

Atividade 3 - Org. Arg. 1

Profa. Dra. Cíntia Borges Margi
(cintia@usp.br)

Guilherme S. Salustiano
(salustiano@usp.br)

Parte 1 - Emulação de ponto flutuante

Contexto

Podem ser encontrado no mercado diversos processadores sem unidade de ponto flutuante, seja por uma questão de custo, área ou limite energético. Entretanto o 'float' e 'double' são definidos no C, assim sendo os códigos contendo esse tipo precisam rodar em todas as arquiteturas. Como isso é possível?

Para investigar vamos tentar compilar o seguinte código para uma arquitetura sem ponto flutuante, no caso do RISC-V isso é representado pela ausência da letra f e d da ISA, por se tratar de uma ISA modular ela tem a base de inteiros (como rv32i ou rv64i), e cada letra representa um novo conjunto de instruções [1].

Para isso vamos usar o [godbolt](#), com o compilador RISC-V (64-bits) gcc com as flags -O2 -march=rv64i -mabi=lp64, em -march definimos a arquitetura alvo como um processador RISC-V sem nenhuma extensão e em -mabi definimos a interface binária para também não usar ponto flutuante.

Agora digitando o seguinte código:

```
float sum(float x, float y) {  
    return x + y;  
}
```

Observamos a seguinte saída:

```
sum(float, float):  
    mv     a5,a1  
    mv     a1,a0  
    addi   sp,sp,-16  
    mv     a0,a5  
    sd     ra,8(sp)  
    call   __addsf3  
    ld     ra,8(sp)  
    addi   sp,sp,16  
    jr     ra
```

Conforme esperado o processador não realiza a soma, e delegou a função __addsf3 que recebe y e x nos registradores a0 e a1 (observe que ele precisa alterar os parâmetros, uma vez que a soma de pontos flutuantes não é cumulativa).

A função __addsf3 faz parte da biblioteca de runtime de C, que permite a portabilidade da linguagem em diversos sistemas. As outras rotinas podem ser encontradas [aqui](#).

Tarefa

Implemente as seguintes funções de uma biblioteca de ponto flutuante:

- mfloat floatsisf (mint i) - converte um inteiro para a representação ponto flutuante
- mint fixsfsi (mfloat a) - converte um ponto flutuante para a representação inteira
- mfloat negsf2 (mfloat a) - retorna o negado de a (Dica: é apenas um bit flip)
- mfloat addsf3 (mfloat a, mfloat b) - retorna a soma entre a e b
- mfloat subsf3 (mfloat a, mfloat b) - retorna a subtração entre a e b (Dica: pode ser definido a partir da combinação de outras duas funções)

Sendo mfloat e mint tipos definidos a partir da inttypes.h para garantir compatibilidade.

```
#include <inttypes.h>
```

```
typedef mint int32_t  
typedef mfloat uint32_t
```

O código não pode conter a palavra reservada float ou double e será checado estaticamente. Palavras derivadas (acrecidas de caracteres antes ou após) como float_valor, valor_float, são permitidas.

Seu código não precisa tratar casos os casos de NaN, Infinity ou underflow.

Entrega final

Ao final, gere um zip atv3.zip com os arquivos.

```
atv3.zip  
└─ float_lib.c
```

Bibliografia

- [1] Wikipedia, “RISC-V --- Wikipedia, the free encyclopedia,” 2023. ([url{http://en.wikipedia.org/w/index.php?title=RISC-V&oldid=1179194505}](http://en.wikipedia.org/w/index.php?title=RISC-V&oldid=1179194505))