中国科学技术大学计算机学院《计算机系统详解》实验报告



实验题目: CSAPP_labs

学生姓名: 胡毅翔

学生学号: PB18000290

完成日期: 2020年6月24日

计算机实验教学中心制 2019 年 09 月

实验目的

1. 结合课上所学,完成 CSAPP 课程实验 Data Lab,Bomb Lab,Cache Lab。

实验环境

- 1. PC 一台
- 2. Windows 系统
- 3. Ubuntu

Data Lab

本实验内容为根据要求用有限的操作实现下列函数:

Name	Description	Rating	Max ops
bitXor(x,y)	x y using only & and ~.	1	14
tmin()	Smallest two's complement integer	1	4
isTmax(x)	True only if x x is largest two's comp. integer.	1	10
allOddBits(x)	True only if all odd-numbered bits in x set to 1.	2	12
negate(x)	Return -x with using - operator.	2	5
isAsciDigit(x)	True if $0 \times 30 \le x \le$.	3	15
conditional	Same as x ? y : z	3	16
isLessOrEqual(x, y)	True if $x \le y$, false otherwise	3	24
logicalNeg(x))	Compute ! x without using ! operator.	4	12
howManyBits(x)	Min. no. of bits to represent x in two's comp.	4	90
floatScale2(uf)	Return bit-level equiv. of 2*f for f.p. arg. f.	4	30
floatFloat2Int(uf)	Return bit-level equiv. of (int) f for f.p. arg. f.	4	30
floatPower2(x)	Return bit-level equiv. of 2.0 x for integer x.	4	30

表 1

核心代码部分详见链接。

Bomb Lab

根据实验要求,需要找到对应的 6 个字符串。而提供的源文件中并不包含所需的信息。 因此,需要通过反汇编来实现。

首先输入"objdump -d bomb > bomb.txt",得到 bomb 的汇编代码。进而通过阅读汇编代码,并使用 gdb 进行调试,完成炸弹的解码过程。

第一个密码,通过下列汇编代码,及 strings_not_equal 的代码可知:密码为存在地址 0x402400 的字符串。

0000000000400ee0 <phase_1>:

400ee0:	48 83 ec 08	sub	\$0x8,%rsp
400ee4:	be 00 24 40 00	mov	\$0x402400,%esi
400ee9:	e8 4a 04 00 00	callq	401338 <strings_not_equal></strings_not_equal>
400eee:	85 c0	test	%eax,%eax
400ef0:	74 05	je	400ef7 <phase_1+0x17></phase_1+0x17>
400ef2:	e8 43 05 00 00	callq	40143a <explode_bomb></explode_bomb>
400ef7:	48 83 c4 08	add	\$0x8,%rsp
400efb:	c3	retq	

在 gdb 中输入 "x/ls 0x402400",得到密码 "Border relations with Canada have never been better."

图 1

第二个密码中,通过"read_six_numbers"可知密码为 6 个数字。再阅读下列代码后可知,密码为首项为 1,公比为 2 的等比数列。

400f0a:	83 3c 24 01	cmpl	\$0x1,(%rsp)
400f0e:	74 20	je	400f30 <phase_2+0x34></phase_2+0x34>
400f10:	e8 25 05 00 00	callq	40143a <explode_bomb></explode_bomb>
400f15:	eb 19	jmp	400f30 <phase_2+0x34></phase_2+0x34>
400f17:	8b 43 fc	mov	-0x4(%rbx),%eax
400f1a:	01 c0	add	%eax,%eax
400f1c:	39 03	cmp	%eax,(%rbx)
400f1e:	74 05	je	400f25 <phase_2+0x29></phase_2+0x29>
400f20:	e8 15 05 00 00	callq	40143a <explode_bomb></explode_bomb>
400f25:	48 83 c3 04	add	\$0x4,%rbx
400f29:	48 39 eb	cmp	%rbp,%rbx
400f2c:	75 e9	jne	400f17 <phase_2+0x1b></phase_2+0x1b>
400f2e:	eb 0c	jmp	400f3c <phase_2+0x40></phase_2+0x40>
400f30:	48 8d 5c 24 04	lea	0x4(%rsp),%rbx
400f35:	48 8d 6c 24 18	lea	0x18(%rsp),%rbp
400f3a:	eb db	jmp	400f17 <phase_2+0x1b></phase_2+0x1b>
11	r		

故密码为12481632.

```
G hyx@DESKTOP-LP1A2G2: ~(Github/CSAPP_labs/bomb) - X

The bomb has blown up,
hyx@DESKTOP-LP1A2G2: *(Github/CSAPP_labs/bomb)$ ./bomb

Welcome to my fiendish little bomb. You have 6 phases with
which to blow yourself up. Have a nice day!
Border relations with Canada have never been better.
Phase 1 defused. How about the next one?
1 2 4 8 16 32

That's number 2. Keep going!
```

对于第三个密码,由下列代码知,输入的值的个数应大于等于 2:

400bf0 <__isoc99_sscanf@plt> 400f5b: e8 90 fc ff ff callq

400f60: 83 f8 01 cmp \$0x1,%eax

400f63: 7f 05 400f6a < phase 3+0x27> jg 400f65: e8 d0 04 00 00 callq 40143a <explode_bomb>

再由下列代码知,第一个数应小于等于7:

400f6a: 83 7c 24 08 07 cmpl \$0x7,0x8(%rsp)

400f6f: 77 3c 400fad <phase 3+0x6a> ja

400f71: 8b 44 24 08 0x8(%rsp),%eax mov 400f75: ff 24 c5 70 24 40 00 impa *0x402470(,%rax,8)

不妨令第一个数为 1, 跳转到 0x400fb9:

400fb9: b8 37 01 00 00 mov \$0x137,%eax 400fbe: 3b 44 24 0c cmp 0xc(%rsp),%eax

400fc2: 74 05 400fc9 <phase_3+0x86> je 400fc4: e8 71 04 00 00 callq 40143a <explode bomb>

400fc9: 48 83 c4 18 \$0x18,%rsp add

400fcd: с3 retq

观察可知,须令第二个数为 0x137 即 311,即可返回。

故第三个密码为1311.

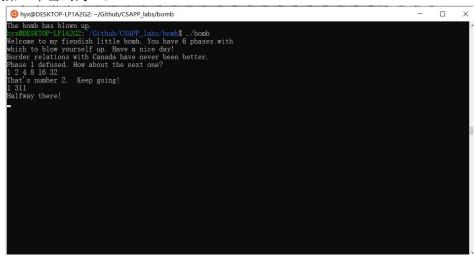


图 3

对于第四个密码,阅读下列代码知,须输入两个数:

401024: e8 c7 fb ff ff 400bf0 <__isoc99_sscanf@plt> callq

401029: 83 f8 02 \$0x2,%eax cmp

40102c: 75 07 401035 < phase 4+0x29> jne

而由如下这段知,第一个数应小于等于14:

40102e: 83 7c 24 08 0e \$0xe,0x8(%rsp) cmpl

401033: 76 05 40103a <phase_4+0x2e> jbe 40143a <explode_bomb> 401035: e8 00 04 00 00 callq

mov

40103a: ba 0e 00 00 00 \$0xe,%edx mov 40103f: be 00 00 00 00 \$0x0,%esi mov 8b 7c 24 08 401044: 0x8(%rsp),%edi 401048: e8 81 ff ff ff callq 400fce <func4>

跳转后进入 func4,由 func4 内代码判断,输入的第一个数应为 7.

而由 phase_4 后半部分代码,知第二个数必须为 0:

40104d: 85 c0 test %eax,%eax

40104f: 75 07 jne 401058 < phase 4+0x4c>

401051: 83 7c 24 0c 00 cmpl \$0x0,0xc(%rsp)

401056: 74 05 je 40105d <phase_4+0x51> 401058: e8 dd 03 00 00 callq 40143a <explode_bomb>

40105d: 48 83 c4 18 add \$0x18,%rsp

401061: c3 retq

图 4

紧接着第5个密码,由下列指令知,字符串长度为6:

40107a: e8 9c 02 00 00 callq 40131b <string_length>

40107f: 83 f8 06 cmp \$0x6,%eax 此处为一个循环,内容为对字符串进行逐字符操作:

40108b: 0f b6 0c 03 movzbl (%rbx,%rax,1),%ecx

 40108f:
 88 0c 24
 mov
 %cl,(%rsp)

 401092:
 48 8b 14 24
 mov
 (%rsp),%rdx

 401096:
 83 e2 0f
 and
 \$0xf,%edx

401099: 0f b6 92 b0 24 40 00 movzbl 0x4024b0(%rdx),%edx

4010a0: 88 54 04 10 mov %dl,0x10(%rsp,%rax,1)

 4010a4:
 48 83 c0 01
 add
 \$0x1,%rax

 4010a8:
 48 83 f8 06
 cmp
 \$0x6,%rax

4010ac: 75 dd jne 40108b <phase 5+0x29>

随后判断操作后的字符串是否与目标字符串相等:

4010bd: e8 76 02 00 00 callq 401338 <strings_not_equal>

4010c2: 85 c0 test %eax,%eax

4010c4: 74 13 je 4010d9 <phase_5+0x77> 4010c6: e8 6f 03 00 00 callq 40143a <explode_bomb>

图 5

图 6

目标字符串为 flyers,而操作后的字符串中每个值,为操作前的值与 0xf 与后得到的值 a 对应于首地址为 0x4024b0 的字符数组的第 a 个值。因此,输入的字符串的 ASCII 码后四位(二进制)分别应为 9, f, e, 5, 6, 7。故 9?>567 可作为第 5 个密码。

```
The bomb has blown up.
hyx@DESKTOP-LPIA2G2: /Github/CSAPP_labs/bomb$ ./bomb
Welcome to my flendish little bomb. You have 6 phases with
which to blow yourself up. Have a nice day!
Border relations with Canada have never been better.
Phase 1 defused. How about the next one?
1 2 4 8 16 32
That's number 2. Keep going!
1 311
Halfway there!
7 0
So you got that one. Try this one.
9/2567
SOOM!!!
The bomb has blown up.
hyx@DESKTOP-LPIA2G2: /Github/CSAPP_labs/bomb$ ./bomb
Welcome to my flendish little bomb. You have 6 phases with
which to blow yourself up. Have a nice day!
Border relations with Canada have never been better.
Phase 1 defused. How about the next one?
1 2 4 8 16 32
That's number 2. Keep going!
1 311
Halfway there!
7 0
So you got that one. Try this one.
9/2567
Good work! On to the next...
```

图 7

最后一个密码,代码量较大,从开头可看出,还是读取6个数字:

401106: e8 51 03 00 00 callq 40145c <read_six_numbers>

读取完后, 便是二重循环, 判断输入的数范围为 1-6, 且各不相同:

40110b:	49 89 e6	mov	%rsp,%r14
40110e:	41 bc 00 00 00 00	mov	\$0x0,%r12d
401114:	4c 89 ed	mov	%r13,%rbp
401117:	41 8b 45 00	mov	0x0(%r13),%eax
40111b:	83 e8 01	sub	\$0x1,%eax
40111e:	83 f8 05	cmp	\$0x5,%eax
401121:	76 05	jbe	401128 <phase_6+0x34></phase_6+0x34>
401123:	e8 12 03 00 00	callq	40143a <explode_bomb></explode_bomb>
401128:	41 83 c4 01	add	\$0x1,%r12d
40112c:	41 83 fc 06	cmp	\$0x6,%r12d
401130:	74 21	je	401153 <phase_6+0x5f></phase_6+0x5f>
401132:	44 89 e3	mov	%r12d,%ebx
401135:	48 63 c3	movslq %ebx,%rax	
401138:	8b 04 84	mov	(%rsp,%rax,4),%eax
40113b:	39 45 00	cmp	%eax,0x0(%rbp)
40113e:	75 05	jne	401145 <phase_6+0x51></phase_6+0x51>
401140:	e8 f5 02 00 00	callq	40143a <explode_bomb></explode_bomb>
401145:	83 c3 01	add	\$0x1,%ebx
401148:	83 fb 05	cmp	\$0x5,%ebx
40114b:	7e e8	jle	401135 <phase_6+0x41></phase_6+0x41>
40114d:	49 83 c5 04	add	\$0x4,%r13
401151:	eb c1	jmp	401114 <phase_6+0x20></phase_6+0x20>

之后对各个数进行 7-x 操作。

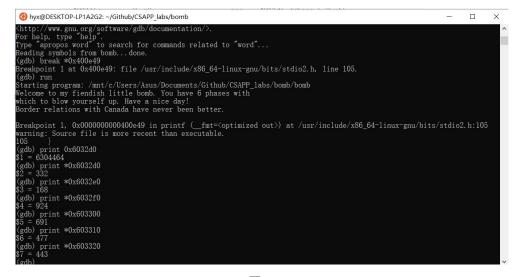


图 8

根据后半部分代码的要求,要使经 7-x 操作过后的数列对应到 0x6032d0 到 0x603320 的顺序为从大到小。而地址 0x6032d0 到 0x603320 上对应的数为,332,168,924,691,477,443。故从大到小顺序为 3,4,5,6,1,2。而 7-x 操作前的数列应为 4,3,2,1,6,5。炸弹拆解成功!

图 9

Cache Lab

本实验的第一部分为设计一个 cache 的模拟器,并根据模拟参数(s(索引位个数), E(行数), b(块的大小)),输出仿真结果。

实验的核心代码详见链接。

编译运行后,结果如下:

图 10

在进行第二部分实验前,需先安装 valgrind,输入 sudo apt install valgrind 后完成安装。 第二部分实验的核心代码详见<u>链接</u>。

之后便可输入测试命令,逐一测试,32*32,64*64,及61*67的矩阵转置,测试结果如下,基本符合实验预期:

```
### Open Control of Co
```

图 11

遇到的问题和解决方案说明

- 1. Data Lab 编译时出现报错,导致未能使用实验自带的测试程序,对实验结果进行测试。解决方案:未找到合适的解决办法,报错信息来自测试程序。
- 2. Cache Lab 运行第二个实验部分时,缺少 valgrind。 解决方案:在 Ubuntu 中输入 sudo apt install valgrind,完善环境。

体会与总结

Data Lab 在中实现的函数,有的与 Leetcode 上的题类似,都是考察对数在计算机中的表示的理解。通过这个实验复习了课上的知识,但因为对操作的限制,实验本身的难度还是不小。

Bomb Lab 是这些实验中最有意思的实验。虽然,汇编代码比较难阅读,但拆弹的过程趣味性十足。而且,通过这种寓教于乐的方式,真正深入了解了汇编语言。相较于组成原理课程中,简单的冒泡排序汇编程序设计,这种形式的实验(引导式的学习)有助于掌握汇编,而非简单的一带而过。此外,这次实验还学会了gdb 的基础操作。因此,我认为这一实验是我目前做过的实验中收获最多的。

Cache Lab 部分与组成原理所学内容比较接近,因此实验的过程并不算太复杂。当然,在矩阵转置优化过程中,也是需要有很多思考的内容。

总体而言,经过一学期这门课的学习,我学到了很多在计算机学院计划课程中所未能学到/未能学精的知识。当然,我也希望学校能更多地借鉴国外的类似 CSAPP 这样的课程,比如 MIT 的 The Missing Semester of Your CS Education。让我们能学到一些在日后科研/工作中应具备的基本技能,基本素养。

本次实验的所有文件都保存在了 Github 仓库中。

实验过程中,我也遇到了不少问题,在 csdn,stack overflow 等社区上也学习到了不少

实验以外的内容。整体而言,通过这门课所学确实超过了我的预期,而且也不至于像某些专业课那样负担大。可以说,这是性价比极高的一门课。