Universidade Federal de Minas Gerais Instituto de Ciências Exatas Departamento de Ciência da Computação Software Básico (DCC008) - 2017/1

Trabalho Prático 1 - Montador

1. Descrição Geral

O principal objetivo deste trabalho é implementar um montador para a máquina *Swombat*, a ser executada no simulador CPUSim, ambos disponibilizados com este documento.

O CPUSim é desenvolvido em linguagem Java, e é capaz de simular uma máquina completa. Ele recebe como entrada um arquivo contendo os dados em linguagem de máquina, o qual é carregado na memória RAM da máquina *Swombat*, deixando-o pronto para execução.

Mais informações sobre o CPUSim podem ser encontradas na página oficial do simulador, em http://www.cs.colby.edu/djskrien/CPUSim/.

2. Informações Importantes

- O trabalho deve ser feito em duplas ou trios, podendo ser discutido entre os colegas desde que não haja cópia ou compartilhamento do código fonte.
- A data de entrega será especificada através de uma tarefa no Moodle.
- Os trabalhos poderão ser entregues até às 23:55 do dia especificado para a entrega. O horário de entrega deve respeitar o relógio do sistema Moodle. Haverá uma tolerância de 5 minutos de atraso, de forma que os alunos podem fazer a entrega até às 0:00. A partir desse horário, os trabalhos já estarão sujeitos a penalidades. A fórmula para desconto por atraso na entrega do trabalho prático é:

 $Desconto = 2^d/0.32 \%$

onde d é o atraso em dias úteis. Note que após 5 dias úteis, o trabalho não pode ser mais entregue.

- O trabalho deve ser implementado obrigatoriamente na linguagem C ou C++.
- Deverá ser entregue o código fonte com os arquivos de dados necessários para a execução e um arquivo *Makefile* que permita a compilação do programa nas máquinas Linux do DCC.
- Além disso, deverá ser entregue uma pequena documentação contendo todas as decisões de projeto que foram tomadas durante a implementação, sobre aspectos não contemplados na especificação, assim como uma justificativa para essas decisões. Esse documento não precisa ser extenso (mínimo 3 e máximo de 6 páginas). A documentação deve indicar o nome dos alunos integrantes do grupo.

- A ênfase do trabalho está no funcionamento do sistema e não em aspectos de programação ou interface com o usuário. Assim, não deverá haver tratamento de erros no programa de entrada.
- Todas as dúvidas referentes ao trabalho serão esclarecidas por meio do fórum disponível no ambiente Moodle da disciplina.
- A entrega do trabalho deverá ser realizada pelo Moodle, na tarefa criada especificamente para tal. A entrega deverá ser feita no seguinte formato:
 - O trabalho a ser entregue deverá estar contido em um único arquivo compactado, em formato ".zip", com o nome no formato "tp1 aluno1 aluno3.zip"
 - O arquivo .zip definido deverá ter três pastas:
 - o "assembler": Essa pasta deverá conter o código-fonte do montador implementado, juntamente do arquivo *Makefile* (OBS.: Não devem ser incluídos arquivos .o nem executáveis nessa pasta.)
 - o "tst": Essa pasta irá conter os arquivos ".a" de entrada desenvolvidos para testar o montador. Mais detalhes desse arquivo estão definidos na Seção 5.
 - o "doc": Essa pasta deverá conter o arquivo da documentação, em formato PDF. Caso o grupo julgue necessário incluir quaisquer outros arquivos a parte, esses deverão ser justificados em um arquivo texto com o nome README.
 - Atenção: Trabalhos que descumprirem o padrão definido acima serão penalizados.

3. Especificação da Máquina Swombat

A máquina a ser utilizada é a *Swombat*, projetada para uso no simulador CPUSim. Seguem as especificações:

- A menor unidade endereçável nessa máquina é uma palavra de 16 bits (um inteiro).
- Os tipos de dados tratados pela máquina são somente inteiros.
- A máquina possui uma memória principal de 256 posições, numeradas de 0 a 255.
- O endereço 254 (última palavra de 16 bits) é um endereço especial para E/S (Entrada e Saída). Quaisquer valores armazenados nesse endereço (através da instrução storei, descrita no Anexo 1 deste documento) serão impressos na tela. Por outro lado, toda instrução de carregamento nesse endereço (loadi, também descrita no Anexo 1) será interpretada como uma solicitação de dados ao usuário, o qual deverá digitar um valor inteiro através de um console do CPUSim. O registrador de destino definido na instrução deverá receber o valor informado.
- A memória é dividida em células de 8 bits. Dessa forma, uma palavra de 16 bits situada no endereço de memória X será seguida por outra palavra no endereço X+2, e assim por diante.
- Existe também uma pilha (127 elementos de 8 bits) acessada por instruções específicas (push e pop)
- Cada registrador pode armazenar uma palavra de 16 bits.
- Existem 8 registradores de uso geral, os quais são nomeados de "R0" a "R7".
- Os registradores de propósito específico são:
 - o **stackpt** (Stack Pointer): O ponteiro de pilhas aponta para o topo da pilha (inicializado com 0).
 - o **sdr:** Armazena um dado recém-lido ou a ser escrito na pilha.
 - o **pc:** (Program Counter): Armazena o contador de programa, ou seja, o endereço de memória da próxima instrução a ser executada. O *pc* inicia a execução do programa

na posição 0 (zero) da memória e é automaticamente incrementado a cada ciclo de instrução de forma que as instruções são normalmente executadas sequencialmente a partir da memória. O pc é afetado também pelas instruções de desvio e chamadas de procedimentos.

- buffer1 e buffer2: Registradores que recebem operandos a serem tratados, como por exemplo, os operandos da instrução *add*.
- o **ir**: Armazena a instrução em execução. É a partir desse registrador que a instrução é decodificada.
- o mar: Armazena um endereço a ser lido ou escrito na memória.
- o **mdr:** Armazena um dado recém-lido (por exemplo, uma instrução recém-carregada) ou a ser escrito na memória.
- o **status:** Armazena uma flag chamada *halt-bit* que, se escrita, interrompe o programa imediatamente, gerando uma mensagem de encerramento na tela.
- Cada instrução pode ter 0, 1 ou 2 operandos, e sempre possui 16 bits de tamanho. Algumas instruções são especificadas com campos não utilizados, e portanto, devem ser escritos 0's (zeros) nessa região. O Anexo 1 deste documento mostra a ordem dos operandos (incluindo os campos não utilizados) em cada instrução de máquina, bem como o número de bits de cada operando. Essa ordem deve ser respeitada pelo montador, caso contrário, o programa não funcionará corretamente.

4. Especificação do Montador

- O montador a ser implementado é um montador de 2 passos, conforme descrito no capítulo 7 do livro utilizado na disciplina.
- O conjunto de instruções é o especificado pela máquina *Swombat*, e está descrito em mais detalhes no Anexo 1 deste documento.
- Cada linha da linguagem de montagem da máquina *Swombat* possui o seguinte formato:
 - _[rótulo:] operador [operando(s)] [;comentário]

Ou seja:

- Se houver algum rótulo (*label*), ele será definido no início da linha. Todo rótulo deverá iniciar com um *underscore* ("_") e finalizar com dois-pontos (":").
 - Exemplo: _label1: add r0 r1 r2 ;soma de dois números
- A presença do operador é obrigatória, pois o mesmo identifica a instrução de máquina a ser executada.
- A presença ou não de operandos depende da instrução, tendo em vista que o número de operandos varia de instrução para instrução. Se houverem dois operandos, estes serão separados por espaços ("").
- Um comentário pode ser incluído opcionalmente no fim da linha, devendo necessariamente começar por um ponto-e-vírgula (;) e devendo ser ignorado pelo montador.
- O endereço de memória indicado nos operandos das instruções, inclusive nas de desvio, é a posição absoluta da memória, ou seja, nenhum pré-tratamento deve ser feito sobre esse endereço.
- Rótulo, operador, operandos e comentário deverão ser separados por espaços (" "), assim como dois operandos mencionados anteriormente. Poderá haver mais de um espaço, o que não deve afetar o funcionamento do montador.

- Não deve haver linhas vazias e linhas contendo apenas comentários ou rótulos.
- A pseudo-instrução .*data* (inclui ponto no nome) deverá ser tratada pelo montador, e sua função será a de reservar uma posição da memória da máquina *Swombat*.
 - Formato: label: .data num_bytes valor_inicial"
 - A instrução servirá para alocar uma região de memória de tamanho 'num_bytes', com "valor inicial". Essa região de memória será identificada pelo rótulo 'label', o qual pode ser usado ao longo do código Assembly para acessar aquela região de memória.
 - A política de alocação de variáveis na memória deverá ser definida pelo montador. A
 documentação deve explicar e justificar como foi feita a alocação (por exemplo o
 início da memória de dados, se a alocação cresce para cima ou para baixo, entre
 outras definições).
- IMPORTANTE: o programa a ser carregado na máquina deverá sempre iniciar na posição 0
 (zero) de memória, uma vez que o contador de programa (PC) da máquina Swombat é
 inicializado por padrão com este valor.

5. Formato de Entrada de Dados no Montador

O programa a ser traduzido pelo montador deverá ser escrito em um arquivo texto simples com formato ".a", sendo que as instruções devem ser dispostas uma por linha no arquivo, e deverão ser lidas pelo montador. Um arquivo exemplo será disponibilizado no Moodle. **Atenção:** O arquivo de exemplo não deve ser utilizado como teste oficial do montador implementado.

Deverão ser escritos ao menos **dois** programas em Assembly que, juntos, executem ao menos **dois terços** das instruções da máquina *Swombat* e ao menos uma chamada de função ou procedimento.

A implementação desses programas de teste será avaliada e, portanto, os seus códigos não devem ser compartilhados entre os grupos. Como mencionado na Seção 2, os arquivos de teste ".a" deverão ser incluídos no diretório "tst" do arquivo de entrega do trabalho, e devidamente explicados na documentação (o que se propõe fazer, qual o resultado esperado e outras informações que possam ser relevantes).

6. Formato de Saída de Dados do Montador

O grupo deverá optar por um dos formatos de saída abaixo, os quais são suportados pelo CPUSim:

- 1. Saída em formato ".hex" (Intel HEX): Para a máquina Swombat, este será um arquivo texto que irá separar as instruções de máquina byte-a-byte, uma vez que cada célula de memória possui um byte. Em outras palavras, cada linha do arquivo nesse formato define uma célula de memória que possuirá um dos dois bytes de uma instrução, em formato hexadecimal. Além desse dado, outras informações são definidas segundo o padrão Intel HEX, como o tipo do dado, o número de bytes da linha (nesse caso, sempre 1 byte), bem como o checksum (soma de verificação). Mais informações de como esse arquivo deve ser escrito podem ser encontrados no site da Wikipédia em https://en.wikipedia.org/wiki/Intel_HEX#Format
- 2. Saída em formato ".mif": Arquivo texto que define cada célula de memória (1 byte no caso da máquina Swombat), em formato binário (Caracteres 1 e 0 em ASCII).

Como forma de auxiliar os grupos na escolha do formato mais adequado, bem como ter uma referência na hora de implementar o montador, os grupos podem utilizar o montador interno do CPUSim para gerar as instruções de máquina em um dos formatos acima. Para isso, execute os seguintes passos:

- 1. Com o CPUSim aberto, vá em File->Open Machine... Selecione a máquina Swombat.
- 2. Vá em *File->Open Text...* e abra o arquivo ".a" desejado (faça isso inicialmente com o arquivo exemplo fornecido).
- 3. Vá em *Execute->Assemble & load*. Esse comando irá executar o montador interno do CPUSim e irá carregar o programa para a memória RAM, situada na direita do simulador. (OBS: Serão apresentadas mensagens de erro caso o CPUSim encontre problemas no código Assembly fornecido, e consequentemente o mesmo não será carregado para a memória RAM.)
- 4. Por fim, vá em *File->Save RAM->from Main...* Escolha o formato (".hex" ou ".mif") e escolha a pasta de destino. **Atenção**: apenas a memória principal (*Main*) será avaliada, já que a pilha estará vazia na inicialização de qualquer programa.
- 5. O arquivo gerado pode ser aberto em qualquer editor de texto. (OBS: O formato do arquivo gerado pelo grupo deve ser o mesmo que o gerado pelo montador do CPUSim. Por isso, o grupo pode utilizar este arquivo como uma forma de orientação na resolução de quaisquer problemas no código fonte do montador.)

7. Formato de Chamada do Simulador

A saída do montador (arquivo no formato ".hex" ou ".mif") deve ser executada na máquina Swombat do CPUSim para garantir que o programa Assembly foi traduzido corretamente. Para isso, execute os seguintes passos:

- 1. Com o CPUSim aberto, vá em File->Open Machine... Selecione a máquina Swombat.
- 2. Vá em *File->Open RAM->into Main....* Selecione o arquivo (".hex" ou ".mif") gerado pelo montador.
- 3. Com o arquivo carregado na memória RAM, vá em *Execute->Run*. O programa carregado na memória RAM será executado.

8. Sobre a Documentação

- Deve conter todas as decisões de projeto.
- Deve conter informações sobre cada programa testado, sobre o que ele faz, entradas de dados, saída esperada, etc.
- Deve conter elementos que comprovem que o montador foi testado (Ex.: imagens das telas de montagem e execução no CPUSim). Quaisquer arquivos relativos a testes devem ser enviados no pacote do trabalho, como mencionado na Seção 2. A documentação deve conter referências a esses arquivos, explicação do que eles fazem e dos resultados obtidos.
- O código fonte não deve ser incluído no arquivo PDF da documentação.

9. Considerações Finais

É obrigatório o cumprimento fiel de todas as especificações descritas neste documento. As decisões de projeto devem fazer parte apenas da estrutura interna do montador, não podendo afetar a interface de entrada e saída.

ANEXO 1 - Conjunto de Instruções da Máquina Swombat

É possível ver uma representação gráfica das instruções dentro do CPUSim. Para tanto, vá em *Modify -> Machine Instructions...* A ordem dos operandos de cada instrução (de cima para baixo) é definida a seguir.

Algumas instruções possuem bits não utilizados, os quais devem ser preenchidos com o valor zero, caso contrário, o programa não funcionará corretamente. A posição desses campos (se houverem) também é definida em cada instrução.

Código da operação (5 bits)	00 - exit
Significado	Encerra o programa, através da escrita da flag halt-bit.
Operandos	● Bits não utilizados (valor zero) – 11 bits
Ação	N/A
Exemplo	exitlabel: exit
Representação gráfica	5 11 op un11

Código da operação (5 bits)	01 - loadi
Significado	Carrega um dado da memória para um registrador. (OBS.: Caso o endereço especificado seja o de E/S, ou seja, 254, será solicitado ao usuário que digite um valor.)
Operandos	 Registrador – 3 bits Endereço de Memória – 8 bits
Ação	Registrador << Memória[endereço]
Exemplo	 loadi R0 150 loadi R0 var ; 'var' alocada na memória com a pseudo-instrução .data label: loadi R0 var
Representação gráfica	5 3 8 op reg addr

Código da operação (5 bits)	02 – storei
Significado	Escreve o valor de um registrador em um endereço de memória. (OBS.: Caso o endereço especificado seja o de E/S, ou seja, 254, o valor contido no Registrador 1 será impresso na tela.)
Operandos	 Registrador – 3 bits Endereço de Memória – 8 bits
Ação	Registrador >> Memória[endereço]
Exemplo	 storei R0 150 storei R0 var ; 'var' alocada na memória com a pseudo-instrução .data label: storei R0 var
Representação gráfica	5 3 8 op reg addr

Código da operação (5 bits)	03 - add
Significado	Faz a soma dos valores de dois registradores e escreve o resultado no primeiro.
Operandos	 Registrador 1 – 3 bits Registrador 2 – 3 bits Bits não utilizados (valor zero) – 5 bits
Ação	Registrador 1 << Registrador 1 + Registrador 2
Exemplo	add R0 R1label: add R0 R1
Representação gráfica	5 3 3 5 op reg reg un5

Código da operação (5 bits)	04 - subtract
Significado	Faz a subtração dos valores de dois registradores e escreve o resultado no primeiro.
Operandos	• Registrador 1 – 3 bits

	 Registrador 2 – 3 bits Bits não utilizados (valor zero) – 5 bits
Ação	Registrador 1 << Registrador 2
Exemplo	subtract R0 R1label: subtract R0 R1
Representação gráfica	5 3 3 5 op reg reg un5

Código da operação (5 bits)	05 - multiply
Significado	Faz a multiplicação dos valores de dois registradores e escreve o resultado no primeiro.
Operandos	 Registrador 1 – 3 bits Registrador 2 – 3 bits Bits não utilizados (valor zero) – 5 bits
Ação	Registrador 1 << Registrador 1 * Registrador 2
Exemplo	multiply R0 R1label: multiply R0 R1
Representação gráfica	5 3 3 5 op reg reg un5
Código da operação (5 bits)	06 - divide
Significado	Faz a divisão dos valores de dois registradores e escreve o resultado no primeiro.
Operandos	 Registrador 1 – 3 bits Registrador 2 – 3 bits Bits não utilizados (valor zero) – 5 bits
Ação	Registrador 1 << Registrador 1 / Registrador 2
Exemplo	divide R0 R1label: divide R0 R1
Representação gráfica	5 3 3 5 op reg reg un5

Código da operação (5 bits)	07 - jump
Significado	Pula para a instrução contida no endereço de memória especificado.
Operandos	 Bits não utilizados (valor zero) – 3 bits Endereço de Memória – 8 bits
Ação	pc << endereço
Exemplo	jump 50jump labellabel: jump 50
Representação gráfica	op un3 addr

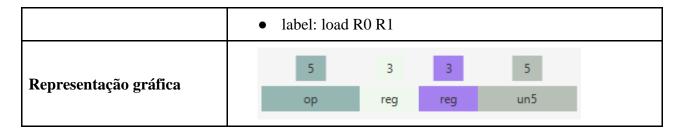
Código da operação (5 bits)	08 - jmpz
Significado	Pula para a instrução contida no endereço de memória especificado, caso o valor contido no registrador seja igual a zero.
Operandos	 Registrador – 3 bits Endereço de Memória – 8 bits
Ação	Se (Registrador = 0) pc << endereço
Exemplo	 jmpz R0 50 jmpz R0 label label: jmpz R0 50
Representação gráfica	op reg addr

Código da operação (5 bits)	09 - jmpn
Significado	Pula para a instrução contida no endereço de memória especificado, caso o valor contido no registrador seja menor que zero.

Operandos	 Registrador – 3 bits Endereço de Memória – 8 bits
Ação	Se (Registrador < 0) pc << endereço
Exemplo	 jmpn R0 50 jmpn R0 label label: jmpn R0 50
Representação gráfica	5 3 8 op reg addr

Código da operação (5 bits)	10 - move
Significado	Copia o conteúdo do Registrador 2 para o Registrador 1.
Operandos	 Registrador 1 – 3 bits Registrador 2 – 3 bits Bits não utilizados (valor zero) – 5 bits
Ação	Registrador 1 << Registrador 2
Exemplo	move R0 R1label: move R0 R1
Representação gráfica	5 3 3 5 op reg reg un5

Código da operação (5 bits)	11 - load
Significado	Carrega o endereço de memória contido no Registrador 2 para o Registrador 1. (OBS.: Caso o endereço especificado seja o de E/S, ou seja, 254, será solicitado ao usuário que digite um valor.)
Operandos	 Registrador 1 – 3 bits Registrador 2 – 3 bits Bits não utilizados (valor zero) – 5 bits
Ação	Registrador 1 << Memória[Registrador 2]
Exemplo	• load R0 R1



Código da operação (5 bits)	12 - store
Significado	Armazena o valor contido no Registrador 1 no endereço de memória contido no Registrador 2. (OBS.: Caso o endereço especificado seja o de E/S, ou seja, 254, o valor contido no Registrador 1 será impresso na tela.)
Operandos	 Registrador 1 – 3 bits Registrador 2 – 3 bits Bits não utilizados (valor zero) – 5 bits
Ação	Registrador 1 >> Memória[Registrador 2]
Exemplo	store R0 R1label: store R0 R1
Representação gráfica	5 3 3 5 op reg reg un5

Código da operação (5 bits)	13 - loadc	
Significado	Carrega uma constante de 8 bits (com sinal) em um registrador. A constante é armazenada nos 8 bits menos significativos do registrador.	
Operandos	 Registrador – 3 bits Constante – 8 bits 	
Ação	Registrador 1[8-15] << Constante	
Exemplo	loadc R0 150label: loadc R0 150	
Representação gráfica	5 3 8 op reg sgn5	

Código da operação (5 bits)	14 - clear
Significado	Zera o valor contido em um registrador.
Operandos	 Registrador – 3 bits Bits não utilizados (valor zero) – 8 bits
Ação	Registrador << 0
Exemplo	clear R0label: clear R0
Representação gráfica	5 3 8 op reg un8

Código da operação (5 bits)	15 - negate
Significado	Inverte o sinal do valor de um registrador e armazena o resultado em outro.
Operandos	 Registrador 1 – 3 bits Registrador 2 – 3 bits Bits não utilizados (valor zero) – 5 bits
Ação	Registrador 1 << - Registrador 2
Exemplo	negative R0 R1label: negative R0 R1
Representação gráfica	5 3 3 5 op reg reg un5

Código da operação (5 bits)	16 - push
Significado	Insere um valor contido em um registrador na pilha.
Operandos	 Registrador – 3 bits Bits não utilizados (valor zero) – 8 bits
Ação	Pilha[stackpt] << R0 stackpt << stackpt + 2

Exemplo	push R0label: push	R0		
Representação gráfica	5	3	8	
	ор	reg	un8	

Código da operação (5 bits)	17 - pop
Significado	Remove o topo da pilha e o coloca em um registrador.
Operandos	 Registrador – 3 bits Bits não utilizados (valor zero) – 8 bits
Ação	stackpt << stackpt - 2 R0 << Pilha[stackpt]
Exemplo	pop R0label: pop R0
Representação gráfica	op reg un8

Código da operação (5 bits)	18 - addi
Significado	Adiciona ao valor de um registrador uma constante (positiva ou negativa).
Operandos	 Registrador de destino – 3 bits Constante (com sinal) – 8 bits
Ação	Registrador << Registrador + constante
Exemplo	 addi R0 2 addi R0 -2 label: addi 2
Representação gráfica	5 3 8 op reg sgn5

Código da operação (5 bits)	19 - call	
Significado	Chama o procedimento que está contido no endereço de memória especificado. Empilha o endereço da próxima instrução a ser chamada (pc, antes do redirecionamento) para ser usado pelo return posteriormente.	
Operandos	 Bits não utilizados (valor zero) – 3 bits Endereço de memória – 8 bits 	
Ação	<pre>push(pc) pc << Memória[endereço]</pre>	
Exemplo	 call proc ;Procedimento identificado pelo label 'proc' label: call proc 	
Representação gráfica	5 3 8 op un3 addr	

Código da operação (5 bits)	20 - return	
Significado	Encerra um procedimento e retorna para endereço especificado pelo valor no topo da pilha. ATENÇÃO: uma função chamada que insere elementos na pilha deve obrigatoriamente removê-los antes do return .	
Operandos	Bits não utilizados (valor zero) – 11 bits	
Ação	<i>pc</i> << pop ()	
Exemplo	• return	
Representação gráfica	5 11 op un11	