# **CESC16 Detailed Instructions**

This document contains an in-depth description of all the machine instructions supported by the CESC16 computer, as well as the macros provided by the assembler.

ASSEMBLER MNEMONIC	Machine code (binary)	Pseudocode (C-style)
--------------------	--------------------------	-------------------------

### **C-style Pseudocode notation:**

- ALU(A, B) ALU operation between A (first operand) and B (second operand)

- P Flags get updated according to ALU result

- FLAGS Direct access to the flags (status) register

- RAM[Addr] Access RAM at a given address (Addr)

- ROM[Addr][W] Access a word W (1=Upper, 0=Lower) of ROM at a given address (Addr)

- memSpace Controls from which memory space (ROM or RAM) instructions will be fetched

push(A)
 Push a given register A to the stack. It's the same as RAM[--sp]=A

A=pop()
 Pop a given register A from the stack. It's the same as A=RAM[sp++]

#### **Macros notation:**

- [mode] Address in any of the 3 addressing modes ([Addr], [rB] or [rA+Imm16])

- OPERAND Either a register (rB), immediate (Imm), or memory address ([mode])

- The mnemonic on the left side of the arrow gets replaced by the instruction(s) on the right side of the arrow: MACRO → Translated Instructions

# ALU Operations:

### **Register operand:**

ALU rD, rA, rB	00000FFFDDDDAAAA XXXXXXXXXXXBBBB	rD = ALU(rA, rB) □
----------------	-------------------------------------	-----------------------

### Immediate operand:

ALU rD, rA, Imm16	00001FFFDDDDAAAA	rD = ALU(rA, Imm16) □
-------------------	------------------	--------------------------

# **Direct addressing:**

ALU rD, rA, [Addr16]	01000FFFDDDDAAAA @@@@@@@@@@@@@@@@@	rD = ALU(rA, RAM[Addr16]) □
----------------------	---------------------------------------	--------------------------------

# **Indirect addressing:**

ALU rD, rA, [rB]	01001FFFDDDDAAAA XXXXXXXXXXXBBBB	rD = ALU(rA, RAM[rB]) □
------------------	-------------------------------------	----------------------------

### Indexed addressing:

# Operations on each clock cycle (Register and Immediate):

Fetch instruction + 1st operand	Fetch argument (2nd operand)	Perform ALU operation and store result in register file
---------------------------------	------------------------------	---

### Operations on each clock cycle (Direct and Indirect):

Fetch instruction +	Fetch argument and	Fetch 2nd operand from	Perform ALU op. and
1st operand	compute address	memory	store result in regfile

### Operations on each clock cycle (Indexed):

Fetch instruction	Fetch argument,	Fetch 1st operand	Fetch 2nd operand	Perform ALU op,
reten instruction	compute address	(rD) from regfile	from memory	store result in reg.

### **Description:**

Performs an ALU operation as indicated by the 3 Funct bits, using the <u>contents of rA</u> as first operand (except in indexed mode) and either the <u>contents of rB</u>, a <u>16 bit immediate</u> or the contents of a <u>memory address</u> as second operand. The result of the operation is stored in rD and the flags are updated accordingly. See table in main documentation for ALU operations, mnemonics and descriptions.

### Remarks about ALU operations:

- Carry and oVerflow flags are undefined after all operations except add, sub, addc and subb.
- The mov instruction doesn't require the first operand (rA), doesn't update the flags (see movf macro for this purpose) and takes only 2 clock cycles (Register and Immediate) or 3 clock cycles (Direct, Indirect and Indexed).
- When using indexed addressing mode, rA is used for computing the address. Therefore, rD is used both as the <u>first argument</u> and the destination register.

### **Examples:**

mov t0, t1 The value stored in t1 gets copied into t0. The value at t1 and the flags are unchanged.

mov t0, 0x1234 The value stored in t0 becomes 0x1234. Flags are preserved.

and t0, t1, t2 Perform a logic AND between the contents of t1 and t2. Store result into t0. The values at t1 and t2 remain unchanged.

sub t0, t1, [123] Fetch the data stored at address 123 (0x7B) and subtract it from the data stored at t1. Store the result in t0 (operands remain unchanged).

Fetch the data pointed by register t2 and add them to the contents of t1. addc t0, t1, [t2] Add 1 to result if Carry bit is set. Store result in t0 (operands are unchanged).

mov a0, [t0+10] Add 10 to the contents of t0 to get a memory address, then load the data stored at that address into a0.

Add 10 to the contents of t0 to get a memory address, then add the data add a0, [t0+10] stored at that address to the contents of a0. Store the result in a0.

### **Macros:**

Negate register (bitwise NOT): not rD, rA xor rD, rA, 0xFFFF

Bitwise NAND, NOR and XNOR:

nand/nor/xnor rD, rA, OPERAND and/or/xor rD, rA, OPERAND not rD, rD

Shift Left with Carry (1 bit): addc rD, rA, rA sllc rD, rA

movf rD, OPERAND add rD, zero, OPERAND Move and update Flags\*:

Compare register to operand\*: cmp rA, OPERAND sub zero, rA, OPERAND

Test masked register\*: mask rA, OPERAND and zero, rA, OPERAND

Test register (or memory): test OPERAND movf zero, OPERAND

Clear flags: movf zero, 0x0001 clrf

Swap registers (no-temp): swap rA, rB xor rA, rA, rB

xor rB, rA, rB

xor rA, rA, rB

Warning: This method saves a temp register, but it's 3 cycles slower and modifies the flags

No operation: mov zero, zero

- There are many alternative expansions for nop. This one is encoded as all zeros (0x0000).

\* It's not possible to implement an indexed version of movf, cmp or mask, since rD also acts as the first operand. However, an indexed version of test can be implemented.

# ALU Operations (destination in memory):

#### **Direct addressing**

ALU [Addr16], rA	01100FFFXXXXAAAA @@@@@@@@@@@@@@@@@	RAM[Addr16] = ALU(RAM[Addr16], rA) □
------------------	---------------------------------------	---

### Indirect addressing:

ALU [rA], rB	01101FFFXXXXAAAA	RAM[rA] =
ALO [IA], ID	XXXXXXXXXXXBBBB	ALU(RAM[rA], rB) □

### Indexed addressing:

ALU [rA+ <mark>Imm16</mark> ], rB	0111XFFFBBBBAAAA	RAM[rA+Imm16] =
ALO [TATIMITO], TO	IIIIIIIIIIIIII	ALU(RAM[rA+Imm16], rB) □

### Operations on each clock cycle (Direct and Indirect):

Fetch Instruction	argument and npute address	Fetch 1st operand from memory	Perform ALU op. and store result in memory
-------------------	----------------------------	-------------------------------	--

# Operations on each clock cycle (Indexed):

### **Description:**

Performs an ALU operation as indicated by the 3 Funct bits, using the contents of a memory address (direct, indirect or indexed addressing) as first operand and a register as second operand.

The result of the operation is stored in the <u>same address as the first operand</u> and the flags are updated accordingly. See table in main documentation for ALU operations, mnemonics and descriptions.

### Remarks about memory ALU operations:

- Carry and oVerflow flags are undefined after all operations except add, sub, addc and subb.
- The decoded memory address is used for both <u>first operand</u> and <u>destination</u>. The second operand must be a register (no Memory-Memory or Memory-Immediate operations). If those restrictions can't be met, considering loading the needed value to a temporary register first.
- The mov instruction <u>doesn't update the flags</u> (see movf macro for this purpose) and takes only 3 clock cycles (4 cycles in indexed mode).

#### **Examples:**

mov [0x1234], zero	The memory contents at address 0x1234 become 0x0000 (the contents of zero get stored in memory). Flags are preserved.
mov [s0+12], a1	Store the contents of a1 into memory. The memory address consists of the contents of s0, plus an offset of 12. Flags are preserved.
xor [6], s1	Perform a logic XOR between the data at memory address 6 and the contents stored in s1. Store the result into address 6 (s1 remains unchanged).
add [sp], a1	Increment the top of the stack by the amount stored in a1. The contents of a1 and sp remain unchanged.
subb [t2+1], t1	Fetch the data pointed by register t2 (plus an offset of 1) and subtract the contents of t1 from it. Subtract 1 to result if Borrow bit is set. Store the result in memory (at the address pointed by t2 plus 1).

### Shifts:

## **Shift Left Logical:**

sll rD, rA, Imm4
------------------

### **Shift Right Logical:**

srl rD, rA, Imm4	0010IIIIDDDDAAAA XXXXXXXXXXXXXXXXX	rD = rA>>Imm4 戸 (unsigned)
------------------	---------------------------------------	-------------------------------

### **Shift Right Arithmetic:**

sra rD, rA, Imm4	0011IIIIDDDDAAAA XXXXXXXXXXXXXXXX	rD = rA>>Imm4 ┌─ (signed)
------------------	--------------------------------------	------------------------------

### Operations on each clock cycle:

Fetch instruction + operand (rA)	Shift 1 position	Shift 1 position		Shift 1 position and store result
----------------------------------	------------------	------------------	--	-----------------------------------

### **Description:**

The contents of rA get shifted (left or right) as many bits as indicated by Imm4.

- s11: bits get shifted to the <u>left</u> and <u>filled with zeros</u>.
- srl: bits get shifted to the right and filled with zeros.
- sra: bits get shifted to the right and the sign is extended.

Flags are updated (Carry and oVerflow flags are undefined) and the result is stored in rD.

### Remarks about shifts:

- Memory contents can't be shifted directly and must be copied to/from a temporary register.
- Bit shifts are the only instructions with variable clock durations. Each shifted bit takes 1 clock cycle, plus 1 extra clock for fetching.
- The ISA allows shifting 0 bits but, since it has no practical use, it can be considered an illegal instruction. The computer will interpret a shift of 0 bits as a NOP.

# Move byte (load from memory):

movb rD, [rA+Imm16]	1000000DDDAAAA	rD = (byte) RAM[rA+Imm16]
---------------------	----------------	---------------------------

#### Operations on each clock cycle:

Fetch instruction	Fetch offset and compute address	Read data memory, extend sign and store in register file
-------------------	----------------------------------	--

### **Description:**

The contents of rA are added to an immediate offset to get a memory address. Then, the lower 8 bits of the data stored at that address are fetched, and the upper 8 bits are sign-extended.

The result is stored in rD.

#### Remarks:

- An unsigned version can be implemented with a mask (set upper 8 bits to 0), see 1bu below.
- The movb instruction only supports the indexed addressing mode, but converting it to direct or indirect addressing is very easy (see macros below).
- Memory access is <u>not</u> Byte-Oriented and Data Memory is 16 bits wide:
  - There is no way to access only the upper 8 bits of the word at a memory address (other than using mov to load all 16 bits into a register and shifting 8 positions with sr1).
  - It doesn't matter if we want to store a byte (8 bit) or a word (16 bit), both will take the same amount of space in memory (1 word).
- Therefore, an instruction to <u>store</u> bytes to memory isn't needed: the only thing that matters is how we interpret the data when we <u>load</u> it (choosing between word, signed byte and unsigned byte). However, in order to increase code clarity, the *Move Byte to memory* alias is available below.

# Macros:

Direct addressing:	movb rD,	[Addr16]	$\rightarrow$	movb rD,	[zero+Addr16]
Indirect addressing:	movb rD,	[rA]	$\rightarrow$	movb rD,	[rA+0]

Move Byte to memory (alias): movb [mode], rB → mov [mode], rB

#### Alternative MIPS/RISC-V syntax:

Alternative will office v syritax	•		
Load Word:	lw rD, Imm16(rA)	$\rightarrow$	mov rD, [rA+Imm16]
Store Word:	sw rB, Imm16(rA)	$\rightarrow$	mov [rA+Imm16], rB
Load Byte:	lb rD, Imm16(rA)	$\rightarrow$	movb rD, [rA+Addr16]
Store Byte:	sb rB, Imm16(rA)	$\rightarrow$	movb [rA+Addr16], rB
Load Byte Unsigned:	lbu rD, Imm16(rA)	$\rightarrow$	movb rD, [rA+Addr16] and rD, rD, 0x00FF

# Swap register with memory:

swap rD, [rA+Imm16]	1000001DDDDAAAA	<pre>temp = rD; rD = RAM[rA+Imm16]; RAM[rA+Imm16] = temp</pre>
---------------------	-----------------	--

#### Operations on each clock cycle:

Fetch instruction	Fetch offset and	Fetch rB	Read data from	Write data to	
reten instruction	compute address	(same as rD)	memory	memory	

### **Description:**

Swaps the contents of rD and the contents stored at an address (with offset) of <u>data memory (RAM)</u>. An indexed mov is performed simultaneously <u>to and from</u> rD.

#### Macros:

Direct addressing: swap rD, [Addr16] → swap rD, [zero+Addr16]

Indirect addressing: swap rD,  $[rA] \rightarrow swap rD$ , [rA+0]

## Peek program memory:

peek rD, [rA+Imm16], W	1000001WDDDDAAAA	rD = ROM[rA+Imm16][W]
------------------------	------------------	-----------------------

### Operations on each clock cycle:

Fetch instruction	Fetch offset and compute address	Read program memory
-------------------	----------------------------------	---------------------

#### **Description:**

Loads into rD the contents <u>from program memory (ROM)</u> at the address contained by rA (plus an offset). Since the program memory is 32 bits wide,  $\overline{W}$  indicates which 16-bit word will be fetched:

- W=1: Most significant bits get fetched (instruction opcode).
- W=0: Least significant bits get fetched (instruction argument).

The assembler uses big endian encoding. Therefore, when peek is used to load 16-bit constants, the most significant bits (W=1) correspond to the <u>first word (lower address)</u> and the least significant bits (W=0) correspond to the <u>second word (higher address)</u>.

### Macros:

Peek upper bits:	peek rD, [rA+Imm16], Up	o → pee	ek rD, [rA+Imm16], 1	
Peek lower bits:	peek rD, [rA+Imm16], Lo	ow → pee	ek rD, [rA+Imm16], 0	
Peek opcode:	peek rD, [rA+Imm16], Op	o → pe∈	ek rD, [rA+Imm16], 1	
Peek argument:	peek rD, [rA+Imm16], Ar	-g → pee	ek rD, [rA+Imm16], 0	

# Stack Push and Pop:

### Push register to Stack:

rusii register to stack.		
push rB	10000100XXXX0001 XXXXXXXXXXBBBB	<pre>RAM[sp] = rB   (push(rB))</pre>
Push immediate to Stack:		
push Imm16	10000101XXXX0001 IIIIIIIIIIIII	push(Imm16)
Push flags to Stack:		
pushf	10000110XXXX0001 XXXXXXXXXXXXXX	push(FLAGS)
Pop register from Stack:		
pop rD	10000111DDDD00001 XXXXXXXXXXXXXXXX	rD = RAM[sp++] (rD = pop())
Pop flags from Stack:		
popf	10001000XXXX0001 XXXXXXXXXXXXXXX	FLAGS = pop()
Operations on each clock cyc	ele:	
Fetch instruction	Fetch and update Stack Pointer	Read/Write data memory

## **Description:**

- push pushes the contents of a register (or an immediate value) into the stack: sp is decremented by 1 and the data is stored at the new address pointed by sp.

Only in push: Fetch argument

- pop pops the top of the stack into rD: loads the contents pointed by sp into rD and then sp is incremented by 1.
- pushf and popf work the same way, but they store and load the flags (status register). This isn't usually needed for regular subroutines, but an interrupt handler <u>must</u> use them to preserve the status of the main program that got interrupted.

<u>Warning</u>: mov instructions can also be used to access the stack without the limitations of push and pop (by using sp as address), but you shouldn't use both methods at once (unless you really know what you are doing) to avoid making errors.

For example, after a push, the variable stored at sp+2 will be stored at sp+3. This small detail can cause many bugs that will be hard to find.

Remarks about interrupt handlers: An interrupt handler *must* push the flags and <u>all</u> registers it's going to use (not just the safe registers). However, all of this is <u>already done by the OS</u> before handing over control to the user's interrupt handler, which <u>can treat the registers and flags as if it was a regular subroutine</u> (that is, it only needs to push and pop *safe* registers).

## Conditional Jumps:

### Jump to register:

JMP rA	1100FFFFXXXXAAAA XXXXXXXXXXXXXXXXX	if(condition) PC = rA
--------	---------------------------------------	-----------------------

### Jump to immediate address:

JMP Addr16	1101FFFFXXXXXXXX @@@@@@@@@@@@@@@@@	if(condition) F	PC = Addr16
------------	---------------------------------------	-----------------	-------------

### Operations on each clock cycle:

Fetch instruction	Check flags.  If condition is true, load new address into PC
-------------------	--

## **Description:**

Checks the condition indicated by the 3 Funct bits, then jumps to an immediate address (or the address stored in rA) only if the condition is true.

Therefore, the next executed instruction is pointed by:

- Addr16 or rA, if the jump condition is met. The jmp instruction is always performed.
- PC+1, if the jump condition is not met (or PC+2 if running from RAM).

The jump condition is checked using the flags, which depend on the <u>last ALU operation</u>.

Conditional jumps (and macros) can be separated in 2 groups:

- Check result of last operation: jz, jnz, jc, jnc, jo, jno, js, jns
- Compare 2 integers (must be executed right after a cmp instruction):

je, ja, jae, jb, jbe, jl, jle, jg, jge

And their negations:

jne, jna, jnae, jnb, jnbe, jnl, jnle, jng, jnge

Most of these mnemonics share opcodes since they perform the same action. See *Jump Conditions* table in <u>DOCS/CESC16.pdf</u> for all the alternative names for each real instruction.

### Macros:

Skip N instructions: JMP skip(N)  $\rightarrow$  JMP pc + N + 1

Skip N instructions (in RAM): JMP skip(N)  $\rightarrow$  JMP pc + 2\*(N + 1)

### Call subroutine:

### Call subroutine in same memory space (address in register):

### Call subroutine in same memory space (immediate address):

call Addr16	11100001XXXX0001 @@@@@@@@@@@@@@@@@	<pre>push(PC+N); PC = Addr16</pre>
-------------	---------------------------------------	------------------------------------

### Call subroutine in ROM (address in register):

syscall rB	11100010XXXX0001 XXXXXXXXXXXBBBB	<pre>push(PC+N); PC = rA; memSpace = ROM</pre>
------------	-------------------------------------	--

# Call subroutine in ROM (immediate address):

syscall Addr16	11100011XXXX0001 @@@@@@@@@@@@@@@@@	<pre>push(PC+N); PC = Addr16; memSpace = ROM</pre>
----------------	---------------------------------------	--

### Call subroutine in RAM (address in register):

enter rB
----------

### Call subroutine in RAM (immediate address):

enter Addr16	11100101XXXX0001 @@@@@@@@@@@@@@@@@@	<pre>push(PC+N); PC = Addr16; memSpace = RAM</pre>
--------------	--	--

### Operations on each clock cycle:

### **Description:**

Those instructions push the address of the <u>next</u> instruction to the stack (PC+1 if running from ROM, PC+2 if running from RAM) before jumping unconditionally to an address. Arbitrary depths of subroutine calls are allowed (as well as recursion).

<u>Instructions can be fetched from ROM or RAM</u>. This family of instructions allows jumping between them:

- call: stays in the current memory space (used for regular subroutines)
- syscall: calls a subroutine stored in ROM (used for system calls)
- enter: calls a subroutine stored in RAM (used for entering user programs)

Memory space BEFORE	Instruction	Memory space AFTER	Gets pushed to stack
	call	ROM	
ROM	syscall	ROM	PC+1
	enter	RAM	
RAM	call	RAM	
	syscall	ROM	PC+2
	enter	RAM	

### Return from subroutine:

#### Return from call:

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
--

### **Return from syscall:**

sysret 111001111XXXX0001 PC = pop() XXXXXXXXXXXXXXXX memSpace = RAM
---

#### Return from enter:

exit	11101000XXXX0001 XXXXXXXXXXXXXXXX	PC = pop() memSpace = ROM
exit		, ,

### Operations on each clock cycle:

Fetch instruction	Fetch and update Stack Pointer	Pop new address from stack to PC, update memory space
-------------------	--------------------------------	---

#### **Description:**

Those instructions pop the top of the stack and jump unconditionally to that address: control is returned to the routine that performed the call instruction.

<u>Warning</u>: Make sure the subroutine has freed all the memory it had allocated in the stack before using the *return* instruction (otherwise sp won't be pointing at the correct return address).

The return instructions are companions of a type of call instruction:

- ret returns from routines that were called using call: stays in the current memory space.
- sysret returns from routines that were called from RAM (using syscall): always jumps to RAM.
- exit exits user programs that were called from ROM (using enter): always jumps to ROM.

For more information, read the "Memory Map" section in <a href="DOCS/CESC16.pdf">DOCS/CESC16.pdf</a>

Memory space BEFORE	Instruction	Memory space AFTER	Intended use: returning from
ROM	ret	ROM	call, syscall*
	sysret	RAM	syscall**
	exit	ROM	[use ret instead]
RAM	ret	RAM	call
	sysret	RAM	[use ret instead]
	exit	ROM	enter

<sup>\*</sup> OS routines can be called from ROM using call, but it's recommended to use syscall instead.

<sup>\*\*</sup> OS routines use ret instead of sysret (even though they are called using syscall), because they don't know if they have been called from ROM or RAM, so they will assume it's been ROM. When calling OS routines from RAM, the CALL\_GATE routine must be used. Read the "Operating System" section in DOCS/CESC16.pdf for more information.