Assembler specification

The assembler must read an assembly file containing a program that follows the ISA spec, and it must output a MIF file with 2048 32-bit words of memory (8192 bytes in total).

Opcode and function mappings

For the following tables, the rows indices represent the most significant bits (MSB) and the column indices are the LSB.

General opcode mapping

	00	01	10	11
00	ALU-R		CMP-R	
01		Store	BRANCH	
10	ALU-I	Load	CMP-I	JAL
11				

ALU-R/ALU-I function mapping

	00	01	10	11
00	ADD/ADDI	SUB/SUBI		
01	AND/ANDI	OR/ORI	XOR/XORI	
10				

11	NAND/NANDI	NOR/NORI	NXOR/NXORI	

CMP-R/CMP-I function mapping

	00	**01**	**10**	**11**
00	F/FI	EQ/EQI	LT/LTI	LTE/LTEI
01				
10	T/TI	NE/NEI	GTE/GTEI	GT/GTI
11				

Instruction format

- ALU-R
  - rd = rs1 op rs2
- CMP-R
  - rd = (rs1 op rs2)? 1: 0
- Store
  - Mem[rs1 + signextension(imm)] = rs2
- Load
  - rd = Mem[rs1 + signextension(imm)]
- ALU-I
  - rd = rs1 op signextension(imm)
- CMP-I
  - rd = (rs1 op signextension(imm)) ? 1 : 0
- BRANCH
  - if (rs1 op rs2) PC = PC + 4 + (signextension(imm) * 4)
- JAL
  - rd = PC + 4
  - PC = rs1 + 4 * signextension(imm)

Assembler syntax

- Instruction opcodes and register names
  - Are reserved words (can't be used as labels)
  - Appear in either lowercase or uppercase
  - If there is a destination register, it is listed first
- Labels
  - Created using a name and then ":" at the start of a line
  - Corresponds to the address where label created
- Immediate operands – number or label
  - If number, hex (C format, e.g. 0xffff) or decimal (can have - sign)
  - If label, just use the name of the lable (without ":")
    - For PC-relative, the immediate field is label_addr-PC-4
    - For other insts, the immediate field is 16 least-significant bits of label_addr
- Each register has multiple names:
  - R0..R3 are also A0..A3 (function arguments, caller saved)
  - R3 is also RV (return value, caller saved)
  - R4..R5 are also T0..T1 (temporaries, caller saved)
  - R6..R8 are also S0..S2 (callee-saved values)
  - R9 reserved for assembler use
  - R10..R11 reserved for system use (we'll see later for what)
  - R12 is GP (global pointer)
  - R13 is FP (frame pointer)
  - R14 is SP (stack pointer)
    - Stack grows down, SP points to lowest in-use address
  - R15 is RA (return address)

Special assembler instructions

- .ORG <number>
  - Changes "current" address to <number>
- .WORD <value>
  - Places 32-bit word <value> at the current address
  - <value> can be a number or a label name
  - If label name, value is the full 32-bit label_addr
- .NAME <name>=<value>
  - Defines a name (label) with a given value (number)