mips汇编转换机器码模拟

一.总体思路

1. 从汇编指令集当中读入一行一行的 mips 指令
2. 通过对指令各个部分的分析，得到相应的机器码
3. 将对应的指令和它的机器码输出到标准输出
4. 同时输出到二进制文件当中

二.实现的指令

R-type：

add rd, rs, rt 100000

addu rd, rs, rt 100001

and rd, rs, rt 100100

nor rd, rs, rt 100111

or rd, rs, rt 100101

slt rd, rs, rt 101010

sltu rd, rs, rt 101011

sub rd, rs, rt 100010

subu rd, rs, rt 100011

xor rd, rs, rt 100110

srav rd, rt, rs 000111

srlv rd, rt, rs 000110

sllv rd, rt, rs 000100

sll rd, rt, sa 000000

srl rd, rt, sa 000010

sra rd, rt, sa 000011

div rs, rt 011010

divu rs, rt 011011

mult rs, rt 011000

multu rs, rt 011001

jalr rd, rs 001001

jr rs 001000

mthi rs 010001

mtlo rs 010011

mfhi rd 010000

mflo rd 010010

syscall 001100

break 001101

I-type

:

addi rt, rs, immediate 001000

addiu rt, rs, immediate 001001

andi	rt, rs, immediate	001100	
ori	rt, rs, immediate	001101	
xori	rt, rs, immediate	001110	
slti	rt, rs, immediate	001010	
sltiu	rt, rs, immediate	001011	
beq	rs, rt, label	000100	
bne	rs, rt, label	000101	
bgez	rs, label	000001	rt = 00001
bgtz	rs, label	000111	rt = 00000
blez	rs, label	000110	rt = 00000
bltz	rs, label	000001	rt = 00000
lui	rt, immediate	001111	
lb	rt, immediate(rs)	100000	
lbu	rt, immediate(rs)	100100	
lh	rt, immediate(rs)	100001	
lhu	rt, immediate(rs)	100101	
lw	rt, immediate(rs)	100011	
lwc1	rt, immediate(rs)	110001	

sb rt, immediate(rs) 101000

sh rt, immediate(rs) 101001

sw rt, immediate(rs) 101011

swc1 rt, immediate(rs) 111001

J-type:

j label 000010 coded address of label

jal label 000011 coded address of label

Pseudo Instructions:

blt $r1 $r2 label → slt $at $r1 $r2 bne $at $zero label

bgt $r1 $r2 label → slt $at $r2 $r1 bne $at $zero label

ble $r1 $r2 label → slt $at $r2 $r1 beq $at $zero label

bge $r1 $r2 label → slt $at $r1 $r2 beq $at, $zero, label

三.使用方法

源文件 mips.c 和测试数据 instructions.txt 在文件夹 src 当中，可直接编译运行。

src 文件夹中还有对应的支持指令集 instruction 和 寄存器集 regs

编译得到的程序不需要输入，要改变测试只要改变 instructions.txt 当中的指令

若在 linux 系统中，可直接使用 make 命令编译得到可执行程序，直接运行可以得到 instructions.txt 当中指令对应的机器码，以及得到名为 out 的二进制文件，内容如同输出

四.已知问题

对 j 命令的 label 的处理可能有问题，整个程序相当于只有一遍编译过程，j 类型命令的立即数仍然采用了将对应十进制转化成 2 进制，而不是绝对地址。

五.实例

instructions.txt 当中的指令对应机器码， 开头为行号。

1 add $s0 $s0 $s1

00000010000100011000000000100000

2 srl $t1 $t2 10

00000000000010100100101010000010

3 addu $t1 $s2 $s3

00000010010100110100100000100001

4 sub $s0 $s0 $s1

00000010000100011000000000100010

5 subu $t1 $s2 $s3

00000010010100110100100000100011

6 sll $t1 $t2 9

00000000000010100100101001000000

7 and $t2 $s4 $s5

00000010100101010101000000100100

8 nor $t3 $6 $7

00000000000000000101100000100111

9 or $t4 $t5 $t6

00000001101011100110000000100101

10 xor $t7 $s3 $s4

00000010011101000111100000100110

11 sra $t1 $t2 11

00000000000010100100101011000011

12

13 addi $s0 $s1

00100010001100000000000000000000

14 addiu $s1 $s0

00100110000100010000000000000000

15 beq $s2 $s3 20

00010010010100110000000000000000

16 bne $s4 $s5 25

00010110100101010000000000000000

17 ori $t1 $t2 100

00110101010010010000000001100100

18 xori $t2 $t1 -52

00111001001010101111111111001100

19 sw $s1 -10($t2)

10101101010100011111111111110110

20

21

22 blt $s1 $s2 15

00000010001100100000100000101010

00010100001000000000000000001111

23 bgt $t1 $t2 20

00000001010010010000100000101010

00010100001000000000000000010100

24 ble $t1 $t2 25

00000001010010010000100000101010

00010000001000000000000000011001

25 bge $s1 $s2 -30

00000010001100100000100000101010

00010000001000001111111111100010

26

27 j 12345

00001000000000000011000000111001

28 jal -12345

00001111111111111100111111000111