EduMIPS64 Documentation

Release 0.5.3

Andrea Spadaccini (and the EduMIPS64 development team)

September 19, 2011

CONTENTS

1 Source files format **3**
- 1.1 The *.data* section . . . 3
- 1.2 The *.code* section . . . 4
- 1.3 The *#include* command . . . 5

2 The instruction set **7**
- 2.1 ALU Instructions . . . 7
- 2.2 Load/Store instructions . . . 11
- 2.3 Flow control instructions . . . 12
- 2.4 The *SYSCALL* instruction . . . 13
- 2.5 Other instructions . . . 14

3 The user interface **17**
- 3.1 The menu bar . . . 17
- 3.2 Frames . . . 18
- 3.3 Dialogs . . . 19
- 3.4 Command line options . . . 20
- 3.5 Running EduMIPS64 . . . 20

4 Code Examples **23**
- 4.1 SYSCALL . . . 23

EduMIPS64 is a MIPS64 Instruction Set Architecture (ISA) simulator. It is designed to be used to execute small programs that use the subset of the MIPS64 ISA implemented by the simulator, allowing the user to see how instructions behave in the pipeline, how stalls are handled by the CPU, the status of registers and memory and much more. It is both a simulator and a visual debugger.

EduMIPS64 is developed by a group of students of the University of Catania (Italy), and started as a clone of WinMIPS64, even if now there are lots of differences between the two simulators.

This manual will introduce you to EduMIPS64, and will cover some details on how to use it.

The first chapter of this manual covers the format of source files accepted by the simulator, describing the data types and the directives, in addition to command line parameters. In the second chapter there's an overview of the subset of the MIPS64 instruction set that is accepted by EduMIPS64, with all the needed parameters and indications to use them. The third chapter is a description of the user interface of EduMIPS64, that explains the purpose of each frame and menu, along with a description of the configuration dialog, the Dinero frontend dialog, the Manual dialog and command line options. The fourth chapter contains some useful examples.

This manual refers to EduMIPS64 version 0.5.3.

CHAPTER

ONE

SOURCE FILES FORMAT

EduMIPS64 tries to follow the conventions used in other MIPS64 and DLX simulators, so that old time users will not be confused by its syntax.

There are two sections in a source file, the *data* section and the *code* section, introduced respectively by the *.data* and the *.code* directives. In the following listing you can see a very basic EduMIPS64 program:

```
; This is a comment
        .data
label:  .word   15      ; This is an inline comment

        .code
        daddi   r1, r0, 0
        syscall 0
```

To distinguish the various parts of each source code line, any combination of spaces and tabs can be used, as the parser ignores multiple spaces and only detects whitespaces to separate tokens.

Comments can be specified using the ";" character, everything that follows that character will be ignored. So a comment can be used "inline" (after the directive) or on a row by itself.

Labels can be used in the code to reference a memory cell or an instruction. They are case insensitive. Only a label for each source code line can be used. The label can be specified one or more rows above the effective data declaration or instruction, provided that there's nothing, except for comments and empty lines, between the label and the declaration.

1.1 The *.data* section

The *data* section contains commands that specify how the memory must be filled before program execution starts. The general form of a *.data* command is:

```
[label:] .datatype value1 [, value2 [, ...]]
```

EduMIPS64 supports different data types, that are described in the following table.

Type	Directive	Bits required
Byte	*.byte*	8
Half word	*.word16*	16
Word	*.word32*	32
Double Word	*.word* or *.word64*	64

Please note that a double word can be introduced either by the *.word* directive or by the *.word64* directive.

There is a big difference between declaring a list of data elements using a single directive or by using multiple directives of the same type. EduMIPS64 starts writing from the next 64-bit double word as soon as it finds a datatype identifier,

so the first *.byte* statement in the following listing will put the numbers 1, 2, 3 and 4 in the space of 4 bytes, taking 32 bits, while code in the next four rows will put each number in a different memory cell, occupying 32 bytes:

```
.data
.byte    1, 2, 3, 4
.byte    1
.byte    2
.byte    3
.byte    4
```

In the following table, the memory is represented using byte-sized cells and each row is 64 bits wide. The address on the left side of each row of the table refers to the right-most memory cell, that has the lowest address of the eight cells in each line.

0	0	0	0	0	4	3	2	1
8	0	0	0	0	0	0	0	1
16	0	0	0	0	0	0	0	2
24	0	0	0	0	0	0	0	3
36	0	0	0	0	0	0	0	4

There are some special directives that need to be discussed: *.space*, *.ascii* and *.asciiz*.

The *.space* directive is used to leave some free space in memory. It accepts as a parameter an integer, that indicates the number of bytes that must be left empty. It is handy when you must save some space in memory for the results of your computations.

The *.ascii* directive accepts strings containing any of the ASCII characters, and some special C-like escaping sequences, that are described in the following table, and puts those strings in memory.

Escaping sequence	Meaning	ASCII code
\0	Null byte	0
\t	Horizontal tabulation	9
\n	Newline character	10
\”	Literal quote character	34
	Literal backslash character	92

The *.asciiz* directive behaves exactly like the *.ascii* command, with the difference that it automatically ends the string with a null byte.

1.2 The *.code* section

The *code* section contains commands that specify how the memory must be filled when the program will start. The general form of a *.code* command is:

```
[label:] instruction [param1 [, param2 [, param3]]]
```

The *code* section can be specified with the *.text* alias.

The number and the type of parameters depends on the instruction itself.

Instructions can take three types of parameters:

- *Registers* a register parameter is indicated by an uppercase or lowercase “r”, or a $, followed by the number of the register (between 0 and 31), as in “r4”, “R4” or “$4”;
- *Immediate values* an immediate value can be a number or a label; the number can be specified in base 10 or in base 16: base 10 numbers are simply inserted by writing the number, while base 16 number are inserted by putting before the number the prefix “0x”

- *Address* an address is composed by an immediate value followed by a register name enclosed in brackets. The value of the register will be used as base, the value of the immediate will be the offset.

The size of immediate values is limited by the number of bits that are available in the bit encoding of the instruction.

You can use standard MIPS assembly aliases to address the first 32 registers, appending the alias to one of the standard register prefixes like "r", "$" and "R". See the next table.

Register	Alias
0	*zero*
1	*at*
2	*v0*
3	*v1*
4	*a0*
5	*a1*
6	*a2*
7	*a3*
8	*t0*
9	*t1*
10	*t2*
11	*t3*
12	*t4*
13	*t5*
14	*t6*
15	*t7*
16	*s0*
17	*s1*
18	*s2*
19	*s3*
20	*s4*
21	*s5*
22	*s6*
23	*s7*
24	*t8*
25	*t9*
26	*k0*
27	*k1*
28	*gp*
29	*sp*
30	*fp*
31	*ra*

1.3 The *#include* command

Source files can contain the *#include filename* command, which has the effect of putting in place of the command row the content of the file *filename*. It is useful if you want to include external routines, and it comes with a loop-detection algorithm that will warn you if you try to do something like "*#include A.s*" in file *B.s* and "*#include B.s*" in file *A.s*.

CHAPTER

TWO

THE INSTRUCTION SET

In this section we will the subset of the MIPS64 instruction set that EduMIPS64 recognizes. We can operate two different taxonomic classification: one based on the functionality of the instructions and one based on the type of the parameters of the instructions.

The first classification divides instruction into three categories: ALU instructions, Load/Store instructions, Flow control instructions. The next three subsections will describe each category and every instruction that belongs to those categories.

The fourth subsection will describe instructions that do not fit in any of the three categories.

2.1 ALU Instructions

The Arithmetic Logic Unit (in short, ALU) is a part of the execution unit of a CPU, that has the duty of doing arithmetical and logic operations. So in the ALU instructions group we will find those instructions that do this kind of operations.

ALU Instructions can be divided in two groups: *R-Type* and *I-Type*.

Four of those instructions make use of two special registers: LO and HI. They are internal CPU registers, whose value can be accessed through the *MFLO* and *MFHI* instructions.

Here's the list of R-Type ALU Instructions.

- *AND rd, rs, rt*

  Executes a bitwise AND between rs and rt, and puts the result into rd.

- *ADD rd, rs, rt*

  Sums the content of 32-bits registers rs and rt, considering them as signed values, and puts the result into rd. If an overflow occurs then trap.

- *ADDU rd, rs, rt*

  Sums the content of 32-bits registers rs and rt, and puts the result into rd. No integer overflow occurs under any circumstances.

- *DADD rd, rs, rt*

  Sums the content of 64-bits registers rs and rt, considering them as signed values, and puts the result into rd. If an overflow occurs then trap.

- *DADDU rd, rs, rt*

  Sums the content of 64-bits registers rs and rt, and puts the result into rd. No integer overflow occurs under any circumstances.

- *DDIV rs, rt*

  Executes the division between 64-bits registers rs and rt, putting the 64-bits quotient in LO and the 64-bits remainder in HI.

- *DDIVU rs, rt*

  Executes the division between 64-bits registers rs and rt, considering them as unsigned values and putting the 64-bits quotient in LO and the 64-bits remainder in HI.

- *DIV rs, rt*

  Executes the division between 32-bits registers rs and rt, putting the 32-bits quotient in LO and the 32-bits remainder in HI.

- *DIVU rs, rt*

  Executes the division between 32-bits registers rs and rt, considering them as unsigned values and putting the 32-bits quotient in LO and the 32-bits remainder in HI.

- *DMULT rs, rt*

  Executes the multiplication between 64-bits registers rs and rt, putting the low-order 64-bits doubleword of the result into special register LO and the high-order 64-bits doubleword of the result into special register HI.

- *DMULTU rs, rt*

  Executes the multiplication between 64-bits registers rs and rt, considering them as unsigned values and putting the low-order 64-bits doubleword of the result into special register LO and the high-order 64-bits doubleword of the result into special register HI.

- *DSLL rd, rt, sa*

  Does a left shift of 64-bits register rt, by the amount specified in the immediate (positive) value sa, and puts the result into 64-bits register rd. Empty bits are padded with zeros.

- *DSLLV rd, rt, rs*

  Does a left shift of 64-bits register rt, by the amount specified in low-order 6-bits of rs threatd as unsigned value, and puts the result into 64-bits register rd. Empty bits are padded with zeros.

- *DSRA rd, rt, sa*

  Does a right shift of 64-bits register rt, by the amount specified in the immediate (positive) value sa, and puts the result into 64-bits register rd. Empty bits are padded with zeros if the leftmost bit of rt is zero, otherwise they are padded with ones.

- *DSRAV rd, rt, rs*

  Does a right shift of 64-bits register rt, by the amount specified in low-order 6-bits of rs threatd as unsigned value, and puts the result into 64-bits register rd. Empty bits are padded with zeros if the leftmost bit of rt is zero, otherwise they are padded with ones.

- *DSRL rd, rs, sa*

  Does a right shift of 64-bits register rs, by the amount specified in the immediate (positive) value sa, and puts the result into 64-bits register rd. Empty bits are padded with zeros.

- *DSRLV rd, rt, rs*

  Does a right shift of 64-bits register rt, by the amount specified in low-order 6-bits of rs threatd as unsigned value, and puts the result into 64-bits register rd. Empty bits are padded with zeros.

- *DSUB rd, rs, rt*

Subtracts the value of 64-bits register rt to 64-bits register rs, considering them as signed values, and puts the result in rd. If an overflow occurs then trap.

- *DSUBU rd, rs, rt*

  Subtracts the value of 64-bits register rt to 64-bits register rs, and puts the result in rd. No integer overflow occurs under any circumstances.

- *MFLO rd*

  Moves the content of the special register LO into rd.

- *MFHI rd*

  Moves the content of the special register HI into rd.

- *MOVN rd, rs, rt*

  If rt is different from zero, then moves the content of rs into rd.

- *MOVZ rd, rs, rt*

  If rt is equal to zero, then moves the content of rs into rd.

- *MULT rs, rt*

  Executes the multiplication between 32-bits registers rs and rt, putting the low-order 32-bits word of the result into special register LO and the high-order 32-bits word of the result into special register HI.

- *MULTU rs, rt*

  Executes the multiplication between 32-bits registers rs and rt, considering them as unsigned values and putting the low-order 32-bits word of the result into special register LO and the high-order 32-bits word of the result into special register HI.

- *OR rd, rs, rt*

  Executes a bitwise OR between rs and rt, and puts the result into rd.

- *SLL rd, rt, sa*

  Does a left shift of 32-bits register rt, by the amount specified in the immediate (positive) value sa, and puts the result into 32-bits register rd. Empty bits are padded with zeros.

- *SLLV rd, rt, rs*

  Does a left shift of 32-bits register rt, by the amount specified in low-order 5-bits of rs threatd as unsigned value, and puts the result into 32-bits register rd. Empty bits are padded with zeros.

- *SRA rd, rt, sa*

  Does a right shift of 32-bits register rt, by the amount specified in the immediate (positive) value sa, and puts the result into 32-bits register rd. Empty bits are padded with zeros if the leftmost bit of rt is zero, otherwise they are padded with ones.

- *SRAV rd, rt, rs*

  Does a right shift of 32-bits register rt, by the amount specified in low-order 5-bits of rs threatd as unsigned value, and puts the result into 32-bits register rd. Empty bits are padded with zeros if the leftmost bit of rt is zero, otherwise they are padded with ones.

- *SRL rd, rs, sa*

  Does a right shift of 32-bits register rs, by the amount specified in the immediate (positive) value sa, and puts the result into 32-bits register rd. Empty bits are padded with zeros.

- *SRLV rd, rt, rs*

  Does a right shift of 32-bits register rt, by the amount specified in low-order 5-bits of rs threatd as unsigned value, and puts the result into 32-bits register rd. Empty bits are padded with zeros.

- *SUB rd, rs, rt*

  Subtracts the value of 32-bits register rt to 32-bits register rs, considering them as signed values, and puts the result in rd. If an overflow occurs then trap.

- *SUBU rd, rs, rt*

  Subtracts the value of 32-bits register rt to 32-bits register rs, and puts the result in rd. No integer overflow occurs under any circumstances.

- *SLT rd, rs, rt*

  Sets the value of rd to 1 if the value of rs is less than the value of rt, otherwise sets it to 0. This instruction performs a signed comparison.

- *SLTU rd, rs, rt*

  Sets the value of rd to 1 if the value of rs is less than the value of rt, otherwise sets it to 0. This instruction performs an unsigned comparison.

- *XOR rd, rs, rt*

  Executes a bitwise exclusive OR (XOR) between rs and rt, and puts the result into rd.

Here's the list of I-Type ALU Instructions.

- *ADDI rt, rs, immediate*

  Executes the sum between 32-bits register rs and the immediate value, putting the result in rt. This instruction considers rs and the immediate value as signed values. If an overflow occurs then trap.

- *ADDIU rt, rs, immediate*

  Executes the sum between 32-bits register rs and the immediate value, putting the result in rt. No integer overflow occurs under any circumstances.

- *ANDI rt, rs, immediate*

  Executes the bitwise AND between rs and the immediate value, putting the result in rt.

- *DADDI rt, rs, immediate*

  Executes the sum between 64-bits register rs and the immediate value, putting the result in rt. This instruction considers rs and the immediate value as signed values. If an overflow occurs then trap.

- *DADDIU rt, rs, immediate*

  Executes the sum between 64-bits register rs and the immediate value, putting the result in rt. No integer overflow occurs under any circumstances.

- *DADDUI rt, rs, immediate*

  Executes the sum between 64-bits register rs and the immediate value, putting the result in rt. No integer overflow occurs under any circumstances.

- *LUI rt, immediate*

  Loads the constant defined in the immediate value in the upper half (16 bit) of the lower 32 bits of rt, sign-extending the upper 32 bits of the register.

- *ORI rt, rs, immediate*

  Executes the bitwise OR between rs and the immediate value, putting the result in rt.

- *SLTI rt, rs, immediate*

  Sets the value of rt to 1 if the value of rs is less than the value of the immediate, otherwise sets it to 0. This instruction performs a signed comparison.

- *SLTUI rt, rs, immediate*

  Sets the value of rt to 1 if the value of rs is less than the value of the immediate, otherwise sets it to 0. This instruction performs an unsigned comparison.

- *XORI rt, rs, immediate*

  Executes a bitwise exclusive OR (XOR) between rs and the immediate value, and puts the result into rt.

2.2 Load/Store instructions

This category contains all the instructions that operate transfers between registers and the memory. All of these instructions are in the form:

```
[label:] instruction rt, offset(base)
```

Where rt is the source or destination register, depending if we are using a store or a load instruction; offset is a label or an immediate value and base is a register. The address is obtained by adding to the value of the register *base* the immediate value *offset*.

The address specified must be aligned according to the data type that is treated. Load instructions ending with "U" treat the content of the register rt as an unsigned value.

List of load instructions:

- *LB rt, offset(base)*

  Loads the content of the memory cell at address specified by offset and base in register rt, treating it as a signed byte.

- *LBU rt, offset(base)*

  Loads the content of the memory cell at address specified by offset and base in register rt, treating it as an unsigned byte.

- *LD rt, offset(base)*

  Loads the content of the memory cell at address specified by offset and base in register rt, treating it as a double word.

- *LH rt, offset(base)*

  Loads the content of the memory cell at address specified by offset and base in register rt, treating it as a signed half word.

- *LHU rt, offset(base)*

  Loads the content of the memory cell at address specified by offset and base in register rt, treating it as an unsigned half word.

- *LW rt, offset(base)*

  Loads the content of the memory cell at address specified by offset and base in register rt, treating it as a signed word.

- *LWU rt, offset(base)*

Loads the content of the memory cell at address specified by offset and base in register rt, treating it as a signed word.

List of store instructions:

- *SB rt, offset(base)*

  Stores the content of register rt in the memory cell specified by offset and base, treating it as a byte.

- *SD rt, offset(base)*

  Stores the content of register rt in the memory cell specified by offset and base, treating it as a double word.

- *SH rt, offset(base)*

  Stores the content of register rt in the memory cell specified by offset and base, treating it as a half word.

- *SW rt, offset(base)*

  Stores the content of register rt in the memory cell specified by offset and base, treating it as a word.

2.3 Flow control instructions

Flow control instructions are used to alter the order of instructions that are fetched by the CPU. We can make a distinction between these instructions: R-Type, I-Type and J-Type.

Those instructions effectively executes the jump in the ID stage, so often an useless fetch is executed. In this case, two instructions are removed from the pipeline, and the branch taken stalls counter is incremented by two units.

List of R-Type flow control instructions:

- *JALR rs*

  Puts the content of rs into the program counter, and puts into R31 the address of the instruction that follows the JALR instruction, the return value.

- *JR rs*

  Puts the content of rs into the program counter.

List of I-Type flow control instructions:

- *B offset*

  Unconditionally jumps to offset

- *BEQ rs, rt, offset*

  Jumps to offset if rs is equal to rt.

- *BEQZ rs, offset*

  Jumps to offset if rs is equal to zero.

- *BGEZ rs, offset*

  If rs is greather than zero, does a PC-relative jump to offset.

- *BNE rs, rt, offset*

  Jumps to offset if rs is not equal to rt.

- *BNEZ rs*

  Jumps to offset if rs is not equal to zero.\

List of J-Type flow control instructions:

- *J target*

  Puts the immediate value target into the program counter.

- *JAL target*

  Puts the immediate value target into the program counter, and puts into R31 the address of the instruction that follows the JAL instruction, the return value.

2.4 The *SYSCALL* instruction

The SYSCALL instruction offers to the programmer an operating-system-like interface, making available six different system calls.

System calls expect that the address of their parameters is stored in register R14, and will put their return value in register R1.

System calls follow as much as possible the POSIX convention.

2.4.1 *SYSCALL 0 - exit()*

SYSCALL 0 does not expect any parameter, nor it returns anything. It simply stops the simulator.

Note that if the simulator does not find SYSCALL 0 in the source code, or any of its equivalents (HALT - TRAP 0), it will be added automatically at the end of the source.

2.4.2 *SYSCALL 1 - open()*

The SYSCALL 1 expects two parameters: a zero-terminated string that indicates the pathname of the file that must be opened, and a double word containing an integer that indicates the flags that must be used to specify how to open the file.

This integer must be built summing the flags that you want to use, choosing them from the following list:

- *O_RDONLY (0x01)* Opens the file in read only mode;
- *O_WRONLY (0x02)* Opens the file in write only mode;
- *O_RDWR (0x03)* Opens the file in read/write mode;
- *O_CREAT (0x04)* Creates the file if it does not exist;
- *O_APPEND (0x08)* In write mode, appends written text at the end of the file;
- *O_TRUNC (0x08)* In write mode, deletes the content of the file as soon as it is opened.

It is mandatory to specify one of the first three modes. The fourth and the fifth modes are exclusive, you can not specify O_APPEND if you specify O_TRUNC (and vice versa).

You can specify a combination of modes by simply adding the integer values of those flags. For instance, if you want to open a file in write only mode and append the written text to the end of file, you should specify the mode 2 + 8 = 10.

The return value of the system call is the new file descriptor associated with the file, that can be further used with the other system calls. If there is an error, the return value will be -1.

2.4.3 *SYSCALL 2 - close()*

SYSCALL 2 expects only one parameter, the file descriptor of the file that is closed.

If the operation ends successfully, SYSCALL 2 will return 0, otherwise it will return -1. Possible causes of failure are the attempt to close a non-existent file descriptor or the attempt to close file descriptors 0, 1 or 2, that are associated respectively to standard input, standard output and standard error.

2.4.4 *SYSCALL 3 - read()*

SYSCALL 3 expects three parameters: the file descriptor to read from, the address where the read data must be put into, the number of bytes to read.

If the first parameter is 0, the simulator will prompt the user for an input, via an input dialog. If the length of the input is greater than the number of bytes that have to be read, the simulator will show again the message dialog.

It returns the number of bytes that have effectively been read, or -1 if the read operation fails. Possible causes of failure are the attempt to read from a non-existent file descriptor, the attempt to read from file descriptors 1 (standard output) or 2 (standard error) or the attempt to read from a write-only file descriptor.

2.4.5 *SYSCALL 4 - write()*

SYSCALL 4 expects three parameters: the file descriptor to write to, the address where the data must be read from, the number of bytes to write.

If the first parameter is two or three, the simulator will pop the input/output frame, and write there the read data.

It returns the number of bytes that have been written, or -1 if the write operation fails. Possible causes of failure are the attempt to write to a non-existent file descriptor, the attempt to write to file descriptor 0 (standard input) or the attempt to write to a read-only file descriptor.

2.4.6 *SYSCALL 5 - printf()*

SYSCALL 5 expects a variable number of parameters, the first being the address of the so-called "format string". In the format string can be included some placeholders, described in the following list: * *%s* indicates a string parameter; * *%i* indicates an integer parameter; * *%d* behaves like *%i*; * %% literal %

For each *%s*, *%d* or *%i* placeholder, SYSCALL 5 expects a parameter, starting from the address of the previous one.

When the SYSCALL finds a placeholder for an integer parameter, it expects that the corresponding parameter is an integer value, when if it finds a placeholder for a string parameter, it expects as a parameter the address of the string.

The result is printed in the input/output frame, and the number of bytes written is put into R1.

If there's an error, -1 is written to R1.

2.5 Other instructions

In this section there are instructions that do not fit in the previous categories.

2.5.1 *BREAK*

The BREAK instruction throws an exception that has the effect to stop the execution if the simulator is running. It can be used for debugging purposes.

2.5.2 *NOP*

The NOP instruction does not do anything, and it's used to create gaps in the source code.

2.5.3 *TRAP*

The TRAP instruction is a deprecated alias for the SYSCALL instruction.

2.5.4 *HALT*

The HALT instruction is a deprecated alias for the SYSCALL 0 instruction, that halts the simulator.

CHAPTER

THREE

THE USER INTERFACE

The GUI of EduMIPS64 is inspired to WinMIPS64 user interface. In fact, the main window is identical, except for some menus.

The EduMIPS64 main window is composed by a menu bar and six frames, showing different aspects of the simulation. There's also a status bar, that has the double purpose to show the content of memory cells and registers when you click them and to notify the user that the simulator is running when the simulation has been started but verbose mode is not selected. There are more details in the following section.

3.1 The menu bar

The menu bar contains six menus:

3.1.1 File

The File menu contains menu items about opening files, resetting or shutting down the simulator, writing trace files.

- *Open...* Opens a dialog that allows the user to choose a source file to open.
- *Open recent* Shows the list of the recent files opened by the simulator, from which the user can choose the file to open
- *Reset* Resets the simulator, keeping open the file that was loaded but resetting the execution.
- *Write Dinero Tracefile...* Writes the memory access data to a file, in xdin format.
- *Exit* Closes the simulator.

The *Write Dinero Tracefile...* menu item is only available when a whole source file has been executed and the end has been already reached.

3.1.2 Execute

The Execute menu contains menu items regarding the execution flow of the simulation.

- *Single Cycle* Executes a single simulation step
- *Run* Starts the execution, stopping when the simulator reaches a *SYSCALL 0* (or equivalent) or a *BREAK* instruction, or when the user clicks the Stop menu item (or presses F9).
- *Multi Cycle* Executes some simulation steps. The number of steps executed can be configured through the Setting dialog.

- *Stop* Stops the execution when the simulator is in "Run" or "Multi cycle" mode, as described previously.

This menu is only available when a source file is loaded and the end of the simulation is not reached. The *Stop* menu item is available only in "Run" or "Multi Cycle" mode.

3.1.3 Configure

The Configure menu provides facilities for customizing EduMIPS64 appearance and behavior.

- *Settings...* Opens the Settings dialog, described in the next sections of this chapter;
- *Change Language* Allows the user to change the language used by the user interface. Currently only English and Italian are supported. This change affects every aspect of the GUI, from the title of the frames to the online manual and warning/error messages.

The *Settings...* menu item is not available when the simulator is in "Run" or "Multi Cycle" mode, because of potential race conditions.

3.1.4 Tools

This menu contains only an item, used to invoke the Dinero Frontend dialog.

- *Dinero Frontend...* Opens the Dinero Frontend dialog.

This menu is not available until you have not executed a program and the execution has reached its end.

3.1.5 Window

This menu contains items related to operations with frames.

- *Tile* Sorts the visible windows so that no more that three frames are put in a row. It tries to maximize the space occupied by every frame.

The other menu items simply toggle the status of each frame, making them visible or minimizing them.

3.1.6 Help

This menu contains help-related menu items.

- *Manual...* Shows the Help dialog.
- *About us...* Shows a cute dialog that contains the names of the project contributors, along with their roles.

3.2 Frames

The GUI is composed by seven frames, six of which are visible by default, and one (the I/O frame) is hidden.

3.2.1 Cycles

The Cycles frame shows the evolution of the execution flow during time, showing for each time slot which instructions are in the pipeline, and in which stage of the pipeline they are located.

3.2.2 Registers

The Registers frame shows the content of each register. By left-clicking on them you can see in the status bar their decimal (signed) value, while double-clicking on them will pop up a dialog that allows the user to change the value of the register.

3.2.3 Statistics

The Statistics frame shows some statistics about the program execution.

3.2.4 Pipeline

The Pipeline frame shows the actual status of the pipeline, showing which instruction is in which pipeline stage. Different colors highlight different pipeline stages.

3.2.5 Memory

The Memory frame shows memory cells content, along with labels and comments taken from the source code. Memory cells content, like registers, can be modified double-clicking on them, and clicking on them will show their decimal value in the status bar. The first column shows the hexadecimal address of the memory cell, and the second column shows the value of the cell. Other columns show additional info from the source code.

3.2.6 Code

The Code window shows the instructions loaded in memory. The first column shows the address of the instruction, while the second column shows the hexadecimal representation of the instructions. Other columns show additional info taken from the source code.

3.2.7 Input/Output

The Input/Output window provides an interface for the user to see the output that the program creates through the SYSCALLs 4 and 5. Actually it is not used for input, as there's a dialog that pops up when a SYSCALL 3 tries to read from standard input, but future versions will include an input text box.

3.3 Dialogs

Dialogs are used by EduMIPS64 to interact with the user in many ways. Here's a summary of the most important dialogs:

3.3.1 Settings

In the Settings dialog various aspects of the simulator can be configured.

The Main Settings tab allow to configure forwarding and the number of steps in the Multi Cycle mode.

The Behavior tab allow to enable or disable warnings during the parsing phase, the "Sync graphics with CPU in multi-step execution" option, when checked, will synchronize the frames' graphical status with the internal status of the simulator. This means that the simulation will be slower, but you'll have an explicit graphical feedback of what is

happening during the simulation. If this option is checked, the "Interval between cycles" option will influence how many milliseconds the simulator will wait before starting a new cycle. Those options are effective only when the simulation is run using the "Run" or the "Multi Cycle" options from the Execute menu.

The last two options set the behavior of the simulator when a synchronous exception is raised. If the "Mask synchronous exceptions" option is checked, the simulator will ignore any Division by zero or Integer overflow exception. If the "Terminate on synchronous exception" option is checked, the simulation will be halted if a synchronous exception is raised. Please note that if synchronous exceptions are masked, nothing will happen, even if the termination option is checked. If exceptions are not masked and the termination option is not checked, a dialog will pop out, but the simulation will go on as soon as the dialog is closed. If exceptions are not masked and the termination option is checked, the dialog will pop out, and the simulation will be stopped as soon as the dialog is closed.

The last tab allows to change the colors that are associated to the different pipeline stages through the frames. It's pretty useless, but it's cute.

3.3.2 Dinero Frontend

The Dinero Frontend dialog allows to feed a DineroIV process with the trace file internally generated by the execution of the program. In the first text box there is the path of the DineroIV executable, and in the second one there must be the parameters of DineroIV.

The lower section contains the output of the DineroIV process, from which you can take the data that you need.

3.3.3 Help

The Help dialog contains three tabs with some indications on how to use the simulator. The first one is a brief introduction to EduMIPS64, the second one contains informations about the GUI and the third contains a summary of the supported instructions.

3.4 Command line options

Three command line options are available. They are described in the following list, with the long name enclosed in round brackets. Long and short names can be used in the same way.

- *-h (–help)* shows a help message containing the simulator version and a brief summary of command line options
- *-f (–file) filename* opens *filename* in the simulator
- *-d (–debug)* enters Debug mode

The *–debug* flag has the effect to activate Debug mode. In this mode, a new frame is available, the Debug frame, and it shows the log of internal activities of EduMIPS64. It is not useful for the end user, it is meant to be used by EduMIPS64 developers.

3.5 Running EduMIPS64

The EduMIPS64 *.jar* file can act both as a stand-alone executable *.jar* file and as an applet, so it can be executed in both ways. Both methods need the Java Runtime Environment, version 5 or later.

To run it as a stand-alone application, the *java* executable must be issued in this way: *java -jar edumips64-version.jar*, where the *version* string must be replaced with the actual version of the simulator. On some systems, you may be able to execute it by just clicking on the *.jar* file.

To embed it in an HTML, the <*applet*> tag must be used. The EduMIPS64 web site contains a page that already contains the applet, so that everyone can execute it without the hassle of using the command line.

CHAPTER

FOUR

CODE EXAMPLES

In this chapter you'll find some sample listings that will be useful in order to understand how EduMIPS64 (version 0.5.3) works.

4.1 SYSCALL

It's important to understand that examples for SYSCALL 1-4 refer to the *print.s* file, that is the example for SYSCALL 5. If you want to run the examples, you should copy the content of that example in a file named *print.s* and include it in your code.

Some examples use an already existing file descriptor, even if it doesn't truly exist. If you want to run those examples, use the SYSCALL 1 example to open a file.

4.1.1 SYSCALL 0

When SYSCALL 0 is called, it stops the execution of the program. Example:

```
.code
daddi   r1, r0, 0    ; saves 0 in R1
syscall 0            ; exits
```

4.1.2 SYSCALL 1

Example program that opens a file:

```
                .data
error_op:       .asciiz     "Error opening the file"
ok_message:     .asciiz     "All right"
params_sys1:    .asciiz     "filename.txt"
                .word64     0xF

                .text
open:           daddi       r14, r0, params_sys1
                syscall     1
                daddi       $s0, r0, -1
                dadd        $s2, r0, r1
                daddi       $a0,r0,ok_message
                bne         r1,$s0,end
                daddi       $a0,r0,error_op
```

```
end:            jal         print_string
                syscall 0

                #include    print.s
```

In the first two rows we write to memory the strings containing the error message and the success message that we will pass to print_string function, and we give them two labels. The print_string function is included in the print.s file.

Next, we write to memory the data required from SYSCALL 1 (row 4, 5), the path of the file to be opened (that must exist if we work in read or read/write mode) and, in the next memory cell, an integer that defines the opening mode.

In this example, the file was opened using the following modes: *O_RDWR* textbar{} *O_CREAT* textbar{} *O_APPEND*. The number 15 (0xF in base 16) comes from the sum of the values of these three modes (3 + 4 + 8).

We give a label to this data so that we can use it later.

In the .text section, we save the address of params_sys1 (that for the compiler is a number) in register r14; next we can call SYSCALL 1 and save the content of r1 in $s2, so that we can use it in the rest of the program (for instance, with other SYSCALL).

Then the print_string function is called, passing error_op as an argument if r1 is equal to -1 (rows 13-14) or else passing ok_message as an argument if everything went smoothly (rows 12 and 16).

4.1.3 SYSCALL 2

Example program that closes a file:

```
                .data
params_sys2:    .space 8
error_cl:       .asciiz     "Error closing the file"
ok_message:     .asciiz     "All right"

                .text
close:          daddi       r14, r0, params_sys2
                sw          $s2, params_sys2(r0)
                syscall     2
                daddi       $s0, r0, -1
                daddi       $a0, r0, ok_message
                bne         r1, $s0, end
                daddi       $a0, r0, error_cl

end:            jal         print_string
                syscall     0

                #include    print.s
```

First we save some memory for the only argument of SYSCALL 2, the file descriptor of the file that must be closed (row 2), and we give it a label so that we can access it later.

Next we put in memory the strings containing the error message and the success message, that will be passed to the print_string function (rows 3, 4).

In the .text section, we save the address of params_sys2 in r14; then we can call SYSCALL 2.

Now we call the print_string function using error_cl as a parameter if r1 yields -1 (row 13), or we call it using ok_message as a parameter if all went smoothly (row 11).

Note: This listing needs that registry $s2 contains the file descriptor of the file to use.

4.1.4 SYSCALL 3

Example program that reads 16 bytes from a file and saves them to memory:

```
                .data
params_sys3:    .space      8
ind_value:      .space      8
                .word64     16
error_3:        .asciiz     "Error while reading from file"
ok_message:     .asciiz     "All right"

value:          .space      30

                .text
read:           daddi       r14, r0, params_sys3
                sw          $s2, params_sys3(r0)
                daddi       $s1, r0, value
                sw          $s1, ind_value(r0)
                syscall     3
                daddi       $s0, r0, -1
                daddi       $a0, r0,ok_message
                bne         r1, $s0,end
                daddi       $a0, r0,error_3

end:            jal         print_string
                syscall     0

                #include    print.s
```

The first 4 rows of the .data section contain the arguments of SYSCALL 3, the file descriptor of the from which we must read, the memory address where the SYSCALL must save the read data, the number of bytes to read. We give labels to those parameters that must be accessed later. Next we put, as usual, the strings containing the error message and the success message.

In the .text section, we save the params_sys3 address to register r14, we save in the memory cells for the SYSCALL parameters the file descriptor (that we suppose to have in $s2) and the address that we want to use to save the read bytes.

Next we can call SYSCALL 3, and then we call the print_string function passing as argument error_3 or ok_message, according to the success of the operation.

4.1.5 SYSCALL 4

Example program that writes to a file a string:

```
                .data
params_sys4:    .space      8
ind_value:      .space      8
                .word64     16
error_4:        .asciiz     "Error writing to file"
ok_message:     .asciiz     "All right"
value:          .space      30

                .text

write:          daddi       r14, r0,params_sys4
                sw          $s2, params_sys4(r0)
                daddi       $s1, r0,value
```

```
                sw          $s1, ind_value(r0)
                syscall     4
                daddi       $s0, r0,-1
                daddi       $a0, r0,ok_message
                bne         r1, $s0,end
                daddi       $a0, r0,error_4

end:            jal         print_string
                syscall     0

                #include    print.s
```

The first 4 rows of the .data section contain the arguments of SYSCALL 4, the file descriptor of the from which we must read, the memory address from where the SYSCALL must read the bytes to write, the number of bytes to write. We give labels to those parameters that must be accessed later. Next we put, as usual, the strings containing the error message and the success message.

In the .text section, we save the params_sys4 address to register r14, we save in the memory cells for the SYSCALL parameters the file descriptor (that we suppose to have in $s2) and the address from where we must take the bytes to weite.

Next we can call SYSCALL 3, and then we call the print_string function passing as argument error_3 or ok_message, according to the success of the operation.

4.1.6 SYSCALL 5

Example program that contains a function that prints to standard output the string contained in $a0:

```
                .data
params_sys5:    .space  8

                .text
print_string:
                sw      $a0, params_sys5(r0)
                daddi   r14, r0, params_sys5
                syscall 5
                jr      r31
```

The second row is used to save space for the string that must be printed by the SYSCALL, that is filled by the first instruction of the .text section, that assumes that in $a0 there's the address of the string to be printed.

The next instruction puts in r14 the address of this string, and then we can call SYSCALL 5 and print the string. The last instruction sets the program counter to the content of r31, as the usual MIPS calling convention states.

4.1.7 A more complex usage example of SYSCALL 5

SYSCALL 5 uses a not-so-simple arguments passing mechanism, that will be shown in the following example:

```
                .data
format_str:     .asciiz   "%dth of %s:\n%s version %i.%i is being tested!"
s1:             .asciiz   "June"
s2:             .asciiz   "EduMIPS64"
fs_addr:        .space    4
                .word     5
s1_addr:        .space    4
s2_addr:        .space    4
```

```
                .word     0
                .word     5
test:
                .code
                daddi     r5, r0, format_str
                sw        r5, fs_addr(r0)
                daddi     r2, r0, s1
                daddi     r3, r0, s2
                sd        r2, s1_addr(r0)
                sd        r3, s2_addr(r0)
                daddi     r14, r0, fs_addr
                syscall   5
                syscall   0
```

The address of the format string is put into R5, whose content is then saved to memory at address fs_addr. The string parameters' addresses are saved into s1_addr and s2_addr. Those two string parameters are the ones that match the two %s placeholders in the format string.

Looking at the memory, it's obvious that the parameters matching the placeholders are stored immediately after the address of the format string: numbers match integer parameters, while addresses match string parameters. In the s1_addr and s2_addr locations there are the addresses of the two strings that we want to print instead of the %s placeholders.

The execution of the example will show how SYSCALL 5 can handle complex format strings like the one stored at format_str.