ARM Assembly Language & Programming

ARM7

Clicker question

Have you installed ARMSim#?

- A. Yes
- B. No

http://armsim.cs.uvic.ca/

B1.5

GNU Assembler Quick Reference

This section summarizes the more useful commands and expressions available with the GNU assembler, gas, when you target this assembler for ARM. Each assembly line has the format

```
{<label>:} {<instruction or directive>} @ comment
```

Unlike the ARM assembler, you needn't indent instructions and directives. Labels are recognized by the following colon rather than their position at the start of the line. The following example shows a simple assembly file defining a function *add* that returns the sum of the two input arguments:

```
.section .text, "x"

.global add @ give the symbol add external linkage

add:

ADD r0, r0, r1 @ add input arguments

MOV pc, lr @ return from subroutine
```

GNU Assembler Directives

Here is an alphabetical list of the more common gas directives.

Assembler directives

```
.align n
.ascii "<string>"
.asciiz "<string>"
.byte <byte1> {, <byte2>, ...}
.data {<addr>}
.global <symbol>
.text {<addr>}
.word <word1> {, <word2>, ...}
```

The ARMSim# User Guide

Table 4. SWI I/O operations (0x00 - 0xFF)

Op	code	Description and Action	Inputs Outputs		EQU
swi	0x00	Display Character on Stdout	r0: the character		SWI_PrChr
swi	0x02	Display String on Stdout	r0: address of a null ter- minated ASCII string	(see also 0x69 below)	
swi	0x11	Halt Execution			SWI_Exit
swi	Allocate Block of Memory on Heap		r0: block size in bytes	r0:address of block	SWI_MeAlloc
swi	0x13	Deallocate All Heap Blocks			SWI_DAlloc
swi	0 x 66	*	r0: file name, i.e. address of a null terminated ASCII string containing the name r1: mode	r0:file handle If the file does not open, a result of -1 is returned	SWI_Open
swi	0x68	Close File	r0: file handle		SWI_Close
swi	0x69	Write String to a File or to Stdout	r0: file handleor Stdout r1: address of a null termi- nated ASCII string		SWI_PrStr

Table 4. SWI I/O operations (0x00 - 0xFF)

Op	code	Description and Action	Inputs	Outputs	EQU
swi	0x6a	Read String from a File		r0: number of bytes stored	SWI_RdStr
swi	0x6b	Write Integer to a File	r0: file handle r1: integer		SWI_PrInt
swi	0x6c	Read Integer from a File	r0: file handle	r0: the integer	SWI_RdInt
swi	0x6d	Get the current time (ticks)	l e	r0: the number of ticks (milliseconds)	SWI_Timer

.data

str: .asciz "Hello World!\n"

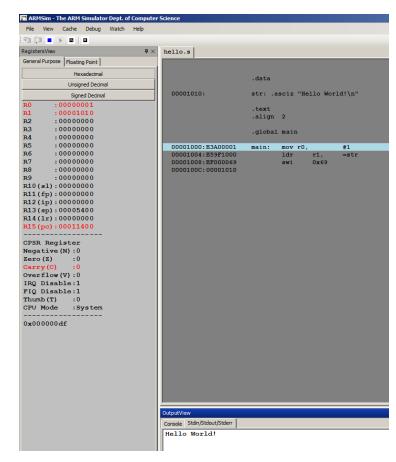
.text .align 2

.global main

main: mov r0, #1

Idr r1, =str

swi 0x69

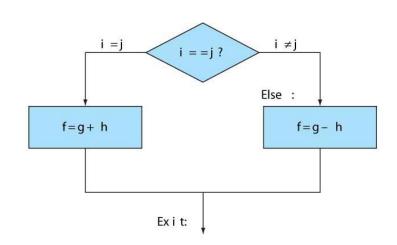


ARM Programming

- ➤ Branch (cmp, beq, bne)
- ♠ "If-then-else"

Example:

if(i==j)
$$f = g + h$$
;
else $f = g - h$;



assume f, g, h, i, and j are in r0, r1, r2, r3, and r4 respectively

CMP r3, r4

BNE Else

ADD r0, r1, r2

B Exit

Else: sub r0, r1, r2

Exit:

A more compact and efficient version:

CMP r3,r4

ADDEQ r0,r1,r2; f = g + h (skipped if $i \neq j$)

SUBNE r0, r1, r2; f = g - h (skipped if i = j)

signed versus unsigned comparison

Suppose

```
r0 = 1111 \ 1111 \ 1111 \ 1111 \ 1111 \ 1111 \ 1111 \ 1111 \ 1111 \ 1111
```

If the following instructions are executed

0x0000 1000: CMP r0, r1 0x0000 1004: BLO L1 0x0000 1008: BLT L2

Where will be the PC at?

A: L1

B: L2

C: 0x0000 100C **D**: 0x0000 1010

Value	Meaning	Value	Meaning
0	EQ (EQual)	8	HI (unsigned Higher)
1	NE (Not Equal)	9	LS (unsigned Lower or Same)
2	HS (unsigned Higher or Same)	10	GE (signed Greater than or Equal)
3	LO (unsigned LOwer)	11	LT (signed Less Than)
4	MI (MInus, <0)	12	GT (signed Greater Than)
5	PL - (PLus, >=0)	13	LE (signed Less Than or Equal)
6	VS (oVerflow Set, overflow)	14	AL (Always)
7	VC (oVerflow Clear, no overflow)	15	NV (reserved)

FIGURE 2.9.5 Encodings of Options for Cond field.

This example is in page 108 - 109, of Chapter 02 COD 4e ARM.pdf).

So, the result of r0 - r1 is computed (see page B1-15 in Appendix_B1.pdf for details) as $r0 + (-r1) = r0 + (\sim r1 + 1)$

```
1111 1111 1111 1111 1111 1111 1111 (r0)
+) 1111 1111 1111 1111 1111 1111 (r1)
```

Δ

11111 1111 1111 1111 1111 1111 1110

Based on this result,

N = 1 (the result is negative, treated as two's complement)

Z = 0 (the result is not zero)

C = 1 (there is carry out of the left-most bit)

V = 0 (there is no overflow)

Therefore, the instruction "BLO" is not taken because of suffix "LO" indicates unsigned lower which is taken when carry bit is clear (See Table B1.2, in Appendix B1.pdf for details). Instead instruction "BLT" is taken when N != V, which is the case as shown above.

Condition Mnemoics

cond	Mnemonic	Name	CondEx
0000	EQ	Equal	Z
0001	NE	Not equal	$ar{Z}$
0010	CS / HS	Carry set / Unsigned higher or same	С
0011	CC / LO	Carry clear / Unsigned lower	Ē
0100	MI	Minus / Negative	N
0101	PL	Plus / Positive of zero	\overline{N}
0110	VS	Overflow / Overflow set	V
0111	VC	No overflow / Overflow clear	$ar{V}$
1000	н	Unsigned higher	ĪC
1001	LS	Unsigned lower or same	Z OR \bar{C}
1010	GE	Signed greater than or equal	$\overline{N \oplus V}$
1011	LT	Signed less than	$N \oplus V$
1100	GT	Signed greater than	$\bar{Z}(\overline{N \oplus V})$
1101	LE	Signed less than or equal	$Z OR (N \oplus V)$
1110	AL (or none)	Always / unconditional	ignored

Value	Meaning	Value	Meaning
0	EQ (EQual)	8	HI (unsigned Higher)
1	NE (Not Equal)	9	LS (unsigned Lower or Same)
2	HS (unsigned Higher or Same)	10	GE (signed Greater than or Equal)
3	LO (unsigned LOwer)	11	LT (signed Less Than)
4	MI (MInus, <0)	12	GT (signed Greater Than)
5	PL - (PLus, >=0)	13	LE (signed Less Than or Equal)
6	VS (oVerflow Set, overflow)	14	AL (Always)
7	VC (oVerflow Clear, no overflow)	15	NV (reserved)

FIGURE 2.9.5 Encodings of Options for Cond field.

TABLE B1.2 ARM condition mnemonics.

<cond></cond>	Instruction is executed when	cpsr condition
{ AL}	ALways	TRUE
EQ	EQual (last result zero)	Z==1
NE	Not Equal (last result nonzero)	Z==0
{CS HS}	Carry Set, unsigned Higher or Same (following a compare)	C==1
{CC LO}	Carry Clear, unsigned LOwer (following a comparison)	C==0
MI	Minus (last result negative)	N==1
PL	PLus (last result greater than or equal to zero)	N==0
VS	V flag Set (signed overflow on last result)	V==1
VC	V flag Clear (no signed overflow on last result)	V==0
HI	unsigned Higher (following a comparison)	C==1 && Z==0
LS	unsigned Lower or Same (following a comparison)	C==0 Z==1
GE	signed Greater than or Equal	N==V
LT	signed Less Than	N!=V
GT	signed Greater Than	N==V && Z==O
LE	signed Less than or Equal	N!=V Z==1
NV	NV NeVer—ARMv1 and ARMv2 only—DO NOT USE	

CMP Compare two 32-bit integers

```
1. CMP<cond> Rn, \#<rotated\_immed> ARMv1
2. CMP<cond> Rn, Rm \{, <shift>\} ARMv1
3. CMP Ln, \#<immed8> THUMBv1
4. CMP Rn, Rm THUMBv1
```

Action

```
    cpsr flags set on the result of (Rn - <rotated_immed>)
    cpsr flags set on the result of (Rn - <shifted_Rm>)
    cpsr flags set on the result of (Ln - <immed8>)
    cpsr flags set on the result of (Rn - Rm)
```

Notes

- In the *cpsr*: $N = \langle Negative \rangle$, $Z = \langle Zero \rangle$, $C = \langle NoUnsigned Overflow \rangle$, $V = \langle Signed Overflow \rangle$. The carry flag is set this way because the subtract x y is implemented as the add $x + \sim y + 1$. The carry flag is one if $x + \sim y + 1$ overflows. This happens when $x \ge y$ (equivalently when $x \hat{A}y$ doesn't overflow).
- If Rn or Rm is pc, then the value used is the address of the instruction plus eight bytes for ARM instructions, or plus four bytes for Thumb instructions.

Example

```
CMP r0, r1, LSR\#2 ; compare r0 with (r1/4) BHS label ; if (r0 >= (r1/4)) goto label;
```

Appendix B1

- > For 2.16.1-2, the values of r0 have a one as their left most
- > bit. Should we assume these values are signed integers or
- > unsigned?

Yes, you assume they are signed integers. The instruction "CMP r0, r1" subtracts value in r0 by value in r1. What matters is the subtraction result, though not saved, based on which the high 4 flag bits NZCV of the current program status register (CPSR) are set accordingly. It is these four bits that determine the behavior of the next conditional instruction.

Loop

Example:

assume i is in r3, k is in r5, and the base of the array is in r6.

Loop: ADD r12, r6, r3, LSL #2

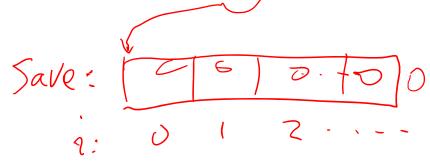
LDR r0, [r12, #0]

CMP r0, r5 BNE Exit

ADD r3, r3, #1

B Loop

Exit:



≻Procedure call and stack

motivation: abstraction and code reusability

e.g., a * b

C code:

while(b != 0) { sum = sum + a; b = b -1; }

ARM Assembly code:

@ assume a is in r0, b is in r1

@ initialize

How do we make use of that piece of code (let's give it a name MUL)?

E.g.,
$$a*b + c$$

solution 1: ("cut-and-paste": embed the code of MUL to where it is needed)
(a valid approach under certain circumstances, inline function in C++)
@ assume a, b, c, are in r0, r1, r7

mov r2, r7 @ move c out of r7, as it is to be used by MUL

@initialize sub r7, r7, r7 r1, #0 Loop: cmp Done \emptyset if b == 0 beq r7, r7, r0 @ sum += a add sub r1, r1, #1 @b = b-1b Loop Done: @add c to a*b add r7, r7, r2

Can we reuse the routine "MUL" multiple times? We don't want to write the same piece of code wherever it is used; we may not know how many times it is to be used when we write the code, e.g., multiply integer *a* by itself for *b* times where b is an integer to be read from the keyboard.

So we want to write the MUL code once and reuse it.

E.g.,
$$a*a + b*b + c*c$$

Solution 2: ("call-by-name": Jump to the MUL code and jump back when it is done)

@ assume a, b, c, are in r2, r3, r4

Done:	mov move b add	r0, r2 r1, r3 MUL r7, r7, r4	@ move a t @ move b t @ call MUL @add c to t	to r1 ., product a*b will be put in r7
MUL:				
	sub	r7, r7, r7	@initialize	
Loop:	cmp	r1, #0		
	beq	Done	@if b == 0	
	add	r7, r7, r0	# sum += a	
	sub	r1, r1, #1	# b = b-1	We want "MUL" to return to
	b	Loop		the next instruction where it is called. In this example, the return address "Done" is

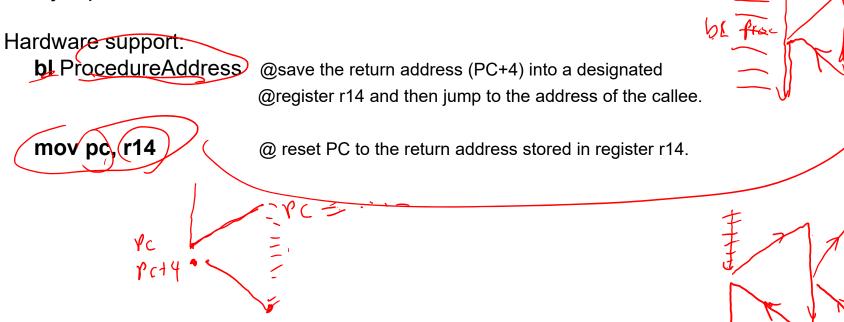
hard coded -- so it can only return to one place.

- @ For example, let's compute a*a + b*b
- @ assume a, b, c, are in r2 and r3

```
mov r0, r2
                             @ copy a to r0
         mov r0, r2
                            @ copy a to r1
         b MUL
                             @ call MUL, product a*a will be in r7
                             @ move a*a to r4
Done:
         mov r4, r7
         mov r0, r3
                            @ copy b to r0
         mov r1, r3
                            @ copy b to r1
         b MUL
                            @ call MUL, product b*b will be in r7
         add r7, r7, r4
                            @ a*a + b*b is put to r7
Done?:
MUL:
                   $v0, $v0, $v0 # initialize
                                                      A hard-coded return address
Loop:
                   $0, $a1, Done # if a == 0
                                                      "Done" is not working well --
                   $v0, $v0, $a0 # sum += a
                                                      it can only return to one
                   $a1, $a1, 1
                                                      place.
                                 # b = b-1
                   Loop
```

Solution:

- 1. save the return address to a register before calling the subroutine
- 2. jump to the saved return address when the subroutine is done.



- @ For example, let's compute a*a + b*b
- @ assume a and b are in r3 and r4

	mov	r0, r3	@ copy a to r0
	mov	r1, r3	@ copy ⊮to r1
	bl	MUL	@ call MUL, address of next instruction is put in r14
~7	mov	r5, r7	@ move a*a to r5
)	mov	r0, r4	@ copy b to r0
	mov	r1, r4	@ copy b to r1
	bl	MUL	@call MUL, address of next instruction is put in r14
	add	r7, r7, r5	@a*a + b*b is put to r7
(MUL:			
	sub	r7, r7, r7	@ initialize
Loop:	cmp	r1, #0	
	beq	Done	@ if a == 0
	add	r7, r7, r0	@ sum += a
	sub	r1, r1, #1	@ $b = b-1$
	b	Loop	
Done:	mov	pc, r14	@ a procedure always ends with this instruction

Six steps

- 1. (caller) place parameters in a location where the procedure (callee) can access them (r0, r1, r2, r3)
- 2. transfer control to the procedure. (BL)

- 3. acquire the storage resources needed for the procedure.
- 4. perform the desired task
- 5. place the result value in a place where the caller can access it.
- 6. return control to the point next to where the program is called. (MOV pc, r14)

ARM conventions:

- r0 - r3, r12: registers for storing arguments or scratch registers to used by the callee (not preserved)

callee

- r4-r11: registers that need to be preserved, if used by callee
- / Ir/ register storing the return address (r14)
- -sp; stack pointer (r13)

Clicker question:

What is a leaf-procedure?

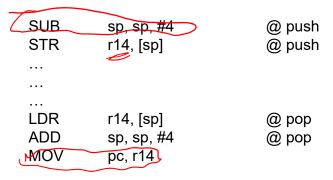
- A. It is a routine you do during the fall to clean up the falling leaves.
- B. It is a procedure that does not call any other procedures
- C. It is a procedure that is not called by any other procedures
- D. It is a procedure that calls other procedures but itself
- E. It is a procedure that calls only itself.

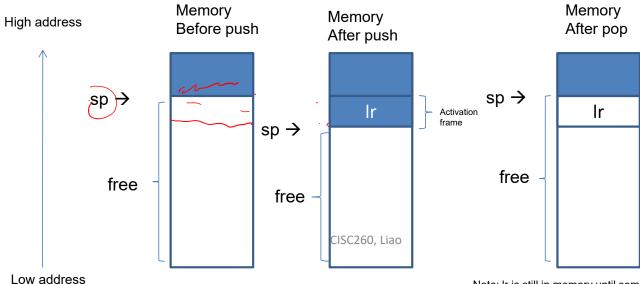
```
What if a procedure calls another procedure?
                                             Addr:
                                    PC \rightarrow
                                                                     proc1
                    proc1(a);
invocation
                    proc1(A) {
                                                   proc1:
                                                            mov
                                                                     pc, r14
                                                                              @r14 has wrong
                                                                                 return address
                                                   proc2:
                    proc2(B) {
                                                                     pc, r14
```

Follow the PC and monitor the value change in r14 CISC260, Liao

Solution: Stack (LIFO) (ast In First Out

Acquire storage space in main memory, the acquired space is called activation frame

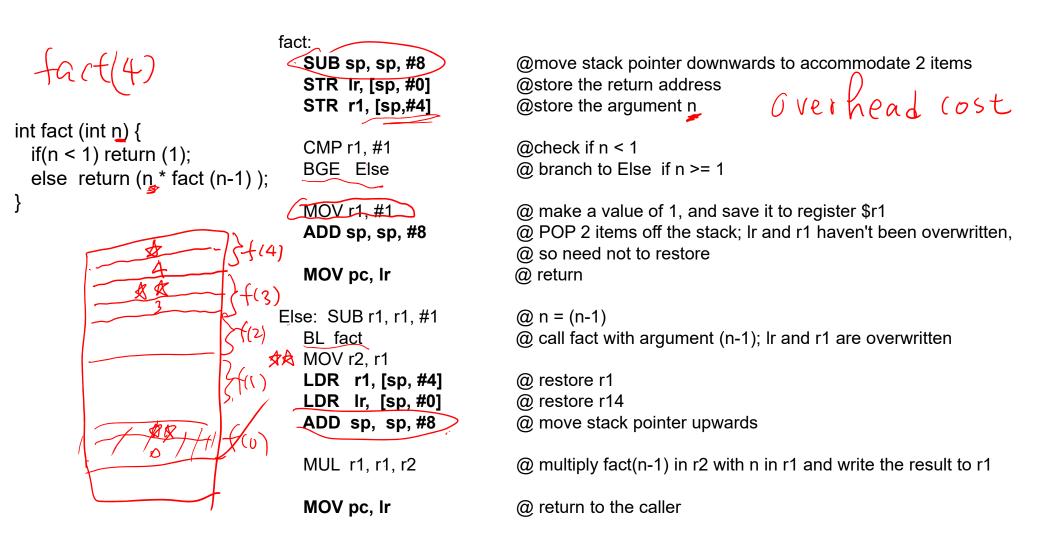




Note: Ir is still in memory until some new value is written in the same location. Potential security loophole.

```
Main() {
  int n = 4;
  int res = fact(n);
  printf(res);
}

int fact (int n) {
  if(n < 1) return (1);
  else return (n * fact (n-1) );
}</pre>
```



Q: The instruction MOV r1, #1 will overwrite what's already in r1. Is that a problem?

A.Yes B.No

```
F(4)
= 4*F(3)
= 4*3*F(2)
= 4*3*2*F(1)
= 4*3*2*1*F(0)
= 4*3*2*1*1
```

Address: 0x 0000 1000 fact: SUB sp, sp, #8 0x 0000 1004 STR Ir, [sp, #0] 0x 0000 1008 STR r1, [sp,#4] 0x 0000 100C CMP r1, #1 BGE Else 0x 0000 1010 MOV r1, #1 0x 0000 1014 0x 0000 1018 ADD sp, sp, #8 0x 0000 101C MOV pc, Ir Else: SUB r1, r1, #1 0x 0000 1020 0x 0000 1024 BL fact 0x 0000 1028 MOV r2, r1 0x 0000 102C LDR r1, [sp, #4] 0x 0000 1030 LDR Ir, [sp, #0] 0x 0000 1034 ADD sp, sp, #8 0x 0000 1038 MUL r1, r1, r2 0x 0000 103C MOV pc, Ir

Note: The following is for the 32-bit ARM7, see Chapter 02_COD 4e ARM

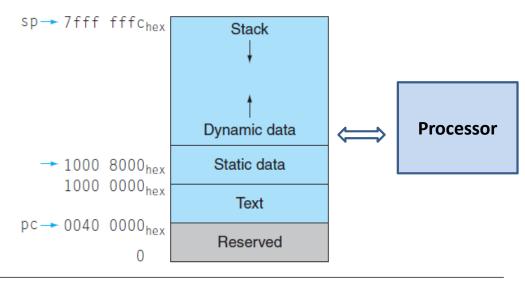
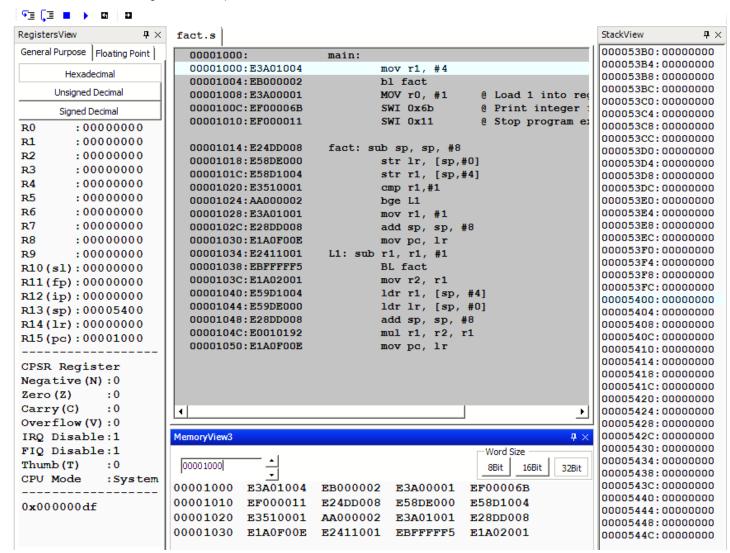
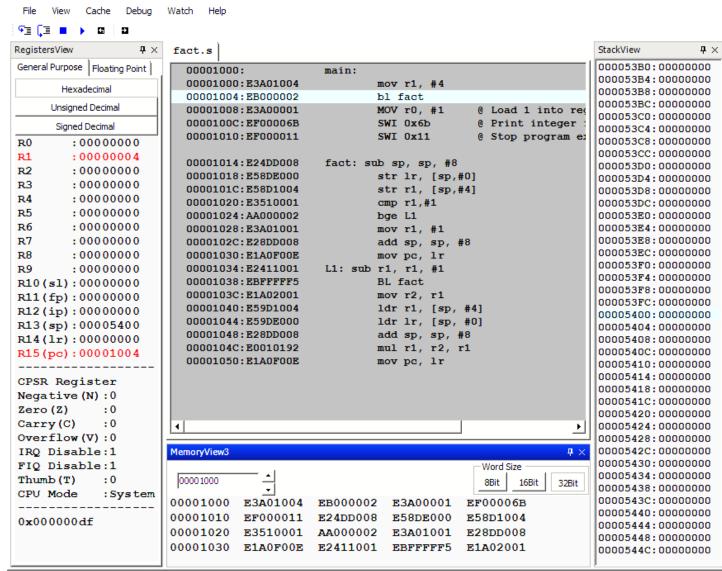
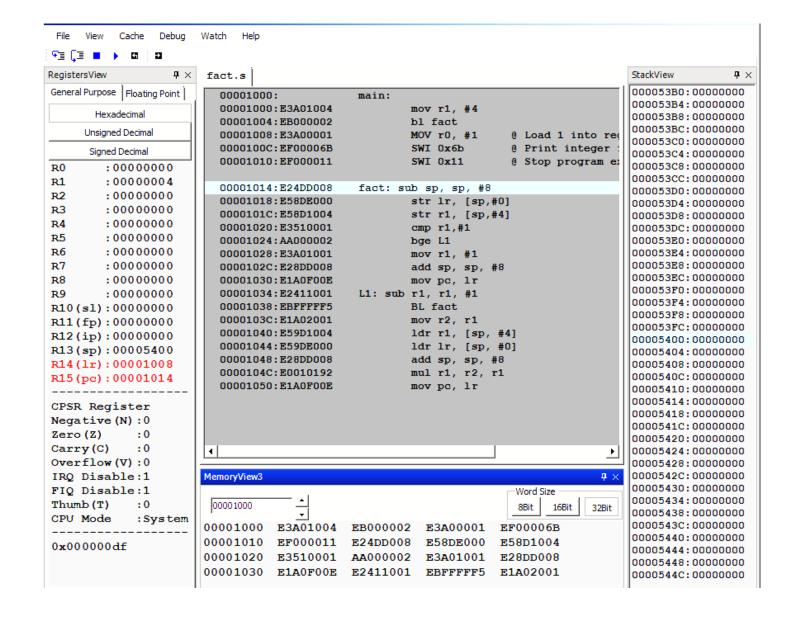


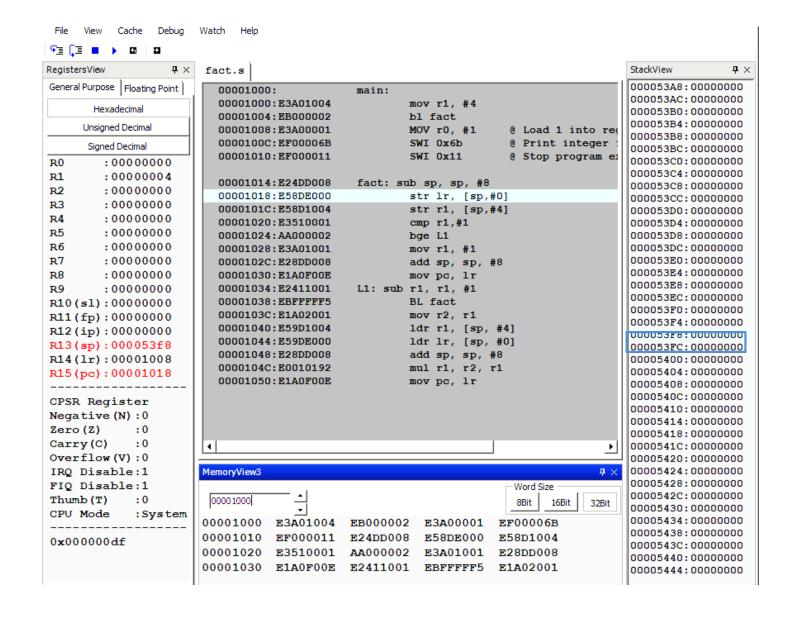
FIGURE 2.13 Typical ARM memory allocation for program and data. These addresses are only a software convention, and not part of the ARM architecture. The stack pointer is initialized to $7fffffc_{hex}$ and grows down toward the data segment. At the other end, the program code ("text") starts at $0040\,0000_{hex}$. The static data starts at $1000\,0000_{hex}$. Dynamic data, allocated by malloc in C and by new in Java, is next. It grows up toward the stack in an area called the heap.

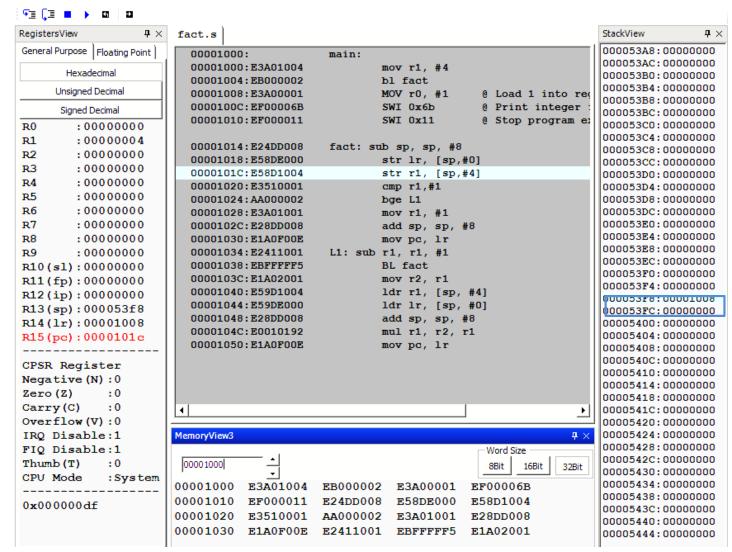
File View Cache Debug Watch Help

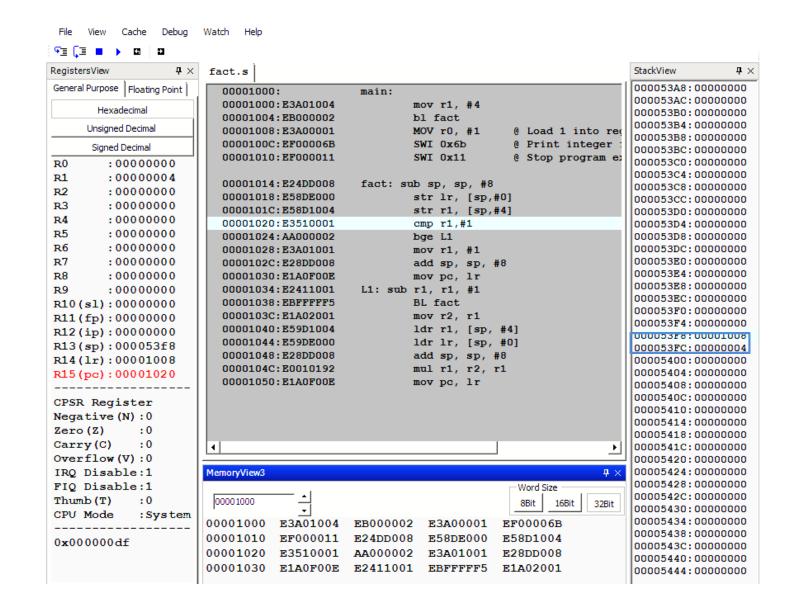


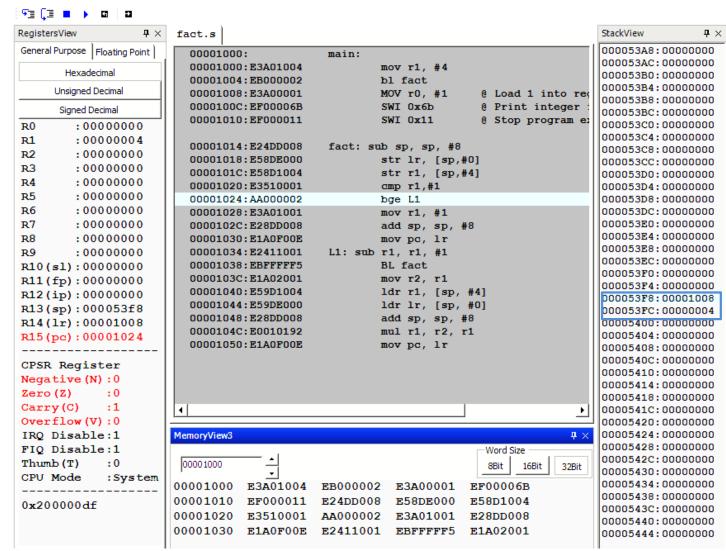


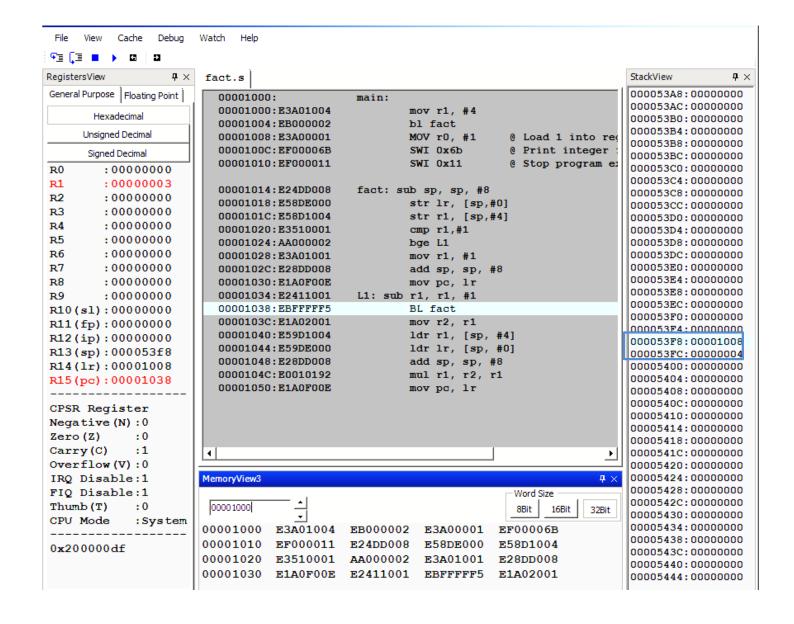


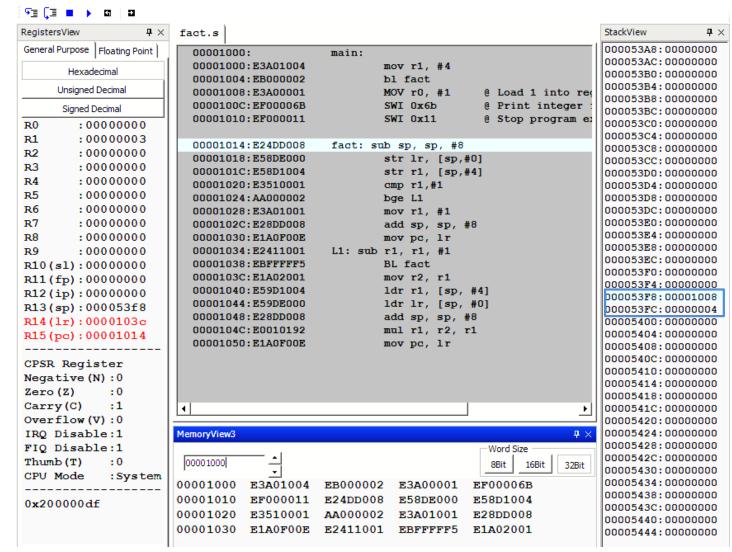


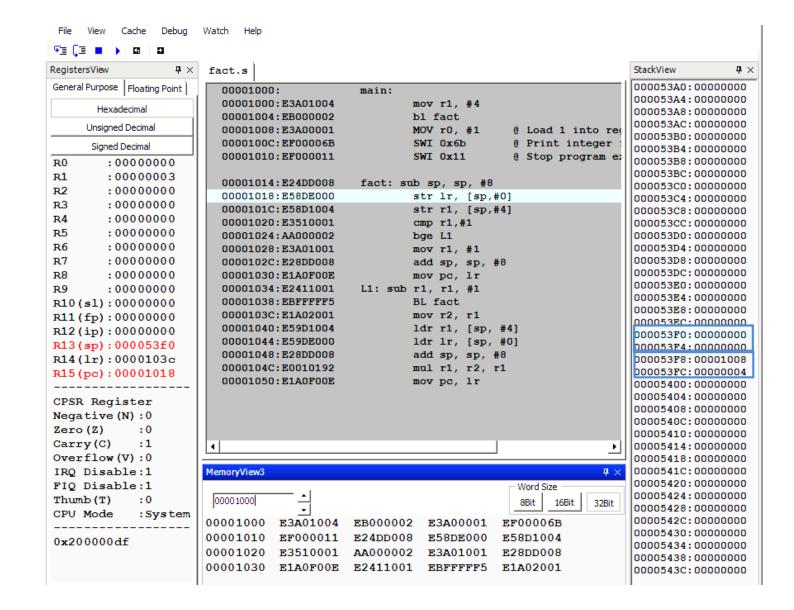


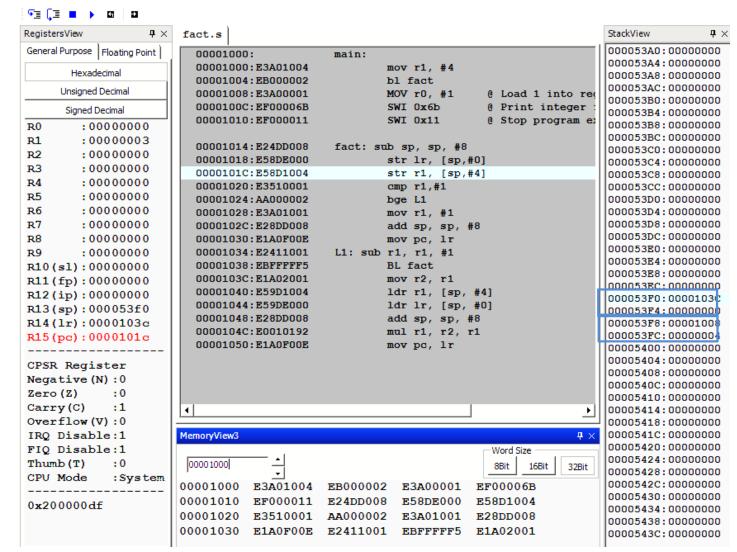


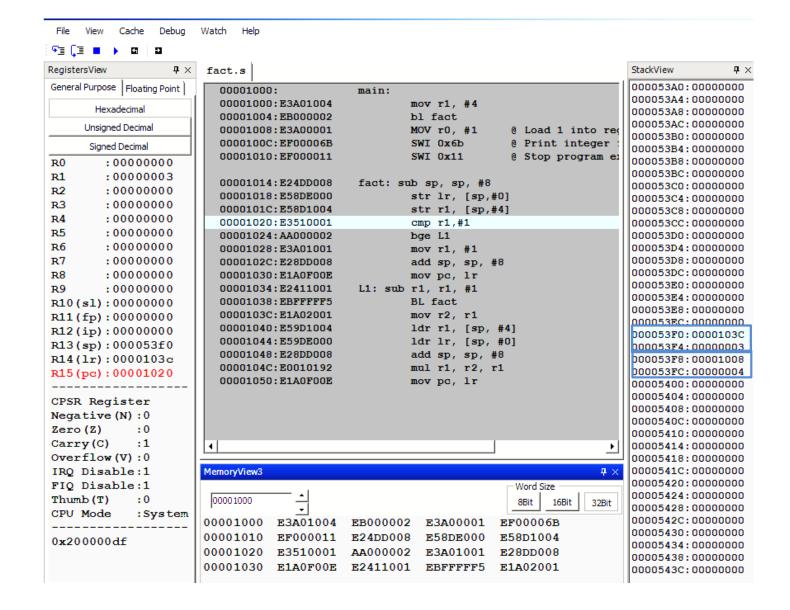


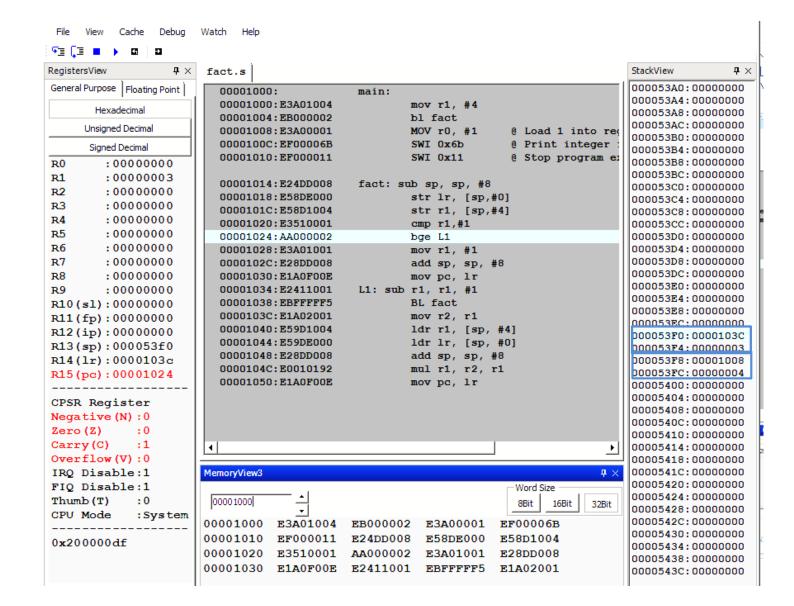


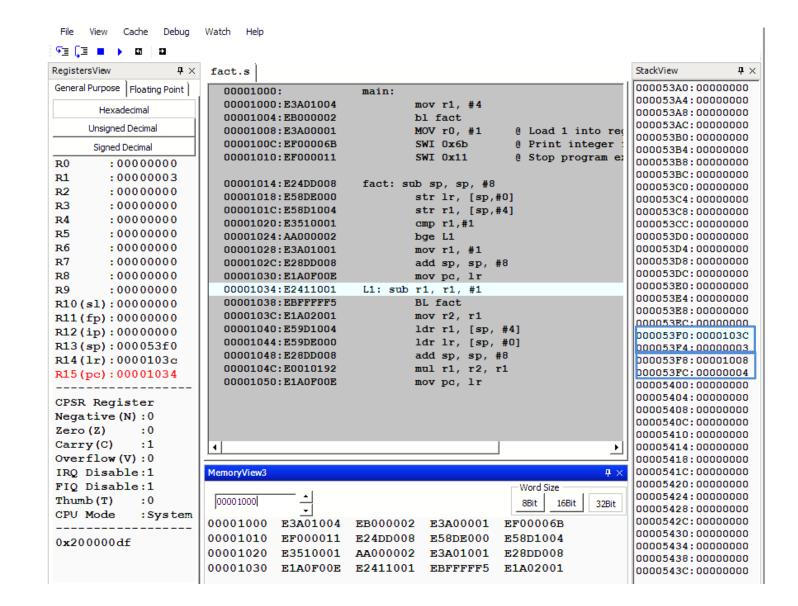










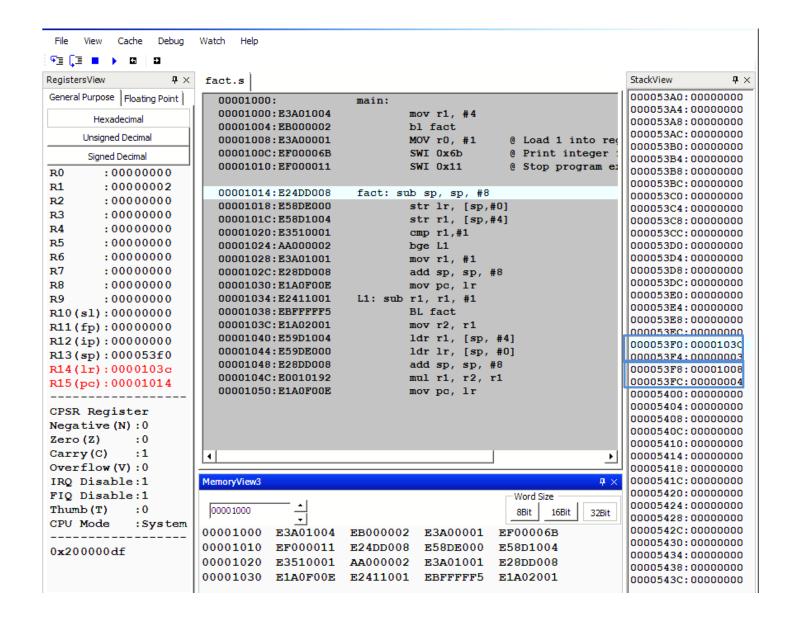


Cache

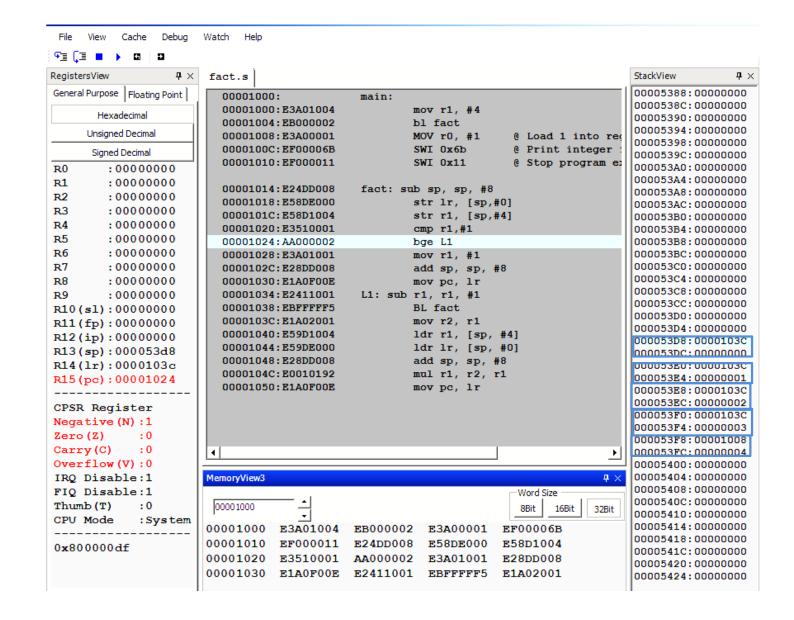
Debug

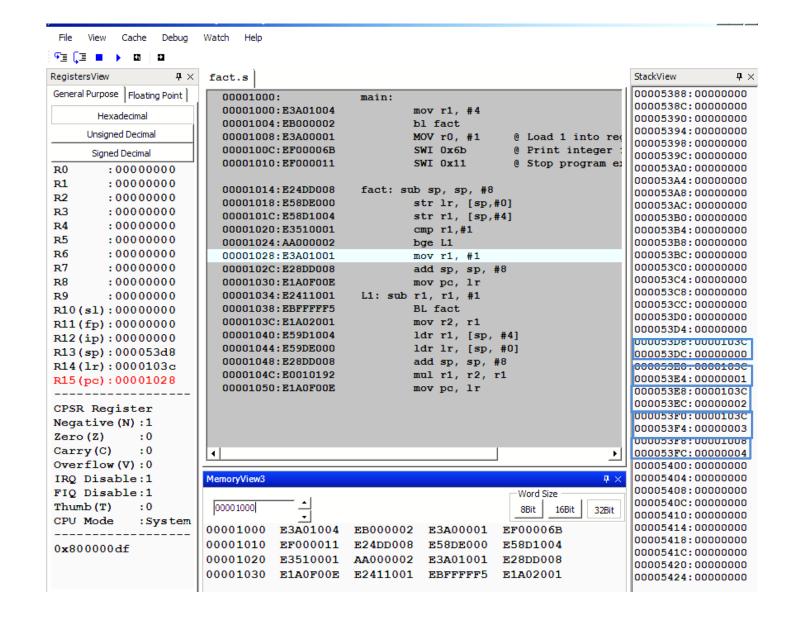
Watch

Help

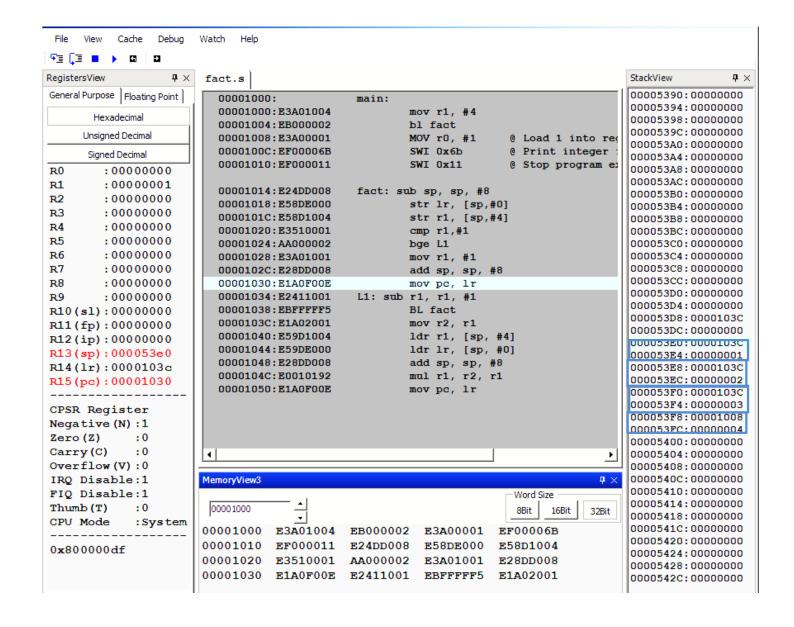


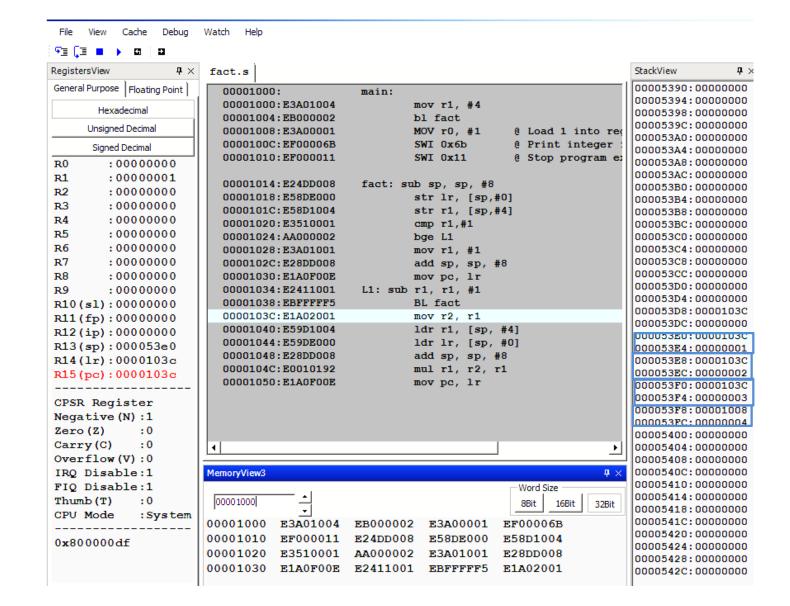
Debug Watch Help RegistersView Ψ× StackView **4** × fact.s General Purpose | Floating Point | 00005388:00000000 00001000: main: 0000538C:00000000 00001000:E3A01004 mov r1, #4 Hexadecimal 00005390:00000000 bl fact 00001004:EB000002 00005394:00000000 Unsigned Decimal 00001008:E3A00001 MOV r0, #1 @ Load 1 into reg 00005398:00000000 0000100C: EF00006B SWI 0x6b @ Print integer Signed Decimal 0000539C:00000000 00001010:EF000011 SWI 0x11 @ Stop program ex 000053A0:00000000 R0 :00000000 000053A4:00000000 R1 :00000000 00001014:E24DD008 fact: sub sp, sp, #8 000053A8:00000000 R2 :00000000 00001018:E58DE000 str lr, [sp,#0] 000053AC: 00000000 R3 :00000000 0000101C:E58D1004 str r1, [sp,#4] 000053B0:00000000 R4 :00000000 00001020:E3510001 cmp r1.#1 000053B4:00000000 R5 :00000000 00001024:AA000002 bae L1 000053B8:00000000 Rб :00000000 000053BC:00000000 00001028:E3A01001 mov r1, #1 **R7** :00000000 000053C0:00000000 0000102C: E28DD008 add sp, sp, #8 000053C4:00000000 R8 :00000000 00001030:E1A0F00E mov pc, lr 000053C8:00000000 L1: sub r1, r1, #1 R9 00001034:E2411001 :00000000 000053CC:00000000 00001038: EBFFFFF5 BL fact R10(s1):00000000 000053D0:00000000 0000103C:E1A02001 mov r2, r1 R11(fp):00000000 000053D4:00000000 00001040:E59D1004 ldr r1, [sp, #4] R12(ip):00000000 000053D8:0000103C 00001044:E59DE000 ldr lr, [sp, #0] R13(sp):000053d8 000053DC:00000000 add sp, sp, #8 00001048:E28DD008 R14(lr):0000103c 000053E0:0000103C mul r1, r2, r1 0000104C:E0010192 000053E4:00000001 R15 (pc): 00001020 00001050:E1A0F00E mov pc, lr 000053E8:0000103C 000053EC:00000002 CPSR Register 000053F0:0000103C Negative (N):0 000053F4:00000003 Zero(Z) :1 000053F8:00001008 Carry (C) :1 4 000053FC:00000004 Overflow (V):0 00005400:00000000 IRQ Disable:1 MemoryView3 00005404:00000000 $\mathbf{p} \times$ 00005408:00000000 FIQ Disable:1 Word Size 0000540C:00000000 Thumb (T) : 0 00001000 8Bit 16Bit 32Bit 00005410:00000000 CPU Mode :System 00005414:00000000 00001000 E3A01004 EB000002 E3A00001 EF00006B 00005418:00000000 00001010 EF000011 E24DD008 E58DE000 E58D1004 0x600000df 0000541C:00000000 00001020 E3510001 AA000002 E3A01001 E28DD008 00005420:00000000 00001030 E1A0F00E E2411001 EBFFFFF5 E1A02001 00005424:00000000

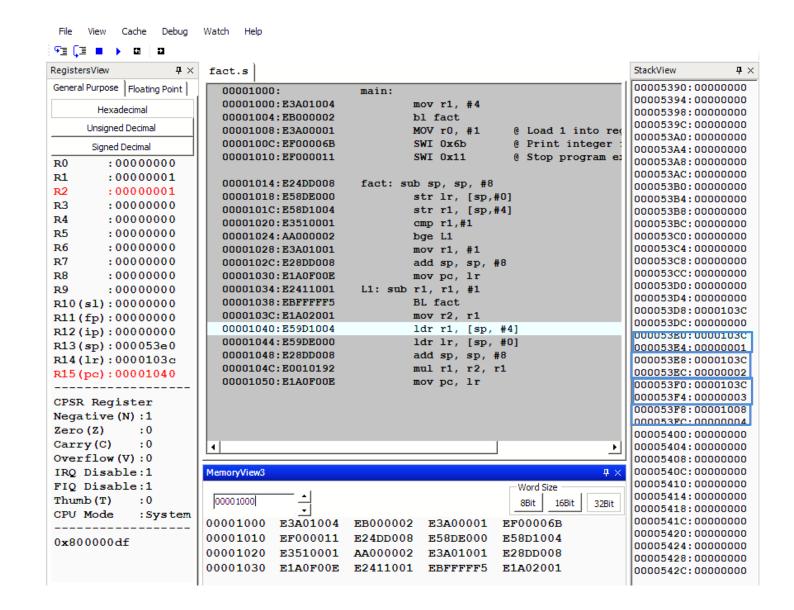


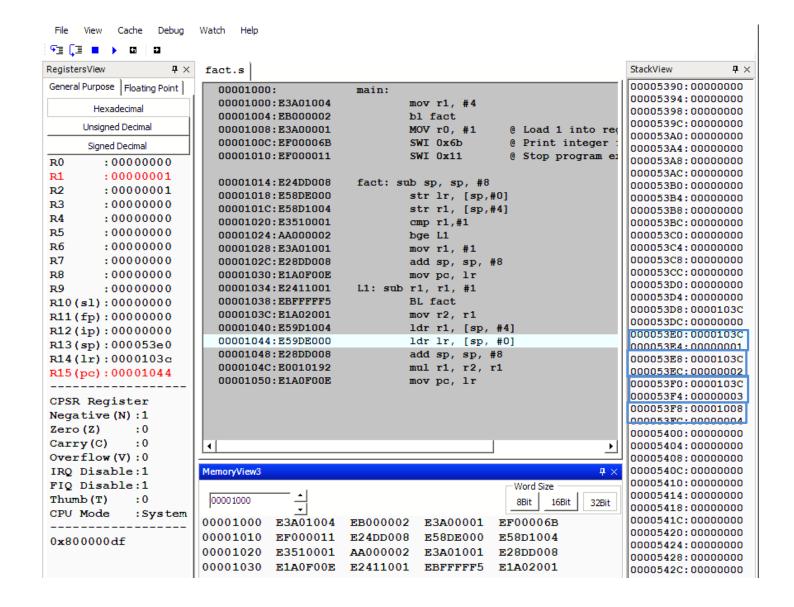


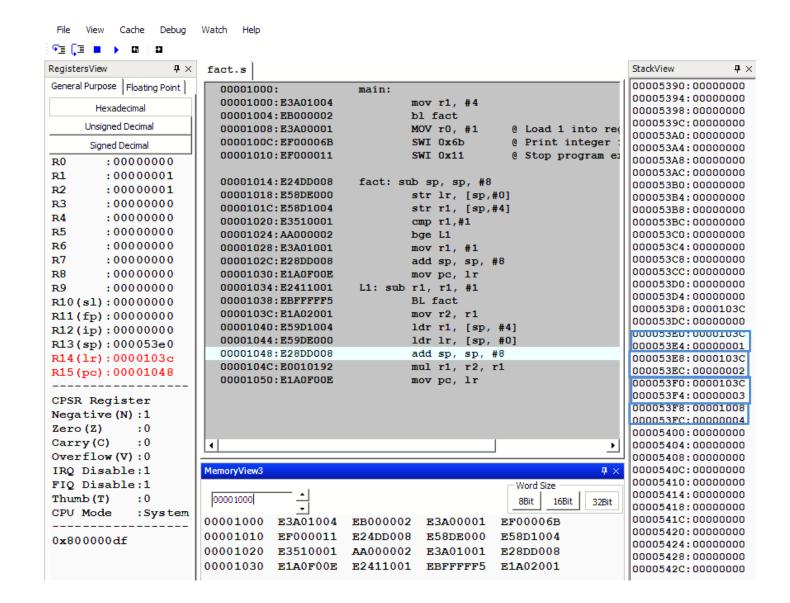
View Cache Debug Watch Help RegistersView $\mathbf{p} \times \mathbf{l}$ StackView $\mathbf{p} \times$ fact.s General Purpose | Floating Point 00005388:00000000 00001000: main: 0000538C:00000000 00001000: E3A01004 mov r1, #4 Hexadecimal 00005390:00000000 00001004:EB000002 bl fact 00005394:00000000 Unsigned Decimal MOV r0, #1 @ Load 1 into red 00001008:E3A00001 00005398:00000000 0000100C:EF00006B SWI 0x6b Print integer Signed Decimal 0000539C:00000000 00001010:EF000011 SWI 0x11 @ Stop program ex 000053A0:00000000 R0:00000000 000053A4:00000000 R1 :00000001 00001014:E24DD008 fact: sub sp, sp, #8 000053A8:00000000 R2 :00000000 str lr, [sp,#0] 00001018:E58DE000 000053AC: 00000000 R3 :00000000 0000101C: E58D1004 str r1, [sp,#4] 000053B0:00000000 R4 :00000000 00001020:E3510001 cmp r1,#1 000053B4:00000000 R5 :00000000 00001024:AA000002 bge L1 000053B8:00000000 R6 :00000000 000053BC:00000000 00001028:E3A01001 mov r1, #1 000053C0:00000000 **R7** :00000000 0000102C:E28DD008 add sp, sp, #8 000053C4:00000000 R8 :00000000 00001030:E1A0F00E mov pc, lr 000053C8:00000000 R9 00001034:E2411001 L1: sub r1, r1, #1 :00000000 000053CC: 00000000 00001038: EBFFFFF5 BL fact R10(s1):00000000 000053D0:00000000 mov r2, r1 0000103C:E1A02001 R11(fp):00000000 000053D4:00000000 00001040:E59D1004 ldr r1, [sp, #4] R12(ip):00000000 D00053D8:0000103C 00001044:E59DE000 ldr lr, [sp, #0] R13(sp):000053d8 000053DC: 00000000 add sp, sp, #8 00001048:E28DD008 R14(lr):0000103c D00053E0:0000103C mul r1, r2, r1 0000104C:E0010192 000053E4 • 00000001 R15 (pc): 0000102c 00001050:E1A0F00E mov pc, lr 000053E8:0000103C 000053EC:00000002 CPSR Register 000053F0:0000103C Negative (N):1 000053F4:00000003 Zero(Z) 000053F8:00001008 Carry (C) : 0 4 000053FC:00000004 Overflow (V):0 00005400:00000000 IRQ Disable:1 MemoryView3 00005404:00000000 00005408:00000000 FIQ Disable:1 Word Size 0000540C:00000000 Thumb (T) 00001000 8Bit 16Bit 32Bit 00005410:00000000 CPU Mode :System E3A01004 00005414:00000000 EB000002 E3A00001 EF00006B 00001000 00005418:00000000 00001010 EF000011 E24DD008 E58DE000 E58D1004 0x800000df 0000541C:00000000 00001020 E3510001 AA000002 E3A01001 E28DD008 00005420:00000000 E2411001 00001030 E1A0F00E EBFFFFF5 E1A02001 00005424:00000000

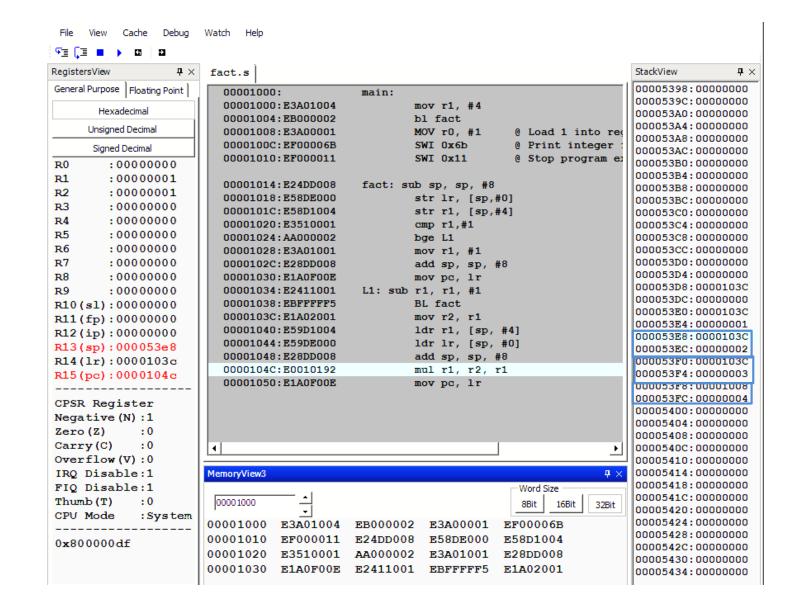


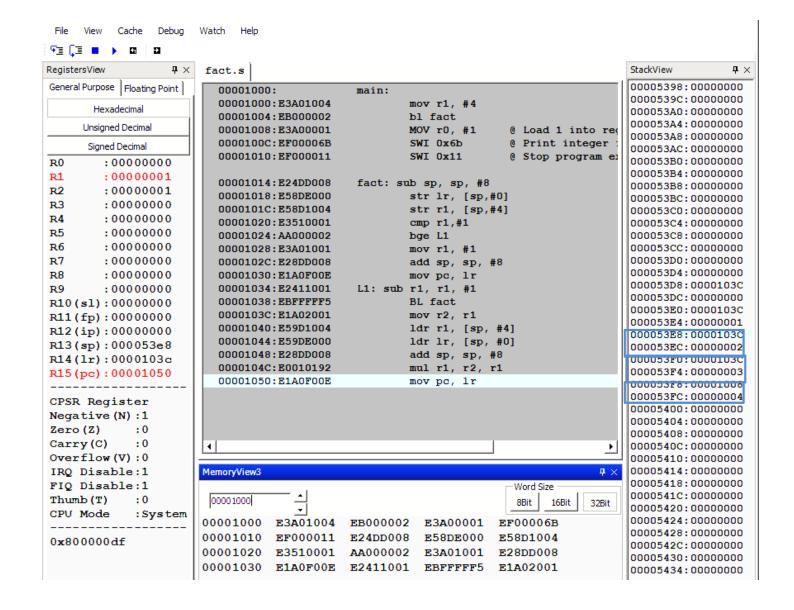


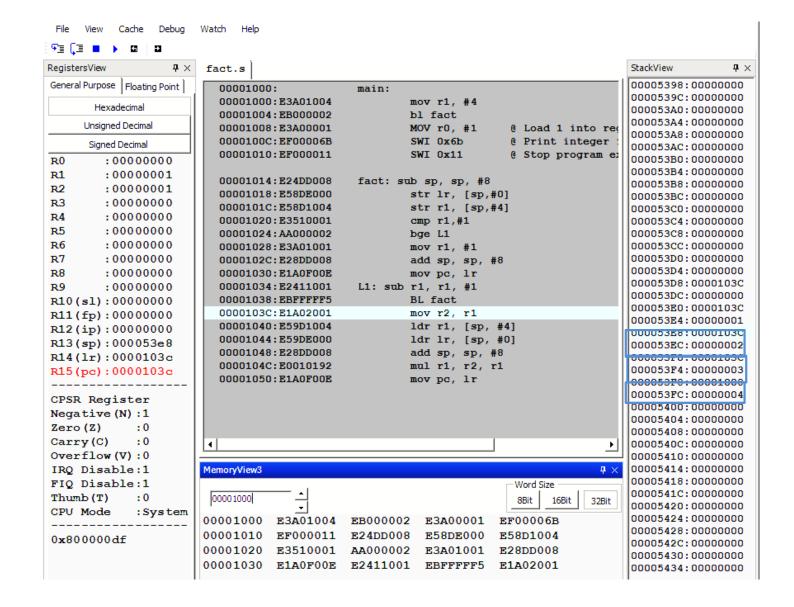


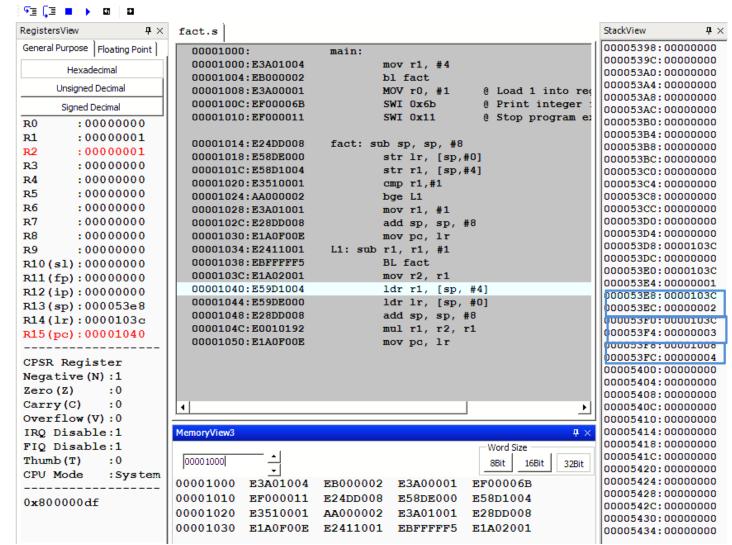


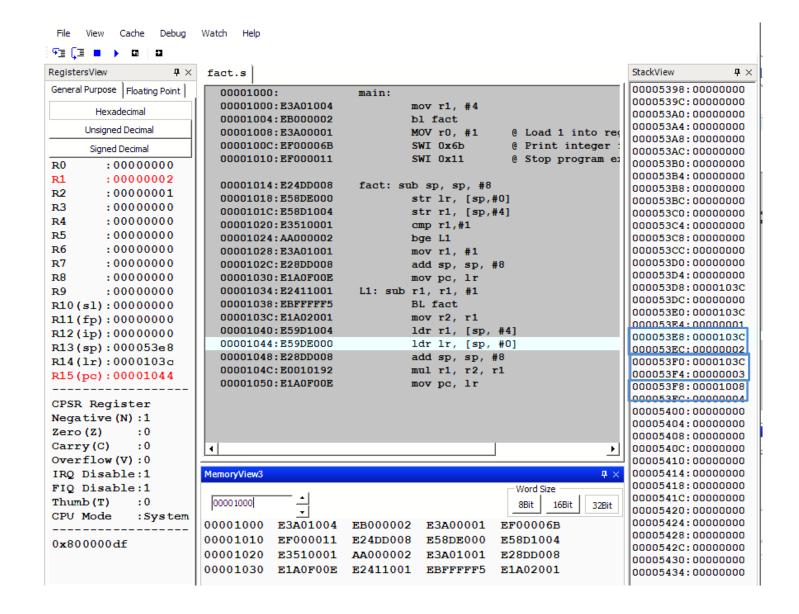


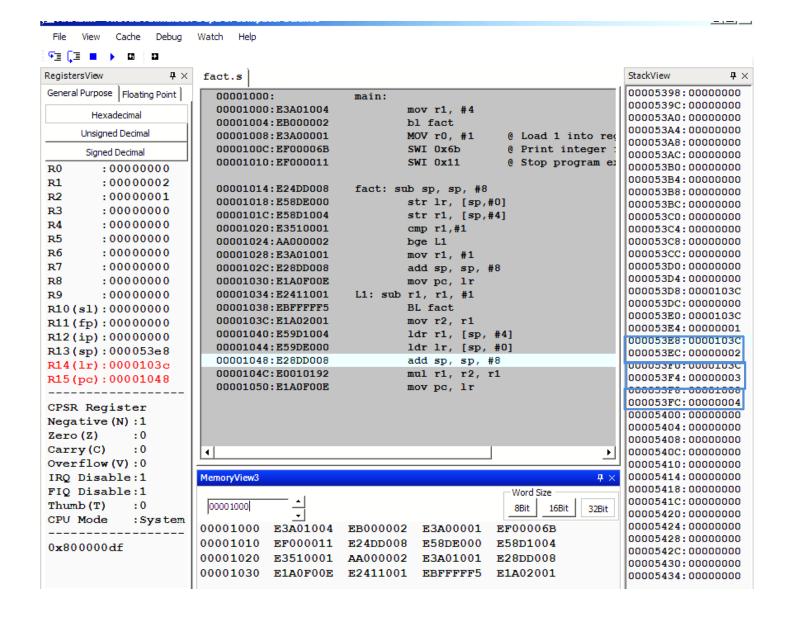


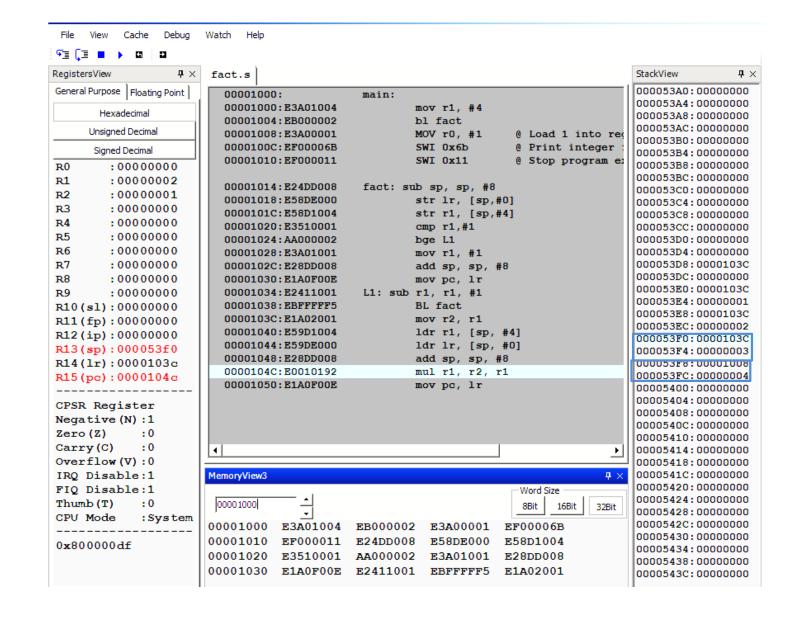


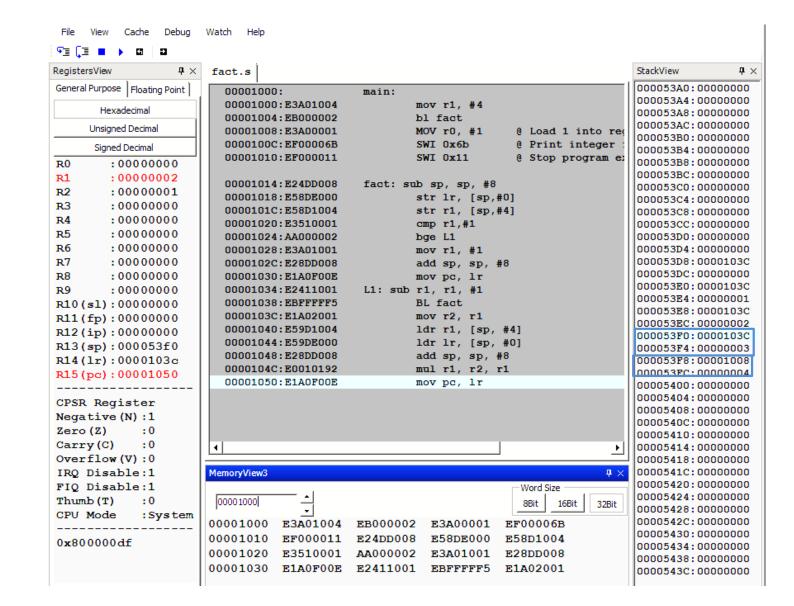


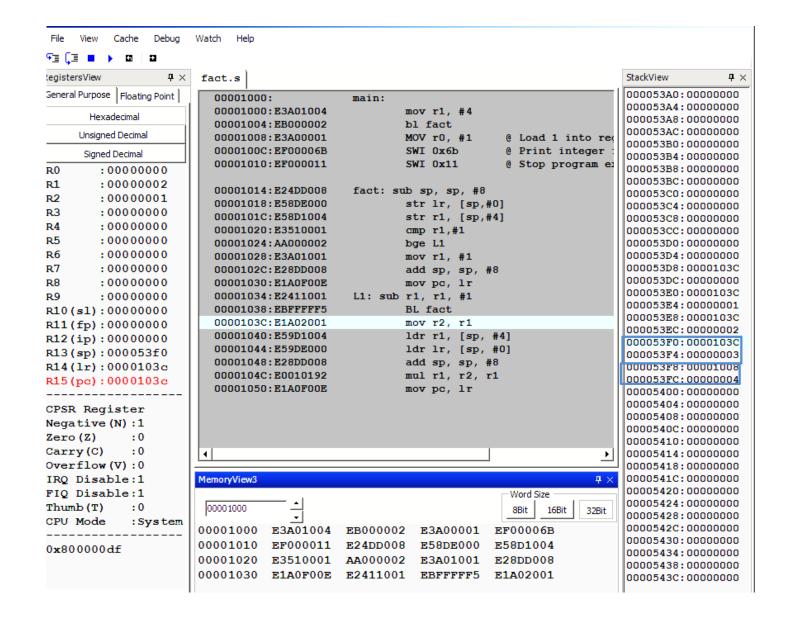












Cache Debug Watch Help FI [I 🔳 RegistersView $\mathbf{p} \times \mathbf{l}$ StackView **4** × fact.s General Purpose | Floating Point | 000053A8:00000000 00001000: main: 000053AC: 00000000 00001000:E3A01004 mov r1, #4 Hexadecimal 000053B0:00000000 00001004:EB000002 bl fact 000053B4:00000000 Unsigned Decimal MOV r0, #1 @ Load 1 into red 00001008:E3A00001 000053B8:00000000 0000100C: EF00006B SWI 0x6b @ Print integer Signed Decimal 000053BC:00000000 00001010:EF000011 SWI 0x11 @ Stop program ex 000053C0:00000000 R0 :00000000 000053C4:00000000 R1:00000003 00001014:E24DD008 fact: sub sp, sp, #8 000053C8:00000000 R2 :00000002 00001018:E58DE000 str lr, [sp,#0] 000053CC: 00000000 R3 :00000000 0000101C: E58D1004 str r1, [sp,#4] 000053D0:00000000 R4 :00000000 cmp r1,#1 00001020:E3510001 000053D4:00000000 R5 :00000000 00001024:AA000002 bge L1 000053D8:0000103C **R6** :00000000 000053DC:00000000 00001028:E3A01001 mov r1, #1 **R7** :00000000 000053E0:0000103C 0000102C: E28DD008 add sp, sp, #8 000053E4:00000001 R8 :00000000 00001030:E1A0F00E mov pc, lr 000053E8:0000103C L1: sub r1, r1, #1 R9 00001034:E2411001 :00000000 000053EC:00000002 00001038: EBFFFFF5 BL fact R10(sl):00000000 000053F0:0000103C 0000103C:E1A02001 mov r2, r1 R11(fp):00000000 000053F4:00000003 00001040:E59D1004 ldr r1, [sp, #4] R12(ip):00000000 000053F8:00001008 00001044:E59DE000 ldr lr, [sp, #0] R13(sp):000053f8 000053FC:00000004 add sp, sp, #8 00001048:E28DD008 R14(lr):0000103c 00005400:00000000 mul r1, r2, r1 0000104C:E0010192 00005404:00000000 R15 (pc): 0000104c 00001050: E1A0F00E mov pc, lr 00005408:00000000 00005400:00000000 CPSR Register 00005410:00000000 Negative (N):1 00005414:00000000 Zero(Z) 00005418:00000000 Carry (C) : 0 4 0000541C:00000000 Overflow (V):0 00005420:00000000 IRQ Disable:1 00005424:00000000 MemoryView3 **4** × 00005428:00000000 FIQ Disable:1 Word Size 0000542C:00000000 Thumb (T) 00001000 8Bit 16Bit 32Bit 00005430:00000000 CPU Mode :System 00005434:00000000 00001000 E3A01004 E3A00001 EB000002 EF00006B 00005438:00000000 00001010 EF000011 E24DD008 E58DE000 E58D1004 0x800000df 0000543C:00000000 00001020 E3510001 AA000002 E3A01001 E28DD008 00005440:00000000

E2411001

EBFFFFF5

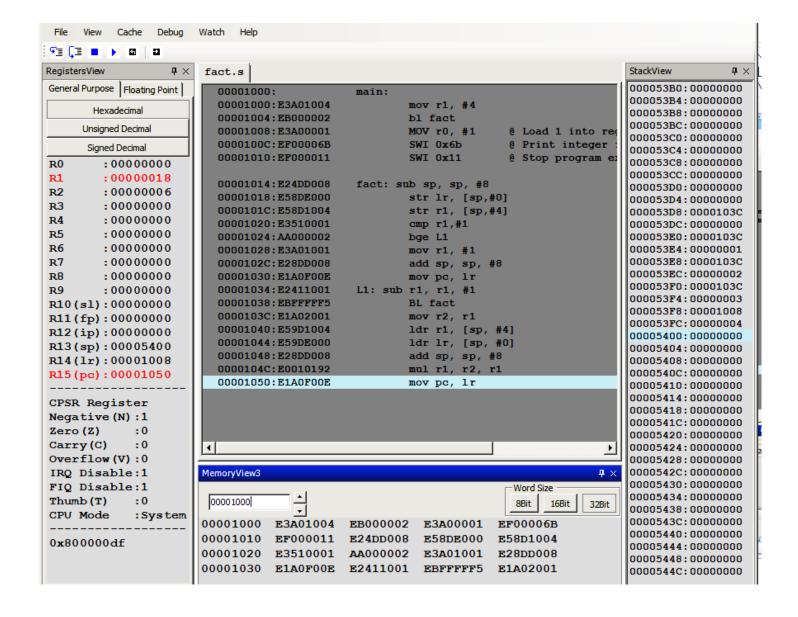
E1A02001

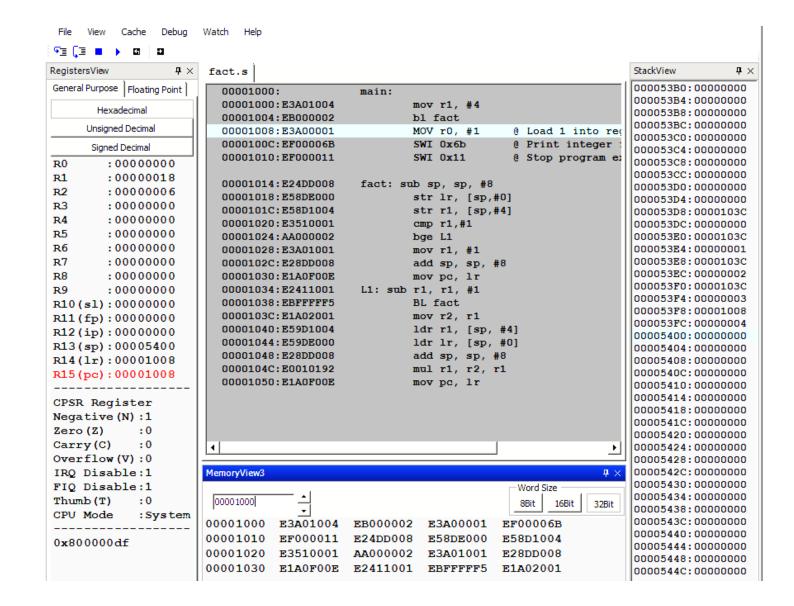
00005444:00000000

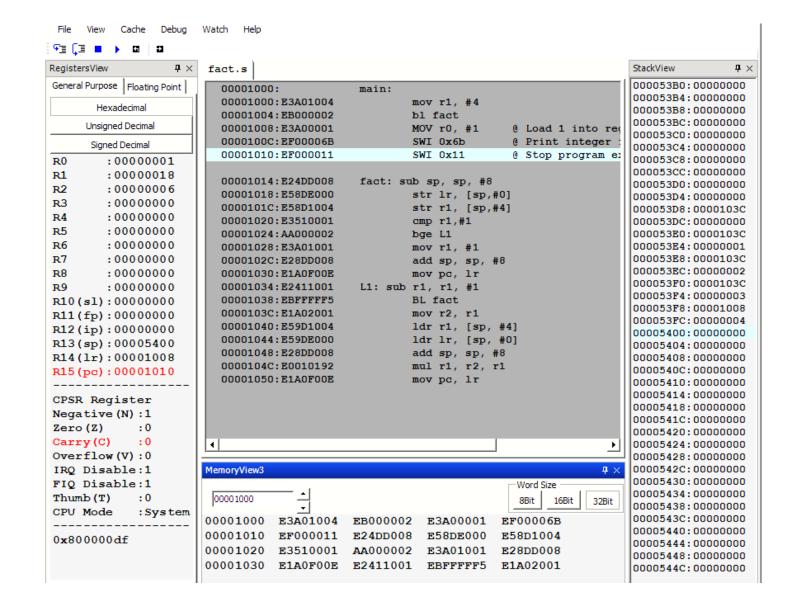
00001030

E1A0F00E

Cache Debua Watch Help Fi [i 🔳 🕨 RegistersView $\mathbf{p} \times$ StackView ŢΧ fact.s General Purpose | Floating Point | 000053A8:00000000 00001000: main: 000053AC: 00000000 00001000: E3A01004 mov r1, #4 Hexadecimal 000053B0:00000000 00001004:EB000002 bl fact 000053B4:00000000 Unsigned Decimal MOV r0, #1 00001008:E3A00001 @ Load 1 into red 000053B8:00000000 SWI 0x6b @ Print integer 0000100C:EF00006B Signed Decimal 000053BC:00000000 SWI 0x11 @ Stop program e: 00001010:EF000011 R0 :00000000 000053C0:00000000 000053C4:00000000 R1 :00000006 00001014:E24DD008 fact: sub sp, sp, #8 000053C8:00000000 R2 :00000002 00001018:E58DE000 str lr, [sp,#0] 000053CC:00000000 R3 :00000000 str r1, [sp,#4] 0000101C:E58D1004 000053D0:00000000 R4 :00000000 00001020:E3510001 cmp r1,#1 000053D4:00000000 R5 :00000000 00001024:AA000002 bge L1 000053D8:0000103C R6 :00000000 000053DC:00000000 00001028:E3A01001 mov r1, #1 R7 000053E0:0000103C :00000000 0000102C: E28DD008 add sp, sp, #8 000053E4:00000001 R8 :00000000 00001030:E1A0F00E mov pc, lr 000053E8:0000103C R9 00001034:E2411001 L1: sub r1, r1, #1 :00000000 000053EC:00000002 R10(sl):00000000 00001038: EBFFFFF5 BL fact 000053F0:0000103C mov r2, r1 0000103C:E1A02001 R11(fp):00000000 000053F4:00000003 ldr r1, [sp, #4] 00001040:E59D1004 R12(ip):00000000 000053F0:00001000 00001044:E59DE000 ldr lr, [sp, #0] R13(sp):000053f8 000053FC:00000004 00001048:E28DD008 add sp, sp, #8 R14(lr):0000103c 00005400:00000000 0000104C:E0010192 mul r1, r2, r1 R15 (pc): 00001050 00005404:00000000 00001050: E1A0F00E mov pc, lr 00005408:00000000 0000540C:00000000 CPSR Register 00005410:00000000 Negative (N):1 00005414:00000000 Zero(Z) : 0 00005418:00000000 Carry (C) : 0 I∢I 0000541C:00000000 Overflow (V):0 00005420:00000000 00005424:00000000 IRQ Disable:1 MemoryView3 **4** × 00005428:00000000 FIQ Disable:1 Word Size 0000542C:00000000 Thumb (T) : 0 00001000 8Bit 16Bit 32Bit 00005430:00000000 CPU Mode :System 00005434:00000000 00001000 E3A01004 EB000002 E3A00001 EF00006B 00005438:00000000 00001010 EF000011 E24DD008 E58DE000 E58D1004 0x800000df 0000543C:00000000 00001020 E3510001 AA000002 E3A01001 E28DD008 00005440:00000000 00001030 E1A0F00E E2411001 EBFFFFF5 E1A02001 00005444:00000000







Note: The following is for the 32-bit ARM7, see Chapter 02_COD 4e ARM

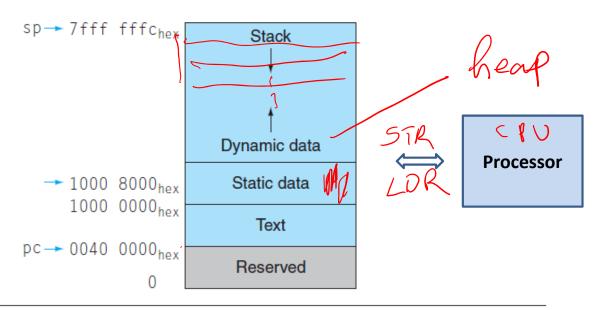
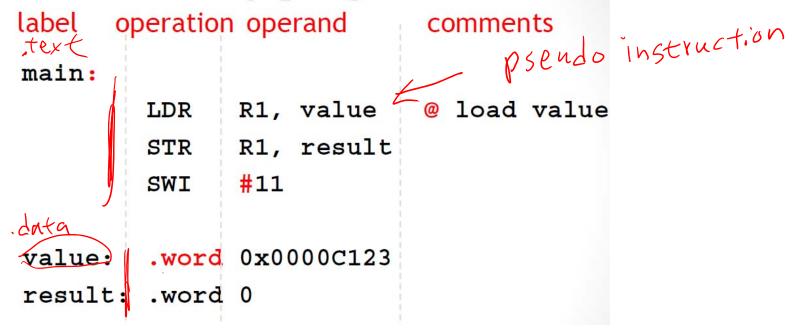
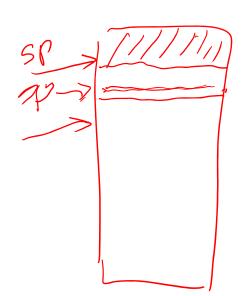


FIGURE 2.13 Typical ARM memory allocation for program and data. These addresses are only a software convention, and not part of the ARM architecture. The stack pointer is initialized to $7fffffc_{hex}$ and grows down toward the data segment. At the other end, the program code ("text") starts at $0040\ 0000_{hex}$. The static data starts at $1000\ 0000_{hex}$. Dynamic data, allocated by malloc in C and by new in Java, is next. It grows up toward the stack in an area called the heap.

ARM assembly program



```
gcd(a, b) {
  if(a==b)          return a;
  else if (a>b)          return gcd(a-b, b);
  else                return gcd(a, b-a);
}
```



```
gcd:
```

sub sp, sp, 4 str kr, [sp, #0] cmp r0, r1 beq Return

If: ble Else sub r0, r0, r1

Else: sub r1, r1, r0

@ recursive call

Re Re

Return: Idr Ir, [sp, #0]
add sp, sp, #4
move r7, r0
mov pc, Ir

@ pop

ged

@ push

@push

```
while (a!=b) {

    if (a > b) {
        a = a - b;
    } else {
        b = b - a;
    }
} return a;
```

"Normal" Assembler

```
gcd cmp r0, r1 ;reached the end?
  beq stop
  blt less ;if r0 > r1
  sub r0, r0, r1 ;subtract r1 from r0
  bal gcd
less sub r1, r1, r0 ;subtract r0 from r1
  bal gcd
stop
```

ARM Conditional Assembler

```
gcd cmp r0, r1 ;if r0 > r1
subgt r0, r0, r1 ;subtract r1 from r0
sublt r1, r1, r0 ;else subtract r0 from r1
bne gcd ;reached the end?
```

Handling large immediate values, label addresses, words, and bytes, ...

.text

@mov r0, #345 @ see this number cannot be used as immediate value

ldr r0, =0x12345678 @ the way to load a large number to register

@ see where the number is and pc-relative addressing

ldr r1, =myByte @ the way to load address of a label to register

ldr r2, [r1] @ see the order of these 4 bytes in memory and in register

str r0, [r1] @ see the 4 bytes in a word are stored in memory (little endian)

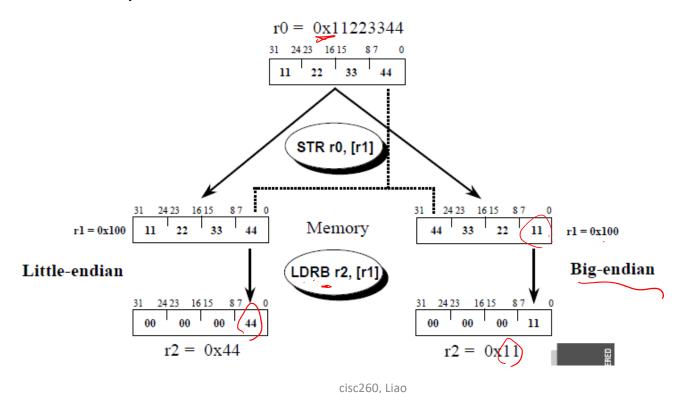
ldrb r4, [r1] @ see which byte in 0x12345678 is loaded back

.data

myByte: .byte 1, 2, 3, 4

Little-endian byte ordering – least-significant byte of word or half-word stored at lower address in memory

Big-endian byte ordering – most-significant byte of word or half-word stored at lower address in memory



 The ARM architecture supports both little-endian and big-endian access to memory. The ARM-Sim# supports only the little-endian format (the same as the Intel architecture which hosts the ARMSim#).

 This is only relevant when data is stored by words and is then accessed by bytes.

Etymology [edit]

In 1726, Jonathan Swift described in his satirical novel *Gulliver's Travels* tensions in Lilliput and Blefuscu: whereas royal edict in Lilliput requires cracking open one's soft-boiled egg at the small end, inhabitants of the rival kingdom of Blefuscu crack theirs at the big end (giving them the moniker *Big-endians*). [2][3] The terms *little-endian* and *endianness* have a similar intent. [4]



Danny Cohen's "On Holy Wars and a Plea for Peace" published in 1980^[3] ends with: "Swift's point is that the difference between breaking the egg at the little-end and breaking it at the big-end is trivial. Therefore, he suggests, that everyone does it in his own preferred way. We agree that the difference between sending eggs with the little- or the big-end first is trivial, but we insist that everyone must do it in the same way, to avoid anarchy. Since the difference is trivial we may choose either way, but a decision must be made."

This trivial difference was the reason for a hundred-years war between the fictional kingdoms. It is widely assumed that Swift was either alluding to the historic War of the Roses or – more likely – parodying through oversimplification the religious discord in England, Ireland and Scotland brought about by the conflicts between the Roman Catholics (Big Endians) on the one side and the Anglicans and Presbyterians (Little Endians) on the other.

Credit: http://en.wikipedia.org/wiki/Endianness

cisc260, Liao