    >r
    r@ planemoving? 0= if                         \ not landing if not moving
        rdrop false exit then

    r@ xy@ field#? if                            \ west direc=6
        r@ direc@ 6 = r@ height@ 0 =
        and r@ dest@ 10 = and 0<> if
            rdrop true exit
        else
            r@ abortlanding?
            rdrop false exit
        then
    then

    r@ ( newxy@ ) xy@ field%? if                  \ north west direc=7
        r@ direc@ 7 =
        r@ height@ 0 =
        and
        r@ dest@ 11 =
        and 0<> if
            rdrop true exit
        else
            r@ abortlanding?
            rdrop false exit
        then
    then
        rdrop false                                \ }

: (landed) ( p -- )
    dup deactivateplane
    ." plane " .plchar ." landed" cr
;
: handleplanestep ( p -- f\t )                   \ checks landing, exit and updates XY {
    >r
    r@ landed? if
        r> (landed) true exit then

```

```

r@ newxy@ boundaries? if          \ if boundaries passed with updated xy
    r@ correctexit?           \ passed at valid exit with correct dir...
    r@ height@ 5 =           \ passes at valid height?
    and if
        r@ deactivateplane
        ." plane " r@ .plchar ." left your area" cr true
    else drop
        r@ deactivateplane      \ debugging
        ." BOUNDARY ERROR plane " r> .plchar cr false exit
    then
then
rdrop true                      \ }

;

: checkplanes ( -- f/t )          \ { check for
#plane 1 do
    i planeactive? if
        i handleplanestep 0= if
            i deactivateplane
            false ( unloop exit ) \ debugging
        then
    then
loop true
;
;

: updateplanes ( -- )             \ {
#plane 1 do
    prop_phase i prop@ + if      \ 0=prop/1=jet -> if 1 or 2 -> do plane
        i planemoving? if        \ only moving planes
            i updatexy
        then
    then
loop
;
\ }

( state_engine ) \ *****
\ (*
\   0 -> not active - goto 0
\   1 -> proceed - goto 1
\   2 -> wait for take-off at airfield - goto 2
\   3 -> prepare for takeoff in prophase=1 - goto 1
\   4 -> prepare for takeoff in prophase=0 - goto 3
\   10 -> turn 45 left - goto 1
\   11 -> turn 90 left - goto 10
\   12 -> turn 135 left - goto 11
\   13 -> turn 180 left - goto 12
\   14 -> turn 45 right - goto 1
\   15 -> turn 90 right - goto 14
\   16 -> turn 135 right - goto 15
\   17 -> turn 180 right - goto 16
\   20 -> at * turn sharp to # airfield - goto 20 or 1
\   21 -> at * turn sharp to % airfield - goto 21 or 1
\   29 -> circle 45 left - goto 29
\   30 -> at * start to circle - goto 30 ( => * not yet reached ) or 29
\   40 -> goto height=0 and start landing - goto 40 - goto 1 on wrong landing
\ *) */

: dolanding ( p -- )
    >r r@ height@ 1- 0 max r@ height! \ lowers height with 1
    0 r> newheight! ;                \ newheight always 0
: (@navbeacon?) ( x y -- t/f )     \ { 5,8 - 18,8
    8 = if dup 5 = swap 18 = or 0<>
    else drop false then ; \ #checked }
: @navbeacon? ( p -- )              \ state=30
    >r r@ xy@ (@navbeacon?) if
        r@ leftdir                  \ start circling
        29 r@ stat!                 \ state engine to 29

```

```

    then rdrop ;
: gotofield# ( p -- )                                \ state=20
    >r r@ xy@ (@navbeacon?) if
        6 r@ direc!
        40 r@ stat!
    then rdrop ;
: gotofield% ( p -- )                                \ state=21
    >r r@ xy@ (@navbeacon?) if
        7 r@ direc!
        40 r@ stat!
    then rdrop ;
: t180l ( p -- )                                     \ { state=13
    >r r@ leftdir 12 r> stat! ;
: t135l ( p -- )                                     \ state=12
    >r r@ leftdir 11 r> stat! ;
: t90l ( p -- )                                      \ state=11
    >r r@ leftdir 10 r> stat! ;
: t45l ( p -- )                                      \ state=10
    >r r@ leftdir 1 r> stat! ;
: t180r ( p -- )                                     \ state=17
    >r r@ rightdir 16 r> stat! ;
: t135r ( p -- )                                     \ state=16
    >r r@ rightdir 15 r> stat! ;
: t90r ( p -- )                                      \ state=15
    >r r@ rightdir 14 r> stat! ;
: t45r ( p -- )                                      \ state=14
    >r r@ rightdir 1 r> stat!
    ;
: updatedoing ( plane -- )                           \ {
    >r r@ stat@
        dup 40 = if drop r> dolanding exit then \ state=40 can never be undone
        dup 30 = if drop r> @navbeacon? exit then \ start circling when at *
        dup 29 = if drop r> leftdir exit then \ continue circling
        dup 20 = if drop r> gotofield# exit then \ if on navbeacon -> direction = west
        dup 21 = if drop r> gotofield% exit then \ if on navbeacon -> direction = west
        dup 13 = if drop r> t180l exit then
        dup 12 = if drop r> t135l exit then
        dup 11 = if drop r> t90l exit then
        dup 10 = if drop r> t45l exit then
        dup 17 = if drop r> t180r exit then
        dup 16 = if drop r> t135r exit then
        dup 15 = if drop r> t90r exit then
        dup 14 = if drop r> t45r exit then
        drop rdrop
    ;
: updatestateengine ( -- )                           \ {
    #plane 1 do
        prop_phase i prop@ + if
            i planeactive? if
                i updatedoing
                i adaptheight
            then
                then
                    loop
    ;
: startmodule ( -- f/t )                            \ {
    win4 cls uartclear
    ." play game? (y/n)"
    key [char] n = if
        backtonormal true exit
    then false win6
    ;
: userinput ( -- f/t )
    win4 cr input#min
    fill$seed gameinit win6
    ;

```

```

: endmodule ( -- f/t )
    win4 uartclear
    cr ." play another game? (y/n)"
    key [char] n = if
        true
    else
        gameinit false
    then win6
    ;                                \
        : .debuginput
    win4
    space ." -> ." pl: " opd4plane .
    ." - opd: " opdracht .
    ." - num: " getal .
    win6
    ;
: checkopdracht ( -- f/t )           \ false on inactive plane
    .debuginput

    opd4plane planeactive? 0= if
        win4 5 spaces ." -----" win6      \ <= "dead air" response
        0 to opd4plane
        false exit
    then

    getal 0 6 within 0= if
        win4 5 spaces ." SAY AGAIN?" win6
        false exit
    then

    opdracht 1 =
    getal 0= and
    opd4plane stat@ 29 = and if
        win4 5 spaces ." UNABLE - HOLDING" win6
        false exit
    then

    opdracht 1 =
    getal 0= and
    opd4plane planemoving? 0= and if      \ no landing if waiting to start
        win4 5 spaces ." WAITING TO START" win6
        false exit
    then

    opdracht 1 =
    getal 0<> and
    opd4plane stat@ 40 = and if          \ cleared for landing
        win4 5 spaces ." UNABLE - LANDING" win6 \ dead-air is also appropriate
        false exit
    then

\ opd4plane stat@ 2 5 between
    opd4plane planemoving? 0=             \ if waiting
    opdracht 1 <> and if                \ no other task than altitude
        win4 5 spaces ." UNABLE - WAITING" win6
        false exit
    then

    win4 space ." ROGER!" win6          \ ROGER mogelijk later printen
    true
    ;
: doopdracht ( -- )
    opdracht 1 = if                    \ 1=altitude
        getal opd4plane newhght!
    opd4plane stat@ 2 = if

```

```

opd4plane prop@ 0=
prop_phase      0= and if
    4 planestat!                                \ if prop=0 and prop_phase=0 -> status=4
    else
        3 planestat!                            \ otherwise status=3
    then
        opd4plane (makenewxy)
    then

getal 0= if
    40 planestat! then
        \ move status to landing

exit
then

opdracht 2 = if
    getal 1 5 within if
        getal 9 + planestat! then
    getal 0 = if
        30 planestat! then                      \ hold at nav-beacon
    getal 5 = if
        20 planestat! then                      \ at nav-beacon turn to airfield #
    exit
then

opdracht 3 = if
    getal 1 5 within if
        getal 13 + planestat! then
    getal 0 = if
        1 planestat! then                      \ proceed
    getal 5 = if
        21 planestat! then                      \ at nav-beacon turn to airfield %
    exit
then

opdracht 4 = if
    30 planestat! then                        \ hold=start circling at nav-beacon

opdracht 5 7 within if
    1 planestat! then                        \ maintain and proceed

\ opdracht 7 = if .status then            \ status

opdracht 8 = if
    20 planestat! then                      \ proceed to airfield #

opdracht 9 = if
    21 planestat! then                      \ proceed to airfield %

;
: entrymodule ( -- )
key? if
    key >caps
    input_handler
    is= 3 = if
        checkopdracht if
            doopdracht
        then
            win4cr
            0 to is=
        then
    then
    ;
: gameloop
    gameinit
    startmodule if exit then
begin

```

```

userinput
0 to prop_phase
uartclear cr
#steps 0 do
    i ." step: " . cr
    i to presentstep
    preplanes                                \ activate plane & update destinations
    .upcommingplanes
    .activeplanes

    checkplanes 0= if
        ( false unloop exit ) then \ debug
    30 0 do                                \ 30 microsteps of ~500ms
        i to microstep
        .planesfield
        resettimer begin
            entrymodule
            gettimer 10 >
        until
        switchscreens
    loop
    updateplanes                            \ destination to source
    updatestateengine                      \ change directions, states, heights etc.
    switchpropphase

    loop
endmodule                                \ f/t
until                                     \ exit on
backtonormal
;

\ end game ****
\ *****

unused -      ." bytes"
word# swap -  ." definitions"

: .pl                                         \ #debug - shows planes
#plane 1 do
    cr 8 spaces i .planel
    2 spaces
    i stat@ .
    i pstart@ .
    i xnew@ .
    i ynew@ .
loop ;

: btn backtonormal ;


```

- end of text -