

アセンブリ言語による整列アルゴリズムに関する実験

1250373 溝口洸熙 *

2022 年 11 月 2 日

目次

1	アセンブリ言語による整列アルゴリズム記述可否の検証	1
1.1	実験の目的	1
1.2	プログラムの仕様	1
1.3	実験	1
1.3.1	実験方法	1
1.3.2	実験結果	2
1.4	考察	3
2	アセンブリ言語と高級言語の実装に関する違い	4
2.1	実験の目的	4
2.2	実験の方法	4
2.3	実験の結果	4
2.4	考察	4
	謝辞	6
	参考文献	6
	付録 A ソースコード	7

* 高知工科大学 情報学群 2 年 清水研究室

1 アセンブリ言語による整列アルゴリズム記述可否の検証

1.1 実験の目的

高級プログラミング言語、Java、C、Python などは、『コンパイラ』と呼ばれる装置を通して機械語に書き換えられ、コンピュータで実行されている。

それに対して、アセンブリ言語は各機械語命令につけられた「意味する名前」(ニーモニック；mnemonic)を使ってプログラムを表記する表記法である。[1] また、アセンブリ言語表記を機械語のビット列に変換する作業をアセンブルと言い、それを行うソフトウェアをアセンブラと言う。つまり、コンパイラとアセンブラは別物であり、アセンブラは機械語の表記を変えたものである故にコンピュータへの命令を1対1で書き換えるものである点がコンパイラと大きく違う点である。

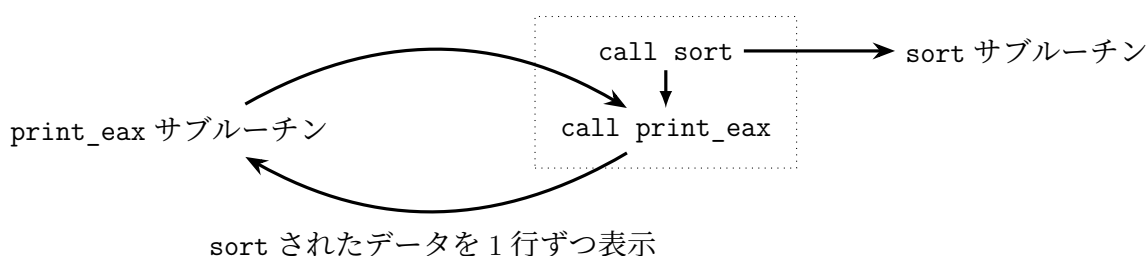
本実験課題の目的は、このようなアセンブリ言語・機械語に対して、コンパイラを使わずに整列アルゴリズムを直接技術することが可能であるか確認することである。

1.2 プログラムの仕様

0 以上 2^{31} 未満の自然数列に対して昇順に整列するアルゴリズムをアセンブリ言語で記述する。その際、test_sort.s ファイルが sort サブルーチンを呼び出して整列を行う。検証したい自然数列は、test_sort.s 内の data1 にダブルワードで、データの個数は ndata1 で定義しており、test_sort.s 実行時 data1 に格納してある自然数列が print_eax サブルーチンによって出力される。

data1 の先頭番地は EBX, ndata1 は ECX に格納する。sort の呼び出し前後で他の汎用レジスタの値は変化しないように設計されている。内部処理の概要を Fig 1.1 に示す。整列アルゴリズムは、バブルソート^{*1}アルゴリズムを元に作成している。

Fig 1.1: 処理概要



1.3 実験

1.3.1 実験方法

src 1.1 に示すコンピュータを利用して、アセンブリ言語で書いた整列プログラムを実行可能ファイルに変換し実行した。実行結果を確認し、整列されているか確認するとともに、Java で記述したプログラムとアセンブリプログラムを比較してどの部分をどのようにアセンブリ言語化したかを確認する。

*1 バブルソートの簡単な説明

src 1.1: 使用したコンピュータ

```
$ uname -a
Linux KUT20VLIN-462 5.4.0-70-generic #78~18.04.1-Ubuntu SMP Sat Mar 20 14:10:07 UTC
2021 x86_64 x86_64 x86_64 GNU/Linux
```

print_eax.s, test_sort.s, sort.s をそれぞれアセンブルして実行する.

src 1.2: 実行したコマンド

```
$ nasm -felf print_eax.s ; nasm -felf test_sort.s ; nasm -felf sort.s
$ ld -m elf_i386 print_eax.o test_sort.o sort.o
$ ./a.out
```

1.3.2 実験結果

実験結果と期待される結果を Tbl 1.1 に示す. これからわかるように, 自然数列を昇順に整列されていることが確認できる.

Tbl 1.1: 実験結果と比較

入力	1	3	5	7	9	2	4	6	8	0	1	2
出力	0	1	1	2	2	3	4	5	6	7	8	9
期待出力	0	1	1	2	2	3	4	5	6	7	8	9

Java で記述したバブルソート (src 1.3) と, アセンブリ言語で記述したバブルソート (src 1.4) を比較して, 言語化の箇所を確認する. 前者は入力データ列を data とし, 後者は EBX に入力データ列の先頭番地, ECX 入力データ列の個数を格納している. ソースコードの各処理の内容と, 2つの言語による対応を Tbl 1.2 に記す.

Tbl 1.2: ソースコードの行対応

処理内容	src 1.3	src 1.4
ループ処理 1 の条件比較・変数処理	3	2 - 3, 25, 26
最大値の更新	4	5
最大値インデックスの更新	5	6
ループ処理 2 の条件比較・変数処理	6	7 - 9, 18, 19
最大値の比較・更新	7 - 11	10 - 17
data[max_index] と data[i] の入れ替え	12 - 14	21 - 24

src 1.3: Java

```

1  int max_index = 0;
2  int max = 0;
3  for (int i = data.length - 1; i > 0; i--) {
4      max = data[0];
5      max_index = 0;
6      for (int j = 1; j <= i; j++) {
7          if (data[j] >= max) {
8              max = data[j];
9              max_index = j;
10         }
11     }
12     int m = data[max_index];
13     data[max_index] = data[i];
14     data[i] = m;
15 }

```

注)

src 1.3, src 1.4 いずれも、整列アルゴリズムの部分のみ掲載している。

src 1.4: アセンブリ

```

1  loop0:
2      cmp     ecx, 0
3      jle     endp
4      mov     edx, [ebx]
5      mov     eax, 0
6      mov     edi, 1
7      loop1:
8          cmp     edi, ecx
9          jg     loop01
10         mov     esi, [ebx + edi*4]
11         cmp     esi, edx
12         jge     then
13         jmp     endif
14     then:
15         mov     edx, [ebx + edi*4]
16         mov     eax, edi
17     endif:
18         inc     edi
19         jmp     loop1
20     loop01:
21         mov     esi, [ebx + eax*4]
22         mov     edi, [ebx + ecx*4]
23         mov     [ebx + eax*4], edi
24         mov     [ebx + ecx*4], esi
25         dec     ecx
26         jmp     loop0

```

1.4 考察

実験結果より自然数列を整列アルゴリズムをアセンブリ言語で記述できることがわかった。

ただ、これはあくまで 0 以上 2^{31} 未満の自然数に限った整列アルゴリズムであるため、負の整数やその他の有理数などを対象にした整列アルゴリズムが記述可能であるかは、この実験では検証できていない。

2 アセンブリ言語と高級言語の実装に関する違い

2.1 実験の目的

先にも述べたように、アセンブリ言語は機械語を 1 対 1 に記した記法である。それに対して高級言語（高水準言語）は、人間の言語・概念に近づけて設計されたプログラム言語である。[2]

本実験の目的は、アセンブリ言語や機械語などの低級言語は Java などの高級言語と比べて、実装に関して大きく異なる点があるかどうかを明らかにすることであり、コードの複雑さやコード量がどうなるかを明らかにすることである。

2.2 実験の方法

実装に関して大きく異なる点を明らかにするため、低級言語と高級言語の一般的な実装手順を書き出し比較をする。低級言語はアセンブリ言語、高級言語は Java を利用する。

さらに、コードの複雑さやコード量がどうなるかを明らかにするため、実際に 2 つの言語で書いた同一のアルゴリズムに対して、行数や条件分岐、ループの回数を比べる。

Java におけるループ回数は for 文の個数、比較回数は if 文の個数とし、アセンブリ言語における比較回数は cmp の個数、ループ回数は一定条件下で上の行のラベルにジャンプする回数とする。評価値は

$$\text{評価値} = \text{行数} + \text{比較回数} + \text{ループ回数}$$

と定義し、評価値とコードの複雑さは比例するものとする。

2.3 実験の結果

付録 A.1, 付録 A.2 は、アセンブリ言語で記述したプログラム、付録 A.3 は Java で記述したプログラムである。いずれも、入力データを受け取りバブルソートアルゴリズムで整列してその整列結果を 1 行ずつ出力するプログラムであり、入力と出力は一致している。それぞれの行数と比較回数を Tbl 2.1 に示す。

Tbl 2.1: 行数とループ・比較回数

記述言語	ファイル名	行数	比較回数	ループ回数	評価値
アセンブリ言語	sort.s	44	3	2	77
	test_sort.s	26	1	1	
Java	sort.java	19	1	2	22

2.4 考察

実験の結果より、両言語の評価値を比べるとアセンブリ言語の評価値の方が Java に比べて 3.5 倍であることが確認できる。1 番目立った違いは行数だろう。アセンブリ言語で記述したものに比べて Java で記述したアルゴリズムは約 1/3 とより簡潔に記述できることが分かる。

その要因として Java のループに使われる for 文は、比較とループ、ループ変数の定義と処理を 1 行で行うことができることが挙げられる。

今回の実験で、低級言語であるアセンブリ言語の方が、高級言語である Java よりも複雑でゴードの量も多くなることが分かった。ただし、この実験では独自の指標でコードの複雑さを測っているため、一般的な複雑の指標である、**サイクロマティック複雑度（循環的複雑度）**で計測できていない。

謝辞

本実験課題は本学情報学群 1250372 三上柊氏と共同で実施した。また本学情報学群の高田 喜朗准教授には、整列アルゴリズムに関する様々な助言をいただいた。これらの方々に深く感謝いたします。

参考文献

- [1] 高田喜朗. 情報学群実験第 2 テキスト, 第 1 章.
<http://www.info.kochi-tech.ac.jp/y-takata/pl2/part1/hellohex.html>, 2022.
- [2] 新村出. 広辞苑. 岩波書店, 第 6 版, 2008.

付録 A ソースコード

付録 A.1: sort.s

```
1      section .text
2      global sort
3      sort:
4          push esi
5          push edi
6          push edx
7          push ecx
8          push ebx
9          push eax
10         dec ecx
11     loop0:
12         cmp ecx, 0 ; ecx = i
13         jle endp
14         mov edx, [ebx] ; max = data[0]
15         mov eax, 0 ; max_index = 0
16         mov edi, 1 ; edi = j
17     loop1:
18         cmp edi, ecx ; j > i?
19         jg loop01
20         mov esi, [ebx + edi*4] ; data[j]
21         cmp esi, edx ; data[j] >= max
22         jge then
23         jmp endif
24     then:
25         mov edx, [ebx + edi*4] ; max = data[j]
26         mov eax, edi ; max_index = j
27     endif:
28         inc edi
29         jmp loop1
30     loop01:
31         mov esi, [ebx + eax*4] ; m = data[max_index]
32         mov edi, [ebx + ecx*4] ; edi = data[i]
33         mov [ebx + eax*4], edi ; data[max_index], data[i]
34         mov [ebx + ecx*4], esi ; data[i] = m
35         dec ecx
36         jmp loop0
37     endp:
38         pop eax
39         pop ebx
40         pop ecx
41         pop edx
42         pop edi
43         pop esi
44         ret
```


付録 A.2: test_sort.java

```

1  section .text
2  global _start
3  extern sort, print_eax
4
5  _start:
6      mov ebx, data
7      mov ecx, ndata
8      call sort      ; ソート
9      mov edi, 0
10 loop:    ; 結果の出力
11      cmp edi, ndata
12      je endp
13      mov eax, [data + edi * 4]
14      call print_eax
15      inc edi
16      jmp loop
17
18
19 endp:
20      mov eax, 1
21      mov ebx, 0
22      int 0x80
23
24  section .data
25  data:  dd 1,3, 5, 7, 9, 2, 4, 6, 8, 0, 1, 2
26  ndata  equ ($ - data) / 4

```

付録 A.3: sort.java

```

1  class sort{
2      public static void main(String[] args){
3          int[] data = {1, 3, 5, 7, 9, 2, 4, 6, 8, 0, 1, 2};
4          bubble(data);
5          for(int d: data){
6              System.out.println(d); // 出力
7          }
8      }
9      private static void bubble(int[] data) {
10         int max_index = 0; int max = 0;
11         for (int i = data.length - 1; i > 0; i--) {
12             max = data[0]; max_index = -0;
13             for (int j = 1; j <= i; j++) {
14                 if (data[j] >= max) { max = data[j]; max_index = j; }
15             }
16             int m = data[max_index]; data[max_index] = data[i]; data[i] = m;
17         }
18     }
19 }

```