

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)
- `⚑` 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	0000FFFDDDDAAAA XXXXXXXXXXXXBBBB	$rD = \text{ALU}(rA, rB)$ P
------------------------	-------------------------------------	--------------------------------

Immediate operand:

ALU rD , rA , Imm16	00001FFFDDDDAAAA IIIIIIIIIIIIIIII	$rD = \text{ALU}(rA, \text{Imm16})$ P
-------------------------	--------------------------------------	------------------------------------------

Direct addressing:

ALU rD , rA , [Addr16]	01000FFFDDDDAAAA @@@@@@@@@@@@@@@@	$rD = \text{ALU}(rA, \text{RAM}[\text{Addr16}])$ P
----------------------------	--------------------------------------	-------------------------------------------------------

Indirect addressing:

ALU rD , rA , [rB]	01001FFFDDDDAAAA XXXXXXXXXXXXBBBB	$rD = \text{ALU}(rA, \text{RAM}[rB])$ P
----------------------------	--------------------------------------	--------------------------------------------

Indexed addressing:

ALU rD , [$rA + \text{Imm16}$]	0101XFFFDDDDAAAA IIIIIIIIIIIIIIII	$rD = \text{ALU}(rD, \text{RAM}[rA + \text{Imm16}])$ P
------------------------------------	--------------------------------------	--------------------------------------------------------

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

Operations on each clock cycle (Indexed):

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

Description:

Performs an ALU operation as indicated by the 3 Funct bits, using the contents of rA as first operand (except in indexed mode) and either the contents of rB , a 16 bit immediate or the contents of a memory address 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 first argument and the destination register.

Examples:

<code>mov t0, t1</code>	The value stored in <code>t1</code> gets copied into <code>t0</code> . The value at <code>t1</code> and the flags are unchanged.
<code>mov t0, 0x1234</code>	The value stored in <code>t0</code> becomes 0x1234. Flags are preserved.
<code>and t0, t1, t2</code>	Perform a logic AND between the contents of <code>t1</code> and <code>t2</code> . Store result into <code>t0</code> . The values at <code>t1</code> and <code>t2</code> remain unchanged.
<code>sub t0, t1, [123]</code>	Fetch the data stored at address 123 (0x7B) and subtract it from the data stored at <code>t1</code> . Store the result in <code>t0</code> (operands remain unchanged).
<code>addc t0, t1, [t2]</code>	Fetch the data pointed by register <code>t2</code> and add them to the contents of <code>t1</code> . Add 1 to result if Carry bit is set. Store result in <code>t0</code> (operands are unchanged).
<code>mov a0, [t0+10]</code>	Add 10 to the contents of <code>t0</code> to get a memory address, then load the data stored at that address into <code>a0</code> .
<code>add a0, [t0+10]</code>	Add 10 to the contents of <code>t0</code> to get a memory address, then add the data stored at that address to the contents of <code>a0</code> . Store the result in <code>a0</code> .

Macros:

Negate register (bitwise NOT):	<code>not rD, rA</code>	→	<code>xor rD, rA, 0xFFFF</code>
Bitwise NAND, NOR and XNOR:	<code>nand/nor/xnor rD, rA, OPERAND</code>	→	<code>and/or/xor rD, rA, OPERAND</code> <code>not rD, rD</code>
Shift Left with Carry (1 bit):	<code>sllc rD, rA</code>	→	<code>addc rD, rA, rA</code>
Move and update Flags*:	<code>movf rD, OPERAND</code>	→	<code>add rD, zero, OPERAND</code>
Compare register to operand*:	<code>cmp rA, OPERAND</code>	→	<code>sub zero, rA, OPERAND</code>
Test masked register*:	<code>mask rA, OPERAND</code>	→	<code>and zero, rA, OPERAND</code>
Test register (or memory):	<code>test OPERAND</code>	→	<code>movf zero, OPERAND</code>
Clear flags:	<code>clrf</code>	→	<code>movf zero, 0x0001</code>
No operation:	<code>nop</code>	→	<code>mov zero, zero</code>
- There are many alternative expansions for <code>nop</code> . 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) \mathcal{P}
------------------	--------------------------------------	-----------------------------------------------------

Indirect addressing:

ALU [rA], rB	01101FFFXXXXAAAA XXXXXXXXXXXXBBBB	RAM[rA] = ALU(RAM[rA], rB) \mathcal{P}
--------------	--------------------------------------	---------------------------------------------

Indexed addressing:

ALU [rA+Imm16], rB	0111XFFFBBBBAAAA IIIIIIIIIIIIIIII	RAM[rA+Imm16] = ALU(RAM[rA+Imm16], rB) \mathcal{P}
--------------------	--------------------------------------	---------------------------------------------------------

Operations on each clock cycle (Direct and Indirect):

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

Operations on each clock cycle (Indexed):

Fetch instruction	Fetch argument, compute address	Fetch 2nd operand (rB) from regfile	Fetch 1st operand from memory	Perform ALU op, store in memory
-------------------	------------------------------------	----------------------------------------	----------------------------------	------------------------------------

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 same address as the first operand 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 first operand and destination. 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 doesn't update the flags (see `movf` macro for this purpose) and takes only 3 clock cycles (4 cycles in indexed mode).

Examples:

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

Shifts:

Shift Left Logical:

<code>sll rD, rA, Imm4</code>	0001IIIIDDDDAAAA XXXXXXXXXXXXXXXXXXXX	$rD = rA \ll Imm4$ ∇
-------------------------------	------------------------------------------	-----------------------------

Shift Right Logical:

<code>srl rD, rA, Imm4</code>	0010IIIIDDDDAAAA XXXXXXXXXXXXXXXXXXXX	$rD = rA \gg Imm4$ ∇ (unsigned)
-------------------------------	------------------------------------------	-------------------------------------------

Shift Right Arithmetic:

<code>sra rD, rA, Imm4</code>	0011IIIIDDDDAAAA XXXXXXXXXXXXXXXXXXXX	$rD = rA \gg Imm4$ ∇ (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.

- `sll`: bits get shifted to the left and filled with zeros.
- `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]	10000000DDDDAAAA IIIIIIIIIIIIIIII	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 `lbu` 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 not 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 `srl`).
 - 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 store bytes to memory isn't needed: the only thing that matters is how we interpret the data when we load 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` \rightarrow `mov [mode], rB`

Alternative MIPS/RISC-V syntax:

Load Word: `lw rD, Imm16(rA) → mov rD, [rA+Imm16]`

Store Word: `sw rB, Imm16(rA)` \rightarrow `mov [rA+Imm16], rB`

Load Byte: `lb rD, Imm16(rA) → movb rD, [rA+Addr16]`

Store Byte: `sb rB, Imm16(rA)` \rightarrow `movb [rA+Addr16], rB`

Load Byte Unsigned:

<code>lbu rD, Imm16(rA)</code>	→	<code>movb rD, [rA+Addr16]</code>
		<code>and rD, rD, 0x00FF</code>

Swap register with memory:

swap rD, [rA+Imm16]	10000001DDDDAAAA IIIIIIIIIIIIIIIIII	temp = rD; rD = RAM[rA+Imm16]; RAM[rA+Imm16] = temp
---------------------	----------------------------------------	-----------------------------------------------------------

Operations on each clock cycle:

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

Description:

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

Macros:

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

Indirect addressing: swap rD, [rA] → swap rD, [rA+0]

Peek program memory:

peek rD, [rA+Imm16], W	10000001WDDDDAAAA IIIIIIIIIIIIIIIIII	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 from program memory (ROM) at the address contained by rA (plus an offset). Since the program memory is 32 bits wide, 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 first word (lower address) and the least significant bits (W=0) correspond to the second word (higher address).

Macros:

Peek upper bits: peek rD, [rA+Imm16], Up → peek rD, [rA+Imm16], 1

Peek lower bits: peek rD, [rA+Imm16], Low → peek rD, [rA+Imm16], 0

Peek opcode: peek rD, [rA+Imm16], Op → peek rD, [rA+Imm16], 1

Peek argument: peek rD, [rA+Imm16], Arg → peek rD, [rA+Imm16], 0

Stack Push and Pop:

Push register to Stack:

push rB	10000100XXXX0001 XXXXXXXXXXXXBBBB	RAM[--sp] = rB (push(rB))
---------	--------------------------------------	------------------------------

Push immediate to Stack:

push Imm16	10000101XXXX0001 IIIIIIIIIIIIIIII	push(Imm16)
------------	--------------------------------------	-------------

Push flags to Stack:

pushf	10000110XXXX0001 XXXXXXXXXXXXXXXX	push(FLAGS)
-------	--------------------------------------	-------------

Pop register from Stack:

pop rD	10000111DDDD0001 XXXXXXXXXXXXXXXX	rD = RAM[sp++] (rD = pop())
--------	--------------------------------------	--------------------------------

Pop flags from Stack:

popf	10001000XXXX0001 XXXXXXXXXXXXXXXX	FLAGS = pop()
------	--------------------------------------	---------------

Operations on each clock cycle:

Fetch instruction	Fetch and update Stack Pointer Only in push: Fetch argument	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.
- 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 must use them to preserve the status of the main program that got interrupted.

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

Conditional Jumps:

Jump to register:

JMP <i>rA</i>	11000FFFXXXXAAAA XXXXXXXXXXXXXXXXXXXX	if(condition) PC = <i>rA</i>
---------------	------------------------------------------	------------------------------

Jump to immediate address:

JMP <i>Addr16</i>	11001FFFXXXXXXXXXX @@@@@@@@@@@@@@@@	if(condition) PC = <i>Addr16</i>
-------------------	----------------------------------------	----------------------------------

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 last ALU operation.

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

- Check result of last operation: *jz*, *jnz*, *jc*, *jnc*, *jb*, *jnb*
- Compare 2 integers (*must* be executed right after a *cmp* instruction):
jeq, *jne*, *jlt*, *jle*, *jltu*, *jleu*

Macros:

Jump if Borrow:	<i>jb addr</i>	→	<i>jnc addr</i>
Jump if Not Borrow:	<i>jnb addr</i>	→	<i>jc addr</i>
Jump if Equal:	<i>jeq addr</i>	→	<i>jz addr</i>
Jump if Not Equal:	<i>jne addr</i>	→	<i>jnz addr</i>
Jump if Less Than (Unsigned):	<i>jltu addr</i>	→	<i>jnc addr</i>
Skip N instructions:	JMP <i>skip(N)</i>	→	JMP <i>pc</i> + N + 1
Skip N instructions (in RAM):	JMP <i>skip(N)</i>	→	JMP <i>pc</i> + 2*(N + 1)

Call subroutine:

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

call rA	11010000XXXXAAAA XXXXXXXXXXXXXXXXXX	push(PC+N); PC = rA
---------	----------------------------------------	---------------------

Call subroutine in same memory space (immediate address):

call Addr16	11010001XXXX0001 @@@@@@@@@@@@@@@@	push(PC+N); PC = Addr16
-------------	--------------------------------------	-------------------------

Call subroutine in ROM (address in register):

syscall rA	11010000XXXXAAAA XXXXXXXXXXXXXXXXXX	push(PC+N); PC = rA; memSpace = ROM
------------	----------------------------------------	----------------------------------------

Call subroutine in ROM (immediate address):

syscall Addr16	11010001XXXX0001 @@@@@@@@@@@@@@@@	push(PC+N); PC = Addr16; memSpace = ROM
----------------	--------------------------------------	--------------------------------------------

Call subroutine in RAM (address in register):

enter rA	11010000XXXXAAAA XXXXXXXXXXXXXXXXXX	push(PC+N); PC = rA; memSpace = RAM
----------	----------------------------------------	----------------------------------------

Call subroutine in RAM (immediate address):

enter Addr16	11010001XXXX0001 @@@@@@@@@@@@@@@@	push(PC+N); PC = Addr16; memSpace = RAM
--------------	--------------------------------------	--------------------------------------------

Operations on each clock cycle:

Fetch instruction	Fetch and update SP Fetch new address	Store PC in stack	Load address into PC, update memory space
-------------------	------------------------------------------	-------------------	----------------------------------------------

Description:

Those instructions push the address of the next 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).

Instructions can be fetched from ROM or RAM. 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
ROM	call	ROM	PC+1
	syscall	ROM	
	enter	RAM	
RAM	call	RAM	PC+2
	syscall	ROM	
	enter	RAM	

Return from subroutine:

Return from call:

ret	11010110XXXX0001 XXXXXXXXXXXXXXXXXX	PC = pop()
-----	----------------------------------------	------------

Return from syscall:

sysret	11010111XXXX0001 XXXXXXXXXXXXXXXXXX	PC = pop() memSpace = RAM
--------	----------------------------------------	------------------------------

Return from enter:

exit	11010111XXXX0001 XXXXXXXXXXXXXXXXXX	PC = pop() memSpace = ROM
------	----------------------------------------	------------------------------

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.

Warning: 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 [DOCS/CESC16.pdf](#)

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

* OS routines can be called from ROM using `call`, but it's recommended to use `syscall` instead.

** 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.